EMPIRICAL FOURIER METHODS FOR INTERVAL CENSORED DATA

Peter G. Hall, W. John Braun* and Thierry Duchesne†

*University of Melbourne, *University of British Columbia and †Université Laval*

Supplementary Material

S1 APPENDIX TO SECTION 2

To appreciate why approaches (a) and (b) are effective, consider the assumption

> there does not exist a nondegenerate interval $\mathcal{J}$ such that $f_X^{\mathrm{Ft}}(u)$ is nonzero for at least one element u of $\mathcal{J}$, and $f_{Z_1,Z_2}^{\mathrm{Ft}}(-s,t) = 0$ for all pairs (s,t) such that $s+t \in \mathcal{J}$. (S1.1)

In view of (2.3), condition (S1.1) is sufficient for f_X to be completely determined once the distributions of (L, R) and (Z_1, Z_2) are known. Since (S1.1) fails only if $f_{Z_1,Z_2}^{\mathrm{Ft}}(-s,t)$ vanishes on a particular line in the plane, then (S1.1) would hold for all parametric models that are likely to be used in practice, and more generally, in a nonparametric setting, (S1.1) would fail only rarely. Indeed, it can be seen from (2.3) that the distributions of

(Z_1, Z_2) and X are both nonparametrically identifiable from data on (L, R) alone if:

$$\begin{aligned}&\text{the characteristic functions } f^{\rm Ft}_{Z_1,Z_2} \text{ and } f^{\rm Ft}_X \text{ satisfy (S1.1), and } f^{\rm Ft}_{Z_1,Z_2} \text{ cannot}\\ &\text{be decomposed in the manner } f^{\rm Ft}_{Z_1,Z_2}(s,t)=\psi(s,t)\phi(t-s), \text{ where } \psi \text{ is}\\ &\text{the characteristic function of a bivariate distribution and the function } \phi,\\ &\text{of a single variable, is the Fourier-Stieltjes transform of a function whose}\\ &\text{variation does not occur at a single point.}\end{aligned} \tag{S1.2}$$

We claim that (S1.2) holds whenever Z_1 and Z_2 are independent, which condition holds when the point processes $\mathcal{T}_i$ are Poisson. To appreciate why the assumption of independent Z_1 and Z_2 is sufficient for (S1.2), note that if $f^{\rm Ft}_{Z_1,Z_2}(s,t) = \psi(s,t)\phi(t-s)$ then (Z_1, Z_2) has distribution function

$$F(z_1, z_2) = \int F_1(z_1 - v, z_2 + v) dG(v),$$

where F_1 is the distribution function of the distribution with characteristic function ψ, and ϕ is the Fourier-Stieltjes transform of G. The function F fails to be the product of its marginals if G is not concentrated at a single point. (The case where G is a distribution function is more familiar. There, F is the distribution function of $(U_1 + V, U_2 - V)$, where (U_1, U_2) has distribution F_1, V has distribution function G, (U_1, U_2) and V are independent, and Z_1 and Z_2 fail to be independent if V is not identically constant.)

S2 APPENDIX TO SECTION 3

The reason for raising the weight function w in (3.7) to the power q is that, when taking $q = \infty$, this gives the criterion

$$S_\infty(\theta, \mathcal{B}, \omega) = \sup_{-\infty < s,t < \infty} |\widehat{f_{LR}^{\mathrm{Ft}}}(s,t) - f_X^{\mathrm{Ft}}(s+t|\mathcal{B},\omega) f_{Z_1,Z_2}^{\mathrm{Ft}}(-s,t|\theta)| w(s,t). \qquad \text{(S2.3)}$$

This is more useful than its analogue where $w \equiv 1$, which we would obtain if in (3.7) we were to replace $w(s,t)^q$ by $w(s,t)$ and let $q \to \infty$.

In some instances it is convenient to modify the criterion in (3.7). For example, under the model at (3.6) we might change $S_q(\theta, \mathcal{B}, \omega)$ to

$$S_q(\lambda, \mathcal{B}, \omega) = \left\{ \int\int \left| (1+\lambda^{-1} is)(1-\lambda^{-1} it) \widehat{f_{LR}^{\mathrm{Ft}}}(s,t) - f_X^{\mathrm{Ft}}(s+t|\mathcal{B},\omega) \right|^q w(s,t)^q ds dt \right\}^{1/q}, \qquad \text{(S2.4)}$$

which when $q = \infty$ reduces to

$$S_\infty(\lambda, \mathcal{B}, \omega) = \sup_{-\infty < s,t < \infty} \left| (1+\lambda^{-1} is)(1-\lambda^{-1} it) \widehat{f_{LR}^{\mathrm{Ft}}}(s,t) - f_X^{\mathrm{Ft}}(s+t|\mathcal{B},\omega) \right| w(s,t).$$

If we are in case (b), and therefore are employing a nonparametric estimator of $g(s,t) = f_{Z_1,Z_2}^{\mathrm{Ft}}(s,t)$, for example $\widehat{g}(s,t)$ at (3.4), then the criterion function changes to:

$$S_q(\mathcal{B}, \omega) = \left\{ \int\int \left| \widehat{f_{LR}^{\mathrm{Ft}}}(s,t) - f_X^{\mathrm{Ft}}(s+t|\mathcal{B},\omega) \widehat{g}(-s,t) \right|^q w(s,t) ds dt \right\}^{1/q} \qquad \text{(S2.5)}$$

The $q = \infty$ version of this quantity has a formula analogous to (S2.3):

$$S_\infty(\mathcal{B}, \omega) = \sup_{-\infty < s,t < \infty} |\widehat{f_{LR}^{\mathrm{Ft}}}(s,t) - f_X^{\mathrm{Ft}}(s+t|\mathcal{B},\omega) \widehat{g}(-s,t)| w(s,t). \qquad \text{(S2.6)}$$

S2.1 Bandwidth selection

An appropriate smoothing parameter, for example a bandwidth in the definition of $\widetilde{f}_X$ at (3.8), can be chosen using the comparison method; see Deheuvels (1977). Specifically, we fit a smooth parametric model $\bar{F}_X$ to $\widehat{F}_X$ (for example, $\bar{F}_X$ might be a Normal distribution with mean and variance estimated from data), and also to the point processes $\mathcal{T}_i$ (here we would generally use a Poisson model), and simulate from both to generate many "resamples" of intervals, $\mathcal{I}_n^* = \{[L_1^*, R_1^*], \ldots, [L_n^*, R_n^*]\}$. Using $\mathcal{I}_n^*$ in place of the original dataset $\mathcal{I}_n = \{[L_1, R_1], \ldots, [L_n, R_n]\}$ we construct an estimator $\widehat{F}_X^*$ of $\bar{F}_X$, using the methodology summarised in section 3.2, and from this distribution estimator we compute the resampled version $\widetilde{f}_X^*$ of $\widetilde{f}_X$. Critically, $\widetilde{f}_X^*$ depends on one or more tuning parameters, for example on the bandwidth h if we are using the kernel method at (3.8):

$$\widetilde{f}_X^*(x|h) = \frac{1}{h} \int_{-\infty}^{\infty} K\left(\frac{x-y}{h}\right) d\widehat{F}_X^*(y).$$

Since the fitted model $\bar{F}_X$ is smooth then it has a well defined density $\bar{f}_X$. We choose the smoothing parameters in the construction of $\bar{f}_X^*$ so as to minimise the average distance, for example mean integrated squared error conditional on the data, from $\bar{f}_X^*$ to $\bar{f}_X$, both of which are known. Finally, we use these smoothing parameters when constructing $\bar{f}_X$ at (3.8).

S2.2 Renewal point processes.

Assume that the point process $\mathcal{T} = \{\ldots, T_j, T_{j+1}, \ldots\}$ is a stationary renewal process, where the common lifetime distribution has probability density $f_{\text{Life}}(z|\theta)$ and characteristic function $f_{\text{Life}}^{\text{Ft}}(t|\theta)$, and θ is a finite vector of unknown parameters. Then the joint distribution of (Z_1, Z_2), where Z_1 and Z_2 are as introduced in section 2, is that of the current life and excess life for the stationary renewal process, and has density

$$f_{Z_1,Z_2}(z_1, z_2) = \mu(\theta)^{-1} f_{\text{Life}}(z_1 + z_2|\theta), \quad 0 < z_1, z_2 < \infty,$$

where $\mu(\theta) = \int_{z>0} z f_{\text{Life}}(z|\theta) dz$ denotes the mean of the lifetime distribution, assumed to be finite. See Karlin and Taylor (1975, pp. 192–4).

The characteristic function $f^{\text{Ft}}_{Z_1,Z_2}$ is therefore given by

$$f^{\text{Ft}}_{Z_1,Z_2}(s,t|\theta) = \{i\mu(\theta)(s-t)\}^{-1}\{f^{\text{Ft}}_{\text{Life}}(s|\theta) - f^{\text{Ft}}_{\text{Life}}(t|\theta)\}.$$

This quantity would be substituted into (3.7) when undertaking inference.

S3 TECHNICAL ARGUMENTS

S3.1 Proof of Theorem 1.

It is convenient to give the proof of part (b) first, and then summarise the minor changes needed to derive part (a).

Step 1: Preparatory lemma for part (b) of Theorem 1. Define

$$s(\mathcal{B},\omega) = \sup_{-\infty<t_1,t_2<\infty} |f^{\text{Ft}}_{LR}(t_1,t_2) - f^{\text{Ft}}_X(t_1+t_2|\mathcal{B},w) f^{\text{Ft}}_{Z_1,Z_2}(-t_1,t_2)| w(t_1,t_2). \quad \text{(S3.7)}$$

Let $(\widehat{\mathcal{B}},\widehat{\omega})$ denote the minimiser of $S_\infty(\mathcal{B},\omega)$ under the constraint (4.12)(ii). (Formally, $F_X(|\widehat{\mathcal{B}},\widehat{\omega})$ is the weak limit of a sequence of distributions with densities $f_X(|\mathcal{B},\omega)$ that satisfy $\int |x|^{C_3} f_X(x|\mathcal{B},\omega)dx \le C_4$ and that, in the limit of the sequence for fixed n, equal the infimum of $S_\infty(\mathcal{B},\omega)$ over all $(\mathcal{B},\omega)$.) Our first step is to prove the following result.

Lemma 1. *Under the assumptions in part (b) of Theorem 1, and for each $B_1 > 0$,*

$$P\left\{s(\widehat{\mathcal{B}},\widehat{\omega}) \le n^{B_1-(1/2)}\right\} \to 1. \quad \text{(S3.8)}$$

Proof of Lemma 1. It can be proved from Bernstein's inequality that

$$\sup_{-\infty<t_1<t_2<\infty} |P\left\{|\widehat{f^{\text{Ft}}_{LR}}(t_1,t_2) - f^{\text{Ft}}_{LR}(t_1,t_2)| > n^{\varepsilon-(1/2)}\right\} = O(n^{-C})$$

for all $C, \varepsilon > 0$. Therefore, if the set $\mathcal{A}_n$ contains $O(n^{C'})$ pairs (t_1, t_2) for some $C' > 0$, then

$$P\left\{\sup_{(t_1,t_2)\in A_n} |\widehat{f_{LR}^{\rm Ft}}(t_1,t_2) - f_{LR}^{\rm Ft}(t_1,t_2)| > n^{\varepsilon-(1/2)}\right\} = O(n^{-C}) \tag{S3.9}$$

for all $C, \varepsilon > 0$. Assumption (4.12)(i) asserts that $E|L|^{C_3} + E|R|^{C_3} < \infty$, where, without loss of generality, $0 < C_3 \le 1$. Therefore,

$$\frac{1}{n}\sum_{j=1}^{n}(|L_j|^{C_3} + |R_j|^{C_3}) = O_p(1), \tag{S3.10}$$

$$\left|f_{LR}^{\rm Ft}(t_1,t_2) - f_{LR}^{\rm Ft}(t_3,t_4)\right| \le (|t_1 - t_3|^{C_3} + |t_2 - t_4|^{C_3})E(|L|^{C_3} + E|R|^{C_3}), \tag{S3.11}$$

$$\begin{aligned}
&\left|\widehat{f_{LR}^{\rm Ft}}(t_1,t_2) - \widehat{f_{LR}^{\rm Ft}}(t_3,t_4)\right| \le \frac{1}{n}\sum_{j=1}^{n} |\exp\{i(t_1-t_3)L_j + i(t_2-t_4)R_j\} - 1| \\
&\le \frac{1}{n}\sum_{j=1}^{n} \min(|t_1-t_3||L_j| + |t_2-t_4||R_j|, 1) \\
&\le (|t_1-t_3|^{C_3} + |t_2-t_4|^{C_3})\frac{1}{n}\sum_{j=1}^{n}(|L_j|^{C_3} + |R_j|^{C_3}).
\end{aligned} \tag{S3.12}$$

Together, (S3.9)–(S3.12) imply that for any $B_1 > 0$, no matter how small, and any $B_2 > 0$, no matter how large,

$$P\left\{\sup_{|t_1|,|t_2|\le n^{B_2}} \left|\widehat{f_{LR}^{\rm Ft}}(t_1,t_2) - f_{LR}^{\rm Ft}(t_1,t_2)\right| \le n^{B_1-(1/2)}\right\} \to 1. \tag{S3.13}$$

Similarly it can be proved that, for the same choice of B_1 and B_2,

$$P\left\{\sup_{|t_1|,|t_2|\le n^{B_2}} |\widehat{g}(t_1,t_2) - f_{Z_1,Z_2}^{\rm Ft}(t_1,t_2)| \le n^{B_1-(1/2)}\right\} \to 1. \tag{S3.14}$$

Assumption (4.12)(vi) implies that we may choose B_2 so large that, whenever $n \neq 2$, $\omega(t_1, t_2) \leq \frac{1}{4} n^{-1/2}$ if (t_1, t_2) is in the set $\mathcal{A}_{1n}$ of all pairs for which the absolute value of at least one component exceeds n^{B_2}. Let $\mathcal{A}_{2n}$ denote the complement of $\mathcal{A}_{1n}$ in $\mathbb{R}^2$, and put

$$s_1(\mathcal{B}, \omega) = \sup_{(t_1,t_2)\in\mathcal{A}_{1n}} \left| f_{LR}^{\mathrm{Ft}}(t_1, t_2) - f_X^{\mathrm{Ft}}(t_1 + t_2|\mathcal{B}, w)\widehat{g}(-t_1, t_2) \right| w(t_1, t_2). \tag{S3.15}$$

$$s_2(\mathcal{B}, \omega) = \sup_{(t_1,t_2)\in\mathcal{A}_{2n}} \left| f_{LR}^{\mathrm{Ft}}(t_1, t_2) - f_X^{\mathrm{Ft}}(t_1 + t_2|\mathcal{B}, w) f_{Z_1,Z_2}^{\mathrm{Ft}}(-t_1, t_2) \right| w(t_1, t_2). \tag{S3.16}$$

Since the quantities within absolute value signs in (S3.15) and (S3.16) are bounded above by 2 then

$$P\{0 \leq s_1(\mathcal{B}, \omega) \leq \frac{1}{2} n^{-1/2}\} = P\{0 \leq s(\mathcal{B}, \omega) - s_2(\mathcal{B}, \omega) \leq \frac{1}{2} n^{-1/2}\} = 1 \tag{S3.17}$$

for $n \geq 2$.

We may assume without loss of generality that $sup \omega \leq 1$. In this case, (S3.17) implies that with probability 1,

$$\begin{aligned}
&|S_\infty(\mathcal{B}, \omega) - s(\mathcal{B}, \omega)| \\
&\leq \left| \sup_{(t_1,t_2)\in\mathcal{A}_{2n}} |\widehat{f_{LR}^{\mathrm{Ft}}}(t_1, t_2) - f_X^{\mathrm{Ft}}(t_1 + t_2|\mathcal{B}, \omega)\widehat{g}(-t_1, t_2)| w(t_1, t_2) - s_2(\mathcal{B}, \omega) \right|. \\
&+ \{s(\mathcal{B}, \omega) - s_2(\mathcal{B}, \omega)\} + s_1(\mathcal{B}, \omega) \\
&\leq \sup_{|t_1|,|t_2|\leq n^{B_2}} \{|\widehat{f_{LR}^{\mathrm{Ft}}}(t_1, t_2) - f_{LR}^{\mathrm{Ft}}(t_1, t_2)| + |\widehat{g}(-t_1, t_2) - f_{Z_1,Z_2}^{\mathrm{Ft}}(t_1, t_2)|\} + n^{-1/2}.
\end{aligned}$$

These inequalities, (S3.13) and (S3.14) imply that

$$P\{|S_\infty(\mathcal{B}, \omega) - s(\mathcal{B}, \omega)| \leq 2n^{B_1-(1/2)} + n^{-1/2} \text{ for all } (\mathcal{B}, \omega)\} \to 1. \tag{S3.18}$$

It follows from the definition of $(\widehat{\mathcal{B}}, \widehat{\omega})$ that $\inf_{\mathcal{B},\omega} S_\infty(\mathcal{B}, \omega) = S_\infty(\widehat{\mathcal{B}}, \widehat{\omega})$, where of course the constraint noted in (4.12)(ii) is imposed. From this property and (S3.18), and the fact that B_1 in (S3.18) denotes an arbitrary positive constant, we deduce that for all $B_1 > 0$,

$$P\{|S_\infty(\widehat{\mathcal{B}},\widehat{\omega}) - \inf_{\mathcal{B},\omega} s(\mathcal{B},\omega)| \le n^{B_1-(1/2)}\} \to 1, \qquad \text{(S3.19)}$$

$$P\{|S_\infty(\widehat{\mathcal{B}},\widehat{\omega}) - s(\widehat{\mathcal{B}},\widehat{\omega})| \le n^{B_1-(1/2)}\} \to 1, \qquad \text{(S3.20)}$$

where in (S3.19) the infimum is taken over pairs $(\mathcal{B},\omega)$ that satisfy (4.12)(ii). Since the distribution of X satisfies that constraint, and

$$f^{\rm Ft}_{Z_1,Z_2}(t_1,t_2) - f^{\rm Ft}_X(t_1+t_2) f^{\rm Ft}_{Z_1,Z_2}(-t_1,t_2),$$

then the quantity $\inf_{\mathcal{B},\omega} s(\mathcal{B},\omega)$ in (S3.19) vanishes. Therefore (S3.19) and (S3.20) imply (S3.8).

Step 2: Completion of proof of part (b) of Theorem 1. The definition of $s(\mathcal{B},\omega)$ at (S3.7) can be written equivalently as

$$s(\mathcal{B},\omega) = \sup_{-\infty<x<\infty} |f^{\rm Ft}_X(t_1+t_2) - f^{\rm Ft}_X(t_1+t_2|\mathcal{B},\omega)| f^{\rm Ft}_{Z_1,Z_2}(-t_1,t_2)| w(t_1,t_2).$$

Taking $t_1 = t$ and $t_2 = 0$ in the supremum we deduce that

$$s(\mathcal{B},\omega) \ge |f^{\rm Ft}_X(t) - f^{\rm Ft}_X(t,\mathcal{B},\omega)||f^{\rm Ft}_{Z_1,Z_2}(-t,0)| w(t,0) \text{ for all real } t. \qquad \text{(S3.21)}$$

Write F_X and $F_X(|\mathcal{B},\omega)$ for the distributions with characteristic functions $f^{\rm Ft}_X$ and $f^{\rm Ft}_X(|\mathcal{B},\omega)$, respectively. Since, by assumption (4.12)(iv), $B_3 = \sup_x f_X(x) < \infty$, then for each $B_4 > \pi^{-1}$ there exists $B_5 > 0$, depending only on B_3 and B_4, such that for all $T > 0$,

$$\sup_{-\infty<x<\infty} |F_X(x|\mathcal{B},\omega) - F_X(x)| \le B_4 \int_0^T t^{-1} |f^{\rm Ft}_X(t|\mathcal{B},\omega) - f^{\rm Ft}_X(t)| dt + \frac{B_5}{T}. \qquad \text{(S3.22)}$$

(This is Esseen's smoothing lemma; see, for example, Petrov (1975, p. 109).) By assumption (4.12)(i), both $\int |x|^{C_3} f_X(x|\mathcal{B},\omega) dx \le C_4$ and $E|X|^{C_3} < C_4$ where, without loss of generality, $0 < C_3 \le 1$. Hence there exists $B_6 > 0$, depending only on C_4, such that $|f_X^{\rm Ft}(t|\mathcal{B},\omega) - 1| \le B_6|t|^{C_3}$ and $|f_X^{\rm Ft}(t) - 1| \le B_6|t|^{C_3}$ for all real t. Combining this property with (S3.21) and (S3.22) we deduce that, if $0 \le a_n \le T = T(n)$,

$$\sup_{-\infty<x<\infty} |F_X(x|\mathcal{B},\omega) - F_X(x)| \le 2B_6 \int_0^{a_n} t^{C_4-1} dt + s(\mathcal{B},\omega)e_n \int_{a_n}^T t^{-1} dt$$
$$= 2B_6 C_4^{-1} a_n^{C_4} + s(\mathcal{B},\omega)e_n \log(T/a_n), \qquad \text{(S3.23)}$$

where

$$e_n^{-1} = \inf_{a_n \le t \le T} |f_{Z_1,Z_2}^{\rm Ft}(-t,0)|w(t,0).$$

Together (S3.8) and (S3.23) imply that, for all $B_1 > 0$,

$$P\Big\{ \sup_{-\infty<x<\infty} |F_X(x|\widehat{\mathcal{B}},\widehat{\omega}) - F_X(x)|$$
$$\le 2B_6 C_4^{-1} a_n^{C_4} + n^{B_1-(1/2)} e_n \log(T/a_n) \Big\} \to 1. \qquad \text{(S3.24)}$$

Assumptions (4.12)(iii) and (4.12)(vi) imply that $e_n \le (C_5C_7)^{-1}(1+T)^{C_6+C_8}$, and so we can choose a_n and T^{-1} to converge to zero at sufficiently slow polynomial rates to give (4.14) as a consequence of (S3.24).

Finally we derive part (a) of the theorem.

Step 1: Preparatory lemma for part (a) of Theorem 1. Define $s(\theta,\mathcal{B},\omega)$ as at (4.9), and write $(\widehat{\theta},\widehat{\mathcal{B}},\widehat{\omega})$ for the minimiser of $S_\infty(\theta,\mathcal{B},\omega)$ at (S2.3). The following result is analogous to Lemma 1.

Lemma 2. *Under the assumptions in part (a) of Theorem 1, and for each $B_1 > 0$,*

$$P\Big\{ s(\widehat{\theta},\widehat{\mathcal{B}},\widehat{\omega}) \le n^{B_1-(1/2)} \Big\} \to 1.$$

To prove the lemma, note that (S3.13) holds as before and leads directly to Lemma 2, using the argument in Step 1 of the proof of part (b) of the theorem. On this occasion the derivation uses (4.11)(i)–(4.11)(iii); the first and last of these assumptions are identical to (4.12)(i) and (4.12)(iii), respectively, and (4.11)(ii) is analogous to (4.11)(ii).

Step 2: Completion of proof of part (a) of Theorem 1. It follows directly from Lemma 2 and (4.10) that for each $B_1 > 0$,

$$P\left\{||\widehat{\theta}-\theta_0|| + \sup_{-\infty<s,t<\infty} |f_X^{\rm Ft}(s+t) - f_X^{\rm Ft}(s+t|\widehat{\mathcal{B}},\widehat{\omega})|w_1(s,t) \le n^{B_1-(1/2C_4)}\right\} \to 1.$$

In particular, $P\left(||\widehat{\theta}-\theta_0|| \le n^{\mathcal{B}_1-(1/2C_4)}\right) \to 1$ for all $B_1 > 0$, which implies the first part of (4.13), and

$$P\left\{\sup_{-\infty<s,t<\infty}\left|f_X^{\rm Ft}(s+t) - f_X^{\rm Ft}(s+t|\widehat{\mathcal{B}},\widehat{\omega})\right| w_1(s,t) \le n^{B_1-(1/2C_4)}\right\} \to 1.$$

which, by paralleling Step 2 in the proof of Theorem 1, can be shown to imply the second part of (4.13).

S3.2 Proof of Theorem 2.

Analogously to the density estimator $\widehat{f_X}$ at (3.8), define the deterministic quantity

$$\bar{f_X}(x) = \frac{1}{h}\int K\left(\frac{x-y}{h}\right) dF_X(x).$$

Now,

$$\begin{aligned} ||\widetilde{f}(x) - \bar{f}(x)||_\infty &= \sup_{-\infty<x<\infty}\left|\frac{1}{h}\int K'(u)\{F_X(x-hu|\widehat{\mathcal{B}},\widehat{\omega}) - F_X(x-hu)\}du\right| \\ &\le \frac{1}{h}\left(\int |K'|\right)\sup_{-\infty<x<\infty}|F_X(x|\widehat{\mathcal{B}},\widehat{\omega}) - F_X(x)| \\ &= O_p(n^{-\varepsilon}h^{-1}), \end{aligned} \tag{S3.25}$$

where the last identity follows from (4.13) and (4.14) in the respective cases (a) and (b). Observe too that, with C_{11} and δ as in (4.15),

$$\begin{aligned} ||\widetilde{f}(x) - f(x)||_\infty &\leq \sup_{-\infty<x<\infty} \int |K(u)||f_X(x-hu) - f_X(x)|du \\ &\leq C_{11} h^\delta \int (1+|u|)^\delta |K(u)|du = O(h^\delta). \end{aligned} \tag{S3.26}$$

Result (4.17) follows from (S3.25) and (S3.26).

S4 NUMERICAL IMPLEMENTATION

Several elements must be specified or tuned in order to implement the proposed method: the histogram bins $\mathcal{B}$, the loss function $S_q(\theta, \mathcal{B}, \omega)$ and the weight function $w(s,t)$. We suggest to use the L_2 loss

$$S_2^2(\theta, \mathcal{B}, \omega) = \int\int \left\{ \widehat{f_{LR}^{\mathrm{Ft}}}(s,t) - f_X^{\mathrm{Ft}}(s+t|\mathcal{B},\omega) f_{Z_1,Z_2}^{\mathrm{Ft}}(-s,t|\theta) \right\}^2 w(s,t)^2 ds dt \tag{S4.27}$$

because it can be minimized explicitly in ω for fixed $\mathcal{B}$; minimizing (S4.27) amounts to solving a linear system of equations in ω. For instance in the case of homogeneous Poisson with rate 1 monitoring times, where

$$f_{Z_1,Z_2}^{\mathrm{Ft}}(-s,t) = \{(1+si)(1-ti)\}^{-1} = \frac{(1+st) + i(t-s)}{(1+st)^2 + (t-s)^2}, \tag{S4.28}$$

straightforward manipulations show that $\widehat{\omega} = M_1^{-1} M_2$ where M_1 is a $K \times K$ matrix with element in position (k,j) given by

$$\begin{aligned} [M_1]_{j,k} = \int\int \frac{w^2(s,t)}{\{(1+s^2)(1+t^2)\}^2} \Big[& \Psi_1(s,t,j)\Upsilon_1(s,t,k)(1+st) - \Psi_2(s,t,j)\Upsilon_1(s,t,k)(t-s) \\ & - \Psi_2(s,t,j)\Upsilon_2(s,t,k)(1+st) - \Psi_1(s,t,j)\Upsilon_2(s,t,k)(t-s) \Big] ds dt \end{aligned}$$

and M_2 is a K-vector with element in position k given by

$$[M_2]_k = \int\int \frac{w^2(s,t)}{(1+s^2)(1+t^2)} \frac{1}{n} \sum_{\ell=1}^{n} \Big\{ \cos(sL_\ell + tR_\ell)\Upsilon_1(s,t,k) - \sin(sL_\ell + tR_\ell)\Upsilon_2(s,t,k) \Big\} ds dt,$$

where

$$\Psi_1(s,t,k) = \frac{\sin\{B_k^U(s+t)\} - \sin\{B_k^L(s+t)\}}{s+t}$$

$$\Psi_2(s,t,k) = \frac{\cos\{B_k^L(s+t)\} - \cos\{B_k^U(s+t)\}}{s+t}$$

$$\Upsilon_1(s,t,k) = \Psi_2(s,t,k)(t-s) - \Psi_1(s,t,k)(1+st)$$

$$\Upsilon_2(s,t,k) = \Psi_2(s,t,k)(1+st) + \Psi_1(s,t,k)(t-s)$$

and B_k^L and B_k^U respectively denote the lower and upper bounds of the k-th bin of $\mathcal{B}$.

The choice of the weight function $w(s,t)$ has to respect the regularity conditions outlined in Section 4. We obtained good results in our trials when using $w(s,t) = \{(1+|s|)(1+|t|)\}^{-p}$ with $p = 5$. Such a high power ensures that sufficient weight is given to observations in the neighborhood of $(s,t) = (0,0)$ where $\widehat{f_{LR}^{\mathrm{Ft}}}(s,t)$ is more accurate.

The specification of the histogram bins $\mathcal{B}$ can either be done on a trial and error basis, or by finding $\mathcal{B}$ (for a fixed number of bins) that minimizes S_2 if one has an efficient algorithm to do so. The former option is relatively simple to implement: a good choice of $\mathcal{B}$ will yield a histogram that looks reasonable while a poor choice will generate some of the histogram heights, ω, to be negative. A compromise that worked well in our trials consists in minimizing S_2 over a small number of bin sets. Naive automated ways of specifying the bin sets include using J bins of identical size or using the quantiles of the midpoints of the observed or innermost intervals; in our trials the first two of these three options yielded good results. For each bin set, we estimated the bin heights $\omega_1, \ldots, \omega_J$ by solving the linear equations that yield the values of the ω's that minimize the discretized loss function

$$S_2^2(\omega, \mathcal{B}) = \sum_{(s,t)\in\mathcal{G}} w^2(s,t) \Bigg\| n^{-1} \sum_{\ell=1}^{n} e^{i(sL_\ell + tR_\ell)} - \left\{ \sum_{k=1}^{J} \omega_k \int_{B_k(\mathcal{B})} e^{i(s+t)u} du \right\} \{(1+is)(1-it)\}^{-1} \Bigg\|^2, \tag{S4.29}$$

where $B_k(\mathcal{B})$ are the bins and $\mathcal{G}$ is a grid of (s,t) points, for instance $\mathcal{G} = \{(s,t) : -5 \leq s,t \leq 5\}$ a grid of size $40,000 = 200 \times 200$ of values of (s,t).

ADDITIONAL REFERENCES

DEHEUVELS, P. (1977). Estimation Nonparamétrique de la Densité par Histogrammes Généralisés. *Revue Statistique Appliquée 25* 5–42.

KARLIN, S. AND TAYLOR, H.M. (1975). *A First Course in Stochastic Processes*, 2nd Edn. Academic Press, New York.

PETROV, V.V. (1975). *Sums of Independent Random Variables.* Springer, Berlin.