NONCONVENTIONAL LIMIT THEOREMS IN DISCRETE AND CONTINUOUS TIME VIA MARTINGALES

By Yuri Kifer[1] and S. R. S. Varadhan[2]

Hebrew University and New York University

We obtain functional central limit theorems for both discrete time expressions of the form $1/\sqrt{N} \sum_{n=1}^{[Nt]} (F(X(q_1(n)), \ldots, X(q_\ell(n))) - \bar{F})$ 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, $\bar{F} = \int F \, d(\mu \times \cdots \times \mu)$, 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 [*Probab. Theory Related Fields* **148** (2010) 71–106], 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 [*Probab. Theory Related Fields* **148** (2010) 71–106], our results hold true when $X_i(n) = T^n f_i$, 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 (cf. [7]) $1/N \sum_{n=1}^{N} T^{q_1(n)} f_1 \cdots T^{q_\ell(n)} f_\ell$, 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 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

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

and for the corresponding continuous time expressions of the form

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

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 and 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 with $q_j(t) = jt$ 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 approximations approach which has 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 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." Once martingale approximations are constructed, it remains only to check convergence of covariances which we do in Section 4, while the whole approach is explained and completed in Section 5.

Unlike the classical situation, our functional central limit theorem yields a process which has Gaussian distributions but not necessarily independent increments and we demonstrate an explicit example of such limiting process with dependent increments. This interesting effect rarely appears in natural models. 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 [2]). 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 [13], 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.

Our results can be applied to large classes of stochastic processes $X(n), n \geq 0$, in particular, to functions of Markov chains satisfying Doeblin's condition or to those which are constructed from sufficiently fast mixing dynamical systems. The continuous time version holds true, in particular, when $X(t)$ is a function of an irreducible continuous time Markov chain or of a nondegenerate diffusion on a compact manifold or of Ornstein–Uhlenbeck type processes.

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

$$(2.1) \quad \begin{aligned} &\varpi_{q,p}(\mathcal{G}, \mathcal{H}) \\ &= \sup\{\|E[g|\mathcal{G}] - E[g]\|_p : g \text{ is } \mathcal{H}\text{-measurable and } \|g\|_q \leq 1\}, \end{aligned}$$

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 [5], Chapter 4)

$$\alpha(\mathcal{G}, \mathcal{H}) = \tfrac{1}{4}\varpi_{\infty,1}(\mathcal{G}, \mathcal{H}), \qquad \rho(\mathcal{G}, \mathcal{H}) = \varpi_{2,2}(\mathcal{G}, \mathcal{H}),$$

$$\phi(\mathcal{G}, \mathcal{H}) = \tfrac{1}{2}\varpi_{\infty,\infty}(\mathcal{G}, \mathcal{H}) \quad \text{and} \quad \psi(\mathcal{G}, \mathcal{H}) = \varpi_{1,\infty}(\mathcal{G}, \mathcal{H}).$$

We set also

$$(2.2) \qquad \varpi_{q,p}(n) = \sup_{k \geq 0} \varpi_{q,p}(\mathcal{F}_{-\infty,k}, \mathcal{F}_{k+n,\infty})$$

and, accordingly,

$$\alpha(n) = \tfrac{1}{4}\varpi_{\infty,1}(n), \qquad \rho(n) = \varpi_{2,2}(n),$$

$$\phi(n) = \tfrac{1}{2}\varpi_{\infty,\infty}(n), \qquad \psi(n) = \varpi_{1,\infty}(n).$$

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

$$(2.3) \qquad C(q, p) = \sum_{n \geq 1} \varpi_{q,p}(n)$$

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

$$(2.4) \qquad \beta(p, r) = \sup_{k \geq 0} \|X(k) - E[X(k)|\mathcal{F}_{k-r,k+r}]\|_p.$$

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

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

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

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

$$(2.6) \qquad \begin{aligned} &\left| F(x_1, \ldots, x_\ell) - F(y_1, \ldots, y_\ell) \right| \\ &\leq K \left[1 + \sum_{j=1}^{\ell} |x_j|^\iota + \sum_{j=1}^{\ell} |y_j|^\iota \right] \sum_{j=1}^{\ell} |x_j - y_j|^\kappa \end{aligned}$$

and

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

To simplify formulas, we assume a centering condition

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

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

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

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 $k + 1 \leq i \leq \ell$,

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

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

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

which implies because of (2.10) that

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

To shorten some of the 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

$$(2.13) \qquad \gamma_\theta^\theta = \|X\|_\theta^\theta = E|X(n)|^\theta = \int |x|^\theta \, d\mu.$$

Our main result relies on the following.

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

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

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

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

REMARK 2.2. The reader willing to reduce technicalities in the first reading can be advised to keep in mind simplified assumptions such as $\ell = k$ [i.e., to consider only linear times $q_j(n) = jn$], bounded and Lipschitz continuous F and $\varpi_{q,p}(n), \beta(q, n)$ decaying exponentially fast in n. Such simplifications save some of our estimates but, otherwise, most of our machinery still should be applied.

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 $F = F(x_1, x_2, \ldots, x_\ell)$ in the form

$$(2.17) \qquad F = F_1(x_1) + \cdots + F_\ell(x_1, x_2, \ldots, x_\ell),$$

where for $i < \ell$,

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

and

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

which ensures, in particular, that

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

These enable us to write

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

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

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

and for $i \geq k+1$,

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

THEOREM 2.3. *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 $t D_{i,i}$, where*

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

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

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

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 nonincreasing in q and nondecreasing 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 [8], Section 9.3, and [6], Section VI.10.11), whenever $\theta \in [0, 1]$, $1 \leq p_0, p_1, q_0, q_1 \leq \infty$ and

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

then

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

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

$$\varpi_{q,p}(n) \le \left(2\alpha(n)\right)^{1/p - 1/q},$$

(2.25) $$\varpi_{q,p}(n) \le 2^{1+1/p-1/q}\left(\rho(n)\right)^{1-1/p+1/q} \quad \text{and}$$

$$\varpi_{q,p}(n) \le 2^{1+1/p}\left(\phi(n)\right)^{1-1/p}.$$

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

(2.26) $$\beta(q, r) \le 2^{1-\alpha}\left[\beta(p, r)\right]^{\alpha}\gamma_{pq(1-\alpha)/(p-q\alpha)}^{1-\alpha}$$

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 $q > p \ge 1$, then

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

and so (2.27) holds true under Assumption 2.1.

Concerning the function $F = F(x_1, \ldots, x_\ell)$, we can take it, for instance, to be a polynomial in $x_1, \ldots, x_\ell$, in particular, $F(x_1, \ldots, x_\ell) = x_1 x_2 \cdots x_\ell$ which leads to a functional central limit theorem for

$$N^{-1/2} \sum_{1 \le n \le [Nt]} X\left(q_1(n)\right) X\left(q_2(n)\right) \cdots X\left(q_\ell(n)\right).$$

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 $X(q_j(n))$, $j = 1, 2, \ldots, \ell$, of the function F there. The realignment in the definition of $\{\xi_{i,N}(t)\}$ for $i \le k$ 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 $i \ge k + 1$, for they will all turn out to be mutually independent in the limit.

The conditions of Theorem 2.3 hold true for many important models. Let, for instance, Υ_n be a Markov chain on a space M satisfying the Doeblin condition (see, e.g., [11], pages 367 and 368) and f_j, $j = 1, \ldots, \ell$, be bounded measurable functions on the space of sequences $x = (x_i, i = 0, 1, 2, \ldots, x_i \in M)$ such that $|f_j(x) - f_j(y)| \le Ce^{-cn}$ provided $x = (x_i)$, $y = (y_i)$ and $x_i = y_i$ for all $i = 0, 1, \ldots, n$, where $c, C > 0$ do not depend on n and j. In fact, some polynomial decay in n will suffice here as well. Let $X(n) = (X_1(n), \ldots, X_\ell(n))$ with $X_j(n) = f_j(\Upsilon_n, \Upsilon_{n+1}, \Upsilon_{n+2}, \ldots)$ and take σ-algebras $\mathcal{F}_{kl}$, $k < l$ generated by $\Upsilon_k, \Upsilon_{k+1}, \ldots, \Upsilon_l$, then our condition will be satisfied considering $\{\Upsilon_n, n \ge 0\}$ 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 an L^2 spectral gap which leads to an exponentially fast decay of the ρ-mixing coefficient.

REMARK 2.4.   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 Theorem 2.3 (as well as its continuous time version Theorem 2.5) 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, that is, $X(n)\Pi = \mu$ 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 taking the unconditional expectation of g 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 [4]), f_j's be either Hölder continuous functions or functions which are constant on elements of a Markov partition and let $X(n) = (X_1(n), \ldots, X_\ell(n))$ 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 $\bigcap_{i=k}^{l} T^i \zeta$. 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 $\varsigma = (\varsigma_i, i \geq 0)$, $\varsigma_i = 1, \ldots, l_0$, 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 bounded Hölder continuous [with respect to the metric $d((\varsigma_i, i \geq 0), (\varsigma_i', i \geq 0)) = \exp(-\min\{j \geq 0 : \varsigma_j \neq \varsigma_j'\})$] 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 [4]). 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 [11] and [10]). The latter enables us to consider the number $N_a(x, n)$, $a = (a_1, \ldots, a_\ell)$ 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, $j = 1, \ldots, \ell$. Then Theorem 2.3 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 systems, in particular, those whose transfer operator has an L^2 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 (including unbounded ones) and dynamical systems with polynomially fast mixing which still satisfy our conditions, but they are more difficult to 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

$$(2.28) \qquad \varpi_{q,p}(t) = \sup_{s \geq 0} \varpi_{q,p}(\mathcal{F}_{-\infty,s}, \mathcal{F}_{s+t,\infty})$$

and

$$(2.29) \qquad \beta(p, t) = \sup_{s \geq 0} \left\| X(s) - E\big[X(s)|\mathcal{F}_{s-t,s+t}\big] \right\|_p,$$

where $\varpi_{q,p}(\mathcal{G}, \mathcal{H})$ is defined by (2.1). We continue to impose Assumption 2.1 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 nonincreasing functions of t. Let $q_1(t) < q_2(t) < \cdots < q_\ell(t)$ be increasing positive functions such that $q_i(t) = it$ for $i = 1, \ldots, k$ 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

$$(2.30) \qquad \lim_{t \to \infty} \big(q_i(t + \gamma) - q_i(t)\big) = \infty \qquad \text{for any } \gamma > 0 \text{ and } i > k.$$

THEOREM 2.5. *Suppose that Assumption 2.1 holds true. Then the distribution of the process*

$$(2.31) \qquad \xi_N(t) = \frac{1}{\sqrt{N}} \int_0^{Nt} F\big(X(q_1(s)), \ldots, X(q_\ell(s))\big) \, ds$$

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 $\eta_i, i > k$ are zero there while $\{\eta_1(t), \ldots, \eta_k(t)\}$ is a k-dimensional Gaussian process having stationary independent increments. The means are 0 and variances and covariances are given by $E[\eta_i(s)\eta_j(t)] = \min(s, t)D_{i,j}, i, j = 1, \ldots, k$. The expressions for these $D_{i,j}$ are provided in Section 6.

The conditions of Theorem 2.5 are satisfied when, for instance, $X(t) = (X_1(t), \ldots, X_\wp(t))$ with $X_j(t) = f_j(\Upsilon_t)$, where Υ_t is either an irreducible continuous time finite state Markov chain or a nondegenerate diffusion process on a compact manifold. Furthermore, Ornstein–Uhlenbeck type processes $X(t)$ produce a class of unbounded processes still satisfying our assumptions. 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 in the proof of conventional central limit theorems the standard tool of suspension flows is usually applied while this does not seem to work in our circumstances and a different approach should be employed here.

REMARK 2.6.   Under stronger mixing and moment conditions it is possible to derive convergence of all moments of $\xi_N(t)$ to the corresponding moments of the limiting Gaussian process $\xi(t)$.

3. **Approximation estimates.**   This section contains estimates which are crucial for our proofs and some of them may also have independent interest beyond this paper. Still, in the first reading the reader can skip this section all together and only refer to general estimates of Corollary 3.6 when needed in what follows. We will make repeated use of the following simple variations of Hölder's inequality.

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

$$\|Z^h D^\kappa\|_a \leq \|Z\|_{a*}^h \|D\|_{b*}^\kappa$$

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

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

provided $\frac{1}{a} \geq \frac{h+\kappa-\alpha}{a*} + \frac{\alpha}{b*}$.
  (ii) *If $f(x, \omega)$ is a measurable function of x and ω such that for almost all ω,*

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

then

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

provided $\frac{1}{a} \geq \frac{1}{p} + \frac{h}{m}$ where γ_m is a bound for $\|X\|_m$.
  (iii) *If $f(x, \omega)$ is a measurable function of x and ω satisfying for almost all ω,*

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

then

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

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

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

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

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

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

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

For (ii), by Hölder's inequality,

$$E\left[\left|f\left(X(\omega), \omega\right)\right|^a\right] \le E\left[\left[C(\omega)\right]^a \left[1 + |X|^h\right]^a\right]$$
$$\le \left[E\left[\left[C(\omega)\right]^p\right]\right]^{a/p} \left[E\left[\left[1 + |X|^h\right]^{p^*/h}\right]\right]^{ah/p^*}$$

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

The assertion (iii) follows similarly from the inequality

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

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

We will need also the following.

LEMMA 3.2. (i) *Let $F(x_1, \ldots, x_{\ell-1}, x_\ell)$ be any function that satisfies* (2.6) *and* (2.7). *Then the functions $F_i(x_1, \ldots, x_i)$ 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*

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

then

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

and

$$\left|G_i(x_1, \ldots, x_{i-1}, \omega) - G_i(y_1, \ldots, y_{i-1}, \omega)\right|$$
$$\le C\left(1 + C(\omega)^\iota + |x|^\iota + |y|^\iota\right)|x - y|^\kappa,$$

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

PROOF. For (i), if

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

then

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

The Hölder property is similar.

Assertion (ii) follows from

$$\left|G_i(x_1, \ldots, x_{i-1}, \omega)\right| \leq E\big[|F_i(x_1, \ldots, x_{i-1}, Z)|\,|\mathcal{G}\big] \leq C_1 E\big[(C_2 + |x|^\iota + |Z|^\iota)|\mathcal{G}\big]$$

and

$$\left|G_i(x_1, \ldots, x_{i-1}, \omega) - G_i(y_1, \ldots, y_{i-1}, \omega)\right|$$
$$\leq E\big[|F_i(x_1, \ldots, x_{i-1}, Z) - F_i(y_1, \ldots, y_{i-1}, Z)|\,|\mathcal{G}\big]$$
$$\leq C E\big[(1 + |x|^\iota + |y|^\iota + 2|Z|^\iota)|\mathcal{G}\big]|x - y|^\kappa. \qquad \square$$

REMARK 3.3.   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 $b_1, b_2, \ldots, b_l \geq 0$ and $\gamma > 0$,

$$(3.2) \qquad \max_{1 \leq i \leq l} b_i^\gamma \leq \sum_{i=1}^l b_i^\gamma \leq l \max_{1 \leq i \leq l} b_i^\gamma \leq l\left(\sum_{i=1}^l b_i\right)^\gamma \leq l^{1+\gamma} \max_{1 \leq i \leq l} b_i^\gamma.$$

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.

THEOREM 3.4.   Let $f(x, \omega)$ be a collection of random variables continuously (or separable) dependent on $x \in \mathbb{R}^d$ for almost all ω and satisfying

$$(3.3) \qquad \begin{aligned} \|f(x, \omega) - f(y, \omega)\|_p &\leq C_1\big(1 + |x|^\iota + |y|^\iota\big)|x - y|^\kappa \quad \text{and} \\ \|f(x, \omega)\|_p &\leq C_2\big(1 + |x|^\iota\big) \end{aligned}$$

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

$$(3.4) \qquad \begin{aligned} |f(x, \omega)| &\leq G(\omega)\big(1 + |x|^{\iota'}\big) \\ \text{a.e. with } \|G(\omega)\|_p &\leq c_0[C_1 + C_2]^{d/(p\theta)} C_2^{1-d/(p\theta)}, \end{aligned}$$

where $c_0 = c_0(d, p, \kappa, \theta, \iota, \iota') > 0$ depends only on parameters in brackets. Since $\kappa \leq 1$ 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 $\|Z\|_m \leq \gamma_m$ and if $\frac{1}{a} \geq \frac{1}{p} + \frac{\iota+1}{m}$, then*

$$\|f(Z(\omega), \omega)\|_a \leq \|G(\omega)(1 + |Z|^{\iota+1})\|_a$$

(3.5)
$$\leq c_0[C_1 + C_2]^{d/(p\theta)} C_2^{1-d/(p\theta)}[1 + \gamma_m^{\iota+1}]$$

$$= c_0 c(\gamma_m)[C_1 + C_2]^{d/(p\theta)} C_2^{1-d/(p\theta)}.$$

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

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

with

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

and the estimate

$$\|f(X_1, X_2, \ldots, X_{i-1}, \omega) - f(Y_1, Y_2, \ldots, Y_{i-1}, \omega)\|_a$$

(3.6)
$$\leq \|H(\omega)[1 + |X|^{\iota+2} + |Y|^{\iota+2}]|X - Y|^\delta\|_a$$

$$\leq \|H\|_p (1 + \gamma_m^{\iota+2}) \sum_{j=1}^{i-1} \|X_j - Y_j\|_q^\delta$$

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

REMARK 3.5. 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 OF THEOREM 3.4. 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 $|x - y| \leq \rho_0 = \frac{\sqrt{d}}{2}$,

$$\|\tilde{f}(x, \omega) - \tilde{f}(y, \omega)\|_p$$

(3.7)
$$\leq \|f(x, \omega) - f(y, \omega)\|_p (1 + |x|^{\iota+1})^{-1}$$

$$+ \|f(y, \omega)\|_p ||y|^{\iota+1} - |x|^{\iota+1}|\eta(x)$$

$$\leq c_1[C_1 + C_2]|x - y|^\kappa \eta(x)$$

and

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

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 $B_w(\rho)$ denote 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 [15], page 60) states that if a continuous (or separable) $g : \mathbb{R}^d \to \mathbb{R}$ satisfies

$$\int_{B_w(\rho) \times B_w(\rho)} \Psi\left(\frac{|g(x) - g(y)|}{\sigma(|x - y|)}\right) dx\, dy \leq Q_{w,\rho}$$

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

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

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

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

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 $x, y \in B_w(\rho)$,

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

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

$$\|Q_{w,\rho}\|_p \leq c_2 v_d (C_1 + C_2) \eta(w) \rho^{(\kappa - \theta)}, \tag{3.11}$$

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) are, in fact, the conclusion of a multidimensional version of the Kolmogorov theorem (see, e.g., [14], 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

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

and so

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

Therefore,

$$\sup_{x\in\mathbb{R}^d}|\tilde{f}(x,\omega)|^p \leq 2^{p-1}\sup_{w\in\mathbb{Z}_h^d}\big[|\tilde{f}(w,\omega)|^p + c_2^p Q_{w,h\rho_0}^p(\omega)\rho_0^{p\theta}h^{p\theta}\big]$$

$$\leq 2^{p-1}\sum_{w\in\mathbb{Z}_h^d}\big[|\tilde{f}(w,\omega)|^p + c_2^p Q_{w,h\rho_0}^p(\omega)\rho_0^{p\theta}h^{p\theta}\big]$$

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

$$E\Big[\big[\sup_{\mathbb{R}^d}|\tilde{f}(x,\omega)|^p\big]\Big] \leq 2^{p-1}\sum_{w\in\mathbb{Z}_h^d}\|\tilde{f}(w,\omega)\|_p^p$$

$$+ 2^{p-1}c_2^p\rho_0^{p\theta}h^{p\theta}\sum_{w\in\mathbb{Z}_h^d}\|Q_{w,h\rho_0}(\omega)\|_p^p$$

$$\leq c_3^p\big[C_2^p + (C_1+C_2)^p h^{p\kappa}\big]\sum_{w\in\mathbb{Z}_h^d}[\eta(w)]^p$$

$$\leq c_5^p\big[C_2^p + (C_1+C_2)^p h^{p\kappa}\big]h^{-d}$$

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

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

Now set

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

Then

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

and so

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

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,

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

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

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

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 $|x-y| \leq 1$,

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

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

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

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

$$\big\|f(X_1,X_2,\ldots,X_{i-1},\omega)-f(Y_1,Y_2,\ldots,Y_{i-1},\omega)\big\|_a$$

$$\leq \big\|H(\omega)\big[1+|X|^{\iota+2}+|Y|^{\iota+2}\big]|X-Y|^\delta\big\|_a$$

$$\leq K\|H\|_p\big(1+\gamma_m^{\iota+2}\big)\sum_{j=1}^{i-1}\|X_j-Y_j\|_q^\delta$$

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

In our nonconventional setup Theorem 3.4 will be applied in the form of the following useful result.

COROLLARY 3.6. *Let* $\mathcal{G}$ *and* $\mathcal{H}_1 \subset \mathcal{H}_2$ *be* σ-*subalgebras on a probability space* $(\Omega,\mathcal{F},P)$, X *and* Y *be* d-*dimensional random vectors and* $f_i = f_i(x,\omega), i=1,2,$ *be collections of random variables that are continuously (or separable) dependent on* $x \in \mathbb{R}^d$ *for almost all* ω, *measurable with respect to* $\mathcal{H}_i$, $i=1,2,$ *respectively, and satisfy*

$$\begin{aligned}(3.12)\quad &\|f_i(x,\omega)-f_i(y,\omega)\|_q \leq C_1\big(1+|x|^\iota+|y|^\iota\big)|x-y|^\kappa \quad and\\ &\|f_i(x,\omega)\|_q \leq C_2\big(1+|x|^\iota\big).\end{aligned}$$

Set $\tilde{f}_i(x,\omega)=E[f_i(x,\cdot)|\mathcal{G}](\omega)$ *and* $g_i(x)=E[f_i(x,\omega)]$.

(i) *Assume that* $q \geq p, 1 \geq \kappa > \theta > \frac{d}{p}$ *and* $\frac{1}{a} \geq \frac{1}{p}+\frac{\iota+1}{m}$. *Then for* $i=1,2,$

$$\begin{aligned}(3.13)\quad &\big\|\tilde{f}_i\big(X(\omega),\omega\big)-g_i(X)\big\|_a\\ &\qquad\leq c\varpi_{q,p}(\mathcal{G},\mathcal{H}_i)(C_1+C_2)^{d/(p\theta)}C_2^{1-d/(p\theta)}\big(1+\|X\|_m^{\iota+1}\big),\end{aligned}$$

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

(ii) *Next, assume that* $\frac{1}{a} \geq \frac{1}{p}+\frac{\iota+2}{m}+\frac{\delta}{q}$. *Then for* $i=1,2,$

$$\begin{aligned}(3.14)\quad &\big\|E\big[f_i(X,\cdot)|\mathcal{G}\big]-g_i(X)\big\|_a\\ &\qquad\leq R+2c(C_1+C_2)\big(1+2\|X\|_m^{\iota+2}\big)\big\|X-E[X|\mathcal{G}]\big\|_q^\delta,\end{aligned}$$

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

(iii) *Furthermore, 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 $f_i(x, \omega) = f_i(v, z, \omega)$ satisfy* (3.12) *in $x = (v, z)$. Set $\tilde{g}_i(v) = E[f_i(v, Z(\omega), \omega)]$. Then for $i = 1, 2$,*

$$
\begin{aligned}
& \left\| E[f_i(V, Z, \cdot)|\mathcal{G}] - \tilde{g}_i(V) \right\|_a \\
& \quad \leq c\left(1 + \|X\|_m^{\iota+2}\right) \\
(3.15) \quad & \qquad \times \left(\varpi_{q,p}(\mathcal{G}, \mathcal{H}_i)(C_1 + C_2)^{d_1/(p\theta)} C_2^{1-d_1/(p\theta)} \right. \\
& \qquad\qquad \left. + \left\| V - E[V|\mathcal{G}] \right\|_q^\delta + \left\| Z - E[Z|\mathcal{H}_i] \right\|_q^\delta\right).
\end{aligned}
$$

(iv) *Finally, for $a, p, q, \iota, m, \delta$ satisfying conditions of* (ii),

$$
\begin{aligned}
(3.16) \quad & \left\| \tilde{f}_1(X(\omega), \omega) - \tilde{f}_2(Y(\omega), \omega) - g_1(X) + g_2(Y) \right\|_a \\
& \quad \leq c\varpi_{q,p}(\mathcal{G}, \mathcal{H}_2)\left(1 + \|X\|_m^{\iota+2} + \|Y\|_m^{\iota+2}\right)\|X - Y\|_q^\delta,
\end{aligned}
$$

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

PROOF. (i) Set $h(x, \omega) = \tilde{f}_i(x, \omega) - g_i(x)$, $K_1 = C_1\varpi_{q,p}(\mathcal{G}, \mathcal{H}_i)$ and $K_2 = C_2\varpi_{q,p}(\mathcal{G}, \mathcal{H}_i)$. Then by (3.12) and the definition of $\varpi_{q,p}$ for all $x, y \in \mathbb{R}^d$ and $q, p \geq 1$,

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

and

$$
(3.18) \quad \left\| h(x, \omega) \right\|_p \leq \varpi_{q,p}(\mathcal{G}, \mathcal{H}_i)\left\| f_i(x, \omega) - g_i(x) \right\|_q \leq 2K_2\left(1 + |x|^\iota\right).
$$

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).

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

$$
\begin{aligned}
& \left\| E[f_i(X, \cdot)|\mathcal{G}] - g_i(X) \right\|_a \\
(3.19) \quad & \quad \leq \left\| E[f_i(\tilde{X}, \cdot)|\mathcal{G}] - g_i(\tilde{X}) \right\|_a \\
& \qquad + \left\| E[f_i(\tilde{X}, \cdot)|\mathcal{G}] - E[f_i(X, \cdot)|\mathcal{G}] \right\|_a + \left\| g_i(\tilde{X}) - g_i(X) \right\|_a \\
& \quad \leq \left\| \tilde{f}_i(\tilde{X}, \cdot) - g_i(\tilde{X}) \right\|_a + \left\| f_i(\tilde{X}, \cdot) - f_i(X, \cdot) \right\|_a + \left\| g_i(\tilde{X}) - g_i(X) \right\|_a.
\end{aligned}
$$

We can estimate the first term in the right-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),

$$
(3.20) \quad \left\| f_i(\tilde{X}, \omega) - f_i(X, \omega) \right\|_a \leq cC_1\left(1 + \gamma_m^{\iota+2}\right)\|\tilde{X} - X\|_q^\delta.
$$

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

$$\left| g_i(x) - g_i(y) \right| \le c \big[1 + |x|^\iota + |y|^\iota \big] |x - y|^\kappa$$

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

$$\left\| g_i(X) - g_i(\tilde{X}) \right\|_a \le c \big(1 + \gamma_m^{\iota+2} \big) \| \tilde{X} - X \|_q^\delta.$$

(iii) Set $\tilde{V} = E[V|\mathcal{G}]$, $\tilde{Z} = E[Z|\mathcal{H}]$, $\tilde{\tilde{g}}_i(v) = E[f_i(v, \tilde{Z}, \cdot)]$ and $\tilde{g}_i(v) = E[f_i(v, Z, \cdot)]$. Then

$$(3.21) \quad \begin{aligned} &\left\| E\big[f_i(V, Z, \cdot)|\mathcal{G} \big] - \tilde{g}_i(V) \right\|_a \\ &\qquad \le \left\| f_i(V, Z, \cdot) - f_i(V, \tilde{Z}, \cdot) \right\|_a \\ &\qquad\quad + \left\| E\big[f_i(V, \tilde{Z}, \cdot)|\mathcal{G} \big] - \tilde{\tilde{g}}_i(V) \right\|_a + \left\| \tilde{\tilde{g}}_i(V) - \tilde{g}_i(V) \right\|_a. \end{aligned}$$

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

$$\begin{aligned} \left| \tilde{\tilde{g}}_i(v) - \tilde{g}_i(v) \right| &\le E\big[\big| f_i(v, \tilde{Z}, \cdot) - f_i(v, Z, \cdot) \big| \big] \\ &\le E\big[(1 + |v|^\iota + |Z|^\iota + |\tilde{Z}|^\iota) |Z - \tilde{Z}|^\kappa \big] \end{aligned}$$

and then substituting V in place of v there.

(iv) Set $\hat{h}(x, \omega) = \tilde{f}_1(x, \omega) - \tilde{f}_2(x, \omega) - g_1(x) + g_2(x)$, $\hat{K}_1 = C_1 \varpi_{q,p}(\mathcal{G}, \mathcal{H}_2)$ and $\hat{K}_2 = C_2 \varpi_{q,p}(\mathcal{G}, \mathcal{H}_2)$. Then by (3.12) and the definition of $\varpi_{q,p}$ for all $x, y \in \mathbb{R}^d$ and $q, p \ge 1$,

$$(3.22) \quad \begin{aligned} &\left\| \hat{h}(x, \omega) - \hat{h}(y, \omega) \right\|_p \\ &\qquad \le \varpi_{q,p}(\mathcal{G}, \mathcal{H}_2) \big\| f_1(x, \omega) - f_1(y, \omega) - g_1(x) + g_1(y) \\ &\qquad\qquad\qquad - f_2(x, \omega) + f_2(y, \omega) + g_2(x) - g_2(y) \big\|_q \\ &\qquad \le 2\hat{K}_1 \big(1 + |x|^\iota + |y|^\iota \big) |x - y|^\kappa \end{aligned}$$

and

$$(3.23) \quad \begin{aligned} \left\| \hat{h}(x, \omega) \right\|_p &\le \varpi_{q,p}(\mathcal{G}, \mathcal{H}_2) \big\| f_1(x, \omega) - f_2(x, \omega) - g_1(x) + g_2(x) \big\|_q \\ &\le 2\hat{K}_2 \big(1 + |x|^\iota \big). \end{aligned}$$

Now (3.22) and (3.23) enable us to apply (3.6), which yields (3.16). $\square$

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

$$(3.24) \qquad \frac{1}{a} \geq \frac{1}{p} + \frac{\iota + 2}{m} + \frac{\delta}{q}.$$

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

4. Limiting covariances. Here and in what follows we set $Y_{i,q_i(n)} = F_i(X(q_1(n)), \ldots, X(q_i(n)))$ and $Y_{i,m} = 0$ if $m \neq q_i(n)$ for any n. Let $F_{i,n,r}(x_1, x_2, \ldots, x_{i-1}, \omega) = E[F_i(x_1, x_2, \ldots, x_{i-1}, X(n))|\mathcal{F}_{n-r,n+r}]$ and $X_r(n) = E[X(n)|\mathcal{F}_{n-r,n+r}]$. We denote also $Y_{i,q_i(n),r} = F_{i,q_i(n),r}(X_r(q_1(n)), \ldots, X_r(q_{i-1}(n)), \omega)$ and $Y_{i,m,r} = 0$ if $m \neq q_i(n)$ for any n. In view of (2.6), we can and will always choose continuous in $(x_1, \ldots, x_{i-1})$ versions of conditional expectations $F_{i,n,r}$ which will enable us to apply Corollary 3.6 when needed.

In this section we will study the asymptotical behavior of covariances

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

of the processes $\{\xi_{i,N}(t)\}$ defined by (2.21) and (2.22). 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.

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

$$\lim_{N \to \infty} E[\xi_{i,N}(s)\xi_{j,N}(t)]$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{\substack{0 \leq in \leq Ns \\ 0 \leq jl \leq Nt}} E[F_i(X(n), X(2n), \ldots, X(in))$$

$$\times F_j(X(l), X(2l), \ldots, X(jl))]$$

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

$$A_{i,j}(x_{i'}, x_{2i'}, \ldots, x_{\upsilon i'}, y_{j'}, y_{2j'}, \ldots, y_{\upsilon j'})$$

$$= \int F_i(x_1, \ldots, x_{i-1}, x_i)$$

$$\times F_j(y_1, \ldots, y_{j-1}, y_j) \prod_{\sigma \notin \{i', 2i', \ldots, \upsilon i'\}} d\mu(x_\sigma) \prod_{\sigma' \notin \{j', 2j', \ldots, \upsilon j'\}} d\mu(y_{\sigma'})$$

and

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

Then

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

where

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

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 $|n - l| \gg 1$, 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 also be used later on.

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

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

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 $i = 1, \ldots, \ell$,

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

PROOF. First, observe that for $i = 1, \ldots, k$,

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

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 $n > l \geq 0$,

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

and so

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

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 $r = \frac{1}{3}s_{i,j}(n,l) = \frac{1}{3}\hat{s}_{i,j}(n,l)$. 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 at the beginning of this section, then the difference between $b_{i,j}(n,l)$ and

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

can be estimated easily using Corollary 3.6(iv) with $\mathcal{H}_2 = \mathcal{F}$, which gives

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

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 Corollary 3.6(i),

$$\begin{aligned}
\left| b_{i,j}^{(r)}(n,l) \right| &= \left| EY_{i,q_i(n),r}Y_{j,q_j(l),r} \right| \\
&= \left| E\left[E[Y_{i,q_i(n),r}|\mathcal{F}_{0,q_i(n)-r}]Y_{j,q_j(l),r} \right] \right| \\
&\leq \left\| F_j\left(X_r(q_1(l)), \ldots, X_r(q_j(l)) \right) \right\|_{L_2(P)} \\
&\quad \times \left\| E[Y_{i,q_i(n),r}|\mathcal{F}_{0,q_i(n)-r}] \right\|_{L_2(P)} \\
&\leq C\varpi_{q,p}(\tfrac{1}{3}s_{i,j}(n,l)).
\end{aligned}$$

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
E\left[|\xi_{i,N}(t) - \xi_{i,N}(s)|^2 \right] &\leq \frac{1}{N}\left(\sum_{Ns \leq l \leq Nt} b_{i,i}(l,l) + 2 \sum_{\substack{Ns \leq l \leq Nt \\ n \geq l+1}} |b_{i,i}(n,l)| \right) \\
&\leq \frac{1}{N} \sum_{Ns \leq l \leq Nt} \left(EY_{i,l}^2 + 2 \sum_{n \geq l+1} h(n - l) \right) \leq Ct
\end{aligned}$$

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

Next, we will need a result which will be formulated in a somewhat more general situation. Let $H(x_1, x_2, \ldots, x_d)$ be a function on $(R^\nu)^d$ that is continuous and satisfies the growth condition $|H(x_1, x_2, \ldots, x_d)| \leq 1 + \sum_i \|x_i\|^\iota$ 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 $\{Y(n_1), Y(n_2), \ldots, Y(n_l)\}$ depends only on the spacings $\{n_i - n_{i-1}\}, l = 2, \ldots, l$ 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

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

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 $n_1 < n_2 < \cdots < n_d$ be a sequence of integers that tend to ∞ with some of the gaps $\{n_{i+1} - n_i\}$ tending to infinity while others are kept fixed. This splits the set of integers $1, 2, \ldots, d$ 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 denote the number of integers in a block B_j and S_j denote the set of spacings in B_j, that is, the 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.

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

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

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 cutoff 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 $B = \{n_{d'+1}, \ldots, n_d\}$ at the end with a rigid spacing S between integers in the block. We will show that

$$(4.7) \qquad \lim_{\substack{n_1,\ldots,n_d \to \infty \\ \mathcal{P} \text{ fixed}}} E\big[\hat{H}(X(n_1), \ldots, X(n_d))\big] = 0,$$

where

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

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

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

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 $\int H(x_1, \ldots, x_d)\, d\mu_{\mathcal{P}}$, which will complete the proof of the lemma.

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

$$\big\| \hat{H}(X(n_1), \ldots, X(n_d)) - \hat{H}(X_r(n_1), \ldots, X_r(n_d)) \big\|$$
$$\leq C \sup_n \| X_r(n) - X(n) \|_{L_1(P)} \to 0$$

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

$$(4.8) \qquad \lim_{r \to \infty} \limsup_{n_1,\ldots,n_d \to \infty} E\big[\hat{H}(X_r(n_1), \ldots, X_r(n_d))\big] = 0.$$

Observe that

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

where

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

To prove (4.8), it is clearly sufficient to show that

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

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

$$\sup_{x_1,\ldots,x_{d'}} \left| G_r(x_1,\ldots,x_{d'},\omega) \right| \leq C \int \left| G_r(x_1,\ldots,x_{d'},\omega) \right| dx_1 \cdots dx_{d'}.$$

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

$$E \sup_{x_1,\ldots,x_{d'}} \left| G_r(x_1,\ldots,x_{d'},\omega) \right| \leq C L^{d'} \sup_{x_1,\ldots,x_{d'}} E \left| G_r(x_1,\ldots,x_{d'},\omega) \right|.$$

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

$$\sup_{x_1,\ldots,x_{d'}} \left\| G_r(x_1,\ldots,x_{d'},\omega) - \hat{H}_r(x_1,\ldots,x_{d'}) \right\|_1$$

$$\leq 2\varpi_{\infty,1}(n_{d'+1} - n_{d'} - 2r)\|H\|_\infty,$$

where

$$\hat{H}_r(x_1,\ldots,x_{d'}) = E\left[\hat{H}(x_1,\ldots,x_{d'}, X_r(n_{d'+1}),\ldots,X_r(n_d)) \right],$$

while

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

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

$$\left| E\left[\hat{H}(x_1,\ldots,x_{d'}, X(n_{d'+1}),\ldots,X(n_d)) \right] \right.$$

$$- E\left[\hat{H}(x_1,\ldots,x_{d'}, X_r(n_{d'+1}),\ldots,X_r(n_d)) \right] \Big|$$

$$\leq C \sup_n E \left| X(n) - X_r(n) \right| = \epsilon(r) \to 0 \qquad \text{as } r \to \infty.$$

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

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

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

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

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

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

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

and 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 $n, 2n, \ldots, in$ and $l, 2l, \ldots, jl$ and ordering them into a single sequence, we obtain by employing Lemma 4.3 that

$$\lim_{\substack{n,l \to \infty \\ in-jl=\upsilon u}} b_{i,j}(n,l)$$

$$= \lim_{\substack{n,l \to \infty \\ in-jl=\upsilon u}} E\big[F_i\big(X(n), X(2n), \ldots, X(in)\big) F_j\big(X(l), X(2l), \ldots, X(jl)\big)\big]$$

$$= a_{i,j}(u, 2u, \ldots, \upsilon u).$$

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 $(\alpha m n, \beta m l)$, $m = 1, \ldots, \upsilon$, with $\alpha m n - \beta m l = mu$ will give υ blocks of size 2. There cannot 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 $\alpha \upsilon n \le Nt$ and $\beta \upsilon l \le Ns$. 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 constraints 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 an 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). $\square$

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

PROPOSITION 4.5. *For $i \ge k + 1$,*

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

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

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

PROOF.   It follows from (4.6) that

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

and so, by (4.2),

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

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 $\limsup$ in the left-hand side above equals zero, that is, 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 $j = 2, 3, \ldots, \ell$ as $n \to \infty$, it follows from Lemma 4.3 that

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

proving (4.12).

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

$$\big|E\xi_{i,N}(t)\xi_{j,N}(s)\big|$$

$$(4.15) \qquad \leq \big|E\xi_{i,N}(\varepsilon T)\xi_{j,N}(s)\big| + \big|E\big(\xi_{i,N}(t) - \xi_{i,N}(\varepsilon T)\big)\xi_{j,N}(s)\big|$$

$$\leq \big(E\xi_{i,N}^2(\varepsilon T)\big)^{1/2}\big(E\xi_{j,N}^2(s)\big)^{1/2} + \frac{1}{N} \sum_{\varepsilon NT \leq n \leq NT, 1 \leq l \leq NT} |b_{i,j}(n, l)|$$

$$\leq CT\sqrt{\varepsilon} + \frac{1}{N} \sum_{\varepsilon NT \leq n \leq NT, 1 \leq l \leq NT} h\big(s_{i,j}(n, l)\big).$$

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 $N \geq N(\varepsilon), n \geq \varepsilon NT, l \leq NT$ and, moreover, by (4.5),

$$s_{i,j}(n, l) = \min(q_i(n) - q_j(l), n)$$

$$\geq \min(q_i(n) - q_i(\varepsilon NT) + \varepsilon^{-1}, n) \geq \min(n - \varepsilon NT + \varepsilon^{-1}, n).$$

Hence,

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

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

5. Proof of the main theorem. The proof of Theorem 2.3 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, e.g., Section 2 in Chapter VIII of [12] and close versions in Theorem 18.2 in [3] and Theorem 4.1 in [9]). For each N let $\mathcal{G}_{N,n}, n = 1, 2, \ldots$, be a filtration of σ-algebras and let $\{U_{N,n} : n \geq 1\}$ be a triangular array of random variables satisfying the following conditions:

B1. For every N, $\{U_{N,n}\}$ is adapted to some $(\Omega_N, \mathcal{G}_{N,n}, P_N), n = 1, 2, \ldots$;

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

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

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

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

where

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

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

THEOREM 5.1. *Under assumptions* B1–B4,

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

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

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

as can be seen from the following result.

LEMMA 5.2. *If for a fixed l the random variables*

$$V_{N,r} = \left(\sum_{r(l-1)+1 \leq n \leq rl} U_{N,n} \right)^2$$

satisfy a uniform law of large numbers in the sense that

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

then (5.1) *implies* B4.

PROOF. We begin with the observation that if $\eta_n, n \geq 1$ 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 $\tilde{\eta}_n = \hat{\eta}_n + \bar{\eta}_n$ with $\|\bar{\eta}_n\|_{L_1(P)} \leq \|\eta_n - \tilde{\eta}_n\|_{L_1(P)}$ and $\hat{\eta}_n$ being a martingale difference with a uniformly bounded second moment.

We will now compare

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

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

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

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 (in r) 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

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

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

On the other hand, $W_{N,n} = U_{N,n} - R_{N,n-1} + R_{N,n}$, where

$$R_{N,n} = \sum_{m \geq n+1} E[U_{N,m}|\mathcal{G}_{N,n}]$$

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. $\square$

REMARK 5.3. Let the filtration $\mathcal{F}_{m,n}$ satisfy any mixing condition, that is, $\varpi_{p,q}(k) \to 0$ 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 (centered) 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.

COROLLARY 5.4. *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 a routine argument by contradiction. 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. $\square$

REMARK 5.5. For each N let $\mathcal{G}_{N,n}, n = 1, 2, \ldots,$ be a filtration of σ-algebras and let $k_N \geq 1, N = 1, 2, \ldots,$ be an integer sequence with $k_N \to \infty$ as $N \to \infty$. One way to generate new triangular arrays for $N = 1, 2, \ldots,$ 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 $(\Omega_N, \tilde{\mathcal{G}}_{N,n}, \tilde{U}_{N,n}, P_{N,B_N}), n = 1, 2, \ldots,$ 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 $k_N \le CN$ 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.

THEOREM 5.6.   *Let $\mathcal{X}$ be a complete separable metric space and for each $N \ge 1$ 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*

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

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

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

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 $A(t)$,

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

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 $\frac{\mu(E \times F)}{\lambda(E)} = \gamma(F)$. □

COROLLARY 5.7. *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,*

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

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 to 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 2.3 into several steps and use the following representations:

$$Y_{i,q_i(n)} = Y_{i,q_i(n),1} + \sum_{r=1}^{\infty} [Y_{i,q_i(n),2^r} - Y_{i,q_i(n),2^{r-1}}],$$

$$\zeta_{i,N,0}(t) = \frac{1}{\sqrt{N}} \sum_{1 \leq n \leq M_i(Nt)} Y_{i,q_i(n),1},$$

(5.2)

$$\zeta_{i,N,r}(t) = \frac{1}{\sqrt{N}} \sum_{1 \leq n \leq M_i(Nt)} [Y_{i,q_i(n),2^r} - Y_{i,q_i(n),2^{r-1}}], \qquad r \geq 1, \quad \text{and}$$

$$\xi_{i,N}(t) = \sum_{r=1}^{\infty} \zeta_{i,N,r}(t),$$

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

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

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

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 $|F| \leq C(1 + \sum |x_i|^\iota)$, it is easily seen that $\{Y_{i,q_i(n),r}\}$ are uniformly square integrable. To control $\|E[Y_{i,q_i(n),r}|\mathcal{F}_{-\infty,l}]\|_2$, we use Corollary 3.6(ii) for $q_{i-1}(n)+r \leq l$, which yields the estimate

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

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

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

whenever $n \geq 2r$ and $n \geq n^* = n^*(i) = \min\{m : q_i(l) - q_{i-1}(l) \geq l \ \forall l \geq m\}$, observing that $n^* < \infty$ by (2.12). Assuming that $q \geq p$, 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

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

If we set

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

then it follows from the above estimates that

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

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. $\quad\square$

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).

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

$$(5.4) \qquad \sum_r \sup_{N \geq 1} \max_{1 \leq i \leq \ell} \left\| \sup_{0 \leq t \leq T} |\zeta_{i,N,r}(t)| \right\|_2 \leq C < \infty.$$

PROOF. Set $\tilde{Y}_{i,n,r} = Y_{i,n,2^r} - Y_{i,n,2^{r-1}}, r \geq 1$ and

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

Estimating conditional expectations here by Corollary 3.6(iv) when $m - n \geq 2^{r+1}$ and by the contraction argument when $n + 1 \leq m \leq n + 2^{r+1}$, and applying Corollary 3.6(iv) after that again, we obtain

$$\|\tilde{R}_{i,n,r}\|_2 \leq 2^{r+1} \sup_n \|\tilde{Y}_{i,n,r}\|_2 + \tilde{C}((\beta(q, 2^r))^{\delta} + (\beta(q, 2^{r-1}))^{\delta})$$

$$(5.5)$$

$$\leq \hat{C}2^r((\beta(q, 2^r))^{\delta} + (\beta(q, 2^{r-1}))^{\delta}),$$

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

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

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

$$
\begin{aligned}
\frac{1}{N} E \sup_{0 \le t \le T} \left| \sum_{1 \le l \le Nt} Z_{i,q_i(l),r} \right|^2 &\le \frac{4}{N} \sum_{1 \le l \le NT} E Z_{i,q_i(l),r}^2 \\
(5.7) &\le 4T \max_{1 \le l \le NT} E Z_{i,q_i(l),r}^2 \\
&\le 12T \left(\sup_n \| \tilde{Y}_{i,n,r} \|_2 + 2 \sup_n \| \tilde{R}_{i,n,r} \|_2 \right).
\end{aligned}
$$

We can estimate also

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

Now collecting (5.5)–(5.8) and applying Corollary 3.6(iv) again to (5.7) and (5.8), we obtain that

$$(5.9) \qquad \sup_{N \ge 1} \left\| \sup_{0 \le t \le T} |\zeta_{i,N,r}(t)| \right\|_2 \le \tilde{\tilde{C}} 2^r \left((\beta(q, 2^r))^\delta + (\beta(q, 2^{r-1}))^\delta \right),$$

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

Next, we deal specifically with the terms $Y_{i,q_i(n)}$, $k + 1 \le i \le \ell$ 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

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

for $1 \le i \le \ell$ will be a Gaussian process with independent increments. The processes $\{\eta_i(\cdot), k + 1 \le i \le \ell\}$ will be mutually independent as well as totally independent of $\{\eta_i(\cdot), 1 \le i \le k\}$, 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 $\sum_{n \le k_N(i)} Y_{i,q_i(n)}$ in the definition of $\xi_{i,N}(t)$, where $k_N(i) = \max\{n : q_i(n) \le$

$q_{i-1}(NT)$} 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 at the beginning of Section 4. 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 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 until Nt or until $Nt + k_N(i)$ 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 $\eta_i, i > k$, of each other and of $\eta_i, i \leq k$, 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 the 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 2.3. 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 $\eta_{k+1}(t), \ldots, \eta_\ell(t)$ are both mutually independent and independent of processes $\eta_1(t), \ldots, \eta_k(t)$. 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 Propositions 4.1 and 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 $(\eta_1(t), \ldots, \eta_k(t))$ is a k-dimensional Gaussian process with independent increments, then the vector increments $(\eta_j(it) - \eta_j((i-1)t), j = 1, 2, \ldots, k)$ for $i = 1, 2, \ldots, k$ 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

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

Then by Proposition 4.1,

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

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 $X(0), X(1), X(2), \dots$ 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 $F(x, y) = x^2 y^2 - 1$. 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

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

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

6. Continuous time case. First, we represent again the function F in the form (2.17) and $\xi_N(t)$ given by (2.31) in the form (2.20) where now

$$(6.1) \qquad \xi_{i,N}(t) = \frac{1}{\sqrt{N}} \int_0^{S_i(Nt)} F_i\big(X(q_1(s)), \dots, X(q_i(s))\big) \, ds$$

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

$$F_{i,r,t} = F_{i,r,t}(x_1, \ldots, x_{i-1}, \omega) = E\big(F_i(x_1, \ldots, x_{i-1}, X(t))|\mathcal{F}_{t-r,t+r}\big),$$

$$X_r(t) = E\big(X(t)|\mathcal{F}_{t-r,t+r}\big),$$

$$Y_i(t) = F_i\big(X(q_1(s)), \ldots, X(q_i(s))\big) \qquad \text{if } t = q_i(s)$$

and

$$Y_i(t) = 0 \qquad \text{if } t \neq q_i(s) \qquad \text{for any } s,$$

$$Y_{i,r}(t) = F_{i,r,t}\big(X_r(q_1(s)), \ldots, X_r(q_i(s))\big) \qquad \text{if } t = q_i(s)$$

and

$$Y_{i,r}(t) = 0 \qquad \text{if } t \neq q_i(s) \qquad \text{for any } s.$$

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

$$(6.2) \qquad \sup_{0 \leq t \leq T} |\xi_{i,N}(t) - \tilde{\xi}_{i,N}(t)| \leq \frac{1}{\sqrt{N}} \max_{0 \leq n \leq NT} Q_i(n),$$

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

$$P\Big\{ \max_{0 \leq n \leq NT} Q_i(n) > \varepsilon\sqrt{N} \Big\} \leq NT \max_{0 \leq n \leq NT} P\{Q_i(n) > \varepsilon\sqrt{N}\}$$

$$\leq \frac{T}{\varepsilon^2} \max_{0 \leq n \leq NT} \int_{\{Q_i(n) > \varepsilon\sqrt{N}\}} Q_i^2(n) \, dP$$

$$\leq (\varepsilon\sqrt{N})^{-\delta} \int Q_i^{2+\delta}(n) \, dP$$

$$\leq (\varepsilon\sqrt{N})^{-\delta} \int_0^1 EY_i^{2+\delta}(q_i(n+s)) \, ds$$

$$\leq C(\varepsilon\sqrt{N})^{-\delta}.$$

Thus, the left-hand side of (6.2) 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 $\tilde{\tilde{\xi}}_{i,N,r}$ of $\tilde{\xi}_{i,N}$ by

$$(6.3) \qquad \tilde{\tilde{\xi}}_{i,N,r}(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{[S_i(Nt)]} I_{i,r}(n),$$

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, similar to Section 5, we obtain that

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

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 $E(\zeta_{i,r,N}(S_i(Nt)))^2$. As in the discrete time case considered in Section 4, in view of (2.17) and (6.1), it suffices to study the asymptotical behavior of

(6.5)
$$D_{i,j}(N,s,t) = E\big[\xi_{i,N}(s)\xi_{j,N}(t)\big]$$
$$= \frac{1}{N} \int_0^{S_j(Nt)} \int_0^{S_i(Ns)} E\big[Y_i(q_i(u))Y_j(q_j(v))\big]\,du\,dv.$$

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,

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

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

(6.7)
$$\frac{1}{N} \int_0^{Nt/j} \int_0^{Ns/i} E\big[Y_i(iu)Y_j(jv)\big]\,du\,dv$$
$$= \frac{\upsilon}{Ni} \int_0^{Nt/j} \int_{-jv/\upsilon}^{(Ns-jv)/\upsilon} E\bigg[Y_i\bigg(\frac{jv + w\upsilon}{i}\bigg)Y_j(jv)\bigg]\,dw\,dv.$$

When υ is large, then the expectation under the integral equals approximately $a_{i,j}(w, 2w, \ldots, \upsilon w)$ 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 $0 \le v \le Nt/j$ and $u = (jv + wv)/i \le Ns/i$, that is, asymptotically for N large $0 \le v \le \frac{N}{j}\min(s, t)$. It follows that the expression in (6.7) is approximately equal as $N \to \infty$ to

$$(6.8) \qquad \frac{v}{ij}\min(s, t)\int_{-\infty}^{\infty} a_{i,j}(w, 2w, \ldots, vw)\, dw$$

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

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

$$(6.9) \qquad \lim_{N \to \infty} D_{i,i}(N, t, t) = 0.$$

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

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

(6.10)

for some $C > 0$ independent of t, N and γ, where we obtain by (2.30) and estimates similar to Lemma 4.2 and Proposition 4.5 that for any $i > k$ and $\gamma > 0$,

$$(6.11) \quad \beta_\gamma^{(i)}(M) = \sup_{u \ge M}\int_{u+\gamma}^{\infty}\big|b_{i,i}(u, v)\big|\, dv < \infty \quad \text{and} \quad \lim_{M \to \infty}\beta_\gamma^{(i)}(M) = 0.$$

So, letting first $N \to \infty$ and then $\gamma \to 0$, we obtain (6.9).

REMARK 6.1. 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 $q_i(t)$, $i = k + 1, \ldots, \ell$ as before. In this situation (6.6) becomes

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

where $\rho_1 < \rho_2 < \cdots < \rho_{n_{ij}} < 1$ and $\alpha_i \rho_l, \alpha_j \rho_l \in \{\alpha_1, \ldots, \alpha_k\}$ for $l = 1, \ldots, n_{ij}$. Then the covariances (6.8) will have the form

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

REFERENCES

[1] BERGELSON, V. (1987). Weakly mixing PET. *Ergodic Theory Dynam. Systems* **7** 337–349. MR0912373

[2] BERGELSON, V., LEIBMAN, A. and MOREIRA, C. G. (2012). From discrete- to continuous-time ergodic theorems. *Ergodic Theory Dynam. Systems* **32** 383–426. MR2901353

[3] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York. MR1700749

[4] BOWEN, R. (1975). *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Math.* **470**. Springer, Berlin. MR0442989

[5] BRADLEY, R. C. (2007). *Introduction to Strong Mixing Conditions*. Kendrick Press, Heber City, UT.

[6] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators. Part I*. Wiley, New York.

[7] FURSTENBERG, H. (1990). Nonconventional ergodic averages. In *The Legacy of John Von Neumann* (*Hempstead*, *NY*, 1988). *Proceedings of Symposia in Pure Mathematics* **50** 43–56. Amer. Math. Soc., Providence, RI. MR1067751

[8] GARLING, D. J. H. (2007). *Inequalities*: *A Journey Into Linear Analysis*. Cambridge Univ. Press, Cambridge. MR2343341

[9] HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*: *Probability and Mathematical Statistics*. Academic Press, New York. MR0624435

[10] HEINRICH, L. (1996). Mixing properties and central limit theorem for a class of non-identical piecewise monotonic C^2-transformations. *Math. Nachr.* **181** 185–214. MR1409076

[11] IBRAGIMOV, I. A. and LINNIK, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen. MR0322926

[12] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877

[13] KIFER, Y. (2010). Nonconventional limit theorems. *Probab. Theory Related Fields* **148** 71–106. MR2653222

[14] KUNITA, H. (1990). *Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Mathematics* **24**. Cambridge Univ. Press, Cambridge. MR1070361

[15] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **233**. Springer, Berlin. MR0532498

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
JERUSALEM 91904
ISRAEL
E-MAIL: kifer@math.huji.ac.il

COURANT INSTITUTE
  FOR MATHEMATICAL STUDIES
NEW YORK UNIVERSITY
251 MERCER ST
NEW YORK, NEW YORK 10012
USA
E-MAIL: varadhan@cims.nyu.edu