A. Proofs of the main results

A.1. Proof of Theorem 1

Using Assumptions 1 and 2 and the fact that $\| \cdot \|_2 \leq \| \cdot \|_s$, since $s \leq 2$, we have that for any $x \in [0, A]^d$, 

$$
\left| \mathbb{E} [\hat{f}(x)] - f(x) \right| \leq c'' A^d d \int_{\mathbb{R}^d} K \left( \frac{y-x}{h} \right) \|y-x\|^s dy
$$

$$
= c'' h^d \sum_{i=1}^{d} \int_{\mathbb{R}^d} |y_i - x_i|^s \prod_{j=1}^{d} K_0 \left( \frac{y_j - x_j}{h} \right) dy_j
$$

$$
= c'' h^d \sum_{i=1}^{d} \int_{\mathbb{R}} K_0 \left( \frac{y_i - x_i}{h} \right) |y_i - x_i|^s dy_i \prod_{j \neq i} \int_{\mathbb{R}} K_0 \left( \frac{y_j - x_j}{h} \right) dy_j
$$

$$
= c'' h^d \sum_{i=1}^{d} \int_{\mathbb{R}} K_0 \left( \frac{y_i - x_i}{h} \right) |y_i - x_i|^s dy_i = c'' h^d \int_{\mathbb{R}} K_0(0) |u|^s du.
$$

Since $0 < s \leq 2$, we have that $s \log |u| \leq \max(0, 2 \log |u|)$, which means that $|u|^s \leq \max(1, |u|^2) \leq 1 + |u|^2$. Therefore, we can write

$$
\int_{\mathbb{R}} K_0(0) |u|^s du \leq \int_{\mathbb{R}} K_0(0) (1 + |u|^2) du \leq 1 + C'.
$$

Thus,

$$
\left| \mathbb{E} [\hat{f}(x)] - f(x) \right| \leq c'' (1 + C') d h^s. \tag{4}
$$

Let $0 < \delta < \frac{1}{2}$. Note that since $K_0$ is non-negative and bounded by $C$, we have that $K \leq C^d$. For $y \in \mathbb{R}^d$, let us write $Y_i = A^d f(X_i) K \left( \frac{y-X_i}{h} \right)$, where $X_i \sim \mathcal{U}([0, A]^d)$. This implies that $|Y_i| \leq c(CA)^d$. Moreover, since $f \leq c$, we have

$$
\mathbb{V}(Y_i) \leq \mathbb{E} [Y_i^2] = \frac{A^{2d}}{A^d} \int_{[0, A]^d} K^2 \left( \frac{y-x}{h} \right) f^2(x) dx
$$

$$
\leq c^2 A^d \int_{\mathbb{R}^d} K^2 \left( \frac{y-x}{h} \right) dx
$$

$$
\leq c^2 A^d h^d \int_{\mathbb{R}^d} K^2(u) du
$$

$$
\leq c^2 C^d A^d h^d \int_{\mathbb{R}^d} K(u) du = c^2 C^d A^d h^d.
$$

Therefore, by Bernstein’s inequality, for any $x \in [0, A]^d$, we know that with probability larger than $1 - \delta$

$$
\left| \mathbb{E} [\hat{f}(x)] - f(x) \right| = \left| \frac{1}{Nh^d} \sum_{k=0}^{n} \left( f(X_i) K \left( \frac{X_i-x}{h} \right) - \mathbb{E} \left[ f(X_i) K \left( \frac{X_i-x}{h} \right) \right] \right) \right|
$$

$$
\leq \frac{1}{Nh^d} \left( 2c \sqrt{C^d A^d h^d N \log(1/\delta) + 2c A^d C^d \log(1/\delta)} \right)
$$

$$
\leq 2c \sqrt{C^d A^d \log(1/\delta) N h^d + 2c A^d C^d \log(1/\delta) N h^d}, \tag{5}
$$

for $n$ large enough with respect to $\delta$. 

By Equations 4 and 5, we have for any $x \in [0, A]^d$, with probability larger than $1 - \delta$, that
\[
\left| \tilde{f}(x) - f(x) \right| \leq 2c\sqrt{C''d} \frac{\log(1/\delta)}{Nh^d} + 2cC''d \frac{\log(1/\delta)}{Nh^d} + c''(1 + C')dh^s.
\] (6)

Therefore, for $h_{=}^{\text{def}} = \frac{1}{\left( \log(NA/\delta) \right) \frac{1}{\log(1/\delta)} N h^d}$, we get that with probability larger than $1 - \delta$,
\[
\left| \tilde{f}(x) - f(x) \right| \leq \left( 2c\sqrt{C''d} + c''(1 + C')d \right) \frac{\log(NA/\delta)}{N} \frac{1}{\log(1/\delta)} + 2cC'd \frac{\log(NA/\delta)}{N} \frac{1}{\log(1/\delta)} .
\] (7)

Let $X$ be a $1/N^v$ covering set in $\| \cdot \|_2$ norm and of minimal cardinality of the hypercube $[0, A]^d$. Its cardinality is at most $(AN^v d/v)^d \leq (ANd)^d(v+1)$, since the covering number of $[0, A]^d$ with small hypercubes of side $N^{-v}/\sqrt{d}$ is smaller than $\left( \frac{A}{\sqrt{d}} \right)^d = (AN^v d/v)^d$, and hypercubes of side $N^{-v}/\sqrt{d}$ are contained in $\ell_2$ balls of diameter $N^{-v}$. By a union bound and Equation 7, it holds that with probability larger than $1 - \delta$, for any $x \in X$, we have
\[
\left| \tilde{f}(x) - f(x) \right| \leq 4d(v + 1) \left( 2c\sqrt{C''d} + c''(1 + C')d \right) \frac{\log(NAd/\delta)}{N} \frac{1}{\log(1/\delta)} + 4cd(v + 1)Ad^d \frac{\log(NAd/\delta)}{N} .
\]

Let $\xi$ be the event of probability larger than $1 - \delta$ where this is satisfied. Let $y \in [0, A]^d$. Then, there exists $x \in X$ such that $\|x - y\|_2 \leq 1/N^v$. Since $K_0$ is $\varepsilon$-Hölder and since $f$ is bounded by $c$, we have that
\[
\left| \tilde{f}(x) - \tilde{f}(y) \right| \leq cN(C''d) \frac{\sqrt{d}}{\log(1/\delta) N^{dev}} \leq (C''d) c \frac{1}{\log(1/\delta)} N^{1+\frac{d}{2n}N^{dev}},
\]
and for $v > 3/\varepsilon \log(1 + 1/(C''c))$, we get
\[
\left| \tilde{f}(x) - \tilde{f}(y) \right| \leq N^{-1}.
\]

In the same way, by Assumption 3, we get also that for $v \geq 2 \log(1 + 1/(c'' + c))/\min(1, s)$,
\[
\left| f(x) - f(y) \right| \leq c''N^{-v} + cN^{-vs} \leq N^{-1}.
\]

This implies that
\[
\left| \tilde{f}(y) - f(y) \right| \leq \left| \tilde{f}(x) - f(x) \right| + 2/N,
\]
which means that on $\xi$, for any $x \in [0, A]^d$ we have
\[
\left| \tilde{f}(x) - f(x) \right| \leq 4d(v + 1) \left( 2\sqrt{cA^dC} + c''(1 + C')d \right) \frac{\log(NAd/\delta)}{N} \frac{1}{\log(1/\delta)} + 4cd(v + 1)Ad^d \frac{\log(NAd/\delta)}{N} \frac{1}{\log(1/\delta)} + 2/N,
\]
where $v = \log \left( 1 + \frac{1}{\varepsilon/(c'' + c)} \right) \frac{\sqrt{d}}{\min(1, s)} + \frac{3}{\varepsilon} \log \left( 1 + \frac{1}{\varepsilon/(c'' + c)} \right)$. Therefore, we get that for any $x \in [0, A]^d$,
\[
\left| \tilde{f}(x) - f(x) \right|
\leq 4d(v + 1) \left( 2\sqrt{cA^dC} + c''(1 + C')d \right) \frac{\log(NAd/\delta)}{N} \frac{1}{\log(1/\delta)} + 4cd(v + 1)Ad^d \frac{\log(NAd/\delta)}{N} \frac{1}{\log(1/\delta)} + 2/N
\leq \left( 2\sqrt{cA^dC} + c''(1 + C')d + c(AC)d \right) \frac{\log(NAd/\delta)}{N} \frac{1}{\log(1/\delta)} \cdot 8d(v + 1) \frac{\log(NAd/\delta)}{N} \frac{1}{\log(1/\delta)} ,
\]
\[
\leq H_0 \left( \frac{\log(NAd/\delta)}{N} \right) \frac{1}{\log(1/\delta)} ,
\]
where $v = \log \left( 1 + \frac{1}{\varepsilon/(c'' + c)} \right) \frac{\sqrt{d}}{\min(1, s)} + \frac{3}{\varepsilon} \log \left( 1 + \frac{1}{\varepsilon/(c'' + c)} \right)$ and $H_0$ a constant that depends on $d, v, c''$, $C, C'$, and $A$. 

Pliable rejection sampling
B. Extension to densities with unbounded support

In this part of the appendix, we extend our method to densities that do not have a compact support. For the sake of clarity, we assume that $f$ is normalized here. In fact, we have already shown how to deal with the unnormalized case in the Section 3, where the normalizaiton constant is estimated by a Monte-Carlo sum over the same samples used to estimate $f$.

Assumption 3 (Assumption on the density). The density $f$, defined on $\mathbb{R}^d$, is sub-Gaussian, i.e., there exist constants $c, c' > 0$ such that the density $f$ satisfies for any $x \in \mathbb{R}^d$

$$ f(x) \leq c \exp \left( -c' \|x\|^2 \right). $$

Moreover, $f$ can be uniformly expanded by a Taylor expansion in any point up to degree $s$ for some $0 < s \leq 2$, i.e., there exists $c'' > 0$ such that, for any $x \in \mathbb{R}^d$, and for any $u \in \mathbb{R}^d$, we have

$$ |f(x + u) - f(x) - \langle \nabla f(x), u \rangle 1 \{s > 1\}| \leq c'' \|u\|_2. $$

The above assumption means that the tails of $f$ are sub-Gaussian, and also that $f$ is in a Hölder ball of smoothness $s$. Note that a bounded function $f$ with a compact support in $\mathbb{R}^d$ is sub-Gaussian. The fact that $f$ is in a Hölder ball of smoothness $s$ is also not very restrictive, in particular for a small $s$.

B.1. Uniform bounds for kernel density estimation

Let $X_1, \ldots, X_N$ be $N$ points generated by $f$. Let us define for $h = h_s(\delta) = \left( \frac{\log(N/\delta)}{N} \right)^{\frac{1}{s+1}}$, $\hat{f}(x) = \frac{1}{Nh^{d}} \sum_{k=1}^{N} K \left( \frac{X_k - x}{h} \right)$, and $\tilde{f}$ be such that

$$ \tilde{f}(x) = \hat{f}(x) 1 \{ \|x\|_2 \leq \log(N) \}. \quad (8) $$

Theorem 3. Assume that Assumptions 2 and 3 hold with $0 < s \leq 2$, $C, C', c''', c', c'' > 0$, and $\varepsilon > 0$. The estimate $\tilde{f}$ is such that with probability larger than $1 - \delta$, for any point $x \in \mathbb{R}^d$,

$$ \left| \tilde{f}(x) - f(x) \right| \leq 8d(v + 1)2\sqrt{cC' + c''(1 + C'')d + C' + C''} \left( \frac{\log(N/\delta)}{N} \right)^{\frac{1}{s+3}} + c \exp \left( -c' \|x\|^2 \right) \mathbb{1}_{\|x\|_2 \geq \log(N)} $n$,

where $v = \log \left( 1 + \frac{1}{c'''} \right) - \frac{2}{\min(1,c')} \log \left( 1 + \frac{1}{c'''} \right)$, and $H_1$ is a constant that depends on $d, v, c', C', C''$.

Theorem 3 provides a uniform bound on the error of $\tilde{f}$ on a large centered ball of radius $\log n$ denoted by $B_d(\log n)$, and bounds the far fluctuations by an upper bound on $f$ itself. Note that the previous bound implies in particular that

$$ \left| \tilde{f}(x) - f(x) \right| \leq O \left( \left( \frac{\log(N/\delta)}{N} \right)^{\frac{1}{s+3}} \right). $$

B.2. Proof of Theorem 3

By Assumptions 2 and 3, similarly to the starting point of proof of Theorem 1 (see Appendix A), we have that for any $x \in \mathbb{R}^d$

$$ \mathbb{E} \left[ \tilde{f}(x) \right] - f(x) \leq c''(1 + C')dh^s. \quad (9) $$
Let $0 < \delta < \frac{1}{2}$. Note that since $K_0$ is non-negative and bounded by $C$, we have that $K \leq C^d$. For $y \in \mathbb{R}^d$, let us write $Y_i = K \left( \frac{y - X_i}{h} \right)$, where $X_i \sim f$. This implies that $|Y_i| \leq C^d$. Moreover, since $f(x) \leq c \exp \left(-c'\|x\|^2\right) \leq c$, we have

$$\nabla(Y_i) \leq \int_{\mathbb{R}^d} K^2 \left( \frac{y - x}{h} \right) f(x) dx \leq c \int_{\mathbb{R}^d} K^2 \left( \frac{y - x}{h} \right) dx \leq c h^d \int_{\mathbb{R}^d} K^2(u) du \leq c C^d h^d \int_{\mathbb{R}^d} K(u) du = c C^d h^d.$$

Therefore, by Bernstein’s inequality, for any $x \in \mathbb{R}^d$, we know that with probability larger than $1 - \delta$

$$| \mathbb{E} \left[ \hat{f}(x) \right] - f(x) | \leq \frac{1}{N h^d} \sum_{k=0}^{n} \left( K \left( \frac{X_i - x}{h} \right) - \mathbb{E} \left[ K \left( \frac{X_i - x}{h} \right) \right] \right) \leq \frac{1}{N h^d} \left( 2 \sqrt{c C^d h^d N \log(1/\delta)} + 2 C^d \log(1/\delta) \right) \leq \frac{2 \sqrt{c C^d \log(1/\delta)} + 2 C^d \log(1/\delta)}{N h^d} = \frac{2 \sqrt{c C^d \log(1/\delta)} + 2 C^d \log(1/\delta)}{N h^d},$$

(10)

for $n$ large enough with respect to $\delta$.

By Equations 9 and 10, we thus know that for any $x \in \mathbb{R}^d$, with probability larger than $1 - \delta$,

$$| \hat{f}(x) - f(x) | \leq 2 \sqrt{c C^d \log(1/\delta)} + 2 C^d \log(1/\delta) + c'(1 + C') dh^s.$$  

(11)

Therefore, for $h = h_s(\delta) = \left( \frac{\log(N/\delta)}{N} \right)^{\frac{1}{d+s}}$, we get that with probability larger than $1 - \delta$,

$$| \hat{f}(x) - f(x) | \leq \left( 4 \sqrt{c C^d + c'(1 + C')d} \frac{\log(N/\delta)}{N} \right)^{\frac{1}{d+s}} + 4 C^d \left( \frac{\log(N/\delta)}{N} \right)^{\frac{1}{d+s}}.$$  

(12)

Now, since the $\Psi_2$ norm of $f$ is bounded by $c'$, we know that for any $x \in \mathbb{R}^d$,

$$f(x) \leq c \exp \left(-c'\|x\|^2\right).$$

This implies in particular that for any $x$ such that $\|x\|_2 \geq \log(N)$, we have

$$f(x) \leq c \exp \left(-c'\|x\|^2\right) 1_{\|x\|_2 \geq \log(N)}.$$  

(13)

Let $\mathcal{X}$ be a $1/N^v$ covering set in $\|\cdot\|_2$ norm and of minimal cardinality of the ball of $\mathbb{R}^d$ of center $0$ and radius $\log N$ that we will denote by $B_d(\log N)$. Its cardinality is at most $(2N^v \log N)^d \leq N^{d(v+1)}$ (by a similar reasoning as in the proof of Thm 1). By a union bound and Equation 12, it holds that with probability larger than $1 - \delta$, for any $x \in \mathcal{X}$, we have

$$| \hat{f}(x) - f(x) | \leq 4(d+1) \left( 4 \sqrt{c C^d + c'(1 + C')d} \frac{\log(N/\delta)}{N} \right)^{\frac{1}{d+s}} + 4(d+1) C^d \left( \frac{\log(N/\delta)}{N} \right)^{\frac{1}{d+s}}.$$  

(14)

Let $\xi$ be the event of probability larger than $1 - \delta$ where this is satisfied. Let $y \in B_d(\log N)$. Then, there exists $x \in \mathcal{X}$ such that $\|x - y\|_2 \leq 1/N^v$. Since $K_0$ is $\varepsilon$-Hölder, we have that

$$| \hat{f}(x) - \hat{f}(y) | \leq N(C'')^d \frac{c^d}{h_s(\delta)^d N^{dv}} \leq (C'')^d c^d N^{1 + \frac{d}{d+s}}.$$  

Let $\xi$ be the event of probability larger than $1 - \delta$ where this is satisfied. Let $y \in B_d(\log N)$. Then, there exists $x \in \mathcal{X}$ such that $\|x - y\|_2 \leq 1/N^v$. Since $K_0$ is $\varepsilon$-Hölder, we have that

$$| \hat{f}(x) - \hat{f}(y) | \leq N(C'')^d \frac{c^d}{h_s(\delta)^d N^{dv}} \leq (C'')^d c^d N^{1 + \frac{d}{d+s}}.$$  

Plausible rejection sampling

and for \( v > 3/\varepsilon \log(1 + 1/(C''c)) \), we get

\[
\left| \hat{f}(x) - f(y) \right| \leq N^{-1}.
\]

In the same way, by Assumption 3, we also get that for \( v \geq 2 \log(1 + 1/(c'' + e))/(\min(1,s)) \),

\[
|f(x) - f(y)| \leq c''N^{-v} + cN^{-s} \leq N^{-1}.
\]

This implies that

\[
\left| \hat{f}(y) - f(y) \right| \leq \left| \hat{f}(x) - f(x) \right| + 2/N,
\]

which means that on \( \xi \), for any \( x \in B_d(\log N) \) we have

\[
\left| \hat{f}(x) - f(x) \right| \leq 4d(v + 1) \left( 2\sqrt{cC} + c''(1 + C')d \right) \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + 4d(v + 1)C^d \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + 2/N,
\]

where \( v = \log \left( 1 + \frac{1}{c'' + c} \right) \frac{2}{\min(1,s)} + \frac{2}{\varepsilon} \log \left( 1 + \frac{1}{c'' + c} \right) \). Combining this with the definition of \( \hat{f} \), and Equation 13, we get that for any \( x \in \mathbb{R}^d \),

\[
\left| \hat{f}(x) - f(x) \right| \\
\leq 4d(v + 1) \left( 2\sqrt{cC} + c''(1 + C')d \right) \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + 4d(v + 1)C^d \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + 2/N + f(x)1_{\|x\|_2 \geq \log N}
\leq 8d(v + 1) \left( 2\sqrt{cC} + c''(1 + C')d \right) \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + 4d(v + 1)C^d \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + f(x)1_{\|x\|_2 \geq \log N}
\leq 8d(v + 1) \left( 2\sqrt{cC} + c''(1 + C')d \right) \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + 4d(v + 1)C^d \left( \frac{\log(N/\delta)}{N} \right)^{\frac{2d}{2d}} + c \exp \left( -c'' \|x\|_2^2 \right) 1_{\|x\|_2 \geq \log N},
\]

where \( v = \log \left( 1 + \frac{1}{c'' + c} \right) \frac{2}{\min(1,s)} + \frac{2}{\varepsilon} \log \left( 1 + \frac{1}{c'' + c} \right) \).

B.3. Extended plausible rejection sampling

Our modified algorithm, extended plausible rejection sampling (EPRS), aims at sampling as many i.i.d. points distributed according to \( f \) as possible with a fixed budget of evaluations of \( f \). It consists of three main steps: (i) a first rejection sampling step where it generates \( \tilde{N} \) initial samples from \( f \) by rejection sampling using an initial proposal. Then, (ii) EPRS uses these samples to estimate \( f \) by a kernel density estimation method. Finally, (iii) EPRS uses the newly obtained estimate, plus a uniform bound on it, as a new extended plausible proposal for rejection sampling. Since this plausible proposal is closer to the target density than the initial proposal, the rejection sampling will reject significantly fewer points by using it. Our EPRS method is described as Algorithm 2.

As mentioned, our method makes use of an initial proposal density \( g \) that must satisfy the following properties with respect to the target density.

Assumption 4 (Assumption on the initial proposal). Let \( M > 0 \). We have

\[
f \leq Mg.
\]

Furthermore, the density \( g \) is sub-Gaussian, i.e., there exist constants \( a, a' > 0 \) such that the density \( g \) defined on \( \mathbb{R}^d \) satisfies

\[
g(x) \leq a \exp \left( -a' \|x\|_2^2 \right).
\]

We set the following constants:

\[
T_s \overset{\text{def}}{=} n^{\frac{2d+4}{2d+2}}, \quad \text{and} \quad \tilde{N} \overset{\text{def}}{=} T_s/M - 2\sqrt{T_s \log(1/\delta)}.
\]
Pliable rejection sampling

$\bar{N}$ is actually a high probability lower bound of $\bar{N}$ given by our algorithm (it is the number of samples obtained by the initial rejection sampling step). $T_s$ is the number of calls needed for the first estimation step that will optimize the number of accepted samples in the second step.

We also define

$$r_N \overset{\text{def}}{=} \mathcal{V}_n H_E \left( \frac{\log(N/\delta)}{N} \right)^{\frac{1}{2\delta}}, \quad \text{and}$$

$$\bar{g}_N \overset{\text{def}}{=} \int_{B_d(\log(N))^C} g(x)dx,$$

where $\mathcal{V}_n = \mathcal{V}(B_d(\log(n)))$ is the volume of $B_d(\log n)$ the centered ball in $\mathbb{R}^d$ of radius $\log n$ and where $H$ is a parameter of the algorithm.

Our method samples most of the samples by rejection sampling according to a pliable proposal that is defined as

$$\tilde{g}^* \overset{\text{def}}{=} \frac{1}{1 + r_N + M\bar{g}_N} \left( \hat{f} + r_N \mathcal{U}_{B_d(\log(n))} + Mg1_{\|x\|_2 \geq \log(N)} \right),$$

where $\mathcal{U}_{B_d(\log(n))}$ is the uniform distribution on $B_d(\log(n))$, and $\hat{f}$ the estimate of $f$ defined in (8).

**Algorithm 2** Extended pliable rejection sampling (EPRS)

**Parameters:** $s$, $n$, $\delta$, $H$, $g$, and $M$.

**Initial sampling**

- Draw $T_s$ samples at random according to $g$, and evaluate $f$ on them

**Estimation of $f$**

- Perform rejection sampling on the samples using $M$ as the constant
- Obtain this way $\bar{N}$ samples from $f$
- Estimate $f$ by $\hat{f}$ on these $\bar{N}$ samples (Section B.1)

**Generating the sample**

- Sample $n - T_s$ samples from the pliable proposal $\tilde{g}^*$
- Perform rejection sampling on these samples using $1 + r_N/\mathcal{V}_n + M\bar{g}_N$ as a constant
- Obtain this way $\bar{n}$ samples from $f$

**Output:** Return the $\bar{n}$ samples.

**Theorem 4.** Assume that Assumptions 2, 3, and 4 hold with $0 < s \leq 2$, $g$, $M > 0$, and that $H_E > 0$ is an upper bound on the constant $H_1$ defined in Theorem 3 (applied to $f$ and $\hat{f}$). The number $\bar{n}$ of samples generated in an i.i.d. way according to $f$ in this way is such that for $n$ large enough, with probability larger than $1 - \delta$,

$$\bar{n} \geq n \left[ 1 - \mathcal{O} \left( \log(\delta/n)^{d+1} n^{-\frac{1}{d+2}} \right) \right].$$

**B.4. Proof of Theorem 4**

By definition of $\bar{g}_N$ and since the $g$ is sub-Gaussian by Assumption 3, we have that

$$\bar{g}_N = \int_{B_d(\log(N))^C} g(x)dx \leq \int_{B_d(\log(N))^C} a \exp(-a'\|x\|_2^2)dx \leq a(N)^{-da'(\log \bar{N})/4} \leq \bar{N}^{-1},$$

for $n$ (and thus $\bar{N}$) large enough.

By definition of rejection sampling, the probability of accepting a sample is exactly $1/M$, where $M$ is the upper bound used in the rejection sampling (provided that $f \leq Mg$). Therefore, and since the first rejection sampling step uses $T_s$ samples, $\bar{N}$ is a sum of $T_s$ independent Bernoulli random variables of parameter $1/M$. Thus, we have with probability larger than $1 - \delta$ that

$$\bar{N} \geq T_s/M - 2\sqrt{T_s \log(1/\delta)} \overset{\text{def}}{=} \bar{N}. \quad (15)$$
Pliable rejection sampling

Let us write $\xi'$ for the event where this happens. On $\xi'$, we have by Theorem 3 (end of the proof) that with probability larger than $1 - \delta$, for any $x \in \mathbb{R}^d$

$$|\tilde{f}(x) - f(x)| \leq \frac{r_{\tilde{N}}}{V_{n}} + f_{1}\mathbb{1}_{\|x\|_2 \geq \log \tilde{N}} \leq \frac{r_{\tilde{N}}}{V_{n}} + M g_{1}\mathbb{1}_{\|x\|_2 \geq \log \tilde{N}}.$$ 

Let $\xi$ be the intersection of $\xi'$ and the event where Equation 15 holds. It has probability larger than $1 - 2\delta$. On $\xi$, we thus have that

$$\tilde{f} + r_{\tilde{N}}M B_{d}(\log n) + M g_{1}\mathbb{1}_{\|x\|_2 \geq \log \tilde{N}} \geq f.$$ 

Therefore, the rejection sampling is going to provide samples that are i.i.d. according to $f$, and $\tilde{n}$ will be a sum of Bernoulli random variables of parameter $\frac{1}{1 + r_{\tilde{N}} + M \tilde{g}_{\tilde{N}}}$. We thus have that on $\xi$, with probability larger than $1 - \delta$,

$$\tilde{n} \geq (n - T_{s}) \frac{1}{1 + r_{\tilde{N}} + M \tilde{g}_{\tilde{N}}} - 2\sqrt{n \log(1/\delta)}.$$ 

This implies, together with the definition of $r_{\tilde{N}}$ and the upper bound on $\tilde{g}_{\tilde{N}}$, that $\tilde{n}$ is with probability larger than $1 - 3\delta$ lower bounded as

$$\tilde{n} \geq (n - T_{s}) \left(1 - \pi^d \log (\tilde{N})^d H \left(\frac{\log(\tilde{N}/\delta)}{N}\right)^{\frac{1}{2 + d}} - M \tilde{N}^{-1}\right) - 2\sqrt{n \log(1/\delta)}$$

$$\geq n \left[1 - \frac{T_{s}}{n} - \pi^d \log (n)^d H \left(\frac{\log(n/\delta)}{T_{s}/M - 2\sqrt{T_{s} \log(1/\delta)}}\right)^{\frac{1}{2 + d}} - \frac{M}{T_{s}/M - 2\sqrt{T_{s} \log(1/\delta)}} - 4\sqrt{\frac{\log(1/\delta)}{n}}\right].$$

Since

$$T_{s} = n^{\frac{2 + d}{\pi^d}},$$

we have that with probability larger than $1 - 3\delta$, for $n$ large enough, there exists a constant $K$ such that

$$\tilde{n} \geq n \left[1 - K \log(n/\delta)^{d+1} n^{-\frac{2 + d}{\pi^d}}\right].$$

(16)