SUPPLEMENTARY MATERIAL TO 'GLOBAL RATES OF CONVERGENCE IN LOG-CONCAVE DENSITY ESTIMATION'

By Arlene K. H. Kim[†] and Richard J. Samworth[*,†]

University of Cambridge[†]

This is the supplementary material for Kim and Samworth (2016), hereafter referred to as the main text.

1. Proof and auxiliary result for Theorem 1 in the main text. We recall that the Kullback–Leibler divergence between two densities f and g on $\mathbb{R}^d$ is given by

$$d_{\mathrm{KL}}^2(f,g) := \int_{\mathbb{R}^d} f \log \frac{f}{g}.$$

Given $\epsilon > 0$ and $\rho \in \{d_{\mathrm{KL}}, h\}$, we write $V(\epsilon, \mathcal{F}_d, \rho)$ for the ϵ-covering number of $\mathcal{F}_d$ with respect to ρ; thus $V(\epsilon, \mathcal{F}_d, \rho)$ is the minimal $N \in \mathbb{N}$ with the property that there exist densities $f_1, \ldots, f_N$ such that for any $f \in \mathcal{F}_d$, we can find $j^* \in \{1, \ldots, N\}$ with $\rho(f, f_{j^*}) \leq \epsilon$. We also write $N(\epsilon, \mathcal{F}_d, h)$ for the ϵ-packing number of $\mathcal{F}_d$ with respect to h; thus $N(\epsilon, \mathcal{F}_d, h)$ is the maximal $N \in \mathbb{N}$ with the property that there exist densities $f_1, \ldots, f_N \in \mathcal{F}_d$ with $h(f_j, f_k) > \epsilon$ for $j \neq k$. Finally, given $0 < \underline{c} \leq \bar{C} < \infty$ and a compact interval $I \subseteq \mathbb{R}$, we write $\mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I}$ for the set of univariate upper semi-continuous densities f that are concave on I, and satisfy $\underline{c} \leq f(x) \leq \bar{C}$ for $x \in I$, and $f(x) = 0$ for $x \notin I$. Note that $\mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I} \subseteq \mathcal{F}_1$.

Proof of Theorem 1. *The case* $d = 1$: There exist $0 < \underline{c} \leq \bar{C} < \infty$, $c_1^* \in (0, \infty)$, a compact interval $I \subseteq \mathbb{R}$ and $\epsilon_0 > 0$ such that, given $\epsilon \in (0, \epsilon_0]$, we can find a subfamily $\bar{\mathcal{F}}_1 := \{f_1, \ldots, f_{N_\epsilon}\} \subseteq \mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I}$ with the following properties:

(i) $\log N_\epsilon \geq c_1^* \epsilon^{-1/2}$;
(ii) $h(f_j, f_k) > \epsilon$ for all $j \neq k$.

Such a construction can be obtained, for instance, by adapting the convex densities of Devroye and Lugosi (2001, Lemma 15.1); alternatively, details can be found in the `arxiv` version of this paper (Kim and Samworth, 2015,

[*]The research of Richard J. Samworth was supported by an EPSRC Early Career Fellowship and a grant from the Leverhulme Trust.

Theorem 1). We therefore have that $\log N(\epsilon, \mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I}, h) \geq c_1^* \epsilon^{-1/2}$. Moreover, we can recall that

$$d^2_{\mathrm{KL}}(f,g) \leq \frac{4\bar{C}}{\underline{c}} h^2(f,g)$$

for $f, g \in \mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I}$. We deduce that, for some $K_1^\dagger \geq 1$,

$$V(\epsilon, \mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I}, d_{\mathrm{KL}}) \leq N\biggl(\frac{\underline{c}^{1/2}\epsilon}{2\bar{C}^{1/2}}, \mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I}, h\biggr) \leq V\biggl(\frac{\underline{c}^{1/2}\epsilon}{4\bar{C}^{1/2}}, \mathcal{F}^{\mathrm{conc}}_{\underline{c},\bar{C},I}, h\biggr) \leq \exp(K_1^\dagger \epsilon^{-1/2}),$$

where the final claim follows because square roots of densities that are concave on I are concave on I, so we can apply the L_2-covering number bounds of Guntuboyina and Sen (2013, Theorem 3.1), or Proposition 4 below. It follows from Yang and Barron (1999, Theorem 1), restated as Lemma 1 below for convenience, that we can take $\epsilon_n = (K_1^\dagger)^{2/5} n^{-2/5}$ and $\underline{\epsilon}_{n,h} = \frac{(c_1^*)^2}{36(K_1^\dagger)^{8/5}} n^{-2/5}$ there to conclude that for n large enough that $\underline{\epsilon}_{n,h} \leq \epsilon_0$,

$$\inf_{\tilde{f}_n \in \tilde{\mathcal{F}}_n} \sup_{f_0 \in \mathcal{F}_1} \mathbb{E}_{f_0}\bigl\{h^2(\tilde{f}_n, f_0)\bigr\} \geq \frac{(c_1^*)^4}{8 \times 36^2 (K_1^\dagger)^{16/5}} n^{-4/5}.$$

The case $d \geq 2$: As mentioned in Section 2 in the main text, $\mathcal{F}_d$ contains the class of uniform densities on closed, convex sets. The result therefore follows by slight modifications of the arguments in, for example, the proof of Brunel (2013, Theorem 5); see also Brunel (2016); Korostelev and Tsybakov (1993); Mammen and Tsybakov (1995). □

The following lemma is a special case of Yang and Barron (1999, Theorem 1).

LEMMA 1 (Yang and Barron (1999), Theorem 1). *Suppose that* $\epsilon_n > 0$ *is such that* $\epsilon_n^2 \geq n^{-1} \log V(\epsilon_n, \mathcal{F}_d, d_{\mathrm{KL}})$ *and that* $\underline{\epsilon}_{n,h} > 0$ *is such that* $\log N(\underline{\epsilon}_{n,h}, \mathcal{F}_d, h) \geq 4n\epsilon_n^2 + 2\log 2$. *Then*

$$\inf_{\tilde{f}_n \in \tilde{\mathcal{F}}_n} \sup_{f_0 \in \mathcal{F}_d} \mathbb{E}_{f_0}\bigl\{h^2(\tilde{f}_n, f_0)\bigr\} \geq \frac{1}{8}\underline{\epsilon}^2_{n,h}.$$

2. Auxiliary results for the proof of Theorem 4 in the main text. We first provide the following entropy bound for convex sets, which is a minor extension of Dudley (1999, Corollary 8.4.2). For a d-dimensional, closed, convex set $D \subseteq \mathbb{R}^d$, we write $\mathcal{A}_d(D)$ for the class of closed, convex subsets of

D. Further, and in a slight abuse of notation, we let $N_{[]}(\epsilon, \mathcal{A}_d(D), L_1)$ denote the ϵ-bracketing number of $\{\mathbb{1}_A : A \in \mathcal{A}_d(D)\}$ in the $L_1 = L_1(\mu_d)$-metric. Recall also that we write $\log_{++}(x) = \max(1, \log x)$.

PROPOSITION 2. *For each $d \in \mathbb{N}$, there exists $K_d \in (0,\infty)$, depending only on d, such that*

$$\log N_{[]}\bigl(\epsilon, \mathcal{A}_d(D), L_1\bigr) \leq K_d \max\biggl\{\log_{++}\Bigl(\frac{\mu_d(D)}{\epsilon}\Bigr), \Bigl(\frac{\mu_d(D)}{\epsilon}\Bigr)^{(d-1)/2}\biggr\}$$

for all $\epsilon > 0$.

PROOF. By Fritz John's theorem (Ball, 1997; John, 1948, p. 13), there exist $A \in \mathbb{R}^{d\times d}$ and $b \in \mathbb{R}^d$ such that $D' := AD + b$ has the property that $d^{-1}\bar{B}_d(0,1) \subseteq D' \subseteq \bar{B}_d(0,1)$. Let $a_d := \mu_d\bigl(\bar{B}_d(0,1)\bigr) = \pi^{d/2}/\Gamma(1 + d/2)$. Now, by Dudley (1999, Corollary 8.4.2) and the remark immediately preceding it, there exists $\epsilon_{20,d} \in \bigl(0, \min(e^{-1}, a_d)\bigr)$ and $\check{\check{K}}_d \in (0,\infty)$ such that

$$\begin{aligned}\log N_{[]}\bigl(\epsilon, \mathcal{A}_d(D'), L_1\bigr) &\leq \log N_{[]}\bigl(\epsilon, \mathcal{A}_d(\bar{B}_d(0,1)), L_1\bigr)\\ &\leq \check{\check{K}}_d \max\{\log(1/\epsilon), \epsilon^{-(d-1)/2}\}\end{aligned}$$

for all $\epsilon \in (0, \epsilon_{20,d}]$. Now set

$$\check{K}_d := \check{\check{K}}_d \frac{\max\{\log(1/\epsilon_{20,d}), \epsilon_{20,d}^{-(d-1)/2}\}}{\max\{\log_{++}(1/a_d), a_d^{-(d-1)/2}\}}.$$

Then, for $\epsilon \in (\epsilon_{20,d}, a_d)$,

$$\begin{aligned}&\log N_{[]}\bigl(\epsilon, \mathcal{A}_d(D'), L_1\bigr) \leq \log N_{[]}\bigl(\epsilon_{20,d}, \mathcal{A}_d(D'), L_1\bigr)\\ &\quad\leq \check{\check{K}}_d \max\{\log(1/\epsilon_{20,d}), \epsilon_{20,d}^{-(d-1)/2}\} = \check{K}_d \max\{\log_{++}(1/a_d), a_d^{-(d-1)/2}\}\\ &\quad\leq \check{K}_d \max\{\log_{++}(1/\epsilon), \epsilon^{-(d-1)/2}\}.\end{aligned}$$

For $\epsilon \geq a_d$, we can use the single bracketing pair $\{\psi^L, \psi^U\}$ with $\psi^L(x) := 0$ and $\psi^U(x) := 1$ for $x \in D'$, noting that $L_1(\psi^U, \psi^L) = \mu_d(D') \leq a_d$. Thus, for $\epsilon \geq a_d$,

$$\log N_{[]}\bigl(\epsilon, \mathcal{A}_d(D'), L_1\bigr) = 0 \leq \check{K}_d \max\{\log_{++}(1/\epsilon), \epsilon^{-(d-1)/2}\}.$$

We can therefore construct an ϵ-bracketing set in L_1 for $\{\mathbb{1}_A : A \in \mathcal{A}_d(D)\}$ as follows: first find an $\frac{\epsilon a_d}{d^d \mu_d(D)}$-bracketing set $\{[\psi_j^L, \psi_j^U] : j = 1, \ldots, N\}$ for $\{\mathbb{1}_A : A \in \mathcal{A}_d(D')\}$, where

$$\log N \leq \check{K}_d \max\biggl\{\log_{++}\Bigl(\frac{d^d \mu_d(D)}{\epsilon a_d}\Bigr), \Bigl(\frac{d^d \mu_d(D)}{\epsilon a_d}\Bigr)^{(d-1)/2}\biggr\}.$$

Now define $\phi_j^L, \phi_j^U : D \to \mathbb{R}$ by $\phi_j^L(x) := \psi_j^L(Ax+b)$ and $\phi_j^U(x) := \psi_j^U(Ax+b)$. Then

$$\begin{aligned} L_1(\phi_j^U, \phi_j^L) &= \int_D |\psi_j^U(Ax+b) - \psi_j^L(Ax+b)| \, d\mu_d(x) \\ &\leq \frac{\epsilon a_d}{|\det A| d^d \mu_d(D)} = \frac{\epsilon a_d}{d^d \mu_d(D')} \leq \frac{\epsilon a_d}{d^d \mu_d\bigl(d^{-1}\bar{B}_d(0,1)\bigr)} = \epsilon. \end{aligned}$$

Since $\log_{++}(a/\epsilon) \leq \bigl\{2 + \frac{2\log_{++}(a)}{\log_{++}(e/a)}\bigr\} \log_{++}(1/\epsilon)$ for all $a, \epsilon > 0$, the result therefore holds with

$$K_d := \check{K}_d \max\biggl\{\biggl(2 + \frac{2\log_{++}(d^d/a_d)}{\log_{++}(ea_d/d^d)}\biggr), \frac{d^{d(d-1)/2}}{a_d^{(d-1)/2}}\biggr\}.$$

□

We now provide a bracketing entropy bound for classes of uniformly bounded concave functions on arbitrary d-dimensional, convex, compact domains in $\mathbb{R}^d$ when $d = 1, 2, 3$. These results build on the work of Guntuboyina and Sen (2013), who study covering (as opposed to bracketing) numbers and rectangular domains, and a recent result of Gao and Wellner (2015), who study various special classes of domains, including d-dimensional simplices. For convenience, we state the result to which we will appeal below.

Recall that we say $\mathcal{S} \subseteq \mathbb{R}^d$ is a *d-dimensional simplex* if there exist affinely independent vectors $u_0, u_1, \ldots, u_d \in \mathbb{R}^d$ such that

$$\mathcal{S} = \biggl\{u_0 + \sum_{j=1}^d \lambda_j u_j : \lambda_1, \ldots, \lambda_d \geq 0, \sum_{j=1}^d \lambda_j \leq 1\biggr\}.$$

A set $D \subseteq \mathbb{R}^d$ can be *triangulated into simplices* if there exist d-dimensional simplices $S_1, \ldots, S_N \subseteq D$ such that $\cup_{j=1}^N S_j = D$ and if $j \neq k$ then there is a common (possibly empty) face F of the boundaries of S_j and S_k with $S_j \cap S_k = F$. For a d-dimensional, closed, convex subset D of $\mathbb{R}^d$, and for $B > 0$, we define $\bar{\Phi}_B(D)$ to be the set of upper semi-continuous, concave functions ϕ with $\mathrm{dom}(\phi) = D$ that are bounded in absolute value by B.

THEOREM 3 (Gao and Wellner (2015), Theorem 1.1(ii)). *For each $d \in \mathbb{N}$, there exists $K_d^{**} \in (0, \infty)$, depending only on d, such that if D is a d-dimensional closed, convex subset of $\mathbb{R}^d$ that can be triangulated into m simplices, then*

$$\log N_{[]}\bigl(2\epsilon, \bar{\Phi}_B(D), L_2\bigr) \leq K_d^{**} m \biggl(\frac{B\mu_d^{1/2}(D)}{\epsilon}\biggr)^{d/2}$$

for all $\epsilon > 0$.

For any d-dimensional, compact, convex set $D \subseteq \mathbb{R}^d$ and any $\eta \geq 0$, let

$$_{\eta}D := \{x \in D : w \in D \text{ for all } \|w - x\| \leq \eta\}, \quad \text{and} \quad D^{\eta]} := D + \eta \bar{B}_d(0,1).$$

We refer to Schenider (2014, Section 3.1) for basic properties of $_{\eta}D$ and $D^{\eta]}$, which we will use in Proposition 4 below.

We are now in a position to state our bracketing entropy bound.

PROPOSITION 4. *There exists* $K_d^{\circ} \in (0,\infty)$*, depending only on* d*, such that for all* d*-dimensional, convex, compact sets* $D \subseteq \mathbb{R}^d$ *and all* $B, \epsilon > 0$*, we have*

$$\log N_{[]}\bigl(2\epsilon, \bar{\Phi}_B(D), L_2\bigr) \leq \begin{cases} K_1^{\circ} \mu_1^{1/4}(D)(B/\epsilon)^{1/2} & \text{if } d = 1 \\ K_2^{\circ} \mu_2^{1/2}(D)(B/\epsilon)\log_{++}^{3/2}(B\mu_2^{1/2}(D)/\epsilon) & \text{if } d = 2 \\ K_3^{\circ} \mu_3(D)(B/\epsilon)^2 & \text{if } d = 3. \end{cases}$$

PROOF. As a preliminary, recall that the Hausdorff distance between two non-empty, compact subsets $A, B \subseteq \mathbb{R}^d$ is given by

$$\mathrm{Haus}(A,B) := \max\biggl\{\sup_{x\in A}\inf_{y\in B}\|x-y\|\,,\,\sup_{y\in B}\inf_{x\in A}\|x-y\|\biggr\}.$$

By the main result of Bronshteyn and Ivanov (1975), there exist $\delta_{\mathrm{BI},d} > 0$ and $C_d > 0$, both depending only on d, such that for every $\delta \in (0, \delta_{\mathrm{BI},d}]$ and every d-dimensional convex, compact set $D \subseteq \bar{B}_d(0,1)$, we can find a (convex) polytope $P \supseteq D$ such that P has at most $C_d\delta^{-(d-1)/2}$ vertices and $\mathrm{Haus}(P,D) \leq \delta$. (Throughout, we follow, e.g., Rockafellar (1997), and define a polytope to be a set formed as the convex hull of finitely many points.) Moreover, by Lemma 8.4.3 of Dudley (1999), there exists $c_0 \in (0, 16\delta_{\mathrm{BI},d}]$, depending only on d (though this dependence is suppressed for notational simplicity), such that for any d-dimensional, closed convex set $D \subseteq \bar{B}_d(0,1)$ and any $\delta > 0$, we have $\mu_d(D \setminus {}_{c_0\delta}D) \leq \delta/16$.

We now begin the main proof in the case $B = 1$, and handle the general case at the end of the whole argument. Fix a d-dimensional, convex, compact set $D \subseteq \mathbb{R}^d$, and, as in the proof of Proposition 2, apply Fritz John's theorem to construct an affine transformation $D' := AD + b$ of D such that $d^{-1}\bar{B}_d(0,1) \subseteq D' \subseteq \bar{B}_d(0,1)$. We initially find bracketing sets for $\bar{\Phi}_1(D')$, and consider different dimensions separately.

The case $d = 1$: This is an extension from metric to bracketing entropy of Theorem 3.1 of Guntuboyina and Sen (2013), and can be found in Doss and Wellner (2016, Proposition 4.1). In particular, these authors show that there exist $\epsilon_1^\circ \in (0,1)$ and $K_{1,1}^\circ > 0$ such that, when $d = 1$,

$$\log N_{[]}\bigl(2\epsilon, \bar{\Phi}_1(D'), L_2\bigr) \leq K_{1,1}^\circ \epsilon^{-1/2}$$

for all $\epsilon \in (0, \epsilon_1^\circ]$.

The case $d = 2$: Set $\epsilon_2^\circ := 1/8$, and fix $\epsilon \in (0, \epsilon_2^\circ]$, noting that $\mu_2(D' \setminus {}_{c_0\epsilon^2}D') \leq \epsilon^2/16$. Applying the result of Bronshteyn and Ivanov (1975), we can find a polytope $P_1 \supseteq {}_{c_0\epsilon^2}D'$ such that P_1 has at most $C_2 c_0^{-1/2}\epsilon^{-1}$ vertices and $\mathrm{Haus}(P_1, {}_{c_0\epsilon^2}D') \leq c_0\epsilon^2$. We deduce that $P_1 \subseteq ({}_{c_0\epsilon^2}D')^{c_0\epsilon^2]} \subseteq D'$. Applying the result of Bronshteyn and Ivanov (1975) recursively, with $M := \bigl\lfloor \log\bigl(\frac{1}{4\epsilon}\bigr)/\log 2 \bigr\rfloor$ (the condition that $\epsilon \leq 1/8$ ensures that $M \in \mathbb{N}$), for each $i = 2, 3, \ldots, M$, there exists a polytope $P_i \supseteq {}_{c_0 4^i\epsilon^2}(P_{i-1})$ with at most $C_2 c_0^{-1/2} 2^{-i}\epsilon^{-1}$ vertices such that $\mathrm{Haus}\bigl(P_i, {}_{c_0 4^i \epsilon^2}(P_{i-1})\bigr) \leq c_0 4^i \epsilon^2$. Observe that the Bronshteyn–Ivanov result can be applied in each case, because for $i = 2, 3, \ldots, M$,

$$c_0 4^i \epsilon^2 \leq c_0 4^M \epsilon^2 \leq \frac{c_0}{16} \leq \delta_{\mathrm{BI},2}.$$

Note moreover that the Bronshteyn–Ivanov construction yields $P_i \subseteq P_{i-1}$. We claim that P_M is a two-dimensional polytope, by our choice of M. In fact,

$$\begin{aligned}
\mu_2(P_M) &= \mu_2(D') - \mu_2(D' \setminus P_1) - \sum_{i=2}^M \mu_2(P_{i-1} \setminus P_i) \\
&\geq \frac{\pi}{4} - \mu_2(D' \setminus P_1) - \sum_{i=2}^M \mu_2(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})) \\
&\geq \frac{\pi}{4} - \frac{\epsilon^2}{16}\sum_{i=1}^M 4^i \geq \frac{\pi}{4} - 4^{M-1}\epsilon^2 \geq \frac{\pi}{8}.
\end{aligned}$$

For $i = 2, 3, \ldots, M$, we now describe how to construct a finite set of simplices (triangles) $S_{i,1}, \ldots, S_{i,N_i}$ that cover $P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})$, so in particular, they cover $P_{i-1} \setminus P_i$. Since ${}_{c_0 4^i\epsilon^2}(P_{i-1})$ is a two-dimensional polyhedral convex set, we can pick two distinct vertices in this set. The line L passing through these two points forms the boundary of two closed halfspaces $\mathcal{H}_1$ and $\mathcal{H}_2$; we show how to triangulate $\mathcal{H}_1 \cap \bigl(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})\bigr)$, with the triangulation of $\mathcal{H}_2 \cap \bigl(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})\bigr)$ being entirely analogous. We claim that, in the terminology of Devadoss and O'Rourke (2011), $\mathcal{H}_1 \cap \bigl(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})\bigr)$ is a polygon, i.e. a closed subset of $\mathbb{R}^2$ bounded by a finite collection of line segments forming a simple closed curve.

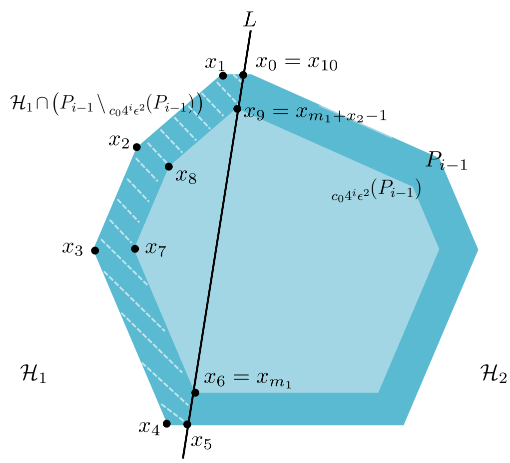

FIG 1. *Illustration of triangulation construction when $d = 2$.*

To see this, observe that the line L intersects $\mathrm{bd}(P_{i-1})$ at precisely two points; let $x_0 \in L \cap \mathrm{bd}(P_{i-1})$ denote the point that is larger in the lexicographic ordering (with respect to the standard Euclidean basis); see Figure 1. Let $m_1 \in \mathbb{N}$ denote the number of vertices of $\mathcal{H}_1 \cap P_{i-1}$. Now, for $j = 1, \ldots, m_1 - 1$, let $x_j \in \mathcal{H}_1 \cap \mathrm{bd}(P_{i-1})$ denote the vertex of the polyhedral convex set $\mathcal{H}_1 \cap P_{i-1}$ that is the unique neighbour of x_{j-1} not belonging to $\{x_0, \ldots, x_{j-1}\}$. Note here that x_{m_1-1} is the other point in $L \cap \mathrm{bd}(P_{i-1})$. Let x_{m_1} denote the closest point of $L \cap {}_{c_0 4^i \epsilon^2}(P_{i-1})$ to x_{m_1-1} (so the line segment joining x_{m_1-1} and x_{m_1} is a subset of L). Let $m_2 \in \mathbb{N}$ denote the number of vertices of $\mathcal{H}_1 \cap {}_{c_0 4^i \epsilon^2}(P_{i-1})$. For $j = 1, \ldots, m_2 - 1$, let $x_{m_1+j} \in \mathcal{H}_1 \cap \mathrm{bd}\big({}_{c_0 4^i \epsilon^2}(P_{i-1})\big)$ denote the vertex of the polyhedral convex set $\mathcal{H}_1 \cap {}_{c_0 4^i \epsilon^2}(P_{i-1})$ that is the unique neighbour of x_{m_1+j-1} not belonging to $\{x_{m_1}, \ldots, x_{m_1+j-1}\}$. Finally, let $x_{m_1+m_2} = x_0$. Let $0 = t_0 < t_1 < \ldots < t_{m_1+m_2} = 1$. The boundary of the set $\mathcal{H}_1 \cap \big(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})\big)$ is parametrised by the closed curve $\gamma : [0, 1] \to \mathbb{R}^2$ given by

$$\gamma(t) := \Big(\frac{t_{j+1} - t}{t_{j+1} - t_j}\Big) x_j + \Big(\frac{t - t_j}{t_{j+1} - t_j}\Big) x_{j+1}$$

for $t \in [t_j, t_{j+1}]$. In fact, we claim that γ is a simple closed curve. To see this, note that P_{i-1} and ${}_{c_0 4^i \epsilon^2}(P_{i-1})$ are polyhedral convex sets in $\mathbb{R}^2$, so their (disjoint) boundaries are simple closed curves; $\gamma(t) \in \mathrm{bd}(P_{i-1})$ for $t \in [0, t_{m_1-1}]$ and $\gamma(t) \in \mathrm{bd}\big({}_{c_0 4^i \epsilon^2}(P_{i-1})\big)$ for $t \in [t_{m_1}, t_{m_1+m_2-1}]$. Moreover, $\gamma(t)$

belongs to the interior of the line segment joining x_{m_1-1} and x_{m_1} (and hence to the interior of $P_{i-1} \setminus_{c_0 4^i \epsilon^2}(P_{i-1}))$ for $t \in (t_{m_1-1}, t_{m_1})$ and to the interior of the line segment joining $x_{m_1+m_2-1}$ and $x_{m_1+m_2}$ for $t \in (t_{m_1+m_2-1}, t_{m_1+m_2})$; these two line segments are themselves disjoint. This establishes that γ is a simple closed curve, and hence that $\mathcal{H}_1 \cap \big(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})\big)$ is a polygon. Note, incidentally, that our reason for introducing the line L was precisely to ensure this fact. We can therefore apply Theorems 1.4 and 1.8 of Devadoss and O'Rourke (2011) to conclude that there exist simplices $S_{i,1}, \ldots, S_{i,N_i}$ that triangulate $P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})$, where $N_i \leq 4C_2 c_0^{-1/2} 2^{-i} \epsilon^{-1}$.

For $i = 2, 3, \ldots, M$ and $j = 1, \ldots, N_i$, let

$$\alpha_{i,j} := \frac{2^{1/2}}{M^{1/2}} \biggl(\frac{\mu_2(S_{i,j})}{\mu_2(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1}))} \biggr)^{1/2}.$$

By Theorem 3, there exists a bracketing set $\{[\phi^L_{i,j,\ell}, \phi^U_{i,j,\ell}] : \ell = 1, \ldots, n_{i,j}\}$ for $\bar{\Phi}_1(S_{i,j})$, where $\log n_{i,j} \leq K_2^{**}\bigl(\frac{\mu_2^{1/2}(S_{i,j})}{\alpha_{i,j}\epsilon}\bigr)$, such that $L_2(\phi^U_{i,j,\ell}, \phi^L_{i,j,\ell}) \leq \alpha_{i,j}\epsilon$. Moreover, by another application of Theorem 3, there exists a bracketing set $\{[\phi^L_{M+1,r}, \phi^U_{M+1,r}] : r = 1, \ldots, n_{M+1}\}$ for $\bar{\Phi}_1(P_M)$, where $\log n_{M+1} \leq 8K_2^{**} C_2 c_0^{-1/2}\bigl(\frac{\mu_2^{1/2}(P_M)}{\epsilon}\bigr)$, such that $L_2(\phi^U_{M+1,r}, \phi^L_{M+1,r}) \leq \epsilon$. This last statement follows, because $2^{-M}\epsilon^{-1} \leq 8$.

We can therefore define a bracketing set for $\bar{\Phi}_1(D')$ as follows: first, for $i = 2, \ldots, M$ and $j = 1, \ldots, N_i$, let

$$\tilde{S}_{i,j} := S_{i,j} \setminus \biggl\{ \biggl(\bigcup_{k=2}^{i-1} \bigcup_{m=1}^{N_k} S_{k,m} \biggr) \bigcup \biggl(\bigcup_{m=1}^{j-1} S_{i,m} \biggr) \biggr\},$$
$$\tilde{P}_M := P_M \setminus \bigcup_{k=2}^{M} \bigcup_{m=1}^{N_k} S_{k,m}.$$

Now, for the array $\boldsymbol{\ell} = (\ell_{i,j})$ where $i \in \{2, \ldots, M\}$, $j \in \{1, \ldots, N_i\}$ and $\ell_{i,j} \in \{1, \ldots, n_{i,j}\}$, and for $r = 1, \ldots, n_{M+1}$, let

(1)
$$\psi^U_{\boldsymbol{\ell},r}(x) := \mathbb{1}_{\{x \in D' \setminus P_1\}} + \sum_{i=2}^{M} \sum_{j=1}^{N_i} \phi^U_{i,j,\ell_{i,j}}(x) \mathbb{1}_{\{x \in \tilde{S}_{i,j}\}} + \phi^U_{M+1,r}(x) \mathbb{1}_{\{x \in \tilde{P}_M\}},$$

(2)
$$\psi^L_{\boldsymbol{\ell},r}(x) := -\mathbb{1}_{\{x \in D' \setminus P_1\}} + \sum_{i=2}^{M} \sum_{j=1}^{N_i} \phi^L_{i,j,\ell_{i,j}}(x) \mathbb{1}_{\{x \in \tilde{S}_{i,j}\}} + \phi^L_{M+1,r}(x) \mathbb{1}_{\{x \in \tilde{P}_M\}},$$

for $x \in D'$. Observe that

$$\begin{aligned} L_2^2(\psi^U_{\boldsymbol{\ell},r}, \psi^L_{\boldsymbol{\ell},r}) &\leq 4\mu_2(D' \setminus P_1) + \sum_{i=2}^M \sum_{j=1}^{N_i} L_2^2(\phi^U_{i,j,\ell_{i,j}}, \phi^L_{i,j,\ell_{i,j}}) \\ &\qquad\qquad\qquad\qquad + L_2^2(\phi^U_{M+1,r}, \phi^L_{M+1,r}) \\ &\leq 4\mu_2(D' \setminus {}_{c_0\epsilon^2}D') + \epsilon^2 \sum_{i=2}^M \sum_{j=1}^{N_i} \alpha_{i,j}^2 + \epsilon^2 \leq 4\epsilon^2. \end{aligned}$$

Moreover, the logarithm of the cardinality of the bracketing set is

$$\begin{aligned} \sum_{i=2}^M \sum_{j=1}^{N_i} \log n_{i,j} + \log n_{M+1} &\leq K_2^{**} \sum_{i=2}^M \sum_{j=1}^{N_i} \frac{\mu_2^{1/2}(S_{i,j})}{\alpha_{i,j}\epsilon} + \frac{8K_2^{**}C_2\mu_2^{1/2}(P_M)}{\epsilon c_0^{1/2}} \\ &\leq \frac{K_2^{**}2^{-1/2}M^{1/2}}{\epsilon} \sum_{i=2}^M N_i \mu_2^{1/2}(P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}P_{i-1}) + \frac{16K_2^{**}C_2c_0^{-1/2}}{\epsilon} \\ &\leq \frac{K_2^{**}C_2c_0^{-1/2}M^{3/2}}{\epsilon} + \frac{16K_2^{**}C_2c_0^{-1/2}}{\epsilon} \leq \frac{32K_2^{**}C_2c_0^{-1/2}M^{3/2}}{\epsilon} \\ &\leq \frac{32K_2^{**}C_2c_0^{-1/2}}{\log^{3/2} 2} \epsilon^{-1} \log^{3/2}\Big(\frac{1}{4\epsilon}\Big). \end{aligned}$$

Defining $K_{1,2}^\circ := \frac{32K_2^{**}C_2}{\log^{3/2} 2}$, we have therefore proved that when $d = 2$,

$$\log N_{[]}\big(2\epsilon, \bar{\Phi}_1(D'), L_2\big) \leq K_{1,2}^\circ \epsilon^{-1} \log^{3/2}\Big(\frac{1}{4\epsilon}\Big)$$

for all $\epsilon \in (0, \epsilon_2^\circ]$.

The case $d = 3$: The proof is similar in spirit to the case $d = 2$, so we emphasise the points of difference, and give fewer details where the argument is essentially the same.

Set $\epsilon_3^\circ := 1/8$, and fix $\epsilon \in (0, \epsilon_3^\circ]$. The Bronshteyn–Ivanov result once again yields a polytope P_1 with ${}_{c_0\epsilon^2}D' \subseteq P_1 \subseteq ({}_{c_0\epsilon^2}D')^{c_0\epsilon^2]} \subseteq D'$ such that P_1 has at most $C_3 c_0^{-1}\epsilon^{-2}$ vertices and $\mathrm{Haus}(P_1, {}_{c_0\epsilon^2}D') \leq c_0\epsilon^2$. Applying the result of Bronshteyn and Ivanov (1975) recursively, with $M := \lfloor \log\big(\frac{1}{4\epsilon}\big)/\log 2 \rfloor$, for each $i = 2, 3, \ldots, M$, there exists a polytope ${}_{c_0 4^i\epsilon^2}(P_{i-1}) \subseteq P_i \subseteq P_{i-1}$ with at most $C_3 c_0^{-1} 4^{-i}\epsilon^{-2}$ vertices such that $\mathrm{Haus}\big(P_i, {}_{c_0 4^i \epsilon^2}(P_{i-1})\big) \leq c_0 4^i \epsilon^2$. Again we claim that P_M is a three-dimensional polytope, since

$$\mu_3(P_M) = \mu_3(D') - \mu_3(D' \setminus P_1) - \sum_{i=2}^M \mu_3(P_{i-1} \setminus P_i) > 0.$$

We can now appeal to the construction of Wang and Yang (2000) (cf. also Chazelle and Shouraboura (1995)), which yields, for each $i = 2, 3, \ldots, M$, simplices $S_{i,1}, \ldots, S_{i,N_i}$, where $N_i \leq 16C_3c_0^{-1}4^{-i}\epsilon^{-2}$ that triangulate $P_{i-1} \setminus {}_{c_0 4^i \epsilon^2}(P_{i-1})$. Set

$$\alpha_{i,j} := \biggl(\frac{2^{-(i-2)/2}}{\sum_{k=2}^M 2^{-k/2}}\biggr)^{1/2}\biggl(\frac{\mu_3(S_{i,j})}{\mu_3(P_{i-1} \setminus {}_{c_0 4^i\epsilon^2}(P_{i-1}))}\biggr)^{1/2}.$$

Applying Theorem 3 again, there exists a bracketing set $\{[\phi^L_{i,j,\ell}, \phi^U_{i,j,\ell}] : \ell = 1, \ldots, n_{i,j}\}$ for $\bar{\Phi}_1(S_{i,j})$, where $\log n_{i,j} \leq K_3^{**}\bigl(\frac{\mu_3^{1/2}(S_{i,j})}{\alpha_{i,j}\epsilon}\bigr)^{3/2}$, such that $L_2(\phi^U_{i,j,\ell}, \phi^L_{i,j,\ell}) \leq \alpha_{i,j}\epsilon$. Moreover, the same result also yields a bracketing set $\{[\phi^L_{M+1,r}, \phi^U_{M+1,r}] : r = 1, \ldots, n_{M+1}\}$ for $\bar{\Phi}_1(P_M)$, where $\log n_{M+1} \leq 64C_3c_0^{-1}K_3^{**}\bigl(\frac{\mu_3^{1/2}(P_M)}{\epsilon}\bigr)^{3/2}$, such that $L_2(\phi^U_{M+1,r}, \phi^L_{M+1,r}) \leq \epsilon$.

Defining brackets $\psi^U_{\boldsymbol{\ell},r}$ and $\psi^L_{\boldsymbol{\ell},r}$ as in (1) and (2), we find again that $L_2^2(\psi^U_{\boldsymbol{\ell},r}, \psi^L_{\boldsymbol{\ell},r}) \leq 4\epsilon^2$, where we have used the fact that

$$\sum_{i=2}^M \sum_{j=1}^{N_i} \alpha_{i,j}^2 = 2.$$

Moreover, the logarithm of the cardinality of the bracketing set is

$$\begin{aligned}
&\sum_{i=2}^M \sum_{j=1}^{N_i} \log n_{i,j} + \log n_{M+1} \\
&\quad \leq K_3^{**} \sum_{i=2}^M \sum_{j=1}^{N_i} \Bigl(\frac{\mu_3^{1/2}(S_{i,j})}{\alpha_{i,j}\epsilon}\Bigr)^{3/2} + 64K_3^{**}C_3c_0^{-1}\Bigl(\frac{\mu_3^{1/2}(P_M)}{\epsilon}\Bigr)^{3/2} \\
&\quad \leq \frac{K_3^{**}}{\epsilon^{3/2}} \sum_{i=2}^M \biggl(\frac{\sum_{k=2}^M 2^{-k/2}}{2^{-(i-2)/2}}\biggr)^{3/4} N_i \mu_3^{3/4}(P_{i-1} \setminus {}_{c_0 4^i\epsilon^2}P_{i-1}) + \frac{256K_3^{**}C_3c_0^{-1}}{\epsilon^{3/2}} \\
&\quad \leq \frac{4K_3^{**}C_3c_0^{-1}}{\epsilon^2} \sum_{i=2}^M 2^{-i/8} + \frac{256K_3^{**}C_3c_0^{-1}}{\epsilon^{3/2}} \leq \frac{512K_3^{**}C_3c_0^{-1}}{\epsilon^2}
\end{aligned}$$

Defining $K^\circ_{1,3} := 512K_3^{**}C_3c_0^{-1}$, we have therefore proved that when $d = 3$,

$$\log N_{[]}\bigl(2\epsilon, \bar{\Phi}_1(D'), L_2\bigr) \leq K^\circ_{1,3}\epsilon^{-2}$$

for all $\epsilon \in (0, \epsilon^\circ_3]$.

For the final steps, we deal with the cases $d = 1, 2, 3$ simultaneously. Let

$$\tilde{h}_d(\epsilon) := \begin{cases} \epsilon^{-1/2} & \text{when } d = 1 \\ \epsilon^{-1} \log_{++}^{3/2}(\frac{1}{4\epsilon}) & \text{when } d = 2 \\ \epsilon^{-2} & \text{when } d = 3. \end{cases}$$

(Thus $\tilde{h}_d$ is defined in almost the same way as h_d from the proof of Theorem 4 in the main text, except for the 4 inside the logarithm when $d = 2$.) Set $K_{2,d}^\circ := K_{1,d}^\circ \tilde{h}_d(\epsilon_d^\circ)/\tilde{h}_d\big(\mu_d^{1/2}(D')\big)$. Then, for $\epsilon \in (\epsilon_d^\circ, \mu_d^{1/2}(D')]$, we have

$$\begin{aligned} \log N_{[]}\big(2\epsilon, \bar{\Phi}_1(D'), L_2\big) &\leq \log N_{[]}\big(2\epsilon_d^\circ, \bar{\Phi}_1(D'), L_2\big) \leq K_{1,d}^\circ \tilde{h}_d(\epsilon_d^\circ) \\ &= K_{2,d}^\circ \tilde{h}_d\big(\mu_d^{1/2}(D')\big) \leq K_{2,d}^\circ \tilde{h}_d(\epsilon). \end{aligned}$$

On the other hand, for $\epsilon > \mu_d^{1/2}(D')$, it suffices to consider a single bracketing pair consisting of the constant functions $\psi^U(x) := 1$ and $\psi^L(x) := -1$ for $x \in D'$. Note that $L_2^2(\psi^U, \psi^L) = 4\mu_d(D')$, so that $\log N_{[]}\big(2\epsilon, \Phi_B(D'), L_2\big) = 0$ for $\epsilon > \mu_d^{1/2}(D')$. We conclude that when D' is a d-dimensional closed, convex subset of $\mathbb{R}^d$ with $d^{-1}\bar{B}_d(0,1) \subseteq D' \subseteq \bar{B}_d(0,1)$,

$$\log N_{[]}\big(2\epsilon, \bar{\Phi}_1(D'), L_2\big) \leq K_{2,d}^\circ \tilde{h}_d(\epsilon)$$

for all $\epsilon > 0$.

Finally, we show how to transform the brackets to the original domain D and rescale their ranges to $[-B, B]$. Recall that $D' = AD + b$. Simplifying our notation from before, given $\epsilon > 0$, we have shown that we can define a bracketing set $\{[\psi_j^L, \psi_j^U] : j = 1, \ldots, N\}$ for $\bar{\Phi}_1(D')$ with $L_2^2(\psi_j^U, \psi_j^L) \leq 4\epsilon^2 |\det A|/B^2$ and $\log N \leq K_{2,d}^\circ \tilde{h}_d(\epsilon|\det A|^{1/2}/B)$. We now define transformed brackets for $\bar{\Phi}_B(D)$ by

$$\tilde{\psi}_j^U(z) := B\psi_j^U(Az + b) \quad \text{and} \quad \tilde{\psi}_j^L(z) := B\psi_j^L(Az + b).$$

Then

$$\begin{aligned} L_2^2(\tilde{\psi}_j^U, \tilde{\psi}_j^L) &= B^2 \int_D \{\psi_j^U(Az + b) - \psi_j^L(Az + b)\}^2 \, d\mu_d(z) \\ &= \frac{B^2}{|\det A|} L_2^2(\psi_j^U, \psi_j^L) \leq 4\epsilon^2. \end{aligned}$$

Now

$$|\det A| = \frac{\mu_d(AD + b)}{\mu_d(D)} \geq \frac{\mu_d(d^{-1}\bar{B}_d(0,1))}{\mu_d(D)} = \frac{d^{-d}\pi^{d/2}}{\Gamma(1 + d/2)\mu_d(D)}.$$

It is convenient for the case $d = 2$ to note that

$$\tilde{h}_2\biggl(\frac{\epsilon|\det A|^{1/2}}{B}\biggr) \leq \tilde{h}_2\biggl(\frac{\epsilon\pi^{1/2}}{2B\mu_2^{1/2}(D)}\biggr) \leq \frac{2}{\pi^{1/2}}h_2\biggl(\frac{\epsilon}{B\mu_2^{1/2}(D)}\biggr).$$

The final result therefore follows, taking $K_1^\circ := K_{2,1}^\circ$, $K_2^\circ := \frac{2}{\pi^{1/2}}K_{2,2}^\circ$ and $K_3^\circ := \frac{81}{4\pi}K_{2,3}^\circ$. $\square$

3. Auxiliary result for the proof of Theorem 5 in the main text.

THEOREM 5 (van de Geer (2000), Theorem 7.4). *Let $\mathcal{F}$ denote a class of (Lebesgue) densities on $\mathbb{R}^d$, let $X_1, X_2, \ldots$ be independent and identically distributed with density $f_0 \in \mathcal{F}$, and let $\hat{f}_n$ denote a maximum likelihood estimator of f_0 based on $X_1, \ldots, X_n$. Write $\bar{\mathcal{F}} := \bigl\{\bigl(\frac{f+f_0}{2}\bigr) : f \in \mathcal{F}\bigr\}$, and let*

$$J_{[]}(\delta, \bar{\mathcal{F}}, h) := \max\biggl\{\int_{\delta^2/2^{13}}^{\delta} \sqrt{\log N_{[]}(u, \bar{\mathcal{F}}, h)}\, du\,,\, \delta\biggr\}.$$

If $\Psi(\delta) \geq J_{[]}(\delta, \bar{\mathcal{F}}, h)$ is such that $\Psi(\delta)/\delta^2$ a non-increasing function of δ and (δ_n) is such that $2^{-16}n^{1/2}\delta_n^2 \geq J_{[]}(\delta_n, \bar{\mathcal{F}}, h)$, then for all $t \geq \delta_n$,

$$\mathbb{P}_{f_0}\{h(\hat{f}_n, f_0) \geq 2^{1/2}t\} \leq 2^{13/2}\sum_{s=0}^{\infty}\exp\biggl(-\frac{2^{2s}nt^2}{2^{27}}\biggr).$$

References.

Ball, K. (1997). An elementary introduction to modern convex geometry. In *Flavors of Geometry* (ed. S. Levy), pp. 1–58, MSRI publications.

Bronshteyn, E. M. and Ivanov, L. D. (1975). The approximation of convex sets by polyhedra. *Siberian Math. J.*, **16**, 852–853.

Brunel, V.-E. (2013). Adaptive estimation of convex polytopes and convex sets from noisy data. *Elect. J. Statist.*, **7**, 1301–1327.

Brunel, V.-E. (2016). Adaptive estimation of convex and polytopal density support. *Probab. Theory. Rel. Fields*, **164**, 1–16.

Chazelle, B. and Shouraboura, N. (1995). Bounds on the size of tetrahedralizations. *Discrete Comput. Geom.*, **14**, 429–444.

Devadoss, S. L. and O'Rourke, J. (2011). *Discrete and Computational Geometry.* Princeton University Press, Princeton, NJ.

Devroye, L. and Lugosi, G. (2001) *Combinatorial Methods in Density Estimation.* Springer-Verlag, New York.

Doss, C. and Wellner, J. A. (2016). Global rates of convergence of the MLEs of log-concave and S-concave densities. *Ann. Statist.*, to appear.

Dudley, R. M. (1999). *Uniform Central Limit Theorems.* Cambridge University Press, Cambridge.

Gao, F. and Wellner, J. A. (2015). Entropy of convex functions on $\mathbb{R}^d$. Available at `http://arxiv.org/abs/1502.01752`.

Guntuboyina, A. and Sen, B. (2013). Covering numbers for convex functions. *IEEE Transactions on Information Theory*, **59**, 1957–1965.

John, F. (1948). Extremum problems with inequalities as subsidiary conditions. In *Studies and essays presented to R. Courant on his 60th birthday*, pp. 187–204, Interscience, New York.

Kim, A. K. H. and Samworth, R. J. (2015). Global rates of convergence in log-concave density estimation. Available at `http://arxiv.org/abs/1404.2298v2`.

Kim, A. K. H. and Samworth, R. J. (2016). Global rates of convergence in log-concave density estimation. *Ann. Statist., to appear.*

Korostelev, A. P. and Tsybakov, A. B. (1993). *Minimax Theory of Image Reconstruction.* Springer-Verlag, New York.

Mammen, E. and Tsybakov, A. B. (1995). Asymptotical minimax recovery of sets with smooth boundaries. *Ann. Statist.*, **23**, 502–524.

Rockafellar, R. T. (1997). *Convex Analysis.* Princeton University Press, Princeton.

Schneider, R. (2014) *Convex bodies: the Brunn–Minkowski theory*, 2nd edition. Cambridge University Press, Cambridge.

van de Geer, S. (2000). *Empirical Processes in M-Estimation.* Cambridge University Press, Cambridge.

Wang, C. A. and Yang, B. (2000). Tetrahedralization of two nested convex polyhedra. In *Lecture Notes in Computer Science* (eds. D.-Z. Du et al.), **1858**, pp. 291–298.

Yang, Y. and Barron, A. (1999). Information-theoretic determination of minimax rates of convergence. *Ann. Statist.*, **27**, 1564–1599.

Statistical Laboratory
Wilberforce Road
Cambridge
CB3 0WB
United Kingdom
E-mail: r.samworth@statslab.cam.ac.uk
E-mail: a.kim@statslab.cam.ac.uk
URL: http://www.statslab.cam.ac.uk/~rjs57
URL: http://sites.google.com/site/kyoungheearlene/home