Proofs of the martingale FCLT[*]

Ward Whitt

Department of Industrial Engineering and Operations Research
Columbia University, New York, NY 10027-6699
e-mail: ww2040@columbia.edu
url: www.columbia.edu/∼ww2040

Abstract: This is an expository review paper elaborating on the proof of the martingale functional central limit theorem (FCLT). This paper also reviews tightness and stochastic boundedness, highlighting one-dimensional criteria for tightness used in the proof of the martingale FCLT. This paper supplements the expository review paper Pang, Talreja and Whitt (2007) illustrating the "martingale method" for proving many-server heavy-traffic stochastic-process limits for queueing models, supporting diffusion-process approximations.

Contents

[*]This is an original survey paper

1. Introduction

In this paper we elaborate upon the proof of the martingale FCLT in §7.1 of Ethier and Kurtz [6]. We also discuss alternative arguments, as in Jacod and Shiryaev [8]. The proof in [6] is correct, but it is concise, tending to require knowledge of previous parts of the book. We aim to make the results more accessible by sacrificing some generality and focusing on the "common case."

In addition to reviewing proofs of the martingale FCLT, we review tightness criteria, which play an important role in the proofs of the martingale FCLT. An important role is played by simple "one-dimensional" criteria for tightness of stochastic processes in D, going beyond the classical criteria in Billingsley [3, 5], reviewed in Whitt [22]. The alternative one-dimensional criteria come from Billingsley [4, 5], Kurtz [12, 13], Aldous [1, 2], Rebolledo [19] and Jacod et al. [7].

The martingale FCLT has many applications, so that this paper may aid in many contexts. However, we were motivated by applications to queueing models. Specifically, this paper is intended to supplement Pang, Talreja and Whitt [17], which is an expository review paper illustrating how to do martingale proofs of many-server heavy-traffic limit theorems for Markovian queueing models, as in Krichagina and Puhalskii [10] and Puhalskii and Reiman [18]. Pang et al. [17] review martingale basics, indicate how martingales arise in the queueing models and show how they can be applied to establish the stochastic-process limits for the queueing models. They apply the martingale method to the elementary $M/M/\infty$ queueing model and a few variations.

The rest of this paper is organized as follows: We start in §2 by stating a version of the martingale FCLT from p 339 of Ethier and Kurtz [6]. Next in §3 we review tightness, focusing especially on criteria for tightness. Then we turn to the proof of the martingale FCLT. We give proofs of tightness in §4; we give proofs of the characterization in §5.

2. The Martingale FCLT

We now state a version of the martingale FCLT for a sequence of local martingales $\{M_n : n \geq 1\}$ in D^k, based on Theorem 7.1 on p. 339 of Ethier and Kurtz [6], hereafter referred to as EK. We shall also make frequent reference to Jacod and Shiryayev [8], hereafter referred to as JS. See Section VIII.3 of JS for related results; see other sections of JS for generalizations.

We will state a special case of Theorem 7.1 of EK in which the limit process is multi-dimensional Brownian motion. However, the framework always produces limits with continuous sample paths and independent Gaussian increments. Most applications involve convergence to Brownian motion. Other situations are covered by JS, from which we see that proving convergence to discontinuous processes evidently is more complicated.

We assume familiarity with martingales; basic notions are reviewed in §3 of Pang et al. [17]. The key part of each condition in the martingale FCLT below is the convergence of the quadratic covariation processes. Condition (i) involves the optional quadratic-covariation (square-bracket) processes $[M_{n,i}, M_{n,j}]$, while condition (ii) involves the predictable quadratic-covariation (angle-bracket) processes $\langle M_{n,i}, M_{n,j} \rangle$. Recall that the square-bracket process is more general, being well defined for any local martingale (and thus any martingale), whereas the associated angle-bracket process is well defined only for any locally square-integrable martingale (and thus any square-integrable martingale); see §3.2 of [17].

Thus the key conditions below are the assumed convergence of the quadratic-variation processes in conditions (3) and (6). The other conditions (2), (4) and (6) are technical regularity conditions. There is some variation in the literature concerning the extra technical regularity conditions; e.g., see Rebolledo [19] and JS [8].

Let $\Rightarrow$ denote convergence in distribution and let $D \equiv D([0, \infty), \mathbb{R})$ be the usual space of right-continuous real-valued functions on the semi-infinite interval $[0, \infty)$ with limits from the left, endowed with the Skorohod [20] J_1 topology; see Billingsley [3, 5], EK, JS and Whitt [22] for background. For a function x in D, let $J(x, T)$ be the absolute value of the maximum jump in x over the interval $[0, T]$, i.e.,

$$J(x, T) \equiv \sup \left\{ |x(t) - x(t-)| : 0 < t \leq T \right\} . \tag{1}$$

Theorem 2.1 (multidimensional martingale FCLT) *For $n \geq 1$, let $M_n \equiv (M_{n,1}, \ldots, M_{n,k})$ be a local martingale in D^k with respect to a filtration $\boldsymbol{F}_n \equiv \{\mathcal{F}_{n,t} : t \geq 0\}$ satisfying $M_n(0) = (0, \ldots, 0)$. Let $C \equiv (c_{i,j})$ be a $k \times k$ covariance matrix, i.e., a nonnegative-definite symmetric matrix of real numbers.*

Assume that one of the following two conditions holds:

(i) *The expected value of the maximum jump in M_n is asymptotically negligible; i.e., for each $T > 0$,*

$$\lim_{n \to \infty} \left\{ E \left[J(M_n, T) \right] \right\} = 0 \tag{2}$$

and, for each pair (i, j) with $1 \leq i \leq k$ and $1 \leq j \leq k$, and each $t > 0$,

$$[M_{n,i}, M_{n,j}] (t) \Rightarrow c_{i,j} t \quad in \quad \mathbb{R} \quad as \quad n \to \infty . \tag{3}$$

(ii) *The local martingale M_n is locally square-integrable, so that the predictable quadratic-covariation processes $\langle M_{n,i}, M_{n,j} \rangle$ can be defined. The expected value of the maximum jump in $\langle M_{n,i}, M_{n,j} \rangle$ and the maximum squared jump of M_n are asymptotically negligible; i.e., for each $T > 0$ and (i, j) with $1 \leq i \leq k$ and $1 \leq j \leq k$,*

$$\lim_{n \to \infty} \left\{ E \left[J \left(\langle M_{n,i}, M_{n,j} \rangle, T \right) \right] \right\} = 0 , \tag{4}$$

$$\lim_{n \to \infty} \left\{ E \left[J \left(M_n, T \right)^2 \right] \right\} = 0 , \tag{5}$$

and

$$\langle M_{n,i}, M_{n,j}\rangle(t) \Rightarrow c_{i,j}t \quad in \quad \mathbb{R} \quad as \quad n \to \infty \tag{6}$$

for each $t > 0$ and for each (i, j).

Conclusion:

If indeed one of the the conditions (i) or (ii) above holds, then

$$M_n \Rightarrow M \quad in \quad D^k \quad as \quad n \to \infty , \tag{7}$$

*where M is a k-**dimensional** $(0, C)$-**Brownian motion**, having mean vector and covariance matrix*

$$E[M(t)] = (0, \dots, 0) \quad and \quad E[M(t)M(t)^{tr}] = Ct, \quad t \geq 0 , \tag{8}$$

where, for a matrix A, A^{tr} is the transpose.

Of course, a common simple case arises when C is a diagonal matrix; then the k component marginal one-dimensional Brownian motions are independent. When $C = I$, the identity matrix, M is a standard k-dimensional Brownian motion, with independent one-dimensional standard Brownian motions as marginals.

At a high level, Theorem 2.1 says that, under regularity conditions, convergence of martingales in D is implied by convergence of the associated quadratic covariation processes. At first glance, the result seems even stronger, because we need convergence of only the one-dimensional quadratic covariation processes for a single time argument. However, that is misleading, because the stronger weak convergence of these quadratic covariation processes in D^{k^2} is actually equivalent to the weaker required convergence in $\mathbb{R}$ for each t, i, j in conditions (3) and (6), as we will show below.

To state the result, let $[[M_n]]$ be the matrix-valued random element of D^{k^2} with $(i, j)^{\text{th}}$ component $[M_{n,i}, M_{n,j}]$; and let $\langle\langle M_n\rangle\rangle$ be the matrix-valued random element of D^{k^2} with $(i, j)^{\text{th}}$ component $\langle M_{n,i}, M_{n,j}\rangle$. Let e be the identity map, so that Ce is the matrix-valued deterministic function in D^{k^2} with elements $\{c_{i,j}t : t \geq 0\}$.

Lemma 2.1 (*modes of convergence for quadratic covariation processes*) *Let $[[M_n]]$ and $\langle\langle M_n\rangle\rangle$ be the matrix-valued quadratic-covariation processes defined above; let Ce be the matrix-valued deterministic limit defined above. Then condition (3) is equivalent to*

$$[[M_n]] \Rightarrow Ce \quad in \quad D^{k^2} , \tag{9}$$

while condition (6) is equivalent to

$$\langle\langle M_n\rangle\rangle \Rightarrow Ce \quad in \quad D^{k^2} , \tag{10}$$

Lemma 2.1 is important, not only for general understanding, but because it plays an important role in the proof of Theorem 2.1. (See the discussion after (37).)

Proof of Lemma 2.1. We exploit the fact that the ordinary quadratic variation processes are nondecreasing, but we need to do more to treat the quadratic covariation processes when $i \neq j$. Since $[M_{n,i}, M_{n,i}]$ and $\langle M_{n,i}, M_{n,i} \rangle$ for each i are nondecreasing and $\{c_{i,i}t : t \geq 0\}$ is continuous, we can apply §VI.2.b and Theorem VI.3.37 of JS to get convergence of these one-dimensional quadratic-variation processes in D for each i from the corresponding limits in $\mathbb{R}$ for each t. Before applying Theorem VI.3.37 of JS, we note that we have convergence of the finite-dimensional distributions, because we can apply Theorem 11.4.5 of Whitt [22]. We then use the representations

$$2[M_{n,i}, M_{n,j}] = [M_{n,i} + M_{n,j}, M_{n,i} + M_{n,j}] - [M_{n,i}, M_{n,i}] - [M_{n,j}, M_{n,j}]$$
$$2\langle M_{n,i}, M_{n,j} \rangle = \langle M_{n,i} + M_{n,j}, M_{n,i} + M_{n,j} \rangle - \langle M_{n,i}, M_{n,i} \rangle - \langle M_{n,j}, M_{n,j} \rangle ,$$

e.g., see §1.8 of Liptser and Shiryaev [15]. First, we obtain the limits

$$[M_{n,i} + M_{n,j}, M_{n,i} + M_{n,j}](t) \Rightarrow 2c_{i,j}t + c_{i,i}t + c_{j,j}t \quad \text{in} \quad \mathbb{R}$$

and

$$\langle M_{n,i} + M_{n,j}, M_{n,i} + M_{n,j} \rangle(t) \Rightarrow 2c_{i,j}t + c_{i,i}t + c_{j,j}t \quad \text{in} \quad \mathbb{R}$$

for each t and $i \neq j$ from conditions (3) and (6). The limits in (3) and (6) for the components extend to vectors and then we can apply the continuous mapping theorem with addition. Since $[M_{n,i}+M_{n,j}, M_{n,i}+M_{n,j}]$ and $\langle M_{n,i}+M_{n,j}, M_{n,i}+M_{n,j} \rangle$ are both nondecreasing processes, we can repeat the argument above for $[M_{n,i}, M_{n,i}]$ and $\langle M_{n,i}, M_{n,i} \rangle$ to get

$$[M_{n,i} + M_{n,j}, M_{n,i} + M_{n,j}] \Rightarrow (2c_{i,j} + c_{i,i} + c_{j,j})e \quad \text{in} \quad D$$

and

$$\langle M_{n,i} + M_{n,j}, M_{n,i} + M_{n,j} \rangle \Rightarrow (2c_{i,j} + c_{i,i} + c_{j,j})e \quad \text{in} \quad D$$

We then get the corresponding limits in D^3 for the vector processes, and apply the continuous mapping theorem with addition again to get

$$[M_{n,i}, M_{n,j}] \Rightarrow c_{i,j}e \quad \text{and} \quad \langle M_{n,i}, M_{n,j} \rangle \Rightarrow c_{i,j}e \quad \text{in} \quad D$$

for all i and j. Since these limits extend to vectors, we finally have derived the claimed limits in (9) and (10). ∎

Outline of the Proof

The broad outline of the proof of Theorem 2.1 is standard. As in both EK and JS, the proof is an application of Corollary 3.3: We first show that the sequence $\{M_n : n \geq 1\}$ is tight in D^k, which implies relative compactness by Theorem 3.1. Having established relative compactness, we show that the limit of any convergent subsequence must be k-dimensional Brownian motion with covariance matrix C. That is, there is a tightness step and there is a characterization step. Both EK and JS in their introductory remarks, emphasize the importance of the characterization step.

Before going into details, we indicate where the key results are in JS. Case (i) in Theorem 2.1 is covered by Theorem VIII.3.12 on p 432 of JS. Condition

3.14 on top of p 433 in JS is implied by condition (2). Condition (b.ii) there in JS is condition (3). The counterexample in Remark 3.19 on p 434 of JS shows that weakening condition 3.14 there can cause problems.

Case (ii) in Theorem 2.1 is covered by Theorem VIII.3.22 on p 435 of JS. Condition 3.23 on p 435 of JS is implied by condition (5). There ν is the predictable random measure, which is part of the characteristics of a semimartingale, as defined in §II.2 of JS and $*$ denotes the operator in 1.5 on p 66 of JS constructing the associated integral process. (We will not use the notions ν and $*$ here.)

3. Tightness

3.1. Basic Properties

We work in the setting of a complete separable metric space (CSMS), also known as a Polish space; see §§13 and 19 of Billingsley [5], §§3.8-3.10 of EK [6] and §§11.1 and 11.2 of Whitt [22]. (The space $D^k \equiv D([0,\infty),\mathbb{R})^k$ is made a CSMS in a standard way and the space of probability measures on D^k becomes a CSMS as well.) Key concepts are: closed, compact, tight, relatively compact and sequentially compact. We assume knowledge of metric spaces and compactness in metric spaces.

Definition 3.1 (tightness) *A set A of probability measures on a metric space S is **tight** if, for all $\epsilon > 0$, there exists a compact subset K of S such that*

$$P(K) > 1 - \epsilon \quad \text{for all} \quad P \in A \ .$$

A set of random elements of the metric space S is tight if the associated set of their probability laws on S is tight. Consequently, a sequence $\{X_n : n \geq 1\}$ of random elements of the metric space S is tight if, for all $\epsilon > 0$, there exists a compact subset K of S such that

$$P(X_n \in K) > 1 - \epsilon \quad \text{for all} \quad n \geq 1 \ .$$

Since a continuous image of a compact subset is compact, we have the following lemma.

Lemma 3.1 (continuous functions of random elements) *Suppose that $\{X_n : n \geq 1\}$ is a tight sequence of random elements of the metric space S. If $f : S \to S'$ is a continuous function mapping the metric space S into another metric space S', then $\{f(X_n) : n \geq 1\}$ is a tight sequence of random elements of the metric space S'.*

Proof. As before, let $\circ$ be used for composition: $(f \circ g)(x) \equiv f(g(x))$. For any function $f : S \to S'$ and any subset A of S, $A \subseteq f^{-1} \circ f(A)$. Let $\epsilon > 0$ be given. Since $\{X_n : n \geq 1\}$ is a tight sequence of random elements of the metric space S, there exists a compact subset K of S such that

$$P(X_n \in K) > 1 - \epsilon \quad \text{for all} \quad n \geq 1 \ .$$

Then $f(K)$ will serve as the desired compact set in S', because

$$P(f(X_n) \in f(K)) = P(X_n \in (f^{-1} \circ f)(K)) \geq P(X_n \in K) > 1 - \epsilon$$

for all $n \geq 1$. ∎

We next observe that on products of separable metric spaces tightness is characterized by tightness of the components; see §11.4 of [22].

Lemma 3.2 (tightness on product spaces) *Suppose that* $\{(X_{n,1}, \ldots, X_{n,k}) : n \geq 1\}$ *is a sequence of random elements of the product space* $S_1 \times \cdots \times S_k$, *where each coordinate space* S_i *is a separable metric space. The sequence* $\{(X_{n,1}, \ldots, X_{n,k}) : n \geq 1\}$ *is tight if and only if the sequence* $\{X_{n,i} : n \geq 1\}$ *is tight for each* i, $1 \leq i \leq k$.

Proof. The implication from the random vector to the components follows from Lemma 3.1 because the component $X_{n,i}$ is the image of the projection map $\pi_i : S_1 \times \cdots \times S_k \to S_i$ taking $(x_1, \ldots, x_k)$ into x_i, and the projection map is continuous. Going the other way, we use the fact that

$$A_1 \times \cdots \times A_k = \bigcap_{i=1}^{k} \pi_i^{-1}(A_i) = \bigcap_{i=1}^{k} \pi_i^{-1} \circ \pi_i(A_1 \times \cdots \times A_k)$$

for all subsets $A_i \subseteq S_i$. Thus, for each i and any $\epsilon > 0$, we can choose K_i such that $P(X_{n,i} \notin K_i) < \epsilon/k$ for all $n \geq 1$. We then let $K_1 \times \cdots \times K_k$ be the desired compact for the random vector. We have

$$
\begin{aligned}
P\left((X_{n,1}, \ldots, X_{n,k}) \notin K_1 \times \cdots \times K_k\right) &= P\left(\bigcup_{i=1}^{k}\{X_{n,i} \notin K_i\}\right) \\
&\leq \sum_{i=1}^{k} P\left(X_{n,i} \notin K_i\right) \leq \epsilon . \quad \blacksquare
\end{aligned}
$$

Tightness goes a long way toward establishing convergence because of Prohorov's theorem. It involves the notions of sequential compactness and relative compactness.

Definition 3.2 (relative compactness and sequential compactness) *A subset* A *of a metric space* S *is* **relatively compact** *if every sequence* $\{x_n : n \geq 1\}$ *from* A *has a subsequence that converges to a limit in* S *(which necessarily belongs to the closure* $\bar{A}$ *of* A*). A subset of* S *is* **sequentially compact** *if it is closed and relatively compact.*

We rely on the following basic result about compactness on metric spaces.

Lemma 3.3 (compactness coincides with sequential compactness on metric spaces) *A subset* A *of a metric space* S *is compact if and only if it is sequentially compact.*

We can now state Prohorov's theorem; see §11.6 of [22]. It relates compactness of sets of measures to compact subsets of the underlying sample space S on which the probability measures are defined.

Theorem 3.1 (Prohorov's theorem) *A subset of probability measures on a CSMS is tight if and only if it is relatively compact.*

We have the following elementary corollaries:

Corollary 3.1 (convergence implies tightness) *If $X_n \Rightarrow X$ as $n \to \infty$ for random elements of a CSMS, then the sequence $\{X_n : n \geq 1\}$ is tight.*

Corollary 3.2 (individual probability measures) *Every individual probability measure on a CSMS is tight.*

As a consequence of Prohorov's Theorem, we have the following method for establishing convergence of random elements:

Corollary 3.3 (convergence in distribution via tightness) *Let $\{X_n : n \geq 1\}$ be a sequence of random elements of a CSMS S. We have*

$$X_n \Rightarrow X \quad in \quad S \quad as \quad n \to \infty$$

if and only if (i) the sequence $\{X_n : n \geq 1\}$ is tight and (ii) the limit of every convergent subsequence of $\{X_n : n \geq 1\}$ is the same fixed random element X (has a common probability law).

In other words, once we have established tightness, it only remains to show that the limits of all converging subsequences must be the same. With tightness, we only need to uniquely determine the limit. When proving Donsker's theorem, it is natural to uniquely determine the limit through the finite-dimensional distributions. Convergence of all the finite-dimensional distributions is not enough to imply convergence on D, but it does uniquely determine the distribution of the limit; see pp 20 and 121 of Billingsley [3] and Example 11.6.1 in Whitt [22].

We will apply this approach to prove the martingale FCLT in this paper. In the martingale setting it is natural to use the martingale characterization of Brownian motion, originally established by Lévy [14] and proved by Ito's formula by Kunita and Watanabe [11]; see p. 156 of Karatzas and Shreve [9], and various extensions, such as to continuous processes with independent Gaussian increments, as in Theorem 1.1 on p. 338 of EK [6]. A thorough study of martingale characterizations appears in Chapter 4 of Liptser and Shiryayev [15] and in Chapters VIII and IX of JS [8].

3.2. Stochastic Boundedness

We now discuss stochastic boundedness because it plays a role in the tightness criteria in the next section. We start by defining stochastic boundedness and relating it to tightness. We then discuss situations in which stochastic boundedness is preserved. Afterwards, we give conditions for a sequence of martingales to be stochastically bounded in D involving the stochastic boundedness of appropriate sequences of $\mathbb{R}$-valued random variables.

3.2.1. Connection to Tightness

For random elements of $\mathbb{R}$ and $\mathbb{R}^k$, stochastic boundedness and tightness are equivalent, but tightness is stronger than stochastic boundedness for random elements of the functions spaces C and D (and the associated product spaces C^k and D^k).

Definition 3.3 (stochastic boundedness for random vectors) *A sequence $\{X_n : n \geq 1\}$ of random vectors taking values in $\mathbb{R}^k$ is **stochastically bounded (SB)** if the sequence is tight, as defined in Definition* 3.1.

The notions of tightness and stochastic boundedness thus agree for random elements of $\mathbb{R}^k$, but these notions differ for stochastic processes. For a function $x \in D^k \equiv D([0, \infty), \mathbb{R})^k$, let

$$\|x\|_T \equiv \sup_{0 \leq t \leq T} \{|x(t)|\} \ ,$$

where $|b|$ is a norm of $b \equiv (b_1, b_2, \ldots, b_k)$ in $\mathbb{R}^k$ inducing the Euclidean topology, such as the maximum norm: $|b| \equiv \max\{|b_1|, |b_2|, \ldots, |b_k|\}$. (Recall that all norms on Euclidean space $\mathbb{R}^k$ are equivalent.)

Definition 3.4 (stochastic boundedness for random elements of D^k) *A sequence $\{X_n : n \geq 1\}$ of random elements of D^k is **stochastically bounded in** D^k if the sequence of real-valued random variables $\{\|X_n\|_T : n \geq 1\}$ is stochastically bounded in $\mathbb{R}$ for each $T > 0$, using Definition* 3.3.

For random elements of D^k, tightness is a strictly stronger concept than stochastic boundedness. Tightness of $\{X_n\}$ in D^k implies stochastic boundedness, but not conversely; see §15 of Billingsely [3].

3.2.2. Preservation

We have the following analog of Lemma 3.2, which characterizes tightness for sequences of random vectors in terms of tightness of the associated sequences of components.

Lemma 3.4 (stochastic boundedness on D^k via components) *A sequence*

$$\{(X_{n,1}, \ldots, X_{n,k}) : n \geq 1\} \quad in \quad D^k \equiv D \times \cdots \times D$$

is stochastically bounded in D^k if and only if the sequence $\{X_{n,i} : n \geq 1\}$ is stochastically bounded in $D \equiv D^1$ for each i, $1 \leq i \leq k$.

Proof. Assume that we are using the maximum norm on product spaces. We can apply Lemma 3.2 after noticing that

$$\|(x_1, \ldots, x_k)\|_T = \max\{\|x_i\|_T : 1 \leq i \leq k\}$$

for each element $(x_1, \ldots, x_k)$ of D^k. Since other norms are equivalent, the result applies more generally. ∎

Lemma 3.5 (stochastic boundedness in D^k for sums) *Suppose that*

$$Y_n(t) \equiv X_{n,1}(t) + \cdots + X_{n,k}(t), \quad t \geq 0,$$

for each $n \geq 1$, where $\{(X_{n,1}, \ldots, X_{n,k}) : n \geq 1\}$ is a sequence of random elements of the product space $D^k \equiv D \times \cdots \times D$. If $\{X_{n,i} : n \geq 1\}$ is stochastically bounded in D for each i, $1 \leq i \leq k$, then the sequence $\{Y_n : n \geq 1\}$ is stochastically bounded in D.

Note that the converse is not true: We could have $k = 2$ with $X_{n,2}(t) = -X_{n,1}(t)$ for all n and t. In that case we have $Y_n(t) = 0$ for all $X_{n,1}(t)$.

3.2.3. Stochastic Boundedness for Martingales

We now provide ways to get stochastic boundedness for sequences of martingales in D from associated sequences of random variables. Our first result exploits the classical submartingale-maximum inequality; e.g., see p. 13 of Karatzas and Shreve [9]. We say that a function $f : \mathbb{R} \to \mathbb{R}$ is *even* if $f(-x) = f(x)$ for all $x \in \mathbb{R}$.

Lemma 3.6 (SB from the maximum inequality) *Suppose that, for each $n \geq 1$, $M_n \equiv \{M_n(t) : t \geq 0\}$ is a martingale (with respect to a specified filtration) with sample paths in D. Also suppose that, for each $T > 0$, there exists an even nonnegative convex function $f : \mathbb{R} \to \mathbb{R}$ with first derivative $f'(t) > 0$ for $t > 0$ (e.g., $f(t) \equiv t^2$), there exists a positive constant $K \equiv K(T, f)$, and there exists an integer $n_0 \equiv n_0(T, f, K)$, such that*

$$E[f(M_n(T))] \leq K \quad \text{for all} \quad n \geq n_0 .$$

Then the sequence of stochastic processes $\{M_n : n \geq 1\}$ is stochastically bounded in D.

Proof. Since any set of finitely many random elements of D is automatically tight, Theorem 1.3 of Billingsley [5], it suffices to consider $n \geq n_0$. Since f is continuous and $f'(t) > 0$ for $t > 0$, $t > c$ if and only if $f(t) > f(c)$ for $t > 0$. Since f is even,

$$E[f(M_n(t))] = E[f(|M_n(t)|)] \leq E[f(|M_n(T)|)] = E[f(M_n(T))] \leq K$$

for all t, $0 \leq t \leq T$. Since these moments are finite and f is convex, the stochastic process $\{f(M_n(t)) : 0 \leq t \leq T\}$ is a submartingale for each $n \geq 1$, so that we can apply the submartingale-maximum inequality to get

$$P(\|M_n\|_T > c) = P(\|f \circ M_n\|_T > f(c)) \leq \frac{E[f(M_n(T))]}{f(c)} \leq \frac{K}{f(c)}$$

for all $n \geq n_0$. Since $f(c) \to \infty$ as $c \to \infty$, we have the desired conclusion. ∎

We now establish another sufficient condition for stochastic boundedness of square-integrable martingales by applying the Lenglart-Rebolledo inequality; see p. 66 of Liptser and Shiryayev [15] or p. 30 of Karatzas and Shreve [9].

Lemma 3.7 (Lenglart-Rebolledo inequality) *Suppose that $M \equiv \{M(t) : t \geq 0\}$ is a square-integrable martingale (with respect to a specified filtration) with predictable quadratic variation $\langle M \rangle \equiv \{\langle M \rangle(t) : t \geq 0\}$, i.e., such that $M^2 - \langle M \rangle \equiv \{M(t)^2 - \langle M \rangle(t) : t \geq 0\}$ is a martingale by the Doob-Meyer decomposition. Then, for all $c > 0$ and $d > 0$,*

$$P \left(\sup_{0 \leq t \leq T} \{|M(t)|\} > c \right) \leq \frac{d}{c^2} + P \left(\langle M \rangle(T) > d \right) . \tag{11}$$

As a consequence we have the following criterion for stochastic boundedness of a sequence of square-integrable martingales.

Lemma 3.8 (SB criterion for square-integrable martingales) *Suppose that, for each $n \geq 1$, $M_n \equiv \{M_n(t) : t \geq 0\}$ is a square-integrable martingale (with respect to a specified filtration) with predictable quadratic variation $\langle M_n \rangle \equiv \{\langle M_n \rangle(t) : t \geq 0\}$, i.e., such that $M_n^2 - \langle M_n \rangle \equiv \{M_n(t)^2 - \langle M_n \rangle(t) : t \geq 0\}$ is a martingale by the Doob-Meyer decomposition. If the sequence of random variables $\{\langle M_n \rangle(T) : n \geq 1\}$ is stochastically bounded in $\mathbb{R}$ for each $T > 0$, then the sequence of stochastic processes $\{M_n : n \geq 1\}$ is stochastically bounded in D.*

Proof. For $\epsilon > 0$ given, apply the assumed stochastic boundedness of the sequence $\{\langle M_n \rangle(T) : n \geq 1\}$ to obtain a constant d such that

$$P \left(\langle M_n \rangle(T) > d \right) < \epsilon/2 \quad \text{for all} \quad n \geq 1 .$$

Then for that determined d, choose c such that $d/c^2 < \epsilon/2$. By the Lenglart-Rebolledo inequality (11), these two inequalities imply that

$$P \left(\sup_{0 \leq t \leq T} \{|M_n(t)|\} > c \right) < \epsilon . \quad \blacksquare \tag{12}$$

3.3. Tightness Criteria

The standard characterization for tightness of a sequence of stochastic processes in D, originally developed by Skorohod [20] and presented in Billingsley [3], involves suprema. Since then, more elementary one-dimensional criteria have been developed; see Billingsley [4], Kurtz [12], Aldous [1, 2], Jacod et al. [7], §§3.8 and 3.9 of EK [6] and §16 of Billingsley [5]. Since these simplifications help in proving the martingale FCLT, we will present some of the results here.

3.3.1. Criteria Involving a Modulus of Continuity

We start by presenting the classical characterization of tightness, as in Theorems 13.2 and 16.8 of Billingsley [5]. For that purpose we define functions w and w' that can serve as a modulus of continuity. For any $x \in D$ and subset A of $[0, \infty)$, let

$$w(x, A) \equiv \sup \{|x(t_1) - x(t_2)| : t_1, t_2 \in A\} . \tag{13}$$

For any $x \in D$, $T > 0$ and $\delta > 0$, let

$$w(x, \delta, T) \equiv \sup \{w(x, [t_1, t_2]) : 0 \leq t_1 < t_2 \leq (t_1 + \delta) \wedge T\} \qquad (14)$$

and

$$w'(x, \delta, T) \equiv \inf_{\{t_i\}} \max_{1 \leq i \leq k} \{w(x, [t_{i-1}, t_i))\} \,, \qquad (15)$$

where the infimum in (15) is over all k and all subsets of $[0, T]$ of size $k + 1$ such that

$$0 = t_0 < t_1 < \cdots < t_k = T \quad \text{with} \quad t_i - t_{i-1} > \delta \quad \text{for} \quad 1 \leq i \leq k - 1 \,.$$

(We do not require that $t_k - t_{k-1} > \delta$.)

The following is a variant of the classical characterization of tightness; see Theorem 16.8 of Billingsley [5].

Theorem 3.2 (classical characterization of tightness) *A sequence of stochastic processes $\{X_n : n \geq 1\}$ in D is tight if and only if*

(i) *The sequence $\{X_n : n \geq 1\}$ is stochastically bounded in D*
and

(ii) *for each $T > 0$ and $\epsilon > 0$,*

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} P(w'(X_n, \delta, T) > \epsilon) = 0 \,.$$

If the modulus $w(X_n, \delta, T)$ is substituted for $w'(X_n, \delta, T)$ in Condition (ii) of Theorem 3.2, then the sequence $\{X_n\}$ is said to be **C-tight**, because then the sequence is again tight but the limit of any convergent subsequence must have continuous sample paths; see Theorem 15.5 of Billingsley [3]. With this modified condition (ii), condition (i) is implied by having the sequence $\{X_n(0)\}$ be tight in $\mathbb{R}$.

Conditions (i) and (ii) in Theorem 3.2 are both somewhat hard to verify because they involve suprema. We have shown how the stochastic-boundedness condition (i) can be made one-dimensional for martingales via Lemmas 3.6 and 3.8. The following is another result, which exploits the modulus condition (ii) in Theorem 3.2; see p. 175 of Billingsley [5]. To state it, let J be the maximum-jump function defined in (1).

Lemma 3.9 (substitutes for stochastic boundedness) *In the presence of the modulus condition (ii) in Theorem 3.2, each of the following is equivalent to the stochastic-boundedness condition (i) in Theorem 3.2:*

(i) *The sequence $\{X_n(t) : n \geq 1\}$ is stochastically bounded in $\mathbb{R}$ for each t in a dense subset of $[0, \infty)$.*

(ii) *The sequence $\{X_n(0) : n \geq 1\}$ is stochastically bounded in $\mathbb{R}$ and, for each $T > 0$, the sequence $\{J(X_n, T) : n \geq 1\}$ is stochastically bounded in $\mathbb{R}$.*

We also mention the role of the maximum-jump function J in (1) in characterizing continuous limits; see Theorem 13.4 of Billingsley [5]:

Lemma 3.10 (identifying continuous limits) *Suppose that $X_n \Rightarrow X$ in D. Then $P(X \in C) = 1$ if and only if $J(X_n, T) \Rightarrow 0$ in $\mathbb{R}$ for each $T > 0$.*

3.3.2. Simplifying the Modulus Condition

Simplifying the modulus condition (ii) in Theorem 3.2 is even of greater interest. We first present results from EK. Conditions (i) and (iii) below are of particular interest because they shows that, for each n and t, we need only consider the process at the two single time points $t + u$ and $t - v$ for $u > 0$ and $v > 0$. As we will see in Lemma 3.11 below, we can obtain a useful sufficient condition involving only the single time point $t + u$ (ignoring $t - v$).

For the results in this section, we assume that the strong stochastic-boundedness condition (i) in Theorem 3.2 is in force. For some of the alternatives to the modulus condition (ii) in Theorem 3.2 it is also possible to simplify the stochastic-boundedness condition, as in Lemma 3.9, but we do not carefully examine that issue.

Theorem 3.3 (substitutes for the modulus condition) *In the presence of the stochastic-boundedness condition (i) in Theorem* 3.2, **each of the following is equivalent to the modulus condition** (ii):

(i) *For each $T > 0$, there exists a constant $\beta > 0$ and a family of nonnegative random variables $\{Z_n(\delta, T) : n \geq 1, \delta > 0\}$ such that*

$$E\left[(1 \wedge |X_n(t + u) - X_n(t)|)^\beta \mid \mathcal{F}_{n,t}\right]\left((1 \wedge |X_n(t) - X_n(t - v)|)^\beta\right)$$
$$\leq E\left[Z_n(\delta, T)|\mathcal{F}_{n,t}\right] \quad w.p.1 \tag{16}$$

for $0 \leq t \leq T$, $0 \leq u \leq \delta$ and $0 \leq v \leq t \wedge \delta$, where $\mathcal{F}_{n,t}$ is the σ-field in the internal filtration of $\{X_n(t) : t \geq 0\}$,

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[Z_n(\delta, T)] = 0 . \tag{17}$$

and

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[(1 \wedge |X_n(\delta) - X_n(0)|)^\beta] = 0 . \tag{18}$$

(ii) *The sequence of stochastic processes $\{\{f(X_n(t)) : t \geq 0\} : n \geq 1\}$ is tight in D for each function $f : \mathbb{R} \to \mathbb{R}$ in a dense subset of all bounded continuous functions in the topology of uniform convergence over bounded intervals.*

(iii) *For each function $f : \mathbb{R} \to \mathbb{R}$ in a dense subset of all bounded continuous functions in the topology of uniform convergence over bounded intervals, and $T > 0$, there exists a constant $\beta > 0$ and a family of nonnegative random variables $\{Z_n(\delta, f, T) : n \geq 1, \delta > 0\}$ such that*

$$E\left[|f(X_n(t + u)) - f(X_n(t))|^\beta \mid \mathcal{F}_{n,f,t}\right]\left(|f(X_n(t)) - f(X_n(t - v))|^\beta\right)$$
$$\leq E\left[Z_n(\delta, f, T)|\mathcal{F}_{n,f,t}\right] \quad w.p.1 \tag{19}$$

for $0 \leq t \leq T$, $0 \leq u \leq \delta$ and $0 \leq v \leq t \wedge \delta$, where $\mathcal{F}_{n,f,t}$ is the σ-field in the internal filtration of $\{f(X_n(t)) : t \geq 0\}$, and

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[Z_n(\delta, f, T)] = 0 . \tag{20}$$

and

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[|f(X_n(\delta)) - f(X_n(0))|^\beta] = 0 . \tag{21}$$

Proof. Condition (i) is condition (b) in Theorem 3.8.6 of EK for the case of real-valued stochastic processes. Condition (ii) is Theorem 3.9.1 of EK. Condition (iii) is a modification of condition (b) in Theorem 3.8.6 of EK, exploiting condition (ii) and the fact that our stochastic processes are real valued. Since the functions f are bounded in (iii), there is no need to replace the usual metric $m(c,d) \equiv |c - d|$ with the bounded metric $q \equiv m \wedge 1$. $\blacksquare$

In applications of Theorem 3.3, it is advantageous to replace Conditions (i) and (iii) in Theorem 3.3 with stronger sufficient conditions, which only involves the conditional expectation on the left. By doing so, we need to consider only conditional distributions of the processes at a single time point $t + u$ with $u > 0$, given the relevant history $\mathcal{F}_{n,t}$ up to time time t. We need to do an estimate for all t and n, but given specific values of t and n, we need to consider only one future time point $t + u$ for $u > 0$. **These simplified conditions are sufficient, but not necessary, for D-tightness**. On the other hand, they are not sufficient for C-tightness; see Remark 3.1 below.

In addition, for condition (iii) in Theorem 3.3, it suffices to specify a specific dense family of functions. As in the Ethier-Kurtz proof of the martingale FCLT (their Theorem 7.1.4), we introduce smoother functions in order to exploit Taylor series expansions. Indeed, we present the versions of Theorem 3.3 actually applied by EK in their proof of their Theorem 7.1.4.

Lemma 3.11 (simple sufficient criterion for tightness) *The following are sufficent, but not necessary, conditions for a sequence of stochastic processes $\{X_n : n \geq 1\}$ in D to be tight:*

(i) The sequence $\{X_n : n \geq 1\}$ is stochastically bounded in D

and either

(ii.a) For each $n \geq 1$, the stochastic process X_n is adapted to a filtration $\boldsymbol{F}_n \equiv \{\mathcal{F}_{n,t} : t \geq 0\}$. In addition, for each function $f : \mathbb{R} \to \mathbb{R}$ belonging to $\mathcal{C}_c^\infty$ (having compact support and derivatives of all orders) and $T > 0$, there exists a family of nonnegative random variables $\{Z_n(\delta, f, T) : n \geq 1, \delta > 0\}$ such that

$$|E\left[f(X_n(t+u)) - f(X_n(t)) \mid \mathcal{F}_{n,t}\right]| \leq E\left[Z_n(\delta, f, T) \mid \mathcal{F}_{n,t}\right] \tag{22}$$

w.p.1 for $0 \leq t \leq T$ and $0 \leq u \leq \delta$ and

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[Z_n(\delta, f, T)] = 0 \ . \tag{23}$$

or

(ii.b) For each $n \geq 1$, the stochastic process X_n is adapted to a filtration $\boldsymbol{F}_n \equiv \{\mathcal{F}_{n,t} : t \geq 0\}$. In addition, for each $T > 0$, there exists a family of nonnegative random variables $\{Z_n(\delta, T) : n \geq 1, \delta > 0\}$ such that

$$\left|E\left[(X_n(t+u) - X_n(t))^2 \mid \mathcal{F}_{n,t}\right]\right| \leq E\left[Z_n(\delta, T) \mid \mathcal{F}_{n,t}\right] \tag{24}$$

w.p.1 for $0 \leq t \leq T$ and $0 \leq u \leq \delta$ and

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[Z_n(\delta, T)] = 0 \ . \tag{25}$$

Proof. We apply Theorem 3.3. For condition (ii.a), we apply Theorem 3.3 (iii). We have specified a natural class of functions f that are dense in the set of all continuous bounded real-valued functions with the topology of uniform convergence over bounded intervals. We have changed the filtration. Theorem 3.3 and Theorem 3.8.6 of EK specify the internal filtration of the process, which here would be the process $\{f(X_n(t)) : t \geq 0\}$, but it is convenient to work with the more refined filtration associated with X_n. By taking conditional expected values, conditional on the coarser filtration generated by $\{f(X_n(t)) : t \geq 0\}$, we can deduce the corresponding inequality conditioned on the coarser filtration. The conditions here follow from condition (16) by taking $\beta = 2$, because

$$E[|f(X_n(t+u)) - f(X_n(t))|^2|\mathcal{F}_{n,t}] = E[f(X_n(t+u))^2 - f(X_n(t))^2|\mathcal{F}_{n,t}]$$
$$-2f(X_n(t))E[f(X_n(t+u)) - f(X_n(t))|\mathcal{F}_{n,t}] ,$$

(see (1.35) on p. 343 of EK), so that

$$E[|f(X_n(t+u)) - f(X_n(t))|^2|\mathcal{F}_{n,t}] \leq E[Z_n(\delta, f^2, T)|\mathcal{F}_{n,t}] + 2\|f\|E[Z_n(\delta, f, T)|\mathcal{F}_{n,t}]$$

under condition (22). Note that f^2 belongs to our class of functions for every function f in the class. Finally, note that under (22)

$$E[|f(X_n(\delta)) - f(X_n(0))|^2] \leq E[Z_n(\delta, f, T)] , \tag{26}$$

so that the limit (21) holds by (23).

Now consider condition (ii.b). Note this is a direct consequence of Theorem 3.3 (i) using $\beta = 2$. Condition (26) is made stronger than (16) by removing the $\wedge 1$. As in (26),

$$E[|X_n(\delta) - X_n(0)|^2] \leq E[Z_n(\delta, T)] , \tag{27}$$

so that the limit (18) holds by (25). ∎

We have mentioned that Theorem 3.3 and Lemma 3.11 come from EK [6]. As indicated there, these in turn come from Kurtz [12]. However, Lemma 3.11 is closely related to a sufficient condition for tightness proposed by Billingsley [4]. Like Lemma 3.11, this sufficient condition is especially convenient for treating Markov processes. We state a version of that here. For that purpose, let $\alpha_n(\lambda, \epsilon, \delta, T)$ be a number such that

$$P(|X_n(u) - X_n(t_m)| > \epsilon|X_n(t_1), \cdots, X_n(t_m)) \leq \alpha_n(\lambda, \epsilon, \delta, T) \tag{28}$$

holds with probability 1 on the set $\{\max_i\{|X_n(t_i)| \leq \lambda\}$ for all m and all m time points t_i with

$$0 \leq t_1 \leq \cdots \leq t_{m-1} \leq t_m < u \leq T \quad \text{and} \quad u - t_m \leq \delta . \tag{29}$$

The following is a variant of Theorem 1 of Billingsley [4]. (He has the stochastic-boundedness condition (i) below replaced by a weaker condition.)

Lemma 3.12 (another simple sufficient condition for tightness) *Alternative sufficent, but not necessary, conditions for a sequence of stochastic processes $\{X_n : n \geq 1\}$ in D to be tight are the following:*

(i) The sequence $\{X_n : n \geq 1\}$ is stochastically bounded in D

and

(ii) For each $T > 0$, $\epsilon > 0$ and $\lambda < \infty$,

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \alpha_n(\lambda, \epsilon, \delta, T) = 0 \ , \tag{30}$$

for $\alpha_n(\lambda, \epsilon, \delta, T)$ defined in (28).

Remark 3.1 (*sufficient for D-tightness but not C-tightness.*) The sufficient conditions in Lemmas 3.11 and 3.12 are sufficient but not necessary for D tightness. On the other hand, these conditions are *not* sufficient for C-tightness. To substantiate these claims, it suffices to consider simple examples. The single function $x(t) = 1_{[1,\infty)}(t)$, $t \geq 0$, is necessarily tight in D, but the conditions in Lemmas 3.11 and 3.12 are not satisfied for it. On the other hand, the stochastic process $X(t) = 1_{[T,\infty)}(t)$, $t \geq 0$, where T is an exponential random variable with mean 1, is D tight, but not C-tight. By the lack of memory property of the exponential distribution, the conditions of Lemmas 3.11 and 3.12 are satisfied for this simple random element of D. ■

3.3.3. Stopping Times and Quadratic Variations

We conclude this section by mentioning alternative criteria for tightness specifically intended for martingales. These criteria involve stopping times and the quadratic-variation processes. The criteria in terms of stopping times started with Aldous [1] and Rebolledo [19]. Equivalence with the conditions in Theorem 3.3 and Lemma 3.11 was shown in Theorems 2.7 of Kurtz [13] and 3.8.6 of EK [6]; See also p 176 of Billingsley [5]. These criteria are very natural for proving tightness of martingales, as Aldous [1, 2] and Rebolledo [19] originally showed. Aldous' [1] proof of (a generalization of) the martingale FCLT from McLeish [16] is especially nice.

Theorem 3.4 (another substitute for the modulus condition) *Suppose that the stochastic-boundedness condition (i) in Theorem 3.2 holds.*

The following is equivalent to the modulus condition in Theorem 3.2, condition (ii) in Theorem 3.3 and thus to tightness:

For each $T > 0$, there exists a constant $\beta > 0$ such that

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} C_n(\delta, T) = 0 \ , \tag{31}$$

where $C_n(\delta, T)$ is the supremum of

$$E\left[(1 \wedge |X_n(\tau + u) - X_n(\tau)|)^\beta \left((1 \wedge |X_n(\tau) - X_n(\tau - v)|)^\beta \right) \right]$$

over $0 \leq u \leq \delta$, $\tau \geq v$ and $\tau \in \mathcal{S}_{n,T}$, with $\mathcal{S}_{n,T}$ being the collection of all finite-valued stopping times with respect to the internal filtration of X_n, bounded by T.

Theorem 3.5 (another substitute for the sufficient condition) *Suppose that the stochastic-boundedness condition* (*i*) *in Theorem 3.2 holds.*

Then the following are equivalent sufficient, but not necessary, conditions for the modulus condition (*ii*) *in Theorem 3.2, condition* (*ii*) *in Theorem 3.3 and thus for tightness:*

(*i*) *For each $T > 0$ and sequence $\{(\tau_n, \delta_n) : n \geq 1\}$, where τ_n is a finite-valued stopping time with respect to the internal filtration generated by X_n, with $\tau_n \leq T$ and δ_n is a positive constant with $\delta_n \downarrow 0$,*

$$X_n(\tau_n + \delta_n) - X_n(\tau_n) \Rightarrow 0 \quad as \quad n \to \infty . \tag{32}$$

(*ii*) *For each $\epsilon > 0$, $\eta > 0$ and $T > 0$, there exist $\delta > 0$, n_0 and two sequences $\{\tau_{n,i} : n \geq n_0\}$, $i = 1, 2$, where $\tau_{n,i}$ is a stopping time with respect to the internal filtration generated by X_n such that $0 \leq \tau_{n,1} < \tau_{n,2} \leq T$ and*

$$P\left(X_n(\tau_{n,2}) - X_n(\tau_{n,1}) > \epsilon, \quad \tau_{n,2} - \tau_{n,1} \leq \delta\right) < \eta \quad for \ all \quad n \geq n_0 . \tag{33}$$

Aldous (1978) showed that the two conditions in Lemma 3.12 imply condition (32), because

$$P(X_n(\tau_n + \delta_n) - X_n(\tau_n) > \epsilon) \leq \alpha_n(\lambda, \epsilon, \delta_n, T) + P(\|X_n\|_T > \lambda) . \tag{34}$$

Related criteria for tightness in terms of quadratic variation processes are presented in Rebolledo [19], Jacod et al. [7] and §§VI.4-5 of JS [8]. The following is a minor variant of Theorem VI.4.13 on p 322 of [8]. Lemma 3.2 shows that C-tightness of the sum $\sum_{i=1}^{k} \langle X_{n,i}, X_{n,i} \rangle$ is equivalent to tightness of the components.

Theorem 3.6 (tightness criterion in terms of the angle-bracket process) *Suppose that $X_n \equiv (X_{n,1}, \ldots, X_{n,1})$ is a locally-square-integrable martingale in D^k for each $n \geq 1$, so that the predictable quadratic covariation processes $\langle X_{n,i}, X_{n,j} \rangle$ are well defined. The sequence $\{X_n : n \geq 1\}$ is C-tight if*

(*i*) *The sequence $\{X_{n,i}(0) : n \geq 1\}$ is tight in $\mathbb{R}$ for each i, and*

(*ii*) *the sequence $\langle X_{n,i}, X_{n,i} \rangle$ is C-tight for each i.*

The following is a minor variant of Lemma 11 of Rebolledo [19]. We again apply Lemma 3.2 to treat the vector case. Since a local martingale with bounded jumps is locally square integrable, the hypothesis of the next theorem is not weaker than the hypothesis of the previous theorem.

Theorem 3.7 (tightness criterion in terms of the square-bracket processes) *Suppose that $X_n \equiv (X_{n,1}, \ldots, X_{n,1})$ is a local martingale in D^k for each $n \geq 1$. The sequence $\{X_n : n \geq 1\}$ is C-tight if*

(*i*) *For all $T > 0$, there exists n_0 such that*

$$J(X_n, T) \equiv \|\Delta X_{n,i}\|_T \equiv \sup_{0 \leq t \leq T} \{|X_n(t) - X_n(t-)|\} \leq b_n \tag{35}$$

for all i and $n \geq n_0$, where b_n are real numbers such that $b_n \to 0$ as $n \to \infty$; and

(ii) The sequence of optional quadratic variation processes $[X_{n,i}]$ is C-tight for each i.

In closing we remark that these quadratic-variation tightness conditions in Theorems 3.6 and 3.7 are often easy to verify because these are nondecreasing processes; see Lemma 2.1.

4. Proofs of Tightness

We now begin the proof of the martingale FCLT in Theorem 2.1. In this section we do the tightness part; in the next section we do the characterization part.

4.1. Quick Proofs of C-Tightness

We can give very quick proofs of C-tightness in both cases, exploiting Theorems 3.6 and 3.7. In case (i) we need to add an extra assumption. We need to assume that the jumps are uniformly bounded and that the bound is asymptotically negligible. In particular, we need to go beyond conditions (2) and assume that, for all $T > 0$, there exists n_0 such that

$$J(M_{n,i}, T) \leq b_n \quad \text{for all} \quad i \quad \text{and} \quad n \geq n_0 \, , \tag{36}$$

where J is the maximum-jump function in (1) and b_n are real numbers such that $b_n \to 0$ as $n \to \infty$, as in Theorem 3.7. We remark that this extra condition is satisfied for many queueing applications, as in Pang et al. [17], because the jumps are at most of size 1 before scaling, so that they become at most of size $1/\sqrt{n}$ after scaling.

Case (i) with bounded jumps. Suppose that the jumps of the martingale are uniformly bounded and that the bound is asymptotically negligible, as in (36). We thus can apply Theorem 3.7. By condition (3) and Lemma 2.1, we have convergence of the optional quadratic-variation (square-bracket) processes: $[M_{n,i}, M_{n,j}] \Rightarrow c_{i,j}e$ in D for each i and j. Thus these sequences of processes are necessarily C-tight. Hence Theorem 3.7 implies that the sequence of martingales $\{M_{n,i} : n \geq 1\}$ is C-tight in D for each i. Lemma 3.2 then implies that $\{M_n : n \geq 1\}$ is C-tight in D^k. ∎

Case (ii). We need no extra condition in case (ii). We can apply Theorem 3.6. Condition (6) and Lemma 2.1 imply that there is convergence of the predictable quadratic-variation (angle-bracket) processes: $\langle M_{n,i}, M_{n,j} \rangle \Rightarrow c_{i,j}e$ in D for each i and j. Thus these sequences of processes are necessarily C-tight, so condition (ii) of Theorem 3.6 is satisfied. Condition (i) is satisfied too, since $M_{n,i}(0) = 0$ for all i by assumption. ∎

4.2. The Ethier-Kurtz Proof in Case (i).

We now give a longer proof for case (i) without making the extra assumption in (36). We follow the proof on pp 341-343 of EK [6], elaborating on several points.

First, if the stochastic processes M_n are local martingales rather than martingales, then we introduce the available stopping times τ_n with $\tau_n \uparrow \infty$ w.p.1 so that $\{M_n(\tau_n \wedge t) : t \geq 0\}$ are martingales for each n. Hence we simply assume that the processes M_n are martingales.

We then exploit the assumed limit $[M_{n,i}, M_{n,j}](t) \Rightarrow c_{i,j}t$ in $\mathbb{R}$ as $n \to \infty$ in (3) in order to act as if $[M_{n,i}, M_{n,j}](t)$ is bounded. In particular, as in (1.23) on p. 341 of EK, we introduce the stopping times

$$\eta_n \equiv \inf\{t \geq 0 : [M_{n,i}, M_{n,i}](t) > c_{i,i}t + 1 \quad \text{for some} \quad i\} . \tag{37}$$

As asserted in EK, by (3), $\eta_n \Rightarrow \infty$, but we provide additional detail about the supporting argument: If we only had the convergence in distribution for each t, it would not follow that $\eta_n \Rightarrow \infty$. For this step, it is critical that we can strengthen the mode of convergence to convergence in distribution in D, where the topology on the underlying space corresponds to uniform convergence over bounded intervals. To do so, we use the fact that the quadratic variation processes $[M_{n,i}, M_{n,i}]$ are monotone and the limit is continuous. In particular, we apply Lemma 2.1. But, indeed, $\eta_n \Rightarrow \infty$ as claimed.

Hence, it suffices to focus on the martingales $\tilde{M}_n \equiv \{M_n(\eta_n \wedge t) : t \geq 0\}$. We then reduce the k initial dimensions to 1 by considering the martingales

$$Y_n \equiv \sum_{i=1}^{k} \theta_i \tilde{M}_{n,i} \tag{38}$$

for an arbitrary non-null vector $\theta \equiv (\theta_1, \ldots, \theta_k)$. The associated optional quadratic variation process is

$$A_{n,\theta}(t) \equiv [Y_n] = \sum_{i=1}^{k}\sum_{j=1}^{k} \theta_i \theta_j \left[\tilde{M}_{n,i}, \tilde{M}_{n,j}\right](t), \quad t \geq 0 . \tag{39}$$

From condition (3) and Lemma 2.1, $A_{n,\theta} \Rightarrow c_\theta e$ in D, where e is the identity function and

$$c_\theta \equiv \sum_{i=1}^{k}\sum_{j=1}^{k} \theta_i \theta_j c_{i,j} . \tag{40}$$

Some additional commentary may help here. The topology on the space $D([0, \infty), \mathbb{R}^k)$ where the functions take values in $\mathbb{R}^k$ is strictly stronger than the topology on the product space $D([0, \infty), \mathbb{R})^k$, but there is no difference on the subset of continuous functions, and so this issue really plays no role here. We can apply Lemma 3.10 and condition (2) to conclude that a limit of any convergent subsequence must have continuous paths. In their use of Y_n defined in (38), EK work to establish the stronger tightness in $D([0, \infty), \mathbb{R}^k)$; sufficiency

is covered by problem 3.22 on p. 153 of EK. On the other hand, for the product topology it would suffice to apply Lemma 3.2, but the form of Y_n in (38) covers that as well. The form of Y_n in (38) is also convenient for characterizing the limiting process, as we will see.

The idea now is to establish the tightness of the sequence of martingales $\{Y_n : n \geq 1\}$ in (38). The approach in EK is to introduce smooth bounded functions $f : \mathbb{R} \to \mathbb{R}$ and consider the associated sequence $\{f(Y_n(t)) : n \geq 1\}$. This step underlies the entire EK book, and so is to be expected. To quickly see the fundamental role of this step in EK, see Chapters 1 and 4 in EK, e.g., the definition of a full generator in (1.5.5) on p 24 and the associated martingale property in Proposition 4.1.7 on p 162. This step is closely related to Ito's formula, as we explain in §5.2 below. It is naturally associated with the Markov-process-centric approach in EK [6] and Stroock and Varadhan [21], as opposed to the martingale-centric approach in Aldous [1, 2], Rebolledo [19] and JS [8].

Consistent with that general strategy, Ethier and Kurtz exploit tightness criteria involving such smooth bounded functions; e.g., see Theorem 3.3 and Lemma 3.11 here. Having introduced those smooth bounded functions, we establish tightness by applying Lemma 3.11 (ii.a). We will exploit the bounds provided by (37). We let the filtrations $\mathbf{F}_n \equiv \{\mathcal{F}_{n,t} : t \geq 0\}$ be the internal filtrations of Y_n. First, the submartingale-maximum inequality in Lemma 3.6 will be used to establish the required stochastic boundedness; see (1.34) on p. 343 of EK. It will then remain to establish the moment bounds in (22).

To carry out both these steps, we fix a function f in $\mathcal{C}_c^\infty$, the set of infinitely differentiable functions $f : \mathbb{R} \to \mathbb{R}$ with compact support (so that the function f and all derivatives are bounded). The strategy is to apply Taylor's theorem to establish the desired bounds. With that in mind, as in (1.27) on p. 341 of EK, we write

$$E\left[f(Y_n(t+s)) - f(Y_n(t)) \mid \mathcal{F}_{n,t}\right]$$
$$= E\left[\sum_{i=0}^{m-1} \left(f(Y_n(t_{i+1})) - f(Y_n(t_i)) - f'(Y_n(t_i))\xi_{n,i}\right) \mid \mathcal{F}_{n,t}\right] , \quad (41)$$

where $0 \leq t = t_0 < t_1 < \cdots < t_m = t + s$ and $\xi_{n,i} = Y_n(t_{i+1}) - Y_n(t_i)$. Formula (41) is justified because $E[\xi_{n,i} | \mathcal{F}_{n,t}] = 0$ for all i since Y_n is an $\mathbf{F_n}$-martingale. Also, there is cancellation in the first two terms in the sum.

Following (1.28)–(1.30) in EK, we write

$$\gamma_n \equiv \max\{j : t_j < \eta_n \wedge (t+s)\} \tag{42}$$

and

$$\zeta_n \equiv \tag{43}$$
$$\max\left\{j : t_j < \eta_n \wedge (t+s), \sum_{i=0}^{j} \xi_{n,i}^2 \leq k \sum_{i=1}^{k} c_{i,i}\theta_i^2(t+s) + 2k \sum_{i=1}^{k} \theta_i^2\right\} .$$

By the definition of η_n, we will have $\gamma_n = \zeta_n$ when there are sufficiently many points so that the largest interval $t_{i+1} - t_i$ is suitably small and n is sufficiently

large. Some additional explanation might help here: To get the upper bound in (43), we exploit the definition of the optional quadratic covariation (equation (24) in Pang et al. [17]) and the simple relation $(a + b)^2 \geq 0$, implying that $|2ab| \leq a^2 + b^2$ for real numbers a and b. Thus

$$2|[M_{n,i}, M_{n,j}](t)\theta_i\theta_j| \leq [M_{n,i}, M_{n,i}](t)\theta_i^2 + [M_{n,j}, M_{n,j}](t)\theta_j^2 \qquad (44)$$

for each i and j, so that

$$|A_{n,\theta}(t)| = |\sum_{i=1}^{k}\sum_{j=1}^{k}\theta_i\theta_j \left[\tilde{M}_{n,i}, \tilde{M}_{n,j}\right](t)| \leq k \sum_{i=1}^{k}\left[\tilde{M}_{n,i}, \tilde{M}_{n,i}\right](t)\theta_i^2 . \qquad (45)$$

In turn, by (37), inequality (45) implies that

$$|A_{n,\theta}(t)| \leq k \sum_{i=1}^{k}(c_{i,i}t + 1)\theta_i^2 = k \sum_{i=1}^{k}c_{i,i}t\theta_i^2 + k \sum_{i=1}^{k}\theta_i^2 . \qquad (46)$$

In definition (43) we increase the target by replacing the constant 1 by 2 in the last term, so that eventually we will have $\zeta_n = \gamma_n$, as claimed.

As in (1.30) on p. 341 of EK, we next extend the expression (41) to include second derivative terms, simply by adding and subtracting. In particular, we have

$$E\left[f(Y_n(t + s)) - f(Y_n(t)) \mid \mathcal{F}_{n,t}\right]$$

$$= E\left[\sum_{i=\zeta_n}^{\gamma_n}(f(Y_n(t_{i+1})) - f(Y_n(t_i)) - f'(Y_n(t_i))\xi_{n,i}) \mid \mathcal{F}_{n,t}\right]$$

$$+ E\left[\sum_{i=0}^{\zeta_n-1}\left(f(Y_n(t_{i+1})) - f(Y_n(t_i)) - f'(Y_n(t_i))\xi_{n,i} - \frac{1}{2}f''(Y_n(t_i))\xi_{n,i}^2\right) \mid \mathcal{F}_{n,t}\right]$$

$$+ E\left[\sum_{i=0}^{\zeta_n-1}\frac{1}{2}f''(Y_n(t_i))\xi_{n,i}^2 \mid \mathcal{F}_{n,t}\right] . \qquad (47)$$

Following p. 342 of EK, we set $\Delta Y_n(u) \equiv Y_n(u) - Y_n(u-)$ and introduce more time points so that the largest difference $t_{i+1} - t_i$ converges to 0. In the limit as we add more time points, we obtain the representation

$$E\left[f(Y_n(t + s)) - f(Y_n(t)) \mid \mathcal{F}_{n,t}\right] = W_{n,1}(t, t + s) + W_{n,2}(t, t + s)$$
$$+ W_{n,3}(t, t + s) , \qquad (48)$$

where

$$W_{n,1}(t, t + s) \equiv E\left[f(Y_n((t + s) \wedge \eta_n)) - f(Y_n((t + s) \wedge \eta_n)-)\right.$$
$$\left. - f'(Y_n((t + s) \wedge \eta_n)-)\Delta Y_n((t + s) \wedge \eta_n)) \mid \mathcal{F}_{n,t}\right]$$

$$W_{n,2}(t, t+s) \equiv E\left[\sum_{t \wedge \eta_n < u < (t+s) \wedge \eta_n} (f(Y_n(u)) - f(Y_n(u-)) - f'(Y_n(u-))\Delta Y_n(u)\right.$$

$$\left. - f'(Y_n(u-))\Delta Y_n(u) - \frac{1}{2}f''(Y_n(u-))(\Delta Y_n(u))^2 \right) | \mathcal{F}_{n,t}\right]$$

$$W_{n,3}(t, t+s) \equiv E\left[\int_{t \wedge \eta_n}^{(t+s) \wedge \eta_n} \frac{1}{2}f''(Y_n(u-))\, dA_{n,\theta}(u) | \mathcal{F}_{n,t}\right] . \tag{49}$$

(We remark that we have inserted $t \wedge \eta_n$ in two places in (1.31) on p 342 of EK. We can write the sum for $W_{n,2}$ that way because Y_n has at most countably many discontinuities.)

Now we can bound the three terms in (48) and (49). As in (1.32) on p. 342 of EK, we can apply Taylor's theorem in the form

$$f(b) = f(a) + f^{(1)}(a)(b-a) + \cdots + f^{(k-1)}(a)\frac{(b-a)^{k-1}}{(k-1)!} + f^{(k)}(c)\frac{(b-a)^k}{k!} \tag{50}$$

for some point c with $a < c < b$ for $k \geq 1$ (using modified derivative notation).

Applying this to the second conditional-expectation term in (48) and the definition of η_n in (37), we get

$$W_{n,2}(t, t+s) \leq \frac{\|f'''\|}{6} E\left[\sup_{t < u < t+s} |\Delta Y_n(u)| k \sum_{i=1}^{k} (c_{i,i}(t+s) + 1)\theta_i^2 | \mathcal{F}_{n,t}\right] .$$

Reasoning the same way for the other two terms, overall we get a minor variation of the bound in (1.33) on p. 342 of EK, namely,

$$E\left[f(Y_n(t+s)) - f(Y_n(t)) | \mathcal{F}_{n,t}\right]$$

$$\leq C_f E\left[\sup_{t < u \leq t+s} |\Delta Y_n(u)| \left(1 + k \sum_{i=1}^{k} (c_{i,i}(t+s) + 1)\theta_i^2\right)\right.$$

$$\left. + A_{n,\theta}((t+s) \wedge \eta_n-) - A_{n,\theta}(t \wedge \eta_n-) | \mathcal{F}_{n,t}\right] , \tag{51}$$

where the constant C_f depends only on the norms of the first three derivatives: $\|f'\|$, $\|f''\|$ and $\|f'''\|$.

We now apply the inequality in (51) to construct the required bounds. First, for the stochastic boundedness needed in condition (i) of Lemma 3.11, we let the function f have the additional properties in Lemma 3.6; i.e., we assume that f is an even nonnegative convex function with $f'(t) > 0$ for $t > 0$. Then there exists a constant K such that

$$E[f(Y_n(T))] \leq C_f E\left[(1 + J(Y_n, T)|) \left(1 + k \sum_{i=1}^{k} (c_{i,i}(T) + 1)\theta_i^2\right)\right] < K \tag{52}$$

for all n by virtue of (2), again using J in (1)). (See (1.34) on p 343 of EK, where t on the right should be T.)

Next, we can let the random variables $Z_n(\delta, f, T)$ needed in condition (ii.a) of Lemma 3.11 be defined, as in (1.36) on p. 343 of EK, by

$$Z_n(\delta, f, T) \equiv C_f \left[J(Y_n, T + \delta) \left(1 + k \sum_{i=1}^{k} (c_{i,i}(t+s) + 1)\theta_i^2 \right) \right.$$
$$\left. + \sup_{0 \le t \le T} \{ A_{n,\theta}((t+\delta) \wedge \eta_n-) - A_{n,\theta}(t \wedge \eta_n-) \} \right] , \quad (53)$$

where C_f is a new constant depending on the function f. Then, by (2) and (3),

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[Z_n(\delta, f, T)] = 0 , \quad (54)$$

as required in condition (ii.a) of Lemma 3.11. That completes the proof of tightness in case (i). ■

4.3. The Ethier-Kurtz Proof in Case (ii).

We already gave a complete proof of tightness for case (ii) in §4.1. We now give an alternative proof, following EK. The proof starts out the same as the proof for case (i). First, if the stochastic processes $M_{n,i,i}^2 - \langle M_{n,i}, M_{n,i} \rangle$ are local martingales rather than martingales, then we introduce the available stopping times τ_n with $\tau_n \uparrow \infty$ w.p.1 so that $\{ M_{n,i,i}^2(\tau_n \wedge t) - \langle M_{n,i}, M_{n,i} \rangle(\tau_n \wedge t) : t \ge 0 \}$ are martingales for each n and i. Hence, we can assume that $M_{n,i,i}^2 - \langle M_{n,i}, M_{n,i} \rangle$ are martingales.

We then exploit the assumed limit $\langle M_{n,i}, M_{n,j} \rangle(t) \Rightarrow c_{i,j}t$ in $\mathbb{R}$ as $n \to \infty$ in (6) in order to act as if $\langle M_{n,i}, M_{n,j} \rangle(t)$ is bounded. In particular, as in (37) and (1.23) on p. 341 of EK, we introduce the additional stopping times

$$\eta_n \equiv \inf \{ t \ge 0 : \langle M_{n,i}, M_{n,i} \rangle(t) > c_{i,i}t + 1 \quad \text{for some} \quad i \} . \quad (55)$$

We then define
$$\tilde{M}_n(t) \equiv M_n(\eta_n \wedge t), \quad t \ge 0 . \quad (56)$$
Again we apply Lemma 2.1 to deduce that we have convergence $\langle M_{n,i}, M_{n,i} \rangle \Rightarrow c_{i,i}e$ in D for each i, which implies that $\eta_n \Rightarrow \infty$ as $n \to \infty$.

Closely paralleling EK, simplify notation by writing

$$A_{n,i,j}(t) \equiv \langle M_{n,i}, M_{n,j} \rangle(t) \quad \text{and} \quad \tilde{A}_{n,i,j}(t) \equiv A_{n,i,j}(t \wedge \eta_n), \quad t \ge 0 . \quad (57)$$

Then
$$\tilde{A}_{n,i,i}(t) \le 1 + c_{i,i}t + J(A_{n,i,i}, t), \quad t \ge 0 , \quad (58)$$

for J in (1), where

$$E[J(A_{n,i,i}, t)] \to 0 \quad \text{as} \quad n \to \infty \quad \text{for all} \quad t > 0 , \quad (59)$$

by condition (4). Since $\tilde{M}_{n,i,i}^2 - \tilde{A}_{n,i,i}$ is a martingale, so is $\sum_{i=1}^{k}(\tilde{M}_{n,i,i}^2 - \tilde{A}_{n,i,i})$. Hence, for $0 \le s < t$,

$$\sum_{i=1}^{k} E\left[\tilde{M}_{n,i}(t+s)^2 - \tilde{M}_{n,i}(t)^2 | \mathcal{F}_{n,t}\right] = \sum_{i=1}^{k} E\left[\tilde{A}_{n,i}(t+s) - \tilde{A}_{n,i}(t) | \mathcal{F}_{n,t}\right] \tag{60}$$

On the other hand,

$$\sum_{i=1}^{k} E\left[\left(\tilde{M}_{n,i}(t+s) - \tilde{M}_{n,i}(t)\right)^2 | \mathcal{F}_{n,t}\right]$$

$$= \sum_{i=1}^{k} E\left[\left(\tilde{M}_{n,i}(t+s)^2 - 2\tilde{M}_{n,i}(t+s)\tilde{M}_{n,i}(t) + \tilde{M}_{n,i}(t)^2\right) | \mathcal{F}_{n,t}\right]$$

$$= \sum_{i=1}^{k} E\left[\left(\tilde{M}_{n,i}(t+s)^2 - \tilde{M}_{n,i}(t)^2\right) | \mathcal{F}_{n,t}\right] . \tag{61}$$

Hence,

$$E\left[\sum_{i=1}^{k}\left(\tilde{M}_{n,i}(t+s) - \tilde{M}_{n,i}(t)\right)^2 | \mathcal{F}_{n,t}\right]$$

$$= E\left[\sum_{i=1}^{k} \tilde{A}_{n,i}(t+s) - \tilde{A}_{n,i}(t) | \mathcal{F}_{n,t}\right] , \tag{62}$$

as in 1.41 on p 344 of EK.

Now we can verify condition (ii.b) in Lemma 3.11: We can define the random variables $Z_n(\delta, T)$ by

$$Z_n(\delta, T) \equiv \sup_{0 \le t \le T}\left\{\sum_{i=1}^{k}(\tilde{A}_{n,i,i}(t+\delta) - \tilde{A}_{n,i,i}(t))\right\} . \tag{63}$$

By condition (6) and Lemma 2.1,

$$Z_n(\delta, T) \Rightarrow \sup_{0 \le t \le T}\left\{\sum_{i=1}^{k}(c_{i,i}(t+\delta) - c_{i,i}(t))\right\} = \delta \sum_{i=1}^{k} c_{i,i} \quad \text{as} \quad n \to \infty . \tag{64}$$

We then have uniform integrability because, by (58) and (63),

$$Z_n(\delta, T) \le \sum_{i=1}^{k}\left(1 + c_{i,i}(T+\delta) + J(A_{n,i,i}, T)\right) , \tag{65}$$

where J is the maximum jump function in (1) and $E\left[J(A_{n,i,i}, T)\right] \to 0$ as $n \to \infty$. Hence,

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} E[Z_n(\delta, T)] = 0 . \tag{66}$$

Finally, we must verify condition (i) Lemma 3.11; i.e., we must show that $\{M_{n,i} : n \geq 1\}$ is stochastically bounded. As for case (i), we can apply Lemma 3.6 for this purpose, here using the function $f(x) \equiv x^2$. Modifying (62), we have, for all $T > 0$, the existence of a constant K such that

$$E\left[\sum_{i=1}^{k} \tilde{M}_{n,i}(T)^2\right] = E\left[\sum_{i=1}^{k} \tilde{A}_{n,i}(T)\right] \leq K \tag{67}$$

for all sufficiently large n. That completes the proof. ∎

5. Characterization of the Limit

5.1. With Uniformly Bounded Jumps.

Since the sequences of martingales $\{Y_n : n \geq 1\}$ and $\{\tilde{M}_n : n \geq 1\}$ are tight, they are relatively compact by Prohorov's theorem (Theorem 3.1). We consider a converging subsequence $\tilde{M}_{n_k} \Rightarrow L$. It remains to show that L necessarily is the claimed k-dimensional $(0, C)$-Brownian motion M. By Corollary 3.3, that will imply that $\tilde{M}_n \Rightarrow M$. Since $\eta_n \Rightarrow \infty$, it then will follow that $M_n \Rightarrow M$ as well. But in this section we will be working with the martingales $\tilde{M}_{n_k}$, for which convergence has been established by the previous tightness proofs.

For this characterization step, we first present a proof based upon JS, which requires that the jumps be uniformly bounded as an extra condition. We consider the proof in EK afterwards in the next section, §5.2.

It seems useful to state the desired characterization result as a theorem. Conditions (i) and (ii) below cover the two cases of Theorem 2.1 under the extra condition (68).

Theorem 5.1 (characterization of the limit) *Suppose that $M_n \Rightarrow M$ in D^k, where $M_n \equiv (M_{n,1}, \ldots, M_{n,k})$ is a k-dimensional local martingale adapted to the filtration $\boldsymbol{F}_n \equiv \{\mathcal{F}_{n,t}\}$ for each n with bounded jumps, i.e., for all $T > 0$, there exists n_0 and K such that*

$$J(M_{n,i}, T) \leq K \quad for \ all \quad n \geq n_0 , \tag{68}$$

where J is the maximum-jump function in (1). Suppose that $M \equiv (M_1, \ldots, M_k)$ is a continuous k-dimensional process with $M(0) = (0, \ldots, 0)$. In addition, suppose that either

$$(i) \quad [M_{n,i}, M_{n,j}](t) \Rightarrow c_{i,j} t \quad as \quad n \to \infty \quad in \quad \mathbb{R} \tag{69}$$

for all $t > 0$, i and j, or (ii) M_n is locally square integrable and

$$\langle M_{n,i}, M_{n,j}\rangle(t) \Rightarrow c_{i,j} t \quad as \quad n \to \infty \quad in \quad \mathbb{R} \tag{70}$$

for all $t > 0$, i and j. Then M is a k-dimensional $(0, C)$-Brownian motion, having time-dependent mean vector and covariance matrix

$$E[M(t)] = (0, \ldots, 0) \quad and \quad E[M(t)M(t)^{tr}] = Ct, \quad t \geq 0 . \tag{71}$$

Theorem 5.1 follows directly from several other basic results. The first is the classical Lévy martingale characterization of Brownian motion; e.g., see p. 156 of Karatzas and Shreve [9] or p. 102 of JS. Ito's formula provides an elegant proof. (Theorem 7.1.2 on p 339 of EK is a form that exploits Ito's formula in the statement.) Recall that for a continuous local martingale both the angle-bracket and square-bracket processes are well defined and equal.

Theorem 5.2 (Lévy martingale characterization of Brownian motion) *Let* $M \equiv (M_1, \ldots, M_k)$ *be a continuous k-dimensional process adapted to a filtration $\boldsymbol{F} \equiv \{\mathcal{F}_t\}$ with $M_i(0) = 0$ for each i. Suppose that each one-dimensional marginal process M_i is a continuous local-$\boldsymbol{F}$-martingale. If either the optional covariation processes satisfy*

$$[M_i, M_j] = c_{i,j}t, \quad t \geq 0 , \tag{72}$$

for each i and j, or if M_i are locally square-integrable martingales, so that the predictable quadratic covariation processes are well defined, and they satisfy

$$\langle M_i, M_j \rangle(t) = c_{i,j}t, \quad t \geq 0 , \tag{73}$$

for each i and j, where $C \equiv (c_{i,j})$ is a nonnegative-definite symmetric real matrix, then M is a k-dimensional $(0, C)$-Brownian motion, i.e.,

$$E[M(t)] = (0, \ldots, 0) \quad and \quad E[M(t)M(t)^{tr}] = Ct, \quad t \geq 0 . \tag{74}$$

We now present two basic preservation results. The first preservation result states that the limit of any convergent sequence of martingales must itself be a martingale, under regularity conditions. The following is Problem 7 on p 362 of EK. The essential idea is conveyed by the bounded case, which is treated in detail in Proposition IX.1.1 on p 481 of JS. JS then go on to show that the boundedness can be replaced by uniform integrability (UI) and then a bounded-jump condition. (Note that this is UI of the sequence $\{M_n(t) : n \geq 1\}$ over n for fixed t, as opposed to the customary UI over t for fixed n.)

Theorem 5.3 (preservation of the martingale property under UI) *Suppose that (i) X_n is a random element of D^k and M_n is a random element of D for each $n \geq 1$, (ii) M_n is a martingale with respect to the filtration generated by (X_n, M_n) for each $n \geq 1$, and (iii) $(X_n, M_n) \Rightarrow (X, M)$ in D^{k+1} as $n \to \infty$. If, in addition, $\{M_n(t) : t \geq 1\}$ is uniformly integrable for each $t > 0$, then M is a martingale with respect to the filtration generated by (X, M) (and thus also with respect to the filtration generated by M).*

Proof.   Let $\{\mathcal{F}_{n,t} : t \geq 0\}$ and $\{\mathcal{F}_t : t \geq 0\}$ denote the filtrations generated by (X_n, M_n) and (X, M), respectively, on their underlying probability spaces. Let the probability space for the limit (X, M) be $(\Omega, \mathcal{F}, P)$. It is difficult to relate these filtrations directly, so we do so indirectly. This involves a rather tricky measurability argument. We accomplish this goal by considering the stochastic processes (X, M) and (X_n, M_n) as maps from the underlying probability spaces to the function space D^{k+1} with its usual sigma-field generated by the coordinate

projections, i.e., with the filtration $\mathbf{D}^{k+1} \equiv \{\mathcal{D}_t^{k+1} : t \geq 0\}$. As discussed in §VI.1.1 of JS and §11.5.3 of Whitt [22], the Borel σ-field on D^{k+1} coincides with the σ-field on D^{k+1} generated by the coordinate projections. A $\mathcal{D}_t^{k+1}$-measurable real-valued function $f(x)$ defined on D^{k+1} depends on the function x in D^{k+1} only through its behavior in the initial time interval $[0, t]$.

As in step (a) of the proof of Proposition IX.1.1 of JS, we consider the map $(X, M) : \Omega \to D^{k+1}$ mapping the underlying probability space into D^{k+1} with the sigma-field $\mathbf{D}^{k+1}$ generated by the coordinate projections. By this step, we are effectively lifting the underlying probability space to D^{k+1}, with $(X, M)(t)$ obtained as the coordinate projection.

Let t_1 and t_2 with $t_1 < t_2$ be almost-sure continuity points of the limit process (X, M). Let $f : D^{k+1} \to \mathbb{R}$ be a bounded continuous $\mathcal{D}_{t_1}^{k+1}$-measurable real-valued function in that setting. By this construction, not only is $f(X, M)$ $\mathcal{F}_{t_1}$-measurable but $f(X_n, M_n)$ is $\mathcal{F}_{n,t_1}$ measurable for each $n \geq 1$ as well. (Note that $f(X_n, M_n)$ is the composition of the maps $(X_n, M_n) : \Omega_n \to D^{k+1}$ and $f : D^{k+1} \to \mathbb{R}$.)

By the continuous-mapping theorem, we have first

$$(f(X_n, M_n), M_n(t_2), M_n(t_1)) \Rightarrow (f(X, M), M(t_2), M(t_1)) \quad \text{in} \quad \mathbb{R}^3 \quad \text{as} \quad n \to \infty$$

and then

$$f(X_n, M_n)(M_n(t_2) - M_n(t_1)) \Rightarrow f(X, M)(M(t_2) - M(t_1)) \quad \text{in} \quad \mathbb{R} \quad \text{as} \quad n \to \infty .$$

By the boundedness of f and the uniform integrability of $\{M_n(t_i)\}$, we then have

$$E\left[f(X_n, M_n)(M_n(t_2) - M_n(t_1))\right] \to E\left[f(X, M)(M(t_2) - M(t_1))\right] \quad \text{as} \quad n \to \infty .$$

Now we exploit the fact that $f(X_n, M_n)$ is actually $\mathcal{F}_{n,t_1}$-measurable for each n. We are thus able to invoke the martingale property for each n and conclude that

$$E\left[f(X_n, M_n)(M_n(t_2) - M_n(t_1))\right] = 0 \quad \text{for all} \quad n .$$

Combining these last two relations, we have the relation

$$E\left[f(X, M)(M(t_2) - M(t_1))\right] = 0 . \tag{75}$$

Now, by a monotone class argument (e.g., see p 496 of EK), we can show that the relation (75) remains true for all bounded $\mathcal{D}_{t_1}^{k+1}$-measurable real-valued functions f, which includes the indicator function of an arbitrary measurable set A in $\mathcal{D}_{t_1}^{k+1}$. Hence,

$$E\left[1_{\{(X,M)\in A\}}(M(t_2) - M(t_1))\right] = 0 .$$

This in turn implies that

$$E\left[1_B(M(t_2) - M(t_1))\right] = 0$$

for all B in $\mathcal{F}_{t_1}$, which implies that

$$E\left[M(t_2) - M(t_1))|\mathcal{F}_{t_1}\right] = 0 ,$$

which is the desired martingale property for these special time points t_1 and t_2. We obtain the result for arbitrary time points t_1 and t_2 by considering limits from the right. ∎

A way to get the uniform-integrability regularity condition for martingales in Theorem 5.3 is to have a local martingale with uniformly bounded jumps. The following is Corollary IX.1.19 to Proposition IX.1.17 on p 485 of JS.

Theorem 5.4 (preservation of the local-martingale property with bounded jumps) *Suppose that conditions* $(i)-(iii)$ *of Theorem* 5.3 *hold, except that* M_n *is only required to be a local martingale for each* n. *If, in addition, for each* $T > 0$, *there exists a positive integer* n_0 *and a constant* K *such that*

$$J(M_n, T) \le K \quad \text{for all} \quad n \ge n_0 \ , \tag{76}$$

then M *is a local martingale with respect to the filtration generated by* X (*and thus also with respect to the filtration generated by* M).

The second preservation result states that, under regularity conditions, the optional quadratic variation $[M]$ of a local martingale is a continuous function of the local martingale. The following is Corollary VI.6.7 on p 342 of JS.

Theorem 5.5 (preservation of the optional quadratic variation) *Suppose that* M_n *is a local martingale for each* $n \ge 1$ *and* $M_n \Rightarrow M$ *in* D. *If, in addition, for each* $T > 0$, *there exists a positive integer* n_0 *and a constant* K *such that*

$$E\left[J(M_n, T)\right] \le K \quad \text{for all} \quad n \ge n_0 \ , \tag{77}$$

then

$$(M_n, [M_n]) \Rightarrow (M, [M]) \quad in \quad D^2 \ . \tag{78}$$

Note that condition (77) is implied by condition (76). More importantly, condition (77) is implied by condition (2).

Proof of Theorem 5.1. From Theorem 5.2, we see that it suffices to focus on one coordinate at a time. To characterize the covariation processes, we can consider the weighted sums

$$M_{n,\theta} \equiv \sum_{i=1}^{k} \theta_i M_{n,i} \tag{79}$$

for arbitrary non-null vector $\theta \equiv (\theta_1, \ldots, \theta_k)$. First suppose that condition (i) of Theorem 5.1 is satisfied. We get the covariations via

$$2[M_{n,i}, M_{n,j}] = [M_{n,i} + M_{n,j}, M_{n,i} + M_{n,j}] - [M_{n,i}, M_{n,i}] - [M_{n,j}, M_{n,j}] \ . \tag{80}$$

Henceforth fix θ. First, condition (i) of Theorem 5.1 implies that

$$[M_{n,\theta}](t) \Rightarrow c_\theta t \quad \text{as} \quad n \to \infty \quad \text{in} \quad \mathbb{R} \quad \text{for all} \quad t > 0 \ . \tag{81}$$

Condition (68) implies condition (76) for $M_{n,\theta}$. Hence, we can apply Theorems 5.4 and 5.5 to deduce that the limit process $M_\theta \equiv \sum_{i=1}^{k} \theta_i M_i$ is a local martingale with optional quadratic variation $[M_\theta] = c_\theta e$, where $c_\theta \equiv \sum_{i=1}^{k} \sum_{j=1}^{k} \theta_i \theta_j c_{i,j}$. Since this is true for all θ, we can apply Theorem 5.2 and (80) to complete the proof.

Now, instead, assume that condition (ii) of Theorem 5.1 is satisfied. First, by the Doob-Meyer decomposition, Theorem 3.1 of Pang et al. [17], we can deduce that $M_{n,\theta}^2 - \langle M_{n,\theta} \rangle$ is a local martingale for each n and θ. Next, it follows from the assumed convergence in (70), Lemma 2.1 and the continuous mapping theorem that

$$M_{n,\theta}^2 - \langle M_{n,\theta} \rangle \Rightarrow M_\theta^2 - c_\theta e \quad \text{in} \quad D \quad \text{as} \quad n \to \infty . \tag{82}$$

By Theorem 5.4 and condition (68), the stochastic process $M_\theta^2 - c_\theta e$ is a local martingale for each vector θ, but that means that $\langle M_\theta \rangle(t) = c_\theta t$ for each vector θ, which implies the predictable quadratic covariation condition in Theorem 5.2. Paralleling (80), we use

$$2\langle M_i, M_j \rangle = \langle M_i + M_j, M_i + M_j \rangle - \langle M_i, M_i \rangle - \langle M_j, M_j \rangle . \tag{83}$$

Hence, we can apply Theorem 5.2 to complete the proof. ∎

5.2. The Ethier-Kurtz Proof in Case (i).

The proof of the characterization step for case (i) on p 343-344 of EK is quite brief. From the pointer to Problem 7 on p. 362 of EK, one might naturally think that we should be applying the result there, which corresponds to Theorem 5.3 here, and represent the limit process as a limit of an appropriate sequence of martingales. A helpful initial observation here is that it is not actually necessary for the converging processes to be martingales. Instead of applying Theorem 5.3, we can follow the proof of Theorem 5.3 in order to obtain the desired martingale property asymptotically.

Accordingly, we directly establish that the limit process is a martingale. In particular, we show that the stochastic process

$$\{f(L(t)) - \frac{c_\theta}{2} \int_0^t f''(L(s-)) \, ds : t \geq 0\} , \tag{84}$$

is a martingale, where L is the limit of the converging subsequence $\{Y_{n_k}\}$. In particular, we will show that

$$E\left[f(L(t+s)) - f(L(t) - \frac{c_\theta}{2} \int_t^{t+s} f''(L(u-)) \, du \,\big|\, \mathcal{F}_t \right] = 0 \tag{85}$$

for $0 < t < t + s$.

Toward that end, EK present their (1.38) and (1.39) on p 343. The claim about convergence in probability uniformly over compact subsets of D in (1.38)

on p 343 of EK seems hard to verify, but it is not hard to prove the desired (1.39), exploiting stronger properties that hold in this particular situation. We now provide additional details about the proof of a variant of (1.39) on p. 343 of EK. The desired variant of that statement is (87) below.

Lemma 5.1 *Under the conditions of Theorem 2.1 in case (i), if L is the limit of a convergent subsequence $\{Y_{n_k} : n \geq 1\}$, then*

$$\int_{t \wedge \eta_{n_k}}^{(t+s) \wedge \eta_{n_k}} \frac{1}{2} f''(Y_{n_k}(u-)) \, dA_{n_k,\theta}(u) \Rightarrow \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \qquad (86)$$

in $D([t, t+s], \mathbb{R})$ as $n_k \to \infty$. As a consequence,

$$E\left[\left| \int_{t \wedge \eta_{n_k}}^{(t+s) \wedge \eta_{n_k}} \frac{1}{2} f''(Y_{n_k}(u-)) \, dA_{n_k,\theta}(u) - \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \right| \right] \to 0 \; ,$$
$$(87)$$

as in (1.39) on p 343 of EK.

Proof of Lemma 5.1. We start by proving (86). First, we use the fact that both $A_{n,\theta}$ and $c_\theta e$ are nondecreasing functions. We also exploit Lemma 2.1 in order to get $A_{n_k,\theta} \Rightarrow c_\theta e$ in D. The topology is uniform convergence over bounded intervals, because $c_\theta e$ is a continuous function. Given that $Y_{n_k} \Rightarrow L$, we can apply Lemma 3.10 and condition (2) to deduce that L and thus $f''(L(t))$ actually have continuous paths, so the topology is again uniform convergence over bounded intervals. Then the limit in (86) is elementary: First, for any $\epsilon > 0$, $P(\|Y_{n_k} - L\|_{t+s} > \epsilon) \to 0$ and $P(\eta_{n_k} \leq t + s) \to 0$ as $n_k \to \infty$. As a consequence, $P(\|f'' \circ Y_{n_k} - f'' \circ L\|_{t+s} > \epsilon) \to 0$ as $n \to \infty$ too. Given that $\|f'' \circ Y_{n_k} - f'' \circ L\|_{t+s} \leq \epsilon$ and $\eta_{n_k} \geq t + s$,

$$\left| \int_{t \wedge \eta_{n_k}}^{(t+s) \wedge \eta_{n_k}} \frac{1}{2} f''(Y_{n_k}(u-)) \, dA_{n_k,\theta}(u) - \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \right|$$

$$\leq \left| \int_t^{(t+s)} \frac{1}{2} f''(Y_{n_k}(u-)) \, dA_{n_k,\theta}(u) - \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, dA_{n_k,\theta}(u) \right|$$

$$+ \left| \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, dA_{n_k,\theta}(u) - \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \right|$$

$$\leq \epsilon (A_{n_k,\theta}(t+s) - A_{n_k,\theta}(t)) + \left| \int_t^{t+s} |\frac{1}{2} f''(L(u-))| \, (dA_{n_k,\theta}(u) - c_\theta \, du) \right|$$

$$\Rightarrow 0 \quad \text{as} \quad n \to \infty \; . \qquad (88)$$

Hence we have the first limit (86). Since f'' and the length of the interval $[t, t+s]$ are both bounded, we also get the associated limit of expectations in (87). ∎

Back to the Characterization Proof. We now apply the argument used in the proof of Theorem 5.3. Let $\{\mathcal{F}_{n_k,t} : t \geq 0\}$ and $\{\mathcal{F}_t : t \geq 0\}$ denote the

filtrations generated by Y_{n_k} and L, respectively, on their underlying probability spaces. Let the probability space for the limit Y be $(\Omega, \mathcal{F}, P)$. As in the proof of Theorem 5.3, we consider a real-valued function from the function space D with its usual sigma-field generated by the coordinate projections, i.e., with the filtration $\mathbf{D} \equiv \{\mathcal{D}_t : t \geq 0\}$. A $\mathcal{D}_t$-measurable function defined on D depends on the function x in D only through its behavior in the initial time interval $[0, t]$.

Let t and $t + s$ be the arbitrary time points with $0 \leq t < t + s$. Let $h : D \to \mathbb{R}$ be a bounded continuous $\mathcal{D}_t$-measurable real-valued function in that setting. By this construction, not only is $h(L)$ $\mathcal{F}_t$-measurable, but $h(Y_{n_k})$ is $\mathcal{F}_{n_k, t}$ measurable for each $k \geq 1$ as well.

By the continuous-mapping theorem, we have first

$$\left(h(Y_{n_k}), f(Y_{n_k}(t)), f(Y_{n_k}(t+s)), \int_{t \wedge \eta_{n_k}}^{(t+s) \wedge \eta_{n_k}} \frac{1}{2} f''(Y_{n_k}(u-)) \, dA_{n_k, \theta}(u) \right)$$
$$\Rightarrow \left(h(L), f(L(t)), f(L(t+s)), \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \right) \qquad (89)$$

in $\mathbb{R}^4$ as $n \to \infty$, and then

$$h(Y_{n_k}) \left(f(Y_{n_k}(t+s)) - f(Y_{n_k}(t)) - \int_{t \wedge \eta_{n_k}}^{(t+s) \wedge \eta_{n_k}} \frac{1}{2} f''(Y_{n_k}(u-)) \, dA_{n_k, \theta}(u) \right)$$
$$\Rightarrow h(L) \left(f(L(t+s)) - f(L(t)) - \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \right) \qquad (90)$$

in $\mathbb{R}$ as $k \to \infty$. On the other hand, as argued by EK, by condition (46) and the boundedness of f and f'', the first two terms $W_{n,1}(t+s)$ and $W_{n,2}(t, t+s)$ on the right in (48) converge to 0 in L_1 as well. Hence, the limit in (90) must be 0. As a consequence, we have shown that

$$E \left[h(L) \left(f(L(t+s)) - f(L(t)) - \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \right) \right] = 0 \qquad (91)$$

for all continuous bounded $\mathcal{D}_t$-measurable real-valued functions h. By the approximation argument involving the monotone class theorem (as described in the proof of Theorem 5.3), we thus get

$$E \left[1_B \left(f(L(t+s)) - f(L(t)) - \int_t^{t+s} \frac{1}{2} f''(L(u-)) \, c_\theta \, du \right) \right] = 0 \qquad (92)$$

for each measurable subset B in $\mathcal{F}_t$. That implies (85), which in turn implies the desired martingale property. Finally, Theorem 7.1.2 on p 339 of EK (playing the role of Theorem 5.2 here) shows that this martingale property for all these smooth functions f implies that the limit L must be Brownian motion, as claimed. ■

5.3. The Ethier-Kurtz Proof in Case (ii).

The proof of characterization in this second case seems much easier for two reasons: First, the required argument is relatively straightforward (compared to the previous section) and, second, EK provides almost all the details.

We mostly just repeat the argument in EK. It suffices to focus on a limit L of a convergent subsequence $\{\tilde{M}_{n_k} : n_k \geq 1\}$ of the relatively compact sequence $\{\tilde{M}_n : n \geq 1\}$ in D^k. By condition (5) and Lemma 3.10, L almost surely has continuous paths. By the arguments in the EK tightness proof in §4.3, starting with (60), we have

$$\sup_{n \geq 1} E[\sum_{i=1}^{k} \tilde{M}_{n,i}(T)^2] < \infty$$

for each $T > 0$. As a consequence, the sequence of random vectors $\{\tilde{M}_n(T) : n \geq 1\}$ is uniformly integrable; see (3.18) on p 31 of Billingsley [5]. Hence, Theorem 5.3 implies that the limit L is a martingale. Moreover, by condition (4) and Lemma 2.1, we have the convergence

$$\tilde{M}_{n_k,i}\tilde{M}_{n_k,j} - \tilde{A}_{n,i,j} \Rightarrow L_i L_j - c_{i,j}e \quad \text{in} \quad D \quad \text{as} \quad n_k \to \infty \ .$$

We can conclude that the limit $L_i L_j - c_{i,j}e$ too is a martingale by applying Theorem 5.3, provided that we can show that the sequence $\{\tilde{M}_{n_k,i}(T)\tilde{M}_{n_k,j}(T) - \tilde{A}_{n_k,i,j}(T) : n_k \geq 1\}$ is uniformly integrable for each $T > 0$.

By (58), condition (4), and the inequality

$$2|A_{n_k,i,j}(T)| \leq |A_{n_k,i,i}(T)| + |A_{n_k,j,j}(T)| \ ,$$

the sequence $\{\tilde{A}_{n_k,i,j}(T) : n_k \geq 1\}$ is uniformly integrable. Hence it suffices to focus on the other sequence $\{\tilde{M}_{n_k,i}(T)\tilde{M}_{n_k,j}(T) : n_k \geq 1\}$, and since

$$2|\tilde{M}_{n_k,i}(T)\tilde{M}_{n_k,j}(T)| \leq |\tilde{M}_{n_k,i}(T)^2| + |\tilde{M}_{n_k,j}(T)^2| \ ,$$

it suffices to consider the sequence $\{\tilde{M}_{n_k,i}(T)^2 : n_k \geq 1\}$ for arbitrary i. The required argument now is simplified because the random variables are non-negative. Since $\tilde{M}_{n_k,i}(T)^2 \Rightarrow L_i(T)^2$ as $n_k \to \infty$, it suffices to show that $E[\tilde{M}_{n_k,i}(T)^2] \to E[L_i(T)^2]$ as $n_k \to \infty$. By (60)–(62), we already know that $E[\tilde{M}_{n_k,i}(T)^2] \to c_{i,i}T$ as $n_k \to \infty$. So what is left to prove is that

$$E[L_i(T)^2] = c_{i,i}T \ , \tag{93}$$

as claimed in (1.45) on p 344 of EK.

For this last step, we introduce stopping times

$$\tau_n(x) \equiv \inf \{t \geq 0 : \tilde{M}_{n,i}(t)^2 > x\} \quad \text{and} \quad \tau(x) \equiv \inf \{t \geq 0 : L_i(t)^2 > x\} \ , \tag{94}$$

for each $x > 0$. We deviate from EK a bit here. It will be convenient to guarantee that these stopping times are finite for all x. (So far, we could have $L(t) = 0$

for all $t \geq 0$.) Accordingly, we define modified stopping times that necessarily are finite for all x. We do so by adding the continuous deterministic function $(t - T)^+$ to the two stochastic processes. In particular, let

$$\tau_n(x)' \equiv \inf \{t \geq 0 : \tilde{M}_{n,i}(t)^2 + (t - T)^+ > x\}$$

and

$$\tau(x)' \equiv \inf \{t \geq 0 : L_i(t)^2 + (t - T)^+ > x\} \quad \text{for} \quad x \geq 0 .$$

With this modification, the stopping times $\tau_n(x)'$ and $\tau(x)'$ are necessarily finite for all x, and yet nothing is changed over the relevant time interval $[0, T]$:

$$\tau_n(x)' \wedge t = \tau_n(x) \wedge t \quad \text{and} \quad \tau(x)' \wedge t = \tau(x) \wedge t \quad \text{for} \quad 0 \leq t \leq T \qquad (95)$$

for all $x \geq 0$.

After having made this minor modification, we can apply the continuous mapping theorem with the inverse function as in §13.6 of Whitt [22], mapping the subset of functions unbounded above in D into itself, which for the cognoscenti will be a simplification. That is, we regard the first-passage times as stochastic processes indexed by $x \geq 0$. For this step, we work with the stochastic processes (random elements of D) $\tau_n \equiv \{\tau_n(x) : x \geq 0\}$ and similarly for the other first passage times. We need to be careful to make this work: We need to use a weaker topology on D on the range. But Theorem 13.6.2 of [22] implies that

$$(\tau'_{n_k}, \tilde{M}_{n_k}) \Rightarrow (\tau', L) \quad \text{in} \quad (D, M_1') \times (D, J_1) \quad \text{as} \quad k \to \infty , \qquad (96)$$

which in turn, by the continuous mapping theorem again, implies that

$$(\tau_{n_k}(x)' \wedge T, \tilde{M}_{n_k}) \Rightarrow (\tau(x)' \wedge T, L) \quad \text{in} \quad \mathbb{R} \times (D, J_1) \quad \text{as} \quad k \to \infty , \qquad (97)$$

for all x except the at most countably many x that are discontinuity points of the limiting stochastic process $\{\tau(x)' : x \geq 0\}$. Since we have restricted the time argument to the interval $[0, T]$, we can apply (95) and use the original stopping times, obtaining

$$(\tau_{n_k}(x) \wedge T, \tilde{M}_{n_k}) \Rightarrow (\tau(x) \wedge T, L) \quad \text{in} \quad \mathbb{R} \times (D, J_1) \quad \text{as} \quad k \to \infty \qquad (98)$$

again for all x except the at most countably many x that are discontinuity points of the limiting stochastic process $\tau' \equiv \{\tau(x)' : x \geq 0\}$. By the continuous mapping theorem once more, with the composition map and the square, e.g., see VI.2.1 (b5) on p 301 of JS, we obtain

$$\tilde{M}_{n_k,i}(t \wedge \tau_{n_k}(x))^2 \Rightarrow L(t \wedge \tau(x))^2 \quad \text{in} \quad \mathbb{R} \quad \text{as} \quad k \to \infty \qquad (99)$$

for all but countably many x. At this point we have arrived at a variant of the conclusion reached on p 345 of EK. (We do not need to restrict the T because the limit process L has been shown to have continuous paths.) It is not necessary to consider the inverse function mapping a subset of D into itself, as we did;

instead we could go directly to (98), but that joint convergence is needed in order to apply the continuous mapping theorem to get (99).

Now we have the remaining UI argument in (1.48) of EK, which closely parallels (65). We observe that the sequence $\{\tilde{M}_{n_k,i}(T \wedge \tau_{n_k}(x))^2 : k \geq 1\}$ is UI, because

$$\tilde{M}_{n_k,i}(T \wedge \tau_{n_k}(x))^2 \leq 2 \left(x + J(\tilde{M}_{n_k,i}^2, T) \right) \tag{100}$$

where

$$E \left[J(\tilde{M}_{n_k,i}^2, T) \right] \to 0 \quad \text{as} \quad k \to \infty \tag{101}$$

by virtue of condition (5), again using J in (1).

As a consequence, as in (1.49) on p 345 of EK, for all but countably many x, we get

$$
\begin{aligned}
E[L_i(T \wedge \tau(x))^2] &= \lim_{k \to \infty} E[\tilde{M}_{n_k,i}(T \wedge \tau_{n_k}(x))^2] \\
&= \lim_{k \to \infty} E[\tilde{A}_{n_k,i,i}(T \wedge \tau_{n_k}(x))] \\
&= E[c_{i,i}(T \wedge \tau(x))] \ .
\end{aligned} \tag{102}
$$

Letting $x \to \infty$ through allowed values, we get $\tau(x) \Rightarrow \infty$ and then the desired (93). ∎

Acknowledgments.

References

[1] Aldous, D. 1978. Stopping times and tightness. *Ann. Probab.* 6, 335–340. MR0474446

[2] Aldous, D. 1989. Stopping times and tightness II. *Ann. Probab.* 17, 586–595. MR0985380

[3] Billingsley, P. 1968. *Convergence of Probability Measures*, Wiley, New York. MR0233396

[4] Billingsley, P. 1974. Conditional distributions and tightness. *Ann. Probab.* 2, 480–485. MR0368095

[5] Billingsley, P. 1999. *Convergence of Probability Measures*, second edition, Wiley, New York. MR1700749

[6] Ethier, S. N. and T. G. Kurtz. 1986. *Markov Processes; Characterization and Convergence*, Wiley. MR0838085

[7] Jacod, J., J Mémin, M. Métivier. 1983. On tightness and stopping times. *Stochastic Proc. Appl.* 14, 109–146. MR0679668

[8] Jacod, J. and A. N. Shiryayev. 1987. *Limit Theorems for Stochastic Processes*, Springer. MR0959133

[9] Karatzas, I, and S. Shreve. 1988. *Brownian Motion and Stochastic Calculus*, Springer. MR0917065

[10] Krichagina, E. V. and A. A. Puhalskii. 1997. A heavy-traffic analysis of a closed queueing system with a GI/∞ service center. *Queueing Systems* 25, 235–280. MR1458591

[11] Kunita, H. and S. Watanabe. 1967. On square-integrble martingales. *Nagoya Math. J.* 30, 209–245. MR0217856

[12] Kurtz, T. 1975. Semigoups of conditioned shifts and approximation of Markov processes. *Ann. Probab.* 3, 618–642. MR0383544

[13] Kurtz, T. 1981. *Approximation of Population Processes*, SIAM, Philadelphia, PA. MR0610982

[14] Lévy, P. 1948. *Processus Stochastiques et Mouvement Borownien*, Gauthiers-Villars, Paris.

[15] Liptser, R. Sh. and A. N. Shiryayev. 1989. *Theory of Martingales*, Kluwer.

[16] McLeish, D. L. 1974. Dependent central limit theorems and invariance principles. *Ann. Probab.* 2, 620–628. MR0358933

[17] Pang, G, R. Talreja and W. Whitt. 2007. Martingale proofs of many-server heavy-traffic limits for Markovian queues. *Probability Surveys*, to appear.

[18] Puhalskii, A. A. and M. I. Reiman. 2000. The mutliclass $GI/PH/N$ queue in the Halfin-Whitt regime. *Adv. Appl. Prob.* 32, 564-595. MR1778580

[19] Rebolledo, R. 1980. Central limit theorems for local martingales. *Zeitschrift Wahrscheinlichkeitstheorie verw. Gebiete* 51, 269–286. MR0566321

[20] Skorohod, A. V. 1956. Limit theorems for stochastic processes. *Prob. Theory Appl.* 1, 261–290. MR0084897

[21] Stroock, D. W. and S. R. S. Varadhan. 1979. *Multidimensional Diffusion Processes*, Springer, New York. MR0532498

[22] Whitt, W. 2002. *Stochastic-Process Limits*, Springer. MR1876437