Supplementary Material: Proofs

To prove Theorem 2.1, we need the following lemma.

**Lemma A.1.** Under Assumptions 1 and 2,

\[
E(\hat{\lambda}_h(t_0)|\lambda(\cdot)) = \int_{-b}^{b} f(r)\lambda(rh + t_0)dr,
\]

\[
\text{var}(\hat{\lambda}_h(t_0)|\lambda(\cdot)) = \frac{1}{h} \int_{-b}^{b} f^2(r)\lambda(rh + t_0)dr.
\]

**Proof of Lemma A.1.** The following fact of inhomogeneous Poisson processes (Daley and Vere-Jones 1988) is very useful:

**Result A.** For a Poisson process with deterministic inhomogeneous arrival rate \( \lambda(t) \), the total number of points \( K \) arriving within the time interval \([0, T] \) has a Poisson distribution with parameter \( \int_0^T \lambda(t)dt \), and conditioning on \( K \), the arrival times \( s_i, i = 1, 2, \ldots, K \), have the same joint distribution as the order statistics \( X_{(1)}, \ldots, X_{(K)} \), where \( X_i \) are i.i.d. with the density \( \lambda(t)/\int_0^T \lambda(s)ds \).

Using Result A and the law of iterated expectation, we have

\[
E(\hat{\lambda}_h(t_0)|\lambda(\cdot)) = E\{E\left(\hat{\lambda}_h(t_0)|\lambda(\cdot), K\right)|\lambda(\cdot)\} = E\left[E\left(\sum_{i=1}^{K} f_h(t_0, s_i)|\lambda(\cdot), K\right)\right]|\lambda(\cdot)
\]

\[
= \sum_{k=0}^{+\infty} \frac{1}{k!} \left( \int_0^T \lambda(s) \right)^k \exp\left[ -\int_0^T \lambda(s)ds \right] \times \frac{k \int_0^T f_h(t_0, s)\lambda(s)ds}{\int_0^T \lambda(s)ds}.
\]

\[
= \int_0^T f_h(t_0, s)\lambda(s)ds = \int_{-b}^{b} f(r)\lambda(rh + t_0)dr
\]

where \( f_h(t, s) \) is defined in (2.1) and a change of variable \( r = (s - t_0)/h \) is used. Similarly, the law of conditional variance gives

\[
\text{var}(\hat{\lambda}_h(t_0)|\lambda(\cdot)) = E\{\text{var}\left(\sum_{i=0}^{K} f_h(t_0, s_i)|\lambda(\cdot), K\right)|\lambda(\cdot)\} + \text{var}\left\{E\left[\sum_{i=0}^{K} f_h(t_0, s_i)|\lambda(\cdot), K\right]\right\}|\lambda(\cdot)
\]

\[
= E\{K \cdot \text{var}(f_h(t_0, s)|\lambda(\cdot), K)|\lambda(\cdot)\} + \text{var}\left\{K \int_0^T f_h(t_0, s)\lambda(s)ds / \int_0^T \lambda(s)ds|\lambda(\cdot)\right\}.
\]

Since \( \text{var}(K|\lambda(\cdot)) = E(K|\lambda(\cdot)) = \int_0^T \lambda(s)ds \) and

\[
\text{var}(f_h(t_0, s)|\lambda(\cdot), K) = \frac{\int_0^T f_h^2(t_0, s)\lambda(s)ds}{\int_0^T \lambda(s)ds} - \left( \frac{\int_0^T f_h(t_0, s)\lambda(s)ds}{\int_0^T \lambda(s)ds} \right)^2,
\]

1
we simplify $\text{var}(\hat{\lambda}_h(t_0)|\lambda(\cdot))$ to

$$\text{var}(\hat{\lambda}_h(t_0)|\lambda(\cdot)) = \int_0^T f_h^2(t, s)\lambda(s)ds = \frac{1}{h} \int_{-b}^b f^2(r)\lambda(\gamma(rh + t))dr. \quad \square$$

**Proof of Theorem 2.1.** The MSE of $\hat{\lambda}_h(t)$ has the decomposition

$$E(\hat{\lambda}_h(t_0) - \lambda(t_0))^2 = E \left\{ \text{var}[\hat{\lambda}_h(t_0)|\lambda(\cdot)] + \left[ E(\hat{\lambda}_h(t_0)|\lambda(\cdot)) - \lambda(t_0) \right]^2 \right\}, \quad (A.1)$$

where $\lambda(\cdot)$ denotes the full realization of the stochastic arrival rate. Using Lemma A.1, we obtain

$$E\{ (\hat{\lambda}_h(t_0) - \lambda(t_0))^2 \} = E[ \int_{-b}^b f(r)\lambda(\gamma(rh + t_0))dr - \lambda(t_0)]^2 + \frac{E[\lambda(0)]}{h} \int_{-b}^b f^2(r)dr.$$

The expectation of the first term equals

$$\int_{-b}^b \int_{-b}^b f(r_1)f(r_2)C(|r_1 - r_2|h)dr_1dr_2 - 2 \int_{-b}^b f(r)C(|r|h)dr + C(0).$$

A Taylor expansion on $C(t)$ around $0^+$ simplifies the above expression to $hC'(0^+)\gamma_f + o(h)$ where $\gamma_f$ is defined in (2.4). This leads to (2.3), from which it is easily seen that the $h$ that minimize the MSE is $h_{opt}$ defined in (2.5). \quad \square

**Proof of Lemma 3.1.** Using Result A and following the proof of Lemma A.1, it is straightforward to establish (3.2). The continuity assumption of $C((r - m)h + t)$ for $t \geq 2bh$ and around 0 for $t \leq 2bh$ yields the subsequent results. \quad \square

**Proof of Lemma 3.2.** First, consider $\text{var}\{ E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) \}$. Using Result A, it is straightforward to show that $E(\hat{\lambda}_h(s + t)\hat{\lambda}_h(s)|\lambda(\cdot)) = \frac{1}{h^2} \int_0^T dl \int_0^T f(\frac{l-s-t}{h})f(\frac{m-s}{h})\lambda(l)\lambda(m)dm + \frac{1}{T} \int_0^T f(\frac{l-s-t}{h})f(\frac{m-s}{h})\lambda(l)dl$, which implies that

$$E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) = \frac{1}{T-2bh-t} \left[ \int_{bh}^{T-bh-t} ds \int_0^T \lambda(l)f(\frac{l-s-t}{h})f(\frac{m-s}{h})dl + \int_{bh}^{T-bh-t} ds \int_0^T dl \int_0^T [\lambda(l) - \mu][\lambda(m) - \mu]f(\frac{l-s-t}{h})f(\frac{m-s}{h}) dm \right].$$

A term-by-term expansion and four change-of-variables give the detailed expression (3.4). Next, we note that the three terms $A_{i,T}^h, B_{i,T}^h$ and $C_{i,T}^h$ are averages of $C, C_3$ and $v$ over $[bh, T - bh - t]^2$. Assumptions 4 and 5 thus guarantee that $\text{var}(E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) \to 0$ as $T \to \infty$. 

2
Now consider $E\{\var(C_{\mu,h}(t)|\lambda(\cdot))\}$. We will demonstrate that in equation (3.5)

\[D_t^h = \text{Int}_{f:[-b,b]^4}^h(E\{\lambda(0)\lambda(t + (l - m)h)\lambda((l - m - l' + m')h)\}) + \text{Int}_{f:[-b,b]^4}^h(E\{\lambda(0)\lambda(t + (l - m)h)\lambda(t + (l' - m')h)\}) + 2 \text{Int}_{f:[-b,b]^4}^h(E\{\lambda(0)\lambda(t + (l - m)h)\lambda(2t + (l - m + l' - m')h)\}) \]

\[-8\mu \int_{[b,-b]^2}^2 C(t + (l - m)h)f(l)f(m)dl \, dm - 4\mu^3, \quad (A.2)\]

\[E_t^h = \int_{-2b}^{2b}^2 dr \int_{[b,-b]^2}^2 [C(t + (l - m)h) + \mu^2]f(l)f(l + r)f(m)f(m + r)dl \, dm + 4 \text{Int}_{f:[-b,b]^3}^h(C(t + (l' - m')h)f(l + \frac{t}{h})) + \text{Int}_{f:[-b,b]^3}^h(C(t + (l - m)h) + \mu^2)f(l - m + m' + \frac{2t}{h}), \quad (A.3)\]

\[F_t^h \quad \mu \int_{[b,-b]^2}^2 f(l)f(l + \frac{t}{h})f(m)f(m + \frac{t}{h})dl \, dm. \quad (A.4)\]

Here for notational ease we have used the short hand notation $\text{Int}_{f:A}^{l,m,l',m'}(\cdot)$, which stands for the integral over region $A$ with respect to kernel $f$. The superscript refers to the dummy variables in the integral. For example, $\text{Int}_{f:[-b,b]^3}^{l,l',m'}(g(l, l', m')) = \int\int\int\int_{[b,-b]^3} g(l, l', m')f(l)f(l')f(m')dl \, dl' \, dm'$ and $\text{Int}_{f:[-b,b]^4}^{l,m,l',m'}(g(l, m, l', m')) = \int\int\int\int\int_{[b,-b]^4} g(l, m, l', m')f(l)f(m)f(l')f(m')dl \, dl' \, dm \, dl' \, dm'$.

To establish these expressions, we first note that

\[\var(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) = \frac{1}{(T - 2bh - t)^2} \int_{[bh,T-bh-t]^2} \text{cov}\{(\hat{\lambda}_h(s + t) - \mu)(\hat{\lambda}_h(s) - \mu)\} ds \, ds'. \quad (A.5)\]

A term-by-term expansion on $\text{cov}\{(\hat{\lambda}_h(s + t) - \mu)(\hat{\lambda}_h(s) - \mu), (\hat{\lambda}_h(s' + t) - \mu)(\hat{\lambda}_h(s') - \mu)\lambda(\cdot)\}$ leads to a lengthy linear combination of $E\{\hat{\lambda}_h(s + t)\hat{\lambda}_h(s)\hat{\lambda}_h(s' + t)\hat{\lambda}_h(s')|\lambda(\cdot)\}$ and the conditional expectations of the product of three $\hat{\lambda}_h$ terms, two $\hat{\lambda}_h$ terms, and so on. The derivation, therefore, boils down to the calculation of $E\{\hat{\lambda}_h(s + t)\hat{\lambda}_h(s)\hat{\lambda}_h(s' + t)\hat{\lambda}_h(s')|\lambda(\cdot)\}$, $E\{\hat{\lambda}_h(s + t)\hat{\lambda}_h(s)\hat{\lambda}_h(s')|\lambda(\cdot)\}$, etc. By definition,

\[E\{\hat{\lambda}_h(s + t)\hat{\lambda}_h(s)\hat{\lambda}_h(s' + t)\hat{\lambda}_h(s')|\lambda(\cdot)\} = \frac{1}{h^2} \sum_{i=1}^k f\left(\frac{s_i - s - t}{h}\right) \sum_{j=1}^k f\left(\frac{s_j - s}{h}\right) \sum_{i' = 1}^k f\left(\frac{s_{i'} - s' - t}{h}\right) \sum_{j' = 1}^k f\left(\frac{s_{j'} - s'}{h}\right) |\lambda(\cdot)|.\]
Depending on whether \( i, j, i' \) and \( j' \) take on distinct values or not, we have

\[
E\{\hat{\lambda}_h(s + t)\hat{\lambda}_h(s')\hat{\lambda}_h(s' + t)|\lambda(\cdot)\} = \frac{1}{h^3} \left\{ \sum_{\text{all distinct}} + \sum_{\text{three distinct}, i = j} + \sum_{\text{three distinct}, i = i'} + \sum_{\text{three distinct}, i = j'} \right.
\]
\[
+ \sum_{\text{three distinct}, j = i'} + \sum_{\text{three distinct}, j = j'} + \sum_{\text{two distinct}, i = i', j = j'} + \sum_{\text{two distinct}, i = j', i' = j} \left. + \sum_{\text{two distinct}, i = i', j = j'} + \sum_{\text{two distinct}, i = j', i' = j'} \right\}
\]

Using Result A, we can straightforwardly simplify each term. For example, the third term

\[
\sum_{\text{three distinct}, i = i'} = \frac{1}{h} \int \int_{[-b,b]^3} \left\{ (\lambda(s + t + lh)\lambda(s + mh)\lambda(s' + m'h)f(l + \frac{s - s'}{h})) \right\}, \quad (A.6)
\]

and the term

\[
\sum_{\text{two distinct}, i = j', i' = j'} = \frac{1}{h^2} \int \int_{[-b,b]^2} \left\{ (\lambda(s + t + lh)\lambda(s + mh)f(m + \frac{s - s' - t}{h})f(l + \frac{s + t - s'}{h})) \right\}. \quad (A.7)
\]

Applying similar treatment (lengthy but straightforward algebra), we obtain the expressions for \( E\{\hat{\lambda}_h(s + t)\hat{\lambda}_h(s')\hat{\lambda}_h(s' + t)|\lambda(\cdot)\} \), \( E\{\hat{\lambda}_h(s + t)\hat{\lambda}_h(s')\hat{\lambda}_h(s')|\lambda(\cdot)\} \), etc. To save space, we omit the detailed long formulas, which are available upon request. Once the terms of \( \text{cov}\{(\hat{\lambda}_h(s + t) - \mu)(\hat{\lambda}_h(s) - \mu), (\hat{\lambda}_h(s' + t) - \mu)(\hat{\lambda}_h(s') - \mu)|\lambda(\cdot)\} \) are fully expanded and evaluated, we observe many cancellations. Finally, taking an expectation on (A.5) gives us the formula of \( E\{\text{var}(\hat{\lambda}_{\mu,h}(l)|\lambda(\cdot))\} \).

To see how the detailed expressions of \( D^h_l, E^h_l \) and \( F^h_l \) arise, let us use (A.6) and (A.7) as an example. Taking expectation on (A.6) and integrating \( s \) and \( s' \) over \([bh, T - bh - t]^2\) gives the expression

\[
\frac{1}{h} \int \int_{[bh, T - bh - t]^2} ds ds' \int \int_{[-b,b]^3} \left\{ (E\{\lambda(0)\lambda(t + (l - m)h)\lambda(s' - s + (m' - m)h)f(l + \frac{s - s'}{h})\}) \right\},
\]

A change of variable \( l' = l + (s - s')/h \) reduces it to

\[
(T - 2bh - t) \int \int_{[-b,b]^4} \left\{ (E\{\lambda(0)\lambda(t + (l - m)h)\lambda((l - m - l' + m')h)\}) \right\} + D_1, \quad (A.8)
\]
where

\[ D_1 = \int_{f_{[-b, b]}^I}^{t, m, m'} (E\{\lambda(0)\lambda(t + (l - m)h)\lambda((l - m - l' + m')h)\}) \cdot \left[ \int_{bh}^{3bh} ds \int_{-b}^{l+\frac{s-t}{h}} f(l') dl' + \int_{T-bh-t}^{T-bh} ds \int_{l+\frac{t-s}{h}}^{b} f(l') dl' - 4bh \right] \]

\[ = O(h). \]

Equation (A.8) corresponds to the first term in (A.2). Likewise, taking expectation on (A.7) and integrating \( s \) and \( s' \) over \([bh, T-bh-t]^2\) gives the expression

\[ \frac{1}{h^2} \int_{[bh, T-bh-t]^2} ds \int_{-[b, b]^2}^{T, m, m'} \left[ (C(t + (l - m)h) + \mu^2)f(l + \frac{s - s' - t}{h}) f(l + \frac{s + t - s'}{h}) \right]. \]

A change of variable \( m' = m + (s - s' - t)/h \) reduces it to

\[ \frac{T - 2bh - t}{h} \int_{f_{[-b, b]}^I}^{t, m, m'} (C(t + (l - m)h) + \mu^2)f(l + \frac{s - s' - t}{h}) f(l + \frac{s + t - s'}{h}) + O(h), \]

which corresponds to the third term in (A.3). All the other terms in (A.2), (A.3) and (A.4) arise via similar calculations. \( \square \)

Proof of Theorem 3.3. From Lemma 3.1, it is transparent that as \( h \to 0 \), the bias of \( \hat{\mu}(t) \) goes to zero for any fixed \( t > 0 \). For the variance, we note that for \( h \) sufficiently small such that \( t \geq 2bh \), \( F_t^h \) in Lemma 3.2 is identically zero (see (A.4)). Therefore, for \( h \) sufficiently small, the variance of \( \hat{\mu}(t) \) is a linear combination of \( O(t/h) \), \( O(1/h^2) \), \( O(1/h^3) \), \( O(T) \), \( O(1/T) \) and \( O(1/T^2) \) terms, each converging to zero as \( Th \to \infty \) and \( h \to 0 \). The desired result thus follows. \( \square \)

Proof of Theorem 3.4. Let us first consider \( E(\hat{\mu} - \mu)^2 \). The law of conditional variance says

\[ E(\hat{\mu} - \mu)^2 = \text{var}(\hat{\mu}) = E(\text{var}(\hat{\mu}|\lambda(\cdot))) + \text{var}(E(\hat{\mu}|\lambda(\cdot))) \]

\[ = \frac{\mu}{T} + \frac{1}{T^2} \int_0^T \int_0^T C(s - s') ds' ds \to 0, \text{ as } T \to \infty. \]

Thus, \( \hat{\mu} - \mu \to 0 \) in probability. Next, consider \( \int_{bh,T-bh-t}^{T-bh-t}(\hat{\lambda}_h(s+t) + \hat{\lambda}_h(s) - 2\mu) ds \). Using Lemma A.1, we know its expectation is zero. Its variance is equal to

\[ \int_{[bh, T-bh-t]^2} ds \int_{-[b, b]^2}^{T, m, m'} \text{cov}(\hat{\lambda}_h(s+t) + \hat{\lambda}_h(s) + \hat{\lambda}_h(s')) \]

\[ = 2 \int_{[bh, T-bh-t]^2} ds ds' \int_{-b, b}^{T, m, m'} C(s - s' + (l - m)h) f(l) f(m) dl dm \]

\[ + 2 \int_{[bh, T-bh-t]^2} ds ds' \int_{-b, b}^{T, m, m'} C(s + t - s' + (l - m)h) f(l) f(m) dl dm \]

\[ + 2 \int_{[bh, T-bh-t]^2} ds ds' \int_{-b}^{b} f(l) f(l + \frac{s - s'}{h}) + f(l + \frac{s + t - s'}{h}) dl. \]
Applying L’Hospital’s rule, we know that \( \text{var}\left(\frac{1}{T-2bh-t} \int_{bh}^{T-bh-t} (\dot{\lambda}_h(s+t) + \dot{\lambda}_h(s) - 2\mu)ds\right) \to 0 \), as \( T \to \infty \), so it, in particular, converges to zero in probability. Using these results together with Theorem 3.3, we know from

\[
\hat{C}_{\mu,h}(t) = \hat{C}_{\mu,h}(t) + (\hat{\mu} - \mu)^2 + \frac{\mu - \hat{\mu}}{T - 2bh - t} \int_{bh}^{T-bh-t} (\dot{\lambda}_h(s+t) + \dot{\lambda}_h(s) - 2\mu)ds.
\]

that \( \hat{C}_{\mu,h}(t) \to C(t) \) in probability. \( \square \)

**Proof of Proposition 3.5.** From the definition of \( A_{t,T}^h \) (Lemma 3.2), it is transparent that

\[
A_{t,T}^h = \mu^4 \int_{bh}^{T-bh-t} ds ds' \int_{[-\hat{b}, \hat{b}]^4} \text{cov}\{ (\lambda_0(mh + s) - 1)(\lambda_0(\hat{l}h + t + s) - 1), (\lambda_0(m'h + s') - 1)(\lambda_0(\hat{l}'h + t + s') - 1) \} \cdot f(l) f(m) f(l') f(m') dl dm dl' dm' = O(\mu^4),
\]

since the law of \( \{\lambda_0(t), t \in \mathbb{R}\} \) is fixed. On the other hand for fixed \( T, h \) and \( t \), it is easily seen that

\[
\text{var}(T - 2bh - t) \hat{C}_{\mu,h}(t)) - A_{t,T}^h = O(\mu^3),
\]

from which (3.7) follows. \( \square \)

To prove Theorem 4.2, we need the following two lemmas.

**Lemma A.2.** Suppose that Assumptions 1, 2, 4 and 5 hold and that the arrival rate \( \lambda(s), s \in \mathbb{R} \) is bounded. Then, for fixed \( t, h \geq 0 \), conditioning on the realization of \( \lambda \), with probability one we have

\[
\sqrt{T} \left[ \hat{C}_{\mu,h}(t) - E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) \right] \xrightarrow{D} N(0, \sigma^2_1(t, h)) \text{ as } T \to \infty,
\]

where the constant \( \sigma^2_1(t, h) = D_t^h + E_t^h/h + F_t^h/h^2 \), defined in Lemma 3.2 (see also (A.2), (A.3) and (A.4)), does not depend on \( \lambda \). Furthermore, for fixed \( t, h \geq 0 \),

\[
\sqrt{T}[\hat{C}_{\mu,h}(t) - E(\hat{C}_{\mu,h}(t))] \xrightarrow{D} N(0, \sigma^2(t, h)) \text{ as } T \to \infty,
\]

where \( \sigma^2(t, h) = \lim_{T \to \infty} T \text{var}(\hat{C}_{\mu,h}(t)) \), if and only if

\[
\sqrt{T}[E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) - E(\hat{C}_{\mu,h}(t))] \xrightarrow{D} N(0, \tau^2(t, h)) \text{ as } T \to \infty,
\]

where

\[
\tau^2(t, h) = \lim_{T \to \infty} T \text{var}(E(\hat{C}_{\mu,h}(t)|\lambda(\cdot))).
\]
Lemma A.2 indicates that the asymptotic distribution of \( \hat{C}_{\mu,h}(t) \) rests on the stochastic properties of the underlying arrival rate \( \{\lambda(t), t \in \mathbb{R}\} \).

**Lemma A.3.** Suppose that Assumptions 1, 2, 4, and 5 hold, and that the arrival rate process \( \{\lambda(t), t \in \mathbb{R}\} \) is finite \( \rho \)-mixing. Then for fixed \( h \) and \( t \geq 0 \), as \( T \to \infty \)

\[
\sqrt{T}[E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) - E(\hat{C}_{\mu,h}(t))] \xrightarrow{D} N(0, \tau^2(t, h)),
\]

where \( \tau^2(t, h) \) is defined in (A.11).

**Proof of Lemma A.2.** We first introduce some notations that will be used in the subsequent proofs. Let

\[
\begin{align*}
g_i(t, h) & = \int_{bh+i-1}^{bh+i} (\hat{\lambda}_h(s) - \mu)(\hat{\lambda}_h(s + t) - \mu)ds, \\
S_n(t, h) & = \sum_{i=1}^{n} g_i(t, h) - E(\sum_{i=1}^{n} g_i(t, h)) = \sum_{i=1}^{n} g_i(t, h) - nE(\hat{C}_{\mu,h}(t)), \\
v_i(t, h) & = E(g_i(t, h)|\lambda(\cdot)) - E(\hat{C}_{\mu,h}(t)), \\
G_n(t, h) & = \sum_{i=1}^{n} v_i(t, h).
\end{align*}
\]

Since we can write

\[
\frac{1}{\sqrt{T - 2bh - t}} \int_{T-bh-t}^{T-bh-t} \{(\hat{\lambda}_h(s) - \mu)(\hat{\lambda}_h(s + t) - \mu) - E(\hat{C}_{\mu,h}(t))\} ds, \tag{A.12}
\]

where \( n = [T - 2bh - t] \), and the reminder term (A.12) converges to zero in probability as \( T \to \infty \), the asymptotic distributions of \( \sqrt{T - 2bh - t}(\hat{C}_{\mu,h}(t) - E(\hat{C}_{\mu,h}(t))) \) and \( S_n(t, h)/\sqrt{n} \) are the same. Similarly, the asymptotic distributions of \( \sqrt{T - 2bh - t}[E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) - E(\hat{C}_{\mu,h}(t))] \) and \( G_n(t, h)/\sqrt{n} \) are the same. Thus, we can focus on the sequences of \( S_n(t, h)/\sqrt{n} \) and \( G_n(t, h)/\sqrt{n} \) in the proof.

Next, we need the following result from Corollary 4 of Herrndorff (1984).

Let \( \beta \in (2, \infty) \) and \( r = 2/\beta \). Assume that a sequence of random variables \( \{X_n\} \) is \( m \)-dependent and satisfies (i) \( E(X_n) = 0 \) and \( E(X_n^2) < \infty \), (ii) \( E(s_n^2/n) \to \sigma^2 > 0 \), where \( s_n = \sum_{i=1}^{n} X_i \), and (iii) \( \sup\{E(s_{m+n} - s_n)^2/n : m, n \in \mathbb{N}\} < \infty \) and \( \|X_n\|_{\beta} = E^{\frac{1}{\beta}}(|X_n|^\beta) = o(n^{(1-r)/2}) \). Then \( s_n/(\sigma \sqrt{n}) \) converges to \( N(0, 1) \) in distribution.

In our case, each \( \hat{\lambda}_h(t) \) for \( t \in [bh, T - bh] \) depends only on the arrivals in the interval \( [t - bh, t + bh] \), so the sequence \( \{g_i, i = 1, 2, \ldots\} \) given \( \lambda(\cdot) \) is \( m \)-dependent (see Herrndorff
(1984) for the definition), where \( m = [t + 1 + 2bh] + 1 \). Since the arrival rates \( \lambda(\cdot) \) are bounded, the requirements (i) and (iii) on the sequence \( \{g_i - E(g_i|\lambda(\cdot)), i = 1, 2, \cdots \} \) given \( \lambda(\cdot) \) are automatically satisfied. We have shown that
\[
\lim_{n \to \infty} E[\var(\frac{S_n(t, h)}{\sqrt{n}}|\lambda(\cdot))] = \sigma_1^2(t, h)
\]
in the proof of Lemma 3.2. Thus, owing to the ergodicity of \( \{\lambda(s), s \in \mathbb{R}\} \), we have
\[
\var(\frac{S_n(t, h)}{\sqrt{n}}|\lambda(\cdot)) \to \sigma_1^2(t, h) \text{ with probability one},
\]
which tells us that the requirement (ii) is also satisfied. Therefore, (A.9) holds.

To prove the second statement, we consider the characteristic function of \( S_n(t, h)/\sqrt{n} \)
\[
E(\exp(i\kappa S_n(t, h)/\sqrt{n})) = E(\exp(i\kappa [S_n(t, h) - G_n(t, h) + G_n(t, h)]/\sqrt{n}))
\]
\[
= E \left\{ E \left[ \exp\{i\kappa(S_n(t, h) - G_n(t, h))/\sqrt{n}\} - \exp\{-\sigma_1^2(t, h)\kappa^2\} |\lambda(\cdot)\right] \exp(i\kappa G_n(t, h)/\sqrt{n}) \right\} + \exp\{-\sigma_1^2(t, h)\kappa^2\} E(\exp(i\kappa G_n(t, h)/\sqrt{n})).
\]
We know from (A.9) that
\[
E \left[ \exp\{i\kappa(S_n(t, h) - G_n(t, h))/\sqrt{n}\} - \exp\{-\sigma_1^2(t, h)\kappa^2\} |\lambda(\cdot)\right] \to 0 \text{ as } n \to \infty.
\]
Since both \( E \left[ \exp\{i(S_n(t, h) - G_n(t, h))\kappa/\sqrt{n}\} - \exp\{-\sigma_1^2(t, h)\kappa^2\} |\lambda(\cdot)\right] \) and \( \exp\{i(G_n(t, h))\kappa/\sqrt{n}\} \) are bounded, using dominated convergence theorem we have
\[
\lim_{n \to \infty} E \left\{ E \left[ \exp\{i\kappa(S_n(t, h) - G_n(t, h))/\sqrt{n}\} - \exp\{-\sigma_1^2(t, h)\kappa^2\} |\lambda(\cdot)\right] \exp(i\kappa G_n(t, h)/\sqrt{n}) \right\} = 0.
\]
Consequently,
\[
\lim_{n \to \infty} E(\exp\{iS_n(t, h)\kappa/\sqrt{n}\}) = \exp\{-\sigma_1^2(t, h)\kappa^2 - \tau^2(t, h)\kappa^2\},
\]
if and only if \( E(\exp\{iG_n(t, h)\kappa/\sqrt{n}\}) = \exp\{-\tau^2(t, h)\kappa^2\} \), since we have already shown in the proof of Lemma 3.2 that
\[
\lim_{n \to \infty} \var(G_n(t, h))/n = \lim_{T \to \infty} (T - 2bh - t)\var[E(\hat{C}_{\mu,h}(t)|\lambda(\cdot))]. \quad \Box
\]

Proof of Lemma A.3. We use the same notations as in the proof of Lemma A.2. Let \( \mathcal{F}'_k = \sigma(\nu_i(t, h) : i \leq k) \), \( \mathcal{G}'_k = \sigma(\nu_i(t, h) : i \geq k) \) and \( \rho'_k = \sup\{E(\xi) : \xi \in \mathcal{F}'_j, E\xi = 0, \|\xi\| \leq \}

1; \eta \in G_{j+k}', E \eta = 0, \| \eta \| \leq 1). Then \mathcal{F}_k' \subset \mathcal{F}_{k+t+2bh}, and \mathcal{G}_k' \subset \mathcal{G}_{k-1}, so \rho_k' \leq \rho_{k-t-2bh-1} \text{ for } k \geq t + 2bh + 1, and

\[ \sum_k \rho_k' \leq \sum_{k=1}^{[t+2bh+1]} \rho_k' + \sum_{k \geq [t+2bh+1]} \rho_k \leq \sum_{k=1}^{[t+2bh+1]} \rho_k' + \int_0^\infty \rho_t dt < +\infty. \]

We thus conclude that if the process \{\lambda(t), t \in \mathbb{R}\} is \rho-mixing, then so is the process \{\nu_i(t, h), i = 1, 2, \ldots\}. Using Theorem 19.2 in Billingsley (1999), the desired result follows. ∎

Proof of Theorem 4.3. It is shown in Billingsley (1999, p. 201) that discrete-time finite-state Markov chains are \rho-mixing. The proof can be straightforwardly extended to continuous-time finite-state Markov chains. Using Lemma A.2, Theorems A.3 and 4.1, the desired result follows. ∎

Proof of Theorem 4.3. To prove this theorem, we need the following result (Theorem 4) of Arcones (1994): Let \( \{X_i = (X_i^{(1)}, \ldots, X_i^{(d)})\}_{i=1}^\infty \) be a \( \mathbb{R}^d \)-valued stationary mean-zero Gaussian sequence, and \( F \) be a function on \( \mathbb{R}^d \) with finite rank \( \tau \geq 1 \). Denote \( r^{(p,q)}(k) = E[X_m^{(p)} X_{m+k}^{(q)}] \) for \( k \in \mathbb{Z} \) and \( 1 \leq p, q \leq d \). If \( \sum_{k=-\infty}^{\infty} |r^{(p,q)}(k)|^\tau < \infty \) for each \( 1 \leq p, q \leq d \), then

\[ n^{-1/2} \sum_{j=1}^n (F(X_j) - E(F(X_j))) \xrightarrow{d} N(0, \sigma^2), \]

where \( \sigma^2 = \text{var}(F(X_1)) + 2 \sum_{k=1}^\infty \text{cov}(F(X_1), F(X_{1+k})) \).

In our case, we take \( X_i = (W((i-1)\epsilon), W(i\epsilon), \ldots, W([2bh+t+i]\epsilon)) = [2bh+t+i] \epsilon \) to be a \( d = [(2bh+t)/\epsilon] + 2 \) dimension vector, and define

\[ \bar{v}_i(t, h) = F(v_i) := E \left[ \int_{bh+(i-1)\epsilon}^{bh+i\epsilon} (\hat{\lambda}_h(s) - \mu)(\hat{\lambda}_h(s + t) - \mu) ds | \lambda(.) \right] \]

\[ = \int_{bh+(i-1)\epsilon}^{bh+i\epsilon} \left[ \int_{-b}^{b} f(l) f(m) (\lambda(lh+s) - \mu)(\lambda(mh+s+t) - \mu) dldmds \right. \]

\[ + \left. \frac{1}{h} \int_{bh+(i-1)\epsilon}^{bh+i\epsilon} \int_{-b}^{b} f(l)f(l + t/h) \lambda(lh + t + s) dlds \right]. \]

Without loss of generality, we assume \( \text{var}(W(i\epsilon)) = 1 \). Since \( g \) is bounded, \( E(F^2(X_i)) < \infty \). In our case, \( r^{(p,q)}(k) = \gamma(\{k + q - p\epsilon\}) \). Since \( g(W(i\epsilon)) \) has bounded second moment, the function \( g(x) \) can be expanded in \( L^2(\mathbb{R}, 1/\sqrt{2\pi} e^{-x^2/2}) \) in terms of Hermite polynomials (Taquq 1975): \( g(x) = \sum_{k=0}^{\infty} \frac{J(k)}{k!} H_k(x) \), where \( J(k) = E(G(Z)H_k(Z)) \) for \( Z \sim N(0, 1) \) and \( H_k(x) \) is the \( k \)th Hermite polynomial. Thus, \( \bar{v}_i(t, h) \), which is a polynomial of \( g(X_i^{(p)}) \), can be expanded as a series of polynomials of \( X_i^{(p)} \). Therefore, we can always find a polynomial \( P(X_i) \) of \( X_i \) such that \( P(X_i) \) is close (in mean square sense) enough to \( F(X_i) - E(F(X_i)) \) to have \( E[(F(X_i) - E(F(X_i)))P(X_i)] > 0 \). Thus, the rank \( \tau \) of the function is finite (see Arcones (1994) for the definition of the rank of a function). Since we know \( \sum_{j=1}^{\infty} |\gamma(j\epsilon)| < \infty \), it follows
that for any $\nu \geq 1$, we also have $\sum_{j=-\infty}^{\infty} |\gamma(j\varepsilon)|^\nu < \infty$. In particular, it holds for $\nu = \tau$. Now using the result that we stated in the beginning of the proof, we know that the asymptotic normality of $\tilde{v}_i(t, h) = F(X_i)$ holds. Let $n = \lfloor (T - 2bh - t)/\varepsilon \rfloor$, then

$$(T - 2bh - t)^{-1/2}(E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) - E(\hat{C}_{\mu,h}(t))) = \sum_{i=1}^{n}(\tilde{v}_i(t, h) - E(\tilde{v}_i(t, h))) +$$

$$+ \frac{1}{(T - 2bh - t)^{1/2}} \left\{ \int_{bh+n\varepsilon}^{T-bh-t} [(\hat{\lambda}_h(s) - \mu)(\hat{\lambda}_h(s+t) - \mu) - E(\hat{C}_{\mu,h}(t))] ds \right\}.$$  

(A.13)

Since term (A.13) converges to zero in probability, $(T - 2bh - t)^{-1/2} \sum_{i=1}^{n}(\tilde{v}_i(t, h) - E(\tilde{v}_i(t, h)))$ is asymptotically equivalent to $(n\varepsilon)^{-1/2} \sum_{i=1}^{n}(\tilde{v}_i(t, h) - E(\tilde{v}_i(t, h)))$ as $n \to \infty$. Therefore, (A.10) holds, which gives us the desired result upon using Lemma A.2.

Additional References
