Submitted to the Annals of Statistics

SUPPLEMENT TO "APPROXIMATION AND ESTIMATION OF S-CONCAVE DENSITIES VIA RÉNYI DIVERGENCES"

By Qiyang Han and Jon A. Wellner*

University of Washington

In this supplement, we present omitted proofs of Lemmas 2.1,2.3, Corollaries 2.7, 2.8, 2.10, Theorem 2.12, Corollary 2.13, Theorems 2.14, 2.16 in Appendix A, Lemmas 3.1-3.3, Theorem 3.4, Lemma 3.5, Theorems 3.7, 3.8 in Appendix B, and Theorem 4.4, Lemma 4.6 in Appendix C. Appendix D is devoted to the proof of Theorem 6.1 due to its length. Some supporting lemmas and auxiliary results from convex analysis are collected Appendix E.

APPENDIX A: SUPPLEMENTARY PROOFS FOR SECTION 2

Proof of Lemma 2.1. Let $Q \in \mathcal{Q}_1$. Then by letting $g(x) := \|x\| + 1$, we have

$$L(Q) \leq L(g, Q) = \int (1 + \|x\|) \ \mathrm{d}Q + \frac{1}{|\beta|} \int \frac{\mathrm{d}x}{(1 + \|x\|)^{-\beta}} < \infty,$$

by noting $Q \in \mathcal{Q}_1$, and $-\beta = -1 - 1/s > d$. Now assume $L(Q) < \infty$. If $Q \notin \mathcal{Q}_1$, i.e. $\int \|x\| \ \mathrm{d}Q = \infty$, then since for each $g \in \mathcal{G}$, we can find some $a, b > 0$ such that $g(x) \geq a\|x\| - b$, we have

$$L(g, Q) = \int g \ \mathrm{d}Q + \frac{1}{|\beta|} \int g^{\beta} \ \mathrm{d}x \geq \int (a\|x\| - b) \ \mathrm{d}Q = \infty,$$

a contradiction. This implies $Q \in \mathcal{Q}_1$. □

Proof of Lemma 2.3. Let g, h be two minimizers for $\mathcal{P}_Q$. Since $\psi_s(x) = \frac{1}{|\beta|} x^{\beta}$ is strictly convex on $[0, \infty)$, $L(t \cdot g + (1-t) \cdot h, Q)$ is strictly convex in $t \in [0, 1]$ unless $g = h$ a.e. with respect to the canonical Lebesgue measure. We claim if two closed functions g, h agree a.e. with respect to the canonical

*Supported in part by NSF Grant DMS-1104832 and NI-AID grant R01 AI029168

MSC 2010 subject classifications: Primary 62G07, 62H12; secondary 62G05, 62G20

Keywords and phrases: s-concavity, consistency, projection, asymptotic distribution, mode estimation, nonparametric estimation, shape constraints

Lebesgue measure, then it must agree everywhere, thus closing the argument. It is easy to see $\mathrm{int}(\mathrm{dom} g) = \mathrm{int}(\mathrm{dom} h)$. Since $\mathrm{int}(\mathrm{dom}(g)) \neq \emptyset$, we have $\mathrm{ri}(\mathrm{dom} g) = \mathrm{int}(\mathrm{dom} g) = \mathrm{int}(\mathrm{dom} h) = \mathrm{ri}(\mathrm{dom} h)$. Also note that a convex function is continuous in the interior of its domain, and hence almost everywhere equality implies everywhere equality within the interior of the domain, i.e. $g\big|_{\mathrm{int}(\mathrm{dom} g)} = h\big|_{\mathrm{int}(\mathrm{dom} h)}$. Now by Corollary 7.3.4 in Rockafellar (1997), and the closedness of g, h, we find that $g = \mathrm{cl} g = \mathrm{cl} h = h$. □

PROOF OF COROLLARY 2.7. It is known by Varadarajan's theorem (cf. Dudley (2002) Theorem 11.4.1), $\mathbb{Q}_n$ converges weakly to Q with probability 1. Further by the strong law of large numbers (SLLN), we know that $\int \|x\| \, \mathrm{d}\mathbb{Q}_n \to_{a.s.} \int \|x\| \mathrm{d}Q$. This verifies all conditions required in Theorem 2.5. □

PROOF OF COROLLARY 2.8. The conclusion follows from Corollary 2.7 if $-1/(d+1) < s < 0$, so suppose $-1/d < s \leq -1/(d+1)$. Since $f \in \mathcal{P}_{s'}$, we may write $f = g^{1/s'}$ where g is convex. If f is unbounded, then $g(x_0) = 0$ for some $x_0 \in \mathbb{R}$. By Lemma E.9 with $r' = -1/s'$, it follows that $\int f = \infty$, contradicting the fact that f is a density. Thus f must necessarily be bounded. To see that f has a finite mean, note that by Lemma 3.5 $f(x) = (b + a\|x\|)^{1/s'}$ where $a, b > 0$ and $r' \equiv -1/s' > d+1$. Thus $\int_{\mathbb{R}^d} \|x\| f(x) dx \leq \int_{\mathbb{R}^d} \|x\| (b + a\|x\|)^{-r'} dx < \infty$. Now note that (2.8) holds by the existence of the Rényi divergence estimator for the empirical measure (cf. Theorem 4.1 in Koenker and Mizera (2010)) and the same argument in the proof of Theorem 2.5. Also note that by the proof of Theorem 3.7, (2.8) would be enough to ensure (2.10). Since f is continuous on the interior of the domain, we see that (2.10) implies weak convergence: let $\hat{Q}_n$ be the measures corresponding to $\hat{f}_n$. Then $\hat{Q}_n \to Q$ weakly as $n \to \infty$. Now the rest follows immediately from Theorems 3.6 and 3.8. □

PROOF OF COROLLARY 2.10. Let $g \equiv g(\cdot|Q)$. Then by Theorem 2.2 and Lemma E.4, we find that there exists some $a, b > 0$ such that $g(x) \geq a\|x\| + b$. Now take $v \in \partial h(0)$, i.e. $h(x) \geq h(0) + v^T x$ holds for all $x \in \mathbb{R}^d$. Hence for $t > 0$, we have

$$g(x) + th(x) \geq a\|x\| + b + t(h(0) + v^T x) \geq (a - t\|v\|)\|x\| + (b + th(0)),$$

which implies that $g + th \in \mathcal{G}$ for $t > 0$ small enough. Now the conclusion follows from the Theorem 2.9. □

PROOF OF THEOREM 2.12. We first note that if F is a distribution function for a probability measure supported on $[X_{(1)}, X_{(n)}]$, and $h : [X_{(1)}, X_{(n)}] \to$

$\mathbb{R}$ an absolutely continuous function, then integration by parts (Fubini's theorem) yields

$$\int h \, \mathrm{d}F = h(X_{(n)}) - \int_{X_{(1)}}^{X_{(n)}} h'(x)F(x) \, \mathrm{d}x. \tag{A.1}$$

First we assume $g_n = \hat{g}_n$. For fixed $t \in [X_{(1)}, X_{(n)}]$, let h_1 be a convex function whose derivative is given by $h_1'(x) = -\mathbf{1}(x \le t)$. Now by Theorem 2.9 we find that $\int h_1 \, \mathrm{d}F_n = \int h_1 \, \mathrm{d}\hat{F}_n \le \int h_1 \, \mathrm{d}\mathbb{F}_n$. Plugging in (A.1) we find that $\int_{X_{(1)}}^t F_n(x) \, \mathrm{d}x \le \int_{X_{(1)}}^t \mathbb{F}_n(x) \, \mathrm{d}x$. For $t \in \mathcal{S}_n(g_n)$, let h_2 be the function with derivative $h_2'(x) = \mathbf{1}(x \le t)$. It is easy to see $g_n + th_2$ is convex for $t > 0$ small enough, whence Theorem 2.9 is valid, thus giving the reverse direction of inequality. This shows the necessity.

For sufficiency, assume g_n satisfies (2.13). In view of the proof of Theorem 2.9, we only have to show (2.12) holds for all functions $h : \mathbb{R} \to \overline{\mathbb{R}}$ which are linear on $[X_{(i)}, X_{(i+1)}](i = 1, \ldots, n-1)$ and $g_n + th$ convex for $t > 0$ small enough. Since g_n is a linear function between two consecutive knots, h must be convex between consecutive knots. This implies that the derivative of such an h can be written as $h'(x) = \sum_{j=2}^n \beta_j \mathbf{1}(x \le X_{(j)})$, with $\beta_2, \ldots, \beta_n$ satisfying $\beta_j \le 0$ if $X_{(j)} \notin \mathcal{S}_n(g_n)$. Now again by (A.1) we have

$$\begin{aligned}\int h \, \mathrm{d}\hat{F}_n &= h(X_n) - \sum_{j=2}^n \beta_j \int_{X_{(1)}}^{X_{(j)}} \hat{F}_n(x) \, \mathrm{d}x \\ &\le h(X_n) - \sum_{j=2}^n \beta_j \int_{X_{(1)}}^{X_{(j)}} \mathbb{F}_n(x) \, \mathrm{d}x = \int h \, \mathrm{d}\mathbb{F}_n,\end{aligned}$$

as desired. □

PROOF OF COROLLARY 2.13. This follows directly from the Theorem 2.12 by noting for $x_1 < x_0 < x_2$ we have

$$\frac{1}{x_2 - x_0} \int_{x_0}^{x_2} \hat{F}_n(x) \, \mathrm{d}x \le \frac{1}{x_2 - x_0} \int_{x_0}^{x_2} \mathbb{F}_n(x) \, \mathrm{d}x,$$

and

$$\frac{1}{x_0 - x_1} \int_{x_1}^{x_0} \hat{F}_n(x) \, \mathrm{d}x \ge \frac{1}{x_0 - x_1} \int_{x_1}^{x_0} \mathbb{F}_n(x) \, \mathrm{d}x.$$

Now let $x_1 \nearrow x_0$ and $x_2 \searrow x_0$ we find that $\hat{F}_n(x_0) \le \mathbb{F}_n(x_0)$ by right continuity and $\hat{F}_n(x_0) \ge \mathbb{F}_n(x_0-) = \mathbb{F}_n(x_0) - \frac{1}{n}$. □

Proof of Theorem 2.14. The proof closely follows the proof of Theorem 2.7 of Dümbgen, Samworth and Schuhmacher (2011). For the reader's convenience we give a full proof here. Let P denote the probability distribution corresponding to F. We first show necessity by assuming $g = g(\cdot|Q)$. By Corollary 2.10 applied to $h(x) = \pm x$, we find by Fubini's theorem that

$$0 = \int_{\mathbb{R}} x \ \mathrm{d}(Q-P)(x) = \int_{\mathbb{R}} (F-G)(t)\mathrm{d}t$$

which proves (1). Now we turn to (2). Since the map $s \mapsto (s-x)_+$ is convex, again by Corollary 2.10, we find

$$0 \le \int_{\mathbb{R}} (s-x)_+ \mathrm{d}(Q-P)(s) = -\int_{-\infty}^{x} (F-G)(t) \ \mathrm{d}t,$$

where in the last equality we used the proved fact that $\int_{\mathbb{R}}(F-G)\mathrm{d}\lambda = 0$. Now we assume $x \in \tilde{\mathcal{S}}(g)$, and discuss two different cases to conclude. If $x \in \partial(\mathrm{dom}(g))$, then let $h(s) = -(s-x)_+$, it is easy to see $g + th \in \mathcal{G}$ for $t > 0$ small enough. Then by Theorem 2.9, we have

$$0 \le \int h(s)\mathrm{d}(Q-P)(s) = \int_{-\infty}^{x} (F-G)(t) \ \mathrm{d}t.$$

If $x \in \mathrm{int}(\mathrm{dom}(g))$, then $g'(x-\delta) < g'(x+\delta)$ for small $\delta > 0$ by definition, and hence we define

$$H_\delta'(u) = -\frac{g'(u) - g'(x-\delta)}{g'(x+\delta) - g'(x-\delta)} \mathbf{1}_{\{u \in [x-\delta, x+\delta]\}} - \mathbf{1}_{\{u > x+\delta\}},$$

whose integral $H_\delta(s) := \int_{-\infty}^{s} H_\delta'(u) \ \mathrm{d}u$ serves as an approximation of $-(s-x)_+$ as $\delta \searrow 0$. Note that

$$\big(g+tH_\delta\big)(s) = g(s) - \frac{t}{g'(x+\delta) - g'(x-\delta)} \int_{s\wedge(x-\delta)}^{s\wedge(x+\delta)} \big(g'(u) - g'(x-\delta)\big) \ \mathrm{d}u - t\big(s-(x+\delta)\big)_+,$$

implying $g + tH_\delta \in \mathcal{G}$ for $t > 0$ small enough (which may depend on δ). Then by Theorem 2.9,

$$0 \le \int H_\delta(s)\mathrm{d}(Q-P)(s) \to -\int (s-x)_+ \mathrm{d}(Q-P)(s) = \int_{-\infty}^{x} (F-G)(t) \ \mathrm{d}t,$$

as $\delta \searrow 0$, where the convergence follows easily from dominated convergence theorem. This proves (2). Now we show sufficiency by assuming (1)-(2).

Consider a Lipschitz continuous function $\Delta(\cdot)$ with Lipschitz constant L. Then

$$\begin{aligned}\int \Delta \mathrm{d}(Q-P) &= \int \Delta'(F-G)\ \mathrm{d}\lambda = -\int (L-\Delta')(F-G)\ \mathrm{d}\lambda \\ &= -\int_{\mathbb{R}} \left(\int_{-L}^{L} \mathbf{1}_{\{s>\Delta'(t)\}} \mathrm{d}s \right) (F-G)(t)\mathrm{d}t \\ &= -\int_{-L}^{L} \int_{A(\Delta',s)} (F-G)(t)\ \mathrm{d}t\mathrm{d}s,\end{aligned}$$

where the second line follows from (1), and $A(\Delta', s) := \{t \in \mathbb{R} : \Delta'(t) < s\}$. Now replace the generic Lipschitz function Δ with $g^{(\epsilon)}$ as defined in Lemma E.2 with Lipschitz constant $L = 1/\epsilon$. Note in this case $A\big((g^{(\epsilon)})', s\big) = (-\infty, a(g, \epsilon))$, where $a(g, s) = \min\{t \in \mathbb{R} : g'(t+) \geq s\}$ and hence $a(g, s) \in \tilde{\mathcal{S}}(g)$. This implies that $\int_{A\left((g^{(\epsilon)})',s\right)} (F-G)(s)\mathrm{d}s = 0$ for all $s \in (-L, L)$ by (2), yielding that $\int g^{(\epsilon)}\ \mathrm{d}(Q-P) = 0$. Similarly we have $\int g_0^{(\epsilon)}\ \mathrm{d}(Q-P) \geq 0$ where $g_0 = g(\cdot|Q)$. Now let $\epsilon \searrow 0$, by monotone convergence theorem we find that $\int g\ \mathrm{d}Q = \int g\ \mathrm{d}P$ and that $\int g_0\ \mathrm{d}Q \geq \int g_0\ \mathrm{d}P$. This yields

$$L(g_0, Q) \geq L(g_0, P) \geq L(g, P) = L(g, Q),$$

where the second inequality follows from the Fisher consistency of functional $L(\cdot, \cdot)$ and the fact that P is the distribution corresponding to g. □

Before we prove Theorem 2.16, we will need an elementary lemma.

LEMMA A.1. *Fix a sequence* $0 < \alpha_n < 1$ *with* $\alpha_n \nearrow 1$. *Let* f_{α_n} *be an* $(\alpha_n - 1)$*-concave density on* $\mathbb{R}$. *Let* $g_{\alpha_n} := f_{\alpha_n}^{\alpha_n - 1}$ *be the underlying convex function. Suppose* $\{g_{\alpha_n}\}$*'s are linear on* $[a, b]$ *with* $\lim_{n\to\infty} f_{\alpha_n}(a) = \gamma_a \in [0, \infty]$ *and* $\lim_{n\to\infty} f_{\alpha_n}(b) = \gamma_b \in [0, \infty]$. *Then for all* $x \in [a, b]$,

$$f_{\alpha_n}(x) \to \exp\left(\frac{\log \gamma_b - \log \gamma_a}{b-a}(x-a) + \log \gamma_a \right) \tag{A.2}$$

where $\exp(-\infty) := 0$ *and* $\exp(\infty) := \infty$.

PROOF OF LEMMA A.1. First assume $\gamma_b \neq \gamma_a$ and $\gamma_a, \gamma_b \in (0, \infty)$. For notational convenience we drop explicit dependence on n and the limit is taken as $\alpha \nearrow 1$. Let $\gamma_{a,\alpha} = f_\alpha(a) = g_\alpha(a)^{1/(\alpha-1)}$ and $\gamma_{b,\alpha} = f_\alpha(b) =$

$g_\alpha(b)^{1/(\alpha-1)}$. For any $x \in [a,b]$,

(A.3)

$$\begin{aligned}\lim_{\alpha\to 1}\log f_\alpha(x) &= \lim_{\alpha\to 1}\frac{1}{\alpha-1}\log\left(\frac{\gamma_{b,\alpha}^{\alpha-1}-\gamma_{a,\alpha}^{\alpha-1}}{b-a}(x-a)+\gamma_{a,\alpha}^{\alpha-1}\right)\\ &= \lim_{\alpha\to 1}\frac{1}{\alpha-1}\log\left(\frac{\gamma_{b}^{\alpha-1}-\gamma_{a}^{\alpha-1}}{b-a}(x-a)\cdot\frac{\gamma_{b,\alpha}^{\alpha-1}-\gamma_{a,\alpha}^{\alpha-1}}{\gamma_{b}^{\alpha-1}-\gamma_{a}^{\alpha-1}}+\gamma_{a,\alpha}^{\alpha-1}\right)\\ &\equiv \log\gamma_a+\lim_{\alpha\to 1}\frac{1}{\alpha-1}\log\left((\gamma_b^{\alpha-1}-\gamma_a^{\alpha-1})\frac{(x-a)}{(b-a)}\cdot\frac{1}{\gamma_{a,\alpha}^{\alpha-1}}\cdot r_\alpha+1\right).\end{aligned}$$

Since $\gamma_{a,\alpha}^{\alpha-1}\to 1$, we claim that it suffices to show that

$$r_\alpha \equiv \frac{\gamma_{b,\alpha}^{\alpha-1}-\gamma_{a,\alpha}^{\alpha-1}}{\gamma_{b}^{\alpha-1}-\gamma_{a}^{\alpha-1}}\to 1 \quad \text{as} \quad \alpha\to 1. \tag{A.4}$$

To see this, assume without loss of generality that $\gamma_a > \gamma_b$ and hence $\gamma_b^{\alpha-1} - \gamma_a^{\alpha-1} > 0$. Suppose that (A.4) holds and let $\epsilon > 0$. Then the second term on right hand side of (A.3) can be bounded from above by

$$\begin{aligned}&\lim_{\alpha\nearrow 1}\frac{1}{\alpha-1}\log\left((\gamma_b^{\alpha-1}-\gamma_a^{\alpha-1})\frac{(x-a)}{(b-a)}(1-\epsilon)+1\right)\\ &=\lim_{\alpha\nearrow 1}\left(\log\gamma_b\cdot\gamma_b^{\alpha-1}-\log\gamma_a\cdot\gamma_a^{\alpha-1}\right)\frac{(x-a)}{(b-a)}(1-\epsilon)\\ &=(\log\gamma_b-\log\gamma_a)\frac{(x-a)}{(b-a)}(1-\epsilon)\end{aligned}$$

where the second line follows from L'Hospital's rule. Similarly we can derive a lower bound:

$$(\log\gamma_b-\log\gamma_a)\frac{(x-a)}{(b-a)}(1+\epsilon).$$

Thus it remains to show that (A.4) holds. But we can rewrite r_α as

$$\begin{aligned} r_\alpha &= \frac{c_\alpha^{\alpha-1}-1}{c^{\alpha-1}-1}\\ &= \frac{c^{\alpha-1}(c_\alpha/c)^{\alpha-1}-(c_\alpha/c)^{\alpha-1}+(c_\alpha/c)^{\alpha-1}-1}{c^{\alpha-1}-1}\\ &= (c_\alpha/c)^{\alpha-1}+\frac{(c_\alpha/c)^{\alpha-1}-1}{c^{\alpha-1}-1}\\ &\to 1+0 \quad \text{as} \quad \alpha\to 1\end{aligned}$$

since $\log((c_\alpha/c)^{\alpha-1}) = (\alpha-1)\log(c_\alpha/c) \to 0\cdot\log 1 = 0$, and where the second limit follows from an upper and lower bound argument using $c_\alpha/c \to 1$. where $c_\alpha := \gamma_{b,\alpha}/\gamma_{a,\alpha}$ and $c = \gamma_b/\gamma_a \neq 1$.

This shows that (A.4) holds, thereby proving the case for $\gamma_a \neq \gamma_b \in (0,\infty)$. For the case $\gamma_b = \gamma_a \in (0,\infty)$, similarly we have

$$\lim_{\alpha\to 1}\log f_\alpha(x) = \log\gamma_a + \lim_{\alpha\to 1}\frac{1}{\alpha-1}\log\left(\frac{c_\alpha^{\alpha-1}-1}{b-a}(x-a)+1\right).$$

The second term is 0 by an argument much as above by observing $c_\alpha = \gamma_{b,\alpha}/\gamma_{a,\alpha} \to \gamma_b/\gamma_a = 1$. Finally, if $\gamma_a \wedge \gamma_b = 0$, then by the first line of (A.3) we see that $\log f_\alpha(x) \to -\infty$; if $\gamma_a \vee \gamma_b = \infty$, then again $\log f_\alpha(x) \to \infty$. $\square$

PROOF OF THEOREM 2.16. In the following, the notation $\sup_\alpha, \inf_\alpha, \lim_\alpha$ is understood as taking corresponding operation over α close to 1 unless otherwise specified. We first show almost everywhere convergence by invoking Lemma E.7. To see this, for fixed $s_0 \in (-1/2, 0)$, let $g_\alpha := f_\alpha^{\alpha-1}$ and $g_\alpha^{(s_0)} := (f_\alpha)^{s_0}$. Then for $\alpha > 1+s_0$, the transformed function $g_\alpha^{(s_0)}$ is convex. We need to check two conditions in order to apply Lemma E.7 as follows:

(C1) The set $(X_{(1)}, X_{(n)}) \subset \{\liminf_\alpha f_\alpha(x) > 0\}$;

(C2) There is a uniform lower bound function $\tilde{g}^{s_0} \in \mathcal{G}$ such that $g_\alpha^{(s_0)} \geq \tilde{g}^{s_0}$ holds for α sufficiently close to 1.

The first assertion can be checked by using the characterization Theorem 2.12. Let F_α be the distribution function of f_α. Then $\int_{X_{(1)}}^t (F_\alpha - \mathbb{F}_n)(x)\,\mathrm{d}x \leq 0$ with equality attained if and only if $t \in \mathcal{S}_n(g_\alpha)$. For $x \in (X_{(1)}, X_{(n)})$ close enough to $X_{(n)}$, we claim that $\liminf_\alpha f_\alpha(x) > 0$. If not, we may assume without loss of generality that $\lim_\alpha f_\alpha(x) = 0$. We first note that there exists some $t \in \{1, \cdots n-1\}$ and some subsequence $\{\alpha(\beta)\}_{\beta\in\mathbb{N}}$ with $\alpha(\beta) \nearrow 1$ for which (1) $X_{(t)}$ is a knot point for $\{g_{\alpha(\beta)}\}$, and (2) $X_{(u)}$ is not a knot point for any $\{g_{\alpha(\beta)}\}$ for $u \geq t+1$, i.e. $g_{\alpha(\beta)}$'s are linear on $[X_{(t)}, X_{(n)}]$. We drop β for notational simplicity and assume without loss of generality that both limits $\lim_\alpha f_\alpha(X_{(n)}), \lim_\alpha f_\alpha(X_{(t)})$ exist. Now Lemma A.1 shows that $\min\{\lim_\alpha f_\alpha(X_{(n)}), \lim_\alpha f_\alpha(X_{(t)})\} = 0$ since we have assumed $\lim_\alpha f_\alpha(x) = 0$ for some $x \in (X_{(t)}, X_{(n)})$. This in turn implies that $\lim_\alpha f_\alpha(x) = 0$ for all $x \in (X_{(t)}, X_{(n)})$. Now we consider the following two cases to derive a contradiction with the fact

$$\int_{X_{(t)}}^{X_{(n)}} F_\alpha(x)\mathrm{d}x = \int_{X_{(t)}}^{X_{(n)}} \mathbb{F}_n(x)\mathrm{d}x \tag{A.5}$$

that follows from Theorem 2.12, thereby proving $\liminf_\alpha f_\alpha(x) > 0$ for x close enough to $X_{(n)}$.

[Case 1.] If $\lim_\alpha f_\alpha(X_{(n)}) = 0$, then the left hand side of (A.5) converges to $X_{(n)} - X_{(t)}$ while the right hand side is no larger than $\frac{n-1}{n}\big(X_{(n)} - X_{(t)}\big)$.

[Case 2.]. If $\lim_\alpha f_\alpha(X_{(n)}) > 0$, then we must necessarily have $\lim_\alpha f_\alpha(x) = 0$ for all $x \in [X_{(1)}, X_{(n)})$ by convexity of g_α: If $\lim_\alpha f_\alpha(x_0) > 0$ for some $x_0 \in [X_{(1)}, X_{(t)}]$, then $\lim_\alpha g_\alpha(x_0) \vee g_\alpha(X_{(n)}) < \infty$ while $\lim_\alpha g_\alpha(x) = \infty$ for all $x \in (X_{(t)}, X_{(n)})$, which is absurd. Note that this also forces $\lim_\alpha f_\alpha(X_{(n)}) = \infty$, otherwise the constraint $\int f_\alpha = 1$ will be invalid eventually. Now the left hand side of (A.5) converges to 0 while the right hand side is bounded from below by $\frac{1}{n}(X_{(n)} - X_{(t)})$.

Similarly we can show $\liminf_\alpha f_\alpha(x) > 0$ for x close to $X_{(1)}$. Now (C1) follows by convexity of f_α.

(C2) can be seen by first noting $M := \sup_\alpha \|f_\alpha\|_\infty < \infty$. This can be verified by Lemma 3.3 combined with the first assertion proved above. This implies that the class $\{g_\alpha^{(s_0)}\}_\alpha$ has a uniform lower bound M^{s_0}. Now (C2) follows by noting that the domain of all $g_\alpha^{(s_0)}$ is $\mathrm{conv}(\underline{X})$. Therefore all conditions needed for Lemma E.7 are valid, and hence we can extract a subsequence $\{g_{\alpha_n}^{(s_0)}\}_{n\in\mathbb{N}}$ such that

$$\begin{aligned} \lim_{n\to\infty, x\to y} g_{\alpha_n}^{(s_0)}(x) &= g^{(s_0)}(y), \quad \text{for all } y \in \mathrm{int}(\mathrm{dom}(g^{(s_0)})); \\ \lim_{n\to\infty, x\to y} g_{\alpha_n}^{(s_0)}(x) &\geq g^{(s_0)}(y), \quad \text{for all } y \in \mathbb{R}^d, \end{aligned}$$

holds for some $g^{(s_0)} \in \mathcal{G}$. This implies $f_{\alpha_n} \to_{a.e.} f^{(s_0)}$ as $n \to \infty$ where $f^{(s_0)} := \big(g^{(s_0)}\big)^{1/s_0}$. Now repeat the above argument with another s_1 with a further extracted subsequence $\{\alpha_{n(k)}\}$, we see that $f_{\alpha_{n(k)}} \to_{a.e.} f^{(s_1)} (k \to \infty)$ for some s_1-concave $f^{(s_1)}$ holds for the subsequence $\{\alpha_{n(k)}\}_{k\in\mathbb{N}}$. This implies that $f^{(s_0)} =_{a.e.} f^{(s_1)}$. Since a convex function is continuous in the interior of the domain, we can choose a version of upper semi-continuous f such that $f = f^{(s)}$ a.e. for all $\{1/2 < s < 0\} \cap \mathbb{Q}$. This implies that f is s-concave for any rational $1/2 < s < 0$ and hence log-concave. Next we show weighted L_1 convergence: For fixed $\kappa > 0$, choose $0 > s_0 > -1/(\kappa+1)$. Since there exists $a, b > 0$ such that $g_{\alpha_n}^{(s_0)} \geq g^{(s_0)} \geq a\|x\| - b$ holds for all $n \in \mathbb{N}$, we have an integrable envelope function:

$$\big(1 + \|x\|\big)^\kappa \big(f_{\alpha_n}(x) \vee f(x)\big) \leq \big(1 + \|x\|\big)^\kappa \Big(\big(a\|x\| - b\big) \vee M\Big)^{1/s_0}.$$

Now an application of the dominated convergence theorem yields the desired weighted L_1 convergence. Similar arguments show weighted convergence is also valid in arbitrary L_p norms ($p \geq 1$).

Finally we show that $f = f_1$ by virtue of Theorem 2.2 in Dümbgen and Rufibach (2009) and Theorem 2.9. We note that by Lemma A.1, f must be log-linear between consecutive data points. Now since f_1 and f are both log-linear between consecutive data points of $\{X_1, \ldots, X_n\}$, we only have to consider test functions h such that h is piecewise linear on consecutive data points. Recall $g_\alpha = f_\alpha^{\alpha-1}$ and $g := -\log f$ are the underlying convex functions for f_α and f. For any such h with the property that, $g + th \in \mathcal{G}$ for t small enough, we wish to argue that such h is also a valid test for f_α(i.e. $g_\alpha + th \in \mathcal{G}$ for $t > 0$ small enough), for a sequence of $\{\alpha_k\}$ converging up to 1 as $k \to \infty$. Thus we only have to argue that for all $X_{(i)} \in \mathcal{S}(g)$, $X_{(i)} \in \mathcal{S}(g_\alpha)$ for a sequence of $\{\alpha_k\}$ going up to 1 as $k \to \infty$. Assume the contrary that $X_{(i)} \notin \mathcal{S}(g_\alpha)$ for all α close enough to 1. Then $\{g_\alpha\}$'s are all linear on a closed interval $I = [a, b]$ containing $X_{(i)}$ for α close to 1. Since $f_\alpha \to f$ uniformly on I by Theorem 3.7, in particular $f_\alpha(a)$ and $f_\alpha(b)$ converges, Lemma A.1 entails that f is log-linear over I, a contradiction to the fact $X_{(i)} \in \mathcal{S}(g)$. Hence we can find a subsequence $\{\alpha_k\}$ going up to 1 as $k \to \infty$ such that for all $X_{(i)} \in \mathcal{S}(g)$, $X_{(i)} \in \mathcal{S}(g_{\alpha_k})$, i.e. for all feasible test function h of f_1, being linear on consecutive data points, is also valid for f_{α_k}. Now combining the fact that f_{α_k} converges in L_2 metric to f and Theorem 2.2 in Dümbgen and Rufibach (2009) we conclude $f_1 = f$. □

APPENDIX B: SUPPLEMENTARY PROOFS FOR SECTION 3

Proof of Lemma 3.1. The proof closely follows the first part of the proof of Proposition 2 Kim and Samworth (2015). Suppose $\dim\big(\mathrm{csupp}(\nu)\big) = d$, we show $\mathrm{csupp}(\nu) \subset \overline{C}$. To see this, we take $x_0 \notin \overline{C}$, then there exists $\delta > 0$ such that $B(x_0, \delta) \subset C^c$, and we claim that

$$\text{(B.1)} \qquad \text{For all } x^* \in B(x_0, \delta) \subset C^c, x^* \notin \mathrm{int}(\mathrm{csupp}(\nu)).$$

If (B.1) holds, then $x_0 \notin \mathrm{csupp}(\nu)$ and hence $\mathrm{csupp}(\nu) \subset \overline{C}$. Now we turn to show (B.1). Since $x^* \notin C = \{\liminf_{n\to\infty} f_n(x) > 0\}$, we can find a subsequence $\{f_{n(k)}\}_{k\in\mathbb{N}}$ of $\{f_n\}_{n\in\mathbb{N}}$ such that $f_{n(k)}(x^*) < \frac{1}{k}$ holds for all $k \in \mathbb{N}$. Hence $x^* \notin \Gamma_k := \{x \in \mathbb{R}^d : f_{n(k)}(x) \geq \frac{1}{k}\}$. Note that Γ_k is a closed convex set, hence by Hyperplane Separation Theorem we can find $b_k \in \mathbb{R}^d$ with $\|b_k\| = 1$ such that $\{x \in \mathbb{R}^d : \langle b_k, x\rangle \leq \langle b_k, x^*\rangle\} \subset (\Gamma_k)^c$. Without loss of generality we may assume $b_k \to b_{x^*}$ as $k \to \infty$ for some $b_{x^*} \in \mathbb{R}^d$ with $\|b_{x^*}\| = 1$. Now for fixed $R > 0$ and $\eta > 0$, define

$$A_{R,\eta} := \{x \in \mathbb{R}^d : \langle b_{x^*}, x\rangle < \langle b_{x^*}, x^*\rangle - \eta, \|x\| \leq R\}.$$

Choose $k_0 \in \mathbb{N}$ large enough such that $\|b_k - b_{x^*}\| \leq \frac{\eta}{2R}$ holds for all $k \geq k_0(x^*, \eta, R)$. Now for $R > \|x^*\|$ and $x \in A_{R,\eta}$, we have

$$\langle b_k, x - x^* \rangle = \langle b_{x^*}, x - x^* \rangle + \langle b_k - b_{x^*}, x - x^* \rangle < -\eta + \frac{\eta}{2R}(\|x\| + \|x^*\|) \leq 0$$

holds for all $k \geq k_0(x^*, \eta, R)$. This implies for $R > \|x^*\|$ and $\eta > 0$,

$$A_{R,\eta} \subset \{x \in \mathbb{R}^d : \langle b_k, x \rangle \leq \langle b_k, x^* \rangle\} \subset (\Gamma_k)^c = \{x \in \mathbb{R}^d : f_{n(k)}(x) < \frac{1}{k}\}.$$

Now note $A_{R,\eta}$ is open, by Portmanteau Theorem we find that

$$\nu(A_{R,\eta}) \leq \liminf_{k\to\infty} \nu_{n(k)}(A_{R,\eta}) = \liminf_{k\to\infty} \int_{A_{R,\eta}} f_{n(k)}(x)\,\mathrm{d}x \leq \liminf_{k\to\infty} \frac{\lambda_d(A_{R,\eta})}{k} = 0.$$

This implies

$$\nu\big(\{x \in \mathbb{R}^d : \langle b_{x^*}, x \rangle < \langle b_{x^*}, x^* \rangle\}\big) = \nu\bigg(\bigcup_{R=1}^{\infty} A_{R,1/R}\bigg) = \lim_{R\to\infty} \nu(A_{R,1/R}) = 0,$$

where the second equality follows from the fact $\{A_{R,1/R}\}$ is an increasing family as R increases. By the assumption that $\dim\big(\mathrm{csupp}(\nu)\big) = d$, we find $x^* \notin \mathrm{int}(\mathrm{csupp}(\nu))$, as we claimed in (B.1).

Now Suppose $\dim C = d$, we claim $\overline{C} \subset \mathrm{csupp}(\nu)$. To see this, we only have to show $C \subset \mathrm{csupp}(\nu)$ by the closedness of $\mathrm{csupp}(\nu)$. Suppose not, then we can find $x_0 \in C \setminus \mathrm{csupp}(\nu)$. This implies that there exists $\delta > 0$ such that $B(x_0, \delta) \cap \mathrm{csupp}(\nu) \neq \emptyset$. By the assumption that $\dim C = d$, we can find $x_1, \ldots, x_d \in B(x_0, \delta) \cap C$ such that $\{x_0, \ldots, x_d\}$ are in general position. By definition of C we can find $\epsilon_0 > 0, n_0 \in \mathbb{N}$ such that $f_n(x_j) \geq \epsilon_0$ for all$j = 0, 1, \ldots, d$ and $n \geq n_0$. By convexity, we conclude that $f_n(x) \geq \epsilon_0$, for all $x \in \mathrm{conv}(\{x_0, \ldots, x_d\})$ and $n \geq n_0$. This gives

$$\begin{aligned}\nu\big(\mathrm{conv}(\{x_0, \ldots, x_d\})\big) &\geq \limsup_{n\to\infty} \nu_n\big(\mathrm{conv}(\{x_0, \ldots, x_d\})\big) \\ &\geq \epsilon_0 \lambda_d\big(\mathrm{conv}(\{x_0, \ldots, x_d\})\big) > 0,\end{aligned}$$

a contradiction with $B(x_0, \delta) \cap \mathrm{csupp}(\nu) \neq \emptyset$, thus completing the proof of the claim. To summarize, we have proved

1. If $\dim\big(\mathrm{csupp}(\nu)\big) = d$, then $\mathrm{csupp}(\nu) \subset \overline{C}$. This in turn implies $\dim C = d$, and hence $\overline{C} \subset \mathrm{csupp}(\nu)$. Now it follows that $\mathrm{csupp}(\nu) = \overline{C}$;
2. If $\dim C = d$, then $\overline{C} \subset \mathrm{csupp}(\nu)$. This in turn implies $\dim\big(\mathrm{csupp}(\nu)\big) = d$, and hence $\mathrm{csupp}(\nu) \subset \overline{C}$. Now it follows that $\mathrm{csupp}(\nu) = \overline{C}$. □

PROOF OF LEMMA 3.2. The proof is essentially the same as the proof of Proposition 2 Cule and Samworth (2010) by exploiting convexity at the level of the underlying basic convex function so we shall omit it. □

PROOF OF LEMMA 3.3. Set $U_{n,t} = \{x \in \mathbb{R}^d : f_n(x) \geq t\}$. We first claim that there exists $n_0 \in \mathbb{N}, \epsilon_0 \in (0,1)$ such that $\lambda_d(U_{n,\epsilon_0}) \geq \epsilon_0$ holds for all $n \geq n_0$. If not, then for all $k \in \mathbb{N}, l \in \mathbb{N}$, there exists $n_{k,l} \in \mathbb{N}$ such that $\lambda_d(U_{n_{k,l},1/l}) \leq \frac{1}{l}$. Note that $\{\liminf_n f_n > 0\} = \cup_{k\in\mathbb{N}} \cup_{l\in\mathbb{N}} \cap_{n\geq k} U_{n,1/l}$. Since $\lambda_d\big(\bigcup_{l\in\mathbb{N}} \bigcap_{n\geq k} U_{n,1/l}\big) = \lim_{l\to\infty} \lambda_d\big(\bigcap_{n\geq k} U_{n,1/l}\big) \leq \lim_{l\to\infty} \lambda_d(U_{n_{k,l},1/l}) = 0$, we find that $C = \{\liminf_n f_n > 0\}$ is a countable union of null set and hence $\lambda_d(C) = 0$, a contradiction to the assumption $\dim C = d$. This shows the claim.

Denote $M_n := \sup_{x\in\mathbb{R}^d} f_n(x), \epsilon_n \in \operatorname{Arg\,max} f_n(x)$. Without loss of generality we assume $M_n \geq \frac{\epsilon_0}{(1+\kappa_s)^{1/s}}$ where $\kappa_s = (1/2)^s - 1 > 0$, and we set $\lambda_n := \frac{\kappa_s M_n^s}{\epsilon_0^s - M_n^s} \in [0,1]$. Now for $x \in U_{n,\epsilon_0}$, by convexity of f_n^s we have

$$f_n^s\left(\epsilon_n + \lambda_n(x - \epsilon_n)\right) \leq \lambda_n f_n^s(x) + (1-\lambda_n) f_n^s(\epsilon_n) \leq \lambda_n \epsilon_0^s + (1-\lambda_n) M_n^s = (M_n/2)^s.$$

This implies $f_n(x) \geq M_n/2 := \Omega_n$, for all $x \in V_{n,\epsilon_0} := \{\epsilon_n + \lambda_n(x - \epsilon_n) : x \in U_{n,\epsilon_0}\}$. Hence $V_{n,\epsilon_0} \subset U_{n,\Omega_n}$ and therefore $\lambda_d(V_{n,\epsilon_0}) = \lambda_d(U_{n,\epsilon_0})\lambda_n^d$, thus

$$\lambda_d(U_{n,\Omega_n}) \geq \lambda_d(V_{n,\epsilon_0}) = \lambda_d(U_{n,\epsilon_0})\lambda_n^d \geq \epsilon_0 \lambda_n^d,$$

holds for all $n \geq n_0$. On the other hand,

$$1 = \int f_n \geq \Omega_n \lambda_d(U_{n,\Omega_n}) \geq \Omega_n \epsilon_0 \lambda_n^d,$$

and suppose the contrary that $M_n \to \infty$ as $n \to \infty$, then

$$1 \geq \Omega_n \epsilon_0 \lambda_n^d = \frac{\epsilon_0 \kappa_s^d}{2(\epsilon_0^s - M_n^s)^d} M_n^{1+sd} \geq c M_n^{1+sd} \to \infty, \quad n \to \infty,$$

since $1 + sd > 0$ by assumption $-1/d < s < 0$. Here $c = \frac{\epsilon_0^{1-sd}\kappa_s^d}{2}$. This gives a contradiction and the proof is complete. □

PROOF OF THEOREM 3.4. We only have to show ν is absolutely continuous with respect to λ_d. To this end, for given $\epsilon > 0$, choose $\delta = \epsilon/2M$, where $M := \sup_n \|f_n\|_\infty < \infty$ by virtue of Lemma 3.3. Now for Borel set $A \subset \mathbb{R}^d$ with $\lambda_d(A) \leq \delta$, we can take an open $A' \supset A$ such that $\lambda_d(A') \leq 2\delta$ by the regularity of Lebesgue measure. Then

$$\nu(A) \leq \nu(A') \leq \liminf_{n\to\infty} \nu_n(A') = \liminf_{n\to\infty} \int_{A'} f_n \leq 2\delta M = \epsilon,$$

as desired. □

Proof of Lemma 3.5. Let $g_n = f_n^s$ and $g = f^s$. Without loss of generality we assume $0 \in \text{int}(\text{dom}(g))$, and choose $\eta > 0$ small enough such that $B_\eta := \overline{B}(0,\eta) \subset \text{int}(\text{dom}(g))$. By the Lemma E.4, we know there exists $a > 0, R > 0$ such that $\frac{g(x)-g(0)}{\|x\|} \geq a$, holds for all $\|x\| \geq \frac{R}{2}$. Now we claim that there exists $n_0 \in \mathbb{N}$ such that $\frac{g_n(x)-g_n(0)}{\|x\|} \geq \frac{a}{8}$, holds for all $\|x\| \geq R$ and $n \geq n_0$. Note for each $n \in \mathbb{N}$, by convexity of $g_n(\cdot)$, we know that for fixed $x \in \mathbb{R}^d$, the quantity $\frac{g_n(\lambda x)-g_n(0)}{\|\lambda x\|}$ is non-decreasing in λ, so we only have to show the claim for $\|x\| = R$ and $n_0 \geq n$. Suppose the contrary, then we can find a subsequence $\{g_{n(k)}\}$ and $\|x_{n(k)}\| = R$ such that $\frac{g_{n(k)}(x_{n(k)})-g_{n(k)}(0)}{\|x_{n(k)}\|} < \frac{a}{8}$. For simplicity of notation we think of $\{g_n\}, \{x_n\}$ as $\{g_{n(k)}\}, \{x_{n(k)}\}$. Now define $A_n := \text{conv}(\{x_n, B_\eta\}); B_n := \{y \in \mathbb{R}^d : \|y - x_n\| \leq R/2\}; C_n := A_n \cap B_n$. By reducing $\eta > 0$ if necessary, we may assume $B_\eta \cap B_n = \emptyset$. It is easy to see C_n is convex and $\lambda_d(C_n) = \lambda_0$ is a constant independent of $n \in \mathbb{N}$. By Lemma 3.2, we know that $g_n \to_{a.e.} g$ on B_η, and hence $\sup_{x \in B_\eta} |g_n(x) - g(x)| \to 0 (n \to \infty)$ by Theorem 10.8, Rockafellar (1997). By further reducing $\eta > 0$ if necessary, we may assume $g_n(y) \leq g(0) + \frac{aR}{8}$, holds for all $y \in B_\eta$ and $n \in \mathbb{N}$. Now for any $x^* \in C_n$, write $x^* = \lambda x_n + (1-\lambda)y$, by noting $R/2 \leq \|x^*\| \leq R$ and convexity of g_n, we get

$$\begin{aligned}
\frac{g_n(x^*) - g_n(0)}{\|x^*\|} &\leq \frac{\lambda g_n(x_n) + (1-\lambda)g_n(y) - g_n(0)}{\|x^*\|} \\
&= \lambda \cdot \frac{g_n(x_n) - g_n(0)}{\|x_n\|} \cdot \frac{\|x_n\|}{\|x^*\|} + (1-\lambda)\frac{g_n(y) - g_n(0)}{\|x^*\|} \\
&\leq \lambda \cdot \frac{a}{8}\frac{R}{R/2} + (1-\lambda)\frac{aR/8}{R/2} = \frac{a}{4}.
\end{aligned}$$

This gives rise to

$$\begin{aligned}
\liminf_{n\to\infty} \int_{C_n} (f_n - f) &\geq \liminf_{n\to\infty} \lambda_0\big((aR/4 + g_n(0))^{1/s} - (aR/2 + g(0))^{1/s}\big) \\
&= \lambda_0\big((aR/4 + g(0))^{1/s} - (aR/2 + g(0))^{1/s}\big) > 0,
\end{aligned}$$

which is a contradiction to Lemma E.10. This establishes our claim. Now by Lemma 3.2, we find that the set $\{\liminf_n f_n(\cdot) > 0\}$ is full-dimensional, and hence by Lemma 3.3 we conclude $g_n(\cdot)$ is uniformly bounded away from zero. Also note by Lemma E.9 we find $g(\cdot)$ must be bounded away from zero, which gives the desired assertion. □

Before the proof of Theorem 3.7, we first state some useful lemmas that give good control of tails with local information of the s-concave densities; the proof can be found in Appendix E.

LEMMA B.1. *Let $x_0, \ldots, x_d$ be $d+1$ points in $\mathbb{R}^d$ such that its convex hull $\Delta = \mathrm{conv}(\{x_0, \ldots, x_d\})$ is non-void. If $f(y) \le \min_j \left(\frac{1}{d}\sum_{i\neq j} f^s(x_i)\right)^{1/s}$, then*

$$f(y) \le f_{\max}\left(1 - \frac{d}{r} + \frac{d}{r} f_{\min} C(1+\|y\|^2)^{1/2}\right)^{-r}.$$

Here the constant $C = \lambda_d(\Delta)(d+1)^{-1/2}\sigma_{\max}(X)^{-1}$ where $X = \begin{pmatrix} x_0 & \ldots & x_d \\ 1 & \ldots & 1\end{pmatrix}$ and $f_{\min} := \min_{0\le j\le d} f(x_j)$, $f_{\max} := \max_{0\le j\le d} f(x_j)$.

LEMMA B.2. *Let ν be a probability measure with s-concave density f. Suppose that $B(0,\delta) \subset \mathrm{int}(\mathrm{dom}(f))$ for some $\delta > 0$. Then for any $y \in \mathbb{R}^d$,*

$$\sup_{x\in B(y,\delta_t)} f(x) \le J_0\left(\frac{1}{t}\left(\left(\frac{\nu(B(ty,\delta_t))}{J_0\lambda_d(B(ty,\delta_t))}\right)^{-1/r} - (1-t)\right)\right)^{-r},$$

where $J_0 := \inf_{v\in B(0,\delta)} f(v)$ and $\delta_t = \delta\frac{1-t}{1+t}$.

Now we are in position to prove Theorem 3.7.

PROOF OF THEOREM 3.7. That the sequence $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly on any compact subset in $\mathrm{int}(\mathrm{dom}(f))$ follows directly from Lemma 3.2 and Theorem 10.8 Rockafellar (1997). Now we show that if f is continuous at $y \in \mathbb{R}^d$ with $f(y) = 0$, then for any $\eta > 0$ there exists $\delta = \delta(y,\eta)$ such that

$$\limsup_{n\to\infty} \sup_{x\in B(y,\delta(y,\eta))} f_n(x) \le \eta. \tag{B.2}$$

Assume without loss of generality that $B(0,\delta_0) \subset \mathrm{int}(\mathrm{dom}(f))$ for some $\delta_0 > 0$. Let $J_0 := \inf_{x\in B(0,\delta_0)} f(x)$. Then uniform convergence of $\{f_n\}$ to f over $B(0,\delta_0)$ entails that

$$\liminf_{n\to\infty} \inf_{x\in B(0,\delta_0)} f_n(x) \ge J_0.$$

Hence with $\delta_t = \delta_0\frac{1-t}{1+t}$, it follows from Lemma B.2 that

$$\begin{aligned}\limsup_{n\to\infty} \sup_{x\in B(y,\delta_t)} f_n(x) &\le J_0\left(\frac{1}{t}\left(\left(\frac{\nu(B(ty,\delta_t))}{J_0\lambda_d(B(ty,\delta_t))}\right)^{-1/r} - (1-t)\right)\right)^{-r} \\ &\le J_0\left(\frac{J_0^{1/r}\left(\sup_{x\in B(ty,\delta_t)} f(x)\right)^{-1/r} - (1-t)}{t}\right)^{-r} \to 0\end{aligned}$$

as $t \nearrow 1$. This completes the proof for (B.2). So far we have shown that

$$\lim_{n\to\infty} \sup_{x\in S\cap B(0,\rho)} |f_n(x) - f(x)| = 0$$

holds for every $\rho \geq 0$, where S is the closed set contained in the continuity points of f. Our goal is to let $\rho \to \infty$ and conclude. Let $\Delta = \mathrm{conv}(\{x_0, \ldots, x_d\})$ be a non-void simplex with $x_0, \ldots, x_d \in \mathrm{int}(\mathrm{dom}(f))$. Note first by a closer look at the proof of Lemma 3.5, $f_n(x) \vee f(x) \leq \big((a\|x\| - b)\big)_+^{1/s}$ holds for all $x \in \mathbb{R}^d$ with some $a, b > 0$. Let $\rho_0 := \inf\{\rho \geq 0 : \big(a\rho - b\big)^{1/s} \leq f_{\min}/2\}$ where $f_{\min} := \min_{0\leq j\leq d} f(x_i) > 0$. Then

$$\begin{aligned}
\{x \in \mathbb{R}^d : \|x\| \geq \rho_0\} &\subset \bigcap_{n\geq 1}\{f_n \leq f_{\min}/2\} \bigcap \{f \leq f_{\min}/2\} \\
&\subset \bigcap_{n\geq n_0} \{f_n \leq (f_n)_{\min}\} \bigcap \{f \leq f_{\min}\} \\
&\subset \bigcap_{n\geq n_0} \{f_n \leq \min_j \big(\frac{1}{d}\sum_{i\neq j} f_n^s(x_i)\big)^{1/s}\} \bigcap \{f \leq \min_j \big(\frac{1}{d}\sum_{i\neq j} f^s(x_i)\big)^{1/s}\},
\end{aligned}$$

where $n_0 \in \mathbb{N}$ is a large constant. The second inclusion follows from the fact that $\lim_{n\to\infty} f_n(x_i) = f(x_i)$ holds for $i = 0, \ldots, d$. By Lemma B.1 we conclude that

$$\begin{aligned}
&\limsup_{n\to\infty} \sup_{x:\|x\|\geq \rho\vee\rho_0} \big(1 + \|x\|\big)^\kappa \big(f_n(x) \vee f(x)\big) \\
&\leq \sup_{x:\|x\|\geq \rho\vee\rho_0} f_{\max}\big(1 + \|x\|\big)^\kappa \bigg(1 - \frac{d}{r} + \frac{d}{r} f_{\min} C\big(1 + \|x\|^2\big)^{1/2}\bigg)^{-r} \to 0,
\end{aligned}$$

as $\rho \to \infty$. This completes the proof. □

Proof of Theorem 3.8. Since $\nabla_\xi f_n(x) = -r g_n(x)^{1/s-1} \nabla_\xi g_n(x)$,

$$\begin{aligned}
&|\nabla_\xi f_n(x) - \nabla_\xi f(x)| \\
&= r \left| g_n(x)^{1/s} \nabla_\xi g_n(x) - g(x)^{1/s} \nabla_\xi g(x) \right| \\
&\leq r \bigg(f_n(x) \left|\nabla_\xi g_n(x) - \nabla_\xi g(x)\right| + |f_n(x) - f(x)| \left|\nabla_\xi g(x)\right| \bigg) \\
&\leq 2r \sup_{x\in T} |f(x)| \left|\nabla_\xi g_n(x) - \nabla_\xi g(x)\right| + r \sup_{x\in T} |f_n(x) - f(x)| \sup_{x\in T} \|\nabla g(x)\|_2
\end{aligned}$$

holds for n large enough by Theorem 3.7. By Theorem 23.4 in Rockafellar (1997), $\nabla_\xi g_n(x) = \tau_x^T \xi$ for some $\tau_x \in \partial g_n(x)$ since $\partial g_n(x)$ is a closed set.

Thus the first term above is further bounded by

$$2r \sup_{x\in T} |f(x)| \sup_{x\in T, \tau\in\partial g_n(x)} \|\tau - \nabla g(x)\|_2,$$

which vanishes as $n \to \infty$ in view of Lemma 3.10 in Seijo and Sen (2011). Note that $\nabla g(\cdot)$ is continuous on T by Corollary 25.5.1 in Rockafellar (1997), and hence $\sup_{x\in T}\|\nabla g(x)\|_2 < \infty$. Now it is easy to see that the second term also vanishes as $n \to \infty$ by virtue of Theorem 3.7. □

APPENDIX C: SUPPLEMENTARY PROOFS FOR SECTION 4

Proof of Theorem 4.4. The proof is essentially the same as that of Theorem 3.6 Balabdaoui, Rufibach and Wellner (2009). □

Lemma C.1. *Assume (A1)-(A4). Then*

$$\int_{-\infty}^{\infty} \tilde{f}_\epsilon(x)\ \mathrm{d}x = 1 + \pi_k \frac{r g^{(k)}(m_0)}{g(m_0)^{r+1}}\epsilon^{k+1} + o(\epsilon^{k+1}),$$

where

$$\pi_k = \frac{1}{(k+1)!}\left[3^{k-1}(2k^2-4k+3)+2k^2-1\right].$$

Proof of Lemma C.1. This is straightforward calculation by Taylor expansion. Note that

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{g}_\epsilon^{-r}(x)\ \mathrm{d}x &= \int_{-\infty}^{\infty} (\tilde{g}_\epsilon^{-r}(x) - g^{-r}(x))\ \mathrm{d}x + 1 \\
&= \int_{m_0-c_\epsilon\epsilon}^{m_0-\epsilon} \Big(\tilde{g}_\epsilon^{-r}(x) - g^{-r}(x)\Big)\ \mathrm{d}x \\
&\quad + \int_{m_0-\epsilon}^{m_0+\epsilon} \Big(\tilde{g}_\epsilon^{-r}(x) - g^{-r}(x)\Big)\ \mathrm{d}x + 1 \\
&:= I + II + 1.
\end{aligned}$$

For $y > x$, we have $x^{-r} - y^{-r} = \sum_{n\geq 1}^{\infty} \binom{-r}{n}(-1)^n (y-x)^n y^{-r-n}$. Now for the first term above, we continue our calculation of its leading term by noting

(C.1)

$$\begin{aligned}
& g(x) - \tilde{g}_\epsilon(x) \\
&= g(x) - g(m_0 - c_\epsilon \epsilon) - (x - m_0 + c_\epsilon \epsilon) g'(m_0 - c_\epsilon \epsilon) \\
&= g(m_0) + \frac{g^{(k)}(m_0)}{k!}(x - m_0)^k - \left[g(m_0) + \frac{g^{(k)}(m_0)}{k!}(-c_\epsilon \epsilon)^k \right] \\
&\qquad - (x - m_0 + c_\epsilon \epsilon) \frac{g^{(k)}(m_0)}{(k-1)!}(-c_\epsilon \epsilon)^{k-1} + \text{ higher order terms} \\
&= \frac{g^{(k)}(m_0)}{k!} \left[(x - m_0)^k - c_\epsilon^k \epsilon^k + k c_\epsilon^{k-1} \epsilon^{k-1}(x - m_0 + c_\epsilon \epsilon) \right] + \text{ higher order terms.}
\end{aligned}$$

Here we used the fact k is an even number, as shown in Lemma D.1. Thus we have

$$\begin{aligned}
& \text{leading term of I} \\
&= \int_{m_0 - c_\epsilon \epsilon}^{m_0 - \epsilon} r \left(g(x) - g(m_0 - c_\epsilon \epsilon) - (x - m_0 + c_\epsilon \epsilon) g'(m_0 - c_\epsilon \epsilon) \right) g(x)^{-r-1} \, \mathrm{d}x \\
&= \frac{r g^{(k)}(m_0)}{k! g(m_0)^{r+1}} \int_{m_0 - c_\epsilon \epsilon}^{m_0 - \epsilon} \left[(x - m_0)^k - c_\epsilon^k \epsilon^k + k c_\epsilon^{k-1} \epsilon^{k-1}(x - m_0 + c_\epsilon \epsilon) \right] \mathrm{d}x + o(\epsilon^{k+1}) \\
&= \alpha_k \frac{r g^{(k)}(m_0)}{g(m_0)^{r+1}} \epsilon^{k+1} + o(\epsilon^{k+1})
\end{aligned}$$

Here

$$\alpha_k = \frac{1}{(k+1)!} \left[3^{k-1}(2k^2 - 4k + 3) - 1 \right].$$

For the second term,

(C.2)

$$\begin{aligned}
& g(x) - \tilde{g}_\epsilon(x) \\
&= g(x) - g(m_0 + \epsilon) - (x - m_0 - \epsilon) g'(m_0 + \epsilon) \\
&= \frac{g^{(k)}(m_0)}{k!} \left[(x - m_0)^k - \epsilon^k - k \epsilon^{k-1}(x - m_0 - \epsilon) \right] + \text{ higher order terms.}
\end{aligned}$$

Now similar calculations yield that the second term $= \beta_k \frac{r g^{(k)}(m_0)}{g(m_0)^{r+1}} \epsilon^{k+1} + o(\epsilon^{k+1})$ with

$$\beta_k = \frac{2k^2}{(k+1)!}.$$

This gives the conclusion. □

PROOF OF LEMMA 4.6. By definition of the Hellinger metric and Lemma C.1, we have

$$
\begin{aligned}
2h^2(f_\epsilon, f) &= \int_{-\infty}^{\infty} \left(\sqrt{f_\epsilon(x)} - \sqrt{f(x)}\right)^2 \,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(\tilde{g}_\epsilon^{-r/2}(x)\left(1 - \frac{\pi_k}{2}\frac{rg^{(k)}(m_0)}{g(m_0)^{r+1}}\epsilon^{k+1} + o(\epsilon^{k+1})\right) - g^{-r/2}(x)\right)^2 \,\mathrm{d}x \\
&\equiv \int_{-\infty}^{\infty} \left(\tilde{g}_\epsilon^{-r/2}(x)(1+\eta_k(\epsilon)) - g^{-r/2}(x)\right)^2 \,\mathrm{d}x
\end{aligned}
$$

since

$$
\begin{aligned}
f_\epsilon(x) &= \tilde{g}_\epsilon^{-r}(x)\left(1 + \pi_k \frac{rg^{(k)}(m_0)}{g(m_0)^{r+1}}\epsilon^{k+1} + o(\epsilon^{k+1})\right)^{-1} \\
&= \tilde{g}_\epsilon^{-r}(x)\left(1 - \pi_k \frac{rg^{(k)}(m_0)}{g(m_0)^{r+1}}\epsilon^{k+1} + o(\epsilon^{k+1})\right).
\end{aligned}
$$

Here $\eta_k(\epsilon) = O(\epsilon^{k+1})$. Splitting two terms apart in the above integral we get

$$
\begin{aligned}
2h^2(f_\epsilon, f) &= \int_{-\infty}^{\infty} \left(\tilde{g}_\epsilon^{-r/2}(x) - g^{-r/2}(x) + \eta_k(\epsilon)\tilde{g}_\epsilon^{-r/2}(x)\right)^2 \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(\tilde{g}_\epsilon^{-r/2}(x) - g^{-r/2}(x)\right)^2 \,\mathrm{d}x + \left(\eta_k(\epsilon)\right)^2 \int_{-\infty}^{\infty} \tilde{g}_\epsilon^{-r}(x) \,\mathrm{d}x \\
&\quad + 2\eta_k(\epsilon) \int_{-\infty}^{\infty} \tilde{g}_\epsilon^{-r/2}(x)\left(\tilde{g}_\epsilon^{-r/2}(x) - g^{-r/2}(x)\right) \mathrm{d}x \\
&= I + II + III.
\end{aligned}
$$

Now for the first term,

$$
\begin{aligned}
I &= \int_{m_0 - c_\epsilon \epsilon}^{m_0+\epsilon} \frac{r^2}{4}\left[g(x) - \tilde{g}_\epsilon(x)\right]^2 g(x)^{-r-2} \,\mathrm{d}x + \text{ higher order terms} \\
&= \frac{r^2}{4g(m_0)^{r+2}} \int_{m_0 - c_\epsilon \epsilon}^{m_0+\epsilon} \left[g(x) - \tilde{g}_\epsilon(x)\right]^2 \,\mathrm{d}x + \text{ higher order terms} \\
&= \frac{r^2}{4g(m_0)^{r+2}}\left(\int_{m_0 - c_\epsilon \epsilon}^{m_0-\epsilon} + \int_{m_0-\epsilon}^{m_0+\epsilon}\right)\left[g(x) - \tilde{g}_\epsilon(x)\right]^2 \,\mathrm{d}x + \text{ higher order terms} \\
&= I_1 + I_2 + \text{ higher order terms.}
\end{aligned}
$$

By (C.1) and (C.2) we see that for $i = 1, 2$,

$$\begin{aligned} I_i &= \frac{r^2}{4g(m_0)^{r+2}} \int_{\mathcal{I}_i} \big[g(x) - \tilde{g}_\epsilon(x)\big]^2 \, \mathrm{d}x \\ &= \zeta_k^{(i)} \frac{r^2 f(m_0) g^{(k)}(m_0)^2}{g(m_0)^2} \epsilon^{2k+1} + o(\epsilon^{2k+1}). \end{aligned}$$

Here $\mathcal{I}_1 = [m_0 - c_\epsilon \epsilon, m_0 - \epsilon]$, $\mathcal{I}_2 = [m_0 - \epsilon, m_0 + \epsilon]$, and

$$\begin{aligned} \zeta_k^{(1)} = \frac{1}{108(k!)^2(k+1)(k+2)(2k+1)} \bigg[&- 4 \cdot 3^{k+2}(2k+1)(3^{k+2} + k^2 + k - 3) \\ &+ (k+1)(k+2)\Big(27(3^{2k+1} - 1) + 2 \cdot 3^{2k}(2k+1)(2k(2k-9) + 27)\Big)\bigg]. \end{aligned}$$

$$\zeta_k^{(2)} = \frac{2k^2(2k^2+1)}{3(k!)^2(k+1)(2k+1)}.$$

On the other hand, $II = O(\epsilon^{(2k+2)}) = o(\epsilon^{2k+1})$ and $|III| \leq O(\epsilon^{k+1} \cdot \epsilon^{(2k+1)/2} \cdot \epsilon^{(2k+2)/2}) = o(\epsilon^{2k+1})$ by Cauchy-Schwarz. This completes the proof. □

APPENDIX D: PROOF OF THEOREM 6.1

We first observe that

LEMMA D.1. *k is an even integer and $g_0^{(k)}(x_0) > 0$.*

PROOF OF LEMMA D.1. By Taylor expansion of g_0'' around x_0, we find that locally for $x \approx x_0$,

$$g_0''(x) = \frac{g_0^{(k)}(x_0)}{(k-2)!}(x - x_0)^{k-2} + o\big((x - x_0)^{k-2}\big).$$

Also note $g_0''(x) \geq 0$ by convexity and local smoothness assumed in (A3). This gives that $k - 2$ is even and $g_0^{(k)}(x_0) > 0$. □

For further technical discussions, we denote throughout this subsection that for fixed k, $r_n := n^{\frac{k+2}{2k+1}}; s_n := n^{-\frac{1}{2k+1}}; x_n(t) := x_0 + s_n t; \boldsymbol{l}_{n,x_0} := [x_0, x_n(t)]$. Let $\tau_n^+ := \inf\{t \in \mathcal{S}_n(\hat{g}_n) : t > x_0\}$, and $\tau_n^- := \sup\{t \in \mathcal{S}_n(\hat{g}_n) : t < x_0\}$. The key step in establishing the limit theory, is to establish a stochastic bound for the gap $\tau_n^+ - \tau_n^-$ as follows.

THEOREM D.2. *Assume (A1)-(A4) hold. Then*

$$\tau_n^+ - \tau_n^- = O_p(s_n).$$

PROOF. Define $\Delta_0(x) := (\tau_n^- - x)\mathbf{1}_{[\tau_n^-,\bar{\tau}]}(x) + (x - \tau_n^+)\mathbf{1}_{[\bar{\tau},\tau_n^+]}(x)$, and $\Delta_1 := \Delta_0 + \frac{\tau_n^+ - \tau_n^-}{4}\mathbf{1}_{[\tau_n^-,\tau_n^+]}$, where $\bar{\tau} =: \frac{\tau_n^- + \tau_n^+}{2}$. Thus we find that

$$\begin{aligned}\int \Delta_1 \ \mathrm{d}(\mathbb{F}_n - F_0) &= \int \Delta_1 \ \mathrm{d}(\mathbb{F}_n - \hat{F}_n) + \int \Delta_1 \ \mathrm{d}(\hat{F}_n - F_0) \\ &\geq -\frac{\tau_n^+ - \tau_n^-}{4}\left|\int_{\tau_n^-}^{\tau_n^+} \mathrm{d}(\mathbb{F}_n - \hat{F}_n)\right| + \int \Delta_1(\hat{f}_n - f_0) \ \mathrm{d}\lambda \\ &\geq -\frac{\tau_n^+ - \tau_n^-}{2n} + \int \Delta_1(\hat{f}_n - f_0) \ \mathrm{d}\lambda,\end{aligned}$$

where the last line follows from Corollary 2.13. Now let $R_{1n} := \int \Delta_1(\hat{f}_n - f_0) \ \mathrm{d}\lambda, R_{2n} := \int \Delta_1 \ \mathrm{d}(\mathbb{F}_n - F_0)$. The conclusion follows directly from the following lemma. □

LEMMA D.3. *Suppose (A1)-(A4) hold. Then* $R_{1n} = O_p(\tau_n^+ - \tau_n^-)^{k+2}$ *and* $R_{2n} = O_p(r_n^{-1})$.

PROOF OF LEMMA D.3. Define $p_n := \hat{g}_n/g_0$ on $[\tau_n^+, \tau_n^-]$. It is easy to see that $\tau_n^+ - \tau_n^- = o_p(1)$, so with large probability, for all $n \in \mathbb{N}$ large enough, $\inf_{x\in[\tau_n^+,\tau_n^-]} f_0(x) > 0$ by (A2).

$$\begin{aligned}R_{1n} &= \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x)\big(\hat{f}_n(x) - f_0(x)\big) \ \mathrm{d}x = \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) f_0(x)\left(\frac{\hat{f}_n(x)}{f_0(x)} - 1\right) \mathrm{d}x \\ &= \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) f_0(x)\left(\sum_{j=1}^{k-1}\binom{-r}{j}(p_n(x) - 1)^j + \binom{-r}{k}\theta_{x,n}^{-r-k}(p_n(x) - 1)^k\right) \mathrm{d}x,\end{aligned}$$

where $\theta_{x,n} \in [1 \wedge \frac{\hat{g}_n(x)}{g_0(x)}, 1 \vee \frac{\hat{g}_n(x)}{g_0(x)}]$. Now define

$$\begin{aligned}S_{nj} &= \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) f_0(x)\binom{-r}{j}(p_n(x) - 1)^j \ \mathrm{d}x, 1 \leq j \leq k - 1, \\ S_{nk} &= \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) f_0(x)\binom{-r}{k}\theta_{x,n}^{-r-k}(p_n(x) - 1)^k \ \mathrm{d}x.\end{aligned}$$

Expand f_0 around $\bar{\tau}$, then we have

$$S_{nj} = \sum_{l=0}^{k-1} \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) \frac{f_0^{(l)}(\bar{\tau})}{l!} (x-\bar{\tau})^l \binom{-r}{j} (p_n(x)-1)^j \, \mathrm{d}x$$

$$+ \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) \frac{f_0^{(l)}(\eta_{n,x,k})}{k!} (x-\bar{\tau})^k \binom{-r}{k} (p_n(x)-1)^k \, \mathrm{d}x,$$

$$S_{nk} = \sum_{l=0}^{k-1} \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) \frac{f_0^{(l)}(\bar{\tau})}{l!} \theta_{x,n}^{-r-k} (x-\bar{\tau})^l \binom{-r}{j} (p_n(x)-1)^k \, \mathrm{d}x$$

$$+ \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) \frac{f_0^{(l)}(\eta_{n,x,k})}{k!} \theta_{x,n}^{-r-k} (x-\bar{\tau})^k \binom{-r}{k} (p_n(x)-1)^k \, \mathrm{d}x.$$

Now we see the dominating term is the first term in S_{n1} since all other terms are of higher orders, and $|\theta_{x,n} - 1| = o_p(1)$ uniformly locally in x in view of Theorem 3.7. We denote this term Q_{n1}. Note that $1/g_0(x_0) = 1/g_0(\tau) + o_p(1)$ uniformly in τ around x_0, and that $\hat{g}_n$ is piecewise linear, yielding

$$\begin{aligned}
\frac{Q_{n1}}{-rf_0(\bar{\tau})} &= \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) \frac{1}{g_0(x)} \big(\hat{g}_n(x) - g_0(x)\big) \, \mathrm{d}x \\
&= \left(\frac{1}{g_0(x_0)} + o_p(1)\right) \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) \big(\hat{g}_n(x) - g_0(x)\big) \, \mathrm{d}x \\
&= \left(\frac{1}{g_0(x_0)} + o_p(1)\right) \Bigg[\big(\hat{g}_n(\bar{\tau}) - g_0(\bar{\tau})\big) \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x) \, \mathrm{d}x \\
&\qquad + \big(\hat{g}_n'(\bar{\tau}) - g_0'(\bar{\tau})\big) \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x)(x-\bar{\tau}) \, \mathrm{d}x \\
&\qquad - \sum_{j=2}^{k} \frac{g_0^{(j)}(\bar{\tau})}{j!} \int_{\tau_n^-}^{\tau_n^+} \Delta_1(x)(x-\bar{\tau})^j \, \mathrm{d}x \\
&\qquad - \int_{\tau_n^-}^{\tau_n^+} \epsilon_n(x) \Delta_1(x)(x-\bar{\tau})^k \, \mathrm{d}x \Bigg],
\end{aligned}$$

where the first two terms in the bracket is zero by construction of Δ_1. Now note that

$$\int_{\tau_n^-}^{\tau_n^+} \Delta_1(x)(x-\bar{\tau})^j \, \mathrm{d}x = \begin{cases} 0 & j = 0, \text{ or } j \text{ is odd;} \\ \frac{j}{2^{j+2}(j+1)(j+2)} \big(\tau_n^+ - \tau_n^-\big)^{j+2} & j > 0, \text{ and } j \text{ is even,} \end{cases}$$

and that $g_0^{(j)}(\bar{\tau}) = \frac{1}{(k-j)!}(g_0^{(k)}(x_0) + o_p(1))\big(\bar{\tau} - x_0\big)^{k-j}$. This means that for

$j \geq 2$ and j even,

$$\frac{g_0^{(j)}(\bar{\tau})}{j!}\int_{\tau_n^-}^{\tau_n^+} \Delta_1(x)(x-\bar{\tau})^j \, \mathrm{d}x = \frac{j(g_0^{(k)}(x_0)+o_p(1))}{(k-j)!(j+2)!2^{j+2}}(\bar{\tau}-x_0)^{k-j}(\tau_n^+-\tau_n^-)^{j+2}$$
$$= \frac{j(g_0^{(k)}(x_0)+o_p(1))}{(k-j)!(j+2)!2^{j+2}}O_p(1)(\tau_n^+-\tau_n^-)^{k+2}.$$

Further note that $\|\epsilon_n\|_\infty = o_p(1)$ as $\tau_n^+ - \tau_n^- \to_p 0$, we get $Q_{n1} = O_p(\tau_n^+ - \tau_n^-)^{k+2}$. This establishes the first claim. The proof for R_{2n} follows the same line as in the proof of Lemma 4.4 Balabdaoui, Rufibach and Wellner (2009) p1318-1319. □

LEMMA D.4. *We have the following:*

$$f_0^{(j)}(x_0) = j!\binom{-r}{j} g_0(x_0)^{-r-j}\big(g_0'(x_0)\big)^j, 1 \leq j \leq k-1;$$
$$f_0^{(k)}(x_0) = k!\binom{-r}{k} g_0(x_0)^{-r-k}\big(g_0'(x_0)\big)^k - r g_0(x_0)^{-r-1} g_0^{(k)}(x_0).$$

PROOF. This follows from direct calculation. □

LEMMA D.5. *For any $M > 0$, we have*

$$\sup_{|t|\leq M}\left|\hat{g}_n'(x_0+s_nt) - \hat{g}_0'(x_0)\right| = O_p(s_n^{k-1});$$
$$\sup_{|t|\leq M}\left|\hat{g}_n(x_0+s_nt) - g_0(x_0) - s_nt g_0'(x_0)\right| = O_p(s_n^k).$$

The proof is identical to Lemma 4.4 in Groeneboom, Jongbloed and Wellner (2001) so we shall omit it.

LEMMA D.6. *Let*

$$\hat{e}_n(u) := \hat{f}_n(u) - \sum_{j=0}^{k-1}\frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j - f_0(x_0)\binom{-r}{k}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^k (u-x_0)^k.$$

Then for any $M > 0$, we have $\sup_{|t|\leq M}|\hat{e}_n(x_0+s_nt)| = O_p(s_n^k)$.

PROOF. Note that

(D.1)

$$\hat{f}_n(u)-f_0(x_0)=f_0(x_0)\left[\frac{\hat{f}_n(u)}{f_0(x_0)}-1\right]=f_0(x_0)\left[\left(\frac{\hat{g}_n(u)}{g_0(x_0)}\right)^{-r}-1\right]$$
$$=f_0(x_0)\left(\sum_{j=1}^{k}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j+\underbrace{\sum_{j\geq k+1}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j}_{=:\hat{\Psi}_{k,n,1}(u)}\right).$$

Define $\hat{\Psi}_{k,n,1}(u):=\sum_{j\geq k+1}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j=\sum_{j\geq k+1}\binom{-r}{j}\frac{1}{g_0(x_0)^j}\big(\hat{g}_n(u)-g_0(x_0)\big)^j$. Note that

$$\begin{aligned}\big(\hat{g}_n(u)-g_0(x_0)\big)^j&=\big(\hat{g}_n(u)-g_0(x_0)-(u-x_0)g_0'(x_0)+(u-x_0)g_0'(x_0)\big)^j\\&=\sum_{l=1}^{j}\binom{j}{l}\big[\hat{g}_n(u)-g_0(x_0)-(u-x_0)g_0'(x_0)\big]^l(u-x_0)^{j-l}g_0'(x_0)^{j-l}\\&\qquad+(u-x_0)^jg_0'(x_0)^j\\&=O_p(s_n^{kl}\cdot s_n^{j-l})+O_p(s_n^j)\\&\qquad\qquad\text{uniformly on } \{u:\lvert u-x_0\rvert\leq Mn^{-1/(2k+1)}\}\\&=O_p(n^{-\frac{j}{2k+1}}),\end{aligned}$$

if $j\geq k+1$. Here the third line follows from Lemma D.5. This implies $\hat{\Psi}_{k,n,1}(u)=o_p(n^{-\frac{k}{2k+1}})$, uniformly on $\{u:\lvert u-x_0\rvert\leq Mn^{-1/(2k+1)}\}$. Using the same expansion in the first term on the right hand side of (D.1), we arrive at

$$\begin{aligned}&\underbrace{\hat{f}_n(u)-f_0(x_0)}_{(1)}\\&=\underbrace{f_0(x_0)\sum_{j=1}^{k}\binom{-r}{j}\frac{1}{[g_0(x_0)]^j}\sum_{r=1}^{j}\binom{j}{r}\big[\hat{g}_n(u)-g_0(x_0)-(u-x_0)g_0'(x_0)\big]^r(u-x_0)^{j-r}g_0(x_0)^{j-r}}_{(2)}\\&\qquad+\underbrace{f_0(x_0)\sum_{j=1}^{k}\binom{-r}{j}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^j(u-x_0)^j}_{(3)}+\underbrace{f_0(x_0)\hat{\Psi}_{k,n,1}(u)}_{(4)}.\end{aligned}$$

By Lemma D.4, we see that $\hat{e}_n(u)=(1)-(3)=(2)+(4)=O_p(s_n^k)$ uniformly on $\{u:\lvert u-x_0\rvert\leq Mn^{-1/(2k+1)}\}$. This yields the desired result. ☐

We are now ready for the proof of Theorem 6.1.

PROOF OF THEOREM 6.1. For the first assertion, note that

$$
\begin{aligned}
&[f_0(x_0)]^{-1}\left(\hat{f}_n(u) - \sum_{j=0}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j\right) \\
=&[f_0(x_0)]^{-1}\left(\hat{f}_n(u) - f_0(x_0) - \sum_{j=1}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j\right) \\
=&[f_0(x_0)]^{-1}\left(f_0(x_0)\left(\sum_{j=1}^{k}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j + \hat{\Psi}_{k,n,1}(u)\right) - \sum_{j=1}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j\right) \\
&\qquad \text{by } (D.1) \\
=&\hat{\Psi}_{k,n,1}(u) + \sum_{j=1}^{k}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j - [f_0(x_0)]^{-1}\sum_{j=1}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j \\
=&\hat{\Psi}_{k,n,1}(u) + \binom{-r}{1}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right) - \frac{1}{f_0(x_0)}f_0'(x_0)(u-x_0) \\
&\quad + \sum_{j=2}^{k}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j - [f_0(x_0)]^{-1}\sum_{j=2}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j \\
=&\hat{\Psi}_{k,n,1}(u) - \frac{r}{g_0(x_0)}\left(\hat{g}_n(u) - g_0(x_0) - g_0'(x_0)(u-x_0)\right) + \sum_{j=2}^{k}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j \\
&\quad - [f_0(x_0)]^{-1}\sum_{j=2}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j \\
=&-\frac{r}{g_0(x_0)}\left(\hat{g}_n(u) - g_0(x_0) - g_0'(x_0)(u-x_0)\right) + \hat{\Psi}_{k,n,2}(u),
\end{aligned}
$$

where

$$
\hat{\Psi}_{k,n,2}(u) := \hat{\Psi}_{k,n,1}(u) + \sum_{j=2}^{k}\binom{-r}{j}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j - [f_0(x_0)]^{-1}\sum_{j=2}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j.
$$

Now we calculate

$$
\begin{aligned}
&\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}\hat{\Psi}_{k,n,2}(u)\mathrm{d}u\mathrm{d}v\\
=&\frac{1}{2}t^2 n^{-\frac{2}{2k+1}}\sup_{u\in\boldsymbol{l}_{n,x_0}}\left|\hat{\Psi}_{k,n,1}(u)\right| + \sum_{j=2}^{k}\binom{-r}{j}\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}\left(\frac{\hat{g}_n(u)}{g_0(x_0)}-1\right)^j \mathrm{d}u\mathrm{d}v\\
&\qquad - [f_0(x_0)]^{-1}\sum_{j=2}^{k-1}\frac{f_0^{(j)}(x_0)}{j!}\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}(u-x_0)^j \mathrm{d}u\mathrm{d}v\\
=&o_p(r_n^{-1}) + \sum_{j=2}^{k}\binom{-r}{j}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^j\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}(u-x_0)^j \mathrm{d}u\mathrm{d}v\\
&\qquad - \sum_{j=2}^{k-1}\binom{-r}{j}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^j\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}(u-x_0)^j \mathrm{d}u\mathrm{d}v\\
&\qquad + \Bigg(\sum_{j=2}^{k}\binom{-r}{j}\frac{1}{[g_0(x_0)]^j}\\
&\qquad\quad \times\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}\sum_{l=1}^{j}\binom{j}{l}\big(\hat{g}_n(u)-g_0(x_0)-g_0'(x_0)(u-x_0)\big)^l(u-x_0)^{j-l}[g_0'(x_0)]^{j-l}\mathrm{d}u\mathrm{d}v\Bigg)\\
=&o_p(r_n^{-1}) + \binom{-r}{k}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^k\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}(u-x_0)^k \mathrm{d}u\mathrm{d}v\\
&+ \Bigg(\sum_{j=2}^{k}\binom{-r}{j}\frac{1}{[g_0(x_0)]^j}\\
&\qquad \times\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}\sum_{l=1}^{j}\binom{j}{l}\big(\hat{g}_n(u)-g_0(x_0)-g_0'(x_0)(u-x_0)\big)^l(u-x_0)^{j-l}[g_0'(x_0)]^{j-l}\mathrm{d}u\mathrm{d}v\Bigg)\\
:=&o_p(r_n^{-1}) + (2) + (1).
\end{aligned}
$$

Consider (1): for each (j,l) satisfying $1\leq l\leq j\leq k$ and $j\geq 2$, we have

$$
\begin{aligned}
(1): r_n&\int_{\boldsymbol{l}_{n,x_0}}\int_{x_0}^{v}\big(\hat{g}_n(u)-g_0(x_0)-g_0'(x_0)(u-x_0)\big)^l(u-x_0)^{j-l}[g_0'(x_0)]^{j-l}\mathrm{d}u\mathrm{d}v\\
&= n^{\frac{k+2}{2k+1}}\cdot O(n^{-\frac{2}{2k+1}})\cdot O_p(n^{-\frac{kl}{2k+1}})\cdot O_p(n^{-\frac{j-l}{2k+1}}) = O_p(n^{-\frac{k(l-1)+(j-l)}{2k+1}}) = o_p(1).
\end{aligned}
$$

Consider (2) as follows:

$$(2) = \binom{-r}{k}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^k \int_{\boldsymbol{l}_{n,x_0}} \int_{x_0}^{v} (u-x_0)^k \, \mathrm{d}u\mathrm{d}v$$
$$= \frac{1}{(k+1)(k+2)}\binom{-r}{k}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^k t^{k+2} r_n^{-1}.$$

Hence we have

$$r_n \int_{\boldsymbol{l}_{n,x_0}} \int_{x_0}^{v} \hat{\Psi}_{k,n,2}(u)\mathrm{d}u\mathrm{d}v = \frac{1}{(k+1)(k+2)}\binom{-r}{k}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^k t^{k+2} + o_p(1).$$

Note by definition we have

$$\mathbb{Y}_n^{\mathrm{locmod}}(t) = \frac{\mathbb{Y}_n^{\mathrm{loc}}(t)}{f_0(x_0)} - r_n \int_{\boldsymbol{l}_{n,x_0}} \int_{x_0}^{v} \hat{\Psi}_{k,n,2}(u)\mathrm{d}u\mathrm{d}v. \tag{D.2}$$

Let $n \to \infty$, by the same calculation in the proof of Theorem 6.2 Groeneboom, Jongbloed and Wellner (2001), we have

$$\mathbb{Y}_n^{\mathrm{locmod}}(t) \to_d \frac{1}{\sqrt{f_0(x_0)}} \int_0^t W(s)\, \mathrm{d}s$$
$$+ \left[\frac{f_0^{(k)}(x_0)}{(k+2)!f_0(x_0)} - \frac{1}{(k+1)(k+2)}\binom{-r}{k}\left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^k\right] t^{k+2}$$
$$= \frac{1}{\sqrt{f_0(x_0)}} \int_0^t W(s)\, \mathrm{d}s - \frac{r g_0^{(k)}(x_0)}{g_0(x_0)(k+2)!} t^{k+2},$$

where the last line follows from Lemma D.4. Now we turn to the second assertion. It is easy to check by the definition of $\hat{\Psi}_{k,n,2}(\cdot)$ that

$$\mathbb{H}_n^{\mathrm{locmod}}(t) = \frac{\mathbb{H}_n^{\mathrm{loc}}(t)}{f_0(x_0)} - r_n \int_{\boldsymbol{l}_{n,x_0}} \int_{x_0}^{v} \hat{\Psi}_{k,n,2}(u)\mathrm{d}u\mathrm{d}v. \tag{D.3}$$

On the other hand, simple calculation yields that $\mathbb{Y}_n^{\mathrm{loc}}(t) - \mathbb{H}_n^{\mathrm{loc}}(t) = r_n\big(\mathbb{H}_n(x_0 + s_n t) - \hat{H}_n(x_0 + s_n t)\big) \geq 0$ where the inequality follows from Theorem 2.12. Combined with (D.2) and (D.3) we have shown the second assertion. Finally we show tightness of $\{\hat{A}_n\}$ and $\{\hat{B}_n\}$. By Theorem D.2, we can find $M > 0$ and $\tau \in \mathcal{S}(\hat{g}_n)$ such that $0 \leq \tau - x_0 \leq M n^{-1/(2k+1)}$ with large probability.

Now note

$$\begin{aligned}\left|\hat{A}_n\right| &\leq r_n s_n \left|\left(\hat{F}_n(x_0) - \hat{F}_n(\tau)\right) - \left(\mathbb{F}_n(x_0) - \mathbb{F}_n(\tau)\right)\right| + \frac{r_n s_n}{n} \\ &\leq r_n s_n \left|\int_{x_0}^{\tau} \left(\hat{f}_n(u) - \sum_{j=0}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j\right) \mathrm{d}u\right| \\ &\quad + r_n s_n \left|\int_{x_0}^{\tau} \left(\sum_{j=0}^{k-1} \frac{f_0^{(j)}(x_0)}{j!}(u-x_0)^j - f_0(u)\right) \mathrm{d}u\right| \\ &\quad + r_n s_n \left|\int_{x_0}^{\tau} \mathrm{d}(\mathbb{F}_n - F_0)\right| + n^{-k/(2k+1)} \\ &=: \hat{A}_{n1} + \hat{A}_{n2} + \hat{A}_{n3} + n^{-k/(2k+1)}.\end{aligned}$$

We calculate three terms respectively.

$$\begin{aligned}\hat{A}_{n1} &\leq r_n s_n \left|\int_{x_0}^{\tau} \hat{e}_n(u)\, \mathrm{d}u\right| + r_n s_n \left|\int_{x_0}^{\tau} f_0(x_0) \binom{-r}{k} \left(\frac{g_0'(x_0)}{g_0(x_0)}\right)^k (u-x_0)^k\, \mathrm{d}u\right| \\ &= O_p(r_n s_n \cdot s_n^{k+1}) + o_p(r_n s_n \cdot s_n^{k+1}) = O_p(1), \qquad \text{by Lemma D.6} \\ \hat{A}_{n2} &\leq r_n s_n \left|\int_{x_0}^{\tau} \frac{f_0^{(k)}(x_0)}{k!}(u-x_0)^k\, \mathrm{d}u\right| + r_n s_n \left|\int_{x_0}^{\tau} (u-x_0)^k \epsilon_n(u)\, \mathrm{d}u\right| \\ &= O_p(1), \qquad \text{since } \|\epsilon_n\|_\infty \to_p 0 \text{ as } x_0 - \tau \to_p 0.\end{aligned}$$

For $\hat{A}_{n3}$, we follow the lines of Lemma 4.1 Balabdaoui, Rufibach and Wellner (2009) again to conclude. Fix $R > 0$, and consider the function class $\mathcal{F}_{x_0,R} := \{\mathbf{1}_{[x_0,y]} : x_0 \leq y \leq x_0 + R\}$. Then $F_{x_0,R}(z) := \mathbf{1}_{[x_0,x_0+R]}(z)$ is an envelop function for $\mathcal{F}_{x_0,R}$, and $\mathbb{E}F^2_{x_0,R} = \int_{x_0}^{x_0+R} \mathrm{d}z = R$. Now let $s = k, d = 1$ in Lemma 4.1 Balabdaoui, Rufibach and Wellner (2009), we have

$$\hat{A}_{n3} = \left|\int_{x_0}^{\tau} \mathrm{d}(\mathbb{F}_n - F_0)(z)\right| \leq |\tau - x_0|^{k+1} + O_p(1) n^{-\frac{k+1}{2k+1}} = O_p(1).$$

This completes the proof for tightness for $\{A_n\}$. $\{B_n\}$ follows from similar argument so we omit the details. □

APPENDIX E: AUXILIARY RESULTS

E.1. Proosf of Lemmas B.1 and B.2.

LEMMA E.1. *Let ν be a probability measure with s-concave density f, and $x_0, \ldots, x_d \in \mathbb{R}^d$ be $d+1$ points such that $\Delta := \mathrm{conv}(\{x_0, \ldots, x_d\})$ is non-void. If $f(x_0) \le \left(\frac{1}{d}\sum_{i=1}^d f^s(x_i)\right)^{1/s}$, then*

$$f(x_0) \le \bar{g}^{-r}\left(1 - \frac{d}{r} + \frac{d}{r}\frac{\lambda_d(\Delta)\bar{g}^{-r}}{\nu(\Delta)}\right)^{-r},$$

where $\bar{g} := \frac{1}{d}\sum_{j=1}^d f^s(x_j)$.

PROOF OF LEMMA E.1. For any point $x \in \Delta$, we can find some $u = (u_1, \ldots, u_d) \in \Delta_d = \{u : \sum_{i=1}^d u_i \le 1\}$ such that $x(u) = \sum_{i=0}^d u_i x_i$. Here $u_0 := 1 - \sum_{i=1}^d u_i \ge 0$. We use the following representation of integration on the unit simplex Δ_d: For any measurable function $h : \Delta_d \to [0, \infty)$, we have $\int_{\Delta_d} h(u)\, \mathrm{d}u = \frac{1}{d!}\mathbb{E}h(B_1, \ldots, B_d)$, where $B_i = E_i / \sum_{j=0}^d E_j$ with independent, standard exponentially distributed random variables $E_0, \ldots, E_d$.

$$\begin{aligned}\frac{\nu(\Delta)}{\lambda_d(\Delta)} &= \frac{1}{\lambda_d(\Delta_d)}\int_{\Delta_d} g\big(x(u)\big)^{-r}\, \mathrm{d}u = \mathbb{E}g\bigg(\sum_{j=0}^d B_j x_j\bigg)^{-r} \\ &\ge \mathbb{E}\bigg(\sum_{j=0}^d B_j g(x_j)\bigg)^{-r} = \mathbb{E}\bigg(B_0 g_0 + (1 - B_0)\sum_{i=1}^d \tilde{B}_i g(x_i)\bigg)^{-r},\end{aligned}$$

where $\tilde{B}_i := E_i / \sum_{j=1}^d E_j$ for $1 \le i \le d$. Following Cule and Dümbgen (2008), it is known that B_0 and $\{\tilde{B}_i\}_{i=1}^d$ are independent, and $\mathbb{E}[\tilde{B}_i] = 1/d$. Hence it follows from Jensen's inequality that

$$\begin{aligned}\frac{\nu(\Delta)}{\lambda_d(\Delta)} &\ge \mathbb{E}\left[\mathbb{E}\bigg(B_0 g_0 + (1 - B_0)\sum_{i=1}^d \tilde{B}_i g(x_i)\bigg)^{-r}\bigg| B_0\right] \\ &\ge \mathbb{E}\bigg(B_0 g_0 + (1 - B_0)\frac{1}{d}\sum_{i=1}^d g(x_i)\bigg)^{-r} \\ &= \mathbb{E}\left(B_0 g_0 + (1 - B_0)\bar{g}\right)^{-r} \\ &= \int_0^1 d(1-t)^{d-1}\big(t g_0 + (1-t)\bar{g}\big)^{-r}\, \mathrm{d}t \\ &= \bar{g}^{-r}\int_0^1 d(1-t)^{d-1}\left(1 - st\left((-1/s)\left(\frac{g_0}{\bar{g}} - 1\right)\right)\right)\, \mathrm{d}t \\ &= \bar{g}^{-r} J_{d,s}\bigg(-\frac{1}{s}\left(\frac{g_0}{\bar{g}} - 1\right)\bigg),\end{aligned}$$

where

$$J_{d,s}(y) = \int_0^1 d(1-t)^{d-1}(1-syt)^{1/s}\ \mathrm{d}t.$$

We claim that

$$J_{d,s}(y) \geq \int_0^1 d(1-t)^{d-1}(1-t)^y \mathrm{d}t = \frac{d}{d+y},$$

holds for $s<0, y>0$. To see this, we write $(1-syt)^{1/s} = (1+yt/r)^{-(r/y)y}$. Then we only have to show $(1+yt/r)^{-r/y} \geq (1-t)$ for $0 \leq t \leq 1$, or equivalently $(1+bt) \leq (1-t)^{-b}$ where we let $b = y/r$. Let $g(t) := (1-t)^{-b} - (1+bt)$. It is easy to verify that $g(0) = 0$, $g'(t) = b(1-t)^{-b-1} - b$ with $g'(0) = 0$, and $g''(t) = b(b+1)(1-t)^{-b-2} \geq 0$. Integrating g'' twice yields $g(t) \geq 0$, and hence we have verified the claim. Now we proceed with the calculation

$$\frac{\nu(\Delta)}{\lambda_d(\Delta)} \geq \bar{g}^{-r} J_{d,s}\left(-\frac{1}{s}\left(\frac{g_0}{\bar{g}}-1\right)\right) \geq \bar{g}^{-r}\frac{d}{d-\frac{1}{s}\left(\frac{g_0}{\bar{g}}-1\right)}.$$

Solving for g_0 and replacing $-1/s = r$ proves the desired inequality. □

Proof of Lemma B.1. For fixed $j \in \{0,\ldots,d\}$, note $|\det(x_i - x_j) : i \neq j| = |\det X|$ where $X = \begin{pmatrix} x_0 & \ldots & x_d \\ 1 & \ldots & 1 \end{pmatrix}$. Also for each $y \in \mathbb{R}^d$, since $\Delta = \mathrm{conv}(\{x_0,\ldots,x_d\})$ is non-void, y must be in the affine hull of Δ and hence we can write $y = \sum_{i=0}^d \lambda_i x_i$ with $\sum_{i=0}^d \lambda_i = 1$ (not necessary non-negative), i.e. $\lambda = X^{-1}\binom{y}{1}$. Let $\Delta_j(y) := \mathrm{conv}(\{x_i : i \neq j\} \cup \{y\})$. Then

$$\begin{aligned}\lambda_d(\Delta_j(y)) &= \frac{1}{d!}\left|\det\begin{pmatrix} x_0 & \ldots & x_{j-1} & y & x_{j+1} & \ldots & x_d \\ 1 & \ldots & 1 & 1 & 1 & \ldots & 1\end{pmatrix}\right| \\ &= \frac{1}{d!}\left|\lambda_j\right|\left|\det X\right| = \left|\lambda_j\right|\lambda_d(\Delta).\end{aligned}$$

Hence,

$$\begin{aligned}\max_{0\leq j\leq d}\lambda_d(\Delta_j(y)) &\geq \lambda_d(\Delta)\max_j\left|\lambda_j\right| = \lambda_d(\Delta)\|X^{-1}\begin{pmatrix} y \\ 1\end{pmatrix}\|_\infty \\ &\geq \lambda_d(\Delta)(d+1)^{-1/2}\|X^{-1}\begin{pmatrix} y \\ 1\end{pmatrix}\| \\ &\geq \lambda_d(\Delta)(d+1)^{-1/2}\sigma_{\max}(X)^{-1}(1+\|y\|^2)^{1/2} = C(1+\|y\|^2)^{1/2}.\end{aligned}$$

Now the conclusion follows from Lemma E.1 by noting

$$f(y) \leq \bar{g}_j^{-r}\left(1-\frac{d}{r}+\frac{d}{r}\frac{\lambda_d(\Delta_j(y))\bar{g}_j^{-r}}{\nu(\Delta_j(y))}\right)^{-r} \leq f_{\max}\left(1-\frac{d}{r}+\frac{d}{r}f_{\min}C(1+\|y\|^2)^{1/2}\right)^{-r},$$

since $\bar{g}_j^{-r} = \left(\frac{1}{d}\sum_{i\neq j} f^s(x_i)\right)^{1/s}$ and hence $f_{\min} \leq \bar{g}_j^{-r} \leq f_{\max}$, and the index j is chosen such that $\lambda_d(\Delta_j(y))$ is maximized. □

PROOF OF LEMMA B.2. The key point that for any point $x \in B(y, \delta_t)$

$$B(ty, \delta_t) \subset (1-t)B(0, \delta) + tx$$

can be shown in the same way as in the proof of Lemma 4.2 Schuhmacher, Hüsler and Dümbgen (2011). Namely, pick any $w \in B(ty, \delta_t)$, let $v := (1-t)^{-1}(w - tx)$, then since

$$\|v\| = (1-t)^{-1}\|w-tx\| = (1-t)^{-1}\|w-ty+t(y-x)\| \leq (1-t)^{-1}(\delta_t+t\delta_t) = \delta,$$

and hence $v \in B(0, \delta)$. This implies that $w = (1-t)v + tx \in (1-t)B(0,\delta) + tx$, as desired. By s-concavity of f, we have

$$\begin{aligned} f(w) &\geq \left((1-t)f(v)^s + tf(x)^s\right)^{1/s} \\ &\geq \left((1-t)J_0^s + tf(x)^s\right)^{1/s} \\ &= J_0\left(1 - t + t\left(\frac{f(x)}{J_0}\right)^s\right)^{1/s}. \end{aligned}$$

Averaging over $w \in B(ty, \delta_t)$ yields

$$\frac{\nu(B(ty,\delta_t))}{\lambda_d(B(ty,\delta_t))} \geq J_0\left(1 - t + t\left(\frac{f(x)}{J_0}\right)^s\right)^{1/s}.$$

Solving for $f(x)$ completes the proof. □

E.2. Auxiliary convex analysis.

LEMMA E.2 (Lemma 4.3, Dümbgen, Samworth and Schuhmacher (2011)). *For any $\varphi(\cdot) \in \mathcal{G}$ with non-empty domain, and $\epsilon > 0$, define*

$$\varphi^{(\epsilon)}(x) := \sup_{(v,c)}(v^T x + c)$$

where the supremum is taken over all pairs of $(v, c) \in \mathbb{R}^d \times \mathbb{R}$ such that

1. $\|v\| \leq \frac{1}{\epsilon}$;
2. $\varphi(y) \geq v^T y + c$ *holds for all* $y \in \mathbb{R}^d$.

Then $\varphi^{(\epsilon)} \in \mathcal{G}$ with Lipschitz constant $\frac{1}{\epsilon}$. Furthermore,

$$\varphi^{(\epsilon)} \nearrow \varphi, \text{ as } \epsilon \searrow 0,$$

where the convergence is pointwise for all $x \in \mathbb{R}^d$.

LEMMA E.3 (Lemma 2.13, Dümbgen, Samworth and Schuhmacher (2011)). *Given $Q \in \mathcal{Q}_0$, a point $x \in \mathbb{R}^d$ is an interior point of csupp(Q) if and only if*

$$h(Q,x) \equiv \sup\{Q(C) : C \subset \mathbb{R}^d \text{ closed and convex}, x \notin \mathrm{int}(C)\} < 1.$$

Moreover, if $\{Q_n\} \subset \mathcal{Q}$ converges weakly to Q, then

$$\limsup_{n\to\infty} h(Q_n,x) \le h(Q,x)$$

holds for all $x \in \mathbb{R}^d$.

LEMMA E.4. *If $g \in \mathcal{G}$, then there exists $a, b > 0$ such that for all $x \in \mathbb{R}^d$, $g(x) \ge a\|x\| - b$.*

PROOF. The proof is essentially the same as for Lemma 1, Cule and Samworth (2010), so we shall omit it. □

Consider the class of functions

$$\mathcal{G}_M := \left\{ g \in \mathcal{G} : \int g^\beta \, \mathrm{d}x \le M \right\}.$$

LEMMA E.5. *For a given $g \in \mathcal{G}_M$, denote $D_r := D(g,r) := \{g \le r\}$ to be the level set of $g(\cdot)$ at level r, and $\epsilon := \inf g$. Then for $r > \epsilon$, we have*

$$\lambda(D_r) \le \frac{M(-s)(r-\epsilon)^d}{(s+1)\int_0^{r-\epsilon} v^d (v+\epsilon)^{1/s} \, \mathrm{d}v},$$

where $\beta = 1 + 1/s$, and $-1 < s < 0$.

PROOF. For $u \in [\epsilon, r]$, by convexity of $g(\cdot)$, we have

$$\lambda(D_u) \ge \left(\frac{u-\epsilon}{r-\epsilon}\right)^d \lambda(D_r).$$

This can be seen as follows: Consider the epigraph Γ_g of $g(\cdot)$, where $\Gamma_g = \{(t,x) \in \mathbb{R}^d \times \mathbb{R} : x \ge g(t)\}$. Let $x_0 \in \mathbb{R}^d$ be a minimizer of g. Consider the convex set $C_r = \mathrm{conv}\big(\Gamma_g \cap \{g = r\}, (x_0, \epsilon)\big) \subset \Gamma_g \cap \{g \le r\}$. where the inclusion follows from the convexity of Γ_g as a subset of $\mathbb{R}^{d+1}$. The claimed inequality follows from

$$\lambda_d(D_u) = \lambda_d\big(\pi_d(\Gamma_g \cap \{g = u\})\big) \ge \lambda_d\big(\pi_d(C_r \cap \{g = u\})\big) = \left(\frac{u-\epsilon}{r-\epsilon}\right)^d \lambda_d(D_r),$$

where $\pi_d : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is the natural projection onto the first component. Now we do the calculation as follows:

$$
\begin{aligned}
M &\geq \int_{D_r} \left(g(x)^{1/s+1} - r^{1/s+1} \right) \mathrm{d}x \\
&= -\left(\frac{1}{s}+1\right) \int_{D_r} \left(\int_{\epsilon}^{r} \mathbf{1}(u \geq g(x)) u^{1/s} \, \mathrm{d}u \right) \mathrm{d}x \\
&= -\left(\frac{1}{s}+1\right) \int_{\epsilon}^{r} u^{1/s} \, \mathrm{d}u \int_{D_r} \mathbf{1}(u \geq g(x)) \, \mathrm{d}x \\
&= -\left(\frac{1}{s}+1\right) \int_{\epsilon}^{r} \lambda(D_u) u^{1/s} \, \mathrm{d}u \\
&\geq -\left(\frac{1}{s}+1\right) \int_{\epsilon}^{r} \left(\frac{u-\epsilon}{r-\epsilon}\right)^d \lambda(D_r) u^{1/s} \, \mathrm{d}u \\
&= \lambda(D_r) \cdot \frac{(s+1) \int_{\epsilon}^{r} (u-\epsilon)^d u^{1/s} \, \mathrm{d}u}{(-s)(r-\epsilon)^d}.
\end{aligned}
$$

By a change of variable in the integral we get the desired inequality. □

Lemma E.6. *Let G be a convex set in $\mathbb{R}^d$ with non-empty interior, and a sequence $\{y_n\}_{n\in\mathbb{N}}$ with $\|y_n\| \to \infty$ as $n \to \infty$. Then there exists $\{x_1, \ldots, x_d\} \subset G$ such that*

$$\lambda_d\big(\mathrm{conv}\,(x_1, \ldots, x_d, y_{n(k)})\big) \to \infty,$$

as $k \to \infty$ where $\{y_{n(k)}\}_{k\in\mathbb{N}}$ is a suitable subsequence of $\{y_n\}_{n\in\mathbb{N}}$.

Proof. Without loss of generality we assume $0 \in \mathrm{int}(\mathrm{dom}(G))$, and we first choose a convergent subsequence $\{y_{n(k)}\}_{k\in\mathbb{N}}$ from $\{y_n/\|y_n\|\}_{n\in\mathbb{N}}$. Now if we let $a := \lim_{k\to\infty} y_{n(k)}/\|y_{n(k)}\|$, then $\|a\| = 1$. Since G has non-empty interior, $\{a^T x = 0\} \cap G$ has non-empty relative interior. Thus we can choose $x_1, \ldots, x_d \subset \{a^T x = 0\} \cap G$ such that $\lambda_{d-1}(K) \equiv \lambda_{d-1}\big(\mathrm{conv}\,(x_1, \ldots, x_d)\big) > 0$. Note that

$$\mathrm{dist}\,\big(y_{n(k)}, \mathrm{aff}(K)\big) = \mathrm{dist}\,\big(y_{n(k)}, \{a^T x = 0\}\big) = \langle y_{n(k)}, a\rangle = \|y_{n(k)}\| \langle y_{n(k)}/\|y_{n(k)}\|, a\rangle \to \infty,$$

as $k \to \infty$. Since

$$\lambda_d\big(\mathrm{conv}\,(x_1, \ldots, x_d, y_{n(k)})\big) = \lambda_d\big(\mathrm{conv}\,(K, y_{n(k)})\big) = c\lambda_{d-1}(K) \cdot \mathrm{dist}\,\big(y_{n(k)}, \mathrm{aff}(K)\big),$$

for some constant $c = c(d) > 0$, the proof is complete as we let $k \to \infty$. □

LEMMA E.7 (Lemma 4.2, Dümbgen, Samworth and Schuhmacher (2011)). *Let $\bar{g}$ and $\{g_n\}_{n\in\mathbb{N}}$ be functions in $\mathcal{G}$ such that $g_n \geq \bar{g}$, for all $n \in \mathbb{N}$. Suppose the set $C := \{x \in \mathbb{R}^d : \limsup_{n\to\infty} g_n(x) < \infty\}$ is non-empty. Then there exist a subsequence $\{g_{n(k)}\}_{k\in\mathbb{N}}$ of $\{g_n\}_{n\in\mathbb{N}}$, and a function $g \in \mathcal{G}$ such that $C \subset \mathrm{dom}(g)$ and*

$$\begin{aligned} &\lim_{k\to\infty, x\to y} g_{n(k)}(x) = g(y), \quad \textit{for all } y \in \mathrm{int}(\mathrm{dom}(g)), \\ &\liminf_{k\to\infty, x\to y} g_{n(k)}(x) \geq g(y), \quad \textit{for all } y \in \mathbb{R}^d. \end{aligned} \tag{E.1}$$

LEMMA E.8. *Let $\{g_n\}$ be a sequence of non-negative convex functions satisfying the following conditions:*

(A1). There exists a convex set G with non-empty interior such that for all $x_0 \in \mathrm{int}(G)$, we have $\sup_{n\in\mathbb{N}} g_n(x_0) < \infty$.
(A2). There exists some $M > 0$ such that $\sup_{n\in\mathbb{N}} \int \big(g_n(x)\big)^\beta \,\mathrm{d}x \leq M < \infty$.

Then there exists $a, b > 0$ such that for all $x \in \mathbb{R}^d$ and $k \in \mathbb{N}$

$$g_{n(k)}(x) \geq a\|x\| - b,$$

where $\{g_{n(k)}\}_{k\in\mathbb{N}}$ is a suitable subsequence of $\{g_n\}_{n\in\mathbb{N}}$.

PROOF. Without loss of generality we may assume G is contained in all $\mathrm{int}(\mathrm{dom}(g_n))$. We first note (A1)-(A2) implies that $\{\widehat{x}_n \in \mathrm{Arg\,min}_{x\in\mathbb{R}^d} g_n(x)\}_{n=1}^\infty$ is a bounded sequence, i.e.

$$\sup_{n\in\mathbb{N}} \|\widehat{x}_n\| < \infty, \tag{E.2}$$

Suppose not, then without loss of generality we may assume $\|\widehat{x}_n\| \to \infty$ as $n \to \infty$. By Lemma E.6, we can choose $\{x_1, \ldots, x_d\} \subset G$ such that $\lambda_d\big(\mathrm{conv}\,(x_1, \ldots, x_d, \widehat{x}_{n(k)})\big) \to \infty$, as $k \to \infty$ for some subsequence $\{\widehat{x}_{n(k)}\} \subset \{\widehat{x}_n\}$. For simplicity of notation we think of $\{\widehat{x}_n\}$ as such an appropriate subsequence. Denote $\epsilon_n := \inf_{x\in\mathbb{R}^d} g_n(x)$, and $M_2 := \sup_{n\in\mathbb{N}} \epsilon_n \leq \sup_{n\in\mathbb{N}} g_n(x_0) < \infty$ by (A1). Again by (A1) and convexity we may assume that

$$\sup_{x\in\mathrm{conv}(x_1,\ldots,x_d,\widehat{x}_n)} g_n(x) \leq M_1,$$

holds for some $M_1 > 0$ and all $n \in \mathbb{N}$. This implies that

$$\int g_n^\beta(x) \,\mathrm{d}x \geq M_1^\beta \lambda_d\big(\mathrm{conv}\,(x_1, \ldots, x_d, \hat{x}_n)\big) \to \infty,$$

as $n \to \infty$, which gives a contradiction to (A2). This shows (E.2).

Now we define $\underline{g}(\cdot)$ be the convex hull of $\tilde{g}(x) := \inf_{n\in\mathbb{N}} g_n(x)$, then $\underline{g} \le g_n$ holds for all $n \in \mathbb{N}$. We claim that $\underline{g}(x) \to \infty$ as $\|x\| \to \infty$. By Lemma E.5, for fixed $\eta > 1$, we have

$$\begin{aligned}\lambda_d\big(D(g_n, \eta M_2)\big) &\le \frac{M(-s)(\eta M_2 - \epsilon_n)^d}{(s+1)\int_0^{\eta M_2 - \epsilon_n} v^d (v+\epsilon_n)^{1/s}\ \mathrm{d}v}\\ &\le \frac{M(-s)(\eta M_2)^d}{(s+1)\int_0^{(\eta-1)M_2} v^d (v+M_2)^{1/s}\ \mathrm{d}v} < \infty,\end{aligned}$$

where $D(g_n, \eta M_2) := \{g_n \le \eta M_2\}$. Hence

$$\sup_{n\in\mathbb{N}} \lambda_d\big(D(g_n, \eta M_2)\big) < \infty. \tag{E.3}$$

holds for every $\eta > 1$. Now combining (E.2) and (E.3), we claim that, for fixed η large enough, it is possible to find $R = R(\eta) > 0$ such that

$$g_n(x) \ge \eta M_2, \tag{E.4}$$

holds for all $x \ge R(\eta)$ and $n \in \mathbb{N}$. If this is not true, then for all $k \in \mathbb{N}$, we can find $n(k) \in \mathbb{N}$ and $\bar{x}_k \in \mathbb{R}^d$ with $\|\bar{x}_k\| \ge k$ such that $g_{n(k)}(\bar{x}_k) \le \eta M_2$. We consider two cases to derive a contradiction.

[Case 1.] If for some $n_0 \in \mathbb{N}$ there exists infinitely many $k \in \mathbb{N}$ with $n(k) = n_0$, then we may assume without loss of generality that we can find some a sequence $\{\bar{x}_k\}_{k\in\mathbb{N}}$ with $\|\bar{x}_k\| \to \infty$ as $k \to \infty$, and $g_{n_0}(\bar{x}_k) \le \eta M_2$. Since the support g_{n_0} has non-empty interior, by Lemma E.6, we can find $x_1, \ldots, x_d \in \mathrm{supp}(g_{n_0})$ such that $\lambda_d\big(\mathrm{conv}(x_1, \ldots, x_d, \bar{x}_{k(j)})\big) \to \infty$ as $j \to \infty$ holds for some subsequence $\{\bar{x}_{k(j)}\}_{j\in\mathbb{N}}$ of $\{\bar{x}_k\}_{k\in\mathbb{N}}$. Let $\bar{M} := \max_{1\le i\le d} g_{n_0}(x_i)$, then we find $\lambda_d\big(D(g_{n_0}, \bar{M} \vee \eta M_2)\big) = \infty$. This contradicts with (E.3).

[Case 2.] If $\#\{k \in \mathbb{N} : n = n(k)\} < \infty$ for all $n \in \mathbb{N}$, then without loss of generality we may assume that for all $k \in \mathbb{N}$, we can find $\bar{x}_k \in \mathbb{R}^d$ with $\|\bar{x}_k\| \ge k$ such that $g_k(x_k) \le \eta M_2$. Recall by assumption (A1) convex set G has non-empty interior, and is contained in the support of g_n for all $n \in \mathbb{N}$. Again by Lemma E.6, we may take $x_1, \ldots, x_d \in C$ such that $\lambda_d\big(\mathrm{conv}(x_1, \ldots, x_d, \bar{x}_{k(j)})\big) \to \infty$ as $j \to \infty$ holds for some subsequence $\{\bar{x}_{k(j)}\}_{j\in\mathbb{N}}$ of $\{\bar{x}_k\}_{k\in\mathbb{N}}$. In view of (A1), we conclude by convexity that $\bar{M} := \max_{1\le i\le d} \sup_{j\in\mathbb{N}} g_{k(j)}(x_i) < \infty$. This implies

$$\lambda_d\big(D(g_{n_{k(j)}}, \bar{M} \vee \eta M_2)\big) \ge \lambda_d\big(\mathrm{conv}(x_1, \ldots, x_d, \bar{x}_{k(j)})\big) \to \infty, \quad j \to \infty,$$

which gives a contradiction.

Combining these two cases we have proved (E.4). This implies that $\tilde{g}(x) \to \infty$ as $\|x\| \to \infty$, whence verifying the claim that $\underline{g}(x) \to \infty$ as $\|x\| \to \infty$. Hence in view of Lemma E.4, we find that there exists $a, b > 0$ such that $g_n(x) \geq a\|x\| - b$ holds for all $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$. □

LEMMA E.9. *Assume $x_0, \ldots, x_d \in \mathbb{R}^d$ are in general position. If $g(\cdot)$ is a non-negative function with $\Delta \equiv \mathrm{conv}(x_0, \ldots, x_d) \subset \mathrm{dom}(g)$, and $g(x_0) = 0$. Then for $r \geq d$, we have $\int_\Delta \big(g(x)\big)^{-r}\, \mathrm{d}x = \infty$.*

PROOF. We may assume without loss of generality that $x_0 = 0, x_i = \mathbf{e}_i \in \mathbb{R}^d$, where $\mathbf{e}_i$ is the unit directional vector with 1 in its i-th coordinate and 0 otherwise. Then $\Delta = \Delta_0 := \{x \in \mathbb{R}^d : \sum_{i=1}^d x_i \leq 1, x_i \geq 0, \forall i = 1, \ldots, d\}$. Denote $a_i = g(x_i) \geq 0$. We may assume there is at least one $a_i \neq 0$. Then by convexity of g we find $g(x) \leq \sum_{i=1}^d a_i x_i$ for all $x \in \Delta_0$. This gives

$$\begin{aligned}\int_{\Delta_0} \big(g(x)\big)^{-r}\, \mathrm{d}x &\geq \int_{\Delta_0} \Big(\sum_{i=1}^d a_i x_i\Big)^{-r}\, \mathrm{d}x \geq \int_{\Delta_0} \frac{1}{(\max_{i=1,\ldots,d} a_i)^r \|x\|_1^r}\, \mathrm{d}x \\ &\geq \frac{1}{(\max_{i=1,\ldots,d} a_i)^r d^{r/2}} \int_{C_0} \frac{1}{\|x\|_2^r}\, \mathrm{d}x = \infty,\end{aligned}$$

where $C_0 := \{\|x\|_2 \leq \frac{1}{\sqrt{d}}\} \cap \{x_i \geq 0, i = 1, \ldots, d\}$. Note we used the fact that $\|x\|_1 \leq \sqrt{d}\|x\|_2$. □

LEMMA E.10 (Theorem 1.11, Bhattacharya and Ranga Rao (1976)). *Let $f_n \to_d f$, and $\mathcal{D}$ be the class of all Borel measurable, convex subsets in $\mathbb{R}^d$. Then $\lim_{n\to\infty} \sup_{D \in \mathcal{D}} \big|\int_D (f_n - f)\big| = 0$.*

REFERENCES

BALABDAOUI, F., RUFIBACH, K. and WELLNER, J. A. (2009). Limit distribution theory for maximum likelihood estimation of a log-concave density. *Ann. Statist.* **37** 1299–1331. MR2509075 (2010h:62290)

BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. John Wiley & Sons, New York-London-Sydney Wiley Series in Probability and Mathematical Statistics. MR0436272 (55 ##9219)

CULE, M. L. and DÜMBGEN, L. (2008). On an auxiliary function for log-density estimation. *arXiv preprint arXiv:0807.4719*.

CULE, M. and SAMWORTH, R. (2010). Theoretical properties of the log-concave maximum likelihood estimator of a multidimensional density. *Electron. J. Stat.* **4** 254–270. MR2645484 (2012f:62071)

DUDLEY, R. M. (2002). *Real Analysis and Probability. Cambridge Studies in Advanced Mathematics* **74**. Cambridge University Press, Cambridge Revised reprint of the 1989 original. MR1932358(2003h:60001)

Dümbgen, L. and Rufibach, K. (2009). Maximum likelihood estimation of a log-concave density and its distribution function: basic properties and uniform consistency. *Bernoulli* **15** 40–68. MR2546798 (2011b:62096)

Dümbgen, L., Samworth, R. and Schuhmacher, D. (2011). Approximation by log-concave distributions, with applications to regression. *Ann. Statist.* **39** 702–730. MR2816336 (2012e:62039)

Groeneboom, P., Jongbloed, G. and Wellner, J. A. (2001). Estimation of a convex function: characterizations and asymptotic theory. *Ann. Statist.* **29** 1653–1698. MR1891742 (2003a:62047)

Kim, A. K. and Samworth, R. J. (2015). Global rates of convergence in log-concave density estimation. *arXiv preprint arXiv:1404.2298v2*.

Koenker, R. and Mizera, I. (2010). Quasi-concave density estimation. *Ann. Statist.* **38** 2998–3027. MR2722462 (2011j:62108)

Rockafellar, R. T. (1997). *Convex Analysis. Princeton Landmarks in Mathematics*. Princeton University Press, Princeton, NJ Reprint of the 1970 original, Princeton Paperbacks. MR1451876 (97m:49001)

Schuhmacher, D., Hüsler, A. and Dümbgen, L. (2011). Multivariate log-concave distributions as a nearly parametric model. *Stat. Risk Model.* **28** 277–295. MR2838319 (2012h:62044)

Seijo, E. and Sen, B. (2011). Nonparametric least squares estimation of a multivariate convex regression function. *Ann. Statist.* **39** 1633–1657. MR2850215 (2012j:62119)

Department of Statistics, Box 354322
University of Washington
Seattle, WA 98195-4322
E-mail: royhan@uw.edu

Department of Statistics, Box 354322
University of Washington
Seattle, WA 98195-4322
E-mail: jaw@stat.washington.edu
URL: http://www.stat.washington.edu/jaw/