Theoréme de la limite centrale pour des fonctionelles non linéaires de la mesure empirique et pour le rééchantillonnage stratifié

Roberta Flenghi

▶ **To cite this version:**

Roberta Flenghi. Theoréme de la limite centrale pour des fonctionelles non linéaires de la mesure empirique et pour le rééchantillonnage stratifié. Mathématiques [math]. Ecole des Ponts, 2023. Français. ⟨NNT : ⟩. ⟨tel-04360933⟩

HAL Id: tel-04360933

https://hal.science/tel-04360933v1

Submitted on 22 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons CC0 1.0 - Universal - International License

École des Ponts
ParisTech

THÈSE DE DOCTORAT
de l'École des Ponts ParisTech

Theoréme de la limite centrale pour des fonctionelles non linéaires de la mesure empirique et pour le rééchantillonnage stratifié

École doctorale MSTIC

Discipline : Mathématiques

Thèse préparée au CERMICS, équipe Inria MATHRISK

Thèse soutenue le 20/12/2023, par
Roberta FLENGHI

Composition du jury:

Nicolas, CHOPIN Professeur, ENSAE	*Rapporteur*
Randal, DOUC Professeur, Télécom SudParis	*Rapporteur*
Pierre, DEL MORAL Directeur de Recherche, INRIA Bordeaux	*Examinateur*
Florence, MERLEVEDE Professeure, Université Gustave Eiffel	*Examinatrice*
Benjamin, JOURDAIN Professeur, École des Ponts	*Directeur de thèse*

Summary

This thesis is dedicated to the central limit theorem which is one of the two fundamental limit theorems in probability theory with the strong law of large numbers.

The central limit theorem which is well known for linear functionals of the empirical measure of independent and identically distributed random vectors, has recently been extended to non-linear functionals. The main tool permitting this extension is the linear functional derivative, one of the notions of derivation on the Wasserstein space of probability measures. We generalize this extension by first relaxing the equal distribution assumption and then the independence property to be able to deal with the successive values of an ergodic Markov chain.

In the second place, we focus on the stratified resampling mechanism. This is one of the resampling schemes commonly used in particle filters. We prove a central limit theorem for the first resampling according to this mechanis m under the assumption that the initial positions are independent and identically distributed and the weights proportional to a positive function of the positions such that the image of their common distribution by this function has a non zero component absolutely continuous with respect to the Lebesgue measure.
This result relies on the convergence in distribution of the fractional part of partial sums of the normalized weights to some random variable uniformly distributed on [0,1]. More generally, we prove the joint convergence in distribution of q variables modulo one obtained as partial sums of a sequence of i.i.d. square integrable random variables multiplied by a common factor given by some function of an empirical mean of the same sequence. The limit is uniformly distributed over $[0,1]^q$. To deal with the coupling introduced by the common factor, we assume that the common distribution of the random variables has a non zero component absolutely continuous with respect to the Lebesgue measure, so that the convergence in the central limit theorem for this sequence holds in total variation distance.
Under the conjecture that the convergence in distribution of fractional parts to some uniform random variable remains valid at the next steps of a particle filter which alternates selections according to the stratified resampling mechanism and mutations according to Markov kernels, we provide an inductive formula for the asymptotic variance of the resampled population after n steps. We perform numerical experiments which support the validity of this formula.

Résumé

Cette thèse porte sur le théorème de la limite centrale, l'un des deux théorèmes limites fondamentaux de la théorie des probabilités avec la loi forte des grands nombres.

Le théorème de la limite centrale usuel qui porte sur des fonctionnelles linéaires de la mesure empirique de vecteurs aléatoires indépendants et identiquement distribués a récemment été étendu à des fonctionnelles non linéaires par l'utilisation de la dérivée fonctionnelle linéaire sur l'espace de Wasserstein des mesures de probabilité. Nous généralisons cette extension à la mesure empirique de vecteurs aléatoires indépendants mais non identiquement distribués d'une part et à la mesure empirique des états successifs d'une chaîne de Markov ergodique d'autre part.

Dans un second temps, nous nous intéressons au rééchantillonnage stratifié qui est couramment utilisé dans les filtres particulaires. Nous prouvons un théorème de la limite centrale pour le premier rééchantillonnage sous l'hypothèse que les positions initiales des particules sont indépendantes et identiquement distribuées et leurs poids sont proportionnels à une fonction positive des positions qui envoie leur loi commune sur une probabilité possédant une composante non nulle absolument continue par rapport à la mesure de Lebesgue. Ce résultat repose sur la convergence en loi de la partie fractionnaire des sommes partielles de poids normalisés vers une variable aléatoire uniforme sur [0,1]. Plus généralement, nous montrons la convergence en loi vers un vecteur aléatoire uniforme sur $[0,1]^q$ de q sommes partielles d'une suite de variables aléatoires i.i.d. de carré intégrable multipliées par une fonction de la moyenne empirique de cette suite. Pour traiter le couplage introduit par ce facteur commun, nous supposons que la loi commune a une composante non nulle absolument continue par rapport à la mesure de Lebesgue, ce qui assure la convergence en variation totale dans le théorème de la limite centrale pour cette suite. Sous l'hypothèse que la convergence en loi de la partie fractionnaire des poids normalisés reste vraie au étapes suivantes d'un filtre particulaire calculé en alternant des étapes de rééchantillonnage suivant le méchanisme stratifié et des mutations suivant des noyaux markoviens, nous obtenons une formule de récurrence pour la variance asymptotique des particules après n étapes. Nous vérifions la validité de cette formule au travers d'expériences numériques.

Acknowledgements

Le plus grand merci va à Benjamin Jourdain. Merci d'avoir été un guide constant au cours de ces trois années. J'espère avoir absorbé un peu de ta rigueur et de ton intégrité intellectuelle. Merci d'avoir toujours su quand me pousser et me fournir le soutien dont j'avais besoin, tant d'un point de vue mathématique qu'humain. Et enfin, merci d'avoir souri (et non pleuré) face au choix "discutable" de mes notations et pas que...
Je tiens aussi à exprimer ma reconnaissance à Nicolas Chopin et à Randal Douc pour avoir accepté d'être les rapporteurs de ma thèse. Je voudrais également remercier les membres du jury qui ont accepté de participer à ma soutenance: Pierre Del Moral et Florence Merlevede.
Merci à Frédéric Meunier et Jean-Philippe Chancelier qui m'ont fait confiance et m'ont permis d'enseigner dans leurs cours.
Un remerciement particulier à Isabelle et Stéphanie, les piliers du CERMICS. Leur efficacité et leur disponibilité ont rendu tout extrêmement facile.
Je tiens aussi à remercier Virginie Ehrlacher et Olga Mula avec qui j'ai eu l'occasion de collaborer pendant le CEMRACS, je les remercie pour nous avoir accompagné et conseillé pour ce projet. Merci à tous les amis du CEMRACS, en particulier à Augustin, Giulia, Maria et Siwar. Une amitié née à Marseille et qui s'est poursuivie à Paris.
Un merci spécial à tous les doctorants du CERMICS qui ont rempli les journées de sourires, de légèreté et de desserts: Albéric, Alfred, Amandine, Andrea, Anton, Camila, Coco, Edoardo, Eloïse, Emanuele, Épiphane, Étienne, Fabian, Faten, Guillaume, Guido, Hélène, Hervé, Jean, Julien, Kacem, Laurent, Léo, Louis, Luca, Mathias, Michel, Mohamad, Nerea, Noé, Oscar, Raian, Regis, Rutger, Seta, Shiva, Simon, Solal, Vitor,Yue, Zoé. J'espère n'avoir oublié personne.
En particulier, je remercie Edo pour son authenticité et sa générosité, merci pour son rire qui emplissait le deuxième étage, et peut-être même le troisième. Merci à Ema qui accueille à bras ouverts tous les nouveaux arrivants, toujours prêt à aider n'importe qui et à manger n'importe quoi.
Merci à Amaury pour sa tendresse. Grazie al mio amico di una vita e per la vita, Marco.
E infine grazie alla mia famiglia, in particolare a mamma che ci ha insegnato la libertà e l'indipendenza.

Contents

1 Introduction **1**
1.1 Théorème de la Limite Centrale : une revue de la littérature . . . 1
1.1.1 Théorème de la Limite Centrale pour les variables aléatoires i.i.d. . . . 1
Convergence en loi . . . 1
Convergence pour la distance en variation totale . . . 2
1.1.2 Théorème de la Limite Centrale pour des variables aléatoires indépendantes non équidistribuées . . . 2
1.1.3 Théorème de la Limite Centrale pour les tableaux de martingale . . . 3
1.1.4 Théorème de la Limite Centrale pour une fonctionnelle non linéaire des mesures empiriques . . . 4
1.1.5 Théorème de la Limite Centrale pour les chaînes de Markov . . . 8
1.1.6 Théorème de la Limite Centrale pour les systèmes de particules en interaction avec mécanisme d'échantillonnage . . . 10
1.2 Convergence des sommes de variables aléatoires modulo 1 vers la distribution uniforme 16
1.3 Contributions de la thèse . . . 18
1.3.1 Théorème de la Limite Centrale pour les fonctionnelles non linéaires des mesures empiriques : au-delà du cadre i.i.d. (voir Chapitre 2) . . . 18
1.3.2 Convergence vers la loi uniforme des vecteurs de sommes partielles modulo 1 avec un facteur commun (voir Chapitre 3) . . . 20
1.3.3 Théorème de la Limite Centrale pour le mécanisme d'échantillonnage stratifié (voir Chapitre 4) . . . 22

2 Central Limit Theorem over non-linear functionals of empirical measures: beyond the iid setting **27**
2.1 Main Results . . . 30
2.1.1 Independent Non-Equidistributed Random Variables . . . 30
The Wasserstein distance and the linear functional derivative . . . 30
Statement of the theorem . . . 32
2.1.2 Markov Chains . . . 35
Markov Chains and the Poisson equation . . . 35
Statement of the theorem . . . 38
2.2 Proof of the Results . . . 43
2.2.1 Markov Chains . . . 43
2.2.2 Independent Non-Equidistributed Random Variables . . . 55

3 Convergence to the uniform distribution of vectors of partial sums modulo one with a common factor **65**
3.1 Introduction . . . 65
3.2 Notation . . . 68
3.3 Main Result . . . 69
3.4 Proof of Proposition 5 . . . 73
3.5 Proof of Proposition 6 and Theorem 29 . . . 74
3.6 Appendix . . . 84

4 Central Limit Theorem for the stratified resampling mechanism **87**
4.1 Introduction . . . 87
4.2 Statement of the Main Result . . . 89
4.3 Asymptotic Variance . . . 91
4.4 Central Limit Theorem . . . 97
4.5 Proof of Proposition 8 . . . 101
4.6 Asymptotic Variance for the Next Steps . . . 106
4.7 Numerical Results . . . 116
4.7.1 Verification of the Conjectures . . . 116
Conjecture 1 . . . 116
Conjecture 2 . . . 117
4.7.2 Focus on the Variance . . . 118
case $n = 0$ . . . 118
case $n = 1$ . . . 120

Bibliography **123**

Chapter 1

Introduction

1.1 Théorème de la Limite Centrale : une revue de la littérature

Le Théorème de la Limite Centrale (abrégé TLC) est un résultat fondamental en théorie des probabilités et en statistiques. La première preuve rigoureuse du TLC a été fournie par Abraham de Moivre au XVIIIe siècle dans *Miscellanea analytica supplementum*. Sa preuve se limitait au cas particulier de variables aléatoires indépendantes identiquement distribuées selon une distribution de Bernoulli de paramètre $p = \frac{1}{2}$. Ce résultat a été étendu par Pierre Simon Laplace en 1812 aux variables aléatoires de Bernoulli de paramètre quelconque dans *Théorie analytique des probabilités*. De nombreux mathématiciens ont continué à développer les idées pionnières introduites par ces deux études. Dans la suite de cette section, nous preséntons une revue non exhaustive des résultats de type Théorème de la Limite Centrale.

1.1.1 Théorème de la Limite Centrale pour les variables aléatoires i.i.d.

Convergence en loi

Étant donnée une suite $(X_i)_{i\geq 1}$ de vecteurs aléatoires i.i.d. à valeurs dans $\mathbb{R}^d$ et une fonction mesurable $f : \mathbb{R}^d \to \mathbb{R}$ telle que $\mathbb{E}(|f(X_1)|) < \infty$, la Loi Forte des Grands Nombres affirme que la moyenne empirique converge presque sûrement vers l'espérance de $f(X_1)$ quand la taille de l'échantillon grandit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(X_i) = \mathbb{E}(f(X_1)) \quad p.s.$$

Le taux de convergence dans la Loi Forte des Grands Nombres est fourni par le Théorème de la Limite Centrale lorsque $\mathbb{E}\left(f^2(X_1)\right) < \infty$. Ainsi, on a

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}(f(X_1))\right) \overset{d}{\Longrightarrow} G \sim \mathcal{N}\left(0, \sigma^2\right) \tag{1.1}$$

où l'on a noté:

- $\overset{d}{\Longrightarrow}$ la convergence en loi,
- σ^2 la variance de $f(X_1)$,
- $\mathcal{N}\left(m, \sigma^2\right)$ la loi normale d'espérance m et de variance σ^2.

Ainsi, lorsque σ est strictement positif, (1.1) implique que l'erreur d'approximation $\varepsilon_N := \frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}(f(X_1))$ satisfait pour tous $c_1 < c_2$

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{c_1}{\sqrt{N}} \leq \varepsilon_N \leq \frac{c_2}{\sqrt{N}}\right) = \int_{c_1}^{c_2} e^{-\frac{x^2}{2\sigma^2}} \frac{dx}{\sigma\sqrt{2\pi}}.$$

En statistique, cela nous permet de définir un intervalle de confiance qui encadre la valeur de ϵ_N. Par exemple, si $\sigma^2 = 1$, puisque $\mathbb{P}(|G| \leq 1.96) \approx 0.95$, pour un N suffisamment grand et avec une probabilité proche de 0.95, on a

$$|\varepsilon_N| \leq \frac{1.96}{\sqrt{N}}.$$

Convergence pour la distance en variation totale

Soit $\mathcal{P}(\mathbb{R}^d)$, l'ensemble des mesures de probabilité sur $\mathbb{R}^d$. Rappelons que la distance en variation totale entre deux mesures $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$ est donnée par la formule

$$d_{TV}(\mu_1, \mu_2) = \sup_{A \in \mathcal{B}(\mathbb{R}^d)} |\mu_1(A) - \mu_2(A)| = \frac{1}{2} \sup_{\|f\|_{\infty,\mathbb{R}} \leq 1} \left| \int_{\mathbb{R}^d} f(x) (\mu_1(dx) - \mu_2(dx)) \right|.$$

Soit $(X_i)_{i \geq 1}$ une suite de vecteurs aléatoires i.i.d. à valeurs dans $\mathbb{R}^d$. Nous supposons que les X_i sont de carré intégrables, qu'ils ont une moyenne nulle et que leur matrice de covariance est la matrice identité. On dit que le TLC (Théorème de la Limite Centrale) a lieu en variation totale si

$$\lim_{N \to \infty} d_{TV}\left(\mu_{\frac{1}{\sqrt{N}} \sum_{k=1}^{N} X_k}, \mathcal{N}(0, I_{d \times d}) \right) = 0$$

où l'on a noté $I_{d \times d}$ la matrice identité de taille d et μ_Y la loi d'un vecteur aléatoire Y à valeurs dans $\mathbb{R}^d$.
En 1952, Prohorov [50] a été le premier à s'intéresser à ce problème : il a montré qu'en dimension $d = 1$ une condition nécessaire et suffisante pour obtenir le résultat est qu'il existe $N_0 \geq 1$ tel que la loi de $\sum_{k=1}^{N_0} X_k$ ait une composante absolument continue. En 1964, Mamatov et Halikov [42] ont étendu ce résultat en dimension quelconque. Plus précisément, ils ont prouvé que si la loi de X_1 a une composante absolument continue, alors

$$\lim_{N \to \infty} d_{TV}\left(\mu_{\frac{1}{\sqrt{N}} \sum_{k=1}^{N} X_k}, \mathcal{N}(0, I_{d \times d}) \right) = 0.$$

Voir également [3]. La généralisation suivante en découle facilement.

Theorem 1. *Soit $(T_i)_{i \geq 1}$ une suite de vecteurs aléatoires i.i.d. de carré intégrables à valeurs dans $\mathbb{R}^d$ telle que la loi de T_1 a une composante absolument continue. Notons Σ la matrice de covariance définie positive de T_1. Alors, la convergence en variation totale a lieu*

$$\lim_{N \to \infty} d_{TV}\left(\mu_{\frac{1}{\sqrt{N}} \sum_{k=1}^{N} (T_k - \mathbb{E}(T_1))}, \mu_G \right) = 0$$

où $G \sim \mathcal{N}(0, \Sigma)$.

1.1.2 Théorème de la Limite Centrale pour des variables aléatoires indépendantes non équidistribuées

Soit $(X_i)_{i \geq 1}$ une suite de vecteurs aléatoires indépendants à valeurs dans $\mathbb{R}^d$ et $f : \mathbb{R}^d \to \mathbb{R}$ une fonction mesurable. Notons la loi de X_i par ν_i.
Rappelons qu'une suite $(f(X_i))_{i \geq 1}$ est Fortement Résiduellement Cesaro β-Intégrable (Strongly Residually Cesaro β-Integrable) pour $\beta > 0$ (abrégé SRCI(β)) si

1. $\sup_{n\geq 1} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(|f(X_i)|) < \infty$

2. $\sum_{i=1}^{\infty} \frac{1}{i}\mathbb{E}\left(\left(|f(X_i)| - i^{\beta}\right) 1_{\{|f(X_i)|>i^{\beta}\}}\right) < \infty.$

Chandra et Goswami [13, Théorème 4.1] ont montré que si $(f(X_i))_{i\geq 1}$ est une suite de variables aléatoires indépendantes deux à deux et vérifie la condition SRCI(β) pour un certain $\beta \in (0,1)$, alors la Loi Forte des Grands Nombres est vérifiée :

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n}(f(X_i) - \mathbb{E}(f(X_i)))}{n} = 0 \quad p.s. \tag{1.2}$$

De plus, Lindeberg (voir par exemple [7]) a prouvé que le Théorème de la Limite Centrale est vérifié. Ainsi, si une suite $(f(X_i))_{i\geq 1}$ de variables aléatoires indépendantes de carré intégrable est telle que $\sigma_n^2 := \sum_{i=1}^{n} Var(f(X_i))$ est positif pour n grand et vérifie la condition de Lindeberg suivante

$$\forall \epsilon > 0 \quad \lim_{n\to\infty} \frac{1}{\sigma_n^2}\sum_{i=1}^{n} \mathbb{E}((f(X_i) - \mathbb{E}(f(X_i)))^2 1_{|f(X_i)-\mathbb{E}(f(X_i))|>\epsilon\sigma_n}) = 0,$$

alors

$$\frac{\sum_{i=1}^{n}(f(X_i) - \mathbb{E}(f(X_i)))}{\sigma_n} \xrightarrow{d} \mathcal{N}(0,1). \tag{1.3}$$

1.1.3 Théorème de la Limite Centrale pour les tableaux de martingale

Nous présentons maintenant le Théorème de la Limite Centrale pour des tableaux de martingale (voir [31], Corollaire 3.1).
Pour chaque $n \geq 1$, soit $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n\}$ une martingale à espérance nulle et de carré intégrable telle que $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$ pour $1 \leq i \leq n$, $n \geq 1$. Nous appelons $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n, n \geq 1\}$ un tableau de martingales.

Theorem 2. *Soit $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq n, n \geq 1\}$ un tableau de martingales à espérance nulle et de carré intégrable. Supposons que les variables $X_{n,i} = S_{n,i} - S_{n,i-1}$, $1 \leq i \leq n$ $(S_{n,0} = 0)$ satisfont les conditions suivantes:*

- *il existe $\eta^2 \in (0,\infty)$ tel que lorsque $n \to \infty$ la convergence en probabilité suivante a lieu*

$$\sum_{i=1}^{n} \mathbb{E}\left(X_{n,i}^2 | \mathcal{F}_{n,i-1}\right) \xrightarrow{P} \eta^2 \quad (\textit{condition du crochet})$$

- *pour tout $\epsilon > 0$ lorsque $n \to \infty$*

$$\sum_{i=1}^{n} \mathbb{E}\left(X_{n,i}^2 1_{\{|X_{n,i}|>\epsilon\}} | \mathcal{F}_{n,i-1}\right) \xrightarrow{P} 0 \quad (\textit{condition de Lindeberg}).$$

Alors,

$$\sum_{i=1}^{n} X_{n,i} \xrightarrow{d} \mathcal{N}\left(0, \eta^2\right).$$

1.1.4 Théorème de la Limite Centrale pour une fonctionnelle non linéaire des mesures empiriques

Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{R}^d$ i.i.d. selon μ. La mesure empirique associée est définie par la formule suivante

$$\mu_N := \frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}.$$

Dans cette section, nous nous intéressons à la convergence en loi de la suite de variables aléatoires réelles

$$\sqrt{N}\left(U\left(\mu_N\right) - U\left(\mu\right)\right)$$

où U est une fonction quelconque définie sur $\mathcal{P}\left(\mathbb{R}^d\right)$. Dans la littérature de statistique mathématique, les fonctionnelles $U\left(\mu_N\right)$ sont connues sous le nom de "fonctions statistiques". Par exemple, sous l'hypothèse que X_1 est de carré intégrable,

$$U(\mu_N) = \int_{\mathbb{R}^d}\left(x - \int_{\mathbb{R}^d} x\mu_N\left(dx\right)\right)^2 \mu_N\left(dx\right)$$

est un estimateur de la variance de la population

$$U(\mu) = \int_{\mathbb{R}^d}\left(x - \int_{\mathbb{R}^d} x\mu\left(dx\right)\right)^2 \mu\left(dx\right).$$

La théorie asymptotique des fonctions statistiques a été étudiée dans la littérature de statistique mathématique, et l'idée commune est de linéariser $\sqrt{N}\left(U\left(\mu_N\right) - U\left(\mu\right)\right)$ en une somme d'un terme auquel s'applique le Théorème de la Limite Centrale classique (voir la Section 1.1.1) et un terme de reste qui tend vers zéro en probabilité.

Von Mises [45] [46] a été le premier à étudier ce problème. En supposant que $(X_i)_{i\geq 1}$ est une famille de variables aléatoires réelles indépendantes et identiquement distribuées ($d = 1$), il a introduit un développement de Taylor de la fonctionnelle $U\left(\mu_N\right)$ à la distribution commune μ des variables aléatoires. Rappellons brièvement la méthod permettant de dériver une fonctionnelle $U\left(\cdot\right)$ définie sur $\mathcal{P}\left(\mathbb{R}\right)$: la différentielle de Gâteaux. Soit $\mu, \nu \in \mathcal{P}\left(\mathbb{R}\right)$ et soit U une fonctionnelle définie sur $\mu + \lambda\left(\nu - \mu\right)$ pour $0 \leq \lambda \leq 1$. Si la limite

$$d_1^G U\left(\mu; \nu - \mu\right) = \lim_{\lambda \to 0^+} \frac{U\left(\mu + \lambda\left(\nu - \mu\right)\right) - U\left(\mu\right)}{\lambda} \tag{1.4}$$

existe, on l'appelle la différentielle de Gâteaux de U en μ dans la direction de ν. Remarquez que $d_1^G U\left(\mu; \nu - \mu\right)$ est simplement la dérivée à droite en $\lambda = 0$ de $\lambda \mapsto U\left(\mu + \lambda\left(\nu - \mu\right)\right)$, $0 \leq \lambda \leq 1$. En général, la différentielle de Gâteaux d'ordre k de U en μ dans la direction de ν est définie, si la limite existe, comme

$$d_k^G U\left(\mu; \nu - \mu\right) = \frac{d^k}{d\lambda^k} U\left(\mu + \lambda\left(\nu - \mu\right)\right)\big|_{\lambda = 0^+}.$$

Dans ce qui suit, nous supposerons que la différentielle de Gâteaux d'ordre 1 est linéaire : il existe une fonction $f_G\left(\mu; x\right), x \in \mathbb{R}$ telle que $\int_{\mathbb{R}} f_G^2\left(\mu; x\right)\sigma\left(dx\right) < \infty$, $\forall \sigma \in \mathcal{P}\left(\mathbb{R}\right)$ telle que

$$d_1^G U\left(\mu; \nu - \mu\right) = \int_{\mathbb{R}} f_G\left(\mu; x\right)\left(\nu\left(dx\right) - \mu\left(dx\right)\right). \tag{1.5}$$

Ainsi, grâce à la différentielle de Gâteaux, Von Mises a introduit le développement de Taylor de $\lambda \mapsto U(\mu + \lambda(\mu_N - \mu))$, $0 \leq \lambda \leq 1$ en $\lambda = 0$ (sous les hypothèses pour lesquelles un développement de Taylor est valable, voir [2] p. 96). Si le premier terme non nul dans le développement de Taylor est le terme linéaire, alors on a

$$\sqrt{N}(U(\mu_N) - U(\mu)) = \sqrt{N} d_1^G U(\mu; \mu_N - \mu) + \sqrt{N} R_{1,N}$$

où $R_{1,N} = \frac{1}{2}\frac{d^2}{d\lambda^2} U(\mu + \lambda(\mu_N - \mu))\big|_{\lambda=\lambda^*}$ pour $0 \leq \lambda^* \leq 1$. La stratégie de Von Mises consistait à prouver que le reste tend vers zéro en probabilité en vérifiant que

$$\sqrt{N} \sup_{0\leq\lambda\leq 1} \left| \frac{d^2}{d\lambda^2} U(\mu + \lambda(\mu_N - \mu)) \right| \xrightarrow{P} 0$$

et en observant que, grâce à (1.5), $d_1^G U(\mu; \mu_N - \mu)$ peut être écrit comme une moyenne de variables aléatoires i.i.d. auxquelles s'applique le Théorème de la Limite Centrale classique.
En résumé, la principale difficulté technique était de prouver que le reste dans le développement de Taylor tend vers zéro. Un inconvénient de cette méthode est qu'elle implique un ordre de différentiation plus élevé que nécessaire. Nous renvoyons au Chapitre 6 de [56] pour plus de détails.

En dimension $d = 1$, Boos et Serfling [8] ont supposé l'existence d'une différentielle pour $U(.)$ en μ dans un sens plus fort que la différentielle de Gâteaux.
Soit $\mathcal{D} := \{\pi : \pi = c(\mu - \nu), c \in \mathbb{R}, \mu, \nu \in \mathcal{P}(\mathbb{R})\}$ équipé d'une norme $\|\cdot\|$. La fonctionnelle U définie sur $\mathcal{P}(\mathbb{R})$ est dite avoir une différentielle en μ par rapport à $\|\cdot\|$ si il existe une fonctionnelle $d_1^N U(\mu; \pi)$

- définie sur $\pi \in \mathcal{D}$ et linéaire en π, c'est-à-dire $d_1^N U(\mu; \sum_{i=1}^k a_i \pi_i) = \sum_{i=1}^k a_i d_1^N U(\mu; \pi_i)$ pour $a_1, \cdots, a_k \in \mathbb{R}$ et $\pi_1, \cdots, \pi_k \in \mathcal{D}$
- telle que

$$U(\mu) - U(\nu) - d_1^N U(\mu; \nu - \mu) = o(\|\nu - \mu\|) \tag{1.6}$$

quand $\|\nu - \mu\| \to 0$.

Remarquons que dans le cas des espaces de Banach, $d_1^N U(\mu; \pi)$ est appelée dérivée de Fréchet. Boos et Serfling supposent, comme c'est généralement le cas dans la plupart de la littérature et comme cela a été fait dans le cas de la différentielle de Gâteaux, que

$$d_1^N U(\mu; \pi) = \int_{\mathbb{R}} f_N(\mu; x) \pi(dx)$$

pour une certaine fonction $f_N(\mu; \cdot)$ telle que $\int_{\mathbb{R}} f_N^2(\mu; x) \sigma(dx) < \infty$, $\forall \sigma \in \mathcal{P}(\mathbb{R})$.
Si U admet une telle différentielle par rapport à la distance de Kolmogorov-Smirnov, on a

$$\sqrt{N}(U(\mu) - U(\mu_N)) = \frac{1}{\sqrt{N}} \sum_{i=1}^N d_1^N U(\mu; \delta_{X_i} - \mu) + \sqrt{N} o\left(\left\| \frac{1}{N} \sum_{i=1}^N 1_{\{X_i \leq \cdot\}} - \mu(-\infty, \cdot] \right\|_\infty \right).$$

La convergence en loi de $\frac{1}{\sqrt{N}} \sum_{i=1}^N d_1^N U(\mu; \delta_{X_i} - \mu)$ découle de l'application du Théorème de la Limite Centrale classique.
En ce qui concerne le reste, Boos et Serfling ont tiré parti des propriétés connues de la distance de Kolmogorov-Smirnov : $\sqrt{N} \left\| \frac{1}{N} \sum_{i=1}^N 1_{\{X_i \leq \cdot\}} - \mu(-\infty, \cdot] \right\|_\infty$ est stochastiquement bornée. Cela permet de prouver que le reste tend vers zéro en probabilité. Cette approche permet de contourner

un reste d'ordre supérieur, mais elle repose sur l'uniformité de l'approximation (1.6) par rapport à la distance de Kolmogorov-Smirnov qui est une hypothèse forte. De plus, il est difficile d'étendre cette approche à plusieurs dimensions.

En 1990, Dudley [21] a généralisé le travail de Boos et Serfling en considérant dans (1.6) une norme différente de celle de Kolmogorov-Smirnov et ne s'est pas limité au cas unidimensionnel. Étant donné $\mathcal{F}$ une famille de fonctions mesurables sur $\mathbb{R}^d$ telles que $\int_{\mathbb{R}^d} |f(x)| \sigma(dx) < \infty$, $\forall f \in \mathcal{F}, \sigma \in \mathcal{P}(\mathbb{R}^d)$, nous définissons pour une mesure signée τ

$$\|\tau\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} \left| \int_{\mathbb{R}^d} f(x) \tau(dx) \right|.$$

Remarquons que dans le cas unidimensionnel, en prenant $\mathcal{F}$ égal à l'ensemble des fonctions indicatrices $1_{(-\infty,t]}, t \in \mathbb{R}$, $\|\cdot\|_{\mathcal{F}}$ coincide avec la distance de Kolmogorov-Smirnov. Ainsi, Dudley a prouvé que si $\mathcal{F}$ est telle qu'un Théorème de la Limite Centrale pour les mesures empiriques a lieu avec une convergence uniforme sur $\mathcal{F}$ et que U est différentiable en μ par rapport à $\|\cdot\|_{\mathcal{F}}$, alors $\sqrt{N}(U(\mu) - U(\mu_N))$ converge en distribution vers une variable aléatoire gaussienne. Évidemment, il y a un équilibre entre la convergence uniforme sur $\mathcal{F}$ et la différentiabilité de U. En effet plus la classe $\mathcal{F}$ est grande, plus la norme $\|\cdot\|_{\mathcal{F}}$ est grossière et donc la différentiabilité par rapport à $\|\cdot\|_{\mathcal{F}}$ devient plus simple, tandis que la convergence uniforme sur $\mathcal{F}$ devient plus forte.

Récemment, Jourdain et Tse [36] ont réexaminé le même problème. L'outil principal est la dérivée fonctionnelle linéaire de U que nous allons maintenant brièvement présenter, ainsi que la notion de distance de Wasserstein.

La distance de Wasserstein et la dérivée fonctionnelle linéaire

Soit $U : \mathcal{P}_\ell(\mathbb{R}^d) \to \mathbb{R}$ où, pour $\ell \geq 0$, nous notons par $\mathcal{P}_\ell(\mathbb{R}^d)$ l'ensemble des mesures de probabilité m sur $\mathbb{R}^d$ telles que $\int_{\mathbb{R}^d} |x|^\ell m(dx) < \infty$. Pour $\ell > 0$, nous considérons la distance de Wasserstein d'ordre ℓ définie pour $\mu_1, \mu_2 \in \mathcal{P}_\ell(\mathbb{R}^d)$ par

$$W_\ell(\mu_1, \mu_2) = \inf \left\{ \left(\int_{\mathbb{R}^{2d}} |x-y|^\ell \rho(dx, dy) \right)^{\frac{1}{\ell \vee 1}} : \rho \in \mathcal{P}(\mathbb{R}^{2d}) \text{ avec } \rho\left(\cdot \times \mathbb{R}^d\right) = \mu_1(\cdot), \rho\left(\mathbb{R}^d \times \cdot\right) = \mu_2(\cdot) \right\}.$$

Si $\mu \in \mathcal{P}_\ell(\mathbb{R}^d)$ et $(\mu_n)_{n \in \mathbb{N}}$ est une suite dans cet espace, alors $\lim_{n\to\infty} W_\ell(\mu_n, \mu) = 0$ si et seulement si $\lim_{n\to\infty} \int_{\mathbb{R}^d} |x|^\ell \mu_n(dx) = \int_{\mathbb{R}^d} |x|^\ell \mu(dx)$ et μ_n converge faiblement vers μ lorsque $n \to \infty$, où nous écrivons $\mu_n \rightharpoonup \mu$ pour indiquer la convergence faible. Alternativement, $\lim_{n\to\infty} W_\ell(\mu_n, \mu) = 0$ si et seulement si, pour toute fonction $\phi : \mathbb{R}^d \to \mathbb{R}$ continue presque partout par rapport à μ et telle que $\sup_{x \in \mathbb{R}^d} \frac{|\phi(x)|}{1+|x|^\ell} < \infty$,

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \phi(x) \mu_n(dx) = \int_{\mathbb{R}^d} \phi(x) \mu(dx). \tag{1.7}$$

Pour $\ell = 0$ et $\mu_1, \mu_2 \in \mathcal{P}_0(\mathbb{R}^d)$, nous considérons

$$W_0(\mu_1, \mu_2) = \inf \left\{ \int_{\mathbb{R}^{2d}} (1 \wedge |x-y|) \rho(dx, dy) : \rho \in \mathcal{P}(\mathbb{R}^{2d}) \text{ avec } \rho\left(\cdot \times \mathbb{R}^d\right) = \mu_1(\cdot), \rho\left(\mathbb{R}^d \times \cdot\right) = \mu_2(\cdot) \right\}.$$

Remarquons que la topologie de la convergence faible est induite par W_0.

Rappelons maintenant la notion de dérivée fonctionnelle linéaire (d'ordre un) associée à U. Pour une description plus détaillée, y compris la définition de la dérivée fonctionnelle linéaire d'ordre supérieur, voir [11].

Definition 1. *Soit $\ell \geq 0$. Une fonction $U : \mathcal{P}_\ell\left(\mathbb{R}^d\right) \to \mathbb{R}$ admet une dérivée fonctionnelle linéaire en $\mu \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$ s'il existe une fonction mesurable $\mathbb{R}^d \ni y \mapsto \frac{\delta U}{\delta m}(\mu, y)$ telle que $\sup_{y\in\mathbb{R}^d} \frac{\left|\frac{\delta U}{\delta m}(\mu,y)\right|}{1+|y|^\ell} < \infty$ et*

$$\forall \nu \in \mathcal{P}_\ell\left(\mathbb{R}^d\right), \frac{d}{d\epsilon}_{|_{\epsilon=0^+}} U\left(\mu + \epsilon(\nu - \mu)\right) = \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu, y)\,(\nu - \mu)(dy).$$

Puisque $(\nu - \mu)\left(\mathbb{R}^d\right) = 0$, $\frac{\delta U}{\delta m}$ est défini à une constante additive près. Nous observons que la dérivée fonctionnelle linéaire est une dérivée de Gâteaux qui est linéaire par définition. Le lemme suivant permet d'exprimer une différence finie de la fonction U sous forme d'une intégrale de la dérivée fonctionnelle linéaire.

Lemma 1. *Soit $\ell \geq 0$, m, $m' \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$ et supposons que la dérivée fonctionnelle linéaire d'une fonction $U : \mathcal{P}_\ell\left(\mathbb{R}^d\right) \to \mathbb{R}$ existe dans le segment $(m_s := sm' + (1-s)m)_{s\in[0,1]}$.*
Alors si $\sup_{(s,y)\in[0,1]\times\mathbb{R}^d} \frac{\left|\frac{\delta U}{\delta m}(m_s,y)\right|}{1+|y|^\ell} < \infty$, on a

$$U(m') - U(m) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}((1-s)m + sm', y)(m' - m)(dy)ds.$$

Jourdain and Tse supposent que la dérivée fonctionnelle linéaire de U existe non seulement en μ, mais aussi sur une boule pour la distance de Wasserstein de rayon strictement positif contenant μ. Il s'agit d'une petite restriction, car lorsque un Théorème de la Limite Centrale s'applique à une fonction statistique, il n'est généralement pas limité à la seule valeur de la distribution commune μ des échantillons. Grâce au Lemme 1, ils linéarisent $\sqrt{N}\left(U\left(\mu_N\right) - U\left(\mu\right)\right)$ et obtiennent la somme du terme suivant

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1} \delta_{X_j} + \frac{N+1-i}{N}\mu, X_i \right) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1} \delta_{X_j} + \frac{N+1-i}{N}\mu, x \right) \mu(dx) \right) \tag{1.8}$$

et d'un reste. Une telle décomposition permet d'appliquer à la somme ci-dessus le Théorème de la Limite Centrale pour les tableaux de martingale (voir Section 1.1.3) et d'étudier les conditions suffisantes pour que le reste converge en probabilité vers zéro. En conclusion, le résultat principal de [36] est le suivant:

Theorem 3. *Soit $\ell \geq 0$ et soit $X_i, i \geq 1$ une suite de variables aléatoires i.i.d. à valeurs dans $\mathbb{R}^d$ de loi $\mu \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$. Supposons les hypothèses suivantes concernant la régularité de la dérivée fonctionnelle U dans un voisinage de μ:*

1. *il existe $r > 0$ tel que U admette une dérivée fonctionnelle linéaire sur la boule $B(\mu, r)$ centrée en μ avec un rayon r pour la métrique W_ℓ*
2. *il existe $C < \infty$ tel que $\forall(\tilde{\mu}, x) \in B(\mu, r) \times \mathbb{R}^d, \left|\frac{\delta U}{\delta m}(\tilde{\mu}, x)\right| \leq C\left(1 + |x|^{\frac{\ell}{2}}\right)$*
3. *$\sup_{x\in\mathbb{R}^d} \frac{|\frac{\delta U}{\delta m}(\tilde{\mu}, x) - \frac{\delta U}{\delta m}(\mu, x)|}{1 + |x|^{\frac{\ell}{2}}}$ converge vers 0 lorsque $W_\ell(\tilde{\mu}, \mu)$ tend vers 0*
4. *il existe $\alpha \in (\frac{1}{2}, 1]$ et $C < \infty$ tel que $\forall \mu_1, \mu_2 \in B(\mu, r), \forall x \in \mathbb{R}^d$*
   $$\left|\frac{\delta U}{\delta m}(\mu_2, x) - \frac{\delta U}{\delta m}(\mu_1, x)\right| \leq C\left(\left(1 + |x|^\ell\right)\left(d_{TV}\left(\mu_1, \mu_2\right)\right)^\alpha + \left(1 + |x|^{\ell(1-\alpha)}\right)\left(\int_{\mathbb{R}^d} |y|^\ell |\mu_2 - \mu_1|(dy) \right)^\alpha \right)$$

où $|\cdot|$ représente la variation totale d'une mesure signée. Alors, la convergence en loi suivante a lieu:

$$\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N} \delta_{X_i}\right) - U\left(\mu\right) \right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, Var\left(\frac{\delta U}{\delta m}\left(\mu, X_1\right)\right)\right).$$

1.1.5 Théorème de la Limite Centrale pour les chaînes de Markov

Considérons $(X_i)_{i\geq 1}$ une chaîne de Markov avec une distribution initiale ν_1 et un noyau de transition P sur $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$. Étant donnée une fonction mesurable $V : \mathbb{R}^d \to [1, \infty)$, nous pouvons introduire l'espace normé suivant

$$\mathcal{V}_V := \left\{ \phi : \mathbb{R}^d \to \mathbb{R} \text{ mesurable} : \sup_{x\in\mathbb{R}^d} \frac{|\phi(x)|}{V(x)} < \infty \right\}$$

équipé de la norme

$$\|\phi\|_V = \sup_{x\in\mathbb{R}^d} \frac{|\phi(x)|}{V(x)}.$$

Étant donné $\mu \in \mathcal{P}\left(\mathbb{R}^d\right)$ et $\phi : \mathbb{R}^d \to \mathbb{R}$ une fonction mesurable qui est soit positive, soit telle que $\int_{\mathbb{R}^d} |\phi(x)|\, \mu(dx) < \infty$, nous notons $\mu(\phi) = \int_{\mathbb{R}^d} \phi(x)\, \mu(dx)$.

Definition 2. *On dit qu'une chaîne de Markov $(X_i)_{i\geq 1}$ est V-uniformément ergodique s'il existe $\mu \in \mathcal{P}\left(\mathbb{R}^d\right)$ tel que*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}^d} \frac{\sup_{\phi:\|\phi\|_V\leq 1} |P^n\phi(x) - \mu(\phi)|}{V(x)} = 0.$$

Rappelons quelques définitions classiques de la théorie des chaînes de Markov. En particulier les notions de chaîne de Markov ψ-irréductible (voir Proposition 4.2.1 [43]), de chaîne de Markov positive et de chaîne de Markov Harris récurrente.

Definition 3. *On dit qu'une chaîne de Markov $(X_i)_{i\geq 1}$ est ψ-irréductible s'il existe une mesure ψ sur $\mathcal{B}\left(\mathbb{R}^d\right)$ telle que pour tout $x \in \mathbb{R}^d$, chaque fois que $\psi(A) > 0$, il existe un certain $n > 0$, éventuellement dépendant à la fois de A et de x, tel que $P^n(x, A) > 0$.*

Definition 4. *Soit $(X_i)_{i\geq 1}$ une chaîne de Markov ψ-irréductible. Alors*

- *$(X_i)_{i\geq 1}$ est appelée une chaîne de Markov positive si elle admet une mesure invariante.*
- *$(X_i)_{i\geq 1}$ est appelée une chaîne de Markov Harris récurrente si chaque ensemble $A \in \mathcal{B}\left(\mathbb{R}^d\right)$ tel que $\psi(A) > 0$ satisfait*

$$\mathbb{P}_{\delta_x}\left(\bigcap_{N=1}^{\infty}\bigcup_{k=N}^{\infty}\{X_k \in A\}\right) = 1, \quad x \in A.$$

La proposition suivante montre qu'en cas de V-uniforme ergodicité, la chaîne de Markov possède plusieurs propriétés : en particulier, elle est positive et Harris récurrente.

Proposition 1. *Supposons que la chaîne de Markov $(X_i)_{i\geq 1}$ est V-uniformément ergodique avec une mesure de probabilité $\mu \in \mathcal{P}\left(\mathbb{R}^d\right)$. Alors $(X_i)_{i\geq 1}$ est une chaîne de Markov positive Harris récurrente avec une mesure de probabilité invariante unique μ et*

$$\mu(V) < \infty.$$

De plus,

- *il existe $r \in (0, 1)$, $R < \infty$ tels que pour tout $n \in \mathbb{N}$*

$$\sup_{x\in\mathbb{R}^d} \frac{\sup_{\phi:\|\phi\|_V\leq 1} |P^n\phi(x) - \mu(\phi)|}{V(x)} \leq Rr^n \tag{1.9}$$

- *il existe* $\chi \in (0,1)$, $D < \infty$ *tels que pour tout* $n \in \mathbb{N}$

$$\sup_{x \in \mathbb{R}^d} \frac{\sup_{\phi:\|\phi\|_{\sqrt{V}} \leq 1} |P^n \phi(x) - \mu(\phi)|}{\sqrt{V(x)}} \leq D\chi^n. \tag{1.10}$$

Nous sommes maintenant prêts à rappeler la Loi Forte des Grands Nombres et le Théorème de la Limite Centrale pour les chaînes de Markov sous la V-uniforme ergodicité. Commençons par la Loi Forte des Grands Nombres (une démonstration est fournie dans le Théorème 17.1.7 [43] sous l'hypothèse que la chaîne de Markov est positive Harris récurrente).

Theorem 4. *Supposons que la chaîne de Markov* $(X_i)_{i \geq 1}$ *est* V*-uniformément ergodique, et soit* μ *sa mesure de probabilité invariante. Alors, pour chaque fonction* $f : \mathbb{R}^d \to \mathbb{R}$ *mesurable telle que* $\mu(|f|) < \infty$,

$$\forall \nu_1 \in \mathcal{P}(\mathbb{R}^d) \quad \mathbb{P}_{\nu_1}\left(\frac{1}{N}\sum_{k=1}^{N} f(X_k) \underset{N\to\infty}{\to} \mu(f)\right) = 1.$$

Puisque, selon la Proposition 1, $\mu(V) < \infty$, alors

$$\sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{V(x)} < \infty$$

est une condition suffisante pour garantir que $\mu(|f|) < \infty$.
Avant de formuler le Théorème de la Limite Centrale pour les chaînes de Markov, nous introduisons l'équation de Poisson. Soit f une fonction telle que $\mu(|f|) < \infty$ où μ, comme précédemment, désigne la mesure de probabilité invariante associée à P. Une fonction $F : \mathbb{R}^d \to \mathbb{R}$ mesurable telle que, pour chaque $x \in \mathbb{R}^d$ et chaque $n \geq 0$, $P|F|(x) < \infty$ est appelée solution de l'équation de Poisson si elle satisfait

$$\forall x \in \mathbb{R}^d, \quad F(x) - PF(x) = f(x) - \mu(f). \tag{1.11}$$

Si la chaîne de Markov $(X_i)_{i \geq 1}$ est V-uniformément ergodique, alors, grâce à (1.9), la série de terme général $(P^n f - \mu(f))_{n \in \mathbb{N}}$ converge dans l'espace $\mathcal{V}_V$ équipé de la norme $\|\|_V$ et

$$F = \sum_{n \in \mathbb{N}} (P^n f - \mu(f))$$

est une solution de l'équation de Poisson (1.11). De plus, toute solution peut être écrite comme $c + \sum_{n \in \mathbb{N}} (P^n f - \mu(f))$ pour une constante $c \in \mathbb{R}$.
En conséquence de (1.10), la proposition suivante est vérifiée.

Proposition 2. *Supposons que la chaîne de Markov* $(X_i)_{i \geq 1}$ *est* V*-uniformément ergodique, et soit* μ *sa mesure de probabilité invariante. Alors si* f *est tel que* $f^2 \in \mathcal{V}_V$, $F = \sum_{n \in \mathbb{N}} (P^n f - \mu(f))$ *converge dans* $\|\cdot\|_{\sqrt{V}}$. F *est une solution de l'équation de Poisson (1.11) et satisfait* $F^2 \in \mathcal{V}_V$. *De plus, pour chaque* $n \in \mathbb{N}$

$$\sup_{x \in \mathbb{R}^d} \frac{|P^n f(x) - \mu(f)|}{\sqrt{V(x)}} \leq D\chi^n \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{\sqrt{V(x)}}$$

avec $\chi \in (0,1)$ *et* D *une constante finie.*

Nous sommes maintenant en mesure d'énoncer le Théorème de la Limite Centrale pour les chaînes de Markov sous la V-uniforme ergodicité (voir Théorème 17.5.4 [43]).

Theorem 5. *Supposons que la chaîne de Markov $(X_i)_{i\geq 1}$ est V-uniformément ergodique, et soit μ sa mesure de probabilité invariante. Alors, pour chaque fonction $f : \mathbb{R}^d \to \mathbb{R}$ mesurable telle que $f^2 \in \mathcal{V}_V$,*

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} f\left(X_i\right) - \int_{\mathbb{R}^d} f(x)\mu(dx)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \mu\left(F^2\right) - \mu((PF)^2)\right)$$

où F est une solution de l'équation de Poisson

$$F(x) - PF(x) = f(x) - \mu(f) \quad x \in \mathbb{R}^d.$$

1.1.6 Théorème de la Limite Centrale pour les systèmes de particules en interaction avec mécanisme d'échantillonnage

Les filtres particulaires, également connu sous le nom de méthodes de Monte Carlo séquentielles (voir le Chapitre 4 de [10] et [15] pour une introduction générale), sont des techniques puissantes pour estimer l'état évolutif d'un système au fil du temps, même lorsque l'état ne peut pas être directement observé mais peut seulement être déduit à partir de mesures ou observations bruitées. Ces méthodes numériques sont devenues très populaires pour la résolution de problèmes d'estimation optimale dans des scénarios non linéaires et non gaussiens. Elle sont utilisées dans des applications en temps réel dans des domaines tels que le génie chimique, la vision par ordinateur, l'économétrie financière, la poursuite de cibles, la robotique et les statistiques (voir entre autres [51] et [20]). L'utilisation des méthodes de Monte Carlo pour les problèmes de filtrage non linéaires remonte à Handschin [32] et Mayne [33]. Ils ont introduit une version séquentielle de la méthode d'échantillonnage préférentiel (voir [22], [52] comme références pour l'échantillonnage préférentiel) et l'algorithme correspondant est connu sous le nom d'échantillonnage préférentiel séquentiel, souvent abrégé SIS en anglais. Un inconvénient de ce dernier a été identifié par Gordon et al. dans [29] : à mesure que le nombre d'itérations augmente, les poids ont tendance à dégénérer (un phénomène généralement connu sous le nom de « dégénérescence des poids »). Cela signifie qu'après un certain nombre d'itérations, certains poids ont tendance à devenir très petits, de sorte que les positions correspondantes ne contribuent plus à l'estimation. Ainsi, Gordon et al. ont introduit l'étape d'rééchantillonnage où, dans le but de stabiliser l'erreur de Monte Carlo au fil du temps, l'idée clé est d'éliminer les particules ayant des poids faibles et de répliquer les particules ayant des poids élevés. Par conséquent, en propageant M particules à travers des étapes de pondération, d'échantillonnage et de mutation (chaque particule évolue de manière aléatoire selon un noyau de transition donné), les systèmes de particules en interaction peuvent être utilisés pour estimer numériquement l'état du système en fonction des observations (voir également [20], [15]).

Nous présentons maintenant une version du problème de filtrage non linéaire et l'algorithme correspondant des systèmes de particules en interaction qui alternent les sélections selon le mécanisme d'échantillonnage et les mutations selon les noyaux de Markov (voir [20]). Le signal d'état $(Z_n)_{n\geq 0}$ est une chaîne de Markov à valeurs dans $\mathbb{R}^d$. Pour toute fonction mesurable bornée $h : \mathbb{R}^d \to \mathbb{R}$, $\mathbb{E}\left(h\left(Z_n\right)|\sigma\left(Z_0, \cdots, Z_{n-1}\right)\right) = \int_{\mathbb{R}^d} h(x)P_n\left(Z_{n-1}, dx\right)$ pour un noyau de transition P_n pouvant dépendre de n. Nous notons la loi de Z_0 par η. La suite d'observations $(T_n)_{n\geq 0}$ prend la forme

$$T_n = h\left(Z_n, W_n\right), \quad n \geq 1$$

pour une fonction mesurable h où $(W_n)_{n\geq 0}$ est une suite de vecteurs aléatoires à valeurs dans $\mathbb{R}^d$ i.i.d indépendants de $(Z_n)_{n\geq 0}$. De plus, supposons que pour tout $x \in \mathbb{R}^d$ et pour tout $n \geq 0$, $h\left(x, W_n\right)$ admet une densité strictement positive et bornée $y \mapsto \bar{g}_n\left(x, y\right)$ par rapport à la mesure

de Lebesgue sur $\mathbb{R}^d$. Le problème de filtrage non linéaire consiste à calculer pour chaque $n \geq 0$ et chaque fonction mesurable f

$$\eta_{n,T}(f) = \mathbb{E}\left(f(Z_n) \mid \sigma\left((T_i)_{1\leq i\leq n-1}\right)\right) \tag{1.12}$$

avec la convention que $\eta_{0,T}$ est la loi de Z_0. Les observations sont considérées comme fixes, c'est-à-dire $T_n = t_n$, $n \in \mathbb{N}$, et pour simplifier la notation, nous écrivons η_n au lieu de $\eta_{n,T}$. Soit $M \geq 1$. Le but est de concevoir une approche de système de particules en interaction pour le calcul numérique de la distribution η_n, c'est-à-dire d'approximer η_n par une suite de mesures empiriques

$$\eta_n^M = \frac{1}{M}\sum_{i=1}^{M} \delta_{X_n^{M,i}} \tag{1.13}$$

associées aux vecteurs aléatoires $\left(X_n^{M,1}, \cdots, X_n^{M,M}\right)$ à valeurs dans $\left(\mathbb{R}^d\right)^M$. Pour ce faire, l'idée principale est de dériver une formulation récursive pour η_n de la forme

$$\eta_n = \phi_n(\eta_{n-1}), \quad n \geq 1 \tag{1.14}$$

pour des applications continues $\phi_n : \mathcal{P}\left(\mathbb{R}^d\right) \to \mathcal{P}\left(\mathbb{R}^d\right)$ où $\mathcal{P}\left(\mathbb{R}^d\right)$ définit l'ensemble de toutes les mesures de probabilité sur $\mathbb{R}^d$. L'équation (1.14) permettra de définir la suite de vecteurs aléatoires $\left(X_n^{M,1}, \cdots, X_n^{M,M}\right)_{n\geq 0}$ prenant des valeurs dans $\left(\mathbb{R}^d\right)^M$ comme une chaîne de Markov avec une distribution initiale $\hat{\eta}_0(dx_1, ..., dx_p) = \prod_{p=1}^{M} \eta(dx_p)$ et des probabilités de transition $\tilde{P}_n$ données par $\tilde{P}_n(z, dx) = \prod_{p=1}^{M} \phi_n(m(z))(dx_p)$ avec $m(z) = \frac{1}{M}\sum_{i=1}^{M} \delta_{z_i}$. L'idée derrière cela est que si η_{n-1}^M est une bonne approximation de η_{n-1}, alors la mesure empirique associée à M variables aléatoires conditionnellement indépendantes distribuées selon $\phi_n\left(\eta_{n-1}^M\right)$, conformément à (1.14), devrait être une bonne approximation de η_{n-1}. En rappelant que les observations $T_n = t_n$ sont fixes, nous pouvons définir

$$g_n(x) = \bar{g}_n(x, t_n)$$

et il est possible de démontrer (voir Section 3.2 dans [20]) que la suite des applications $(\phi_n)_{n\geq 1}$ dans (1.14) est donnée par

$$\psi_n(\eta)(f) = \frac{\eta(fg_n)}{\eta(g_n)}, \quad n \geq 0$$
$$\phi_n(\eta)(f) = \psi_{n-1}(\eta) P_n f, \quad n \geq 1.$$

Selon l'expression ci-dessus, chaque étape de p à $p+1$ de l'algorithme des filtres particulaires peut être divisée en deux parties : une partie de sélection et une partie de mutation

$$\left(X_p^{M,i}\right)_{1\leq i\leq M} \overset{sélection}{\longrightarrow} \left(Y_{p+1}^{M,i}\right)_{1\leq i\leq M} \overset{mutation}{\longrightarrow} \left(X_{p+1}^{M,i}\right)_{1\leq i\leq M}$$

que nous allons maintenant décrire. Tout d'abord, les variables $\left(X_0^{M,i}\right)_{1\leq i\leq M}$ sont tirées indépendamment selon la loi initiale η. Par convention, $Y_0^{M,i} = X_0^{M,i}$ pour $i = 1, \cdots, M$. Ensuite, par induction sur $0 \leq p \leq n-1$, on a

Sélection : Après avoir calculé les poids comme

$$w_p^{M,j} = g_p\left(X_p^{M,j}\right) / \sum_{i=1}^{M} g_p\left(X_p^{M,i}\right),$$

effectuons un échantillonnage pour obtenir les nouvelles variables aléatoires $\left(Y_{p+1}^{M,i}\right)_{1\leq i\leq M}$ conditionnellement indépendantes étant donné $\sigma\left(\left(X_s^{M,i},Y_s^{M,i}\right)_{1\leq i\leq M,0\leq s\leq p}\right)$ et vérifiant la propriété de sélection

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_{p+1}^{M,m}}\;\middle|\;\sigma\left(\left(X_s^{M,i},Y_s^{M,i}\right)_{1\leq i\leq M,0\leq s\leq p}\right)\right)=\sum_{m=1}^{M}w_p^{M,m}\delta_{X_p^{M,m}}. \tag{1.15}$$

Après avoir décrit la deuxième étape, nous revisiterons plusieurs approches d'échantillonnage vérifiant (1.15).

Mutation : étant donné $\sigma\left(\left(X_s^{M,i}\right)_{1\leq i\leq M,0\leq s\leq p},\left(Y_s^{M,i}\right)_{1\leq i\leq M,0\leq s\leq p+1}\right)$, les variables aléatoires $X_{p+1}^{M,i},1\leq i\leq M$ sont générées conditionnellement indépendantes et distribuées selon la mesure de probabilité $P_{p+1}\left(Y_{p+1}^{M,i},\cdot\right)$.

La mesure approximée est définie comme dans (1.13). Del Moral (voir Corollaire 7.4.2 dans [47]) a prouvé qu'une Loi Forte des Grands Nombres est vérifiée.

Theorem 6. *Soit $g_n:\mathbb{R}^d\to(0,\infty)$, $n\geq 0$ une famille de fonctions mesurables telles que $0<\inf_{x\in\mathbb{R}^d}g_n(x)\leq\sup_{x\in\mathbb{R}^d}g_n(x)<\infty$ et soit f une fonction mesurable bornée. Alors, pour tout temps $n\geq 0$,*

$$\lim_{M\to\infty}\eta_n^M(f)=\eta_n(f)\quad p.s.$$

Concentrons-nous désormais sur l'étape de sélection en introduisant différentes méthodes d'échantillonnage vérifiant la propriété de sélection (1.15). Étant donné une suite de vecteurs aléatoires $(X_m)_{1\leq m\leq M}$ à valeurs dans $\mathbb{R}^d$ avec des poids aléatoires associés $(w_m^M)_{1\leq m\leq M}$ tels que $w_m^M\geq 0$ et $\sum_{m=1}^{M}w_m^M=1$, une méthode d'échantillonnage définit la suite rééchantillonnée $(Y_m^M)_{1\leq m\leq M}$ de telle manière que

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m^M}\;\middle|\;\mathcal{G}\right)=\sum_{m=1}^{M}w_m^M\delta_{X_m} \tag{1.16}$$

où $\mathcal{G}:=\sigma\left((X_m,w_m^M)_{1\leq m\leq M}\right)$. Les méthodes d'échantillonnage [19] permettent de remplacer la mesure de probabilité $\sum_{m=1}^{M}w_m^M\delta_{X_m}$ avec des poids non égaux par une mesure empirique $\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m^M}$ ayant la même espérance conditionnelle donnée $\mathcal{G}$. Selon la définition utilisée pour les variables aléatoires Y_m^M, plusieurs méthodes d'échantillonnage peuvent être considérées.
Les méthodes d'échantillonnage les plus courantes sont les suivantes : multinomiale, résiduelle, stratifiée [37] et systématique [12]. Voir [19] pour une brève description de ces méthodes. L'approche la plus simple est l'échantillonnage multinomial. Introduisons le ainsi que l'échantillonnage résiduel.

échantillonnage multinomial

Il consiste à tirer, conditionnellement à $\mathcal{G}$, les nouvelles positions $\left(Y_m^M\right)_{1\leq m\leq M}$ indépendamment suivant la probabilité $\sum_{j=1}^{M}w_j^M\delta_{X_j}$.

échantillonnage résiduel

Le schéma d'échantillonnage résiduel consiste à choisir

$$Y_m^M = X_\ell$$

pour $\sum_{j=1}^{\ell-1} \lfloor M w_j^M \rfloor < m \leq \sum_{j=1}^{\ell} \lfloor M w_j^M \rfloor$ (nous répliquons $\lfloor M w_\ell^M \rfloor$ fois la valeur X_ℓ) et, si $\sum_{j=1}^{M} \lfloor M w_j^M \rfloor < M$, à tirer, conditionnellement à $\mathcal{G}$, $r = M - \sum_{j=1}^{M} \lfloor M w_j^M \rfloor = \sum_{j=1}^{M} \left\{ M w_j^M \right\}$ variables aléatoires $Y^M_{\sum_{j=1}^{M} \lfloor M w_j^M \rfloor + 1}, \cdots, Y_M^M$ indépendamment suivant

$$\frac{1}{M - \sum_{j=1}^{M} \lfloor M w_j^M \rfloor} \sum_{\ell=1}^{M} \left\{ M w_\ell^M \right\} \delta_{X_\ell}.$$

échantillonnage stratifié

Le comportement du schéma d'échantillonnage stratifié n'est pas immédiat à comprendre en raison de la structure compliquée dans la définition des Y_m^M, qui, cependant, restent conditionnellement indépendantes étant donné $\mathcal{G}$. À partir de M variables aléatoires $(U_m)_{1 \leq m \leq M}$ distribuées de manière i.i.d. selon la loi uniforme sur $(0, 1)$ et indépendamment de $\mathcal{G}$, la suite $(Y_m^M)_{1 \leq m \leq M}$ est définie comme suit

$$Y_m^M = \sum_{\ell=1}^{M} 1_{\left\{ \sum_{j=1}^{\ell-1} w_j^M < \frac{m - U_m}{M} \leq \sum_{j=1}^{\ell} w_j^M \right\}} X_\ell \text{ pour } m \in \{1, \cdots, M\}. \tag{1.17}$$

échantillonnage systématique

Le comportement du schéma d'échantillonnage systématique est encore plus compliqué à comprendre car les Y_m^M ne sont plus conditionnellement indépendantes étant donné $\mathcal{G}$. Dans la formule précédente, nous prenons les variables U_m toutes égales à une seule variable U distribuée de manière uniforme sur $(0, 1)$.

Les schémas d'échantillonnage ont été largement étudiés dans la littérature. Nous présentons maintenant une sélection des résultats obtenus. Le comportement asymptotique du schéma d'échantillonnage multinomial a été largement étudié dans [47]. Nous résumerons le résultat à la fin de cette section. Douc et al. dans [19] ont montré que l'échantillonnage résiduel et l'échantillonnage stratifié améliorent l'échantillonnage multinomial dans le sens où ils ont une variance conditionnelle plus faible par rapport à $\mathcal{G}$. Ils ont également prouvé, au moyen d'un contre-exemple, que cette propriété ne s'applique pas au échantillonnage systématique. De plus, ils ont établi un Théorème de la Limite Centrale pour le schéma d'échantillonnage résiduel (ce qui est également présenté à la fin de cette section), suggérant qu'un résultat similaire devrait être obtenu pour le schéma d' échantillonnage stratifié. L'une des dernières contributions concernant les schémas d'échantillonnage est donnée par Gerber et al. [28]. En utilisant la notion d'*association négative* pour les variables aléatoires [35], ils ont d'abord fourni un résultat général de convergence pour les schémas d'échantillonnage. Une application de ce théorème donne la convergence faible presque sûre de $\frac{1}{M} \sum_{m=1}^{M} \delta_{Y_m^M}$ dans la méthode d'échantillonnage stratifié. De plus, ils ont fourni un contre-exemple de convergence faible presque sûre pour la méthode d'échantillonnage systématique. Plus récemment, Chopin et al. [16] ont

étudié les schémas d'échantillonnage pour les filtres particulaires avec des observations faiblement informatives. Les preuves empiriques indiquent que lorsque les poids utilisés dans l'échantillonnage présentent une grande variabilité, le choix de la stratégie d'échantillonnage tend à avoir un impact faible. Cependant, dans les cas où les poids sont proches d'être uniformes, les différences de performance entre les différentes méthodes d'échantillonnage peuvent être importantes. En gardant M fixe, ils ont examiné le comportement asymptotique des schémas d'échantillonnage à mesure que les poids deviennent de moins en moins informatifs. Une telle situation se présente naturellement en considérant des discrétisations du modèle de Feynman-Kac en temps continu. Ils ont introduit une condition générale pour les schémas d'échantillonnage garantissant qu'une limite en temps continu a lieu et ils ont introduit des versions modifiées du échantillonnage stratifié et systématique qui satisfont à cette condition. Voir également [41], [23], [38], [14] pour des références supplémentaires.

Théorème de la Limite Centrale pour le mécanisme de sélection multinomial

Del Moral [47] a étudié le comportement asymptotique des systèmes de particules en interaction lorsque la méthode d'échantillonnage multinomial est utilisée. Rappelons que la mesure η_n^M est définie dans (1.13). Il a fourni un Théorème de la Limite Centrale pour la différence entre la mesure empirique des particules η_n^M et la mesure limite correspondante η_n (voir Section 9.4.2 [47]). La preuve repose sur des techniques de martingale et de semi-groupe.

Theorem 7. *Soit $g_n : \mathbb{R}^d \to (0, \infty)$, $n \geq 0$, une famille de fonctions mesurables telle que $0 < \inf_{x \in \mathbb{R}^d} g_n(x) \leq \sup_{x \in \mathbb{R}^d} g_n(x) < \infty$, et soit $f : \mathbb{R}^d \to \mathbb{R}$ une fonction bornée mesurable. Alors, il existe $\mu \in \mathbb{R}$ et $\sigma_n^2 > 0$ tels que lorsque $M \to \infty$, la convergence en loi suivante a lieu*

$$\sqrt{M}\left(\eta_n^M(f) - \eta_n(f)\right) \Longrightarrow^{d} \mathcal{N}\left(\mu, \sigma_n^2\right).$$

Théorème de la Limite Centrale pour le mécanisme de sélection résiduelle

Douc et al. [19] ont étudié le comportement asymptotique des systèmes de particules en interaction lorsque la méthode d'échantillonnage résiduelle est utilisée. Ils ont tout d'abord adapté deux résultats donnés au Chapitre 9 [10] sous des hypothèses plus générales. On a les résultats suivants.

Theorem 8. *Soit f une fonction bornée mesurable. Supposons que $\left(X_n^{M,i}\right)_{1 \leq i \leq M}$ satisfait les hypothèses suivantes :*

A. lorsque $M \to \infty$, il existe une densité de probabilité ν_n telle que

$$\frac{1}{M}\sum_{i=1}^{M} f\left(X_n^{M,i}\right) \xrightarrow{P} \int_{\mathbb{R}^d} f(x)\nu_n(x)dx$$

et, en notant $\sigma^2(f)$ la variance asymptotique,

$$\sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M} f\left(X_n^{M,i}\right) - \int_{\mathbb{R}^d} f(x)\nu_n(x)dx\right) \Longrightarrow^{d} \mathcal{N}\left(0, \sigma^2(f)\right).$$

B. Les poids sont donnés par $w_n^{M,i} = \frac{g_n\left(X_n^{M,i}\right)}{\sum_{i=1}^{M} g_n\left(X_n^{M,i}\right)}$, où $g_n(x) = \frac{\mu_n(x)}{\nu_n(x)}$ pour une densité de probabilité μ_n. De plus, g_n est bornée par le dessus.

Si les $Y_{n+1}^{M,i}$ *sont conditionnellement indépendants étant donné* $\mathcal{G}^n := \sigma\left(\left(Y_s^{M,i}, X_s^{M,i}\right)_{1\leq i\leq M, 0\leq s\leq n}\right)$ *et s'il existe un nombre réel* $\kappa(f) < \infty$ *tel que*

$$Var\left(\frac{1}{\sqrt{M}}\sum_{i=1}^{M} f\left(Y_{n+1}^{M,i}\right)\middle|\mathcal{G}^n\right) \xrightarrow{P} \kappa(f), \quad M\to\infty, \tag{1.18}$$

alors $\left(Y_{n+1}^{M,i}\right)_{1\leq i\leq M}$ *satisfait*

$$\frac{1}{M}\sum_{i=1}^{M} f\left(Y_{n+1}^{M,i}\right) \xrightarrow{P} \int_{\mathbb{R}^d} f(x)\mu_n(x)dx$$

et

$$\sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M} f\left(Y_{n+1}^{M,i}\right) - \int_{\mathbb{R}^d} f(x)\mu_n(x)dx\right) \xRightarrow{d} \mathcal{N}\left(0, \tilde{\sigma}^2(f)\right)$$

avec variance asymptotique

$$\tilde{\sigma}^2(f) = \kappa(f) + \sigma^2\left(\frac{\mu_n}{\nu_n}\left(f - \int_{\mathbb{R}^d} f(x)\mu_n(x)dx\right)\right).$$

Theorem 9. *Soit* f *une fonction bornée mesurable. Supposons l'existence d'une densité de probabilité* μ_n *telle que*

$$\frac{1}{M}\sum_{i=1}^{M} f\left(Y_{n+1}^{M,i}\right) \xrightarrow{P} \int_{\mathbb{R}^d} f(x)\mu_n(x)dx$$

et, en notant $\tilde{\sigma}^2(f)$ *la variance asymptotique,*

$$\sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M} f\left(Y_{n+1}^{M,i}\right) - \int_{\mathbb{R}^d} f(x)\mu_n(x)dx\right) \xRightarrow{d} \mathcal{N}\left(0, \tilde{\sigma}^2(f)\right).$$

Alors, $\left(X_{n+1}^{M,i}\right)_{1\leq i\leq M}$ *satisfait*

$$\frac{1}{M}\sum_{i=1}^{M} f\left(X_{n+1}^{M,i}\right) \xrightarrow{P} \int_{\mathbb{R}^d} f(x)\int_{\mathbb{R}^d} P_{n+1}(y,dx)\,\mu_n(y)dy$$

et

$$\sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M} f\left(X_{n+1}^{M,i}\right) - \int_{\mathbb{R}^d} f(x)\int_{\mathbb{R}^d} P_{n+1}(y,dx)\,\mu_n(y)dy\right) \xRightarrow{d} \mathcal{N}\left(0, \bar{\sigma}^2(f)\right)$$

avec

$$\bar{\sigma}^2(f) = \int_{\mathbb{R}^d}\left(P_{n+1}f^2(x) - \left(P_{n+1}f\right)^2(x)\right)\mu_n(x)dx + \tilde{\sigma}^2\left(P_{n+1}f\right).$$

Ainsi, la combinaison du Théorème 8 et du Théorème 9 fournit un Théorème de la Limite Centrale pour les filtres particulaires à chaque itération.

Remark 10. *Remarquons que pour l'échantillonnage multinomial, il est facile de vérifier (1.18). En effet, selon la définition du schéma d'échantillonnage multinomial donnée ci-dessus, on a*

$$Var\left(\frac{1}{\sqrt{M}}\sum_{i=1}^{M} f\left(Y_{n+1}^{M,i}\right)\middle|\mathcal{G}^n\right) = \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\left(f^2\left(Y_{n+1}^{M,i}\right)\middle|\mathcal{G}^n\right) - \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\left(f\left(Y_{n+1}^{M,i}\right)\middle|\mathcal{G}^n\right)^2$$
$$= \sum_{i=1}^{M} w_n^{M,i} f^2\left(X_n^{M,i}\right) - \left(\sum_{i=1}^{M} w_n^{M,i} f\left(X_n^{M,i}\right)\right)^2.$$

Ainsi, en supposant que $0 < \inf_{x\in\mathbb{R}^d} g_n(x)$*, le côté droit converge presque sûrement vers* $\frac{\eta_n\left(f^2 g_n\right)}{\eta_n(g_n)} - \left(\frac{\eta_n(f g_n)}{\eta_n(g_n)}\right)^2$ *d'après le Théorème 6.*

Dans le Corollaire 5 [19], Douc et al. ont montré que l'échantillonnage résiduel vérifie également (1.18) et donc, sous les hypothèses A et B, il satisfait un Théorème de la Limite Centrale selon le Théorème 8.

À notre connaissance, le Théorème de la Limite Centrale n'a été démontré que pour l'échantillonnage multinomial et l'échantillonnage résiduel.

1.2 Convergence des sommes de variables aléatoires modulo 1 vers la distribution uniforme

Nous notons par $\lfloor x\rfloor$ l'entier j tel que $j \le x < j+1$ et par $\{x\} = x - \lfloor x\rfloor$ la partie fractionnaire de $x \in \mathbb{R}$. Soit $(Y_i)_{i\ge1}$ une suite de variables aléatoires réelles. Le problème de la convergence en loi de la suite

$$(V_M)_{M\ge1} = (\{Y_1 + \cdots + Y_M\})_{M\ge1}$$

de sommes de variables aléatoires définies modulo 1 vers la distribution uniforme sur $[0,1]$ a été étudié par de nombreux chercheurs. Dans son ouvrage Feller [24] a montré qu'une fonction de répartition sur le cercle est uniquement déterminée par ses coefficients de Fourier. En effet, l'analyse de Fourier est l'outil approprié pour traiter les problèmes concernant les sommes de variables aléatoires définies modulo 1. Nous allons maintenant fournir un cas particulier où cette convergence est immédiate, mais d'abord, rappelons le critère de Weyl. Étant donné un vecteur aléatoire réel de $\mathbb{R}^d$ Y, nous notons ϕ_Y sa fonction caractéristique définie par $\phi_Y(u) = \mathbb{E}(e^{iu^T Y}) = \int_{\mathbb{R}^d} e^{iu^T y}\mu_Y(dy)$, $u \in \mathbb{R}^d$.

Theorem 11. *Soit* $(B_M)_{M\ge1}$ *une suite de variables aléatoires. Alors, lorsque* $M\to\infty$*, la suite* $(\{B_M\})_{M\ge1}$ *converge en distribution vers* $U \sim \mathcal{U}(0,1)$ *si et seulement si pour tout* $k \in \mathbb{Z}\setminus\{0\}$

$$\lim_{M\to\infty} \phi_{B_M}(2k\pi) = 0.$$

Si les variables aléatoires $(Y_i)_{i\ge1}$ sont i.i.d. telles que la loi de Y_1 n'est pas une variable aléatoire discrète, pour chaque $k \in \mathbb{Z}\setminus\{0\}$, on a

$$\phi_{Y_1+\cdots+Y_M}(2k\pi) = \mathbb{E}\left(e^{i2k\pi Y_1}\right)^M.$$

Il existe $\psi \ge 0, \rho \in [0, 2\pi)$ tels que $\mathbb{E}\left(e^{i2k\pi Y_1}\right) = \psi e^{i\rho}$. Par conséquent,

$$\mathbb{E}\left(e^{i(2k\pi Y_1-\rho)}\right) = \mathbb{E}\left(\mathfrak{Re}\left(e^{i(2k\pi Y_1-\rho)}\right)\right) = \psi.$$

$\mathfrak{Re}\left(e^{i(2k\pi Y_1-\rho)}\right) \leq 1$ presque sûrement, et, lorsque $k \neq 0$, étant donné que Y_1 n'est pas une variable aléatoire discrète, $\mathbb{P}\left(\mathfrak{Re}\left(e^{i(2k\pi Y_1-\rho)}\right)=1\right) < 1$. Par conséquent, lorsque $k \neq 0$, $\psi < 1$ et donc $\lim_{M\to\infty} \mathbb{E}\left(e^{i2k\pi Y_1}\right)^M = 0$. Ainsi, selon le critère de Weyl, $\{Y_1 + \cdots + Y_M\} \xrightarrow{d} U \sim \mathcal{U}(0,1)$.

Donnons un aperçu succinct de la littérature portant sur la dérivation de la loi de Benford et ses généralisations. En 1939, Lévy [39] a donné des conditions nécessaires et suffisantes pour cette convergence lorsque les Y_i sont i.i.d. En 1986, Störmer [58] a fourni des conditions suffisantes en termes de fonctions de répartition des Y_i pour que la convergence soit maintenue lorsque les Y_i sont simplement indépendants. En 2007, en supposant que les Y_i sont indépendantes et absolument continues, Miller et Nigrini [44] ont caractérisé la convergence de $(V_M)_{M\geq 1}$ dans $L^1([0,1])$. En 2010, Szewczak [59] a généralisé les résultats ci-dessus en l'hypothèse d'indépendance des Y_i. En particulier, il a prouvé que si la valeur absolue de la fonction caractéristique de $Y_1+\cdots+Y_M$ satisfait une certaine condition de croissance et si l'ensemble $\left\{n \in \mathbb{N}\setminus\{0\} : \left|\mathbb{E}\left(e^{2\pi i n Y_1}\right)\right| = 1\right\}$ est vide, alors $(V_M)_{M\geq 1}$ converge en loi vers une variable aléatoire uniforme sur $[0,1]$.

Nous présentons maintenant formellement la loi de Benford. La loi de Benford en base $b > 1$ est la mesure de probabilité μ_b sur l'intervalle $[1,b)$ définie par

$$\mu_b([1,a)) = \log_b a, \qquad \forall a \in [1,b).$$

La mantisse en base b d'un nombre réel positif x est le nombre unique $\mathcal{M}_b(x)$ dans $[1,b)$ tel que $x = \mathcal{M}_b(x) \times b^{\lfloor \log_b(x)\rfloor}$.

Une suite de nombres positifs $a_n, n \geq 1$ satisfait la loi de Benford en base b si $\forall a \in [1,b)$

$$\lim_{M\to\infty} \frac{\#\{n \leq M : 1 \leq \mathcal{M}_b(a_n) \leq a\}}{M} = \log_b(a).$$

Ainsi, si nous fixons $b = 10$, la probabilité que le premier chiffre soit a est $\log_{10}(a+1) - \log_{10}(a) = \log_{10}\left(1+\frac{1}{a}\right)$. Les chiffres plus petits comme $1, 2$, et 3 apparaissent plus fréquemment en tant que premier chiffre, tandis que les chiffres plus grands comme 8 et 9 surviennent moins fréquemment. Le premier à avoir observé ce phénomène statistique était Newcomb dans les années 1880. Il a remarqué que les pages dans les tables logarithmiques correspondant à des nombres commençant par 1 étaient significativement plus usées que les pages correspondant à des nombres avec un premier chiffre plus élevé. Depuis lors, de nombreux systèmes divers ont été démontrés pour satisfaire à la loi de Benford (voir [5], [48]). Cette loi trouve des applications dans divers domaines, notamment la comptabilité légale, la détection de fraudes, l'analyse de données et le contrôle qualité. De nos jours, elle est particulièrement utile pour identifier des irrégularités ou des anomalies dans les ensembles de données, car des écarts par rapport à la distribution attendue peuvent indiquer une manipulation ou des erreurs dans les données.

Nous nous intéressons au problème de la distribution des premiers chiffres des produits de variables aléatoires (voir par exemple [18]). Étant donnée une suite de variables aléatoires positives $(X_i)_{i\geq 1}$, de nombreux chercheurs ([54], [9], [55], [40]) se sont intéressés à l'étude de la convergence faible lorsque $M \to \infty$ de la loi de $\mathcal{M}_b\left(\prod_{i=1}^M X_i\right)$ vers μ_b. Comme $\log_b \mathcal{M}_b(x) = \{\log_b(x)\}$ pour chaque nombre réel positif x, on a

$$\log_b \mathcal{M}_b\left(\prod_{i=1}^M X_i\right) = \left\{\sum_{i=1}^M \log_b(X_i)\right\},$$

et donc la convergence faible de la loi de $\mathcal{M}_b\left(\prod_{i=1}^M X_i\right)$ vers μ_b est équivalente à la convergence en loi des sommes partielles modulo 1 des variables aléatoires $(\log_b(X_i))_{i\geq 1}$ en raison de la continuité

de $[1,b) \ni z \mapsto \log_b(z)$ et de son inverse.

1.3 Contributions de la thèse

Cette section rassemble les résultats principaux obtenus pendant la thèse.

1.3.1 Théorème de la Limite Centrale pour les fonctionnelles non linéaires des mesures empiriques : au-delà du cadre i.i.d. (voir Chapitre 2)

Notre objectif est de généraliser ce qui a été fait par Jourdain et Tse [36] (voir Section 1.1.4, Théorème 3). Nous relaxons d'abord l'hypothèse d'équidistribution en considérant une suite de variables aléatoires indépendantes X_i à valeurs dans $\mathbb{R}^d$ avec des lois $\nu_i \in \mathcal{P}_\ell\left(\mathbb{R}^d\right), i \geq 1$. Comme les variables ne sont plus équidistribuées, nous supposons que

$$\bar{\nu}_N = \mathbb{E}\left(\mu_N\right) = \frac{1}{N}\sum_{i=1}^{N} \nu_i$$

converge vers une limite $\mu \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$ pour la distance de Wasserstein d'ordre ℓ. En outre, en supposant que

$$\textbf{TX} \text{ il existe } \beta \in (0,1) \text{ tel que } \sum_{i=1}^{\infty} \frac{1}{i}\mathbb{E}\left(\left(|X_i|^\ell - i^\beta\right) 1_{\left\{|X_i|^\ell > i^\beta\right\}}\right) < \infty,$$

ce qui n'est autre que le deuxième point de la condition $SRCI\left(\beta\right)$ appliquée à $f\left(x\right) = |x|^\ell$ (voir Section 1.1.2), il est possible de prouver que, en utilisant le résultat de Wellner [60], $\mu_N = \frac{1}{N}\sum\limits_{i=1}^{N} \delta_{X_i}$ converge également vers μ pour la distance de Wasserstein d'ordre ℓ. Étant donné $U : \mathcal{P}_\ell\left(\mathbb{R}^d\right) \to \mathbb{R}$, nous considérons la décomposition suivante

$$\sqrt{N}\left(U\left(\mu_N\right) - U\left(\mu\right)\right) = \sqrt{N}\left(U\left(\mu_N\right) - U\left(\bar{\nu}_N\right)\right) + \sqrt{N}\left(U\left(\bar{\nu}_N\right) - U\left(\mu\right)\right).$$

Nous fournissons des conditions suffisantes pour que le second terme du côté droit de l'égalité converge vers une constante lorsque $N \to \infty$. Nous appliquons ensuite l'approche de Jourdain-Tse pour étudier le premier terme. En particulier, grâce au Lemme 1, ce premier terme s'écrit comme un reste plus

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j} + \frac{1}{N}\sum_{j=i}^{N}\nu_j, X_i\right) - \int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j} + \frac{1}{N}\sum_{j=i}^{N}\nu_j, x\right)\nu_i(dx)\right).$$

Nous fournissons des conditions suffisantes pour que le reste converge en probabilité vers zéro, et nous appliquons à la somme au-dessus le Théorème de la Limite Centrale pour les tableaux de martingale (voir Section 1.1.3). En appliquant ce dernier, en particulier nous devons vérifier la "condition du crochet". Nous sommes alors confrontés au problème de comprendre la convergence de

$$\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\mu, x\right)\frac{\delta U}{\delta m}\left(\mu, y\right)\nu_i\left(dx\right)\nu_i\left(dy\right).$$

C'est pourquoi, pour obtenir le Théorème de la Limite Centrale, nous devons supposer la convergence lorsque $N \to \infty$ de $\frac{1}{N}\sum_{i=1}^{N}\nu_i\left(dx\right)\nu_i\left(dy\right)$ vers une mesure de probabilité limite η sur $\mathbb{R}^d \times \mathbb{R}^d$

dont les lois marginales sont égales à μ.

Pour $\ell \geq 0$, soit $\mathcal{M}_\ell(\mathbb{R}^d)$ l'espace des mesures signées τ sur $\mathbb{R}^d$ telles que $\int_{\mathbb{R}^d} |x|^\ell |\tau|(dx) < \infty$. Pour chaque $\tau \in \mathcal{M}_\ell(\mathbb{R}^d)$, nous définissons la norme $\|\tau\|_\ell := \sup_{f:|f(x)|\leq 1+|x|^\ell} \int_{\mathbb{R}^d} f(x)\tau(dx)$.
En conclusion, nous prouvons le théorème suivant.

Theorem 12. *Soit $\ell \geq 0$ et soit X_i, $i \geq 1$ une suite de variables aléatoires indépendantes à valeurs dans $\mathbb{R}^d$ de loi $\nu_i \in \mathcal{P}_\ell(\mathbb{R}^d)$. Supposons **TX** et les hypothèses **1-4** du Théorème 3 sur la régularité de la dérivée fonctionnelle U. De plus, supposons que*

- $W_\ell\left(\frac{1}{N}\sum_{i=1}^N \nu_i(dx)\nu_i(dy), \eta(dx,dy)\right) \underset{N\to\infty}{\longrightarrow} 0$ *pour une mesure $\eta \in \mathcal{P}_\ell(\mathbb{R}^d \times \mathbb{R}^d)$ avec $\mu(\cdot) = \eta(\mathbb{R}^d, \cdot) = \eta(\cdot, \mathbb{R}^d)$*
- $\|\sqrt{N}(\bar{\nu}_N - \mu) - \sigma\|_\ell \underset{N\to\infty}{\longrightarrow} 0$ *pour une mesure σ dans $\mathcal{M}_\ell(\mathbb{R}^d)$.*

Alors, la convergence en loi suivante est vérifiée

$$\sqrt{N}\left(U(\mu_N) - U(\mu)\right) \overset{d}{\Longrightarrow}$$
$$\mathcal{N}\left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu, x)\,\sigma(dx), \int_{\mathbb{R}^d}\left(\frac{\delta U}{\delta m}(\mu,x)\right)^2 \mu(dx) - \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu,x)\frac{\delta U}{\delta m}(\mu,y)\eta(dx,dy)\right).$$

Dans un deuxième temps, nous relâchons également l'hypothèse d'indépendance en supposant que la suite de vecteurs aléatoires $(X_i)_{i\geq 1}$ est une chaîne de Markov à valeurs dans $\mathbb{R}^d$ avec une distribution initiale ν_1 et un noyau de transition P sur $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$. De plus, nous supposons que la chaîne de Markov est V-uniformément ergodique (voir Section 1.1.5) et nous notons μ son unique mesure de probabilité invariante. Comme dans (1.8), nous linéarisons à nouveau $\sqrt{N}\left(U(\mu_N) - U(\mu)\right)$ en la somme de

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \left(\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j} + \frac{N+1-i}{N}\mu, X_i\right) - \int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j} + \frac{N+1-i}{N}\mu, x\right)\mu(dx)\right) \tag{1.19}$$

et un reste qui tend vers zéro en probabilité lorsque $M \to \infty$. Afin d'appliquer le Théorème Central Limite pour les tableaux de martingale, nous introduisons pour $m \in B(\mu, r)$ (la boule pour la métrique W_ℓ où U admet une dérivée fonctionnelle linéaire) l'équation de Poisson suivante

$$F(m,x) - PF(m,x) = \frac{\delta U}{\delta m}(m,x) - \int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(m,y)\,\mu(dy), \quad x \in \mathbb{R}^d. \tag{1.20}$$

Sous les hypothèses qu'il existe $C < \infty$ tel que $\forall x \in \mathbb{R}^d$, $\left|\frac{\delta U}{\delta m}(m,x)\right| \leq C\left(1 + |x|^{\frac{\ell}{2}}\right)$ et

L3 $C_\ell := \sup_{x\in\mathbb{R}^d} \frac{|x|^\ell}{V(x)} < \infty$,

selon la Proposition 2, (1.20) admet une solution donnée par $F(m,\cdot) = \sum_{n\in\mathbb{N}}(P^n \frac{\delta U}{\delta m}(m,\cdot) - \mu(\frac{\delta U}{\delta m}(m,\cdot)))$. Par conséquent, nous pouvons réécrire (1.19) comme

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \left(F\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j} + \frac{N+1-i}{N}\mu, X_i\right) - PF\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j} + \frac{N+1-i}{N}\mu, X_{i-1}\right)\right)$$

à laquelle nous appliquons le Théorème de la Limite Centrale pour les tableaux de martingale, plus un reste qui converge vers 0 en probabilité. Afin d'appliquer le Théorème de la Limite Centrale pour les tableaux de martingale, nous devons vérifier que la somme ci-dessus est de carré intégrable et satisfait la condition de Lindeberg. Pour ce faire, nous avons besoin de la propriété d'intégrabilité de la fonction V par rapport à la distribution initiale ν_1. Pour être en mesure de couvrir toute choix de la distribution initiale de la chaîne de Markov, nous prouvons qu'il est possible de relaxer cette hypothèse.

En conclusion, nous démontrons le théorème suivant.

Theorem 13. *Soit $\ell \geq 0$ et soit X_i, $i \geq 1$ une chaîne de Markov avec une loi initiale ν_1 et un noyau de transition P. Nous supposons que $(X_i)_{i\geq 1}$ est V-uniformément ergodique avec V satisfaisant à **L3**. Soit μ son unique mesure de probabilité invariante. De plus, supposons les hypothèses **1-3** du Théorème 3 concernant la régularité de la dérivée fonctionnelle U et que*

- $\exists \alpha \in (\frac{1}{2}, 1], \exists C < \infty, \forall \mu_1, \mu_2 \in B(\mu, r)$

$$\sup_{x\in\mathbb{R}^d} \frac{\left|\frac{\delta U}{\delta m}(\mu_2, x) - \frac{\delta U}{\delta m}(\mu_1, x)\right|}{1 + |x|^{\frac{\ell}{2}}} \leq C \left(\int_{\mathbb{R}^d} (1 + |y|^{\frac{\ell}{2\alpha}}) |\mu_2 - \mu_1|(dy) \right)^{\alpha}.$$

Alors

$$\sqrt{N} \left(U\left(\mu_N\right) - U\left(\mu\right) \right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \mu\left(F^2\left(\mu, \cdot\right)\right) - \mu\left(\left(PF\right)^2\left(\mu, \cdot\right)\right)\right)$$

où $F\left(\mu, \cdot\right)$ désigne l'unique solution (à une constante additive près) de l'équation de Poisson

$$F(\mu, x) - PF(\mu, x) = \frac{\delta U}{\delta m}\left(\mu, x\right) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}\left(\mu, y\right) \mu(dy), \quad x \in \mathbb{R}^d.$$

1.3.2 Convergence vers la loi uniforme des vecteurs de sommes partielles modulo 1 avec un facteur commun (voir Chapitre 3)

Étant donné $(Y_i)_{i\geq 1}$ une suite de variables aléatoires réelles de carré intégrables i.i.d. et une fonction réelle mesurable ϕ, nous donnons des conditions suffisantes pour la convergence en loi vers la loi uniforme sur $[0, 1]$ de

$$\left\{ \phi\left(\frac{1}{M} \sum_{i=1}^{\beta_M} Y_i \right) \left(Y_1 + \cdots + Y_M\right) \right\}, \tag{1.21}$$

pour $(\beta_M)_{M\geq 1} \subseteq \mathbb{N}^*$ tels que $\lim_{M\to\infty} \sqrt{M}\left(\frac{\beta_M}{M} - \beta\right) = 0$ avec $\beta > 1$. Notons que cela généralise le cas d'une fonction constante ϕ présentée dans la Section 1.2 et a également une application dans le contexte de la loi de Benford. L'introduction d'une fonction ϕ non constante dans notre travail permet de traiter le choix d'une mantisse dépendant des données, par exemple donnée par $\hat{b}_M = e^{\frac{1}{\beta M} \sum_{i=1}^{\beta_M} \ln(X_i)}$. En effet,

$$\log_{\hat{b}_M} \mathcal{M}_{\hat{b}_M}\left(\prod_{i=1}^{M} X_i \right) = \left\{ \phi\left(\frac{1}{M} \sum_{i=1}^{\beta_M} \ln(X_i) \right) \sum_{i=1}^{M} \ln(X_i) \right\} \text{ où } \phi(x) = \frac{1}{\beta x}.$$

Cependant, l'étude de la convergence de (1.21) est principalement motivée par la dérivation d'un Théorème de la Limite Centrale pour le mécanisme d'échantillonnage stratifié que nous allons décrire en détail dans la Section 1.3.3, où la motivation pour étudier (1.21) sera explicite.

La principale difficulté dans l'étude de la convergence de (1.21) vient du couplage entre les variables introduit par le facteur commun $\phi\left(\frac{1}{M}\sum_{i=1}^{\beta_M} Y_i\right)$. L'étude de la convergence conjointe de plusieurs sommes partielles modulo 1 avec un facteur commun et une composante supplémentaire satisfaisant un Théorème de la Limite Centrale n'ajoute pas de difficultés significatives supplémentaires. Par conséquent, nous considérons directement la convergence en loi de

$$\left(\left(\left\{\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}} Y_m\right)\left(Y_1+\cdots+Y_{\beta_M^i}\right)\right\}\right)_{1\le i\le q}, \sqrt{M}\left(\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}} Y_m\right)\frac{1}{M}\sum_{i=1}^{\beta_M^{q+1}} Z_i-\theta\right)\right) \tag{1.22}$$

où

- $(\beta_M^1,\cdots,\beta_M^{q+1})_{M\ge1}\subseteq\mathbb{N}^{q+1}$ tels que $\lim_{M\to\infty}\sqrt{M}\left(\frac{\beta_M^i}{M}-\beta^i\right)=0$ avec $0<\beta^1<\cdots<\beta^{q+1}$
- $(Z_i)_{i\ge1}$ est une suite de variables aléatoires réelles telle que $((Y_i,Z_i))_{i\ge1}$ est une suite de variables aléatoires de carré intégrables i.i.d.
- la dernière composante où θ est une constante (typiquement égale à $\beta^{q+1}\phi\left(\beta^{q+1}\mathbb{E}(Y_1)\right)\mathbb{E}(Z_1)$) converge en distribution vers T lorsque $M\to\infty$.

Dans ce qui suit, nous posons $m_Y=\mathbb{E}(Y_1)$ et $m_Z=\mathbb{E}(Z_1)$. Notre résultat principal est le suivant.

Theorem 14. *Soit $(Y_i,Z_i)_{i\ge1}$ une suite de vecteurs aléatoires de carré intégrable i.i.d. à valeurs dans $\mathbb{R}^2$ tels que la loi de Y_i a une composante absolument continue. De plus, soit $\phi:\mathbb{R}\to\mathbb{R}$ une fonction mesurable différentiable en $\beta^{q+1}m_Y$ telle que $\phi(\beta^{q+1}m_Y)\neq0$. S'il existe $\tilde{M}\in\mathbb{N}$ tel que*

$$\int_{\mathbb{R}}\frac{1}{|\phi(\beta^{q+1}m_Y+y)|^{q+1}}e^{-\tilde{M}y^2}dy<\infty, \tag{1.23}$$

alors le vecteur aléatoire (1.22) converge en loi lorsque $M\to\infty$ vers (U,T) où U est une variable aléatoire uniforme sur $[0,1]^q$ indépendante de T.

Pour traiter le problème du couplage entre les variables, le Théorème 1 sera fondamental. De plus, pour comprendre la convergence en loi du vecteur ci-dessus, nous fournissons le critère de Weyl suivant, qui est une généralisation de celui présenté dans la Section 1.2.

Theorem 15 (Critère de Weyl)**.** *Soit $(B_M)_{M\ge1}$ une suite de vecteurs aléatoires à valeurs dans $\mathbb{R}^q$ et soit $(H_M)_{M\ge1}$ une suite de vecteurs aléatoires à valeurs dans $\mathbb{R}^{q'}$ qui converge en loi vers un vecteur aléatoire H à valeurs dans $\mathbb{R}^{q'}$. Alors lorsque $M\to\infty$, la suite $\left(\left\{B_M^1\right\},\cdots,\left\{B_M^q\right\},H_M\right)$ converge en loi vers (U,H) où U est indépendant de H et uniformément distribué sur $[0,1]^q$ si et seulement si pour tout $k\in\mathbb{Z}^q\backslash\{0\}$ et $u\in\mathbb{R}^{q'}$*

$$\lim_{M\to\infty}\phi_{(B_M,H_M)}(2k\pi,u)=0. \tag{1.24}$$

Décrivons brièvement les idées principales de la preuve du Théorème 14. Définissons

$$R_M:=\left(\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}} Y_m\right)\left(Y_1+\cdots+Y_{\beta_M^1}\right),\cdots,\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}} Y_m\right)\left(Y_{\beta_M^{q-1}+1}+\cdots+Y_{\beta_M^q}\right)\right)$$

$$K_M:=\sqrt{M}\left(\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}} Y_m\right)\frac{1}{M}\sum_{i=1}^{\beta_M^{q+1}} Z_i-\theta\right)$$

ÉTAPE 1. Nous commençons par prouver le théorème sous l'hypothèse renforcée que la loi de (Y_i,Z_i) a une composante absolument continue.

1. Soit $\gamma = \left(\beta^1, \beta^2 - \beta^1, \cdots, \beta^q - \beta^{q-1}\right)$. De plus, soit

$$a_M := \frac{1}{\sqrt{M}},\ b_M := -m_Y \phi\left(\beta^{q+1} m_Y\right) \gamma \sqrt{M}.$$

Grâce au Théorème 1, nous " réécrivons asymptotiquement" le vecteur $(a_M R_M + b_M, K_M)$ en termes d'un vecteur aléatoire normal T_M :

$$\lim_{M\to\infty} d_{TV}\left(\mu_{a_M R_M + b_M}, \mu_{T_M}\right) = 0.$$

2. Soit $\mu_{T_M} = \mu_{T_M,c} + \mu_{T_M,s}$ la décomposition de μ_{T_M} en une partie $\mu_{T_M,c}$ absolument continue par rapport à la mesure de Lebesgue et une partie singulière $\mu_{T_M,s}$. Soit p_{T_M} une densité de $\mu_{T_M,c}$ par rapport à la mesure de Lebesgue. Nous vérifions que

$$\liminf_{M\to\infty} p_{T_M}(x) \geq p(x)\, dx \ p.s.$$

où p désigne la densité de $Y \sim \mathcal{N}(0, \Gamma)$, $\Gamma \in \mathbb{R}^{(q+1)\times(q+1)}$ définie positive. Cela implique que

$$\lim_{M\to\infty} d_{TV}\left(\mu_{(a_M R_M + b_M, K_M)}, \mu_Y\right) = 0. \tag{1.25}$$

3. Avec le lemme de Riemann-Lebesgue, (1.25) implique que $\forall (u_1, u_2) \in \{\mathbb{R}^q \setminus \{0\}\} \times \mathbb{R}$

$$\lim_{M\to\infty} \phi_{(R_M, K_M)}(u_1, u_2) = 0.$$

Puisque cela équivaut au fait que $\forall (u_1, u_2) \in \{\mathbb{R}^q \setminus \{0\}\} \times \mathbb{R}$

$$\lim_{M\to\infty} \phi\left(\left(\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}} Y_m\right)\left(Y_1 + \cdots + Y_{\beta_M^i}\right)\right)_{1\leq i\leq q}, K_M\right)(u_1, u_2) = 0,$$

le Théorème 15 permet de conclure la preuve.

ÉTAPE 2. Nous prouvons le théorème sous l'hypothèse que (Y_i, Z_i) est telle que la loi de Y_i a une composante absolument continue :

1. Soit $(\xi_i)_{i\geq 1}$ une suite de variables aléatoires réelles de carré intégrables i.i.d. à moyenne nulle, absolument continues et indépendantes de $(Y_i, Z_i)_{i\geq 1}$. Pour $n \geq 1$, nous considérons la suite de vecteurs aléatoires i.i.d. de carré intégrables $\left(Y_i, Z_i + \frac{\xi_i}{n}\right)_{i\geq 1}$.

2. Nous montrons que la loi de $\left(Y_i, Z_i + \frac{\xi_i}{n}\right)$ a une composante absolument continue et nous appliquons l'**ÉTAPE 1** à la suite $\left(Y_i, Z_i + \frac{\xi_i}{n}\right)_{i\geq 1}$ pour chaque n.

3. Nous déduisons le théorème pour la suite (Y_i, Z_i).

1.3.3 Théorème de la Limite Centrale pour le mécanisme d'échantillonnage stratifié (voir Chapitre 4)

Première partie

Nous nous concentrons maintenant sur le schéma d'échantillonnage stratifié introduit dans la Section 1.1.6 pour les poids $w_m^M = g(X_m) \Big/ \sum_{\ell=1}^{M} g(X_\ell)$ où g est une fonction mesurable strictement positive.

Notre objectif est d'étudier le comportement asymptotique de la méthode lorsque le nombre de particules M tend vers l'infini. Étant donné f une fonction mesurable bornée réelle et sous l'hypothèse que les X_i sont i.i.d. de telle sorte que la loi de $g(X_i)$ a une composante absolument continue, nous explicitons la limite de $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f\left(Y_m^M\right)\right)$ et prouvons que $\frac{1}{\sqrt{M}}\sum_{m=1}^{M}\left(f(Y_m^M) - \frac{\mathbb{E}(f(X_1)g(X_1))}{\mathbb{E}(g(X_1))}\right)$ converge en loi vers une variable aléatoire gaussienne centrée dont la variance est égale à la limite. Décrivons brièvement les principales idées dans l'étude respectivement de la variance asymptotique et du Théorème de la Limite Centrale afin que le lien avec la section précédente soit clair.

Variance asymptotique : principales idées

Nous rappelons qu'à partir de M variables aléatoires $(U_m)_{1\leq m\leq M}$ i.i.d. distribuées selon la loi uniforme sur $(0,1)$ et indépendantes de $(X_m)_{1\leq m\leq M}$, la suite $(Y_m^M)_{1\leq m\leq M}$ est définie comme suit

$$Y_m^M = \sum_{i=1}^{M} 1_{\{\bar{S}_{i-1}^M < m-U_m \leq \bar{S}_i^M\}} X_i \text{ pour } m\in\{1,\cdots,M\} \text{ avec } \bar{S}_j^M = M\sum_{m=1}^{j} w_m^M,$$

avec la convention $\bar{S}_0^M = 0$. Pour $f:\mathbb{R}^d\to\mathbb{R}$ mesurable, nous avons

$$\mathbb{E}\left(f(Y_m^M)|\mathcal{F}\right) = \sum_{i=1}^{M} f(X_i)\int_{m-1}^{m} 1_{\{\bar{S}_{i-1}^M < u\leq \bar{S}_i^M\}} du \tag{1.26}$$

où $\mathcal{F} = \sigma\left((X_m)_{1\leq m\leq M}\right)$. De plus, la propriété suivante, commune à tous les schémas d'échantillonnage, est vérifiée

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m^M}\,\middle|\,\mathcal{F}\right) = \sum_{m=1}^{M} w_m^M \delta_{X_m}. \tag{1.27}$$

En utilisant une décomposition standard de la variance et l'indépendance conditionnelle de $(Y_m^M)_{1\leq m\leq M}$ donnée $\mathcal{F}$, nous obtenons

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\right)\\
&= \mathrm{Var}\left(\mathbb{E}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right)\\
&= \mathrm{Var}\left(\sqrt{M}\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{m=1}^{M} g(X_m)}\right) + \mathbb{E}\left(\frac{\sum_{m=1}^{M} g(X_m)f^2(X_m)}{\sum_{m=1}^{M} g(X_m)}\right) - \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left(f(Y_m^M)\,\middle|\,\mathcal{F}\right)^2\right).
\end{aligned} \tag{1.28}$$

Étant donné que le comportement asymptotique lorsque $M\to\infty$ des deux premiers termes de la dernière ligne peut être analysé à l'aide d'arguments standard, nous nous concentrons sur le troisième terme. En utilisant (1.26) et $\bar{S}_M^M = M$, nous obtenons que $\sum_{m=1}^{M}\mathbb{E}\left(f(\xi_m^M)|\mathcal{F}\right)^2$ est égal à

$$\sum_{k=0}^{M-1}\left(1+1_{\{k\geq 1\}}\right)\sum_{i=1}^{M-k} f(X_i)f(X_{i+k})\sum_{m\geq\lfloor\bar{S}_{i-1}^M\rfloor+1}\int_{m-1}^{m} 1_{\{\bar{S}_{i-1}^M<u\leq\bar{S}_i^M\}}du\int_{m-1}^{m} 1_{\{\bar{S}_{i+k-1}^M<u\leq\bar{S}_{i+k}^M\}}du.$$

Étant donné que $\frac{1}{M}\sum_{i=1}^{M}\gamma_i = \int_0^1 \gamma_{\lceil\alpha M\rceil}d\alpha$, nous en déduisons que $\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left(f(\xi_m^M)|\mathcal{F}\right)^2$ est égal à la somme sur $k\in\mathbb{N}$ de

$$\left(1+1_{\{k\geq 1\}}\right)\int_0^1 1_{\{\lceil \alpha M\rceil \leq M-k\}} f(X_{\lceil \alpha M\rceil}) f(X_{\lceil \alpha M\rceil+k}) \psi_k\left(\left\{\bar{S}^M_{\lceil \alpha M\rceil-1}\right\}, \frac{Mg\left(X_{\lceil \alpha M\rceil}\right)}{\sum_{m=1}^M g(X_m)}, \cdots, \frac{Mg\left(X_{\lceil \alpha M\rceil+k}\right)}{\sum_{m=1}^M g(X_m)}\right) d\alpha$$

où

$$\psi_k(u_0, w_1, \cdots, w_{k+1}) = \sum_{m\geq 1} \int_{m-1}^m 1_{\{u_0<u\leq u_0+w_1\}} du \int_{m-1}^m 1_{\left\{u_0+\sum_{\ell=1}^k w_\ell < u \leq u_0+\sum_{\ell=1}^{k+1} w_\ell\right\}} du.$$

C'est pourquoi, afin de calculer la variance asymptotique de $\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_m^M)$, il est très utile de comprendre le comportement lorsque $M\to\infty$ de

$$\left\{\bar{S}^M_{\lceil \alpha M\rceil-1}\right\}$$

qui est égal au premier composant de (1.22) lorsque $q=1$, $Y_i = g(X_i), \phi(x) = \frac{1}{x}$, $\beta_M^1 = \lceil \alpha M\rceil - 1$ et $\beta_M^2 = M$.

Théorème de la Limite Centrale : principales idées

En ce qui concerne le Théorème de la Limite Centrale, la fonction caractéristique de

$$\frac{1}{\sqrt{M}}\sum_{m=1}^M \left(f(Y_m^M) - \frac{\mathbb{E}\left(f(X_1)g(X_1)\right)}{\mathbb{E}(g(X_1))}\right)$$

s'écrit pour $u\in\mathbb{R}$

$$\mathbb{E}\left(e^{iu\sqrt{M}\left(\frac{\sum_{m=1}^M g(X_m)f(X_m)}{\sum_{m=1}^M g(X_m)} - \frac{\mathbb{E}(f(X_1)g(X_1))}{\mathbb{E}(g(X_1))}\right)} \prod_{m=1}^M \mathbb{E}\left(e^{\frac{iu}{\sqrt{M}}\left(f\left(Y_m^M\right)-\mathbb{E}\left(f\left(Y_m^M\right)|\mathcal{F}\right)\right)} \,\middle|\, \mathcal{F}\right)\right).$$

En utilisant le développement en série de Taylor de $e^{\frac{iu}{\sqrt{M}}\left(f\left(Y_m^M\right)-\mathbb{E}\left(f\left(Y_m^M\right)|\mathcal{F}\right)\right)}$, on peut approximer le produit par une somme d'intégrales de fonctions de

$$\left(\left\{\bar{S}^M_{\lceil \alpha_1 M\rceil-1}\right\}, \left\{\bar{S}^M_{\lceil \alpha_2 M\rceil-1}\right\}, \cdots, \left\{\bar{S}^M_{\lceil \alpha_q M\rceil-1}\right\}\right)$$

par rapport à $1_{\{0<\alpha_1<\alpha_2<\cdots<\alpha_q<1\}} d\alpha_1 d\alpha_2\cdots d\alpha_q$. Cela relie étroitement le comportement asymptotique de la fonction caractéristique à celui des vecteurs (1.22) pour le choix de $Y_i = g(X_i)$, $\phi(x)=\frac{1}{x}$, $Z_i = g(X_i)f(X_i)$, $\beta_M^i = \lceil \alpha_i M\rceil - 1$ pour $i\in\{1,\cdots,q\}$ et $\beta_M^{q+1} = M$.

Deuxième partie

Dans cette partie du travail, nous examinons l'ensemble de l'algorithme des filtres particulaires associé au mécanisme d'échantillonnage stratifié. Voir la Section 1.1.6 pour une description de l'algorithme.

De manière similaire à ce que nous avons fait dans (1.28), nous pouvons réécrire la variance $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m})\right)$ pour $n\geq 0$ de la manière suivante

$$\mathrm{Var}\left(\sqrt{M}\frac{\sum_{m=1}^M g_n(X_n^{M,m})f(X_n^{M,m})}{\sum_{m=1}^M g_n(X_n^{M,m})}\right) + \mathbb{E}\left(\frac{\sum_{m=1}^M g_n(X_n^{M,m})f^2(X_n^{M,m})}{\sum_{m=1}^M g_n(X_n^{M,m})}\right) \tag{1.29}$$

$$-\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^M \mathbb{E}\left(f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)^2\right) \tag{1.30}$$

où $\mathcal{F}^n := \sigma\left(\left(X_s^{M,i}, Y_s^{M,i}\right)_{1\leq i\leq M, 0\leq s\leq n}\right)$. Sous une conjecture qui permet d'étudier la convergence de (1.29) et sous la conjecture que la convergence en distribution du vecteur (1.22) peut être généralisée de manière similaire aux étapes suivantes d'un algorithme des filtres particulaires où, par exemple, l'indépendance des variables n'est plus valable, nous fournissons une formule inductive pour la variance asymptotique $\lim_{M\to\infty} \mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_{n+1}^{M,m})\right)$.

Chapter 2

Central Limit Theorem over non-linear functionals of empirical measures: beyond the iid setting

The material for this chapter has been released in [26].

The central limit theorem is, with the strong law of large numbers, one of the two fundamental limit theorems in probability theory. Benjamin Jourdain and Alvin Tse have extended to non-linear functionals of the empirical measure of independent and identically distributed random vectors the central limit theorem which is well known for linear functionals. The main tool permitting this extension is the linear functional derivative, one of the notions of derivation on the Wasserstein space of probability measures that have recently been developed. The purpose of this work is to relax first the equal distribution assumption made by Jourdain and Tse and then the independence property to be able to deal with the successive values of an ergodic Markov chain.

Introduction

In this work we are interested in the convergence in distribution of

$$\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U\left(\mu\right)\right), \tag{2.1}$$

where U is a general function defined on some Wasserstein space of probability measures on $\mathbb{R}^d$ and $(X_i)_{i\geq 1}$ is a given sequence of $\mathbb{R}^d$-valued random vectors.
In the mathematical statistics literature, functionals $U(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i})$ are known as "statistical functions". Their asymptotic distribution theory was first studied by Von Mises [45] [46], under the assumption that $(X_i)_{i\geq 1}$ is a family of independent and identically distributed real-valued random variables. Through the use of the Gâteaux differential, he introduced a Taylor expansion of $U(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i})$ around $U(\mu)$ where μ is the common distribution of the random variables. He proved that if the linear term is the first nonvanishing term in the Taylor expansion of the functional $U(.)$ at μ, the limit distribution is normal (under the usual restrictions corresponding to the central limit theorem). The main technical difficulty was proving that the remainder in the Taylor expansion goes to zero. We refer to Chapter 6 of [56] for more details.
In dimension $d=1$, Boos and Serfling [8] assumed the existence of a differential for $U(.)$ at μ in a sense stronger than the Gâteaux differential. They assumed the existence of $\frac{d}{d\epsilon}_{|\epsilon=0^+}U\left(\mu+\epsilon\left(\nu-\mu\right)\right)=$

$dU(\mu, \nu-\mu)$ at μ that is linear in $\nu-\mu$ (for ν any probability measure on the real line) and such that

$$U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)=\frac{1}{N}\sum_{i=1}^{N}dU(\mu,\delta_{X_i}-\mu)+o\left(\left\|\frac{1}{N}\sum_{i=1}^{N}1_{\{X_i\leq\cdot\}}-\mu(-\infty,\cdot]\right\|_{\infty}\right). \quad (2.2)$$

Taking the advantage of some well known stochastic properties of the Kolmogorov-Smirnov distance, in particular the boundness in probability of $\left(\sqrt{N}\left(\left\|\frac{1}{N}\sum_{i=1}^{N}1_{\{X_i\leq\cdot\}}-\mu(-\infty,\cdot]\right\|_{\infty}\right)\right)_{N\geq 1}$, they conclude that $\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)\right)$ converges in distribution to a centered Gaussian random variable with asymptotic variance equal to the common variance of the independent and identically distributed random variables $dU(\mu,\delta_{X_i}-\mu)$ when they are square integrable and centered. In addition to the limitation of dimension $d=1$ in their approach, it relies on the uniformity in (2.2) of the approximation with respect to the Kolmogorov-Smirnov distance which is a strong assumption almost amounting to Fréchet differentiability of U at μ for the Kolmogorov norm. When μ is a probability measure on any measurable space, Dudley [21] obtained central limit theorems for $\sqrt{N}(U(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i})-U(\mu))$ under Fréchet differentiability of U at μ with respect to $\|\nu-\mu\|:=\sup_{f\in\mathcal{F}}\left|\int f(x)(\nu-\mu)(dx)\right|$. The class $\mathcal{F}$ of measurable functions is such that a central limit theorem for empirical measures holds with respect to uniform convergence over $\mathcal{F}$. Clearly the requirements on $\mathcal{F}$ impose some balance: the larger the class $\mathcal{F}$, the simpler the Fréchet differentiability becomes, but the harder the uniform convergence over $\mathcal{F}$ becomes.
Recently Jourdain and Tse [36] reconsidered the same problem under the assumption of independent and identically distributed $\mathbb{R}^d$-valued random vectors $(X_i)_{i\geq 1}$. By means of the notion of the linear functional derivative of U that is a Gâteaux differential with the property that $dU(\mu,\nu-\mu)=\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,y)(\nu-\mu)(dy)$ for some measurable real valued function $\mathbb{R}^d\ni y\to\frac{\delta U}{\delta m}(\mu,y)$ with some polynomial growth assumption in y, they linearize $\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)\right)$ into the sum of

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j}+\frac{N+1-i}{N}\mu,X_i\right)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j}+\frac{N+1-i}{N}\mu,x\right)\mu(dx)\right) \quad (2.3)$$

and a remainder. Such a decomposition allows to apply to the above sum the Central Limit Theorem for arrays of martingale increments and to investigate sufficient conditions for the remainder to vanish in probability. The main result of [36] is the following convergence

$$\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)\right)\overset{d}{\Longrightarrow}\mathcal{N}\left(0,Var\left(\frac{\delta U}{\delta m}(\mu,X_1)\right)\right).$$

They replace the uniformity leading to Fréchet differentiability required in the statistical literature, by supposing that the linear functional derivative exists not only at μ but on a Wasserstein ball with positive radius containing μ. This is a mild restriction, since when a central limit theorem holds for some statistical functional, it is in general not limited to a single value of the common distribution μ of the samples.

The aim of this work is to generalize what has been done by Jourdain and Tse. We will first relax the equidistribution assumption by studying the convergence of (2.1) by assuming that the $\mathbb{R}^d$-valued random vectors $(X_i)_{i\geq 1}$ are independent and non-equidistributed. Since in our case we do not assume the equidistribution of the random variables, we need to give sufficient conditions for $\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}$ and for $\frac{1}{N}\sum_{i=1}^{N}\nu_i$ to converge to a common limit μ (for a distance that will be specified later) where ν_i denotes the law of X_i. We will split $\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)\right)$ into the sum

$$\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\nu_i\right)-U(\mu)\right)+\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U\left(\frac{1}{N}\sum_{i=1}^{N}\nu_i\right)\right), \quad (2.4)$$

and we will provide sufficient conditions for the first component to converge to a constant, whose value will be specified later. To prove the convergence in distribution to a Gaussian random variable of the second component, we use the linear functional derivative to rewrite it as a remainder plus

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j}+\frac{1}{N}\sum_{j=i}^{N}\nu_j,X_i\right)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j}+\frac{1}{N}\sum_{j=i}^{N}\nu_j,x\right)\nu_i(dx)\right).$$

The increment of the bracket corresponding to the i-th term in the sum involves the square of an integral with respect to $\frac{\nu_i}{\sqrt{N}}$ or equivalently an integral with respect to $\frac{\nu_i\otimes\nu_i}{N}$. That is why, to obtain the central limit theorem, we need to assume the convergence as $N\to\infty$ of $\frac{1}{N}\sum_{i=1}^{N}\nu_i(dx)\nu_i(dy)$ to some limiting probability measure η on $\mathbb{R}^d\times\mathbb{R}^d$ (with both marginals equal to μ) and the limiting variance is expressed in terms of η. Note that for the empirical mean of independent square integrable random variables retrieved with a linear functional U, the normalisation in the central limit theorem is given by the square root of the variance of their sum. Its asymptotic behaviour may be far from the one of the sequence $(\sqrt{N})_{N\geq 1}$ which is the only possible choice to deal with a wide class of nonlinear functionals U.

The next step is to drop the independence hypothesis by assuming that the sequence of random vectors $(X_i)_{i\geq 1}$ to be a $\mathbb{R}^d$-valued Markov chain with transition kernel P and unique invariant probability measure μ. By using the linear functional derivative $\frac{\delta U}{\delta m}$, we again linearize $\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)\right)$ into the sum of (2.3) and a remainder. Under assumptions on the Markov kernel P that ensure that the central limit theorem holds for linear functionals and by giving sufficient conditions for the remainder to vanish in probability as N goes to infinity, we conclude that

$$\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)\right)\overset{d}{\Longrightarrow}\mathcal{N}\left(0,\mu\left(PF^2(\mu,\cdot)\right)-\mu\left((PF)^2(\mu,\cdot)\right)\right)$$

with F solution of the Poisson equation

$$F(\mu,x)-PF(\mu,x)=\frac{\delta U}{\delta m}(\mu,x)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,y)\,\mu(dy),\quad x\in\mathbb{R}^d.$$

The main contribution in the linearization of $\sqrt{N}\left(U\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}\right)-U(\mu)\right)$ is now

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(F\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j}+\frac{N+1-i}{N}\mu,X_i\right)-PF\left(\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j}+\frac{N+1-i}{N}\mu,X_{i-1}\right)\right)$$

in place of (2.3). Of course, to prove that the remainder still vanishes, one needs to transfer the regularity of the functional U into regularity of the solution $F(\mu,x)$ to the Poisson equation in its measure argument μ. This is achieved under the V-uniform ergodicity hypothesis thanks to Lemma 4 below. To be able to cover any choice of the initial distribution of the Markov chain, we also relax using Lemma 3 below the integrability property of the function V with respect to this initial distribution. This integrability property is convenient to check that the above sum is square integrable and satisfies the Lindeberg condition in order to apply the central limit theorem for martingale increments.

As far as we know, such generalizations to nonlinear functionals of non i.i.d. random vectors X_i appear for the first time in the literature. Even if they are not straightforward and some additional technical difficulties appear compared to the i.i.d. setting, these difficulties may be handled thanks to the linear functional derivative. This illustrates that this notion of derivative is a versatile tool.

In the first section, we will provide the statement of the two results. Each of them will be preceeded by relevant reminders useful for its understanding. Together with the independent non-equidistributed case we will recall the notions of Wasserstein distance and linear functional derivative and together with the Markov chains case we will recall definitions and general facts about Markov chains and the Poisson equation. In the second section, the proofs of the two results are given.

2.1 Main Results

2.1.1 Independent Non-Equidistributed Random Variables

Let us observe that the case of a linear functional U, that is $U(m) = \int_{\mathbb{R}^d} f(x)m(dx)$ with $f : \mathbb{R}^d \to \mathbb{R}$ a measurable function, has been largely studied in this context.
We recall that a sequence $(f(X_i))_{i\geq 1}$ is Strongly Residually Cesaro β-Integrable for some $\beta > 0$ (SRCI(β), in short) if

1. $\sup_{N\geq 1} \frac{1}{N} \sum_{i=1}^N \mathbb{E}(|f(X_i)|) < \infty$
2. $\sum_{i=1}^\infty \frac{1}{i} \mathbb{E}\left(\left(|f(X_i)| - i^\beta \right) 1_{\{|f(X_i)|>i^\beta\}} \right) < \infty.$

Chandra and Goswami [13, Theorem 4.1] proved that if $(f(X_i))_{i\geq 1}$ is a sequence of random variables pairwise independent and verifying the condition SRCI(β) for some $\beta \in (0, 1)$, then the Strong Law of Large Numbers holds:

$$\lim_{N\to\infty} \frac{\sum_{i=1}^N (f(X_i) - \mathbb{E}(f(X_i)))}{N} = 0 \quad a.s. \tag{2.5}$$

Moreover Lindeberg proved (see for instance [7]) that, also in this case, the Central Limit Theorem holds. More in detail, consider a sequence $(f(X_i))_{i\geq 1}$ of square-integrable independent random variables such that $\lim_{N\to\infty} \frac{\sum_{i=1}^N Var(f(X_i))}{N} = \sigma^2$ where $\sigma^2 > 0$. If moreover the Lindeberg's condition is satified

$$\forall \epsilon > 0 \quad \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \mathbb{E}((f(X_i) - \mathbb{E}(f(X_i)))^2 1_{|f(X_i)-\mathbb{E}(f(X_i))|>\epsilon\sqrt{N}}) = 0,$$

then

$$\frac{\sum_{i=1}^N (f(X_i) - \mathbb{E}(f(X_i)))}{\sqrt{N}} \overset{d}{\Longrightarrow} \mathcal{N}(0, \sigma^2).$$

Let us first review some concepts related to the Wasserstein distance and the linear functional derivative before presenting the statement of the Central Limit Theorem for general U, in the case of independent non-equidistributed random variables.

The Wasserstein distance and the linear functional derivative

Let $U : \mathcal{P}_\ell\left(\mathbb{R}^d\right) \to \mathbb{R}$ where for $\ell \geq 0$ we denote by $\mathcal{P}_\ell\left(\mathbb{R}^d\right)$ the set of probability measures m on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |x|^\ell m(dx) < \infty$. For $\ell > 0$, we consider the ℓ-Wasserstein metric defined for $\mu_1, \mu_2 \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$ by

$$W_\ell(\mu_1, \mu_2) = \inf\left\{ \left(\int_{\mathbb{R}^{2d}} |x - y|^\ell \rho(dx, dy) \right)^{\frac{1}{\ell\vee 1}} : \rho \in \mathcal{P}(\mathbb{R}^{2d}) \, with \, \rho\left(\cdot \times \mathbb{R}^d\right) = \mu_1(\cdot), \rho\left(\mathbb{R}^d \times \cdot\right) = \mu_2(\cdot) \right\}. \tag{2.6}$$

If $\mu \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$ and $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in this space, then $\lim_{n\to\infty} W_\ell(\mu_n,\mu) = 0$ if and only if $\lim_{n\to\infty}\int_{\mathbb{R}^d}|x|^\ell\mu_n(dx) = \int_{\mathbb{R}^d}|x|^\ell\mu(dx)$ and μ_n converges weakly to μ as $n\to\infty$ where we write $\mu_n \rightharpoonup \mu$ to denote the weak convergence. Alternatively $\lim_{n\to\infty} W_\ell(\mu_n,\mu) = 0$ if and only if $\forall\phi : \mathbb{R}^d \to \mathbb{R}$ continous μ-almost everywhere and such that $\sup_{x\in\mathbb{R}^d}\frac{|\phi(x)|}{1+|x|^\ell} < \infty$,

$$\lim_{n\to\infty}\int_{\mathbb{R}^d}\phi(x)\mu_n(dx) = \int_{\mathbb{R}^d}\phi(x)\mu(dx). \tag{2.7}$$

For $\ell = 0$ and $\mu_1,\mu_2 \in \mathcal{P}_0\left(\mathbb{R}^d\right)$, we consider

$$W_0(\mu_1,\mu_2) = \inf\left\{\int_{\mathbb{R}^{2d}}(1\wedge|x-y|)\rho(dx,dy) : \rho\in\mathcal{P}(\mathbb{R}^{2d}) \text{ with } \rho\left(\cdot\times\mathbb{R}^d\right) = \mu_1(\cdot), \rho\left(\mathbb{R}^d\times\cdot\right) = \mu_2(\cdot)\right\}.$$

Notice that W_0 metricizes the topology of weak convergence. For $\ell \geq 0$ we can also consider $\mathcal{M}_\ell(\mathbb{R}^d)$, the space of the signed measures τ on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d}|x|^\ell|\tau|(dx) < \infty$ where $|\cdot|$ denotes the total variation of a signed measure. For each $\tau\in\mathcal{M}_\ell(\mathbb{R}^d)$ we will define the norm $\|\tau\|_\ell = \sup_{f:|f(x)|\leq 1+|x|^\ell}\int_{\mathbb{R}^d} f(x)\tau(dx)$ where the supremum is computed over the set of the measurable functions satisfying the growth condition and it can be proved that given $\tau\in\mathcal{M}_\ell(\mathbb{R}^d)$ and $(\tau_n)_{n\in\mathbb{N}}$ a sequence in this space such that $\|\tau_n-\tau\|_\ell \to 0$, then $\||\tau_n|-|\tau|\|_\ell \to 0$.
Let us observe that in $\mathcal{P}(\mathbb{R}^d)$ the convergence with respect to $\|\cdot\|_\ell$ implies the convergence with respect to W_ℓ. Moreover if $\ell = 0$, for $\mu_1,\mu_2\in\mathcal{P}(\mathbb{R}^d)$ we have

$$\begin{aligned}\|\mu_1-\mu_2\|_0 &= \sup_{\|f\|_\infty\leq 1}\left|\int_{\mathbb{R}^d} f(x)(\mu_1(dx)-\mu_2(dx))\right| = 2d_{TV}(\mu_1,\mu_2)\\ &= 2\sup_{A\in\mathcal{B}\left(\mathbb{R}^d\right)}|\mu_1(A)-\mu_2(A)| = |\mu_1-\mu_2|\left(\mathbb{R}^d\right)\end{aligned}$$

where $\mathcal{B}\left(\mathbb{R}^d\right)$ denotes the Borel σ-algebra of $\mathbb{R}^d$ and d_{TV} the total variation distance between μ_1 and μ_2. Let us now recall the notion of (first order) linear functional derivative associated to U. For a more detailed description, including the definition of the linear functional derivative of a superior order, see [11].

Definition 5. *Let $\ell\geq 0$. A function $U:\mathcal{P}_\ell\left(\mathbb{R}^d\right)\to\mathbb{R}$ admits a linear functional derivative at $\mu\in\mathcal{P}_\ell\left(\mathbb{R}^d\right)$ if there exists a measurable function $\mathbb{R}^d\ni y\mapsto\frac{\delta U}{\delta m}(\mu,y)$ such that $\sup_{y\in\mathbb{R}^d}\frac{\left|\frac{\delta U}{\delta m}(\mu,y)\right|}{1+|y|^\ell}<\infty$ and*

$$\forall\nu\in\mathcal{P}_\ell\left(\mathbb{R}^d\right), \frac{d}{d\epsilon}_{|_{\epsilon=0^+}} U(\mu+\epsilon(\nu-\mu)) = \int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,y)(\nu-\mu)(dy).$$

Since $(\nu-\mu)\left(\mathbb{R}^d\right) = 0$, $\frac{\delta U}{\delta m}$ is defined up to an additive constant. The next lemma allows to express a finite difference of the function U as an integral of the functional linear derivative.

Lemma 2. *Let $\ell\geq 0$, m, $m'\in\mathcal{P}_\ell\left(\mathbb{R}^d\right)$ and suppose that the linear functional derivative of a function $U:\mathcal{P}_\ell\left(\mathbb{R}^d\right)\to\mathbb{R}$ exists in the segment $(m_s := sm'+(1-s)m)_{s\in[0,1]}$.*
Then if $\sup_{(s,y)\in[0,1]\times\mathbb{R}^d}\frac{\left|\frac{\delta U}{\delta m}(m_s,y)\right|}{1+|y|^\ell}<\infty$, one has

$$U(m')-U(m) = \int_0^1\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}((1-s)m+sm',y)(m'-m)(dy)ds. \tag{2.8}$$

Statement of the theorem

Given a measure $\mu \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$, we will consider the following group of hypotheses about the regularity of the functional derivative U (**RU**) in a neighborhood of μ:

RU1 there exists $r > 0$ such that U admits a linear functional derivative on the ball $B(\mu, r)$ centered at μ with radius r for the metric W_ℓ

RU2 $\exists C < \infty, \forall(\tilde{\mu}, x) \in B(\mu, r) \times \mathbb{R}^d, \left|\frac{\delta U}{\delta m}(\tilde{\mu}, x)\right| \leq C\left(1 + |x|^{\frac{\ell}{2}}\right)$

RU3 $\sup_{x\in\mathbb{R}^d} \frac{|\frac{\delta U}{\delta m}(\tilde{\mu}, x) - \frac{\delta U}{\delta m}(\mu, x)|}{1+|x|^{\frac{\ell}{2}}}$ converges to 0 when $W_\ell(\tilde{\mu}, \mu)$ goes to 0

RU4 $x \mapsto \frac{\delta U}{\delta m}(\mu, x)$ is continuous μ-almost everywhere

RU5 $\exists \alpha \in (\frac{1}{2}, 1], \exists C < \infty, \forall \mu_1, \mu_2 \in B(\mu, r), \forall x \in \mathbb{R}^d$

$$\left|\frac{\delta U}{\delta m}(\mu_2, x) - \frac{\delta U}{\delta m}(\mu_1, x)\right| \leq C\left(\left(1 + |x|^\ell\right) \|\mu_2 - \mu_1\|_0^\alpha + \left(1 + |x|^{\ell(1-\alpha)}\right)\left(\int_{\mathbb{R}^d} |y|^\ell |\mu_2 - \mu_1|(dy)\right)^\alpha\right).$$

Moreover we will consider the following assumption about the tails of the random vectors $(X_i)_{i\geq 1}$:

TX there exists $\beta \in (0, 1)$ such that $\sum_{i=1}^\infty \frac{1}{i}\mathbb{E}\left(\left(|X_i|^\ell - i^\beta\right) 1_{\{|X_i|^\ell > i^\beta\}}\right) < \infty$.

Let us observe that it coincides with condition (ii) of the Strongly Residually Cesaro $\beta-$Integrability with the choice of $f(x) = x^\ell$. We are now ready to state respectively the Strong Law of Large Numbers and the Central Limit Theorem.

Theorem 16 (LLN for independent non-equidistributed r.v.)**.** *Let $\ell \geq 0$ and let X_i, $i \geq 1$ be a sequence of independent random variables on $\mathbb{R}^d$ with law $\nu_i \in \mathcal{P}_\ell(\mathbb{R}^d)$ and let us define*

$$\mu_N := \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$$

and

$$\bar{\nu}_N := \mathbb{E}\left(\mu_N\right) = \frac{1}{N}\sum_{i=1}^N \nu_i.$$

*Let us assume **TX** and the existence of $\mu \in \mathcal{P}_\ell(\mathbb{R}^d)$ such that $\lim_{N\to\infty} W_\ell(\bar{\nu}_N, \mu) = 0$. Then*

$$W_\ell\left(\mu_N, \mu\right) \underset{N}{\longrightarrow} 0 \quad a.s.$$

Proof. Thanks to the existence of $\mu \in \mathcal{P}_\ell(\mathbb{R}^d)$ such that $\lim_{N\to\infty} W_\ell(\bar{\nu}_N, \mu) = 0$, by the characterization of the Wasserstein convergence, one has $\bar{\nu}_N \rightharpoonup \mu$ and so by Wellner paper [60] $\mu_N \rightharpoonup \mu$ a.s.. If $\ell = 0$ the proof is completed, let us therefore suppose that $\ell > 0$ and let us check the convergence of the ℓ-th moment. Since the $|X_i|^\ell$'s are independent, Assumption **TX** holds and $\sup_{N\geq 1} \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left(|X_i|^\ell\right) = \sup_{N\geq 1} \int_{\mathbb{R}^d} |x|^\ell \bar{\nu}_N(dx) < \infty$ (using again the characterisation of the Wasserstein convergence), we can apply (2.5) and obtain that

$$\int_{\mathbb{R}^d} |x|^\ell \mu_N(dx) - \int_{\mathbb{R}^d} |x|^\ell \bar{\nu}_N(dx)$$
$$= \frac{|X_1|^\ell + \cdots + |X_N|^\ell - \left(\mathbb{E}\left(|X_1|^\ell\right) + \cdots + \mathbb{E}\left(|X_N|^\ell\right)\right)}{N} \underset{N\to\infty}{\longrightarrow} 0 \qquad a.s.$$

Therefore

$$\lim_{N\to\infty} \int_{\mathbb{R}^d} |x|^\ell \mu_N(dx) - \int_{\mathbb{R}^d} |x|^\ell \mu(dx)$$
$$= \lim_{N\to\infty} \int_{\mathbb{R}^d} |x|^\ell \mu_N(dx) - \int_{\mathbb{R}^d} |x|^\ell \bar{\nu}_N(dx) + \int_{\mathbb{R}^d} |x|^\ell \bar{\nu}_N(dx) - \int_{\mathbb{R}^d} |x|^\ell \mu(dx) = 0 \qquad a.s.$$

□

Theorem 17 (CLT for independent non-equidistributed r.v.)**.** *Using the same notations of Theorem 16, let us assume that*

1. $W_\ell\left(\frac{1}{N}\sum_{i=1}^N \nu_i(dx)\nu_i(dy), \eta(dx,dy)\right) \underset{N\to\infty}{\longrightarrow} 0$ *for some measure* $\eta \in \mathcal{P}_\ell(\mathbb{R}^d \times \mathbb{R}^d)$ *with* $\mu(\cdot) = \eta(\mathbb{R}^d, \cdot) = \eta(\cdot, \mathbb{R}^d)$
2. $\|\sqrt{N}(\bar{\nu}_N - \mu) - \sigma\|_\ell \underset{N\to\infty}{\longrightarrow} 0$ *for some measure* σ *in* $\mathcal{M}_\ell(\mathbb{R}^d)$

If moreover ***RU1-5*** *and* ***TX*** *hold then*

$$\sqrt{N}\left(U(\mu_N) - U(\mu)\right)$$
$$\overset{d}{\Longrightarrow} \mathcal{N}\left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu, x)\,\sigma(dx), \int_{\mathbb{R}^d} \left(\frac{\delta U}{\delta m}(\mu, x)\right)^2 \mu(dx) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{\delta U}{\delta m}(\mu, x)\frac{\delta U}{\delta m}(\mu, y)\eta(dx, dy)\right).$$

Remark 18. *Since the marginals of* $\frac{1}{N}\sum_{i=1}^N \nu_i(dx)\nu_i(dy)$ *and* $\eta(dx,dy)$ *are respectively* $\bar{\nu}_N$ *and* μ*, one has*

$$W_\ell(\bar{\nu}_N, \mu) \le W_\ell\left(\frac{1}{N}\sum_{i=1}^N \nu_i(dx)\nu_i(dy), \eta(dx,dy)\right)$$

and so Assumption 1 implies that $W_\ell(\bar{\nu}_N, \mu) \underset{N\to\infty}{\longrightarrow} 0$.

Remark 19. *According to the theorem, the asymptotic variance is given by*

$$\int_{\mathbb{R}^d} \left(\frac{\delta U}{\delta m}(\mu, x)\right)^2 \mu(dx) - \int_{\mathbb{R}^{2d}} \frac{\delta U}{\delta m}(\mu, x)\frac{\delta U}{\delta m}(\mu, y)\eta(dx, dy).$$

By Jensen's inequality

$$\frac{1}{N}\sum_{i=1}^N \int_{\mathbb{R}^{2d}} \frac{\delta U}{\delta m}(\mu, x)\frac{\delta U}{\delta m}(\mu, y)\nu_i(dx)\nu_i(dy) = \frac{1}{N}\sum_{i=1}^N \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu, x)\nu_i(dx)\right)^2$$
$$\ge \left(\frac{1}{N}\sum_{i=1}^N \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu, x)\nu_i(dx)\right)^2.$$

Thanks to Hypothesis 1, Hypothesis ***RU2*** *and Hypothesis* ***RU4****, by taking the limit over* $N \to \infty$*, one deduces that*

$$\int_{\mathbb{R}^{2d}} \frac{\delta U}{\delta m}(\mu, x)\frac{\delta U}{\delta m}(\mu, y)\eta(dx, dy) \ge \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu, x)\mu(dx)\right)^2.$$

Therefore the following "variance" inequality holds

$$\int_{\mathbb{R}^d}\left(\frac{\delta U}{\delta m}(\mu,x)\right)^2\mu(dx)-\int_{\mathbb{R}^{2d}}\frac{\delta U}{\delta m}(\mu,x)\frac{\delta U}{\delta m}(\mu,y)\eta(dx,dy)$$
$$\leq\int_{\mathbb{R}^d}\left(\frac{\delta U}{\delta m}(\mu,x)\right)^2\mu(dx)-\left(\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\mu(dx)\right)^2.$$

In what follows, we provide an example where Hypotheses 1 and 2 are verified.

Example 1. *Let $\theta_0,\theta_1,\cdots,\theta_{m-1}\in\mathcal{P}_\ell(\mathbb{R}^d)$ for $m\in\mathbb{N}^*$ and define $\nu_i(dx):=\theta_{(i-1)\,mod\,m}(dx)$ for $i\geq 1$. We are now going to verify that Hypothesis 1 and 2 are satisfied with $\eta(dx,dy)=\frac{1}{m}\sum_{i=0}^{m-1}\theta_i(dx)\theta_i(dy)$ and $\sigma=0$.*
If we prove that

$$\lim_{N\to\infty}\sqrt{N}\left\|\frac{1}{N}\sum_{i=1}^{N}\theta_{(i-1)\,mod\,m}(dx)\theta_{(i-1)\,mod\,m}(dy)-\frac{1}{m}\sum_{i=0}^{m-1}\theta_i(dx)\theta_i(dy)\right\|_\ell=0,\tag{2.9}$$

then using that the convergence with respect to $\|\cdot\|_\ell$ implies the convergence with respect to W_ℓ, Hypothesis 1 will follow. Moreover since the marginals of $\frac{1}{N}\sum_{i=1}^{N}\theta_{(i-1)\,mod\,m}(dx)\theta_{(i-1)\,mod\,m}(dy)$ and $\frac{1}{m}\sum_{i=0}^{m-1}\theta_i(dx)\theta_i(dy)$ are respectively $\frac{1}{N}\sum_{i=1}^{N}\theta_{(i-1)\,mod\,m}$ and $\frac{1}{m}\sum_{i=0}^{m-1}\theta_i$, Hypothesis 2 will follow too. Let us therefore prove (2.9). For each $N>0$, there exist $k_N\in\mathbb{N}$ and $r_N\in\mathbb{N}$ with $0\leq r_N<m$ such that $N=k_Nm+r_N$. Let f be a function on $\mathbb{R}^{2d}$ such that $|f(x,y)|\leq 1+|x|^\ell+|y|^\ell$. One has

$$\begin{aligned}
&\sqrt{N}\left|\int_{\mathbb{R}^{2d}}f(x,y)\frac{1}{N}\sum_{i=1}^{N}\theta_{i-1\,mod\,m}(dx)\theta_{i-1\,mod\,m}(dy)-\frac{1}{m}\sum_{i=0}^{m-1}\int_{\mathbb{R}^{2d}}f(x,y)\theta_i(dx)\theta_i(dy)\right|\\
&=\sqrt{N}\left|\left(\frac{k_N}{N}-\frac{1}{m}\right)\sum_{i=0}^{m-1}\int_{\mathbb{R}^{2d}}f(x,y)\theta_i(dx)\theta_i(dy)+\frac{1}{N}\sum_{i=0}^{r_N-1}\int_{\mathbb{R}^{2d}}f(x,y)\theta_i(dx)\theta_i(dy)\right|\\
&=\left|-\frac{r_N}{\sqrt{N}m}\sum_{i=0}^{m-1}\int_{\mathbb{R}^{2d}}f(x,y)\theta_i(dx)\theta_i(dy)+\frac{1}{\sqrt{N}}\sum_{i=0}^{r_N-1}\int_{\mathbb{R}^{2d}}f(x,y)\theta_i(dx)\theta_i(dy)\right|.
\end{aligned}$$

Taking the superior limit over $N\to\infty$, the right hand-side converges to 0 which completes the proof.

In the following example Hypothesis 2 is verified with $\sigma=0$.

Example 2. *Let $\mu\in\mathcal{P}_\ell(\mathbb{R}^d)$ and $\nu_i\in\mathcal{P}_\ell(\mathbb{R}^d)$ such that $\|\nu_i-\mu\|_\ell\leq\frac{c}{i^\alpha}$ for $i\geq 1$ with $c<\infty$ and $\alpha>\frac{1}{2}$. Then*

$$\begin{aligned}
&\sqrt{N}\sup_{f:|f(x)|\leq 1+|x|^\ell}\left|\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}f(x)\nu_i(dx)-\int_{\mathbb{R}^d}f(x)\mu(dx)\right|\\
&\leq\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\|\nu_i-\mu\|_\ell\leq\frac{c}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{i^\alpha}=\begin{cases}O\left(\frac{1}{\sqrt{N}}\right) & \text{if } \alpha>1\\ O\left(\frac{\ln N}{\sqrt{N}}\right) & \text{if } \alpha=1\\ O\left(N^{\frac{1}{2}-\alpha}\right) & \text{if } \alpha\in(0,1)\end{cases}
\end{aligned}$$

Therefore, since $\alpha>\frac{1}{2}$, taking the superior limit over $N\to\infty$ we have $\sqrt{N}\|\bar{\nu}_N-\mu\|_\ell\underset{N\to\infty}{\longrightarrow}0$.

2.1.2 Markov Chains

Markov Chains and the Poisson equation

We are now going to state the Central Limit Theorem for Markov chains. Let us first recall some facts useful to it. Let us consider X_i, $i \geq 1$ a Markov chain with initial distribution ν_1 and transition kernel P on $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$. Given a measurable function $V : \mathbb{R}^d \to [1, \infty)$, we can introduce the following normed space

$$\mathcal{V}_V := \left\{ \phi : \mathbb{R}^d \to \mathbb{R} \text{ measurable} : \sup_{x \in \mathbb{R}^d} \frac{|\phi(x)|}{V(x)} < \infty \right\}$$

equipped with the norm

$$\|\phi\|_V = \sup_{x \in \mathbb{R}^d} \frac{|\phi(x)|}{V(x)}.$$

Definition 6. *We say that a Markov chain $(X_i)_{i \geq 1}$ is V-uniformly ergodic if there exists $\mu \in \mathcal{P}\left(\mathbb{R}^d\right)$ such that*

$$\lim_{n \to \infty} \sup_{x \in \mathbb{R}^d} \frac{\sup_{\phi : \|\phi\|_V \leq 1} \left| P^n \phi\left(x\right) - \mu\left(\phi\right) \right|}{V\left(x\right)} = 0.$$

We now recall respectively the notions of $\psi-$irreducible Markov chains (see Proposition 4.2.1 [43]), positive Markov chain and Harris recurrent Markov chain.

Definition 7. *We say that a Markov chain $(X_i)_{i \geq 1}$ is ψ-irreducible if there exists a measure ψ on $\mathcal{B}\left(\mathbb{R}^d\right)$ such that for all $x \in \mathbb{R}^d$, whenever $\psi(A) > 0$, there exists some $n > 0$, possibly depending on both A and x, such that $P^n\left(x, A\right) > 0$.*

Definition 8. *Let $(X_i)_{i \geq 1}$ be a ψ-irreducible Markov chain. Then*

- *$(X_i)_{i \geq 1}$ is called a positive Markov chain if it admits an invariant probability measure.*
- *$(X_i)_{i \geq 1}$ is called Harris recurrent if every set $A \in \mathcal{B}\left(\mathbb{R}^d\right)$ s.t. $\psi\left(A\right) > 0$ satisfies*

$$\mathbb{P}_{\delta_x}\left(\bigcap_{N=1}^{\infty} \bigcup_{k=N}^{\infty} \left\{ X_k \in A \right\} \right) = 1, \quad x \in A.$$

The following proposition shows that under the V-uniformly ergodicity, the Markov chain has several properties: in particular it is positive Harris recurrent.

Proposition 3. *Let us assume that the Markov chain $(X_i)_{i \geq 1}$ is V-uniformly ergodic with probability measure $\mu \in \mathcal{P}\left(\mathbb{R}^d\right)$. Then $(X_i)_{i \geq 1}$ is a positive Harris recurrent Markov chain with unique invariant probability measure μ and*

$$\mu\left(V\right) < \infty. \tag{2.10}$$

Moreover

- *there exist $r \in (0, 1)$, $R < \infty$ such that for all $n \in \mathbb{N}$*

$$\sup_{x \in \mathbb{R}^d} \frac{\sup_{\phi : \|\phi\|_V \leq 1} \left| P^n \phi\left(x\right) - \mu\left(\phi\right) \right|}{V\left(x\right)} \leq R r^n \tag{2.11}$$

- *there exist* $\chi \in (0,1)$, $D < \infty$ *such that for all* $n \in \mathbb{N}$

$$\sup_{x \in \mathbb{R}^d} \frac{\sup_{\phi:\|\phi\|_{\sqrt{V}} \leq 1} \left|P^n \phi(x) - \mu(\phi)\right|}{\sqrt{V(x)}} \leq D\chi^n. \tag{2.12}$$

- *there exist* $K_1, K_2 \in \mathbb{R}_+$ *such that* $\forall x \in \mathbb{R}^d$

$$PV(x) \leq K_1 V(x) + K_2. \tag{2.13}$$

Proof of Proposition 3. Let us first observe that the hypothesis of V-uniformly ergodicity implies that $\forall x \in \mathbb{R}^d$, $\forall A \in \mathcal{B}\left(\mathbb{R}^d\right)$, $\left|P^n(x,A) - \mu(A)\right| \underset{n\to\infty}{\longrightarrow} 0$ and so according to Definition 7, the Markov chain $(X_i)_{i\geq 1}$ is μ-irreducible. Moreover the Markov chain is aperiodic.
Thus by Theorem 16.0.1 [43], the V-uniform ergodicity implies (2.11) and the existence of a real valued function $V_0 : \mathbb{R}^d \to [1,\infty)$ equivalent to V in the sense that $c^{-1}V \leq V_0 \leq cV$ for some $c \geq 1$, a petite set C (see Section 5.5.2 [43] for the definition of petite sets and their properties) and constants $H \in (0,1]$, $b \in [0,\infty)$ such that

$$PV_0(x) \leq (1-H)V_0(x) + b1_C(x), \quad x \in \mathbb{R}^d. \tag{2.14}$$

Since V_0 is equivalent to V, (2.13) easily follows and by Lemma 15.2.2 [43], V_0 is unbounded off petite sets that is $\left\{y \in \mathbb{R}^d : V_0(y) \leq n\right\}$ is petite for all $n \in \mathbb{N}$ (see Section 8.4.2 [43]). The above estimation has multiple consequences. By Theorem 14.0.1 [43], (2.14) implies that μ is an invariant probability measure satisfying

$$\mu(V) = \int_{\mathbb{R}^d} V(x)\,\mu(dx) < \infty$$

and so, by Definition 8, $(X_i)_{i\geq 1}$ is a positive Markov chain.
Moreover, since (2.14) obviously implies that $PV_0(x) \leq V_0(x)$ if $x \notin C$, by Theorem 9.1.8 [43] $(X_i)_{i\geq 1}$ is Harris recurrent. By the Aperiodic Ergodic Theorem (see Theorem 13.0.1 [43]), (2.14) implies also that the invariant measure μ is unique.
To conclude the proof, we just need to prove (2.12). Jensen's inequality and (2.14) imply that

$$\begin{aligned} P\sqrt{V_0}(x) &\leq \sqrt{PV_0(x)} \leq \sqrt{(1-H)V_0(x) + b1_C(x)} \\ &\leq \sqrt{1-H}\sqrt{V_0(x)} + \sqrt{b}1_C(x) \end{aligned}$$

By Theorem 16.1.2 [43], the Markov chain $(X_i)_{i\geq 1}$ is $\sqrt{V_0}$ - uniformly ergodic and, using again that $\sqrt{c^{-1}}\sqrt{V} \leq \sqrt{V_0} \leq \sqrt{c}\sqrt{V}$, it is also $\sqrt{V}$ - uniformly ergodic.
The application of Theorem 16.0.1 [43] concludes the proof of (2.12). □

Remark 20. *Let us observe that (2.13) and Jensen's inequality imply that*

$$P\sqrt{V}(x) \leq \sqrt{PV(x)} \leq \sqrt{K_1 V(x) + K_2} \leq \sqrt{K_1}\sqrt{V(x)} + \sqrt{K_2}. \tag{2.15}$$

We are now ready to recall respectively the Strong Law of Large Numbers and the Central Limit Theorem for Markov chains under V-uniform ergodicity. Let us start with the Strong Law of Large Numbers, a result that will be largely used in what follows (a proof is provided in Theorem 17.1.7 [43] under the assumption that the Markov chain is positive Harris recurrent).

Theorem 21. *Let us assume that the Markov chain* $(X_i)_{i\geq 1}$ *is V-uniformly ergodic and let* μ *be its unique invariant probability measure. Then for each function* $f : \mathbb{R}^d \to \mathbb{R}$ *measurable and such that* $\mu(|f|) < \infty$,

$$\forall \nu_1 \in \mathcal{P}(\mathbb{R}^d) \quad \mathbb{P}_{\nu_1}\left(\frac{1}{N}\sum_{k=1}^{N} f(X_k) \underset{N\to\infty}{\to} \mu(f)\right) = 1.$$

Since, by Proposition 3, $\mu(V) < \infty$ then

$$\sup_{x\in\mathbb{R}^d} \frac{|f(x)|}{V(x)} < \infty \tag{2.16}$$

is a sufficient condition to ensure that $\mu(|f|) < \infty$.
Before enunciating the Central Limit Theorem for Markov chains, we need to introduce some facts about the Poisson equation. For a fixed f such that $\mu(|f|) < \infty$, a function $F : \mathbb{R}^d \to \mathbb{R}$ measurable and such that $\forall x \in \mathbb{R}^d$, $P|F|(x) < \infty$ is called solution of the Poisson equation if it satisfies

$$\forall x \in \mathbb{R}^d, \quad F(x) - PF(x) = f(x) - \mu(f) \tag{2.17}$$

where μ as above denotes the invariant probability measure associated to P.
If the Markov chain $(X_i)_{i\geq 1}$ is V-uniformly ergodic, since by (2.11) $\sup_{x\in\mathbb{R}^d} \frac{|P^n f(x)-\mu(f)|}{V(x)} \leq \|f\|_V Rr^n$ with $r \in (0,1)$, the series of general term $(P^n f - \mu(f))_{n\in\mathbb{N}}$ converges in the space $\mathcal{V}_V$ equipped with the norm $\|\|_V$ and

$$F = \sum_{n\in\mathbb{N}} (P^n f - \mu(f)) \tag{2.18}$$

is a solution of the Poisson equation (2.17). Moreover any solution can be written as $c + \sum_{n\in\mathbb{N}} (P^n f - \mu(f))$ for a constant $c \in \mathbb{R}$.

As a consequence of (2.12), the following proposition holds.

Proposition 4. *Let us assume that the Markov chain $(X_i)_{i\geq 1}$ is V-uniformly ergodic and let μ be its unique invariant probability measure. Then if f is such that $f^2 \in \mathcal{V}_V$, $F = \sum_{n\in\mathbb{N}}(P^n f - \mu(f))$ converges in $\|\|_{\sqrt{V}}$. F is solution of the Poisson equation (2.17) and satisfies $F^2 \in \mathcal{V}_V$. Moreover for each $n \in \mathbb{N}$*

$$\sup_{x\in\mathbb{R}^d} \frac{|P^n f(x) - \mu(f)|}{\sqrt{V(x)}} \leq D\chi^n \sup_{x\in\mathbb{R}^d} \frac{|f(x)|}{\sqrt{V(x)}}$$

with $\chi \in (0,1)$ and D a finite constant.

We are now ready to enunciate the Central Limit Theorem for Markov chains in the linear case under the V-uniform ergodicity (see Thereom 17.5.4 [43]).

Theorem 22. *Let us assume that the Markov chain $(X_i)_{i\geq 1}$ is V-uniformly ergodic and let μ be its unique invariant probability measure. Then for each function $f : \mathbb{R}^d \to \mathbb{R}$ measurable and such that $f^2 \in \mathcal{V}_V$,*

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \int_{\mathbb{R}^d} f(x)\mu(dx)\right) \xRightarrow{d} \mathcal{N}\left(0, \mu\left(F^2\right) - \mu((PF)^2)\right)$$

with F solution of the Poisson equation

$$F(x) - PF(x) = f(x) - \mu(f) \quad x \in \mathbb{R}^d.$$

We are now going to provide some conditions on P under which a Markov chain is V- uniformly ergodic for some measurable $V : \mathbb{R}^d \to [1, \infty)$.

Example 3. *We say that P verifies the **Lyapunov Condition** if:*

L1 *$\exists v : \mathbb{R}^d \to \mathbb{R}_+$ measurable, $\exists K \in \mathbb{R}_+$, $\exists \gamma \in (0,1)$, $\forall x \in \mathbb{R}^d$, $Pv(x) \leq \gamma v(x) + K$*

L2 *$\exists R > \frac{2K}{1-\gamma}$, $\exists \rho \in (0,1]$, $\forall x, y \in \mathbb{R}^d$ such that $v(x) + v(y) \leq R$, $P(x,\cdot) \wedge P(y,\cdot)(\mathbb{R}^d) \geq \rho$.*

*In [30], Hairer and Mattingly proved that under stronger conditions on P: it verifies **L1** and there exists a constant $q \in (0,1)$ and a probability measure ζ so that*

$$\inf_{x \in C} P(x, \cdot) \geq q\zeta(\cdot) \tag{2.19}$$

with $C = \left\{x \in \mathbb{R}^d : v(x) \leq R\right\}$ for some $R > \frac{2K}{1-\gamma}$, the transition kernel P admits a unique invariant probability measure μ. Moreover

$$\mu(v) \leq \frac{K}{1-\gamma} \tag{2.20}$$

and $\forall \beta > 0, \forall \sigma \in \mathcal{P}\left(\mathbb{R}^d\right), \forall n \in \mathbb{N}$,

$$\sup_{\phi : \|\phi\|_{1+\beta v} \leq 1} |(\sigma P^n)(\phi) - \mu(\phi)| \leq (\chi(\rho, \beta, \gamma, K, R))^n \sup_{\phi : \|\phi\|_{1+\beta v} \leq 1} |\sigma(\phi) - \mu(\phi)| \tag{2.21}$$

where $\chi(\rho, \beta, \gamma, K, R) = (1 - \rho + \beta K) \vee \frac{2+\beta\gamma R + 2\beta K}{2+\beta R} \in (0,1)$ if $\beta \in \left(0, \frac{\rho}{K}\right)$.
*It is possible to adapt the proofs in [30] and obtain the same results by assuming that P verifies **L2** in place of (2.19). In particular, in the proof of Theorem 3.1 [30] the same estimation in the case of x and y such that $v(x) + v(y) \leq R$ holds under the Lyapunov condition.*
Let $\beta \in \left(0, \frac{\rho}{K}\right)$ and let us observe that if we take $\sigma = \delta_x, x \in \mathbb{R}^d$ in (2.21) we obtain

$$\begin{aligned}
\frac{1}{\min(\beta, 1)} \sup_{\phi : \|\phi\|_{1+v} \leq 1} |P^n\phi(x) - \mu(\phi)| &\leq \sup_{\phi : \|\phi\|_{1+\beta v} \leq 1} |P^n\phi(x) - \mu(\phi)| \\
&\leq (\chi(\rho, \beta, \gamma, K, R))^n \sup_{\phi : \|\phi\|_{1+\beta v} \leq 1} |\phi(x) - \mu(\phi)| \\
&\leq (\chi(\rho, \beta, \gamma, K, R))^n \max\left(2 + \frac{\beta K}{1-\gamma}, \beta\right)(1 + v(x))
\end{aligned}$$

where to obtain the last inequality we apply (2.20). Thus

$$\frac{1}{\min(\beta, 1)} \sup_{x \in \mathbb{R}^d} \frac{\sup_{\phi : \|\phi\|_{1+v} \leq 1} |P^n\phi(x) - \mu(\phi)|}{1 + v(x)} \leq (\chi(\rho, \beta, \gamma, K, R))^n \max\left(2 + \frac{\beta K}{1-\gamma}, \beta\right) \underset{n \to \infty}{\longrightarrow} 0.$$

This proves that if P verifies the Lyapunov Condition, then the Markov chain is $1 + v$-uniformly ergodic.

Statement of the theorem

As in the statement of Theorem 17, for a given measure $\mu \in \mathcal{P}_\ell\left(\mathbb{R}^d\right)$, let us consider the Hyphoteses **RU1-3**, together with

RU6 $\exists \alpha \in (\frac{1}{2}, 1], \exists C < \infty, \forall \mu_1, \mu_2 \in B(\mu, r)$

$$\sup_{x \in \mathbb{R}^d} \frac{\left|\frac{\delta U}{\delta m}(\mu_2, x) - \frac{\delta U}{\delta m}(\mu_1, x)\right|}{1 + |x|^{\frac{\ell}{2}}} \leq C\left(\int_{\mathbb{R}^d} (1 + |y|^{\frac{\ell}{2\alpha}}) |\mu_2 - \mu_1|(dy)\right)^\alpha.$$

L3 $C_\ell := \sup_{x \in \mathbb{R}^d} \frac{|x|^\ell}{V(x)} < \infty$.

Remark 23. *Let us observe that by (2.10) and **L3**, we have*

$$\int_{\mathbb{R}^d} |x|^\ell \mu(dx) < \infty. \tag{2.22}$$

We are now ready to provide the Strong Law of Large Numbers and the Central Limit Theorem in this context.

Theorem 24 (LLN for Markov chains)**.** *Let $\ell \geq 0$ and let X_i, $i \geq 1$ a Markov chain with initial law ν_1 and transition kernel P. We assume that $(X_i)_{i\geq 1}$ is V-uniformly ergodic with V satisfying **L3**. Let μ denote its unique invariant probability measure and let us define*

$$\mu_N := \frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}.$$

Then

$$W_\ell\left(\mu_N, \mu\right) \underset{N}{\longrightarrow} 0 \quad a.s.$$

Proof. Let us first prove that $\mu_N \rightharpoonup \mu$ a.s. It can be proved that for probability measures it is possible to test the weak convergence over the continuos functions with compact support (see for instance Corollary 30.9 [4]). Therefore

$$\mathbb{P}\left(\mu_N(f) \underset{N}{\to} \mu(f)\ \forall f \in C_b(\mathbb{R}^d)\right) = \mathbb{P}\left(\mu_N(f) \underset{N}{\to} \mu(f)\ \forall f \in C_c(\mathbb{R}^d)\right)$$

where $C_b(\mathbb{R}^d)$ denotes the set of bounded continuous functions and $C_c(\mathbb{R}^d)$ the set of continuous functions with compact support. Since $C_c(\mathbb{R}^d)$ is separable with respect to the infinity norm, we can apply Theorem 21 to deduce that the right hand-side is equal to 1 and so the almost sure weak convergence of μ_N to μ is proved. If $\ell > 0$, to conclude the proof of the Wasserstein convergence we need to prove the convergence of the ℓth moment. Since (2.22) holds, we can again apply Theorem 21 and deduce that

$$\lim_{N\to\infty}\int_{\mathbb{R}^d}|x|^\ell \mu_N(dx) = \int_{\mathbb{R}^d}|x|^\ell\mu(dx) \quad a.s..$$

□

Theorem 25 (CLT for Markov chains)**.** *Under the same notations of Theorem 24, let us assume that the Markov chain satisfies the V-uniform ergodicity with V satisfying **L3**. Let us moreover assume **RU1-3**, **RU6**.*
Then

$$\sqrt{N}\left(U\left(\mu_N\right) - U\left(\mu\right)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \mu\left(F^2\left(\mu,\cdot\right)\right) - \mu\left(\left(PF\right)^2\left(\mu,\cdot\right)\right)\right)$$

where $F\left(\mu,\cdot\right)$ denotes the unique (up to an additive constant) solution of the Poisson equation

$$F(\mu,x) - PF(\mu,x) = \frac{\delta U}{\delta m}(\mu,x) - \int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,y)\,\mu(dy), \quad x\in\mathbb{R}^d.$$

In what follows, we provide an example of functional satisfying the Hypotheses **RU1-3**, **RU6**.

Example 4. *(U-statistics) Let $\ell > 0$ and $n \in \mathbb{N}\setminus\{0,1\}$. The function on $\mathcal{P}_\ell(\mathbb{R}^d)$ defined by*

$$U(\mu) := \int_{(\mathbb{R}^d)^n}\phi(x_1,\cdots,x_n)\mu(dx_1)\cdots\mu(dx_n) \tag{2.23}$$

where $\phi : (\mathbb{R}^d)^n \to \mathbb{R}$ is a symmetric function such that the above integral is well defined, is called a U-statistics. Hoeffding [34] in 1948 provided a Central Limit Theorem for the U-statistics in the case of independent equidistributed random variables. Several papers then tried to extend the limit theorems for U-statistics established in the i.i.d. setup to the weakly dependent framework and this was mainly done by means of the coupling techniques. For instance Denker and Keller in [17] proved a Central Limit Theorem for particular stationary processes under a Lipschitz condition on

ϕ. Moreover in [1], in the case of stationary ergodic processes, the strong law of large numbers is established.

Bertail and Clémençon [6] developed an alternative to the coupling methodology in the Markovian framework, the regenerative method (see [57]). In particular, they focused on the Harris Markov chains: sample paths of such chains may be divided in i.i.d. regeneration blocks. Thanks to this property they investigated the asymptotic behaviour of U-statistics of positive Harris Markov chains by defining an approximant "regenerative U-statistics" to which the results established in the i.i.d. framework apply. Under some technical assumptions concerning the integrability of ϕ involving the hitting time of the regeneration set, a Central Limit Theorem is provided. We acknowledge that verifying such hypotheses could pose challenges. Let us now study the asymptotic behaviour of the U-statistics as a consequence of our result.
By assuming that ϕ is a continuous functions such that

$$\lim_{|x_1|\to\infty} \sup_{(x_2,\cdots,x_n)\in(\mathbb{R}^d)^{n-1}} \frac{|\phi(x_1,x_2,\cdots,x_n)|}{\prod_{i=1}^n (1+|x_i|^{\frac{\ell}{2}})} = 0, \tag{2.24}$$

*we are now going to verify that (2.23) satisfies **RU2-3** and **RU6**. Therefore, by considering a V-uniformly ergodic Markov chain with V satisfying **L3**, a Central Limit Theorem holds.*
It is possible to prove (see Example 2.7 [36]) that the linear functional derivative of (2.23) is defined for each $\mu \in W_\ell(\mathbb{R}^d)$ and it is given, by choosing the constant equal to zero, by

$$\frac{\delta U}{\delta m}(\mu, x_1) = n \int_{(\mathbb{R}^d)^{n-1}} (\phi(x_1,x_2,\cdots,x_n) - \phi(0,x_2,\cdots,x_n))\mu(dx_2)\cdots\mu(dx_n),\ x_1 \in \mathbb{R}^d.$$

By the symmetry of ϕ and (2.24), $\forall \epsilon > 0\ \exists M_\epsilon$ such that for $|x_i| > M_\epsilon$

$$\sup_{(x_1,,\cdots,x_{i-1},x_{i+1},\cdots,x_n)\in(\mathbb{R}^d)^{n-1}} \frac{|\phi(x_1,x_2,\cdots,x_n)|}{\prod_{j=1}^n (1+|x_j|^{\frac{\ell}{2}})} \le \epsilon. \tag{2.25}$$

Moreover by (2.25) and the continuity of ϕ, $\forall\ (x_1,\cdots,x_n) \in (\mathbb{R}^d)^n$ one has

$$|\phi(x_1,x_2,\cdots,x_n)| = |\phi(x_1,x_2,\cdots,x_n)|\, 1_{\|x\|_\infty > M_1} + |\phi(x_1,x_2,\cdots,x_n)|\, 1_{\|x\|_\infty \le M_1} \tag{2.26}$$

$$\le \prod_{i=1}^n (1+|x_i|^{\frac{\ell}{2}}) + C \tag{2.27}$$

for some positive constant $C < \infty$.

Let us fix $\mu \in \mathcal{P}_\ell(\mathbb{R}^d)$ and $r > 0$. We recall that $B(\mu, r)$ denotes the ball centered at μ with radius r for the metric W_ℓ.

RU2. *Let $(\tilde{\mu}, x_1) \in B(\mu, r) \times \mathbb{R}^d$. Using the above estimation we obtain*

$$\begin{aligned}
\left|\frac{\delta U}{\delta m}(\tilde{\mu}, x_1)\right| &\le n \int_{(\mathbb{R}^d)^{n-1}} (|\phi(x_1,x_2,\cdots,x_n)| + |\phi(0,x_2,\cdots,x_n)|)\, \tilde{\mu}(dx_2)\cdots\tilde{\mu}(dx_n) \\
&\le n \int_{(\mathbb{R}^d)^{n-1}} \left(\prod_{i=1}^n (1+|x_i|^{\frac{\ell}{2}}) + \prod_{i=2}^n (1+|x_i|^{\frac{\ell}{2}})\right) \tilde{\mu}(dx_2)\cdots\tilde{\mu}(dx_n) + 2Cn \\
&\le D(1+|x_1|^{\frac{\ell}{2}})
\end{aligned}$$

for some positive constant $D < \infty$.

RU3. *If $W_\ell(\tilde{\mu},\mu)$ goes to 0, let $\pi \in \mathcal{P}(\mathbb{R}^{2d})$ with $\pi\left(\cdot \times \mathbb{R}^d\right) = \tilde{\mu}(\cdot)$, $\pi\left(\mathbb{R}^d \times \cdot\right) = \mu(\cdot)$. One has*

$$\sup_{x_1\in\mathbb{R}^d} \frac{|\frac{\delta U}{\delta m}(\tilde{\mu},x_1) - \frac{\delta U}{\delta m}(\mu,x_1)|}{1+|x_1|^{\frac{\ell}{2}}} \tag{2.28}$$

$$\leq n \sup_{x_1\in\mathbb{R}^d} \frac{\int_{(\mathbb{R}^d)^{2(n-1)}} |\phi(x_1,x_2,\cdots,x_n) - \phi(x_1,s_2,\cdots,s_n)| \prod_{i=2}^n \pi(dx_i,ds_i)}{1+|x_1|^{\frac{\ell}{2}}} \tag{2.29}$$

$$+ n \int_{(\mathbb{R}^d)^{2(n-1)}} |\phi(0,x_2,\cdots,x_n) - \phi(0,s_2,\cdots,s_n)| \prod_{i=2}^n \pi(dx_i,ds_i) \tag{2.30}$$

$$\leq 2n \sup_{x_1\in\mathbb{R}^d} \frac{\int_{(\mathbb{R}^d)^{2(n-1)}} |\phi(x_1,x_2,\cdots,x_n) - \phi(x_1,s_2,\cdots,s_n)| \prod_{i=2}^n \pi(dx_i,ds_i)}{1+|x_1|^{\frac{\ell}{2}}}. \tag{2.31}$$

Let $\epsilon > 0$. By observing that for $|y| \geq \frac{1}{\epsilon}$, $\frac{1}{|y|} \leq \epsilon$, let us define $\tilde{M}_\epsilon := \max(\frac{1}{\epsilon}, M_\epsilon)$. The function ϕ is uniformly continuous on $B = \left\{x \in (\mathbb{R}^d)^n : \|x\|_\infty \leq \tilde{M}_\epsilon\right\}$. Therefore $\exists \eta_{\epsilon,\tilde{M}_\epsilon} > 0$ such that for $x, \tilde{x} \in B$ satisfying $\max_{2\leq i\leq n} |x_i - \tilde{x}_i| \leq \eta_{\epsilon,\tilde{M}_\epsilon}$,

$$|\phi(x_1,x_2,\cdots,x_n) - \phi(x_1,\tilde{x}_2,\cdots,\tilde{x}_n)| \leq \epsilon. \tag{2.32}$$

If we denote the vectors $(x_2,\cdots,x_n)$ and $(s_2,\cdots,s_n)$ respectively by $x_{2:n}$ and $s_{2:n}$, it is possible to rewrite the integral in (2.31) in the following way ($\star$)

$$\begin{aligned}
&\int |\phi(x_1,x_{2:n}) - \phi(x_1,s_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|x_{2:n}-s_{2:n}\|_\infty \leq \eta_{\epsilon,\tilde{M}_\epsilon}} \prod_{i=2}^n \pi(dx_i,ds_i) \\
&+ 2\int |\phi(x_1,x_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|x_{2:n}-s_{2:n}\|_\infty > \eta_{\epsilon,\tilde{M}_\epsilon}} \prod_{i=2}^n \pi(dx_i,ds_i) \\
&+ \int |\phi(x_1,x_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i,ds_i) \\
&+ \int |\phi(x_1,s_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i,ds_i) \\
&+ \int |\phi(x_1,x_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i,ds_i) \\
&+ \int |\phi(x_1,s_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i,ds_i).
\end{aligned}$$

Let us now study one by one the terms appearing in the above expression ($\star$).

1. *$\int |\phi(x_1,x_{2:n}) - \phi(x_1,s_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|x_{2:n}-s_{2:n}\|_\infty \leq \eta_{\epsilon,\tilde{M}_\epsilon}} \prod_{i=2}^n \pi(dx_i,ds_i) \leq \epsilon$ by (2.32).*
2. *By the continuity of ϕ and Markov's inequality*

$$\begin{aligned}
&2\int |\phi(x_1,x_{2:n})|\, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|x_{2:n}-s_{2:n}\|_\infty > \eta_{\epsilon,\tilde{M}_\epsilon}} \prod_{i=2}^n \pi(dx_i,ds_i) \\
&\leq 2 \cdot \sup_{\|x\|_\infty \leq \tilde{M}_\epsilon} |\phi(x)| \cdot \sum_{i=2}^n \int 1_{|x_i-s_i|>\eta_{\epsilon,\tilde{M}_\epsilon}} \pi(dx_i,ds_i) \\
&\leq \frac{2}{\eta^\ell_{\epsilon,\tilde{M}_\epsilon}} \cdot \sup_{\|x\|_\infty \leq \tilde{M}_\epsilon} |\phi(x)| \cdot \sum_{i=2}^n \int |x_i - s_i|^\ell \pi(dx_i,ds_i).
\end{aligned}$$

3. *By (2.27), Markov's inequality, Cauchy-Schwarz inequality and the fact that* $\frac{1}{\tilde{M}_\epsilon} \leq \epsilon$

$$\begin{aligned}
&\int |\phi(x_1, x_{2:n})| \, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i, ds_i) \\
&\leq \sum_{i=2}^n \int (C + \prod_{j=1}^n (1 + |x_j|^{\frac{\ell}{2}})) 1_{|s_i| > \tilde{M}_\epsilon} \prod_{j=2}^n \pi(dx_j, ds_j) \\
&\leq \epsilon^\ell (n-1) C \int |s|^\ell \mu(ds) + \epsilon^{\frac{\ell}{2}} (n-1)(1 + |x_1|^{\frac{\ell}{2}}) \left(\int (1 + |x|^{\frac{\ell}{2}})^2 \tilde{\mu}(dx) \right)^{\frac{n-1}{2}} \left(\int |s|^\ell \mu(ds) \right)^{\frac{1}{2}}.
\end{aligned}$$

4. *Similarly to the previous point*

$$\begin{aligned}
&\int |\phi(x_1, s_{2:n})| \, 1_{\|(x_1,x_{2:n})\|_\infty \leq \tilde{M}_\epsilon} 1_{\|(x_1,s_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i, ds_i) \\
&\leq \sum_{i=2}^n \int (C + (1 + |x_1|^{\frac{\ell}{2}}) \prod_{j=2}^n (1 + |s_j|^{\frac{\ell}{2}})) 1_{|s_i| > \tilde{M}_\epsilon} \prod_{j=2}^n \pi(dx_j, ds_j) \\
&\leq \epsilon^\ell (n-1) C \int |s|^\ell \mu(ds) + \epsilon^{\frac{\ell}{2}} (n-1)(1 + |x_1|^{\frac{\ell}{2}}) \left(\int (1 + |s|^{\frac{\ell}{2}})^2 \mu(ds) \right)^{\frac{n-1}{2}} \left(\int |s|^\ell \mu(ds) \right)^{\frac{1}{2}}
\end{aligned}$$

5. *Finally by (2.25)*

$$\begin{aligned}
&\int |\phi(x_1, x_{2:n})| \, 1_{\|(x_1,x_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i, ds_i) + \int |\phi(x_1, s_{2:n})| \, 1_{\|(x_1,x_{2:n})\|_\infty > \tilde{M}_\epsilon} \prod_{i=2}^n \pi(dx_i, ds_i) \\
&\leq \epsilon (1 + |x_1|^{\frac{\ell}{2}}) \left(\int \left(1 + |x|^{\frac{\ell}{2}}\right) \tilde{\mu}(dx) \right)^{n-1} + \epsilon (1 + |x_1|^{\frac{\ell}{2}}) \left(\int \left(1 + |s|^{\frac{\ell}{2}}\right) \mu(ds) \right)^{n-1}.
\end{aligned}$$

Therefore plugging all these estimations in (2.31), choosing π *as the optimal* W_ℓ *coupling between* $\tilde{\mu}$ *and* μ*, we obtain*

$$\begin{aligned}
&\sup_{x_1 \in \mathbb{R}^d} \frac{|\frac{\delta U}{\delta m}(\tilde{\mu}, x_1) - \frac{\delta U}{\delta m}(\mu, x_1)|}{1 + |x_1|^{\frac{\ell}{2}}} \\
&\leq D \left(\sup_{\|x\|_\infty \leq \tilde{M}_\epsilon} |\phi(x)| \cdot \frac{W_\ell^{\ell \wedge 1}(\tilde{\mu}, \mu)}{\eta^\ell_{\epsilon, \tilde{M}_\epsilon}} + (\epsilon + \epsilon^{\frac{\ell}{2}} + \epsilon^\ell) \left(\left(\int (1 + |x|^{\frac{\ell}{2}})^2 \tilde{\mu}(dx) \right)^{\frac{n-1}{2}} + 1 \right) \right)
\end{aligned}$$

for a positive constant D *depending neither on* ϵ *nor on* $\tilde{\mu}$*. For fixed* ϵ*, let* $W_\ell(\tilde{\mu}, \mu)$ *converge to* 0*. Then* $\int |x|^\ell \tilde{\mu}(dx)$ *converges to* $\int |s|^\ell \mu(ds)$ *and* $\int (1 + |x|^{\frac{\ell}{2}})^2 \tilde{\mu}(dx)$ *converges to* $\int (1 + |s|^{\frac{\ell}{2}})^2 \mu(ds)$*. It is then possible to conclude that the left-hand side goes to* 0 *by letting* ϵ *go to* 0*.*

RU6. *Let* $\mu_1, \mu_2 \in B(\mu, r)$*. By (2.27) we have*

$$\begin{aligned}
&\sup_{x_1 \in \mathbb{R}^d} \frac{\left| \frac{\delta U}{\delta m}(\mu_2, x_1) - \frac{\delta U}{\delta m}(\mu_1, x_1) \right|}{1 + |x_1|^{\frac{\ell}{2}}} \\
&\leq 2n \sup_{x_1 \in \mathbb{R}^d} \frac{\int_{(\mathbb{R}^d)^{n-1}} |\phi(x_1, x_2, \cdots, x_n)| \, |\mu_2(dx_2) \cdots \mu_2(dx_n) - \mu_1(dx_2) \cdots \mu_1(dx_n)|}{1 + |x_1|^{\frac{\ell}{2}}} \\
&\leq 2n \sup_{x_1 \in \mathbb{R}^d} \frac{\sum_{k=2}^n \int_{(\mathbb{R}^d)^{n-1}} \left(C + \prod_{i=1}^n (1 + |x_i|^{\frac{\ell}{2}}) \right) \prod_{j=k+1}^n \mu_2(dx_j) \, |\mu_2 - \mu_1| \, (dx_k) \prod_{j=2}^{k-1} \mu_1(dx_j)}{1 + |x_1|^{\frac{\ell}{2}}} \\
&\leq \tilde{C} \int_{\mathbb{R}^d} (1 + |x|^{\frac{\ell}{2}}) \, |\mu_2 - \mu_1| \, (dx)
\end{aligned}$$

for a constant $\tilde{C} < \infty$ not depending on (μ_1, μ_2).

2.2 Proof of the Results

In this section we will provide the proofs of Theorem 17 and Theorem 25. We will start by proving the Markov chains case since it is more complex. Afterwards, we will present the proof for the independent and non-equidistributed case, avoiding repetition of common parts whenever possible.

2.2.1 Markov Chains

In the proof of the theorem we will need that the integral of V respect to ν_1 is finite. The next Lemma permits to suppose it in the proof.

Lemma 3. *If Theorem 25 holds for each $\nu_1 \in \mathcal{P}(\mathbb{R}^d)$ such that $\nu_1(V) < \infty$, then it holds for all $\nu_1 \in \mathcal{P}(\mathbb{R}^d)$.*

Proof. If $\nu_1(V) = \infty$, for $K > \inf_{x\in\mathbb{R}^d} V(x)$ let us consider $x_0 \in \mathbb{R}^d$ such that $V(x_0) \leq K$ and define $\phi_K(x) = x1_{\{V(x)\leq K\}} + x_0 1_{\{V(x)>K\}}$.
If we now consider the measure $\nu_1^K(A) := \nu_1(\phi_K^{-1}(A))$ for $A \in \mathcal{B}(\mathbb{R}^d)$, then

$$\begin{aligned}\nu_1^K(V) &= \int_{\mathbb{R}^d} V(x)\nu_1^K(dx) = \int_{\mathbb{R}^d} V(\phi_K(x))\nu_1(dx)\\ &= \int_{\mathbb{R}^d} \left(V(x)1_{\{V(x)\leq K\}} + V(x_0)1_{\{V(x)>K\}}\right)\nu_1(dx) \leq K.\end{aligned}$$

Moreover

$$\begin{aligned}d_{TV}\left(\nu_1^K, \nu_1\right) &= \sup_{\psi:\|\psi\|_\infty\leq\frac{1}{2}} \left|\int_{\mathbb{R}^d} \psi(x)\nu_1^K(dx) - \int_{\mathbb{R}^d} \psi(x)\nu_1(dx)\right|\\ &= \sup_{\psi:\|\psi\|_\infty\leq\frac{1}{2}} \left|\int_{\mathbb{R}^d} (\psi(x_0) - \psi(x))\, 1_{\{V(x)>K\}}\nu_1(dx)\right| \leq \nu_1\left(\{x : V(x) > K\}\right).\end{aligned}$$

Let us consider now the Markov chain with initial law ν_1^K and transition kernel P on $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$ and denote by $\mathbb{P}_{\nu_1^K}$ the law of the process ($\mathbb{P}_{\nu_1}$ will denote the law of the original Markov chain). Since the initial law of a Markov chain, together with its transition kernel, determines the law of the entire process, one has

$$\begin{aligned}d_{TV}\left(\mathbb{P}_{\nu_1}, \mathbb{P}_{\nu_1^K}\right) &= \sup_{F\in\mathcal{B}\left(\mathbb{R}^d\right)^{\otimes\mathbb{N}}:\|F\|_\infty\leq\frac{1}{2}} \left|\int_{\left(\mathbb{R}^d\right)^{\otimes\mathbb{N}}} F(x)\mathbb{P}_{\nu_1}(dx) - \int_{\left(\mathbb{R}^d\right)^{\otimes\mathbb{N}}} F(x)\mathbb{P}_{\nu_1^K}(dx)\right|\\ &= \sup_{F\in\mathcal{B}\left(\mathbb{R}^d\right)^{\otimes\mathbb{N}}:\|F\|_\infty\leq\frac{1}{2}} \left|\int_{\mathbb{R}^d} \mathbb{P}_{\delta_x}(F)\nu_1(dx) - \int_{\mathbb{R}^d} \mathbb{P}_{\delta_x}(F)\nu_1^K(dx)\right|\\ &= \sup_{F\in\mathcal{B}\left(\mathbb{R}^d\right)^{\otimes\mathbb{N}}:\|F\|_\infty\leq\frac{1}{2}} \left|\int_{\mathbb{R}^d} f_F(x)\nu_1(dx) - \int_{\mathbb{R}^d} f_F(x)\nu_1^K(dx)\right| \leq d_{TV}\left(\nu_1, \nu_1^K\right)\end{aligned}$$

where we introduce the notation $f_F(x) = \mathbb{P}_{\delta_x}(F)$ for each $F \in \mathcal{B}\left(\mathbb{R}^d\right)^{\otimes\mathbb{N}}$. Since it is straightforward to demonstrate that $d_{TV}\left(\mathbb{P}_{\nu_1}, \mathbb{P}_{\nu_1^K}\right) \geq d_{TV}\left(\nu_1, \nu_1^K\right)$, the equality $d_{TV}\left(\mathbb{P}_{\nu_1}, \mathbb{P}_{\nu_1^K}\right) = d_{TV}\left(\nu_1, \nu_1^K\right)$ holds. Therefore if we are able to prove that for all bounded, continuous functions f

$$\lim_{N\to\infty} \mathbb{E}_{\nu_1^K}\left(f\left(\sqrt{N}\left(U(\mu_N) - U(\mu)\right)\right)\right) = \mathbb{E}\left(f(\mathcal{G})\right)$$

where $\mathcal{G}$ is a Gaussian random variable not depending on K, then

$$\begin{aligned}
&\left|\mathbb{E}_{\nu_1}\left(f\left(\sqrt{N}\left(U(\mu_N)-U(\mu)\right)\right)\right)-\mathbb{E}\left(f(\mathcal{G})\right)\right| \\
&\quad\leq\left|\mathbb{E}_{\nu_1}\left(f\left(\sqrt{N}\left(U(\mu_N)-U(\mu)\right)\right)\right)-\mathbb{E}_{\nu_1^K}\left(f\left(\sqrt{N}\left(U(\mu_N)-U(\mu)\right)\right)\right)\right| \\
&\qquad+\left|\mathbb{E}_{\nu_1^K}\left(f\left(\sqrt{N}\left(U(\mu_N)-U(\mu)\right)\right)\right)-\mathbb{E}\left(f(\mathcal{G})\right)\right| \\
&\quad\leq 2\left\|f\right\|_\infty d_{TV}\left(\mathbb{P}_{\nu_1},\mathbb{P}_{\nu_1^K}\right)+\left|\mathbb{E}_{\nu_1^K}\left(f\left(\sqrt{N}\left(U(\mu_N)-U(\mu)\right)\right)\right)-\mathbb{E}\left(f(\mathcal{G})\right)\right| \\
&\quad\leq 2\left\|f\right\|_\infty d_{TV}\left(\nu_1,\nu_1^K\right)+\left|\mathbb{E}_{\nu_1^K}\left(f\left(\sqrt{N}\left(U(\mu_N)-U(\mu)\right)\right)\right)-\mathbb{E}\left(f(\mathcal{G})\right)\right|.
\end{aligned}$$

Since $d_{TV}\left(\nu_1,\nu_1^K\right)\leq\nu_1\left(\{x:V(x)>K\}\right)$, we can first take the superior limit over $N\to\infty$ so to obtain that

$$\limsup_{N\to\infty}\left|\mathbb{E}_{\nu_1}\left(f\left(\sqrt{N}\left(U(\mu_N)-U(\mu)\right)\right)\right)-\mathbb{E}\left(f(\mathcal{G})\right)\right|\leq 2\left\|f\right\|_\infty\nu_1\left(\{x:V(x)>K\}\right).$$

Then we can conclude that the left hand-side is equal to 0 by letting K to infinity. □

The proof of the theorem relies on the Poisson equation whose explicit solution and norm estimations are studied in the following Lemma.

Lemma 4. *Under the assumptions of the previous theorem, for a given $m\in B(\mu,r)$ let us consider the following Poisson equation*

$$F(m,x)-PF(m,x)=\frac{\delta U}{\delta m}(m,x)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(m,y)\,\mu(dy),\quad x\in\mathbb{R}^d. \tag{2.33}$$

Then (2.33) admits a solution given by $F(m,\cdot)=\sum_{n\in\mathbb{N}}(P^n\frac{\delta U}{\delta m}(m,\cdot)-\mu(\frac{\delta U}{\delta m}(m,\cdot)))$. Moreover

- *For $m\in B(\mu,r)$*
$$\left\|F(m,\cdot)\right\|_{\sqrt{V}}\leq\bar{C}_{\frac{1}{2}} \tag{2.34}$$
- *For $m_1,m_2\in B(\mu,r)$*
$$\left\|F(m_1,\cdot)-F(m_2,\cdot)\right\|_{\sqrt{V}}\leq\bar{C}_{\frac{1}{2}}\sup_{x\in\mathbb{R}^d}\frac{\left|\frac{\delta U}{\delta m}(m_1,x)-\frac{\delta U}{\delta m}(m_2,x)\right|}{\sqrt{V(x)}} \tag{2.35}$$

  and so for $\alpha\in(\frac{1}{2},1]$
$$\left\|F(m_1,\cdot)-F(m_2,\cdot)\right\|_{\sqrt{V}}\leq\bar{C}_{\frac{1}{2}}\left(\int_{\mathbb{R}^d}(1+|y|^{\frac{\ell}{2\alpha}})|m_2-m_1|(dy)\right)^\alpha \tag{2.36}$$

  with $\bar{C}_{\frac{1}{2}}$ a finite constant not depending on m, m_1 and m_2.

Proof. By Proposition 4, to ensure the existence of the solution, it is sufficient to check that $\left(\frac{\delta U}{\delta m}(m,\cdot)\right)^2\in\mathcal{V}_V$. By Hypothesis **RU2** and **L3**

$$\sup_{x\in\mathbb{R}^d}\frac{\left(\frac{\delta U}{\delta m}(m,x)\right)^2}{V(x)}\leq 2C^2\sup_{x\in\mathbb{R}^d}\frac{1+|x|^\ell}{V(x)}\leq 2C^2\left(1+C_\ell\right)<\infty.$$

Let us now estimate the norm $\|\|\|_{\sqrt{V}}$ of $F(m,\cdot)$. By Proposition 4, one has

$$\begin{aligned}\|F(m,\cdot)\|_{\sqrt{V}} &= \lim_{n\to\infty}\left\|\sum_{k=0}^{n} P^k \frac{\delta U}{\delta m}(m,\cdot) - \mu(\frac{\delta U}{\delta m}(m,\cdot))\right\|_{\sqrt{V}} \\ &\leq \sum_{k=0}^{\infty}\left\|P^k \frac{\delta U}{\delta m}(m,\cdot) - \mu(\frac{\delta U}{\delta m}(m,\cdot))\right\|_{\sqrt{V}}.\end{aligned}$$

Then again by Proposition 4 and Hypothesis **RU2**

$$\begin{aligned}\left\|P^k \frac{\delta U}{\delta m}(m,\cdot) - \mu(\frac{\delta U}{\delta m}(m,\cdot))\right\|_{\sqrt{V}} &= \sup_{x\in\mathbb{R}^d}\frac{\left|P^k \frac{\delta U}{\delta m}(m,x) - \mu(\frac{\delta U}{\delta m}(m,\cdot))\right|}{\sqrt{V(x)}} \leq D\chi^k \sup_{x\in\mathbb{R}^d}\frac{\left|\frac{\delta U}{\delta m}(m,x)\right|}{\sqrt{V(x)}} \\ &\leq DC\chi^k\left(1+\sup_{x\in\mathbb{R}^d}\frac{|x|^{\frac{\ell}{2}}}{\sqrt{V(x)}}\right) \leq DC\chi^k\left(1+\sqrt{C_\ell}\right)\end{aligned}$$

where to obtain the last inequality we used **L3**.
Therefore we have obtained that

$$\|F(m,\cdot)\|_{\sqrt{V}} \leq \frac{1}{1-\chi}DC\left(1+\sqrt{C_\ell}\right).$$

Finally let $m_1, m_2 \in B(\mu, r)$ and proceeding as above, by Proposition 4 we have

$$\begin{aligned}&\|F(m_1,\cdot) - F(m_2,\cdot)\|_{\sqrt{V}} \\ &\leq \sum_{k=0}^{\infty}\left\|P^k\left(\frac{\delta U}{\delta m}(m_1,\cdot) - \frac{\delta U}{\delta m}(m_2,\cdot)\right) - \mu\left(\frac{\delta U}{\delta m}(m_1,\cdot) - \frac{\delta U}{\delta m}(m_2,\cdot)\right)\right\|_{\sqrt{V}}\end{aligned}$$

By Proposition 4 and Hypothesis **RU6**

$$\begin{aligned}&\left\|P^k\left(\frac{\delta U}{\delta m}(m_1,\cdot) - \frac{\delta U}{\delta m}(m_2,\cdot)\right) - \mu\left(\frac{\delta U}{\delta m}(m_1,\cdot) - \frac{\delta U}{\delta m}(m_2,\cdot)\right)\right\|_{\sqrt{V}} \\ &= \sup_{x\in\mathbb{R}^d}\frac{\left|P^k\left(\frac{\delta U}{\delta m}(m_1,x) - \frac{\delta U}{\delta m}(m_2,x)\right) - \mu\left(\frac{\delta U}{\delta m}(m_1,\cdot) - \frac{\delta U}{\delta m}(m_2,\cdot)\right)\right|}{\sqrt{V(x)}} \\ &\leq D\chi^k \sup_{x\in\mathbb{R}^d}\frac{\left|\frac{\delta U}{\delta m}(m_1,x) - \frac{\delta U}{\delta m}(m_2,x)\right|}{\sqrt{V(x)}} \\ &\leq DC\chi^k \sup_{x\in\mathbb{R}^d}\frac{1+|x|^{\frac{\ell}{2}}}{\sqrt{V(x)}}\left(\int_{\mathbb{R}^d}(1+|y|^{\frac{\ell}{2\alpha}})|m_2-m_1|(dy)\right)^{\alpha} \\ &= DC\chi^k\left(1+\sqrt{C_\ell}\right)\left(\int_{\mathbb{R}^d}(1+|y|^{\frac{\ell}{2\alpha}})|m_2-m_1|(dy)\right)^{\alpha}\end{aligned}$$

where we used **L3** to obtain the last equality. Therefore we have

$$\|F(m_1,\cdot) - F(m_2,\cdot)\|_{\sqrt{V},1} \leq DC\frac{1}{1-\chi}\left(1+\sqrt{C_\ell}\right)\left(\int_{\mathbb{R}^d}(1+|y|^{\frac{\ell}{2\alpha}})|m_2-m_1|(dy)\right)^{\alpha}.$$

□

Proof of Theorem 25. **FIRST STEP**
Let us preliminary recall that by Theorem 24, $W_\ell(\mu_N,\mu) \underset{N}{\longrightarrow} 0$ a.s. To study the limit distribution of $\sqrt{N}(U(\mu_N) - U(\mu))$, let us define for $i = 1,\cdots,N$ and $s \in [0,1]$

$$\mu_N^{i,s} = \frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j} + \frac{s}{N}\delta_{X_i} + \left(1 + \frac{1-i-s}{N}\right)\mu.$$

We have $\mu_N^{N,1} = \mu_N$ and $\mu_N^{1,0} = \mu$. Moreover, since for $i = 1, \cdots, N-1$

$$\mu_N^{i,1} = \mu_N^{i+1,0},$$

one has

$$\sqrt{N}\left(U\left(\mu_N\right) - U\left(\mu\right)\right) = \sqrt{N}\sum_{i=1}^{N}\left(U(\mu_N^{i,1}) - U(\mu_N^{i,0})\right).$$

Let us now show that almost surely $\mu_N^{i,s}$ converges uniformly to μ with respect to the distance W_ℓ that is

$$\max_{1\le i\le N}\sup_{s\in[0,1]} W_\ell(\mu_N^{i,s},\mu) \underset{N\to\infty}{\to} 0 \quad a.s.. \tag{2.37}$$

For $i = 1, \cdots, N$ and $s \in [0,1]$

$$\mu_N^{i,s} = s\mu_N^{i,1} + (1-s)\mu_N^{i-1,1}$$

under the convention $\mu_N^{0,1} = \mu_N^{1,0}$. Let $\pi \in \Pi(\mu_N^{i,1},\mu)$ and $\tilde{\pi} \in \Pi(\mu_N^{i-1,1},\mu)$ where $\Pi(\mu_1,\mu_2) = \left\{\mu \in \mathcal{P}(\mathbb{R}^{2d}) : \mu(\cdot\times\mathbb{R}^d) = \mu_1(\cdot), \mu(\mathbb{R}^d\times\cdot) = \mu_2(\cdot)\right\}$ for $\mu_1,\mu_2 \in \mathcal{P}_\ell(\mathbb{R}^d)$ and define

$$\bar{\pi}\left(dx,dy\right) = s\pi\left(dx,dy\right) + (1-s)\tilde{\pi}\left(dx,dy\right).$$

Then

$$\begin{aligned}
\bar{\pi}\left(dx,\mathbb{R}^d\right) &= s\pi\left(dx,\mathbb{R}^d\right) + (1-s)\tilde{\pi}\left(dx,\mathbb{R}^d\right) = s\mu_N^{i,1} + (1-s)\mu_N^{i-1,1} = \mu_N^{i,s}\\
\bar{\pi}\left(\mathbb{R}^d,dy\right) &= s\pi\left(\mathbb{R}^d,dy\right) + (1-s)\tilde{\pi}\left(\mathbb{R}^d,dy\right) = s\mu + (1-s)\mu = \mu.
\end{aligned}$$

By supposing that $\ell > 0$, we have:

$$W_\ell^{\ell\vee 1}\left(\mu_N^{i,s},\mu\right) \le \int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\bar{\pi}\left(dx,dy\right) = s\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\pi\left(dx,dy\right) + (1-s)\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\tilde{\pi}\left(dx,dy\right). \tag{2.38}$$

Taking the infimum over π and $\tilde{\pi}$, we conclude that

$$W_\ell^{\ell\vee 1}\left(\mu_N^{i,s},\mu\right) \le sW_\ell^{\ell\vee 1}\left(\mu_N^{i,1},\mu\right) + (1-s)W_\ell^{\ell\vee 1}\left(\mu_N^{i-1,1},\mu\right) \le W_\ell^{\ell\vee 1}\left(\mu_N^{i,1},\mu\right) \vee W_\ell^{\ell\vee 1}\left(\mu_N^{i-1,1},\mu\right)$$

and so

$$\max_{1\le i\le N}\sup_{s\in[0,1]} W_\ell\left(\mu_N^{i,s},\mu\right) = \max_{0\le i\le N} W_\ell\left(\mu_N^{i,1},\mu\right). \tag{2.39}$$

Now since for $i = 1, \cdots, N$

$$\mu_N^{i,1} = \frac{1}{N}\sum_{j=1}^{i}\delta_{X_j} + \left(1-\frac{i}{N}\right)\mu = \frac{i}{N}\mu_i + \left(1-\frac{i}{N}\right)\mu,$$

let $\gamma \in \Pi(\mu_i,\mu)$ and define $\tilde{\gamma}(dx,dy) = \frac{i}{N}\gamma(dx,dy) + \left(1-\frac{i}{N}\right)\mu\left(dx\right)\delta_x(dy)$. Then

$$\tilde{\gamma}(dx,\mathbb{R}^d)=\frac{i}{N}\gamma(dx,\mathbb{R}^d)+\left(1-\frac{i}{N}\right)\mu\left(dx\right)\delta_x(\mathbb{R}^d)=\frac{i}{N}\mu_i(dx)+\left(1-\frac{i}{N}\right)\mu\left(dx\right)=\mu_N^{i,1}(dx)$$
$$\tilde{\gamma}(\mathbb{R}^d,dy)=\frac{i}{N}\gamma(\mathbb{R}^d,dy)+\left(1-\frac{i}{N}\right)\mu\left(dy\right)\delta_y(\mathbb{R}^d)=\frac{i}{N}\mu(dy)+\left(1-\frac{i}{N}\right)\mu(dy)=\mu(dy).$$

Therefore

$$W_\ell^{\ell\vee 1}\left(\mu_N^{i,1},\mu\right)\leq\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\tilde{\gamma}(dx,dy)=\frac{i}{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\gamma(dx,dy). \tag{2.40}$$

Taking the infimum over γ, one has

$$W_\ell\left(\mu_N^{i,1},\mu\right)\leq\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}}W_\ell\left(\mu_i,\mu\right).$$

Finally let $\alpha\in(0,1)$ so that

$$\begin{aligned}\max_{0\leq i\leq N}W_\ell\left(\mu_N^{i,1},\mu\right)&\leq\max_{1\leq i\leq N}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}}W_\ell\left(\mu_i,\mu\right)\\&=\max_{1\leq i\leq\llcorner\alpha N\lrcorner}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}}W_\ell\left(\mu_i,\mu\right)+\max_{\llcorner\alpha N\lrcorner<i\leq N}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}}W_\ell\left(\mu_i,\mu\right)\\&\leq\alpha^{\frac{1}{\ell\vee 1}}\sup_i W_\ell\left(\mu_i,\mu\right)+\max_{\llcorner\alpha N\lrcorner<i\leq N}W_\ell\left(\mu_i,\mu\right).\end{aligned}$$

Since we have proved that $\lim_{N\to\infty}W_\ell(\mu_N,\mu)=0$, for fixed α the last term goes to 0 as N goes to infinity while the first one is arbitrarily small for α small and so (2.37) is proved.

By replacing $|x-y|^\ell$ with $1\wedge|x-y|$ in (2.38) and (2.40), we can do exactly the same for $\ell=0$.

SECOND STEP Let us define

$$I_N=\min\left\{1\leq i\leq N:\exists s\in[0,1]:W_\ell\left(\mu_N^{i,s},\mu\right)\geq r\right\}\wedge(N+1)$$

and let us introduce the filtration $(\mathcal{F}_i=\sigma\left(X_1,\cdots,X_i\right))_{i\geq 1}$ for which I_N is a stopping time. According to the first step, under the convention $\min\emptyset=+\infty$, I_N is $a.s.$ equal to $N+1$ for each $N\geq N^*$ for a random variable N^* taking integers values. This stopping time allows to introduce in the proof the linear functional derivative associated to U since by Hypothesis **RU1** it is well defined in the ball of radius r and center μ.

For $N\geq N^*$, by Lemma 2, we have

$$U\left(\mu_N\right)-U\left(\mu\right)=\sum_{i=1}^N\left(U(\mu_N^{i,1})-U(\mu_N^{i,0})\right)=\sum_{i=1}^N\int_0^1 ds\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu_N^{i,s},x)\frac{(\delta_{X_i}-\mu)(dx)}{N}.$$

Setting

$$Q_N:=\frac{1}{N}\sum_{i=1}^N\left(\frac{\delta U}{\delta m}(\mu_N^{i\wedge I_N,0},X_i)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu_N^{i\wedge I_N,0},x)\mu(dx)\right),$$

we deduce that for $N\geq N^*$, $U\left(\mu_N\right)-U\left(\mu\right)-Q_N$ coincides with

$$R_N=\frac{1_{N\geq N^*}}{N}\sum_{i=1}^N\int_0^1 ds\int_{\mathbb{R}^d}\left(\frac{\delta U}{\delta m}(\mu_N^{i,s},x)-\frac{\delta U}{\delta m}(\mu_N^{i,0},x)\right)(\delta_{X_i}-\mu)(dx).$$

Let us therefore consider the following decomposition:

$$\sqrt{N}\left(U\left(\mu_N\right)-U\left(\mu\right)\right)=\sqrt{N}\left(U\left(\mu_N\right)-U\left(\mu\right)-Q_N\right)+\sqrt{N}Q_N.$$

We will see that the first term will go to 0 in probability (**third step**) while the second one will converge in distribution to a normal random variable (**fourth step**).

THIRD STEP Let us first prove the convergence in probability of $\sqrt{N}\left(U\left(\mu_N\right)-U\left(\mu\right)-Q_N\right)$ to 0. By definition we need to prove that $\forall\epsilon>0$

$$\lim_{N\to\infty}\mathbb{P}\left(\sqrt{N}|U\left(\mu_N\right)-U\left(\mu\right)-Q_N|\geq\epsilon\right)=0.$$

One has

$$\left\{\sqrt{N}|U\left(\mu_N\right)-U\left(\mu\right)-Q_N|\geq\epsilon\right\}$$

$$=\left(\left\{\sqrt{N}|U\left(\mu_N\right)-U\left(\mu\right)-Q_N|\geq\epsilon\right\}\bigcap\{N\geq N^*\}\right)\dot{\bigcup}\left(\left\{\sqrt{N}|U\left(\mu_N\right)-U\left(\mu\right)-Q_N|\geq\epsilon\right\}\bigcap\{N<N^*\}\right)$$

$$\subseteq\left(\left\{\sqrt{N}|U\left(\mu_N\right)-U\left(\mu\right)-Q_N|\geq\epsilon\right\}\bigcap\{N\geq N^*\}\right)\dot{\bigcup}\{N<N^*\}.$$

Therefore

$$\mathbb{P}\left(\sqrt{N}|U\left(\mu_N\right)-U\left(\mu\right)-Q_N|\geq\epsilon\right)\leq\mathbb{P}\left(N<N^*\right)+\mathbb{P}\left(\sqrt{N}1_{N\geq N^*}|U\left(\mu_N\right)-U\left(\mu\right)-Q_N|\geq\epsilon\right).$$

Since $\lim_{N\to\infty}\mathbb{P}\left(N<N^*\right)=0$, it is sufficient to prove the almost sure convergence of $\sqrt{N}R_N$ to 0. One has

$$\begin{aligned}|R_N|&\leq\frac{1_{N\geq N^*}}{N}\sum_{i=1}^{N}\int_0^1 ds\int_{\mathbb{R}^d}\left|\frac{\delta U}{\delta m}(\mu_N^{i,s},x)-\frac{\delta U}{\delta m}(\mu_N^{i,0},x)\right||\delta_{X_i}-\mu|(dx)\\&\leq\frac{1_{N\geq N^*}}{N}\sum_{i=1}^{N}\int_0^1 ds\int_{\mathbb{R}^d}\left|\frac{\delta U}{\delta m}(\mu_N^{i,s},x)-\frac{\delta U}{\delta m}(\mu_N^{i,0},x)\right|(\delta_{X_i}+\mu)(dx).\end{aligned}$$

By Assumption **RU6**, $\exists C<\infty$, $\exists\alpha\in\left(\frac{1}{2},1\right]$ such that for $N\geq N^*$

$$\left|\frac{\delta U}{\delta m}(\mu_N^{i,s},x)-\frac{\delta U}{\delta m}(\mu_N^{i,0},x)\right|\leq C(1+|x|^{\frac{\ell}{2}})\left(\int_{\mathbb{R}^d}\left(1+|y|^{\frac{\ell}{2\alpha}}\right)|\mu_N^{i,s}-\mu_N^{i,0}|(dy)\right)^{\alpha}$$

with

$$|\mu_N^{i,s}-\mu_N^{i,0}|(dy)\leq\frac{s}{N}(\delta_{X_i}+\mu)(dy).$$

Substituting the above quantity and using the subadditivity of $x\mapsto x^{\alpha}$ one obtains

$$\begin{aligned}\left|\frac{\delta U}{\delta m}(\mu_N^{i,s},x)-\frac{\delta U}{\delta m}(\mu_N^{i,0},x)\right|&\leq C\left((1+|x|^{\frac{\ell}{2}})\left(2+\int_{\mathbb{R}^d}|y|^{\frac{\ell}{2\alpha}}(\delta_{X_i}+\mu)(dy)\right)^{\alpha}\left(\frac{s}{N}\right)^{\alpha}\right)\\&\leq\frac{C}{N^{\alpha}}(1+|x|^{\frac{\ell}{2}})\left(2^{\alpha}+|X_i|^{\frac{\ell}{2}}+\left(\int_{\mathbb{R}^d}|y|^{\frac{\ell}{2\alpha}}\mu(dy)\right)^{\alpha}\right)\\&\leq\frac{C^*}{N^{\alpha}}\left(1+|x|^{\frac{\ell}{2}}\right)\left(1+|X_i|^{\frac{\ell}{2}}\right)\end{aligned}$$

where we used (2.22) to obtain the last inequality with C^* a finite constant. Therefore

$$
\begin{aligned}
|R_N| &\leq \frac{C^*}{N^{\alpha+1}} \sum_{i=1}^{N} \int_{\mathbb{R}^d} \left(1 + |x|^{\frac{\ell}{2}}\right) \left(1 + |X_i|^{\frac{\ell}{2}}\right) (\delta_{X_i} + \mu)(dx) \\
&= \frac{C^*}{N^{\alpha+1}} \sum_{i=1}^{N} \left(1 + |X_i|^{\frac{\ell}{2}}\right) \left(2 + |X_i|^{\frac{\ell}{2}} + \int_{\mathbb{R}^d} |x|^{\frac{\ell}{2}} \mu(dx)\right) \\
&= \frac{C^*}{N^{\alpha+1}} \sum_{i=1}^{N} \left(2 + \int_{\mathbb{R}^d} |x|^{\frac{\ell}{2}} \mu(dx) + \left(3 + \int_{\mathbb{R}^d} |x|^{\frac{\ell}{2}} \mu(dx)\right) |X_i|^{\frac{\ell}{2}} + |X_i|^{\ell}\right) \\
&\leq \frac{C_1}{N^{\alpha}} + \frac{C_2}{N^{\alpha}} \int_{\mathbb{R}^d} |y|^{\ell} \mu_N(dy)
\end{aligned}
$$

for some positive constants C_1 and C_2. Therefore one has

$$
\sqrt{N}|R_N| \leq \frac{C_1}{N^{\alpha-\frac{1}{2}}} + \frac{C_2}{N^{\alpha-\frac{1}{2}}} \int_{\mathbb{R}^d} |y|^{\ell} \mu_N(dy)
$$

and since $\alpha > \frac{1}{2}$ and $\lim_{N\to\infty} \int_{\mathbb{R}^d} |y|^{\ell} \mu_N(dy) = \int_{\mathbb{R}^d} |y|^{\ell} \mu(dy)$, we can conclude that the left-hand side goes to 0 as N goes to infinity and so the third step is concluded.

FOURTH STEP As previously mentioned, we will now proceed to prove the convergence in distribution of $\sqrt{N}Q_N$ to a Gaussian random variable.
Remember that

$$
\sqrt{N}Q_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\delta U}{\delta m}(\mu_N^{i\wedge I_N,0}, X_i) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\mu_N^{i\wedge I_N,0}, x)\mu(dx)\right)
$$

and let us consider for a given $m \in B(\mu, r)$ the following Poisson equation

$$
F(m,x) - PF(m,x) = \frac{\delta U}{\delta m}(m,x) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m,y)\,\mu(dy), \quad x \in \mathbb{R}^d. \tag{2.41}
$$

By Proposition 4 we know that it admits a solution. Thanks to that, we are able to rewrite $\sqrt{N}Q_N$ as

$$
\begin{aligned}
\sqrt{N}Q_N &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(F\left(\mu_N^{i\wedge I_N,0}, X_i\right) - PF\left(\mu_N^{i\wedge I_N,0}, X_i\right)\right) \\
&= \frac{1}{\sqrt{N}} \left(\sum_{i=1}^{N} F\left(\mu_N^{i\wedge I_N,0}, X_i\right) - \sum_{i=2}^{N} F\left(\mu_N^{(i-1)\wedge I_N,0}, X_i\right)\right) \\
&\quad + \frac{1}{\sqrt{N}} \left(\sum_{i=2}^{N} F\left(\mu_N^{(i-1)\wedge I_N,0}, X_i\right) - \sum_{i=2}^{N+1} PF\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right)\right) \\
&= K_{0,N} + K_{1,N} + K_{2,N} + K_{3,N}
\end{aligned}
$$

with

$$
\begin{aligned}
K_{0,N} &= \frac{F(\mu, X_1)}{\sqrt{N}}, \\
K_{1,N} &= \frac{1}{\sqrt{N}} \sum_{i=2}^{N} \left(F\left(\mu_N^{i\wedge I_N,0}, X_i\right) - F\left(\mu_N^{(i-1)\wedge I_N,0}, X_i\right)\right), \\
K_{2,N} &= -\frac{PF\left(\mu_N^{N\wedge I_N,0}, X_N\right)}{\sqrt{N}}, \\
K_{3,N} &= \frac{1}{\sqrt{N}} \sum_{i=2}^{N} \left(F\left(\mu_N^{(i-1)\wedge I_N,0}, X_i\right) - PF\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right)\right).
\end{aligned}
$$

The idea now is to study the convergence of $K_{i,N}$ for $i = 0, \cdots, 3$. We will see that $K_{0,N}, K_{1,N}$ and $K_{2,N}$ go to 0 in probability as N goes to infinity while $K_{3,N}$ is the term providing the convergence in distribution to a Gaussian random variable. By Slutsky's theorem, we can therefore obtain the limit of $\sqrt{N}Q_N$ (same as the limit of $K_{3,N}$) and conclude the proof.

Convergence of $K_{0,N} + K_{1,N} + K_{2,N}$ to 0 in probability

The almost sure convergence (and so in probability) of $K_{0,N}$ to 0 is immediate.
For what concerns the convergence in probability of $K_{1,N}$ to 0, following the same idea used in the third step, it is sufficient to prove the almost sure convergence of $1_{N\geq N^*}|K_{1,N}|$ to 0. Therefore

$$
\begin{aligned}
1_{N\geq N^*}|K_{1,N}| &= \left| \frac{1_{N\geq N^*}}{\sqrt{N}} \sum_{i=2}^{N} \left(F\left(\mu_N^{i\wedge I_N,0}, X_i\right) - F\left(\mu_N^{(i-1)\wedge I_N,0}, X_i\right)\right)\right| && (2.42)\\
&\leq \frac{1_{N\geq N^*}}{\sqrt{N}} \sum_{i=2}^{N} \left| F\left(\mu_N^{i,0}, X_i\right) - F\left(\mu_N^{i-1,0}, X_i\right)\right| && (2.43)\\
&\leq \frac{1_{N\geq N^*}}{\sqrt{N}} \sum_{i=2}^{N} \left\| F\left(\mu_N^{i,0}, \cdot\right) - F\left(\mu_N^{i-1,0}, \cdot\right)\right\|_{\sqrt{V}} \sqrt{V(X_i)}. && (2.44)
\end{aligned}
$$

By (2.36), observing that $\mu_N^{i,0} - \mu_N^{i-1,0} = \frac{1}{N}\left(\delta_{X_{i-1}} - \mu\right)$, one obtains

$$
\begin{aligned}
1_{N\geq N^*} \left\| F\left(\mu_N^{i,0}, \cdot\right) - F\left(\mu_N^{i-1,0}, \cdot\right)\right\|_{\sqrt{V}} &\leq 1_{N\geq N^*} \bar{C}_{\frac{1}{2}} \left(\int_{\mathbb{R}^d} \left(1 + |y|^{\frac{\ell}{2\alpha}}\right) |\mu_N^{i-1,0} - \mu_N^{i,0}|(dy)\right)^{\alpha} \\
&\leq 1_{N\geq N^*} \bar{C}_{\frac{1}{2}} \frac{1}{N^{\alpha}} \left(2 + \int_{\mathbb{R}^d} |y|^{\frac{\ell}{2\alpha}} \delta_{X_{i-1}}(dy) + \int_{\mathbb{R}^d} |y|^{\frac{\ell}{2\alpha}} \mu(dy)\right)^{\alpha} \\
&\leq 1_{N\geq N^*} \bar{C}^*_{\frac{1}{2}} \frac{1}{N^{\alpha}} \left(|X_{i-1}|^{\frac{\ell}{2}} + 1\right).
\end{aligned}
$$

where we used (2.22) and the subadditivity of $x \mapsto x^{\alpha}$ to obtain the last inequality with $\bar{C}^*_{\frac{1}{2}}$ a finite constant. Therefore plugging this inequality in (2.44), we obtain the following estimation for $1_{N\geq N^*}|K_{1,N}|$

$$
\begin{aligned}
1_{N\geq N^*}|K_{1,N}| &\leq \frac{1_{N\geq N^*}\bar{C}^*_{\frac{1}{2}}}{N^{\alpha+\frac{1}{2}}} \sum_{i=2}^{N} \left(|X_{i-1}|^{\frac{\ell}{2}} + 1\right) \sqrt{V(X_i)} \\
&= 1_{N\geq N^*} \bar{C}^*_{\frac{1}{2}} \left(\frac{1}{N^{\alpha+\frac{1}{2}}} \sum_{i=2}^{N} |X_{i-1}|^{\frac{\ell}{2}} \sqrt{V(X_i)} + \frac{1}{N^{\alpha-\frac{1}{2}}} \times \frac{1}{N} \sum_{i=2}^{N} \sqrt{V(X_i)}\right) \\
&\leq 1_{N\geq N^*} \bar{C}^*_{\frac{1}{2}} \left(\frac{1}{2N^{\alpha-\frac{1}{2}}} \times \frac{1}{N} \sum_{i=2}^{N} |X_{i-1}|^{\ell} + \frac{1}{N^{\alpha-\frac{1}{2}}} \times \frac{1}{N} \sum_{i=2}^{N} V(X_i) + \frac{N-1}{2N^{\alpha+\frac{1}{2}}}\right).
\end{aligned}
$$

Since $\alpha > \frac{1}{2}$, $\frac{1}{N}\sum_{i=2}^{N}|X_{i-1}|^{\ell}$ converges a.s. to $\int_{\mathbb{R}^d}|x|^{\ell}\mu(dx)$ and $\frac{1}{N}\sum_{i=2}^{N}V(X_i)$ converges a.s. to $\mu(V)$, we can conclude that $K_{1,N}$ goes to 0 in probability.
Finally let us prove the convergence of $K_{2,N}$ to 0 in probability. Still in this case it is sufficient to prove that $1_{N\geq N^*}|K_{2,N}|$ converges in probability to 0. By (2.34) we have

$$\begin{aligned}1_{N\geq N^*}|K_{2,N}| &= \frac{1_{N\geq N^*}}{\sqrt{N}}\left|PF\left(\mu_N^{N,0},X_N\right)\right| \leq \frac{1_{N\geq N^*}}{\sqrt{N}}\left\|F\left(\mu_N^{N,0},\cdot\right)\right\|_{\sqrt{V}}\int_{\mathbb{R}^d}\sqrt{V(x)}P(X_N,dx)\\ &= \frac{1_{N\geq N^*}}{\sqrt{N}}\left\|F\left(\mu_N^{N,0},\cdot\right)\right\|_{\sqrt{V}}P\sqrt{V}(X_N) \leq \frac{1_{N\geq N^*}}{\sqrt{N}}\bar{C}_{\frac{1}{2}}\left(\sqrt{K_1}\sqrt{V(X_N)}+\sqrt{K_2}\right)\end{aligned}$$

where we applied (2.15) to obtain the last inequality. We can therefore conclude that the right-hand side goes to 0 as N goes to infinity by observing that by Theorem 21, $\bar{V}_N := \frac{1}{N}\sum_{i=1}^{N}V(X_i)$ converges almost surely to $\mu(V)$ and consequently

$$\sqrt{\frac{V(X_N)}{N}} = \sqrt{\bar{V}_N - \frac{N-1}{N}\bar{V}_{N-1}} \longrightarrow 0 \ a.s.$$

Convergence in distribution of $K_{3,N}$

To study the convergence in distribution of $K_{3,N}$, we apply the Central Limit Theorem for martingales (see Corollary 3.1 [31]). First we need to prove a square integrable martingale property and then that the Bracket condition and the Lindeberg condition hold.
Let us recall the expression of $K_{3,N}$

$$K_{3,N} = \frac{1}{\sqrt{N}}\sum_{i=2}^{N}\left(F\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)-PF\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)\right).$$

For $i=2,\cdots,N$ let

$$Y_{N,i} = \frac{1}{\sqrt{N}}\left(F\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)-PF\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)\right).$$

We will start by **checking that** $\mathbb{E}(Y_{N,i}|\mathcal{F}_{i-1}) = 0$**.** Since $\mu_N^{i\wedge I_N,0} = \sum_{j=1}^{i-1}1_{\{I_N=j\}}\mu_N^{j,0} + 1_{\{I_N>i-1\}}\mu_N^{i,0}$ is $\mathcal{F}_{i-1}$- measurable, one has

$$\begin{aligned}\mathbb{E}(Y_{N,i}|\mathcal{F}_{i-1}) &= \frac{1}{\sqrt{N}}\mathbb{E}\left(F\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)-PF\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)|\mathcal{F}_{i-1}\right)\\ &= \frac{1}{\sqrt{N}}\left(\mathbb{E}\left(F\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)|\mathcal{F}_{i-1}\right)-\mathbb{E}\left(F\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)|\mathcal{F}_{i-1}\right)\right)=0.\end{aligned}$$

For what concerns the **square integrability**, it is sufficient to check that

$$\mathbb{E}\left(F^2\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)\right)<\infty.$$

By (2.34)

$$\mathbb{E}\left(F^2\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)\right) \leq \mathbb{E}\left(\left\|F\left(\mu_N^{(i-1)\wedge I_N,0},\cdot\right)\right\|_{\sqrt{V}}^2\left(\sqrt{V(X_i)}\right)^2\right) \leq \bar{C}_{\frac{1}{2}}^2\mathbb{E}(V(X_i))$$

where $\mathbb{E}(V(X_i)) \leq \gamma^{i-1}\nu_1(V)+K\sum_{\ell=0}^{i-2}\gamma^{\ell}$ and so by Lemma 3, the right-hand side is finite.

Let us now study the **convergence of** $\sum_{i=2}^{N}\mathbb{E}\left(Y_{N,i}^2|\mathcal{F}_{i-1}\right)$:

$$\begin{aligned}
&\sum_{i=2}^{N} \mathbb{E}\left(Y_{N,i}^2 | \mathcal{F}_{i-1}\right) \\
&= \frac{1}{N}\sum_{i=2}^{N} \mathbb{E}\left(F^2\left(\mu_N^{(i-1)\wedge I_N,0}, X_i\right) + (PF)^2\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right) | \mathcal{F}_{i-1}\right) \\
&\quad - \frac{1}{N}\sum_{i=2}^{N} \mathbb{E}\left(2F\left(\mu_N^{(i-1)\wedge I_N,0}, X_i\right) PF\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right) | \mathcal{F}_{i-1}\right) \\
&= \frac{1}{N}\sum_{i=2}^{N} PF^2\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right) - \frac{1}{N}\sum_{i=2}^{N} (PF)^2\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right)
\end{aligned}$$

Before studying the behavior of the above quantity, let us observe that the convergence of $\sup_{x\in\mathbb{R}^d} \frac{|\frac{\delta U}{\delta m}(\tilde{\mu},x) - \frac{\delta U}{\delta m}(\mu,x)|}{1+|x|^{\frac{\ell}{2}}}$ to 0 when $W_\ell(\tilde{\mu},\mu)$ goes to 0 together with the a.s. convergence of $\max_{1\le i\le N} W_\ell(\mu_N^{i,0},\mu)$ to 0 imply the existence of a sequence of random variables $(\epsilon_N)_{N\ge 0}$ converging a.s. to 0 as $N\to\infty$ such that

$$\forall 1\le i\le N \qquad \left|\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0}, x\right) - \frac{\delta U}{\delta m}(\mu,x)\right| \le \left(1+|x|^{\frac{\ell}{2}}\right)\epsilon_N. \tag{2.45}$$

With (2.35) and **L3**, we deduce that

$$\left\|F\left(\mu_N^{i\wedge I_N,0},\cdot\right) - F(\mu,\cdot)\right\|_{\sqrt{V}} \le \bar{C}_{\frac{1}{2}}\epsilon_N \sup_{x\in\mathbb{R}^d}\frac{1+|x|^{\frac{\ell}{2}}}{\sqrt{V(x)}} \le \bar{C}_{\frac{1}{2}}\left(1+\sqrt{C_\ell}\right)\epsilon_N. \tag{2.46}$$

First Component We can rewrite the first component in the following way

$$\begin{aligned}
&\frac{1}{N}\sum_{i=2}^{N} PF^2\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right) \\
&\quad = \frac{1}{N}\sum_{i=2}^{N}\left(PF^2\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right) - PF^2(\mu, X_{i-1})\right) + \frac{1}{N}\sum_{i=2}^{N} PF^2(\mu, X_{i-1}).
\end{aligned}$$

Since by (2.13)

$$\left\|PF^2(\mu,\cdot)\right\|_V \le \left\|F^2(\mu,\cdot)\right\|_V \sup_{x\in\mathbb{R}^d}\frac{PV(x)}{V(x)} < \infty, \tag{2.47}$$

by Theorem 21 we obtain that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=2}^{N} PF^2(\mu, X_{i-1}) = \mu(PF^2(\mu,\cdot)) = \mu(F^2(\mu,\cdot)) \quad a.s.$$

where for the last equality we used the invariance of μ with respect to P. On the other hand

$$\begin{aligned}
&\left|\frac{1}{N}\sum_{i=2}^{N}\left(PF^2\left(\mu_N^{(i-1)\wedge I_N,0}, X_{i-1}\right) - PF^2(\mu, X_{i-1})\right)\right| \\
&\quad \le \frac{1}{N}\sum_{i=2}^{N}\int_{\mathbb{R}^d}\left|F^2\left(\mu_N^{(i-1)\wedge I_N,0}, x\right) - F^2(\mu,x)\right| P(X_{i-1}, dx)
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{N}\sum_{i=2}^{N}\int_{\mathbb{R}^d}\left(F\left(\mu_N^{(i-1)\wedge I_N,0},x\right)-F\left(\mu,x\right)\right)^2 P\left(X_{i-1},dx\right)\\
&\quad+\frac{2}{N}\sum_{i=2}^{N}\int_{\mathbb{R}^d}\left|F\left(\mu,x\right)\right|\left|F\left(\mu_N^{(i-1)\wedge I_N,0},x\right)-F\left(\mu,x\right)\right| P\left(X_{i-1},dx\right)\\
&\leq \frac{1}{N}\sum_{i=2}^{N}\left\|F\left(\mu_N^{(i-1)\wedge I_N,0},\cdot\right)-F\left(\mu,\cdot\right)\right\|_{\sqrt{V}}^2 PV\left(X_{i-1}\right)\\
&\quad+\frac{2}{N}\sum_{i=2}^{N}\left\|F\left(\mu,\cdot\right)\right\|_{\sqrt{V}}\left\|F\left(\mu_N^{(i-1)\wedge I_N,0},\cdot\right)-F\left(\mu,\cdot\right)\right\|_{\sqrt{V}} PV\left(X_{i-1}\right)\\
&\leq \epsilon_N\left(\bar{C}_{\frac{1}{2}}^2\left(1+\sqrt{C_\ell}\right)^2\epsilon_N+2\bar{C}_{\frac{1}{2}}^2\left(1+\sqrt{C_\ell}\right)\right)\frac{1}{N}\sum_{i=2}^{N}\left(K_1V(X_{i-1})+K_2\right)
\end{aligned}
$$

where we used (2.46), (2.34) and (2.13). Therefore the right-hand side goes to 0 since $\frac{1}{N}\sum_{i=2}^{N}V\left(X_{i-1}\right)$ converges to $\mu(V)$ and ϵ_N goes to 0 a.s..

Second Component As before, we can rewrite the second component in the following way

$$
\begin{aligned}
&\frac{1}{N}\sum_{i=2}^{N}\left(PF\right)^2\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)\\
&\quad=\frac{1}{N}\sum_{i=2}^{N}\left(\left(PF\right)^2\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)-\left(PF\right)^2\left(\mu,X_{i-1}\right)\right)+\frac{1}{N}\sum_{i=2}^{N}\left(PF\right)^2\left(\mu,X_{i-1}\right).
\end{aligned}
$$

By (2.47) and Jensen's inequality $\left\|\left(PF\right)^2\left(\mu,\cdot\right)\right\|_V<\infty$, therefore by Theorem 21

$$
\frac{1}{N}\sum_{i=2}^{N}\left(PF\right)^2\left(\mu,X_{i-1}\right)=\mu(\left(PF\right)^2\left(\mu,\cdot\right))\quad a.s..
$$

On the other hand

$$
\begin{aligned}
&\left|\frac{1}{N}\sum_{i=2}^{N}\left(\left(PF\right)^2\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)-\left(PF\right)^2\left(\mu,X_{i-1}\right)\right)\right|\\
&\quad\leq\frac{1}{N}\sum_{i=2}^{N}\left(PF\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)-PF\left(\mu,X_{i-1}\right)\right)^2\\
&\qquad+\frac{2}{N}\sum_{i=2}^{N}\left|PF\left(\mu,X_{i-1}\right)\right|\left|PF\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)-PF\left(\mu,X_{i-1}\right)\right|\\
&\quad\leq\frac{1}{N}\sum_{i=2}^{N}\left(\left\|F\left(\mu_N^{(i-1)\wedge I_N,0},\cdot\right)-F\left(\mu,\cdot\right)\right\|_{\sqrt{V}}\left(P\sqrt{V}(X_{i-1})\right)\right)^2\\
&\qquad+\frac{2}{N}\sum_{i=2}^{N}\left\|F\left(\mu,\cdot\right)\right\|_{\sqrt{V}}\left\|F\left(\mu_N^{(i-1)\wedge I_N,0},\cdot\right)-F\left(\mu,\cdot\right)\right\|_{\sqrt{V}}\left(P\sqrt{V}(X_{i-1})\right)^2\\
&\quad\leq\epsilon_N\left(\bar{C}_{\frac{1}{2}}^2\left(1+\sqrt{C_\ell}\right)^2\epsilon_N+2\bar{C}_{\frac{1}{2}}^2\left(1+\sqrt{C_\ell}\right)\right)\frac{1}{N}\sum_{i=2}^{N}\left(\sqrt{K_1}\sqrt{V\left(X_{i-1}\right)}+\sqrt{K_2}\right)^2.
\end{aligned}
$$

The right-hand side goes to 0 as N goes to infinity since ϵ_N goes to 0 a.s. and, by Theorem 21, $\frac{1}{N}\sum_{i=2}^{N} V(X_{i-1})$ converges to $\mu(V)$ and $\frac{1}{N}\sum_{i=2}^{N}\sqrt{V(X_{i-1})}$ converges to $\mu(\sqrt{V})$.
In conclusion we have proved that almost surely

$$\frac{1}{N}\sum_{i=2}^{N}\left(PF^2\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)-(PF)^2\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)\right)\underset{N\to\infty}{\longrightarrow}\mu\left(F^2(\mu,\cdot)\right)-\mu\left((PF)^2(\mu,\cdot)\right). \tag{2.48}$$

To finally conclude that $K_{3,N}=\sum_{i=2}^{N}Y_{N,i}\overset{d}{\Longrightarrow}\mathcal{N}\left(0,\mu\left(F^2(\mu,\cdot)\right)-\mu\left((PF)^2(\mu,\cdot)\right)\right)$, it remains just to verify that the **Lindeberg condition** holds.
We need to check that for $\epsilon>0$, $\sum_{i=2}^{N}\mathbb{E}\left(Y_{N,i}^2 1_{\{Y_{N,i}^2>\epsilon\}}|\mathcal{F}_{i-1}\right)$ goes to 0 in probability as N goes to infinity. By Jensen's inequality, (2.13) and (2.34) we obtain

$$\begin{aligned}
NY_{N,i}^2&=\left(F\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)-PF\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)\right)^2\\
&\le 2F^2\left(\mu_N^{(i-1)\wedge I_N,0},X_i\right)+2PF^2\left(\mu_N^{(i-1)\wedge I_N,0},X_{i-1}\right)\\
&\le 2\left\|F\left(\mu_N^{(i-1)\wedge I_N,0},\cdot\right)\right\|_{\sqrt{V}}^2 V(X_i)+2\left\|F\left(\mu_N^{(i-1)\wedge I_N,0},\cdot\right)\right\|_{\sqrt{V}}^2 PV(X_{i-1})\\
&\le 2\bar{C}_{\frac{1}{2}}^2\left(V(X_i)+K_1V(X_{i-1})+K_2\right)\le 2L\bar{C}_{\frac{1}{2}}^2\left(1+V(X_i)+V(X_{i-1})\right)
\end{aligned}$$

where L is a finite constant. As, for $a,b,c,g\in\mathbb{R}_+$

$$\begin{aligned}
(a+b+c)\,1_{\{a+b+c\ge g\}}&=(a+b+c)\left(1_{\{a>b,a>c,a+b+c\ge g\}}+1_{\{b\ge a,b>c,a+b+c\ge g\}}+1_{\{c\ge a,c\ge b,a+b+c\ge g\}}\right)\\
&\le 3a1_{\{a>b,a>c,a+b+c\ge g\}}+3b1_{\{b\ge a,b>c,a+b+c\ge g\}}+3c1_{\{c\ge a,c\ge b,a+b+c\ge g\}}\\
&\le 3a1_{\left\{a\ge\frac{g}{3}\right\}}+3b1_{\left\{b\ge\frac{g}{3}\right\}}+3c1_{\left\{c\ge\frac{g}{3}\right\}},
\end{aligned}$$

it is enough to check that for each $\epsilon>0$

$$\frac{N-1}{N}1_{\left\{\frac{1}{N}>\epsilon\right\}}+\sum_{i=2}^{N}\mathbb{E}\left(\frac{V(X_i)}{N}1_{\left\{\frac{V(X_i)}{N}>\epsilon\right\}}|\mathcal{F}_{i-1}\right)+\sum_{i=2}^{N}\frac{V(X_{i-1})}{N}1_{\left\{\frac{V(X_{i-1})}{N}>\epsilon\right\}}$$

goes to 0 in probability as N goes to infinity.
It is immediate to prove that the **first component** goes to 0. For the **second component** let us observe that

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=2}^{N}\mathbb{E}\left(\frac{V(X_i)}{N}1_{\left\{\frac{V(X_i)}{N}>\epsilon\right\}}|\mathcal{F}_{i-1}\right)\right)&=\sum_{i=2}^{N}\mathbb{E}\left(\frac{V(X_i)}{N}1_{\left\{\frac{V(X_i)}{N}>\epsilon\right\}}\right)\\
&=\sum_{i=2}^{N}\int_{\mathbb{R}^d}\frac{V(x)}{N}1_{\left\{\frac{V(x)}{N}>\epsilon\right\}}\nu_1P^{i-1}(dx)\\
&=\sum_{i=2}^{N}\int_{\mathbb{R}^d}\frac{V(x)}{N}1_{\left\{\frac{V(x)}{N}>\epsilon\right\}}\left(\nu_1P^{i-1}(dx)-\mu(dx)\right)\\
&\quad+\frac{N-1}{N}\int_{\mathbb{R}^d}V(x)1_{\left\{\frac{V(x)}{N}>\epsilon\right\}}\mu(dx)\\
&=\sum_{i=2}^{N}\int_{\mathbb{R}^d}\nu_1(dy)\int_{\mathbb{R}^d}\frac{V(x)}{N}1_{\left\{\frac{V(x)}{N}>\epsilon\right\}}\left(P^{i-1}(y,dx)-\mu(dx)\right)\\
&\quad+\frac{N-1}{N}\int_{\mathbb{R}^d}V(x)1_{\left\{\frac{V(x)}{N}>\epsilon\right\}}\mu(dx).
\end{aligned}$$

We can apply (2.11) and deduce that

$$\sum_{i=2}^{N}\int_{\mathbb{R}^d}\nu_1(dy)\int_{\mathbb{R}^d}\frac{V(x)}{N}1_{\left\{\frac{V(x)}{N}>\epsilon\right\}}\left(P^{i-1}(y,dx)-\mu(dx)\right)$$
$$\leq R\left\|\frac{V(\cdot)}{N}1_{\left\{\frac{V(\cdot)}{N}>\epsilon\right\}}\right\|_V\nu_1(V)\times\sum_{i=2}^{N}r^{i-1}\leq\frac{R}{N}\frac{\nu_1(V)}{1-r}.$$

Thus

$$\mathbb{E}\left(\sum_{i=2}^{N}\mathbb{E}\left(\frac{V(X_i)}{N}1_{\left\{\frac{V(X_i)}{N}>\epsilon\right\}}|\mathcal{F}_{i-1}\right)\right)\leq\frac{R}{N}\frac{\nu_1(V)}{1-r}+\frac{N-1}{N}\int_{\mathbb{R}^d}V(x)1_{\left\{\frac{V(x)}{N}>\epsilon\right\}}\mu(dx).$$

By Lemma 3, it is enough to suppose that $\nu_1(V)<\infty$ and the first term goes to 0. The second term goes to 0 by Lebesgue's theorem since $\mu(V)<\infty$. Finally let us study the **third component**. By Theorem 21 we know that $\overline{V}_N:=\frac{1}{N}\sum_{i=1}^{N}V(X_i)$ converges almost surely to $\mu(V)$. Therefore

$$\frac{V(X_N)}{N}=\overline{V}_N-\frac{N-1}{N}\overline{V}_{N-1}\underset{N\to\infty}{\longrightarrow}0\quad a.s.$$

that implies $\forall\epsilon>0$

$$V(X_N)1_{\left\{\frac{V(X_N)}{N}>\epsilon\right\}}\underset{N\to\infty}{\longrightarrow}0\quad a.s.$$

and taking the Cesaro mean we deduce that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}V(X_i)1_{\left\{\frac{V(X_i)}{N}>\epsilon\right\}}\leq\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}V(X_i)1_{\left\{\frac{V(X_i)}{i}>\epsilon\right\}}=0.$$

□

2.2.2 Independent Non-Equidistributed Random Variables

Before providing the proof of Theorem 17, let us observe that Hypothesis **TX** implies the Lindeberg condition.

Lemma 5. *Hypothesis **TX** implies that*

$$\forall\epsilon>0,\ \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(|X_i|^{\ell}1_{\{|X_i|^{\ell}>N\epsilon\}}\right)=0.\tag{2.49}$$

Proof. An application of Kronecker's lemma shows that **TX** implies

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\left(|X_i|^{\ell}-i^{\beta}\right)1_{\{|X_i|^{\ell}>i^{\beta}\}}\right)=0.\tag{2.50}$$

Let $\epsilon>0$, then exists $\bar{N}_{\epsilon,\beta}$ such that for $N\geq\bar{N}_{\epsilon,\beta}$, $N^{\beta}\leq\frac{\epsilon N}{2}$. Therefore if $N\geq\bar{N}_{\epsilon,\beta}$, the following chain of inequalities holds for each $i=1,\cdots,N$

$$\left(|X_i|^{\ell}-i^{\beta}\right)1_{\{|X_i|^{\ell}>i^{\beta}\}}\geq\left(|X_i|^{\ell}-\frac{\epsilon N}{2}\right)1_{\{|X_i|^{\ell}>\epsilon N\}}$$
$$\geq\frac{|X_i|^{\ell}}{2}1_{\{|X_i|^{\ell}>\epsilon N\}}.$$

Taking the expectation and using (2.50) we can obtain (2.49).

□

Proof of Theorem 17. To study the limit distribution of $\sqrt{N}\left(U\left(\mu_N\right)-U\left(\mu\right)\right)$, let us consider the following decomposition

$$\sqrt{N}\left(U\left(\mu_N\right)-U\left(\mu\right)\right)=\sqrt{N}\left(U\left(\mu_N\right)-U\left(\bar{\nu}_N\right)\right)+\sqrt{N}\left(U\left(\bar{\nu}_N\right)-U\left(\mu\right)\right). \tag{2.51}$$

We will prove that

$$\sqrt{N}\left(U\left(\mu_N\right)-U\left(\bar{\nu}_N\right)\right)\overset{d}{\Rightarrow}\mathcal{N}\left(0,\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\mu,x\right)^2\mu(dx)-\int_{\mathbb{R}^d\times\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\frac{\delta U}{\delta m}(\mu,y)\eta(dx,dy)\right)$$

and $\sqrt{N}\left(U\left(\bar{\nu}_N\right)-U\left(\mu\right)\right)\underset{N\to\infty}{\longrightarrow}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\sigma(dx)$.
Let us start by studying the second term in the above decomposition.

Limit of $\sqrt{N}\left(U\left(\bar{\nu}_N\right)-U\left(\mu\right)\right)$

Let us first observe that by Remark 18, Assumption 1 implies $W_\ell\left(\bar{\nu}_N,\mu\right)\underset{N\to\infty}{\longrightarrow}0$. Therefore one has

$$\lim_{N\to\infty}\sup_{s\in[0,1]}W_\ell\left(\mu+s(\bar{\nu}_N-\mu),\mu\right)=0.$$

For N bigger than a fixed N_r, by Lemma 2 we can rewrite

$$\sqrt{N}\left(U\left(\bar{\nu}_N\right)-U\left(\mu\right)\right)=\sqrt{N}\int_0^1 ds\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu+s(\bar{\nu}_N-\mu),x)(\bar{\nu}_N-\mu)(dx).$$

Let us now prove that the above quantity tends to $\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\sigma(dx)$ where σ is the measure such that $\|\sqrt{N}\left(\bar{\nu}_N-\mu\right)-\sigma\|_\ell\underset{N\to\infty}{\longrightarrow}0$ (see Hypothesis 2). By the triangle inequality, for $N\geq N_r$, one has

$$\begin{aligned}
&\left|\sqrt{N}\int_0^1 ds\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu+s(\bar{\nu}_N-\mu),x)(\bar{\nu}_N-\mu)(dx)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\sigma(dx)\right|\\
&\quad\leq\left|\sqrt{N}\int_0^1 ds\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu+s(\bar{\nu}_N-\mu),x)(\bar{\nu}_N-\mu)(dx)-\sqrt{N}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)(\bar{\nu}_N-\mu)(dx)\right|\\
&\qquad+\left|\sqrt{N}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)(\bar{\nu}_N-\mu)(dx)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\sigma(dx)\right|.
\end{aligned}$$

Recalling that $\|\tau\|_\ell=\sup_{f:|f(x)|\leq 1+|x|^\ell}\int_{\mathbb{R}^d}f(x)\tau(dx)$ for $\tau\in\mathcal{M}_\ell(\mathbb{R}^d)$, the second component of the right-hand side goes to 0 thanks to Assumption 2 and Assumption **RU2**.
For what concerns the first component of the right-hand side, we have by Hypothesis **RU5**

$$\begin{aligned}
&\left|\sqrt{N}\int_0^1 ds\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu+s(\bar{\nu}_N-\mu),x)(\bar{\nu}_N-\mu)(dx)-\sqrt{N}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)(\bar{\nu}_N-\mu)(dx)\right|\\
&\leq\int_0^1 ds\int_{\mathbb{R}^d}\left|\frac{\delta U}{\delta m}(\mu+s(\bar{\nu}_N-\mu),x)-\frac{\delta U}{\delta m}(\mu,x)\right|\left|\sqrt{N}\left(\bar{\nu}_N-\mu\right)\right|(dx)\\
&\leq\int_0^1 ds\int_{\mathbb{R}^d}C\left(\left(1+|x|^\ell\right)\|s(\bar{\nu}_N-\mu)\|_0^\alpha+\left(1+|x|^{\ell(1-\alpha)}\right)\left(\int_{\mathbb{R}^d}|y|^\ell|s(\bar{\nu}_N-\mu)|(dy)\right)^\alpha\right)\left|\sqrt{N}\left(\bar{\nu}_N-\mu\right)\right|(dx)\\
&\leq C\left(|\bar{\nu}_N-\mu|\left(\mathbb{R}^d\right)\right)^\alpha\cdot\int_{\mathbb{R}^d}\left(1+|x|^\ell\right)\left|\sqrt{N}\left(\bar{\nu}_N-\mu\right)\right|(dx)\\
&\quad+C\left(\int_{\mathbb{R}^d}|y|^\ell|\bar{\nu}_N-\mu|(dy)\right)^\alpha\int_{\mathbb{R}^d}\left(1+|x|^{\ell(1-\alpha)}\right)\left|\sqrt{N}\left(\bar{\nu}_N-\mu\right)\right|(dx).
\end{aligned}$$

Since $\|\sqrt{N}\left(\bar{\nu}_N-\mu\right)-\sigma\|_\ell\underset{N\to\infty}{\longrightarrow}0$ implies $\|\left|\sqrt{N}\left(\bar{\nu}_N-\mu\right)\right|-|\sigma|\|_\ell\underset{N\to\infty}{\longrightarrow}0$, we can deduce that:

1. $\int_{\mathbb{R}^d}(1+|x|^\ell)|\sqrt{N}\,(\bar{\nu}_N-\mu)\,|(dx) \underset{N\to\infty}{\longrightarrow} \int_{\mathbb{R}^d}(1+|x|^\ell)\sigma(dx)<\infty$

2. $\int_{\mathbb{R}^d}(1+|x|^{\ell(1-\alpha)})|\sqrt{N}\,(\bar{\nu}_N-\mu)\,|(dx) \underset{N\to\infty}{\longrightarrow} \int_{\mathbb{R}^d}(1+|x|^{\ell(1-\alpha)})\sigma(dx)<\infty$

3. $|\bar{\nu}_N-\mu|\,(\mathbb{R}^d)=\dfrac{|\sqrt{N}\,(\bar{\nu}_N-\mu)\,|(\mathbb{R}^d)}{\sqrt{N}} \underset{N\to\infty}{\longrightarrow} 0$

4. $\int_{\mathbb{R}^d}|y|^\ell|\bar{\nu}_N-\mu|(dy)\leq\int_{\mathbb{R}^d}(1+|y|^\ell)|\bar{\nu}_N-\mu|(dy)=\dfrac{\int_{\mathbb{R}^d}(1+|y|^\ell)|\sqrt{N}(\bar{\nu}_N-\mu)|(dy)}{\sqrt{N}} \underset{N\to\infty}{\longrightarrow} 0.$

To conclude the proof, we have to study the convergence of $\sqrt{N}\,(U\,(\mu_N)-U\,(\bar{\nu}_N))$ in the decomposition (2.51).

Limit distribution of $\sqrt{N}\,(U\,(\mu_N)-U\,(\bar{\nu}_N))$

FIRST STEP Let us preliminary recall that by Remark 18, Theorem 16 implies $W_\ell\,(\mu_N,\mu)\underset{N}{\longrightarrow}0$ a.s.
Let us now define for $i=1,\cdots,N$ and $s\in[0,1]$

$$\mu_N^{i,s}=\frac{1}{N}\sum_{j=1}^{i-1}\delta_{X_j}+\frac{s}{N}\delta_{X_i}+\frac{(1-s)}{N}\nu_i+\frac{1}{N}\sum_{j=i+1}^{N}\nu_j.$$

We have $\mu_N^{N,1}=\mu_N$ and $\mu_N^{1,0}=\bar{\nu}_N$. Moreover since for $i=1,\cdots,N-1$

$$\mu_N^{i,1}=\mu_N^{i+1,0},$$

one has

$$\sqrt{N}\,(U\,(\mu_N)-U\,(\bar{\nu}_N))=\sqrt{N}\sum_{i=1}^{N}\left(U(\mu_N^{i,1})-U(\mu_N^{i,0})\right).$$

Let us now show that

$$\max_{1\leq i\leq N}\sup_{s\in[0,1]}W_\ell(\mu_N^{i,s},\mu)\underset{N\to\infty}{\longrightarrow}0\quad a.s.$$

Since for $i=1,\cdots,N$ and $s\in[0,1]$

$$\mu_N^{i,s}=s\mu_N^{i,1}+(1-s)\mu_N^{i-1,1}$$

under the convention $\mu_N^{0,1}=\mu_N^{1,0}$, we can proceed as in the first step of the proof of Theorem 25 to prove that

$$\max_{1\leq i\leq N}\sup_{s\in[0,1]}W_\ell\left(\mu_N^{i,s},\mu\right)=\max_{0\leq i\leq N}W_\ell\left(\mu_N^{i,1},\mu\right). \tag{2.52}$$

By the triangle inequality,

$$W_\ell\left(\mu_N^{i,1},\mu\right)\leq W_\ell\left(\mu_N^{i,1},\bar{\nu}_N\right)+W_\ell\left(\bar{\nu}_N,\mu\right) \tag{2.53}$$

where $W_\ell\,(\bar{\nu}_N,\mu)$ goes to 0 as N goes to infinity.
Therefore it is sufficient to prove that $\max_{0\leq i\leq N}W_\ell\left(\mu_N^{i,1},\bar{\nu}_N\right)\underset{N\to\infty}{\longrightarrow}0$ a.s. which we are now going to do.

For $i=1,\cdots,N$

$$\mu_N^{i,1} = \frac{i}{N}\mu_i + \frac{1}{N}\sum_{j=i+1}^{N}\nu_j.$$

Hence let $\gamma \in \Pi(\mu_i, \bar{\nu}_i)$ and define $\tilde{\gamma}(dx,dy) = \frac{i}{N}\gamma(dx,dy) + \frac{1}{N}\sum_{j=i+1}^{N}\nu_j(dx)\delta_x(dy)$: one has $\tilde{\gamma}(dx,\mathbb{R}^d) = \mu_N^{i,1}(dx)$ and $\tilde{\gamma}(\mathbb{R}^d,dy) = \bar{\nu}_N(dy)$. If $\ell > 0$ we have

$$\begin{aligned} W_\ell^{\ell\vee 1}\left(\mu_N^{i,1},\bar{\nu}_N\right) &\leq \int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\tilde{\gamma}(dx,dy) \\ &= \frac{i}{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\gamma(dx,dy) + \frac{1}{N}\sum_{j=i+1}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\nu_j(dx)\delta_x(dy) \\ &= \frac{i}{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^\ell\gamma(dx,dy). \end{aligned}$$

Taking the infimum over γ, one has

$$W_\ell\left(\mu_N^{i,1},\bar{\nu}_N\right) \leq \left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu_i,\bar{\nu}_i\right).$$

Finally let $\alpha \in (0,1)$ so that

$$\begin{aligned} \max_{0\leq i\leq N} W_\ell\left(\mu_N^{i,1},\bar{\nu}_N\right) &= \max_{1\leq i\leq N} W_\ell\left(\mu_N^{i,1},\bar{\nu}_N\right) \leq \max_{1\leq i\leq N}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu_i,\bar{\nu}_i\right) \\ &\leq \max_{1\leq i\leq N}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu_i,\mu\right) + \max_{1\leq i\leq N}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu,\bar{\nu}_i\right) \\ &\leq \max_{1\leq i\leq \lfloor\alpha N\rfloor}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu_i,\mu\right) + \max_{\lfloor\alpha N\rfloor<i\leq N}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu_i,\mu\right) \\ &\quad + \max_{1\leq i\leq \lfloor\alpha N\rfloor}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu,\bar{\nu}_i\right) + \max_{\lfloor\alpha N\rfloor<i\leq N}\left(\frac{i}{N}\right)^{\frac{1}{\ell\vee 1}} W_\ell\left(\mu,\bar{\nu}_i\right) \\ &\leq \alpha^{\frac{1}{\ell\vee 1}}\sup_i W_\ell\left(\mu_i,\mu\right) + \alpha^{\frac{1}{\ell\vee 1}}\sup_i W_\ell\left(\mu,\bar{\nu}_i\right) \\ &\quad + \max_{\lfloor\alpha N\rfloor<i\leq N} W_\ell\left(\mu_i,\mu\right) + \max_{\lfloor\alpha N\rfloor<i\leq N} W_\ell\left(\mu,\bar{\nu}_i\right). \end{aligned}$$

For fixed α, the sum of the two last terms goes to 0 as N goes to infinity while the sum of the first two terms is arbitrarily small for α small. Then it is possible to conclude that the left-hand side goes to 0 almost surely.
As explained in the first step of the Theorem 25, we can adapt the reasoning to the case $\ell = 0$.

SECOND STEP We can now reconsider the following stopping time

$$I_N = \min\left\{1\leq i\leq N : \exists s\in[0,1] : W_\ell\left(\mu_N^{i,s},\mu\right)\geq r\right\}\wedge(N+1)$$

for the filtration $(\mathcal{F}_i = \sigma(X_1,\cdots,X_i))_{i\geq 1}$ where by convention $\min\emptyset = +\infty$. Proceeding exactly as in the second step of the proof of the Theorem 25 and keeping the same notations, we can obtain the following decomposition

$$\sqrt{N}\left(U\left(\mu_N\right) - U\left(\bar{\nu}_N\right)\right) = \sqrt{N}\left(U\left(\mu_N\right) - U\left(\bar{\nu}_N\right) - Q_N\right) + \sqrt{N}Q_N$$

where

$$Q_N := \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\delta U}{\delta m}(\mu_N^{i\wedge I_N,0},X_i) - \int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu_N^{i\wedge I_N,0},x)\nu_i(dx)\right)$$

and for $N \geq N^*$, $U(\mu_N) - U(\bar{\nu}_N) - Q_N$ coincides with

$$R_N = \frac{1_{N\geq N^*}}{N} \sum_{i=1}^{N} \int_0^1 ds \int_{\mathbb{R}^d} \left(\frac{\delta U}{\delta m}(\mu_N^{i,s}, x) - \frac{\delta U}{\delta m}(\mu_N^{i,0}, x) \right) (\delta_{X_i} - \nu_i)(dx).$$

THIRD STEP Let us first prove the convergence in L^1 of $\sqrt{N}R_N$ to 0. By applying the same argument as in the proof of Theorem 25 this is sufficient to deduce that $\sqrt{N}(U(\mu_N) - U(\bar{\nu}_N) - Q_N)$ goes to 0 in probability.
By Assumption **RU5**, $\exists C < \infty$, $\exists \alpha \in \left(\frac{1}{2}, 1\right]$ such that for $N \geq N^*$

$$\begin{aligned}
&\left| \frac{\delta U}{\delta m}(\mu_N^{i,s}, x) - \frac{\delta U}{\delta m}(\mu_N^{i,0}, x) \right| \\
&\quad \leq C \left((1 + |x|^\ell) \left\| \mu_N^{i,s} - \mu_N^{i,0} \right\|_0^\alpha + (1 + |x|^{\ell(\alpha-1)}) \left(\int_{\mathbb{R}^d} |y|^\ell |\mu_N^{i,s} - \mu_N^{i,0}|(dy) \right)^\alpha \right)
\end{aligned}$$

with

$$|\mu_N^{i,s} - \mu_N^{i,0}|(dy) \leq \frac{s}{N}(\delta_{X_i} + \nu_i)(dy)$$

and

$$\left\| \mu_N^{i,s} - \mu_N^{i,0} \right\|_0 = |\mu_N^{i,s} - \mu_N^{i,0}|(\mathbb{R}^d) = \frac{s}{N}|\delta_{X_i} - \nu_i|(\mathbb{R}^d) \leq \frac{s}{N}(\delta_{X_i} + \nu_i)(\mathbb{R}^d) = \frac{2s}{N}.$$

Substituting the above quantities and using Young's inequality, one obtains

$$\begin{aligned}
&\left| \frac{\delta U}{\delta m}(\mu_N^{i,s}, x) - \frac{\delta U}{\delta m}(\mu_N^{i,0}, x) \right| \\
&\leq C \left((1 + |x|^\ell) \left(\frac{2s}{N} \right)^\alpha + (1 + |x|^{\ell(1-\alpha)}) \left(\int_{\mathbb{R}^d} |y|^\ell (\delta_{X_i} + \nu_i)(dy) \right)^\alpha \left(\frac{s}{N} \right)^\alpha \right) \\
&\leq \frac{2^\alpha C}{N^\alpha} \left((1 + |x|^\ell) + \alpha \left(|X_i|^\ell + \int_{\mathbb{R}^d} |y|^\ell \nu_i(dy) \right) + (1 - \alpha) + \alpha \left(|X_i|^\ell + \int_{\mathbb{R}^d} |y|^\ell \nu_i(dy) \right) + |x|^\ell (1 - \alpha) \right) \\
&\leq \frac{2^\alpha C}{N^\alpha} \left((1 + |x|^\ell) + \left(|X_i|^\ell + \int_{\mathbb{R}^d} |y|^\ell \nu_i(dy) \right) + \frac{1}{2} + \left(|X_i|^\ell + \int_{\mathbb{R}^d} |y|^\ell \nu_i(dy) \right) + \frac{|x|^\ell}{2} \right) \\
&= \frac{2^\alpha C}{N^\alpha} \left(\frac{3}{2}(1 + |x|^\ell) + 2|X_i|^\ell + 2 \int_{\mathbb{R}^d} |y|^\ell \nu_i(dy) \right).
\end{aligned}$$

Therefore

$$\begin{aligned}
|R_N| &\leq \frac{1_{N\geq N^*}}{N} \sum_{i=1}^{N} \int_0^1 ds \int_{\mathbb{R}^d} \left| \frac{\delta U}{\delta m}(\mu_N^{i,s}, x) - \frac{\delta U}{\delta m}(\mu_N^{i,0}, x) \right| (\delta_{X_i} + \nu_i)(dx) \\
&\leq \frac{2^\alpha C}{N^{\alpha+1}} \sum_{i=1}^{N} \int_{\mathbb{R}^d} \left(\frac{3}{2}(1 + |x|^\ell) + 2|X_i|^\ell + 2 \int_{\mathbb{R}^d} |y|^\ell \nu_i(dy) \right) (\delta_{X_i} + \nu_i)(dx) \\
&= \frac{2^\alpha C}{N^{\alpha+1}} \sum_{i=1}^{N} \left(3 + \frac{11}{2} \int_{\mathbb{R}^d} |y|^\ell \delta_{X_i}(dy) + \frac{11}{2} \int_{\mathbb{R}^d} |y|^\ell \nu_i(dy) \right) \\
&= \frac{C_1}{N^\alpha} + \frac{C_2}{N^\alpha} \int_{\mathbb{R}^d} |y|^\ell \mu_N(dy) + \frac{C_3}{N^\alpha} \int_{\mathbb{R}^d} |y|^\ell \bar{\nu}_N(dy)
\end{aligned}$$

for some positive constants C_1, C_2 and C_3.
Finally

$$\mathbb{E}(\sqrt{N}|R_N|) \leq \frac{C_1}{N^{\alpha-\frac{1}{2}}} + \frac{C_2}{N^{\alpha-\frac{1}{2}}} \int_{\mathbb{R}^d} |y|^\ell \bar{\nu}_N(dy) + \frac{C_3}{N^{\alpha-\frac{1}{2}}} \int_{\mathbb{R}^d} |y|^\ell \bar{\nu}_N(dy)$$

and since $\alpha > \frac{1}{2}$ and $\lim_{N\to\infty} \int_{\mathbb{R}^d} |y|^\ell \bar{\nu}_N(dy) = \int_{\mathbb{R}^d} |y|^\ell \mu(dy)$, we can conclude that the left-hand side goes to 0 as N goes to infinity and so the third step is concluded.

FOURTH STEP We are now going to prove the convergence in distribution of $\sqrt{N}Q_N$ to a Gaussian random variable. In this case, we can immediately apply the Central Limit Theorem for martingales while in Theorem 25 we had to introduce first the Poisson equation before applying it to $K_{3,N}$. For $1 \leq i \leq N$ let

$$Y_{N,i} = \frac{1}{\sqrt{N}} \left(\frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, X_i \right) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right) \nu_i(dx) \right).$$

Let us start by **checking that** $\mathbb{E}\left(Y_{N,i}|\mathcal{F}_{i-1}\right) = 0$**.**

Since X_i is independent of $\mathcal{F}_{i-1}$ and $\mu_N^{i\wedge I_N,0} = \sum_{j=1}^{i-1} 1_{\{I_N=j\}} \mu_N^{j,0} + 1_{\{I_N>i-1\}} \mu_N^{i,0}$ is $\mathcal{F}_{i-1}$- measurable, applying the Freezing Lemma one has

$$\begin{aligned}
\mathbb{E}\left(Y_{N,i}|\mathcal{F}_{i-1}\right) &= \mathbb{E}\left(\frac{1}{\sqrt{N}} \left(\frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, X_i \right) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right) \nu_i(dx) \right) |\mathcal{F}_{i-1} \right) \\
&= \frac{1}{\sqrt{N}} \left(\mathbb{E}\left(\frac{\delta U}{\delta m}(t, X_i) \right)_{t=\mu_N^{i\wedge I_N,0}} - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right) \nu_i(dx) \right) = 0.
\end{aligned}$$

For what concerns the **square integrability**, let us check that both the components are squared integrable by applying Hypothesis **RU2** and the fact that $\nu_i \in \mathbb{P}_\ell\left(\mathbb{R}^d\right)$.

- $$\mathbb{E}\left(\frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, X_i \right)^2 \right) \leq C^2 \mathbb{E}\left(\left(1 + |X_i|^{\frac{\ell}{2}} \right)^2 \right) \leq 2C^2 \mathbb{E}\left(1 + |X_i|^\ell \right) < \infty$$

- $$\begin{aligned}
\mathbb{E}\left(\left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right) \nu_i(dx) \right)^2 \right) &\leq \mathbb{E}\left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right)^2 \nu_i(dx) \right) \\
&\leq 2C^2 \left(1 + \int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) \right) < \infty.
\end{aligned}$$

Let us now study the **convergence of** $\sum_{i=1}^N \mathbb{E}\left(Y_{N,i}^2|\mathcal{F}_{i-1}\right)$:

$$\begin{aligned}
\sum_{i=1}^N \mathbb{E}\left(Y_{N,i}^2|\mathcal{F}_{i-1}\right) &= \frac{1}{N} \sum_{i=1}^N \mathbb{E}\left(\left(\frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, X_i \right) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right) \nu_i(dx) \right)^2 |\mathcal{F}_{i-1} \right) \\
&= \frac{1}{N} \sum_{i=1}^N \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right)^2 \nu_i(dx) - \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right) \nu_i(dx) \right)^2 \right).
\end{aligned}$$

As in the fourth step of the proof of the Theorem 25, thanks to Hypothesis **RU3** and thanks to the a.s. convergence of $\max_{1\leq i\leq N} W_\ell(\mu_N^{i,0}, \mu)$ to 0, we obtain the existence of a sequence of random variables $(\epsilon_N)_{N\geq 0}$ converging a.s. to 0 such that

$$\forall 1 \leq i \leq N \qquad \left| \frac{\delta U}{\delta m} \left(\mu_N^{i\wedge I_N,0}, x \right) - \frac{\delta U}{\delta m}(\mu, x) \right| \leq \left(1 + |x|^{\frac{\ell}{2}} \right) \epsilon_N. \tag{2.54}$$

First Component Let us first show that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)^2\nu_i(dx)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)^2\,\mu(dx)=0\ \ a.s.$$

We can rewrite the difference in the following way

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)^2\nu_i(dx)-\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)^2\,\nu_i(dx)\\
&\quad+\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)^2\,\nu_i(dx)-\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)^2\,\mu(dx)\\
&=\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}\left(\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)^2-\frac{\delta U}{\delta m}(\mu,x)^2\right)\nu_i(dx)\\
&\quad+\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)^2\,(\bar\nu_N-\mu)\,(dx).
\end{aligned}$$

Since $\lim_{N\to\infty}W_\ell(\bar\nu_N,\mu)=0$, by the characterization (2.7), we can deduce that the second term of the sum goes to 0 thanks to Assumption **RU2** and Assumption **RU4**. For what concerns the first term for $i=1,\cdots,N$, using Assumption **RU2** and (2.54), one has

$$\begin{aligned}
\left|\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)^2-\frac{\delta U}{\delta m}(\mu,x)^2\right|&\le\left(\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)-\frac{\delta U}{\delta m}(\mu,x)\right)^2\\
&\quad+2\left|\frac{\delta U}{\delta m}(\mu,x)\right|\left|\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)-\frac{\delta U}{\delta m}(\mu,x)\right|\\
&\le\left(1+|x|^{\frac{\ell}{2}}\right)^2\epsilon_N^2+2C\left(1+|x|^{\frac{\ell}{2}}\right)^2\epsilon_N\\
&\le 2\left(1+|x|^{\ell}\right)\epsilon_N\left(\epsilon_N+2C\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\left|\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}\left(\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)^2-\frac{\delta U}{\delta m}(\mu,x)^2\right)\nu_i(dx)\right|&\le 2\epsilon_N\left(\epsilon_N+2C\right)\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d}\left(1+|x|^{\ell}\right)\nu_i(dx)\\
&=2\epsilon_N\left(\epsilon_N+2C\right)\left(1+\int_{\mathbb{R}^d}|x|^{\ell}\bar\nu_N(dx)\right)
\end{aligned}$$

where the right-hand side goes to 0 a.s. since $\lim_{N\to\infty}\int_{\mathbb{R}^d}|x|^{\ell}\bar\nu_N(dx)=\int_{\mathbb{R}^d}|x|^{\ell}\mu(dx)$.

Second Component We are going to prove that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)\nu_i(dx)\right)^2-\int_{\mathbb{R}^d\times\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\frac{\delta U}{\delta m}(\mu,y)\eta(dx,dy)=0\ a.s.$$

Let us rewrite the difference in the following way

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\left(\left(\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}\left(\mu_N^{i\wedge I_N,0},x\right)\nu_i(dx)\right)^2-\left(\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\,\nu_i(dx)\right)^2\right)\\
&+\left(\frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\frac{\delta U}{\delta m}(\mu,y)\,\nu_i(dx)\nu_i(dy)-\int_{\mathbb{R}^d\times\mathbb{R}^d}\frac{\delta U}{\delta m}(\mu,x)\frac{\delta U}{\delta m}(\mu,y)\eta(dx,dy)\right)
\end{aligned}$$

Since Hypothesis 1 holds, again by the characterization (2.7), we can deduce that the second term of the sum goes to 0 a.s. thanks to Hypothesis **RU2** and Hypothesis **RU4**. As far as the first term is concerned, using Hypothesis **RU2** and (2.54), one has that for $i = 1, \cdots, N$

$$
\begin{aligned}
&\left| \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right) \nu_i(dx) \right)^2 - \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} (\mu, x) \, \nu_i(dx) \right)^2 \right| \\
&\leq \left(\int \left(\frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right) - \frac{\delta U}{\delta m} (\mu, x) \right) \nu_i(dx) \right)^2 \\
&\quad + 2 \left| \int \frac{\delta U}{\delta m} (\mu, x) \, \nu_i(dx) \right| \left| \int \left(\frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right) - \frac{\delta U}{\delta m} (\mu, x) \right) \nu_i(dx) \right| \\
&\leq \int \left(\frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right) - \frac{\delta U}{\delta m} (\mu, x) \right)^2 \nu_i(dx) \\
&\quad + 2 \int \left| \frac{\delta U}{\delta m} (\mu, x) \right| \nu_i(dx) \int_{\mathbb{R}^d} \left| \left(\frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right) - \frac{\delta U}{\delta m} (\mu, x) \right) \right| \nu_i(dx) \\
&\leq 2\epsilon_N^2 \int_{\mathbb{R}^d} (1 + |x|^\ell) \nu_i(dx) + 2C\epsilon_N \int_{\mathbb{R}^d} \left(1 + |x|^{\frac{\ell}{2}} \right) \nu_i(dx) \int_{\mathbb{R}^d} \left(1 + |x|^{\frac{\ell}{2}} \right) \nu_i(dx) \\
&\leq 2\epsilon_N^2 \int_{\mathbb{R}^d} (1 + |x|^\ell) \nu_i(dx) + 4C\epsilon_N \int_{\mathbb{R}^d} \left(1 + |x|^\ell \right) \nu_i(dx) = \left(2\epsilon_N^2 + 4C\epsilon_N \right) \left(1 + \int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) \right).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&\frac{1}{N} \sum_{i=1}^N \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right) \nu_i(dx) \right)^2 - \frac{1}{N} \sum_{i=1}^N \left(\int_{\mathbb{R}^d} \frac{\delta U}{\delta m} (\mu, x) \, \nu_i(dx) \right)^2 \\
&\qquad \leq \left(2\epsilon_N^2 + 4C\epsilon_N \right) \left(1 + \int_{\mathbb{R}^d} |x|^\ell \, \bar{\nu}_N(dx) \right)
\end{aligned}
$$

and the left-hand side goes to 0 since $\lim_{N \to \infty} \int_{\mathbb{R}^d} |x|^\ell \bar{\nu}_N(dx) = \int_{\mathbb{R}^d} |x|^\ell \mu(dx)$.

To finally apply the Central Limit Theorem for martingales (see Corollary 3.1 [31]) to conclude that

$$
\sqrt{N} Q_N = \sum_{i=1}^N Y_{N,i} \underset{d}{\Longrightarrow} \mathcal{N} \left(0, \int_{\mathbb{R}^d} \frac{\delta U}{\delta m} (\mu, x)^2 \, \mu(dx) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{\delta U}{\delta m} (\mu, x) \frac{\delta U}{\delta m} (\mu, y) \eta(dx, dy) \right),
$$

it remains just to verify that the **Lindeberg condition** holds.

We need to check that for $\epsilon > 0$, $\sum_{i=1}^N \mathbb{E} \left(Y_{N,i}^2 1_{\{Y_{N,i}^2 > \epsilon\}} | \mathcal{F}_{i-1} \right)$ goes to 0 in probability as $N \to \infty$. By Jensen's inequality and Hypothesis **RU2**, one has

$$
\begin{aligned}
N Y_{N,i}^2 &= \left(\frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, X_i \right) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right) \nu_i(dx) \right)^2 \\
&\leq 2 \frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, X_i \right)^2 + 2 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m} \left(\mu_N^{i \wedge I_N, 0}, x \right)^2 \nu_i(dx) \\
&\leq 4C^2 \left(1 + |X_i|^\ell \right) + 4C^2 \int_{\mathbb{R}^d} \left(1 + |x|^\ell \right) \nu_i(dx) = 4C^2 \left(2 + |X_i|^\ell + \int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) \right).
\end{aligned}
$$

Therefore, as in the fourth step of the proof of the Theorem 25, it is enough to check that for each $\epsilon > 0$

$$2 \times 1_{\{\frac{2}{N}>\epsilon\}} + \sum_{i=1}^{N} \mathbb{E}\left(\frac{|X_i|^\ell}{N} 1_{\{|X_i|^\ell > N\epsilon\}} \,\middle|\, \mathcal{F}_{i-1} \right) + \sum_{i=1}^{N} \frac{\int_{\mathbb{R}^d} |x|^\ell \nu_i(dx)}{N} 1_{\{\int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) > N\epsilon\}}$$

goes to 0 as N goes to infinity (a.s.).
It is immediate to prove that the **first component** goes to 0 while the **second component** goes to 0 by the independence of X_i by $\mathcal{F}_{i-1}$ and (2.49). For what concerns the **third component**, we have that $\bar{y}_N := \frac{1}{N} \sum_{i=1}^{N} \int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) \to \int_{\mathbb{R}^d} |x|^\ell \mu(dx)$. Therefore

$$\frac{\int_{\mathbb{R}^d} |x|^\ell \nu_N(dx)}{N} = \bar{y}_N - \frac{N-1}{N} \bar{y}_{N-1} \underset{N\to\infty}{\longrightarrow} 0$$

that implies $\forall \epsilon > 0$

$$\int_{\mathbb{R}^d} |x|^\ell \nu_N(dx) 1_{\{\int_{\mathbb{R}^d} |x|^\ell \nu_N(dx) > N\epsilon\}} \underset{N\to\infty}{\longrightarrow} 0$$

and taking the Cesaro mean we deduce that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) 1_{\{\int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) > N\epsilon\}} \leq \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) 1_{\{\int_{\mathbb{R}^d} |x|^\ell \nu_i(dx) > i\epsilon\}} = 0$$

and so the proof is concluded.

□

Chapter 3

Convergence to the uniform distribution of vectors of partial sums modulo one with a common factor

The material for this chapter has been released in [27].

In this work, we prove the joint convergence in distribution of q variables modulo one obtained as partial sums of a sequence of i.i.d. square integrable random variables multiplied by a common factor given by some function of an empirical mean of the same sequence. The limit is uniformy distributed over $[0,1]^q$. To deal with the coupling introduced by the common factor, we assume that the joint distribution of the random variables has a non zero component absolutely continuous with respect to the Lebesgue measure, so that the convergence in the central limit theorem for this sequence holds in total variation distance. While our result provides a generalization of Benford's law to a data adapted mantissa, our main motivation is the derivation of a central limit theorem for the stratified resampling mechanism, which is performed in the companion paper [25].

3.1 Introduction

Given $(Y_i)_{i\geq 1}$ a sequence of i.i.d. square integrable random variables and a measurable real valued function ϕ, we are going to give sufficient conditions for the convergence in distribution to the uniform law on $[0,1]$ of

$$\left\{\phi\left(\frac{1}{M}\sum_{i=1}^{\beta_M} Y_i\right)(Y_1+\cdots+Y_M)\right\}, \tag{3.1}$$

for $(\beta_M)_{M\geq 1}\subseteq \mathbb{N}^*$ such that $\lim_{M\to\infty}\sqrt{M}\left(\frac{\beta_M}{M}-\beta\right)=0$ with $\beta>1$. Here $\{x\}$ denotes the fractional part of the real number x given by $\{x\}=x-\lfloor x\rfloor$ where $\lfloor x\rfloor$ is the integer such that $\lfloor x\rfloor\leq x<\lfloor x\rfloor+1$ and we also define $\lceil x\rceil$ as the integer such that $\lceil x\rceil-1<x\leq\lceil x\rceil$. Our main motivation for considering (3.1) comes from the derivation of a central limit theorem for the stratified resampling mechanism. Before giving more details about this particular application, let us review the existing literature which addresses the case of a constant function ϕ with the derivation of Benford's law as a common motivation.

The convergence in distribution of the sequence $(V_M=\{Y_1+\cdots+Y_M\})_{M\geq 1}$ of sums of random variables defined modulo 1 to the uniform distribution on $[0,1]$ has been studied by many researchers using Fourier analysis. In 1939, Lévy [39] gave necessary and sufficient conditions for this convergence when the Y_i are i.i.d.. In 1986, Störmer [58] provided sufficient conditions in terms of the distribution functions of the Y_i for the convergence to hold when the Y_i are merely independent. In

2007, under the assumption of independent absolutely continuous Y_i, Miller and Nigrini [44] characterized the convergence of $(V_M)_{M\geq 1}$ in $L^1([0,1])$. In 2010, Szewczak [59] generalized the above results by getting rid of the hypothesis of independence of the Y_i. In particular he proved that, if the absolute value of the characteristic function of $Y_1+\cdots+Y_M$ satisfies a certain growth condition and if the set $\left\{n\in\mathbb{N}\setminus\{0\} : \left|\mathbb{E}\left(e^{2\pi i n Y_1}\right)\right|=1\right\}$ is empty, then $(V_M)_{M\geq 1}$ converges in distribution to a uniform random variable on $[0,1]$.

Let us now briefly introduce Benford's law and the problem of the distribution of the leading digits of products of random variables (see for instance [18]). Benford's law in base $b>1$ is the probability measure μ_b on the interval $[1,b)$ defined by

$$\mu_b\left([1,a)\right)=\log_b a, \qquad \forall a\in[1,b).$$

The mantissa in base b of a positive real number x is the unique number $\mathcal{M}_b(x)$ in $[1,b)$ such that $x=\mathcal{M}_b(x)\times b^{\lfloor \log_b(x)\rfloor}$. Given a sequence of positive random variables $(X_i)_{i\geq 1}$, many researchers ([54], [9], [55], [40]) were interested in studying the weak convergence as $M\to\infty$ of the law of $\mathcal{M}_b\left(\prod_{i=1}^{M} X_i\right)$ to μ_b. Since $\log_b \mathcal{M}_b(x)=\{\log_b(x)\}$ for each positive real number x, one has

$$\log_b \mathcal{M}_b\left(\prod_{i=1}^{M} X_i\right)=\left\{\sum_{i=1}^{M}\log_b(X_i)\right\},$$

and this convergence is equivalent to the weak convergence of the partial sums modulo 1 of the random variables $(\log_b(X_i))_{i\geq 1}$ by continuity of $[1,b)\ni z\mapsto \log_b(z)$ and its inverse.

The introduction of a non-constant function ϕ in our work permits to address the choice of a data dependent mantissa for instance given by the geometric mean $\hat{b}_M=\exp\left(\frac{1}{\beta M}\sum_{i=1}^{\beta_M}\ln(X_i)\right)$. Indeed,

$$\log_{\hat{b}_M}\mathcal{M}_{\hat{b}_M}\left(\prod_{i=1}^{M}X_i\right)=\left\{\frac{\sum_{i=1}^{M}\ln(X_i)}{\frac{1}{\beta M}\sum_{i=1}^{\beta_M}\ln(X_i)}\right\}=\left\{\phi\left(\frac{1}{M}\sum_{i=1}^{\beta_M}\ln(X_i)\right)\sum_{i=1}^{M}\ln(X_i)\right\}$$

where $\phi(x)=\frac{1}{\beta x}$. Compared to previous works, the main difficulty that we have to address comes from the coupling between the variables introduced through the common factor $\phi\left(\frac{1}{M}\sum_{i=1}^{\beta_M}Y_i\right)$. To overcome this difficulty, we assume that the Y_i are i.i.d. according to a common distribution with a non-zero component absolutely continuous with respect to the Lebesgue measure. This allows us to apply one of the key results for our proof, the convergence in total variation in the central limit theorem, that we now recall. Let F be a centered square-integrable random vector in $\mathbb{R}^n$ with identity covariance matrix and let $(F_i)_{i\geq 1}$ independent copies of F. Under the assumption that the law of F has an absolutely continuous component, Prohorov [50] in the one-dimensional case $n=1$ and Mamatov et Halikov [42] in the multidimensional case, proved that the total variation distance between the distribution of $\frac{1}{\sqrt{M}}\sum_{i=1}^{M}F_i$ and the standard Gaussian law in $\mathbb{R}^n$ goes to 0 as $M\to\infty$.

The study of the joint convergence of several partial sums modulo 1 with a common factor and an additional component satisfying a central limit theorem does not add further significant difficulties and is useful in the derivation of the central limit theorem for the stratified resampling mechanism. That is why we address the convergence in distribution of

$$\left(\left(\left\{\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}}Y_m\right)\left(Y_1+\cdots+Y_{\beta_M^i}\right)\right\}\right)_{1\leq i\leq q}, \sqrt{M}\left(\phi\left(\frac{1}{M}\sum_{m=1}^{\beta_M^{q+1}}Y_m\right)\frac{1}{M}\sum_{i=1}^{\beta_M^{q+1}}Z_i-\theta\right)\right) \tag{3.2}$$

for $(\beta_M^1, \cdots, \beta_M^{q+1})_{M\geq 1} \subseteq \mathbb{N}^{q+1}$ such that $\lim_{M\to\infty} \sqrt{M}\left(\frac{\beta_M^i}{M} - \beta^i\right) = 0$ with $0 < \beta^1 < \cdots < \beta^{q+1}$ and $(Z_i)_{i\geq 1}$ such that the sequence $((Y_i, Z_i))_{i\geq 1}$ also is i.i.d. and the last component where θ is a constant (typically equal to $\beta^{q+1}\phi\left(\beta^{q+1}\mathbb{E}(Y_1)\right)\mathbb{E}(Z_1)$) converges in distribution to T as $M \to \infty$. We give mild conditions ensuring that the full vectors converge in distribution to (U, T) where U is uniformly distributed on $[0,1]^q$ and independent of T.

Main motivation

Let us now come back to our main motivation. We are interested in providing a central limit theorem for the stratified resampling scheme under the simplifying assumption that the $\mathbb{R}^d$-valued initial draws $(X_m)_{m\geq 1}$ are independent and identically distributed and weighted proportionally to their image by some measurable function $g : \mathbb{R}^d \to (0, +\infty)$. For $M \geq 1$, resampling schemes (see for instance [19]) permit to replace the probability measure $\frac{\sum_{m=1}^M g(X_m)\delta_{X_m}}{\sum_{m=1}^M g(X_m)}$ with non equal weights by some empirical measure $\frac{1}{M}\sum_{m=1}^M \delta_{\xi_m^M}$ with the same conditional expectation given $\mathcal{F} = \sigma((X_m)_{m\geq 1})$. For $(U_m)_{m\geq 1}$ an independent sequence of independent random variables uniformly distributed on $(0,1)$, the stratified resampling scheme consists in setting

$$\xi_m^M = \sum_{i=1}^M 1_{\{\bar{S}_{i-1}^M < m - U_m \leq \bar{S}_i^M\}} X_i \text{ for } m \in \{1, \cdots, M\}$$

with $\bar{S}_j^M = \frac{M\sum_{m=1}^j g(X_m)}{\sum_{m=1}^M g(X_m)}$ for $j \in \{1, \cdots, M\}$, under the convention $\bar{S}_0^M = 0$. For $f : \mathbb{R}^d \to \mathbb{R}$ measurable, we have

$$\mathbb{E}\left(f(\xi_m^M)|\mathcal{F}\right) = \sum_{i=1}^M f(X_i) \int_{m-1}^m 1_{\{\bar{S}_{i-1}^M < u \leq \bar{S}_i^M\}} du. \tag{3.3}$$

The central limit theorem deals with the asymptotic behaviour of $\frac{1}{\sqrt{M}}\sum_{m=1}^M \left(f(\xi_m^M) - \frac{\mathbb{E}(f(X_1)g(X_1))}{\mathbb{E}(g(X_1))}\right)$ as $M \to \infty$. Let us explain how the computation of the asymptotic variance is related to our main result. By a standard decomposition of the variance and the conditional independence of $(\xi_m^M)_{1\leq m\leq M}$ given $\mathcal{F}$,

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(\xi_m^M)\right) = \mathrm{Var}\left(\mathbb{E}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(\xi_m^M) \,\middle|\, \mathcal{F}\right)\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(\xi_m^M) \,\middle|\, \mathcal{F}\right)\right) \\
&= \mathrm{Var}\left(\sqrt{M}\frac{\sum_{m=1}^M g(X_m)f(X_m)}{\sum_{m=1}^M g(X_m)}\right) + \mathbb{E}\left(\frac{\sum_{m=1}^M g(X_m)f^2(X_m)}{\sum_{m=1}^M g(X_m)}\right) - \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^M \mathbb{E}\left(f(\xi_m^M) \,\middle|\, \mathcal{F}\right)^2\right).
\end{aligned}$$

Since the asymptotic behaviour as $M \to \infty$ of the first two terms in the right-hand side can be analysed using standard arguments, we focus on that of the third term. Using (3.3) and $\bar{S}_M^M = M$, we get that $\sum_{m=1}^M \mathbb{E}\left(f(\xi_m^M)|\mathcal{F}\right)^2$ is equal to

$$\sum_{k=0}^{M-1}\left(1 + 1_{\{k\geq 1\}}\right)\sum_{i=1}^{M-k} f(X_i)f(X_{i+k}) \sum_{m\geq \lfloor \bar{S}_{i-1}^M \rfloor + 1} \int_{m-1}^m 1_{\{\bar{S}_{i-1}^M < u \leq \bar{S}_i^M\}} du \int_{m-1}^m 1_{\{\bar{S}_{i+k-1}^M < u \leq \bar{S}_{i+k}^M\}} du.$$

Since $\frac{1}{M}\sum_{i=1}^M \gamma_i = \int_0^1 \gamma_{\lceil \alpha M \rceil} d\alpha$, we deduce that $\frac{1}{M}\sum_{m=1}^M \mathbb{E}\left(f(\xi_m^M)|\mathcal{F}\right)^2$ is equal to the sum over $k \in \mathbb{N}$ of

$$\left(1 + 1_{\{k\geq 1\}}\right)\int_0^1 1_{\{\lceil \alpha M \rceil \leq M-k\}} f(X_{\lceil \alpha M \rceil})f(X_{\lceil \alpha M \rceil + k})\psi_k\left(\left\{\bar{S}_{\lceil \alpha M \rceil - 1}^M\right\}, \frac{Mg\left(X_{\lceil \alpha M \rceil}\right)}{\sum_{m=1}^M g(X_m)}, \cdots, \frac{Mg\left(X_{\lceil \alpha M \rceil + k}\right)}{\sum_{m=1}^M g(X_m)}\right) d\alpha$$

where $\psi_k(u_0, w_1, \cdots, w_{k+1}) = \sum_{m\geq 1} \int_{m-1}^{m} 1_{\{u_0<u\leq u_0+w_1\}} du \int_{m-1}^{m} 1_{\left\{u_0+\sum_{\ell=1}^{k} w_\ell < u \leq u_0+\sum_{\ell=1}^{k+1} w_\ell\right\}} du$. That is why, in order to compute the asymptotic variance of $\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(\xi_m^M)$, it is very useful to understand the behaviour as $M \to \infty$ of $\left\{\bar{S}^M_{\lceil \alpha M\rceil -1}\right\}$ which is equal to the first component in (3.2) when $q = 1$, $Y_i = g(X_i), \phi(x) = \frac{1}{x}$, $\beta_M^1 = \lceil \alpha M\rceil - 1$ and $\beta_M^2 = M$.

Regarding the Central Limit Theorem, the characteristic function of

$$\frac{1}{\sqrt{M}} \sum_{m=1}^{M} \left(f(\xi_m^M) - \frac{\mathbb{E}\left(f(X_1) g(X_1)\right)}{\mathbb{E}(g(X_1))} \right)$$

writes for $u \in \mathbb{R}$

$$\mathbb{E}\left(e^{iu\sqrt{M}\left(\frac{\sum_{m=1}^{M} g(X_m) f(X_m)}{\sum_{m=1}^{M} g(X_m)} - \frac{\mathbb{E}(f(X_1)g(X_1))}{\mathbb{E}(g(X_1))}\right)} \prod_{m=1}^{M} \mathbb{E}\left(e^{\frac{iu}{\sqrt{M}}\left(f(\xi_m^M) - \mathbb{E}\left(f(\xi_m^M)|\mathcal{F}\right)\right)} \,\middle|\, \mathcal{F}\right)\right).$$

Using some Taylor expansion of $e^{\frac{iu}{\sqrt{M}}\left(f(\xi_m^M) - \mathbb{E}\left(f(\xi_m^M)|\mathcal{F}\right)\right)}$, one may approximate the product by a sum of integrals of functions of

$$\left(\left\{\bar{S}^M_{\lceil \alpha_1 M\rceil -1}\right\}, \left\{\bar{S}^M_{\lceil \alpha_2 M\rceil -1}\right\}, \cdots, \left\{\bar{S}^M_{\lceil \alpha_q M\rceil -1}\right\}\right)$$

with respect to $1_{\{0<\alpha_1<\alpha_2<\cdots<\alpha_q<1\}} d\alpha_1 d\alpha_2 \cdots d\alpha_q$. This closely relates the asymptotic behaviour of the characteristic function to the one of the vectors (3.2) for the choice $Y_i = g(X_i)$, $\phi(x) = \frac{1}{x}$, $Z_i = g(X_i) f(X_i)$, $\beta_M^i = \lceil \alpha_i M\rceil - 1$ for $i \in \{1, \cdots, q\}$ and $\beta_M^{q+1} = M$. In the companion paper [25], using the main result of the present paper, we compute the asymptotic variance and prove the associated central limit theorem.

3.2 Notation

We denote by $\mathbb{N}^*$ the set of natural numbers without 0 and by $\mathbb{R}_+$ the set of non negative real numbers. We denote by $\lfloor x\rfloor$ the integer j such that $j \leq x < j+1$ and by $\{x\} = x - \lfloor x\rfloor$ the fractional part of $x \in \mathbb{R}$.
We denote by μ_Y the law of a $\mathbb{R}^d$-valued random vector $Y = \left(Y^1, \cdots, Y^d\right)$ where $d \in \mathbb{N}^*$ and by ϕ_Y its characteristic function given by $\phi_Y(u) = \mathbb{E}(e^{iu^T Y}) = \int_{\mathbb{R}^d} e^{iu^T y} \mu_Y(dy)$, $u \in \mathbb{R}^d$. The convergence in distribution is denoted by $\stackrel{d}{\Longrightarrow}$. Moreover let $\mu_Y = \mu_{Y,c} + \mu_{Y,s}$ denote the decomposition of μ_Y into a part $\mu_{Y,c}$ absolutely continuous with respect to the Lebesgue measure and a singular part $\mu_{Y,s}$. Notice that there exists A, a Borel subset of $\mathbb{R}^d$, such that $\mu_{Y,s}(A) = 0$ and $\int_{\mathbb{R}^d} 1_{\{x\notin A\}} dx = 0$. Let p_Y denote a density of $\mu_{Y,c}$ with respect to the Lebesgue measure. In what follows, we will always consider absolute continuity with respect to the Lebesgue measure and so we avoid to write it everytime. Moreover, when we write that the law of a random variable has an absolutely continuous component we mean a non zero absolutely continuous component. We denote by $m_Y = \left(\mathbb{E}\left(Y^1\right), \cdots, \mathbb{E}\left(Y^d\right)\right)$ the expected value vector of Y.
For the total variation distance between the measures μ_1 and μ_2, we write

$$d_{TV}(\mu_1, \mu_2) = \sup_{A\in\mathcal{B}\left(\mathbb{R}^d\right)} |\mu_1(A) - \mu_2(A)| = \frac{1}{2} \sup_{\|f\|_{\infty,\mathbb{R}}\leq 1} \left| \int_{\mathbb{R}^d} f(x) \left(\mu_1(dx) - \mu_2(dx)\right)\right| \tag{3.4}$$

$$= \frac{1}{2} \sup_{\|f\|_{\infty,\mathbb{C}}\leq 1} \left| \int_{\mathbb{R}^d} f(x) \left(\mu_1(dx) - \mu_2(dx)\right)\right| \tag{3.5}$$

where the first supremum is taken over the real-valued measurable functions and the second one is taken over the complex-valued measurable functions. The latter formulation is less usual so we will provide the proof of the third equality in the Appendix (see Lemma 10).
Given two $\mathbb{R}^d$-valued random vectors X and Y and a measurable function $g : \mathbb{R}^d \to \mathbb{R}^{d'}$ where $d' \in \mathbb{N}^*$, the following inequality holds

$$d_{TV}(\mu_{g(X)}, \mu_{g(Y)}) \leq d_{TV}(\mu_X, \mu_Y). \tag{3.6}$$

In particular,

$$d_{TV}(\mu_{(0,X)}, \mu_{(0,Y)}) = d_{TV}(\mu_X, \mu_Y) \tag{3.7}$$

and given $A \in \mathbb{R}^{d\times d}$ an invertible matrix

$$d_{TV}(\mu_{AX}, \mu_{AY}) = d_{TV}(\mu_X, \mu_Y). \tag{3.8}$$

Additionally we introduce the following notation: given a real sequence $(x_j)_{j\geq 1}$ and a real number $c \in \mathbb{R}$ we define for any integer $M \geq 1$

$$\bar{x}_M^c := \frac{x_1 + \cdots + x_M + c}{M}.$$

3.3 Main Result

Let $(Y_i, Z_i)_{i\geq 1}$ be a sequence of square-integrable i.i.d. random vectors in $\mathbb{R}^2$ where Y_i is not constant. Moreover let $q \in \mathbb{N}^*$ and consider a sequence of vectors of integers $(\beta_M^1, \cdots, \beta_M^{q+1})_{M\geq 1} \subseteq \mathbb{N}^{q+1}$ such that

$$\lim_{M\to\infty} \sqrt{M}\left(\frac{\beta_M^i}{M} - \beta^i\right) = 0$$

with

$$0 < \beta^1 < \cdots < \beta^{q+1}.$$

Let us observe that for M big enough $0 < \beta_M^1 < \cdots < \beta_M^{q+1}$. Given $(x, z, y_1, \cdots, y_q) \in \mathbb{R}^{q+2}$ and ϕ a measurable real-valued function, we are interested in studying the convergence in distribution as $M \to \infty$ of the following random vector

$$\left(\left\{R_{\beta_M}^{1,y_1}\right\}, \cdots, \left\{R_{\beta_M}^{q,y_q}\right\}, \sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right) \times \frac{\beta_M^{q+1}}{M}\overline{Z}^z_{\beta_M^{q+1}} - \theta\right)\right) \tag{3.9}$$

where

$$R_{\beta_M}^{i,y_i} := \phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\left(Y_1 + \cdots + Y_{\beta_M^i} + y_i\right) \tag{3.10}$$

and

$$\theta = \phi\left(\beta^{q+1}m_Y\right)\beta^{q+1}m_Z. \tag{3.11}$$

Remark 26. *Let us observe that if ϕ is differentiable at $\beta^{q+1}m_Y$, the application of the delta method provides the convergence in distribution of the last component of (3.9):*

$$\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right) \times \frac{\beta_M^{q+1}}{M}\overline{Z}^z_{\beta_M^{q+1}} - \theta\right) \xRightarrow{d} T \sim \mathcal{N}\left(0, \sigma_T^2\right) \tag{3.12}$$

where $\sigma_T^2 = \beta^{q+1}\left(\phi'(\beta^{q+1}m_Y)\beta^{q+1}m_Z, \phi(\beta^{q+1}m_Y)\right)^T \Sigma_{(Y_1,Z_1)}\left(\phi'(\beta^{q+1}m_Y)\beta^{q+1}m_Z, \phi(\beta^{q+1}m_Y)\right)$ being $\Sigma_{(Y_1,Z_1)}$ the covariance matrix of (Y_1, Z_1).

We are now ready to state the main result of this work.

Theorem 27. *Let $(Y_i, Z_i)_{i\geq 1}$ be a sequence of square-integrable i.i.d. random vectors in $\mathbb{R}^2$ such that the law of Y_i has an absolutely continuous component. Moreover let $(x, z, y_1, \cdots, y_q) \in \mathbb{R}^{q+2}$ and let $\phi : \mathbb{R} \to \mathbb{R}$ be a measurable function differentiable at $\beta^{q+1} m_Y$ such that $\phi(\beta^{q+1} m_Y) \neq 0$. If there exists $\tilde{M} \in \mathbb{N}$ such that*

$$\int_{\mathbb{R}} \frac{1}{|\phi(\beta^{q+1} m_Y + y)|^{q+1}} e^{-\tilde{M} y^2} dy < \infty, \tag{3.13}$$

then the random vector

$$\left(\left\{ R^{1,y_1}_{\beta_M} \right\}, \cdots, \left\{ R^{q,y_q}_{\beta_M} \right\}, \sqrt{M} \left(\phi \left(\frac{\beta_M^{q+1}}{M} \overline{Y}_{\beta_M^{q+1}}^{x} \right) \times \frac{\beta_M^{q+1}}{M} \overline{Z}_{\beta_M^{q+1}}^{z} - \theta \right) \right)$$

converges in distribution as $M \to \infty$ to (U, T) where T has been introduced in (3.12) and where U is a uniform random variable on $[0,1]^q$ independent of T.

Remark 28. *Let us observe that we can apply Theorem 27 to $(Y_i, 0)_{i\geq 1}$ where $(Y_i)_{i\geq 1}$ is a sequence of square-integrable i.i.d. real-valued random variables such that the law of Y_i has an absolutely continuous component. In particular we have*

$$\left(\left\{ R^{1,y_1}_{\beta_M} \right\}, \cdots, \left\{ R^{q,y_q}_{\beta_M} \right\} \right) \overset{d}{\Longrightarrow} U$$

where U is a uniform random variable on $[0,1]^q$. It is possible to prove that in this case the hypothesis (3.13) can be replaced by the following slightly weaker hypothesis: $\exists \tilde{M} \in \mathbb{N}$ such that

$$\int_{\mathbb{R}} \frac{1}{|\phi(\beta^{q+1} m_Y + y)|^{q}} e^{-\tilde{M} y^2} dy < \infty. \tag{3.14}$$

We are now going to provide the statement of Theorem 27 in the particular case where the law of (Y_1, Z_1) has an absolutely continuous component and $(z, y_1, \cdots, y_q) = (0, \cdots, 0)$. This, together with the lemma that immediately follows, will allow to prove Theorem 27.

Proposition 5. *Theorem 27 holds under the reinforced hypotheses that the law of (Y_i, Z_i) has an absolutely continuous component and $(z, y_1, \cdots, y_q) = (0, \cdots, 0)$.*

Lemma 6. *Let Y and Z be two real-valued random variables such that the law of Y has an absolutely continuous component. If ξ is an absolutely continuous real-valued random variable independent of (Y, Z), the law of $(Y, Z + \xi)$ has an absolutely continuous component.*

We provide the proof of Proposition 5 in Section 3.4 and the proof of Lemma 6 in the Appendix.

Proof of Theorem 27. The proof consists of two steps. In step *(i)* we are going to check that the conclusion still holds when the hypothesis of existence of an absolutely continuous component for (Y_1, Z_1) made in Proposition 5 is weakened to the existence of an absolutely continuous component for Y_1. Moreover we suppose that $(z, y_1, \cdots, y_q) = (0, \cdots, 0)$. In step *(ii)* we deal with the case when the vector $(z, y_1, \cdots, y_q) \neq (0, \cdots, 0)$.

(i) To simplify the notation in what follows we write $R^{i}_{\beta_M}$ instead of $R^{i,0}_{\beta_M}$ and $\overline{Z}_{\beta_M^{q+1}}$ instead of $\overline{Z}^{0}_{\beta_M^{q+1}}$.

Let $(\xi_i)_{i\geq 1}$ be a sequence of zero-mean absolutely continuous square-integrable i.i.d. real-valued random variables independent of $(Y_i, Z_i)_{i\geq 1}$. For each $n \geq 1$, let $\xi_i^n = \frac{\xi_i}{n}$ and let us consider the sequence $(Y_i, Z_i + \xi_i^n)_{i\geq 1}$ of square-integrable i.i.d. random vectors in $\mathbb{R}^2$ such that by Lemma 6 the law of $(Y_i, Z_i + \xi_i^n)$ has an absolutely continuous component. Thus we can apply Proposition 5 and obtain that for each $n \geq 1$ the random vector

$$\left(\left\{R^1_{\beta_M}\right\},\cdots,\left\{R^q_{\beta_M}\right\},\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\left(\overline{Z}_{\beta_M^{q+1}}+\overline{\xi^n}_{\beta_M^{q+1}}\right)-\theta\right)\right)\overset{d}{\Longrightarrow}(U,T_n) \tag{3.15}$$

where $\theta=\phi\left(\beta^{q+1}m_Y\right)\beta^{q+1}m_Z$, U is a uniform random variable on $[0,1]^q$ independent of T_n and by Remark 26, $T_n\sim\mathcal{N}\left(0,\sigma^2_{T_n}\right)$ with

$$\sigma^2_{T_n}=\beta^{q+1}\left(\phi'(\beta^{q+1}m_Y)\beta^{q+1}m_Z,\phi(\beta^{q+1}m_Y)\right)^T\Sigma_{(Y_1,Z_1+\xi^n_1)}\left(\phi'(\beta^{q+1}m_Y)\beta^{q+1}m_Z,\phi(\beta^{q+1}m_Y)\right).$$

Since $(Y_1,Z_1+\xi^n_1)$ converges in L^2 to (Y_1,Z_1) as n goes to infinity, $\Sigma_{(Y_1,Z_1+\xi^n_1)}\underset{n\to\infty}{\longrightarrow}\Sigma_{(Y_1,Z_1)}$ and $\sigma^2_{T_n}$ converges as $n\to\infty$ to

$$\sigma^2_{T}=\beta^{q+1}\left(\phi'(\beta^{q+1}m_Y)\beta^{q+1}m_Z,\phi(\beta^{q+1}m_Y)\right)^T\Sigma_{(Y_1,Z_1)}\left(\phi'(\beta^{q+1}m_Y)\beta^{q+1}m_Z,\phi(\beta^{q+1}m_Y)\right).$$

Let us now prove that this implies the convergence in distribution as $M\to\infty$ of (3.9) to (U,T) where $T\sim\mathcal{N}\left(0,\sigma^2_T\right)$ independent of $U=(U_1,\cdots,U_q)$.
Let $(u_1,u_2)\in\mathbb{R}^q\times\mathbb{R}$ and let us denote the random vector $\left(\left\{R^1_{\beta_M}\right\},\cdots,\left\{R^q_{\beta_M}\right\}\right)$ by $\left\{R^i_{\beta_M}\right\}_{i\le q}$.
For $n\ge1$ one has

$$\begin{aligned}
&\left|\mathbb{E}\left(e^{iu_1^T\left(\left\{R^i_{\beta_M}\right\}_{i\le q}\right)+iu_2\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)}\right)-\mathbb{E}\left(e^{iu_1^TU}\right)e^{-\frac12\sigma^2_Tu_2^2}\right|\\
&\le\left|\mathbb{E}\left(e^{iu_1^T\left(\left\{R^i_{\beta_M}\right\}_{i\le q}\right)}e^{iu_2\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)}\left(1-e^{iu_2\sqrt{M}\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\frac{\beta_M^{q+1}}{M}\overline{\xi^n}_{\beta_M^{q+1}}}\right)\right)\right|\\
&\quad+\left|\mathbb{E}\left(e^{iu_1^T\left(\left\{R^i_{\beta_M}\right\}_{i\le q}\right)+iu_2\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\left(\overline{Z}_{\beta_M^{q+1}}+\overline{\xi^n}_{\beta_M^{q+1}}\right)-\theta\right)}\right)-\mathbb{E}\left(e^{iu_1^TU}\right)e^{-\frac12\sigma^2_{T_n}u_2^2}\right|\\
&\quad+\left|\mathbb{E}\left(e^{iu_1^TU}\right)e^{-\frac12\sigma^2_{T_n}u_2^2}-\mathbb{E}\left(e^{iu_1^TU}\right)e^{-\frac12\sigma^2_Tu_2^2}\right|\\
&\le\mathbb{E}\left(\left|1-e^{iu_2\sqrt{M}\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\frac{\beta_M^{q+1}}{M}\overline{\xi^n}_{\beta_M^{q+1}}}\right|\right)\\
&\quad+\left|\mathbb{E}\left(e^{iu_1^T\left(\left\{R^i_{\beta_M}\right\}_{i\le q}\right)+iu_2\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\left(\overline{Z}_{\beta_M^{q+1}}+\overline{\xi^n}_{\beta_M^{q+1}}\right)-\theta\right)}\right)-\mathbb{E}\left(e^{iu_1^TU}\right)e^{-\frac12\sigma^2_{T_n}u_2^2}\right|\\
&\quad+\left|e^{-\frac12\sigma^2_{T_n}u_2^2}-e^{-\frac12\sigma^2_Tu_2^2}\right|.
\end{aligned}$$

Let us now study the first term of the right-hand side.
By observing that $H_M=\sqrt{M}\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\frac{\beta_M^{q+1}}{M}\overline{\xi}_{\beta_M^{q+1}}\overset{d}{\Longrightarrow}\mathcal{N}\left(0,\beta^{q+1}\phi^2\left(\beta^{q+1}m_Y\right)\mathrm{Var}\left(\xi\right)\right)$, this in particular implies that the sequence H_M is tight: $\forall\epsilon>0$ there exists $K_\epsilon>0$ such that $\sup_{M\ge1}\mathbb{P}\left(|H_M|>K_\epsilon\right)\le\epsilon$. One has

$$\mathbb{E}\left(\left|1-e^{iu_2\sqrt{M}\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\frac{\beta_M^{q+1}}{M}\overline{\xi^n}_{\beta_M^{q+1}}}\right|\right)=\mathbb{E}\left(\left|1-e^{i\frac{u_2}{n}H_M}\right|\right)$$
$$=\mathbb{E}\left(\left|1-e^{i\frac{u_2}{n}H_M}\right|1_{|H_M|\leq K_\epsilon}\right)+\mathbb{E}\left(\left|1-e^{i\frac{u_2}{n}H_M}\right|1_{|H_M|>K_\epsilon}\right)$$
$$\leq\frac{|u_2|}{n}K_\epsilon+2\epsilon$$

where to obtain the last inequality we use that $\left|1-e^{ix}\right|\leq|x|$, $\forall x\in\mathbb{R}$.
Therefore we have obtained that $\forall\epsilon>0$, $\forall n>0$

$$\left|\mathbb{E}\left(e^{iu_1^T\left(\left\{R^i_{\beta_M}\right\}_{i\leq q}\right)+iu_2\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)}\right)-\mathbb{E}\left(e^{iu_1^TU}\right)e^{-\frac{1}{2}\sigma_T^2u_2^2}\right|$$
$$\leq\frac{|u_2|}{n}K_\epsilon+2\epsilon+\left|e^{-\frac{1}{2}\sigma^2_{T_n}u_2^2}-e^{-\frac{1}{2}\sigma_T^2u_2^2}\right|$$
$$+\left|\mathbb{E}\left(e^{iu_1^T\left(\left\{R^i_{\beta_M}\right\}_{i\leq q}\right)+iu_2\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\left(\overline{Z}_{\beta_M^{q+1}}+\overline{\xi^n}_{\beta_M^{q+1}}\right)-\theta\right)}\right)-\mathbb{E}\left(e^{iu_1^TU}\right)e^{-\frac{1}{2}\sigma^2_{T_n}u_2^2}\right|.$$

To conclude this part of the proof it is sufficient to take in the above expression first the superior limit as $M\to\infty$ so that the last term converges to 0 by (3.15), then the superior limit as $n\to\infty$ so that $\frac{|u_2|}{n}K_\epsilon\to0$ and $\left|e^{-\frac{1}{2}\sigma_T^2u_2^2}-e^{-\frac{1}{2}\sigma^2_{T_n}u_2^2}\right|\to0$ and finally the superior limit as $\epsilon\to0$.

(ii) Let us now take $(z,y_1,\cdots,y_q)\in\mathbb{R}^{q+2}$ different from the zero vector. One has

$$\left(\left\{R^{1,y_1}_{\beta_M}\right\},\cdots,\left\{R^{1,y_q}_{\beta_M}\right\},\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}^z_{\beta_M^{q+1}}-\theta\right)\right) \tag{3.16}$$
$$=\left(\left(\left\{\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)y_i+\left\{R^i_{\beta_M}\right\}\right\}\right)_{1\leq i\leq q},\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)\right)$$
$$+\left(0,\cdots,0,\frac{z}{\sqrt{M}}\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\right).$$

By the Strong Law of Large Numbers and the hypothesis of continuity of ϕ at $\beta^{q+1}m_Y$, $\lim_{M\to\infty}\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)y_i=\phi\left(\beta^{q+1}m_Y\right)y_i$. We can therefore apply the previous step and Slutsky's theorem to deduce the following convergence in distribution

$$\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^{x_2}_{\beta_M^{q+1}}\right)y_1+\left\{R^1_{\beta_M}\right\},\cdots,\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)y_q+\left\{R^q_{\beta_M}\right\},\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)\right)$$
$$\overset{d}{\Longrightarrow}\left(\phi\left(\beta^{q+1}m_Y\right)y_1+U_1,\cdots,\phi\left(\beta^{q+1}m_Y\right)y_q+U_q,T\right).$$

Now by observing that the set of the points of discontinuity of the function $(x_1,\cdots,x_q,y)\longmapsto(\{x_1\},\cdots,\{x_q\},y)$ has a zero measure with respect to the law of

$$\left(\phi\left(\beta^{q+1}m_Y\right)y_1+U_1,\cdots,\phi\left(\beta^{q+1}m_Y\right)y_q+U_q,T\right)$$

and applying the continuous mapping theorem we can deduce that

$$\left(\left(\left\{\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)y_i+\left\{R^i_{\beta_M}\right\}\right\}\right)_{1\leq i\leq q},\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)\right)$$
$$\overset{d}{\Longrightarrow}\left(\left\{\phi\left(\beta^{q+1}m_Y\right)y_1+U_1\right\},\cdots,\left\{\phi\left(\beta^{q+1}m_Y\right)y_q+U_q\right\},T\right)\overset{d}{=}\left(U_1,\cdots,U_q,T\right).$$

This combined with the fact that $\left(0,\cdots,0,\frac{z}{\sqrt{M}}\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\right)$ converges as M goes to infinity to the zero vector allows to conclude that (3.16) converges in distribution to (U,T).

□

3.4 Proof of Proposition 5

Before proving Proposition 5, we need some preliminary results. As done in the proof of Theorem 27, in what follows we write $R^i_{\beta_M}$ instead of $R^{i,0}_{\beta_M}$ and $\overline{Z}_{\beta_M^{q+1}}$ instead of $\overline{Z}^0_{\beta_M^{q+1}}$. Let us observe that for each $j=1,\cdots,q$

$$R^j_{\beta_M}=\sum_{\ell=1}^{j}R^{\ell-1:\ell}_{\beta_M}$$

where we have introduced the notation $R^{\ell-1:\ell}_{\beta_M}:=\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\left(Y_{\beta_M^{\ell-1}+1}+\cdots+Y_{\beta^\ell_M}\right)$ with $\beta^0_M:=0$ for each M by convention. Let us therefore study the asymptotic behaviour of the vector

$$(R_M,K_M) \tag{3.17}$$

where

$$R_M:=\left(R^{0:1}_{\beta_M},R^{1:2}_{\beta_M},\cdots,R^{q-1:q}_{\beta_M}\right)$$

and

$$K_M:=\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)=\frac{\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\beta_M^{q+1}\overline{Z}_{\beta_M^{q+1}}-M\theta}{\sqrt{M}}$$

with θ as in (3.11).

The proof of the following proposition is given in Section 3.5.

Proposition 6. *Let $\gamma=\left(\beta^1,\beta^2-\beta^1,\cdots,\beta^q-\beta^{q-1}\right)$. Under the assumptions of Proposition 5 the following convergence in total variation holds*

$$d_{TV}\left(\mu_{\left(\frac{R_M-m_Y\phi(\beta^{q+1}m_Y)M\gamma}{\sqrt{M}},K_M\right)},\mathcal{N}(0,\Gamma)\right)\underset{M\to\infty}{\longrightarrow}0 \tag{3.18}$$

where $\Gamma\in\mathbb{R}^{(q+1)\times(q+1)}$ is a positive definite covariance matrix.

Let us now state the Weyl criterion concerning the convergence in distribution to a vector composed of a uniform random variable on $[0,1]^q$ and an independent vector.

Theorem 29 (Weyl criterion)**.** *Let $(B_M)_{M\geq1}$ be a sequence of $\mathbb{R}^q$-valued random vectors and let $(H_M)_{M\geq1}$ be a sequence of $\mathbb{R}^{q'}$-valued random vectors that converges in distribution to a $\mathbb{R}^{q'}$-valued random vector H. Then as $M\to\infty$ the sequence $\left(\left\{B^1_M\right\},\cdots,\left\{B^q_M\right\},H_M\right)$ converges in distribution to (U,H) where U is independent of H and uniformly distributed on $[0,1]^q$ if and only if for every $k\in\mathbb{Z}^q\backslash\{0\}$ and $u\in\mathbb{R}^{q'}$*

$$\lim_{M\to\infty}\phi_{(B_M,H_M)}(2k\pi,u)=0. \tag{3.19}$$

For the sake of completeness, the proof of Theorem 29 is provided in Section 3.5. We are now ready to prove Proposition 5.

Proof of Proposition 5. By Theorem 29, it is sufficient to prove that for each $k \in \mathbb{Z}^q \setminus \{0\}$, $u \in \mathbb{R}$

$$\lim_{M\to\infty} \phi_{\left(R^1_{\beta_M},\cdots,R^q_{\beta_M},\sqrt{M}\left(\phi\left(\frac{\beta_M^{q+1}}{M}\overline{Y}^x_{\beta_M^{q+1}}\right)\times\frac{\beta_M^{q+1}}{M}\overline{Z}_{\beta_M^{q+1}}-\theta\right)\right)}(2k\pi,u)=0.$$

Let us observe that it is equivalent to prove that for each $k \in \mathbb{Z}^q \setminus \{0\}$, $u \in \mathbb{R}$

$$\lim_{M\to\infty} \phi_{(R_M,K_M)}(2k\pi,u)=0 \tag{3.20}$$

where (R_M, K_M) has been introduced in (3.17).

By Proposition 6,

$$\lim_{M\to\infty} d_{TV}(\mu_{(a_M R_M+b_M,K_M)},\mu_Y)=0 \tag{3.21}$$

where $a_M = \frac{1}{\sqrt{M}}$, $b_M = -m_Y\phi(\beta^{q+1}m_Y)\sqrt{M}\gamma$ and $Y \sim \mathcal{N}(0,\Gamma)$, Γ positive definite.

For $u := (u_1, u_2) \in \{\mathbb{R}^q \setminus \{0\}\} \times \mathbb{R}$

$$\begin{aligned}\left|\phi_{(R_M,K_M)}(u_1,u_2)\right| &= \left|\phi_{(a_M R_M+b_M,K_M)}\left(\frac{u_1}{a_M},u_2\right)e^{-\frac{iu_1^T b_M}{a_M}}\right| = \left|\phi_{(a_M R_M+b_M,K_M)}\left(\frac{u_1}{a_M},u_2\right)\right| \\ &\leq \left|\phi_{(a_M R_M+b_M,K_M)}\left(\frac{u_1}{a_M},u_2\right)-\phi_Y\left(\frac{u_1}{a_M},u_2\right)\right| + \left|\phi_Y\left(\frac{u_1}{a_M},u_2\right)\right|.\end{aligned}$$

Since by (3.5) the first term of the right-hand side can be bounded from above by $2d_{TV}(\mu_{(a_M R_M+b_M,K_M)},\mu_Y)$, we deduce that

$$\limsup_{M\to\infty}\left|\phi_{(R_M,K_M)}(u_1,u_2)\right| \leq \limsup_{M\to\infty}\left|\phi_Y\left(\frac{u_1}{a_M},u_2\right)\right|.$$

Since the law of Y is absolutely continuous and $\lim_{M\to\infty} a_M = 0$, the right-hand side goes to 0 as $M \to \infty$ by the Riemann Lebesgue lemma. In particular (3.20) is true.

□

3.5 Proof of Proposition 6 and Theorem 29

In this section we are first going to prove Proposition 6. Let $(Y_i, Z_i)_{i\geq 1}$ be a sequence of square-integrable i.i.d. random vectors in $\mathbb{R}^2$ such that the law of (Y_i, Z_i) has an absolutely continuous component. We denote by Σ its covariance matrix which has rank 2.

The proof of Proposition 6 strongly relies on the following result.

Theorem 30. *Let $(Y_i, Z_i)_{i\geq 1}$ be a sequence of square-integrable i.i.d. random vectors in $\mathbb{R}^2$ such that the law of (Y_i, Z_i) has an absolutely continuous component. Under the notations introduced above*

$$\lim_{M\to\infty} d_{TV}\left(\mu_{\frac{1}{\sqrt{M}}\sum_{k=1}^{M}(Y_k-m_Y,Z_k-m_Z)},\mu_G\right)=0$$

where $G \sim \mathcal{N}(0,\Sigma)$.

Let F be a centered square-integrable random variable in $\mathbb{R}^n$ with identity covariance matrix and let F_k, $k \in \mathbb{N}^*$, independent copies of F. The main instrument that we will use in the proof of Theorem 30 is the result about the convergence in total variation for the CLT that is $\lim_{M\to\infty} d_{TV}\left(\mu_{\frac{1}{\sqrt{M}}\sum_{k=1}^{M}F_k},\mathcal{N}(0,I_{n\times n})\right)=0$ where $I_{n\times n}$ denotes the identity matrix of size n. Prohorov [50] in 1952 was the first to give his contribution to the problem: he proved that, in dimension 1, a necessary and sufficient condition in order to get the result is that there exists

M_0 such that the law of $\sum_{k=1}^{M_0} F_k$ has an absolutely continuous component. In 1964, Mamatov et Halikov [42] contribute in the same direction by extending the result to any dimension. They prove that under the assumption that the law of F has an absolutely continuous component, $\lim_{M\to\infty} d_{TV}\left(\mu_{\frac{1}{\sqrt{M}}\sum_{k=1}^{M} F_k}, \mathcal{N}(0, I_{n\times n})\right) = 0$. See also [3].

We are now ready to prove Theorem 30.

Proof of Theorem 30. Let $O \in \mathbb{R}^{2\times 2}$ be an orthogonal matrix $\left(O^T O = I\right)$ that diagonalizes Σ that is $\Sigma = O^T D O$ where D is the diagonal matrix containing the eigenvalues λ_1, λ_2 of Σ. Given $G \sim \mathcal{N}(0, \Sigma)$ and by introducing the following notation $D^{-\frac{1}{2}} = \text{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}\right)$, we have

$$d_{TV}\left(\mu_{\frac{1}{\sqrt{M}}\sum_{k=1}^{M}(Y_k - m_Y, Z_k - m_Z)}, \mu_G\right) = d_{TV}\left(\mu_{\frac{D^{-\frac{1}{2}}}{\sqrt{M}}\sum_{k=1}^{M} O(Y_k - m_Y, Z_k - m_Z)}, \mu_{D^{-\frac{1}{2}}OG}\right)$$

where to obtain the above equality we apply (3.8) with $A = D^{-\frac{1}{2}}O$. By the hypothesis that the law of (Y_1, Z_1) has an absolutely continuous component, the right-hand side converges to 0 as $M \to \infty$ by the result of convergence in total variation mentioned above. □

We now recall a result obtained by Parthasarathy and Steerneman in the second section of [49] regarding the behavior of the total variation convergence with respect to the sum and the multiplication by a real sequence.

Lemma 7. *Let $(B_M)_{M\geq 1}$ and $(T_M)_{M\geq 1}$ be two independent sequences of $\mathbb{R}^d$-valued random variables such that $\lim_{M\to\infty} d_{TV}(\mu_{B_M}, \mu_B) = 0$ and $T_M \xrightarrow{d} T$ for $\mathbb{R}^d$-valued random variables B, T. If μ_B is absolutely continuous with respect to the Lebesgue measure, then $\lim_{M\to\infty} d_{TV}(\mu_{B_M+T_M}, \mu_{B+T}) = 0$. Moreover under the same condition, $\lim_{M\to\infty} d_{TV}(\mu_{c_M B_M}, \mu_{c_0 B}) = 0$ if $(c_M)_{M\geq 1}$ is a deterministic real sequence converging to $c_0 \in \mathbb{R}^*$.*

The first step in the proof of Proposition 6 consists in applying Theorem 30 so that we can work with gaussian random vectors. Thus the new problem becomes to study the convergence in total variation of the law of a given function of a normal random vector. The following lemma deals with this problem and its proof is given after the proof of Proposition 6.

Lemma 8. *Let*

- $(\eta_{M,1}, \cdots, \eta_{M,q+2})_{M\geq 1} \subseteq \mathbb{R}^{q+2}$ *such that for* $i = 1, \cdots, q+2$ $\lim_{M\to\infty} \sqrt{M}\left(\frac{\eta_{M,i}}{M} - \eta_i\right) = 0$ *with* $(\eta_1, \cdots, \eta_{q+2}) \in \mathbb{R}^{q+2}$
- $(W_1, \cdots, W_{q+2})$ *a zero-mean normal vector with a positive definite covariance matrix* $\Sigma_1 \in \mathbb{R}^{(q+2)\times(q+2)}$
- $\phi : \mathbb{R} \to \mathbb{R}$ *be a measurable function differentiable at* η_{q+2} *and such that* $\phi(\eta_{q+2}) \neq 0$.

If there exists $\tilde{M} \in \mathbb{N}$ *such that*

$$\int_{\mathbb{R}} \frac{1}{|\phi(\eta_{q+2} + y)|^{q+1}} e^{-\tilde{M}y^2} dy < \infty, \tag{3.22}$$

the law of the following random vector

$$\left(\frac{\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{W_{q+2}}{\sqrt{M}}\right)\left(\sqrt{M}W_i + \eta_{M,i}\right) - M\phi(\eta_{q+2})\eta_i}{\sqrt{M}}\right)_{1\leq i\leq q+1} \tag{3.23}$$

converges in total variation as $M \to \infty$ to the law of $(\phi(\eta_{q+2})W_i + \eta_i\phi'(\eta_{q+2})W_{q+2})_{1\leq i\leq q+1}$.

Remark 31. *Let us observe that it is not difficult to prove the pointwise convergence of (3.23) to*

$$\left(\phi(\eta_{q+2})W_i + \eta_i\phi'(\eta_{q+2})W_{q+2}\right)_{1\leq i\leq q+1}.$$

Indeed for $i=1,\cdots,q+1$, using that $\lim_{M\to\infty}\sqrt{M}\left(\frac{\eta_{M,j}}{M}-\eta_j\right)=0$ and the hypothesis of differentiability of ϕ at η_{q+2}, one has

$$\begin{aligned}
&\frac{\phi\left(\frac{\eta_{M,q+2}}{M}+\frac{W_{q+2}}{\sqrt{M}}\right)\left(\sqrt{M}W_i+\eta_{M,i}\right)-M\phi(\eta_{q+2})\eta_i}{\sqrt{M}}\\
&=\phi(\eta_{q+2})\left(W_i+\sqrt{M}\left(\frac{\eta_{M,i}}{M}-\eta_i\right)\right)+\phi'(\eta_{q+2})\left(W_{q+2}+\sqrt{M}\left(\frac{\eta_{M,q+2}}{M}-\eta_{q+2}\right)\right)\left(\frac{W_i}{\sqrt{M}}+\frac{\eta_{M,i}}{M}\right)+o(1)\\
&\underset{M\to\infty}{\longrightarrow}\phi(\eta_{q+2})W_i+\eta_i\phi'(\eta_{q+2})W_{q+2}.
\end{aligned}$$

The proof of the convergence in total variation will require more effort.

Remark 32. *Let us observe that the covariance matrix of the zero-mean normal random vector $(\phi(\eta_{q+2})W_i+\eta_i\phi'(\eta_{q+2})W_{q+2})_{1\leq i\leq q+1}$ is $A_1\Sigma_1A_1^T$ with $A_1\in\mathbb{R}^{q+1\times(q+2)}$ the $q+1$-rank matrix given by*

$$A_1=\begin{pmatrix}
\phi(\eta_{q+2}) & 0 & 0 & 0 & \cdots & 0 & \eta_1\phi'(\eta_{q+2})\\
0 & \phi(\eta_{q+2}) & 0 & \cdots & 0 & 0 & \eta_2\phi'(\eta_{q+2})\\
0 & 0 & \phi(\eta_{q+2}) & 0 & 0 & 0 & \cdots\\
\cdots & \cdots & \cdots & \cdots & 0 & \cdots & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
\cdots & \cdots & \cdots & \cdots & \phi(\eta_{q+2}) & 0 & \cdots\\
0 & 0 & 0 & 0 & 0 & \phi(\eta_{q+2}) & \eta_{q+1}\phi'(\eta_{q+2})
\end{pmatrix}. \tag{3.24}$$

Therefore the covariance matrix is positive definite.

The combination of Theorem 30, Lemma 7 and Lemma 8 allows to prove Proposition 6.

Proof of Proposition 6. Let

$$\begin{aligned}
\bar\gamma&:=\left(\beta^1,\beta^2-\beta^1,\cdots,\beta^q-\beta^{q-1},\beta^{q+1}-\beta^q\right)\\
&=\left(\beta^1-\beta^0,\beta^2-\beta^1,\cdots,\beta^q-\beta^{q-1},\beta^{q+1}-\beta^q\right)=\left(\gamma,\beta^{q+1}-\beta^q\right)
\end{aligned}$$

with $\beta^0:=0$ by convention. For $j=1,\cdots,q+1$ let us consider the following independent $\mathbb{R}^2$-valued random vectors

$$V_{M,j}:=\left(V_{M,j}^1,V_{M,j}^2\right)=\frac{1}{\sqrt{\bar\gamma_jM}}\sum_{i=\beta_M^{j-1}+1}^{\beta_M^j}(Y_i-m_Y,Z_i-m_Z)$$

and let us rewrite $\left(\frac{R_M-m_Y\phi(\beta^{q+1}m_Y)M\gamma}{\sqrt{M}},K_M\right)$ in terms of $V_M:=(V_{M,j})_{1\leq j\leq q+1}$. One has

$$\left(\frac{R_M-m_Y\phi(\beta^{q+1}m_Y)M\gamma}{\sqrt{M}},K_M\right)=\left(\left(\frac{R_{\beta_M^{\ell-1:\ell}}-m_Y\phi(\beta^{q+1}m_Y)\bar\gamma_\ell M}{\sqrt{M}}\right)_{1\leq\ell\leq q},K_M\right)=g_M\left(V_M\right)$$

with the component $g_{M,\ell}$ of $g_M : \mathbb{R}^{2(q+1)} \to \mathbb{R}^{q+1}$ for $\ell = 1, \cdots, q$ given by: $\forall c = \left(c_j^1, c_j^2\right)_{1\le j\le q+1} \in \mathbb{R}^{2(q+1)}$

$$g_{M,\ell}(c) = \frac{\phi\left(\frac{m_Y\beta_M^{q+1}}{M} + \frac{x}{M} + \frac{\sum_{j=1}^{q+1}\sqrt{\bar\gamma_j}c_j^1}{\sqrt{M}}\right)\left(\sqrt{\bar\gamma_\ell M}c_\ell^1 + m_Y\left(\beta_M^\ell - \beta_M^{\ell-1}\right)\right) - m_Y\phi(\beta^{q+1}m_Y)\bar\gamma_\ell M}{\sqrt{M}}, \tag{3.25}$$

and $\forall c = \left(c_j^1, c_j^2\right)_{1\le j\le q+1} \in \mathbb{R}^{2(q+1)}$

$$g_{M,q+1}(c) = \frac{\phi\left(\frac{m_Y\beta_M^{q+1}}{M} + \frac{x}{M} + \frac{\sum_{j=1}^{q+1}\sqrt{\bar\gamma_j}c_j^1}{\sqrt{M}}\right)\left(\sqrt{M}\sum_{j=1}^{q+1}\sqrt{\bar\gamma_j}c_j^2 + m_Z\beta_M^{q+1}\right) - m_Z\phi\left(\beta^{q+1}m_Y\right)\beta^{q+1}M}{\sqrt{M}}. \tag{3.26}$$

The purpose now is to "asymptotically rewrite" the vector $\left(\frac{R_M - m_Y\phi(\beta^{q+1}m_Y)\gamma M}{\sqrt{M}}, K_M\right) = g_M(V_M)$ in terms of a normal random vector so to apply Lemma 8. By recalling that we denote by Σ the covariance matrix of (Y_i, Z_i), let $G = (G_1, \cdots, G_{q+1})$ where the two dimensional vectors $G_j = \left(G_j^1, G_j^2\right)$ $1 \le j \le q+1$ are i.i.d. according to $\mathcal{N}(0, \Sigma)$. By the independence of the random vectors $V_{M,j}$, the independence of the random vectors G_j and the well known fact that $d_{TV}\left(\prod_{i=1}^\ell \tilde\nu_i, \prod_{i=1}^\ell \nu_i\right) \le \sum_{i=1}^\ell d_{TV}\left(\tilde\nu_i, \nu_i\right)$ with $\tilde\nu_i, \nu_i$ probability measures for $i = 1, \cdots, \ell$, we have

$$d_{TV}(\mu_{V_M}, \mu_G) \le \sum_{j=1}^{q+1} d_{TV}(\mu_{V_{M,j}}, \mu_{G_j}). \tag{3.27}$$

We are now going to prove that the right-hand side converges to 0 as $M \to \infty$. Let us observe that for $j = 1, \cdots, q+1$

$$V_{M,j} = \frac{\sqrt{\beta_M^j - \beta_M^{j-1}}}{\sqrt{\bar\gamma_j M}} \times \frac{1}{\sqrt{\beta_M^j - \beta_M^{j-1}}} \sum_{i=\beta_M^{j-1}+1}^{\beta_M^j} (Y_i - m_Y, Z_i - m_Z) \tag{3.28}$$

with $\lim_{M\to\infty} \frac{\sqrt{\beta_M^j - \beta_M^{j-1}}}{\sqrt{\bar\gamma_j M}} = 1$.

Thanks to Theorem 30 and the hypothesis that the law of (Y_i, Z_i) has an absolutely continuous component, $\forall j = 1, \cdots, q+1$ we have that the law of $\frac{1}{\sqrt{\beta_M^j - \beta_M^{j-1}}} \sum_{i=\beta_M^{j-1}+1}^{\beta_M^j} (Y_i - m_Y, Z_i - m_Z)$ converges in total variation as $M \to \infty$ to the law of G_j and, by Lemma 7 and (3.28), this implies

$$\lim_{M\to\infty} d_{TV}\left(\mu_{V_{M,j}}, \mu_{G_j}\right) = 0, \quad j = 1, \cdots, q+1.$$

Hence we have proved the right-hand side of (3.27) converges to 0 as $M \to \infty$. Applying (3.6) with $g = g_M$, we deduce that

$$d_{TV}\left(\mu_{g_M(V_M)}, \mu_{g_M(G)}\right) \le d_{TV}(\mu_{V_M}, \mu_G) \underset{M\to\infty}{\longrightarrow} 0.$$

We are now going to prove the convergence in total variation of the law of $g_M(G)$ and to do so we apply Lemma 8. Recalling that $G_j \sim \mathcal{N}(0,\Sigma)$ where Σ is positive definite, G_j i.i.d., it is possible to prove that the random variables $\sum_{j=1}^{q+1}\sqrt{\bar{\gamma}_j}G_j^1, \sum_{j=1}^{q+1}\sqrt{\bar{\gamma}_j}G_j^2, \sqrt{\bar{\gamma}_1}G_1^1, \cdots, \sqrt{\bar{\gamma}_q}G_q^1$ are linearly independent.

Thus by using that $\lim_{M\to\infty}\sqrt{M}\left(\frac{\beta_M^i}{M}-\beta^i\right)=0$ for $i=1,\cdots,q+1$ and (3.13), we can apply Lemma 8 with

- $(\eta_{M,1},\cdots,\eta_{M,q},\eta_{M,q+1},\eta_{M,q+2}) = \left(m_Y\left(\beta_M^1-\beta_M^0\right),\cdots,m_Y\left(\beta_M^q-\beta_M^{q-1}\right),m_Z\beta_M^{q+1},\frac{m_Y\beta_M^{q+1}}{M}+\frac{x}{M}\right)$ and $(\eta_1,\cdots,\eta_q,\eta_{q+1},\eta_{q+2}) = \left(m_Y\bar{\gamma}_1,\cdots,m_Y\bar{\gamma}_q,m_Z\beta^{q+1},m_Y\beta^{q+1}\right)$
- $(W_1,\cdots,W_{q+2}) = \left(\sqrt{\bar{\gamma}_1}G_1^1,\cdots,\sqrt{\bar{\gamma}_q}G_q^1,\sum_{j=1}^{q+1}\sqrt{\bar{\gamma}_j}G_j^2,\sum_{j=1}^{q+1}\sqrt{\bar{\gamma}_j}G_j^1\right)$

to conclude that the law of $g_M(G)$ converges in total variation as $M\to\infty$ to the law of $Y\sim\mathcal{N}(0,\Gamma)$ where $\Gamma\in\mathbb{R}^{(q+1)\times(q+1)}$ is positive definite by Remark 32.

□

The following lemma contains the key result to prove Lemma 8 . Its proof is provided in the Appendix.

Lemma 9. *Let $(Y_M)_{M\geq 1}$ be a sequence of $\mathbb{R}^d$-valued random vectors such that*

$$\liminf_{M\to\infty} p_{Y_M}(x)\geq p(x)\quad dx\ a.e. \tag{3.29}$$

with p the density of an absolutely continuous $\mathbb{R}^d$-valued random variable Y. Then

1. $\lim_{M\to\infty}\int_{\mathbb{R}^d}|p_{Y_M}(x)-p(x)|\,dx=0$
2. $\lim_{M\to\infty} d_{TV}(\mu_{Y_M},\mu_Y)=0$.

Proof of Lemma 8. According to Lemma 9, to prove the convergence in total variation of the law of the random vector (3.23) it is enough to check that its density with respect to the Lebesgue measure converges pointwise to that of its limit.

Let us preliminary observe that the density of the zero-mean normal vector $(W_1,\cdots,W_{q+2})$ is given by

$$p(x_{1:q+2}) = \frac{1}{(2\pi)^{\frac{q+2}{2}}\times det(\Sigma_1)^{1/2}}e^{-\frac{1}{2}x_{1:q+2}^T\Sigma_1^{-1}x_{1:q+2}},\quad x_{1:q+2}\in\mathbb{R}^{q+2} \tag{3.30}$$

where we denote the vector $(x_1,\cdots,x_{q+2})$ by $x_{1:q+2}$. Let us moreover observe that by the property of positive definiteness of Σ_1^{-1}, the smallest eigenvalue λ_1 of Σ_1^{-1} is positive and we have

$$p(x_{1:q+2})\leq\frac{1}{(2\pi)^{\frac{q+2}{2}}\times det(\Sigma_1)^{1/2}}e^{-\frac{1}{2}\lambda_1\left(\sum_{i=1}^{q+2}x_i^2\right)},\quad x_{1:q+2}\in\mathbb{R}^{q+2}. \tag{3.31}$$

Let us now compute the density of (3.23). Let $f:\mathbb{R}^{q+1}\to\mathbb{R}_+$ be a non-negative measurable function. Writing the expectation

$$\mathbb{E}\left(f\left(\left(\frac{\phi\left(\frac{\eta_{M,q+2}}{M}+\frac{W_{q+2}}{\sqrt{M}}\right)\left(\sqrt{M}W_i+\eta_{M,i}\right)-M\phi(\eta_{q+2})\eta_i}{\sqrt{M}}\right)_{1\leq i\leq q+1}\right)\right) \tag{3.32}$$

as an integral with respect to the density p and then applying the change of variable

$$\xi_i = \frac{\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\left(\sqrt{M}x_i + \eta_{M,i}\right) - M\phi(\eta_{q+2})\eta_i}{\sqrt{M}} \quad i = 1, \cdots, q+1$$

with inverse

$$x_{M,i} := \frac{1}{\sqrt{M}}\left(\frac{\sqrt{M}\xi_i + M\phi(\eta_{q+2})\eta_i}{\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)} - \eta_{M,i}\right) \quad i = 1, \cdots, q+1 \tag{3.33}$$

for $x_{q+2} \in \mathbb{R}$ outside the set $\left\{t \in \mathbb{R} : \phi\left(\frac{\eta_{M,q+2}}{M} + \frac{t}{\sqrt{M}}\right) = 0\right\}$ which is Lebesgue negligible since by a change of variable and (3.22)

$$\int_{\mathbb{R}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{t}{\sqrt{M}}\right)\right|^q + 1} e^{-\tilde{M}\left(\frac{\eta_{M,q+2}}{M} - \eta_{q+2} + \frac{t}{\sqrt{M}}\right)^2} dt = \sqrt{M}\int_{\mathbb{R}} \frac{1}{|\phi(\eta_{q+2}+y)|^{q+1}} e^{-\tilde{M}y^2} dy < \infty,$$

we get

$$\int_{\mathbb{R}} dx_{q+2} \int_{\mathbb{R}^{q+1}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} f(\xi_{1:q+1}) p\left(x_{M,1}, \cdots, x_{M,q+1}, x_{q+2}\right) d\xi_{1:q+1}.$$

We have therefore obtained that the density of $\left(\frac{\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{W_{q+2}}{\sqrt{M}}\right)\left(\sqrt{M}W_i + \eta_{M,i}\right) - M\phi(\eta_{q+2})\eta_i}{\sqrt{M}}\right)_{1 \le i \le q+1}$ at $\xi \in \mathbb{R}^{q+1}$ is given by

$$\int_{\mathbb{R}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} p\left(x_{M,1}, \cdots, x_{M,q+1}, x_{q+2}\right) dx_{q+2}. \tag{3.34}$$

Since, by hypothesis, ϕ is continuous at η_{q+2} with $\phi(\eta_{q+2}) \neq 0$, there exists $\delta > 0$ such that $\forall t \in \mathbb{R} : |t| \le \delta$

$$|\phi(\eta_{q+2}+t)| \ge \frac{|\phi(\eta_{q+2})|}{2}. \tag{3.35}$$

Let us now study the pointwise convergence of (3.34). Let $M \ge 1$ and let us rewrite the integral as

$$\int_{\mathbb{R}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} p\left(x_{M,1:q+1}, x_{q+2}\right) dx_{q+2} \tag{3.36}$$

$$= \int_{|x_{q+2}| \le \frac{\delta}{2}\sqrt{M}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} p\left(x_{M,1:q+1}, x_{q+2}\right) dx_{q+2} \tag{3.37}$$

$$+ \int_{|x_{q+2}| > \frac{\delta}{2}\sqrt{M}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} p\left(x_{M,1:q+1}, x_{q+2}\right) dx_{q+2}. \tag{3.38}$$

Let us start studying the convergence as $M \to \infty$ of (3.37). Thanks to (3.35) and (3.31) and for M big enough so that $\left|\frac{\eta_{M,q+2}}{M} - \eta_{q+2}\right| \le \frac{\delta}{2}$, we have

$$1_{\left\{|x_{q+2}| \le \frac{\delta}{2}\sqrt{M}\right\}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} p\left(x_{M,1:q+1}, x_{q+2}\right) \le \frac{2^{q+1}}{(2\pi)^{\frac{q+2}{2}} \times det(\Sigma_1)^{\frac{1}{2}} |\phi(\eta_{q+2})|^{q+1}} e^{-\frac{1}{2}\lambda_1 x_{q+2}^2}$$

and since by hypothesis ϕ is continuous at η_{q+2}, for each $x_{q+2} \in \mathbb{R}$ we have

$$\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right) \underset{M\to\infty}{\longrightarrow} \phi(\eta_{q+2}). \tag{3.39}$$

Moreover for each $x_{q+2} \in \mathbb{R}$, using that $\lim_{M\to\infty} \sqrt{M}\left(\frac{\eta_{M,i}}{M} - \eta_i\right) = 0$ for $i = 1, \cdots, q+2$ and the hypothesis of differentiability of ϕ one has

$$\begin{aligned} x_{M,i} &= \frac{1}{\sqrt{M}}\left(\frac{\sqrt{M}\xi_i + M\phi(\eta_{q+2})\eta_i}{\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)} - \eta_{M,i}\right) \\ &= \left(\frac{\xi_i + \phi(\eta_{q+2})\sqrt{M}\left(\eta_i - \frac{\eta_{M,i}}{M}\right) - \phi'(\eta_{q+2})\frac{\eta_{M,i}}{\sqrt{M}}\left(\frac{x_{q+2}}{\sqrt{M}} + \frac{\eta_{M,q+2}}{M} - \eta_{q+2}\right) + o(1)}{\phi(\eta_{q+2}) + \phi'(\eta_{q+2})\left(\frac{x_{q+2}}{\sqrt{M}} + \frac{\eta_{M,q+2}}{M} - \eta_{q+2}\right) + o\left(\frac{1}{\sqrt{M}}\right)}\right) \\ &\underset{M\to\infty}{\longrightarrow} \frac{\xi_i - \phi'(\eta_{q+2})\eta_i x_{q+2}}{\phi(\eta_{q+2})}. \end{aligned} \tag{3.40}$$

We can therefore apply Lebesgue's theorem to (3.37) and obtain that for each $\xi \in \mathbb{R}^{q+1}$

$$\begin{aligned} &\lim_{M\to\infty} \int_{|x_{q+2}|\le \frac{\delta}{2}\sqrt{M}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} p\left(x_{M,1:q+1}, x_{q+2}\right) dx_{q+2} \\ &= \frac{1}{|\phi(\eta_{q+2})|^{q+1}} \int_{\mathbb{R}} p\left(\left(\frac{\xi_i - \phi'(\eta_{q+2})\eta_i x_{q+2}}{\phi(\eta_{q+2})}\right)_{1\le i\le q+1}, x_{q+2}\right) dx_{q+2} \\ &= \int_{\mathbb{R}} p_{\left((\phi(\eta_{q+2})W_i + \phi'(\eta_{q+2})\eta_i W_{q+2})_{1\le i\le q+1}, W_{q+2}\right)}(\xi_{1:q+1}, x_{q+2}) dx_{q+2} \\ &= p_{\left((\phi(\eta_{q+2})W_i + \phi'(\eta_{q+2})\eta_i W_{q+2})_{1\le i\le q+1}\right)}(\xi_{1:q+1}). \end{aligned}$$

Let us now prove that (3.38) converges to 0 as $M \to \infty$. By applying a change of variable and by choosing M big enough so that $\left|\frac{\eta_{M,q+2}}{M} - \eta_{q+2}\right| \le \frac{\delta}{4}$, we obtain

$$\begin{aligned} &\int_{|x_{q+2}|>\frac{\delta}{2}\sqrt{M}} \frac{1}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + \frac{x_{q+2}}{\sqrt{M}}\right)\right|^{q+1}} p\left(x_{M,1:q+1}, x_{q+2}\right) dx_{q+2} \\ &= \int_{|y|>\frac{\delta}{2}} \frac{\sqrt{M}}{\left|\phi\left(\frac{\eta_{M,q+2}}{M} + y\right)\right|^{q+1}} p\left(x_{M,1:q+1}, \sqrt{M}y\right) dy \\ &\le \int_{|z|>\frac{\delta}{4}} \frac{\sqrt{M}}{|\phi(\eta_{q+2} + z)|^{q+1}} p\left(x_{M,1:q+1}, \sqrt{M}z + \sqrt{M}\left(\eta_{q+2} - \frac{\eta_{M,q+2}}{M}\right)\right) dz. \end{aligned}$$

Therefore by (3.31) and by using that $\forall x_1, x_2 \in \mathbb{R}\ (x_1 - x_2)^2 \ge \frac{x_1^2}{2} - x_2^2$, one has

$$\begin{aligned} &1_{\{|z|>\frac{\delta}{4}\}} \frac{\sqrt{M}}{|\phi(\eta_{q+2} + z)|^{q+1}} p\left(x_{M,1:q+1}, \sqrt{M}z + \sqrt{M}\left(\eta_{q+2} - \frac{\eta_{M,q+2}}{M}\right)\right) \\ &\le 1_{\{|z|>\frac{\delta}{4}\}} \frac{\sqrt{M}}{(2\pi)^{\frac{q+2}{2}} \times det(\Sigma_1)^{\frac{1}{2}} |\phi(\eta_{q+2} + z)|^{q+1}} e^{-\frac{1}{2}\lambda_1\left(\sqrt{M}z - \sqrt{M}\left(\frac{\eta_{M,q+2}}{M} - \eta_{q+2}\right)\right)^2} \\ &\le 1_{\{|z|>\frac{\delta}{4}\}} \frac{\sqrt{M}}{(2\pi)^{\frac{q+2}{2}} \times det(\Sigma_1)^{\frac{1}{2}} |\phi(\eta_{q+2} + z)|^{q+1}} e^{-\frac{1}{4}\lambda_1 M z^2} e^{\frac{1}{2}\lambda_1\left(\sqrt{M}\left(\frac{\eta_{M,q+2}}{M} - \eta_{q+2}\right)\right)^2} \\ &\le C 1_{\{|z|>\frac{\delta}{4}\}} \frac{\sqrt{M}}{(2\pi)^{\frac{q+2}{2}} \times det(\Sigma_1)^{\frac{1}{2}} |\phi(\eta_{q+2} + z)|^{q+1}} e^{-\frac{1}{4}\lambda_1 M z^2} \end{aligned}$$

where for the last inequality we use that the sequence $\left(\sqrt{M}\left(\frac{\eta_{M,q+2}}{M}-\eta_{q+2}\right)\right)_{M\geq 1}$ converges to 0 and so in particular $e^{\frac{1}{2}\lambda_1\left(\sqrt{M}\left(\frac{\eta_{M,q+2}}{M}-\eta_{q+2}\right)\right)^2}$ is bounded by a positive constant $C<\infty$. The right-hand side converges to 0 as $M\to\infty$.
Let us now observe that for $|z|>\frac{\delta}{4}$ and $t\geq\frac{32}{\delta^2\lambda_1}$,

$$\partial_t(\sqrt{t}e^{-\frac{1}{4}\lambda_1 tz^2})=\frac{e^{-\frac{1}{4}\lambda_1 z^2 t}}{4}\left(\frac{2}{\sqrt{t}}-z^2\sqrt{t}\lambda_1\right)\leq 0.$$

Therefore if $M\geq\bar{M}:=\max\left(\left\lceil\frac{4\tilde{M}}{\lambda_1}\right\rceil,\left\lceil\frac{32}{\delta^2\lambda_1}\right\rceil\right)$, where $\tilde{M}$ is defined in the statement of the Lemma 8,

$$1_{\{|z|>\frac{\delta}{4}\}}\frac{\sqrt{M}}{|\phi\left(\eta_{q+2}+z\right)|^{q+1}}e^{-\frac{1}{4}\lambda_1 Mz^2}\leq 1_{\{|z|>\frac{\delta}{4}\}}\frac{\sqrt{\bar{M}}}{|\phi\left(\eta_{q+2}+z\right)|^{q+1}}e^{-\frac{1}{4}\lambda_1\bar{M}z^2} \tag{3.41}$$

$$\leq 1_{\{|z|>\frac{\delta}{4}\}}\frac{\sqrt{\bar{M}}}{|\phi\left(\eta_{q+2}+z\right)|^{q+1}}e^{-\tilde{M}z^2} \tag{3.42}$$

where the last term is integrable by hypothesis (3.22). By Lebesgue's theorem we can therefore conclude that $\int_{|z|>\frac{\delta}{4}}\frac{\sqrt{M}}{|\phi(\eta_{q+2}+z)|^{q+1}}p\left(x_{M,1:q+1},\sqrt{M}z+\sqrt{M}\left(\eta_{q+2}-\frac{\eta_{M,q+2}}{M}\right)\right)dz$ tends to 0 as $M\to\infty$. In conclusion, we have obtained that the density of $\left(\frac{\phi\left(\frac{\eta_{M,q+2}}{M}+\frac{W_{q+2}}{\sqrt{M}}\right)\left(\sqrt{M}W_i+\eta_{M,i}\right)-M\phi(\eta_{q+2})\eta_i}{\sqrt{M}}\right)_{1\leq i\leq q+1}$ converges pointwise as $M\to\infty$ to the density of $\left((\phi(\eta_{q+2})W_i+\phi'(\eta_{q+2})\eta_i W_{q+2})_{1\leq i\leq q+1}\right)$. □

We are now going to prove Theorem 29.

Proof of Theorem 29. We denote $(\{B_M\},H_M):=\left(\left\{B_M^1\right\},\cdots,\left\{B_M^q\right\},H_M\right)$. Let us first observe that for every $k\in\mathbb{Z}^q,u\in\mathbb{R}^{q'}$ one has

$$\phi_{(\{B_M\},H_M)}(2\pi k,u)=\mathbb{E}\left(e^{i2\pi k^T\{B_M\}}e^{iu^TH_M}\right)=\mathbb{E}\left(e^{i2\pi k^TB_M}e^{-i2\pi\sum_{j=1}^q k_j\lfloor B_M^j\rfloor}e^{iu^TH_M}\right)$$

$$=\phi_{(B_M,H_M)}(2\pi k,u).$$

If $(\{B_M\},H_M)$ converges in distribution as $M\to\infty$ to (U,H) where U is independent of H and uniformly distributed on $[0,1]^q$, $\forall(k,u):=(k_1,\cdots,k_q,u)\in\{\mathbb{Z}^q\setminus\{0\}\}\times\mathbb{R}^{q'}$ one has

$$\lim_{M\to\infty}\phi_{(B_M,H_M)}(2\pi k,u)=\lim_{M\to\infty}\phi_{(\{B_M\},H_M)}(2\pi k,u)=\phi_{(U,H)}(2k\pi,u)$$

$$=\mathbb{E}\left(e^{i2\pi kU}\right)\mathbb{E}\left(e^{iuH}\right)=0.$$

Conversely, assume that (3.19) holds for every $(k,u)\in\{\mathbb{Z}^q\setminus\{0\}\}\times\mathbb{R}^{q'}$ and let us prove that $\forall(u_1,u_2)\in\mathbb{R}^q\times\mathbb{R}^{q'}$,

$$\lim_{M\to\infty}\phi_{(\{B_M\},H_M)}(u_1,u_2)=\phi_U(u_1)\phi_H(u_2)$$

with (U,H) as above.
Fix $(u_1,u_2)\in\{\mathbb{R}^q\setminus\{0\}\}\times\mathbb{R}^{q'}$. Given $0<\epsilon<\frac{1}{4}$, let us define for $j=1,\cdots,q$ a complex $\lceil\frac{q+1}{2}\rceil$-times continuously differentiable function $\phi_{\epsilon,j}$ defined on $[0,1]$ that coincides with $x_j\mapsto e^{iu_1^jx_j}$ on $[0,1-\epsilon]$ and such that $\phi_{\epsilon,j}(1)=\phi_{\epsilon,j}(0)$. It is possible to choose $\phi_{\epsilon,j}$ such that $\sup_{x\in[0,1]}|\phi_{\epsilon,j}(x)|\leq C^{\frac{1}{q}}$, with $C<\infty$ not depending on ϵ and j. If we now consider the function $\phi_\epsilon(x_1,\cdots,x_q):=\phi_{\epsilon,1}(x_1)\cdot$

$\cdot\cdot\,\phi_{\epsilon,q}\left(x_q\right)\ \forall\left(x_1,\cdots,x_q\right)\in[0,1]^q$, then it is $\lceil\frac{q+1}{2}\rceil$-times continuously differentiable, coincides with $(x_1,\cdots,x_q)\mapsto e^{iu_1^1x_1}\cdots e^{iu_1^qx_q}$ on $[0,1-\epsilon]^q$ and $\forall j,\ \forall\left(x_1,\cdots,x_{j-1},x_{j+1},\cdots,x_q\right)\in[0,1]^{q-1}$

$$\phi_\epsilon(x_1,\cdots,x_{j-1},0,x_{j+1},\cdots,x_q)=\phi_\epsilon(x_1,\cdots,x_{j-1},1,x_{j+1},\cdots,x_q).$$

By the theory of Fourier's series (see for instance the Section Sobolev Spaces in Chapter 5 of [53]), there exists $M^\epsilon\in\mathbb{N}$ such that

$$\sup_{x\in[0,1]^q}\left|\phi_\epsilon(x)-\sum_{\max\limits_{j=1,\cdots,q}|k_j|\leq M^\epsilon}c_k^\epsilon e^{i2\pi k^Tx}\right|=\sup_{x\in[0,1]^q}\left|\phi_\epsilon(x)-\sum_{|k_1|\leq M^\epsilon}\cdots\sum_{|k_q|\leq M^\epsilon}c_k^\epsilon e^{i2\pi k^Tx}\right|\leq\epsilon \quad (3.43)$$

where for $k\in\mathbb{Z}^q$

$$c_k^\epsilon=\int_{[0,1]^q}\phi_\epsilon(x)e^{-i2\pi k^Tx}dx$$

satisfies $|c_k^\epsilon|\leq C$. Let ν_M denote the law of $\left(\{B_M\},H_M\right)$ (having support in $[0,1]^q\times\mathbb{R}^{q'}$). Given $M\geq1$ one has

$$\begin{aligned}
&\left|\phi_{(\{B_M\},H_M)}(u_1,u_2)-\phi_U(u_1)\phi_H(u_2)\right|\\
&\leq\left|\int_{[0,1]^q\times\mathbb{R}^{q'}}\left(e^{iu_1^Tx}-\phi_\epsilon(x)\right)e^{iu_2^Ty}\nu_M(dx,dy)\right|\\
&\quad+\left|\int_{[0,1]^q\times\mathbb{R}^{q'}}\phi_\epsilon(x)e^{iu_2^Ty}\nu_M(dx,dy)-\sum_{\max\limits_{j=1,\cdots,q}|k_j|\leq M^\epsilon}c_k^\epsilon\int_{[0,1]^q\times\mathbb{R}^{q'}}e^{i2\pi k^Tx}e^{iu_2^Ty}\nu_M(dx,dy)\right|\\
&\quad+\left|\sum_{\max\limits_{j=1,\cdots,q}|k_j|\leq M^\epsilon,k\neq0}c_k^\epsilon\int_{[0,1]^q\times\mathbb{R}^{q'}}e^{i2\pi k^Tx}e^{iu_2^Ty}\nu_M(dx,dy)\right|+\left|c_0^\epsilon\phi_{H_M}(u_2)-c_0^\epsilon\phi_H(u_2)\right|\\
&\quad+\left|c_0^\epsilon\phi_H(u_2)-\phi_U(u_1)\phi_H(u_2)\right|\\
&\leq(1+C)\int_{[0,1]^q\setminus[0,1-\epsilon]^q}\nu_M(dx,\mathbb{R}^{q'})+\epsilon+\left|\sum_{\max\limits_{j=1,\cdots,q}|k_j|\leq M^\epsilon,k\neq0}c_k^\epsilon\phi_{(B_M,H_M)}(2k\pi,u_2)\right|\\
&\quad+C\left|\phi_{H_M}(u_2)-\phi_H(u_2)\right|+\left|\phi_U(u_1)-c_0^\epsilon\right|.
\end{aligned}$$

Let us look in more detail at the first term of the right-hand side. Let $\bar\phi_\epsilon$ be a real continuously differentiable function defined on $[0,1]$ that is equal to 0 on $[2\epsilon,1-2\epsilon]$ and equal to 1 on $[0,\epsilon]$ and $[1-\epsilon,1]$. By the theory of Fourier's series, there exists $\bar M^\epsilon\in\mathbb{N}$ such that

$$\sup_{x\in[0,1]}\left|\bar\phi_\epsilon(x)-\sum_{|\ell|\leq\bar M^\epsilon}d_\ell^\epsilon e^{i2\pi\ell x}\right|\leq\epsilon \quad (3.44)$$

where for $\ell\in\mathbb{Z}$

$$d_\ell^\epsilon=\int_{[0,1]}\bar\phi_\epsilon(x)e^{-i2\pi\ell x}dx.$$

One has

$$\begin{aligned}\int_{[0,1]^q\setminus[0,1-\epsilon]^q}\nu_M(dx,\mathbb{R}^{q'}) &\le \sum_{j=1}^q\int_{[1-\epsilon,1]}\mu_{\{B_M^j\}}(dx) \le \sum_{j=1}^q\int_{[0,1]}\bar\phi_\epsilon(x)\,\mu_{\{B_M^j\}}(dx)\\ &\le \sum_{j=1}^q\int_{[0,1]}\left|\bar\phi_\epsilon(x)-\sum_{|\ell|\le\bar M^\epsilon}d_\ell^\epsilon e^{i2\pi\ell x}\right|\mu_{\{B_M^j\}}(dx)\\ &\quad+\sum_{j=1}^q\left|\sum_{|\ell|\le\bar M^\epsilon,\ell\neq0}d_\ell^\epsilon\int_{[0,1]}e^{i2\pi\ell x}\mu_{\{B_M^j\}}(dx)\right|+q\,|d_0^\epsilon|\\ &\le q\epsilon+\sum_{j=1}^q\left|\sum_{|\ell|\le\bar M^\epsilon,\ell\neq0}d_\ell^\epsilon\phi_{B_M^j}(2\pi\ell)\right|+q\,|d_0^\epsilon|\,.\end{aligned}$$

Thus we have obtained that

$$\begin{aligned}&\left|\phi_{(\{B_M\},H_M)}(u_1,u_2)-\phi_U(u_1)\phi_H(u_2)\right|\\ &\le(1+C)\left(q\epsilon+\sum_{j=1}^q\left|\sum_{|\ell|\le\bar M^\epsilon,\ell\neq0}d_\ell^\epsilon\phi_{B_M^j}(2\pi\ell)\right|+q\,|d_0^\epsilon|\right)+\epsilon\\ &\quad+\left|\sum_{\max_{j=1,\cdots,q}|k_j|\le M^\epsilon,k\neq0}c_k^\epsilon\phi_{(B_M,H_M)}(2\pi k,u_2)\right|+C\,|\phi_{H_M}(u_2)-\phi_H(u_2)|+|\phi_U(u_1)-c_0^\epsilon|\,.\end{aligned}$$

Given $0<\epsilon<\frac14$, we can first take the superior limit as $M\to\infty$ of the right-hand side. By (3.19) we have

$$\lim_{M\to\infty}\sum_{j=1}^q\left|\sum_{|\ell|\le\bar M^\epsilon,\ell\neq0}d_\ell^\epsilon\phi_{B_M^j}(2\pi\ell)\right|=0=\lim_{M\to\infty}\left|\sum_{\max_{j=1,\cdots,q}|k_j|\le M^\epsilon,k\neq0}c_k^\epsilon\phi_{(B_M,H_M)}(2\pi k,u_2)\right|$$

and by the hypothesis that $H_M\overset{d}{\Rightarrow}H$ we have $\lim_{M\to\infty}|\phi_{H_M}(u_2)-\phi_H(u_2)|=0$. We therefore obtain

$$\limsup_{M\to\infty}\left|\phi_{(\{B_M\},H_M)}(u_1,u_2)-\phi_U(u_1)\phi_H(u_2)\right|\le q(1+C)\,(\epsilon+|d_0^\epsilon|)+\epsilon+|\phi_U(u_1)-c_0^\epsilon|\,. \tag{3.45}$$

Now since

$$|d_0^\epsilon|=d_0^\epsilon=\int_{[0,1]}\bar\phi_\epsilon(x)dx\le4\epsilon$$

and

$$|\phi_U(u_1)-c_0^\epsilon|=\left|\int_{[0,1]^q}(e^{iu_1^Tx}-\phi_\epsilon(x))dx\right|\le(C+1)\int_{[0,1]^q\setminus[0,1-\epsilon]^q}dx$$

with $\int_{[0,1]^q\setminus[0,1-\epsilon]^q}dx=1-(1-\epsilon)^q\underset{\epsilon\to0}{\longrightarrow}0$, the limit as $\epsilon\to0$ of the right-hand side of (3.45) is 0. If $u_1=0$ and $u_2\in\mathbb{R}$, by the hypothesis that $H_M\overset{d}{\Rightarrow}H$, we directly obtain

$$\phi_{(\{B_M\},H_M)}(0,u_2)=\phi_{H_M}(u_2)\underset{M\to\infty}{\longrightarrow}\phi_H(u_2)=\phi_U(0)\phi_H(u_2).$$

. □

3.6 Appendix

Lemma 10. *Given the measures μ_1 and μ_2 on $\mathbb{R}^d$, one has*

$$\sup_{\|f\|_{\infty,\mathbb{C}}\leq 1}\left|\int_{\mathbb{R}^d} f(x)\left(\mu_1(dx)-\mu_2(dx)\right)\right| = 2d_{TV}\left(\mu_1,\mu_2\right) \tag{3.46}$$

where the supremum is taken over the complex-valued measurable functions.

Proof of Lemma 10. By observing that the set of complex-valued measurable functions contains the set of real-valued measurable functions and that $\|f\|_{\infty,\mathbb{C}} = \|f\|_{\infty,\mathbb{R}}$ when f is real-valued, thanks to (3.4) we can easily obtain that

$$\sup_{\|f\|_{\infty,\mathbb{C}}\leq 1}\left|\int_{\mathbb{R}^d} f(x)\left(\mu_1(dx)-\mu_2(dx)\right)\right| \geq 2d_{TV}\left(\mu_1,\mu_2\right).$$

For what concerns the other inequality, let f be a complex-valued measurable function such that $\|f\|_{\infty,\mathbb{C}} \leq 1$. Then there exist $\psi \geq 0$ and $\rho \in [0, 2\pi)$ such that

$$\int_{\mathbb{R}^d} f(x)\left(\mu_1(dx)-\mu_2(dx)\right) = \psi e^{i\rho}.$$

Note that $\left\|\mathfrak{Re}(e^{-i\rho}f)\right\|_{\infty,\mathbb{R}} \leq \left\|e^{-i\rho}f\right\|_{\infty,\mathbb{C}} = \|f\|_{\infty,\mathbb{C}} \leq 1$. Hence

$$\begin{aligned}\left|\int_{\mathbb{R}^d} f(x)\left(\mu_1(dx)-\mu_2(dx)\right)\right| &= \int_{\mathbb{R}^d} \mathfrak{Re}(e^{-i\rho}f(x))\left(\mu_1(dx)-\mu_2(dx)\right)\\ &\leq \sup_{\|g\|_{\infty,\mathbb{R}}\leq 1}\left|\int_{\mathbb{R}^d} g(x)\left(\mu_1(dx)-\mu_2(dx)\right)\right| = 2d_{TV}\left(\mu_1,\mu_2\right)\end{aligned}$$

where for the last equality we use (3.4). Since f is arbitrary, we can conclude the proof. □

Proof of Lemma 6. Let ξ be an absolutely continuous random variable independent of (Y, Z) where the law of Y has an absolutely continuous component and let us observe that $\mu_{(Y,Z)}(dy, dz) = \mu_{Z|Y=y}(dz)\mu_Y(dy)$ where $\mu_{Z|Y=y}(dz)$ denotes the conditional law of Z given $Y = y$. Given $f : \mathbb{R}^2 \to \mathbb{R}_+$ a non-negative measurable function, one has

$$\begin{aligned}&\mathbb{E}\left(f\left(Y, Z+\xi\right)\right) = \int_{\mathbb{R}^3} f(y, z+\epsilon)\mu_{(Y,Z,\xi)}\left(dy, dz, d\epsilon\right) = \int_{\mathbb{R}^3} f(y, z+\epsilon)p_\xi\left(\epsilon\right)\mu_{(Y,Z)}\left(dy, dz\right)d\epsilon\\ &= \int_{\mathbb{R}^3} f(y, z+\epsilon)p_\xi\left(\epsilon\right)\mu_{Z|Y=y}(dz)\mu_Y(dy)d\epsilon = \int_{\mathbb{R}^3} f(y, x)p_\xi\left(x-z\right)\mu_{Z|Y=y}(dz)\mu_Y(dy)dx\\ &= \int_{\mathbb{R}^2} f(y, x)\int_{\mathbb{R}} p_\xi\left(x-z\right)\mu_{Z|Y=y}(dz)p_Y(y)dydx + \int_{\mathbb{R}^2} f(y, x)\int_{\mathbb{R}} p_\xi\left(x-z\right)\mu_{Z|Y=y}(dz)\mu_{Y,s}(dy)dx\end{aligned}$$

where $\int_{\mathbb{R}^2}\int_{\mathbb{R}} p_\xi\left(x-z\right)\mu_{Z|Y=y}(dz)p_Y(y)dydx$ is positive since Y has an absolutely continuous component and ξ is an absolutely continuous random variable. Thus the law of $(Y, Z+\xi)$ has an absolutely continuous component. □

Proof of Lemma 9. Let us first observe that by Fatou's lemma, (3.29) and the fact that p is a probability density, one has

$$\liminf_{M\to\infty}\int_{\mathbb{R}^d} p_{Y_M}(x)dx \geq \int_{\mathbb{R}^d}\liminf_{M\to\infty} p_{Y_M}(x)dx \geq \int_{\mathbb{R}^d} p(x)dx = 1.$$

Moreover $\int_{\mathbb{R}^d} p_{Y_M}(x)dx \leq 1$ for each $M \in \mathbb{N}^*$. Therefore we can conclude that

$$\lim_{M\to\infty}\int_{\mathbb{R}^d} p_{Y_M}(x)dx = 1. \tag{3.47}$$

As an immediate consequence of (3.47) we have that

$$\lim_{M\to\infty} \mu_{Y_M,s}(\mathbb{R}^d) = 1 - \lim_{M\to\infty} \int_{\mathbb{R}^d} p_{Y_M}(x)dx = 0. \tag{3.48}$$

Let us now prove the L^1 convergence. For each $M \in \mathbb{N}^*$,

$$\int_{\mathbb{R}^d} |p(x) - p_{Y_M}(x)|\, dx = 2\int_{\mathbb{R}^d} (p(x) - p_{Y_M}(x))^+\, dx - \int_{\mathbb{R}^d} (p(x) - p_{Y_M}(x))\, dx$$

where the second component of the right-hand side goes to 0 as $M \to \infty$ by (3.47) and the fact that p is a probability density while the first component goes to 0 as $M \to \infty$ by Lebesgue theorem and the fact that $0 \leq \liminf_{M\to\infty} (p(x) - p_{Y_M}(x))^+ \leq \limsup_{M\to\infty} (p(x) - p_{Y_M}(x))^+ = (\limsup_{M\to\infty} (p(x) - p_{Y_M}(x)))^+ = 0$.
Hence $\lim_{M\to\infty} \int_{\mathbb{R}^d} |p(x) - p_{Y_M}(x)|\, dx = 0$.
Let us now prove the second point of the Lemma. For a fixed $M \in \mathbb{N}^*$, one has

$$d_{TV}\left(\mu_{Y_M}, \mu_Y\right) = \sup_{A\in\mathcal{B}\left(\mathbb{R}^d\right)} \left|\mu_{Y_M}(A) - \mu_Y(A)\right| = \sup_{A\in\mathcal{B}\left(\mathbb{R}^d\right)} \left|\mu_{Y_M,c}(A) + \mu_{Y_M,s}(A) - \mu_Y(A)\right|$$

$$\leq \sup_{A\in\mathcal{B}\left(\mathbb{R}^d\right)} \left|\mu_{Y_M,c}(A) - \mu_Y(A)\right| + \mu_{Y_M,s}\left(\mathbb{R}^d\right) \leq \int_{\mathbb{R}^d} |p_{Y_M}(x) - p(x)|\, dx + \mu_{Y_M,s}\left(\mathbb{R}^d\right).$$

and the right-hand side tends to 0 as $M \to \infty$ thanks to what has been proved in the previous steps. □

Chapter 4

Central Limit Theorem for the stratified resampling mechanism

The material for this chapter has been released in [25].

The stratified resampling mechanism is one of the resampling schemes commonly used in the resampling steps of particle filters. In the present paper, we prove a central limit theorem for this mechanism under the assumption that the initial positions are independent and identically distributed and the weights proportional to a positive function of the positions such that the image of their common distribution by this function has a non zero component absolutely continuous with respect to the Lebesgue measure. This result relies on the convergence in distribution of the fractional part of partial sums of the normalized weights to some random variable uniformly distributed on $[0, 1]$, which is established in the companion paper [27] by overcoming the difficulty raised by the coupling through the normalization. Under the conjecture that a similar convergence in distribution remains valid at the next steps of a particle filter which alternates selections according to the stratified resampling mechanism and mutations according to Markov kernels, we provide an inductive formula for the asymptotic variance of the resampled population after n steps. We perform numerical experiments which support the validity of this formula.

4.1 Introduction

Particle filtering, also known as Sequential Monte Carlo methods (see Chapter 4 in [10] and [15] for a general introduction), is a powerful method to estimate the evolving state of a system over time, even when the state cannot be directly observed but can only be inferred through noisy measurements or observations. It has become a very popular class of numerical methods for the solution of optimal estimation problems in non-linear non-Gaussian scenarios. This kind of method is used in real-time applications appearing in fields such as chemical engineering, computer vision, financial econometrics, target tracking, robotics and statistics (see among others [51] and [20]). The use of Monte Carlo methods for nonlinear filtering problems can be traced back to Handschin [32] and Mayne [33]. They introduced a sequential version of the importance sampling method (see [22], [52] as references for the importance sampling) and the corresponding algorithm is known as sequential importance sampling, often abbreviated SIS. A drawback of the latter was identified by Gordon et al. in [29]: as the number of iterations increases, the importance weights tend to degenerate (a phenomenon usually known as *weight degeneracy*). This means that after a certain number of iterations, some weights tend to become very small so that the corresponding positions no longer contribute to the estimation. Thus, Gordon et al. introduced the resampling step where, in view of stabilizing the Monte Carlo error over time, the key idea is to eliminate the particles having low weights and to replicate the particles having high weights. Therefore, by propagating M particles through weighting, resampling and mutation steps (each particle evolves randomly according to a given transition probability kernel), particle filters can be used to numerically estimate the state of

the system given the observations (see also [20], [15]).

Let us henceforth concentrate on the resampling step. Let $M \geq 1$. Given a sequence of $\mathbb{R}^d$-valued random vectors $(X_m)_{1\leq m\leq M}$ with associated random weights $(w_m^M)_{1\leq m\leq M}$ such that $w_m^M > 0$ and $\sum_{m=1}^{M} w_m^M = 1$, a resampling scheme defines the resampled sequence $(Y_m)_{1\leq m\leq M}$ such that

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m}\,\middle|\,\mathcal{G}\right) = \sum_{m=1}^{M} w_m^M \delta_{X_m}.$$

where $\mathcal{G} := \sigma\left((X_m, w_m^M)_{1\leq m\leq M}\right)$. Resampling schemes [19] permit to replace the probability measure $\sum_{m=1}^{M} w_m^M \delta_{X_m}$ with non equal weights by some empirical measure $\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m}$ with the same conditional expectation given $\mathcal{G}$. Depending on the definition used for the random variables Y_m, several resampling schemes can be considered. The most common resampling techniques are of the following types: multinomial, residual , stratified [37] and systematic [12]. See [19] for a brief description of such methods. The simplest approach is the multinomial resampling. It consists in drawing, conditionally upon $\mathcal{G}$, the new positions $(Y_m)_{1\leq m\leq M}$ independently from the common distribution $\sum_{m=1}^{M} w_m^M \delta_{X_m}$. While the residual resampling consists in replicating $\lfloor M w_m^M \rfloor$-times X_m for $m \in \{1, \cdots, M\}$ and the remaining $M - \sum_{m=1}^{M} \lfloor M w_m^M \rfloor$ variables Y_m are drawn, conditionally upon $\mathcal{G}$, independently from the common distribution $\frac{1}{M-\sum_{m=1}^{M}\lfloor M w_m^M\rfloor}\sum_{\ell=1}^{M}\left\{M w_\ell^M\right\}\delta_{X_\ell}$. Concerning the stratified resampling, it is not straightforward to understand its behaviour, notably due to the complicated structure in the definition of the Y_m which, however, continue to be conditionally independent given $\mathcal{G}$. The systematic resampling is even more complicated to understand since the Y_m are no longer conditionally independent given $\mathcal{G}$.

Resampling schemes have been largely studied in the literature, we now present a selection of such results. The asymptotic behaviour for the multinomial resampling scheme has been extensively studied in [47] (see Corollary 7.4.2 and Section 9.4.2). Douc et al. in [19] showed that residual and stratified resampling improve over multinomial resampling in the sense that they have a lower conditional variance (with respect to the σ-algebra generated by $(X_m)_{1\leq m\leq M}$). They also proved, by means of a counter-example, that the same property does not hold for systematic resampling. Furthermore, they established a central limit theorem for the residual resampling approach suggesting that a similar result should be obtained for the stratified resampling scheme. One of the last contributions concerning resampling schemes, is given by Gerber et al. [28]. Using the notion of negative association [35], they first provided a general consistency result for resampling. An application of this theorem gives the proof of almost sure weak convergence of $\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m}$ in the stratified resampling method. Moreover they provide a counter-example to almost sure weak convergence for the systematic resampling method. More recently, Chopin et al. [16] studied the resampling schemes for particle filters with weakly informative observations. Empirical evidence indicates that when the weights used in resampling exhibit high variability, the selection of the resampling strategy tends to have a weak impact. However, in cases where the weights are close to being uniform, the performance differences between the different resampling methods can be substantial. By keeping M fixed, they also considered the asymptotic behaviour of the resampling schemes as the weights become less and less informative. See also [41], [23], [38], [14] for additional references.

In this paper, we focus on the stratified resampling scheme for the weights $w_m^M = g(X_m)/\sum_{\ell=1}^{M} g(X_\ell)$ where g is a positive measurable function. Our purpose is to study the asymptotic behaviour of the method as the number of particles M goes to ∞. Let f be a real-valued measurable function. To

compute the asympotic variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m)$, it turns out to be essential to understand the behaviour as $M \to \infty$ of

$$\left\{ \frac{M \sum_{\ell=1}^{\lceil \alpha M \rceil - 1} g(X_\ell)}{\sum_{\ell=1}^{M} g(X_\ell)} \right\} \tag{4.1}$$

where $0 < \alpha < 1$ and where $\{x\}$ denotes the fractional part of $x \in \mathbb{R}$. Under the assumption that the X_i are i.i.d. such that the law of $g(X_i)$ has an absolutely continuous component, we prove in the companion paper [27] the convergence in distribution of (4.1) to a random variable uniformly distributed on $[0,1]$. Under this assumption, we explicit the limit of $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m)\right)$ and prove that $\frac{1}{\sqrt{M}}\sum_{m=1}^{M}\left(f(Y_m^M) - \frac{\mathbb{E}(f(X_1)g(X_1))}{\mathbb{E}(g(X_1))}\right)$ converges in distribution to a centered Gaussian random variable with variance equal to the limit. The proof of this Central Limit Theorem relies on the asymptotic behaviour as $M \to \infty$ of a vector composed of $\sqrt{M}\left(\frac{\sum_{m=1}^{M} g(X_m) f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \mathbb{E}(f(X_1)\, g(X_1))\right)$ and the fractional parts (4.1) for $\alpha \in \{\alpha_1, \cdots, \alpha_s\}$ with $0 < \alpha_1 < \cdots < \alpha_s < 1$. In the companion paper [27], we check that this vector converges in distribution to a vector with centered Gaussian first component and independent s last components uniformly distributed on $[0,1]^s$. Under the conjecture that a similar convergence in distribution remains valid at the next steps of a particle filter which alternates selections according to the stratified resampling mechanism and mutations according to Markov kernels, we provide an inductive formula for the asymptotic variance of the resampled population after n steps.

The paper is organized as follows. In Section 4.2, we recall the definition of the stratified sampling scheme and the statement of the main result (Theorem 34) is given. In Section 4.3, the asymptotic variance is derived and in Section 4.4, the proof of the Central Limit Theorem is given. In Section 4.6, we consider a particle filter which alternates selections according to the stratified resampling mechanism and mutations according to Markov kernels and we provide an inductive formula for the asymptotic variance of the resampled sequence. We perform numerical experiments which support the validity of this formula.

Notation We denote by $\lfloor x \rfloor$ the integer j such that $j \leq x < j+1$ and by $\{x\} = x - \lfloor x \rfloor$ the fractional part of $x \in \mathbb{R}$. We denote the set of real-valued bounded measurable functions on $\mathbb{R}^d$ by $\mathcal{B}_b(\mathbb{R}^d)$. Given μ a positive measure on $\mathbb{R}^d$ endowed with the Borel sigma algebra and $\phi : \mathbb{R}^d \to \mathbb{R}$ a measurable function that is either positive or such that $\int_{\mathbb{R}^d} |\phi(x)|\, \mu(dx) < \infty$, we denote $\mu(\phi) = \int_{\mathbb{R}^d} \phi(x)\, \mu(dx)$.

4.2 Statement of the Main Result

Given $M > 0$, let $(X_m)_{1 \leq m \leq M}$ be a sequence of i.i.d. $\mathbb{R}^d$-valued random vectors following the law η and let $g : \mathbb{R}^d \to (0, \infty)$ be a measurable function such that $0 < \inf_{x \in \mathbb{R}^d} g(x) \leq \sup_{x \in \mathbb{R}^d} g(x) < \infty$. In what follows we denote $\bar{g} := \sup_{x \in \mathbb{R}^d} g(x)$ and $\underline{g} := \inf_{x \in \mathbb{R}^d} g(x)$.
We now generate the sequence $(Y_m^M)_{1 \leq m \leq M}$ according to the selection step of the stratified sampling. We recall that starting from M random variables $(U_m)_{1 \leq m \leq M}$ i.i.d. distributed according to the uniform law on $(0,1)$ and independent of $(X_m)_{1 \leq m \leq M}$, the sequence $(Y_m^M)_{1 \leq m \leq M}$ is defined in the following way

$$Y_m^M = \sum_{\ell=1}^{M} 1_{\left\{ \sum_{j=1}^{\ell-1} w_j^M < m - U_m \leq \sum_{j=1}^{\ell} w_j^M \right\}} X_\ell \ \textit{for}\, m \in \{1, \cdots, M\} \tag{4.2}$$

where $w_m^M = \frac{Mg(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)}$ for $m = 1, \cdots, M$. Since the weights are preserved when multiplied by a positive constant, up to dividing g by $\mathbb{E}(g(X_1))$, without loss of generality we may suppose that

$$\mathbb{E}(g(X_1)) = 1.$$

Let us observe that the random vectors $(Y_m^M)_{1\leq m\leq M}$ are conditionally independent given $\mathcal{F}$, the σ-algebra generated by the sequence $(X_m)_{1\leq m\leq M}$. Moreover one has

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{Y_m^M}|\mathcal{F}\right) = \frac{1}{\sum_{m=1}^{M} g(X_m)}\sum_{m=1}^{M} g(X_m)\delta_{X_m}. \tag{4.3}$$

Given $f : \mathbb{R}^d \to \mathbb{R}$ a measurable function, our purpose is to provide a central limit theorem for $\frac{1}{M}\sum_{m=1}^{M} f(Y_m^M)$. If we start looking at the variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)$, using (4.3) one has

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\right) && (4.4)\\
&= \mathrm{Var}\left(\mathbb{E}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right) && (4.5)\\
&= \mathrm{Var}\left(\sqrt{M}\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)}\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right). && (4.6)
\end{aligned}$$

Thus, in particular, we will prove respectively the convergence of the first and second term in (4.6). The first term is common to all the resampling schemes while the second one really depends on the considered resampling scheme.

Let us consider the following hypotheses:

I1 $\sup_{x\in\mathbb{R}^d} |f(x)| < \infty$

I2 the law of $g(X_1)$ has an absolutely continuous component with respect to the Lebesgue measure on $\mathbb{R}$

Before providing the statement of the central limit theorem, to express the asymptotic variance we introduce the real-valued functions β_0 and β_1 respectively given by

$$\beta_0(x, y_1) = \{x + y_1\}(1 - \{x + y_1\}) + x(1 - x) - 2x(1 - x - y_1)1_{\{y_1<1-x\}} \tag{4.7}$$

$$\begin{aligned}
\beta_1(x, y_1, y_2, y_3) = 2\Big(&\{x + y_1\}\Big(1 - \{x + y_1\} - y_2\Big)1_{\{y_2<1-\{x+y_1\}\}} \\
&- \{x + y_1\}\Big(1 - \{x + y_1\} - y_2 - y_3\Big)1_{\{y_2+y_3<1-\{x+y_1\}\}} \\
&- x\Big(1 - x - y_1 - y_2\Big)1_{\{y_1+y_2<1-x\}} \\
&+ x\,(1 - x - y_1 - y_2 - y_3)\,1_{\{y_1+y_2+y_3<1-x\}}\Big).
\end{aligned} \tag{4.8}$$

Remark 33. *Let us observe that β_0 is continuous. Indeed, the fractional part is continuous apart from the integers where its left-hand limit is equal to 1 and its right-hand limit is equal to 0. The composition with the function $f(z) = z(1-z)$ which is such that $f(0) = f(1) = 0$, allows us to conclude that $(x, y_1) \mapsto \{x + y_1\}(1 - \{x + y_1\})$ is continuous. Moreover $(x, y_1) \mapsto x(1 - x - y_1)1_{\{y_1<1-x\}}$ is continuous since the function $\bar{f}(x, y_1) = 1-x-y_1$ is equal to 0 on the set $\{(x, y_1) \in \mathbb{R}^2 : y_1 = 1 - x\}$ of discontinuity points of the indicator function. Similarly it is possible to prove that also β_1 is a continuous function on $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+$.*

The following result holds.

Theorem 34. *Under the notation introduced above and under **I1-2** we have*

$$\lim_{M\to\infty} Var\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_m^M)\right) = \sigma_1^2(f) + \sigma_2^2(f) \tag{4.9}$$

where

$$\sigma_1^2(f) = \eta\left((g(f - \eta(fg)))^2\right) \tag{4.10}$$

and

$$\sigma_2^2(f) := \sum_{k=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \mathbb{E}(F_k) \tag{4.11}$$

with $(F_k)_{k\in\mathbb{N}}$ given by

$$F_k = \begin{cases} f^2(X_1)\beta_0(U_1, g(X_1)) & k = 0 \\ -f(X_1)f(X_{k+1})\beta_1\left(U_1, g(X_1), \sum_{\ell=2}^{k} g(X_\ell), g(X_{k+1})\right) & k > 0 \end{cases} \tag{4.12}$$

where $U_1 \sim \mathcal{U}(0,1)$ is independent of $X_1, \cdots, X_{k+1}$.

Moreover the following convergence in distribution holds

$$\sqrt{M}\left(\frac{1}{M}\sum_{m=1}^{M} f(Y_m^M) - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \sigma_1^2(f) + \sigma_2^2(f)\right). \tag{4.13}$$

We will split the proof of the theorem into two parts: in Section 4.3, we are going to prove the result about the asymptotic variance and in Section 4.4, the proof of the Central Limit Theorem is provided.

4.3 Asymptotic Variance

The following proposition provides the asymptotic behaviour of the first term in (4.6). We provide its proof for the sake of completeness.

Proposition 7. *Under **I1**, the following convergence in distribution holds*

$$\sqrt{M}\left(\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \sigma_1^2(f)\right) \tag{4.14}$$

where $\sigma_1^2(f)$ has been defined in (4.10). Moreover

$$\lim_{M\to\infty} Var\left(\sqrt{M}\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)}\right) = \sigma_1^2(f). \tag{4.15}$$

Proof. We recall that we suppose that $\eta(g) = 1$. Let us define $h := g\,(f - \eta(fg))$. By observing that $\eta(h) = 0$ and

$$\sqrt{M}\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \sqrt{M}\eta(fg) = \frac{M}{\sum_{m=1}^{M} g(X_m)} \times \frac{1}{\sqrt{M}} \sum_{m=1}^{M} h(X_m), \tag{4.16}$$

by Slutsky's theorem we have

$$\sqrt{M}\left(\frac{\sum_{m=1}^{M} g(X_m)f(X_m)}{\sum_{\ell=1}^{M} g(X_\ell)} - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \eta(h^2)\right).$$

Let us now prove (4.15). Let us preliminary study the following quantity

$$\mathbb{E}\left(\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h(X_m)\right)^4\right) = \frac{1}{M^2}\sum_{m_1,m_2,m_3,m_4=1}^{M} \mathbb{E}\left(\prod_{i=1}^{4} h(X_{m_i})\right).$$

Among the M^4 expectations appearing in the sum on the right-hand side, the ones where one index is different from the other three are equal to 0 by the independence of the X_i and the fact that $\eta(h) = 0$. Therefore

$$\mathbb{E}\left(\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h(X_m)\right)^4\right) = \frac{1}{M}\eta(h^4) + \frac{3(M-1)}{M}\eta(h^2)^2 \leq D, \tag{4.17}$$

where D is a finite constant depending on g and f but not on M.
Since (4.16) holds, to prove (4.15) we can study the convergence as M goes to infinity of the variance of $\sqrt{M}\dfrac{1}{\bar{g}_M}\bar{h}_M$ given by

$$\mathbb{E}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2\right) - \mathbb{E}\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2$$

where we denote by $\bar{g}_M$ the empirical mean $\dfrac{1}{M}\sum_{m=1}^{M} g(X_m)$ and by $\bar{h}_M$ the empirical mean $\dfrac{1}{M}\sum_{m=1}^{M} h(X_m)$.

By the Cauchy–Schwarz inequality we have

$$\begin{aligned}
\mathbb{E}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M} - \sqrt{M}\bar{h}_M\right)^2\right) &= \mathbb{E}\left(\left(\sqrt{M}\bar{h}_M\right)^2\left(\frac{1}{\bar{g}_M} - 1\right)^2\right) \\
&\leq \mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\bar{h}_M\right)^4\right)\mathbb{E}^{\frac{1}{2}}\left(\left(\frac{1}{\bar{g}_M} - 1\right)^4\right) \\
&\leq D^{\frac{1}{2}} \cdot \mathbb{E}^{\frac{1}{2}}\left(\left(\frac{1}{\bar{g}_M} - 1\right)^4\right),
\end{aligned}$$

where the right-hand side converges to 0 as M goes to infinity by the Strong Law of Large Numbers and by Lebesgue's theorem. With $\eta(h) = 0$, this in particular implies that

$$0 = \lim_{M\to\infty}\left|\mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2\right) - \mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\bar{h}_M\right)^2\right)\right| = \lim_{M\to\infty}\left|\mathbb{E}^{\frac{1}{2}}\left(\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)^2\right) - \eta(h^2)^{\frac{1}{2}}\right|$$

and

$$0 = \lim_{M\to\infty}\left|\mathbb{E}\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right) - \mathbb{E}\left(\sqrt{M}\bar{h}_M\right)\right| = \lim_{M\to\infty}\left|\mathbb{E}\left(\sqrt{M}\frac{\bar{h}_M}{\bar{g}_M}\right)\right|$$

and so the proof is complete. □

Let us now study the second term in the expression (4.6). Our purpose is to prove the following result:

Theorem 35. *Under **I1-2**, the following convergence holds*

$$\lim_{M\to\infty} \mathbb{E}\left(Var\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) | \mathcal{F} \right) \right) = \sigma_2^2(f) \tag{4.18}$$

where $\sigma_2^2(f)$ has been defined in (4.11).

To prove Theorem 35 we need some preliminary results the proofs of which are given in Section 4.5. By the conditional independence of the random vectors $(Y_m^M)_{1\le m\le M}$ with respect to the σ-algebra $\mathcal{F}$, we have

$$\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) \,\middle|\, \mathcal{F} \right) = \frac{1}{M} \sum_{m=1}^{M} \mathrm{Var}\left(f(Y_m^M) \,\middle|\, \mathcal{F} \right) \tag{4.19}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}\left(f^2(Y_m^M) \,\middle|\, \mathcal{F} \right) - \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}\left(f(Y_m^M) \,\middle|\, \mathcal{F} \right)^2. \tag{4.20}$$

We introduce the following notation that will be useful in what follows: for $i = 1, \cdots, M$, let us denote $u_i^M := \left\{ w_1^M + \cdots + w_i^M \right\}$, $\mu_i^M := \lfloor w_1^M + \cdots + w_i^M \rfloor + 1$ where $w_i^M = \frac{Mg(X_i)}{\sum_{\ell=1}^{M} g(X_\ell)}$ has been introduced in (4.2) and by convention $u_0^M = 0$, $\mu_0^M = 1$.

We now rewrite in a more explicit way the conditional variance $\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) | \mathcal{F} \right)$. The proof is given in the Section 4.5.

Proposition 8. *Under **I1**, we have*

$$Var\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) \,\middle|\, \mathcal{F} \right)$$

$$= \frac{1}{M} \sum_{i=1}^{M} f^2(X_i) \beta_0(u_{i-1}^M, w_i^M) \tag{4.21}$$

$$- \frac{1}{M} \sum_{k=1}^{(M-1)\wedge \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \sum_{i=1}^{M-k} f(X_i) f(X_{i+k}) \beta_1(u_{i-1}^M, w_i^M, \sum_{\ell=2}^{k} w_{i+\ell-1}^M, w_{i+k}^M) \tag{4.22}$$

where β_0 and β_1 are respectively defined in (4.7) and (4.8).

In what follows, without specifying it every time, we will suppose that $M \ge 1 + \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil$ so that $(M-1) \wedge \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil = \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil$. Let us observe that it is possible to rewrite (4.21) and (4.22) in the following way

$$\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_m^M) \,\middle|\, \mathcal{F} \right) = \sum_{k=0}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \int_0^1 F_{k,\alpha_1}^M d\alpha_1 \tag{4.23}$$

where F_{k,α_1}^M is given by $f^2(X_{\lceil \alpha_1 M \rceil}) \beta_0(u_{\lceil \alpha_1 M \rceil - 1}^M, w_{\lceil \alpha_1 M \rceil}^M)$ if $k = 0$ and by

$$-1_{\{\lceil \alpha_1 M \rceil \le M-k\}} f(X_{\lceil \alpha_1 M \rceil}) f(X_{\lceil \alpha_1 M \rceil + k}) \beta_1(u_{\lceil \alpha_1 M \rceil - 1}^M, w_{\lceil \alpha_1 M \rceil}^M, \sum_{\ell=2}^{k} w_{\lceil \alpha_1 M \rceil + \ell - 1}^M, w_{\lceil \alpha_1 M \rceil + k}^M)$$

if $k > 0$. Given $0 < \alpha_1 < 1$ and $k \in \mathbb{N}$, F^M_{k,α_1} is a function of the vector

$$T^M_{\alpha_1,k} := \left(f\left(X_{\lceil \alpha_1 M \rceil}\right), f\left(X_{\lceil \alpha_1 M \rceil + k}\right), u^M_{\lceil \alpha_1 M \rceil - 1}, w^M_{\lceil \alpha_1 M \rceil}, \sum_{\ell=2}^{k} w^M_{\lceil \alpha_1 M \rceil + \ell - 1}, w^M_{\lceil \alpha_1 M \rceil + k} \right)$$

which is well defined for M big enough. Therefore to study the limit behavior as $M \to \infty$ of the expectation of F^M_{k,α_1}, we are going to study the limit behavior of the vector $T^M_{\alpha_1,k}$. To obtain a much more general formulation useful in the next section, we introduce also the following family of vectors well defined for M big enough. Let $s \geq 1$ and $0 < \alpha_1 < \cdots < \alpha_s < 1$. For $k := (k_1, \cdots, k_s) \in \mathbb{N}^s$, let us define $\forall i = 1, \cdots, s$

$$T^M_{\alpha_i,k_i} := \left(f\left(X_{\lceil \alpha_i M \rceil}\right), f\left(X_{\lceil \alpha_i M \rceil + k_i}\right), u^M_{\lceil \alpha_i M \rceil - 1}, w^M_{\lceil \alpha_i M \rceil}, \sum_{\ell=2}^{k_i} w^M_{\lceil \alpha_i M \rceil + \ell - 1}, w^M_{\lceil \alpha_i M \rceil + k_i} \right). \tag{4.24}$$

Moreover let

$$H^M := \sqrt{M} \left(\frac{\sum_{m=1}^{M} (fg)(X_m)}{\sum_{m=1}^{M} g(X_m)} - \eta(fg) \right) \tag{4.25}$$

that thanks to (4.14) converges in distribution to a centered Gaussian random variable with variance $\sigma_1^2(f)$. The purpose of the next proposition is therefore to study the convergence in distribution of $\left(H^M, T^M_{\alpha_1,k_1}, \cdots, T^M_{\alpha_s,k_s}\right)$.

Proposition 9. *Assume **I1-2**. Let $s \in \mathbb{N}^*$ and $k := (k_1, \cdots, k_s) \in \mathbb{N}^s$. We set $K_i := k_1 + \cdots + k_{i-1} + i - 1$ for $i \in \{2, \cdots, s\}$ and $K_1 := 0$. Let $(U_j)_{1 \leq j \leq s}$ be a sequence of i.i.d random variables distributed according to the uniform law on $(0,1)$ and independent of $H \sim \mathcal{N}\left(0, \sigma_1^2(f)\right)$ with $(U_1, \cdots, U_s, H)$ independent from $(X_j)_{1 \leq j \leq K_{s+1}}$. For $0 < \alpha_1 < \cdots < \alpha_s < 1$, as M goes to infinity, the following convergence in distribution holds*

$$\left(H^M, T^M_{\alpha_1,k_1}, \cdots, T^M_{\alpha_s,k_s}\right) \overset{d}{\Longrightarrow} (H, T_{1,k}, \cdots, T_{s,k}) \tag{4.26}$$

where

$$T_{i,k} := \left(f\left(X_{K_i+1}\right), f\left(X_{K_{i+1}}\right), U_i, g(X_{K_i+1}), \sum_{\ell=K_i+2}^{K_{i+1}-1} g(X_\ell), g(X_{K_{i+1}}) \right) \quad \forall i = 1, \cdots, s. \tag{4.27}$$

Proof of Proposition 9. Let $M \geq \max\left(\frac{2}{\alpha_1}, \frac{k_s}{1-\alpha_s}, \max_{1 \leq i \leq s-1} \frac{k_i+1}{\alpha_{i+1}-\alpha_i} \right)$ so that $2 \leq \lceil \alpha_i M \rceil \leq M - k_i$ $\forall\, i = 1, \cdots, s$ and $k_i < \lceil \alpha_{i+1} M \rceil - \lceil \alpha_i M \rceil$ $\forall\, i = 1, \cdots, s-1$. The last condition allows us to separate the variables: $\left(X_{\lceil \alpha_i M \rceil}, \cdots, X_{\lceil \alpha_i M \rceil + k_i}\right)$ is independent of $\left(X_{\lceil \alpha_j M \rceil}, \cdots, X_{\lceil \alpha_j M \rceil + k_j}\right)$ for i, j distinct elements of $\{1, \cdots, s\}$.
By defining $J^M_{\alpha_i,k_i} := \{\lceil \alpha_i M \rceil, \cdots, \lceil \alpha_i M \rceil + k_i\}$ for $i = 1, \cdots, s$, let $J^M := \bigcup_{i=1}^s J^M_{\alpha_i,k_i}$.
Moreover let $(\tilde{X}_m)_{m \geq 1}$ be a sequence of i.i.d. $\mathbb{R}^d$-valued random vectors following the law η independent of $(X_m)_{1 \leq m \leq M}$. Given $x := (x_j)_{1 \leq j \leq K_{s+1}} \in \mathbb{R}^{d(K_{s+1})}$, we define

$$H^M_x := \sqrt{M} \left(\frac{\sum_{j=1}^{K_{s+1}} (fg)(x_j) + \sum_{j \notin J^M_s} (fg)(X_j)}{\sum_{j=1}^{K_{s+1}} g(x_j) + \sum_{j \notin J^M_s} g(X_j)} - \eta(fg) \right)$$

and

$$\tilde{H}_x^M := \sqrt{M}\left(\frac{\sum_{j=1}^{K_{s+1}}(fg)(x_j)+\sum_{j=1}^{M-K_{s+1}}(fg)(\tilde{X}_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}-\eta(fg)\right).$$

Similarly for $i=1,\cdots,s$ we define the vector $T^M_{\alpha_i,k_i,x}$ by

$$\left(f(x_{K_i+1}),f\left(x_{K_{i+1}}\right),\left\{M\times\frac{\sum_{j=1}^{\lceil\alpha_i M\rceil-1}g(X_j)+\sum_{j=1}^{K_i}g(x_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)}\right\},\frac{Mg(x_{K_i+1})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)},\frac{M\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)},\frac{Mg(x_{K_{i+1}})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j\notin J_s^M}g(X_j)}\right)$$

and the vector $\tilde{T}^M_{\alpha_i,k_i,x}$ by $\left(f\left(x_{K_i+1}\right),f\left(x_{K_{i+1}}\right),\tilde{T}^M_{\alpha_i,k_i,x,f}\right)$ with

$$\tilde{T}^M_{\alpha_i,k_i,x,f}=\left(\left\{M\times\frac{\sum_{j=1}^{\lceil\alpha_i M\rceil-K_i-1}g(\tilde{X}_j)+\sum_{j=1}^{K_i}g(x_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}\right\},\frac{Mg(x_{K_i+1})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)},\frac{M\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)},\frac{Mg(x_{K_{i+1}})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}\right).$$

We notice that

$$\left(H_x^M,T^M_{\alpha_1,k_1,x},\cdots,T^M_{\alpha_s,k_s,x}\right)\overset{\mathcal{L}}{=}\left(\tilde{H}_x^M,\tilde{T}^M_{\alpha_1,k_1,x},\cdots,\tilde{T}^M_{\alpha_s,k_s,x}\right). \tag{4.28}$$

Let us observe that by hypothesis **I2** the law of $g\left(\tilde{X}_i\right)$ has a component absolutely continuous. Moreover since $\int_{\mathbb{R}}|1+y|^{s+1}e^{-y^2}dy<\infty\ \forall s\geq 1$, we can apply Theorem 1 of [27] with $q=s$ and

- (Y_i,Z_i) equal to $\left(g\left(\tilde{X}_i\right),(fg)(\tilde{X}_i)\right)$
- $\left(\beta_M^1,\cdots,\beta_M^s,\beta_M^{s+1}\right)$ equal to $(\lceil\alpha_1 M\rceil-K_1-1,\cdots,\lceil\alpha_s M\rceil-K_s-1,M-K_{s+1})$
- $\phi(x)=\frac{1}{x}$
- x equal to $\sum_{j=1}^{K_{s+1}}g(x_j)$, z equal to $\sum_{j=1}^{K_{s+1}}(fg)(x_j)$, θ equal to $\eta(fg)$ and $(y_1,\cdots,y_s)$ equal to $\left(\sum_{j=1}^{K_1}g(x_j),\cdots,\sum_{j=1}^{K_s}g(x_j)\right)$
- $T:=H\sim\mathcal{N}\left(0,\sigma_1^2(f)\right)$ by (4.14)

and deduce the following convergence in distribution as $M\to\infty$

$$\left(\tilde{H}_x^M,\left(\left\{M\times\frac{\sum_{j=1}^{\lceil\alpha_i M\rceil-K_i-1}g(\tilde{X}_j)+\sum_{j=1}^{K_i}g(x_j)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}\right\}\right)_{1\leq i\leq s}\right)\overset{d}{\Longrightarrow}(H,U_1,\cdots,U_s). \tag{4.29}$$

Moreover, by the Strong Law of Large Numbers

$$\lim_{M\to\infty}\frac{Mg(x_{K_i+1})}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}=g(x_{K_i+1}),$$

$$\lim_{M\to\infty}\frac{M\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)}{\sum_{j=1}^{K_{s+1}}g(x_j)+\sum_{j=1}^{M-K_{s+1}}g(\tilde{X}_j)}=\sum_{\ell=K_i+2}^{K_{i+1}-1}g(x_\ell)$$

and

$$\lim_{M\to\infty} \frac{Mg(x_{K_{i+1}})}{\sum_{j=1}^{K_{s+1}} g(x_j) + \sum_{j=1}^{M-K_{s+1}} g(\tilde{X}_j)} = g(x_{K_{i+1}}).$$

Therefore by Slutsky's theorem, for each $x := (x_j)_{1\le j\le K_{s+1}} \in \mathbb{R}^{d(K_{s+1})}$ we can conclude that the following random vector

$$\left(\tilde{H}_x^M, \tilde{T}^M_{\alpha_1,k_1,x}, \cdots, \tilde{T}^M_{\alpha_s,k_s,x}\right)$$

converges in distribution as M goes to infinity to

$$(H, T_{1,k,x}, \cdots, T_{s,k,x}) \tag{4.30}$$

where the vector $T_{i,k,x}$ is given by

$$\left(f\left(x_{K_i+1}\right), f\left(x_{K_{i+1}}\right), U_i, g(x_{K_i+1}), \sum_{\ell=K_i+2}^{K_{i+1}-1} g(x_\ell), g(x_{K_{i+1}})\right).$$

In particular, given $\phi : \mathbb{R}^{1+6s} \to \mathbb{R}$ a continuous bounded function one has

$$\lim_{M\to\infty} \mathbb{E}\left(\phi\left(\tilde{H}_x^M, \tilde{T}^M_{\alpha_1,k_1,x}, \cdots, \tilde{T}^M_{\alpha_s,k_s,x}\right)\right) = \mathbb{E}\left(\phi\left(H, T_{1,k,x}, \cdots, T_{s,k,x}\right)\right). \tag{4.31}$$

We are now ready to prove (4.26). By conditioning with respect to $\sigma\left((X_i)_{i\in J_s^M}\right)$ and by applying the Freezing Lemma, (4.28) and (4.31), one has

$$\begin{aligned}
&\mathbb{E}\left(\phi\left(H^M, T^M_{\alpha_1,k_1}, \cdots, T^M_{\alpha_s,k_s}\right)\right) \\
&\quad= \int_{\mathbb{R}^{d(k_1+\cdots+k_s+s)}} \mathbb{E}\left(\phi\left(H_x^M, T^M_{\alpha_1,k_1,x}, \cdots, T^M_{\alpha_s,k_s,x}\right)\right) \prod_{i=1}^{K_{s+1}} \eta(dx_i) \\
&\quad= \int_{\mathbb{R}^{d(k_1+\cdots+k_s+s)}} \mathbb{E}\left(\phi\left(\tilde{H}_x^M, \tilde{T}^M_{\alpha_1,k_1,x}, \cdots, \tilde{T}^M_{\alpha_s,k_s,x}\right)\right) \prod_{i=1}^{K_{s+1}} \eta(dx_i) \\
&\quad\underset{M\to\infty}{\longrightarrow} \mathbb{E}\left(\phi\left(H, T_{1,k}, \cdots, T_{s,k}\right)\right)
\end{aligned}$$

where $T_{i,k}$, $i = 1, \cdots, s$ is defined in (4.27) and so the proof is complete. □

We are now ready to prove Theorem 35 .

Proof of Theorem 35. By (4.23), it is enough to study the convergence as M goes to infinity of

$$\sum_{k=0}^{\left\lceil \frac{\bar{g}}{g} \right\rceil} \int_0^1 \mathbb{E}\left(F^M_{k,\alpha_1}\right) d\alpha_1. \tag{4.32}$$

Let us observe that we can replace β_0 and β_1 (introduced in (4.7) and (4.8)) in the definition of F^M_{k,α_1} respectively by the bounded functions

$$\beta_0\left(0 \vee x \wedge 1, 0 \vee y_1\right), \beta_1\left(0 \vee x \wedge 1, 0 \vee y_1, 0 \vee y_2, 0 \vee y_3\right)$$

that to simplify the notation we will keep calling β_0 and β_1.

For what has been said there exists a constant $\tilde{C} < \infty$ such that

$$\sup_{(x,y_1)\in\mathbb{R}^2} |\beta_0(x,y_1)| \vee \sup_{(x,y_1,y_2,y_3)\in\mathbb{R}^4} |\beta_1(x,y_1,y_2,y_3)| \leq \tilde{C}. \tag{4.33}$$

By Proposition 9 applied with $s=1$ and by Remark 33, we have that for each $\alpha_1 \in (0,1)$

$$F^M_{0,\alpha_1} \overset{d}{\Longrightarrow} F_0$$

with F_0 defined in (4.12). Moreover by the hypothesis of boundedness of f and (4.33) we have

$$\lim_{M\to\infty} \mathbb{E}\left(F^M_{0,\alpha_1}\right) = \mathbb{E}\left(F_0\right).$$

By Lebesgue's theorem, $\int_0^1 \mathbb{E}\left(F^M_{0,\alpha_1}\right) d\alpha_1$ will converge to the same limit.

Let us now study the convergence of $\sum_{k=1}^{\left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil} \int_0^1 \mathbb{E}\left(F^M_{k,\alpha_1}\right) d\alpha_1$. By Proposition 9 applied with $s=1$ and by Remark 33, we have that for each $\alpha_1 \in (0,1)$ and $k \geq 1$

$$F^M_{k,\alpha_1} \overset{d}{\Longrightarrow} F_k$$

with F_k defined in (4.12). Moreover by the hypothesis of boundedness of f and (4.33),

$$\lim_{M\to\infty} \mathbb{E}\left(F^M_{k,\alpha_1}\right) = \mathbb{E}\left(F_k\right).$$

By Lebesgue's theorem, $\lim_{M\to\infty} \int_0^1 \mathbb{E}\left(F^M_{k,\alpha_1}\right) d\alpha_1 = \mathbb{E}\left(F_k\right)$ for each $k = 1, \cdots, \left\lceil \frac{\bar{g}}{\underline{g}} \right\rceil$ and so the proof is complete. □

We are now ready to provide a central limit theorem for $\frac{1}{M}\sum_{m=1}^{M} f(Y^M_m)$ knowing its asymptotic variance.

4.4 Central Limit Theorem

Theorem 36. *Let assume **I1-2**. The following convergence in distribution holds*

$$\sqrt{M}\left(\frac{1}{M}\sum_{m=1}^{M} f(Y^M_m) - \eta(fg)\right) \overset{d}{\Longrightarrow} \mathcal{N}\left(0, \sigma_1^2(f) + \sigma_2^2(f)\right) \tag{4.34}$$

where $\sigma_1^2(f)$ and $\sigma_2^2(f)$ are respectively defined in (4.10) and (4.11).

Proof of Theorem 36. Let $u \in \mathbb{R}$. By introducing the notation

$$a^M_m := e^{\frac{iu}{\sqrt{M}}\left(f(Y^M_m) - \mathbb{E}\left(f(Y^M_m)|\mathcal{F}\right)\right)} \text{ for } m = 1, \cdots, M \tag{4.35}$$

and by applying (4.3), we have

$$\begin{aligned}
\mathbb{E}\left(e^{\frac{iu}{\sqrt{M}}\sum_{m=1}^{M}\left(f(Y^M_m)-\eta(fg)\right)}\right) &= \mathbb{E}\left(e^{iu\sqrt{M}\left(\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} f(Y^M_m)\middle|\mathcal{F}\right)-\eta(fg)\right)}\mathbb{E}\left(\prod_{m=1}^{M} a^M_m \,\middle|\, \mathcal{F}\right)\right) \\
&= \mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M} a^M_m \,\middle|\, \mathcal{F}\right)\right)
\end{aligned} \tag{4.36}$$

where H^M is defined in (4.25). The purpose of what follows is therefore to study the convergence as M goes to infinity of (4.36).

First step

We are first going to study the conditional expectation $\mathbb{E}\left(\prod_{m=1}^{M} a_m^M \,\middle|\, \mathcal{F}\right)$.

By using that for each $u, y \in \mathbb{R}$ $\left|e^{iuy} - 1 - iuy + \frac{(uy)^2}{2}\right| \leq \frac{|uy|^3}{6}$, one has

$$\left|a_m^M - 1 - \frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) + \frac{u^2\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right)^2}{2M}\right|$$
$$\leq \frac{|u|^3\left|f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right|^3}{6M^{\frac{3}{2}}} \leq \frac{4\|f\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}}.$$

Therefore for $m = 1, \cdots, M$, a_m^M can be rewritten as

$$a_m^M = 1 + \frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) - \frac{u^2\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right)^2}{2M} + r_m^M$$

with $\left|r_m^M\right| \leq \frac{4\|f\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}}$.

By defining for $m = 1, \cdots, M$

$$b_m^M := 1 - \frac{u^2\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right)^2}{2M}, \tag{4.37}$$

let us now study the difference

$$\left|\mathbb{E}\left(\prod_{m=1}^{M} a_m^M \,\middle|\, \mathcal{F}\right) - \mathbb{E}\left(\prod_{m=1}^{M} b_m^M \,\middle|\, \mathcal{F}\right)\right|. \tag{4.38}$$

By rewriting $\prod_{m=1}^{M} a_m^M - \prod_{m=1}^{M} b_m^M = \sum_{m=1}^{M} \prod_{\ell=1}^{m-1} a_\ell^M \left(a_m^M - b_m^M\right) \prod_{\ell=m+1}^{M} b_\ell^M$, where by convention the empty product is equal to one, by using that the random vectors $(Y_m^M)_{1\leq m\leq M}$ are conditionally independent given $\mathcal{F}$ and that $\left|r_m^M\right| \leq \frac{4\|f\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}}$, $\left|a_\ell^M\right| \leq 1$ and $\left|b_\ell^M\right| \leq 1 + \frac{2\|f\|_\infty^2 u^2}{M}$, (4.38) becomes

$$\left|\mathbb{E}\left(\prod_{m=1}^{M} a_m^M \,\middle|\, \mathcal{F}\right) - \mathbb{E}\left(\prod_{m=1}^{M} b_m^M \,\middle|\, \mathcal{F}\right)\right|$$
$$= \left|\sum_{m=1}^{M} \mathbb{E}\left(\prod_{\ell=1}^{m-1} a_\ell^M \left(\frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) + r_m^M\right) \prod_{\ell=m+1}^{M} b_\ell^M \,\middle|\, \mathcal{F}\right)\right|$$
$$= \left|\sum_{m=1}^{M} \mathbb{E}\left(\prod_{\ell=1}^{m-1} a_\ell^M \,\middle|\, \mathcal{F}\right) \mathbb{E}\left(\frac{iu}{\sqrt{M}}\left(f(Y_m^M) - \mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)\right) + r_m^M \,\middle|\, \mathcal{F}\right) \mathbb{E}\left(\prod_{\ell=m+1}^{M} b_\ell^M \,\middle|\, \mathcal{F}\right)\right|$$
$$= \left|\sum_{m=1}^{M} \mathbb{E}\left(\prod_{\ell=1}^{m-1} a_\ell^M \,\middle|\, \mathcal{F}\right) \mathbb{E}\left(r_m^M \,\middle|\, \mathcal{F}\right) \mathbb{E}\left(\prod_{\ell=m+1}^{M} b_\ell^M \,\middle|\, \mathcal{F}\right)\right|$$
$$\leq M \times \frac{4\|f\|_\infty^3 |u|^3}{3M^{\frac{3}{2}}} \times \left(1 + \frac{2\|f\|_\infty^2 u^2}{M}\right)^M \underset{M\to\infty}{\longrightarrow} 0.$$

This implies that

$$\limsup_{M\to\infty}\left|\mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M}a_m^M\,\middle|\,\mathcal{F}\right)\right)-\mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M}b_m^M\,\middle|\,\mathcal{F}\right)\right)\right|=0. \tag{4.39}$$

Second step

With b_m^M defined in (4.37), by the previous step, it is enough to study the limit behaviour as $M\to\infty$ of $\mathbb{E}\left(e^{iuH^M}\mathbb{E}\left(\prod_{m=1}^{M}b_m^M\,\middle|\,\mathcal{F}\right)\right)$. By using the general fact that

$$\prod_{m=1}^{M}(1-c_m)=\sum_{s=0}^{M}(-1)^s\sum_{1\le m_1<\cdots<m_s\le M}\prod_{j=1}^{s}c_{m_j}$$

for any real sequence $(c_m)_{1\le m\le M}$, we have

$$\begin{aligned}\mathbb{E}\left(\prod_{m=1}^{M}b_m^M\,\middle|\,\mathcal{F}\right)&=\sum_{s=0}^{M}\left(\frac{-u^2}{2M}\right)^s\sum_{1\le m_1<\cdots<m_s\le M}\mathbb{E}\left(\prod_{j=1}^{s}\left(f(Y_{m_j}^M)-\mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2\,\middle|\,\mathcal{F}\right)\\&=\sum_{s=0}^{M}\left(\frac{-u^2}{2M}\right)^s\sum_{1\le m_1<\cdots<m_s\le M}\prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M)-\mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2\,\middle|\,\mathcal{F}\right)\end{aligned} \tag{4.40}$$

where for the last equality we use that the random vectors $(Y_m^M)_{1\le m\le M}$ are conditionally independent given $\mathcal{F}$. Let us introduce $B_{m_1,\cdots,m_s}:=\prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M)-\mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2\,\middle|\,\mathcal{F}\right)$. Since $B_{m_1,\cdots,m_s}$ is symmetric in its indexes and satisfies $|B_{m_1,\cdots,m_s}|\le\left(4\|f\|_\infty^2\right)^s$, for each $s\ge1$ we have

$$\left|\frac{1}{M^s}\left(\frac{1}{s!}\sum_{m_1,\cdots,m_s=1}^{M}B_{m_1,\cdots,m_s}-\sum_{1\le m_1<\cdots<m_s\le M}B_{m_1,\cdots,m_s}\right)\right|\le\frac{\left(4\|f\|_\infty^2\right)^s}{M^s}\left(\frac{M^s}{s!}-\binom{M}{s}\right).$$

It is therefore possible to rewrite (4.40) as

$$\begin{aligned}\mathbb{E}\left(\prod_{m=1}^{M}b_m^M\,\middle|\,\mathcal{F}\right)&=\sum_{s=0}^{M}\left(-\frac{u^2}{2M}\right)^s\sum_{1\le m_1<\cdots<m_s\le M}\prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M)-\mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2\,\middle|\,\mathcal{F}\right)\\&=\sum_{s=0}^{M}\left(-\frac{u^2}{2M}\right)^s\frac{1}{s!}\sum_{m_1,\cdots,m_s=1}^{M}\prod_{j=1}^{s}\mathbb{E}\left(\left(f(Y_{m_j}^M)-\mathbb{E}\left(f(Y_{m_j}^M)\mid\mathcal{F}\right)\right)^2\,\middle|\,\mathcal{F}\right)+\bar{r}^M\\&=\sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\left(\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left(\left(f(Y_m^M)-\mathbb{E}\left(f(Y_m^M)\mid\mathcal{F}\right)\right)^2\,\middle|\,\mathcal{F}\right)\right)^s+\bar{r}^M\\&=\sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M}f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right)^s+\bar{r}^M\end{aligned}$$

with $\bar{r}^M$ such that

$$|\bar{r}^M|\le\sum_{s\ge1}\frac{\left(2u^2\|f\|_\infty^2\right)^s}{s!}\left(1-\frac{M(M-1)\cdots(M-s+1)}{M^s}\right)\underset{M\to\infty}{\longrightarrow}0.$$

With $\left|e^{iuH^M}\right| \leq 1$, (4.36) and (4.39), this implies that

$$\left|\mathbb{E}\left(e^{\frac{iu}{\sqrt{M}}\sum_{m=1}^{M}\left(f(Y_m^M)-\eta(fg)\right)}\right) - \mathbb{E}\left(e^{iuH^M}\sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M}f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right)^s\right)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Third step

The purpose of what follows is therefore to study the convergence as M goes to infinity of

$$\mathbb{E}\left(e^{iuH^M}\sum_{s=0}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M}f(Y_m^M)\,\middle|\,\mathcal{F}\right)\right)^s\right). \tag{4.41}$$

By (4.23), (4.41) can be rewritten as

$$\begin{aligned}&\mathbb{E}\left(e^{iuH^M}\right) + \sum_{s=1}^{M}\left(-\frac{u^2}{2}\right)^s\frac{1}{s!}\sum_{k_1,\cdots,k_s=0}^{\left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil}\int_{[0,1]^s}\mathbb{E}\left(e^{iuH^M}F^M_{k_1,\alpha_1}\cdots F^M_{k_s,\alpha_s}\right)d\alpha_1\cdots d\alpha_s\\ &= \mathbb{E}\left(e^{iuH^M}\right) + \sum_{s\geq 1}1_{\{s\leq M\}}z(s,M)\end{aligned} \tag{4.42}$$

where for $s \geq 1$ and $M \geq 1$

$$z(s,M) = \left(-\frac{u^2}{2}\right)^s\sum_{k_1,\cdots,k_s=0}^{\left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil}\int_{[0,1]^s}1_{\{0<\alpha_1<\cdots<\alpha_s<1\}}\mathbb{E}\left(e^{iuH^M}F^M_{k_1,\alpha_1}\cdots F^M_{k_s,\alpha_s}\right)d\alpha_1\cdots d\alpha_s$$

and where we recall that $F^M_{k_i,\alpha_i}$ for $i = 1,\cdots,s$ is given by $f^2(X_{\lceil\alpha_i M\rceil})\beta_0(u^M_{\lceil\alpha_i M\rceil-1}, w^M_{\lceil\alpha_i M\rceil})$ if $k_i = 0$ and by

$$-1_{\{\lceil\alpha_i M\rceil\leq M-k_i\}}f(X_{\lceil\alpha_i M\rceil})f(X_{\lceil\alpha_i M\rceil+k_i})\beta_1(u^M_{\lceil\alpha_i M\rceil-1}, w^M_{\lceil\alpha_i M\rceil}, \sum_{\ell=2}^{k_i}w^M_{\lceil\alpha_i M\rceil+\ell-1}, w^M_{\lceil\alpha_i M\rceil+k_i})$$

if $k_i > 0$. Let us now study the convergence as $M \to \infty$ of (4.42): $\lim_{M\to\infty}\mathbb{E}\left(e^{iuH^M}\right) = e^{-\frac{u^2\sigma_1^2(f)}{2}}$ by (4.14) and to study the convergence of the sum we will apply Lebesgue's theorem.
Given $s \geq 1$, we first compute the pointwise convergence as $M \to \infty$ of $z(s,M)$ and then we show the existence of a function $\tilde{z}$ independent of M that dominates $|z|$ and such that $\sum_{s\geq 1}\tilde{z}(s) < \infty$. Let us start by computing the pointwise convergence. Let $s \geq 1$ and $0 \leq k_1, k_2, \cdots, k_s \leq \left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil$. By applying the reasoning based on Proposition 9 done in the proof of Theorem 35, $\forall\ 0 < \alpha_1 < \cdots < \alpha_s < 1$ we check that

$$e^{iuH^M}F^M_{k_1,\alpha_1}\cdots F^M_{k_s,\alpha_s} \overset{d}{\Longrightarrow} e^{iuH}F_{1,k}\cdots F_{s,k}$$

where, recalling that k denotes the multiindex $(k_1,\cdots,k_s)$ and $K_i := k_1 + \cdots + k_{i-1} + i - 1$ $\forall i = 1,\cdots,s$ with the convention $K_1 := 0$, $F_{i,k}$ for $i = 1,\cdots,s$ is given by

$$F_{i,k} = \begin{cases} f^2(X_{K_i+1})\beta_0\left(U_i, g(X_{K_i+1})\right) & k_i = 0\\ -f(X_{K_i+1})f(X_{K_{i+1}})\beta_1\left(U_i, g(X_{K_i+1}), \sum_{\ell=K_i+2}^{K_{i+1}-1}g(X_\ell), g(X_{K_{i+1}})\right) & k_i > 0\end{cases}$$

and $H \sim \mathcal{N}\left(0, \sigma_1^2(f)\right)$ independent of $(U_j)_{1\le j\le s}$ and $(X_j)_{1\le j\le K_{s+1}}$.
Moreover by the hypothesis of boundedness of f, (4.33) and the fact that $\left|e^{iuH^M}\right| \le 1$,

$$\lim_{M\to\infty} \mathbb{E}\left(e^{iuH^M} F_{k_1,\alpha_1}^M \cdots F_{k_s,\alpha_s}^M\right) = \mathbb{E}\left(e^{iuH} F_{1,k}\cdots F_{s,k}\right).$$

By Lebesgue's theorem and by observing that $F_{1,k},\cdots,F_{s,k}$ are independent, respectively distributed as $F_{k_1},\cdots,F_{k_s}$ and independent of H, $z(s,M)$ converges to

$$\begin{aligned}
&\left(-\frac{u^2}{2}\right)^s \sum_{k_1,\cdots,k_s=0}^{\left\lceil \frac{\bar g}{\underline g}\right\rceil} \int_{[0,1]^s} 1_{\{0<\alpha_1<\cdots<\alpha_s<1\}} \mathbb{E}\left(e^{iuH} F_{1,k}\cdots F_{s,k}\right) d\alpha_1\cdots d\alpha_s \\
&\quad = e^{-\frac{u^2\sigma_1^2(f)}{2}} \frac{1}{s!}\left(-\frac{u^2}{2}\right)^s \sum_{k_1,\cdots,k_s=0}^{\left\lceil \frac{\bar g}{\underline g}\right\rceil} \mathbb{E}\left(F_{k_1}\right)\cdots\mathbb{E}\left(F_{k_s}\right) \\
&\quad = e^{-\frac{u^2\sigma_1^2(f)}{2}} \frac{1}{s!}\left(-\frac{u^2}{2}\right)^s \left(\sum_{k_1=0}^{\left\lceil \frac{\bar g}{\underline g}\right\rceil} \mathbb{E}\left(F_{k_1}\right)\right)^s = e^{-\frac{u^2\sigma_1^2(f)}{2}} \frac{1}{s!}\left(-\frac{u^2}{2}\right)^s \left(\sigma_2^2(f)\right)^s.
\end{aligned}$$

Let us now show that z is dominated by a summable function $\tilde z$. By using that f, β_0 and β_1 are bounded functions, for $s \ge 1$ one has

$$\begin{aligned}
|z(s,M)| &= \left|\left(-\frac{u^2}{2}\right)^s \sum_{k_1,\cdots,k_s=0}^{\left\lceil \frac{\bar g}{\underline g}\right\rceil} \int_{[0,1]^s} 1_{\{0<\alpha_1<\cdots<\alpha_s<1\}} \mathbb{E}\left(e^{iuH} F_{1,k}\cdots F_{s,k}\right) d\alpha_1\cdots d\alpha_s\right| \\
&\le C\frac{1}{s!}\left(\frac{u^2\left(\left\lceil \frac{\bar g}{\underline g}\right\rceil + 1\right)}{2}\right)^s
\end{aligned}$$

for a given finite constant C. It is now sufficient to observe that $\sum_{s\ge1} \frac{1}{s!}\left(\left(\left\lceil \frac{\bar g}{\underline g}\right\rceil + 1\right) \times \frac{u^2}{2}\right)^s = e^{\left(\left\lceil \frac{\bar g}{\underline g}\right\rceil+1\right)\frac{u^2}{2}} < \infty$. In conclusion we have proved that (4.42) converges as $M\to\infty$ to

$$e^{-\frac{u^2\sigma_1^2(f)}{2}} \sum_{s\ge0} \frac{1}{s!}\left(-\frac{u^2}{2}\right)^s \left(\sigma_2^2(f)\right)^s = e^{-\frac{u^2\left(\sigma_1^2(f)+\sigma_2^2(f)\right)}{2}}.$$

□

4.5 Proof of Proposition 8

We recall that for $i = 1,\cdots,M$ we denote $u_i^M := \left\{w_1^M + \cdots + w_i^M\right\}$ and $\mu_i^M := \lfloor w_1^M + \cdots + w_i^M \rfloor + 1$, where by convention $u_0^M = 0$, $\mu_0^M = 1$. The following technical result holds.

Lemma 11. *If $1 \le p < \ell \le M$ then*

1. *$\mu_p^M < \mu_\ell^M$ if and only if $w_{p+1}^M + \cdots + w_\ell^M \ge 1 - u_p^M$*
2. *$\mu_p^M = \mu_\ell^M$ if and only if $w_{p+1}^M + \cdots + w_\ell^M < 1 - u_p^M$*
3. *$\mu_p^M = \mu_\ell^M$ if and only if $u_\ell^M = u_p^M + w_{p+1}^M + \cdots + w_\ell^M$.*

Proof of Lemma 11. Let us observe that 2. follows directly from 1. and 3. is a direct consequence of the definition of the integer and fractional part. Therefore it is enough to prove 1.
If $\mu_p^M < \mu_\ell^M$, then

$$w_1^M + \cdots + w_p^M - u_p^M + 1 = \mu_p^M \leq \mu_\ell^M - 1 \leq w_1^M + \cdots + w_\ell^M$$

that implies

$$1 - u_p^M \leq w_{p+1}^M + \cdots + w_\ell^M.$$

Let us now prove the other implication. If $1 - u_p^M \leq w_{p+1}^M + \cdots + w_\ell^M$, then

$$\mu_p^M = w_1^M + \cdots + w_p^M + 1 - u_p^M \leq w_1^M + \cdots + w_\ell^M < \mu_\ell^M.$$

□

The following lemma provides an explicit expression for the conditional expectation $\mathbb{E}\left(f(Y_m^M) \mid \mathcal{F}\right)$ for $m = 1, \cdots, M$ which appears in (4.20) and so it allows to prove Proposition 8.

Lemma 12. *Given $\varrho \in \mathcal{B}_b\left(\mathbb{R}^d\right)$, for $m = 1, \cdots, M$ we have*

$$\mathbb{E}\left(\varrho(Y_m^M) \mid \mathcal{F}\right) = \sum_{i=1}^{M} \varrho(X_i) q_{m,i}^M 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}}$$

where for $i = 1, \cdots, M$

$$q_{m,i}^M = 1_{\{\mu_{i-1}^M < \mu_i^M\}}(1_{\{\mu_{i-1}^M < m < \mu_i^M\}} + 1_{\{m = \mu_{i-1}^M\}}(1 - u_{i-1}^M) + 1_{\{m = \mu_i^M\}} u_i^M) + 1_{\{\mu_{i-1}^M = m = \mu_i^M\}} w_i^M. \tag{4.43}$$

Proof of Lemma 12. For $m = 1, \cdots, M$, by (1.17), by observing that $m - U_m$ is uniformly distributed on $[m-1, m]$ and by the Freezing Lemma, we have

$$\mathbb{E}\left(\varrho(Y_m^M) \mid \mathcal{F}\right) = \sum_{i=1}^{M} \varrho(X_i) \int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du. \tag{4.44}$$

By observing that for $i = 1, \cdots, M$

$$1 = 1_{\{m < \mu_{i-1}^M\}} + 1_{\{m > \mu_i^M\}} + 1_{\{\mu_{i-1}^M < \mu_i^M\}} \left(1_{\{m = \mu_{i-1}^M\}} + 1_{\{m = \mu_i^M\}} + 1_{\{\mu_{i-1}^M < m < \mu_i^M\}}\right) + 1_{\{m = \mu_{i-1}^M = \mu_i^M\}},$$

let us study the value of the integral in (4.44) according to this partition. If $m < \mu_{i-1}^M$ then $m \leq \mu_{i-1}^M - 1 \leq w_1^M + \cdots + w_{i-1}^M$ and if $m > \mu_i^M$ then $m \geq \mu_i^M + 1 > w_1^M + \cdots + w_i^M + 1$ and so in both cases the integral is zero. Let us now suppose that $\mu_{i-1}^M < \mu_i^M$. If $\mu_{i-1}^M < m < \mu_i^M$, then $m \geq \mu_{i-1}^M + 1 > w_1^M + \cdots + w_{i-1}^M + 1$ and $m \leq w_1^M + \cdots + w_i^M$. Therefore

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = 1.$$

If $m = \mu_{i-1}^M$:

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = \int_{w_1^M + \cdots + w_{i-1}^M}^{\mu_{i-1}^M} du = 1 - u_{i-1}^M.$$

If $m = \mu_i^M$:

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = \int_{\mu_i^M - 1}^{w_1^M + \cdots + w_i^M} du = u_i^M.$$

Finally if $m = \mu_i^M = \mu_{i-1}^M$:

$$\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du = \int_{w_1^M + \cdots + w_{i-1}^M}^{w_1^M + \cdots + w_i^M} du = w_i^M.$$

To sum up we have obtained that for $i = 1, \cdots, M$

$$\begin{aligned}
&\int_{m-1}^{m} 1_{\left\{\sum_{j=1}^{i-1} w_j^M < u \leq \sum_{j=1}^{i} w_j^M\right\}} du \\
&\quad = 1_{\{\mu_{i-1}^M < \mu_i^M\}} (1_{\{\mu_{i-1}^M < m < \mu_i^M\}} + 1_{\{m = \mu_{i-1}^M\}} (1 - u_{i-1}^M) + 1_{\{m = \mu_i^M\}} u_i^M) + 1_{\{\mu_{i-1}^M = m = \mu_i^M\}} w_i^M.
\end{aligned}$$

□

Proof of Proposition 8. Let us rewrite (4.20). By Lemma 12 we have

$$\begin{aligned}
&\frac{1}{M} \sum_{m=1}^{M} \mathbb{E}\left(f^2(Y_m^M)|\mathcal{F}\right) - \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}\left(f(Y_m^M)|\mathcal{F}\right)^2 \\
&\quad = \frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{M} f^2(X_i) q_{m,i}^M 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} - \frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{M} f^2(X_i) (q_{m,i}^M)^2 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} \\
&\qquad - \frac{2}{M} \sum_{m=1}^{M} \sum_{1 \leq i < j \leq M} f(X_i) f(X_j) q_{m,i}^M q_{m,j}^M 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} 1_{\{\mu_{j-1}^M \leq m \leq \mu_j^M\}} \\
&\quad = \frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{M} f^2(X_i) \left(q_{m,i}^M - (q_{m,i}^M)^2\right) 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} - \frac{2}{M} \sum_{m=1}^{M} \sum_{1 \leq i < j \leq M} f(X_i) f(X_j) q_{m,i}^M q_{m,j}^M 1_{\{m = \mu_i^M = \mu_{j-1}^M\}}.
\end{aligned} \tag{4.45}$$

We are now going first to rewrite the first component of the right-hand side and then the second one.

First term:

Let us observe that by Lemma 11 applied to the couple (p, l) equal to $(i-1, i)$, we have $1_{\{\mu_i^M > \mu_{i-1}^M\}} = 1_{\{w_i^M \geq 1 - u_{i-1}^M\}}$ and $1_{\{\mu_i^M = \mu_{i-1}^M\}} = 1_{\{w_i^M < 1 - u_{i-1}^M\}}$ for $i = 1 \cdots, M$. Therefore by Lemma 12 and by using that $\mu_i^M - \mu_{i-1}^M = u_{i-1}^M - u_i^M + w_i^M$, we have

$$\begin{aligned}
&\sum_{m=1}^{M} q_{m,i}^M 1_{\{\mu_{i-1}^M \leq m \leq \mu_i^M\}} \\
&= \sum_{m=1}^{M} (1_{\{w_i^M \geq 1 - u_{i-1}^M\}} (1_{\{\mu_{i-1}^M < m < \mu_i^M\}} + 1_{\{m = \mu_{i-1}^M\}} (1 - u_{i-1}^M) \\
&\qquad + 1_{\{m = \mu_i^M\}} u_i^M) + 1_{\{w_i^M < 1 - u_{i-1}^M\}} 1_{\{\mu_{i-1}^M = m = \mu_i^M\}} w_i^M) \\
&= 1_{\{w_i^M \geq 1 - u_{i-1}^M\}} \left(\mu_i^M - \mu_{i-1}^M - u_{i-1}^M + u_i^M\right) + 1_{\{w_i^M < 1 - u_{i-1}^M\}} w_i^M \\
&= w_i^M 1_{\{w_i^M \geq 1 - u_{i-1}^M\}} + w_i^M 1_{\{w_i^M < 1 - u_{i-1}^M\}} = w_i^M.
\end{aligned}$$

Let us observe that since, by (4.3), $\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\left(f^2(Y_m^M)|\mathcal{F}\right)=\frac{1}{M}\sum_{i=1}^{M}f^2(X_i)w_i^M$, we got the expected result. Similarly, by using that $\mu_i^M-\mu_{i-1}^M=u_{i-1}^M-u_i^M+w_i^M$, we have

$$\begin{aligned}
&\sum_{m=1}^{M}(q_{m,i}^M)^2 1_{\{\mu_{i-1}^M\le m\le\mu_i^M\}}\\
&\quad=1_{\{w_i^M\ge 1-u_{i-1}^M\}}(\mu_i^M-\mu_{i-1}^M-1+(1-u_{i-1}^M)^2+(u_i^M)^2)+1_{\{w_i^M<1-u_{i-1}^M\}}(w_i^M)^2\\
&\quad=(w_i^M-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\{w_i^M\ge 1-u_{i-1}^M\}}+(w_i^M)^2 1_{\{w_i^M<1-u_{i-1}^M\}}\\
&\quad=w_i^M-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M)-(w_i^M(1-w_i^M)-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\{w_i^M<1-u_{i-1}^M\}}.
\end{aligned}$$

Substrating the two quantities and using Lemma 11 to rewrite u_i^M as $w_i^M+u_{i-1}^M$ if $w_i^M<1-u_{i-1}^M$, we have

$$\begin{aligned}
&\sum_{m=1}^{M}q_{m,i}^M 1_{\{\mu_{i-1}^M\le m\le\mu_i^M\}}-\sum_{m=1}^{M}(q_{m,i}^M)^2 1_{\{\mu_{i-1}^M\le m\le\mu_i^M\}}\\
&=u_i^M(1-u_i^M)+u_{i-1}^M(1-u_{i-1}^M)+(w_i^M(1-w_i^M)-u_i^M(1-u_i^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\{w_i^M<1-u_{i-1}^M\}}\\
&=u_i^M(1-u_i^M)+u_{i-1}^M(1-u_{i-1}^M)+(w_i^M(1-w_i^M)-(w_i^M+u_{i-1}^M)(1-w_i^M-u_{i-1}^M)-u_{i-1}^M(1-u_{i-1}^M))1_{\{w_i^M<1-u_{i-1}^M\}}\\
&=u_i^M(1-u_i^M)+u_{i-1}^M(1-u_{i-1}^M)-(2u_{i-1}^M(1-u_{i-1}^M)-2w_i^M u_{i-1}^M)1_{\{w_i^M<1-u_{i-1}^M\}}.
\end{aligned}$$

Therefore by observing that for $i=1\cdots,M$ $u_i^M=\left\{u_{i-1}^M+w_i^M\right\}$, the first term can be rewritten as

$$\frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{M}f^2(X_i)\left(q_{m,i}^M-(q_{m,i}^M)^2\right)1_{\{\mu_{i-1}^M\le m\le\mu_i^M\}}=\frac{1}{M}\sum_{i=1}^{M}f^2(X_i)\beta_0(u_{i-1}^M,w_i^M) \tag{4.46}$$

with β_0 defined in (4.7).

Second term:

By Lemma 12 and by rewriting $1_{\{\mu_{i-1}^M<\mu_i^M\}}=1-1_{\{\omega_i^M<1-u_{i-1}^M\}}$, we can rewrite the right-hand side of (4.45) in the following way

$$\begin{aligned}
&\frac{2}{M}\sum_{m=1}^{M}\sum_{1\le i<j\le M}f(X_i)f(X_j)q_{m,i}^M q_{m,j}^M 1_{\{m=\mu_i^M=\mu_{j-1}^M\}}\\
&=\frac{2}{M}\sum_{1\le i<j\le M}f(X_i)f(X_j)\left(1_{\{\mu_{i-1}^M<\mu_i^M\}}u_i^M+1_{\{\mu_{i-1}^M=\mu_i^M\}}w_i^M\right)\\
&\qquad\qquad\times\left(1_{\{\mu_{j-1}^M<\mu_j^M\}}(1-u_{j-1}^M)+1_{\{\mu_{j-1}^M=\mu_j^M\}}w_j^M\right)1_{\{\mu_i^M=\mu_{j-1}^M\}}\\
&=\frac{2}{M}\sum_{k=1}^{M-1}\sum_{i=1}^{M-k}f(X_i)f(X_{i+k})\left(\rho_{i,k}^M+\tilde{\rho}_{i,k}^M\right)
\end{aligned}$$

where for $k=1,\cdots,M-1$ and $i=1,\cdots,M-k$

$$\begin{aligned}
\rho_{i,k}^M&=u_i^M\left(1_{\{\mu_{i+k-1}^M<\mu_{i+k}^M\}}(1-u_{i+k-1}^M)+1_{\{\mu_{i+k-1}^M=\mu_{i+k}^M\}}w_{i+k}^M\right)1_{\{\mu_i^M=\mu_{i+k-1}^M\}}\\
\tilde{\rho}_{i,k}^M&=(w_i^M-u_i^M)\left(1_{\{\mu_{i+k-1}^M<\mu_{i+k}^M\}}(1-u_{i+k-1}^M)+1_{\{\mu_{i+k-1}^M=\mu_{i+k}^M\}}w_{i+k}^M\right)1_{\{\mu_{i-1}^M=\mu_i^M\}}1_{\{\mu_i^M=\mu_{i+k-1}^M\}}.
\end{aligned}$$

Let us now rewrite $\rho^M_{i,k}$ and $\tilde\rho^M_{i,k}$. To simplify the notation, we will denote $\sum_{\ell=j_1}^{j_2} w^M_\ell$ for $1 \le j_1 \le j_2 \le M$ by $s^{j_2}_{j_1}$. We can apply Lemma 11 to the couples

1. $(p,\ell) = (i, i+k-1)$ so that $1_{\{\mu^M_i = \mu^M_{i+k-1}\}} = 1_{\{s^{i+k-1}_{i+1} < 1-u^M_i\}}$
2. $(p,\ell) = (i-1, i)$ so that $1_{\{\mu^M_{i-1} = \mu^M_i\}} = 1_{\{\omega^M_i < 1-u^M_{i-1}\}}$
3. $(p,\ell) = (i+k-1, i+k)$ so that $1_{\{\mu^M_{i+k-1} = \mu^M_{i+k}\}} = 1_{\{\omega^M_{i+k} < 1-u^M_{i+k-1}\}}$ and $1_{\{\mu^M_{i+k-1} < \mu^M_{i+k}\}} = 1_{\{\omega^M_{i+k} \ge 1-u^M_{i+k-1}\}}$

so that, by observing that when $s^{i+k-1}_{i+1} < 1-u^M_i$ we can rewrite u^M_{i+k-1} as $u^M_i + s^{i+k-1}_{i+1}$, one has

$$\begin{aligned}
\rho^M_{i,k} &= u^M_i \left(1_{\{\omega^M_{i+k} \ge 1-u^M_{i+k-1}\}}(1-u^M_{i+k-1}) + w^M_{i+k} 1_{\{\omega^M_{i+k} < 1-u^M_{i+k-1}\}}\right) 1_{\{s^{i+k-1}_{i+1} < 1-u^M_i\}} \\
&= u^M_i \left(1_{\{\omega^M_{i+k} \ge 1-u^M_i - s^{i+k-1}_{i+1}\}}(1-u^M_i - s^{i+k-1}_{i+1}) + w^M_{i+k} 1_{\{\omega^M_{i+k} < 1-u^M_i - s^{i+k-1}_{i+1}\}}\right) 1_{\left\{s^{i+k_1-1}_{i+1} < 1-u^M_i\right\}} \\
&= u^M_i (1-u^M_i - s^{i+k-1}_{i+1}) 1_{\{s^{i+k-1}_{i+1} < 1-u^M_i \le s^{i+k}_{i+1}\}} + u^M_i w^M_{i+k} 1_{\{s^{i+k}_{i+1} < 1-u^M_i\}} \\
&= u^M_i (1-u^M_i - s^{i+k-1}_{i+1}) 1_{\{s^{i+k-1}_{i+1} < 1-u^M_i\}} - u^M_i (1-u^M_i - s^{i+k}_{i+1}) 1_{\{s^{i+k}_{i+1} < 1-u^M_i\}}
\end{aligned}$$

and, by observing that when $w^M_i < 1-u^M_{i-1}$ and $s^{i+k-1}_{i+1} < 1-u^M_i$ we can rewrite $u^M_i = u^M_{i-1} + w^M_i$ and $u^M_{i+k-1} = u^M_{i-1} + s^{i+k-1}_i$, one has

$$\begin{aligned}
\tilde\rho^M_{i,k} &= \left(w^M_i - u^M_i\right) \left(1_{\{\omega^M_{i+k} \ge 1-u^M_{i+k-1}\}}(1-u^M_{i+k-1}) + w^M_{i+k} 1_{\{\omega^M_{i+k} < 1-u^M_{i+k-1}\}}\right) 1_{\{\omega^M_i < 1-u^M_{i-1}\}} 1_{\{s^{i+k-1}_{i+1} < 1-u^M_i\}} \\
&= -\left(u^M_{i-1} + w^M_i\right) \left(1_{\{s^{i+k}_i \ge 1-u^M_{i-1}\}}(1-u^M_{i-1} - s^{i+k-1}_i) + w^M_{i+k} 1_{\{s^{i+k}_i < 1-u^M_{i-1}\}}\right) 1_{\{s^{i+k-1}_i < 1-u^M_{i-1}\}} \\
&\quad + w^M_i \left(1_{\{s^{i+k}_i \ge 1-u^M_{i-1}\}}(1-u^M_{i-1} - s^{i+k-1}_i) + w^M_{i+k} 1_{\{s^{i+k}_i < 1-u^M_{i-1}\}}\right) 1_{\{s^{i+k-1}_i < 1-u^M_{i-1}\}} \\
&= -u^M_{i-1}(1-u^M_{i-1} - s^{i+k-1}_i) 1_{\{s^{i+k-1}_i < 1-u^M_{i-1} \le s^{i+k}_i\}} - u^M_{i-1} w^M_{i+k} 1_{\{s^{i+k}_i < 1-u^M_{i-1}\}} \\
&= -u^M_{i-1}(1-u^M_{i-1} - s^{i+k-1}_i) 1_{\{s^{i+k-1}_i < 1-u^M_{i-1}\}} + u^M_{i-1}(1-u^M_{i-1} - s^{i+k}_i) 1_{\{s^{i+k}_i < 1-u^M_{i-1}\}}.
\end{aligned}$$

Therefore we have obtained that

$$\begin{aligned}
&\frac{1}{M} \sum_{m=1}^M \sum_{1 \le i < j \le M} f(X_i) f(X_j) q^M_{m,i} q^M_{m,j} 1_{\{m = \mu^M_i = \mu^M_{j-1}\}} \\
&\qquad = \frac{1}{M} \sum_{k=1}^{M-1} \sum_{i=1}^{M-k} f(X_i) f(X_{i+k}) \beta_1 \left(u^M_{i-1}, w^M_i, \sum_{\ell=2}^k w^M_{i+\ell-1}, w^M_{i+k}\right)
\end{aligned}$$

with β_1 defined as in (4.8). Let us now observe that there exists a finite constant C_1 such that for $k = 1, \cdots, M-1$, $i = 1, \cdots, M-k$

$$\left|\beta_1 \left(u^M_{i-1}, w^M_i, \sum_{\ell=2}^k w^M_{i+\ell-1}, w^M_{i+k}\right)\right| \le C_1 1_{\left\{\sum_{\ell=2}^k w^M_{i+\ell-1} < 1\right\}} \le C_1 1_{\left\{k < 1 + \frac{\bar g}{\underline g}\right\}}. \tag{4.47}$$

Therefore

$$\frac{1}{M}\sum_{k=1}^{M-1}\sum_{i=1}^{M-k} f(X_i)f(X_{i+k})\beta_1\left(u_{i-1}^M, w_i^M, \sum_{\ell=2}^{k} w_{i+\ell-1}^M, w_{i+k}^M\right)$$
$$= \frac{1}{M}\sum_{k=1}^{(M-1)\wedge\left\lceil\frac{\bar{g}}{\underline{g}}\right\rceil}\sum_{i=1}^{M-k} f(X_i)f(X_{i+k})\beta_1\left(u_{i-1}^M, w_i^M, \sum_{\ell=2}^{k} w_{i+\ell-1}^M, w_{i+k}^M\right).$$

□

4.6 Asymptotic Variance for the Next Steps

What we have seen so far is the study of the stratified sampling selection step: it is actually part of a more general algorithm where each step consists of a selection part and a mutation part that we are now going to describe more in details (see for instance [47] for a more general version of the algorithm).

For each step $n \geq 0$, we are going to recursively define the selection sequence $\left(Y_n^{M,m}\right)_{1\leq m\leq M}$ and the mutation sequence $\left(X_n^{M,m}\right)_{1\leq m\leq M}$. Let $g_n : \mathbb{R}^d \to (0,\infty), n \geq 0$ be a family of measurable functions such that $0 < \inf_{x\in\mathbb{R}^d} g_n(x) \leq \sup_{x\in\mathbb{R}^d} g_n(x) < \infty$. In what follows we denote $\bar{g}_n := \sup_{x\in\mathbb{R}^d} g_n(x)$ and $\underline{g}_n := \inf_{x\in\mathbb{R}^d} g_n(x)$. Let for $n \geq 0$ and $k \geq 0$

$$\phi_n(k) := \left\lceil \frac{\bar{g}_n}{\underline{g}_n}(1+k)\right\rceil. \tag{4.48}$$

Moreover let $(Z_k)_{k\geq 0}$ be a $\mathbb{R}^d$-valued time-inhomogeneous Markov chain that is for each bounded measurable function $h : \mathbb{R}^d \to \mathbb{R}$, $\mathbb{E}\left(h\left(Z_k\right)|\sigma\left(Z_0,\cdots,Z_{k-1}\right)\right) = \int_{\mathbb{R}^d} h(x)P_k\left(Z_{k-1},dx\right)$ for a transition kernel P_k possibly depending on k. We denote the law of Z_0 by η.

Initialization: we generate a sequence $(X_0^{M,m})_{1\leq m\leq M}$ of i.i.d. $\mathbb{R}^d$-valued random vectors following the law η. By convention $Y_0^{M,m} := X_0^{M,m}$ for $1 \leq m \leq M$.

n → n+1: the transition from $\left(Y_n^{M,m}, X_n^{M,m}\right)_{1\leq m\leq M}$ to $\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)_{1\leq m\leq M}$ for $n \geq 0$ consists of two steps.

Selection: we generate the random vectors $\left(Y_{n+1}^{M,m}\right)_{1\leq m\leq M}$ conditionally independent given $\mathcal{F}^n := \sigma\left(\left(X_i^{M,m}, Y_i^{M,m}\right)_{1\leq m\leq M, 0\leq i\leq n}\right)$ by

$$Y_{n+1}^{M,m} = \sum_{\ell=1}^{M} 1_{\left\{\frac{\sum_{j=1}^{\ell-1} g_n\left(X_n^{M,j}\right)}{\sum_{j=1}^{M} g_n\left(X_n^{M,j}\right)} < m - U_n^m \leq \frac{\sum_{j=1}^{\ell} g_n\left(X_n^{M,j}\right)}{\sum_{j=1}^{M} g_n\left(X_n^{M,j}\right)}\right\}} X_n^{M,\ell} \quad m = 1,\cdots,M \tag{4.49}$$

where $(U_n^m)_{1\leq m\leq M}$ is a sequence of i.i.d. random variables independent of $\mathcal{F}^n$ and distributed according to the uniform law on $(0,1)$.

Mutation: given $\mathcal{G}^{n+1} := \sigma\left(\left(X_i^{M,m}, Y_i^{M,m}\right)_{1\leq m\leq M, 0\leq i\leq n}, \left(Y_{n+1}^{M,m}\right)_{1\leq m\leq M}\right)$ we generate the random vectors $X_{n+1}^{M,m}, 1 \leq m \leq M$ conditionally independent and respectively distributed according to the probability measure $P_{n+1}\left(Y_{n+1}^{M,m}, \cdot\right)$

For each $n \geq 0$ we denote the weights by

$$w_n^{M,m} = \frac{M g_n(X_n^{M,m})}{\sum_{\ell=1}^{M} g_n(X_n^{M,\ell})} \quad m = 1, \cdots, M.$$

Moreover let us observe that the selection property for $n = 0$ (4.3) remains valid for each n: $\forall n \geq 0$

$$\mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} \delta_{Y_{n+1}^{M,m}} | \mathcal{F}^n\right) = \sum_{m=1}^{M} \frac{w_n^{M,m}}{M} \delta_{X_n^{M,m}}. \tag{4.50}$$

The study of the asymptotic variance of the selection part of the first step that is $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f\left(Y_1^{M,m}\right)\right)$ for a given bounded measurable function $f : \mathbb{R}^d \to \mathbb{R}$ was the object of Section 4.3. Our purpose now is to generalize this result by studying the asymptotic variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f\left(Y_{n+1}^{M,m}\right)$ for $n \geq 1$. Before doing it let us begin with some notation: for $n \geq 0$, $M \geq 1$ and $k = 0, \cdots, M-1$ let

$$\widetilde{\eta}_n^{k,M} := \frac{1}{M}\sum_{i=1}^{M-k} \delta_{\left(Y_n^{M,i}, \cdots, Y_n^{M,i+k}\right)}$$

$$\bar{\eta}_n^{k,M} := \frac{1}{M}\sum_{i=1}^{M-k} \delta_{\left(X_n^{M,i}, \cdots, X_n^{M,i+k}\right)}.$$

Remark 37. *Let $f \in \mathcal{B}_b\left(\mathbb{R}^d\right)$. If $k = 0$ and $n = 0$, by the classical Strong Law of Large Numbers we have*

$$\lim_{M\to\infty} \widetilde{\eta}_0^{0,M}(f) = \lim_{M\to\infty} \bar{\eta}_0^{0,M}(f) = \bar{\eta}_0^0(f) := \eta(f) = \mathbb{E}\left(f\left(Z_0\right)\right) \quad a.s.$$

It is possible to prove that a Strong Law of Large Numbers holds (see [47, Corollary 7.4.2]) also in the case $k = 0$ and $n \geq 1$:

$$\lim_{M\to\infty} \widetilde{\eta}_n^{0,M}(f) = \lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} f\left(Y_n^{M,i}\right) = \widetilde{\eta}_n^0(f) := \frac{\mathbb{E}\left(f\left(Z_{n-1}\right)\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)}{\mathbb{E}\left(\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)} \quad a.s.$$

$$\lim_{M\to\infty} \bar{\eta}_n^{0,M}(f) = \lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} f\left(X_n^{M,i}\right) = \bar{\eta}_n^0(f) := \frac{\mathbb{E}\left(f\left(Z_n\right)\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)}{\mathbb{E}\left(\prod_{p=0}^{n-1} g_p\left(Z_p\right)\right)} \quad a.s..$$

We are now going to prove that given $h : \mathbb{R}^{2d} \to \mathbb{R}$ a bounded measurable function, the study of the asymptotic variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)$ depends on the study of the asymptotic variance of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f\left(Y_{n+1}^{M,m}\right)$ for $f : \mathbb{R}^d \to \mathbb{R}$ a bounded measurable function.

Proposition 10. *Let $h \in \mathcal{B}_b\left(\mathbb{R}^{2d}\right)$. The following convergence holds*

$$\left| Var\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)\right) - Var\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right) - \frac{\bar{\eta}_n^0\left(g_n\left(P_{n+1}h^2 - (P_{n+1}h)^2\right)\right)}{\bar{\eta}_n^0(g_n)}\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Proof of Proposition 10. By using that the $X_{n+1}^{M,m}$ are conditionally independent given $\mathcal{G}^{n+1}$ and that the $Y_{n+1}^{M,m}$ are $\mathcal{G}^{n+1}$ measurable, one has

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} \mathbb{E}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right) + \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} \mathrm{Var}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} \mathbb{E}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right) + \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\left(h^2\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)\right)\\
&\quad - \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\left(h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right) \,\middle|\, \mathcal{G}^{n+1}\right)^2\right).
\end{aligned}$$

Since $\mathcal{L}\left(X_{n+1}^{M,m} \mid \mathcal{G}^{n+1}\right) \sim P_{n+1}\left(Y_{n+1}^{M,m}, \cdot\right)$, for each bounded measurable function $\tilde{h} : \mathbb{R}^{2d} \to \mathbb{R}$ one has

$$\mathbb{E}\left(\tilde{h}(Y_{n+1}^{M,m}, X_{n+1}^{M,m}) \mid \mathcal{G}^{n+1}\right) = P_{n+1}\tilde{h}\left(Y_{n+1}^{M,m}\right) := \int_{\mathbb{R}^d} \tilde{h}(Y_{n+1}^{M,m}, x) P_{n+1}\left(Y_{n+1}^{M,m}, dx\right). \tag{4.51}$$

Therefore we obtain that

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} h\left(Y_{n+1}^{M,m}, X_{n+1}^{M,m}\right)\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right) + \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M} P_{n+1}h^2\left(Y_{n+1}^{M,m}\right)\right) - \mathbb{E}\left(\frac{1}{M}\sum_{m=1}^{M}\left(P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right)^2\right)\\
&= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} P_{n+1}h\left(Y_{n+1}^{M,m}\right)\right) + \mathbb{E}\left(\sum_{m=1}^{M} \frac{w_n^{M,m}}{M} P_{n+1}h^2\left(X_n^{M,m}\right)\right) - \mathbb{E}\left(\sum_{m=1}^{M} \frac{w_n^{M,m}}{M}\left(Ph\left(X_n^{M,m}\right)\right)^2\right)
\end{aligned}$$

where to obtain the last equality we use the selection property (4.50). By Remark 37,

$$\begin{aligned}
&\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g_n\left(X_n^{M,m}\right) P_{n+1}h^2\left(X_n^{M,m}\right) = \bar{\eta}_n^0(g_n P_{n+1}h^2) \quad a.s.,\\
&\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g_n\left(X_n^{M,m}\right)\left(P_{n+1}h\left(X_n^{M,m}\right)\right)^2 = \bar{\eta}_n^0(g_n\left(P_{n+1}h\right)^2) \quad a.s.,\\
&\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g_n\left(X_n^{M,m}\right) = \bar{\eta}_n^0(g_n) \quad a.s.
\end{aligned}$$

Thus by Lebesgue theorem,

$$\lim_{M\to\infty} \mathbb{E}\left(\sum_{m=1}^{M} \frac{w_n^{M,m}}{M} P_{n+1}h^2\left(X_n^{M,m}\right)\right) = \frac{\bar{\eta}_n^0(g_n P_{n+1}h^2)}{\bar{\eta}_n^0(g_n)},$$
$$\lim_{M\to\infty} \mathbb{E}\left(\sum_{m=1}^{M} \frac{w_n^{M,m}}{M} \left(P_{n+1}h\left(X_n^{M,m}\right)\right)^2\right) = \frac{\bar{\eta}_n^0(g_n\left(P_{n+1}h\right)^2)}{\bar{\eta}_n^0(g_n)}$$

and this concludes the proof. □

We are now ready to study for $n \geq 1$ and $f \in \mathcal{B}_b\left(\mathbb{R}^d\right)$ the convergence as $M \to \infty$ of

$$V_{n+1}^M(f) := \mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f\left(Y_{n+1}^{M,m}\right)\right).$$

With the aim of studying $\lim_{M\to\infty} V_{n+1}^M(f)$, we will assume the following conjectures:

Conjecture 1. *For each $f \in \mathcal{B}_b\left(\mathbb{R}^d\right)$ one has*

$$\left| Var\left(\sqrt{M}\frac{\bar{\eta}_n^{0,M}(g_n f)}{\bar{\eta}_n^{0,M}(g_n)}\right) - \frac{V_n^M(P_n f_n)}{\left(\bar{\eta}_n^0(g_n)\right)^4} - \frac{\bar{\eta}_{n-1}^0\left(g_{n-1}\left(P_n f_n^2 - (P_n f_n)^2\right)\right)}{\left(\bar{\eta}_n^0(g_n)\right)^4 \bar{\eta}_{n-1}^0(g_{n-1})}\right| \underset{M\to\infty}{\longrightarrow} 0$$

where

$$f_n := g_n\left(\bar{\eta}_n^0(g_n) f - \bar{\eta}_n^0(g_n f)\right).$$

Remark 38. *In Section 4.3 we proved that in the case $n = 0$, by observing that $Var\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}(g_0 f)}{\bar{\eta}_0^{0,M}(g_0)}\right) = Var\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)} f\right)}{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)}\right)}\right)$, one has*

$$\left| Var\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}(g_0 f)}{\bar{\eta}_0^{0,M}(g_0)}\right) - \frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}\right| \underset{M\to\infty}{\longrightarrow} 0.$$

For $n \geq 1$ the idea underlying the Conjecture 1 is the following: by considering

$$Var\left(\sqrt{M}\left(\frac{\bar{\eta}_n^{0,M}(g_n f)}{\bar{\eta}_n^{0,M}(g_n)} - \frac{\bar{\eta}_n^0(g_n f)}{\bar{\eta}_n^0(g_n)}\right)\right) = Var\left(\sqrt{M}\frac{\bar{\eta}_n^{0,M}(f_n)}{\bar{\eta}_n^{0,M}(g_n)\,\bar{\eta}_n^0(g_n)}\right),$$

we asymptotically replace in the denominator of the right-hand side $\bar{\eta}_n^{0,M}(g_n)$ by $\bar{\eta}_n^0(g_n)$. Thus we have

$$\left| Var\left(\sqrt{M}\frac{\bar{\eta}_n^{0,M}(f_n)}{\bar{\eta}_n^{0,M}(g_n)\,\bar{\eta}_n^0(g_n)}\right) - \frac{1}{\left(\bar{\eta}_n^0(g_n)\right)^4} Var\left(\sqrt{M}\bar{\eta}_n^{0,M}(f_n)\right)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

We then rewrite $Var\left(\sqrt{M}\bar{\eta}_n^{0,M}(f_n)\right)$ following the same reasoning used in the proof of Proposition 10.

Conjecture 2. *Let $t \in \mathbb{N}$. Given $\psi \in \mathcal{B}_b\left(\mathbb{R}^{t+2}\right)$ continuous and $h \in \mathcal{B}_b\left(\mathbb{R}^{d(t+1)}\right)$, the following convergence holds*

$$\left| \frac{1}{M} \sum_{m=1}^{M-t} \mathbb{E}\left(h\left(X_n^{M,m}, \cdots, X_n^{M,m+t}\right) \psi\left(u_n^{M,m-1}, w_n^{M,m}, \cdots, w_n^{M,m+t}\right)\right) \right.$$
$$\left. - \frac{1}{M} \sum_{m=1}^{M-t} \mathbb{E}\left(h\left(X_n^{M,m}, \cdots, X_n^{M,m+t}\right) \int_0^1 \psi\left(u, \tilde{g}_n(X_n^{M,m}), \cdots, \tilde{g}_n(X_n^{M,m+t})\right) du\right) \right| \underset{M\to\infty}{\longrightarrow} 0$$

where $u_n^{M,m-1} = \left\{ \frac{M \sum_{i=1}^{m-1} g_n\left(X_n^{M,i}\right)}{\sum_{i=1}^{M} g_n\left(X_n^{M,i}\right)} \right\}$ *and* $\tilde{g}_n(x) = \frac{g_n(x)}{\bar{\eta}_n^0(g_n)}$.

Remark 39. *In Section 4.3 we have provided a formal proof of Conjecture 2 in the case $n = 0$. In that case we strongly used the fact that the random variables we are working with $\left(\left(X_0^{M,m}\right)_{m\geq 1}\right)$ are i.i.d. and the fact that the law of $g_0\left(X_0^{M,m}\right)$ has an absolutely continuous component to prove that is possible to asymptotically replace $u_0^{M,i-1}$ with a uniformly distributed random variable independent of $\mathcal{F}^0$ and $w_0^{M,i}$ with $\frac{g_0\left(X_0^{M,i}\right)}{\bar{\eta}_0^0(g_0)}$.*

Before providing the main result of this section, let us introduce some notation.
For $k \geq 0$ let $P_n^{\otimes(k+1)} : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right)$ be defined by

$$P_n^{\otimes(k+1)} h\left(x_0, \cdots, x_k\right) := \int_{\mathbb{R}^{d(k+1)}} h\left(y_0, \cdots, y_k\right) P_n\left(x_0, dy_0\right) \cdots P_n\left(x_k, dy_k\right)$$

and let for $s_k = 0, \cdots, \phi_{n-1}(k)$, $T_{n-1}^{k\to s_k} : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(s_k+1)}\right)$ given by

$$T_{n-1}^{k\to s_k} f\left(x_0, x_1, \cdots, x_{s_k}\right)$$
$$= \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k} f\left(x_0, x_{s_1}, \cdots, x_{s_k}\right) \int_0^1 \psi_{s_1:s_k}\left(u, \tilde{g}_{n-1}(x_0), \tilde{g}_{n-1}(x_1), \cdots, \tilde{g}_{n-1}(x_{s_k})\right) du$$

and where given $0 \leq s_k \leq \phi_{n-1}(k)$ and $s_0 := 0 \leq s_1 \leq s_2 \leq \cdots \leq s_k$:

$$\psi_{s_1:s_k}\left(u, y_0, y_1, \cdots, y_{s_k}\right) = \sum_{i=1}^{\lceil 1+\bar{g}_n/\underline{g}_n \rceil} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{u+\sum_{j=0}^{s_q-1} y_j < u' \leq u+\sum_{j=0}^{s_q} y_j\right\}} du'.$$

Finally let $\mathcal{T}_{n-1}^k : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(\phi_{n-1}(k)+1)}\right)$ defined by

$$\mathcal{T}_{n-1}^k h\left(x_0, \cdots, x_{\phi_{n-1}(k)}\right) = \sum_{s_k=0}^{\phi_{n-1}(k)} T_{n-1}^{k\to s_k}\left(P_n^{\otimes(k+1)} h\right)\left(x_0, x_1, \cdots, x_{s_k}\right).$$

Theorem 40. *Let us assume Conjecture 1 and Conjecture 2. Then $\forall n \geq 1$ and $\forall f \in \mathcal{B}\left(\mathbb{R}^d\right)$, $V_n^M(f)$ converges as M goes to infinity and we denote its limit by $V_n(f)$. Moreover, by defining $\phi_{m:n}(0) := \phi_m\left(\phi_{m+1}\left(\cdots\left(\phi_n(0)\right)\right)\right)$ for $m \leq n$, $m, n \in \mathbb{N}$, $V_n(f)$ is defined by the following recursive formula:*

$$V_{n+1}(f) = \frac{V_n\left(P_n f_n\right)}{\left(\bar{\eta}_n^0\left(g_n\right)\right)^4} + \frac{\bar{\eta}_{n-1}^0\left(g_{n-1}\left(P_n f_n^2 - \left(P_n f_n\right)^2\right)\right)}{\left(\bar{\eta}_n^0\left(g_n\right)\right)^4 \bar{\eta}_{n-1}^0\left(g_{n-1}\right)}$$
$$+ \int \mathcal{T}_0^{\phi_{1:n}(0)} \cdots \mathcal{T}_{n-2}^{\phi_{n-1:n}(0)} \mathcal{T}_{n-1}^{\phi_{n:n}(0)} \bar{f}_n(x)\, \eta^{\otimes\phi_{0:n}(0)}(dx)$$

where $\bar{f}_n \in \mathcal{B}_b\left(\mathbb{R}^{d(1+\phi_n(0))}\right)$ *is given by*

$$\bar{f}_n\left(x_0,\cdots,x_{\phi_n(0)}\right) := \sum_{k=0}^{\phi_n(0)} f\left(x_0\right) f\left(x_k\right) \int_0^1 \bar{\beta}_k\left(u, \tilde{g}_n\left(x_0\right),\cdots,\tilde{g}_n\left(x_k\right)\right) du$$

with $\bar{\beta}_k\left(u, y_0,\cdots,y_k\right) = \beta_0(u,y_0)1_{\{k=0\}} - \beta_1\left(u, y_0, \sum_{\ell=1}^{k-1} y_\ell, y_k\right) 1_{\{k\neq 0\}}$ *where* β_0 *and* β_1 *are respectively defined in (4.7) and (4.8).*

Remark 41. *Let us observe that by Theorem 34,* $\lim_{M\to\infty} V_1^M(h) = V_1(h) = \sigma_1^2(h) + \sigma_2^2(h)\ \forall h \in \mathcal{B}_b\left(\mathbb{R}^d\right)$.

The proof of Theorem 40 relies on the following proposition the proof of which is provided after the proof of the theorem.

Proposition 11. *Let us assume Conjecture 2. Given* $n \geq 1$ *and* $h \in \mathcal{B}_b\left(\mathbb{R}^{d(\phi_n(0)+1)}\right)$ *one has*

$$\left|\mathbb{E}\left(\bar{\eta}_n^{\phi_n(0),M}(h)\right) - \int \mathcal{T}_0^{\phi_{1:n}(0)}\cdots\mathcal{T}_{n-2}^{\phi_{n-1:n}(0)}\mathcal{T}_{n-1}^{\phi_{n:n}(0)} h\left(x\right)\eta^{\otimes\phi_{0:n}(0)}\left(dx\right)\right| \underset{M\to\infty}{\longrightarrow} 0.$$

Proof of Theorem 40. Similarly to what we have done in (4.5) and (4.6), we can rewrite the variance in the following way

$$\begin{aligned}
V_{n+1}^M(f) &= \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m})\right)\\
&= \mathrm{Var}\left(\mathbb{E}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right)\\
&= \mathrm{Var}\left(\sqrt{M}\frac{\sum_{m=1}^M g_n(X_n^{M,m}) f(X_n^{M,m})}{\sum_{m=1}^M g_n(X_n^{M,m})}\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right) \qquad (4.52)
\end{aligned}$$

where to obtain the last equality we use the selection property (4.50). Conjecture 1 gives the asymptotic behavior of the first term of (4.52).
Let us then study the second term of (4.52). Using the definition of $Y_{n+1}^{M,m}$ and following the same reasoning done to prove Proposition 8, we can rewrite the expression inside the expectation: for $M \geq 1 + \phi_n(0)$

$$\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right) \qquad (4.53)$$

$$= \frac{1}{M}\sum_{k=0}^{\phi_n(0)}\sum_{i=1}^{M-k} f(X_n^{M,i}) f(X_n^{M,i+k}) \bar{\beta}_k\left(u_n^{M,i-1}, w_n^{M,i},\cdots,w_n^{M,i+k}\right) \qquad (4.54)$$

with $\bar{\beta}_k\left(u, y_0,\cdots,y_k\right) := \beta_0(u,y_0)1_{\{k=0\}} - \beta_1\left(u, y_0, \sum_{\ell=1}^{k-1} y_\ell, y_k\right) 1_{\{k\neq 0\}}$ where we recall that β_0 and β_1 are respectively defined in (4.7) and (4.8). Given $k = 0,\cdots,\phi_n(0)$, we can apply Conjecture 2 with $t = k$, $h\left(x_0,\cdots,x_k\right) = f\left(x_0\right) f\left(x_k\right)$ and $\psi\left(u, y_0,\cdots,y_k\right) = \bar{\beta}_k\left(u, y_0,\cdots,y_k\right)$ and obtain that

$$\left|\mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^M f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right) - \sum_{k=0}^{\phi_n(0)} \mathbb{E}\left(\bar{\eta}_n^{k,M}\left(\tilde{f}_k\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0 \qquad (4.55)$$

where for $k = 0, \cdots, \phi_n(0)$, $\tilde{f}_k(x_0, \cdots, x_k) = f(x_0) f(x_k) \int_0^1 \bar{\beta}_k(u, \tilde{g}_n(x_0), \cdots, \tilde{g}_n(x_k)) du$. We now observe that by defining $\bar{f}_n\left(x_0, \cdots, x_{\phi_n(0)}\right) = \sum_{k=0}^{\phi_n(0)} \tilde{f}_k(x_0, \cdots, x_k)$, one has

$$\begin{aligned}
&\left| \sum_{k=0}^{\phi_n(0)} \bar{\eta}_n^{k,M}\left(\tilde{f}_k\right) - \bar{\eta}_n^{\phi_n(0),M}\left(\bar{f}_n\right) \right| \\
&= \frac{1}{M} \left| \sum_{k=0}^{\phi_n(0)} \left(\sum_{i=1}^{M-k} \tilde{f}_k\left(X_n^{M,i}, \cdots, X_n^{M,i+k}\right) - \sum_{i=1}^{M-\phi_n(0)} \tilde{f}_k\left(X_n^{M,i}, \cdots, X_n^{M,i+k}\right) \right) \right| \\
&\leq \frac{C}{M} \sum_{k=0}^{\phi_n(0)} (\phi_n(0) - k) \underset{M\to\infty}{\longrightarrow} 0
\end{aligned}$$

for a finite constant C. Thus we can combine it with (4.55) and obtain that

$$\left| \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n \right) \right) - \mathbb{E}\left(\bar{\eta}_n^{\phi_n(0),M}\left(\bar{f}_n\right) \right) \right| \underset{M\to\infty}{\longrightarrow} 0.$$

We can now apply Proposition 11 with $h = \bar{f}_n$ and obtain that

$$\left| \mathbb{E}\left(\bar{\eta}_n^{\phi_n(0),M}\left(\bar{f}_n\right) \right) - \int \mathcal{T}_0^{\phi_{1:n}(0)} \cdots \mathcal{T}_{n-2}^{\phi_{n-1:n}(0)} \mathcal{T}_{n-1}^{\phi_{n:n}(0)} \bar{f}_n(x) \, \eta^{\otimes \phi_{0:n}(0)}(dx) \right| \underset{M\to\infty}{\longrightarrow} 0.$$

Thus we have obtained that $\forall n \geq 1$, $\forall f \in \mathcal{B}_b\left(\mathbb{R}^d\right)$

$$\begin{aligned}
&\left| V_{n+1}^M(f) - \frac{V_n^M(P_n f_n)}{\left(\bar{\eta}_n^0(g_n)\right)^4} - \frac{\bar{\eta}_{n-1}^0\left(g_{n-1}\left(P_n f_n^2 - (P_n f_n)^2\right)\right)}{\left(\bar{\eta}_n^0(g_n)\right)^4 \bar{\eta}_{n-1}^0(g_{n-1})} \right. \\
&\qquad \left. - \int \mathcal{T}_0^{\phi_{1:n}(0)} \cdots \mathcal{T}_{n-2}^{\phi_{n-1:n}(0)} \mathcal{T}_{n-1}^{\phi_{n:n}(0)} \bar{f}_n(x) \, \eta^{\otimes \phi_{0:n}(0)}(dx) \right| \underset{M\to\infty}{\longrightarrow} 0.
\end{aligned}$$

To conclude the proof, it is now sufficient to observe that by Theorem 34, $\forall h \in \mathcal{B}_b\left(\mathbb{R}^d\right)$ $V_1^M(h)$ converges as M goes to infinity and its limit is given by $V_1(h) = \sigma_1^2(h) + \sigma_2^2(h)$. □

The proof of Proposition 11 is a direct consequence of the following Lemma.

Lemma 13. *Let $n \geq 0$ and $k \geq 0$. Given $h \in \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right)$ one has*

$$\left| \mathbb{E}\left(\bar{\eta}_{n+1}^{k,M}(h) \right) - \mathbb{E}\left(\bar{\eta}_n^{\phi_n(k),M}(\mathcal{T}_n^k h) \right) \right| \underset{M\to\infty}{\longrightarrow} 0.$$

Proof of Lemma 13. Since the $X_{n+1}^{M,m}$ are conditionally independent given $\mathcal{G}^{n+1}$ and $\mathcal{L}\left(X_{n+1}^{M,m} \mid \mathcal{G}^{n+1}\right) \sim P_{n+1}\left(Y_{n+1}^{M,m}, \cdot\right)$, for each bounded measurable function $h : \mathbb{R}^{d(k+1)} \to \mathbb{R}$ one has

$$\mathbb{E}\left(h(X_{n+1}^{M,i}, \cdots, X_{n+1}^{M,i+k}) \mid \mathcal{G}^{n+1} \right) = P_{n+1}^{\otimes(k+1)} h\left(Y_{n+1}^{M,i}, \cdots, Y_{n+1}^{M,i+k}\right) \tag{4.56}$$

$$:= \int_{\mathbb{R}^{d(k+1)}} h(x_0, \cdots, x_k) P_{n+1}\left(Y_{n+1}^{M,i}, dx_0\right) \cdots P_{n+1}\left(Y_{n+1}^{M,i+k}, dx_k\right). \tag{4.57}$$

Therefore for $M \geq k+1$

$$\mathbb{E}\left(\bar{\eta}_{n+1}^{k,M}(h)\right) = \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(h\left(X_{n+1}^{M,i},\cdots,X_{n+1}^{M,i+k}\right)\right) \tag{4.58}$$

$$= \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(P_{n+1}^{\otimes(k+1)}h\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\right) = \mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(P_{n+1}^{\otimes(k+1)}h)\right). \tag{4.59}$$

Thus our purpose now becomes to study the asymptotic behaviour of $\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right)$ for a given $f \in \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right)$. Recalling the definition (4.49) of $Y_{n+1}^{M,m}$, let us observe that if we denote by ℓ_m the random index in $\{1,\cdots,M\}$ such that $1_{\left\{\sum_{j=1}^{\ell_{m-1}} w_n^{M,j} < m - U_n^m \leq \sum_{j=1}^{\ell_m} w_n^{M,j}\right\}} = 1$ so that $Y_{n+1}^{M,m} = X_n^{M,\ell_m}$, one has that $m \mapsto \ell_m$ is non decreasing.
Therefore given $k \geq 0$, $M \geq k+1$ and $1 \leq i \leq M-k$ one has:

$$\begin{aligned}&\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\\ &= \sum_{\ell_0=1}^{M}\cdots\sum_{\ell_k=\ell_{k-1}}^{M}\left(X_n^{M,\ell_0},\cdots,X_n^{M,\ell_k}\right)\prod_{q=0}^{k}1_{\left\{\sum_{j=1}^{\ell_{q-1}} w_n^{M,j} < i+q-U_n^{i+q} \leq \sum_{j=1}^{\ell_q} w_n^{M,j}\right\}}.\end{aligned}$$

We can now apply the following change of variables $m = \ell_0, s_1 = \ell_1 - \ell_0, s_2 = \ell_2 - \ell_0, \cdots, s_k = \ell_k - \ell_0$ and set $s_0 := 0$ so that the above expression becomes

$$\sum_{0\leq s_k\leq M-1}\sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k}\sum_{m=1}^{M-s_k}\left(X_n^{M,m},X_n^{M,m+s_1},\cdots,X_n^{M,m+s_k}\right)\prod_{q=0}^{k}1_{\left\{\sum_{j=1}^{m+s_{q-1}} w_n^{M,j} < i+q-U_n^{i+q} \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}}.$$

Therefore one has

$$\begin{aligned}\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) &= \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(f\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\right) = \frac{1}{M}\sum_{i=1}^{M-k}\mathbb{E}\left(\mathbb{E}\left(f\left(Y_{n+1}^{M,i},\cdots,Y_{n+1}^{M,i+k}\right)\mid\mathcal{F}^n\right)\right)\\ &= \frac{1}{M}\sum_{0\leq s_k\leq M-1}\sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k}\sum_{m=1}^{M-s_k}\mathbb{E}\Bigg(f\left(X_n^{M,m},X_n^{M,m+s_1},\cdots,X_n^{M,m+s_k}\right)\\ &\qquad\times\sum_{i=1}^{M-k}\prod_{q=0}^{k}\int_{i+q-1}^{i+q}1_{\left\{\sum_{j=1}^{m+s_{q-1}} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}}du\Bigg)\end{aligned}$$

where to obtain the last equality we apply the Freezing Lemma and the fact that the sequence $(U_n^m)_{1\leq m\leq M}$ is independent of $\mathcal{F}^n$ and the sequence $\left(X_n^{M,m}\right)_{1\leq m\leq M}$ is $\mathcal{F}^n$ - measurable.
Given $k \geq 0$, $M \geq k+1$, $0 \leq s_k \leq M-1$, $0 \leq s_1 \leq s_2 \leq \cdots \leq s_k$ and $1 \leq m \leq M - s_k$, let us now focus on the sum over i appearing in the above expression:

$$\sum_{i=1}^{M-k}\prod_{q=0}^{k}\int_{i+q-1}^{i+q}1_{\left\{\sum_{j=1}^{m+s_{q-1}} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}}du.$$

Let us first observe that if the quantity inside the sum is different from zero, then in particular $\int_{i-1}^{i}1_{\left\{\sum_{j=1}^{m-1} w_n^{M,j} < u \leq \sum_{j=1}^{m} w_n^{M,j}\right\}}du \neq 0$ and $\int_{i+k-1}^{i+k}1_{\left\{\sum_{j=1}^{m+s_k-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_k} w_n^{M,j}\right\}}du \neq 0$. This implies that $\sum_{j=1}^{m} w_n^{M,j} > i-1$ and $\sum_{j=1}^{m+s_k-1} w_n^{M,j} < i+k$. Thus

$$\frac{\underline{g}_n}{\bar{g}_n}(s_k-1) \leq \sum_{j=m+1}^{m+s_k-1} w_n^{M,j} = \sum_{j=1}^{m+s_k-1} w_n^{M,j} - \sum_{j=1}^{m} w_n^{M,j} < i+k-i+1 = k+1$$

and so we have obtained an upper bound for s_k: $s_k \leq \left\lceil \frac{\bar{g}_n}{\underline{g}_n}(k+1) \right\rceil = \phi_n(k)$ where we use the notation introduced in (4.48).
Moreover we observe that we can replace the finite sum over i with an infinite sum: if $i \geq M-k+1$

$$\int_{i+k-1}^{i+k} 1_{\left\{\sum_{j=1}^{m+s_k-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_k} w_n^{M,j}\right\}} du = 0$$

since $\sum_{j=1}^{m+s_k} w_n^{M,j} \leq \sum_{j=1}^{M} w_n^{M,j} = M$. Therefore after all one has

$$\sum_{i=1}^{M-k} \int_{i-1}^{i} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{\sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}} du$$

$$= 1_{\{s_k \leq \phi_n(k)\}} \sum_{i \geq 1} \prod_{q=1}^{k} \int_{i+q-1}^{i+q} 1_{\left\{\sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}} du.$$

We can now apply the change of variable $u := u' + \left\lfloor \sum_{j=1}^{m-1} w_n^{M,j} \right\rfloor$ in each of the above integrals so to obtain

$$\sum_{i \geq 1} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{\sum_{j=1}^{m+s_q-1} w_n^{M,j} < u \leq \sum_{j=1}^{m+s_q} w_n^{M,j}\right\}} du$$

$$= \sum_{i \geq 1} \prod_{q=0}^{k} \int_{i+q-1-\left\lfloor \sum_{j=1}^{m-1} w_n^{M,j} \right\rfloor}^{i+q-\left\lfloor \sum_{j=1}^{m-1} w_n^{M,j} \right\rfloor} 1_{\left\{u_n^{M,m-1} + \sum_{j=m}^{m+s_q-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m+s_q} w_n^{M,j}\right\}} du'$$

$$= \sum_{i \geq 1} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{u_n^{M,m-1} + \sum_{j=m}^{m+s_q-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m+s_q} w_n^{M,j}\right\}} du' \tag{4.60}$$

where to obtain the last equality we use the fact that if $i < 1$,

$$\int_{i-1}^{i} 1_{\left\{u_n^{M,m-1} + \sum_{j=m}^{m-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m} w_n^{M,j}\right\}} du' = 0.$$

We observe that if the quantity inside the sum is different from zero, then in particular

$$\int_{i-1}^{i} 1_{\{u_n^{M,m-1} < u' \leq u_n^{M,m-1} + w_n^{M,m}\}} du' \neq 0.$$

This implies that $i-1 < u_n^{M,m-1} + w_n^{M,m} \leq 1 + \frac{\bar{g}_n}{\underline{g}_n}$ and so $i \leq \left\lceil 1 + \frac{\bar{g}_n}{\underline{g}_n} \right\rceil$. In conclusion, (4.60) can be rewritten as

$$\sum_{i=1}^{\lceil 1+\bar{g}_n/\underline{g}_n \rceil} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{u_n^{M,m-1} + \sum_{j=m}^{m+s_q-1} w_n^{M,j} < u' \leq u_n^{M,m-1} + \sum_{j=m}^{m+s_q} w_n^{M,j}\right\}} du'. \tag{4.61}$$

We denote (4.61) by $\psi_{s_1:s_k}\left(u_n^{M,m-1}, w_n^{M,m}, w_n^{M,m+1}, \cdots, w_n^{M,m+s_k}\right)$ where $\psi_{s_1:s_k} : \mathbb{R}^{s_k+2} \to \mathbb{R}$. Therefore if $M \geq \max(1+k, 1+\phi_n(k))$ we have obtained that

$$\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) \tag{4.62}$$

$$= \sum_{s_k=0}^{\phi_n(k)} \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k} \frac{1}{M} \sum_{m=1}^{M-s_k} \mathbb{E}\Bigg(f\left(X_n^{M,m}, X_n^{M,m+s_1}, \cdots, X_n^{M,m+s_k}\right) \tag{4.63}$$

$$\times \psi_{s_1:s_k}\left(u_n^{M,m-1}, w_n^{M,m}, w_n^{M,m+1}, \cdots, w_n^{M,m+s_k}\right)\Bigg). \tag{4.64}$$

We can therefore apply Conjecture 2 with $t = s_k$, $h(x_0, x_1, \cdots, x_{s_k}) = f(x_0, x_{s_1}, x_{s_2}, \cdots, x_{s_k})$, $\psi(u, y_0, \cdots, y_{s_k}) = \psi_{s_1:s_k}(u, y_0, \cdots, y_{s_k})$.

$$\left|\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) - \sum_{s_k=0}^{\phi_n(k)} \mathbb{E}\left(\bar{\eta}_n^{s_k,M}(T_n^{k\to s_k} f)\right)\right| \underset{M\to\infty}{\longrightarrow} 0 \tag{4.65}$$

where $T_n^{k\to s_k} : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(s_k+1)}\right)$ is defined by

$$T_n^{k\to s_k} f(x_0, x_1, \cdots, x_{s_k}) = \sum_{0\leq s_1\leq s_2\leq\cdots\leq s_k} f(x_0, x_{s_1}, \cdots, x_{s_k}) \int_0^1 \psi_{s_1:s_k}(u, \tilde{g}_n(x_0), \tilde{g}_n(x_1), \cdots, \tilde{g}_n(x_{s_k}))\, du$$

and where given $0 \leq s_k \leq \phi_n(k)$ and $0 \leq s_1 \leq s_2 \leq \cdots \leq s_k$:

$$\psi_{s_1:s_k}(u, y_0, y_1, \cdots, y_{s_k}) = \sum_{i=1}^{\lceil 1+\bar{g}_n/\underline{g}_n\rceil} \prod_{q=0}^{k} \int_{i+q-1}^{i+q} 1_{\left\{u+\sum_{j=0}^{s_q-1} y_j < u' \leq u+\sum_{j=0}^{s_q} y_j\right\}} du'.$$

We now observe that by defining $T_n^k : \mathcal{B}_b\left(\mathbb{R}^{d(k+1)}\right) \to \mathcal{B}_b\left(\mathbb{R}^{d(\phi_n(k)+1)}\right)$ by

$$T_n^k f\left(x_0, \cdots, x_{\phi_n(k)}\right) = \sum_{s_k=0}^{\phi_n(k)} T_n^{k\to s_k} f(x_0, \cdots, x_{s_k}),$$

one has

$$\left|\sum_{s_k=0}^{\phi_n(k)} \bar{\eta}_n^{s_k,M}(T_n^{k\to s_k} f) - \bar{\eta}_n^{\phi_n(k),M}(T_n^k f)\right|$$

$$= \frac{1}{M}\left|\sum_{s_k=0}^{\phi_n(k)} \left(\sum_{i=1}^{M-s_k} \left(T_n^{k\to s_k} f\right)\left(X_n^{M,i}, \cdots, X_n^{M,i+s_k}\right) - \sum_{i=1}^{M-\phi_n(k)} \left(T_n^{k\to s_k} f\right)\left(X_n^{M,i}, \cdots, X_n^{M,i+s_k}\right)\right)\right|$$

$$\leq \frac{1}{M} \sum_{s_k=0}^{\phi_n(k)} \left\|T_n^{k\to s_k} f\right\|_\infty (\phi_n(k) - s_k) \underset{M\to\infty}{\longrightarrow} 0.$$

Combining the above estimate with (4.65) we obtain that

$$\left|\mathbb{E}\left(\tilde{\eta}_{n+1}^{k,M}(f)\right) - \mathbb{E}\left(\bar{\eta}_n^{\phi_n(k),M}(T_n^k f)\right)\right| \underset{M\to\infty}{\longrightarrow} 0. \tag{4.66}$$

In conclusion, combining (4.59) with (4.66) by taking $f = P_{n+1}^{\otimes(k+1)} h$, we obtain that

$$\left|\mathbb{E}\left(\bar{\eta}_{n+1}^{k,M}(h)\right) - \mathbb{E}\left(\bar{\eta}_n^{\phi_n(k),M}(\mathcal{T}_n^k h)\right)\right| \underset{M\to\infty}{\longrightarrow} 0$$

where $\mathcal{T}_n^k = T_n^k P_{n+1}^{\otimes(k+1)}$ and this concludes the proof. □

4.7 Numerical Results

In what follows we fix $d = 1$ and given a sequence $(W_n)_{n\geq 1}$ of independent real-valued random variables distributed according to the uniform law on $(0,1)$, let $Z_{n+1} = Z_n + W_{n+1}$ for $n \geq 0$ with Z_0 distributed according to the uniform law on $(0,1)$ $\left(\eta(dx) = 1_{[0,1]}(x)\, dx\right)$. Thus in this case the transition kernel is given by $P(x, dy) = 1_{[x,x+1]}(y)dy$. Moreover we will fix $g_n(x) = f(x) = e^x$ $\forall n \geq 0$, $\forall x \in \mathbb{R}$.

In this section, we are first going to numerically verify the Conjecture 1 in the case $n = 1$ and the Conjecture 2 in the case $n = 1$ and $n = 2$. In the second place, we will test the two conjectures together by directly studying the asymptotic variance. As done in the theory, we study numerically the asymptotic behaviour of following expression:

$$V_{n+1}^M(f) = \mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_n^{0,M}(g_n f)}{\bar{\eta}_n^{0,M}(g_n)}\right) + \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_{n+1}^{M,m}) \,\middle|\, \mathcal{F}^n\right)\right). \tag{4.67}$$

We will consider separately the case $n = 0$ and the case $n = 1$. We recall that the first case has been fully studied in Section 4.3 without the need to introduce any conjecture.

4.7.1 Verification of the Conjectures

Conjecture 1

We recall the notation

$$f_1 := g_1\left(\bar{\eta}_1^0(g_1)\, f - \bar{\eta}_1^0(g_1 f)\right).$$

Since we have fixed $g_n(x) = f(x) = e^x$ $\forall n \geq 0$, $\forall x \in \mathbb{R}$, one has $\bar{\eta}_0^0(g_0) = e - 1$, $\bar{\eta}_1^0(g_1) = \frac{e^2-1}{2}$, $\bar{\eta}_1^0(g_1 f) = \frac{1}{6}(e^3 - 1)(e + 1)$. Moreover $Pf_1(x) = \bar{\eta}_1^0(g_1)\frac{e^{2(x+1)}-e^{2x}}{2} - \bar{\eta}_1^0(fg_1)\left(e^{x+1} - e^x\right)$ and

$$Pf_1^2(x) = \left(\bar{\eta}_1^0(g_1)\right)^2 \frac{e^{4(x+1)} - e^{4x}}{4} + \left(\bar{\eta}_1^0(fg_1)\right)^2 \frac{e^{2(x+1)} - e^{2x}}{2} - 2\bar{\eta}_1^0(g_1)\,\bar{\eta}_1^0(fg_1)\,\frac{e^{3(x+1)} - e^{3x}}{3}$$

and

$$\begin{aligned}\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right) = {} & \left(\bar{\eta}_1^0(g_1)\right)^2\left(\frac{e^5-1}{5}\right)\left(\frac{e^2-1}{2}\right) \\ & + \left(\bar{\eta}_1^0(g_1 f)\right)^2\left(\frac{e^3-1}{3}\right)\left(\frac{e^2-1}{2} - (e-1)^2\right) \\ & - \bar{\eta}_1^0(g_1)\,\bar{\eta}_1^0(g_1 f)\left(\frac{e^4-1}{4}\right)\left(\frac{2}{3}\left(e^3-1\right) - \left(e^2-1\right)(e-1)\right).\end{aligned}$$

We first observe that given a sequence $(T_i)_{i\geq 1}$ of square integrable i.i.d. random variables, by using the delta method it is possible to prove that as M goes to infinity the following convergence in distribution holds:

$$\sqrt{M}\left(\left(\frac{1}{M}\sum_{i=1}^{M} T_i^2 - \left(\frac{1}{M}\sum_{i=1}^{M} T_i\right)^2\right) - \mathrm{Var}(T_1)\right) \xRightarrow{d} \mathcal{N}\left(0, \mathrm{Var}\left((T_1 - \mathbb{E}(T_1))^2\right)\right). \tag{4.68}$$

The general strategy will be the following: we fix $M = 10000$ and

1. we simulate $n_1 = 10^7$ independent samples T_j of $\sqrt{M}\frac{\bar{\eta}_1^{0,M}(g_1 f)}{\bar{\eta}_1^{0,M}(g_1)} = \sqrt{M}\frac{\sum_{i=1}^{M}(g_1 f)(X_1^{M,i})}{\sum_{i=1}^{M} g_1(X_1^{M,i})}$ and we compute, by using (4.68), the estimator

$$\hat{v}_1 := \frac{1}{n_1}\sum_{i=1}^{n_1} T_i^2 - \left(\frac{1}{n_1}\sum_{i=1}^{n_1} T_i\right)^2 \tag{4.69}$$

of $\mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_1^{0,M}(g_1 f)}{\bar{\eta}_1^{0,M}(g_1)}\right)$ with relative 95% confidence interval $[a_1, b_1]$ given by

$$a_1 := \hat{v}_1 - \frac{1.96}{\sqrt{n_1}}\sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^4 - \left(\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^2\right)^2} \tag{4.70}$$

$$b_1 := \hat{v}_1 + \frac{1.96}{\sqrt{n_1}}\sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^4 - \left(\frac{1}{n_1}\sum_{i=1}^{n_1}\left(T_i - \frac{1}{n_1}\sum_{j=1}^{n_1} T_j\right)^2\right)^2}. \tag{4.71}$$

2. we simulate $n_1 = 10^7$ independent samples H_j of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})$ and we compute the estimator $\hat{v}_2$ of $V_1^M(Pf_1) = \mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})\right)$ with relative 95% confidence interval $[a_2, b_2]$ with $\hat{v}_2$, a_2, and b_2 respectively defined as in (4.69), (4.70) and (4.71).
Therefore

$$\frac{\hat{v}_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$$

is an estimator of $\frac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ with relative 95% confidence interval

$$\left[\frac{a_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}, \frac{b_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}\right].$$

3. we check that $\hat{v}_1$ is close to $\frac{\hat{v}_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ as expected

n=1	value	CI (95%)
$\hat{v}_1$	2.8021446	[2.7775619 , 2.8267273]
$\frac{\hat{v}_2}{\left(\bar{\eta}_1^0(g_1)\right)^4} + \frac{\bar{\eta}_0^0\left(g_0\left(P_1 f_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$	2.7932862	[2.7831461 , 2.8034263]

Conjecture 2

Let $n \in \{1, 2\}$ and $t \in \{1, 2\}$. Moreover let $h(x_0, \cdots, x_t) = x_0 + \cdots + x_t$ and $\psi(u_0, w_0, \cdots, w_{t+1}) = u_0 + w_0 + \cdots + w_{t+1}$. The general strategy will be the following: we fix $M = 10000$ and

1. we simulate $n_1 = 10^5$ independent samples T_j of

$$\frac{1}{M}\sum_{m=1}^{M-t} h\left(X_n^{M,m},\cdots,X_n^{M,m+t}\right)\psi\left(u_n^{M,m-1},w_n^{M,m},\cdots,w_n^{M,m+t}\right)$$

and we compute the standard estimator

$$\hat{v}_1 := \frac{1}{n_1}\sum_{j=1}^{n_1} T_j \tag{4.72}$$

of $\frac{1}{M}\sum_{m=1}^{M-t}\mathbb{E}\left(h\left(X_n^{M,m},\cdots,X_n^{M,m+t}\right)\psi\left(u_n^{M,m-1},w_n^{M,m},\cdots,w_n^{M,m+t}\right)\right)$ with relative 95% confidence interval $[a_1,b_1]$ given by

$$a_1 := \hat{v}_1 - \frac{1.96}{\sqrt{n_1}} * \sqrt{\frac{1}{n_1}\sum_{j=1}^{n_1}(T_j-\hat{v}_1)^2} \tag{4.73}$$

$$b_1 := \hat{v}_1 + \frac{1.96}{\sqrt{n_1}} * \sqrt{\frac{1}{n_1}\sum_{j=1}^{n_1}(T_j-\hat{v}_1)^2} \tag{4.74}$$

2. we simulate $n_1 = 10^5$ independent samples H_j of

$$\frac{1}{M}\sum_{m=1}^{M-t} h\left(X_n^{M,m},\cdots,X_n^{M,m+t}\right)\psi\left(U,\tilde{g}_n(X_n^{M,m}),\cdots,\tilde{g}_n(X_n^{M,m+t})\right)$$

where U is a random variable uniformly distributed on $(0,1)$ independent of $\mathcal{F}^n$ and $\tilde{g}_n = \frac{g_n}{\eta_n^0(g_n)}\left(\eta_1^0(g_1)=\frac{e^2-1}{2},\eta_2^0(g_2)=\frac{e^3-1}{3}\right)$. We then compute the standard estimator $\hat{v}_2$ of

$$\frac{1}{M}\sum_{m=1}^{M-t}\mathbb{E}\left(h\left(X_n^{M,m},\cdots,X_n^{M,m+t}\right)\psi\left(U,\tilde{g}_n(X_n^{M,m}),\cdots,\tilde{g}_n(X_n^{M,m+t})\right)\right)$$

with relative 95% confidence interval $[a_2,b_2]$ with $\hat{v}_2,a_2,b_2$ respectively defined as in (4.72), (4.73) and (4.74).

3. we check that $\hat{v}_1$ is close to $\hat{v}_2$ as expected

4.7.2 Focus on the Variance

case $n=0$

We recall the notation $\bar{\eta}_0^0(h) := \eta(h)\ \forall h\in\mathcal{B}_b(\mathbb{R})$ and $f_0 := g_0\left(\bar{\eta}_0^0(g_0)f - \bar{\eta}_0^0(g_0 f)\right)$.

By (4.15) and by observing that $\mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}(g_0 f)}{\bar{\eta}_0^{0,M}(g_0)}\right) = \mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)}f\right)}{\bar{\eta}_0^{0,M}\left(\frac{g_0}{\eta(g_0)}\right)}\right)$, one has

$$\left|\mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_0^{0,M}(g_0 f)}{\bar{\eta}_0^{0,M}(g_0)}\right) - \frac{\eta(f_0^2)}{\eta(g_0)^4}\right| \underset{M\to\infty}{\longrightarrow} 0$$

n=1,t=1	value	CI (95%)
$\hat{v}_1$	5.7510732	[5.7509322, 5.7512143]
$\hat{v}_2$	5.7509738	[5.7470892, 5.7548583]
n=1,t=2		
$\hat{v}_1$	11.8853516	[11.8850608, 11.8856425]
$\hat{v}_2$	11.8835799	[11.8777417, 11.8894182]
n=2,t=1		
$\hat{v}_1$	9.2154319	[9.2152408, 9.215623]
$\hat{v}_2$	9.2150201	[9.2087834, 9.2212568]
n=2,t=2		
$\hat{v}_1$	19.0901421	[19.0897513 ,19.0905329]
$\hat{v}_2$	19.0895089	[19.0801501, 19.0988676]

where with the choices made $\bar{\eta}_0^0(g_0) = e-1$, $\bar{\eta}_0^0(g_0 f) = \frac{e^2-1}{2}$, $\eta\left(f_0^2\right) = \left(\bar{\eta}_0^0(g_0)\right)^2 \frac{e^4-1}{4} + \left(\bar{\eta}_0^0(g_0 f)\right)^2 \frac{e^2-1}{2} - 2\bar{\eta}_0^0(g_0)\,\bar{\eta}_0^0(g_0 f)\,\frac{e^3-1}{3}$. The application of Theorem 35 with $g = \tilde{g}_0 = \frac{g_0}{\bar{\eta}_0^0(g_0)}$ gives

$$\left| \mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}} \sum_{m=1}^{M} f(Y_1^{M,m}) \,\middle|\, \mathcal{F}^0 \right) \right) - \sum_{k=0}^{\phi_0(0)} \mathbb{E}(F_k) \right| \underset{M\to\infty}{\longrightarrow} 0$$

with $(F_k)_{k\in\mathbb{N}}$ given by $f^2(X_0^{M,1})\beta_0\left(U_1, \tilde{g}_0(X_0^{M,1})\right)$ if $k = 0$ and by

$$-f(X_0^{M,1}) f(X_0^{M,k+1}) \beta_1\left(U_1, \tilde{g}_0(X_0^{M,1}), \sum_{\ell=2}^{k} \tilde{g}_0(X_0^{M,\ell}), \tilde{g}_0(X_0^{M,k+1})\right)$$

if $k > 0$ with $U_1 \sim \mathcal{U}(0,1)$ independent of $X_0^{M,1}, \cdots, X_0^{M,k+1}$ and β_0 and β_1 respectively defined in (4.7) and (4.8). By using that when U_1 is uniformly distributed on $(0,1)$, $\{U_1 + r\}$ is uniformly distributed on $(0,1)$ for each $r \geq 0$, we can apply the Freezing Lemma to rewrite $\mathbb{E}(F_k)$ $\forall k = 0, \cdots, \phi_0(0)$:

$$\sum_{k=0}^{\phi_0(0)} \mathbb{E}(F_k) = \sum_{k=0}^{\phi_0(0)} \mathbb{E}\left(\phi_{k,\tilde{g}_0}\left(X_0^{M,1}, \cdots, X_0^{M,k+1}\right)\right) \tag{4.75}$$

for measurable functions $\phi_{k,\tilde{g}_0} : \mathbb{R}^{(k+1)} \to \mathbb{R}$ given by

$$\phi_{0,\tilde{g}_0}(x) = \frac{1}{3}\left(f^2(x)\left(1 - 1_{\{\tilde{g}_0(x)<1\}}\left(1 - \tilde{g}_0(x)\right)^3\right)\right) \tag{4.76}$$

and for each $k \geq 1$,

$$\phi_{k,\tilde{g}_0}(x) = f(x_1) f(x_{k+1}) \psi_{k,\tilde{g}_0}(x) \tag{4.77}$$

where $\psi_{k,\tilde{g}_0}(x)$ is given by

$$\begin{aligned}
&\frac{1}{2}\left(1_{\left\{\sum_{i=2}^{k}\tilde{g}_0(x_i)<1\right\}}1_{\left\{\sum_{i=1}^{k+1}\tilde{g}_0(x_i)<1\right\}}\tilde{g}_0(x_1)\tilde{g}_0(x_{k+1})\left(2-2\sum_{i=2}^{k}\tilde{g}_0(x_i)-(\tilde{g}_0(x_1)+\tilde{g}_0(x_{k+1}))\right)\right)\\
&+\frac{1}{6}\left(1_{\left\{\sum_{i=2}^{k}\tilde{g}_0(x_i)<1\right\}}1_{\left\{\sum_{i=1}^{k+1}\tilde{g}_0(x_i)\geq 1\right\}}\left(1-\sum_{i=2}^{k}\tilde{g}_0(x_i)\right)^3\right)\\
&-\frac{1}{6}\left(1_{\left\{\sum_{i=1}^{k}\tilde{g}_0(x_i)<1\right\}}1_{\left\{\sum_{i=1}^{k+1}\tilde{g}_0(x_i)\geq 1\right\}}\left(1-\sum_{i=1}^{k}\tilde{g}_0(x_i)\right)^3\right)\\
&-\frac{1}{6}\left(1_{\left\{\sum_{i=2}^{k+1}\tilde{g}_0(x_i)<1\right\}}1_{\left\{\sum_{i=1}^{k+1}\tilde{g}_0(x_i)\geq 1\right\}}\left(1-\sum_{i=2}^{k+1}\tilde{g}_0(x_i)\right)^3\right).
\end{aligned}$$

Thus, using numerical methods we are going to check what we already know theoretically that is

$$\left|V_1^M(f)-\frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}-\sum_{k=0}^{\phi_0(0)}\mathbb{E}\left(\phi_{k,\tilde{g}_0}\left(X_0^{M,1},\cdots,X_0^{M,k+1}\right)\right)\right|\underset{M\to\infty}{\longrightarrow}0.$$

The general strategy will be the following: we fix $M=10000$ and

1. we simulate $n_1=10^7$ independent samples T_j of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M}f(Y_1^{M,m})$ and we compute the estimator $\hat{v}_1$ of $V_1^M(f)=\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M}f(Y_1^{M,m})\right)$ with relative 95% confidence interval $[a_1,b_1]$ with $\hat{v}_1,a_1,b_1$ respectively defined as in (4.69), (4.70) and (4.71). Therefore $\hat{v}_1-\frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$ is an estimator of $V_1^M(f)-\frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$ with relative 95% confidence interval $\left[a_1-\frac{\eta\left(f_0^2\right)}{\eta(g_0)^4},b_1-\frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}\right]$.
2. We simulate $n_2=10^5$ independent samples Z_j of $\sum_{k=0}^{\phi_0(0)}\phi_{k,\tilde{g}_0}\left(X_0^{M,1},\cdots,X_0^{M,k+1}\right)$ and we compute the standard estimator $\hat{v}_2$ of $\sum_{k=0}^{\phi_0(0)}\mathbb{E}\left(\phi_{k,\tilde{g}_0}\left(X_0^{M,1},\cdots,X_0^{M,k+1}\right)\right)$ with relative 95% confidence interval $[a_2,b_2]$ with $\hat{v}_2,a_2,b_2$ respectively defined as in (4.72), (4.73) and (4.74).
3. We check that $\hat{v}_1-\frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$ is close to $\hat{v}_2$ as expected.

n=0	value	CI (95%)
$\hat{v}_1-\frac{\eta\left(f_0^2\right)}{\eta(g_0)^4}$	0.07943	[0.079127 , 0.079733]
$\hat{v}_2$	0.0793412	[0.0790773 , 0.0796051]

case $n=1$

In this case the asymptotic behaviour of the first term of the right-hand side of (4.67) is given by Conjecture 1

$$\left|\mathrm{Var}\left(\sqrt{M}\frac{\bar{\eta}_1^{0,M}(g_1 f)}{\bar{\eta}_1^{0,M}(g_1)}\right)-\frac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4}-\frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2-(Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4\bar{\eta}_0^0(g_0)}\right|\underset{M\to\infty}{\longrightarrow}0$$

and by (4.55) the asymptotic behaviour of the second term of the right-hand side of (4.67) is given by

$$\left|\mathbb{E}\left(\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_2^{M,m}) \;\middle|\; \mathcal{F}^1\right)\right) - \sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\tilde{f}_k\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0 \tag{4.78}$$

where for $k = 0, \cdots, \phi_1(0)$,

$$\tilde{f}_k(x_0, \cdots, x_k) = f(x_0) f(x_k) \int_0^1 \bar{\beta}_k(u, \tilde{g}_1(x_0), \cdots, \tilde{g}_1(x_k))\, du$$

with $\bar{\beta}_k(u, y_0, \cdots, y_k) := \beta_0(u, y_0) 1_{\{k=0\}} - \beta_1\left(u, y_0, \sum_{\ell=1}^{k-1} y_\ell, y_k\right) 1_{\{k\neq 0\}}$. We recall that β_0 and β_1 are respectively defined in (4.7) and (4.8) and $\tilde{g}_1 = \frac{g_1}{\bar{\eta}_1^0(g_1)}$. We now observe that

$$\sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\tilde{f}_k\right)\right) = \sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\phi_{k,\tilde{g}_1}\right)\right) \tag{4.79}$$

where the $\phi_{k,\tilde{g}_1} : \mathbb{R}^{(k+1)} \to \mathbb{R}$ are measurable functions defined in (4.76) and (4.77).

Thus using numerical methods we are going to check that

$$\left|V_2^M(f) - \frac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)} - \sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}\left(\phi_{k,\tilde{g}_1}\right)\right)\right| \underset{M\to\infty}{\longrightarrow} 0$$

so that Conjecture 1 and Conjecture 2 are numerically verified at the same time. The general strategy will be the following:

1. (a) we simulate $n_1 = 10^7$ independent samples T_j of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_2^{M,m})$ and we compute the estimator $\hat{v}_{1,1}$ of $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} f(Y_2^{M,m})\right)$ with relative 95% confidence interval $[a_{1,1}, b_{1,1}]$ with $\hat{v}_{1,1}, a_{1,1}$, and $b_{1,1}$ respectively defined as in (4.69), (4.70) and (4.71).

   (b) we simulate $n_1 = 10^7$ independent samples H_j of $\frac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})$ independent of $(T_j)_{1\le j\le n_1}$. We compute the estimator $\hat{v}_{1,2}$ of $\mathrm{Var}\left(\frac{1}{\sqrt{M}}\sum_{m=1}^{M} Pf_1(Y_1^{M,m})\right)$ with relative 95% confidence interval $[a_{1,2}, b_{1,2}]$ with $\hat{v}_{1,2}$, $a_{1,2}$, and $b_{1,2}$ respectively defined as in (4.69), (4.70) and (4.71).

   (c) we compute the estimator $\hat{v}_{1,1} - \frac{\hat{v}_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2-(Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ of

$$V_2^M(f) - \frac{V_1^M(Pf_1)}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$$

   with 90% confidence interval $[A, B]$ with $A = a_{1,1} - \frac{b_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2-(Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ and $B = b_{1,1} - \frac{a_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2-(Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$.

2. We simulate $n_2 = 10^5$ independent samples Z_j of $\sum_{k=0}^{\phi_1(0)} \bar{\eta}_1^{k,M}(\phi_{k,\tilde{g}_1})$ and we compute the standard estimator $\hat{v}_2$ of $\sum_{k=0}^{\phi_1(0)} \mathbb{E}\left(\bar{\eta}_1^{k,M}(\phi_{k,\tilde{g}_1})\right)$ with relative 95% confidence interval $[a_2, b_2]$ with $\hat{v}_2, a_2, b_2$ respectively defined as in (4.72), (4.73) and (4.74).

3. We check that $\hat{v}_{1,1} - \frac{\hat{v}_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$ is close to $\hat{v}_2$ as expected.

n=1	value	CI
$\hat{v}_{1,1} - \frac{\hat{v}_{1,2}}{\left(\bar{\eta}_1^0(g_1)\right)^4} - \frac{\bar{\eta}_0^0\left(g_0\left(Pf_1^2 - (Pf_1)^2\right)\right)}{\left(\bar{\eta}_1^0(g_1)\right)^4 \bar{\eta}_0^0(g_0)}$	0.4729737	[0.4690806 , 0.4768669]
$\hat{v}_2$	0.4725217	[0.4724719 , 0.4725714]

Bibliography

[1] J. Aaronson et al. "Strong Laws for L- and U-Statistics". In: *Trans. Amer. Math. Soc.* 348 (1996), 2845–2866.

[2] T.M. Apostol. *Mathematical Analysis*. Reading, Mass: Addison-Wesley, 1957.

[3] V. Bally and L. Caramellino. "Asymptotic development for the CLT in total variation distance". In: *Bernoulli* 22 (4) (2016), pp. 2442 –2485.

[4] H. Bauer. *Measure and Integration Theory*. De Gruyter, 2001.

[5] F. Benford. "The law of anomalous numbers". In: *Proceedings of the American Philosophical Society* 78 (1938), 551–572.

[6] P. Bertail and S. Clémençon. "A Renewal Approach to Markovian U -Statistics". In: *Mathematical Methods of Statistics* 20(2) (2011), 79–105.

[7] P. Billingsley. *Probability and Measure*. Wiley series in probability, mathematical statistics: probability, and mathematical statistics, 3rd edn. A Wiley-Interscience Publication, 1995.

[8] D. D. Boos and R. J. Serfling. "A note on differentials and the CLT and LIL for statistical functions, with application to M-estimates". In: *Ann. Statist.* 8(3) (1980), pp. 618–624.

[9] J. Boyle. "An application of Fourier series to the most significant digit problem". In: *Amer. Math.* 101 (1994), 879–886.

[10] O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer, 2005.

[11] P. Cardaliaguet, F. Delarue, J.M. Lasry, and P.L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games:(AMS-201), volume 201*. Princeton University Press, 2019.

[12] J. Carpenter, P. Clifford, and P. Fearnhead. "Improved particle filter for nonlinear problems". In: *IEE Proceedings Radar, Sonar Navigation* 146 (1) (1999), pp. 2–7.

[13] T. K. Chandra and A. Goswami. "Cesàro α-Integrability and Laws of Large Numbers-II". In: *Journal of Theoretical Probability* 19 (2006), pp. 789–816.

[14] N. Chopin. "Central limit theorem for sequential monte carlo methods and its application to bayesian inference". In: *The Annals of Statistics* 32 (6) (2004), 2385–2411.

[15] N. Chopin and O. Papaspiliopoulos. *An introduction to sequential Monte Carlo*. Springer Series in Statistics. Springer, 2020.

[16] N. Chopin, S.S. Singh, T. Soto, and M. Vihola. "On resampling schemes for particle filters with weakly informative observations". In: *Annals of Statistics* 50 (6) (2022), pp. 3197–3222.

[17] M. Denker and G. Keller. "Rigorous Statistical Procedures for Data from Dynamical Systems". In: *J. Statist. Phys.* 44 (1986), 67–93.

[18] P. Diaconis. "The distribution of leading digits and uniform distribution mod 1". In: *Ann. Probab.* 5 (1979), pp. 72–81.

[19] R. Douc, O. Cappé, and E. Moulines. "Comparison of resampling schemes for particle filtering". In: *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis* (2005), pp. 64–69.

[20] A. Doucet, N. de Freitas, and N.J. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.

[21] R. M. Dudley. "Nonlinear functionals of empirical measures and the bootstrap". In: *Probability in Banach spaces* 7 (1988), pp. 63–82.

[22] M. Evans and T. Swartz. "Methods for approximating integrals in Statistics with special emphasis on Bayesian integration problems". In: *Statist. Sci.* 10 (1995), 254–272.

[23] P. Fearnhead. *Sequential Monte Carlo methods in filter theory*. PhD thesis, University of Oxford, 1998.

[24] W. Feller. *An Introduction to Probability Theory and Its Applications. vol. II*. New York: Wiley, 1971.

[25] R. Flenghi and B. Jourdain. "Central limit theorem for the stratified selection mechanism". In: *ArXiv:2308.02186* (2023).

[26] R. Flenghi and B. Jourdain. "Central limit theorem over non-linear functionals of empirical measures: beyond the iid setting". In: *ArXiv:2204.06482* (2022).

[27] R. Flenghi and B. Jourdain. "Convergence to the uniform distribution of vectors of partial sums modulo one with a common factor". In: *ArXiv:2308.01874* (2023).

[28] M. Gerber, N. Chopin, and N. Whiteley. "Negative association, ordering and convergence of resampling methods". In: *The Annals of Statistics* 47 (4) (2019), 2236–2260.

[29] N. Gordon, D. Salmond, and A. Smith. "Novel approach to nonlinear/non-Gaussian Bayesian state estimation". In: *IEE Proceedings F (Radar and Signal Processing)* 140 (2) (1993), pp. 107–113.

[30] M. Hairer and J. C. Mattingly. "Yet another look at Harris' ergodic theorem for Markov chains". In: *Seminar on Stochastic Analysis, Random Fields and Applications VI* 63 (2011), pp. 109–117.

[31] P. Hall and C. C. Heyde. *Martingale limit theory and its applications*. Academic press, 2014.

[32] J. Handschin. "Monte Carlo techniques for prediction and filtering of nonlinear stochastic processes". In: *Automatica* 6 (1970), 555–563.

[33] J. Handschin and D. Mayne. "Monte Carlo techniques to estimate the conditionnal expectation in multi-stage non-linear filtering". In: *Int. J. Control* 9 (1969), 547–559.

[34] W. Hoeffding. "A Class of Statistics with Asymptotically Normal Distribution". In: *Ann. Math. Statist.* 19(3) (1948), pp. 293–325.

[35] K. Joag-Dev and F. Proschan. "Negative association of random variables with applications". In: *The Annals of Statistics* 11 (1) (1983), pp. 286–295.

[36] B. Jourdain and A. Tse. "Central limit theorem over non-linear functionals of empirical measures with applications to the mean-field fluctuation of interacting particle systems". In: *Electronic Journal of Probability* 26 (2021), pp. 1–34.

[37] G. Kitagawa. "Monte Carlo filter and smoother for non-gaussian state space models". In: *Journal of Computational and Graphical Statistics* 5 (1) (1996), pp. 1–25.

[38] H. R. Künsch. "Recursive Monte-Carlo filters: algorithms and theoretical analysis". In: *The Annals of Statistics* 33 (5) (2005), 1983–2021.

[39] M.P. Lévy. "L'addition des variables aléatoires définies sur une circonférence". In: *Bulletin de la Société Mathématique de France* 67 (1939), pp. 1–41.

[40] E. Ley. "On the peculiar distribution of the U.S. Stock Indices Digits". In: *The American Statistician* 50 (4) (1996), 311–313.

[41] J. Liu and R. Chen. "Sequential Monte-Carlo methods for dynamic systems". In: *Journal of the Royal Statistical Society B* 93 (1998), 1032–1044.

[42] M. Mamatov and M.K. Halikov. "Global limit theorems for distribution functions in the higher-dimensional case". In: *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk* 1 (1964), 13–21.

[43] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, 1993.

[44] S. J. Miller and M. J. Nigrini. "The Modulo 1 Central Limit Theorem and Benford's Law for Products". In: *International Journal of Algebra* 2 (3) (2008), pp. 119 –130.

[45] R. De Misés. "Les lois de probabilité pour les fonctions statistiques". In: *Ann. Inst. H. Poincaré* 6 (1936), pp. 185–212.

[46] R. De Misés. "On the asymptotic distribution of differentiable statistical functions". In: *Ann. Math. Statistics* 18 (1947), pp. 309–348.

[47] P. Del Moral. *Feynman-Kac formulae. Genealogical and interacting particle systems with applications*. Probability and its Applications. Springer, 2004.

[48] M. Nigrini and S. J. Miller. "Benford's Law applied to hydrology data results and relevance to other geophysical data". In: *Mathematical Geology* 5 (2007), 469–490.

[49] K.R. Parthasarathy and T. Steerneman. "A Tool in Establishing Total Variation Convergence". In: *Proceedings of the American Mathematical Society* 95 (4) (1985), pp. 626–630.

[50] Yu.V. Prohorov. "A local theorem for densities". In: *Doklady Akad. Nauk SSSR (N.S.)* 83 (1952), 797–800.

[51] B. Ristic, M. Arulampalam, and A. Gordon. *Beyond Kalman Filters: Particle Filters for Target Tracking*. Artech House, 2004.

[52] C. P. Robert and G.Casella. *Monte Carlo Statistical Methods (2nd ed.)* Springer, 2004.

[53] J. Roe. *Elliptic operators, topology and asymptotic methods (2nd ed.)* Chapman and Hall/CRC, 1999.

[54] P. Schatte. "On sums modulo 2π of independent random variables". In: *Math. Nachr.* 110 (1983), 243–261.

[55] P. Schatte. "On the asymptotic logarithmic distribution of the floating-point mantissas of sums". In: *Math. Nachr.* 115 (1984), 275–281.

[56] R. J. Serfling. *Approximation theorems of mathematical statistics*. Wiley Series in Probability and Mathematical Statistics, 1980.

[57] W. L. Smith. "Regenerative Stochastic Processes". In: *Proc. Roy. Statist. Soc.* 232 (1955), 6–31.

[58] H. Störmer. "Limit theorems for distributions of sums reduced modulo a". In: *Stochastic Processes and their Applications* 22 (1986), pp. 343–347.

[59] Z. S. Szewczak. "A limit theorem for random sums modulo 1". In: *Statistics and Probability Letters* 80 (2010), 747–751.

[60] J. A. Wellner. "A Glivenko-Cantelli theorem for empirical measures of independent but not identically distributed random variables". In: *Stochastic Processes and their Applications* 11 (1981).