On the functional CLT via martingale approximation

Mikhail Gordin\textsuperscript{a} and Magda Peligrad\textsuperscript{b}\textsuperscript{*}

\textsuperscript{a} POMI (Saint Petersburg Department of the Steklov Institute of Mathematics), 27 Fontanka emb., Saint Petersburg 191023, RUSSIA. E-mail address: gordin@pdmi.ras.ru

\textsuperscript{b} Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA. E-mail address: peligrm@math.uc.edu

\textbf{AMS 2000 subject classifications}. Primary 60F17; secondary 60J10.

\textbf{Key words and phrases}. Conditional functional central limit theorem, martingale approximation, mixing sequences, reversible Markov chain.

\textbf{Abstract}

In this paper we develop necessary and sufficient conditions for the validity of a martingale approximation for the partial sums of a stationary process in terms of the maximum of consecutive errors. Such an approximation is useful for transferring from the martingale to the original process the conditional functional central limit theorem. The condition found is simple and well adapted to a variety of examples, leading to a better understanding of the structure of several stochastic processes and their asymptotic behavior. The approximation brings together many disparate examples in probability theory. It is valid for classes of variables defined by familiar projection conditions such as Maxwell-Woodroofe condition, various classes of mixing processes including the large class of strongly mixing processes and for additive functionals of Markov chains with normal or symmetric Markov operators.

\section{Introduction and Results}

The objective of this paper is to find a characterization of stationary stochastic processes that can be studied via a martingale approximation in order to derive the functional central limit theorem for processes associated to partial sums.
There are several ways to present the results, since stationary processes can be introduced in several equivalent ways. We assume that $(\xi_n)_{n \in \mathbb{Z}}$ denotes a stationary Markov chain defined on a probability space $(\Omega, \mathcal{F}, P)$ with values in a measurable space $(S, \mathcal{A})$. The marginal distribution and the transition kernel are denoted by $\pi(A) = P(\xi_0 \in A)$ and $Q(\xi_0, A) = P(\xi_1 \in A | \xi_0)$. In addition $Q$ denotes the operator acting via $(Qf)(\xi) = \int_S f(s)Q(\xi, ds)$. Next let $L^2_0(\pi)$ be the set of functions on $S$ such that $\int f^2 d\pi < \infty$, and $\int f d\pi = 0$. Denote by $\mathcal{F}_k$ the $\sigma$-field generated by $\xi_i$ with $i \leq k$, $X_i = f(\xi_i)$.

$$S_n = \sum_{i=0}^{n-1} X_i \quad (i.e. \quad S_0 = 0, S_1 = X_0, S_2 = X_0 + X_1, \ldots).$$

For any integrable variable $X$ we denote $\mathbb{E}_k(X) = \mathbb{E}(X | \mathcal{F}_k)$. In our notation $\mathbb{E}_0(X_1) = Qf(\xi_0) = \mathbb{E}(X_1 | \xi_0)$.

Everywhere in the paper we assume $f \in L^2_0(\pi)$, in other words we assume $||X||_2 = (\mathbb{E}[X^2])^{1/2} < \infty$ and $\mathbb{E}[X_1] = 0$.

Notice that any stationary sequence $(Y_k)_{k \in \mathbb{Z}}$ can be viewed as a function of a Markov process $\xi_k = (Y_i; i \leq k)$, for the function $g(\xi_k) = Y_k$.

The stationary stochastic processes may be also introduced in the following alternative way. Let $T : \Omega \mapsto \Omega$ be a bijective bi-measurable transformation preserving the probability. Let $\mathcal{F}_0$ be a $\sigma$-algebra of $\mathcal{F}$ satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$. We then define the nondecreasing filtration $(\mathcal{F}_i)_{i \in \mathbb{Z}}$ by $\mathcal{F}_i = T^{-i}(\mathcal{F}_0)$. Let $X_0$ be a random variable which is $\mathcal{F}_0$-measurable. We define the stationary sequence $(X_i)_{i \in \mathbb{Z}}$ by $X_i = X_0 \circ T^i$.

In this paper we shall use both frameworks. In order to analyze the asymptotic behavior of the partial sums $S_n = \sum_{i=0}^{n-1} X_i$, Gordin in [15] proposed to decompose the sums related to the original stationary sequence into the sum

$$S_n = M_n + R_n \quad ,$$

(1)

of a square integrable martingale $M_n = \sum_{i=0}^{n-1} D_i$ adapted to $\mathcal{F}_n$ whose martingale differences $(D_i)$ are stationary, and a so-called coboundary $R_n$, i.e., a telescoping sum of random variables, with the basic property that $\text{sup}_n \mathbb{E}(R_n^2) < \infty$. More precisely, $X_n = D_n + Z_n - Z_{n-1}$, where $Z_n$ is another stationary sequence in $L^2$. Then, the limiting properties of the martingales, can be transported from martingale to the general sequence. In the context of Markov chains, the existence of such a decomposition is equivalent to the solvability of the Poisson equation in $L^2$.

For proving CLT for stationary sequences, a weaker form of martingale approximation was pointed out by many authors (see for instance
Merlevêde-Peligrad-Utev [21] for a survey). Recently, two interesting papers, one by Dedecker-Merlevêde-Volný [9] and the other by Zhao-Woodroofe [32], provided necessary and sufficient conditions for martingale approximation with an error term in (1) satisfying
\[ E((S_n - M_n)^2)/n \to 0. \]  

This decomposition is strong enough for transporting the conditional central limit theorem from sums of stationary ergodic martingale differences in \( L_2 \) to \( S_n/\sqrt{n} \). By conditional CLT, as discussed in Dedecker and Merlevêde [5], we understand in this context, that for any function \( f \) such that \( |f(x)|/(1 + x^2) \) is bounded and for any \( k \)
\[ E_k((f(S_n/\sqrt{n})) \to E(f(||D_0||_2N)) \text{ in } L_1, \]  
where \( N \) is a standard normal variable (here and everywhere in the paper we denote by \( ||.||_p \) the norm in \( L_p \)).

An important extension of this theory is to consider the conditional central limit theorem in its functional form. For \( t \in [0,1] \) define
\[ W_n(t) = \frac{S_{[nt]}}{n^{1/2}}, \]  
where \([x]\) denotes the integer part of \( x \). Notice that \( W_n(t) \) is an element of \( D(0,1) \), the space of all functions on \([0,1]\) which have left-hand limits and are continuous from the right. Then, by the CLT in the functional form, we understand that for any function \( f \) continuous and bounded on \( D(0,1) \) and for any \( k \),
\[ E_k(f(W_n)) \to E(f(cW)) \text{ in probability}; \]  
here \( c > 0 \) is a certain positive constant, \([x]\) denotes the integer part of \( x \) and \( W \) is the standard Brownian motion on \([0,1]\).

It is well known that a martingale with stationary and ergodic differences in \( L_2 \) satisfies this type of behavior with \( c = ||D_0||_2 \), that is at the heart of many statistical procedures. This conditional form of the invariance principle is a stable type of convergence that makes possible the change of measure with another absolutely continuous measure, as discussed in Billingsley [1], Rootzén [27], and Hall-Heyde [11].

With such a result in mind the question is now to find necessary and sufficient conditions for a martingale decomposition with the error term satisfying
\[ E(\max_{1 \leq j \leq n} (S_j - M_j)^2)/n \to 0. \]
In order to state our martingale approximation result, for \( m \) fixed we consider the stationary sequence

\[
Y_0^m = \frac{1}{m} \mathbb{E}_0(X_1 + \ldots + X_m), \quad Y_k^m = Y_0^m \circ T^k.
\]

(7)

In the Markov operators language we then have

\[
Y_0^m = \frac{1}{m}(Qf + \ldots + Q^m f)(\xi_0).
\]

It is convenient to introduce a semi-norm notation namely

\[
||Y_0^m||_{M+} = \limsup_{n \to \infty} \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \left| \sum_{j=1}^k Y_j^m \right|_2.
\]

**Theorem 1** Assume \((X_k)_{k \in \mathbb{Z}}\) is a stationary sequence of centered square integrable random variables. Then,

\[
||Y_0^m||_{M+} \to 0 \quad \text{as} \quad m \to \infty
\]

(8)

if and only if there exists a martingale with stationary increments satisfying (6). Such martingale is unique if it exists.

As a consequence of the proof of Theorem 1 we also obtain the following result that adds a new equivalent condition to the characterizations by Dedecker-Merlevède-Volný [9] and Zhao-Woodroofe [32]. With \((Y_k^m)_{k \in \mathbb{Z}}\) defined by (7) and the semi-norm notation

\[
||Y_0^m||_+ = \limsup_{n \to \infty} \frac{1}{\sqrt{n}} \left| \sum_{j=1}^n Y_j^m \right|_2
\]

we have the following characterization:

**Theorem 2** Assume \((X_k)_{k \in \mathbb{Z}}\) is as in Theorem 1. Then,

\[
||Y_0^m||_+ \to 0 \quad \text{as} \quad m \to \infty
\]

(9)

if and only if there exists a stationary martingale satisfying (2). Such a martingale is unique if it exists.
Our approach is constructive. If the stationary sequence is supposed to be ergodic, the constructed martingale differences are also ergodic and therefore the conditional theorems (3) or (5) can be easily transported to the original processes satisfying (9) or (8) respectively.

A natural and useful question is to provide classes of stochastic processes that have a martingale decomposition with an error term satisfying (6), in other words to provide sharp sufficient conditions for such a decomposition. Obviously, a maximal inequality is needed in order to verify this condition. We shall combine our approach with several maximal inequalities. One is due to Rio [26], page 53, formula (3.9); for related inequalities see Peligrad [23] and Dedecker-Rio [6].

- For any stationary process with centered variables in $\mathbb{L}_2$,

\[
\mathbb{E}( \max_{1 \leq i \leq n} S_i^2 ) \leq n\mathbb{E}(X_0^2) + 12 \sum_{k=1}^{n} \mathbb{E}|X_0\mathbb{E}_0(S_k)| . \tag{10}
\]

Another inequality comes from Peligrad-Utev [24], Proposition (2.3); see also Theorem 1 in Peligrad-Utev-Wu [25] for the inequality in $\mathbb{L}_p$.

- For any stationary process with centered variables in $\mathbb{L}_2$

\[
\mathbb{E}( \max_{1 \leq i \leq n} S_i^2 ) \leq n \left( 2\|X_0\|_2 + 3 \sum_{j=0}^{r-1} \frac{\|\mathbb{E}_0(S_{2^j})\|_2}{2^{j/2}} \right)^2 \leq n \left( 2\|X_0\|_2 + 80 \sum_{j=1}^{n} \frac{\|\mathbb{E}_0(S_j)\|_2}{j^{3/2}} \right)^2 , \tag{11}
\]

where $n \leq 2^r$.

The following maximal inequality is a particular case of Proposition 6 in Dedecker–Merlevêde [5]. See Theorem 1 in Wu [35] for the inequality in $\mathbb{L}_p$.

- For any stationary process with centered variables in $\mathbb{L}_2$ such that $E(X_0|\mathcal{F}_{-\infty}) = 0$ almost surely, we have

\[
\mathbb{E}( \max_{1 \leq i \leq n} S_i^2 ) \leq 2n \sum_{i=0}^{\infty} \|\mathbb{E}_{-i}(X_0) - \mathbb{E}_{-i-1}(X_0)\|_2 \tag{12}
\]
Another inequality we use for additive functionals of stationary reversible Markov chains is a consequence of Wu [33], Corollary 2.7 and relation (2.5) in the same paper (notice there is a typo in this relation, namely a square should be added to the norm); see also Sethuraman-Varadhan-Yau [31]:

• Assume \((\xi_n)_{n \in \mathbb{Z}}\) is a stationary, ergodic, reversible Markov chain. Then for every \(n \geq 1\)

\[
\mathbb{E}( \max_{1 \leq i \leq n} S_i^2 ) \leq (24n + 3) \sum_{n=0}^{\infty} \mathbb{E}(X_0 X_n) \tag{13}
\]

provided the series on the right hand side is convergent.

By combining the martingale decomposition in Theorem 1 with these maximal inequalities we point out various classes of stochastic processes for which conditional functional limit theorem holds. They include mixing processes and classes of Markov chains.

2 Proof of Theorem 1

The proof of this theorem has several steps.

Step 1. Construction of the approximating martingale.

The construction of the martingale decomposition is based on averages. It was introduced by Wu and Woodroofe [34] (see their definition (6) on the page 1677) and further developed in Zhao and Woodroofe [32], extending the construction in Heyde [12] and Gordin and Lifshitz [16]; see also Theorem 8.1 in Borodin and Ibragimov [3], and Kipnis and Varadhan [18]. We give the martingale construction here for completeness.

We introduce a parameter \(m \geq 1\) (kept fixed for the moment), and define the stationary sequence of random variables:

\[
\theta_0^m = \frac{1}{m} \sum_{i=1}^{m} \mathbb{E}_0(S_i), \quad \theta_k^m = \theta_0^m \circ T^k.
\]

Set

\[
D_k^m = \theta_{k+1}^m - \mathbb{E}_k(\theta_{k+1}^m); \quad M_n^m = \sum_{k=0}^{n-1} D_k^m. \tag{14}
\]
Then, \((D_k^m)_{k \in \mathbb{Z}}\) is a stationary martingale difference sequence and \((M_n^m)_{n \geq 0}\) is a martingale. So we have

\[X_k = D_k^m + \theta_k^m - \theta_{k+1}^m + \frac{1}{m} \mathbb{E}_k (S_{k+m+1} - S_{k+1})\]

and therefore

\[S_k = M_k^m + \theta_0^m - \theta_k^m + \sum_{j=1}^k \frac{1}{m} \mathbb{E}_{j-1} (S_{j+m} - S_j) \tag{15}\]

\[= M_k^m + \theta_0^m - \theta_k^m + R_k^m, \]

where we implemented the notation

\[R_k^m = \sum_{j=1}^k \frac{1}{m} \mathbb{E}_{j-1} (S_{j+m} - S_j). \tag{16}\]

Observe that

\[R_k^m = \sum_{j=0}^{k-1} Y_j^m. \tag{16}\]

With the notation

\[R_k^m = \theta_0^m - \theta_k^m + R_k^m \tag{17}\]

we have

\[S_k = M_k^m + R_k^m. \tag{18}\]

**Step 2. Sufficiency.**

We show that \(|Y_0^m||_{M^+} \to 0\) as \(m \to \infty\) is sufficient for (6).

We construct the martingale differences as in (14). By the martingale property and (18) for every positive integers \(m'\) and \(m''\) we have

\[||D_0^{m'} - D_0^{m''}||_2 = \frac{1}{\sqrt{n}}||M_n^{m'} - M_n^{m''}||_2 = \frac{1}{\sqrt{n}}||P_n^{m'} - P_n^{m''}||_2.\]

Now we let \(n \to \infty\). By the structure of the rest (17) and stationarity it follows that

\[
\limsup_{n \to \infty} \frac{1}{\sqrt{n}} ||R_n^{m'} - R_n^{m''}||_2 = \limsup_{n \to \infty} \frac{1}{\sqrt{n}} ||\overline{R}_n^{m'} - \overline{R}_n^{m''}||_2
\leq \limsup_{n \to \infty} \frac{1}{\sqrt{n}} (||\overline{R}_n^{m'}||_2 + ||\overline{R}_n^{m''}||_2).
\]
By (8), the limit when \( m' \) and \( m'' \) both tend to \( \infty \) is then 0, giving that \( (D_0^m) \) is Cauchy in \( L_2 \), therefore convergent. Denote its limit by \( D_0 \). Then \( M_n = \sum_{k=0}^{n-1} D_k \) is a martingale with the desired properties. To see this we start from the decomposition in relation (15) and obtain

\[
|S_k - M_k| \leq |M_k^m - M_k| + |\theta_k^m - \theta_k^0| + |R_k^m|.
\]

Then,

\[
\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |S_k - M_k|_2 \leq \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |M_k^m - M_k|_2 + \\
\frac{1}{\sqrt{n}} ||\theta_0^m||_2 + \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |\theta_k^m|_2 + \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |R_k^m|_2.
\]

By Doob's maximal inequality for martingales and by stationarity we conclude that

\[
\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |M_k^m - M_k|_2 \leq ||D_0^m - D_0||_2.
\]

For \( m \) fixed, since \((\theta_k^m)_{k \in \mathbb{Z}}\) is a stationary sequence of square integrable random variables, for any \( A > 0 \) we have

\[
\frac{1}{n} \mathbb{E}[ \max_{1 \leq k \leq n} |\theta_k^m|^2 ] \leq \frac{A^2}{n} + \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[|\theta_k^m|^2 I(|\theta_k^m| > A)]
\]

\[
= \frac{A^2}{n} + \mathbb{E}[|\theta_0^m|^2 I(|\theta_0^m| > A)]
\]

and then clearly

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E}[ \max_{1 \leq k \leq n} |\theta_k^m|^2 ] = 0.
\]

Then, taking into account (16), we easily obtain

\[
\limsup_{n \to \infty} \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |S_k - M_k|_2 \leq ||D_0^m - D_0||_2 + ||Y^m||_M^+,
\]

and the result follows by letting \( m \to \infty \), from the the fact that \( D_0^m \to D_0 \) in \( L_2 \). It is easy to see that the martingale is unique.

Assume that the martingale approximation (6) holds. With the notation $R_n = S_n - M_n$, we then have

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |R_k|_2 = 0 .$$

In particular this approximation implies

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |E(S_k | \mathcal{F}_0)|_2 = 0 . \quad (20)$$

From

$$||R_n||_2 \leq ||E(S_n | \mathcal{F}_0)||_2$$

we deduce

$$||R_n||_2 = ||\theta_n^n - \theta_n||_2 + ||\bar{R}_n^n||_2 \leq 2||\theta_n||_2 + ||\bar{R}_n^n||_2 \leq 3 \max_{1 \leq k \leq n} ||E(S_k | \mathcal{F}_0)||_2,$$

whence, by (20), it follows that

$$\lim_{n \to \infty} \frac{||R_n^n||_2}{\sqrt{n}} = 0.$$

As a consequence we obtain

$$\mathbb{E}(D_0^n - D_0)^2 = \frac{\mathbb{E}(M_n^n - M_n)^2}{n} = \frac{\mathbb{E}(R_n^n - R_n)^2}{n} \to 0 \text{ as } n \to \infty .$$

This shows that $D_0^n \to D_0$ in $\mathbb{L}_2$. By the triangle inequality, followed by Doob inequality, for any positive integer $m$ we have

$$\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |R_k|^2 \leq \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |R_k|^2 + \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |M_k^m - M_k|^2 \leq \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |R_k|^2 + ||D_0^m - D_0|| .$$

Now, letting $n \to \infty$ followed by $m \to \infty$, we obtain

$$\lim_{m \to \infty} \limsup_{n \to \infty} \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |R_k^m|^2 = 0 . \quad (21)$$
Observe now that by (17) \( R_m^m - \overline{R}_m^m = \theta_0^m - \theta_n^m \) where, for every fixed \( m \), by (19)

\[
\frac{1}{\sqrt{n}} \| \max_{1 \leq k \leq n} |\theta_0^m - \theta_k^m| \|_2 \to 0.
\]

Thus, we conclude from (21) that

\[
\lim_{m \to \infty} ||Y_0^m||_{M^+} = 0,
\]

and the necessity follows. ♦

3 Applications

3.1 Applications using projective criteria.

The first application involves the class of variables satisfying the Maxwell-Woodroofe condition [19].

**Proposition 3** Assume

\[
\Delta(X_0) = \sum_{k=1}^{\infty} \frac{||E_0(S_k)||_2}{k^{3/2}} < \infty.
\]

Then the martingale representation (6) holds.

**Proof.** In order to verify condition (8) of Theorem 1 we apply inequality (11) to the stationary sequence \((Y^m_k)_{k \in \mathbb{Z}}\) defined by (7). Then, \[
\left\| \max_{1 \leq j \leq n} \sum_{k=0}^{j-1} Y^m_k \right\|_2 \leq n^{1/2} \left( 2 ||Y_0^m||_2 + 80 \Delta(Y^m_0) \right).
\]

Notice first that by Proposition 2.5 in Peligrad-Utev [24] we know that condition (22) implies that \( ||Y_0^m||_2 \to 0 \). We finish the proof by showing

\[
\Delta(Y_0^m) \to 0.
\]

Since \( ||Y_0^m||_2 \to 0 \), by triangle inequality and stationarity, every term of the series in the right hand side of the equality

\[
\Delta(Y_0^m) = \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} ||E_0(Y_0^m + \ldots + Y^m_{k-1})||_2
\]
tends to 0 as \( m \to \infty \). Furthermore, because
\[
||\mathbb{E}_0(Y^m_0 + \ldots + Y^m_{k-1})||_2 = \left\| \mathbb{E}_0 \left( \frac{1}{m} \sum_{l=1}^{m} \sum_{i=0}^{k-1} \mathbb{E}_l(X_{i+l}) \right) \right\|_2 \\
\leq ||\mathbb{E}_0(X_0 + \ldots + X_{k-1})||_2,
\]
each term in \( \Delta(Y^m_0) \) is dominated by the corresponding term in \( \Delta(X_0) \), the latter being independent of \( m \). The result follows from the above considerations along with Lebesgue dominated convergence theorem for counting measure. ♦

For the sake of applications we give the following corollary

**Corollary 4** Assume
\[
\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} ||\mathbb{E}_0(X_n)||_2 < \infty .
\] (23)

Then the martingale representation (6) holds.

The fact that (23) implies (22) was verified in Maxwell and Woodroofe [19].

We shall combine now Theorem 1 with Rio’s maximal inequality (10) to obtain the following proposition:

**Proposition 5** Assume that for any \( j \geq 0 \)
\[
\Gamma_j = \sum_{k \geq j} ||X_j\mathbb{E}_0(X_k)||_1 < \infty , \text{ and } \frac{1}{m} \sum_{j=0}^{m-1} \Gamma_j \to 0 \quad \text{as } m \to \infty .
\] (24)

Then the martingale representation (6) holds.

Proof. In order to verify condition (8) we apply now the maximal inequality (10) to \( (Y^m_k)_{k \geq 1} \) defined by (7). We conclude that for \( n \geq m \)
\[
|| \max_{1 \leq j \leq n} \sum_{k=0}^{j-1} Y^m_k ||_2^2 \leq n||Y^m_0||_2^2 + 12 \sum_{j=1}^{n-1} ||Y^m_0\mathbb{E}_0(Y^m_1 + \ldots + Y^m_j)||_1 \\
\leq n(12m + 1)||Y^m_0||_2^2 + 12 \sum_{j=m+1}^{n-1} ||Y^m_0\mathbb{E}_0(Y^m_{m+1} + \ldots + Y^m_j)||_1
\]

11
where in the last sum we implemented a decomposition into two terms to deal with overlapping blocks. So, for an absolute constant $C$,

$$\frac{1}{n} \left\| \max_{1 \leq j \leq n} \sum_{k=0}^{j-1} Y_k^m \right\|^2 \leq C \left( \frac{\left\| \mathbb{E}_0(S_m) \right\|^2}{m} + \frac{1}{n} \sum_{l=m+1}^{n-1} \left\| Y_0^m \mathbb{E}_0(Y_{m+1}^m + \ldots + Y_l^m) \right\|_1 \right).$$

Since for any $l > m$

$$\left\| Y_0^m \mathbb{E}_0(Y_{m+1}^m + \ldots + Y_l^m) \right\|_1 \leq \frac{1}{m} \sum_{j=1}^{m} \sup_{i > m} \left\| (\mathbb{E}_0(X_j))^{\mathbb{E}_0(X_i + \ldots + X_{i+l})} \right\|_1$$

$$\leq \frac{1}{m} \sum_{j=1}^{m} \sum_{k \geq m} \left\| (\mathbb{E}_0(X_j))^{\mathbb{E}_0(X_k)} \right\|_1$$

and also

$$\left\| \mathbb{E}_0(S_m) \right\|^2 \leq 2 \sum_{j=0}^{m-1} \sum_{k=j}^{m} \left\| \mathbb{E}_0(X_j) \mathbb{E}_0(X_k) \right\|_1,$$

we then obtain by the properties of conditional expectations that for a certain absolute constant $C'$

$$\frac{1}{n} \left\| \max_{1 \leq j \leq n} \sum_{k=0}^{j-1} Y_k^m \right\|^2 \leq \frac{1}{m} \sum_{j=1}^{m} \sum_{k \geq m} \left\| X_j \mathbb{E}_0(X_k) \right\|_1$$

$$\leq \frac{C'}{m} \sum_{j=0}^{m} \sum_{k \geq j} \left\| X_j \mathbb{E}_0(X_k) \right\|_1,$$

and the result follows by condition (24) letting first $n \to \infty$ followed by $m \to \infty$. \(\diamondsuit\)

The projective criteria in the next proposition were studied in Heyde [11], Hannan [13], Gordin [17] among others.

**Proposition 6** Assume

$$E(X_0|\mathcal{F}_{-\infty}) = 0 \quad \text{almost surely and} \quad \sum_{i=1}^{\infty} \left\| \mathbb{E}_{-i}(X_0) - \mathbb{E}_{-i-1}(X_0) \right\|_2 < \infty.$$  

Then the martingale approximation (6) holds.
Proof. The validity of this proposition easily follows by verifying condition (8) via maximal inequality (12) applied to \((Y^m_k)_{k \geq 1}\) defined by (7). Indeed, by (12), triangle inequality and stationarity

\[
\frac{1}{n} \left\| \max_{1 \leq j \leq n} \left| \sum_{k=0}^{j-1} Y^m_k \right| \right\|^2_2 \leq 2 \sum_{i=1}^{\infty} \left\| \mathbb{E}_{-i}(Y^m_0) - \mathbb{E}_{-i-1}(Y^m_0) \right\|_2 \leq \frac{2}{m} \sum_{i=1}^{\infty} \sum_{k=1}^{m} \left\| \mathbb{E}_{-i}(X_k) - \mathbb{E}_{-i-1}(X_k) \right\|_2
\]

Now by stationarity, change of order of summation and change of variable

\[
\frac{1}{n} \left\| \max_{1 \leq j \leq n} \left| \sum_{k=0}^{j-1} Y^m_k \right| \right\|^2_2 \leq \frac{2}{m} \sum_{k=1}^{m} \sum_{j=k}^{\infty} \left\| \mathbb{E}_{-j}(X_0) - \mathbb{E}_{-j-1}(X_0) \right\|_2.
\]

To verify condition (8) we let \(n \to \infty\) followed by \(m \to \infty\). Notice that the term in the right hand side of the previous inequality tends to 0 as \(m \to \infty\) by (25).

### 3.2 Application to mixing sequences.

The results in the previous section can be immediately applied to mixing sequences leading to the sharpest possible results and providing additional information about the structures of these processes. Examples include various classes of Markov chains or Gaussian processes.

We shall also introduce the following mixing coefficients: For any two \(\sigma\)-algebras \(\mathcal{A}\) and \(\mathcal{B}\) define the strong mixing coefficient \(\alpha(\mathcal{A}, \mathcal{B})\):

\[\alpha(\mathcal{A}, \mathcal{B}) = \sup \{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| ; A \in \mathcal{A}, B \in \mathcal{B} \}\]

and the \(\rho\)--mixing coefficient, known also under the name of maximal coefficient of correlation \(\rho(\mathcal{A}, \mathcal{B})\):

\[\rho(\mathcal{A}, \mathcal{B}) = \sup \{ \text{Cov}(X, Y) / \|X\|_2 \|Y\|_2 : X \in L_2(\mathcal{A}), Y \in L_2(\mathcal{B}) \}\]

For the stationary sequence of random variables \((X_k)_{k \in \mathbb{Z}}\), we also define \(\mathcal{F}_m^n\) the \(\sigma\)--field generated by \(X_i\) with indices \(m \leq i \leq n\), \(\mathcal{F}_n^n\) denotes the \(\sigma\)--field generated by \(X_i\) with indices \(i \geq n\), and \(\mathcal{F}_m^n\) denotes the \(\sigma\)--field
generated by $X_i$ with indices $i \leq m$. The sequences of coefficients $\alpha(n)$ and $\rho(n)$ are then defined by

$$\alpha(n) = \alpha(\mathcal{F}_0, \mathcal{F}_n), \text{ and } \rho(n) = \rho(\mathcal{F}_0, \mathcal{F}_n).$$

Equivalently, (see Bradley [2], ch. 4)

$$\rho(n) = \sup\{\|E(Y|\mathcal{F}_0)\|_2/\|Y\|_2 : Y \in L_2(\mathcal{F}_n), \ E(Y) = 0\}.$$

Finally we say that the stationary sequence is strongly mixing if $\alpha(n) \to 0$ as $n \to \infty$, and $\rho-$mixing if $\rho(n) \to 0$ as $n \to \infty$.

An interesting application of Proposition 3 is to $\rho-$mixing sequences. It is well known that the central limit theorem and its invariance principle hold for stationary centered sequences with finite second moments under the minimal assumption

$$\sum_{k=1}^{\infty} \rho(2^k) < \infty. \quad (26)$$

where $\rho(n) = \rho(\mathcal{F}_0, \mathcal{F}_n)$. Let us recall that the central limit theorem is due to Ibragimov [14], while the invariance principle is found in [22], [28] and [29], [30]. The fact that the condition (26) is sharp in this context is due to Bradley ( [2], Vol 1p. 367 and Vol 3 ch. 34). Bradley’s example shows that if (26) fails, then $S_n/stdev(S_n)$ might have as weak limit points nondegenerate non-normal distributions.$\diamond$

As a corollary of Proposition 3 we obtain the conditional invariance principle for $\rho-$mixing sequences.

**Proposition 7** Assume $\sum_{k=1}^{\infty} \rho(2^k) < \infty$. Then the martingale representation (6) holds.

Proof. As in Merlevède-Peligrad-Utev [21], for a positive constant $C$, we have

$$\sum_{r=0}^{\infty} \frac{\|E(S^r|\mathcal{F}_0)\|_2}{2^{r/2}} \leq C \sum_{j=0}^{\infty} \rho(2^j).$$

$\diamond$

For getting optimal results for strongly mixing sequence, we shall use Proposition 5.
According to Doukhan-Massart-Rio [10], a condition that is optimal for CLT or invariance principle for strongly mixing sequences is:

\[\sum_{k \geq 1} \mathbb{E} |X_0^2| I(|X_0| \geq Q_{|X_0|}(2\alpha_k)) < \infty \]  \hspace{1cm} (27)

where \(Q_{|X_0|}\) denotes the cadlag inverse of the function \(t \to P(|X_0| > t)\). Also under this condition we add the additional information given by Theorem 1.

**Proposition 8** Assume condition (27) is satisfied. Then the martingale representation (6) holds.

We shall just verify the condition of Proposition 5. Note that, on the set \([0, P(|Y| > 0)]\), the function \(H_Y : x \to \int_0^x Q_Y(u)du\) is an absolutely continuous and increasing function with values in \([0, \mathbb{E}|Y|]\). Denote by \(G_Y\) the inverse of \(H_Y\). The following inequality is taken from Merlevède-Peligrad [20] generalization on Dedecker-Doukhan [7]

If \(Q_X = Q_Y = Q\),

\[\mathbb{E}|X\mathbb{E}(Y|\mathcal{M})| \leq 3 \int_0^{||\mathbb{E}(Y|\mathcal{M})||_1} Q \circ G(u)du.\]

where \(Y\) is centered. By this inequality

\[||X_j\mathbb{E}(X_k|\mathcal{F}_0)||_1 \leq 3 \int_0^{||\mathbb{E}(X_k|\mathcal{F}_0)||_1} Q_{|X_0|} \circ G(u)du\]

and then, we majorate the right hand side in previous inequality by Proposition 1 in Dedecker-Doukhan [7],

\[||X_j\mathbb{E}(X_k|\mathcal{F}_0)||_1 \leq 6 \int_0^{2\alpha(k)} Q_{|X_0|}^2 du.\]

Therefore

\[\sum_{k \geq j} ||X_j\mathbb{E}_0(X_k)||_1 \leq 6 \sum_{k \geq j} \int_0^{2\alpha(k)} Q_{|X_0|}^2 du \leq 6 \sum_{k \geq j} \mathbb{E}X_0^2 I(|X_0| \geq Q_{|X_0|}(2\alpha_k)) \rightarrow 0 \hspace{0.5cm} \text{as} \hspace{0.5cm} j \rightarrow \infty\]

\[\Diamond\]
Notice that the coefficient $\alpha(k)$ is defined by using only one variable in the future. Moreover, by Cauchy Schwarts inequality condition (27) is satisfied if the variables have finite moments of order $2 + \delta$ for a $\delta > 0$ and

$$\sum_{k \geq j} \alpha(k)^{\delta/(2+\delta)} < \infty.$$  

An excellent source of information for classes of mixing sequences and classes of Markov chains satisfying mixing conditions is the book by Bradley [2]. Further applications can be obtained by using the coupling coefficients in Dedecker and Prieur [8].

### 3.3 Application to additive functionals of reversible Markov chains.

For reversible Markov processes (i.e. $Q = Q^*$) the invariance principle under optimal condition is known since Kipnis and Varadhan [18]. Here is a formulation in terms of martingale approximation.

**Proposition 9** Let $(\xi_i)_{i \in \mathbb{Z}}$ denotes a stationary and ergodic reversible Markov chain and $f \in L^2_0(\pi)$, with the property

$$\lim_{n \to \infty} \frac{\text{var}(S_n)}{n} \to \sigma_f^2 < \infty.$$

Then the martingale representation (6) holds.

We have to verify condition (8). Define $Y_0^m$ by (7). By the maximal inequality (13), we have

$$\frac{1}{n} \mathbb{E}(\max_{1 \leq j \leq n} | \sum_{k=0}^{j-1} Y_k^m |^2) \leq 27 \sum_{k \geq 0} \mathbb{E}(Y_0^m Y_k^m).$$

Denote by $\rho_f$ the spectral measure of $f$ corresponding to self-adjoint operator $Q$ on $L_2(\pi)$. Spectral calculus for self-adjoint operator $Q$ gives

$$\sum_{k \geq 0} \mathbb{E}(Y_0^m Y_k^m) \leq \frac{1}{m^2} \int_{-1}^{1} \frac{(1 + t + \ldots + t^{m-1})^2}{(1 - t)^2} \rho_f(dt)$$
and therefore for every positive integer \( m > 0 \),
\[
\| Y_0^m \|_{M^+}^2 \leq 27 \int_{-1}^{1} \frac{(1 + t + \ldots + t^{m-1})^2}{m^2(1 - t)} \rho_f(dt) .
\]
It is well known that the assumption \( \lim_{n \to \infty} \frac{\text{var}(S_n)}{n} \to \sigma_f^2 \) is equivalent to \( \int_{-1}^{1} (1 - t)^{-1} \rho_f(dt) < \infty \) (see [18]) and then
\[
\lim_{m \to \infty} \| Y_0^m \|_{M^+}^2 = 0 .
\]

Similar results are expected to hold for other classes of stationary and ergodic Markov chains when \( Q \) is not necessarily self-adjoint but satisfies instead a quasi-symmetry or strong sector condition, or it is symmetrized. See Wu [33] and Sethuraman-Varadhan-Yau [31] for these related processes. We shall just point out the extension to normal operators.

### 3.4 Application to additive functional of normal Markov chains

For additive functionals of normal Markov chains \((QQ^* = Q^*Q)\) the central limit theorem below is a result of Gordin and Lifshitz [16]. As an application of Theorem 2 we give an alternative proof. We also point out a condition that implies invariance principle.

Let \( \rho_f \) be the spectral measure on the unit disk \( D \) corresponding to function \( f \in L_0^2(\pi) \).

**Proposition 10** For any ergodic and normal Markov chain assume
\[
\int_D \frac{1}{|1 - z|} \rho_f(dz) < \infty .
\]  
(28)

Then, the representation (2) holds.

Proof. According to Theorem 2 we have to check condition (9). By using spectral calculus as in [3, chap. 4], after some computations, we get
\[
\limsup_{n \to \infty} \frac{1}{n} \left\| \sum_{k=0}^{n-1} Y_k^m \right\|_2^2 \leq 4 \int_D \frac{|1 + z + \ldots + z^{m-1}|^2}{m^2|1 - z|} \rho_f(dz)
\]
and the condition (9) is therefore satisfied.

Condition (28) has an interesting equivalent form in terms of conditional moments, that it is in the spirit and implies Mawxell-Woodroofe condition (22).

**Remark 11**  Condition (28) is equivalent to

\[
\sum_{k=1}^{\infty} \frac{\|E_0(S_k)\|_2^2}{k^2} < \infty .
\]  (29)

Condition (29) is further implied by

\[
\sum_{k=1}^{\infty} \|E_0(X_k)\|_2^2 < \infty .
\]  (30)

The equivalence from this Remark is contained in Lemma 2.1 in [4]. The fact that (30) implies (29) is easily established as in the proof of (23) implies (22).
4 Acknowledgement

Mikhail Gordin was supported in part by a Charles Phelps Taft Memorial Fund grant and RFBR grant 09-01-91331-DFG.a.

Magda Peligrad was supported in part by a Charles Phelps Taft Memorial Fund grant and NSA grant H98230-09-1-0005.

The authors are grateful to the referees for carefully reading the paper and for numerous suggestions that improved the presentation of the paper.

References


