The material in this paper has appeared in “Limit theorems for workload input models.”

**Limit theorems for workload input models**

Thomas G. Kurtz
Departments of Mathematics and Statistics
University of Wisconsin - Madison
Madison, WI 53706-1388
USA
February 6, 1996

**Abstract**

General models for workload input into service systems are considered. Scaling limit theorems appropriate for the formulation of fluid and heavy traffic approximations for systems driven by these inputs are given. Under appropriate assumptions, it is shown that fractional Brownian motion can be obtained as the limiting workload input process. Motivation for these results comes from data on communication network traffic exhibiting scaling properties similar to those for fractional Brownian motion.

1 **Discrete source models.**

We consider models for the input of work into a system from a large number of sources. Each source “turns on” at a random time and inputs work into the system for some period of time. Associated with each active period of a source is a cumulative input process, that is, a nondecreasing stochastic process $X$ such that $X(t)$ is the cumulative work input into the system during the first $t$ units of time during the active period. One representation of the process is as follows. Let $N(t)$ denote the number of source activations up to time $t$, and for the $i$th activation, let $X_i(s)$ denote the cumulative work input into the system during the first $s$ units of time that the source is on. We will refer to $N$ as the source activation process. The total work input into the system up to time $t$ is then given by

$$U(t) = \int_0^t X_{N(s)}(t-s) dN(s).$$

(1.1)

In some settings, it is useful to model the length of time $\tau_i$ that a source remains active separately from $X_i$. The total work input up to time $t$ then becomes

$$U(t) = \int_0^t X_{N(s)}(\tau_{N(s)} \wedge (t-s)) dN(s).$$

(1.2)
Of course these two approaches are equivalent since for the second model, we can always define \( \tilde{X}_i(s) = X_i(\tau_i \wedge s) \) and obtain a model of the first form with the same total work input. Note that many sources may be active at the same time, and we let \( L(t) \) denote the number of sources active at time \( t \), that is,

\[
L(t) = \int_0^t I_{[0, \tau_N(s) \wedge s]}(t - s) dN(s) .
\]

We assume that the active periods are identical in the sense that the \( (X_i, \tau_i) \) are i.i.d. with distribution \( \nu \) on \( D_{\mathbb{R}}[0, \infty) \times [0, \infty) \), although the methods we use can be extended to multiple source classes. Let \( \lambda : D_{\mathbb{R}}^3[0, \infty) \to D_{[0, \infty)^2} \) be nonanticipating in the sense that \( \lambda(v, t) = \lambda(v^t, t) \) for all \( t \) where \( v^t(s) = v(t \wedge s), \ v \in D_{\mathbb{R}}^3[0, \infty) \). Assume that \( N \) is a counting process with an intensity \( \lambda(N, U, L, \cdot) \), that is, \( N(t) - \int_0^t \lambda(N, U, L, s) ds \) is a martingale with respect to the filtration given by \( \mathcal{F}_t = \sigma(N(s), U(s), L(s) : s \leq t) \). In particular, \( \lambda(N, U, L, t) \) depends on values of \( N, U, \) and \( L \) only at times \( s \leq t \).

We can represent the process \( (N, U, L) \) as the solution of a stochastic equation involving a Poisson random measure. Recall that a Poisson random measure \( \xi \) with mean measure \( \eta \) on a measurable space \( (E, \mathcal{B}) \) is a random measure on \( \mathcal{B} \) such that for each \( \Gamma \in \mathcal{B} \), \( \xi(\Gamma) \) has a Poisson distribution with mean \( \eta(\Gamma) \) and for disjoint \( \Gamma_1, \Gamma_2, \ldots \), \( \xi(\Gamma_1), \xi(\Gamma_2), \ldots \) are independent.

Let \( \xi \) be a Poisson random measure on \( [0, \infty) \times D_{\mathbb{R}}^3 \) with mean measure \( \eta = m \times \nu \) (\( m \) being Lebesgue measure). \( \xi \) can be represented as

\[
\xi = \sum_{i=1}^{\infty} \delta_{(S_i, X_i, \tau_i)}
\]

where \( 0 < S_1 < S_2 < \cdots \) are the jump times of a unit Poisson process and the \( (X_i, \tau_i) \) are as above. Let \( N, U, \) and \( L \) satisfy the system of equations

\[
\begin{align*}
N(t) &= \xi(B(t)) \\
U(t) &= \int_{B(t)} u(r \wedge \gamma(t)) \xi(ds \times du \times dr) \quad (1.3) \\
L(t) &= N(t) - \xi(A(t))
\end{align*}
\]

where

\[
\begin{align*}
B(t) &= \{(s, u, r) : s \leq \int_0^t \lambda(N, U, L, z) dz\} \\
A(t) &= \{(s, u, r) : s \leq \int_0^{t-r} \lambda(N, U, L, z) dz\}
\end{align*}
\]

\[
\int_0^{\beta(t)} \lambda(N, U, L, s) ds = t
\]
\[ \gamma_t(s) = t - \beta(s). \]

Note that \( \beta(S_i) \) is the \( i \)-th activation time, that is, the \( i \)-th jump time of \( N(t) \); for \( S_i \leq t \), \( \gamma_t(S_i) \) is the elapsed time since the \( i \)-th activation; \( T_i = \beta(S_i) + \tau_i \) is the end of the \( i \)-th active period and satisfies \( \tau_i = \gamma_{T_i}(S_i) \). With these interpretations, \( X_\tau(\gamma_t(S_i)) \) is the work input by the \( i \)-th activated source up to time \( t \) (giving the equation for \( U \) in (1.3)), and \( \xi(A(t)) \) is the number of active periods that have completed at or before time \( t \) (giving the identity for \( L \) in (1.3)). It is easy to see that the solution of the above system exists and is unique up to \( \sigma_\infty = \inf \{ t : N(t) = \infty \} \). This representation is essentially that used in Kurtz (1983) and Solomon (1987) for population models and by Foley (1982) in the particular case of the \( M/G/\infty \) queue.

Example 1.1 Queue-dependent arrivals. Consider the simple model for workload processing in which work is processed at rate 1 as long as there is work in the system. The queued work is given by

\[ Q(t) = Q(0) + U(t) - t + \Lambda(t) \]

where \( \Lambda \) denotes that idle-time of the processor. Note that \( Q \) is uniquely determined by \( U \), so an arrival intensity \( \lambda(Q(t)) \) is of the form considered above.

Example 1.2 On/off sources. Let \( L_0 \) be a positive integer and set

\[ \lambda(N, U, L, t) = \lambda_0(L_0 - L(t)). \]

This model would correspond to \( L_0 \) fixed sources, each of which alternates between active and inactive periods. The active periods have lengths \( \tau \) as above, and the inactive periods are exponentially distributed with parameter \( \lambda_0 \).

Section 2 considers scaling limits of the above models that give deterministic “fluid approximations” and corresponding central limit theorems. The proofs of these results extend ideas developed in Kurtz (1983), Solomon (1987), and Garcia (1995). Section 3 considers models in which the activation times occur at a constant rate (that is, form a Poisson process). Extending the definition of the process to the time interval \(( -\infty, \infty)\), the input process has stationary increments. Under “heavy traffic” scaling, we obtain Gaussian limit processes having stationary increments. In Section 4, we show that fractional Brownian motion can be obtained as a limit under the heavy traffic scaling. Section 5 suggests two variations
of the model indicating how a type of control can be introduced and how models with a fixed number of regenerative sources can be developed. Finally, Section 6 is an appendix containing several technical lemmas.

Iglehart (1973) considers functional limit theorems for models of the form (1.1) in which $N$ is a renewal process. His results have substantial overlap with the results of Section 3. In particular, his proof of relative compactness is similar to that used here. Models of this form have also been considered by Klüppelberg and Mikosch (1995a,b) in the context of risk analysis, where $X$ is interpreted as the cumulative payout of an insurance claim. They assume that $N$ is a Poisson process. Their results also have substantial overlap with the results of Section 3, although the approach is quite different.

Our motivation comes from the analysis that has been carried out on network traffic data (see Willinger, Taqqu, Leland, and Wilson (1995)) which indicates that the traffic exhibits “long-term dependence” that is not consistent with simpler compound-Poisson models of workload input. In this context, Willinger, Taqqu, Sherman and Wilson (1995) obtain a limit theorem for an on/off source model similar to the results of Section 3. Their model is not covered by the class of models considered in detail here (except in the case of exponential off periods as in Example 1.2). Their model is included in a class of regenerative source models that can be studied using techniques similar to those developed here. These models are described briefly in Section 5, but the details of the convergence results will be given elsewhere.

2 Law of large numbers and central limit theorem.

A variety of limit theorems can be proved for solutions of systems of the above form based on the law of large numbers and the central limit theorem for the underlying Poisson random measure $\xi$. For example, suppose that $\lambda_n(v, t) = n\lambda(n^{-1}v, t)$ (where $v \in D_{\mathbb{R}}^3[0, \infty)$) and that $\xi_n$ is a Poisson random measure with mean measure $nm \times \nu$. Let $N_n$ be the activation process, $U_n$ the work input process, and $L_n$ the active source process, and define

$$X_n(t) = \frac{1}{n}N_n(t), \quad Y_n(t) = \frac{1}{n}U_n(t), \quad Z_n(t) = \frac{1}{n}L_n(t).$$

Then we can write

$$X_n(t) = \frac{1}{n}\xi_n(B_n(t))$$
$$Y_n(t) = \int_{B_n(t)} u(r \wedge \gamma_n^u(s)) \frac{1}{n}\xi_n(ds \times du \times dr)$$
$$Z_n(t) = X_n(t) - \frac{1}{n}\xi_n(A_n(t))$$

where

$$B_n(t) = \{(s, u, r) : s \leq \int_0^t \lambda(X_n, Y_n, Z_n, z)dz\}$$
$$A_n(t) = \{(s, u, r) : s \leq \int_0^{t-r} \lambda(X_n, Y_n, Z_n, z)dz\}$$
\[ \int_{0}^{\beta_n(s)} \lambda(X_n, Y_n, Z_n, z)dz = s \]

and

\[ \gamma^n_i(s) = t - \beta_n(s) . \]

Using the fact that \( \frac{1}{n} \xi_n(A) \rightarrow m \times \nu(A) \) in probability for each Borel set \( A \) and uniformity results of Stute (1976), under appropriate uniqueness conditions, it is not difficult to show that \((X_n, Y_n, Z_n)\) converges to the solution of the system

\[
\begin{align*}
X(t) &= \int_{0}^{t} \lambda(X, Y, Z, s)ds \\
Y(t) &= \int_{0}^{t} \mu(t - s)\lambda(X, Y, Z, s)ds \\
Z(t) &= \int_{0}^{t} \lambda(X, Y, Z, s)\nu_r(t - s, \infty)ds
\end{align*}
\]

(2.2)

where \( \mu(t) = \int_{D_{\mathbb{R}[0,\infty) \times [0,\infty)}} u(r \wedge t) \nu(du \times dr) = E[X_i(\tau_i \wedge t)] \) and \( \nu_r \) is the distribution of \( \tau_i \), so that \( \nu_r(t, \infty) = P\{\tau_i > t\} = \nu(D_{\mathbb{R}[0,\infty) \times (t, \infty)}) \). In particular, we have the following theorem.

**Theorem 2.1** Suppose that there exists a finite, nondecreasing function \( K \) such that

\[
|\lambda(x_1, y_1, z_1, t) - \lambda(x_2, y_2, z_2, t)| \\
\leq K(t) \sup_{s \leq t} (|x_1(s) - x_2(s)| + |y_1(s) - y_2(s)| + |z_1(s) - z_2(s)|).
\]

Then there exists a unique solution to (2.2) and

\[
\sup_{s \leq t}(|X_n(s) - X(s)| + |Y_n(s) - Y(s)| + |Z_n(s) - Z(s)|) \rightarrow 0
\]

in probability.

**Proof.** Observing, for example, that \( m \times \nu(B_n(t)) = \int_{0}^{t} \lambda(X_n, Y_n, Z_n, s)ds \), we can rewrite (2.1) as

\[
\begin{align*}
X_n(t) &= \frac{1}{n} \tilde{\xi}_n(B_n(t)) + \int_{0}^{t} \lambda(X_n, Y_n, Z_n, s)ds \\
Y_n(t) &= \int_{B_n(t)} u(r \wedge \gamma^n_i(s)) \frac{1}{n} \tilde{\xi}_n(ds \times du \times dr) \\
&\quad + \int_{0}^{t} \mu(t - s)\lambda(X_n, Y_n, Z_n, s)ds \\
Z_n(t) &= \frac{1}{n} \tilde{\xi}_n(B_n(t) - A_n(t)) + \int_{0}^{t} \nu_r(t - s, \infty)\lambda(X_n, Y_n, Z_n, s)ds,
\end{align*}
\]

where \( \tilde{\xi}_n(B) = \xi_n(B) - m \times \nu(B) \). The terms involving \( \tilde{\xi}_n \) go to zero uniformly in bounded time intervals, and the theorem follows by the Lipschitz assumption and Gronwall’s inequality. \( \square \)
To derive the corresponding central limit theorem, define \( \tilde{X}_n(t) = \sqrt{n}(X_n(t) - X(t)) \), \( \tilde{Y}_n(t) = \sqrt{n}(Y_n(t) - Y(t)) \), \( \tilde{Z}_n(t) = \sqrt{n}(Z_n(t) - Z(t)) \), and

\[
\Xi_n(A) = \frac{1}{\sqrt{n}}(\xi(A) - nm \times \nu(A)).
\]

Note that \( \Xi_n(A) \Rightarrow \Xi(A) \) where \( \Xi \) is Gaussian white noise with \( E[\Xi(A)\Xi(B)] = m \times \nu(A \cap B). \)

Then

\[
\tilde{X}_n(t) = \Xi_n(B_n(t)) + \int_0^t \sqrt{n}(\lambda(X_n,Y_n,Z_n, s) - \lambda(X,Y,Z, s))ds
\]
\[
\tilde{Y}_n(t) = \int_{B_n(t)} u(r \wedge \gamma^n_t(s))\Xi_n(ds \times du \times dr)
+ \int_0^t \mu(t-s)\sqrt{n}(\lambda(X_n,Y_n,Z_n, s) - \lambda(X,Y,Z, s))ds
\]
\[
\tilde{Z}_n(t) = \Xi_n(B_n(t) - A_n(t))
+ \int_0^t \nu_r(t-s, \infty)\sqrt{n}(\lambda(X_n,Y_n,Z_n, s) - \lambda(X,Y,Z, s))ds
\]

Assume that there exist \( K_1 \) and \( K_2 \) such that \( \lambda(v,t) \leq K_1 + K_2 \sup_{s \leq t} |v(s)| \) and that \( \lambda \) is differentiable in an appropriate sense. For example, we can assume that \( D\lambda : D_{\mathbb{R}}[0, \infty) \rightarrow D_{\mathbb{R}}[0, \infty) \) satisfies

\[
|\lambda(v + \epsilon \tilde{v}, t) - \lambda(v, t) - \epsilon D\lambda(v, \tilde{v}, t)| \leq \epsilon^2 K_3 \sup_{s \leq t} |\tilde{v}(s)|^2.
\]

In particular, if \( \lambda(v, t) = h(v(t)) \), then \( D\lambda(v, \tilde{v}, t) = \nabla h(v(t)) \cdot \tilde{v}(t) \).

**Theorem 2.2** Suppose that for each \( T > 0 \), \( \int (u(r \wedge T))^2 \nu(du \times dr) < \infty \) and there exists \( C > 0 \) such that for \( 0 \leq h \leq 1 \),

\[
\sup_{t \leq T} \int (u(r \wedge (t + h)) - u(r \wedge t))^2(1 + u(r \wedge t)^2) \nu(du \times dr) \leq Ch
\]

and

\[
\sup_{h \leq t \leq T} \int (u(r \wedge (t + h)) - u(r \wedge t))^2(u(r \wedge t) - u(r \wedge (t - h)))^2 \nu(du \times dr) \leq Ch^2.
\]

Suppose that \( \lambda \) is bounded and satisfies (2.6) and that \( D\lambda \) satisfies the Lipschitz condition

\[
|D\lambda(v_1, \tilde{v}_1, t) - D\lambda(v_2, \tilde{v}_2, t)| \leq K(t) \sup_{s \leq t}(|v_1(s) - v_2(s)| + |\tilde{v}_1(s) - \tilde{v}_2(s)|).
\]

Then \( (\tilde{X}_n, \tilde{Y}_n, \tilde{Z}_n) \Rightarrow (\tilde{X}, \tilde{Y}, \tilde{Z}) \) where

\[
\tilde{X}(t) = \Xi(B(t)) + \int_0^t D\lambda(X,Y,Z, \tilde{X}, \tilde{Y}, \tilde{Z}, z)dz
\]
\[
\tilde{Y}(t) = \int_{B(t)} u(r \wedge \gamma_t(s))\Xi(ds \times du \times dr)
+ \int_0^t \mu(t-z)D\lambda(X,Y,Z, \tilde{X}, \tilde{Y}, \tilde{Z}, z)dz
\]
\[
\tilde{Z}(t) = \Xi(B(t) - A(t)) + \int_0^t \nu(t-z, \infty)D\lambda(X,Y,Z, \tilde{X}, \tilde{Y}, \tilde{Z}, z)dz
\]
Proof. The proof is similar to the arguments in Kurtz (1983) and Solomon (1987). Note that (2.5 can be viewed as a system of integral equations driven by the random processes $R^1_n(t) = \Xi_n(B_n(t))$, $R^2_n(t) = \int_{B_n(t)} u(r \wedge \gamma^n_t(s))\Xi_n(ds \times du \times dr)$, and $R^3_n(t) = \Xi_n(B_n(t) - A_n(t))$.

By the Lipschitz assumptions on $D\lambda$ and Gronwall’s inequality, the result will follow by the continuous mapping theorem, provided we verify the functional convergence of $(R^1_n, R^2_n, R^3_n)$. Let $E = D\mathbb{P}[0, \infty) \times [0, \infty)$, $\Lambda_n(t) = \int_0^t \lambda(X_n,Y_n,Z_n,s)ds$, and $\Lambda(t) = \int_0^t \lambda(X,Y,Z,s)ds$. If we take $V(s,u,r) = I_{[0,\Lambda_n(t)]}(s) - I_{[0,\Lambda(t)]}(s)$, then

$$\Xi_n(B_n(t)) - \Xi_n(B(t)) = \int_{[0,t]} V(s-,u,r)\Xi_n(ds \times du \times dr).$$

and by (6.3),

$$E[(\Xi_n(B_n(t)) - \Xi_n(B(t)))^2] = E[m \times \nu(B_n(t) \triangle B(t))] = E[|\Lambda_n(t) - \Lambda(t)|].$$

We also have

$$E \left[ \left( \int_{B_n(t)} u(r \wedge \gamma^n_t(s))\Xi_n(ds \times du \times dr) - \int_{B(t)} u(r \wedge \gamma_t(s))\Xi_n(ds \times du \times dr) \right)^2 \right]$$

$$\leq \int_E u^2(r \wedge t)\nu(ds \times du \times dr)E[|\Lambda_n(t) - \Lambda(t)|]$$

$$+ E \left[ \int_{[0,\infty]} I_{[0,\Lambda(t)]}(s)\left( u(r \wedge \gamma^n_t(s)) - u(r \wedge \gamma_t(s)) \right)^2 m \times \nu(ds \times du \times dr) \right]$$

with a similar inequality holding for $\Xi_n(B_n(t) - A_n(t)) - \Xi_n(B(t) - A(t))$. The right sides of these inequalities go to zero by the law of large numbers, Theorem 2.1, and the convergence of the finite dimensional distributions for $(R^1_n, R^2_n, R^3_n)$ follows, that is, replace $B_n(t)$ $A_n(t)$ and $\gamma^n_t$ by their deterministic limits and apply the ordinary central limit theorem. (See Theorem 6.1.)

Since the limits are continuous, to obtain relative compactness of the sequence $\{(R^1_n, R^2_n, R^3_n)\}$, it is sufficient to check relative compactness of each component. We employ results of Chenčov (1956) (see Ethier and Kurtz (1986), Theorem 3.8.8) summarized in Theorem 6.2.

We restrict our attention to

$$R^2_n(t) = \int_{B_n(t)} u(r \wedge \gamma^n_t(s))\Xi_n(ds \times du \times dr)$$

$$= \int_{[0,t]} I_{[0,\Lambda_n(t)]}(s)u(r \wedge \gamma^n_t(s))\Xi_n(ds \times du \times dr).$$

The estimates in the other cases are similar. We apply Corollary 6.8 with $U(t) = R^2_n(t + h) - R^2_n(t)$, $V(t) = R^2_n(t) - R^2_n(t - h)$, and $C(X,Y,M) = c\sqrt{h}$, for appropriate $c$. Note that
\[ M = \Xi_n \quad \text{and} \quad \eta_k = n^{1-k/2} \nu, \quad \text{so} \]
\[
\int_0^t \int_0^1 (I_{[0,\Lambda_n(t+h)]}(s)u(r \wedge \gamma^n_{t+h}(s)) - I_{[0,\Lambda_n(t)]}(s)u(r \wedge \gamma^n_t(s)))^2 \nu(du \times dr)ds
\]
\[
\leq \int_0^t \int_0^1 I_{(\Lambda_n(t),\Lambda_n(t+h))}(s)u^2(r \wedge \gamma^n_{t+h}(s))\nu(du \times dr)ds
\]
\[
\quad + \int_0^t \int_0^1 I_{[0,\Lambda_n(t)]}(s)(u(r \wedge \gamma^n_t(s) - u(r \wedge \gamma^n_{t+h}(s)))^2 \nu(du \times dr)ds
\]
\[
\leq \int_0^t u^2(r \wedge t)\nu(du \times dr)(\Lambda_n(t+h) - \Lambda_n(t)) + tCh
\]
\[
\leq \left( \int_0^t u^2(r \wedge t)\nu(du \times dr)\|\lambda\|_\infty + tC \right) h
\]
gives (6.6) for \( i = 2 \) and \( j = 0 \). The calculation for \( i = 0 \) and \( j = 2 \) is the same. For \( i = j = 2 \), we have
\[
\int_0^t \int_0^1 (I_{[0,\Lambda_n(t+h)]}(s)u(r \wedge \gamma^n_{t+h}(s)) - I_{[0,\Lambda_n(t)]}(s)u(r \wedge \gamma^n_t(s)))^2
\]
\[
\otimes (I_{[0,\Lambda_n(t)]}(s)(u(r \wedge \gamma^n_t(s) - u(r \wedge \gamma^n_{t+h}(s))))^2 \nu(du \times dr)ds
\]
\[
\leq \int_0^t \int_0^1 I_{(\Lambda_n(t-h),\Lambda_n(t))}(s)(u(r \wedge \gamma^n_{t+h}(s) - u(r \wedge \gamma^n_t(s)))^2
\]
\[
\quad \otimes (u(r \wedge \gamma^n_t(s) - u(r \wedge \gamma^n_{t-h}(s)))^2 \nu(du \times dr)ds
\]
\[
\leq (\Lambda_n(t) - \Lambda_n(t-h))Ch + tCh^2
\]
\[
\leq (C\| \lambda \|_\infty + tC) h^2.
\]
The estimates for \( i = 1 \) and \( j = 2 \) and for \( i = 2 \) and \( j = 1 \) follow as in Remark 6.9, and Corollary 6.8 gives the inequalities
\[
E[(R^n_2(h) - R^n_2(0))^2] \leq c_0 h
\]
and
\[
E[(R^n_2(t+h) - R^n_2(t))^2(R^n_2(t) - R^n_2(t-h))^2] \leq c_0 h^2
\]
for some \( c_0 > 0 \) independent of \( n \). Similar estimates hold for \( R^n_1 \) and \( R^n_3 \), and Lemma 6.2 gives the relative compactness of \( \{R^n_1\} \). As noted, the asymptotic continuity then ensures the relative compactness of \( \{(R^n_1, R^n_2, R^n_3)\} \), and convergence of \( \{(\tilde{X}_n, \tilde{Y}_n, \tilde{Z}_n)\} \) follows by the continous mapping theorem and the uniqueness of the solution of the limiting system (2.7).

\[ \square \]

3 Models with stationary increments

We now assume that the arrival rate \( \lambda \) is a constant and that the input process has been running forever. Let \( \xi \) be the Poisson random measure on \( (-\infty, \infty) \times D_{\mathbb{R}}[0, \infty) \times [0, \infty) \)
with mean measure $\lambda m \times \nu$. For $t_1 < t_2$, let $U(t_1, t_2)$ be the work input into the system during the time interval $(t_1, t_2]$. Then, letting $E = D_{\mathbb{R}}[0, \infty) \times [0, \infty)$, we can write

$$U(t_1, t_2) = \int_{(\infty, t_1) \times E} (u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s))) \xi(ds \times du \times dr) + \int_{(t_1, t_2) \times E} u(r \wedge (t_2 - s)) \xi(ds \times du \times dr),$$

where the first term gives the work input into the system in the time interval $(t_1, t_2]$ by sources that activated at or before time $t_1$ and the second term is work input by sources that activate during the time interval $(t_1, t_2]$. Note that $U$ is additive in the sense that $U(t_1, t_3) = U(t_1, t_2) + U(t_2, t_3)$ for $t_1 < t_2 < t_3$, and the amount of work input into the system has stationary increments in the sense that the distribution of $U(t_1 + t, t_2 + t)$ does not depend on $t$. It follows that $E[U(t_1, t_2)] = \alpha(t_2 - t_1)$ where $\alpha = \lambda E_u(r) \nu (du \times dr)$. We can simplify notation if we define $u(s) = 0$ for $s < 0$, and we have

$$U(t_1, t_2) - \alpha(t_2 - t_1) = \int_{(\infty, t_2) \times E} (u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s))) \tilde{\xi}(ds \times du \times dr)$$

where $\tilde{\xi}(A) = \xi(A) - \lambda m \times \nu(A)$.

In a simple service model with a single server working at “rate” one and employing $U$ as the work input, queued work would satisfy

$$Q(t) = Q(0) + U(0, t) - t + \Lambda(t)$$

(3.1)

where $\Lambda$ measures the idle time of the server, and assuming the server works at maximal rate whenever there is work to be done,

$$\Lambda(t) = 0 \lor \sup_{s \leq t} (t - U(0, s) - Q(0)).$$

See, for example, Harrison (1985), Section 2.2. A heavy traffic limit for such a model involves rescaling the work measurements and the arrival and service rates under the assumption that the workload arrival and service rates are asymptotically the same. Letting $\lambda$ and $\nu$ (and hence $\alpha$) depend on $n$, (3.1) becomes

$$\frac{Q_n(t)}{\sigma_n} = \frac{Q_n(0)}{\sigma_n} + \frac{1}{\sigma_n} (U_n(0, t) - \alpha_n t) + \frac{\alpha_n - \beta_n}{\sigma_n} t + \frac{1}{\sigma_n} \Lambda_n(t),$$

and $\sigma_n^{-1} Q_n$ converges in distribution provided $\sigma_n^{-1} (\alpha_n - \beta_n)$ converges and $\sigma_n^{-1} (U_n(0, t) - \alpha_n t)$ converges in distribution.

Consequently, we consider the behavior of

$$V_n(t_1, t_2) = \sigma_n^{-1} (U_n(t_1, t_2) - \alpha_n (t_2 - t_1))$$

$$= \int_{(-\infty, t_2) \times E} (u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s))) \Xi_n(ds \times du \times dr)$$
where $\Xi_n(A) = \sigma_n^{-1}(\xi_n(A) - \lambda_n m \times \nu_n(A))$ and $\sigma_n$ and $\lambda_n$ tend to infinity. Note that $\text{Var}(\Xi_n(A)) = \sigma_n^{-2}\lambda_n m \times \nu_n(A)$. We assume that $\nu_n(D_{\mathbb{R}}[0, \infty) \times \{0\}) = 0$ and that
\[
\lim_{n \to \infty} \sigma_n^{-2}\lambda_n \nu_n = \zeta
\] (3.2)
in a “somewhat vague” topology on $\mathcal{M}(D_{\mathbb{R}}[0, \infty) \times (0, \infty))$, that is,
\[
\sigma_n^{-2}\lambda_n \int_{D_{\mathbb{R}}[0,\infty) \times (0,\infty)} h(u, r) \nu_n(du \times dr) \to \int_{D_{\mathbb{R}}[0,\infty) \times (0,\infty)} h(u, r)\zeta(du \times dr)
\] (3.3)
for all bounded continuous functions on $D_{\mathbb{R}}[0, \infty) \times (0, \infty)$ for which there exists an $r_0 > 0$ such that $h(u, r) = 0$ for all $r < r_0$. In addition, we assume that the measures on $E$ defined by
\[
\gamma_n(B) = \sigma_n^{-2}\lambda_n \int_B u(r)^2 \wedge 1 \nu_n(du \times dr)
\]
converge weakly to the measure
\[
\gamma(B) = \int_B u(r)^2 \wedge 1 \zeta(du \times dr).
\]

Formally, these assumptions imply that $V_n$ converges in distribution to $V$ defined by
\[
V(t_1, t_2) = \int_{(-\infty, t_2] \times E} (u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s))) W(ds \times du \times dr)
\] (3.4)
where $W$ is Gaussian white noise on $(-\infty, \infty) \times E$ corresponding to $m \times \zeta$. The variance of $V(t_1, t_2)$ is
\[
E[V(t_1, t_2)^2] = \int_{(-\infty, t_2] \times E} (u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s)))^2 ds \zeta(du \times dr)
\] (3.5)
again recalling our convention that $u(s) = 0$ for $s < 0$.

**Theorem 3.1** Let $\varphi$ be nonnegative and convex on $[0, \infty)$ with $\varphi(0) = 0$ and $\lim_{x \to \infty} \varphi(x)/x = \infty$. Suppose that for each $T > 0$,
\[
\sup_n \lambda_n \int_E \varphi(\sigma_n^{-2} u(r \wedge T)^2) \nu_n(du \times dr) < \infty,
\] (3.6)
\[
\lim_{t \to \infty} \sup_n \sigma_n^{-2}\lambda_n \int_{D_{\mathbb{R}}[0,\infty) \times \{t\} \times E} \int_t^t (u(r \wedge (T + s)) - u(r \wedge s))^2 ds \nu_n(du \times dr) = 0,
\] (3.7)
and there exists $C > 0$ such that for $0 \leq h \leq 1$,
\[
\sup_n \int_E \int_0^t (u(r \wedge (h + s)) - u(r \wedge s))^2 ds \sigma_n^{-2}\lambda_n \nu_n(du \times dr) \leq Ch
\] (3.8)
and
\[
\sup_n \int_E \int_0^t (u(r \wedge (s + h)) - u(r \wedge s))^2 \otimes (u(r \wedge s) - u(r \wedge (s - h)))^2 ds \sigma_n^{-4}\lambda_n \nu_n(du \times dr)
\]
\[
\leq Ch^2.
\] (3.9)

Suppose that the mapping $(s, u, r) \mapsto u(r \wedge s)$ from $[0, \infty) \times D_{\mathbb{R}}[0, \infty) \times [0, \infty)$ is continuous a.e. $m \times \zeta$. Then $V_n(0, \cdot) \Rightarrow V(0, \cdot)$ in the Skorohod topology on $D_{\mathbb{R}}[0, \infty)$.
Remark 3.2 For fixed \( t_1 \) and \( t_2 \), convergence in distribution of \( V_n(t_1, t_2) \) is essentially a central limit theorem for a “triangular array”. Conditions (3.6) and (3.7) give the “uniform integrability” that such results require. (See Theorem 6.1.) In particular, if \( \lambda_n = n \), \( \sigma_n = \sqrt{n} \), and \( \nu_n \equiv \nu \), then \( \int u(r)^2 \nu(dr \times dr) < \infty \) implies (3.6) and (3.7) for appropriate choice of \( \varphi \). If, in addition, the input process has finite second moments and stationary, independent increments, the remaining conditions follow.

Proof. Let \( t_1 \leq t_2 \). Note that \( E[V_n(0, t_1) V_n(0, t_2)] = \int E \int (u(r \land (t_2 + s)) - u(r \land s))(u(r \land (t_1 + s)) - u(r \land s)) ds \sigma_n^2 \lambda_n \nu_n (du \times dr) \rightarrow \int E \int (u(r \land (t_2 + s)) - u(r \land s))(u(r \land (t_1 + s)) - u(r \land s)) ds \zeta (du \times dr) \) where the convergence follows from the convergence assumption (3.2) and the convergence of \( \gamma_n \) to \( \gamma \), the a.e. continuity assumption on the mapping \((s, u, r) \rightarrow u(r \land s)\), and the uniform integrability assumptions (3.6) and (3.7). The convergence of the finite dimensional distributions of \( V_n(0, \cdot) \) then follows from Theorem 6.1. The relative compactness of \( \{V_n(0, \cdot)\} \) follows from Theorem 6.2 by essentially the same argument as in Theorem 2.2. □

4 A representation of fractional Brownian motion.

Fractional Brownian motion is a mean-zero Gaussian process \( Z \) with stationary increments satisfying \( E[(Z(t_2) - Z(t_1))^2] = c(t_2 - t_1)^\alpha \) for some \( c > 0 \) and \( 1 < \alpha < 2 \). (Of course, if \( \alpha = 1 \), \( Z \) is an ordinary Brownian motion.) If we assume the source in the previous section broadcasts at rate 1, that is, \( X(s) \equiv s \), \( s \geq 0 \), and if we let \( W \) be Gaussian white noise on \((-\infty, \infty) \times [0, \infty)\) corresponding to

\[
ds \zeta(dr) = \lambda \frac{1}{r^\beta} ds \ dr,
\]

where \( 2 < \beta < 3 \), then (3.4) becomes

\[
V(t_1, t_2) = \int_{(-\infty, t_1] \times [0, \infty)} (r \land (t_2 - s) - r \land (t_1 - s)) W(ds \times dr) + \int_{(t_1, t_2] \times [0, \infty)} (r \land (t_2 - s)) W(ds \times dr)
\]

(4.1)

and (3.5) gives

\[
E[V(t_1, t_2)^2] = \int_{-\infty}^{t_1} \int_{0}^{\infty} (r \land (t_2 - s) - r \land (t_1 - s))^2 \lambda r^{-\beta} dr ds + \int_{t_1}^{t_2} \int_{0}^{\infty} (r \land (t_2 - s))^2 \lambda r^{-\beta} dr ds
\]

\[
= \int_{0}^{\infty} \int_{0}^{\infty} (r \land (t_2 - t_1 + u) - r \land u)^2 \lambda r^{-\beta} dr du + \int_{t_2 - t_1}^{t_2} \int_{0}^{\infty} (r \land u)^2 \lambda r^{-\beta} dr du
\]
\[
= \int_0^\infty \int_{t_2-t_1+u}^t (r-u)^2 \lambda r^{-\beta} dr du \\
+ \int_0^\infty \int_{t_2-t_1+u}^t (t_2-t_1)^2 \lambda r^{-\beta} dr du \\
+ \int_{t_2-t_1}^u \int_0^u \lambda r^2 dr du + \int_{t_2-t_1}^\infty u^2 \lambda r^{-\beta} dr du
\]

\[
= \left( \frac{\lambda}{(\beta-1)(4-\beta)} + \frac{\lambda}{(\beta-1)(\beta-2)} + \frac{\lambda}{(3-\beta)(4-\beta)} \right) (t_2-t_1)^{4-\beta}
\]

and hence \( V \) is fractional Brownian motion with \( \alpha = 4 - \beta \).

To see that \( \zeta \) can arise as a limit as in (3.2), let \( \lambda_n = n^\alpha \), with \( \alpha > \beta - 1 \), \( \nu_n(dr) = (\beta-1)n(nr+1)^{-\beta} dr \), and \( \sigma_n = n^{\frac{\alpha-\beta+1}{2}} \). It is easy to check that (3.2) holds and that \( \gamma_n \Rightarrow \gamma \).

For (3.6), take \( \varphi(x) = x^2 \). The integral on the left of (3.7) is bounded by

\[
\int_t^\infty T^2 (\beta-1)r^{-\beta+1} dr
\]

which is convergent since \( \beta > 2 \). Since \( r \land (h+s) - r \land s \leq r \land h \), the left side of (3.8) is bounded by

\[
(\beta-1) \int_0^h r^{-\beta+3} dr + h^2 (\beta-1) \int_h^\infty r^{-\beta+1} dr = \left( \frac{\beta-1}{4-\beta} + \frac{\beta-1}{\beta-2} \right) h^{4-\beta},
\]

which, since \( 4 - \beta > 1 \), implies the relative compactness in \( C_\mathbb{R}[0,\infty) \) by the Kolmogorov criterion without checking (3.9).

For a more complex example, let \( \nu \) be the distribution of \((X,\tau)\) where \( X \) has stationary, ergodic increments (for example, \( X \) can be a compound Poisson process) and \( \tau \) is independent of \( X \) and satisfies

\[
\lim_{r \to \infty} r^{\beta-1} P\{\tau > r\} = c,
\]

that is, \( \tau \) is in the domain of attraction of a stable law of index \( \beta - 1 \). Assume that \( \sup_{t \geq 1} E[\varphi(t^{-2}(X(t))^2)] < \infty \), where \( \varphi \) is as in Theorem 3.1 and that for some \( C > 0 \), \( Var(X(t)) < Ct \) for \( 0 \leq t \leq 1 \) and

\[
E[(X(t_3) - X(t_2))^2(X(t_2) - X(t_1))^2] \leq C(t_3 - t_2)(t_2 - t_1).
\]

Let \( \nu_n \) be the distribution of \((\frac{1}{n} X(n \cdot), \frac{1}{n} \tau)\). Note that this transformation corresponds to rescaling time and workload measurements and that \( \frac{1}{n} X(nt) \to mt \) a.s. and in \( L^2 \). Then with \( h \) as in (3.3) and \( \sigma_n^{-2} \lambda_n = n^{\beta-1} \),

\[
\lim_{n \to \infty} \sigma_n^{-2} \lambda_n \int h(u, r) \nu_n(du \times dr) = \lim_{n \to \infty} \sigma_n^{-2} \lambda_n E[h(\frac{1}{n} X(n \cdot), \frac{1}{n} \tau)]
\]

\[
= \int_0^\infty h(m, r) \frac{c(\beta-1)}{r^{\beta}} dr,
\]

and the estimates in Theorem 3.1 are easy to check.
5 Variations of the model.

5.1 Admission control.

The estimates used in proving relative compactness in the convergence theorems, Theorems 2.2 and 3.1, depend on the integrand being “predictable” with respect to the filtration generated by the Poisson random measures. Similar estimates will also hold for more general predictable integrands. For example, we can associate with each source another variable \( Z_i \), say with values in \([0, 1]\). Now let \( \xi \) be the Poisson random measure on \([0, \infty) \times D_R[0, \infty) \times [0, 1] \) with mean measure \( \lambda m \times \hat{\nu} \). Let \( U^R \) satisfy

\[
U^R(t_1, t_2) = \int_{(\xi)} \left( u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s)) \right)
\]

\[
\otimes I_{[0, R(s)]}(z)(ds \times du \times dr \times dz)
\]

where \( R(t) \in [0, 1] \) depends on the process up to time \( t \) (for example, on the number of active sources and/or the workload backlog). Note that if the \( i \)th (attempted) activation occurs at time \( S_i \), but \( R(S_i) < Z_i \), then the source contributes no work to the system. The introduction of \( R \) allows one to model an admission control policy. The value of \( Z_i \) can be thought of as specifying the priority level of the source. If \( R(s) = 1 \), any source that attempts to activate at time \( s \) is allowed to input work into the system. If \( R(s) = 0 \), any source that attempts to activate at time \( s \) is rejected. Otherwise, whether or not a source is accepted depends on its “priority level” \( Z_i \).

For simplicity, let \( \tilde{\nu}_n = \nu_n \times m \). As before, letting \( \tilde{R}_n = 1 - \frac{1}{\beta_n} \tilde{R}_n \) and setting \( \Xi_n(A) = \sigma_n^{-1}(\xi_n(A) - \lambda_n m \times \tilde{\nu}_n(A)) \), we have

\[
V^R_n(t_1, t_2) = \sigma_n^{-1}(U^R_n(t_1, t_2) - \alpha_n(t_2 - t_1))
\]

\[
= \int_{(\xi)} \left( u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s)) \right)
\]

\[
\otimes I_{[0, R_n(s)]}(z)\Xi_n(ds \times du \times dr \times dz)
\]

\[
- \int_{(\xi)} \left( u(r \wedge (t_2 - s)) - u(r \wedge (t_1 - s)) \right)
\]

\[
\otimes \tilde{R}_n(s) \frac{\lambda_n}{\beta_n \alpha_n} \nu_n(ds \times dr).
\]

For example, under the conditions of Remark 3.2, if \( \beta_n = \sqrt{n} \) and \( \tilde{R}_n \) is a continuous function of the normalized number of active sources, a limit theorem analogous to Theorem 3.1 will hold.

5.2 Regenerative sources.

As noted previously, the models considered above do not include many models of on/off sources. Suppose that there is a fixed number \( n \) of potential sources and that each of the \( n \) sources inputs work over a period of time of length distributed as \( \tau \) and that during that period the cumulative work is distributed as \( X \). We assume that at the end of each such
time period, the source immediately begins a new period with the same statistics and that the behavior in different periods is independent (hence the regenerative terminology). An on/off source that inputs work at a constant rate \( c \) during the on period satisfies
\[
X(t) = c(t \wedge \tau^{(1)}), \quad \tau = \tau^{(1)} + \tau^{(2)}
\]
where \( \tau^{(1)} \) is the length of the on (active) portion of the cycle and \( \tau^{(2)} \) is the length of the off (inactive) portion of the cycle.

Let \( \{(X_i, \tau_i)\} \) be iid with the same distribution as \((X, \tau)\). Define \( \xi_n([0, z] \times A) = \sum_{i=1}^{[nz]} \delta_{(X_i, \tau_i)} \). For \( k = 1, \ldots, n \), let \( \tau^0_k \) denote the first cycle completion (regeneration) time after time 0 for the \( k \)th source, and let \( X^0_k(t), 0 \leq t \leq \tau^0_k \), denote the cumulative work input by the \( k \)th source in the time interval \([0, t]\). Let
\[
U_n(t) = \sum_{k=1}^{n} X^0_k(t \wedge \tau^0_k) + \int_0^t X^0_{N_n(s)}(t \wedge (t - s))dN_n(s)
\]
and
\[
N_n(t) = \sum_{k=1}^{n} I_{[\tau^0_k, \infty)}(t) + \int_0^t I_{[\tau^0_{N_n(s)}(t), \infty)}(t - s)dN_n(s)
\]
where \( N_n(t) \) counts the total number of cycle completions in the time interval \([0, t]\) (each source may have completed more than one cycle) and \( U_n(t) \) denotes the total work input during the time interval \([0, t]\). \( U_n \) and \( N_n \) can be represented in terms of \( \xi_n \) and results analogous to Theorems 2.2 and 3.1 can be proved using the central limit theorem for \( \xi_n \). These results will be discussed elsewhere.

6 Appendix.

The central limit theorem for integrals against Poisson random measures plays an important role in the limit theorems considered in this paper. If \( \{\xi_n\} \) is a sequence of Poisson random measures on \( U \) with mean measures \( \eta^n, \bar{\xi}_n(A) = \xi_n(A) - \eta^n(A) \) and \( f^n \) is a sequence of functions on \( U \), then the characteristic function for
\[
Z^n = \int_U f^n(u)\bar{\xi}_n(du)
\]
is
\[
E[e^{i\theta Z^n}] = \exp\{\int_U (e^{i\theta f^n(u)} - 1 - i\theta f^n(u))\eta^n(du)\}
\]
and the following version of the Lindeberg-Feller central limit theorem holds.

**Theorem 6.1** For \( k = 1, \ldots, d, \ n = 1, 2, \ldots, \) let \( f^n_k \in L^2(\eta^n) \) and
\[
Z^n_k = \int_U f^n_k(u)\bar{\xi}_n(du).
\]
Suppose
\[ \lim_{n \to \infty} \int_U f^n_k(u) f^n_l(u) \eta^n(du) = \sigma_{kl}^2 \]
and for each \( \epsilon > 0 \) and \( k \),
\[ \lim_{n \to \infty} \int_U I_{\{|f^n(u)| > \epsilon\}}(f^n_k(u))^2 \eta^n(du) = 0. \]
Then \( Z^n \Rightarrow Z \), where \( Z \) is normal with mean zero and covariance matrix \( \sigma^2 = (\sigma_{kl}^2) \).

Because the processes of interest in Sections 2 and 3 involve stochastic convolutions, it is difficult or impossible to employ standard martingale estimates in order to verify relative compactness in the proofs of convergence in distribution. Consequently, we exploit classical results of Chenčov (1956) (see Ethier and Kurtz (1986), Theorem 3.8.8) which involve estimating moments of products of increments of the processes. The following simplified version of those results is sufficient for our purposes.

**Theorem 6.2** Let \( \{X_n\} \) be a sequence of cadlag \( \mathbb{R}^d \)-valued processes. Suppose that for each \( t \geq 0 \), \( \{X_n(t)\} \) is relatively compact (in the sense of convergence in distribution), that for each \( \epsilon > 0 \),
\[ \limsup_{t \to 0} P\{|X_n(t) - X_n(0)| > \epsilon\} = 0, \]
and that for each \( T > 0 \) there exists a \( C_T > 0 \) and \( \theta > 1 \) such that for \( 0 \leq h \leq 1 \) and \( h \leq t \leq T \)
\[ E[(X_n(t + h) - X_n(t))^2(X_n(t) - X_n(t - h))^2] \leq C_T h^\theta. \]
Then \( \{X_n\} \) is relatively compact in the sense of convergence in distribution in the Skorohod topology.

**Remark 6.3** Note that to verify the relative compactness of \( \{X_n(t)\} \), it is sufficient to verify \( \sup_n E[|X_n(t)|^a] < \infty \) for some \( a > 0 \), and to verify (6.1), it is sufficient to show that \( \lim_{t \to 0} \sup_n E[|X_n(t) - X_n(0)|^a] = 0. \)

The necessary moment estimates can be obtained using results on integrals against orthogonal martingale random measures. Let \( E \) be a complete, separable metric space with Borel sets \( \mathcal{B}(E) \), and let \( \mathcal{U} \subset \mathcal{B}(E) \) be closed under finite unions and satisfy \( A \in \mathcal{U} \) and \( B \in \mathcal{B}(E) \) implies \( A \cap B \in \mathcal{U} \). \( M \) is an \( \{\mathcal{F}_t\} \)-orthogonal martingale random measure on \( E \) if for each \( A \in \mathcal{U}, \ M(A, \cdot) \) is an \( \{\mathcal{F}_t\} \)-martingale and for disjoint \( A_1, \ldots, A_m \in \mathcal{U}, \ M(A_1, \cdot), \ldots, M(A_m, \cdot) \) are orthogonal martingales (that is, the product of any two is a martingale) and \( \sum_{i=1}^m M(A_i, t) = M(\cup_{i=1}^m A_i, t) \) a.s. If \( M(A, \cdot) \) is square integrable for each \( A \in \mathcal{U} \), then there exists a a random measure \( \tilde{\eta} \) on \( E \times [0, \infty) \) such that the Meyer process for \( M(A, \cdot) \) and \( M(B, \cdot) \) is given by \( \langle M(A, \cdot), M(B, \cdot) \rangle_t = \tilde{\eta}(A \cap B \times [0, t]) \). For simplicity, we assume that
\[ \tilde{\eta}(A \times [0, t]) = \int_0^t \eta(s, A) ds. \]
where \( \eta \) is an \( \mathcal{M}(E) \)-valued process adapted to \( \{ \mathcal{F}_t \} \). Note that if 

\[
M(A, t) = \frac{N(A \times [0, t]) - \nu(A)t}{\sigma}
\]

where \( N \) is a Poisson random measure on \( E \times [0, \infty) \) with mean measure \( \nu \times m \), then \( M \) is an orthogonal martingale random measure with \( \mathcal{U} = \{ A \in \mathcal{B}(E) : \nu(A) < \infty \} \) and \( \tilde{\eta} = \sigma^{-2} \nu \times m \), that is, \( \eta(s, A) = \sigma^{-2} \nu(A) \), and if \( W \) is Gaussian white noise on \( E \times [0, \infty) \) with \( \text{Var}(W(C)) = \nu \times m(C) \), then \( M(A, t) = W(A \times [0, t]) \) is an orthogonal martingale random measure with \( \mathcal{U} = \{ A \in \mathcal{B}(E) : \nu(A) < \infty \} \) and \( \tilde{\eta} = \nu \times m \). For details on integration against orthogonal martingale random measures see Walsh (1986) or Kurtz and Protter (1995).

**Lemma 6.4** Let \( \eta \) be a deterministic measure on \( E \) and \( \mathcal{U} = \{ A \in \mathcal{B}(E) : \eta(A) < \infty \} \), and let \( M \) be an orthogonal martingale random measure on \( E \) adapted to \( \{ \mathcal{F}_t \} \) such that \( M(A, \cdot) \) is square integrable for each \( A \in \mathcal{U} \) and \( \langle M(A, \cdot), M(B, \cdot) \rangle_t = \eta(A \cap B)t \), \( A, B \in \mathcal{U} \). For \( l > 2 \), let \( [M(A, \cdot)]_t^l = \sum_{s \leq t} (M(A, s) - M(A, s-))^l \) (that is, the sum is over all discontinuities) and suppose that for \( 2 < l \leq l_0 \), there exist measures \( \eta_l \) such that

\[
[M(A, \cdot)]_t^l - \eta_l(A)t
\]

is a martingale, and define \( \eta_2 \equiv \eta \). (Recall that for any \( A \) with \( \eta(A) < \infty \), \( [M(A, \cdot)]_t - \eta(A)t \) is a martingale.)

Fix \( m \), \( 2 \leq m \leq l_0 \), and let \( X \) and \( Y \) be cadlag, \( L^2(\eta) \)-valued processes such that

\[
E\left[ \left( \int_0^t \int_E (|X(u, s)|^k + |Y(u, s)|^k)\eta_k(du)ds \right)^m \right] < \infty \quad (6.2)
\]

for \( 2 \leq k \leq m \). Define

\[
U(t) = \int_{E \times [0, t]} X(u, s-)M(du \times ds), \quad V(t) = \int_{E \times [0, t]} Y(u, s-)M(du \times ds).
\]

Then for \( k + l \leq m \)

\[
E[U^k(t)V^l(t)] = \sum_{i+j \leq k, j \leq l} \binom{k}{i} \binom{l}{j} E\left[ \int_{E \times [0, t]} U^{k-i}(s-)V^{l-j}(s-)X^i(u, s-) \otimes Y^j(u, s-)\eta_{i+j}(du)ds \right].
\]

In particular,

\[
E[U(t)V(t)] = E\left[ \int_{E \times [0, t]} X(u, s-)Y(u, s-)\eta(du)ds \right] \quad (6.3)
\]

**Remark 6.5** If \( M \) is a centered Poisson process, then \( \eta_k = \eta \) for all \( k \geq 2 \). If \( M(A, \cdot) \) is continuous for all \( A \in \mathcal{U} \), then \( \eta_k = 0 \) for \( k > 2 \).
Proof. $U$ and $V$ are martingales, and Itô’s formula implies

$$U^k(t)V^l(t) = \int_0^t kU^{k-1}(s-)V^l(s-)dU(s) + \int_0^t lU^k(s-)V^{l-1}(s-)dV(s) + \frac{1}{2} \int_0^t k(k-1)U^{k-2}(s-)V^l(s-)d[U]_s$$

$$+ \int_0^t klU^{k-1}(s-)V^{l-2}(s-)d[U,V]_s$$

$$+ \int_0^t l(l-1)U^k(s-)V^{l-2}(s-)d[V]_s$$

$$+ \sum_{i\leq k\leq l\leq j} \binom{k}{i} \binom{l}{j} \sum_{s\leq t} U^{k-i}(s-)V^{l-j}(s-)(\Delta U(s))^i(\Delta V(s))^j,$$

(6.4)

where $\Delta U(s) = U(s) - U(s-)$. For $2 < i + j \leq m$, the orthogonality and the moment assumptions imply

$$\sum_{s\leq t} (\Delta U(s))^i(\Delta V(s))^j - \int_{E\times[0,t]} X^i(u, s-)Y^j(u, s-)\eta_{i+j}(du)ds$$

is a martingale as are

$$[U]_t - \int_{E\times[0,t]} X^2(u, s-)\eta(du)ds, \quad [U, V]_t - \int_{E\times[0,t]} X(u, s-)Y(u, s-)\eta(du)ds,$$

and the analogous centering of $[V]_t$. The desired identity then follows by taking expectations of both side of (6.4). For the moment estimates needed to justify this argument, see Section 2 of Kurtz and Protter (1995).

□

Corollary 6.6 Let $k$ be even, and let

$$C(X, M) = \max_{2 \leq i \leq k} \mathbb{E}\left[\left(\int_0^t |\int_E X^i(s-, u)\eta_i(du)|ds\right)^{\frac{k}{i}}\right],$$

and let $K$ be the (unique) positive solution of

$$K = \sum_{i=2}^k \binom{k}{i} K^{\frac{k-i}{i}}.$$

Then

$$\mathbb{E}[U^k(t)] \leq KC(X, M).$$

(6.5)

Remark 6.7 For a similar estimate on $\mathbb{E}[|U(t)|^k]$ for $k$ odd, see Section 2 of Kurtz and Protter (1995).
Remark 6.9 Note that

\[ E[U^k(t)] \leq \sum_{2 \leq i \leq k} \binom{k}{i} E[|U^{k-i}(t)|] \int_0^t \int_E X^i(s, u) \eta_i(du) |ds| \]

\[ \leq \sum_{2 \leq i \leq k} \binom{k}{i} E[|U^k(t)|] \frac{k}{k-i} E[\left( \int_0^t \int_E X^i(s, u) \eta_i(du) |ds| \right)^{\frac{k}{k-i}}]^{\frac{k}{i}} \]

\[ \leq \sum_{2 \leq i \leq k} \binom{k}{i} E[|U^k(t)|] \frac{k}{k-i} C(X, M)^{\frac{k}{i}}. \]

Dividing both sides of this inequality by \( C(X, M) \) we see that (6.5) must hold. \( \square \)

Corollary 6.8 Suppose in addition to (6.2), for \( i + j \geq 2, i, j \leq 2, \)

\[ \int_0^t \int_E X^i(s, -) Y^j(s, -) \eta_{i+j}(du) |ds| \leq C^{i+j}(X, Y, M) \quad \text{a.s.} \]  \( (6.6) \)

Then \( E[|U(t)|], E[|V(t)|] \leq C(X, Y, M), E[U^2(t)], E[V^2(t)] \leq C^2(X, Y, M) \) and

\[ E[U^2(t)V^2(t)] \leq 11C^4(X, Y, M) \]

Remark 6.9 Note that

\[ \int_0^t \int_E X^2(s, -) Y^1(s, -) \eta_3(du) |ds| \]

\[ \leq \int_0^t \int_E X^2(s, -) \eta_3(du) \int_E X^2(s - u) Y^2(s, -) \eta_3(du) |ds| \]

\[ \leq \sqrt{\int_0^t \int_E X^2(s, -) \eta_3(du) |ds|} \sqrt{\int_0^t \int_E X^2(s - u) Y^2(s, -) \eta_3(du) |ds|} \]

so, for example, in the Poisson case, the estimates for \( (i, j) = (2, 0), (0, 2) \) and \( (2, 2) \) imply the estimates for \( (2, 1) \) and \( (1, 2) \).

Proof. We have

\[ E[U^2(t)] = E\left[ \int_{E \times [0, t]} X^2(s, -) \eta_2(du) |ds| \right] \leq C^2(X, Y, M), \]

and similarly for \( E[V^2(t)] \), which in turn implies the inequalities for \( E[|U(t)|] \) and \( E[|V(t)|] \). Using the fact that \( |U(t)|, |V(t)|, U^2(t) \) and \( V^2(t) \) are submartingales and that \( |U(s)V(s)| \leq \)
\( \frac{1}{2}(U^2(s) + V^2(s)) \), we have

\[
E[U^2(t)V^2(t)] = \sum_{i \leq j \leq 2} \binom{2}{i} \binom{2}{j} E[\int_{E \times [0,t]} U^{2-i}(s-)V^{2-j}(s-) \otimes X^i(s-, u)Y^j(s-, u)\eta_{i+j}(du)ds] .
\]

\[
\leq \sum_{j=0}^{2} \binom{2}{j} E[|V^{2-j}(t)|] \int_{0}^{t} |\int_{E} X^2(s-, u)Y^j(s-, u)\eta_{2+j}(du)|ds
\]

\[
+ \sum_{i=0}^{1} \binom{2}{i} E[|U^{2-i}(t)|] \int_{0}^{t} |\int_{E} X^i(s-, u)Y^2(s-, u)\eta_{i+2}(du)|ds
\]

\[
+ 2E[(U^2(t) + V^2(t))] \int_{0}^{t} |\int_{E} X^1(s-, u)Y^1(s-, u)\eta_{2}(du)|ds
\]

\[
\leq 11C^4(X, Y, M).
\]

\[\square\]

**Corollary 6.10** If \( XY = 0 \) and

\[
\max\{\int_{0}^{t} |\int_{E} X^2(s, u)\eta(du)|ds, \int_{0}^{t} |\int_{E} Y^2(s, u)\eta(du)|ds\} \leq \tilde{C}(X, Y, M)
\]

then

\[
E[U^2(t)V^2(t)] \leq 2\tilde{C}^2(X, Y, M).
\]

**Proof.** Note that \( E[U^2(t)], E[V^2(t)] \leq \tilde{C}(X, Y, M) \). By Lemma 6.4 and the argument used in the proof above

\[
E[U^2(t)V^2(t)] = E[\int_{E \times [0,t]} U^2(s-)Y^2(s-, u)\eta(du)ds]
\]

\[
+ E[\int_{E \times [0,t]} V^2(s-)X^2(s-, u)\eta(du)ds]
\]

\[
\leq E[U^2(t)]\tilde{C}(X, Y, M) + E[V^2(t)]\tilde{C}(X, Y, M)
\]

which gives the desired result. \[\square\]

**References**

Chenčov, N. N. (1956). Weak convergence of stochastic processes whose trajectories have no discontinuities of the second kind and the heuristic approach to the Kolmogorov-Smirnov tests. *Theory Probab. Appl.* 1, 140-149.


20