Supplementary Material: Proofs

To prove Theorem 2.1, we need the following lemma.

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

$$
\begin{aligned}
E(\hat{\lambda}_h(t_0)|\lambda(\cdot)) &= \int_{-b}^{b} f(r)\lambda(rh+t_0)dr, \\
\mathrm{var}(\hat{\lambda}_h(t_0)|\lambda(\cdot)) &= \frac{1}{h}\int_{-b}^{b} f^2(r)\lambda(rh+t_0)dr.
\end{aligned}
$$

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

$$
\begin{aligned}
E(\hat{\lambda}_h(t_0)|\lambda(\cdot)) &= E[E\left(\hat{\lambda}_h(t_0)|\lambda(\cdot),K\right)|\lambda(\cdot)] = E[E\left(\sum_{i=1}^{K} f_h(t_0,s_i)|\lambda(\cdot),K\right)|\lambda(\cdot)] \\
&= \sum_{k=0}^{+\infty} \frac{1}{k!}\left(\int_0^T \lambda(s)\right)^k \exp[-\int_0^T \lambda(s)ds] \times \frac{k\int_0^T f_h(t_0,s)\lambda(s)ds}{\int_0^T \lambda(s)ds}\cdot \\
&= \int_0^T f_h(t_0,s)\lambda(s)ds = \int_{-b}^{b} f(r)\lambda(rh+t_0)dr
\end{aligned}
$$

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

$$
\begin{aligned}
&\mathrm{var}(\hat{\lambda}_h(t_0)|\lambda(\cdot)) = E\{\mathrm{var}[\sum_{i=0}^{K} f_h(t_0,s_i)|\lambda(\cdot),K]|\lambda(\cdot)\} + \mathrm{var}\{E[\sum_{i=0}^{K} f_h(t_0,s_i)|\lambda(\cdot),K)]|\lambda(\cdot)\} \\
= &\; E\{K\cdot \mathrm{var}(f_h(t_0,s)|\lambda(\cdot),K)|\lambda(\cdot)\} + \mathrm{var}\{K\int_0^T f_h(t_0,s)\lambda(s)ds/\int_0^T \lambda(s)ds|\lambda(\cdot)\}.
\end{aligned}
$$

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

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

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

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

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

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

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

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

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 γ_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). $\square$

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 $C(|(r-m)h+t|)$ around t for $t \geq 2bh$ and around 0 for $t \leq 2bh$ yields the subsequent results. $\square$

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 $E(\hat{\lambda}_h(s+t)\hat{\lambda}_h(s)|\lambda(\cdot)) = \frac{1}{h^2}\int_0^T dl \int_0^T f(\frac{l-s-t}{h})f(\frac{m-s}{h})\lambda(l)\lambda(m)dm +$
$\frac{1}{h^2}\int_0^T f(\frac{l-s-t}{h})f(\frac{l-s}{h})\lambda(l)dl$, which implies that
$E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) = \frac{1}{T-2bh-t}\frac{1}{h^2}\{\int_{bh}^{T-bh-t} ds \int_0^T \lambda(l)f(\frac{l-s-t}{h})f(\frac{l-s}{h})dl +$
$\int_{bh}^{T-bh-t} ds \int_0^T dl \int_0^T [\lambda(l)-\mu][\lambda(m)-\mu]f(\frac{l-s-t}{h})f(\frac{m-s}{h})dm\}$. A term-by-term expansion and four change-of-variables give the detailed expression (3.4). Next, we note that the three terms $A_{t,T}^h$, $B_{t,T}^h$ and $C_{t,T}^h$ 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 $T \to \infty$.

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

$$\begin{aligned} D_t^h &= \mathrm{Int}_{f;[-b,b]^4}^{l,m,l',m'} \langle E\{\lambda(0)\lambda(t+(l-m)h)\lambda((l-m-l'+m')h)\}\rangle \\ &\quad + \mathrm{Int}_{f;[-b,b]^4}^{l,m,l',m'} \langle E\{\lambda(0)\lambda(t+(l-m)h)\lambda(t+(l'-m')h)\}\rangle \\ &\quad + 2\,\mathrm{Int}_{f;[-b,b]^4}^{l,m,l',m'} \langle E\{\lambda(0)\lambda(t+(l-m)h)\lambda(2t+(l-m+l'-m')h)\}\rangle \\ &\quad -8\mu \iint_{[-b,b]^2} C(t+(l-m)h) f(l) f(m) dl\, dm - 4\mu^3, \qquad \text{(A.2)} \\ E_t^h &= \int_{-2b}^{2b} dr \iint_{[-b,b]^2} [C(t+(l-m)h)+\mu^2] f(l) f(l+r) f(m) f(m+r) dl\, dm \\ &\quad +4\,\mathrm{Int}_{f;[-b,b]^3}^{l,l',m'} \langle C(t+(l'-m')h) f(l+\frac{t}{h})\rangle \\ &\quad + \mathrm{Int}_{f;[-b,b]^3}^{l,m,m'} \langle [C(t+(l-m)h)+\mu^2] f(l-m+m'+\frac{2t}{h})\rangle, \qquad \text{(A.3)} \\ F_t^h &= \mu \iint_{[-b,b]^2} f(l) f(l+\frac{t}{h}) f(m) f(m+\frac{t}{h}) dl\, dm. \qquad \text{(A.4)} \end{aligned}$$

Here for notational ease we have used the short hand notation $\mathrm{Int}_{f;A}^{l,m,l',m'}\langle\cdot\rangle$, 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, $\mathrm{Int}_{f;[-b,b]^3}^{l,l',m'}\langle g(l,l',m')\rangle = \iiint_{[-b,b]^3} g(l,l',m') f(l) f(l') f(m') dl\, dl'\, dm'$ and $\mathrm{Int}_{f;[-b,b]^4}^{l,m,l',m'}\langle g(l,m,l',m')\rangle = \iiiint_{[-b,b]^4} g(l,m,l',m') f(l) f(m) f(l') f(m') dl\, dm\, dl'\, dm'$.

To establish these expressions, we first note that

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

A term-by-term expansion on $\mathrm{cov}\{(\hat{\lambda}_h(s+t)-\mu)(\hat{\lambda}_h(s)-\mu), (\hat{\lambda}_h(s'+t)-\mu)(\hat{\lambda}_h(s')-\mu)|\lambda(\cdot)\}$ 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 $E\{\hat{\lambda}_h(s+t)\hat{\lambda}_h(s)\hat{\lambda}_h(s'+t)\hat{\lambda}_h(s')|\lambda(\cdot)\}$, $E\{\hat{\lambda}_h(s+t)\hat{\lambda}_h(s)\hat{\lambda}_h(s'+t)|\lambda(\cdot)\}$, etc. By definition,

$$\begin{aligned} & E\{\hat{\lambda}_h(s+t)\hat{\lambda}_h(s)\hat{\lambda}_h(s'+t)\hat{\lambda}_h(s')|\lambda(\cdot)\} \\ = \; & \frac{1}{h^4} E\{\sum_{i=1}^k f(\frac{s_i-s-t}{h}) \sum_{j=1}^k f(\frac{s_j-s}{h}) \sum_{i'=1}^k f(\frac{s_{i'}-s'-t}{h}) \sum_{j'=1}^k f(\frac{s_{j'}-s'}{h})|\lambda(\cdot)\}. \end{aligned}$$

Depending on whether i, j, i' and j' take on distinct values or not, we have

$$\begin{aligned} & E\{\hat{\lambda}_h(s+t)\hat{\lambda}_h(s)\hat{\lambda}_h(s'+t)\hat{\lambda}_h(s')|\lambda(\cdot)\} \\ = \; & \frac{1}{h^4}\{ \sum_{\text{all distinct}} + \sum_{\text{three distinct, } i=j} + \sum_{\text{three distinct, } i=i'} + \sum_{\text{three distinct, } i=j'} \\ & + \sum_{\text{three distinct, } j=i'} + \sum_{\text{three distinct, } j=j'} + \sum_{\text{three distinct, } i'=j'} + \sum_{\text{two distinct, } i=j,\, i'=j'} \\ & + \sum_{\text{two distinct, } i=i',\, j=j'} + \sum_{\text{two distinct, } i=j',\, i'=j} + \sum_{\text{two distinct, } i=j=i'} + \sum_{\text{two distinct, } i=j=j'} \\ & + \sum_{\text{two distinct, } i=i'=j'} + \sum_{\text{two distinct, } j=i'=j'} + \sum_{i=j=i'=j'} \}. \end{aligned}$$

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

$$\sum_{\text{three distinct, } i=i'} = \frac{1}{h}\,\mathrm{Int}_{f;[-b,b]^3}^{l,m,m'}\langle \lambda(s+t+lh)\lambda(s+mh)\lambda(s'+m'h)f(l+\frac{s-s'}{h})\rangle, \tag{A.6}$$

and the term

$$\sum_{\text{two distinct, } i=j',\, i'=j} = \frac{1}{h^2}\,\mathrm{Int}_{f;[-b,b]^2}^{l,m}\langle \lambda(s+t+lh)\lambda(s+mh)f(m+\frac{s-s'-t}{h})f(l+\frac{s+t-s'}{h})\rangle. \tag{A.7}$$

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

To see how the detailed expressions of D_t^h, E_t^h and F_t^h 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

$$\frac{1}{h}\iint_{[bh,T-bh-t]^2} ds\, ds'\,\mathrm{Int}_{f;[-b,b]^3}^{l,m,m'}\langle E\{\lambda(0)\lambda(t+(l-m)h)\lambda(s'-s+(m'-m)h)\}f(l+\frac{s-s'}{h})\rangle.$$

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

$$(T-2bh-t)\,\mathrm{Int}_{f;[-b,b]^4}^{l,m,l',m'}\langle E\{\lambda(0)\lambda(t+(l-m)h)\lambda((l-m-l'+m')h)\}\rangle + D_1, \tag{A.8}$$

where

$$\begin{aligned} D_1 &= \mathrm{Int}^{l,m,m'}_{f;[-b,b]^3}\langle E\{\lambda(0)\lambda(t+(l-m)h)\lambda((l-m-l'+m')h)\}\cdot \\ &\quad \cdot[\int_{bh}^{3bh} ds \int_{-b}^{l+\frac{s-bh}{h}} f(l')dl' + \int_{T-3bh-t}^{T-bh-t} ds \int_{l+\frac{s-(T-bh-t)}{h}}^{b} f(l')dl' - 4bh]\rangle \\ &= O(h). \end{aligned}$$

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

$$\frac{1}{h^2}\iint_{[bh,T-bh-t]^2} ds\, ds'\, \mathrm{Int}^{l,m}_{f;[-b,b]^2}\langle[C(t+(l-m)h)+\mu^2]f(m+\frac{s-s'-t}{h})f(l+\frac{s+t-s'}{h})\rangle.$$

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

$$\frac{T-2bh-t}{h}\mathrm{Int}^{l,m,m'}_{f;[-b,b]^3}\langle[C(t+(l-m)h)+\mu^2]f(l-m+m'+\frac{2t}{h})\rangle + O(h),$$

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. □

Proof of Theorem 3.3. From Lemma 3.1, it is transparent that as $h \to 0$, the bias of $\hat{C}_{\mu,h}(t)$ 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 $\hat{C}_{\mu,h}(t)$ is a linear combination of $O(\frac{1}{T^2})$, $O(\frac{1}{T^2h})$, $O(\frac{1}{T^2h^2})$, $O(\frac{1}{T})$, $O(\frac{1}{Th})$ and $O(\frac{h}{T^2})$ terms, each converging to zero as $Th \to \infty$ and $h \to 0$. The desired result thus follows. □

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

$$\begin{aligned} E(\hat{\mu}-\mu)^2 &= \mathrm{var}(\hat{\mu}) = E(\mathrm{var}(\hat{\mu}|\lambda(\cdot))) + \mathrm{var}(E(\hat{\mu}|\lambda(\cdot))) \\ &= \frac{\mu}{T} + \frac{1}{T^2}\int_0^T\int_0^T C(s-s')ds'ds \to 0, \text{ as } T\to\infty. \end{aligned}$$

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

$$\begin{aligned} &\iint_{[bh,T-bh-t]^2} ds\, ds'\, \mathrm{cov}(\hat{\lambda}_h(s+t)+\hat{\lambda}_h(s), \hat{\lambda}_h(s'+t)+\hat{\lambda}_h(s')) \\ =\; & 2\iint_{[bh,T-bh-t]^2} ds\, ds' \iint_{[-b,b]^2} C(s-s'+(l-m)h)f(l)f(m)dl\, dm \\ & +2\iint_{[bh,T-bh-t]^2} ds\, ds' \iint_{[-b,b]^2} C(s+t-s'+(l-m)h)]f(l)f(m)dl\, dm \\ & +2\iint_{[bh,T-bh-t]^2} ds\, ds' \int_{-b}^{b} f(l)[f(l+\frac{s-s'}{h})+f(l+\frac{s+t-s'}{h})]dl. \end{aligned}$$

Applying L'Hospital's rule, we know that $\mathrm{var}(\frac{1}{T-bh-t}\int_{bh}^{T-bh-t}(\hat{\lambda}_h(s+t)+\hat{\lambda}_h(s)-2\mu)ds) \to 0$, as $T \to \infty$, so it, in particular, converges to zero in probability. Using these results together with Theorem 3.3, we know from

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

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

Proof of Proposition 3.5. From the definition of $A^h_{t,T}$ (Lemma 3.2), it is transparent that

$$\begin{aligned} A^h_{t,T} &= \mu^4 \iint_{[bh,T-bh-t]^2} ds\, ds' \iiiint_{[-b,b]^4} \cdot \\ &\quad \mathrm{cov}\{(\lambda_0(mh+s)-1)(\lambda_0(lh+t+s)-1),(\lambda_0(m'h+s')-1)(\lambda_0(l'h+t+s')-1)\} \cdot \\ &\quad f(l)f(m)f(l')f(m')dl\, dm\, dl'\, dm' = O(\mu^4), \end{aligned}$$

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

$$\mathrm{var}((T-2bh-t)\hat{C}_{\mu,h}(t)) - A^h_{t,T} = O(\mu^3),$$

from which (3.7) follows. □

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 $\lambda(s), s \in \mathsf{R}$ is bounded. Then, for fixed $t, h \geq 0$, conditioning on the realization of λ, with probability one we have*

$$\sqrt{T}\left[\hat{C}_{\mu,h}(t) - E(\hat{C}_{\mu,h}(t)|\lambda(\cdot))\right] \Big| \ \lambda(\cdot) \xrightarrow{D} N(0,\sigma_1^2(t,h)) \text{ as } T \to \infty, \tag{A.9}$$

where the constant $\sigma_1^2(t,h) = D^h_t + E^h_t/h + F^h_t/h^2$, defined in Lemma 3.2 (see also (A.2), (A.3) and (A.4)), does not depend on λ. 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 $\sigma^2(t,h) = \lim_{T\to\infty} T\ \mathrm{var}(\hat{C}_{\mu,h}(t))$, if and only if

$$\sqrt{T}[E(\hat{C}_{\mu,h}(t)|\lambda(\cdot)) - E(\hat{C}_{\mu,h}(t))] \xrightarrow{D} N(0,\tau^2(t,h)) \text{ as } T \to \infty, \tag{A.10}$$

where

$$\tau^2(t,h) = \lim_{T\to\infty} T\ \mathrm{var}(E(\hat{C}_{\mu,h}(t)|\lambda(\cdot))). \tag{A.11}$$

Lemma A.2 indicates that the asymptotic distribution of $\hat{C}_{\mu,h}(t)$ rests on the stochastic properties of the underlying arrival rate $\{\lambda(t), t \in \mathsf{R}\}$.

Lemma A.3. *Suppose that Assumptions 1, 2, 4, and 5 hold, and that the arrival rate process* $\{\lambda(t), t \in \mathsf{R}\}$ *is finite* ρ*-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

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

Since we can write

$$\sqrt{T-2bh-t}(\hat{C}_{\mu,h}(t) - E(\hat{C}_{\mu,h}(t))) = \frac{S_n(t,h)}{\sqrt{T-2bh-t}} + \frac{1}{\sqrt{T-2bh-t}} \int_{bh+n}^{T-bh-t} \{(\hat{\lambda}_h(s) - \mu)(\hat{\lambda}_h(s+t) - \mu) - E(\hat{C}_{\mu,h}(t))\}ds, \quad \text{(A.12)}$$

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) $\sup\{E(s_{m+n} - s_n)^2/n : m, n \in \mathsf{N}\} < \infty$ 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 $\{g_i, i = 1, 2, \cdots\}$ 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 $\{\lambda(s), s \in \mathsf{R}\}$, 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 $S_n(t,h)/\sqrt{n}$

$$\begin{aligned} & E(\exp(i\kappa S_n(t,h)/\sqrt{n})) = E(\exp(i\kappa[S_n(t,h) - G_n(t,h) + G_n(t,h)]/\sqrt{n})) \\ = \;& E\left\{E\left[\exp\{i\kappa(S_n(t,h) - G_n(t,h))/\sqrt{n}\} - \exp\{-\sigma_1^2(t,h)\kappa^2\}\,|\lambda(\cdot)\right]\exp(i\kappa G_n(t,h)/\sqrt{n})\right\} \\ + \;& \exp\{-\sigma_1^2(t,h)\kappa^2\}E(\exp\{i\kappa G_n(t,h)/\sqrt{n}\}). \end{aligned}$$

We know from (A.9) that

$$E\left[\exp\{i\kappa(S_n(t,h) - G_n(t,h))/\sqrt{n}\} - \exp\{-\sigma_1^2(t,h)\kappa^2\}\,|\lambda(\cdot)\right] \xrightarrow{P} 0 \text{ as } n \to \infty.$$

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

$$\lim_{n\to\infty} \text{var}(G_n(t,h))/n = \lim_{T\to\infty}(T - 2bh - t)\text{var}[E(\hat{C}_{\mu,h}(t)|\lambda(\cdot))]. \quad \square$$

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

$1; \eta \in \mathcal{G}'_{j+k}, E\eta = 0, \|\eta\| \leq 1\}$. Then $\mathcal{F}'_k \subset \mathcal{F}_{k+t+2bh}$, and $\mathcal{G}'_k \subset \mathcal{G}_{k-1}$, so $\rho'_k \leq \rho_{k-t-2bh-1}$ 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 $\{\lambda(t), t \in \mathsf{R}\}$ is ρ-mixing, then so is the process $\{\upsilon_i(t,h), i = 1, 2, \ldots\}$. 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 ρ-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 $\{X_i = (X_i^{(1)}, \cdots, X_i^{(d)})\}_{i=1}^\infty$ be a R^d-valued stationary mean-zero Gaussian sequence, and F be a function on R^d with finite rank $\tau \geq 1$. Denote $r^{(p,q)}(k) = E[X_m^{(p)} X_{m+k}^{(q)}]$ for $k \in Z$ 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 $\sigma^2 = \mathrm{var}(F(X_1)) + 2\sum_{k=1}^\infty \mathrm{cov}(F(X_1), F(X_{1+k}))$.

In our case, we take $X_i = (W((i-1)\varepsilon), W(i\varepsilon), \cdots, W([\frac{2bh+t}{\varepsilon} + i]\varepsilon))^T$ to be a $d = [(2bh + t)/\varepsilon] + 2$ dimension vector, and define

$$\begin{aligned} \tilde{\upsilon}_i(t,h) &= F(X_i) := E\left[\int_{bh+(i-1)\varepsilon}^{bh+i\varepsilon} (\hat{\lambda}_h(s) - \mu)(\hat{\lambda}_h(s+t) - \mu) ds | \lambda(\cdot)\right] \\ &= \int_{bh+(i-1)\varepsilon}^{bh+i\varepsilon} \int_{-b}^{b} \int_{-b}^{b} f(l) f(m) (\lambda(lh+s) - \mu)(\lambda(mh+s+t) - \mu) dl dm ds \\ &+ \frac{1}{h} \int_{bh+(i-1)\varepsilon}^{bh+i\varepsilon} \int_{-b}^{b} f(l) f(l + \frac{t}{h}) \lambda(lh+t+s) dl ds. \end{aligned}$$

Without loss of generality, we assume $var(W(i\varepsilon)) = 1$. Since g is bounded, $E(F^2(X_i)) < \infty$. In our case, $r^{(p,q)}(k) = \gamma(|k+q-p|\varepsilon)$. Since $g(W(i\varepsilon))$ has bounded second moment, the function $g(x)$ can be expanded in $L^2(\mathsf{R}, 1/\sqrt{2\pi} e^{-x^2/2})$ in terms of Hermite polynomials (Taqqu 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 kth Hermite polynomial. Thus, $\tilde{\upsilon}_i(t,h)$, 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 $E\left[(F(X_i) - E(F(X_i)))P(X_i)\right] > 0$. Thus, the rank τ of the function is finite (see Arcones (1994) for the definition of the rank of a function). Since we know $\sum_{j=1}^\infty |\gamma(j\varepsilon)| < \infty$, 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 $n = [(T - 2bh - t)/\varepsilon]$, then

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

$$+\frac{1}{(T-2bh-t)^{1/2}}\{\int_{bh+n\varepsilon}^{T-bh-t} [(\hat{\lambda}_h(s)-\mu)(\hat{\lambda}_h(s+t)-\mu) - E(\hat{C}_{\mu,h}(t))]ds\}. \quad \text{(A.13)}$$

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

[1] Arcones, M.A. (1994) Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Annals of Probability*, 22, 2242-2274.

[2] Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Annals of Probability*, 12, 141-153.