SUPPLEMENTARY MATERIALS: An Order-Theoretic Perspective on Modes and Maximum A Posteriori Estimation in Bayesian Inverse Problems*

Hefin Lambley[†] and T. J. Sullivan[‡]

SM1. Technical supporting results.

SM1.1. Radial cumulative distribution functions.

Lemma SM1.1 (Properties of RCDFs). *Let X be a metric space and let $\mu \in \mathscr{P}(X)$.*

(a) *For each $r > 0$, $x \mapsto \mu(B_r(x))$ is upper semicontinuous.*

(b) *For each $x \in X$, $r \mapsto \mu(B_r(x))$ is monotonically increasing, is continuous from the right, has limits from the left, and is upper semicontinuous. Furthermore, $r \mapsto \mu(B_r(x))$ is differentiable λ^1-a.e.*

Proof. For (a), fix $r > 0$ and let $(x_n)_{n\in\mathbb{N}}$ converge in X to some $x \in X$. Then

$$\begin{aligned}
\mu(B_r(x)) &= \lim_{n\to\infty} \mu(B_{r+d(x,x_n)}(x)) && \text{since } B_r(x) = \bigcap_{n\in\mathbb{N}} B_{r+d(x,x_n)}(x) \\
&= \limsup_{n\to\infty} \mu(B_{r+d(x,x_n)}(x)) \\
&\geqslant \limsup_{n\to\infty} \mu(B_r(x_n)) && \text{since } B_{r+d(x,x_n)}(x) \supseteq B_r(x_n).
\end{aligned}$$

For (b), monotonicity follows from the monotonicity of probability. To examine continuity, fix $x \in X$ and let $(r_n)_{n\in\mathbb{N}}$ be a convergent sequence in $[0,\infty)$ with limit $r \geqslant 0$. If $(r_n)_{n\in\mathbb{N}}$ is decreasing, then $\bigcap_{n\in\mathbb{N}} B_{r_n}(x) = B_r(x)$ and so the continuity of probability along monotone sequences implies that $\mu(B_{r_n}(x)) \searrow \mu(B_r(x))$, which establishes continuity from the right. If $(r_n)_{n\in\mathbb{N}}$ is increasing, then $\bigcup_{n\in\mathbb{N}} B_{r_n}(x) = \mathring{B}_r(x)$, and continuity of probability implies that $\mu(B_{r_n}(x)) \nearrow \mu(\mathring{B}_r(x)) \leqslant \mu(B_r(x))$, and this establishes existence of a limit from the left. Now let $r_n \to r$, and make no assumption that this convergence is monotone. By the above,

$$\limsup_{n\to\infty} \mu(B_{r_n}(x)) \in \{\mu(\mathring{B}_r(x)), \mu(B_r(x))\},$$

i.e. the lim sup is at most $\mu(B_r(x))$, which establishes upper semicontinuity. Finally, a.e.-differentiability of $r \mapsto \mu(B_r(x))$ follows from monotonicity and Lebesgue's theorem on differentiability of monotone functions. ∎

*Supplementary material for JUQ MS#M154243.
https://doi.org/10.1137/22M154243X

[†]Mathematics Institute, University of Warwick, Coventry, CV4 7AL, United Kingdom (hefin.lambley@warwick.ac.uk).

[‡]Mathematics Institute and School of Engineering, University of Warwick, Coventry, CV4 7AL, United Kingdom, and Alan Turing Institute, 96 Euston Road, London, NW1 2DB, United Kingdom (t.j.sullivan@warwick.ac.uk).

Corollary SM1.2. *Let X be a separable metric space, let $\mu \in \mathscr{P}(X)$, and fix $r > 0$. Then*

$$M_r := \sup_{x \in X} \mu(B_r(x)) > 0$$

and every sequence $(x_n)_{n\in\mathbb{N}}$ such that $\mu(B_r(x_n)) \to M_r$ as $n \to \infty$ is bounded.

Proof. The separability of X implies that $\operatorname{supp}(\mu) \neq \varnothing$ [SM2, Theorem 12.14], and so there must exist at least one $x \in X$ with $\mu(B_r(x)) > 0$. Hence, $M_r > 0$.

Now let $(x_n)_{n\in\mathbb{N}}$ be any sequence such that $\mu(B_r(x_n)) \to M_r$ as $n \to \infty$. Then there must exist $N \in \mathbb{N}$ such that

$$n \geqslant N \implies \mu(B_r(x_n)) \geqslant {}^{M_r}\!/_2 > 0,$$

i.e. x_n eventually lies in $\{x \in X \mid \mu(B_r(x_n)) \geqslant {}^{M_r}\!/_2\}$, which is a bounded set by Lemma 4.2(c), and so $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence. ■

Definition SM1.3. *Let X be a metric space. A probability measure $\mu \in \mathscr{P}(X)$ will be called* spherically non-atomic *if every metric sphere has zero μ-mass, i.e., for all $r \geqslant 0$ and all $x \in X$, $\mu(\mathring{B}_r(x)) = \mu(B_r(x))$.*

Corollary SM1.4 (RCDFs of spherically non-atomic measures). *Let X be a metric space and assume that $\mu \in \mathscr{P}(X)$ is spherically non-atomic.*

(a) *For each $r > 0$, $x \mapsto \mu(B_r(x))$ is continuous.*
(b) *For each $x \in X$, $r \mapsto \mu(B_r(x))$ is monotonically increasing and continuous.*
(c) *$(r, x) \mapsto \mu(B_r(x))$ is continuous.*
(d) *For each $x \in X$, $\mu(\{x\}) = \mu(B_0(x)) = \lim_{r\to 0} \mu(B_r(x)) = 0$.*

Proof. Easy modification of the proof of Lemma SM1.1 shows that

- for each $r > 0$, $x \mapsto \mu(\mathring{B}_r(x))$ is lower semicontinuous;
- for each $x \in X$, $r \mapsto \mu(\mathring{B}_r(x))$ is monotonically increasing, is continuous from the left, has limits from the right, and is lower semicontinuous.

For a spherically non-atomic measure μ, each occurrence of $\mu(\mathring{B}_r(x))$ can be replaced with $\mu(B_r(x))$, and this together with the original statement of Lemma SM1.1 proves parts (a) and (b).

An easy modification of the classical theorem of [SM9] on the joint continuity of separately continuous functions (see e.g. [SM7, Theorem 3.1]) establishes (c).

Finally, (d) follows from $\mu(\{x\}) = \mu(B_0(x)) = \mu(\mathring{B}_0(x)) = \mu(\varnothing) = 0$; the claim regarding the limit follows from the continuity of $r \mapsto \mu(B_r(x))$, as proven in (b). ■

SM1.2. Radius-r modes in sequence spaces. Given $p \in [1, \infty)$ and $\alpha \in \mathbb{R}^{\mathbb{N}}_{>0}$, we define the corresponding weighted ℓ^p space and its norm by

$$\ell^p_\alpha := \left\{ x = (x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \,\middle|\, \|x\|_{\ell^p_\alpha} := \left(\sum_{n\in\mathbb{N}} \left| \frac{x_n}{\alpha_n} \right|^p \right)^{1/p} < \infty \right\}.$$

We also equip $\mathbb{R}^{\mathbb{N}}$ and its subspaces with the finite-dimensional projections

$$P_n \colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^n, \quad x = (x_k)_{k\in\mathbb{N}} \mapsto (x_1, \dots, x_n),$$

and denote the ball of radius $r > 0$ centred at $x \in \mathbb{R}^n$ by

$$B_r^n(x) := \left\{ y \in \mathbb{R}^n \;\middle|\; \left(\sum_{k=1}^{n} \frac{|y_k - x_k|^p}{\alpha_k^p} \right)^{1/p} \leqslant r \right\}.$$

Lemma SM1.5. *Let $X = \ell_\alpha^p$ for some $p \in [1, \infty)$, $\alpha \in \mathbb{R}_{>0}^{\mathbb{N}}$ and let $\mu \in \mathscr{P}(X)$. Define the set function $\mu_n(A) := (\mu \circ P_n^{-1})(P_n A)$ (This function is not necessarily a measure.)*

(a) *For any $n \in \mathbb{N}$, $x \in X$, and $r > 0$, the projection maps satisfy $P_n B_r(x) = B_r^n(P_n x)$.*

(b) *For any $x \in X$ and $r > 0$,*

$$\bigcap_{n \in \mathbb{N}} P_n^{-1}(P_n B_r(x)) = B_r(x).$$

(c) *For any $n \in \mathbb{N}$ and $A \in \mathscr{B}(X)$, the projection maps satisfy $P_{n+1}^{-1}\big(P_{n+1} A\big) \subseteq P_n^{-1}\big(P_n A\big)$.*

(d) *For any $n \in \mathbb{N}$ and $A \in \mathscr{B}(X)$, the set functions μ_n satisfy $\mu_n(A) \geqslant \mu(A)$.*

(e) *For any $x \in X$ and $r > 0$, one has $\lim_{n \to \infty} \mu_n(B_r(x)) = \mu(B_r(x))$.*

Proof. (a) Use that

$$\begin{aligned}
P_n B_r(x) &= \left\{ y \in \mathbb{R}^n \;\middle|\; \text{there exists } \tilde{y} \in X \text{ such that } \|\tilde{y} - x\|_{\ell_\alpha^p} \leqslant r \text{ and } y = P_n \tilde{y} \right\} \\
&= \left\{ y \in \mathbb{R}^n \;\middle|\; \text{there exists } \tilde{y} \in X \text{ such that } \sum_{k=1}^{\infty} \frac{|\tilde{y}_k - x_k|^p}{\alpha_k^p} \leqslant r^p \text{ and } y = P_n \tilde{y} \right\} \\
&= \left\{ y \in \mathbb{R}^n \;\middle|\; \sum_{k=1}^{n} \frac{|y_k - x_k|^p}{\alpha_k^p} \leqslant r^p \right\} = B_r^n(P_n x).
\end{aligned}$$

(b) Observe that, by (a),

$$\begin{aligned}
\bigcap_{n \in \mathbb{N}} P_n^{-1}(P_n B_r(x)) &= \bigcap_{n \in \mathbb{N}} P_n^{-1}(B_r^n(P_n x)) \\
&= \{ y \in X \mid P_n y \in B_r^n(P_n x) \text{ for all } n \in \mathbb{N} \} \\
&= \left\{ y \in X \;\middle|\; \sum_{k=1}^{n} \frac{|y_k - x_k|^p}{\alpha_k^p} \leqslant r^p \text{ for all } n \in \mathbb{N} \right\} = B_r(x).
\end{aligned}$$

(c) This is a straightforward consequence of the definitions.

(d) This follows from the inclusion $A \subseteq P_n^{-1}(P_n A)$ and monotonicity of μ.

(e) As $\big(P_n^{-1}(P_n B_r(x))\big)_{n \in \mathbb{N}}$ is a decreasing sequence of sets, it follows that

$$\lim_{n \to \infty} \mu_n(A) = \lim_{n \to \infty} \mu\big(P_n^{-1}(P_n B_r(x))\big) = \mu\left(\bigcap_{n \in \mathbb{N}} P_n^{-1}(P_n B_r(x)) \right) = \mu(B_r(x))$$

by continuity of measure. ∎

Lemma SM1.6 (Spherical non-atomicity and weak upper semicontinuity in sequence spaces). *Let $X = \ell^p_\alpha$, $1 \leqslant p < \infty$, $\alpha \in \mathbb{R}^{\mathbb{N}}_{>0}$, and let $\mu \in \mathscr{P}(X)$. Suppose that $\mu \circ P_n^{-1} \in \mathscr{P}(\mathbb{R}^n)$ is spherically non-atomic for each $n \in \mathbb{N}$. Then, for each fixed $r > 0$, the map $x \mapsto \mu(B_r(x))$ is weakly upper semicontinuous.*

Proof. Suppose that $x_k \rightharpoonup x^\star$ as $k \to \infty$, and let $\mu_n(A) \coloneqq (\mu \circ P_n^{-1})(P_n A)$. Since $\mu_n(B_r(x^\star)) \searrow \mu(B_r(x^\star))$ (Lemma SM1.5), it follows that, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that, for all $n \geqslant N$, $\mu_n(B_r(x^\star)) - \mu(B_r(x^\star)) < \varepsilon$. Using this and the inequality $\mu_n(B_r(x_k)) \geqslant \mu(B_r(x_k))$ for any $k \in \mathbb{N}$, we obtain

$$\mu(B_r(x_k)) - \mu(B_r(x^\star)) \leqslant \mu_n(B_r(x_k)) - \mu_n(B_r(x^\star)) + \varepsilon. \tag{SM1.1}$$

By hypothesis, $x_k \rightharpoonup x^\star$, so $P_n x_k \to P_n x^\star$ as $k \to \infty$. As $\mu \circ P_n^{-1}$ is assumed to be spherically non-atomic, $x \mapsto (\mu \circ P_n^{-1})(B_r^n(x))$ is continuous (Corollary SM1.4). Hence,

$$\begin{aligned}
\lim_{k\to\infty} (\mu \circ P_n^{-1})(P_n B_r(x_k)) &= \lim_{k\to\infty} (\mu \circ P_n^{-1})(B_r^n(P_n x_k)) && \text{(Lemma SM1.5(a))}\\
&= (\mu \circ P_n^{-1})(B_r^n(P_n x^\star)) && \text{(by continuity)}\\
&= (\mu \circ P_n^{-1})(P_n B_r(x^\star)) && \text{(Lemma SM1.5(a))}.
\end{aligned}$$

Hence, $\lim_{k\to\infty} \mu_n(B_r(x_k)) = \mu_n(B_r(x^\star))$. Taking limits as $k \to \infty$ in (SM1.1) yields that

$$\limsup_{k\to\infty} \mu(B_r(x_k)) - \mu(B_r(x^\star)) \leqslant \lim_{k\to\infty} \mu_n(B_r(x_k)) - \mu_n(B_r(x^\star)) + \varepsilon = \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this shows that $x \mapsto \mu(B_r(x))$ is weakly upper semicontinuous. ∎

We now state an explicit version of Anderson's inequality following the inequalities of [SM5, Lemma 3.6] for Gaussian measures and [SM1, Lemma 6.2] for Besov measures with $p = 1$.

Fix parameters[1] $1 \leqslant p < \infty$, $s \in \mathbb{R}$, and $d \in \mathbb{N}$; the (*sequence space*) *Besov space* X_p^s is defined to be ℓ^p_γ for the weighting sequence $\gamma_k \coloneqq k^{-(s/d+1/2)+1/p}$, and the (*sequence space*) *Besov measure* B_p^s is defined to be the countable product measure $\bigotimes_{k\in\mathbb{N}} \mu_k$, where $\mu_k \in \mathscr{P}(\mathbb{R})$ has Lebesgue density proportional to $\exp(-|x_k/\gamma_k|^p)$. It is known that B_p^s charges X_p^t with full mass when $t = s - (1+\eta)d/p$ and $\eta > 0$ [SM8, Lemma 2].

Lemma SM1.7 (Explicit Anderson inequality for Besov-p priors, $1 \leqslant p \leqslant 2$). *Let $s \in \mathbb{R}$, $d \in \mathbb{N}$, $\eta > 0$ and let $t \coloneqq s - (1+\eta)d/p$. Suppose that $X = X_p^t$ and let $\mu = B_p^s \in \mathscr{P}(X)$ be a sequence-space Besov measure. Then, for any $0 < r < \|x\|_{X_p^t}$ and $x \in X$,*

$$\frac{\mu(B_r(x))}{\mu(B_r(0))} \leqslant \exp\left(-\frac{1}{2}\left(\|x\|_{X_p^t} - r\right)^p\right). \tag{SM1.2}$$

Proof. The space X_p^t can be written as the sequence space ℓ^p_δ with the weighting sequence $\delta_k = k^{-(s/d+1/2)+(2+\eta)/p} > \gamma_k = k^{-(s/d+1/2)+1/p}$. The formula for the unnormalised marginal density of the Besov measure then yields

[1] In the original setting of real analysis, s and d were interpreted as smoothness and spatial dimension respectively, but for us only the ratio s/d is important.

$$
\begin{aligned}
&\frac{\mu_n(B_r(x))}{\mu_n(B_r(0))} \\
&\quad = \frac{\int_{P_n B_r(x)} \exp(-\sum_{i=1}^n |u_i/\gamma_i|^p)\,\mathrm{d}u}{\int_{P_n B_r(0)} \exp(-\sum_{i=1}^n |u_i/\gamma_i|^p)\,\mathrm{d}u} \\
&\quad \leqslant \frac{\sup_{y\in P_n B_r(x)} \exp(-\frac{1}{2}\sum_{i=1}^n |y_i/\delta_i|^p) \int_{P_n B_r(x)} \exp\bigl(-\sum_{i=1}^n |u_i/\gamma_i|^p + \frac{1}{2}\sum_{i=1}^n |u_i/\delta_i|^p\bigr)\,\mathrm{d}u}{\int_{P_n B_r(0)} \exp\bigl(-\sum_{i=1}^n |u_i/\gamma_i|^p + \frac{1}{2}\sum_{i=1}^n |u_i/\delta_i|^p\bigr)\,\mathrm{d}u} \\
&\quad \leqslant \sup_{y\in P_n B_r(x)} \exp\left(-\frac{1}{2}\sum_{i=1}^n |y_i/\delta_i|^p\right),
\end{aligned}
$$

where the ratio of integrals is bounded above by 1 using Anderson's inequality [SM3]. Hence, as $\lim_{n\to\infty} \mu_n(B_r(x)) = \mu(B_r(x))$ (Lemma SM1.5),

$$
\frac{\mu(B_r(x))}{\mu(B_r(0))} = \lim_{n\to\infty} \frac{\mu_n(B_r(x))}{\mu_n(B_r(0))} \leqslant \lim_{n\to\infty} \sup_{y\in P_n B_r(x)} \exp\left(-\frac{1}{2}\sum_{i=1}^n \left|\frac{y_i}{\delta_i}\right|^p\right) = \exp\left(-\frac{1}{2}\Bigl(\|x\|_{X_p^t} - r\Bigr)^p\right),
$$

which establishes (SM1.2). ∎

Theorem SM1.8 (Radius-r modes for product measures on weighted ℓ^p spaces). *Let $X = \ell^p_\alpha$, $1 < p < \infty$, $\alpha \in \mathbb{R}^{\mathbb{N}}_{>0}$. Let $\mu_0 = \bigotimes_{n\in\mathbb{N}} \mu_n \in \mathscr{P}(X)$ with each $\mu_n \ll \lambda^1$ on $\mathbb{R}$. If $\mu \ll \mu_0$, then μ has a radius-r mode for any $r > 0$.*

Proof. As μ_0 is a product of the measures μ_n, which are all absolutely continuous with respect to λ^1, the pushforward measures $\mu_0 \circ P_n^{-1}$ are absolutely continuous with respect to λ^n. As $\mu \ll \mu_0$, it follows that $\mu \circ P_n^{-1} \ll \mu_0 \circ P_n^{-1}$, so the pushforwards of μ are also absolutely continuous with respect to λ^n. Hence, the measure μ has spherically non-atomic pushforwards $\mu \circ P_n^{-1}$, and so the map $x \mapsto \mu(B_r(x))$ is weakly upper semicontinuous for any $r > 0$ (Lemma SM1.6). As any sequence $(x_n)_{n\in\mathbb{N}}$ with $\mu(B_r(x_n)) \nearrow M_r$ is bounded (Corollary SM1.2), there must exist a weakly convergent subsequence $(x_{n_k})_{k\in\mathbb{N}} \rightharpoonup x^\star$ by the reflexivity of ℓ^p_α, $p > 1$. The weak upper semicontinuity of $x \mapsto \mu(B_r(x))$ implies that $x^\star$ is a radius-r mode, because $M_r = \lim_{k\to\infty} \mu(B_r(x_{n_k})) \leqslant \mu(B_r(x^\star))$. ∎

Corollary SM1.9. *Suppose that $X = \ell^p_\alpha$, $1 < p < \infty$, $\alpha \in \mathbb{R}^{\mathbb{N}}_{>0}$. If $\mu \ll \mu_0$ and $\mu_0 = \bigotimes_{n\in\mathbb{N}} \mu_n$ is*

(a) *a Gaussian measure;*
(b) *a Besov measure; or*
(c) *a Cauchy measure,*

then μ has a radius-r mode for any $r > 0$.

SM1.3. Small-ball probabilities for the countable dense antichain. The measure in Theorem 5.11 places variants of the prototype densities $\rho_{k,m}$ at each dyadic rational. While a variety of constructions are possible (see Remark 5.12), we choose to use the dyadic rationals in $[0,1]$ as the dense set for simplicity. The advantage of using the dyadic rationals is that one can exploit the natural "level" structure, writing $D_\ell := \bigl\{(2i-1)2^{-\ell} \bigm| 1 \leqslant i \leqslant 2^{\ell-1}\bigr\}$ for those dyadic rationals which, in their simplest form, can be written as $c2^{-\ell}$. From this level structure, one can explicitly compute the distance between terms and bound the support of the densities $\rho_{k(\ell,i),m(\ell)}$ centred at points in D_ℓ.

As the dyadic rationals are precisely the points in $[0,1]$ with a finite binary expansion, the behaviour of the RCDF $\mu(B_r(x))$ at an arbitrary point $x \in [0,1]$ depends on a quantity which we call the *dyadic irrationality exponent*, and denote $\beta_2(x)$, which can be thought of as a quantitative estimate on the length of runs of 0s or 1s in the binary expansion of x. This quantity is very much analogous to the number-theoretic *irrationality measure* $\varphi(x,n) := \min_{1<p<q, q\leqslant n} |x - p/q|$ and corresponding *irrationality exponent* $\beta(x)$ [SM6]. We choose the notation $\beta(x)$ for the irrationality exponent and not the more usual $\mu(x)$ to avoid confusion with the measure μ.

Definition SM1.10. (a) *The* dyadic irrationality measure *of $x \in [0,1]$ is given by*

$$\varphi_2(x,\ell) := \min_{q \in \biguplus_{i=1}^{\ell} D_i} |x - q|.$$

(b) *The* dyadic irrationality exponent *of $x \in [0,1] \setminus D$ is given by*

$$\beta_2(x) := \inf\left\{ \beta \geqslant 1 \,\middle|\, \liminf_{\ell\to\infty} \frac{\varphi_2(x,\ell)}{2^{-\beta\ell}} > 0 \right\} = \sup\left\{ \beta \geqslant 1 \,\middle|\, \liminf_{\ell\to\infty} \frac{\varphi_2(x,\ell)}{2^{-\beta\ell}} < \infty \right\}.$$

The dyadic irrationality exponent $\beta_2(x)$ is well defined, and indeed

$$\liminf_{\ell\to\infty} \frac{\varphi_2(x,\ell)}{2^{-\beta\ell}} = \begin{cases} 0, & \beta < \beta_2(x), \\ +\infty, & \beta > \beta_2(x). \end{cases} \tag{SM1.3}$$

In general, it is not possible to say anything about the limit in (SM1.3) in the critical case $\beta = \beta_2(x)$; the value could be anything in the range $[0, +\infty]$. Furthermore, as $\varphi(x, 2^\ell) \leqslant \varphi_2(x,\ell)$, it immediately follows that $\beta_2(x) \leqslant \beta(x)$, but the quantities are not equal in general — for example, any irrational number must satisfy $\beta(x) \geqslant 2$ by Dirichlet's approximation theorem, but one can construct irrational numbers with $\beta_2(x) = 1$.

Lemma SM1.11 (Properties of the measure in Theorem 5.11). *Let $\mu \in \mathscr{P}(\mathbb{R})$ be the measure in Theorem 5.11 and fix $\ell \in \mathbb{N}$.*

(a) *Given $q_{\ell,i} \in D_\ell$, the density $\rho_{k(\ell,i),m(\ell)}(\,\cdot - q_{\ell,i})$ is supported within $B_{2^{-4\ell+3}}(q_{\ell,i})$.*

(b) *For distinct $q_{\ell,i}, q_{\ell,i'} \in D_\ell$, the densities $\rho_{k(\ell,i),m(\ell)}(\,\cdot - q_{\ell,i})$ and $\rho_{k(\ell,i'),m(\ell)}(\,\cdot - q_{\ell,i'})$ have disjoint support, and the supports are a distance at least $2^{-\ell} - 2^{-4\ell+4}$ apart.*

(c) *Fix $\delta, r > 0$ and $x \in [0,1]$, and suppose that $\inf_{q\in D_\ell} |x - q| > \delta + r$. Then*

$$\sum_{i=1}^{2^{\ell-1}} \int_{x-r}^{x+r} \rho_{k(\ell,i),m(\ell)}(t - q_{\ell,i}) \,\mathrm{d}t \leqslant 2r\delta^{-1/2}.$$

Proof. (a) By construction, $\rho_{k(\ell,i),m(\ell)}$ has mass $m(\ell) = 2^{-2\ell+1}$. Hence, the truncation radius of this singularity is at most $2m(\ell)^2$ (Proposition 5.10(e)) and therefore the support is contained in a ball of radius $2 \times 2^{-4\ell+2} \leqslant 2^{-4\ell+3}$.

(b) Distinct points in D_ℓ must be a distance at least $2^{-\ell}$ apart, and by (a) the supports of the densities $\rho_{k(\ell,i),m(\ell)}$ and $\rho_{k(\ell,i'),m(\ell)}$ are contained in a ball of radius $2^{-4\ell+3}$. Hence, their supports must be at least a distance $2^{-\ell} - 2 \times 2^{-4\ell+3}$ apart.

(c) By Proposition 5.10(f), outside of $B_\delta(q_{\ell,i})$, the density $\rho_{k(\ell,i),m(\ell)}$ is bounded above by $\delta^{-1/2}$, and the supports of the densities are disjoint, so the upper bound follows immediately. ∎

Lemma SM1.12 (Behaviour of RCDFs in Theorem 5.11).
Let $\mu \in \mathscr{P}(\mathbb{R})$ be the measure in Theorem 5.11.

(a) *Suppose that $q_{\ell,i} \in D_\ell$. Then $\mu(B_r(q_{\ell,i})) \sim \mu_{k(\ell,i),m(\ell)}(B_r(0))$ as $r \to 0$.*

(b) *Suppose that $x \in [0,1] \setminus D$ and that $\beta_2(x) < 4$. Then, for any $\beta \in (\beta_2(x), 4)$, it follows that $\mu(B_r(x)) \in O(r^{\min\{1, 2/\beta\}})$ as $r \to 0$, and in particular $\mu(B_r(x)) \in o(r^{1/2})$.*

(c) *Suppose that $x \in [0,1] \setminus D$. Then, for any $q \in D$,*

$$\liminf_{r \to 0} \frac{\mu(B_r(x))}{\mu(B_r(q))} < 1.$$

(d) *Suppose that $x \in [0,1] \setminus D$ and that $\beta_2(x) > 4$. Then, for any $q \in D$,*

$$\limsup_{r \to 0} \frac{\mu(B_r(x))}{\mu(B_r(q))} > 1,$$

and therefore $x \parallel_0 q$.

Proof. (a) For any $r < 2^{-\ell}$, the ball $B_r(q_{\ell,i})$ does not contain any element of $\biguplus_{i=1}^{\ell} D_i$ except $q_{\ell,i}$. Furthermore, for $m \in \mathbb{N}$, if $r < 2^{-m} - 2^{-4m+3}$, then $B_r(q_{\ell,i})$ does not intersect the support of any singularity centred at $q' \in D_m$. (Lemma SM1.11(a)). As $2^{-m} - 2^{-4m+3} \in \Omega(2^{-m})$ as $m \to \infty$, there exists $M \in \mathbb{N}$ and $c > 0$ such that

$$2^{-m} - 2^{-4m+3} \geqslant c2^{-m} \text{ for all } m \geqslant M.$$

Picking $\ell_1(r) := \lfloor -\log_2(r/c) \rfloor$, we observe that $B_r(q_{\ell,i})$ is disjoint from the supports of any singularities in $\biguplus_{i=M}^{\ell_1(r)} D_i$. Hence, we bound the mass from the first M levels using Lemma SM1.11(c), then note that there is no contribution from levels $M, \dots, \ell_1(r)$, and finally bound the total mass from level $\ell_1(r)+1$ onwards crudely. Fix $\delta := \inf_{q' \in \biguplus_{i=1}^{M} D_i} |q_{\ell,i} - q'|$ and suppose that $r \leqslant \delta/2$; then

$$\begin{aligned}
&\mu(B_r(q_{\ell,i})) \\
&\quad \leqslant \mu_{k(\ell,i),m(\ell)}(B_r(0)) \\
&\qquad + \sum_{i=1}^{M} \sum_{j=1}^{2^{i-1}} \int_{x-r}^{x+r} \rho_{k(i,j),m(i)}(t - q_{i,j}) \, \mathrm{d}t + \sum_{i=\ell_1(r)+1}^{\infty} 2^{-i} \\
&\quad \leqslant \mu_{k(\ell,i),m(\ell)}(B_r(0)) + 2M(\delta/2)^{-1/2} r + \sum_{i=\ell_1(r)+1}^{\infty} 2^{-i} \qquad \text{(Lemma SM1.11(c))} \\
&\quad = \mu_{k(\ell,i),m(\ell)}(B_r(0)) + O(r) \text{ as } r \to 0.
\end{aligned}$$

As $\mu_{k(\ell,i),m(\ell)}(B_r(0)) \in \Theta(r^{1/2})$ as $r \to 0$ (Proposition 5.10(c)), the $O(r)$ term is negligible and hence $\mu(B_r(q_{\ell,i})) \sim \mu_{k(\ell,i),m(\ell)}(B_r(0))$.

(b) Take $\beta \in (\beta_2(x), 4)$; (SM1.3) implies that

$$\liminf_{\ell\to\infty} \frac{\inf_{q\in\biguplus_{i=1}^{\ell} D_i}|x-q|}{2^{-\beta\ell}} = \infty.$$

Hence, for $r_\ell \coloneqq 2^{-4\ell+3}$, it follows that $\inf_{q\in\biguplus_{i=1}^{\ell} D_i}|x-q| - r_\ell \in \Omega(2^{-\beta\ell})$ as $\ell\to\infty$. Furthermore, as the supports of the densities centred at distinct elements of D_ℓ are disjoint and at least a distance $2^{-\ell}-2^{-4\ell+4} \in \Omega(2^{-\ell})$ apart (Lemma SM1.11(b)), there must exist $L\in\mathbb{N}$ and $c>0$ such that, for all $\ell \geqslant L$,

$$\inf_{q\in\biguplus_{i=1}^{\ell} D_i}|x-q| - r_\ell > c2^{-\beta\ell}$$
$$2^{-\ell} - 2^{-4\ell+4} > c2^{-\ell}.$$

Defining $\ell_1(r) \coloneqq \lfloor -\frac{1}{\beta}\log_2(r/c)\rfloor$ and $\ell_2(r) \coloneqq \lfloor -\log_2(r/c)\rfloor$, we see that if $L\leqslant \ell \leqslant \ell_1(r)$, then $B_r(x)$ is disjoint from the support of every density centred at a point of D_ℓ, and if $\ell_1(r) < \ell \leqslant \ell_2(r)$, then $B_r(x)$ intersects the support of at most one density centred at a point in D_ℓ. For $\ell > \ell_2(r)$, it is sufficient to bound $\mu(B_r(x))$ by counting the total mass added in the ℓ^{th} level.
Hence, let $\delta \coloneqq \inf_{q\in\biguplus_{i=1}^{L} D_i}|x-q| > 0$ and pick $r < \delta/2$ so that we may bound the mass from the first L levels using Lemma SM1.11(c). Using this and the claims above,

$$\begin{aligned}
&\mu(B_r(x))\\
&\quad\leqslant \sum_{\ell=1}^{L}\sum_{i=1}^{2^{\ell-1}}\int_{x-r}^{x+r}\rho_{k(\ell,i),m(\ell)}(t-q_{\ell,i})\,\mathrm{d}t + \sum_{\ell=\ell_1(r)+1}^{\ell_2(r)} m(\ell)\\
&\qquad+\sum_{\ell=\ell_2(r)+1}^{\infty} 2^{-\ell}\\
&\quad\leqslant 2L(\delta/2)^{-1/2}r + \sum_{\ell=\ell_1(r)+1}^{\ell_2(r)} 2^{-2\ell+1} + 2^{-\ell_2(r)} &&\text{(Lemma SM1.11(c))}\\
&\quad\leqslant 2L(\delta/2)^{-1/2}r + \frac{8}{3}2^{-2(\ell_1(r)+1)} + \frac{2r}{c}\\
&\quad\leqslant 2L(\delta/2)^{-1/2}r + \frac{8}{3}\left(\frac{r}{c}\right)^{2/\beta} + \frac{2r}{c} \in O(r^{\min\{1,2/\beta\}}) \text{ as } r\to 0.
\end{aligned}$$

(c) The case $\beta_2(x) < 4$ follows from (b). Hence, without loss of generality, suppose that $\beta_2(x) > 1$ and pick $\beta \in (1, \beta_2(x))$; (SM1.3) implies that

$$\liminf_{\ell\to\infty} \frac{\inf_{q\in\biguplus_{i=1}^{\ell} D_i}|x-q|}{2^{-\beta\ell}} = 0.$$

Hence, there must exist a sequence $(\ell_k)_{k\in\mathbb{N}} \nearrow \infty$ and a sequence $(q_{\ell_k})_{k\in\mathbb{N}}$ with $q_{\ell_k} \in D_{\ell_k}$ such that $|x-q_{\ell_k}| < 2^{-\beta\ell_k-1}$. This implies that any $q_{\ell_k} \neq q \in \biguplus_{i=1}^{\ell_k} D_i$ must satisfy

$|x-q| > 2^{-\ell_k} - 2^{-\beta\ell_k-1}$. As it suffices to bound $\mu(B_r(x))$ at the radii $s_k \coloneqq 2^{-\beta\ell_k-1} \searrow 0$, we proceed by bounding the mass contributed by the first ℓ_k levels by the total mass from the density centred at q_{ℓ_k} plus a $\Theta(r)$ term given by Lemma SM1.11(c).

For $\ell > \ell_k$, by a similar argument to that used above, any $q \in D_\ell$ satisfies $|x-q| > 2^{-\ell} - 2^{-\beta\ell_k-1}$. As the density centred at q is truncated at a radius at most $r_\ell \coloneqq 2^{-4\ell+3}$, and $2^{-\ell} - 2^{-4\ell+3} \in \Omega(2^{-\ell})$ as $\ell \to \infty$, there must exist $L \in \mathbb{N}$ and $c \in (0,1)$ such that for $L \leqslant \ell \leqslant \beta\ell_k$,

$$|x-q| - r_\ell > 2^{-\ell} - 2^{-\beta\ell_k-1} - 2^{-4\ell+3} > c2^{-\ell} - 2^{-\beta\ell_k-1}.$$

So, $B_{s_k}(x)$ does not intersect the support of any density centred at a point of D_ℓ if $s_k < c2^{-\ell} - 2^{-\beta\ell_k-1}$; hence, if $L \leqslant \ell < \beta\ell_k + \log_2(c)$, then $B_{s_k}(x)$ does not intersect the support of any density in the ℓ^{th} level.

Combining these two claims and taking k large enough that $\ell_k \geqslant L+1$ yields the bound

$$\begin{aligned}
\mu(B_{s_k}(x)) &\leqslant \sum_{\ell=1}^{\ell_k-1} \sum_{i=1}^{2^{\ell-1}} \int_{x-r}^{x+r} \rho_{k(\ell,i),m(\ell)}(t-q_{\ell,i})\,\mathrm{d}t \\
&\quad + \mu(B_{s_k}(q_{\ell_k})) + \sum_{\ell=\lfloor \beta\ell_k + \log_2(c)\rfloor}^{\infty} 2^{-\ell} \\
&\leqslant 2\ell_k\Big(2^{-\ell_k} - 2^{-\beta\ell_k}\Big)^{-1/2} s_k + \mu(B_{s_k}(q_{\ell_k})) + \frac{8s_k}{c} \qquad \text{(Lemma SM1.11(c))} \\
&\leqslant 2\ell_k\Big((2s_k)^{1/\beta} - 2s_k\Big)^{-1/2} s_k + \mu(B_{s_k}(q_{\ell_k})) + \frac{8s_k}{c}.
\end{aligned}$$

As $(2s_k)^{1/\beta} - 2s_k \in \Omega(s_k^{1/\beta})$ as $k \to \infty$, we may pick k sufficiently large that $(2s_k)^{1/\beta} - 2s_k \geqslant Cs_k^{1/\beta}$ for some $C > 0$. Hence, using that $\ell_k = -\frac{1}{\beta}(\log_2(s_k) - 1)$,

$$\begin{aligned}
\mu(B_{s_k}(x)) &\leqslant \mu(B_{s_k}(q_{\ell_k})) + 2\ell_k(Cs_k^{1/\beta})^{-\frac{1}{2}} s_k + \frac{8s_k}{c} \\
&\leqslant \mu(B_{s_k}(q_{\ell_k})) - \frac{2C^{-1/2}}{\beta}\big(\log_2(s_k) - 1\big) s_k^{1-1/2\beta} + \frac{8s_k}{c} \\
&= \mu(B_{s_k}(q_{\ell_k})) + o(s_k^{1/2}) \text{ as } k \to \infty.
\end{aligned}$$

The claim follows because

$$\liminf_{r\to 0} \frac{\mu(B_r(x))}{\mu(B_r(q))} \leqslant \liminf_{k\to\infty} \frac{\mu(B_{s_k}(q_{\ell_k}))}{\mu(B_{s_k}(q))} < 1,$$

and the RCDFs at distinct dyadic rationals are chosen so that their ratio oscillates on either side of unity.

(d) Take $\beta \in (4, \beta_2(x))$. Then, by (SM1.3), there exists a sequence of levels $(\ell_k)_{k\in\mathbb{N}} \nearrow \infty$ and a sequence $(q_{\ell_k})_{k\in\mathbb{N}}$ with $q_{\ell_k} = q_{\ell_k,i_k} \in D_{\ell_k}$ with $|x - q_{\ell_k}| < c2^{-\beta\ell_k}$. Ignoring the contribution from densities centred at points other than q_{ℓ_k}, we observe that

$$\mu(B_r(x)) \geqslant \int_{x-r}^{x+r} \rho_{k(\ell_k,i_k),m(\ell_k)}(t - q_{\ell_k})\,\mathrm{d}t.$$

Either $x > q_{\ell_k}$ or $x < q_{\ell_k}$; we deal with the first case as the second is almost identical. Fix $s_k := 2^{-4\ell_k+3}$. By translating the density, we see that

$$\begin{aligned}\mu(B_{s_k}(x)) &\geqslant \int_{-s_k+(x-q_{\ell_k})}^{s_k} \rho_{k(\ell_k,i_k),m(\ell_k)}(t)\,\mathrm{d}t \\ &= \int_{-s_k}^{s_k} \rho_{k(\ell_k,i_k),m(\ell_k)}(t)\,\mathrm{d}t - \int_{-s_k}^{-s_k+(x-q_{\ell_k})} \rho_{k(\ell_k,i_k),m(\ell_k)}(t)\,\mathrm{d}t.\end{aligned}$$

Indeed, as the density $\rho_{k(\ell_k,i_k),m(\ell_k)}$ is truncated at a radius at most $2^{-4\ell+3}$ (Proposition 5.10(e)), and as $|x - q_{\ell_k}| \in o(s_k)$, one sees that the ball mass around x asymptotically approaches the ball mass around the approximant q_{ℓ_k}. By a similar argument to (c),

$$\limsup_{r\to 0} \frac{\mu(B_r(x))}{\mu(B_r(q))} \geqslant \limsup_{k\to\infty} \frac{\mu(B_{s_k}(q_{\ell_k}))}{\mu(B_{s_k}(q))} > 1,$$

i.e. $x \not\preccurlyeq_0 q$, because the ratio of RCDFs at two distinct dyadic rationals oscillates on either side of one. The claim on incomparability then follows because $q \not\preccurlyeq_0 x$ (Lemma SM1.12(c)). ■

SM2. Alternative small-radius preorders. This section briefly outlines some alternatives to Definition 5.1 of $\preccurlyeq_0$ and their shortcomings.

The main difficulty that one encounters with alternative definitions is that the corresponding relation may not be transitive. We claim that transitivity is an essential property for any small-radius relation: without transitivity, it is not meaningful to talk about maximal and greatest elements, and so the characterisation of modes as greatest elements of an order fails.

Of course, for a small-radius preorder to be relevant to us, its greatest elements must have some natural interpretation as "points of maximum probability". In some sense, determining what characterises a point of maximum probability is the main challenge, but, motivated by the examples considered throughout the paper, we believe that none of the alternative small-radius preorders are a significant improvement on preorder $\preccurlyeq_0$.

It seems natural to define an ordering on X by taking limits of the positive-radius preorders $\preccurlyeq_r$ as $r \to 0$. As any binary relation can be viewed as a subset of the Cartesian product $X \times X$, where $(x, x') \in \preccurlyeq_r$ precisely when $x \preccurlyeq_r x'$, we define some candidate limiting orderings using set-theoretic limits of the net $(\preccurlyeq_r)_{r>0}$. The corresponding limit set need not be a preorder in general, but we show that certain set-theoretic limits do always yield a preorder.

Indeed, the *set-theoretic limits inferior and superior* of a net $(A_r)_{r>0}$ of subsets of X are defined by

$$\begin{aligned}\liminf_{r\to 0} A_r &:= \bigcup_{R>0} \bigcap_{r<R} A_r = \{y \in X \mid y \in A_r \text{ for all } r < R(y)\}, \\ \limsup_{r\to 0} A_r &:= \bigcap_{R>0} \bigcup_{r<R} A_r = \{y \in X \mid \text{for some null sequence } (r_n)_{n\in\mathbb{N}},\, y \in A_{r_n} \text{ for all } n \in \mathbb{N}\},\end{aligned}$$

and the *Kuratowski lower and upper limits* of $(A_r)_{r>0}$ are defined by

$$\operatorname*{Li}_{r\to 0} A_r := \{y \in X \mid y \text{ is a limit point of the net } (A_r)_{r>0}\},$$
$$\operatorname*{Ls}_{r\to 0} A_r := \{y \in X \mid y \text{ is a cluster point of the net } (A_r)_{r>0}\}.$$

The following is a useful equivalent characterisation of the Kuratowski limits:

Lemma SM2.1 ([SM4, Lemmas 5.2.7 and 5.2.8]). *Let X be any metric space and let $(A_r)_{r>0}$ be a net of subsets of X.*

(a) *$x \in \mathrm{Li}_{r\to 0} A_r$ if and only if there exists a net $(x_r)_{r>0}$ converging to x with $x_r \in A_r$.*
(b) *$x \in \mathrm{Ls}_{r\to 0} A_r$ if and only if there exists a decreasing null sequence $(r_n)_{n\in\mathbb{N}}$ and a sequence $(x_{r_n})_{n\in\mathbb{N}}$ converging to x with $x_{r_n} \in A_{r_n}$.*

The set limits described above give four different approaches to taking the limit of the sets $(\preccurlyeq_r)_{r>0}$, which we denote

$$\preccurlyeq_0^{\liminf} := \liminf_{r\to 0} \preccurlyeq_r, \qquad \preccurlyeq_0^{\limsup} := \limsup_{r\to 0} \preccurlyeq_r,$$
$$\preccurlyeq_0^{\mathrm{Li}} := \operatorname*{Li}_{r\to 0} \preccurlyeq_r, \qquad \preccurlyeq_0^{\mathrm{Ls}} := \operatorname*{Ls}_{r\to 0} \preccurlyeq_r.$$

Proposition SM2.2. *Let X be a metric space and let $\mu \in \mathscr{P}(X)$.*

(a) *$\preccurlyeq_0^{\liminf}$ is a preorder;*
(b) *$\preccurlyeq_0^{\liminf}$ is a subset of $\preccurlyeq_0$;*
(c) *$x \preccurlyeq_0^{\liminf} x' \implies x \preccurlyeq_0 x'$.*

Proof. For (a), it is routine to check that $\preccurlyeq_0^{\liminf}$ is a preorder: reflexivity is obvious, and if $x \preccurlyeq_0^{\liminf} y$ and $y \preccurlyeq_0^{\liminf} z$ then there exists $R > 0$ such that, for all $r < R$, $x \preccurlyeq_r y$ and $y \preccurlyeq_r z$, giving $x \preccurlyeq_r z$ by transitivity of $\preccurlyeq_r$.

For (b), $(x, x') \in \preccurlyeq_0^{\liminf}$ implies that, for some $R > 0$ and all $r < R$, $x \preccurlyeq_r x'$. Hence,

$$\limsup_{r\to 0} \frac{\mu(B_r(x))}{\mu(B_r(x'))} \leqslant 1,$$

so $x \preccurlyeq_0 x'$ by definition. Claim (c) follows immediately from (b): if $(x, x') \in \preccurlyeq_0^{\liminf}$, then $(x, x') \in \preccurlyeq_0$, so $x \preccurlyeq_0 x'$. ∎

As a consequence of (b), any $\preccurlyeq_0$-antichain is also a $\preccurlyeq_0^{\liminf}$-antichain. Hence, Theorem 5.11 gives an example of a countable dense $\preccurlyeq_0^{\liminf}$-antichain; this demonstrates that $\preccurlyeq_0^{\liminf}$ does not have better behaviour in this regard than $\preccurlyeq_0$.

The set-theoretic ordering $\preccurlyeq_0^{\liminf}$ can also be criticised as unnecessarily strict in cases where $x' \prec_0 x$ for any $r > 0$, but

$$\limsup_{r\to 0} \frac{\mu(B_r(x))}{\mu(B_r(x'))} = 1.$$

Example 4.9 gives a measure where the $\preccurlyeq_0^{\liminf}$-greatest elements and the $\preccurlyeq_0$-greatest elements differ: under $\preccurlyeq_0^{\liminf}$ only $+1$ is greatest, whereas both -1 and $+1$ are $\preccurlyeq_0$-greatest. While $\preccurlyeq_0^{\liminf}$-greatest elements are reasonable candidates for modes, they do not seem to correspond exactly to any of the established definitions of modes. To be more precise, while Proposition SM2.2(b) implies that they are always weak modes, it is not clear whether or not they are strong modes, and not all weak modes are $\preccurlyeq_0^{\liminf}$-greatest.

Example SM2.3 ($\preccurlyeq_0^{\limsup}$ is not necessarily transitive). The essential idea is even if $x \preccurlyeq_{r_n} y$ for some null sequence $(r_n)_{n\in\mathbb{N}}$, and $y \preccurlyeq_{r'_n} z$ for some null sequence $(r'_n)_{n\in\mathbb{N}}$, it is possible that $x \not\preccurlyeq_r z$ for any $r > 0$. For a concrete example of this situation, let

$$X := \left\{-2 \pm 2^{-3n+2} \,\middle|\, n \in \mathbb{N}\right\} \cup \left\{2 \pm 2^{-3n+2} \,\middle|\, n \in \mathbb{N}\right\} \cup \left\{\pm 2^{-3n+2} \,\middle|\, n \in \mathbb{N}\right\}$$

with its usual Euclidean metric. Define the "target RCDFs"

$$f(2^{-3n+2}) := 2^{-3n+2}, \qquad g(2^{-3n+2}) := 2^{-3n+1},$$
$$h(2^{-3n+2}) := \begin{cases} 2^{-3n+3}, & \text{if } n \text{ is odd,} \\ 2^{-3n}, & \text{if } n \text{ is even,} \end{cases}$$

and let

$$\begin{aligned} \mu := {} & \frac{1}{Z}\sum_{n\in\mathbb{N}} \frac{1}{2}\big(f(2^{-3n+2}) - f(2^{-3n-1})\big)\big(\delta_{-2-2^{-3n+2}} + \delta_{-2+2^{-3n+2}}\big) \\ & + \frac{1}{Z}\sum_{n\in\mathbb{N}} \frac{1}{2}\big(g(2^{-3n+2}) - g(2^{-3n-1})\big)\big(\delta_{2-2^{-3n+2}} + \delta_{2+2^{-3n-1}}\big) \\ & + \frac{1}{Z}\sum_{n\in\mathbb{N}} \frac{1}{2}\big(h(2^{-3n+2}) - h(2^{-3n-1})\big)\big(\delta_{-2^{-3n+2}} + \delta_{2^{-3n-1}}\big), \end{aligned}$$

where Z is a normalisation constant chosen to ensure that $\mu \in \mathscr{P}(X)$.

The construction of μ ensures that the RCDFs at -2, $+2$ and 0 are $\frac{1}{Z}f(r)$, $\frac{1}{Z}g(r)$ and $\frac{1}{Z}h(r)$ for $r \leqslant 2^{-1}$. Then

$$\frac{\mu(B_{2^{-3n+2}}(-2))}{\mu(B_{2^{-3n+2}}(0))} = \begin{cases} 2^{-1}, & \text{if } n \text{ is odd,} \\ 2^{2}, & \text{if } n \text{ is even,} \end{cases}$$
$$\frac{\mu(B_{2^{-3n+2}}(0))}{\mu(B_{2^{-3n+2}}(+2))} = \begin{cases} 2^{2} & \text{if } n \text{ is odd,} \\ 2^{-1} & \text{if } n \text{ is even,} \end{cases}$$
$$\frac{\mu(B_{2^{-3n+2}}(-2))}{\mu(B_{2^{-3n+2}}(+2))} = 2.$$

It follows that $-2 \asymp_0^{\limsup} 0$, because there are null sequences $(r_n)_{n\in\mathbb{N}}$ such that $-2 \preccurlyeq_{r_n} 0$ and vice versa; the same argument shows that $0 \asymp_0^{\limsup} +2$. But $-2 \prec_r 2$ for any $r > 0$, and hence $-2 \not\asymp_0^{\limsup} 2$. This violates transitivity.

Example SM2.4 ($\preccurlyeq_0^{\mathrm{Li}}$ and $\preccurlyeq_0^{\mathrm{Ls}}$ are not necessarily transitive). Let $\mu \in \mathscr{P}(\mathbb{R})$ be the measure with Lebesgue density $\rho(x) := \mathbb{1}\big[x \in [0,1]\big]$. We first verify that:

(a) $x \preccurlyeq_0^{\mathrm{Li}} 1$ for any $x \in \mathbb{R}$;
(b) $1 \preccurlyeq_0^{\mathrm{Li}} x$ for any $x \in \mathbb{R}$;

(c) $1/2 \not\preccurlyeq_0^{\mathrm{Ls}} x$ for any $x \in \mathbb{R} \setminus [0,1]$; and
(d) $1/2 \not\preccurlyeq_0^{\mathrm{Li}} x$ for any $x \in \mathbb{R} \setminus [0,1]$.

For (a), observe that $x \preccurlyeq_r 1-r$ for all small r, so $(x, 1-r) \in \preccurlyeq_r$. Hence, $(x,1) \in \preccurlyeq_0^{\mathrm{Li}}$ by Lemma SM2.1.

For (b), use that $1+r \preccurlyeq_r x$ for all small r, so $(1+r, x) \in \preccurlyeq_r$. This implies that $(1,x) \in \preccurlyeq_0^{\mathrm{Li}}$.

For (c), suppose that $(x'_{r_n}, x_{r_n}) \to (1/2, x)$, and $(x'_{r_n}, x_{r_n}) \in \preccurlyeq_{r_n}$. Let $\varepsilon = \min\{|x|, |x-1|\} > 0$. There exists $N_1 \in \mathbb{N}$ such that, for all $n \geqslant N_1$, $\min\{|x_{r_n}|, |x_{r_n} - 1|\} > \varepsilon/2$. As $(r_n)_{n\in\mathbb{N}}$ is a decreasing null sequence, there exists $N_2 \in \mathbb{N}$ such that, for all $n \geqslant N_2$, $r_n < \varepsilon/2$. Picking $N \coloneqq \max\{N_1, N_2\}$, we have $\mu(B_{r_n}(x_{r_n})) = 0$ for $n \geqslant N$, because $B_{r_n}(x_{r_n}) \cap [0,1] = \varnothing$. It is easy to see that if $x'_{r_n} \to 1/2$, then for sufficiently large n $\mu(B_{r_n}(x'_{r_n})) > 0$. Hence, for all sufficiently large n, $x'_{r_n} \not\preccurlyeq_{r_n} x_{r_n}$. This is a contradiction.

Claim (d) is a corollary of (c), because $\preccurlyeq_0^{\mathrm{Li}} \subseteq \preccurlyeq_0^{\mathrm{Ls}}$.

Now we prove that $\preccurlyeq_0^{\mathrm{Li}}$ and $\preccurlyeq_0^{\mathrm{Ls}}$ are not transitive. Suppose for contradiction that they are: then (a) and (b) imply that every point $x \in \mathbb{R}$ is equivalent to 1, and so all points in $\mathbb{R}$ are equivalent by transitivity. As $\preccurlyeq_0^{\mathrm{Li}} \subseteq \preccurlyeq_0^{\mathrm{Ls}}$, this implies that all points are also $\preccurlyeq_0^{\mathrm{Ls}}$-equivalent. However, (c) and (d) show that not all points in $\mathbb{R}$ are $\preccurlyeq_0^{\mathrm{Ls}}$-equivalent or $\preccurlyeq_0^{\mathrm{Li}}$-equivalent.

REFERENCES

[1] S. Agapiou, M. Burger, M. Dashti, and T. Helin, *Sparsity-promoting and edge-preserving maximum a posteriori estimators in non-parametric Bayesian inverse problems*, Inverse Probl., 34 (2018), 045002, 37pp., https://doi.org/10.1088/1361-6420/aaacac.

[2] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, Berlin, third ed., 2006, https://doi.org/10.1007/3-540-29587-9.

[3] T. W. Anderson, *The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities*, Proc. Amer. Math. Soc., 6 (1955), pp. 170–176, https://doi.org/10.2307/2032333.

[4] G. Beer, *Topologies on Closed and Closed Convex Sets*, vol. 268 of Mathematics and Its Applications, Springer, Dordrecht, 1993, https://doi.org/10.1007/978-94-015-8149-3.

[5] M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss, *MAP estimators and their consistency in Bayesian nonparametric inverse problems*, Inverse Probl., 29 (2013), no. 095017, 27pp., https://doi.org/10.1088/0266-5611/29/9/095017.

[6] N. I. Fel′dman and Yu. V. Nesterenko, *Number Theory* IV, vol. 44 of Encyclopaedia of Mathematical Sciences, Springer, Berlin, Heidelberg, 1998, https://doi.org/10.1007/978-3-662-03644-0.

[7] Ya. I. Grushka, *On monotonous separately continuous functions*, Appl. Gen. Topol., 20 (2019), pp. 75–79, https://doi.org/10.4995/agt.2019.9817.

[8] M. Lassas, E. Saksman, and S. Siltanen, *Discretization-invariant Bayesian inversion and Besov space priors*, Inverse Probl. Imaging, 3 (2009), pp. 87–122, https://doi.org/10.3934/ipi.2009.3.87.

[9] W. Young, *A note on monotone functions*, Q. J. Pure Appl. Math. (Oxf.), 41 (1910), pp. 79–87.