The Annals of Probability
1991, Vol. 19, No. 2, 587–604

ON THE CENTRAL LIMIT THEOREM FOR MARKOV CHAINS IN RANDOM ENVIRONMENTS

BY ROBERT COGBURN

University of New Mexico

A functional central limit theorem is established for Markov chains in random environments under the assumption of existence of a finite invariant, ergodic measure and a mixing condition. These conditions are always satisfied when the state space is finite.

1. Introduction and basic results. Let $\{P(\theta), \theta \in \Theta\}$ be a family of stochastic matrices acting on a finite or denumerable space $\mathscr{X}$. We call Θ the set of environments and the (x, y) entry of $P(\theta)$, $P(\theta; x, y)$, denotes the probability of moving from x to y in one step in environment θ. Let $\mathscr{B}$ be a σ-field in Θ such that $P(\cdot; x, y)$ is $\mathscr{B}$ measurable for each $x, y \in \mathscr{X}$. Let $\vec{\Theta} = \Theta^{\mathbf{Z}}$ be the product space of doubly infinite sequences $\{\theta_n\}$ and $\vec{\mathscr{B}} = \mathscr{B}^{\mathbf{Z}}$ be its product σ-field, T be the sequence shift operator on $\vec{\Theta}$ and π be a shift invariant probability on $(\vec{\Theta}, \vec{\mathscr{B}})$. Now let $X_0, X_1, \dots$ be a sequence in $\mathscr{X}$ such that

$$(1.1) \qquad P\Big(X_{n+1} = y \,|\, X_n = x, X_{n-1}, \dots, X_0; \vec{\theta}\,\Big) = P(\theta_n; x, y) \quad \text{a.s.}$$

for all $x, y \in \mathscr{X}$ and $n \geq 0$. This two-level stochastic sequence is called a *Markov chain in a random environment* (MCRE).

Let $\mathbf{S} = \mathscr{X} \times \vec{\Theta}$, σ-field $\mathscr{F} = \mathscr{A} \times \vec{\mathscr{B}}$, where $\mathscr{A} = 2^{\mathscr{X}}$, and measure $\mu = \kappa \times \pi$, where κ is counting measure on $\mathscr{X}$. Define a transition probability P on S by

$$(1.2) \qquad P\Big((x, \vec{\theta}\,), \big(y, T\vec{\theta}\,\big)\Big) = P(\theta_0; x, y).$$

This formulation allows the L_1 approach to Markov processes of Hopf (see Foguel [7] for a general introduction) to be applied. We let $\mathscr{F}_i$ denote the *σ-field of invariant sets* in S: $F \in \mathscr{F}_i$ if $P(I_F) = I_F$ a.e. Of course, F is *closed* if $P(I_F) \geq I_F$ a.e. For $F \subset S$, $(F)_x$ and $(F)^{\vec{\theta}}$ denote the sections of F at x and $\vec{\theta}$, respectively.

Given the sequence $\vec{\theta}$, the $\{X_n\}$ sequence evolves as a nonhomogeneous Markov chain and we will call these sequences the *$\vec{\theta}$-chains*. The transition probability from time m to $n > m$ for the $\vec{\theta}$-chain is $P(\theta_m) \cdots P(\theta_{n-1})$ and we will write $P(\theta_m \cdots \theta_{n-1})$ for this product. The most interesting results in the theory of MCREs are those for the $\vec{\theta}$-chains; these are results for nonhomogeneous Markov chains in which the one-step transition probabilities are

Received April 1989; revised April 1990.
AMS 1980 *subject classifications*. 60F05, 60J10, 60J99.
Key words and phrases. Markov chains in random environments, central limit theorem, mixing.

selected by a stochastic process. The assumption that this process is stationary is the stochastic analog of the time homogeneous assumption in the classical theory.

Let $P_{x,\vec{\theta}}$ ($P_{\nu,\vec{\theta}}$) denote the distribution of the $\vec{\theta}$-chain $\{X_n\}_{n=0}^{\infty}$ when $X_0 = x$ (X_0 has distribution ν) and $\vec{\theta}$ is the environmental sequence. Also P_{φ} denotes the distribution on the Markovian sequence $\{(X_n, T^n\vec{\theta})\}$ in $\mathbf{S}$ when $(X_0, \vec{\theta})$ has initial density φ. Expectations on these spaces are denoted by E with the corresponding subscript.

We assume the existence of an invariant probability density φ on S:

$$(1.3) \qquad \sum_{y \in \mathscr{X}} \varphi(y, T^{-n}\vec{\theta}) P(\theta_{-n} \cdots \theta_{-1}; y, x) = \varphi(x, \vec{\theta}).$$

An invariant distribution always exists when $\mathscr{X}$ is finite (see [3]). The distribution P_{φ} is ergodic if and only if its support $F_{\varphi} = \{(x, \vec{\theta}): \varphi(x, \vec{\theta}) > 0\}$ is an atom of $\mathscr{F}_i$ and in any case $\mathscr{F}_i$ is atomic provided π is ergodic (see [3]). To avoid unnecessary complications *we also assume π and P_{φ} are ergodic.*

Let $\|\cdot\|$ be total variation norm and given two distributions λ, ν on $\mathscr{X}$, let

$$(1.4) \qquad \delta_n(\lambda, \nu, \vec{\theta}) = \|(\lambda - \nu) P(\theta_0 \cdots \theta_{n-1})\|.$$

When $\lambda(x) = 1$, write $\delta_n(x, \nu, \vec{\theta})$, and if $\nu(y) = 1$ as well, then write $\delta(x, y, \vec{\theta})$. The δ_n are nonincreasing and we let $\delta = \lim_{n \to \infty} \delta_n$.

Let M denote the maximal support of a finite invariant measure on $\mathscr{X}$. Note that M is a well-defined, closed set in $\mathscr{X}$ (see [3]). We say that $(x, \vec{\theta})$ *meets* $(y, \vec{\theta})$, denoted $(x, \vec{\theta}) \leftrightarrow (y, \vec{\theta})$, if $\delta(x, y, \vec{\theta}) < 2$. It is established in [4] that $\leftrightarrow$ is an equivalence relation on M and we let $[(x, \vec{\theta})]$ denote the equivalence class of all $(y, \vec{\theta})$ that meet $(x, \vec{\theta})$. The number of equivalence classes in $(F_{\varphi})^{\vec{\theta}}$ is a finite constant c (not depending on $\vec{\theta}$) and if $(x, \vec{\theta}) \leftrightarrow (y, \vec{\theta})$, then $\delta(x, y, \vec{\theta}) = 0$, for π-a.e. $\vec{\theta}$, as is shown in [4]. *We exclude the null set of $\vec{\theta}$'s such that the above properties fail for any $\vec{\theta}, T\vec{\theta}, \dots$ from further consideration.* It is also established in [4] that for the process started at $(x, \vec{\theta})$, with probability 1 the process will be in an equivalence class $D_n = [(X_n, T^n\vec{\theta})]$ in n steps. Let $D_0 = [(x, \vec{\theta})]$ and for $n \geq 0$ set

$$(1.5) \qquad \varphi_n(y) = c\varphi(y, T^n\vec{\theta}) I_{D_n}(y).$$

Invariance of φ implies $\varphi_0 P(\theta_0 \cdots \theta_{n-1}) = \varphi_n$ for each n. Moreover it is shown in [4] that $\delta(x, \varphi_0, \vec{\theta}) = 0$. We let $\overline{P}_{x,\vec{\theta}}$ be the distribution of $\{X_n\}_{n=0}^{\infty}$ when X_0 has distribution $\varphi_0(y) = c\varphi(y, \vec{\theta}) I_{[(x,\vec{\theta})]}(y)$ and let $\overline{E}_{x,\vec{\theta}}$ be the corresponding expectation.

Now let $\mathscr{A}_n^{\infty}$ be the σ-field generated by $X_n, X_{n+1}, \dots$ and note that

$$\sup_{A \in \mathscr{A}_n^{\infty}} |P_{x,\vec{\theta}}(A) - P_{y,\vec{\theta}}(A)| = \delta_n(x, y, \vec{\theta})/2.$$

Let

$$(1.6) \qquad \rho_n(\vec{\theta}) = \sup_{x, y \in (F_{\varphi})^{\vec{\theta}}: (x, \vec{\theta}) \leftrightarrow (y, \vec{\theta})} \sup_{A \in \mathscr{A}_n^{\infty}} |P_{x,\vec{\theta}}(A) - P_{y,\vec{\theta}}(A)|.$$

Consider the hypothesis

(U) $$\rho_n(\vec{\theta}) \to 0 \quad \text{as } n \to \infty \text{ for } \pi - \text{a.e. } \vec{\theta}.$$

Note the convergence of $\delta_n(x, y, \vec{\theta})$ to 0 does not imply the uniform convergence required by (U) in general, but it does imply (U) when $\mathscr{X}$ is finite.

By a standard argument (e.g., see [12]),

$$\rho_{m+n}(\vec{\theta}) \leq \rho_m(\vec{\theta})\rho_n(T^m\vec{\theta}). \tag{1.7}$$

It follows that, if $\lim_{n\to\infty}\rho_n(\vec{\theta}) = 0$ for a π-positive set of $\vec{\theta}$'s, then the convergence holds for π-a.e. $\vec{\theta}$ when π is ergodic.

Under condition (U), we can pick a set Γ with $\pi\Gamma > 0$ and an n_0 such that $\sup_{\vec{\theta}\in\Gamma}\rho_{n_0}(\vec{\theta}) = \hat{\rho} < 1$. Let $t_0 = 0$ and $t_n = t_n(\vec{\theta})$ be the first time t after t_{n-1} that $T^t\vec{\theta} \in \Gamma$ for $n \geq 1$. Then it is easy to see that (1.7) implies

$$\rho_{t_n}(\vec{\theta}) \leq \hat{\rho}^{n/n_0 - 1} = \alpha\rho^n, \tag{1.8}$$

where $\rho = \hat{\rho}^{1/n_0}$ and $\alpha = \hat{\rho}^{-1}$. Since the ergodic theorem implies $t_n/n \to 1/\pi(\Gamma)$ for a.e. $\vec{\theta}$, we have that for any $\varepsilon > 0$ there exists a finite $n(\vec{\theta})$ such that $t_{[n\pi(\Gamma)(1-\varepsilon)]+1} \leq n$ for $n \geq n(\vec{\theta})$. But then

$$\rho_n(\vec{\theta}) \leq \alpha\rho^{n\pi(\Gamma)(1-\varepsilon)} \quad \text{for } n \geq n(\vec{\theta}). \tag{1.9}$$

Note that, while the $t_n(\vec{\theta})$ are random times in general, when the distribution on sequence space is $P_{x,\vec{\theta}}$ or $\overline{P}_{x,\vec{\theta}}$, $\vec{\theta}$ is fixed so the t_n's are constants, though they are irregularly spaced.

Define a distribution for the MCRE started on $\mathscr{X}\times\Gamma$ by $\Psi(F) = \Phi(F \cap (\mathscr{X}\times\Gamma))/\pi(\Gamma)$, $F \in \mathscr{F}$, and let P_ψ and E_ψ be the corresponding probability and expectation on sequences in $\mathbf{S}$. The blocks $(X_{t_k}, \dots, X_{t_{k+1}-1})$ are Markovian and under the distribution P_ψ they are stationary and ergodic since P_ϕ is ergodic (e.g., see Petersen [18] on recurrence and ergodicity) and this property will be key to ensuing discussion.

The principle results are stated next; proofs are given in Section 2.

Let f be a real valued measurable function of $(\mathbf{S}, \mathscr{F})$ and let

$$\bar{f}(x,\vec{\theta}) = f(x,\vec{\theta}) - \overline{E}_{x,\vec{\theta}}\big(f(x,\vec{\theta})\big) \quad \text{and} \quad Y_n = \sum_{k=t_n}^{t_{n+1}-1} \bar{f}\big(X_k, T^k\vec{\theta}\big).$$

PROPOSITION 1. *Suppose $E_\psi(Y_0^2) < \infty$. Let*

$$\sigma^2 = E_\psi(Y_0^2) + 2\sum_{n=1}^{\infty} E_\psi(Y_0 Y_n). \tag{1.10}$$

Then the series defining σ^2 converges absolutely and

$$\lim_{n\to\infty} E_\psi\left[\left(\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} Y_k\right)^2\right] = \sigma^2.$$

In the following results, interpret $\mathcal{N}(0,0)$ as the law degenerate at 0.

THEOREM 1. *Let* (U) *hold and let* $E_\psi(Y_0^2) < \infty$. *Then for* Φ-*a.e.* $(x, \vec{\theta})$ *and the distribution* $P_{x,\vec{\theta}}$ *on sequence space,*

$$(1.11) \qquad \frac{1}{\sqrt{n}} \sum_{k=0}^{t_n - 1} \left(f\left(X_k, T^k\vec{\theta}\right) - E_{x,\vec{\theta}} f\left(X_k, T^k\vec{\theta}\right)\right) \to_{\mathscr{D}} \mathcal{N}(0, \sigma^2)$$

as $n \to \infty$, *where* σ^2 *is given by* (1.10) *and* $0 \le \sigma^2 < \infty$.

We will need to refer to the following moment conditions:

$$\text{(M1)} \qquad E_\phi\left[\left(\sum_{k=0}^{t_1 - 1} \bar{f}\left(X_k, T^k\vec{\theta}\right)\right)^2\right] < \infty,$$

$$\text{(M2)} \qquad E_\psi\left[\left(\sum_{k=0}^{t_1 - 1} \left|\bar{f}\left(X_k, T^k\vec{\theta}\right)\right|\right)^2\right] < \infty,$$

$$\text{(M3)} \qquad E_\psi\left[\max_{0 < l \le t_1}\left(\sum_{k=0}^{l - 1} \bar{f}\left(X_k, T^k\vec{\theta}\right)\right)^2\right] < \infty.$$

THEOREM 2. *Let* (U) *hold and any of* (M1), (M2) *or* (M3) *hold. Then for* Φ-*a.e.* $(x, \vec{\theta})$ *and the distribution* $P_{x,\vec{\theta}}$ *on sequence space,*

$$(1.12) \qquad \frac{1}{\sqrt{n}} \sum_{k=0}^{n - 1} \left(f\left(X_k, T^k\vec{\theta}\right) - E_{x,\vec{\theta}} f\left(X_k, T^k\vec{\theta}\right)\right) \to_{\mathscr{D}} \mathcal{N}(0, \sigma^2\pi(\Gamma))$$

as $n \to \infty$, *where* σ^2 *is given by* (1.10) *and* $0 \le \sigma^2 < \infty$.

The variance $\sigma^2\pi(\Gamma)$ evidently is a property of the function f and must be the same for any Γ such that one of the moment conditions is satisfied. Therefore we introduce the notation $\sigma^2(f)$ for this limit variance. The following results provide more direct information on the value of $\sigma^2(f)$.

THEOREM 3. *Let* (U) *hold. If* $E_\psi(Y_0^2) < \infty$, *then for* Ψ-*a.e.* $(x, \vec{\theta})$

$$(1.13) \qquad E_{x,\vec{\theta}}\left[\frac{1}{n}\left(\sum_{k=0}^{t_n - 1} \left(f\left(X_k, T^k\vec{\theta}\right) - E_{x,\vec{\theta}} f\left(X_k, T^k\vec{\theta}\right)\right)\right)^2\right] \to \sigma^2(f)$$

as $n \to \infty$. *If* (M1) *holds, then for* Φ-*a.e.* $(x, \vec{\theta})$,

$$(1.14) \qquad E_{x,\vec{\theta}}\left[\frac{1}{n}\left(\sum_{k=0}^{n - 1} \left(f\left(X_k, T^k\vec{\theta}\right) - E_{x,\vec{\theta}} f\left(X_k, T^k\vec{\theta}\right)\right)\right)^2\right] \to \sigma^2(f)$$

as $n \to \infty$ *and if* (M2) *or* (M3) *holds, then* (1.14) *holds for* Ψ-*a.e.* $(x, \vec{\theta})$.

THEOREM 4. *Let* (U) *hold*, $\int t_1\, d\pi < \infty$ *and* f *be bounded. Then condition* (M2) *is satisfied*,

$$(1.15)\quad \sigma^2(f) = E_\phi\Big(\bar{f}\big(X_0,\vec{\theta}\,\big)^2\Big) + 2\sum_{n=1}^{\infty} E_\phi\Big[\bar{f}\big(X_0,\vec{\theta}\,\big)\bar{f}\big(X_n,T^n\vec{\theta}\,\big)\Big]$$

and

$$(1.16)\quad E_{x,\vec{\theta}}\left[\frac{1}{n}\left(\sum_{k=0}^{n-1}\Big(f\big(X_k,T^k\vec{\theta}\,\big) - E_{x,\vec{\theta}}f\big(X_k,T^k\vec{\theta}\,\big)\Big)\right)^2\right] \to \sigma^2(f)$$

as $n \to \infty$ *for* Φ-*a.e.* $(x,\vec{\theta}\,)$. *Moreover*, (1.16) *remains valid when* $E_{x,\vec{\theta}}$ *is replaced by* E_ϕ *or* E_ψ.

Let $\bar{f}^{(b)} = \bar{f}I_{[|\bar{f}|\le b]}$.

COROLLARY 1. *Let* (U) *hold*, $\int t_1\, d\pi < \infty$ *and* f *be unbounded but satisfy* (M2) *and* $E_\phi(f(X_0,\vec{\theta}\,)^2) < \infty$. *Then*

$$(1.17)\quad \sigma^2(f) = E_\phi\Big(\bar{f}\big(X_0,\vec{\theta}\,\big)^2\Big) + 2\lim_{b\to\infty}\sum_{n=1}^{\infty} E_\phi\Big(\bar{f}^{(b)}\big(X_0,\vec{\theta}\,\big)\bar{f}^{(b)}\big(X_n,T^n\vec{\theta}\,\big)\Big).$$

The study of the central limit theorem for nonhomogeneous Markov chains began with A. A. Markov in 1910. A review and refinement of the work on this problem over the ensuing half century is provided in the classic papers of Dobrushin [5]. Quite general results are obtained in that study, but at the price of very strong mixing conditions.

There has been much further work on the central limit theorem for Markov chains. Of those treatments that use a uniform strong mixing condition (ϕ-mixing), see Ibragimov and Linnik [10] and Nagaev [13], [14]. Studies mostly using weaker (α-type) mixing include [8], [9], [11], [15], [16], [17], [19], [20], [21].

Our condition (U) is essentially a ϕ-mixing condition and it is to be hoped that further work will find a way to weaken this hypothesis. Previous studies suggest trying to break the chain into blocks between returns to a state x or set G, but this approach has the problem that the resulting blocks are not independent and it appears difficult to get a mixing coefficient. What makes the approach used here work is that the blocks determined by returns to Γ have nonrandom duration in time for the $P_{x,\vec{\theta}}$ distributions.

2. Proofs.

We will need the standard inequalities.

LEMMA 1. *Let* U, V *be real or complex valued random variables with* U *a function of* $(X_0,\dots,X_m;\vec{\theta}\,)$ *and* V *a function of* $(X_{m+n},X_{m+n+1},\dots;\vec{\theta}\,)$. *Then* *for* $r, s > 1$ *with* $1/r + 1/s = 1$,

$$|E_{x,\vec{\theta}}(UV) - E_{x,\vec{\theta}}UE_{x,\vec{\theta}}V| \le 2\rho_n(T^m\vec{\theta})^{1/r} E_{x,\vec{\theta}}^{1/r}(|U|^r) E_{x,\vec{\theta}}^{1/s}(|V|^s) \tag{2.1}$$

and

$$E_{x,\vec{\theta}}|E_{x,\vec{\theta}}(V|X_m) - E_{x,\vec{\theta}}V| \le 2\rho_n(T^m\vec{\theta})^{1/r} E_{x,\vec{\theta}}^{1/s}(|V - E_{x,\vec{\theta}}V|^s). \tag{2.2}$$

Moreover, these inequalities remain valid when $E_{x,\vec{\theta}}$ *is replaced by* $\overline{E}_{x,\vec{\theta}}$ *throughout.*

The first inequality is essentially Lemma 7.1, Chapter V of Doob [6] and the second follows from the first upon replacing V by $V - E_{x,\vec{\theta}}V$ and U by $\operatorname{sign}(E_{x,\vec{\theta}}(V|X_m) - E_{x,\vec{\theta}}V)$.

LEMMA 2. *Let f be a real valued measurable function on* $(\mathbf{S}, \mathscr{F})$ *such that* $\int f(x,\vec{\theta})^2\, d\Phi(x,\vec{\theta}) < \infty$. *Then for* Φ-*a.e.* $(x,\vec{\theta})$,

$$\sum_{n=0}^{\infty} \left|E_{x,\vec{\theta}} f(X_n, T^n\vec{\theta}) - \overline{E}_{x,\vec{\theta}} f(X_n, T^n\vec{\theta})\right| < \infty. \tag{2.3}$$

PROOF. With φ_n defined by (1.5), since $\varphi_n = \varphi_0 P(\theta_0 \cdots \theta_{n-1})$, we have

$$\overline{E}_{x,\vec{\theta}} \bar{f}(X_n, T^n\vec{\theta}) = \sum_{y\in\mathscr{X}} \varphi_n(y) \bar{f}(y, T^n\vec{\theta}) = 0$$

and $E_{x,\vec{\theta}}\bar{f}(X_n, T^n\vec{\theta}) = \overline{E}_{x,\vec{\theta}}(\bar{f}(X_n, T^n\vec{\theta})|X_0 = x)$ so by (2.2) with $\overline{E}_{x,\vec{\theta}}$ and $V = \bar{f}(X_n, T^n\vec{\theta})$,

$$\overline{E}_{x,\vec{\theta}}\left|E_{x,\vec{\theta}}\bar{f}(X_n, T^n\vec{\theta})\right| \le 2\rho_n(\vec{\theta})^{1/2}\overline{E}_{x,\vec{\theta}}^{1/2}\left(\bar{f}(X_n, T^n\vec{\theta})^2\right).$$

Hence

$$\begin{aligned}
&\overline{E}_{x,\vec{\theta}} \sum_{n=0}^{\infty} \left|E_{x,\vec{\theta}} f(X_n, T^n\vec{\theta}) - \overline{E}_{x,\vec{\theta}} f(X_n, T^n\vec{\theta})\right| \\
&\quad = \sum_{n=0}^{\infty} \overline{E}_{x,\vec{\theta}}\left|E_{x,\vec{\theta}}\bar{f}(X_n, T^n\vec{\theta})\right| \\
&\quad \le 2\sum_{n=0}^{\infty} \rho_n(\vec{\theta})^{1/2}\overline{E}_{x,\vec{\theta}}^{1/2}\left(\bar{f}(X_n, T^n\vec{\theta})^2\right) \\
&\quad \le 2\left(\sum_{n=0}^{\infty}(n+1)^2\rho_n(\vec{\theta})\right)^{1/2}\left(\sum_{n=0}^{\infty}\overline{E}_{x,\vec{\theta}}\left(\bar{f}(X_n, T^n\vec{\theta})^2\right)/(n+1)^2\right)^{1/2}.
\end{aligned} \tag{2.4}$$

The inequality in (1.10) implies the first series is finite, while

$$\begin{aligned}
&\int\left(\sum_{n=0}^{\infty}\overline{E}_{x,\vec{\theta}}\left(\bar{f}(X_n, T^n\vec{\theta})^2\right)/(n+1)^2\right)d\Phi(x,\vec{\theta}) \\
&\qquad = \int\bar{f}(x,\vec{\theta})^2\, d\Phi\sum_{n=0}^{\infty}\frac{1}{(n+1)^2} < \infty.
\end{aligned}$$

Thus the series in (2.4) is finite for Φ-a.e. $(x,\vec{\theta})$. The assertion follows since $\varphi_0(x) = c\varphi(x,\vec{\theta}) > 0$ for any given $(x,\vec{\theta}) \in F_\varphi$. □

Of course, $P_\psi(A) = \int P_{x,\vec\theta}(A)\, d\Psi(x,\vec\theta)$ and using the result that

$$\sum_{y:\,(y,\vec\theta)\leftrightarrow(x,\vec\theta)} \varphi(y,\vec\theta) = 1/c$$

for Φ-a.e. $(x,\vec\theta)$ (see formula (2.5) of [4]), it is also true that

$$\begin{aligned}
&\int \overline{P}_{x,\vec\theta}(A)\, d\Psi(x,\vec\theta)\\
&= \frac{1}{\pi(\Gamma)}\int_\Gamma \sum_{x\in\mathscr{X}} c \sum_{y:\,(y,\vec\theta)\leftrightarrow(x,\vec\theta)} \varphi(y,\vec\theta) P_{y,\vec\theta}(A)\varphi(x,\vec\theta)\, d\pi(\vec\theta)\\
&= \frac{1}{\pi(\Gamma)}\int_\Gamma \sum_{y\in\mathscr{X}} \varphi(y,\vec\theta) P_{y,\vec\theta}(A) c \sum_{x:\,(x,\vec\theta)\leftrightarrow(y,\vec\theta)} \varphi(x,\vec\theta)\, d\pi(\vec\theta)\\
&= \frac{1}{\pi(\Gamma)}\int_\Gamma \sum_{y\in\mathscr{X}} \varphi(y,\vec\theta) P_{y,\vec\theta}(A)\, d\pi(\vec\theta) = P_\psi(A).
\end{aligned} \tag{2.5}$$

Proof of Proposition 1. By Lemma 1 and formula (1.8), noting that $\alpha > 1$,

$$|\overline{E}_{x,\vec\theta}(Y_0Y_n)| \le 2\alpha\rho^{n/2}\overline{E}^{1/2}_{x,\vec\theta}(Y_0^2)\overline{E}^{1/2}_{x,\vec\theta}(Y_n^2).$$

Using (2.5) and this inequality, then applying Schwarz's inequality,

$$\begin{aligned}
|E_\psi(Y_0Y_n)| &= \left|\int \overline{E}_{x,\vec\theta}(Y_0Y_n)\, d\Psi(x,\vec\theta)\right|\\
&\le 2\alpha\rho^{n/2}\left(\int \overline{E}_{x,\vec\theta}(Y_0^2)\, d\Psi(x,\vec\theta)\right)^{1/2}\left(\int \overline{E}_{x,\vec\theta}(Y_n^2)\, d\Psi(x,\vec\theta)\right)^{1/2}\\
&= 2\alpha\rho^{n/2}E_\psi(Y_0^2),
\end{aligned} \tag{2.6}$$

since the Y_n's are stationary under P_ψ. The first assertion follows from this estimate and the second follows from the first as in the proof of Lemma 7.3, Chapter V of Doob [6]. □

Let $Z_n = \sum_{k=1}^{n-1} Y_k$, $Z_{m,n} = Z_n - Z_m$ and $\chi_{m,n}(u) = E_{x,\vec\theta}(e^{iuZ_{m,n}})$ be the characteristic function of $Z_{m,n}$.

Lemma 3. *For any positive integers* L, l, m *and real* u,

$$\begin{aligned}
&\left|\chi_{L,L+ml}(u) - \prod_{k=1}^{m}\chi_{L+(k-1)l,L+kl}(u)\right|\\
&\le Bu^2\left(\sum_{k=1}^{m-1} E_{x,\vec\theta}\left(Z^2_{L+(k-1)l,L+kl}\right)\right)^{1/2}\\
&\quad\times\left(\sum_{k=1}^{m-1}\sum_{j=0}^{\infty}\rho^{j/2}E_{x,\vec\theta}\left(Y^2_{L+kl+j}\right)\right)^{1/2},
\end{aligned} \tag{2.7}$$

where $B = 2\alpha(\sum_{j=0}^\infty \rho^{j/2})^{1/2}$.

PROOF. For $1 \le k < m$, let $J = (m-k)l$. Then

$$
\begin{aligned}
&\left|\chi_{L+(k-1)l,\,L+ml}(u) - \chi_{L+(k-1)l,\,L+kl}(u)\chi_{L+kl,\,L+ml}(u)\right| \\
&\quad = \Bigg|\sum_{j=0}^{J-1}\Big(E_{x,\vec\theta}\big[\big(e^{iuZ_{L+(k-1)l,L+kl}}-1\big)\big(e^{iuY_{L+kl+j}}-1\big)e^{iuZ_{L+kl+j+1,L+ml}}\big] \\
&\qquad\qquad - E_{x,\vec\theta}\big[e^{iuZ_{L+(k-1)l,L+kl}}-1\big] \\
&\qquad\qquad \times E_{x,\vec\theta}\big[\big(e^{iuY_{L+kl+j}}-1\big)e^{iuZ_{L+kl+j+1,L+ml}}\big]\Big)\Bigg| \qquad (2.8)\\
&\quad \le \sum_{j=0}^{J-1} 2\alpha\rho^{j/2}u^2 E_{x,\vec\theta}^{1/2}\big(Z^2_{L+(k-1)l,\,L+kl}\big)E_{x,\vec\theta}^{1/2}\big(Y^2_{L+kl+j}\big) \\
&\quad \le 2\alpha u^2 E_{x,\vec\theta}^{1/2}\big(Z^2_{L+(k-1)l,\,L+kl}\big)\sum_{j=0}^{\infty}\rho^{j/2}E_{x,\vec\theta}^{1/2}\big(Y^2_{L+kl+j}\big),
\end{aligned}
$$

where the first inequality uses Lemma 1 and $|e^{iux}-1| \le |ux|$.

Using the convention that $\prod_{j=1}^{0} a_j = 1$ and (2.8),

$$
\begin{aligned}
&\left|\chi_{L,\,L+ml}(u) - \prod_{k=1}^{m}\chi_{L+(k-1)l,\,L+kl}(u)\right| \\
&\quad = \Bigg|\sum_{k=1}^{m-1}\big(\chi_{L+(k-1)l,\,L+ml}(u) - \chi_{L+(k-1)l,\,L+kl}(u)\chi_{L+kl,\,L+ml}(u)\big) \\
&\qquad\qquad \times \prod_{j=1}^{k-1}\chi_{L+(j-1)l,\,L+jl}(u)\Bigg| \\
&\quad \le 2\alpha u^2\left(\sum_{k=1}^{m-1}E_{x,\vec\theta}\big(Z^2_{L+(k-1)l,\,L+kl}\big)\right)^{1/2}\left(\sum_{k=1}^{m-1}\left(\sum_{j=0}^{\infty}\rho^{1/2}E_{x,\vec\theta}^{1/2}\big(Y^2_{L+kl+j}\big)\right)^2\right)^{1/2} \\
&\quad \le 2\alpha u^2\left(\sum_{j=0}^{\infty}\rho^{j/2}\right)^{1/2}\left(\sum_{k=1}^{m-1}E_{x,\vec\theta}\big(Z^2_{L+(k-1)l,\,L+kl}\big)\right)^{1/2} \\
&\qquad \times\left(\sum_{k=1}^{m-1}\sum_{j=0}^{\infty}\rho^{j/2}E_{x,\vec\theta}\big(Y^2_{L+kl+j}\big)\right)^{1/2},
\end{aligned}
$$

applying Schwarz's inequality to the sum over j at the last step. □

We need the following result which holds for any stationary Markov sequence. It is necessary since the dominated convergence theorem for conditional expectations does not apply in general if domination of the convergent sequence by an integrable variable is replaced by uniform integrability.

LEMMA 4. *Let $W_0, W_1, \dots$ be a stationary Markov sequence on a measurable state space $(\mathscr{W}, \mathscr{E})$. Let h be a real valued measurable function of the sequence $(W_0, W_1, \dots)$ and be integrable. Then*

$$(2.9)\qquad \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} E\big(h(W_k, W_{k+1}, \dots) \mid W_0\big) = E\big(h(W_0, W_1, \dots) \mid \mathscr{I}\big)$$

a.s. and in L_1, where $\mathscr{I}$ is the σ-field of shift invariant sets on the $(W_0, W_1, \dots)$ sequence.

Moreover, if, for some $r > 0$, $E(|h(W_0, W_1, \dots)|^r) < \infty$, then

$$(2.10)\qquad \lim_{n\to\infty} \max_{k\le n} \left(\frac{1}{n} \big| E\big(h(W_k, W_{k+1}, \dots) \mid W_0\big)\big|\right)^r = 0 \quad a.s.$$

PROOF. 1. Let $U_k = E(h(W_k, W_{k+1}, \dots) \mid W_0)$, $U_\infty = E(h(W_0, W_1, \dots) \mid \mathscr{I})$. The L_1 convergence holds since by Jensen's inequality,

$$E\left|\frac{1}{n} \sum_{k=0}^{n-1} U_k - U_\infty\right| \le \left\|\frac{1}{n} \sum_{k=0}^{n-1} h(W_k, W_{k+1}, \dots) - U_\infty\right\|_1 \to 0$$

as $n \to \infty$ by the ergodic theorem.

2. Let λ be the distribution of W_0 on $(\mathscr{W}, \mathscr{E})$ and for $g \in L_1 = L_1(\mathscr{W}, \mathscr{E}, \lambda)$, set $Qg(w) = E(g(W_1) \mid W_0 = w)$. Note that $g(W_k) = E(g(W_{k+1} \mid W_k))$ by stationarity and

$$Q^2 g(w) = E\big(E(g(W_2) \mid W_1) \mid W_0\big) = E\big(g(W_2) \mid W_0 = w\big)$$

by the Markov property. Similarly, $Q^n g(w) = E(g(W_n) \mid W_0 = w)$. Now $(\mathscr{W}, \mathscr{E}, \lambda, Q)$ is an L_1 Markov process and λ is invariant for Q. The L_1 ergodic theorem for Markov processes then asserts that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} E\big(g(W_k) \mid W_0\big) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} Q^k g(W_0)$$

exists a.s. and in L_1 (see [7]).

Let $g(w) = E(h(W_0, W_1, \dots) \mid W_0 = w)$. Since h is integrable, $g \in L_1$. By stationarity, $g(w) = E(h(W_k, W_{k+1}, \dots) \mid W_k = w)$ and using the Markov dependence,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} U_k = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} E\big(E(h(W_k, W_{k+1}, \dots) \mid W_k) \mid W_0\big)$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} E\big(g(W_k) \mid W_0\big)$$

exists a.s. and in L_1. The value of the limit is then given by part 1 of the proof.

3. For the second assertion, suppose first that $r = 1$. By (2.9), we have

$$\frac{1}{n} U_n = \frac{n+1}{n} \frac{1}{n+1} \sum_{k=0}^{n} U_k - \frac{1}{n} \sum_{k=0}^{n-1} U_k \to 0$$

a.s. and in L_1 as $n \to \infty$. Then for any fixed m,

$$\frac{1}{n}\max_{k\le n}|U_k| \le \frac{1}{n}\max_{k\le m}|U_k| + \max_{m<k\le n}\frac{1}{k}|U_k| \to 0$$

as $n \to \infty$, then $m \to \infty$.

For any $r > 0$, use the fact that $E^{1/r}(|V|^r)$ is an increasing function of r for any random variable V to get

$$\frac{1}{n}\Big(\max_{k\le n}|U_k|\Big)^r = \frac{1}{n}\max_{k\le n}|U_k|^r \le \frac{1}{n}\max_{k\le n} E(|h(W_k, W_{k+1},\dots)|^r \mid W_0),$$

then apply the above argument with h replaced by $|h|^r$. □

Proof of Theorem 1. 1. For each fixed l, the sequence $W_k^{(l)} = (X_{t_{kl}}, T^{t_{kl}}\vec{\theta})$, $k = 0, 1, \dots$ is stationary Markov under P_ψ. Moreover $\{W_k^{(1)}\}_{k=0}^\infty$ is ergodic under P_ψ since P_ϕ is ergodic, but $\{W_k^{(l)}\}$ may have sets of period l, so $\{W_k^{(l)}\}_{k=0}^\infty$ is not ergodic when $l > 1$. Let F_ψ be the support of Ψ. If there is a nontrivial period l set, then there is a partition $F_0, \dots, F_{l-1}$ of F_ψ such that $P(W_1^{(1)} \in F_{k+1 (\mathrm{mod}\, l)} \mid W_0^{(1)} \in F_k) = 1$. Then the invariant σ-field for $\{W_k^{(l)}\}$ is $\mathscr{I}_l$, the σ-field generated by the events $[W_0 \in F_0], \dots, [W_0 \in F_{l-1}]$ in F_ψ.

Let

$$h_l = h_l(W_0^{(l)}) = E_\psi(Z_l^2 \mid W_0^{(l)}).$$

Note that $W_0^{(l)} = (X_0, \vec{\theta})$ so $h_l(x, \vec{\theta}) = E_{x,\vec{\theta}}(Z_l^2)$ and

$$E_\psi(h_l(W_0^{(l)})) = E_\psi(Z_l^2) \le l^2 E_\psi(Y_0^2) < \infty,$$

by hypothesis. Applying Lemma 4,

$$\begin{aligned} &\frac{1}{n}\sum_{k=0}^{n-1} E_{x,\vec{\theta}}\big(Z_{kl,(k+1)l}^2\big) \\ (2.11)\qquad &= \frac{1}{n}\sum_{k=0}^{n-1} E_\psi\Big(E_\psi\big(Z_{kl,(k+1)l}^2 \mid W_k^{(l)}\big) \mid W_0^{(l)} = (x, \vec{\theta})\Big) \\ &= \frac{1}{n}\sum_{k=1}^{n} E_\psi\Big(h_l\big(W_k^{(l)}\big) \mid W_0^{(l)} = (x,\vec{\theta})\Big) \to E_\psi\big(Z_l^2 \mid \mathscr{I}_l\big) \end{aligned}$$

as $n \to \infty$ a.s.-P_ψ.

Now set $\Psi_k(F) = l\Psi(F \cap F_k)$, $k = 0, \dots, l-1$ and let P_{ψ_k}, E_{ψ_k} be the corresponding probability and expectation. Then on atom F_k of $\mathscr{I}_l$, $E_\psi(h(W_0) \mid \mathscr{I}_l) = E_{\psi_k}(h(W_0))$. Now if W_0 has distribution Ψ_k, then $W_j^{(l)}$ has distribution $\Psi_{k+j (\mathrm{mod}\, l)}$. It follows that the distribution of $(Y_0, \dots, Y_{l-1})$ has the same distribution under P_{ψ_k} as $(Y_{l-1}, Y_0, \dots, Y_{l-2})$ has under $P_{\psi_{k+1(\mathrm{mod}\, l)}}$, hence $E_{\psi_k}(Z_l^2) = E_{\psi_k}((\sum_{j=0}^{l-1} Y_j)^2)$ is the same for all k and the limit in (2.11) is $E_\psi(Z_l^2)$ for Ψ-a.e. $(x, \vec{\theta})$.

Similarly, letting $H = \sum_{j=0}^\infty \rho^{j/2} Y_j^2$ and $h(x, \vec{\theta}) = E_{x,\vec{\theta}}(H)$, stationarity of the Y_j and the monotone convergence theorem yield

$$E_\psi(h(W_0^{(l)})) = E_\psi(H) = E_\psi(Y_0^2)/(1 - \rho^{1/2}) < \infty.$$

By Lemma 4 and the monotone convergence theorem,

$$\frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{\infty}\rho^{1/2}E_{x,\vec{\theta}}\left(Y_{kl+j}^2\right) \to E_{\psi}(H \mid \mathscr{I}_l)$$

as $n \to \infty$ a.s.-P_ψ. In this case we cannot argue that $E_\psi(H \mid \mathscr{I}_l)$ is constant, but since $E_\psi(H) = (1/l)\sum_{L=0}^{l-1} E_{\psi_L}(H)$, there exists an $L = L(l)$ with $0 \le L < l$ such that $E_{\psi_L}(H) \le E_\psi(H)$.

2. Since $E_\psi(Y_0^2) < \infty$, we have $\max_{k \le n}|Y_k|/\sqrt{n} \to 0$ a.s.-P_ψ. Now let $m = [n/l]$ and $0 \le L < l$. Then $(Z_n - Z_{L,L+ml})/\sqrt{n} \to 0$ a.s.-P_ψ as $n \to \infty$ since the difference involves at most $2(l-1)$ terms of Y_k's. Hence

$$\chi_{0,n}\left(\frac{u}{\sqrt{n}}\right) - \chi_{L,L+ml}\left(\frac{u}{\sqrt{n}}\right) \to 0 \tag{2.12}$$

as $n \to \infty$ for every real u.

3. Now apply Lemma 3:

$$\begin{aligned}
&\left|\chi_{L,L+ml}\left(\frac{u}{\sqrt{n}}\right) - \prod_{k=1}^{m}\chi_{L+(k-1)l,L+kl}\left(\frac{u}{\sqrt{n}}\right)\right| \\
&\quad \le B\frac{u^2}{n}m\left(\frac{1}{m}\sum_{k=1}^{m-1}E_{x,\vec{\theta}}\left(Z_{L+(k-1)l,L+kl}^2\right)\right)^{1/2} \\
&\quad\quad \times\left(\frac{1}{m}\sum_{k=1}^{m-1}\sum_{j=0}^{\infty}\rho^{j/2}E_{x,\vec{\theta}}\left(Y_{L+kl+j}^2\right)\right)^{1/2} \\
&\quad \le B\frac{u^2}{l}\left(\frac{1}{m}\sum_{k=1}^{m-1}E_{x,\vec{\theta}}\left(Z_{L+(k-1)l,L+kl}^2\right)\right)^{1/2}\left(\frac{1}{m}\sum_{k=1}^{m-1}\sum_{j=0}^{\infty}E_{x,\vec{\theta}}\left(Y_{L+kl+j}^2\right)\right)^{1/2}.
\end{aligned}$$

Choose $0 \le L < l$, based on the initial $(x, \vec{\theta})$, to minimize the limit of the last square root expression in the above bound. Letting $n \to \infty$, by part 1 of the proof the lim sup of this bound is at most

$$B\frac{u^2}{l}\left(E_\psi(Z_l^2)\right)^{1/2}\left(E_\psi H\right)^{1/2} \tag{2.13}$$

for Ψ-a.e. $(x, \vec{\theta})$. Letting $l \to \infty$, $E_\psi(Z_l^2) = O(l)$ by Proposition 1 so (2.13) is $O(l^{1/2})$. It follows that

$$\chi_{0,n}\left(\frac{u}{\sqrt{n}}\right) - \prod_{k=1}^{m}\chi_{L+(k-1)l,L+kl}\left(\frac{u}{\sqrt{n}}\right) \to 0.$$

The characteristic function on the right is that of the sum of m independent random variables and by Lindeberg's theorem it converges to $e^{-u^2\sigma^2/2}$ if, as $n \to \infty$, then $l \to \infty$,

(i) $E_{x,\vec{\theta}}(Z_{L,L+ml})/\sqrt{n} \to 0$,
(ii) $(1/n)\sum_{k=1}^{m} E_{x,\vec{\theta}}(Z_{L+(k-1)l,L+kl}^2) \to \sigma^2$,

(iii) for each $\varepsilon > 0$,

$$\frac{1}{n}\sum_{k=1}^{m}\int_{[|Z_{L+(k-1)l,\,L+kl}|>\varepsilon\sqrt{n}]} Z^2_{L+(k-1)l,\,L+kl}\,dP_{x,\vec{\theta}} \to 0.$$

Now (i) holds by Lemma 2 and (ii) holds by part 1 of the proof and the proposition, both for Φ-a.e. $(x,\vec{\theta})$. Finally, for the Lindeberg condition (iii), note first that for any fixed integer N, applying the arguments of part 1 of the proof,

$$\lim_{n\to\infty}\frac{1}{m}\sum_{k=1}^{m} E_{x,\vec{\theta}}\Big(Z^2_{L+(k-1)l,\,L+kl}I_{[|Z_{L+(k-1)l,\,L+kl}|>N]}\Big) = E_{\psi}\Big(Z_l^2 I_{[|Z_l|>N]}\Big)$$

and this holds for all N for Ψ-a.e. $(x,\vec{\theta})$. As $N\to\infty$, the right-hand side converges to 0, and since $\varepsilon\sqrt{n} > N$ for all n sufficiently large and any $\varepsilon > 0$, it follows that (iii) holds for Ψ-a.e. $(x,\vec{\theta})$.

This proves the theorem for Ψ-a.e. $(x,\vec{\theta})$. If $\vec{\theta}\notin\Gamma$, then $Z_1/\sqrt{n}\to 0$ so $Z_n/\sqrt{n}$ and $(Z_n - Z_1)/\sqrt{n}$ have the same limiting distribution. But then

$$\lim_{n\to\infty} P_{x,\vec{\theta}}\left[\frac{Z_n}{\sqrt{n}} < z\right]$$

$$= \lim_{n\to\infty} E_{x,\vec{\theta}}\left(P_{X_{t_1},T^{t_1}\vec{\theta}}\left[\frac{(Z_n - Z_1)}{\sqrt{n}} < z\right]\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-x^2/2}\,dx$$

as $n\to\infty$ for Φ-a.e. $(x,\vec{\theta})$. □

Proof of Theorem 2. 1. For a given n, let $t_m < n \le t_{m+1}$. If (M1) holds, then consider the decomposition

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\big(f(X_k) - E_{x,\vec{\theta}}f(X_k)\big)$$

$$= \frac{1}{\sqrt{n}}Z_{m+1} - \frac{1}{\sqrt{n}}\sum_{k=n}^{t_{m+1}-1}\bar{f}(X_k) - \frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}\big(E_{x,\vec{\theta}}f(X_k) - \bar{E}_{x,\vec{\theta}}f(X_k)\big). \tag{2.14}$$

If $n = t_{m+1}$, then interpret the middle term to be 0. The third term is $O(1/\sqrt{n})$ by Lemma 2. For the first term note that, since $t_m/m \to 1/\pi(\Gamma)$ for Φ-a.e. $(x,\vec{\theta})$, it follows that $t_{m+1}/t_m \to 1$ and $t_{m+1}/n \to 1$ for n in the above range, so

$$\frac{1}{\sqrt{n}}Z_{m+1} = \frac{1}{\sqrt{m+1}}Z_{m+1}\sqrt{\frac{m+1}{t_{m+1}}}\sqrt{\frac{t_{m+1}}{n}} \to_{\mathscr{D}} \mathscr{N}(0,\sigma^2\pi(\Gamma))$$

as $m\to\infty$, hence as $n\to\infty$. Let ν_n be the first k such that $t_k \ge n$ and set $V_k = \sum_{k=n}^{t_{\nu_n}-1}\bar{f}(X_k)$. The V_n are stationary under P_ϕ and (M1) states that

$E_\phi(V_0^2) < \infty$, hence for any $\varepsilon > 0$,

$$E_\phi\left(\sum_{n=0}^{\infty} P_{x,\vec{\theta}}\left[|V_n| > \varepsilon\sqrt{n}\right] = \sum_{n=0}^{\infty} P_\phi\left[|V_0| > \varepsilon\sqrt{n}\right]\right) < \infty.$$

It follows that

$$\sum_{n=0}^{\infty} P_{x,\vec{\theta}}\left[|V_n| > \varepsilon\sqrt{n}\right] = E_{x,\vec{\theta}}\left(\sum_{n=0}^{\infty} I_{[|V_n| > \varepsilon\sqrt{n}]}\right) < \infty$$

for all $(x, \vec{\theta})$ outside a Φ-null set N_ε. Letting $N = \bigcup_{k=1}^{\infty} N_{1/k}$, it follows by Cantelli's lemma that $V_n/\sqrt{n} \to 0$ a.s.-$P_{x,\vec{\theta}}$ for all $(x, \vec{\theta}) \notin N$. The theorem follows in this case.

2. Suppose now that (M2) or (M3) holds and note that (M2) implies (M3). Again take $t_m < n \le t_{m+1}$ and consider the decomposition

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \left(f(X_k) - E_{x,\vec{\theta}} f(X_k)\right) \\ &= \frac{1}{\sqrt{n}} Z_m + \frac{1}{\sqrt{n}} \sum_{k=t_m}^{n-1} \bar{f}(X_k) - \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \left(E_{x,\vec{\theta}} f(X_k) - \overline{E}_{x,\vec{\theta}} f(X_k)\right). \end{aligned} \tag{2.15}$$

The first and third terms converge to 0 as in part 1 of the proof. Letting $\tilde{V}_m = \max_{t_m \le l < t_{m+1}} |\sum_{k=t_m}^{l} \bar{f}(X_k)|$, condition (M3) implies that $E_\psi(\tilde{V}_0^2) < \infty$. The middle term of (2.15) is bounded by $\tilde{V}_m/\sqrt{m}$ and this converges to 0 a.s.-$P_{x,\vec{\theta}}$ for Ψ-a.e. $(x, \vec{\theta})$ by an argument that parallels that given in part 1. Finally, an argument like that at the end of the proof of theorem 1 extends the result to Φ-a.e. $(x, \vec{\theta})$. □

Proof of Theorem 3. 1. Let $h_j(x, \vec{\theta}) = E_{x,\vec{\theta}}(Y_0 Y_j)$ and $g_m(x, \vec{\theta}) = E_{x,\vec{\theta}}(\sum_{l=m}^{\infty} \rho^{1/2} Y_l^2)$. Then as in the proof of Theorem 1, Lemma 4 yields

$$\frac{1}{n} \sum_{k=0}^{n-1} E_{x,\vec{\theta}}(Y_k Y_{k+j}) \to E_\psi(Y_0 Y_j), \tag{2.16}$$

$$\frac{1}{n} \sum_{k=0}^{n-1} \sum_{l=m}^{\infty} \rho^{1/2} E_{x,\vec{\theta}}(Y_{k+l}^2) \to E_\psi(Y_0^2) \sum_{l=m}^{\infty} \rho^{1/2} \tag{2.17}$$

for all j, m and Ψ-a.e. $(x, \vec{\theta})$. Let $c_j = E_\psi(Y_0 Y_j)$ and consider the decomposition

$$\begin{aligned} E_{x,\vec{\theta}}\left[\frac{1}{n} \left(\sum_{k=0}^{n-1} Y_k \right)^2 \right] - \sigma^2 &= \frac{1}{n} \sum_{k=0}^{n-1} E_{x,\vec{\theta}}(Y_k^2) - c_0 \\ &+ \frac{2}{n} \sum_{j=1}^{m-1} \sum_{k=0}^{n-1-j} \left\{ E_{x,\vec{\theta}}(Y_k Y_{j+k}) - c_j \right\} \\ &+ \frac{2}{n} \sum_{j=m}^{n-1} \sum_{k=0}^{n-1-j} E_{x,\vec{\theta}}(Y_k Y_{j+k}) - \frac{2}{n} \sum_{j=1}^{m-1} j c_j - 2 \sum_{j=m}^{\infty} c_j \\ &=: \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V}. \end{aligned} \tag{2.18}$$

For fixed m, I and II converge to 0 as $n \to \infty$ by (2.16), while V converges to 0 by Proposition 1 and in any case IV converges to 0. Finally by Lemma 1,

$$
\begin{aligned}
|\mathrm{III}| &\le \frac{4}{n}\sum_{k=0}^{n-1}\sum_{j=m}^{\infty}\rho^{j/2}E_{x,\vec{\theta}}^{1/2}(Y_k^2)E_{x,\vec{\theta}}^{1/2}(Y_{j+k}^2)\\
&\le \frac{2}{n}\sum_{k=0}^{n-1}E_{x,\vec{\theta}}(Y_k^2)\sum_{j=m}^{\infty}\rho^{j/2}+\frac{2}{n}\sum_{k=0}^{n-1}\sum_{j=m}^{\infty}\rho^{j/2}E_{x,\vec{\theta}}(Y_{j+k}^2)\to 0
\end{aligned}
$$

as $n \to \infty$, then $m \to \infty$ by (2.16) and (2.17). Thus $E_{x,\vec{\theta}}(\sum_{k=0}^{n-1} Y_k)^2/n \to \sigma^2$. Now

$$
\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{k=0}^{t_n-1}\left(f(X_k,T^k\vec{\theta})-E_{x,\vec{\theta}}f(X_k,T^k\vec{\theta})\right)\\
&\quad=\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1}Y_k-\frac{1}{\sqrt{n}}\sum_{k=0}^{t_n-1}\left(E_{x,\vec{\theta}}f(X_k,T^k\vec{\theta})-\bar{E}_{x,\vec{\theta}}f(X_k,T^k\vec{\theta})\right).
\end{aligned}
$$

For a given $(x,\vec{\theta})$, the second term is nonrandom and converges to 0 as $n \to \infty$ and the first assertion in Theorem 3 follows.

2. Suppose (M1) holds, let ν_n be as in the proof of Theorem 2 and consider the decomposition (2.14). For given $(x,\vec{\theta})$, the third term is nonrandom and converges to 0, the second term is $-V_n/\sqrt{n}$ and it follows by the second part of Lemma 4 that $E_{x,\vec{\theta}}(V_n^2)/n \to 0$ as $n \to \infty$ for Φ-a.e. $(x,\vec{\theta})$. For the first term, by part 1 of this proof,

$$
\frac{1}{n}E_{x,\vec{\theta}}(Z_{m+1}^2)=\frac{t_{m+1}}{n}\frac{m+1}{t_{m+1}}\frac{1}{m+1}E_{x,\vec{\theta}}(Z_{m+1}^2)\to\sigma^2\pi(\Gamma)=\sigma^2(f).
$$

The second assertion follows from these relations.

3. If (M3) holds, then consider the decomposition (2.15). The middle term is bounded by $\tilde{V}_m/\sqrt{m}$ and $E_{x,\vec{\theta}}(\tilde{V}_m^2)/m \to 0$ as $n \to \infty$ for Ψ-a.e. $(x,\vec{\theta})$ by the second part of Lemma 4. □

The next result generalizes an equality established in [2] which in turn is a generalization of Kac's formula: $\int_\Lambda \tau_1\, d\pi = 1$, where Λ is an arbitrary set in $\mathscr{B}$ with $\pi(\Lambda) > 0$ and τ_1 denotes the first return time to Λ. Let $\tau_0 = 0$ and τ_n be the first time after τ_{n-1} that $T^n\vec{\theta} \in \Lambda$. Also let $g(n)$ be a nondecreasing function of integers $n \ge 1$ and let $\Delta g(n) = g(n+1) - g(n)$.

LEMMA 5. *For any set $\Lambda \in \mathscr{B}$ with $\pi(\Lambda) > 0$ and any integer $n \ge 1$,*

$$
\int_\Lambda g(\tau_n)\,d\pi=\int_\Lambda\left(g(\tau_{n-1})+\Delta g(\tau_{n-1})\right)d\pi+\int_{\vec{\Theta}-\Lambda}\Delta g(\tau_n)\,d\pi. \tag{2.19}
$$

In particular, taking $g(n) = n^2$,

$$
\begin{aligned}
2\int\tau_n\,d\pi&=\int_\Lambda\left(\tau_n^2-\tau_{n-1}^2\right)d\pi+1\\
&=\int_\Lambda\tau_{n-1}(\tau_n-\tau_{n-1})\,d\pi+\int_\Lambda\tau_1^2\,d\pi+1
\end{aligned}
\tag{2.20}
$$

and

$$\int(\tau_{n+1}-\tau_n)\,d\pi \le \int_\Lambda \tau_1^2\,d\pi = 2\int \tau_1\,d\pi - 1. \tag{2.21}$$

PROOF. 1. Let $A_{k,n} = \{\vec{\theta}\colon\ \vec{\theta} \in \Lambda, \tau_k = n\}$ and $A^*_{k,n} = \{\vec{\theta}\colon\ \vec{\theta} \notin \Lambda, \tau_k = n\}$. Then using stationarity,

$$\begin{aligned}\pi[\tau_k = n] &= \pi A_{k,n} + \pi A^*_{k,n}\\ &= \pi\left[T\vec{\theta} \in \Lambda, \tau_k = n\right] + \pi\left[T\vec{\theta} \notin \Lambda, \tau_k = n\right]\\ &= \pi A_{k-1,n-1} + \pi A^*_{k,n-1}.\end{aligned}$$

Hence

$$\pi A_{k,n} = \pi A_{k-1,n-1} + \pi A^*_{k,n-1} - \pi A^*_{k,n}$$

and

$$\sum_{m=l+1}^{\infty} \pi A_{k,m} = \sum_{m=l}^{\infty} \pi A_{k-1,m} + \pi A^*_{k,l},$$

since the $A^*_{k,m}$ are disjoint so $\pi A^*_{k,m} \to 0$ as $m \to \infty$. Now setting $g(0) = 0$, we have

$$\begin{aligned}\int_\Lambda g(\tau_n)\,d\pi &= \sum_{m=n}^{\infty} g(m)\pi A_{n,m} = \sum_{m=n}^{\infty} \pi A_{n,m} \sum_{l=0}^{m-1} \Delta g(l)\\ &= \sum_{l=0}^{\infty} \Delta g(l) \sum_{m=l+1}^{\infty} \pi A_{n,m}\\ &= \sum_{l=0}^{\infty} \Delta g(l)\pi A_{n-1,l} + \sum_{l=0}^{\infty} \Delta g(l) \sum_{m=l+1}^{\infty} \pi A_{n-1,m}\\ &\quad + \sum_{l=0}^{\infty} \Delta g(l)\pi A^*_{n,l}\\ &= \sum_{l=0}^{\infty} \Delta g(l)\pi A_{n-1,l} + \sum_{m=1}^{\infty} g(m)\pi A_{n-1,m} + \sum_{l=0}^{\infty} \Delta g(l)\pi A^*_{n,l}\\ &= \int_\Lambda (g(\tau_{n-1}) + \Delta g(\tau_{n-1}))\,d\pi + \int_{\vec{\Theta}-\Lambda} \Delta g(\tau_n)\,d\pi.\end{aligned}$$

The interchanges of order of summation are valid since all the quantities except possible $\Delta g(0)$ are nonnegative.

2. If $g(n) = n^2$, then $\Delta g(n) = 2n + 1$ and (2.19) gives

$$\begin{aligned}\int_\Lambda \tau_n^2\,d\pi &= \int_\Lambda (\tau_{n-1}^2 + 2\tau_{n-1})\,d\pi + 2\int_{\vec{\Theta}-\Lambda} \tau_n\,d\pi + 1\\ &= \int_\Lambda \tau_{n-1}^2\,d\pi + 2\int \tau_n\,d\pi - 1,\end{aligned}$$

since $\int_\Lambda \tau_n\,d\pi = \int_\Lambda \tau_{n-1}\,d\pi + 1$ by Kac's formula. Using stationarity, this leads to

$$2\int \tau_n \, d\pi = \int_\Lambda (\tau_n^2 - \tau_{n-1}^2) \, d\pi + 1$$

$$= 2\int_\Lambda \tau_{n-1}(\tau_n - \tau_{n-1}) \, d\pi + \int_\Lambda \tau_1^2 \, d\pi + 1.$$

Hence

$$\int (\tau_{n+1} - \tau_n) \, d\pi = \int_\Lambda (\tau_n(\tau_{n+1} - \tau_n) - \tau_{n-1}(\tau_n - \tau_{n-1})) \, d\pi$$

$$= \int_\Lambda (\tau_n(\tau_{n+1} - \tau_n) - (\tau_n - \tau_1)(\tau_{n+1} - \tau_n)) \, d\pi$$

$$= \int_\Lambda \tau_1(\tau_{n+1} - \tau_n) \, d\pi$$

using stationarity again at the second equality. Finally, by Schwarz's inequality, stationarity and (2.20) with $n = 1$,

$$\int (\tau_{n+1} - \tau_n) \, d\pi \le \int_\Lambda \tau_1^2 \, d\pi = 2\int \tau_1 \, d\pi - 1. \qquad \square$$

LEMMA 6. *If $\int t_1 \, d\pi < \infty$, then for any $r > 0$,*

$$\int \sum_{n=1}^{\infty} \rho_n(\vec{\theta})^r \, d\pi < \infty.$$

PROOF. Since $\rho_n(\vec{\theta}) \le \alpha\rho^k$ for $n \ge t_k$ by formula (1.8), Lemma 5 implies that

$$\int \sum_{n=1}^{\infty} \rho_n(\vec{\theta})^r \, d\pi = \int \sum_{k=0}^{\infty} \sum_{n=t_k}^{t_{k+1}-1} \rho_n(\vec{\theta})^r \, d\pi$$

$$\le \int \sum_{k=0}^{\infty} \alpha\rho^{kr}(t_{k+1} - t_k) \, d\pi \le \sum_{k=0}^{\infty} \alpha\rho^{kr}\left(2\int t_1 \, d\pi - 1\right) < \infty.$$

$\square$

PROOF OF THEOREM 4. To simplify notation let $\bar{f}_k = \bar{f}(X_k, T^k\vec{\theta})$. By hypothesis, $|f| \le b < \infty$ so $|\bar{f}_k| \le 2b$. Since $(\sum_{n=0}^{t_1-1} |\bar{f}_k|)^2 \le 4b^2 t_1^2$, (M2) holds provided $\int_\Gamma t_1^2 \, d\pi < \infty$ and by Lemma 5 this holds since $\int t_1 \, d\pi < \infty$.

We have $\overline{E}_{x,\vec{\theta}} \bar{f}_k = 0$ and $\bar{f}_k \le 4b^2$, so by Lemma 1,

$$\left|\overline{E}_{x,\vec{\theta}}(\bar{f}_0 \bar{f}_n)\right| \le 2\rho_n(\vec{\theta})^{1/2} \overline{E}_{x,\vec{\theta}}^{1/2}(\bar{f}_0^2)\overline{E}_{x,\vec{\theta}}^{1/2}(\bar{f}_n^2) \le 8b^2\rho_n(\vec{\theta})^{1/2}$$

and

$$\left|E_\phi(\bar{f}_0 \bar{f}_n)\right| = \left|\int \overline{E}_{x,\vec{\theta}}(\bar{f}_0 \bar{f}_n) \, d\Phi(x, \vec{\theta})\right| \le 8b^2 \int \rho_n(\vec{\theta})^{1/2} \, d\pi.$$

Then Lemma 6 implies the series in (1.15) converges absolutely. That $E_\phi(\sum_{k=0}^{n-1} \bar{f}_k / \sqrt{n})^2$ converges to this limit follows by an argument similar to that in the proof of Lemma 7.3, Chapter V of Doob [6]. To establish (1.16) for $E_{x,\vec{\theta}}$, let

$$h_j(x,\vec{\theta}) = E_{x,\vec{\theta}}(\bar{f}_0 \bar{f}_k),$$

$$g_m(x,\vec{\theta}) = E_{x,\vec{\theta}}\left(\bar{f}_0^2 \sum_{l=m}^{\infty} \rho_l(\vec{\theta})^{1/2}\right),$$

$$g'_m(x,\vec{\theta}) = E_{x,\vec{\theta}}\left(\sum_{l=m}^{\infty} \rho_l(\vec{\theta})^{1/2} \bar{f}_l^2\right).$$

These functions are all P_ϕ integrable, as follows easily from Lemma 6 since f is bounded. As in the proof of Theorem 1, Lemma 4 then yields

$$\frac{1}{n}\sum_{k=0}^{n-1} E_{x,\vec{\theta}}(\bar{f}_k \bar{f}_{k+l}) \to E_\phi(\bar{f}_0 \bar{f}_l),$$

$$\frac{1}{n}\sum_{k=0}^{n-1}\sum_{l=m}^{\infty} \rho_l(\vec{\theta})^{1/2} E_{x,\vec{\theta}}(\bar{f}_k^2) \to E_\phi\left(\bar{f}_0^2 \sum_{l=m}^{\infty} \rho_l(\vec{\theta})^{1/2}\right),$$

$$\frac{1}{n}\sum_{k=0}^{n-1}\sum_{l=m}^{\infty} \rho_l(\vec{\theta})^{1/2} E_{x,\vec{\theta}}(\bar{f}_{k+l}^2) \to E_\phi\left(\sum_{l=m}^{\infty} \rho_l(\vec{\theta})^{1/2} \bar{f}_l^2\right)$$

for all j, m and Φ-a.e. $(x,\vec{\theta})$. Letting $E_\phi(\bar{f}_0 \bar{f}_j)$ replace c_j and $\bar{f}_k$ replace Y_k, we can make the same decomposition as in (2.18). Terms I, II, IV and V converge to 0 as before. For the middle term,

$$|\mathrm{III}| \le 16b^2 \frac{1}{n}\sum_{j=m}^{n-1}\sum_{k=0}^{n-1} \rho_j(T^k\vec{\theta})^{1/2} \le 16b^2 \frac{1}{n}\sum_{k=0}^{n-1}\sum_{j=m}^{\infty} \rho_j(T^k\vec{\theta})^{1/2}.$$

By the ergodic theorem this converges to $16b^2\sum_{j=m}^{\infty} \int \rho_j(\vec{\theta})^{1/2}\, d\pi$ as $n \to \infty$, then this limit converges to 0 as $m \to \infty$ by Lemma 6. The assertion follows. □

Proof of Corollary 1. By Theorem 4, $\bar{f}^{(b)}$ satisfies (M2), hence by Theorem 2,

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} \bar{f}_k^{(b)} \to_{\mathscr{D}} \mathscr{N}(0, \sigma^2(\bar{f}^{(b)}))$$

as $n \to \infty$, where $\bar{f}_k^{(b)} = \bar{f}^{(b)}(X_k, T^k\vec{\theta})$. Similarly, $\bar{f} - \bar{f}^{(b)}$ satisfies (M2) since $\bar{f}$ does by hypothesis and by Theorem 2,

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} (\bar{f}_k - \bar{f}_k^{(b)}) \to_{\mathscr{D}} \mathscr{N}(0, \sigma^2(\bar{f} - \bar{f}^{(b)})).$$

But by Lemma 1, Proposition 1 and Theorem 1, letting σ_b^2 be the quantity in (1.10) when f is replaced by $\bar{f} - \bar{f}^{(b)}$,

$$\sigma^2(\bar{f} - \bar{f}^{(b)}) = \pi(\Gamma)\sigma_b^2 \le E_\psi\left(\sum_{k=0}^{t_1-1} (\bar{f}_k - \bar{f}_k^{(b)})^2 \left(1 + 4\sum_{n=1}^{\infty} \rho^{n/2}\right)\right).$$

This quantity converges to 0 as $b \to \infty$ by the dominated convergence theorem. It follows that $\sigma^2(\bar{f}^{(b)}) \to \sigma^2(\bar{f})$ and from the definition it is clear that $\sigma^2(\bar{f}) = \sigma^2(f)$ and (1.17) follows directly from this. □

REFERENCES

[1] COGBURN, R. (1972). The central limit theorem for Markov processes. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 485–512. Univ. California Press.

[2] COGBURN, R. (1975). A uniform theory for sums of Markov chain transition probabilities. *Ann. Probab.* **3** 191–214.

[3] COGBURN, R. (1984). The ergodic theory of Markov chains in random environments. *Z. Wahrsch. Verw. Gebiete* **66** 109–128.

[4] COGBURN, R. (1990). On direct convergence and periodicity for transition probabilities of Markov chains in random environments. *Ann. Probab.* **18** 642–654.

[5] DOBRUSHIN, R. L. (1956). Central limit theorem for nonstationary Markov chains. I, II. *Theory Probab. Appl.* **1** 65–80, 329–383.

[6] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[7] FOGUEL, S. R. (1969). *The Ergodic Theory of Markov Processes.* Van Nostrand, Princeton, N.J.

[8] GRIGORESCU, S. and OPRISAN, G. (1976). Limit theorems for J–X processes with a general state space. *Z. Wahrsch. Verw. Gebiete* **35** 65–73.

[9] IBRAGIMOV, I. A. (1975). A note on the central limit theorem for dependent random variables. *Theory Probab. Appl.* **20** 135–141.

[10] IBRAGIMOV, I. A. and LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters–Noordhoff, Groningen.

[11] LIFSHITS, B. A. (1978). On the central limit theorem for Markov chains. *Theory Probab. Appl.* **23** 279–296.

[12] LOÈVE, M. (1978). *Probability Theory* **2**, 4th ed. Springer, New York.

[13] NAGAEV, S. V. (1957). Some limit theorems for stationary Markov chains. *Theory Probab. Appl.* **2** 378–406.

[14] NAGAEV, S. V. (1961). More exact statements of limit theorems for homogeneous Markov chains. *Theory Probab. Appl.* **6** 62–81.

[15] NAGAEV, S. V. (1962). The central limit theorem for Markov chains with discrete time. *Publ. Akad. Nauk Uzbek. SSSR Phys.-Math. Ser.* **2** 12–20.

[16] NAGAEV, S. V. (1962). Some questions on the theory of Markov processes with discrete time. *The All-Soviet Conference on Probability Theory, Vilnus* 145–147.

[17] NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-negative Operators.* Cambridge Univ. Press, New York.

[18] PETERSEN, K. (1983). *Ergodic Theory.* Cambridge Univ. Press, New York.

[19] ROSENBLATT, M. (1971). *Markov Processes. Structure and Asymptotic Behavior.* Springer, New York.

[20] SIL'VESTROV, D. S. (1970). Limit theorems for functionals of a process with piecewise constant sums of random variables determined on a semi-Markov process with a finite set of states. *Soviet Math. Dokl.* **11** 1623–1625.

[21] SIL'VESTROV, D. S. (1972). Remarks on the limit of a composite random function. *Theory Probab. Appl.* **17** 669–677.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEW MEXICO
ALBUQUERQUE, NEW MEXICO 87131