NONCONVENTIONAL LIMIT THEOREMS IN DISCRETE AND CONTINUOUS TIME VIA MARTINGALES

YURI KIFER AND S.R.S. VARADHAN

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY, JERUSALEM, ISRAEL

COURANT INSTITUTE, NEW YORK UNIVERSITY, NEW YORK, USA

ABSTRACT. We obtain functional central limit theorems for both discrete time expressions of the form $1/\sqrt{N}\sum_{n=1}^{[Nt]}\big(F\big(X(q_1(n)),\dots,X(q_\ell(n))\big)-\bar F\big)$ and similar expressions in the continuous time where the sum is replaced by an integral. Here $X(n), n\geq 0$ is a sufficiently fast mixing vector process with some moment conditions and stationarity properties, F is a continuous function with polynomial growth and certain regularity properties, $\bar F=\int Fd(\mu\times\cdots\times\mu)$, μ is the distribution of $X(0)$ and $q_i(n)=in$ for $i\leq k\leq\ell$ while for $i>k$ they are positive functions taking on integer values on integers with some growth conditions which are satisfied, for instance, when q_i's are polynomials of increasing degrees. These results decisively generalize [18] whose method was only applicable to the case $k=2$ under substantially more restrictive moment and mixing conditions and which could not be extended to convergence of processes and to the corresponding continuous time case. As in [18] our results hold true when $X_i(n)=T^nf_i$ where T is a mixing subshift of finite type, a hyperbolic diffeomorphism or an expanding transformation taken with a Gibbs invariant measure, as well, as in the case when $X_i(n)=f_i(\Upsilon_n)$ where Υ_n is a Markov chain satisfying the Doeblin condition considered as a stationary process with respect to its invariant measure. Moreover, our relaxed mixing conditions yield applications to other types of dynamical systems and Markov processes, for instance, where a spectral gap can be established. The continuous time version holds true when, for instance, $X_i(t)=f_i(\xi_t)$ where ξ_t is a nondegenerate continuous time Markov chain with a finite state space or a nondegenerate diffusion on a compact manifold. A partial motivation for such limit theorems is due to a series of papers dealing with nonconventional ergodic averages.

1. Introduction

Nonconventional ergodic theorems known also after [1] as polynomial ergodic theorems studied the limits of expressions having the form (see [9], [11], [10]) $1/N\sum_{n=1}^{N}T^{q_1(n)}f_1\cdots T^{q_\ell(n)}f_\ell$ where T is a weakly mixing measure preserving transformation, f_i's are bounded measurable functions and q_i's are polynomials taking on integer values on the integers. Originally, these results were motivated

Date: December 31, 2011.

2000 *Mathematics Subject Classification.* Primary: 60F17 Secondary: 60G42, 37D99, 60G15.

Key words and phrases. limit theorems, martingale approximation, mixing, Markov processes, hyperbolic diffeomorphisms.

Yu. Kifer was supported by ISF grants 130/06 and 82/10 and S.R.S. Varadhan was supported by NSF grants OISE 0730136 and DMS 0904701 .

by applications to multiple recurrence for dynamical systems, the functions f_i being indicators of some measurable sets.

After an ergodic theorem (or in the probabilistic language: the law of large numbers) is established it is natural to inquire whether a corresponding central limit theorem holds true, as well, though as usual under stronger conditions. In this paper we prove the functional central limit theorem (invariance principle) for expressions of the form

$$\frac{1}{\sqrt{N}}\sum_{n=1}^{[Nt]}\big(F\big(X(q_1(n)),...,X(q_\ell(n))\big)-\bar{F}\big) \tag{1.1}$$

and for the corresponding continuous time expressions of the form

$$\frac{1}{\sqrt{N}}\int_0^{[Nt]}\big(F\big(X(q_1(t)),...,X(q_\ell(t))\big)-\bar{F}\big)dt \tag{1.2}$$

where $\{X(n), n\geq 0\}$, (or $\{X(t)\}, t\geq 0$) is a sufficiently fast mixing vector valued process with some stationarity properties satisfying certain moment conditions, F is a continuous function with polynomial growth with certain regularity properties. $\bar{F}=\int F\,d(\mu\times\cdots\times\mu)$, where μ is the common distribution of $X(n)$, $\{q_j(t)\}$ are positive functions, (taking on integer values on integers in the discrete time case), $q_j(t)=jt$ for $j\leq k$ and for $j>k$ they satisfy certain growth conditions. For instance, it would be enough if $\{q_j(t)\}$ are polynomials of increasing degrees, though we actually do not need any polynomial structure of functions q_j, $j>k$ which was crucial in papers dealing with nonconventional ergodic theorems cited above.

Our methods rely on a martingale approach which have played a decisive role in most proofs of the central limit theorem during the last 50 years. In view of strong dependence on the future of summands in (1.1) application of martingales in our setup does not seem plausible on the first sight. It turns out, somewhat surprisingly, that an appropriately modified martingale approach still works well in our situation if we construct the filtration of σ-algebras so that in some sense "future becomes present". Unlike the classical situation our functional central limit theorem yields a process which has Gaussian distributions but not necessarily independent increments. This interesting effect rarely appears in natural models. Under stronger mixing and moment conditions we derive also convergence of moments of the above expressions to the corresponding moments of the limiting Gaussian distribution. We obtain also a functional central limit theorem in the corresponding continuous time case which only recently was treated in the sense of nonconventional ergodic theorems (see [5]). It turns out that the limiting process in the continuous time case has a somewhat different structure than in the discrete time setup. These results generalize [18] where the partition into blocks and the direct use of characteristic functions showed applicability only to the case $k=2$ under more restrictive conditions and neither the functional central limit theorem nor the continuous time case could be dealt with by the method employed there.

As in [18] our results hold true when, for instance, $X(n)=T^nf$ where $f=(f_1,...,f_\wp)$, T is a mixing subshift of finite type, a hyperbolic diffeomorphism or an expanding transformation taken with a Gibbs invariant measure and some other dynamical systems, as well, as in the case when $X(n)=f(\Upsilon_n)$, $f=(f_1,...,f_\wp)$ where Υ_n is a Markov chain satisfying the Doeblin condition (see [15]) considered as a stationary process with respect to its invariant measure. In the dynamical systems case each f_i should be either Hölder continuous or piecewise constant on elements of Markov partitions (see [3]). As an application we can consider $F(x_1,...,x_\ell)=x_1^{(1)}\cdots x_\ell^{(\ell)},\ x_j=(x_j^{(1)},...,x_j^{(\ell)}),\ X(n)=(X_1(n),...,X_\ell(n)),\ X_j(n)=\mathbb{I}_{A_j}(T^nx)$ in the dynamical systems case and $X_j(n)=\mathbb{I}_{A_j}(\Upsilon_n)$ in the Markov chains case where $\mathbb{I}_A$ is the indicator of a set A. If $N(n)$ is the number of l's between 0 and n for which $T^{jl}x\in A_j$ for $j=0,1,...,k$ with $k=\ell$ (or $\Upsilon_{jl}\in A_j$ in the Markov chains case) then $N(n)$ is the number of k-tuples of return times to A_j's (either by T^lx or by Υ_l) which form an arithmetic progression of length k having a difference between 0 and n. Our result implies in this case that $n^{-1/2}\big(N([tn])-nt(\mu(A))^k\big)$, $t\in[0,1]$ (where μ is a corresponding invariant measure of T or ξ_n, respectively) weakly converges as $n\to\infty$ to a Gaussian process having not necessarily independent increments. Substantially more general than [18] setup of the present paper enables us to apply results to additional classes of dynamical systems and Markov processes, in particular, to those having a spectral gap.

The continuous time version holds true, in particular, when $X_i(t)=f_i(\Upsilon_t)$ where Υ_t is an irreducible continuous time Markov chain or a nondegenerate diffusion on a compact manifold.

2. Preliminaries and main results

Our discrete time setup consists of a $\wp$-dimensional stochastic process $\{X(n),\,n=0,1,...\}$ on a probability space $(\Omega,\mathcal{F},P)$ and of a family of σ-algebras $\mathcal{F}_{kl}\subset\mathcal{F}$, $-\infty\le k\le l\le\infty$ such that $\mathcal{F}_{kl}\subset\mathcal{F}_{k'l'}$ if $k'\le k$ and $l'\ge l$. It is often convenient to measure the dependence between two sub σ-algebras $\mathcal{G},\mathcal{H}\subset\mathcal{F}$ via the quantities

$$\varpi_{q,p}(\mathcal{G},\mathcal{H})=\sup\{\|E\big[g|\mathcal{G}\big]-E[g]\|_p:\ g\text{ is }\mathcal{H}-\text{measurable and }\|g\|_q\le 1\},\tag{2.1}$$

where the supremum is taken over real functions and $\|\cdot\|_r$ is the $L^r(\Omega,\mathcal{F},P)$-norm. Then more familiar α,ρ,ϕ and ψ-mixing (dependence) coefficients can be expressed via the formulas (see [4], Ch. 4),

$$\alpha(\mathcal{G},\mathcal{H})=\tfrac14\varpi_{\infty,1}(\mathcal{G},\mathcal{H}),\ \rho(\mathcal{G},\mathcal{H})=\varpi_{2,2}(\mathcal{G},\mathcal{H})$$
$$\phi(\mathcal{G},\mathcal{H})=\tfrac12\varpi_{\infty,\infty}(\mathcal{G},\mathcal{H})\text{ and }\psi(\mathcal{G},\mathcal{H})=\varpi_{1,\infty}(\mathcal{G},\mathcal{H}).$$

We set also

$$\varpi_{q,p}(n)=\sup_{k\ge0}\varpi_{q,p}(\mathcal{F}_{-\infty,k},\mathcal{F}_{k+n,\infty})\tag{2.2}$$

and accordingly

$$\alpha(n)=\frac14\varpi_{\infty,1}(n),\ \rho(n)=\varpi_{2,2}(n),\ \phi(n)=\frac12\varpi_{\infty,\infty}(n),\ \psi(n)=\varpi_{1,\infty}(n).$$

We will impose mixing rates, i.e. rates of decay of $\varpi_{q,p}(n)$ requiring that

$$C(q,p)=\sum_{n\ge1}\varpi_{q,p}(n)\tag{2.3}$$

is finite for some choices of p and q. Our setup includes also conditions on the approximation rate

$$\beta(p,r)=\sup_{k\ge0}\|X(k)-E\big[X(k)|\mathcal{F}_{k-r,k+r}\big]\|_p.\tag{2.4}$$

In what follows we can always extend the definitions of $\mathcal{F}_{kl}$ given only for $k, l \geq 0$ to negative k by defining $\mathcal{F}_{kl} = \mathcal{F}_{0l}$ for $k < 0$ and $l \geq 0$. Furthermore, we do not require stationarity of the process $X(n), n \geq 0$ assuming only that the distribution of $X(n)$ does not depend on n and the joint distribution of $\{X(n), X(n')\}$ depends only on $n - n'$ which we write for further references by

$$X(n) \overset{d}{\sim} \mu \text{ and } (X(n), X(n')) \overset{d}{\sim} \mu_{n-n'} \text{ for all } n, n' \tag{2.5}$$

where $Y \overset{d}{\sim} \mu$ means that Y has μ for its distribution.

Next, let $F = F(x_1, ..., x_\ell)$, $x_j \in \mathbb{R}^\wp$ be a function on $\mathbb{R}^{\wp\ell}$ such that for some $\iota, K > 0, \kappa \in (0, 1]$ and all $x_i, y_i \in \mathbb{R}^\wp, i = 1, ..., \ell$, we have

$$|F(x_1, ..., x_\ell) - F(y_1, ..., y_\ell)| \leq K\big[1 + \sum_{j=1}^{\ell} |x_j|^\iota + \sum_{j=1}^{\ell} |y_j|^\iota\big] \sum_{j=1}^{\ell} |x_j - y_j|^\kappa \tag{2.6}$$

and

$$|F(x_1, ..., x_\ell)| \leq K\big[1 + \sum_{j=1}^{\ell} |x_j|^\iota\big]. \tag{2.7}$$

To simplify formulas we assume a centering condition

$$\bar{F} = \int F(x_1, ..., x_\ell)\, d\mu(x_1) \cdots d\mu(x_\ell) = 0 \tag{2.8}$$

which is not really a restriction since we can always replace F by $F - \bar{F}$. Our goal is to prove a functional central limit theorem for

$$\xi_N(t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nt]} F\big(X(q_1(n)), ..., X(q_\ell(n))\big) \text{ and } t \in [0, T] \tag{2.9}$$

where $q_1(n) < q_2(n) < \cdots < q_\ell(n)$ are increasing functions taking on integer values on integers and such that for $j \leq k$, $q_j(n) = jn$, whereas the remaining ones grow faster in n. We assume that for $k + 1 \leq i \leq \ell$,

$$\lim_{n\to\infty} (q_i(n + 1) - q_i(n)) = \infty \tag{2.10}$$

and for $i \geq k$ and any $\epsilon > 0$,

$$\liminf_{n\to\infty} (q_{i+1}(\epsilon n) - q_i(n)) > 0 \tag{2.11}$$

which implies because of (2.10) that

$$\lim_{n\to\infty} (q_{i+1}(\epsilon n) - q_i(n)) = \infty. \tag{2.12}$$

To shorten some of arguments we assumed that $q_i(n)$ is increasing in both n and i but, in fact, (2.10) and (2.11) imply already that this holds true for all n large enough which suffices for our purposes. For each $\theta > 0$ set

$$\gamma_\theta^\theta = \|X\|_\theta^\theta = E|X(n)|^\theta = \int |x|^\theta d\mu. \tag{2.13}$$

Our main result relies on

2.1**. Assumption.** With $d=(\ell-1)\wp$ there exist $\infty>p,q\geq 1$ and $\delta,m>0$ with $\delta<\kappa-\frac{d}{p}$ satisfying

$$\sum_{n=0}^{\infty}\varpi_{q,p}(n)=\theta(p,q)<\infty, \tag{2.14}$$

$$\sum_{r=0}^{\infty}[\beta(q,r)]^{\delta}<\infty, \tag{2.15}$$

$$\gamma_m<\infty,\ \gamma_{2q\iota}<\infty \text{ with } \frac{1}{2}\geq\frac{1}{p}+\frac{\iota+2}{m}+\frac{\delta}{q}. \tag{2.16}$$

In order to give a detailed statement of our main result as well as for its proof it will be essential to represent the function $F=F(x_1,x_2,\dots,x_\ell)$ in the form

$$F=F_1(x_1)+\cdots+F_\ell(x_1,x_2,\dots,x_\ell) \tag{2.17}$$

where for $i<\ell$,

$$\begin{aligned}F_i(x_1,\dots,x_i)=&\textstyle\int F(x_1,x_2,\dots,x_\ell)\ d\mu(x_{i+1})\cdots d\mu(x_\ell)\\ &-\textstyle\int F(x_1,x_2,\dots,x_\ell)\,d\mu(x_i)\cdots d\mu(x_\ell)\end{aligned} \tag{2.18}$$

and

$$F_\ell(x_1,x_2,\dots,x_\ell)=F(x_1,x_2,\dots,x_\ell)-\int F(x_1,x_2,\dots,x_\ell)\,d\mu(x_\ell)$$

which ensures, in particular, that

$$\int F_i(x_1,x_2,\dots,x_{i-1},x_i)\,d\mu(x_i)\equiv 0\quad\forall\quad x_1,x_2,\dots,x_{i-1}. \tag{2.19}$$

These enable us to write

$$\xi_N(t)=\sum_{i=1}^{k}\xi_{i,N}(it)+\sum_{i=k+1}^{\ell}\xi_{i,N}(t) \tag{2.20}$$

where for $1\leq i\leq k$,

$$\xi_{i,N}(t)=\frac{1}{\sqrt{N}}\sum_{n=1}^{[\frac{Nt}{i}]}F_i(X(n),X(2n),\dots,X(in)) \tag{2.21}$$

and for $i\geq k+1$,

$$\xi_{i,N}(t)=\frac{1}{\sqrt{N}}\sum_{n=1}^{[Nt]}F_i(X(q_1(n)),\dots,X(q_i(n))). \tag{2.22}$$

2.2**. Theorem.** *Suppose that Assumption 2.1 holds true. Then the ℓ-dimensional process $\{\xi_{i,N}(t):1\leq i\leq\ell\}$ converges in distribution as $N\to\infty$ to a Gaussian process $\{\eta_i(t):1\leq i\leq\ell\}$ with stationary independent increments. The means are 0 and the covariances are given by $E[\eta_i(s)\eta_j(t)]=\min(s,t)D_{i,j}$. For $i,j\leq k$, $D_{i,j}$ is given by Proposition 4.1. Moreover $D_{i,j}=0$ if $i\neq j$, and either i or j is at least $k+1$, making the processes $\{\eta_i(\cdot),\ i\geq k+1\}$ independent of each other and of $\{\eta_j(\cdot):j\leq k\}$. For $i\geq k+1$, the variance of $\eta_i(t)$ is given by $tD_{i,i}$ where*

$$D_{i,i}=\int|F_i(x_1,x_2,...,x_i)|^2d\mu(x_1)d\mu(x_2)\cdots d\mu(x_i).$$

Finally, the distribution of the process $\xi_N(\cdot)$ converges to the Gaussian process $\xi(\cdot)$ which can be represented in the form

$$\xi(t)=\sum_{i=1}^{k}\eta_i(it)+\sum_{i=k+1}^{\ell}\eta_i(t). \tag{2.23}$$

If $k\geq 2$, then the process $\xi(t)$ may not have independent increments.

In order to understand our assumptions observe that $\varpi_{q,p}$ is clearly non-increasing in q and non-decreasing in p. Hence, for any pair $p,q\geq 1$,

$$\varpi_{q,p}(n)\leq\psi(n).$$

Furthermore, by the real version of the Riesz–Thorin interpolation theorem or the Riesz convexity theorem (see [12], Section 9.3 and [7], Section VI.10.11) whenever $\theta\in[0,1]$, $1\leq p_0,p_1,q_0,q_1\leq\infty$ and

$$\frac{1}{p}=\frac{1-\theta}{p_0}+\frac{\theta}{p_1},\ \frac{1}{q}=\frac{1-\theta}{q_0}+\frac{\theta}{q_1}$$

then

$$\varpi_{q,p}(n)\leq 2(\varpi_{q_0,p_0}(n))^{1-\theta}(\varpi_{q_1,p_1}(n))^{\theta}. \tag{2.24}$$

In particular, using the obvious bound $\varpi_{q_1,p_1}\leq 2$ valid for any $q_1\geq p_1$ we obtain from (2.24) for pairs $(\infty,1)$, $(2,2)$ and (∞,∞) that for all $q\geq p\geq 1$,

$$\varpi_{q,p}(n)\leq(2\alpha(n))^{\frac{1}{p}-\frac{1}{q}},\ \varpi_{q,p}(n)\leq 2^{1+\frac{1}{p}-\frac{1}{q}}(\rho(n))^{1-\frac{1}{p}+\frac{1}{q}}$$
$$\text{and }\varpi_{q,p}(n)\leq 2^{1+\frac{1}{p}}(\phi(n))^{1-\frac{1}{p}}. \tag{2.25}$$

We observe also that by the Hölder inequality for $q\geq p\geq 1$ and $\alpha\in(0,p/q)$,

$$\beta(q,r)\leq 2^{1-\alpha}[\beta(p,r)]^{\alpha}\gamma^{1-\alpha}_{\frac{pq(1-\alpha)}{p-q\alpha}} \tag{2.26}$$

with γ_θ defined in (2.13). Thus, we can formulate Assumption 2.1 in terms of more familiar α, ρ, ϕ, and ψ–mixing coefficients and with various moment conditions. It follows also from (2.24) that if $\varpi_{q,p}(n)\to 0$ as $n\to\infty$ for some $q>p\geq 1$ then

$$\varpi_{q,p}(n)\to 0 \text{ as } n\to\infty \text{ for all } q>p\geq 1, \tag{2.27}$$

and so (2.27) holds true under Assumption 2.1.

The key point of our proof will be construction of martingale approximations for the processes $\xi_{i,N}(t)$'s where we will have to overcome problems imposed by strong dependencies between terms in the sum (2.9), as well as between arguments $X(q_j(n))$, $j=1,2,...,\ell$ of the function F there. The realignment in the definition of $\{\xi_{i,N}(t)\}$ for $i\leq k$ will also be important since it makes the collection a process with independent increments in the limit. Otherwise, in the limit, increments of $\{\xi_i(t)\}$ will be correlated with the increments of $\{\xi_j(t)\}$ at different time points. It will not matter for $i\geq k+1$, for they will all turn out to be mutually independent in the limit.

We observe that the estimates of Section 3 not only enable us to rely on martingale approximations and martingale limit theorems but, actually, Corollary 3.6 yields a possibility to represent $F_i(X(q_1(n)),\dots,X(q_i(n)))-EF_i(X(q_1(n)),\dots,X(q_i(n)))$, $n\geq 1$ as a mixingale sequence studied in [21], and so in order to prove weak convergence of $\xi_{i,N}$ to Gaussian processes with independent increments we could take advantage of the (weak) invariance principle obtained in

[21]. Still, in order to derive Theorem 2.2 in its full strength we would need most of estimates of Section 3 and limiting covariances identification of Section 4 together with some of arguments of Section 5 enabling us to conclude about limiting behavior of the sum process ξ_N which has nothing to do with [21]. Anyway, we find it more beneficial for a reader to proceed here relying on the more familiar invariance principle for martingale differences.

Under stronger mixing and moment conditions we are able to derive convergence of all moments of $\xi_N(t)$ to the corresponding moments of $\eta(t)$.

2.3. **Assumption.** For every p there exists $\infty > q \geq p$ such that

$$\sum_{n=0}^{\infty} \varpi_{q,p}(n) = \theta(p,q) < \infty. \tag{2.28}$$

For every $q, \delta > 0$,

$$\sum_{r=0}^{\infty} [\beta(q,r)]^{\delta} < \infty \tag{2.29}$$

and the moments $\gamma_m < \infty$ exist for every $m > 0$.

2.4. **Theorem.** *Suppose that Assumption 2.3 holds true. Then for any $k = 1, 2, ...$ and $t \geq 0$,*

$$\lim_{N\to\infty} E[\xi_N^k(t)] = E[\eta^k(t)]. \tag{2.30}$$

This result will follow from Theorem 2.2 together with additional moment estimates.

The conditions of Theorems 2.2 and 2.4 hold true for many important models. Let, for instance, Υ_n be a Markov chain on a space M satisfying the Doeblin condition (see, for instance, [15], p.p. 367–368) and f_j, $j = 1, ..., \ell$ be bounded measurable functions on the space of sequences $x = (x_i,\, i = 0, 1, 2, ...,\, x_i \in M)$ such that $|f_j(x) - f_j(y)| \leq Ce^{-cn}$ provided $x = (x_i)$, $y = (y_i)$ and $x_i = y_i$ for all $i = 0, 1, ..., n$ where $c, C > 0$ do not depend on n and j. In fact, some polynomial decay in n will suffice here, as well. Let $X(n) = (X_1(n), ..., X_\ell(n))$ with $X_j(n) = f_j(\Upsilon_n, \Upsilon_{n+1}, \Upsilon_{n+2}, ...)$ and take σ-algebras $\mathcal{F}_{kl}$, $k < l$ generated by $\Upsilon_k, \Upsilon_{k+1}, ..., \Upsilon_l$ then our condition will be satisfied considering $\{\Upsilon_n,\, n \geq 0\}$ with its invariant measure as a stationary process. In fact, our conditions hold true for a more general class of processes, in particular, for Markov chains whose transition operator has a spectral gap which leads to an exponentially fast decay of the ρ-mixing coefficient.

2.5. **Remark.** Formally, (2.5) requires some stationarity and, for instance, if we consider a Markov chain ξ_n satisfying the Doeblin condition but whose initial distribution differs from its invariant measure then (2.5) does not hold true for $X(n) = f(\xi_n)$. Still, a slight modification makes our method to work so that Theorems 2.2 and 2.4 (as well, as their continuous time version Theorem 2.6) remain valid. In order to do this we consider another probability measure Π on the space $(\Omega, \mathcal{F})$ and require the weak stationarity (2.5) with respect to Π, i.e. $X(n)\Pi = \mu$ and $(X(n), X(n'))\Pi = \mu_{n-n'}$. In addition, we modify the definition of the dependence coefficient $\varpi_{q,p}$ in (2.1) taking the conditional expectation of g there with respect to the probability P while the unconditional expectation of g taking with respect to Π. It is easy to see that under the same assumptions as above but with modified (2.1) and (2.5) our proof will still go through.

Important classes of processes satisfying our conditions come from dynamical systems. Let T be a C^2 Axiom A diffeomorphism (in particular, Anosov) in a neighborhood of an attractor or let T be an expanding C^2 endomorphism of a Riemannian manifold M (see [3]), f_j's be either Hölder continuous functions or functions which are constant on elements of a Markov partition and let $X(n) = (X_1(n),...,X_\ell(n))$ with $X_j(n) = f_j(T^n x)$. Here the probability space is $(M,\mathcal{B},\mu)$ where μ is a Gibbs invariant measure corresponding to some Hölder continuous function and $\mathcal{B}$ is the Borel σ-field. Let ζ be a finite Markov partition for T then we can take $\mathcal{F}_{kl}$ to be the finite σ-algebra generated by the partition $\cap_{i=k}^{l} T^i\zeta$. In fact, we can take here not only Hölder continuous f_j's but also indicators of sets from $\mathcal{F}_{kl}$. A related example corresponds to T being a topologically mixing subshift of finite type which means that T is the left shift on a subspace Ξ of the space of one-sided sequences $\varsigma = (\varsigma_i, i \geq 0), \varsigma_i = 1,...,l_0$ such that $\varsigma \in \Xi$ if $\pi_{\varsigma_i\varsigma_{i+1}} = 1$ for all $i \geq 0$ where $\Pi = (\pi_{ij})$ is an $l_0 \times l_0$ matrix with 0 and 1 entries and such that Π^n for some n is a matrix with positive entries. Again, we have to take in this case f_j to be Hölder continuous bounded functions on the sequence space above, μ to be a Gibbs invariant measure corresponding to some Hölder continuous function and to define $\mathcal{F}_{kl}$ as the finite σ-algebra generated by cylinder sets with fixed coordinates having numbers from k to l. The exponentially fast ψ-mixing is well known in the above cases (see [3]). Among other dynamical systems with exponentially fast ψ-mixing we can mention also the Gauss map $Tx = \{1/x\}$ (where $\{\cdot\}$ denotes the fractional part) of the unit interval with respect to the Gauss measure G (see [13]). The latter enables us to consider the number $N_a(x,n)$, $a = (a_1,...,a_\ell)$ of m's between 0 and n such that the $q_j(m)$-th digit of the continued fraction of x equals certain integer $a_j, j = 1,...,\ell$. Then Theorem 2.2 implies a central limit theorem for $N_a(x,n)$ considered as a random variable on the probability space $((0,1],\mathcal{B},G)$. In fact, our results rely only on sufficiently fast α or ρ-mixing which holds true for wider classes of dynamical system, in particular, those with a spectral gap (such as many one dimensional not necessarily uniformly expanding maps) which ensures an exponentially fast ρ-mixing. Of course, there are many stationary processes and dynamical systems with polynomially fast mixing which still satisfy our conditions but they are more difficult describe in short.

Next, we discuss a continuous time version of our theorem. Our continuous time setup consists of a $\wp$-dimensional process $X(t)$, $t \geq 0$ on a probability space $(\Omega,\mathcal{F},P)$ and of a family of σ-algebras $\mathcal{F}_{st} \subset \mathcal{F}$, $-\infty \leq s \leq t \leq \infty$ such that $\mathcal{F}_{st} \subset \mathcal{F}_{s't'}$ if $s' \leq s$ and $t' \geq t$. We assume that the distribution of $X(t)$ is independent of t and denote it by μ. The joint distribution of $\{X(t), X(t+s)\}$ is assumed to depend only on s and is denoted by μ_s. For all $t \geq 0$ we set

$$\varpi_{q,p}(t) = \sup_{s\geq 0} \varpi_{q,p}(\mathcal{F}_{-\infty,s}, \mathcal{F}_{s+t,\infty}) \tag{2.31}$$

and

$$\beta(p,t) = \sup_{s\geq 0} \|X(s) - E\big[X(s)|\mathcal{F}_{s-t,s+t}\big]\|_p. \tag{2.32}$$

where $\varpi_{q,p}(\mathcal{G},\mathcal{H})$ is defined by (2.1). We continue to impose Assumtions 2.1 and 2.3 on the decay rates of $\varpi_{q,p}(t)$ and $\beta(p,t)$. Although they only involve integer values of t, it will suffice since they are non-increasing functions of t . Let $q_1(t) < q_2(t) < \cdots < q_\ell(t)$ be increasing positive functions such that $q_i(t) = i\,t$ for $i = 1,...,k$ while $q_i(t)$, $i > k$ grow faster in t. We assume that these functions satisfy the conditions

(2.11) and (2.12) (with t in place of n) while (2.10) is replaced by

$$\lim_{t\to\infty}(q_i(t+\gamma)-q_i(t))=\infty \text{ for any } \gamma>0 \text{ and } i>k. \tag{2.33}$$

2.6**. Theorem.** *Suppose that Assumption 2.1 holds true. Then the distribution of the process*

$$\xi_N(t)=\frac{1}{\sqrt{N}}\int_0^{Nt}F\big(X(q_1(s)),...,X(q_\ell(s))\big)ds \tag{2.34}$$

on $C[0,T]$, converges to the distribution of a Gaussian process $\xi(t)$ which has the representation (2.23) but unlike in the discrete time case all processes η_i, $i>k$ are zero there while $\{\eta_1(t),...,\eta_k(t)\}$ is a k-dimensional Gaussian process having stationary independent increments. The means are 0 *and variances and covariances are given by $E[\eta_i(s)\eta_j(t)]=\min(s,t)D_{i,j}$, $i,j=1,...,k$. The expressions for these $D_{i,j}$ are provided in Section 7. If Assumption 2.3 holds true then the moments of $\xi_N(t)$ converge to corresponding moments of $\xi(t)$.*

The conditions of Theorem 2.6 are satisfied when, for instance, $X(t)=(X_1(t),...,X_\wp(t))$ with $X_j(t)=f_j(\xi_t)$ where ξ_t either an irreducible continuous time finite state Markov chain or a nondegenerate diffusion process on a compact manifold. On the other hand, these conditions do not usually hold true for important classes of continuous time dynamical systems (flows) having rich probabilistic properties such as Axiom A (in particular, Anosov) flows where the standard tool of suspension flows is usually applied while it does not seem to work in our circumstances and a different approach should be employed here. In fact, mixing in the framework of Markov families on unstable manifolds considered in [6] seems to be an appropriate framework in order to obtain a version of Theorem 2.6 when $X(t)=T^tf$ where T^t is an Anosov flow and f is a Hölder continuous vector function. In this case when, for instance, $l=k$ we can derive Theorem 2.6 from Theorem 2.2 representing

$$\int_0^N F(T^ty,T^{2t}y,...,T^{kt}y)dt=\sum_{n=0}^{N-1}\tilde F(T^ny,T^{2n}y,...,T^{kn}y)$$

where $\tilde F(y_1,...,y_k)=\int_0^1 F(T^ty_1,T^{2t}y_2,...,T^{kt}y_k)dt$.

3. Approximation Estimates

This section contains estimates which are crucial for our proofs. We will make repeated use of the following simple variations of Hölder's inequality.

3.1**. Lemma.** *(i) For any two random variables Z,D*

$$\|Z^hD^\kappa\|_a\le\|Z\|^h_{a^*}\|D\|^\kappa_{b^*}$$

provided $\frac{1}{a}\ge\frac{h}{a^*}+\frac{\kappa}{b^*}$. If, in addition, $|D|\le|Z|$ a.e. (almost everywhere), we can replace κ by $\alpha\le\kappa$ and change h to $h+\kappa-\alpha$ obtaining

$$\|Z^hD^\kappa\|_a\le\|Z^{h+\kappa-\alpha}D^\alpha\|_a\le\|Z\|^{h+\kappa-\alpha}_{a^*}\|D\|^\alpha_{b^*}$$

provided $\frac{1}{a}\ge\frac{h+\kappa-\alpha}{a^*}+\frac{\alpha}{b^*}$.

(ii) If $f(x,\omega)$ is a function of x and ω such that for almost all ω,

$$|f(x,\omega)|\le C(\omega)[1+|x|^h]$$

then

$$\|f(X(\omega),\omega)\|_a \le (1+\gamma_m^h)\|C(\omega)\|_p$$

provided $\frac{1}{a} \ge \frac{1}{p} + \frac{h}{m}$ *where* γ_m *is a bound for* $\|X\|_m$.

(iii) If $f(x,\omega)$ *is a function of* x *and* ω *satisfying for almost all* ω,

$$|f(x,\omega) - f(y,\omega)| \le H(\omega)[1+|x|^h + |y|^h]|x-y|^\delta$$

then

$$\|f(X(\omega),\omega) - f(Y(\omega),\omega)\|_a \le (1+2\gamma_m^h)\|H(\omega)\|_p\|X-Y\|_q^\delta \tag{3.1}$$

provided $\frac{1}{a} \ge \frac{1}{p} + \frac{h}{m} + \frac{\delta}{q}$ *where* γ_m *is a bound for* $\|X\|_m$ *and* $\|Y\|_m$.

Proof. For (i), by Hölder's inequality

$$\|Z^h D^\kappa\|_a = \big[E[Z^{ah}D^{a\kappa}]\big]^{\frac{1}{a}} \le \|Z\|_{a^*}^h \|D\|_{b^*}^\kappa$$

provided $\frac{1}{a} \ge \frac{h}{a^*} + \frac{\kappa}{b^*}$. If $|D| \le |Z|$ and $0 \le \alpha \le \kappa$,

$$\|D^\kappa Z^h\|_a \le \|D^\alpha Z^{h+\kappa-\alpha}\|_a \le \|Z\|_{a^*}^{(h+\kappa-\alpha)}\|D\|_{b^*}^\alpha$$

provided $\frac{1}{a} \ge \frac{h+\kappa-\alpha}{a^*} + \frac{\alpha}{b^*}$.

For (ii), by Hölder's inequality

$$\begin{aligned} E[|f(X(\omega),\omega)|^a] &\le E\big[[C(\omega)]^a\,[1+|X|^h]^a\big] \\ &\le \big[E\big[[C(\omega)]^p\big]\big]^{\frac{a}{p}}\big[E\big[[1+|X|^h]^{\frac{p^*}{h}}\big]\big]^{\frac{ah}{p^*}} \end{aligned}$$

provided $\frac{1}{a} \ge \frac{1}{p} + \frac{h}{p^*}$.

The assertion (iii) follows similarly from the inequality

$$E[|XYZ|] \le \|X\|_{s_1}\|Y\|_{s_2}\|Z\|_{s_3}$$

if $1 \ge \frac{1}{s_1} + \frac{1}{s_2} + \frac{1}{s_3}$. □

We will need also

3.2. **Lemma.** *(i) Let* $F(x_1,\dots,x_{\ell-1},x_\ell)$ *be any function that satisfies (2.6) and (2.7). Then the functions* $F_i(x_1,\dots,x_i)$ *defined in (2.18) will inherit similar properties from* F.

(ii) Let Z *be a random vector in* $L_\iota(P)$ *with* $\|Z\|_\iota \le \gamma_\iota$ *and* $\mathcal{G} \subset \mathcal{F}$ *be a sub* σ*-field. If*

$$G_i(x_1,\dots,x_{i-1},\omega) = E[F_i(x_1,\dots,x_{i-1},Z(\omega))|\mathcal{G}]$$

then

$$|G_i(x_1,\dots,x_{i-1},\omega)| \le C(1+C(\omega)^\iota + |x|^\iota)$$

and

$$\begin{aligned} &|G_i(x_1,\dots,x_{i-1},\omega) - G_i(y_1,\dots,y_{i-1},\omega)| \\ &\le C\big(1+C(\omega)^\iota + |x|^\iota + |y|^\iota\big)|x-y|^\kappa \end{aligned}$$

where $C>0$ *is a constant,* $C(\omega) = \big(2E[|Z|^\iota|\mathcal{G}]\big)^{\frac{1}{\iota}}$ *and* $\|C(\omega)\|_\iota^\iota \le 2\gamma_\iota^\iota$.

Proof. For (i), if

$$|F(x_1,x_2,\dots,x_i)|\le C_1(C_2+|x|^\iota)$$

then

$$|\int F(x_1,\dots,x_{i-1},x_i)d\mu(x_i)|\le\int|F(x_1,\dots,x_{i-1},x_i)|d\mu(x_i)$$
$$\le C_1\big(C_2+|x|^\iota+\gamma_\iota^\iota\big).$$

The Hölder property is similar.

The assertion (ii) follows from

$$|G_i(x_1,\dots,x_{i-1},\omega)|\le E\big[|F_i(x_1,\dots,x_{i-1},Z)|\big|\mathcal{G}\big]\le C_1E[(C_2+|x|^\iota+|Z|^\iota)|\mathcal{G}]$$

and

$$|G_i(x_1,\dots,x_{i-1},\omega)-G_i(y_1,\dots,y_{i-1},\omega)|\le E\big[|F_i(x_1,\dots,x_{i-1},Z)$$
$$-F_i(y_1,\dots,y_{i-1},Z)|\big|\mathcal{G}\big]\le CE\big[(1+|x|^\iota+|y|^\iota+2|Z|^\iota)\big|\mathcal{G}\big]|x-y|^\kappa.$$

□

3.3**. Remark.** Here and in what follows it is sometimes more convenient to use together with (2.6) and (2.7) also slightly different looking conditions for growth and Hölder continuity of functions we are dealing with (i.e. considering $|x|^\iota$ in place of $\sum_{j=1}^{\ell}|x_j|^\iota$, $x\in\mathbb{R}^{\ell\wp}$) but, in fact, these sets of conditions are equivalent since for any $b_1,b_2,...,b_l\ge 0$ and $\gamma>0$,

$$\max_{1\le i\le l}b_i^\gamma\le\sum_{i=1}^{l}b_i^\gamma\le l\max_{1\le i\le l}b_i^\gamma\le l(\sum_{i=1}^{l}b_i)^\gamma\le l^{1+\gamma}\max_{1\le i\le l}b_i^\gamma. \tag{3.2}$$

We will need the following result which will serve as a base for our estimates and is, in fact, an extended multidimensional version of the standard Kolmogorov theorem on the Hölder continuity of sample paths.

3.4**. Theorem.** *Let $f(x,\omega)$ be a collection of random variables defined for $x\in\mathbb{R}^d$, satisfying*

$$\|f(x,\omega)-f(y,\omega)\|_p\le C_1(1+|x|^\iota+|y|^\iota)|x-y|^\kappa\ \text{and}\ \|f(x,\omega)\|_p\le C_2(1+|x|^\iota) \tag{3.3}$$

with $\kappa>\frac{d}{p}$. Then for any $\iota'>\iota+\frac{d}{p}$ and θ such that $\kappa>\theta>\frac{d}{p}$ there is a random variable $G(\omega)$ such that

$$|f(x,\omega)|\le G(\omega)(1+|x|^{\iota'})\ \text{a.e. with}\ \|G(\omega)\|_p\le c_0[C_1+C_2]^{\frac{d}{p\theta}}C_2^{1-\frac{d}{p\theta}} \tag{3.4}$$

where $c_0=c_0(d,p,\kappa,\theta,\iota,\iota')>0$ depends only on parameters in brackets. Since $\kappa\le 1$ and $p\kappa>d$, it follows that $p>d$, and therefore we can always take $\iota'=\iota+1$. Furthermore, if $Z\in L_m(P)$ is a random variable with values in $\mathbb{R}^d$ satisfying $\|Z\|_m\le\gamma_m$ and if $\frac{1}{a}\ge\frac{1}{p}+\frac{\iota+1}{m}$ then

$$\|f(Z(\omega),\omega)\|_a\le\|G(\omega)(1+|Z|^{\iota+1})\|_a \tag{3.5}$$
$$\le c_0[C_1+C_2]^{\frac{d}{p\theta}}C_2^{1-\frac{d}{p\theta}}[1+\gamma_m^{\iota+1}]=c_0c(\gamma_m)[C_1+C_2]^{\frac{d}{p\theta}}C_2^{1-\frac{d}{p\theta}}.$$

If $p(\kappa-\delta)>d$, then we can have an almost sure Hölder estimate

$$|f(x,\omega)-f(y,\omega)|\le H(\omega)[1+|x|^{\iota+2}+|y|^{\iota+2}]|x-y|^\delta$$

with

$$\|H(\omega)\|_p\le c(\kappa,\theta,d,p,\delta,\iota)\,(C_1+C_2)$$

and the estimate

$$\|f(X_1,X_2,\dots,X_{i-1},\omega)-f(Y_1,Y_2,\dots,Y_{i-1},\omega)\|_a \le \|H(\omega)[1+|X|^{\iota+2}+|Y|^{\iota+2}]|X-Y|^\delta\|_a \le \|H\|_p(1+\gamma_m^{\iota+2})\sum_{j=1}^{i-1}\|X_j-Y_j\|_q^\delta \tag{3.6}$$

provided $\frac{1}{a}\ge\frac{1}{p}+\frac{\iota+2}{m}+\frac{\delta}{q}$ *where* $X=(X_1,...,X_{i-1})$, $Y=(Y_1,...,Y_{i-1})\in\mathbb{R}^d$ *and* X_j,Y_j, $j=1,...,i-1$ *are random vectors with* $\|X\|_m,\|Y\|_m\le\gamma_m$.

3.5**. Remark.** There are several types of constants that we need to keep track of. Constants C,K will be absolute and may change from line to line. Constants c will depend on other parameters like moments and will be denoted by $c(\cdot)$ to indicate this dependence.

Proof. For $\iota'=\iota+1>\iota+\frac{d}{p}$ set

$$\tilde f(x,\omega)=f(x,\omega)(1+|x|^{\iota+1})^{-1}.$$

Then by (3.3), if $|x-y|\le\rho_0=\frac{\sqrt{d}}{2}$,

$$\|\tilde f(x,\omega)-\tilde f(y,\omega)\|_p\le\|f(x,\omega)-f(y,\omega)\|_p(1+|x|^{\iota+1})^{-1} +\|f(y,\omega)\|_p\big||y|^{\iota+1}-|x|^{\iota+1}\big|\eta(x)\le c_1[C_1+C_2]|x-y|^\kappa\eta(x) \tag{3.7}$$

and

$$\|\tilde f(x,\omega)\|_p\le C_2\eta(x) \tag{3.8}$$

where $\eta(x)=(1+|x|^\iota)(1+|x|^{\iota+1})^{-1}$ and $c_1=c_1(\iota,\kappa,d)<\infty$ is a constant depending only on the parameters in brackets. Let $B_w(\rho)$ denotes an open unit ball of radius ρ centered at $w\in\mathbb{R}^d$. A multivariate generalization of a result of Garsia, Rodemich and Rumsey (see [22], p.60]) states that if $g:\mathbb{R}^d\to\mathbb{R}$ satisfies

$$\int_{B_w(\rho)\times B_w(\rho)}\Psi\big(\frac{|g(x)-g(y)|}{\sigma(|x-y|)}\big)dxdy\le Q_{w,\rho}$$

for some continuous strictly increasing functions Ψ,σ with $\sigma(0)=\Psi(0)=0$ then for any $x,y\in B_w(\rho)$,

$$|g(x)-g(y)|\le 8\int_0^{2|x-y|}\Psi^{-1}\big(\frac{4^{d+1}Q_{w,\rho}}{k_du^{2d}}\big)d\sigma(u) \tag{3.9}$$

where $k_d=\inf_{\substack{a\in B_w(\rho)\\0<u\le 2}}\frac{|B_a(u)\cap B_0(1)|}{u^d}$. Choose here $\Psi(z)=|z|^p$ and $\sigma(u)=u^{\theta+\frac{2d}{p}}$ with $0<\theta<\kappa-\frac{d}{p}$ and set

$$[Q_{w,\rho}(\omega)]^p=\int_{B_w(\rho)\times B_w(\rho)}\frac{|\tilde f(x,\omega)-\tilde f(y,\omega)|^p}{|x-y|^{p\theta+2d}}dxdy.$$

Then by the result above together with (3.7) we derive that there exists $c_2=c_2(\iota,\iota',\kappa,\theta,p,d)>0$ such that for any $x,y\in B_w(\rho)$,

$$|\tilde f(x,\omega)-\tilde f(y,\omega)|\le c_2\,Q_{w,\rho}(\omega)|x-y|^\theta \tag{3.10}$$

and for $0<\rho\le\rho_0$,

$$\|Q_{w,\rho}\|_p\le c_2v_d(C_1+C_2)\eta(w)\rho^{(\kappa-\theta)} \tag{3.11}$$

where

$$v_d^p = \int_{B_0(1)\times B_0(1)} |x-y|^{\kappa p - p\theta - 2d}\, dx\, dy < \infty$$

provided $p(\kappa-\theta) > d$. Observe that (3.10) and (3.11) is, in fact, the conclusion of a multidimensional version of the Kolmogorov theorem (see, for instance, [19], Theorem 1.4.1) but our argument relies also on the specific estimate (3.11).

Let $\mathbb{Z}_h^d$ be the lattice in $\mathbb{R}^d$ with spacing h. The maximum distance of any point in $\mathbb{R}^d$ from $\mathbb{Z}_h^d$ is $h\frac{\sqrt{d}}{2} = h\rho_0$. Therefore in the cube of side h centered around $w \in \mathbb{Z}_h^d$ we have

$$|\tilde{f}(x,\omega)| \le |\tilde{f}(w,\omega)| + c_2 Q_{w,h\rho_0}(\omega)\rho_0^\theta h^\theta,$$

and so

$$|\tilde{f}(x,\omega)|^p \le 2^{p-1}[|\tilde{f}(w,\omega)|^p + c_2^p Q^p_{w,h\rho_0}(\omega)\rho_0^{p\theta} h^{p\theta}]$$

Therefore,

$$\begin{aligned}\sup_{x\in\mathbb{R}^d} |\tilde{f}(x,\omega)|^p &\le 2^{p-1} \sup_{w\in\mathbb{Z}_h^d} \big[|\tilde{f}(w,\omega)|^p + c_2^p Q^p_{w,h\rho_0}(\omega)\rho_0^{p\theta} h^{p\theta}\big] \\ &\le 2^{p-1} \sum_{w\in\mathbb{Z}_h^d} \big[|\tilde{f}(w,\omega)|^p + c_2^p Q^p_{w,h\rho_0}(\omega)\rho_0^{p\theta} h^{p\theta}\big]\end{aligned}$$

and, using (3.11) together with the estimate $\sum_{w\in\mathbb{Z}_h^d}[\eta(w)]^p \le c_4^p(d,i,p)h^{-d}$,

$$\begin{aligned}E\big[[\sup\nolimits_{\mathbb{R}^d} |\tilde{f}(x,\omega)|^p]\big] &\le 2^{p-1} \textstyle\sum_{w\in\mathbb{Z}_h^d} \|\tilde{f}(w,\omega)\|_p^p \\ &\quad + 2^{p-1} c_2^p \rho_0^{p\theta} h^{p\theta} \textstyle\sum_{w\in\mathbb{Z}_h^d} \|Q_{w,h\rho_0}(\omega)\|_p^p \\ &\le c_3^p[C_2^p + (C_1+C_2)^p h^{p\kappa}] \textstyle\sum_{w\in\mathbb{Z}_h^d}[\eta(w)]^p \le c_5^p[C_2^p + (C_1+C_2)^p h^{p\kappa}]h^{-d}\end{aligned}$$

with a constant $c_5 = c_5(d,p,\iota,\kappa,\theta) > 0$. Making the choice of $h = [\frac{C_2}{C_1+C_2}]^{\frac{1}{\kappa}} \le 1$,

$$E\big[[\sup_{\mathbb{R}^d} |\tilde{f}(x,\omega)|^p]\big] \le c_6^p C_2^{p-\frac{d}{\kappa}} [C_1+C_2]^{\frac{d}{\kappa}}$$

Now set

$$\Phi(\omega) = \sup_{x\in\mathbb{R}^d} |\tilde{f}(x,\omega)|.$$

Then

$$|f(x,\omega)| \le \Phi(\omega)(1+|x|^{\iota+1}),$$

and so

$$|f(Z(\omega),\omega)| \le \Phi(\omega)|(1+|Z(\omega)|^{\iota+1}).$$

These yield (3.4) and (3.5) follows by a routine application of the Hölder inequality (see Lemma 3.1).

We now proceed to obtain a Hölder estimate on $f(x,\omega)$. If $p(\kappa-\delta) > d$ then by (3.10) and (3.11) in the same way as above for x, y in a cube of side 1,

$$|\tilde{f}(x,\omega) - \tilde{f}(y,\omega)| \le C_\delta(\omega)|x-y|^\delta$$

with $\|C_\delta(\omega)\|_p \le c(\kappa,d,\delta)(C_1+C_2)$. For such a cube D centered at z we obtain that

$$|f(x,\omega) - f(y,\omega)| \le \tilde{C}_\delta(z,\omega)|x-y|^\delta$$

with $\|\tilde{C}_\delta(z,\omega)\|_p \leq c_7(\kappa,d,\delta,\iota)(1+|z|^{\iota+1})(C_1+C_2)$. It follows that whenever $|x-y|\leq 1$,

$$|f(x,\omega)-f(y,\omega)|\leq C^*(\omega)[1+|x|^{\iota+1}+|y|^{\iota+1}]|x-y|^\delta$$

where $\|C^*\|_p \leq c_8(\kappa,d,\delta,\iota)(C_1+C_2)$. Then for some $H(\omega)=c_9(\delta,\iota)C^*(\omega)$ we obtain the global estimate

$$|f(x,\omega)-f(y,\omega)|\leq H(\omega)[1+|x|^{\iota+2-\delta}+|y|^{\iota+2-\delta}]|x-y|^\delta$$

for all x,y. In particular, by Lemma 3.1,

$$\|f(X_1,X_2,\ldots,X_{i-1},\omega)-f(Y_1,Y_2,\ldots,Y_{i-1},\omega)\|_a\leq \|H(\omega)[1+|X|^{\iota+2} \\ +|Y|^{\iota+2}]|X-Y|^\delta\|_a \leq K\,\|H\|_p(1+\gamma_m^{\iota+2})\sum_{j=1}^{i-1}\|X_j-Y_j\|_q^\delta$$

provided $\frac{1}{a}\geq\frac{1}{p}+\frac{\iota+2}{m}+\frac{\delta}{q}$. □

In our nonconventional setup Theorem 3.4 is applied usually in the form of the following useful result which in several respects improves Lemma 3.1 from [17] used in the proof of [18].

3.6. **Corollary.** *Let $\mathcal{G}$ and $\mathcal{H}$ be σ-subalgebras on a probability space $(\Omega,\mathcal{F},P)$, X and Y be d-dimensional random vectors and $f=f(x,\omega)$, $x\in\mathbb{R}^d$ be a collection of random variables that are measurable with respect to $\mathcal{H}$ and satisfy*

(3.12)
$$\|f(x,\omega)-f(y,\omega)\|_q\leq C_1(1+|x|^\iota+|y|^\iota)|x-y|^\kappa \text{ and } \|f(x,\omega)\|_q\leq C_2(1+|x|^\iota).$$

Set $\tilde{f}(x,\omega)=E[f(x,\cdot)|\mathcal{G}](\omega)$ and $g(x)=E[f(x,\omega)]$. Then, assuming that $q\geq p$, $1\geq\kappa>\theta>\frac{d}{p}$ and $\frac{1}{a}\geq\frac{1}{p}+\frac{\iota+1}{m}$, we have

$$\|\tilde{f}(X(\omega),\omega)-g(X)\|_a\leq c\,\varpi_{q,p}(\mathcal{G},\mathcal{H})(C_1+C_2)^{\frac{d}{p\theta}}C_2^{1-\frac{d}{p\theta}}(1+\|X\|_m^{\iota+1}). \tag{3.13}$$

where $c=c(\iota,\kappa,\theta,p,q,a,\delta,d)>0$ depends only on the parameters in brackets.

Next, assuming $\frac{1}{a}\geq\frac{1}{p}+\frac{\iota+2}{m}+\frac{\delta}{q}$,

$$\|E[f(X,\cdot)|\mathcal{G}]-g(X)\|_a\leq R+2c(C_1+C_2)(1+2\|X\|_m^{\iota+2})\|X-E[X|\mathcal{G}]\|_q^\delta \tag{3.14}$$

where R denotes the right hand side of (3.13).

Moreover, let $x=(v,z)$ and $X=(V,Z)$, where V and Z are d_1 and $d-d_1$-dimensional random vectors, respectively, and let $f(x,\omega)=f(v,z,\omega)$ satisfy (3.12) in $x=(v,z)$. Set $\tilde{g}(v)=E[f(v,Z(\omega),\omega)]$. Then

$$\|E[f(V,Z,\cdot)|\mathcal{G}]-\tilde{g}(V)\|_a\leq c\,(1+\|X\|_m^{\iota'}) \\ \times\big(\varpi_{q,p}(\mathcal{G},\mathcal{H})(C_1+C_2)^{\frac{d_1}{p\theta}}C_2^{1-\frac{d_1}{p\theta}}+\|V-E[V|\mathcal{G}]\|_q^\delta+\|Z-E[Z|\mathcal{H}]\|_q^\delta\big). \tag{3.15}$$

Furthermore,

$$\|\tilde{f}(X(\omega),\omega)-\tilde{f}(Y(\omega),\omega)-g(X)+g(Y)\|_a \\ \leq c\,\varpi_{q,p}(\mathcal{G},\mathcal{H})\big(1+\|X\|_m^{\iota+2}+\|Y\|_m^{\iota+2}\big)\|X-Y\|_q^\delta. \tag{3.16}$$

Proof. We start with (3.13). Set $h(x,\omega)=\tilde f(x,\omega)-g(x)$, $K_1=C_1\varpi_{q,p}(\mathcal{G},\mathcal{H})$ and $K_2=C_2\varpi_{q,p}(\mathcal{G},\mathcal{H})$. Then by (3.3) and the definition of $\varpi_{q,p}$ for all $x,y\in\mathbb{R}^d$ and $q,p\geq 1$,

$$\begin{aligned}(3.17)\quad \|h(x,\omega)-h(y,\omega)\|_p &\leq \varpi_{q,p}(\mathcal{G},\mathcal{H})\|(f(x,\omega)-f(y,\omega))-(g(x)-g(y))\|_q\\ &\leq 2K_1(1+|x|^\iota+|y|^\iota)|x-y|^\kappa\end{aligned}$$

and

$$(3.18)\qquad \|h(x,\omega)\|_p\leq \varpi_{q,p}(\mathcal{G},\mathcal{H})\|f(x,\omega)-g(x)\|_q\leq 2K_2(1+|x|^\iota).$$

These inequalities enable us to apply Theorem 3.4 to $h(x,\omega)$ (in place of $f(x,\omega)$ there) and (3.13) follows from (3.5).

In order to obtain (3.14) we note that since $1>\frac{d}{q}$ it follows that $\tilde f(x,\omega)$ has an almost surely continuous modification and taking into account that $\tilde X=E[X|\mathcal{G}]$ is $\mathcal{G}$-measurable we obtain that $E[f(\tilde X,\cdot)|\mathcal{G}]=\tilde f(\tilde X,\cdot)$. Therefore

$$\begin{aligned}(3.19)\qquad \|E\big[f(X,\cdot)|\mathcal{G}]-g(X)\|_a &\leq \|E[f(\tilde X,\cdot)|\mathcal{G}]-g(\tilde X)\|_a\\ &+\|E[f(\tilde X,\cdot)|\mathcal{G}]-E[f(X,\cdot)|\mathcal{G}])\|_a+\|g(\tilde X)-g(X)\|_a\\ &\leq \|\tilde f(\tilde X,\cdot)-g(\tilde X)\|_a+\|f(\tilde X,\cdot)-f(X,\cdot)\|_a+\|g(\tilde X)-g(X)\|_a.\end{aligned}$$

We can estimate the first term on the hand side of (3.19) by (3.13), with $\tilde X$ replacing X and noting that $\|\tilde X\|_m\leq\|X\|_m$. The second term is estimated by (3.6),

$$(3.20)\qquad \|f(\tilde X,\omega)-f(X,\omega)\|_a\leq c\,C_1(1+\gamma_m^{\iota+2})\|\tilde X-X\|_q^\delta.$$

The third term is easily estimated taking into account that by (3.12) and Lemma 3.2,

$$|g(x)-g(y)|\leq c[1+|x|^\iota+|y|^\iota]|x-y|^\kappa$$

and since $0<\delta<\kappa\leq 1$, it follows from Hölder's inequality that

$$\|g(X)-g(\tilde X)\|_a\leq c(1+\gamma_m^{\iota+2})\|\tilde X-X\|_q^\delta.$$

For (3.15) we have with $\tilde V=E[V|\mathcal{G}]$, $\tilde Z=E[Z|\mathcal{H}]$ and $\tilde{\tilde g}(v)=E[f(v,\tilde Z,\cdot)]$ and $\tilde g(v)=E[f(v,Z,\cdot)]$,

$$\begin{aligned}(3.21)\qquad \|E[f(V,Z,\cdot)|\mathcal{G}]-\tilde g(V)\|_a &\leq \|f(V,Z,\cdot)-f(V,\tilde Z,\cdot)\|_a\\ &+\|E[f(V,\tilde Z,\cdot)|\mathcal{G}]-\tilde{\tilde g}(V)\|_a+\|\tilde{\tilde g}(V)-\tilde g(V)\|_a.\end{aligned}$$

The first term in the right hand side of (3.21) is estimated by (3.6) similar to (3.20). Observe that $f(v,\tilde Z,\cdot)$ is $\mathcal{H}$-measurable, and so we can estimate the second term in the right hand side of (3.21) by (3.14) with V, d_1, $\tilde f(v,\omega)$ and $\tilde{\tilde g}(v)$ in place of X, d, $f(x,\omega)$ and $g(x)$, respectively. The third term in the right hand side of (3.21) is estimated by first using (3.12) to obtain

$$|\tilde{\tilde g}(v)-\tilde g(v)|\leq E[|f(v,\tilde Z,\cdot)-f(v,Z,\cdot)|]\leq E[(1+|v|^\iota+|Z|^\iota+|\tilde Z|^\iota)|Z-\tilde Z|^\kappa]$$

and then substituting V in place of v there. Finally, by (3.17) and (3.18) we can apply (3.6) which yields (3.16). □

3.7. **Remark.** We will always work with a,m,p,δ,q that satisfy $p(\kappa-\delta)>d=(\ell-1)\wp$ and

$$(3.22)\qquad \frac{1}{a}\geq\frac{1}{p}+\frac{\iota+2}{m}+\frac{\delta}{q}$$

Note that $m \geq \frac{ap(\iota+2)}{p-a}$ and $m \geq \frac{aq(\iota+2)}{q-a\delta}$.

Next, we recollect some notations relevant to what follows.

$$
\begin{aligned}
F_{i,n,r} &= F_{i,n,r}(x_1, x_2, \dots, x_{i-1}, \omega) \\
&= E[F(x_1, x_2, \dots, x_{i-1}, X(n))|\mathcal{F}_{n-r,n+r}], \\
F^l_{i,n,r} &= F^l_{i,n,r}(x_1, x_2, \dots, x_{i-1}, \omega) \\
&= E[F_{i,n,r}(x_1, x_2, \dots, x_{i-1}, X(n))|\mathcal{F}_{0,l}], \\
F^l_{i,n,r,r'} &= F^l_{i,n,r} - F^l_{i,n,r'}, \quad X_r(n) = E[X(n)|\mathcal{F}_{n-r,n+r}], \\
\eta^l_{i,q_i(n),r} &= F^l_{i,r,q_i(n)}(X_r(q_1(n)), \dots, X_r(q_{i-1}(n)), \omega), \\
Y_{i,q_i(n)} &= F_i(X(q_1(n)), \dots, X(q_i(n))) \quad \text{and} \\
Y_{i,m} &= 0 \quad \text{if} \quad m \neq q_i(n) \quad \text{for any } n, \\
Y_{i,q_i(n),r} &= F_{i,q_i(n),r}(X_r(q_1(n)), \dots, X_r(q_{i-1}(n)), \omega) \quad \text{and} \\
Y_{i,m,r} &= 0 \quad \text{if} \quad m \neq q_i(n) \quad \text{for any } n.
\end{aligned} \tag{3.23}
$$

In particular if $l \geq q_{(i-1)}(n) + r$,

$$
E[Y_{i,q_i(n),r}|\mathcal{F}_{-\infty,l}] = \eta^l_{i,q_i(n),r}. \tag{3.24}
$$

In order to exploit our mixing conditions it is important to approximate the random variables $X(n)$ by $X_r(n)$ and random variables $Y_{i,n}$ by $Y_{i,n,r}$. In order to justify the approximations, we need to estimate the following quantities

$$
\|Y_{i,n,r}\|_a,\ \|E[Y_{i,n,r}|\mathcal{F}_{-\infty,l}]\|_a,\ \|Y_{i,n,r} - Y_{i,n,r'}\|_a,\ \text{and}\ \|E[Y_{i,n,r} - Y_{i,n,r'}|\mathcal{F}_{-\infty,l}]\|_a \tag{3.25}
$$

which can be done either directly via Theorem 3.4 or relying on Corollary 3.6 but in order to apply either of them we will need the following estimates.

3.8**. Lemma.** *For $x = (x_1, \dots, x_{i-1})$, $y = (y_1, \dots, y_{i-1})$ and any choice of p, q and $l + r \leq n$,*

$$
\|F_{i,n,r}(x, \omega)\|_p \leq K(1 + \|x\|^\iota + \gamma^\iota_{p\iota}), \tag{3.26}
$$

$$
\|F_{i,n,r}(x, \omega) - F_{i,n,r}(y, \omega)\|_p \leq K(1 + \|x\|^\iota + \|y\|^\iota + \gamma^\iota_{p\iota})\|x - y\|^\kappa, \tag{3.27}
$$

$$
\|F^l_{i,n,r}(x, \omega)\|_p \leq K(1 + \|x\|^\iota + \gamma^\iota_{q\iota})\varpi_{q,p}(n - l - r), \tag{3.28}
$$

and

$$
\begin{aligned}
&\|F^l_{i,n,r}(x, \omega) - F^l_{i,n,r}(y, \omega)\|_p \\
&\leq K(1 + \|x\|^\iota + \|y\|^\iota + \gamma^\iota_{q\iota})\|x - y\|^\kappa \varpi_{q,p}(n - l - r).
\end{aligned} \tag{3.29}
$$

Proof. These estimates follow immediately from the definition (2.2) of the mixing rates $\varpi_{q,p}$ and Lemma 3.2. □

3.9**. Lemma.** *Let p, a, m and q be related as in Remark 3.7. Then for any $n \geq r$,*

$$
\|Y_{i,n,r}\|_a \leq Kc(d, p, \kappa, \theta, \iota, \iota')c(\gamma_m).
$$

Proof. We combine (3.5) with (3.26) and (3.27) to derive the above estimate. □

We can now apply Theorem 3.4 to conclude that

3.10**. Lemma.** *Let p, a, m and q be related as in Remark 3.7. If $q \geq 1$ and $q_i(n) \geq r + l$ then*
$$\|\eta^l_{i,q_i(n),r}\|_a \leq c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota})\varpi_{q,p}(q_i(n)-r-l).$$

Proof. We apply Theorem 3.4 relying on the mixing estimates (3.28) and (3.29) to conclude that
$$|F^l_{i,q_i(n),r}(x,\omega)| \leq C(\omega)(1+|x|^{\iota+1}) \text{ a.e.}$$
with $\|C(\omega)\|_p \leq c(\gamma_m,\gamma_{q\iota})\,\varpi_{q,p}(q_i(n)-r-l)$ and then use Lemma 3.1 to complete the proof. □

3.11**. Lemma.** *Let $n \geq l + r \vee r'$ and suppose that $\alpha < 1$, $\alpha \leq \kappa$ and $2\alpha < q$. Then with $\delta = \alpha(1-\frac{d}{p\kappa})$ and $\frac{1}{a} \geq \frac{1}{p} + \frac{\iota+2}{m} + \frac{\delta}{q}$,*
$$\|\eta^l_{i,q_i(n),r} - \eta^l_{i,q_i(n),r'}\|_a \leq c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{2q\iota},\gamma_{\frac{2q(\iota+1)}{q-2\alpha}})$$
$$\times\varpi_{q,p}(q_i(n)-r\vee r'-l) \times \big[[\beta(q,r)]^\delta + [\beta(q,r')]^\delta\big].$$

Proof. We reduce the problem to the estimation of $\|Z_1\|_a + \|Z_2\|_a$ where
$$Z_1 = F^l_{i,q_i(n),r'}(X_{r'}(q_1(n)),\dots,X_{r'}(q_{i-1}(n)),\omega)$$
$$- F^l_{i,q_i(n0,r}(X_{r'}(q_1(n)),\dots,X_{r'}(q_{i-1}(n)),\omega)$$
and
$$Z_2 = F^l_{i,q_i(n),r}(X_{r'}(q_1(n)),\dots,X_{r'}(q_{i-1}(n)),\omega)$$
$$- F^l_{i,q_i(n),r}(X_r(q_1(n)),\dots,X_r(q_{i-1}(n)),\omega).$$
For the first term Z_1 we note that by the mixing property,
$$\|F^l_{i,q_i(n),r,r'}(x,\omega)\|_p \leq \varpi_{q,p}(q_i(n)-r\vee r'-l)\,\|F_{i,q_i(n),r,r'}(x,\omega)\|_q$$
where $x = (x_1,...,x_{i-1})$, and by the triangle inequality
$$\|F_{i,q_i(n),r,r'}(x,\omega)\|_q \leq \|F_{i,q_i(n),r}(x,\omega) - F_i(x,X(q_i(n))\|_q$$
$$+ \|F_{i,q_i(n),r'}(x,\omega) - F_i(x,X(q_i(n)))\|_q.$$
If Y is $\mathcal{G}$ measurable then for any $q \geq 1$,
$$\|[X - E[X|\mathcal{G}]\|_q = \|X - Y - E[X|\mathcal{G}] + Y\|_q$$
$$\leq \|X-Y\|_q + \|E[X-Y|\mathcal{G}]\|_q \leq 2\|X-Y\|_q.$$
Therefore from the Hölder condition on F and Lemma 3.1,
$$\|F_{i,q_i(n),r}(x,\omega) - F_i(x,X(q_i(n)))\|_q \leq 2\|F_i(x,X_r(q_i(n)))$$
$$-F_i(x,X(q_i(n))\|_q \leq \|K[1+|x|^\iota + |X_r(q_i(n))|^\iota + |X(q_i(n))|^\iota]$$
$$\times|X_r(q_i(n)) - X(q_i(n))|^\kappa\|_q \leq K(1+\gamma_{\frac{q(\iota+\kappa-\alpha)}{1-\alpha}})[\beta(q,r)]^\alpha.$$

Finally,
$$\|F^l_{i,q_i(n),r,r'}(x,\omega)\|_p \leq K(1+\|x\|^\iota + c(\gamma_{2q\iota},\gamma_{\frac{q(\iota+\kappa-\alpha)}{1-\alpha}})$$
$$\times\varpi_{q,p}(q_i(n)-r\vee r'-l)\big[[\beta(q,r)]^\alpha + \beta(q,r')]^\alpha\big]$$
and
$$\|F^l_{i,q_i(n),r,r'}(x,\omega) - F^l_{i,q_i(n),r,r}(y,\omega)\|_p$$
$$\leq K(1+\|x\|^\iota + \|y\|^\iota + \gamma^\iota_{q\iota})\varpi_{q,p}(q_i(n)-r\vee r'-l)\|x-y\|^\kappa.$$

Observing that $\kappa-\alpha\leq 1$ we can now apply Theorem 3.4 with

$$C_1=c(\gamma_{q\iota})\varpi_{q,p}(q_i(n)-r\vee r'-l)\ \text{ and}$$
$$C_2=c(\gamma_{2q\iota},\gamma_{\frac{q(i+1)}{1-\alpha}})\varpi_{q,p}(q_i(n)-r\vee r'-l)\big[\beta(q,r)]^\alpha]+[\beta(q,r')]^\alpha\big]$$

to conclude that

$$\|Z_1\|_a\leq c(\gamma_{2q\iota},\gamma_m,\gamma_{\frac{q(i+1)}{1-\alpha}})\varpi_{q,p}(q_i(n)-r\vee r'-l)\big[[\beta(q,r)]^\alpha]+[\beta(q,r')]^\alpha]\big]^{1-\frac{d}{p\kappa}}$$

provided $\frac{1}{a}\geq\frac{1}{p}+\frac{\iota+1}{m}$. Now we turn to Z_2 obtaining similarly to the above that

$$\begin{aligned}\|F^l_{i,q_i(n),r}(x,\omega)-F^l_{i,q_i(n),r}(y,\omega)\|_p&\\ \leq\varpi_{q,p}(q_i(n)-r-l)\|F_{i,q_i(n),r}(x,\omega)-F_{i,q_i(n),r}(y,\omega)\|_q&\\ \leq K\varpi_{q,p}(q_i(n)-r-l)(1+\|x\|^\iota+\|y\|^\iota+\gamma_{q\iota})|x-y|^\kappa&\end{aligned}$$

and

$$\|F^l_{i,q_i(n),r}(x,\omega)\|_p\leq K\varpi_{q,p}(q_i(n)-r-l)(1+\|x\|^\iota+\gamma_{q\iota}).$$

In view of (3.1) of Theorem 3.4, if $\frac{1}{a}\geq\frac{1}{p}+\frac{\iota+2}{m}+\frac{\delta}{q}$, $p(\kappa-\delta)>d$ then

$$\begin{aligned}\|Z_2\|_a=&\|F^l_{i,q_i(n),r}(X_r(q_1(n)),\dots,X_r(q_{i-1}(n)),\omega)\\ &-F^l_{i,q_i(n),r}(X(q_1(n)),\dots,X(q_{i-1}(n)),\omega)\|_a\\ \leq& K\varpi_{q,p}(q_i(n)-r\vee r'-l)c(\gamma_m,\gamma_{2q\iota})[\beta(q,r)]^\delta.\end{aligned}$$

Therefore with the choice of $\delta=\alpha(1-\frac{d}{p\kappa})$,

$$\begin{aligned}&\|Z_1\|_a+\|Z_2\|_a\\ \leq& c(\gamma_m,\gamma_{2q\iota},\gamma_{\frac{q(i+1)}{1-\alpha}})\,\varpi_{q,p}(q_i(n)-r\vee r'-l)\,\big[[\beta(q,r)]^\delta+\beta(q,r')]^\delta\big].\end{aligned}$$

□

3.12**. Lemma.** *Let $\delta=\alpha(1-\frac{d}{p\kappa})<1-\frac{d}{p\kappa}$ and α, a satisfy conditions of Lemma 3.11. Then*

(3.30) $$\|Y_{i,q_i(n)}-Y_{i,q_i(n),r}\|_a\leq c(\gamma_{\frac{aq(\iota+1)}{q-a\delta}},\gamma_{\frac{pa(\iota+1)}{p-a}})[\beta(q,r)]^\delta.$$

Proof. We decompose the difference into two pieces,

$$\begin{aligned}&\|Y_{i,q_i(n)}-Y_{i,q_i(n),r}\|_a\leq\|Y_{i,q_i(n)}-F_i(X_r(q_1(n)),\dots,X(q_i(n))\|_a\\ &+\|F_i(X_r(q_i(n)),\dots,X(q_i(n)))-Y_{i,q_i(n),r}\|_a=T_1+T_2.\end{aligned}$$

From the Hölder continuity of F_i we obtain

$$T_1\leq C\|[1+\sum_{j=1}^i|X(q_j(n))|^\iota+\sum_{j=1}^i|X_r(q_j(n))|^\iota]\sum_{j=1}^i|X(q_j(n))-X_r(q_j(n))|^\kappa\|_a.$$

We can now use Lemma 3.1 to conclude that for $\delta\leq\kappa$ and $a\delta<q$,

$$T_1\leq c(\gamma_{\frac{aq(\iota+1)}{q-a\delta}})[\beta(q,r)]^\delta.$$

On the other hand,

$$T_2=\|G_{i,q_i(n),r}(X_r(q_1(n)),...,X_r(q_{i-1}(n),\omega))\|_a$$

where

$$G_{i,m,r}(x,\omega)=F_i(x,X(m))-E[F_i(x,X(m))|\mathcal{F}_{m-r,m+r}].$$

Now,

$$\|G_{i,q_i(n),r}(x,\omega)\|_p \le 2\|F_i(x,X(q_i(n)) - F_i(x,X_r(q_i(n)))\|_p$$
$$\le 2K\|(1+|x|^\iota + |X(q_i(n))|^\iota + |X_r(q_i(n))|^\iota)|X(q_i(n)) - X_r(q_i(n))|^\kappa\|_p$$
$$\le K(1+c(\gamma_{\frac{pq(\iota+1)}{q-p\alpha}}) + |x|^\iota)[\beta(q,r)]^\alpha$$

provided $\alpha \le \kappa$ and $q > p\alpha$. On the other hand,

$$\|G_{i,n,r}(x,\omega) - G_{i,n,r}(y,\omega)\|_p \le K(1+|x|^\iota + |y|^\iota + \gamma_{p\iota}^\iota)|x-y|^\kappa.$$

From Theorem 3.4, if $p\kappa > d$, this provides the estimate

$$T_2 = \|G_{i,q_i(n),r}(X_r(q_1(n)),\dots,X_r(q_{i-1}(n)),\omega))\|_a$$
$$\le c(\gamma_{\frac{pa(\iota+1)}{p-a}}, \gamma_{\frac{aq(\iota+1)}{q-a\delta}})[\beta(q,r)]^{\alpha(1-\frac{d}{p\kappa})}$$
$$= c(\gamma_{\frac{pa(\iota+1)}{p-a}}, \gamma_{\frac{aq(\iota+1)}{q-a\delta}})[\beta(q,r)]^\delta.$$

□

4. Limiting covariances

In this section we will study the asymptotical behavior of covariance

$$D_{i,j}(N,s,t) = E[\xi_{i,N}(s)\xi_{j,N}(t)] = \frac{1}{N}\sum_{1\le n\le Ns}\sum_{1\le l\le Nt} E[Y_{i,q_i(n)}Y_{j,q_j(l)}]$$

of the processes $\{\xi_{i,N}(t)\}$ defined by (2.21) and (2.22) where $Y_{i,q_i(n)}$ is given by (3.23). We will show that the limits

$$D_{i,j}(s,t) = \lim_{N\to\infty} D_{i,j}(N,s,t)$$

exist and $D_{i,j}(s,t) = \min(s,t)\, D_{i,j}$ where the matrix $\{D_{i,j}\}$ is determined by the results below.

4.1**. Proposition.** *For any $i,j = 1,2,...,k$ and $s,t>0$ the limit*

$$\lim_{N\to\infty} E\big[\xi_{i,N}(s)\xi_{j,N}(t)\big] =$$
$$\lim_{N\to\infty}\frac{1}{N}\sum_{\substack{0\le in\le Ns\\ 0\le jl\le Nt}} E\big[F_i(X(n),X(2n),\dots,X(in))F_j(X(l),X(2l),\dots,X(jl))\big]$$

exists and equals $D_{i,j}\min(s,t)$ which is calculated as follows. Let υ be the greatest common divisor of i and j with $i = \upsilon i'$, $j = \upsilon j'$ and i', j' being coprime. Set

$$A_{i,j}(x_{i'}, x_{2i'},\dots,x_{\upsilon i'}, y_{j'}, y_{2j'}\dots,y_{\upsilon j'}) = \int F_i(x_1,\dots,x_{i-1},x_i)$$
$$\times F_j(y_1,\dots,y_{j-1},y_j)\prod_{\sigma\notin\{i',2i',\dots,\upsilon i'\}} d\mu(x_\sigma)\prod_{\sigma'\notin\{j',2j',\dots,\upsilon j'\}} d\mu(y_{\sigma'})$$

and

$$a_{i,j}(n_1,n_2,\dots,n_\upsilon) = \int A_{i,j}(x_1,\dots,x_\upsilon,y_1,\dots,y_\upsilon)\prod_{\sigma=1}^{\upsilon} d\mu_{n_\sigma}(x_\sigma,y_\sigma). \tag{4.1}$$

Then

$$D_{i,j} = \frac{\upsilon}{ij}\sum_{u=-\infty}^{\infty} a_{i,j}(u,2u,\dots,\upsilon u)$$

where

$$a_{i,j}(0,0,\dots,0)=\int A_{i,j}(x_1,\dots,x_\upsilon,x_1,\dots,x_\upsilon)\prod_{\sigma=1}^{\upsilon}d\mu(x_\sigma)$$

and the series for $D_{i,j}$ converges absolutely.

This is essentially a straightforward but long computation carried out in a few steps, each one formulated as a lemma. We will first derive some uniform bounds on $D_{i,i}(N,t,t)$ A key step is to get for any pair i,j an estimate on

$$b_{i,j}(n,l)=E[Y_{i,q_i(n)}Y_{j,q_j(l)}].$$

If $|n-l|>>1$ then either $q_i(n)$ or $q_j(l)$ will be much bigger than all other $q_i(m)$ and $q_j(m)$ which together with the mean 0 condition on F_i,F_j and estimates of Section 3 will make then this expectation small as shown in the following result which will be used also later on.

4.2**. Lemma.** *There exists a nonincreasing function $h(m)\geq 0$, with $\sum_{m=1}^\infty h(m)<\infty$, such that for any $i,j=1,2,...,\ell$,*

$$\sup_{n,l:\, s_{i,j}(n,l)\geq m}|b_{i,j}(n,l)|\leq h(m) \tag{4.2}$$

where $s_{i,j}(n,l)=\max(\hat s_{i,j}(n,l),\hat s_{j,i}(l,n))$ and $\hat s_{i,j}(n,l)=\min(q_i(n)-q_j(l),\,n)$. Furthermore, there exists a constant $C>0$ such that for all $t\geq s\geq 0$ and $i=1,...,\ell$,

$$\sup_{N\geq 1}E|\xi_{i,N}(t)-\xi_{i,N}(s)|^2\leq C\,(t-s). \tag{4.3}$$

Proof. First, observe that for $i=1,...,k$,

$$q_i(n)-q_{i-1}(n)=n\quad\text{and}\quad s_{i,i}(n,l)=\min(i|n-l|,\max(n,l))\geq|n-l| \tag{4.4}$$

where in the first equality we set $q_0(n)=0$. On the other hand, if $i\geq k+1$ then it follows from (2.10)–(2.12) that for any $\varepsilon>0$ there exists n_ε such that for all $n\geq n_\varepsilon$ and $n>l\geq 0$,

$$q_i(n)-q_{i-1}(n)\geq n+\varepsilon^{-1},\ q_i(n)-q_i(l)\geq n-l+\varepsilon^{-1}, \tag{4.5}$$

and so

$$s_{i,i}(n,l)\geq\min(n-l+\varepsilon^{-1},n)\geq n-l. \tag{4.6}$$

Now, assume that $q_i(n)-q_j(l)\geq 0$ and $n\geq n_1$ so that we will use here (4.4)–(4.6) with $\varepsilon=1$ while only in Proposition 4.5 these estimates will be needed for all positive ε. Set $r=\frac13 s_{i,j}(n,l)=\frac13\hat s_{i,j}(n,l)$. If we replace $Y_{i,q_i(n)}$ and $Y_{j,q_j(l)}$ by $Y_{i,q_i(n),r}$ and $Y_{j,q_j(l),r}$ defined in (3.23) then the difference between $b_{i,j}(n,l)$ and

$$b^{(r)}_{i,j}(n,l)=E[Y_{i,q_i(n),r}Y_{j,q_j(l),r}]$$

can be estimated easily using the approximation estimate of Lemma 3.12 which gives

$$|b^{(r)}_{i,j}(n,l)-b_{i,j}(n,l)|\leq c(\gamma_m,\gamma_{\frac{2p(\iota+1)}{2-p\alpha}})[\beta(q,r)]^\delta.$$

On the other hand, by (4.4) and (4.5) we see that in our circumstances $\min(q_i(n)-q_j(l), q_i(n)-q_{i-1}(n)) \geq \hat{s}_{i,j}(n,l)$, and so by Lemma 3.10,

$$|b^{(r)}_{i,j}(n,l)| = |EY_{i,q_i(n),r}Y_{j,q_j(l),r}|$$
$$= \big|E\big[E\big[Y_{i,q_i(n),r}|\mathcal{F}_{0,q_i(n)-r}\big]Y_{j,q_j(l),r}\big]\big| \leq \|F_j(X_r(q_1(l)),\dots,X_r(q_j(l)))\|_{L_2(P)}$$
$$\times\|E[Y_{i,q_i(n),r}|\mathcal{F}_{0,q_i(n)-r}]\|_{L_2(P)} \leq C\varpi_{q,p}(\tfrac{1}{3}s_{i,j}(n,l)).$$

We can always estimate $|b_{i,j}(n,l)|$ by $|b^{(r)}_{i,j}(n,l)-b_{i,j}(n,l)|+|b^{(r)}_{i,j}(n,l)|$, so that

$$|b_{i,j}(n,l)| \leq C\big(\varpi_{q,p}(\frac{1}{3}s_{i,j}(n,l)) + [\beta(q,\frac{1}{3}s_{i,j}(n,l)))^{\delta}\big).$$

Now, observe that if $n<n_1$ and $q_i(n)-q_j(l)\geq 0$ then

$$s_{i,j}(n,l) \leq L_1 = \max_{n<n_1,\, i\leq \ell} q_i(n) \text{ and } l \leq n_1+L_1.$$

Hence, in order to satisfy (4.2) we can take

$$h(m) = \max_{0\leq n,l\leq n_1+L_1,\, 1\leq i,j\leq \ell} |b_{i,j}(n,l)|$$

for $m\leq L_1$ while for $m>L_1$ we define

$$h(m) = C\big(\varpi_{q,p}([\frac{1}{3}m]) + (\beta(q,[\frac{1}{3}m]))^{\delta}\big).$$

Finally, by (4.4) and (4.6) for $t\geq s\geq 0$,

$$E[|\xi_{i,N}(t)-\xi_{i,N}(s)|^2] \leq \tfrac{1}{N}\big(\textstyle\sum_{Ns\leq l\leq Nt} b_{i,i}(l,l) + 2\sum_{\substack{Ns\leq l\leq Nt\\ n\geq l+1}} |b_{i,i}(n,l)|\big)$$
$$\leq \tfrac{1}{N}\textstyle\sum_{Ns\leq l\leq Nt}\big(EY^2_{i,l} + 2\sum_{n\geq l+1} h(n-l)\big) \leq Ct$$

provided $N(t-s)\geq 1$ and the result follows. □

Next, we will need a result which will be formulated in a somewhat more general situation. Let $H(x_1,x_2,\dots,x_d)$ be a function on $(R^\nu)^d$ that is continuous and satisfies the growth condition $|H(x_1,x_2,\dots,x_d)| \leq 1+\sum_i \|x_i\|^\iota$ for some $\iota\geq 1$. Suppose that $\{Y(n): n\geq 1\}$ is a stochastic process with values in R^ν and there exists an integer $m\geq 1$ such that for any $l\leq m$ the distribution of $\{Y(n_1),Y(n_2),\cdots Y(n_l)\}$ depends only on the spacings $\{n_i-n_{i-1}\}$, $l=2,...,l$ between them. For $l\geq 2$, we denote this distribution by μ_S where S is a set of $l-1$ positive integers prescribing the spacings between the l integers. We assume that all $\{Y(n),\, n\geq 1\}$ have a common distribution μ and that the integrability condition $\int \|x\|^\iota d\mu < \infty$ holds true. For some $p,q\geq 1$ and a nested family of sub σ-fields $\mathcal{F}_{m,n}$ as above assume the mixing condition

$$\varpi_{q,p}(l) = \sup_{m-n\geq l} \varpi_{q,p}(\mathcal{F}_{-\infty,m},\mathcal{F}_{n,\infty}) \to 0 \text{ as } l\to\infty,$$

and the localization condition

$$\lim_{r\to\infty}\sup_n \|Y(n)-E[Y(n)|\mathcal{F}_{n-r,n+r}]\|_{L_1(P)} = 0.$$

Let $n_1<n_2<\ldots<n_d$ be a sequence of integers that tend to ∞ with some of gaps $\{n_{i+1}-n_i\}$ tending to infinity while others are kept fixed. This splits the set of integers $1,2,\dots,d$ into a partition $\mathcal{P}$ consisting of blocks B_j of different sizes. The pairwise distances between integers in each block B_j remain fixed (so it can be viewed as rigid) while the distances between different blocks tend to ∞. We assume that each block B_j consists of at most m integers. Let m_j denotes the number of

integers in a block B_j and S_j denotes the set of spacings in B_j, i.e. sequence of $m_j - 1$ positive integers representing pairwise distances between successive integers in S_j. Let the distribution $\mu_{\mathcal{P}}$ on $(R^l)^d$ be the product measure

$$\mu_{\mathcal{P}} = \Pi_j \mu_{S_j}$$

over successive blocks.

4.3**. Lemma.** *Assume that $\{n_j\}$ goes to infinity with rigid blocks determined by $\mathcal{P}$. Then*

$$\lim_{n_1,\dots,n_d\to\infty} E[H(X(n_1),\dots,X(n_d))] = \int H(x_1,\dots,x_d)d\mu_{\mathcal{P}}$$

where the limit is taken so that the sets S_j of spacings in each block B_j remain fixed while the gaps between different blocks tend to infinity.

Proof. First we note that because of the growth and integrability conditions we can replace H by $H\phi$ where ϕ is a continuous cut off function with compact support. The error is uniformly controlled on either side. We can then approximate H uniformly by a smooth function. In other words we can assume without loss of generality that H is a bounded continuous function supported on some ball of radius L with a bounded gradient. We prove the lemma by reducing the number of blocks by one at each step. The last gap that tends to ∞ cuts off a block $B = \{n_{d'+1},\dots,n_d\}$ at the end with a rigid spacing S between integers in the block. We will show that

$$\lim_{\substack{n_1,\dots,n_d\to\infty\\ \mathcal{P}\ \text{fixed}}} E[\widehat{H}(X(n_1),\dots,X(n_d))] = 0 \tag{4.7}$$

where

$$\widehat{H}(x_1,x_2,\dots,x_d) = H(x_1,x_2,\dots,x_d)$$
$$-\int H(x_1,x_2,\dots,x_{d'},x_{d'+1},\dots,x_d)d\mu_S(x_{d'+1},\dots,x_d).$$

This will reduce the number of blocks by one, replacing H by

$$H_1(x_1,x_2,\dots,x_{d'}) = \int H(x_1,x_2,\dots,x_{d'},x_{d'+1},\dots,x_d)d\mu_S(x_{d'+1},\dots,x_d).$$

The step by step reduction will end when only the first block B_1 with spacings S_1 remains and since it is rigid we can integrate it out with μ_{S_1} and end up with $\int H(x_1,\dots,x_d)d\mu_{\mathcal{P}}$ which will complete the proof of the lemma.

The function $\widehat{H}$ is also bounded with a bounded gradient. Therefore,

$$\|\widehat{H}(X(n_1),\dots,X(n_d)) - \widehat{H}(X_r(n_1),\dots,X_r(n_d))\|$$
$$\le C\sup_n \|X_r(n) - X(n)\|_{L_1(P)} \to 0$$

uniformly over all $n_1,\dots,n_d$ as $r\to\infty$. To establish (4.7), it is therefore sufficient to prove that

$$\lim_{r\to\infty}\limsup_{n_1,\dots n_d\to\infty} E[\widehat{H}(X_r(n_1),\dots,X_r(n_d))] = 0. \tag{4.8}$$

Observe that

$$E[\widehat{H}(X_r(n_1),\dots,X_r(n_d))] = E[E[\widehat{H}(X_r(n_1),\dots,X_r(n_d))|\mathcal{F}_{-\infty,n_{d'}+r}]]$$
$$= E[G_r(X_r(n_1),\dots,X_r(n_{d'}),\omega)]$$

where

$$G_r(x_1,\dots,x_{d'},\omega)=E[\widehat{H}(x_1,\dots,x_{d'},X_r(n_{d'+1}),...,X_r(n_d))|\mathcal{F}_{-\infty,n_{d'}+r}].$$

To prove (4.8) is clearly sufficient to show that

$$\lim_{r\to\infty}E[\sup_{x_1,\dots,x_{d'}}|G_r(x_1,\dots,x_{d'},\omega)|]=0.$$

Since $\|\nabla_x G_r\|_\infty\le\|\nabla_x\widehat{H}\|_\infty\le\|\nabla_x H\|_\infty$, there is a uniform bound on $\|\nabla G_r\|$. We can therefore estimate

$$\sup_{x_1,\dots,x_{d'}}|G_r(x_1,\dots,x_{d'},\omega)|\le C\int|G_r(x_1,\dots,x_{d'},\omega)|dx_1\cdots dx_{d'}.$$

Taking expectations and observing that G_r vanishes outside a ball of radius L,

$$E\sup_{x_1,\dots,x_{d'}}|G_r(x_1,\dots,x_{d'},\omega)|\le C\,L^{d'}\sup_{x_1,\dots,x_{d'}}E|G_r(x_1,\dots,x_{d'},\omega)|.$$

If $n_{d'+1}-n_{d'}>2r$ then by the definition (2.1) of the dependence coefficients ϖ,

$$\sup_{x_1,\dots,x_{d'}}\|G_r(x_1,\dots,x_{d'},\omega)-\widehat{H}_r(x_1,\dots,x_{d'})\|_1\le 2\,\varpi_{\infty,1}(n_{d'+1}-n_{d'}-2r)\|H\|_\infty$$

where

$$\widehat{H}_r(x_1,\dots,x_{d'})=E[\widehat{H}(x_1,\dots,x_{d'},X_r(n_{d'+1}),\dots,X_r(n_d))]$$

while

$$\widehat{H}(x_1,\dots,x_{d'})=E[\widehat{H}(x_1,\dots,x_{d'},X(n_{d'+1}),\dots,X(n_d))]\equiv 0.$$

Since $\widehat{H}$ has a bounded gradient,

$$\big|E[\widehat{H}(x_1,\dots,x_{d'},X(n_{d'+1}),\dots,X(n_d))]-E[\widehat{H}(x_1,\dots,x_{d'},$$
$$X_r(n_{d'+1}),\dots,X_r(n_d))]\big|\le C\sup_n E|X(n)-X_r(n)|=\epsilon(r)\to 0\text{ as }r\to\infty.$$

Taking into account that $\varpi_{\infty,1}(l)\le\varpi_{p,q}(l)\to 0$ the lemma follows from the above estimates. □

4.4. **Lemma.** *For any $i,j\le k$ and $s,t>0$ and integer u, the limit*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\substack{0\le in\le Ns\\0\le jl\le Nt\\in-jl=u}}b_{i,j}(n,l)=\frac{\upsilon\min(s,t)}{ij}\,c_{i,j}(u)\tag{4.9}$$

exists where υ is the greatest common divisor of i and j. For any multiple of υ,

$$c_{i,j}(\upsilon u)=a_{i,j}(u,2u,...,\upsilon u)\tag{4.10}$$

with $a_{i,j}$ defined by (4.1). If u is not a multiple of υ then $c_{i,j}(u)=0$. Furthermore,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\substack{0\le in\le Ns\\0\le jl\le Nt}}b_{i,j}(n,l)=\frac{\upsilon\min(s,t)}{ij}\sum_{-\infty<u<\infty}c_{i,j}(u)\tag{4.11}$$

amd the series in the right hand side converges absolutely.

Proof. It is clear that if u is not a multiple of υ there are no solutions of the equation $in'-jl'=u$ so we can replace u by υu. Combining the indices $n,2n,\dots,in$ and $l,2l,\dots,jl$ and ordering them into a single sequence we obtain employing Lemma 4.3 that

$$\lim_{\substack{n,l\to\infty\\in-jl=\upsilon u}}b_{i,j}(n,l)=\lim_{\substack{n,l\to\infty\\in-jl=\upsilon u}}E[F_i(X(n),X(2n),\dots,X(in))$$
$$\times F_j(X(l),X(2l),\dots,X(jl))]=a_{i,j}(u,2u,\dots,\upsilon u).$$

If υ is the greatest common divisor of i and j then $i=\upsilon\alpha$ and $j=\upsilon\beta$ with α and β being coprime. Since all the gaps in either sequence above go to ∞, we can have blocks of size more than one only by pairing two members from different sequences and therefore the rigid blocks of Lemma 4.3 can be of size one and two only. If we start with (n,l) such that $\alpha n-\beta l=u$, their multiples $(\alpha mn,\beta ml)$, $m=1,...,\upsilon$ with $\alpha mn-\beta ml=mu$ will give υ blocks of size 2. There can not be any other. Indeed, if (a,b) is a pair of integers which is not an integer multiple of (α,β) then taking into account that α and β are coprimes we conclude that $|an-bl|\to\infty$ when $n\to\infty$ preserving $\alpha n-\beta l=u$ fixed. To complete the proof of the lemma we need to count the number of integer solutions of $in-jl=\upsilon u$ or $\alpha n-\beta l=u$ with $\alpha\upsilon n\leq Nt$ and $\beta\upsilon l\leq Ns$. The set of solutions for any u is obtained by shifting the set of solutions of the homogeneous equation $\alpha n-\beta l=0$ by a fixed solution of the above nonhomogeneous one. Therefore, with our constrains their numbers can differ at most by a constant. In the homogeneous case the solutions are precisely those $m=in=jl$ that are multiples of $\upsilon\alpha\beta$. Their number is integral value of $\frac{N\min\{t,s\}}{\upsilon\alpha\beta}=\frac{N\upsilon\min\{s,t\}}{ij}$. This proves (4.9) while Lemma 4.2 and (4.9) imply (4.11). □

Finally we turn to $\xi_{i,N}(t)$ with $k+1\leq i\leq\ell$. We will see in the next section that, in fact, their limits in distribution $\{\eta_i(\cdot);i\geq k+1\}$ are mutually independent processes which are also independent of the processes $\{\eta_i(\cdot);1\leq i\leq k\}$ but here we deal only with their variances and covariances.

4.5. **Proposition.** *For $i\geq k+1$,*

$$\lim_{N\to\infty}E\big(\xi_{i,N}(s)\xi_{i,N}(t)\big)\tag{4.12}$$
$$=\min(s,t)\int\big(F_i(x_1,x_2,\dots,x_i)\big)^2d\mu(x_1)d\mu(x_2)\cdots d\mu(x_i).$$

Moreover, for any t,s and $j<i$, $i>k$,

$$\lim_{N\to\infty}E\big(\xi_{i,N}(t)\xi_{j,N}(s)\big)=0.\tag{4.13}$$

Proof. It follows from (4.6) that

$$s_{i,i}(n,l)\geq\min(|n-l|+\varepsilon^{-1},\max(n,l))\quad\text{if}\quad\max(n,l)\geq n_\varepsilon\ \text{ and }\ n\neq l,$$

and so by (4.2),

$$b_{i,i}(n,l)\to 0\ \text{ as }\ \max(n,l)\to\infty\ \text{ so that }\ |n-l|\geq 1.$$

Therefore for any fixed $L\geq n_1$,

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{1\leq n,l\leq TN,n\neq l}|b_{i,i}(n,l)|\leq 2T\sum_{m\geq L}h(m)$$
$$+\limsup_{N\to\infty}\frac{1}{N}\sum_{\substack{1\leq|n-l|\leq L\\ n,l\leq TN}}|b_{i,i}(n,l)|=2T\sum_{m\geq L}h(m).$$

We now let $L\to\infty$ and since $\sum_m h(m)<\infty$ it follows that lim sup in the left hand side above equals zero, i.e. the off-diagonal terms do not contribute in (4.12). It remains to deal with the diagonal terms $b_{i,i}(n,n)$. Since $q_j(n)-q_{j-1}(n)\to\infty$ for $j=2,3,...,\ell$ as $n\to\infty$ it follows from Lemma 4.3 that

$$\lim_{n\to\infty}b_{i,i}(n,n)=\int\big(F_i(x_1,...,x_i)\big)^2d\mu(x_1)...d\mu(x_i)\tag{4.14}$$

proving (4.12).

Next, we deal with (4.13). Relying on Lemma 4.2 we can estimate for any $\varepsilon > 0$,

$$
\begin{aligned}
(4.15)\qquad & |E\xi_{i,N}(t)\xi_{j,N}(s)| \\
&\le |E\xi_{i,N}(\varepsilon T)\xi_{j,N}(s)| + |E\big(\xi_{i,N}(t) - \xi_{i,N}(\varepsilon T)\big)\xi_{j,N}(s)| \\
&\le \big(E\xi^2_{i,N}(\varepsilon T)\big)^{1/2}\big(E\xi^2_{j,N}(s)\big)^{1/2} + \tfrac{1}{N}\textstyle\sum_{\varepsilon NT\le n\le NT, 1\le l\le NT}|b_{i,j}(n,l)| \\
&\le CT\sqrt{\varepsilon} + \tfrac{1}{N}\textstyle\sum_{\varepsilon NT\le n\le NT, 1\le l\le NT} h(s_{i,j}(n,l)).
\end{aligned}
$$

Since $i > j$ and $i > k$ then by (2.12) we can choose $N(\varepsilon) > \varepsilon^{-1}T^{-1}n_\varepsilon$ such that $q_i(n) - q_j(l) > \varepsilon^{-1}$ whenever $N \ge N(\varepsilon)$, $n \ge \varepsilon NT$, $l \le NT$ and, moreover, by (4.5),

$$
\begin{aligned}
& s_{i,j}(n,l) = \min(q_i(n) - q_j(l), n) \\
&\ge \min(q_i(n) - q_i(\varepsilon NT) + \varepsilon^{-1}, n) \ge \min(n - \varepsilon NT + \varepsilon^{-1}, n).
\end{aligned}
$$

Hence,

$$\frac{1}{N}\sum_{\varepsilon NT\le n\le NT, 1\le l\le NT} h(s_{i,j}(n,l)) \le T\sum_{m\ge\min(\varepsilon^{-1},\varepsilon NT)} h(m)$$

and letting, first, $N\to\infty$ and then $\varepsilon\to 0$ we derive (4.13) from (4.15). □

5. Proof of the main theorem

The proof of Theorem 2.2 relies on martingale approximations and martingale limit theorems but we will need several modifications in our situation. We begin with the following result which can be found in various forms in the literature (see, for instance, Section 2 in Ch. VIII of [16] and close versions in Theorem 18.2 in [2] and Theorem 4.1 in [14]). Let $\{U_{N,n} : n \ge 1\}$ be a triangular array of random variables satisfying the following conditions,

B1. $\{U_{N,n}\}$ is adapted to some $(\Omega_N, \mathcal{G}_{N,n}, P_N)$;

B2. $\{U_{N,n}\}$ are uniformly square integrable;

B3. $\|E[U_{N,m}|\mathcal{G}_{N,n}]\|_2 \le c(m-n)$ for all N, $n \le m$ and for some sequence $c(k)$ satisfying $\sum_{k=0}^\infty c(k) = C < \infty$.

B4. For some increasing function $A(t)$,

$$\lim_{N\to\infty}\|\frac{1}{N}\sum_{1\le n\le Nt} W^2_{N,n} - A(t)\|_{L_1(P)} = 0$$

where

$$W_{N,n} = U_{N,n} + \sum_{m\ge n+1} E[U_{N,m}|\mathcal{G}_{N,n}] - \sum_{m\ge n} E[U_{N,m}|\mathcal{G}_{N,n-1}].$$

Observe, that $W_{N,n}$, $n \ge 1$ is a martingale differences sequence provided **B1**–**B3** hold true.

5.1**. Theorem.** *Under assumptions* **B1-B4**,

$$\xi_N(t) = \frac{1}{\sqrt{N}}\sum_{1\le n\le Nt} U_{N,n}$$

converges in distribution on $D[[0,T];R]$ to a Gaussian process $\xi(t)$ with independent increments such that $\xi(t) - \xi(s)$ has mean 0 and variance $A(t) - A(s)$.

We need however to strengthen the theorem a little bit in our context. First we note that the condition **B4** can be replaced by the weaker condition

$$\lim_{N\to\infty}\frac{1}{N}\sum_{1\le n\le Nt}E[W_{N,n}^2]=A(t) \tag{5.1}$$

as can be seen from the following result.

5.2. **Lemma.** *If for a fixed l the random variables $V_{N,r}=(\sum_{r(l-1)+1\le n\le rl}U_{N,n})^2$ satisfy a uniform law of large numbers in the sense that*

$$\lim_{r\to\infty}\sup_N\sup_n E\big[|\frac{1}{r}\sum_{j=1}^{r}\big[V_{N,n+j}-E[V_{N,n+j}]\big]|\big]=0,$$

then (5.1) implies **B4**.

Proof. We begin with the observation that if η_n, $n\ge 1$ are martingale differences adapted to any filtration $\mathcal{G}_n$ and they are uniformly integrable, then $\frac{1}{N}\sum_{n=1}^{N}\eta_n\to 0$ in $L_1(P)$. To see this, we approximate η_n in $L_1(P)$ by $\tilde\eta_n$ that are uniformly bounded. The latter may not be a martingale difference but it can be written as $\widetilde\eta_n=\widehat\eta_n+\bar\eta_n$ with $\|\bar\eta_n\|_{L_1(P)}\le\|\eta_n-\widetilde\eta_n\|_{L_1(P)}$ and $\widehat\eta_n$ being a martingale difference with a uniformly bounded second moment. We will now compare

$$A_N(t,\omega)=\frac{1}{N}\sum_{n\le[Nt]}(\eta_n)^2$$

with block sums over $B_r=\{n: rl+1\le n\le(r+1)l\}$,

$$A_N^l(t,\omega)=\frac{1}{N}\sum_{r:B_r\subset[0,Nt]}(\sum_{n\in B_r}\eta_n)^2$$

The difference involves the cross terms

$$A_N^l(t,\omega)-A_N(t,\omega)=\frac{2}{N}\sum_{r:B_r\subset[0,Nt]}\sum_{\substack{n>m\\ n,m\in B_r}}\eta_n\eta_m.$$

It is easy to see that the sum

$$\sum_{\substack{n>m\\ n,m\in B_r}}\eta_n\eta_m$$

is a martingale difference adapted to $\mathcal{G}_{rl}$ and therefore for fixed l,

$$\lim_{N\to\infty}\|A_N^l(t,\omega)-A_N(t,\omega)\|_{L_1(P)}=0.$$

Since $E^P[A_N(t,\omega)]=E^P[A_N^l(t,\omega)]$, it follows immediately that

$$\limsup_{N\to\infty}\|A_N(t,\omega)-E^P[A_N(t,\omega)]\|_{L_1(P)}$$
$$\le\limsup_{N\to\infty}\|A_N^l(t,\omega)-E^P[A_N(t,\omega)]\|_{L_1(P)}.$$

On the other hand, $W_{N,n}=U_{N,n}-R_{N,n-1}+R_{N,n}$, where $R_{N,n}=\sum_{m\ge n+1}E[U_{N,m}|\mathcal{G}_{N,n}]$, and

$$\sum_{n\in B_r}W_{N,n}=\sum_{n\in B_r}U_{N,n}-R_{N,jl}+R_{N,(j+1)l}.$$

By our assumption the squares of the block sums $V_{N,r}=(\sum_{n\in B_r}U_{N,n})^2$ satisfy a uniform law of large numbers in $L_1(P)$. The differences between the two block sums come from the correction term and their second moments are uniformly controlled. Therefore their contribution is at most $\frac{C}{l}$. Hence,

$$\limsup_{l\to\infty}\limsup_{N\to\infty}\|A_N^l(t,\omega)-E^P[A_N(t,\omega)]\|_{L_1(P)}=0$$

and the lemma follows. □

5.3**. Remark.** Let the filtration $\mathcal{F}_{m,n}$ satisfiy any mixing condition, i.e. $\varpi_{p,q}(k)\to 0$ as $k\to\infty$. Then any collection of uniformly integrable random variables $\{f_n(\omega)\}$, with f_n being $\mathcal{F}_{n+k,n-k}$ measurable for some fixed k, are easily seen to satisfy the (centralized) law of large numbers. It is obvious for uniformly bounded $\{f_n\}$ and we can always approximate our $\{f_n\}$ uniformly in L_1 by uniformly bounded ones.

5.4**. Corollary.** *If we have a family of triangular arrays and the conditions of Theorem 5.1 are valid uniformly over the family then the limit theorem is also valid uniformly over the family.*

Proof. The proof is by a routine contradiction argument. If the family is indexed by α and the limit theorem is not valid uniformly, then for some choice α_N that depends on N the limit theorem fails to hold. But this is just another triangular array and, by the uniform validity of the assumptions, the limit theorem has to hold. □

5.5**. Remark.** One way to generate new triangular arrays for $N=1,2,...$ is to take a sequence of sub σ-fields, $\mathcal{G}_{N,k_N}$, a sequence of sets $B_N\in\mathcal{G}_{N,k_N}$ with $P_N(B_N)\geq\delta>0$ and to consider $(\Omega_N,\tilde{\mathcal{G}}_{N,n},\tilde{U}_{N,n},P_{N,B_N})$, $n=1,2,...$ where $\tilde{\mathcal{G}}_{N,n}=\mathcal{G}_{N,k_N+n}$, $\tilde{U}_{N,n}=U_{N,k_N+n}$ and the measure P_{N,B_N} is defined by

$$P_{N,B_N}(\Gamma)=\frac{P_N(\Gamma\cap B_N)}{P_N(B_N)}.$$

It is easy to see that $\tilde{U}_n$ are again martingale differences, for each fixed $\delta>0$ uniform integrability under P_{N,B_N} is inherited from the same property under P_N and the condition **B3** of Theorem 5.1 holds uniformly over this family, as well, provided $k_N\leq CN$ for some C. Otherwise, it has to be checked again. The limit $A(t)$ will of course vary depending on the behavior of $\frac{k_N}{N}$. If $\frac{k_N}{N}\to t_0$ then $A(t)$ gets replaced by $A(t+t_0)-A(t_0)$.

This observation leads to the following theorem.

5.6**. Theorem.** *Let $\mathcal{X}$ be a complete separable metric space and for each $N\geq 1$ let $F_N(\omega)$ be a $\mathcal{X}$-valued and $\mathcal{G}_{N,k_N}$-measurable random variable. Suppose that the distribution λ_N of F_N under P_N converges weakly as $N\to\infty$ to λ on $\mathcal{X}$ and $\frac{k_N}{N}\to t_0$. Let the conditions of Theorem 5.1 hold true and set*

$$\xi_{N,k_N}(t)=\frac{1}{\sqrt{N}}\sum_{k_N+1\leq n\leq k_N+Nt}U_{N,n}.$$

Then the joint distribution of the pair $(F_N,\xi_{N,k_N}(\cdot))$ converges on $\mathcal{X}\times D[0,T]$ to the product of λ and the distribution γ of a Gaussian process with independent increments having mean 0 and variance $A(t+t_0)-A(t_0)$. In particular, any limit in distribution of

$$\xi_N(t)=\frac{1}{\sqrt{N}}\sum_{1\leq n\leq Nt}U_{N,n}$$

is always a process with independent increments. We can drop the assumption that $\frac{k_N}{N} \to t_0$ *provided we can verify that for some* $A(t)$*,*

$$\lim_{N\to\infty} \|\frac{1}{N} \sum_{k_N+1\le n\le k_N+Nt} W_{N,n}^2 - A(t)\|_{L_1(P_N)} = 0$$

Proof. Since the conditions of Theorem 5.1 are satisfied here, ξ_{N,k_N} converges in distribution as $N \to \infty$ to a Gaussian process with independent increments whose distribution we denote by γ. Now, if μ_N denotes the joint distribution of F_N and $\xi_{N,k_N}(\cdot)$ the convergence of the marginals implies the tightness of μ_N. Taking a subsequence if necessary, we can assume that μ_N has a limit μ with marginals λ and γ. We need to prove that $\mu = \lambda \times \gamma$. It is enough to prove that if $E \subset \mathcal{X}$ and $F \subset D[0,T]$ are continuity sets of λ and γ, respectively, then $\mu(E\times F) = \lambda(E) \times \gamma(F)$. We can assume without loss of generality that $\lambda(E) > 0$. Set $B_N = \{\omega : F_N(\omega) \in E\}$ then $P_N(B_N) \to \lambda(E)$, and so $P_N(B_N) \ge \frac{1}{2}\lambda(E) > 0$ for N large enough. In view of Remark 5.5, $\xi_{N,k_N}(\cdot)$ converges in distribution under P_{N,B_N} as $N \to \infty$ to a Gaussian process with independent increments and since, clearly, under P_{N,B_N} we have convergence in **B4** to the same $\tilde{A}(t) = A(t+t_0)-A(t_0)$ as under P_N, it follows that the distribution of $\xi_{N,k_N}(\cdot)$ under P_{N,B_N} converges to γ. In particular, since F is a continuity set,

$$P_{N,B_N}\{\omega : \xi_{N,k_N}(\cdot) \in F\} = \frac{\mu_N(E\times F)}{P_N(B_N)} \to \gamma(F).$$

Since $E \times F$ is a continuity set of μ, this proves that $\frac{\mu(E\times F)}{\lambda(E)} = \gamma(F)$. □

5.7. **Corollary.** *Assume that we have a triangular array consisting of* $\mathcal{G}_{N,n}$*-measurable random vectors* $U_{N,n} : \Omega \to R^d$ *and that each linear combination* $\langle \lambda, U_{N,n}\rangle$ *satisfies the assumptions* **B1**-**B4**. *In particular,*

$$\lim_{N\to\infty} \|[\frac{1}{N} \sum_{1\le n\le Nt} \langle \lambda, W_{N,n}\rangle^2] - \langle \lambda, A(t)\lambda\rangle\|_{L_1(P)} = 0.$$

Then

$$\xi_N(t) = \frac{1}{\sqrt{N}} \sum_{k_N+1\le n\le k_N+Nt} U_{N,n}$$

converges in distribution on the Skorokhod space $D[[0,T]; R^d]$ *to the Gaussian process* $\eta(t)$ *with independent increments taking values in* R^d*, having mean* 0 *and covariance*

$$E[\langle \lambda(\eta(t) - \eta(s))\rangle^2] = \langle \lambda, (A(t) - A(s))\lambda\rangle.$$

Proof. By the results for the scalar case, the distribution of $\langle u, \xi_N(t)\rangle$ converges to a Gaussian process with independent increments. This implies compactness of the distributions of the vector process $\xi_N(\cdot)$. Let Q be a limit point of distributions of ξ_N and let η be the corresponding limiting vector process. By the above for each constant vector u the distribution of the increments $\langle u, \eta(t) - \eta(s)\rangle$ must be Gaussian and, therefore, by the Cramér-Wold argument, $\eta(t)-\eta(s)$ has under Q the d-dimensional Gaussian distribution with mean 0 and a covariance matrix $\{A_{i,j}(t)-A_{i,j}(s)\}$. Moreover, by Theorem 5.6, under Q the random variable $\langle u, \eta(t) - \eta(s)\rangle$ is independent of $\{\eta(\tau) : \tau \le s\}$ for every $t > s$ and $u \in R^d$. This is sufficient to determine Q as the distribution of a Gaussian process $\eta(t)$ with independent increments taking values in R^d having mean 0 and covariance

$$E[(\eta_i(t)-\eta_i(s))(\eta_j(t)-\eta_j(s))]=A_{i,j}(t)-A_{i,j}(s)$$

and establish that the distribution of

$$\xi_N(t)=\frac{1}{\sqrt{N}}\sum_{k_N+1\le n\le k_N+Nt}U_{N,n}$$

converges to Q on the Skorokhod space $D[[0,T];R^d]$. □

Next, we break the proof of Theorem 2.2 into several steps and use the following representations

$$\begin{gathered}Y_{i,q_i(n)}=Y_{i,q_i(n),1}+\sum_{r=1}^{\infty}[Y_{i,q_i(n),2^r}-Y_{i,q_i(n),2^{r-1}}],\\ \zeta_{i,N,0}(t)=\tfrac{1}{\sqrt{N}}\sum_{1\le n\le M_i(Nt)}Y_{i,q_i(n),1},\\ \zeta_{i,N,r}(t)=\tfrac{1}{\sqrt{N}}\sum_{1\le n\le M_i(Nt)}[Y_{i,q_i(n),2^r}-Y_{i,q_i(n),2^{r-1}}],\, r\ge 1\\ \text{and }\ \xi_{i,N}(t)=\sum_{r=1}^{\infty}\zeta_{i,N,r}(t)\end{gathered}\tag{5.2}$$

where $M_i(u)=u$ if $i\ge k+1$ and $M_i(u)=u/i$ for $i=1,...,i$. First, we establish

5.8**. Proposition.** *For each fixed u, as N goes to ∞, the partial sums*

$$\xi^u_{i,N}(t)=\sum_{r=1}^{u}\zeta_{i,N,r}(t)=\sum_{1\le n\le M_i(Nt)}Y_{i,q_i(n),2^u}$$

form a tight family of processes on the Skorokhod space $D[[0,t];R^k]$. All the limit points are Gaussian processes with independent increments. The second moments are uniformly integrable so that the covariance of the limiting Gaussian process can be identified as the limit of the covariances of the corresponding approximating processes along the subsequence.

Proof. We note that $Y_{i,q_i(n),r}$ is $\mathcal{F}_{-\infty,q_i(n)+r}$ measurable. In order to apply Theorem 5.1 with $\mathcal{G}_{N,n}=\mathcal{F}_{-\infty,q_i(n)+r}$ we need to verify the conditions **B1-B4**. With such choice of $\mathcal{G}_{N,n}$, **B1** is clearly fulfilled. To verify the uniform square integrability of $\{Y_{i,q_i(n),r}\}$ we observe that the uniform square integrability of any family $\{Z_\alpha\}$ implies the uniform integrability of $\{E[Z_\alpha|\mathcal{G}]\}$ as α and $\mathcal{G}$ vary. The distribution of $\{X(n)\}$ is the same for all n and therefore by our moment condition, $|X(n)|^{2\iota}$ are uniformly integrable. Using the bound $|F|\le C(1+\sum|x_i|^\iota)$ it is easily seen that $\{Y_{i,q_i(n),r}\}$ are uniformly square integrable. To control $\|E[Y_{i,q_i(n),r}|\mathcal{F}_{-\infty,l}]\|_{L_2(P)}$ note that by (3.24) we can use Lemma 3.10 for $q_{i-1}(n)+r\le l$ which yields the estimate

$$\|E[Y_{i,q_i(n),r}|\mathcal{F}_{-\infty,l}]\|_2=\|\eta^l_{i,q_i(n),r}\|_2\le c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota})\varpi_{q,p}(q_i(n)-r-l)$$

provided $q_i(n)\ge l+r$. On the other hand, if $q_{i-1}(n)+r\ge l$, we can write

$$\begin{aligned}\|E[Y_{i,q_i(n),r}|\mathcal{F}_{-\infty,l}]\|_2&\le\|E[Y_{i,q_i(n),r}|\mathcal{F}_{-\infty,q_{i-1}(n)+r}]\|_2\\ &\le c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota})\varpi_{q,p}(q_i(n)-q_{i-1}(n)-2r)\\ &\le c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota})\varpi_{q,p}(n-2r)\end{aligned}$$

whenever $n\ge 2r$ and $n\ge n^*=n^*(i)=\min\{m:\ q_i(l)-q_{i-1}(l)\ge l\ \forall l\ge m\}$ observing that $n^*<\infty$ by (2.12). Assuming that $q\ge p$, we can always bound $\varpi_{p,q}$ by 1. Therefore, choosing $c(n)=1$ for small values of n (there are at most n^*+2r

of them) and estimating $c(n)$ by either $c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota})\varpi_{q,p}(q_i(n)-r-l)$ or by $c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota})\varpi_{q,p}(n-2r)$ we arrive at **B3** with the estimate

$$\sum_{n=0}^{\infty} c(k) \le [n^*+2r+2\sum_{n=1}^{\infty}\varpi_{p,q}(n)]c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota}).$$

If we set

$$R_{i,m,r}=\sum_{n\ge m-r} E[Y_{i,n,r}|\mathcal{F}_{-\infty,m}]$$

then it follows from the above estimates that

$$\sup_{i,l}\|R_{i,l,r}\|_2 \le 2(n^*+r+\theta(p,q))c(d,p,\kappa,\iota)c(\gamma_m,\gamma_{q\iota}) \tag{5.3}$$

where $\theta(p,q)$ is given by (2.14). It is now clear that $W_{i,n,r}=Y_{i,n-r,r}+R_{i,n+1,r}-R_{i,n,r}$ is a martingale difference and is uniformly square integrable. While **B4** may not hold, the limit will exist along suitable subsequences. The uniform bound on $\|W_{i,n,r}\|_2$ ensures that limits $A(t)$ will be Lipschitz continuous functions of t and the convergence is uniform in t. □

In order to obtain convergence of processes $\xi_{i,N}$ and not only their approximations $\xi_{i,N,r}$ we will need uniform bounds in the representations (5.2).

5.9. **Proposition.** *The differences* $\{\zeta_{i,N,r}(t)\}$ *satisfy*

$$\sum_r \sup_{N\ge1}\max_{1\le i\le \ell}\|\sup_{0\le t\le T}|\zeta_{i,N,r}(t)|\|_2 \le C<\infty. \tag{5.4}$$

Proof. Set $\tilde Y_{i,n,r}=Y_{i,n,2^r}-Y_{i,n,2^{r-1}}$, $r\ge1$ and

$$\tilde R_{i,n,r}=\sum_{m\ge n+1}E(\tilde Y_{i,m,r}|\mathcal{F}_{-\infty,n+2^r}).$$

Estimating conditional expectations here by Lemma 3.11 when $m-n\ge 2^{r+1}$ and by the contraction argument when $n+1\le m\le n+2^{r+1}$, and applying Lemma 3.11 after that again we obtain

$$\begin{aligned}\|\tilde R_{i,n,r}\|_2 &\le 2^{r+1}\sup_n\|\tilde Y_{i,n,r}\|_2+\tilde C\big((\beta(q,2^r))^\delta+(\beta(q,2^{r-1}))^\delta\big)\\ &\le \hat C 2^r\big((\beta(q,2^r))^\delta+(\beta(q,2^{r-1}))^\delta\big)\end{aligned} \tag{5.5}$$

where $\tilde C,\hat C>0$ do not depend on i,n,r. Now observe that

$$\zeta_{i,N,r}=\frac{1}{\sqrt N}\sum_{1\le m\le M_i(NT)}Z_{i,q_i(m),r}-\frac{1}{\sqrt N}\big(\tilde R_{i,q_i([M_i(NT)]),r}-\tilde R_{i,0,r}\big) \tag{5.6}$$

where $Z_{i,n,r}=\tilde Y_{i,n,r}+\tilde R_{i,n,r}-\tilde R_{i,n-1,r}$, $n\ge1$ is a martingale differences sequence with respect to the filtration $\{\mathcal{G}_n,\, n\ge1\}$ with $\mathcal{G}_n=\mathcal{F}_{-\infty,n+2^r}$. By the Doob inequality for martingales

$$\begin{aligned}&\tfrac1N E\sup_{0\le t\le T}\big|\textstyle\sum_{1\le l\le Nt}Z_{i,q_i(l),r}\big|^2\le \tfrac4N\textstyle\sum_{1\le l\le NT}EZ^2_{i,q_i(l),r}\\ &\le 4T\max_{1\le l\le NT}EZ^2_{i,q_i(l),r}\le 12T\big(\sup_n\|\tilde Y_{i,n,r}\|_2+2\sup_n\|\tilde R_{i,n,r}\|_2\big).\end{aligned} \tag{5.7}$$

We can estimate also

$$\begin{aligned}&\tfrac1N E\max_{0\le l\le NT}|\tilde R_{i,q_i(l),r}-\tilde R_{i,0,r}|^2\\ &\le \tfrac4N\textstyle\sum_{1\le l\le NT}E\tilde R^2_{i,q_i(l),r}\le 4\max_{0\le l\le NT}R\tilde R^2_{i,q_i(l),r}.\end{aligned} \tag{5.8}$$

Now collecting (5.5)–(5.8) and applying Lemma 3.11 again to (5.7) and (5.8) we obtain that

$$\sup_{N\geq 1}\|\sup_{0\leq t\leq T}|\zeta_{i,N,r}(t)|\|_2\leq\tilde{\tilde{C}}2^r\big((\beta(q,2^r))^\delta+(\beta(q,2^{r-1}))^\delta\big) \tag{5.9}$$

where $\tilde{\tilde{C}}>0$ does not depend on r. Since $\sum_{r\geq 1}(\beta(q,r))^\delta$ converges by our assumption (2.15) then $\sum_{r\geq 1}2^r(\beta(q,2^r))^\delta$ converges, as well, and so the right hand side of (5.9) is summable implying (5.4). □

Next, we deal specifically with the terms $Y_{i,q_i(n)}$, $k+1\leq i\leq\ell$ which satisfy (2.10), (2.11) and (2.12). By Propositions 5.8 and 5.9 any possible limit $\eta_i(t)$ in distribution of

$$\xi_{i,N}(t)=\frac{1}{\sqrt{N}}\sum_{n\leq Nt}Y_{i,q_i(n)}$$

for $1\leq i\leq\ell$ will be a Gaussian process with independent increments. The processes $\{\eta_i(\cdot),\,k+1\leq i\leq\ell\}$ will be mutually independent as well as totally independent of $\{\eta_i(\cdot),\,1\leq i\leq k\}$ which is proved by successive application of Theorem 5.6. We note that it is enough to show that for any $T<\infty$ we can ignore $\sum_{n\leq k_N(i)}Y_{i,q_i(n)}$ in the definition of $\xi_{i,N}(t)$ where $k_N(i)=\max\{n:\,q_i(n)\leq q_{i-1}(NT)\}$ so that Theorem 5.6 will be applicable then to the approximations

$$\xi_{i,N,r}(t)=\frac{1}{\sqrt{N}}\sum_{k_N(i)+1\leq n\leq Nt}Y_{i,q_i(n),r}$$

with $Y_{i,q_i(n),r}$ defined in (3.23). At the end, relying on Proposition 5.9 we can let $r\to\infty$ and complete the proof. From (2.12), for any $\epsilon>0$, $q_i(N\epsilon)\geq q_{i-1}(NT)$ for large N which implies that the the initial terms are at most $N\epsilon$ in number. Since ϵ is arbitrary we see that $N^{-1}k_N(i)\to 0$ as $N\to\infty$. By (4.3) of Lemma 4.2 we obtain that the contribution of initial $k_N(i)$ terms in the sum for $\xi_{i,N}$ is negligible. Similarly we conclude that it does not matter whether we take the sum for $\xi_{i,N,r}(t)$ above till Nt or till $Nt+k_N(i)$ as in Theorem 5.6. By Proposition 4.5 we have also that the limiting variance $A_{i,i}(t)$ of each $\xi_{i,N}(t)$, $i>k$ exists and is given by (4.12).

We observe that independency of processes η_i, $i>k$ of each other and of η_i, $i\leq k$ can be proved in an alternative way without using Theorem 5.6. Namely, we can rely on Theorem 5.1 showing that linear combinations of processes $\xi_{i,N,r}$ converge to Gaussian processes deriving similarly to above via uniform estimates of Proposition 5.9 that linear combination of processes η_i are Gaussian and concluding the proof via the vanishing covariances assertion (4.13) of Proposition 4.5.

Now, we are able to complete the proof of Theorem 2.2. First, we conclude from Propositions 5.8 and 5.9 together with Corollary 5.7 that the k-dimensional process $\{\xi_{i,N}(t):\,1\leq i\leq k\}$ converges in distribution as $N\to\infty$ to a Gaussian process $\{\eta_i(t):\,1\leq i\leq k\}$ with stationary independent increments whose covariances are given by Proposition 4.1. As explained above, when $i\geq k+1$, the process $\xi_{i,N}(t)$ converges in distribution to a Gaussian process $\eta_i(t)$ with stationary independent increments and $\eta_{k+1}(t),...,\eta_\ell(t)$ are both mutually independent and independent of processes $\eta_1(t),...,\eta_k(t)$. It follows that the ℓ-dimensional process $\{\xi_{i,N}(t):\,1\leq i\leq\ell\}$ converges in distribution as $N\to\infty$ to the Gaussian process $\{\eta_i(t):\,1\leq i\leq\ell\}$ with stationary independent increments whose covariances are given by Proposition 4.1 and Proposition 4.5 taking into account independency of processes $\eta_i(t)$ with $i\geq k+1$ of other processes $\eta_j(t)$ with $j\neq i$.

It remains to show that the process $\xi_N(t)$ given by (2.20) converges in distribution as $N\to\infty$ to a Gaussian process $\xi(t)$ given by (2.23). The convergence itself is clear since each $\xi_{i,N}$ converges to the corresponding η_i. In order to show that ξ is a Gaussian process it suffices to prove the same for $\zeta(t)=\sum_{i=1}^k\eta_i(it)$ since $\tilde\zeta(t)=\sum_{i=k+1}^\ell\eta_i(t)$ is a Gaussian process (as a sum of independent Gaussian processes) independent of ζ, and so $\zeta(t)+\tilde\zeta(t)$ is a Gaussian process if $\zeta(t)$ is. Since $(\eta_1(t),...,\eta_k(t))$ is a k-dimensional Gaussian process with independent increments then the vector increments $\big(\eta_j(it)-\eta_j((i-1)t),\ i=1,2,...,k\big)$ for $j=1,2,...,k$ are mutually independent k-dimensional Gaussian processes, and so

$$\zeta_\lambda(t)=\sum_{i=1}^k\sum_{j=1}^k\lambda_{ij}\big(\eta_j(it)-\eta_j((i-1)t)\big)=\sum_{j=1}^k\sum_{i=1}^k\lambda_{ij}\big(\eta_j(it)-\eta_j((i-1)t)\big)$$

is a Gaussian process for any choice of constants λ_{ij} and we recall that $\eta_j(0)=\xi_{j,N}(0)=0$. Now observe that choosing $\lambda_{ij}=1$ if $i\le j$ and $\lambda_{ij}=0$, otherwise, we obtain that $\zeta_\lambda(t)=\zeta(t)$ completing the proof.

As to our claim that increments of $\xi(t)$ may not be independent if $k\ge 2$ consider, for instance, the case $k=\ell=2$ and

$$\xi(t)-\xi(t/2)=\eta_1(t)+\eta_2(2t)-\eta_1(t/2)-\eta_2(t)\ \text{and}\ \xi(t/2)=\eta_1(t/2)+\eta_2(t).$$

Then by Proposition 4.1,

$$E\big(\xi(t/2)(\xi(t)-\xi(t/2)\big)=D_{2,1}t/2$$

where

$$D_{2,1}=\frac{1}{2}\sum_{u=-\infty}^{\infty}a_{2,1}(u)$$

and

$$a_{2,1}(u)=\int F_2(x,y)F_1(z)d\mu(x)d\mu_u(y,z).$$

Assume, for instance, that $X(0),X(1),X(2),...$ is a sequence of independent identically distributed random variables then $\mu_u=\mu\times\mu$ if $u\ne 0$, and so $a_{2,1}(u)=0$ if $u\ne 0$ while

$$a_{2,1}(0)=\int F_2(x,y)F_1(y)d\mu(x)d\mu(y).$$

Now suppose that $EX(0)=0$, $EX^2(0)=1$ and choose $F(x,y)=x^2y^2-1$. Then $\int F(x,y)d\mu(x)d\mu(y)=0$, $F_2(x,y)=x^2(y^2-1)$, $F_1(x)=x^2-1$, and so

$$D_{2,1}=\frac{1}{2}a_{2,1}(0)=\int(y^2-1)^2d\mu(y)\ne 0$$

unless $X^2(0)=1$ with probability one. □

6. Higher moments

It is sometimes of interest to know that there is convergence of higher moments. Once the limit theorem is established it is only a question of obtaining uniform bounds for them. Indeed, if we have bounds of the type

$$E|\xi_N(t)|^m\le C_m(t)$$

for some $C_m(t)>0$ valid for all N and m then

$$E|\xi_N(t)|^m\mathbb{I}_{\xi_N(t)>M}\le M^{-1}E|\xi_N(t)|^{m+1}\mathbb{I}_{\xi_N(t)>M}\le C_{m+1}(t)M^{-1}$$

where $\mathbb{I}_A$ denotes the indicator of an event A, which provides uniform in N integrability of all powers $\{\xi_N^m(t)\}$. Since $\xi_N(t)$ converges in distribution this uniform integrability yields convergence of all moments of $\xi_N(t)$ to the corresponding moments of the limiting distribution.

We start obtaining required estimates with a general result on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration of $\sigma-$fields $\mathcal{G}_j$. Suppose that random variables X_j are $\mathcal{G}_j$ measurable and for all $p < \infty$ satisfy

$$\gamma_p = \sup_j \|X_j\|_p \le \sup_i \sum_{j\ge i} \|E[X_j|\mathcal{G}_i]\|_p = A_p < \infty. \tag{6.1}$$

We will explore the behavior of higher order moments for sums $S_n = \sum_{i=1}^n X_i$ obtaining estimates of the form $E[S_n^{2l}] \le C_{2l} n^l$ with some control on dependence of constants C_{2l} on γ_{2l} and A_{2l}.

6.1. **Lemma.** *Suppose $\{a_n\}$ is a sequence of nonnegative numbers such that for some integer $l \ge 1$ and any integer $n \ge 1$,*

$$a_{n+1} \le c \sum_{j=1}^n \sum_{r=2}^{2l} C^r a_j^{\frac{2l-r}{2l}}.$$

Then

$$a_n \le A\, n^l$$

with $A = \max\{2^l c^l C^{2l}, C^{2l}, a_1\}$.

Proof. We derive the above inequality by induction. It is clearly valid for $n = 1$. Assume it is valid for $j = 1, 2, \dots, n$. Then

$$\begin{aligned} a_{n+1} &\le c \textstyle\sum_{j=1}^n \sum_{r=2}^{2l} C^r (Aj^l)^{\frac{2l-r}{2l}} \\ &\le c\, C^2\, A^{1-\frac{1}{l}} \textstyle\sum_{r=2}^{2l} C^{r-2} A^{-\frac{r-2}{2l}} \ \sum_{j=1}^n j^{l-1} \le A' \frac{(n+1)^l}{l} \end{aligned}$$

where

$$A' = c\, C^2\, A^{1-\frac{1}{l}} \sum_{r=0}^{2l-2} C^r A^{-\frac{r}{2l}}$$

and we need to pick A so that $\frac{A'}{l} \le A$. In particular, $A = \max\{2^l c^l C^{2l}, C^{2l}, a_1\}$ will do because $CA^{-\frac{1}{2l}} \le 1$, $2\, c\, C^2 A^{-\frac{1}{l}} \le 1$ and

$$c\, C^2\, A^{1-\frac{1}{l}} \sum_{r=0}^{2l-2} C^r A^{-\frac{r}{2l}} \le c\, C^2\, A^{1-\frac{1}{l}} (2l-1) \le c\, C^2\, A^{1-\frac{1}{l}} 2l \le l\, A.$$

□

6.2. **Lemma.** *Let the sequence $\{X_i\}$ of random variables satisfy (6.1). Then for each $l \ge 1$ there is a constant c_l depending only on l such that*

$$E[S_n^{2l}] \le c_l\, A_{2l}^{2l}\, n^l.$$

Proof. We begin by expanding $S_{j+1}^{2l} = (S_j + X_{j+1})^{2l}$ by the binomial theorem,

$$S_{j+1}^{2l} = S_j^{2l} + 2l S_j^{2l-1} X_{j+1} + \sum_{r=2}^{2l} \binom{2l}{r} S_j^{2l-r} X_{j+1}^r$$

and expressing

$$S_j^{2l-1}=\sum_{i=1}^{j}(S_i^{2l-1}-S_{i-1}^{2l-1})=\sum_{1\le i\le j}X_i\sum_{r=0}^{2l-2}S_i^rS_{i-1}^{2l-2-r}.$$

This enables us to rewrite

$$S_{j+1}^{2l}=S_j^{2l}+2l\sum_{1\le i\le j}Z_iX_{j+1}+\sum_{r=2}^{2l}\binom{2l}{r}S_j^{2l-r}X_{j+1}^r$$

where $Z_i=X_i\sum_{r=0}^{2l-2}S_i^rS_{i-1}^{2l-2-r}$. Then,

$$\begin{aligned}ES_{n+1}^{2l}&=EX_1^{2l}+2l\sum_{1\le i\le j\le n}EZ_iX_{j+1}+\sum_{j=1}^{n}\sum_{r=2}^{2l}\binom{2l}{r}ES_j^{2l-r}X_{j+1}^r\\&=2l\sum_{1\le i\le n}EZ_iW_i+\sum_{j=1}^{n}\sum_{r=2}^{2l}\binom{2l}{r}ES_j^{2l-r}X_{j+1}^r\end{aligned}$$

where $W_i=\sum_{j=i}^nE[X_{j+1}|\mathcal{F}_i]$. We note that $\|X_i\|_{2l}\le\gamma_{2l}\le A_{2l}$ and $\|W_i\|_{2l}\le A_{2l}$. Hence,

$$\begin{aligned}E[|Z_iW_i|]&\le\|\sum_{r=0}^{2l-2}S_i^rS_{i-1}^{2l-2-r}\|_{\frac{l}{l-1}}\|X_i\|_{2l}\|W_i\|_{2l}\\&\le c_lA_{2l}^2[[E[S_i^{2l}]]^{\frac{l-1}{l}}+[E[S_{i-1}^{2l}]]^{\frac{l-1}{l}}].\end{aligned}$$

Next, for $r\ge2$,

$$|E[S_j^{2l-r}X_{j+1}^r]|\le\|S_j\|_{2l}^{2l-r}\|X_{j+1}\|_{2l}^r\le A_{2l}^r\|S_j\|_{2l}^{2l-r}.$$

It follows that

$$\begin{aligned}E[S_{n+1}^{2l}]&\le c_l\Big[\sum_{j=1}^n\big[\sum_{r=2}^{2l}A_{2l}^r\|S_j\|_{2l}^{2l-r}+A_{2l}^2\big\|S_j\|_{2l}^{2l-2}+A_{2l}^2\|S_{j-1}\|_{2l}^{2l-2}\big]\Big]\\&\le c_l\Big[\sum_{j=1}^n\sum_{r=2}^{2l}A_{2l}^r\|S_j\|_{2l}^{2l-r}\Big]\end{aligned}$$

where c_l is an absolute constant which depends only on l. The sequence $a_n=E[S_n^{2l}]$ satisfies the condition of Lemma 6.1 with $c=c_l$, $C=A_{2l}$ and $a_1\le\gamma_{2l}^{2l}$ and the result follows. □

6.3**. Proposition.** *Suppose that Assumption 2.3 holds true. Then for any integer* $p\ge1$,

$$\sup_N E\big[[\xi_{i,N}(t)]^{2p}\big]<\infty.$$

Proof. Observe that $Y_{i,n,r}$ defined in (3.23) is $\mathcal{F}_{-\infty,n+r}$ measurable. If $l+\max\{r,r'\}\le q_{i-1}(n)+\max\{r,r'\}\le q_i(n)-\max\{r,r'\}$ then by the definition of the mixing coefficient $\varpi_{q,p}$,

$$\|E[Y_{i,q_i(n),r}-Y_{i,q_i(n),r'}|\mathcal{F}_{-\infty,l+\max\{r,r'\}}]\|_p$$
$$\le \|E[Y_{i,q_i(n),r}-Y_{i,q_i(n),r'}|\mathcal{F}_{-\infty,q_{i-1}(n)+\max\{r,r'\}}]\|_p$$
$$\le \varpi_{q,p}(q_i(n)-q_{i-1}(n))-2\max\{r,r'\})\|Y_{i,q_i(n),r}-Y_{i,q_i(n),r}\|_q.$$

On the other hand, if $q_{i-1}(n)+\max\{r,r'\}\le l+\max\{r,r'\}\le q_i(n)-\max\{r,r'\}$ then

$$\|E[Y_{i,q_i(n),r}-Y_{i,q_i(n),r')}|\mathcal{F}_{-\infty,l+\max\{r,r'\}}]\|_p$$
$$\le \varpi_{q,p}(q_i(n)-l-2\max\{r,r'\})\|Y_{i,q_i(n),r}-Y_{i,q_i(n),r'}\|_q.$$

The conditions for these estimates are valid when $n\ge l+2\max(r,r')$. We observe that there are at most $2\max(r,r')+1$ terms with $q_i(n)$ between l and $l+2\max(r,r')$. We estimate the latter terms just by contraction of conditional expectations. Replacing r by 2^r and r' by 2^{r+1} we conclude from the above estimate combined with Lemma 3.12 that

$$\sup_l \sum_{q_i(n)\ge l}\|E[Y_{i,q_i(n),2^r}-Y_{i,q_i(n),2^{r+1}}|\mathcal{F}_{l+2^{r+1}}]\|_p \le C[2^{r+2}+\theta(q,p)]\sup_n\|Y_{i,n,2^r}-Y_{i,n,2^{r+1}}\|_q\le \tilde{C}2^r\beta^\delta(q,2^r) \tag{6.2}$$

for some $C,\tilde{C}>0$ where $\theta(q,p)$ is the same as in (2.14). The right hand side of (6.2) is bounded, even tends to zero as $r\to\infty$ in view of (2.29). Hence, we can apply Lemma 6.2 to $S_n=\sqrt{n}\zeta_{i,n,r}$, with $\zeta_{i,n,r}(t)$ defined in (5.2), which yields that for any positive integer l,

$$E\big(\xi_{i,N,2^r}(t)-\xi_{i,N,2^{r+1}}(t)\big)^{2l}\le c_l\tilde{C}^{2l}2^{rl}\beta^{\delta l}(q_l,2^r)$$

where $q=q_l$ depends on l. Now the assertion of Proposition 6.3 follows by writing

$$\|\xi_{i,N}(t)\|_{2l}\le\|\xi_{i,N,1}(t)\|_{2l}$$
$$+\textstyle\sum_{r=0}^\infty\|\xi_{i,N,2^{r+1}}(t)-\xi_{i,N,2^r}(t)\|_{2l}\le\tilde{\tilde{C}}_l\sum_{r=0}^\infty 2^r\beta^\delta(q_l,2^r)<\infty,$$

where $\tilde{\tilde{C}}_l>0$ depends only on l and convergence in the right hand side is ensured by (2.29). □

Proposition 6.3 gives required uniform bounds on even moments of $\xi_{i,N}(t)$ and bounds on odd absolute moments are obtained by the Cauchy-Schwarz inequality $E|\xi_{i,N}(t)|^p\le(E\xi^{2p}_{i,N}(t))^{1/2}$. The corresponding bounds on moments of $\xi_N(t)$ follow by

$$\|\xi_N(t)\|_l\le\sum_{i=1}^k\|\xi_{i,N}(it)\|_l+\sum_{i=k+1}^{\ell}\|\xi_{i,N}(t)\|_l$$

which together with the argument at the beginning of this section completes the proof of Theorem 2.4. □

We observe that convergence of moments of $\xi_N(t)$ to the corresponding moments of an appropriate Gaussian distribution can be derived also directly without relying on convergence of distributions so that the latter would follow from the former by the method of moments (see, for instance, [20]) but under stronger assumptions of Theorem 2.4. Namely, the moment $E\xi_N^{2m}(t)$ is the sum of terms $N^{-m}\prod_{1\le i\le\ell}\prod_{0\le n\le Nt}Y^{l_{i,n}}_{i,q_i(n)}$ with $\sum_{i,n}l_{i,n}=2m$. Relying on combinatorial counting arguments together with estimates of Corollary 3.6 it is not difficult to see that asymptotically as $N\to\infty$ the contribution to the sum comes only from terms where the product can be arranged into product of pairs $Y_{i,q_i(n)}Y_{j,q_j(l)}$ with $|q_i(n)-q_j(l)|$ not too large while for different pairs with, say, $q_i(n),q_j(l)$ and $q_{i'}(n'),q_{j'}(l')$ the quantities $|q_i(n)-q_{i'}(n')|$ and $|q_j(l)-q_{j'}(l')|$ will be much larger and will tend to ∞ as $N\to\infty$. Relying on Corollary 3.6 this enables us to treat different pairs as if they were independent of each other and applying combinatorial counting arguments similar to estimates of moments of sums of independent random variables we obtain appropriate limits of moments.

7. Continuous time case

First, we represent again the function F in the form (2.17) and $\xi_N(t)$ given by (2.34) in the form (2.20) where now

$$\xi_{i,N}(t)=\frac{1}{\sqrt{N}}\int_0^{S_i(Nt)}F_i\big(X(q_1(s)),...,X(q_i(s))\big)ds \tag{7.1}$$

with $S_i(u)=u/i$ if $i\le k$ and $S_i(u)=u$ if $i\ge k+1$. Set

$$\begin{gathered}F_{i,r,t}=F_{i,r,t}(x_1,...,x_{i-1},\omega)=E\big(F(x_1,...,x_{i-1},X(t)|\mathcal{F}_{t-r,t+r}\big),\\ X_r(t)=E(X(t)|\mathcal{F}_{t-r,t+r}),\ Y_i(t)=F_i\big(X(q_1(s)),...,X(q_i(s))\big)\text{ if }t=q_i(s)\\ \text{and }Y_i(t)=0\text{ if }t\ne q_i(s)\text{ for any }s,\ Y_{i,r}(t)=F_{i,r,t}\big(X_r(q_1(s)),...,X_r(q_i(s))\big)\\ \text{if }t=q_i(s)\text{ and }Y_{i,r}(t)=0\text{ if }t\ne q_i(s)\text{ for any }s.\end{gathered}$$

In order to use fully our discrete time technique it will be convenient to pass from $\xi_{i,N}$ to $\tilde\xi_{i,N}$ given by

$$\tilde\xi_{i,N}(t)=\frac{1}{\sqrt{N}}\sum_{n=0}^{[S_i(Nt)]}I_i(n)$$

where $I_i(n)=\int_n^{n+1}Y_i(q_i(s))ds$. The error of such transition is estimated by

$$\sup_{0\le t\le T}|\xi_{i,N}(t)-\tilde\xi_{i,N}(t)|\le\frac{1}{\sqrt{N}}\max_{0\le n\le NT}Q_i(n) \tag{7.2}$$

where $Q_i(n)=\int_0^1|Y_i(q_i(n+s))|ds$. Now for any $\delta>0$,

$$\begin{gathered}P\{\max_{0\le n\le NT}Q_i(n)>\varepsilon\sqrt{N}\}\le NT\max_{0\le n\le NT}P\{Q_i(n)>\varepsilon\sqrt{N}\}\\ \le\tfrac{T}{\varepsilon^2}\max_{0\le n\le NT}\int_{\{Q_i(n)>\varepsilon\sqrt{N}\}}Q_i^2(n)dP\\ \le(\varepsilon\sqrt{N})^{-\delta}\int Q_i^{2+\delta}(n)dP\le(\varepsilon\sqrt{N})^{-\delta}\int_0^1EY_i^{2+\delta}(q_i(n+s))ds\le C(\varepsilon\sqrt{N})^{-\delta}.\end{gathered}$$

Thus, the left hand side of (7.2) tends to 0 in probability as $N\to\infty$, and so it suffices to prove our functional central limit theorem for $\tilde\xi_{i,N}$ in place of $\xi_{i,N}$.

Introduce the approximations $\tilde\xi_{i,N,r}$ of $\tilde\xi_{i,N}$ by

$$\tilde\xi_{i,N,r}(t)=\frac{1}{\sqrt{N}}\sum_{n=0}^{[S_i(Nt)]}I_{i,r}(n) \tag{7.3}$$

where $I_{i,r}(n)=\int_n^{n+1}Y_{i,r}(q_i(s))ds$. Now set

$$R_{i,r}(m)=\sum_{l=m+1}^{\infty}E\big(I_{i,r}(l)|\mathcal{F}_{-\infty,m+r}\big)$$

and $Z_{i,r}(m)=I_{i,r}(m)+R_{i,r}(m)-R_{i,r}(m-1)$. Then $E(Z_{i,r}(m)|\mathcal{F}_{-\infty,m-1+r})=0$, and so $\{Z_m,\mathcal{G}_m\}_{m\geq 0}$ with $Z_m=Z_{i,r}(m)$ and $\mathcal{G}_m=\mathcal{F}_{-\infty,m+r}$ turns out to be a martingale differences sequence. We saw already above that $\{Q_i^2(n)\}$ is uniformly integrable. Then both $\{I_i^2(n)\}$ and $\{I_{i,r}^2(n)\}$ are uniformly integrable and like in the proof of Proposition 5.8 we conclude that both $\{R_{i,r}^2(n)\}$ and $\{Z_{i,r}^2(n)\}$ are uniformly integrable, as well. Set

$$\zeta_{i,N,r}(t)=\frac{1}{\sqrt{N}}\sum_{n=0}^{[S_i(Nt)]}Z_{i,r}(n).$$

Then similarly to Section 5 we obtain that

$$\sup_{0\leq t\leq T}|\tilde{\xi}_{i,N,r}(t)-\zeta_{i,N,r}(t)|\to 0 \text{ in probability as } N\to\infty, \tag{7.4}$$

and so in order to obtain a central limit theorem for $\tilde{\xi}_{i,r,N}(t)$ it suffices to prove it for the normalized martingal $\zeta_{i,r,N}(t)$.

In order to invoke martingale limit theorems we have to study next the asymptotical behavior as $N\to\infty$ of normalized variances $E\big(\zeta_{i,r,N}(S_i(Nt))\big)^2$. As in the discrete time case considered in Section 4, in view of (2.17) and (7.1) it suffices to study the asymptotical behavior of

$$D_{i,j}(N,s,t)=E[\xi_{i,N}(s)\xi_{j,N}(t)] \tag{7.5}$$
$$=\tfrac{1}{N}\int_0^{S_j(Nt)}\int_0^{S_i(Ns)}E[Y_i(q_i(u))Y_j(q_j(v))]dudv.$$

We treat first the case when $1\leq i,j\leq k$ similarly to Proposition 4.1. Let υ be the greatest common divisor of i and j then similarly to the argument in Lemma 4.4 we obtain that for any integer w,

$$\lim_{u,v\to\infty,\, iu-jv=w\upsilon}E[Y_i(iu)Y_j(jv)]=a_{i,j}(w,2w,...,\upsilon w) \tag{7.6}$$

with $a_{i,j}$ defined in Proposition 4.1. Now, changing variables we have

$$\tfrac{1}{N}\int_0^{Nt/j}\int_0^{Ns/i}E[Y_i(iu)Y_j(jv)]dudv \tag{7.7}$$
$$=\tfrac{\upsilon}{Ni}\int_0^{Nt/j}\int_{-jv/\upsilon}^{(Ns-jv)/\upsilon}E[Y_i\big(\tfrac{jv+w\upsilon}{i}\big)Y_j(jv)]dwdv.$$

When v is large then the expectation under the integral equals approximately $a_{i,j}(w,2w,...,\upsilon w)$ and taking into account that the latter is absolutely integrable in w from $-\infty$ to ∞ we can approximate the interior integral in w by the integral $\int_{-\infty}^{\infty}$. Next we integrate in v within constraints $0\leq v\leq Nt/j$ and $u=(jv+w\upsilon)/i\leq Ns/i$, i.e. asymptotically for N large $0\leq v\leq\frac{N}{j}\min(s,t)$. It follows that the expression in (7.7) is approximately equal as $N\to\infty$ to

$$\frac{\upsilon}{ij}\min(s,t)\int_{-\infty}^{\infty}a_{i,j}(w,2w,...,\upsilon w)dw \tag{7.8}$$

and we obtain the same covariances as in the discrete time case.

Next, we claim that for each $i=k+1,...,\ell$ and $t>0$,

$$\lim_{N\to\infty}D_{i,i}(N,t,t)=0. \tag{7.9}$$

Indeed, set again $b_{i,j}(u,v)=E\big(Y_i(q_i(u))Y_j(q_j(v)\big)$. Then

$$\begin{aligned}(7.10)\qquad & \frac{1}{N}\int_0^{Nt}\int_0^{Nt}|b_{i,i}(u,v)|dudv\le\frac{2}{N}\int_0^{Nt}du\int_u^{u+\gamma}|b_{i,i}(u,v)|dudv\\ &+\frac{2}{N}\int_0^{N\gamma}du\int_{u+\gamma}^{Nt}|b_{i,i}(u,v)|dudv+\frac{2}{N}\int_{N\gamma}^{Nt}du\int_{u+\gamma}^{Nt}|b_{i,i}(u,v)|dudv\\ &\le C(t\gamma+\gamma+t\beta_\gamma^{(i)}(N\gamma))\end{aligned}$$

for some $C>0$ independent of t, N and γ where we obtain by (2.33) and estimates similar to Lemma 4.2 and Proposition 4.5 that for any $i>k$ and $\gamma>0$,

$$(7.11)\qquad \beta_\gamma^{(i)}(M)=\sup_{u\ge M}\int_{u+\gamma}^{\infty}|b_{i,i}(u,v)|dv<\infty \text{ and } \lim_{M\to\infty}\beta_\gamma^{(i)}(M)=0.$$

So, letting first $N\to\infty$ and then $\gamma\to 0$ we obtain (7.9).

Convergence of moments in the continuous time case is obtained similiraly to Section 6 by obtaining appropriate bound on moments of $\xi_{i,N}$. In order to employ directly the technique of Section 6 we can represent $\xi_{i,N}$ as a sum

$$\xi_{i,N}(t)=\frac{1}{\sqrt{N}}\sum_{n=1}^{[Nt]}\tilde{Y}_{i,n}+\frac{1}{\sqrt{N}}\int_{[Nt]}^{S_i(Nt)}Y_i(q_i(s))ds,$$

where $\tilde{Y}_{i,n}=\int_{n-1}^n Y_i(q_i(s))ds$. Relying on Lemma 6.2 for the sum $\sqrt{N}\xi_{i,N}(t)$ written above we obtain essentially the same estimates for moments as in Section 6.

7.1. **Remark.** In fact, in the continuous time case we can take $q_i(t)=\alpha_i t$ for arbitrary $0<\alpha_1<\alpha_2<\cdots<\alpha_k$ in place of $1<2<\cdots<k$ while leaving $q_i(t)$, $i=k+1,...,\ell$ as before. In this situation (7.6) becomes

$$\lim_{u,v\to\infty,\,\alpha_i u-\alpha_j v=z}E[Y_i(\alpha_i u)Y_j(\alpha_j v)]=a_{i,j}(\rho_1 z,\rho_2 z,...,\rho_{n_{ij}}z,z)$$

where $\rho_1<\rho_2<\cdots<\rho_{n_{ij}}<1$ and $\alpha_i\rho_l,\alpha_j\rho_l\in\{\alpha_1,...,\alpha_k\}$ for $l=1,...,n_{ij}$. Then the covariances (7.8) will have the form

$$\frac{1}{\alpha_i\alpha_j}\min(s,t)\int_{-\infty}^{\infty}a_{i,j}(\rho_1 w,\rho_2 w,...,\rho_{n_{ij}}w,w)dw.$$

References

[1] V. Bergelson, *Weakly mixing PET*, Ergod. Th.& Dynam. Sys. 7, 337–349 (1987).

[2] P. Billingsley, *Convergence of Probability Measures*, 2nd ed. Wiley, New York, 1999.

[3] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Math. 470, Springer–Verlag, Berlin, 1975.

[4] R.C. Bradley, *Introduction to Strong Mixing Conditions,* Kendrick Press, Heber City, 2007.

[5] V. Bergelson, A. Leibman and C.G. Moreira, *From discrete to continuous time ergodic theorems*, Preprint.

[6] D. Dolgopyat, *Limit theorems for partially hyperbolic systems,* Trans. Amer. Math. Soc. 356, 1637–1689 (2004).

[7] N. Dunford and J.T. Schwartz, Linear Operators, Part I, Wiley, New York, 1958.

[8] K. Fukuyama, *The central limit theorem for* $\sum f(\theta^n x)g(\theta^{n^2}x)$, Ergod. Th.& Dynam. Sys. 20, 1335–1353 (2000).

[9] H. Furstenberg, *Nonconventional ergodic averages*, Proc. Symp. Pure Math. 50, 43–56 (1990).

[10] N. Frantzikinakis and B. Kra, *Polynomial averages converge to the product of integrals*, Israel J. Math. 148, 267–276 (2005).

[11] H. Furstenberg and B. Weiss, *A mean ergodic theorem for* $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$, Convergence in Ergodic Theory and Probability, p.p. 193–227, de Gruyter, Berlin (1996).

[12] D.J.H. Garling, *Inequalities: a Journey into Linear Analysis*, Cambridge Univ. Press, Cambridge (2007).

[13] L. Heinrich, *Mixing properties and central limit theorem for a class of non-identical piecewise monotonic* C^2*-transformations*, Mathematische Nachricht. 181, 185–214 (1996).

[14] P. Hall and C.C. Heyde, *Martingale Limit Theory and its Application*, Acad. Press, New York (1980).

[15] I.A. Ibragimov and Yu.V. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters–Noordhoff, Groningen (1971).

[16] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd. ed., Springer–Verlag, Berlin, 2003.

[17] Yu. Kifer, *Optimal stopping and strong approximation theorems*, Stochastics 79, 253–273 (2007).

[18] Yu. Kifer, *Nonconventional limit theorems*, Probab. Th. Rel. Fields, 148, 71–106 (2010).

[19] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge Univ. Press, Cambridge 1990.

[20] M. Loève, *Probability Theory One*, 4th ed., Springer-Verlag, New York, 1977.

[21] D.L. McLeish, *On the invariance principle for nonstationary mixingales*, Ann. Probab. 5 (1977), 616–621.

[22] D.W. Stroock and S.R.S. Varadhan, *Multidimensional Diffusion Processes*, Springer–Verlag, Berlin, 1979.

Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel
E-mail address: kifer@math.huji.ac.il

Courant Institute for Mathematical Studies, New York University, 251 Mercer St, New York, NY 10012, USA
E-mail address: varadhan@cims.nyu.edu