Online Appendix to 'Is Newey-West Optimal Among First-Order Kernels?'

Thomas Kolokotrones* James H. Stock† Christopher D. Walker‡

June 22, 2022

1 Proofs of Propositions and Lemmas in Main Text

1.1 Results for Section 2

1.1.1 Proof of Lemma 1 in the Main Text

Proof. The proof has several steps. For that reason, we provide an outline of the steps and then prove them in detail.

Outline. The proof has three steps. **Step 1** demonstrates that if f is a symmetric, positive-definite function, then $f(0) \geq 0$, $|f(t)| \leq f(0)$, and $|f(t+h) - f(t)| \leq \sqrt{2f(0)|f(t) - f(0)|}$. **Step 2** demonstrates that if $f^{[2]}(0)$ and $f'(t)$ exist, then $|f'(t)| \leq \sqrt{2f(0)f^{[2]}(0)}$. **Step 3** establishes that continuity of $f(t)$ at $t = 0$ implies everywhere uniform continuity. **Step 4** establishes that if $f''(t)$ exists as well, then $|f''(t)| \leq 2f^{[2]}(0)$.

Step 1. Let f be a symmetric and positive-semidefinite function, M_n be the $n \times n$ matrix with entries $M_{n,ij} = f(t_j - t_j)$, and $\{e_i\}_{i=1}^n$ be the standard basis for R^n. Assume that $n \geq 3$. Since f is positive-semidefinite, the matrix M_n is positive-semidefinite. Consequently, $v^t M_n v \geq 0$ for any $v \in R^n$. Setting $v = e_1$, we obtain $e_i^t M_n e_i = f(t_i - t_i) = f(0) \geq 0$. Next, set $v = e_2 \pm e_3$. We obtain $v^t M_n v = 2f(0) \pm f(t_2 - t_3) \pm f(t_3 - t_2) = 2\left(f(0) \pm f(t_2 - t_3)\right)$ (by symmetry). As $0 \leq v^t M_n v$, it follows that $f(0) \pm f(t_2 - t_3) \geq 0$ and therefore $|f(t_2 - t_3)| \leq f(0)$. Set $t_2 - t_3 = t$ to obtain that $|f(t)| \leq f(0)$.

Since M_n is positive-semidefinite and symmetric, it possesses a (symmetric, positive-semidefinite) square root, $M_n^{\frac{1}{2}}$, so a version of the Cauchy-Schwartz Inequality holds for the bilinear form associated with M_n. For any $u, v \in R^n$,

$$|u^t M_n v| = \left|\left(M_n^{\frac{1}{2}} u\right)^t \left(M_n^{\frac{1}{2}} v\right)\right| \leq \sqrt{\left(M_n^{\frac{1}{2}} u\right)^t \left(M_n^{\frac{1}{2}} u\right)} \sqrt{\left(M_n^{\frac{1}{2}} v\right)^t \left(M_n^{\frac{1}{2}} v\right)} = \left(u^t M_n u\right)^{\frac{1}{2}} \left(v^t M_n v\right)^{\frac{1}{2}}$$

*Harvard University, Departments of Biostatistics and Epidemiology.
†Harvard University, Department of Economics.
‡Harvard University, Department of Economics.

where the inequality is the standard Cauchy-Schwarz inequality. Setting $u = e_1$ and $v = e_2 - e_3$, we obtain

$$\begin{aligned}|f(t_2 - t_1) - f(t_3 - t_1)| = |e_1^t M_n(e_2 - e_3)| &\leq \left(e_1^t M_n e_1\right)^{\frac{1}{2}} \left((e_2 - e_3)^t M_n(e_2 - e_3)\right)^{\frac{1}{2}} \\ &= \sqrt{2f(0)\,(f(0) - f(t_3 - t_2))}\end{aligned}$$

Setting $t_1 = 0, t_2 = t, t_3 = t+h$ yields $|f(t+h)-f(t)| \leq \sqrt{2f(0)\,(f(0) - f(h))} = \sqrt{2f(0)\,|f(h) - f(0)|}$ because $f(0) \geq f(h)$.

Step 2. Suppose that $f'(t)$ and $f^{[2]}(0)$ exist, then

$$|f'(t)| = \left|\lim_{h\to 0} \frac{f(t+h) - f(t)}{h}\right| \leq \sqrt{2f(0)\left|\lim_{h\to 0} \frac{f(0) - f(h)}{|h|^2}\right|} = \sqrt{2f(0)f^{[2]}(0)}.$$

Step 3. Assume that $f(0) > 0$. This is without loss of generality because $f(0) = 0$ and $|f(t)| \leq f(0)$ implies that $f(t) = 0$ for all t and therefore trivially uniformly continuous. In the case where $f(0) > 0$ and f is continuous at 0, we have that for all $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(h) - f(0)| < 2(f(0))^{-1}\varepsilon^2$ for $|h| < \delta$. Then, for $|h| < \delta$, $|f(t+h) - f(t)| < \varepsilon$, so f is uniformly continuous because δ does not depend on t.

Step 4. Finally,

$$\begin{aligned}&|f(t_3 - t_1) - f(t_4 - t_1) - f(t_3 - t_2) + f(t_4 - t_2)| = \left|(e_1 - e_2)^t M_n (e_3 - e_4)\right| \\ &\leq \left((e_1 - e_2)^t M_n (e_1 - e_2)\right)^{\frac{1}{2}} \left((e_3 - e_4)^t M_n (e_3 - e_4)\right)^{\frac{1}{2}} \\ &= 2\left(f(0) - f(t_2 - t_1)\right)^{\frac{1}{2}} \left(f(0) - f(t_4 - t_3)\right)^{\frac{1}{2}}\end{aligned}$$

Taking $t_1 = 0, t_2 = h, t_3 = t+h, t_4 = t+h+l$ yields,

$$|(f(t+h+l) - f(t+h)) - (f(t+l) - f(t))| \leq 2(f(0) - f(h))^{\frac{1}{2}}(f(0) - f(l))^{\frac{1}{2}}$$

Then, if $f''(t)$ and $f^{[2]}(0)$ exist,

$$\begin{aligned}\left|f''(t)\right| &= \left|\lim_{h\to 0} \frac{f'(t+h) - f'(t)}{h}\right| = \lim_{h\to 0} |h|^{-1} \left|\lim_{l\to 0} \frac{f(t+h+l) - f(t+h)}{l} - \lim_{l\to 0} \frac{f(t+l) - f(h)}{l}\right| \\ &\leq 2\left(\lim_{h\to 0} \frac{f(0) - f(h)}{|h|^2}\right)^{\frac{1}{2}} \left(\lim_{l\to 0} \frac{f(0) - f(l)}{|l|^2}\right)^{\frac{1}{2}} = 2f^{[2]}(0)\end{aligned}$$

This completes the proof. □

1.1.2 Proof of Lemma 2 in the Main Text

Proof. The proof requires verifying each of the statements directly.

1. We consider two cases: $q < \infty$ and $q = \infty$.

   (a) If $q < \infty$, $f^{[q]}(0) < +\infty$. Then, for any $r < q$, $f^{[r]}(0) = \lim_{t\to 0} |t|^{-r} (f(0) - f(t)) = \lim_{t\to 0} |t|^{q-r} \cdot \lim_{t\to 0} |t|^{-q} (f(0) - f(t)) = 0 \cdot f^{[q]}(0) = 0$.

(b) If $q = \infty$, then, by definition, for all r, $f^{[r]}(0) < \infty$. Let $r' < r$, then $f^{[r']}(0) = \lim_{t \to 0} |t|^{r-r'} \cdot \lim_{t \to 0} |t|^{-r} \left(f(0) - f(t) \right) = 0 \cdot f^{[r]}(0) = 0$. Since r, r' are arbitrary, for any $r < q = \infty$, $f^{[r]}(0) = 0$

2. Suppose that f is constant. Then $f^{[r]}(0) = \lim_{t \to 0} |t|^{-r} \left(f(0) - f(t) \right) = 0$ for all r and therefore for $r = 2$. If $f^{[2]}(0) = 0$, we can conclude from Lemma 1 that $\limsup_{h \to 0} \left| \frac{f(t+h) - f(t)}{h} \right| \leq 2^{\frac{1}{2}} f(0)^{\frac{1}{2}} f^{[2]}(0)^{\frac{1}{2}} = 0$. Thus, f' exists for all t and $f' = 0$, so f is constant. Therefore, f is constant if and only if $f^{[2]}(0) = 0$.

3. If f has characteristic exponent $q > 0$, then $\lim_{h \to 0} |f(h) - f(0)| = \lim_{h \to 0} |h|^q |h|^{-q} |f(h) - f(0)| = 0 \cdot f^{[q]}(0) = 0$, so f is continuous at 0, and thus by Lemma 1, f is uniformly continuous.

4. If $q = 2$, $f^{[2]}(0)$ is defined and finite. If $f^{[2]}(0) = 0$, then f is constant by Part 2 of the lemma so $f^{[r]}(0) = 0$ for all r, and, by definition $q = \infty > 2$. Thus, if $q = 2$, $f^{[2]}(0) > 0$.

□

1.1.3 Proof of Proposition 3 in the Main Text

Proof. If $q > 2$, then $f^{[2]}(0) = 0$ by Part 1 of Lemma 2. Then, we apply Part 2 of Lemma 2 to conclude that f is a constant function. If f is constant, then $f^{[r]}(0) = \lim_{t \to 0} |t|^{-r}(f(0) - f(t)) = 0$ for all r and, thus $q = \infty$. Trivially, if $q = \infty$, $q > 2$. Thus, the statements, $q > 2$, $q = \infty$, and f is constant are all equivalent. Therefore, either $q \leq 2$ or $q = \infty$. Additionally, f is constant if and only if $q = \infty$. □

1.2 Proof of Proposition 6 in the Main Text

Proof. The proof of the proposition is long. For that reason, we outline the argument and then complete all of the arguments rigorously.

Outline. The proof has four steps. In **Step 1**, we show that $k^{[1]}(0) = \frac{1}{2} \left(\phi^2(0) + \phi^2(1) \right)$ thereby verifying the first statement in the proposition. **Step 2** then uses mathematical induction to establish expressions for $\frac{d^n k}{dt^n}(t)$ for $t > 0$ and $t < 0$ if $\phi \in \mathcal{C}^n[0, 1]$. In **Step 3**, we use the result from **Step 2** to characterize $\lim_{t \to 0^+} \frac{d^n k}{dt^n}(t)$ and $\lim_{t \to 0^-} \frac{d^n k}{dt^n}(t)$ for n odd and n even. **Step 4** establishes that if $k \in \mathcal{C}^{n-1}[-1, 1]$, considers one-sided derivatives at 0 and verifies the third statement of the proposition. Finally, **Step 5** verifies the claim about second-order kernels.

Step 1. We first show that $k^{[1]}(0) = \frac{1}{2}(\phi^2(0) + \phi^2(1))$. We show the result using the direct method. That is, we establish $\lim_{t \to 0^+} |t|^{-1}(1 - k(t)) = \lim_{t \to 0^-} |t|^{-1}(1 - k(t)) = \frac{1}{2}(\phi^2(1) + \phi^2(0))$. To that end, let $\frac{d}{dt^\pm}$ denote the right and left sided derivatives, respectively, so that $\frac{df}{dt^\pm} = \lim_{h\to 0^\pm} \frac{f(t+h)-f(t)}{h}$ then

$$\begin{aligned}
\lim_{t\to 0^+} |t|^{-1}(1-k(t)) &= \lim_{t\to 0^+} |t|^{-1}(k(0)-k(t)) = -\lim_{t\to 0^+} t^{-1}(k(t)-k(0)) = -\frac{dk}{dt^+}(0) \\
&= -\frac{d}{dt^+}\left[\int_{\max(0,t)}^{\min(1,1+t)} \phi(u)\phi(u-t)du\right]_{t=0} \\
&= -\frac{d}{dt^+}\left[\int_t^1 \phi(u)\phi(u-t)du\right]_{t=0} \\
&= -\left[-\phi(t)\phi(0) + \int_t^1 -\phi(u)\phi'(u-t)du\right]_{t=0} \\
&= \phi(0)^2 + \int_0^1 \phi(u)\phi'(u)du = \phi(0)^2 + \frac{1}{2}\phi(u)^2\Big|_0^1 \\
&= \phi(0)^2 + \frac{1}{2}\left(\phi(1)^2 - \phi(0)^2\right) \\
&= \frac{1}{2}\left(\phi(0)^2 + \phi(1)^2\right)
\end{aligned}$$

and

$$\begin{aligned}
\lim_{t\to 0^-} |t|^{-1}(1-k(t)) &= \lim_{t\to 0^-} |t|^{-1}(k(0)-k(t)) = \lim_{t\to 0^-} (-t)^{-1}(k(0)-k(t)) \\
&= \lim_{t\to 0^-} t^{-1}(k(t)-k(0)) = \frac{dk}{dt^-}(0) \\
&= \frac{d}{dt^-}\left[\int_{\max(0,t)}^{\min(1,1+t)} \phi(u)\phi(u-t)du\right]_{t=0} \\
&= \frac{d}{dt^-}\left[\int_0^{1+t} \phi(u)\phi(u-t)du\right]_{t=0} \\
&= \left[\phi(1+t)\phi(1) + \int_0^{1+t} -\phi(u)\phi'(u-t)du\right]_{t=0} \\
&= \phi(1)^2 - \int_0^1 \phi(u)\phi'(u)du = \phi(1)^2 - \frac{1}{2}\phi(u)^2\Big|_0^1 \\
&= \phi(1)^2 - \frac{1}{2}\left(\phi(1)^2 - \phi(0)^2\right) \\
&= \frac{1}{2}\left(\phi(0)^2 + \phi(1)^2\right).
\end{aligned}$$

So the limit exists and $k^{[1]}(0) = \lim_{t\to 0} |t|^{-1}(1-k(t)) = \frac{1}{2}(\phi(0)+\phi(1))^2$.

Step 2. We use induction to find expressions for $\frac{d^n k}{dt^n}(t)$ for $t > 0$ and $t < 0$ when $\phi \in \mathcal{C}^n[0,1]$. First assume that $t > 0$ and consider the hypothesis that

$$\frac{d^n k}{dt^n}(t) = \sum_{i=1}^{n} (-1)^i \phi^{(n-i)}(t)\phi^{(i-1)}(0) + (-1)^n \int_t^1 \phi(u)\phi^{(n)}(u-t)du.$$

From the above, this clearly holds for $n = 1$. Now assume that it holds for all $n' < n$, then,

$$\begin{aligned}
\frac{d^n k}{dt^n}(t) = \frac{d}{dt}\frac{d^{n-1}k}{dt^{n-1}} &= \frac{d}{dt}\left[\sum_{i=1}^{n-1}(-1)^i\phi^{(n-1-i)}(t)\phi^{(i-1)}(0) + (-1)^{n-1}\int_t^1 \phi(u)\phi^{(n-1)}(u-t)du\right] \\
&= \sum_{i=1}^{n-1}(-1)^i\phi^{(n-i)}(t)\phi^{(i-1)}(0) + (-1)^{n-1}\left[\frac{d}{dt}\int_t^1 \phi(u)\phi^{(n-1)}(u-t)du\right] \\
&= \sum_{i=1}^{n-1}(-1)^i\phi^{(n-i)}(t)\phi^{(i-1)}(0) + (-1)^{n-1}\left[-\phi(t)\phi^{(n-1)}(0) - \int_t^1 \phi(u)\phi^{(n)}(u-t)du\right] \\
&= \sum_{i=1}^{n-1}(-1)^i\phi^{(n-i)}(t)\phi^{(i-1)}(0) + (-1)^n\phi^{(n-n)}(t)\phi^{(n-1)}(0) + (-1)^n\int_t^1 \phi(u)\phi^{(n)}(u-t)du \\
&= \sum_{i=1}^{n}(-1)^i\phi^{(n-i)}(t)\phi^{(i-1)}(0) + (-1)^n\int_t^1 \phi(u)\phi^{(n)}(u-t)du
\end{aligned}$$

Thus, the hypothesis holds for n, so, by induction, the theorem holds for all $n \in N$, with $t > 0$. Now assume that $t < 0$ and consider the hypothesis that

$$\frac{d^n k}{dt^n}(t) = \sum_{i=1}^{n}(-1)^{i-1}\phi^{(n-i)}(1+t)\phi^{(i-1)}(1) + (-1)^n\int_0^{1+t}\phi(u)\phi^{(n)}(u-t)du.$$

From the above, this clearly holds for $n = 1$. Now assume that it holds for all $n' < n$, then,

$$\begin{aligned}
\frac{d^n k}{dt^n}(t) = \frac{d}{dt}\frac{d^{n-1}k}{dt^{n-1}} &= \frac{d}{dt}\left[\sum_{i=1}^{n-1}(-1)^{i-1}\phi^{(n-1-i)}(1+t)\phi^{(i-1)}(1) + (-1)^{n-1}\int_0^{1+t}\phi(u)\phi^{(n-1)}(u-t)du\right] \\
&= \sum_{i=1}^{n-1}(-1)^{i-1}\phi^{(n-i)}(1+t)\phi^{(i-1)}(1) + (-1)^{n-1}\left[\frac{d}{dt}\int_0^{1+t}\phi(u)\phi^{(n-1)}(u-t)du\right] \\
&= \sum_{i=1}^{n-1}(-1)^{i-1}\phi^{(n-i)}(1+t)\phi^{(i-1)}(1) + (-1)^{n-1}\left[\phi(1+t)\phi^{(n-1)}(1) - \int_0^{1+t}\phi(u)\phi^{(n)}(u-t)du\right] \\
&= \sum_{i=1}^{n-1}(-1)^{i-1}\phi^{(n-i)}(1+t)\phi^{(i-1)}(1) + (-1)^{n-1}\phi^{(n-n)}(1+t)\phi^{(n-1)}(1) + (-1)^n\int_0^{1+t}\phi(u)\phi^{(n)}(u-t)du \\
&= \sum_{i=1}^{n}(-1)^{i-1}\phi^{(n-i)}(1+t)\phi^{(i-1)}(1) + (-1)^n\int_0^{1+t}\phi(u)\phi^{(n)}(u-t)du
\end{aligned}$$

Thus, the hypothesis holds for n, so, by induction, it holds for for all $n \in N$ with $t < 0$.

Step 3. Next, we compute $\lim_{t\to 0^+}\frac{d^n k}{dt^n}(t)$ and $\lim_{t\to 0^-}\frac{d^n k}{dt^n}(t)$ for n odd and n even. Observe that

$$\begin{aligned}
\lim_{t\to 0^+}\frac{d^n k}{dt^n}(t) &= \lim_{t\to 0^+}\left[\sum_{i=1}^{n}(-1)^i\phi^{(n-i)}(t)\phi^{(i-1)}(0) + (-1)^n\int_t^1\phi(u)\phi^{(n)}(u-t)du\right] \\
&= \sum_{i=1}^{n}(-1)^i\phi^{(n-i)}(0)\phi^{(i-1)}(0) + (-1)^n\int_0^1\phi(u)\phi^{(n)}(u)du
\end{aligned}$$

and

$$\lim_{t\to 0^-}\frac{d^nk}{dt^n}(t) = \lim_{t\to 0^-}\left[\sum_{i=1}^{n}(-1)^{i-1}\phi^{(n-i)}(1+t)\phi^{(i-1)}(1) + (-1)^n\int_0^{1+t}\phi(u)\phi^{(n)}(u-t)du\right]$$
$$= \sum_{i=1}^{n}(-1)^{i-1}\phi^{(n-i)}(1)\phi^{(i-1)}(1) + (-1)^n\int_0^{1}\phi(u)\phi^{(n)}(u)du.$$

We will use integration by parts in order to reexpress $\int_0^1 \phi(u)\phi^{(n)}(u)du$.

$$\begin{aligned}\int_0^1\phi(u)\phi^{(n)}(u)du &= \phi(u)\phi^{(n-1)}(u)\Big|_0^1 - \int_0^1\phi^{(1)}(u)\phi^{(n-1)}(u)du\\
&= \phi(1)\phi^{(n-1)}(1) - \phi(0)\phi^{(n-1)}(0) - \int_0^1\phi^{(1)}(u)\phi^{(n-1)}(u)du\\
&= \phi(1)\phi^{(n-1)}(1) - \phi(0)\phi^{(n-1)}(0) - \phi^{(1)}(u)\phi^{(n-2)}(u)\Big|_0^1 + \int_0^1\phi^{(2)}(u)\phi^{(n-2)}(u)du\\
&= \sum_{i=1}^{2}(-1)^{i-1}\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right] + (-1)^2\int_0^1\phi^{(2)}(u)\phi^{(n-2)}(u)du\\
&\dots\\
&= \sum_{i=1}^{n}(-1)^{i-1}\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right] + (-1)^n\int_0^1\phi^{(n)}(u)\phi(u)du\end{aligned}$$

which implies

$$\left(1-(-1)^n\right)\int_0^1\phi(u)\phi^{(n)}(u)du = \sum_{i=1}^{n}(-1)^{i-1}\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right]$$

For n odd,

$$\int_0^1\phi(u)\phi^{(n)}(u)du = \frac{1}{2}\sum_{i=1}^{n}(-1)^{i-1}\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right]$$

so

$$\begin{aligned}\lim_{t\to 0^+}\frac{d^nk}{dt^n}(t) &= \sum_{i=1}^{n}(-1)^{i}\phi^{(n-i)}(0)\phi^{(i-1)}(0) + (-1)^n\int_0^1\phi(u)\phi^{(n)}(u)du\\
&= \sum_{i=1}^{n}(-1)^{i}\phi^{(n-i)}(0)\phi^{(i-1)}(0) - \frac{1}{2}\sum_{i=1}^{n}(-1)^{i-1}\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right]\\
&= \sum_{i=1}^{n}(-1)^{i}\phi^{(n-i)}(0)\phi^{(i-1)}(0) + \frac{1}{2}\sum_{i=1}^{n}(-1)^{i}\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right]\\
&= \frac{1}{2}\sum_{i=1}^{n}(-1)^{i}\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) + \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right]\end{aligned}$$

$$
\begin{aligned}
\lim_{t\to 0^-} \frac{d^n k}{dt^n}(t) &= \sum_{i=1}^{n} (-1)^{i-1} \phi^{(n-i)}(1)\phi^{(i-1)}(1) + (-1)^n \int_0^1 \phi(u)\phi^{(n)}(u)du \\
&= \sum_{i=1}^{n} (-1)^{i-1} \phi^{(n-i)}(1)\phi^{(i-1)}(1) - \frac{1}{2}\sum_{i=1}^{n} (-1)^{i-1} \left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right] \\
&= \frac{1}{2}\sum_{i=1}^{n} (-1)^{i-1} \left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) + \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right] \\
&= -\frac{1}{2}\sum_{i=1}^{n} (-1)^{i} \left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) + \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right] \\
&= -\lim_{t\to 0^+} \frac{d^n k}{dt^n}(t)
\end{aligned}
$$

For n even,

$$
\begin{aligned}
0 &= \sum_{i=1}^{n} (-1)^{i-1} \left[\phi^{(i-1)}(1)\phi^{(n-i)}(1) - \phi^{(i-1)}(0)\phi^{(n-i)}(0)\right] \\
&= \sum_{i=1}^{n} (-1)^{i-1} \phi^{(i-1)}(1)\phi^{(n-i)}(1) + \sum_{i=1}^{n} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0)
\end{aligned}
$$

so

$$
\sum_{i=1}^{n} (-1)^{i-1} \phi^{(i-1)}(1)\phi^{(n-i)}(1) = -\sum_{i=1}^{n} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0)
$$

Note that, for n even,

$$
\begin{aligned}
&\sum_{i=1}^{n} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0) \\
&= \sum_{i=1}^{\frac{n}{2}} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0) + \sum_{i=\frac{n}{2}+1}^{n} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0) \\
&= \sum_{i=1}^{\frac{n}{2}} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0) + \sum_{i=\frac{n}{2}+1}^{n} (-1)(-1)(-1)^{n}(-1)^{-i} \phi^{(n-(n-i+1))}(0)\phi^{((n-i+1)-1)}(0) \\
&= \sum_{i=1}^{\frac{n}{2}} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0) - \sum_{i=\frac{n}{2}+1}^{n} (-1)^{n-i+1} \phi^{(n-(n-i+1))}(0)\phi^{((n-i+1)-1)}(0) \\
&= \sum_{i=1}^{\frac{n}{2}} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0) - \sum_{j=1}^{\frac{n}{2}} (-1)^{j} \phi^{(n-j)}(0)\phi^{(j-1)}(0) \\
&= 0
\end{aligned}
$$

where $j = n - i + 1$. Thus,

$$
\sum_{i=1}^{n} (-1)^{i-1} \phi^{(i-1)}(1)\phi^{(n-i)}(1) = -\sum_{i=1}^{n} (-1)^{i} \phi^{(i-1)}(0)\phi^{(n-i)}(0) = 0
$$

so

$$\lim_{t\to 0^+}\frac{d^n k}{dt^n}(t)=\lim_{t\to 0^-}\frac{d^n k}{dt^n}(t)=\int_0^1 \phi(u)\phi^{(n)}(u)du.$$

for n even. In summary, for n odd,

$$\lim_{t\to 0^+}\frac{d^n k}{dt^n}(t)=\frac{1}{2}\sum_{i=1}^{n}(-1)^i\left[\phi^{(i-1)}(1)\phi^{(n-i)}(1)+\phi^{(i-1)}(0)\phi^{(n-i)}(0)\right]=-\lim_{t\to 0^-}\frac{d^n k}{dt^n}(t)$$

and, for n even,

$$\lim_{t\to 0^+}\frac{d^n k}{dt^n}(t)=\lim_{t\to 0^-}\frac{d^n k}{dt^n}(t)=\int_0^1 \phi(u)\phi^{(n)}(u)du.$$

This verifies the expressions in the statement of the proposition.

Step 4. Suppose that $k\in\mathcal{C}^{n-1}[-1,1]$ and $\phi\in\mathcal{C}^n[0,1]$. Then the above results also hold for the one-sided n^{th} derivatives at 0, as well. In particular, since $\phi\in\mathcal{C}^n[0,1]$, then $\frac{d^n k}{dt^n}(t)$ will be continuous on $[-1,1]$, except, perhaps, at 0; since ϕ and its first n derivatives are continuous, and, thus uniformly bounded on $[0,1]$, this follows from an application of the Dominated Convergence Theorem to the two expressions above. Thus, if n is even, $\phi\in\mathcal{C}^n[0,1]$, and $k\in\mathcal{C}^{n-1}[-1,1]$, then, from the above, $\frac{d^n k}{dt^{n\pm}}(0)=\lim_{t\to 0^\pm}\frac{d^n k}{dt^n}(t)$ so $k\in\mathcal{C}^n[-1,1]$ and $\frac{d^n k}{dt^n}(0)=\int_0^1\phi(u)\phi^{(n)}(u)du$.

Step 5. Finally, suppose that ϕ generates a second-order kernel. Lemma 2 establishes that $k^{[1]}(0)=0$, so the left and right derivatives of k are both 0 at 0 and $k\in\mathcal{C}^1[-1,1]$. Then, $k\in\mathcal{C}^2[-1,1]$ and $k^{[2]}(0)=-\frac{1}{2}\frac{d^2k}{dt^2}(0)=-\frac{1}{2}\int_0^1\phi(u)\phi^{(2)}(u)du$. This establishes the last part of the proposition. □

2 Additional Theoretical Discussion

2.1 Structure of Positive-Semidefinite Functions

2.1.1 General Decomposition of Positive-Semidefinite Kernels

We now use Fourier Analysis in order to more fully characterize the structure of symmetric, positive-semidefinite kernels. Henceforth, we will only consider continuous kernels, but, as noted previously, due to Lemmas 1 and 2 in the main text, this will occur whenever the kernel is continuous at the origin or when $q>0$, which is the case for all kernels of interest, so this restriction has no practical impact. We will also assume that all kernels are symmetric and normalized with $k(0)=1$, even if this is not explicitly stated. Recall that $\mathcal{F}$ denotes the Fourier transform with $\hat{f}(\omega)=(\mathcal{F}f)(\omega)=(2\pi)^{-1}\int_{-\infty}^{\infty}e^{-i\omega t}f(t)dt$ if $f\in L^1$, so that, if $\hat{f}\in L^1$, $f(t)=\left(\mathcal{F}^{-1}\hat{f}\right)(t)=\int_{-\infty}^{\infty}e^{i\omega t}\hat{f}(\omega)d\omega$. Several basic results from Fourier analysis are widely used in the classical spectral estimation literature (e.g., Priestley [1981]), such as the fact that, if $f,f'\in L^1$, then $\mathcal{F}(f')(\omega)=i\omega\hat{f}(\omega)$, which can be easily shown using integration by parts and can be iterated to give $\mathcal{F}(f^{(n)})(\omega)=(i\omega)^n\hat{f}(\omega)$. For completeness, we present similar results for the (inverse) Fourier Transform of a finite measure μ, since they are less commonly encountered. In order to simplify notation in what follows, if f is μ-measurable, we will sometimes denote the measure corresponding to $d\nu=fd\mu$ as $\nu=f\mu$ instead.

Lemma 7. *Let μ be a finite Borel measure (possibly signed or complex). Then, if $k = \mathcal{F}^{-1}\mu$, k is uniformly continuous.*

Proof. The result follows from an application of the Dominated Convergence Theorem. Using the definition of the inverse fourier transform, it follows that

$$\begin{aligned} |\mathcal{F}^{-1}(\mu)(t+h) - \mathcal{F}^{-1}(\mu)(t)| &= \left| \int e^{i\omega(t+h)} d\mu(\omega) - \int e^{i\omega t} d\mu(\omega) \right| \\ &= \left| \int e^{i\omega t} \left(e^{i\omega h} - 1 \right) d\mu(\omega) \right| \\ &\leq \int |e^{i\omega t}| \left| e^{i\omega h} - 1 \right| d|\mu|(\omega) \\ &= \int \left| e^{i\omega h} - 1 \right| d|\mu|(\omega) \end{aligned}$$

where the least equality holds because $|e^{i\omega t}| = 1$ and $|\mu|$ is the total variation measure associated for μ. If μ is a positive Borel measure then $|\mu| = \mu$; if μ is a signed measure with Jordan decomposition $\mu = \mu_+ - \mu_-$, then $|\mu| = \mu_+ + \mu_-$. Now, $|e^{i\omega h} - 1| \leq |e^{i\omega h}| + 1 = 2$ for all h, and $e^{i\omega h} - 1 \to 0$ as $h \to 0$. Consequently, we apply the Dominated Convergence Theorem to conclude that $\lim_{h\to 0} \int \left| e^{i\omega h} - 1 \right| d|\mu|(\omega) = 0$. As such, for any $\varepsilon > 0$, we can find a $\delta > 0$ such that for $|h| < \delta$, $|\mathcal{F}^{-1}(\mu)(t+h) - \mathcal{F}^{-1}(\mu)(t)| < \varepsilon$. Thus, $\mathcal{F}^{-1}(\mu)$ is continuous and, since the choice of δ only depends on ε, it is uniformly continuous. $\square$

Lemma 8. *Let μ be a finite Borel measure (possibly signed or complex) with $\omega \in L^1(\mu)$, then, if $k = \mathcal{F}^{-1}\mu$, k is continuously differentiable with $k' = i\mathcal{F}^{-1}(\omega\mu)$.*

Proof. Using the definition of $\mathcal{F}^{-1}(\mu)$ and the fact that $k = \mathcal{F}^{-1}(\mu)$, we have that

$$\frac{k(t+h) - k(t)}{h} = \frac{\mathcal{F}^{-1}(\mu)(t+h) - \mathcal{F}^{-1}(\mu)(t)}{h} = \int h^{-1} e^{i\omega t} \left(e^{i\omega h} - 1 \right) d\mu(\omega).$$

We want to apply the Dominated Convergence Theorem. To that end, we note that

$$\begin{aligned} \left| h^{-1} e^{i\omega t} \left(e^{i\omega h} - 1 \right) \right| &= |h|^{-1} |e^{i\omega t}| |e^{i\omega h} - 1| \\ &= |h|^{-1} |e^{i\omega h} - 1| \\ &= |h|^{-1} \sqrt{2 - e^{i\omega h} - e^{-i\omega h}|} \\ &= |h|^{-1} \sqrt{2 - 2\cos(\omega h)} \\ &= |h|^{-1} \sqrt{\left(2 - 2\left(1 - \frac{1}{2}\cos(\omega h^*)(\omega h)^2 \right) \right)} \\ &= |h|^{-1} \sqrt{\cos(\omega h^*)(\omega h)^2} \leq |h|^{-1} (1 \cdot (\omega h)^2)^{\frac{1}{2}} \\ &= |h|^{-1} |\omega h| \\ &= |\omega|, \end{aligned}$$

where the fourth equality is due to Taylor's Theorem, with $h^* \in [0, h]$, and the inequality is because $|\cos(x)| \leq 1$. Thus, $|h^{-1}e^{i\omega t}(e^{i\omega h} - 1)| \leq |\omega|$. Next,

$$\lim_{h\to 0} \frac{e^{i\omega(t+h)} - e^{i\omega t}}{h} = \frac{d}{dt}e^{i\omega t} = i\omega e^{i\omega t}$$

Consequently, Dominated Convergence Theorem implies that

$$\lim_{h\to 0} \int h^{-1}e^{i\omega t}\left(e^{i\omega h} - 1\right) d\mu(\omega) = i\int \omega e^{i\omega t} d\mu(\omega) = i\mathcal{F}^{-1}(\omega\mu)(t).$$

which implies that $k = \mathcal{F}^{-1}(\mu)$ is differentiable with derivative equal to $i\mathcal{F}^{-1}(\omega\mu)$. To establish that $\mathcal{F}^{-1}(\omega\mu)$ is continuous, we note that, since $\omega \in L^1(\mu)$, $\omega\mu$ is a finite (possibly signed or complex) Borel measure. Therefore, Lemma 7 shows that $\mathcal{F}^{-1}(\omega\mu)$ is uniformly continuous. □

Proposition 9. *Let μ be a finite Borel measure (possibly signed or complex) and $\omega^n \in L^1(\mu)$, then, if $k = \mathcal{F}^{-1}\mu$, $k \in \mathcal{C}^n$ and $k^{(n)} = i^n\mathcal{F}^{-1}(\omega^n\mu)$.*

Proof. For $1 \leq m \leq n$, $|\omega| \leq 1$ for $|\omega| \leq 1$, $|\omega^m| \leq |\omega^n|$ for $|\omega| \geq 1$, so $|\omega^m| \leq 1 + |\omega^n|$. Thus, $\omega^m \in L^1(\mu)$ because $1, \omega^n \in L^1(\mu)$. Consequently, $\omega^m\mu$ is a (possibly signed or complex) finite Borel measure. The result then follows by induction.

Consider the induction hypothesis $k^{(m)} = i^m\mathcal{F}^{-1}(\omega^m\mu)$ and $k^{(m)} \in \mathcal{C}$. For $m = 1$, the hypothesis holds by Lemma 8 since $\omega \in L^1(\mu)$. Now assume that it holds for all $m < n$. Since $\omega^n \in L^1(\mu)$, $\omega \in L^1(\omega^{n-1}\mu)$. Then $k^{(n)}(t) = D_t k^{(n-1)}(t) = D_t i^{n-1}\mathcal{F}^{-1}(\omega^{n-1}\mu) = i^{n-1} \cdot i\mathcal{F}^{-1}(\omega \cdot \omega^{n-1}\mu) = i^n\mathcal{F}^{-1}(\omega^n\mu)(t)$ where the thrid equality holds from Lemma 8 with $\omega^{n-1}\mu$ as the measure. Additionally, by Lemma 8, $k^{(n-1)} \in C^1$, so $k^{(n)} \in \mathcal{C}$. Thus, the induction hypothesis holds so $k^{(n)} = i^n\mathcal{F}^{-1}(\omega^n\mu)$ and $k \in \mathcal{C}^n$, as desired. □

Using these results, we can now prove several facts about the relationship between the properties of a probability distribution and the behavior of its associated kernel. In particular, we present several conditions on the probability measure that guarantee a second-order kernel.

Proposition 10. *Let μ be a symmetric Borel probability measure such that $\mu \neq \delta$, where δ is the Dirac measure, and $\omega^2 \in L^1(\mu)$. Then $k = \mathcal{F}^{-1}\mu$ is a second-order, continuous, symmetric, positive-semidefinite kernel.*

Proof. From Proposition 9, $k = \mathcal{F}^{-1}\mu \in \mathcal{C}^2$ and $k''(t) = -1\mathcal{F}(\omega^2\mu)$. Since μ is a symmetric Borel probability measure, by (the converse of) Bochner's Theorem, k is a continuous, symmetric, positive-definite kernel with $k(0) = 1$.

Since $k \in \mathcal{C}^2$, Taylor's Theorem guarantees that $k(t) = k(0) + k'(0)t + \frac{1}{2}k''(0)t^2 + h_2(t)t^2$, where $\lim_{t\to 0} h_2(t) = 0$. Since $k' \in \mathcal{C}^1$ and k is symmetric, then $k'(0) = 0$. Thus, $k(t) = 1 + \frac{1}{2}k''(0)t^2 + h_2(t)t^2$. As such, we have that

$$k^{[2]}(0) = \lim_{t\to 0} t^{-2}(1 - k(t)) = -\frac{1}{2}k''(0) - \lim_{t\to 0} h_2(t) = -\frac{1}{2}k''(0) = \frac{1}{2}\int \omega^2 d\mu(\omega)$$

where the last equality holds because $k''(0) = -\int \omega^2 d\mu(\omega)$. Since $\mu \neq \delta$, $\int \omega^2 d\mu > 0$ and, since $\omega^2 \in L^1(\mu)$, $\int \omega^2 d\mu < \infty$. Therefore $0 < k^{[2]}(0) < \infty$ and k is a second-order, continuous symmetric, positive-semidefinite kernel. □

Corollary 11. *Let μ be a symmetric Borel probability measure with compact support such that $\mu \neq \delta$, then $k = \mathcal{F}^{-1}\mu$ is a second-order, continuous, positive semidefinite kernel.*

Proof. Let $M = \sup \omega$ over the support of μ. Then, $\int \omega^n d\mu \leq M^n < \infty$, so the conditions of the above lemma are satisfied and k is a second-order, continuous, symmetric, positive-semidefinite kernel. □

We can now prove the following, somewhat surprising, result.

Theorem 12. *If k is a nonconstant, q^{th}-order, continuous, symmetric, positive-semidefinite kernel with $k(0) = 1$ and $k^{[q]}(0) > 0$, then k can be decomposed as the (nontrivial) weighted sum of q^{th} and second-order, continuous, symmetric, positive-semidefinite kernels. Further, there are infinitely many such decompositions.*

Proof. If $q = 2$, we can decompose k as $k = \varepsilon k + (1-\varepsilon)k$ for any $\varepsilon \in (0,1)$. If $q < 2$, then we proceed as follows. Since k is a continuous, symmetric, positive-semidefinite kernel with $k(0) = 1$, by Bochner's Theorem, there exists a symmetric Borel probability measure μ, such that $k = \mathcal{F}^{-1}\mu$. Let a be sufficiently large that $\mu([-a,a]) > 0$ and define $\mu_{i,a}$ for $i = 1,2$ by $d\mu_{1,a} = \mu([-a,a]^c)^{-1} I_{[-a,a]^c} d\mu$ and $d\mu_{2,a} = \mu([-a,a]^c)^{-1} I_{[-a,a]} d\mu$. Note that both are well defined since $\mu([-a,a]) > 0$, by assumption, and $\mu([-;a]^c) > 0$, since otherwise μ would have compact support, but then, by Corollary 11, k would have characteristic exponent $q = 2$, contrary to our assumption that $q < 2$. Thus, the $\mu_{i,a}$s are probability measures. Define $k_{i,a} = \mathcal{F}^{-1}\mu_{i,a}$. By (the converse to) Bochner's Theorem, the $k_{i,a}$s are continuous, symmetric, positive-semidefinite kernels. Further, since $\mu_{2,a}$ has compact support, by Corollary 11, $k_{2,a}$ is a second-order kernel. $\mu = \mu([-a,a]^c)\mu_{1,a} + \mu([-a;a]^c)\mu_{2,a}$ so $k = \mu([-a,a]^c)k_{1,a} + \mu([-a,a])k_{2,a}$. Since, $q < 2$, $k_{2,a}^{[q]} = 0$ so $k_{1,a}^{[q]}(0) = \mu([-a,a]^c)^{-1} k^{[q]}(0) > 0$ because $k^{[q]}(0) > 0$. Thus, $k_{1,a}$ has characteristic exponent q. Therefore, as claimed, k can be decomposed into the (nontrivial) weighted sum of q^{th} and second-order, uniformly continuous, symmetric, positive-semidefinite kernels. Since a is arbitrary, there are infinitely many such decompositions. □

Note that, while the decomposition used in the proof for the $q < 2$ case can typically also be used to decompose second-order kernels, some care must be taken when the support of μ is compact to ensure that both $\mu([-a,a]) > 0$ and $\mu([-a,a]^c) > 0$. In particular, if $\mu = \delta_{-b} + \delta_b$ where δ_b is the Dirac measure with mass at b, such a decomposition is impossible and a decomposition of the form $\varepsilon k + (1-\varepsilon)k$, $\varepsilon \in (0,1)$, as used in the proof of the $q = 2$ case, is the only type available.

All of the above results use μ directly. However, we can still perform this decomposition when we only know k and not μ, by using a form of the Fourier Inversion Theorem typically encountered in abstract harmonic analysis. Specialized to our setting, it says that, if k is continuous, positive-semidefinite, and normalized so that $k(0) = 1$, and $k \in L^1$, then, $\hat{k} \in L^1$ as well, and the probability measure μ, of which k

is the inverse Fourier Transform, has density $\hat{k}$, so that, $d\mu = \hat{k}d\omega$. Thus, we can use the Fourier Transform to recover μ, and then use the above results to decompose k. We note that, since any k with compact support is integrable, we can always perform a decomposition of the kernel in that important special case.

Theorem 12 tells us that any nonconstant, q^{th}-order, positive-semidefinite kernel that satisfies very mild regularity conditions, can be (nontrivially) decomposed into another q^{th}-order, positive-semidefinite kernel and a second-order, positive-semidefinite kernel, so there is no natural notion of a minimal or irreducible q^{th}-order kernel for $q < 2$. This means that apparently artificial kernels, such as the one we exhibited to show that no optimal first-order kernel exists, are, in some sense, no more unnatural than the Bartlett kernel, which appears at first glance to be a 'pure' first-order kernel in a way that such a composite kernel does not. Also of note is the fact that it is the measure (μ) of sets at infinity that determines the order of the kernel; the measure of sets within any closed interval (e.g. in $[-a, a]$) has no effect. If μ has a density, K, we can restate this as saying that the order of the kernel is determined by K's tails; K's value on finite sets has no effect.

2.1.2 Kernel Decompositions

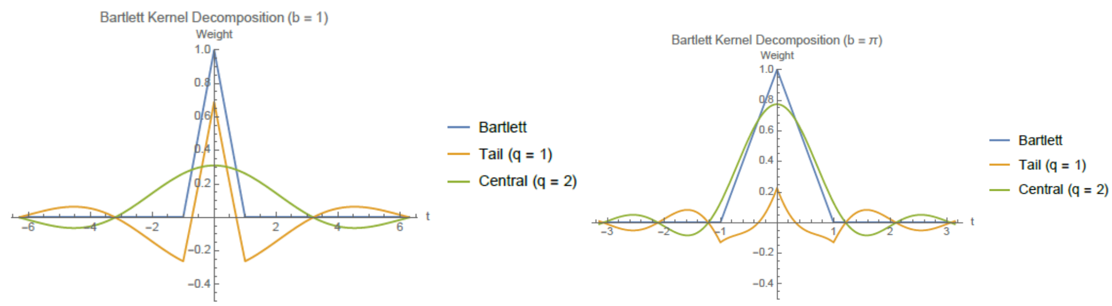
As a demonstration of this decomposition, we will separate the Bartlett kernel into a first-order piece, which comes from the tails of $K = \hat{k}$, and a second-order piece, which is derived from the center of K. We being by computing the Fourier transform of the Bartlett kernel $k(t) = (1 - |t|)I_{[-1,1]}(t)$, $K(\omega) = \pi^{-1}\omega^{-2}(1 - \cos(\omega))$. We then split K at some $a \geq 0$ into two functions $K_{1,a}(\omega) = \pi^{-1}\omega^{-2}(1 - \cos(\omega))I_{[-a,a]^c}(\omega)$ and $K_{2,b}(\omega) = \pi^{-1}\omega^{-2}(1 - \cos(\omega))I_{[-a,a]}(\omega)$ and apply the inverse Fourier transform to obtain $k_{1,a}$ and $k_{2,a}$. Note that these kernels are not properly normalized, so $k_i(0) \neq 1$, but they instead satisfy $k = k_{1,a} + k_{2,a}$, which makes it easier to see how each kernel contributes to the whole. As seen in Figure **??**, the Tail Kernel, $k_{1,a}$, is a scaled first-order kernel and contributes the sharp point at 0, while the Central Kernel, $k_{2,a}$, is a smooth, scaled second-order kernel. Also of note, as a increases, the relative contribution of $k_{1,a}$ decreases; additionally, $k_{2,a}$ becomes narrower and more pointed at its central peak, although, even for arbitrarily large a, $k_{2,a}$ will still remain second-order with a rounded, rather than sharp, peak.

2.2 Oddities of q

We present a kernel that does not possess a characteristic exponent, meaning that there is no maximal value of r for which $k^{[r]}(0)$ is finite, but $k^{[r]}(0)$ is not defined and finite for all r. Let

$$K_{q,p}(\omega) = c_{q,p}\left(I_{[-e,e]}(\omega) + e^{q+1}|\omega|^{-(q+1)}\log(|\omega|)^{-p} \cdot I_{[-e,e]^c}\right),$$

where $c_{q,p} = \left(2e + 2\int_e^\infty e^{q+1}\omega^{-(q+1)}\log(\omega)^{-p}d\omega\right)^{-1}$ so that $\int K_{q,p}(\omega)d\omega = 1$. For $q > 0$ or $q = 0, p > 0$ $c_{p,q}$ is finite, so $K_{q,p}$ is well defined. Since the $K_{q,p}$ are probability densities, their (inverse) Fourier Transforms $k_{q,p} = \mathcal{F}^{-1}K_{q,p}$, are positive-semidefinite kernels with $k_{q,p}(0) = 1$. We now consider $k_{q,-1}$ for $0 < q < 2$. It is possible to explicitly compute $k_{q,-1}$ using generalized hypergeometric and other special functions. Thus, we can directly calculate $k^{[r]}_{q,-1}(0)$. For $r < q, k^{[r]}_{q,-1}(0) = 0$, while for $r \geq q, k^{[r]}_{q,-1}(0) = \infty$. Thus, $k_{q,-1}$ does not have a well defined characteristic exponent.



(a) (b)

Figure 1: Decomposition of the Bartlett Kernel into (nonunique) first and second-order components. (A) is when $a = 1$ and (B) is when $a = \pi$.

The issue is the assumption that all kernels behave like $|t|^q$ to lowest order for some value of q near 0. In fact, their local behavior may be considerably more complex. For $0 < q < 2$, $\lim_{t\to 0} \frac{1-k_{q,-1}(t)}{-|t|^q \log|t|} = \left(q^2+q+1\right)^{-1} e^q \cos\left(\frac{q\pi}{2}\right)\Gamma(1-q)$, showing that, near 0, $k_{q,-1}(t)$ behaves like $-|t|^q \log|t|$, to lowest order. Another example is, $k_{2,0}$ for which $k_{2,0}^{[r]}(0) = 0$ for $r < 2$, $k_{2,0}^{[r]}(0) = \infty$ for $r \geq 2$, and $\lim_{t\to 0} \frac{1-k_{2,0}(t)}{-t^2 \log|t|} = \frac{1}{3}e^2$.

We can also ask whether there are kernels for which the characteristic exponent is defined, but $k^{[q]}(0) = 0$. We again consider the family of kernels $k_{q,p}$. Let $q \geq 0, p > 0$. Then,

$$\begin{aligned}(k_{q,p}(0) - k_{q,p}(t)) &= \int K_{q,p}(\omega)d\omega - \int e^{i\omega t}K_{q,p}(\omega)d\omega = \int \left(1-e^{i\omega t}\right)K_{q,p}(\omega)d\omega \\ &= \int (1-\cos(\omega t))K_{q,p}(\omega)d\omega \\ &= c_{q,p}\int_{-e}^{e}(1-\cos(\omega t))d\omega + 2c_{q,p}\int_e^\infty (1-\cos(\omega t))e^{q+1}\omega^{-(q+1)}\log(\omega)^{-p}d\omega\end{aligned}$$

For $\epsilon > 0$, define $d_{q,p,\epsilon} = \inf_{\omega\geq e} e^{-\epsilon}\omega^\epsilon \log(\omega)^{-p}$ so that $d_{q,p,\epsilon}e^\epsilon\omega^{-\epsilon} \leq \log(\omega)^{-p}$ for $\omega \geq e$. This is always possible since $D_\omega e^{-\epsilon}\omega^\epsilon \log(\omega)^{-p} = e^{-\epsilon}\omega^{\epsilon-1}\log(\omega)^{-p}\left(\epsilon - p\log(\omega)^{-1}\right)$ is positive for sufficiently large ω and has, at most, a single zero (occurring at $e^{\epsilon^{-1}p}$) on $[e,\infty)$ so that $d_{q,p,\epsilon} \geq e^{-\epsilon}e^p\left(\frac{p}{\epsilon}\right)^{-p} = e^{p-\epsilon}\left(\frac{p}{\epsilon}\right)^{-p} > 0$. Also note that, since $e^{-\epsilon}e^\epsilon\log(e)^{-p} = 1$, $d_{q,p,\epsilon} \leq 1$. Then,

$$\begin{aligned}d_{q,p,\epsilon}\int_e^\infty (1-\cos(\omega t))e^{q+\epsilon+1}\omega^{-(q+\epsilon+1)}d\omega &\leq \int_e^\infty (1-\cos(\omega t))e^{q+1}\omega^{-(q+1)}\log(\omega)^{-p}d\omega \\ &\leq \int_e^\infty (1-\cos(\omega t))e^{q+1}\omega^{-(q+1)}d\omega\end{aligned}$$

Thus,

$$\begin{aligned}
d_{q,p,\epsilon}c_{q+\epsilon,0}^{-1}\left(k_{q+\epsilon,0}(0)-k_{q+\epsilon,0}(t)\right) &= d_{q,p,\epsilon}c_{q+\epsilon,0}^{-1}\int(1-\cos(\omega t))K_{q+\epsilon,0}(\omega)d\omega \\
&= d_{q,p,\epsilon}\left(\int_{-e}^{e}(1-\cos(\omega t))d\omega+2\int_{e}^{\infty}(1-\cos(\omega t))e^{q+\epsilon+1}\omega^{-(q+\epsilon+1)}d\omega\right) \\
&\leq \left(\int_{-e}^{e}(1-\cos(\omega t))d\omega+2\int_{e}^{\infty}(1-\cos(\omega t))\omega^{-(q+1)}\log(\omega)^{-p}d\omega\right) \\
&= c_{q,p}^{-1}\left(k_{q,p}(0)-k_{q,p}(t)\right) \\
&\leq \left(\int_{-e}^{e}(1-\cos(\omega t))d\omega+2\int_{e}^{\infty}(1-\cos(\omega t))\omega^{-(q+1)}d\omega\right) \\
&= c_{q,0}^{-1}\left(k_{q,0}(0)-k_{q,0}(t)\right)
\end{aligned}$$

Thus,

$$d_{q,p,\epsilon}c_{q,p}c_{q+\epsilon,0}^{-1}\left(k_{q+\epsilon,0}(0)-k_{q+\epsilon,0}(t)\right)\leq\left(k_{q,p}(0)-k_{q,p}(t)\right)\leq c_{q,p}c_{q,0}^{-1}\left(k_{q,0}(0)-k_{q,0}(t)\right)$$

so, if $k_{q,p}^{[r]}(0)$ exists,

$$d_{q,p}c_{q,p}c_{q+\epsilon,0}^{-1}k_{q+\epsilon,0}^{[r]}(0)\leq k_{q,p}^{[r]}(0)\leq c_{q,p}c_{q,0}^{-1}k_{q,0}^{[r]}(0)$$

For $0<q<2, k_{q,0}$ can be written explicitly in terms of special functions, so it is straightforward to compute $k_{q,0}^{[r]}$ for any r. In particular, for $r<q, k_{q,0}^{[r]}(0)=0, k_{q,0}^{[q]}(0)=(q+1)^{-1}e^{q}\Gamma(1-q)\cos\left(\frac{q\pi}{2}\right)$, and, for $r>q, k_{q,0}^{[r]}(0)=\infty$, so the characteristic exponent is q, just as the notation suggests. Then, the above inequality shows that, for $r<q, k_{q,p}^{[r]}(0)=0, 0\leq k_{q,p}^{[q]}(0)<\infty$, and, for $r>q, k_{q,p}^{[r]}(0)=\infty$ (since the inequality holds for all $\epsilon>0$), so, if $k_{q,p}^{[q]}(0)$ exists, the characteristic exponent of $k_{q,p}$ is q, as well. Finally, we can use the Dominated Convergence Theorem to compute $k_{q,p}^{[q]}(0)$. Note that, for $q\geq 0, p>0, \omega^{-(q+1)}\log(|\omega|)^{-p}$ is integrable on $[e,\infty)$, for any $0\leq r\leq 2, |1-\cos(\omega t)|\leq c_r|\omega t|^r$, for some $c_r>0$, and $\lim_{t\to 0}|t|^{-r}(1-\cos(\omega t))=\lim_{t\to 0}|t|^{-r}\frac{1}{2}\omega^2t^2=\frac{1}{2}\omega^2\lim_{t\to 0}|t|^{2-r}$; if $r<2$ the limit is 0 , while if $r=2$ it is $\frac{1}{2}\omega^2$. Then,

$$\begin{aligned}
|t|^{-r}\int_{e}^{\infty}(1-\cos(\omega t))|\omega|^{-(q+1)}\log(|\omega|)^{-p}d\omega &\leq |t|^{-r}\int_{e}^{\infty}c_r|\omega t|^{r}|\omega|^{-(q+1)}\log(|\omega|)^{-p}d\omega \\
&= c_r\int_{e}^{\infty}|\omega|^{-(q-r+1)}\log(|\omega|)^{-p}d\omega
\end{aligned}$$

If $p>0$, the last integral is finite for all $r\leq q$, so we can apply the Dominated Convergence Theorem, so $\lim_{t\to 0}|t|^{-r}\int_{e}^{\infty}(1-\cos(\omega t))|\omega|^{-(q+1)}\log(|\omega|)^{-p}d\omega=0$. Another application of the Dominated Convergence Theorem shows that $\lim_{t\to 0}|t|^{-r}\int_{-e}^{e}(1-\cos(\omega t))d\omega=0$ for $r<2$. Then, $k_{q,p}^{[r]}(0)=\lim_{t\to 0}|t|^{-r}\int(1-\cos(\omega t))K_{p,q}(\omega)d\omega=0$. Thus, $k_{q,p}^{[q]}(0)=0$, even though the characteristic exponent is q.

This is in contrast to the case in which $q\geq 2$. For $r<2, k_{2,0}^{[r]}(0)=0$ and, for $r\geq 2$, $k_{2,0}^{[r]}(0)=\infty$. While, if $q>2$, for $r<2, k_{q,0}^{[r]}(0)=0, k_{q,0}^{[2]}(0)=(6(q-2))^{-1}e^{2}q$, and, for $r>2, k_{q,0}^{[r]}(0)=\infty$. Thus, if $q\geq 2$, for $r<2, k_{q,0}^{[r]}(0)=0$, and, for $r>2, k_{q,0}^{[r]}(0)=\infty$ So, if $k_{q,p}$ has a characteristic exponent, it must be equal to 2 . Since $|1-\cos(\omega t)| \leq \frac{1}{2}|\omega t|^2$ and $\lim_{t\to 0}|t|^{-2}(1-\cos(\omega t)) = \frac{1}{2}\omega^2$, applying the Dominated Convergence Theorem shows that $\lim_{t\to 0}|t|^{-2}\int_e^\infty(1-\cos(\omega t))|\omega|^{-(q+1)}\log(|\omega|)^{-p}d\omega = \frac{1}{2}\int_e^\infty \omega^2|\omega|^{-(q+1)}\log(|\omega|)^{-p}d\omega = \frac{1}{2}\int_e^\infty |\omega|^{1-q}\log(|\omega|)^{-p}d\omega > 0$ (which is finite, since $1-q \leq -1$, and can be expressed in terms of special functions). Combined with $\lim_{t\to 0}|t|^{-2}\int_{-e}^{e}(1-\cos(\omega t))d\omega = \frac{1}{3}e^3$, we see that $0 < k_{q,p}^{[2]} < \infty$, so that $k_{q,p}$ is a second-order kernel and $k_{q,p}^{[2]}(0) > 0$, which should not be a surprise given Theorem 4 and Lemma 2 in the main text.

3 Additional Results Related to Weighted Orthogonal Series Estimators

In this section, we first derive the optimal weights for a first-order orthogonal series estimator. After that, we specialize the results to Legendre polynomials and Haar wavelets.

3.1 Optimal Weights for a First-Order Orthogonal Series Estimator

We derive the optimal weights for a first-order orthogonal series. Recall that every Weighted Orthogonal Series (WOS) has a limiting implied mean kernel $k_w = \sum_{i=b_0}^B w_i k_i$, where $\{w_i\}_{i=b_0}^B$ is a sequence of weights and k_i is the limiting implied mean kernel of the i^{th} term in the series. Then, $k_w^{[q]} = \sum_{i=b_0}^B w_i k_i^{[q]}$, so, for first-order WOS estimators, $I_q[k_w] = \left(k_w^{[q]}(0)\right)^{\frac{1}{q}} \sum_i w_i^2 = \sum_{i=b_0}^B w_i k_i^{[1]}(0) \sum_{i=b_0}^B w_i^2$. We will now minimize this quantity subject to the adding up condition $\sum_{i=b_0}^B w_i = 1$ in order to obtain the optimal sequence of weights. We begin by constructing the Lagrangian $\mathcal{L}$, setting its first derivative equal to zero, and then using this to solve for the Lagrange multiplier λ. The Lagrangian is

$$\mathcal{L} = \sum_{i=b_0}^B w_i k_i^{[1]}(0) \cdot \sum_{i=b_0}^B w_i^2 + \lambda\left(1 - \sum_{i=b_0}^B w_i\right)$$

and the first-order conditions are

$$0 = \frac{\partial \mathcal{L}}{\partial w_i} = k_i^{[1]}(0) \cdot \sum_{j=b_0}^B w_j^2 + 2w_i \cdot \sum_{j=b_0}^B w_j k_j^{[1]}(0) - \lambda \;\forall\; i = b_0, ..., B$$

Rearranging the first-order conditions, we obtain that the Lagrange multiplier λ should satisfy

$$\lambda = k_i^{[1]}(0) \cdot \sum_{j=b_0}^B w_j^2 + 2w_i \cdot \sum_{j=b_0}^B w_j k_j^{[1]}(0) \;\forall\; i = b_0, ..., B.$$

Summing over i, we obtain

$$\lambda = (B - b_0 + 1)^{-1} \sum_{i=b_0}^{B} \left[k_i^{[1]}(0) \cdot \sum_{j=b_0}^{B} w_j^2 + 2w_i \cdot \sum_{j=b_0}^{B} w_j k_j^{[1]}(0) \right]$$
$$= (B - b_0 + 1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \cdot \sum_{j=b_0}^{B} w_j^2 + 2 \sum_{i=b_0}^{B} w_i \cdot \sum_{j=b_0}^{B} w_j k_j^{[1]}(0) \right)$$
$$= (B - b_0 + 1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \cdot \sum_{j=b_0}^{B} w_j^2 + 2 \cdot \sum_{j=b_0}^{B} w_j k_j^{[1]}(0) \right),$$

where the last equality holds because the adding up constraint requires that $\sum_{j=b_0}^{B} w_j = 1$. Substituting λ into the first order conditions we obtain

$$0 = k_i^{[1]}(0) \cdot \sum_{j=b_0}^{B} w_j^2 + 2w_i \cdot \sum_{j=b_0}^{B} w_j k_j^{[1]}(0) - \lambda$$
$$= k_i^{[1]}(0) \cdot \sum_{j=b_0}^{B} w_j^2 + 2w_i \cdot \sum_{j=b_0}^{B} w_j k_j^{[1]}(0) - (B - b_0 + 1)^{-1} \left(\sum_{j=b_0}^{B} k_i^{[1]}(0) \cdot \sum_{j=b_0}^{B} w_j^2 + 2 \cdot \sum_{j=b_0}^{B} w_j k_j^{[1]}(0) \right)$$
$$= \left(k_i^{[1]}(0) - (B - b_0 + 1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) \cdot \sum_{j=b_0}^{B} w_j^2 + \left(w_i - (B - b_0 + 1)^{-1} \right) \cdot 2 \sum_{j=b_0}^{B} w_j k_j^{[1]}(0)$$

We can then rewrite this as,

$$w_i = (B - b_0 + 1)^{-1} - \left(k_i^{[1]}(0) - (B - b_0 + 1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) \cdot \sum_{j=b_0}^{B} w_j^2 \cdot \left(2 \sum_{j=b_0}^{B} w_j k_j^{[1]}(0) \right)^{-1}$$

Let $A = \frac{1}{2} \left(\sum_{j=b_0}^{B} w_j k_j^{[1]}(0) \right)^{-1} \sum_{j=b_0}^{B} w_j^2$, then

$$w_i = (B - b_0 + 1)^{-1} - \left(k_i^{[1]}(0) - (B - b_0 + 1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A.$$

We can feed this expression back into the definition of A in order to obtain a quadratic equation, in terms of A, which we can then solve to obtain A as a function of the $k_i^{[1]}$s:

$$
\begin{aligned}
2A = & \left(\sum_{i=b_0}^{B} \left(\left((B-b_0+1)^{-1} - \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right) k_i^{[1]}(0) \right) \right)^{-1} \\
& \times \sum_{i=b_0}^{B} \left((B-b_0+1)^{-1} - \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right)^2 \\
= & \left(\sum_{i=b_0}^{B} \left(\left((B-b_0+1)^{-1} - \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right) k_i^{[1]}(0) \right) \right)^{-1} \\
& \times \sum_{i=b_0}^{B} \left((B-b_0+1)^{-2} - 2\,(B-b_0+1)^{-1} \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right. \\
& \left. + \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 A^2 \right) \\
= & \left((B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) - \left(\sum_{j=b_0}^{B} k_j^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 \right) A \right)^{-1} \\
& \times \left((B-b_0+1)^{-1} - 2\,(B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) - \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right. \\
& \left. + \sum_{i=b_0}^{B} \left(k_i^{[1]}(0)^2 - 2k_i^{[1]}(0)\,(B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) + (B-b_0+1)^{-2} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 \right) A^2 \right) \\
= & \left((B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) - \left(\sum_{j=b_0}^{B} k_j^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 \right) A \right)^{-1} \\
& \times \left((B-b_0+1)^{-1} - 2\,(B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) - \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right. \\
& \left. + \left(\sum_{j=b_0}^{B} k_j^{[1]}(0)^2 - 2\,(B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 + (B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 \right) A^2 \right) \\
= & \left((B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) - \left(\sum_{j=b_0}^{B} k_j^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 \right) \right)^{-1} \\
& \times \left((B-b_0+1)^{-1} + \left(\sum_{j=b_0}^{B} k_j^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 \right) A^2 \right)
\end{aligned}
$$

which can be rearranged to obtain

$$2A\left((B-b_0+1)^{-1}\sum_{j=b_0}^{B}k_j^{[1]}(0)-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-(B-b_0+1)^{-1}\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)A\right)$$
$$=(B-b_0+1)^{-1}+\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-(B-b_0+1)^{-1}\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)A^2.$$

Moving all terms to the right-hand side and combining like terms reveals that we must solve a quadratic expression in A:

$$0=3\left((B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)A^2-2\sum_{j=b_0}^{B}k_j^{[1]}(0)\cdot A+1$$

We can now use the quadratic formula to write A in terms of the weights and $k_i^{(1)}(0)$s:

$$A=\frac{2\sum_{j=b_0}^{B}k_j^{[1]}(0)\pm\sqrt{\left(2\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2-4\cdot3\left((B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)}}{2\cdot3\left((B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)}$$
$$=\frac{\sum_{j=b_0}^{B}k_j^{[1]}(0)\pm\sqrt{\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2-3\left((B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)}}{3\left((B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)}$$
$$=\frac{\sum_{j=b_0}^{B}k_j^{[1]}(0)\pm\sqrt{4\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2-3\,(B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2}}{3\left((B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)}$$

Using this, we can now write the weights as functions of the $k_i^{[1]}(0)$s,

$$w_i=(B-b_0+1)^{-1}-\left(k_i^{[1]}(0)-(B-b_0+1)^{-1}\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)A$$
$$=(B-b_0+1)^{-1}-\left(\left(k_i^{[1]}(0)-\left((B-b_0+1)^{-1}\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)\right)\right.$$
$$\left.\times\left(\frac{\sum_{j=b_0}^{B}k_j^{[1]}(0)\pm\sqrt{4\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2-3\,(B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2}}{3\left((B-b_0+1)\sum_{j=b_0}^{B}k_j^{[1]}(0)^2-\left(\sum_{j=b_0}^{B}k_j^{[1]}(0)\right)^2\right)}\right)\right)$$

Note that there is an unresolved $\pm$ in this expression. Thus, it is necessary to compute the weights in both cases, check that they are nonnegative, and then use the set of (valid) weights which gives the lowest value of $I_q\left[k_w\right]$. We can use these weights to compute both $k^{[1]}(0)$ and $\sum_{i=b_0}^{B}w_i^2$

We now compute the expressions for $k^{(1)}(0)$, $\sum_{i=b_0}^{B} w_i^2$, and $k^{(1)}(0)\sum_{i=b_0}^{B} w_i^2$ under the optimal weights.

$$
\begin{aligned}
k^{[1]}(0) &= \sum_{i=b_0}^{B} w_i k_i^{[1]}(0) \\
&= \sum_{i=b_0}^{B} \left((B-b_0+1)^{-1} - \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right) k_i^{[1]}(0) \\
&= (B-b_0+1)^{-1} \sum_{i=b_0}^{B} k_i^{[1]}(0) - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 \right) A \\
&= (B-b_0+1)^{-1} \sum_{i=b_0}^{B} k_i^{[1]}(0) - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 \right) \\
&\times \frac{\sum_{i=b_0}^{B} k_i^{[1]}(0) \pm \sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3\,(B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2}}{3\left((B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2\right)} \\
&= (B-b_0+1)^{-1} \sum_{i=b_0}^{B} k_i^{[1]}(0) \\
&- \frac{1}{3}(B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \pm \sqrt{4 \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 - 3\,(B-b_0+1) \sum_{i=b_0}^{B} k_i^{[1]}(0)^2 } \right) \\
&= \frac{2}{3}(B-b_0+1)^{-1} \sum_{i=b_0}^{B} k_i^{[1]}(0) \\
&\mp \frac{1}{3}(B-b_0+1)^{-1} \sqrt{4} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 - 3\,(B-b_0+1) \sum_{i=b_0}^{B} k_i^{[1]}(0)^2,
\end{aligned}
$$

$$
\begin{aligned}
\sum_{i=b_0}^{B} w_i^2 &= \sum_{i=b_0}^{B} \left((B-b_0+1)^{-1} - \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right)^2 \\
&= \sum_{i=b_0}^{B} \left((B-b_0+1)^{-2} - 2\,(B-b_0+1)^{-1} \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right) A \right. \\
&\quad \left. + \left(k_i^{[1]}(0) - (B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 A^2 \right) \\
&= (B-b_0+1)^{-1} - 2\,(B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) - \sum_{i=b_0}^{B} k_i^{[1]}(0) \right) A \\
&\quad + \sum_{i=b_0}^{B} \left(k_i^{[1]}(0)^2 - 2\,(B-b_0+1)^{-1} \sum_{j=b_0}^{B} k_j^{[1]}(0) \cdot k_i^{[1]}(0) + (B-b_0+1)^{-2} \left(\sum_{j=b_0}^{B} k_j^{[1]}(0) \right)^2 \right) A^2
\end{aligned}
$$

$$
\begin{aligned}
&= (B-b_0+1)^{-1} - 2\cdot 0\cdot A \\
&\quad + \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - 2\,(B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 + (B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 \right) A^2 \\
&= (B-b_0+1)^{-1} + \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 \right) A^2 \\
&= (B-b_0+1)^{-1} + \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - (B-b_0+1)^{-1} \left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 \right) \\
&\quad \times \left(\frac{\sum_{i=b_0}^{B} k_i^{[1]}(0) \pm \sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3\,(B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2}}{3\left((B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2\right)} \right)^2 \\
&= (B-b_0+1)^{-1} + \frac{\left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \pm \sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3\,(B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2}\right)^2}{9\,(B-b_0+1)\left((B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2\right)},
\end{aligned}
$$

and

$$
\begin{aligned}
k^{[1]}(0)\sum_{i=b_0}^{B} w_i^2 = &\left(\frac{2}{3}\,(B-b_0+1)^{-1} \sum_{i=b_0}^{B} k_i^{[1]}(0) \right. \\
&\left. \mp \frac{1}{3}\,(B-b_0+1)^{-1} \sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \right)^2 - 3\,(B-b_0+1) \sum_{i=b_0}^{B} k_i^{[1]}(0)^2 } \right)^2 \\
&\times (B-b_0+1)^{-1} + \frac{\left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \pm \sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3\,(B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2}\right)^2}{9\,(B-b_0+1)\left((B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2\right)}
\end{aligned}
$$

3.2 Application to Legendre Polynomials

The Legendre Polynomials, $\{P_i\}$, are a set of orthogonal polynomials, typically defined on $[-1, 1]$ that satisfy, $P_i(0) = (-1)^i$, $P_i(1) = 1$ and are normalized to have square norm $2(2i+1)^{-\frac{1}{2}}$, for the i^{th} Legendre polynomial. Remapping their domain to the interval $[0, 1]$ will result in dividing their norms by 2, but will not alter their orthogonality. Therefore, if we define $\phi_i(x) = (2i+1)^{\frac{1}{2}} P_i\left(\frac{1}{2}(x+1)\right)$, then the this will define an orthonormal series on $[0, 1]$.

Using Proposition 6 in the main text, $k_i^{[1]}(0) = \frac{1}{2}((2i+1) + (2i+1)) = 2i+1$. Then,

$$
\sum_{i=1}^{B} k_i^{[1]}(0) = \sum_{i=1}^{B}(2i+1) = \left(2\cdot\frac{1}{2}B(B+1)\right) + B = B(B+2)
$$

and

$$\sum_{i=1}^{B} k_i^{[1]}(0)^2 = \sum_{i=1}^{B}(2i+1)^2 = \sum_{i=1}^{B} 4i^2+4i+1 = 4\cdot\frac{1}{6}B(B+1)(2B+1)+4\cdot\frac{1}{2}B(B+1)+B$$
$$= \frac{2}{3}B(B+1)(2B+1)+2B(B+1)+B = \frac{1}{3}\left(4B^3+12B^2+11B\right)$$

Using the formula for optimal weights (with $b_0 = 1$), we obtain

$$\begin{aligned}
w_i = &(B-b_0+1)^{-1} - \left(k_i^{[1]}(0) - (B-b_0+1)^{-1}\sum_{j=b_0}^{B} k_j^{[1]}(0)\right)\\
&\times \frac{\sum_{i=b_0}^{B} k_j^{[1]}(0) \pm \sqrt{4\left(\sum_{j=b_0}^{B} k_j^{[1]}(0)\right)^2 - 3B\sum_{j=b_0}^{B} k_j^{[1]}(0)^2}}{3\left(B\sum_{j=b_0}^{B} k_j^{[1]}(0)^2 - \left(\sum_{j=b_0}^{B} k_j^{[1]}(0)\right)^2\right)}\\
=&B^{-1} - \left(k_i^{[1]}(0) - B^{-1}B(B+2)\right)\\
&\times \frac{B(B+2) \pm \sqrt{4(B(B+2))^2 - 3B\cdot\frac{1}{3}\left(4B^3+12B^2+11B\right)}}{3\left(B\cdot\frac{1}{3}\left(4B^3+12B^2+11B\right) - (B(B+2))^2\right)}\\
=&B^{-1} - \left(k_i^{[1]}(0) - B - 2\right)\frac{B(B+2)\pm\sqrt{4B^3+5B^2}}{3\cdot\frac{1}{3}\left(B^4-B^2\right)}\\
=&B^{-1} - \left(k_i^{[1]}(0) - B - 2\right)\frac{B+2\pm\sqrt{4B+5}}{B\left(B^2-B\right)},
\end{aligned}$$

$$\begin{aligned}
k^{[1]}(0) =&\frac{2}{3}(B-b_0+1)^{-1}\sum_{i=b_0}^{B} k_i^{[1]}(0)\\
&\mp \frac{1}{3}(B-b_0+1)^{-1}\sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3(B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2}\\
=&\frac{2}{3}B^{-1}\cdot B(B+2) \mp \frac{1}{3}B^{-1}\sqrt{4B^3+5B^2}\\
=&\frac{2}{3}(B+2) \mp \frac{1}{3}\sqrt{4B+5},
\end{aligned}$$

$$\begin{aligned}
\lim_{B\to\infty} B^{-1}k^{[1]}(0) &= \frac{2}{3}\lim_{B\to\infty} B^{-1}(B+2) \mp \frac{1}{3}\lim_{B\to\infty} B^{-1}\sqrt{4B+5}\\
&= \frac{2}{3},\\
\sum_{i=1}^{B} w_i^2 &= B^{-1} + \frac{\left(B(B+2)\pm\sqrt{4B^3+5B^2}\right)^2}{9B\cdot\frac{1}{3}\left(B^4-B^2\right)}\\
&= B^{-1} + \frac{(B+2\pm\sqrt{4B+5})^2}{3B\left(B^2-1\right)},
\end{aligned}$$

and

$$\lim_{B\to\infty} B\sum_{i=1}^{B} w_i^2 = 1 + \lim_{B\to\infty} \frac{(B+2\pm\sqrt{4B+5})^2}{3\,(B^2-1)} = 1+\frac{1}{3} = \frac{4}{3}$$

Then, the asymptotic value of $I_q[k]$ is $\lim_{B\to\infty} I_q[k] = \lim_{B\to\infty} k^{[1]}(0)\sum_{i=1}^{B} w_i^2 = \frac{2}{3}\cdot\frac{4}{3} = \frac{8}{9}$. If we instead used equal weights, so that $w_i = B^{-1}$ for $i \le B$, we would get $k^{[1]}(0) = \sum_{i=1}^{B} B^{-1}(2i+1) = B+2$, $\sum_{i=1}^{B} w_i^2 = \sum_{i=1}^{B} B^{-2} = B^{-1}$, so $I_q[k] = 1+2B^{-1}$ and $\lim_{B\to\infty} I_q[k] = 1 > \frac{8}{9}$.

3.3 Application to Haar Wavelets

We begin by computing the limiting implied kernel for each term in the system. Haar wavelets are defined by the wavelet function $\psi(x) = I_{[0,\frac{1}{2})}(x) - I_{[\frac{1}{2},1)}(x)$, so that the Haar basis functions are given by 1 and $\psi_{n,\ell}(x) = 2^{\frac{n}{2}}\psi(2^n x-\ell)$ with $n,\ell \in Z_+$, $0 \le \ell < 2^n$. Then, the limiting implied kernel corresponding to the wavelet $\psi_{n,\ell}$ is given by

$$\begin{aligned}
k_{n,\ell}(t) &= \int_{\max(0,t)}^{\min(1,1+t)} \psi_{n,\ell}(u)\psi_{n,\ell}(u-t)du = \int_{\max(0,t)}^{\min(1,1+t)} 2^{\frac{n}{2}}\psi\left(2^n u-\ell\right)2^{\frac{n}{2}}\psi\left(2^n(u-t)-\ell\right)du \\
&= 2^n \int_{\max(0,t)}^{\min(1,1+t)} \left[I_{[0,\frac{1}{2})}\left(2^n u-\ell\right) - I_{[\frac{1}{2},1)}\left(2^n u-\ell\right)\right]\left[I_{[0,\frac{1}{2})}\left(2^n(u-t)-\ell\right) - I_{[\frac{1}{2},1)}\left(2^n(u-t)-\ell\right)\right]du \\
&= 2^n \int_0^1 I_{[0,\frac{1}{2})}\left(2^n u-\ell\right) I_{[0,\frac{1}{2})}\left(2^n(u-t)-\ell\right) - I_{[0,\frac{1}{2})}\left(2^n u-\ell\right) I_{[\frac{1}{2},1)}\left(2^n(u-t)-\ell\right) \\
&- I_{[\frac{1}{2},1)}\left(2^n u-\ell\right) I_{[0,\frac{1}{2})}\left(2^n(u-t)-\ell\right) + I_{[\frac{1}{2},1)}\left(2^n u-\ell\right) + I_{[\frac{1}{2},1)}\left(2^n(u-t)-\ell\right)du \\
&= 2^n \int_0^1 I\left[2^{-n}\ell \le u, u-t < 2^{-n}\left(\ell+2^{-1}\right)\right] \\
&- I\left[\left(2^{-n}\ell \le u < 2^{-n}\left(\ell+2^{-1}\right)\right) \wedge \left(2^{-n}\left(\ell+2^{-1}\right) \le u-t < 2^{-n}(\ell+1)\right)\right] \\
&- I\left[\left(2^{-n}\left(\ell+2^{-1}\right) \le u < 2^{-n}(\ell+1)\right) \wedge \left(2^{-n}\ell \le u-t < 2^{-n}\left(\ell+2^{-1}\right)\right)\right] \\
&+ I\left[2^{-n}\left(\ell+2^{-1}\right) \le u, u-t < 2^{-n}(\ell+1)\right]du \\
&= 2^n\Big[2^{-(n+1)}\left(1-2^{n+1}|t|\right) I_{\left(-2^{-(n+1)},2^{-(n+1)}\right)}(t) - 2^{-(n+1)}\left(1-\left|1+2^{n+1}t\right|\right) I_{(-2^{-n},0]}(t) \\
&-2^{-(n+1)}\left(1-\left|1-2^{n+1}t\right|\right) I_{[0,2^{-n})}(t) + 2^{-(n+1)}\left(1-2^{n+1}|t|\right) I_{\left(-2^{-(n+1)},2^{-(n+1)}\right)}(t)\Big] \\
&= 2^{-1}\cdot 2\left(1-2^{n+1}|t|\right) I_{\left[0,2^{-(n+1)}\right)}(|t|) - 2^{-1}\left(1-\left|1-2^{n+1}\right||t|\right) I_{[0,2^{-n})}(|t|) \\
&= \left(1-2^{n+1}|t|\right) I_{\left[0,2^{-(n+1)}\right)}(|t|) - 2^{-1}\left(2^{n+1}|t|\right) I_{\left[0,2^{-(n+1)}\right)}(|t|) \\
&- 2^{-1}\left(2-2^{n+1}|t|\right) I_{\left[2^{-(n+1)},2^{-n}\right.}(|t|) \\
&= (1-3\cdot 2^n|t|) I_{[0,2^{-(n+1)})}(|t|) - (1-2^n|t|) I_{[2^{-(n+1)},2^{-n})}(|t|)
\end{aligned}$$

where the fourth equality is due to the fact that the indicators in the wavelet function automatically enforce that the integrand is 0 unless $\max(0,t) \le u \le \min(1,1+t)$, as can be seen explicitly when the indicator functions are rewritten in the fifth equality.

Note that this is independent of ℓ, since, within each level of the hierarchy, the basis functions are simply translations of each other, so we will simply write,

$$k_n(t) = (1 - 3 \cdot 2^n|t|)I_{[0,2^{-(n+1)})}(|t|) - (1 - 2^n|t|)I_{[2^{-(n+1)},2^{-n})}(|t|)$$

Then, it is clear that k_n is a first-order kernel with

$$k_n^{[1]}(0) = \lim_{t \to 0} |t|^{-1}(1 - (1 - 3 \cdot 2^n|t|) = 3 \cdot 2^n$$

$$\begin{aligned}
\int k_n^2(t)dt =& \int (1 - 3 \cdot 2^n|t|)^2 \, I_{\left[0,2^{-(n+1)}\right)}(|t|) + (1 - 2^n|t|)^2 \, I_{\left[2^{-(n+1)},2^{-n}\right)}(|t|)dt \\
=&2 \int_0^{2^{-(n+1)}} (1 - 3 \cdot 2^n t)^2 \, dt + 2 \int_{2^{-(n+1)}}^{2^{-n}} (1 - 2^n t)^2 \, dt \\
=&2 \left[\left[t - \frac{1}{2} 2 \cdot 3 \cdot 2^n t^2 + \frac{1}{3} 9 \cdot 2^{2n} t^3 \right]_0^{2^{-(n+1)}} + \left[t - \frac{1}{2} 2 \cdot 2^n t^2 + \frac{1}{3} \cdot 2^{2n} t^3 \right]_{2^{-(n+1)}}^{2^{-n}} \right] \\
=&2 \left[\left[t - 3 \cdot 2^n t^2 + 3 \cdot 2^{2n} t^3 \right]_0^{2^{-(n+1)}} + \left[t - 2^n t^2 + \frac{1}{3} \cdot 2^{2n} t^3 \right]_{2^{-(n+1)}}^{2^{-n}} \right] \\
=&2 \left[2^{-(n+1)} - 3 \cdot 2^n 2^{-2(n+1)} + 3 \cdot 2^{2n} 2^{-3(n+1)} \right] \\
&+ 2 \left[2^{-n} - 2^n 2^{-2n} + \frac{1}{3} \cdot 2^{2n} 2^{-3n} - \left(2^{-(n+1)} - 2^n 2^{-2(n+1)} + \frac{1}{3} \cdot 2^{2n} 2^{-3(n+1)} \right) \right] \\
=&2 \left[2^{-n-1} - 3 \cdot 2^{-n-2} + 3 \cdot 2^{-n-3} \right] \\
&+ 2 \left[2^{-n} - 2^{-n} + \frac{1}{3} 2^{-n} - \left(2^{-n-1} - 2^{-n-2} + \frac{1}{3} \cdot 2^{-n-3} \right) \right] \\
=&2^{-n} \left[1 - 3 \cdot 2^{-1} + 3 \cdot 2^{-2} \right] + 2^{-n} \left[\frac{2}{3} - \left(1 - 2^{-1} + \frac{1}{3} 2^{-2} \right) \right] \\
=&2^{-n} \frac{1}{4} + 2^{-n} \left[\frac{2}{3} - \frac{1}{2} - \frac{1}{12} \right] = 2^{-n} \left[\frac{1}{4} + \frac{1}{12} \right] \\
=&\frac{1}{3} \cdot 2^{-n}
\end{aligned}$$

It is interesting to note that, for an individual kernel, $I_q[k_n] = k_n^{[1]}(0) \cdot \int k_n^2(t)dt = 3 \cdot 2^n \frac{1}{3} 2^{-n} = 1$. Let

$n' > n$, then,

$$\int k_n(t)k_{n'}(t)dt = \int \Big((1-3\cdot 2^n|t|)\, I_{\left[0,2^{-(n+1)}\right)}(|t|) - (1-2^n|t|)\, I_{\left[2^{-(n+1)},2^{-n}\right)}(|t|)\Big)$$
$$\times \left(\left(1-3\cdot 2^{n'}|t|\right) I_{\left[0,2^{-(n'+1)}\right)}(|t|) - \left(1-2^{n'}|t|\right) I_{\left[2^{-(n'+1)},2^{-n'}\right)}(|t|)\right) dt$$
$$= \int \left[\left(1-3\cdot\left(2^n+2^{n'}\right)|t| + 9\cdot 2^{n+n'}t^2\right) I_{[0,2-(n'+1)}(|t|)\right.$$
$$\left. - \left(1-\left(3\cdot 2^n+2^{n'}\right)|t| + 3\cdot 2^{n+n'}t^2\right) I_{\left[2^{-(n'+1)},2^{-n'}\right)}(|t|)\right] dt$$
$$= 2\int_0^{2^{-(n'+1)}} 1-3\cdot\left(2^n+2^{n'}\right)t + 9\cdot 2^{n+n'}t^2 dt$$
$$- 2\int_{2^{-(n'+1)}}^{2^{-n'}} 1-\left(3\cdot 2^n+2^{n'}\right)t + 3\cdot 2^{n+n'}t^2 dt$$
$$= 2\left[t-\frac{3}{2}\cdot\left(2^n+2^{n'}\right)t^2 + 3\cdot 2^{n+n'}t^3\right]_0^{2^{-(n'+1)}}$$
$$- 2\left[t-\frac{1}{2}\left(3\cdot 2^n+2^{n'}\right)t^2 + 2^{n+n'}t^3\right]_{2^{-(n'+1)}}^{2^{-n'}}$$
$$= 2\left[2^{-(n'+1)} - \frac{3}{2}\cdot\left(2^n+2^{n'}\right)2^{-2(n'+1)} + 3\cdot 2^{n+n'}2^{-3(n'+1)}\right]$$
$$- 2\left[2^{-n'} - \frac{1}{2}\left(3\cdot 2^n+2^{n'}\right)2^{-2n'} + 2^{n+n'}2^{-3n'}\right]$$
$$+ 2\left[2^{-(n'+1)} - \frac{1}{2}\left(3\cdot 2^n+2^{n'}\right)2^{-2(n'+1)} + 2^{n+n'}2^{-3(n'+1)}\right]$$
$$= \left(2^{-n'} - 2^{-(n'-1)} + 2^{-n'}\right) - \left(3\cdot 2^{-2} - 3 + 3\cdot 2^{-2}\right)2^n 2^{-2n'}$$
$$- \left(3\cdot 2^{-2} - 1 + 2^{-2}\right)2^{n'}2^{-2n'} + \left(3\cdot 2^{-2} - 2 + 2^{-2}\right)2^{n'+n}2^{-3n'}$$
$$= 0 - 3\left(2^{-1}-1\right)2^{n-2n'} - 0\cdot 2^{-n'} + (1-2)2^{n-2n'}$$
$$= \frac{3}{2}\cdot 2^{n-2n'} - 2^{n-2n'} = \frac{1}{2}\cdot 2^{n-2n'} = 2^{n-2n'-1}$$
$$= 2^{(n-n')-n'-1}$$

Summing this over $n' > n$ gives,

$$\sum_{n'>n} 2^{(n-n')-n'-1} = 2^{-n-3}\sum_{n'>n} 2^{-2(n'-n-1)} = 2^{-n-3}\sum_{i=0}^{\infty} 2^{-2i} = 2^{-n-3}\left(1-2^{-2}\right)^{-1}$$
$$= 2^{-n-3}\left(\frac{3}{4}\right)^{-1} = \frac{4}{3}\cdot 2^{-n-3} = \frac{1}{3}\cdot 2^{-(n+1)} = \frac{1}{2}\cdot\frac{1}{3}\cdot 2^{-n}$$
$$= \frac{1}{2}\int k_n^2(t)dt$$

Since $\sum_{n,n'}\int k_n(t)k_{n'}(t)dt$ contains two coplies of $\int k_n(t)k_{n'}(t)dt$ for $n' \neq n$, the sum of all terms of the form $\int k_n(t)k_{n'}(t)dt$ with $n' > n$ is $2\cdot\frac{1}{2}\int k_n^2(t)dt = \int k_n^2(t)dt = \frac{1}{3}\cdot 2^{-n}$, so the sum of the cross terms is equal to the sum of the diagonal terms $\int k_n^2(t)dt$. We can now compute $k_w^{[1]}(0)$, $\int k_w^2(t)dt$, and $I_q[k_w]$.

$$\begin{aligned}
k_w^{[1]}(0) &= \lim_{t\to 0} |t|^{-1}\left(1-k_w(t)\right) = \lim_{t\to 0}|t|^{-1}\left(1-\sum_{n=1}^{N} w_n k_n(t)\right) = \lim_{t\to 0}|t|^{-1}\sum_{n=1}^{N} w_n\left(1-k_n(t)\right)\\
&= \sum_{n=1}^{N} w_n\left(\lim_{t\to 0}|t|^{-1}\left(1-k_n(t)\right)\right) = \sum_{n=1}^{N} w_n k_n^{(1)}(0) = \sum_{n=1}^{N} w_n\cdot 3\cdot 2^n\\
&= 3\sum_{n=1}^{N} w_n\cdot 2^n\\
\int k_w^2(t)dt &= \int\left(\sum_{n=1}^{N} w_n k_n(t)\right)^2 dt = \sum_{n,n'=1}^{N}\int w_n w_{n'} k_n(t)k_{n'}(t)dt\\
&= \sum_{n=1}^{N}\left(w_n^2\int k_n^2(t)dt + 2w_n\sum_{n'<n} w_{n'}\int k_n(t)k_{n'}(t)dt\right)\\
&= \sum_{n=1}^{N}\left(\frac{1}{3}\cdot 2^{-n}\cdot w_n^2 + 2w_n\sum_{n'<n} w_{n'}\cdot 2^{n'-2n-1}\right)\\
&= \sum_{n=1}^{N} 2^{-n}\left(\frac{1}{3}\cdot w_n^2 + w_n\cdot 2^{-n}\sum_{n'<n} w_{n'}\cdot 2^{n'}\right)\\
I_q[k_w] &= k_w^{[1]}(0)\int k_w^2(t)dt\\
&= 3\sum_{n} w_n\cdot 2^n\cdot\sum_{n} 2^{-n}\left(\frac{1}{3}\cdot w_n^2 + w_n\cdot 2^{2n+1}\sum_{n'>n} w_n'\cdot 2^{-(2n'+1)}\right)
\end{aligned}$$

Using equal weights for each basis function, and noting that for each level, n, there are 2^n basis functions, $w_n = \left(\sum_{n=1}^{N} 2^n\right)^{-1} 2^n = \left(2^{N+1}-1\right)^{-1} 2^n$. Then we get:

$$\begin{aligned}
k_w^{[1]}(0) &= \sum_{n=0}^{N} w_n k_n^{(1)}(0) = 3\sum_{n} w_n\cdot 2^n = 3\sum_{n=0}^{N}\left(2^{N+1}-1\right)^{-1} 2^n\cdot 2^n = 3\left(2^{N+1}-1\right)^{-1}\sum_{n=0}^{N} 2^{2n}\\
&= 3\left(2^{N+1}-1\right)^{-1}\sum_{n=0}^{N} 4^n = 3\left(2^{N+1}-1\right)^{-1}\frac{4^{N+1}-1}{4-1} = 3\left(2^{N+1}-1\right)^{-1}\frac{4^{N+1}-1}{3}\\
&= \frac{4^{N+1}-1}{2^{N+1}-1} = \frac{\left(2^{N+1}\right)^2-1}{2^{N+1}-1} = 2^{N+1}+1
\end{aligned}$$

and

$$\int k_w^2(x)dx = \int \left(\sum_{n=0}^{N} w_n k_n(x)\right)^2 dx = \sum_{n=0}^{N} 2^{-n}\left(\frac{1}{3}\cdot w_n^2 + w_n \cdot 2^{-n} \sum_{n'<n} w_{n'} \cdot 2^{-n'}\right)$$
$$= \sum_{n=0}^{N} 2^{-n}\left(\frac{1}{3}\cdot\left(\left(2^{N+1}-1\right)^{-1} 2^n\right)^2 + \left(2^{N+1}-1\right)^{-1} 2^n \cdot 2^{-n} \sum_{n'=0}^{n-1} \left(2^{N+1}-1\right)^{-1} 2^{n'} \cdot 2^{n'}\right)$$
$$= \left(2^{N+1}-1\right)^{-2} \sum_{n=0}^{N} 2^{-n}\left(\frac{1}{3}\cdot 2^{2n} + 2^n \cdot 2^{-n} \sum_{n'=0}^{n-1} 2^{2n'}\right)$$
$$= \left(2^{N+1}-1\right)^{-2} \sum_{n=0}^{N} \left(\frac{1}{3}\cdot 2^{n} + 2^{-n} \sum_{n'=0}^{n-1} 4^{n'}\right)$$
$$= \left(2^{N+1}-1\right)^{-2} \sum_{n=0}^{N} \left(\frac{1}{3}\cdot 2^{n} + 2^{-n}\frac{4^n-1}{4-1}\right) = \left(2^{N+1}-1\right)^{-2} \sum_{n=0}^{N} \left(\frac{1}{3}\cdot 2^{n} + \frac{1}{3}\cdot 2^{-n}\left(4^n-1\right)\right)$$
$$= \frac{1}{3}\cdot\left(2^{N+1}-1\right)^{-2} \sum_{n=0}^{N} \left(2^n + \left(2^n - 2^{-n}\right)\right) = \frac{1}{3}\cdot\left(2^{N+1}-1\right)^{-2} \sum_{n=0}^{N} \left(2^{n+1} - 2^{-n}\right)$$
$$= \frac{1}{3}\left(2^{N+1}-1\right)^{-2}\left(2\cdot\frac{2^{N+1}-1}{2-1} - \frac{1-2^{-(N+1)}}{1-2^{-1}}\right)$$
$$= \frac{2}{3}\left(2^{N+1}-1\right)^{-2}\left(2\cdot\left(2^{N+1}-1\right) - 2\cdot 2^{-(N+1)}\left(2^{N+1}-1\right)\right)$$
$$= \frac{2}{3}\cdot\frac{1-2^{-(N+1)}}{2^{N+1}-1}$$
$$= \frac{2}{3}\cdot 2^{-(N+1)}.$$

Putting these together, we obtain

$$I_q[k_w] = k_2^{[1]}(0)\int k_w^2(t)dt = \left(2^{N+1}+1\right)\cdot\frac{2}{3}\cdot 2^{-(N+1)} = \frac{2}{3}\left(1+2^{-(N+1)}\right).$$

Finally, we can also compute the sum of squared weights, for the implied kernel,

$$\sum_{n=0}^{N} w_n^2 = \sum_{n=0}^{N}\left(2^{N+1}-1\right)^{-2} 2^{2n} = \left(2^{N+1}-1\right)^{-2}\sum_{n=0}^{N} 4^n = \left(2^{N+1}-1\right)^{-2}\frac{4^{N+1}-1}{4-1}$$
$$= 3^{-1}\frac{2^{N+1}+1}{2^{N+1}-1}.$$

Note that this is actually not the value that is used when computing $I_q[k]$ for the orthogonal series estimator. For that, we count each basis function separately, which gives,

$$\sum_{n=0}^{N} w_{n,s}^2 = \sum_{n=0}^{N} 2^n\left(2^{N+1}-1\right)^{-2} = \left(2^{N+1}-1\right)^{-2}\sum_{n=0}^{N} 2^n = \left(2^{N+1}-1\right)^{-2}\frac{2^{N+1}-1}{2-1}$$
$$= \left(2^{N+1}-1\right)^{-1}$$

Then, for the equal weighted orthogonal series, we get,

$$I_q[k_s] = k_w^{[1]}(0) \sum_{n=0}^{N} w_{n,s}^2 = \frac{2^{N+1}+1}{2^{N+1}-1}$$

To summarize, for the orthogonal series generated by the Haar system using equal weights for each basis function,

$$\begin{aligned}
k_w^{[1]}(0) &= 2^{N+1} + 1 \\
\int k_w^2(t)dt &= \frac{2}{3} \cdot 2^{-(N+1)} \\
I_q\left[k_w\right] &= \frac{2}{3}\left(1 + 2^{-(N+1)}\right) \\
\sum_{n=0}^{N} w_n^2 &= 3^{-1} \frac{2^{N+1}+1}{2^{N+1}-1} \\
\sum_{n=0}^{N} w_{n,s}^2 &= \left(2^{N+1}-1\right)^{-1} \\
I_q\left[k_s\right] &= \frac{2^{N+1}+1}{2^{N+1}-1}
\end{aligned}$$

We can also use this to compute the optimal weights for the Haar orthogonal series. Since each level n has 2^n wavelets within it, we need to include this factor in all sums

$$\begin{aligned}
\sum_{n=0}^{N} 2^n k_n^{[1]}(0) &= \sum_{n=0}^{N} 2^n \cdot 3 \cdot 2^n = 3 \sum_{n=0}^{N} 4^n = 3 \cdot \frac{4^{N+1}-1}{4-1} = 3 \cdot \frac{1}{3}\left(4^{N+1}-1\right) \\
&= 4^{N+1} - 1 \\
\sum_{n=0}^{N} 2^n \left(k_n^{[1]}(0)\right)^2 &= \sum_{n=0}^{N} 2^n \left(3 \cdot 2^n\right)^2 = \sum_{n=0}^{N} 2^n \cdot 9 \cdot 4^n = 9 \sum_{n=0}^{N} 8^n = 9 \cdot \frac{8^{N+1}-1}{8-1} \\
&= \frac{9}{7}\left(8^{N+1}-1\right)
\end{aligned}$$

We can now compute the weights,

$$
\begin{aligned}
w_n &= (B-b_0+1)^{-1} - \left(k_i^{[1]}(0) - (B-b_0+1)^{-1}\sum_{i=b_0}^{B} k_i^{[1]}(0)\right) \\
&\times \frac{\sum_{i=b_0}^{B} k_i^{[1]}(0) \pm \sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3\,(B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2}}{3\left((B-b_0+1)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2\right)} \\
&= \left(2^{N+1}-1\right)^{-1} - \left(k_n^{[1]}(0) - \left(2^{N+1}-1\right)^{-1}\left(4^{N+1}-1\right)\right) \\
&\times \frac{\left(4^{N+1}-1\right) \pm \sqrt{4\,(4^{N+1}-1)^2 - 3\,(2^{N+1}-1)\,\frac{9}{7}\,(8^{N+1}-1)}}{3\left((2^{N+1}-1)\,\frac{9}{7}\,(8^{N+1}-1) - (4^{N+1}-1)^2\right)} \\
&= \left(2^{N+1}-1\right)^{-1} - \left(k_n^{(1)}(0) - \left(2^{N+1}-1\right)^{-1}\left(4^{N+1}-1\right)\right) \\
&\times \frac{\left(4^{N+1}-1\right) \pm \sqrt{4\,(4^{N+1}-1)^2 - \frac{27}{7}\,(2^{N+1}-1)\,(8^{N+1}-1)}}{3\left(\frac{9}{7}\,(2^{N+1}-1)\,(8^{N+1}-1) - (4^{N+1}-1)^2\right)}
\end{aligned}
$$

$$
\begin{aligned}
k^{[1]}(0) &= \sum_{i=b_0}^{B} w_i k_i^{[1]}(0) \\
&= \frac{2}{3}\left(B-b_0+1\right)^{-1}\sum_{i=b_0}^{B} k_i^{[1]}(0) \\
&\quad \mp \frac{1}{3}\left(B-b_0+1\right)^{-1}\sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3\left(B-b_0+1\right)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2} \\
&= \frac{2}{3}\left(2^{N+1}-1\right)^{-1}\left(4^{N+1}-1\right) \\
&\quad \mp \frac{1}{3}\left(2^{N+1}-1\right)^{-1}\sqrt{4\left(4^{N+1}-1\right)^2 - 3\left(2^{N+1}-1\right)\cdot\frac{9}{7}\left(8^{N+1}-1\right)} \\
&= \frac{2}{3}\frac{\left(2^{N+1}\right)^2-1}{2^{N+1}-1} \mp \frac{1}{3}\sqrt{4\left(\frac{4^{N+1}-1}{2^{N+1}-1}\right)^2 - \frac{27}{7}\cdot\frac{8^{N+1}-1}{2^{N+1}-1}} \\
&= \frac{2}{3}\left(2^{N+1}+1\right) \mp \frac{1}{3}\sqrt{4\left(2^{N+1}+1\right)^2 - \frac{27}{7}\cdot\frac{8^{N+1}-1}{2^{N+1}-1}} \\
\lim_{n\to\infty}\left(2^{N+1}-1\right)^{-1}k^{[1]}(0) &= \lim_{n\to\infty}\left[\frac{2}{3} \mp \frac{1}{3}\sqrt{4\left(\frac{2^{N+1}+1}{2^{N+1}-1}\right)^2 - \frac{27}{7}\cdot\frac{8^{N+1}-1}{\left(2^{N+1}-1\right)^3}}\right] \\
&= \frac{2}{3} \mp \frac{1}{3}\sqrt{4\cdot 1^2 - \frac{27}{7}\cdot 1} \\
&= \frac{2}{3} \mp \frac{1}{3}\sqrt{7^{-1}} \\
&= \frac{2 \mp 7^{-\frac{1}{2}}}{3}
\end{aligned}
$$

$$
\begin{aligned}
\sum_{i=b_0}^{B} w_i^2 &= \left(B-b_0+1\right)^{-1} + \frac{\left(\sum_{i=b_0}^{B} k_i^{[1]}(0) \pm \sqrt{4\left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2 - 3\left(B-b_0+1\right)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2}\right)^2}{9\left(B-b_0+1\right)\left(\left(B-b_0+1\right)\sum_{i=b_0}^{B} k_i^{[1]}(0)^2 - \left(\sum_{i=b_0}^{B} k_i^{[1]}(0)\right)^2\right)} \\
&= \left(2^{N+1}-1\right)^{-1} + \frac{\left(\left(4^{N+1}-1\right) \pm \sqrt{4\left(4^{N+1}-1\right)^2 - 3\left(2^{N+1}-1\right)\cdot\frac{9}{7}\left(8^{N+1}-1\right)}\right)^2}{9\left(2^{N+1}-1\right)\left(\left(2^{N+1}-1\right)\cdot\frac{9}{7}\left(8^{N+1}-1\right) - \left(4^{N+1}-1\right)^2\right)} \\
&= \left(2^{N+1}-1\right)^{-1} + \frac{\left(\left(4^{N+1}-1\right) \pm \sqrt{4\left(4^{N+1}-1\right)^2 - \frac{27}{7}\left(2^{N+1}-1\right)\left(8^{N+1}-1\right)}\right)^2}{9\left(2^{N+1}-1\right)\left(\frac{9}{7}\left(2^{N+1}-1\right)\left(8^{N+1}-1\right) - \left(4^{N+1}-1\right)^2\right)} \\
&= \left(2^{N+1}-1\right)^{-1} + \frac{\left(1 \pm \sqrt{4 - \frac{27}{7}\left(4^{N+1}-1\right)^{-2}\left(2^{N+1}-1\right)\left(8^{N+1}-1\right)}\right)^2}{9\left(2^{N+1}-1\right)\left(\frac{9}{7}\left(4^{N+1}-1\right)^{-2}\left(2^{N+1}-1\right)\left(8^{N+1}-1\right) - 1\right)}
\end{aligned}
$$

$$\lim_{N\to\infty}\left(2^{N+1}-1\right)\sum_{i=b_0}^{B} w_i^2 = \lim_{N\to\infty}\left[1+\frac{\left(1\pm\sqrt{4-\frac{27}{7}\left(4^{N+1}-1\right)^{-2}\left(2^{N+1}-1\right)\left(8^{N+1}-1\right)}\right)^2}{9\left(\frac{9}{7}\left(4^{N+1}-1\right)^{-2}\left(2^{N+1}-1\right)\left(8^{N+1}-1\right)-1\right)}\right]$$

$$= 1+\frac{\left(1\pm\sqrt{4-\frac{27}{7}\cdot 1}\right)^2}{9\left(\frac{9}{7}\cdot 1-1\right)} = 1+\frac{\left(1\pm 7^{-\frac{1}{2}}\right)^2}{9\left(2\cdot 7^{-1}\right)} = 1+\frac{\left(7^{\frac{1}{2}}\pm 1\right)^2}{18}$$

Then,

$$\lim_{N\to\infty} I_q\left[k_{w,s}\right] = \left(\frac{2\mp 7^{\frac{1}{2}}}{3}\right)\left(1+\frac{\left(7^{\frac{1}{2}}\pm 1\right)^2}{18}\right)\approx .94, .91$$

References

M. B. Priestley. *Spectral analysis and time series.* London: Academic Press, 1981.