IBM Research Report

Norm-induced Densities and Testing the Boundedness of a Convex Set

Alexandre Belloni
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

IBM

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

Norm-induced Densities and Testing the Boundedness of a Convex Set

Alexandre Belloni

IBM T.J. Watson Research Center, Building 800, 32-221, 1101 Kitchawan Road, Yorktown Heights, New York, 10598, USA

email: belloni@mit.edu   `http://web.mit.edu/belloni/www`

In this paper we explore properties of a family of probability density functions, called norm-induced densities, defined as

$$f_t(x) = \begin{cases} \dfrac{e^{-t\|x\|^p} dx}{\int_K e^{-t\|y\|^p} dy}, & x \in K \\ 0, & x \notin K, \end{cases}$$

where K is a n-dimensional convex set that contains the origin, parameters $t > 0$ and $p > 0$, and $\|\cdot\|$ is any norm. We also develop connections between these densities and geometric properties of K such as diameter, width of the recession cone, and others. Since f_t is log-concave only if $p \geq 1$, this framework also covers non-logconcave densities.

Moreover, we establish a new set inclusion characterization for convex sets. This leads to a new concentration of measure phenomena for unbounded convex sets. More precisely, we show that most points of an unbounded convex set are contained in any enlargement of the set's recession cone. In turn, when the recession cone has positive width, most points are actually in the recession cone itself.

Finally, these properties are used to develop an efficient probabilistic algorithm to test whether a convex set, represented only by membership oracles (a membership oracle for K and a membership oracle for its recession cone), is bounded or not, where the algorithm reports an associated certificate of boundedness or unboundedness.

Key words: convex sets; unbounded sets; concentration of measure; homotopy; random walk; membership oracle

MSC2000 Subject Classification: Primary: 52A20 , 90C60 ; Secondary: 90C25

OR/MS subject classification: Primary: Convexity ; Secondary: Random Walk

1. Introduction.   The geometry of convex sets has been extensively studied during the past half-century. More recently, the interplay between convex geometry and probability theory has been investigated. Among many possible research directions, log-concave probability measures provide an interesting framework that generalizes uniform densities on convex sets but preserves many interesting properties, see [1, 12]. Here we will focus on a family of density functions whose properties relate to geometric properties of the convex set defined by its support. We say that a probability density function $f_t : \mathbb{R}^n \to \mathbb{R}_+$ is called *norm-induced* if

$$f_t(x) = \begin{cases} \dfrac{e^{-t\|x\|^p} dx}{\int_K e^{-t\|y\|^p} dy}, & x \in K \\ 0, & x \notin K \end{cases} \tag{1}$$

where K is a convex set in $\mathbb{R}^n$ that contains the origin, $t > 0$ and $p > 0$ are parameters, and $\|\cdot\|$ is a fixed norm. Note that a probability density function of the form (1) is always proper. Moreover, f_t is log-concave only if $p \geq 1$ therefore our framework also covers non-logconcave densities.

We will explore the connection between these densities and geometric properties of the convex set K itself, usually as a function of the parameter t. The geometry of unbounded convex sets plays an important role in our analysis. For instance, given an arbitrary unbounded convex set we show that most of its points are contained in any *enlargement* of its recession cone. This simple geometric phenomenon motivates many of our results.

In the case of $p \geq 1$, a random variable whose density distribution function is f_t can be efficiently simulated (at least approximately) by geometric random walk algorithms [11]. In turn, theoretical results on f_t can be used to construct (implementable) algorithms to test properties of K itself.

We also develop an algorithm to test if a given convex set $K \subset \mathbb{R}^n$ is bounded or unbounded. In either case the algorithm will construct an associated certificate of boundedness or unboundedness based on the properties established in Section 3. We emphasize that algorithms for this problem are closely related to the representation used to describe the set K. Our interest lies in cases in which the convex set is given only by membership oracles (a membership oracle for K and a membership oracle for its recession cone). This (minimal assumption) framework is the standard framework in several applications in the computer science literature. Furthermore, it covers many other problems of interest such as convex feasibility.

The decision problem of testing for boundedness has a variety of interesting consequences. In recent years, several probabilistic methods have been proposed to compute quantities like centroid, volume [14], convex optimization [7], and many others [2], in the context of convex bodies. In all these cases, boundedness of a convex set is a fundamental assumption for whose testing our algorithm provides a constructive approach. Khachiyan established the equivalence between a strongly polynomial algorithm for linear programming and a strongly polynomial algorithm for testing unboundedness of a convex set associated with a system of linear inequalities [8]. Moreover, linear homogeneous conic feasibility problems of the form

$$\begin{cases} Ax & \in & C \\ x & \in & \mathbb{R}^n\backslash\{0\} \end{cases} \tag{2}$$

(where C is closed convex cone) can be converted into our framework by defining

$$K = \{x \in \mathbb{R}^n : Ax + h \in C\} \tag{3}$$

for any $h \in \textbf{int } C$. In this case, $0 \in \textbf{int } K$, and the recession cone of K coincides with the set of feasible solutions of the original system (2). Moreover, K is bounded only if (2) is infeasible. Finally, a membership oracle for the cone C suffices to construct a membership oracle for K and its recession cone C_K.

The implementability of our algorithm relies on the ability to sample random variables distributed according to a probability density f_t. Over the last decade many important developments on sampling from log-concave densities, most notably via geometric random walks, have been observed. In particular, the *hit-and-run* random walk has been extensively analyzed and polynomial rates of convergence have been established for this particular random walk under the log-concavity assumption [9, 12, 13, 11]. Besides, the homotopy analysis proposed here is similar to the analysis done by Lovász and Vempala in [14] of the algorithm they called *reverse annealing*, which was applied to the problem of computing the volume of a (bounded) convex body. However, our approach differs from [14] with in respects: by using a different density family, and by dealing explicitly with the possible unboundedness of K.

In the presence of additional structure, other algorithms are available in the literature. For example, assuming that a *self-concordant barrier* function is available for K, minimizing such function leads to appropriate certificates of boundedness or unboundedness (note that the minimum is finite only if K is bounded). That idea was used first by de Ghellink and Vial in [4] for linear programming and more recently by Nesterov, Todd and Ye [16] for nonlinear programming problems. Moreover, if K is given explicitly by a set of linear inequalities, one can identify an element of the recession cone by solving a linear programming problem.

We emphasize that none of these approaches extends to the membership oracle framework. In fact, negative results do exist for approximating the diameter of a bounded convex set, which is a closely related problem. Lovász and Simonovits [10] show that no polynomial time algorithm (deterministic or probabilistic) can approximate the diameter of a convex set within a factor of $\sqrt{n}$ in polynomial time under the membership oracle framework. Thus, it is notable that, as we show, testing if a convex set is unbounded is solvable in polynomial time.

An outline of this paper is as follows. Section 2 illustrates the geometric intuition underlying many of the results and proves a new inclusion representation for K which leads to a concentration phenomena for unbounded convex sets. We establish many properties relating the density functions (1) and the convex set K in Section 3. The algorithm to test boundedness is presented in Section 4 and its analysis is presented in the following sections. Finally, Appendix A contains the details on how to implement the hit-and-run geometric random walk efficiently for the density functions used in the algorithm.

1.1 Preliminaries, definitions, and notation. Recall that a real-valued function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_+$ is said to be a norm if:

   (i) $\|x\| = 0$ only if $x = 0$;
   (ii) $\|tx\| = |t|\|x\|$ for any $t \in \mathbb{R}$;
   (iii) $\|x + y\| \le \|x\| + \|y\|$ for any $x, y \in \mathbb{R}^n$.

For a given norm, we can define a unit ball

$$B_{\|\cdot\|}(x, r) = B(x, r) = \{y \in \mathbb{R}^n : \|y - x\| \le r\}, \tag{4}$$

and a unit sphere $S^{n-1} = \{y \in \mathbb{R}^n : \|y\| = 1\}$.

The Euclidean inner product is denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|_2 = \sqrt{\langle \cdot, \cdot \rangle}$ denotes the Euclidean norm induced by it. For $x \in \mathbb{R}^n$ and $r > 0$, let $B_2(x, r)$ denote the Euclidean ball centered at x with radius r, i.e., $B_2(x, r) = \{y \in \mathbb{R}^n : \|x - y\|_2 \le r\}$. The unit Euclidean sphere of $\mathbb{R}^n$ is denoted by S_2^{n-1}, i.e., $S_2^{n-1} = \{y \in \mathbb{R}^n : \|y\|_2 = 1\}$.

The dual norm of $\| \cdot \|$ induced by $\langle \cdot, \cdot \rangle$ is defined by

$$\|s\|_* = \max\{\langle s, x \rangle : x \in B(0, 1)\}, \tag{5}$$

for which we can also define a unit ball $B_*(s, r) = \{w \in \mathbb{R}^n : \|w - s\|_* \le r\}$ and a unit sphere S_*^{n-1}. By definition we have that $|\langle s, x \rangle| \le \|s\|_* \|x\|$. The dual norm completely defines the original norm, since we have

$$\|x\| = \max\{\langle s, x \rangle : s \in B_*(0, 1)\}. \tag{6}$$

That is, the dual norm of the dual norm is the original norm. Recall that the dual norm of the Euclidean norm is also the Euclidean norm, which is said to be self-dual.

A set S is convex if $x, y \in S$ implies $\alpha x + (1 - \alpha)y \in S$ for every $\alpha \in [0, 1]$. C is a cone if $x \in C$ implies $\alpha x \in C$ for every $\alpha \ge 0$. If C is a cone, the *width* of C is given by

$$\tau_C = \max_{x, \tau}\{\tau : B(x, \tau) \subset C, \ \|x\| = 1\}, \tag{7}$$

the radius of the largest ball contained in C centered at a unit vector (in the appropriate norm). L is a subspace of $\mathbb{R}^n$ if $x, y \in L$ implies that $\alpha x + \beta y \in L$ for every $\alpha, \ \beta \in \mathbb{R}$.

For a set S, the operations $\mathbf{conv}(S)$, $\mathbf{cone}(S)$, $\mathbf{ext}(S)$, $\mathbf{int}\ (S)$, $\mathbf{cl}\ (S)$, $\mathbf{diam}(S)$, and $\mathsf{Vol}(S)$ denote, respectively, the convex hull, conic hull, extreme points, interior, closure, diameter, and volume of S (see [17] for complete definitions). $\mathsf{Vol}_{n-1}(A)$ denotes the $n-1$-dimensional volume of the surface A (see [15] for details and relation with cone measure). Also, for $x \in \mathbb{R}^n$, let $\mathbf{dist}(x, S) = \inf\{\|x - y\| : y \in S\}$ denote the distance from x to S. For a scalar u, set $(u)_+ = \max\{0, u\}$, and for a matrix M denote by $\lambda_{\max}(M)$ (respectively $\lambda_{\min}(M)$) its maximum (respectively minimum) singular value. Moreover, given a vector v, v' denotes its transpose.

A *membership oracle* for a set S is any algorithm, that given any point $x \in \mathbb{R}^n$, correctly identifies if $x \in S$ or not. Let 1_S denote the indicator function of the set S, that is, $1_S(x) = 1$ if $x \in S$ and $1_S(x) = 0$ otherwise.

With respect to complexity notation, $g(n)$ is said to be $O(f(n))$ if there exists a constant M such that $g(n) \le Mf(n)$, while $g(n, \mu)$ is $O^*(f(n, \mu))$ if there exists constants M and k such that $g(n, \mu) \le Mf(n, \mu)\ln^k n$ (that is, the O^* notation omits logarithmic factors of the dimension but still shows the dependence on other possible condition measure μ).

1.2 Logconcave densities: concepts and notation. We define π_f as the probability measure associated with a probability density function f (i.e, $\pi_f(S) = \int_S f(x)dx$), $E_f[\cdot]$ as the expectation with respect to f, and z_f as the mean of a random variable whose probability density function is f. The following class of functions plays a central role in the sampling literature.

DEFINITION 1.1 *A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is logconcave if for any two points $x, \ y \in \mathbb{R}^n$ and any $\lambda \in (0, 1)$,*

$$f(\lambda x + (1 - \lambda)y) \ge f(x)^\lambda f(y)^{1-\lambda}.$$

Definition 1.1 implies that $\ln f$ is a concave function and, in particular, the support of f is a convex set. We say that a random variable is logconcave if its probability density function is a logconcave function. Gaussians, exponential and uniform densities on convex sets are examples of logconcave densities.

There are a variety of metrics available for probability densities. Here, we will make use of two of them: the *total variation* norm, defined as

$$\|f - g\|_{TV} = \frac{1}{2}\int_{\mathbb{R}^n} |f(x) - g(x)|dx,$$

and the L^2 *norm of f with respect to g*, defined as

$$\|f/g\| = E_f\left[\frac{f(x)}{g(x)}\right] = \int_{\mathbb{R}^n} \frac{f(x)}{g(x)} f(x)dx = \int_{\mathbb{R}^n} \left(\frac{f(x)}{g(x)}\right)^2 g(x)dx.$$

The following useful concept is associated with the covariance matrix induced by f.

DEFINITION 1.2 *A density function f with expectation z_f is said to be C-isotropic if for every vector v,*

$$\frac{\|v\|_2^2}{C} \le \int_{\mathbb{R}^n} \langle v, x - z_f \rangle^2 f(x)dx \le \|v\|_2^2 C,$$

equivalently, any eigenvalue λ of the covariance matrix of f satisfies: $\quad \dfrac{1}{C} \le \lambda \le C.$

A function f is said to be in isotropic position if it is 1-isotropic, that is, its covariance matrix is the identity. Thus, any density can be made isotropic by a linear transformation of the space.

2. ϵ-Enlargements of the Recession Cone. In this section we revisit a classical representation theorem for closed convex sets and we provide a new set inclusion characterization for such sets which will be key in our analysis.

Let $K \subset \mathbb{R}^n$ be a closed convex set that contains the origin. As a matter of convenience, assume K is full dimensional, as one can always work within the affine hull of K, but at considerable notational and expositional expense. For the sake of exposition, in this section we will assume that K contains no lines (in the upcoming sections we do not make such assumption).

As is standard in convex theory, the set of all directions of half-lines contained in K defines the *recession cone* of K denoted by C_K, i.e.,

$$C_K = \{d \in \mathbb{R}^n : K + d \subseteq K\}, \tag{8}$$

which is a closed convex cone ([17] Corollary 8.3.2). Moreover, it is well-known that K is unbounded only if $C_K \ne \{0\}$ ([17] Theorem 8.4).

Under this framework, K contains at least one extreme point ([17] Corollary 18.5.3). Thus, the following representation theorem for closed, line-free, convex sets applies to K.

THEOREM 2.1 *Any closed line-free convex set $K \subset \mathbb{R}^n$ can be decomposed into the sum of two sets, the recession cone of K and the convex hull of its extreme points. That is,*

$$K = C_K + \mathbf{conv}\Big(\mathbf{ext}(K)\Big). \tag{9}$$

In order to develop intuition on the relation between high-dimensional convex sets and volume, we need to understand how to draw pictures of what high-dimensional convex sets look like. The intuition for convex bodies (bounded convex sets with nonempty interior) was first suggested by Gromov and Milman in [5]. The fact that the volume of parallel intersections of half-spaces with K decays exponentially fast after passing the median level must be taken into account. As suggested by Gromov and Milman, small dimensional pictures of a high-dimensional convex body should have a hyperbolic form, see Figure 1.

However, our concern here is to extend such intuition to unbounded convex sets. In this context a similar (concentration) phenomenon will also be observed. Assuming that the recession cone has positive width, "most of the points" of the set are in the recession cone. (Note that one needs to be careful when quantifying "most of the points", since the volume is infinite.) Again small dimensional pictures of high-dimensional unbounded convex sets must have a hyperbolic form, see Figure 2.

In fact, even if the recession cone has zero width, "most of the points" of K will be contained in any ϵ-*enlargement* of the recession cone, where the latter is formally defined as follows.

DEFINITION 2.1 *The ϵ-enlargement of the recession cone of a closed convex set K is defined as*

$$C_K^\epsilon = \mathbf{cone}\left\{S^{n-1} \cap (C_K + B(0, \epsilon))\right\}. \tag{10}$$
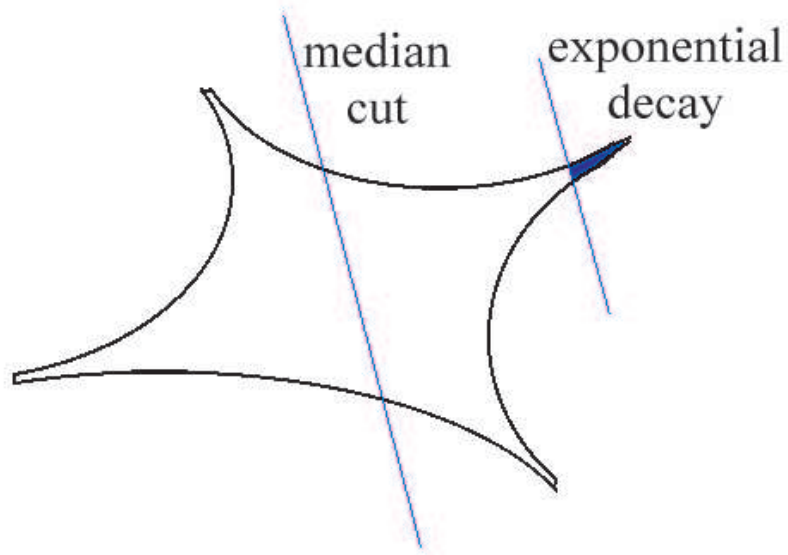
Figure 1: High-dimensional bounded convex sets and volume. Exponential decay as we move away the median cut.
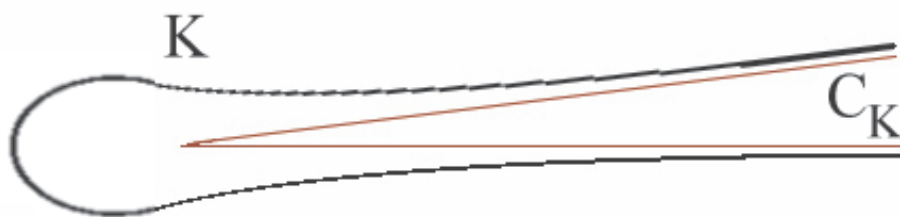

Figure 2: High-dimensional unbounded convex sets and volume. Most of the mass concentrates around the recession cone.

The following properties of the ϵ-enlargement of the recession cone follow directly from the definition.

LEMMA 2.1 *Let K be a convex set that contains no lines and let C_K denote its recession cone. For $\epsilon \in (0,1)$, we have that:*

 (i) *C_K^ϵ is a closed cone;*
 (ii) *if $C_K = \{0\}$, then $C_K^\epsilon = \{0\}$.*
 (iii) *if $C_K \neq \{0\}$, then $\tau_{C_K^\epsilon} \geq \max\{\tau_{C_K}, \ \epsilon/3\}$.*

Moreover, we can take $\epsilon > 0$ sufficient small such that:

 (iv) *C_K^ϵ contains no line.*

PROOF. C_K^ϵ is a closed cone by definition. If $C_K = \{0\}$, for $\epsilon < 1$ we have $S^{n-1} \cap (C_K + B(0,\epsilon)) = \emptyset$ and $C_K^\epsilon = \{0\}$ and (ii) follows. To prove (iii), observe that $C_K \subset C_K^\epsilon$ implies that $\tau_{C_K^\epsilon} \geq \tau_{C_K}$. Moreover, if $z \in C_K$, $\|z\| = 1$, we will show that $B(z, \epsilon/3) \subset C_K^\epsilon$. Take $x \in B(z, \epsilon/3)$ and let $w := x/\|x\|$. Since $\|z\| = 1$, we have $1 + \epsilon/3 \geq \|x\| \geq 1 - \epsilon/3$. Then

$$
\begin{aligned}
\left\| \tfrac{x}{\|x\|} - z \right\| &= \left\| \tfrac{x-z}{\|x\|} + \tfrac{z}{\|x\|} - z \right\| \leq \tfrac{\|x-z\|}{\|x\|} + \|z\| \left(\tfrac{1}{\|x\|} - 1 \right) \\
&\leq \tfrac{\epsilon}{3} \tfrac{1}{1-\epsilon/3} + \tfrac{1}{1-\epsilon/3} - 1 = \tfrac{2\epsilon}{3-\epsilon} < \epsilon,
\end{aligned}
$$

where the last inequality follows since $\epsilon < 1$. Since $x/\|x\| \in S^{n-1}$ and $x/\|x\| \in B(z, \epsilon)$, $x \in C_K^\epsilon$. This implies that $\tau_{C_K^\epsilon} \geq \epsilon/3$.

Since C_K is a pointed cone, the Pompeiu-Hausdorff distance of the cone C_K to any cone C that contains a line is bounded away from zero, $\delta(C_K, C) = \sup_{\|z\| \leq 1} |\mathbf{dist}(z, C_K) - \mathbf{dist}(z, C)| > c > 0$ (see [6] for details). Moreover, $\delta(C_K, C_K^\epsilon) \leq \epsilon$. Thus we can choose $\epsilon < c$ and C_K^ϵ does not contain lines proving (iv).

$\square$

Now we are in position to obtain a set inclusion characterization of the aforementioned geometric phenomena, which will motivate most of the analysis in the sections to come.

THEOREM 2.2 *Let K be a convex set. Then, for any $\epsilon > 0$, there exists a positive scalar R_ϵ such that*

$$
K \subseteq C_K^\epsilon + B(0, R_\epsilon). \tag{11}
$$

PROOF.   Without loss of generality, assume $0 \in K$. Suppose that there exists a sequence $\{x^j\}$ of elements of K such that $\mathbf{dist}(x^j, C_K^\epsilon) > j$. The normalized sequence has a convergent subsequence, $\{d^{k_j} = x^{k_j}/\|x^{k_j}\|\}$, to a point d. Since K is convex and closed, $0 \in K$, and $\|x^{k_j}\| > k_j > 1$, d^{k_j} and d are in K. In fact, since $C_K = \cap_{\delta > 0} \delta K = \lim_{k_j \to \infty} K/k_j$ (Corollary 8.3.2 of [17]) and $d^{k_j} \in K/k_j$, we have $d \in C_K$ after taking limits.

Next note that for any $\epsilon > 0$, there exists a number j_0 such that

$$
\|d^{k_j} - d\| < \epsilon \quad \text{for all } j > j_0.
$$

Therefore, we have $d^{k_j} \in C_K^\epsilon$ for $j > j_0$. In addition, $x^{k_j} \in C_K^\epsilon$, since C_K^ϵ is a cone, contradicting $\mathbf{dist}(x^j, C_K^\epsilon) > j$ for every j. $\square$

It is well-known that for the polyhedral case Theorem 2.2 holds with $\epsilon = 0$ but this is not the case in general.

EXAMPLE 2.1 *Consider $K = \{(x,y) \in \mathbb{R}^2 : x > -1, \ \ln(1+x) \geq y\}$. Here the recession cone is $C_K = \mathbb{R}_+ \times \mathbb{R}_-$ and $\tau_{C_K} = 1/\sqrt{2}$. Note that for all $R > 1$, the point $(R, \ln R) \in K$ and $\mathbf{dist}((R, \ln R), C_K) = \ln R$. Consider $\epsilon < 1/5$ which implies that C_K^ϵ contains no lines. For any point $(x,y) \in K$ such that $x \geq \frac{1}{\epsilon} \ln^2 \frac{1}{\epsilon}$, we have that $y \leq \ln x \leq \epsilon x$ which implies that $(x,y) \in C_K^\epsilon$. Thus, Theorem 2.2 holds with $R_\epsilon := \frac{2}{\epsilon} \ln^2 \frac{1}{\epsilon}$.*

The ϵ-enlargement of the recession cone captures all points of K except for a bounded subset. Thus, assuming that K is unbounded, such geometric phenomenon implies that "most points" of K will be

contained in $K \cap C_K^\epsilon$. Moreover, if $\tau_{C_K} > 0$, most points of K will actually be contained in C_K itself. This is quantified in the next results. Note that these results are valid for an arbitrary norm since both the unit sphere and the ϵ-enlargement depend on the particular norm.

LEMMA 2.2 *Let C_K be a convex cone with strictly positive width τ_{C_K}. Then for every $\epsilon > 0$*

$$\mathsf{Vol}_{n-1}\left(C_K^\epsilon \cap S^{n-1}\right) \leq \left(1 + \frac{2\epsilon}{\tau_{C_K}}\right)^n \mathsf{Vol}_{n-1}\left(C_K \cap S^{n-1}\right).$$

PROOF. By definition of τ_{C_K}, $B(z, \tau_{C_K}) \subset C_K$ for some $z \in S^{n-1}$. Since C_K and C_K^ϵ are cones,

$$\frac{\mathsf{Vol}_{n-1}\left(C_K^\epsilon \cap S^{n-1}\right)}{\mathsf{Vol}_{n-1}\left(C_K \cap S^{n-1}\right)} = \frac{\mathsf{Vol}\left(C_K^\epsilon \cap B(0,1)\right)}{\mathsf{Vol}\left(C_K \cap B(0,1)\right)} \leq \frac{\mathsf{Vol}\left((C_K \cap B(0,1)) + B(0,\epsilon)\right)}{\mathsf{Vol}\left(C_K \cap B(0,1)\right)},$$

since $\left(C_K^\epsilon \cap B(0,1)\right) \subset \left(C_K \cap B(0,1) + B(0,\epsilon)\right)$. Next, we have $B(z/2, \tau_{C_K}/2) \subset \left(C_K \cap B(0,1)\right)$, so that

$$\left(C_K \cap B(0,1)\right) + B(0,\epsilon) \subset \left(C_K \cap B(0,1)\right) + \frac{2\epsilon}{\tau_{C_K}}\left(C_K \cap B(0,1) - \frac{z}{2}\right),$$

and the result follows. □

In the next theorem we relate the probability of a norm-induced random variable being in the recession cone of K and in being in the ϵ-enlargement (note that Theorem 2.3 allows for K_Y to be a non-convex set).

THEOREM 2.3 *Let X_t be norm-induced random variable defined by a norm $\|\cdot\|$, parameters $t > 0$ and $p > 0$, and a convex set K that contains the origin. Moreover, let Y_t also be a norm-induced random variable defined by the same norm $\|\cdot\|$, the same parameters $t > 0$ and $p > 0$, but with a different support set K_Y such that*

$$K \subseteq K_Y \subseteq C_K^\epsilon + B(0, R_\epsilon).$$

Then

$$P\left(X_t \in C_K\right) \geq P\left(Y_t \in C_K\right) \geq \left(1 - \frac{2\epsilon}{\tau_{C_K}}\right)^n P\left(Y_t \in C_K^\epsilon\right).$$

PROOF. By the Total Probability Theorem we have that

$$P\left(Y_t \in C_K\right) = P\left(Y_t \in C_K | Y_t \in K\right) P\left(Y_t \in K\right) + P\left(Y_t \in C_K | Y_t \notin K\right) P\left(Y_t \notin K\right).$$

Next, note that $P\left(Y_t \in C_K | Y_t \notin K\right) = 0$ and $P\left(Y_t \in C_K | Y_t \in K\right) = P\left(X_t \in C_K\right)$ since $C_K \subset K \subset K_Y$. These relations imply the first inequality.

For simplicity, let $A = \{x \in \mathbb{R}^n : \|x\| = \rho\}$ denote the sphere of radius ρ. To prove the second inequality, note that for any $\rho \geq 0$

$$\begin{aligned}
P\left(Y_t \in C_K \mid \|Y_t\| = \rho\right) &= \frac{\mathsf{Vol}_{n-1}(C_K \cap A)}{\mathsf{Vol}_{n-1}(K_Y \cap A)} \geq \left(1 - \frac{2\epsilon}{\tau_{C_K}}\right)^n \frac{\mathsf{Vol}_{n-1}(C_K^\epsilon \cap A)}{\mathsf{Vol}_{n-1}(K_Y \cap A)} \\
&\geq \left(1 - \frac{2\epsilon}{\tau_{C_K}}\right)^n \frac{\mathsf{Vol}_{n-1}(C_K^\epsilon \cap A \cap K_Y)}{\mathsf{Vol}_{n-1}(K_Y \cap A)} \\
&= \left(1 - \frac{2\epsilon}{\tau_{C_K}}\right)^n P\left(Y_t \in C_K^\epsilon \mid \|Y_t\| = \rho\right)
\end{aligned}$$

where the first inequality follows from Lemma 2.2 and the second because C_K^ϵ might not be contained in K_Y. Since this holds for every $\rho \geq 0$, the result follows by integrating over ρ.

□

Taken together, Theorems 2.2 and 2.3 with $\epsilon \leq \tau_{C_K}/4n$ imply that we have at least a constant fraction (independent of the dimension) of the points in C_K^ϵ in C_K. Thus, finding points in C_K is easily achieved by finding random points in $C_K^\epsilon \cap K$.

3. Geometry of norm-induced densities. In this section we establish several properties of the random variable X_t, whose density function f_t is given by (1). Many properties of X_t can be related with a variety of geometric quantities/properties of the set K itself. These results are partially motivated by known properties of uniform distributions over convex sets. Nonetheless, norm-induced densities may fail to be log-concave which is a key property used in similar results [12]. Moreover, for algorithmic reasons, we are also interested in relating explicitly the dependence between geometric quantities of K and the parameter t of the norm-induced density. For the sake of exposition, we will assume the following:

ASSUMPTION 3.1 *K is a closed convex set that contains the origin and has nonempty interior.*

Assumption 3.1 is needed to ensure that K has positive n-dimensional volume (possibly infinite). Nonetheless, one can always work with the affine hull of K if one uses the appropriate lower dimensional volume. As expected, the origin could be replaced by any other point in K. All results could be restated for that particular point (if we translate the density f_t appropriately).

Before we proceed, we mention that the inclusion representation of $K \subset C_K^\epsilon + B(0, R_\epsilon)$ for any $\epsilon > 0$ established in Theorem 2.2 will be used throughout the section. Although stated as an "assumption" in some of the results, it holds for any convex set and $\epsilon > 0$. One particularly interesting case is when $\epsilon < \tau_{C_K}/4n$ which implies that the volume of C_K and C_K^ϵ are comparable (Lemma 2.2).

We start with a simple result of the levels sets of the norm-induced densities.

THEOREM 3.1 *The upper level sets of a density function f_t defined as in (1) are convex. Moreover, consider the function $g(x) = \|\nabla \ln f_t(x)\|_*$, which gives the dual norm of the gradient of the logarithm of f_t at x. This function is well defined even if f_t is non-differentiable, and we have two cases:*
(i) if $p \geq 1$ the lower level sets of g are convex;
(ii) if $0 < p \leq 1$ the upper level sets of g are convex.

PROOF. The upper level sets of f_t are defined by

$$\{x \in K : f_t(x) \geq c\} = \left\{ x \in K : t\|x\|^p \leq -\ln\left(c \int f_t(z)dz \right) \right\}.$$

Since the parameters t and p are positive, such set is either a scaling of the unit ball intersected with K or the empty set (both convex sets).

For the second part of the theorem, we first compute $g(x)$. Recall that if $s \in \partial\|x\|$ (s is in the subdifferential of $\|\cdot\|$ at x), we have that $s \in S_*^{n-1}$. Then we have

$$g(x) = \| -tp\|x\|^{p-1}s\|_* = tp\|x\|^{p-1}\|s\|_* = tp\|x\|^{p-1}.$$

The result follows since g is convex for $p \geq 1$ and concave for $0 < g \leq 1$. $\qquad \square$

The next lemma shows that the norm-induced densities exhibit a concentration phenomena similar to log-concave densities.

LEMMA 3.1 *Let $X_t \sim f_t$ as in (1). For $D \geq 4$ we have*

$$P\left(\|X_t\| \geq \left(\frac{n}{tp}D\right)^{1/p} \right) \leq \frac{1}{3}e^{-\frac{n}{p}D/3}$$

PROOF. Let $M := \left(\frac{n}{tp}D\right)^{1/p}$. Define $K_Y := \mathbf{cone}\{MS^{n-1} \cap K\}$ (a possibly non-convex set) and let $Y_t \sim e^{-t\|y\|^p}$ over K_Y (note that if $K_Y = \emptyset$ we have $K \subset B(0, M)$ and $P(\|X_t\| \geq M) = 0$). Then we have

$$P(\|X_t\| \geq M) \leq P(\|Y_t\| \geq M),$$

since $\{y \in K_Y : \|y\| \leq M\} \subset K$ and $\{y \in K : \|y\| \geq M\} \subset K_Y$.

Then for Y_t we have

$$P(\|Y_t\| \geq M) = \frac{\int_{y \in K_Y : \|Y_t\| \geq M} e^{-t\|y\|^p}dy}{\int_{y \in K_Y} e^{-t\|y\|^p}dy} = \frac{\int_M^\infty v^{n-1}e^{-tv^p}dv}{\int_0^\infty v^{n-1}e^{-tv^p}dv}.$$

Let $u := tv^p$, $c := tM^p = \frac{n}{p}D$, and we obtain

$$P(\|Y_t\| \geq M) = \frac{\int_c^\infty u^{\frac{n}{p}-1}e^{-u}du}{\int_0^\infty u^{\frac{n}{p}-1}e^{-u}du} \leq \frac{\int_c^\infty u^{\lceil \frac{n}{p}-1\rceil}e^{-u}du}{\int_0^\infty u^{\lceil \frac{n}{p}-1\rceil}e^{-u}du} = e^{-c}\sum_{i=0}^{\lceil \frac{n}{p}-1\rceil}\frac{c^i}{i!}.$$

Next recall that $i! \geq (i/e)^i$ (from the Stirling's formula, and using the convention that $0^0 = 1$) and $\ln(1 + h) \leq h$. Then

$$
\begin{aligned}
e^{-c}\sum_{i=0}^{\lceil \frac{n}{p}-1\rceil}\frac{c^i}{i!} &\leq e^{-\frac{n}{p}D}\sum_{i=0}^{\lceil \frac{n}{p}-1\rceil}\frac{(\frac{n}{p}D)^i}{(i/e)^i} = e^{-\frac{n}{p}D}\sum_{i=0}^{\lceil \frac{n}{p}-1\rceil}e^{i\ln(\frac{n/p}{i}De)} \\
&= e^{-\frac{n}{p}D}\sum_{i=0}^{\lceil \frac{n}{p}-1\rceil}e^{i\ln(1+\frac{(n/p)-i}{i})+i\ln(De)} \leq e^{-\frac{n}{p}D}\sum_{i=0}^{\lceil \frac{n}{p}-1\rceil}e^{(n/p)-i+i\ln(De)} \\
&= e^{-\frac{n}{p}(D-1)}\sum_{i=0}^{\lceil \frac{n}{p}-1\rceil}e^{i\ln(D)} = e^{-\frac{n}{p}(D-1)}\frac{e^{\lceil \frac{n}{p}-1\rceil\ln(D)}-1}{e^{\ln(D)}-1} \\
&\leq \frac{1}{3}e^{-\frac{n}{p}(D-1-\ln(D))} \leq \frac{1}{3}e^{-\frac{n}{p}D/3}.
\end{aligned}
$$

$\square$

As anticipated, we will show that the probability of the event $\{X_t \in C_K\}$ will be large for "small" values of t. To do so, we exploit the spherical symmetry of f_t (i.e., $f_t(x) = f_t(y)$ if $\|x\| = \|y\|$) and the geometric phenomena induced by the representation of Theorem 2.2. That symmetry will allow us to connect the volume of relevant sets with the probability of the event $\{X_t \in C_K\}$. Lemma 3.2 below properly quantifies this notion.

LEMMA 3.2 *Suppose that $K \subseteq C_K^\epsilon + B(0, R_\epsilon)$ for some $\epsilon > 0$ where K is an unbounded convex set. Let $X_t \sim f_t$ as in (1), and let $\rho > 6R/\tau_{C_K}$. Then*

$$P(X_t \in C_K \mid \|X_t\| = \rho) \geq \left(1 - \frac{2\epsilon}{\tau_{C_K}}\right)^n\left(1 - \frac{6R_\epsilon}{\tau_{C_K}\cdot\rho}\right)^n.$$

PROOF. Note that restricted on $\|X_t\| = \rho$, the density f_t is constant. Thus, we have

$$
\begin{aligned}
P(X_t \in C_K \mid \|X_t\| = \rho) &= \frac{\mathsf{Vol}_{n-1}(C_K \cap (\rho S^{n-1}))}{\mathsf{Vol}_{n-1}(K \cap (\rho S^{n-1}))} \geq \frac{\mathsf{Vol}_{n-1}(C_K \cap (\rho S^{n-1}))}{\mathsf{Vol}_{n-1}([C_K^\epsilon + B(0, R_\epsilon)] \cap (\rho S^{n-1}))} \\
&\geq \left(1 - \frac{2\epsilon}{\tau_{C_K}}\right)^n\frac{\mathsf{Vol}_{n-1}(C_K^\epsilon \cap (\rho S^{n-1}))}{\mathsf{Vol}_{n-1}([C_K^\epsilon + B(0, R_\epsilon)] \cap (\rho S^{n-1}))},
\end{aligned}
$$

where we used Lemma 2.2 in the last inequality.

For any a, b, and a set S, define the sets $S_{[a,b]} := \{y \in S : a \leq \|y\| \leq b\}$ and $S_a := S_{[a,a]} = (aS^{n-1})\cap S$. Set $\bar{\rho} = \rho/3 \geq 2R_\epsilon/\tau_{C_K}$, and consider the following sets

$$I_{[\bar{\rho},\rho]} := \{y \in C_K^\epsilon : \bar{\rho} \leq \|y\| \leq \rho\} \subset O_{[\bar{\rho},\rho]} := \{y \in C_K^\epsilon + B(0, R_\epsilon : \bar{\rho} \leq \|y\| \leq \rho\}.$$

We first show that $O_{[\bar{\rho},\rho]} \subset I_{[\bar{\rho},\rho]} + B(0, 2R_\epsilon)$. For $y \in O_{[\bar{\rho},\rho]}$, we have $y = v + w$, where $\|v\| \leq R_\epsilon$, $w \in C_K^\epsilon$, $\|v + w\| \in [\bar{\rho}, \rho]$, and hence $\|w\| \in [\bar{\rho} - R_\epsilon, \rho + R_\epsilon]$. Therefore we have $w \in I_{[\bar{\rho}-R_\epsilon,\rho+R_\epsilon]}$. Next we will show that $I_{[\bar{\rho}-R_\epsilon,\rho+R_\epsilon]} \subset I_{[\bar{\rho},\rho]} + B(0, R_\epsilon)$, and the result follows since $v \in B(0, R_\epsilon)$. Take $w \in I_{[\bar{\rho}-R_\epsilon,\rho+R_\epsilon]}$. We can assume that $\|w\| < \bar{\rho}$ (the case of $\|w\| > \rho$ is similar and if $\bar{\rho} \leq \|w\| \leq \rho$ we automatically have $w \in I_{[\bar{\rho},\rho]}$). Note that C_K^ϵ is a cone and $w \in C_K^\epsilon$, $\frac{\bar{\rho}}{\|w\|}w \in C_K^\epsilon$. In fact, we have $\frac{\bar{\rho}}{\|w\|}w \in I_{[\bar{\rho},\rho]}$. Hence $\|w - \frac{\bar{\rho}}{\|w\|}w\| = \bar{\rho} - \|w\| \leq R_\epsilon$, i.e., $w \in I_{[\bar{\rho},\rho]} + B(0, R_\epsilon)$.

Now, take $z \in C_K^\epsilon$, $\|z\| = 1$, such that $B(z, \tau_{C_K}) \subset C_K^\epsilon$ (there is no convexity assumption on C_K^ϵ; we know that it has width at least τ_{C_K}, since it contains C_K). Thus, since C_K^ϵ is a cone, we have

$$B\left(\frac{(\bar{\rho}+\rho)}{2}z, \frac{(\bar{\rho}+\rho)}{2}\tau_{C_K}\right) \subset C_K^\epsilon.$$

Observe that $\rho \geq 3\bar{\rho}$ implies that $\bar{\rho} \leq \frac{(\bar{\rho}+\rho)}{2} - \frac{(\bar{\rho}+\rho)}{4}$ and $\frac{(\bar{\rho}+\rho)}{2} + \frac{(\bar{\rho}+\rho)}{4} \leq \rho$, so we have

$$B\left(\frac{(\bar{\rho}+\rho)}{2}z, \frac{(\bar{\rho}+\rho)}{4}\tau_{C_K}\right) = B\left(\frac{(\bar{\rho}+\rho)}{2}z, 2\kappa R_\epsilon\right) \subset I_{[\bar{\rho},\rho]}, \tag{12}$$

where $\kappa = \frac{(\bar{\rho}+\rho)}{8}\frac{\tau_{C_K}}{R_\epsilon}$. Thus, $B(w, 2R_\epsilon) \subset \frac{1}{\kappa}I_{[\bar{\rho},\rho]}$ for $w = \frac{\bar{\rho}+\rho}{2\kappa}z$.

$$\frac{\mathsf{Vol}(I_{[\bar{\rho},\rho]})}{\mathsf{Vol}(O_{[\bar{\rho},\rho]})} \geq \frac{\mathsf{Vol}(I_{[\bar{\rho},\rho]})}{\mathsf{Vol}(I_{[\bar{\rho},\rho]} + B(0, 2R_\epsilon))}$$

$$\geq \frac{\mathsf{Vol}(I_{[\bar{\rho},\rho]})}{\mathsf{Vol}(I_{[\bar{\rho},\rho]} + \frac{1}{\kappa} I_{[\bar{\rho},\rho]})} = \frac{1}{\left(1 + \frac{1}{\kappa}\right)^n} \geq \left(1 - \frac{1}{\kappa}\right)^n.$$

We will complete the proof in three steps. For $s \in [\bar{\rho}, \rho]$, consider the sets I_s and O_s.

First note that

$$\frac{\mathsf{Vol}_{n-1}(I_s)}{\mathsf{Vol}_{n-1}(O_s)} = \frac{\mathsf{Vol}_{n-1}(C_K^\epsilon \cap sS^{n-1})}{\mathsf{Vol}_{n-1}([C_K^\epsilon + B(0, R_\epsilon)] \cap sS^{n-1})}$$

is a nondecreasing function of s. Second, observe that

$$\frac{\mathsf{Vol}(I_{[\bar{\rho},\rho]})}{\mathsf{Vol}(O_{[\bar{\rho},\rho]})} = \frac{\int_{\bar{\rho}}^{\rho} \mathsf{Vol}_{n-1}(I_s)ds}{\int_{\bar{\rho}}^{\rho} \mathsf{Vol}_{n-1}(O_s)ds}.$$

Next, we will make use of the following remark.

Remark 2. For any $a < b$ and any two positive functions g and h such that

$$\frac{g(s)}{h(s)} \text{ is nondecreasing for } s \in [a, b], \text{ we have } \frac{g(a)}{h(a)} \leq \frac{\int_a^b g(s)}{\int_a^b h(s)}ds \leq \frac{g(b)}{h(b)}.$$

Third, applying Remark 3 with $g(s) = \mathsf{Vol}_{n-1}(I_s)$, $h(s) = \mathsf{Vol}_{n-1}(O_s)$, $a = \bar{\rho}$ and $b = \rho$ to obtain

$$\frac{\mathsf{Vol}_{n-1}(C_K^\epsilon \cap (\rho S^{n-1}))}{\mathsf{Vol}_{n-1}([C_K^\epsilon + B(0, R_\epsilon)] \cap (\rho S^{n-1}))} = \frac{\mathsf{Vol}_{n-1}(I_\rho)}{\mathsf{Vol}_{n-1}(O_\rho)} \geq \frac{\mathsf{Vol}(I_{[\bar{\rho},\rho]})}{\mathsf{Vol}(O_{[\bar{\rho},\rho]})} \geq \left(1 - \frac{1}{\kappa}\right)^n.$$

Finally, since $\bar{\rho} = \rho/3$ we have $\kappa = \frac{\tau_{C_K}\rho}{6R_\epsilon}$ and the result follows. $\square$

PROOF OF REMARK 3. Simply note that $\frac{g(s)}{h(s)} \leq \frac{g(b)}{h(b)}$ for $s \in [a, b]$. Then

$$\int_a^b g(s)ds = \int_a^b h(s)\frac{g(s)}{h(s)}ds \leq \left(\frac{g(b)}{h(b)}\right)\int_a^b h(s)ds,$$

and a similar argument yields the lower bound.

Since the bound obtained in Lemma 3.2 is monotone in ρ and trivially true for bounded convex sets ($\tau_{C_K} = 0$), we have the following useful corollary.

COROLLARY 3.1 *For any convex set $K \subseteq C_K^\epsilon + B(0, R_\epsilon)$ and $\rho \geq 6R_\epsilon/\tau_{C_K}$ we have*

$$P(X_t \in C_K \mid \|X_t\| \geq \rho) \geq \left(1 - \frac{2\epsilon}{\tau_{C_K}}\right)^n \left(1 - \frac{6R_\epsilon}{\tau_{C_K} \cdot \rho}\right)^n.$$

The previous results used conditioning on the event that the random points have large norm. It is natural to bound the probability of these conditional events as well.

LEMMA 3.3 *Suppose that K is unbounded, and that f_t is given by (1). For a random variable X_t distributed according to f_t we have*
$$P(\|X_t\| \geq \rho) \geq 1 - 4e\rho t^{1/p}.$$

PROOF. We can assume that $\rho t^{1/p} < 1/4e$, since the bound is trivial otherwise. We use the notation introduced in the proof of Lemma 3.2, $K_{[a,b]} = \{y \in K : a \leq \|y\| \leq b\}$.

Assuming that K is unbounded, there exists $z \in C_K$, $\|z\| = 1$. Then,

$$K_{[0,\rho]} + (\kappa + 2\rho)z \subset K_{[\kappa+\rho,\kappa+3\rho]} \quad \text{for any} \quad \kappa \in \mathbb{R}_+, \tag{13}$$

since for each $x \in K_{[0,\rho]}$, we have

$$
\begin{aligned}
\|x + (\kappa + 2\rho)z\| &\leq \|x\| + (\kappa + 2\rho)\|z\| &\leq \kappa + 3\rho, \\
\|x + (\kappa + 2\rho)z\| &\geq (\kappa + 2\rho)\|z\| - \|x\| &\geq \kappa + \rho,
\end{aligned}
$$

and $x + (\kappa + 2\rho)z \in K$ (the latter follows from $z \in C_K$).

So, for any odd integer $m \geq 3$, $K_{[\rho,m\rho]} = \bigcup_{i=1}^{(m-1)/2} K_{[(2i-1)\rho,(2i+1)\rho]}$, where the union is disjoint (except on a set of measure zero), and by (13) we have

$$
\mathsf{Vol}(K_{[\rho,m\rho]}) = \sum_{i=1}^{(m-1)/2} \mathsf{Vol}(K_{[(2i-1)\rho,(2i+1)\rho]}) \geq \frac{(m-1)}{2}\mathsf{Vol}(K_{[0,\rho]}).
$$

Now, define $m := \frac{1}{\rho t^{1/p}}$. By assumption we have $\rho t^{1/p} < 1/4e$, which implies $m \geq 4e$ (again, we assume that m in an odd integer for convenience). Thus,

$$
\begin{aligned}
P(\|X_t\| \geq \rho) &= \frac{\int_{K_{[\rho,\infty]}} e^{-t\|x\|^p}dx}{\int_{K_{[0,\rho]}} e^{-t\|x\|^p}dx + \int_{K_{[\rho,\infty]}} e^{-t\|x\|^p}dx} \\[2ex]
&\geq \frac{\int_{K_{[\rho,m\rho]}} e^{-t\|x\|^p}dx}{\int_{K_{[0,\rho]}} e^{-t\|x\|^p}dx + \int_{K_{[\rho,m\rho]}} e^{-t\|x\|^p}dx} \\[2ex]
&\geq \frac{\mathsf{Vol}(K_{[\rho,m\rho]})e^{-t(m\rho)^p}}{\mathsf{Vol}(K_{[0,\rho]}) + \mathsf{Vol}(K_{[\rho,m\rho]})e^{-t(m\rho)^p}} \\[2ex]
&\geq \frac{(m-1)\mathsf{Vol}(K_{[\rho,m\rho]})e^{-t(m\rho)^p}}{2\mathsf{Vol}(K_{[\rho,m\rho]}) + (m-1)\mathsf{Vol}(K_{[\rho,m\rho]})e^{-t(m\rho)^p}} \\[2ex]
&\geq \frac{(m-1)e^{-t(m\rho)^p}}{2 + (m-1)e^{-t(m\rho)^p}}.
\end{aligned}
$$

Using the definition of m, we have

$$
\begin{aligned}
P(\|X_t\| \geq \rho) &\geq \frac{(\frac{1}{\rho t^{1/p}} - 1)e^{-1}}{2 + (\frac{1}{\rho t^{1/p}} - 1)e^{-1}} &= \frac{(\frac{1}{\rho t^{1/p}} - 1)}{2e + (\frac{1}{\rho t^{1/p}} - 1)} \\[2ex]
&\geq 1 - \frac{2e}{\frac{1}{\rho t^{1/p}} - 1} &= 1 - \frac{2e\rho t^{1/p}}{1 - \rho t^{1/p}} \quad \geq 1 - 4e\rho t^{1/p}.
\end{aligned}
$$

$\square$

Lemma 3.2 and Corollary 3.1 quantify the geometric notion mentioned earlier (motivated by Figure 2), that most points in K outside a compact set are in C_K if its width is positive. On the other hand, Lemma 3.3 shows that the norm of X_t is likely to be greater than $0.01/t^{1/p}$. Taken together, they lead to a lower bound on the probability of the event $\{X_t \in C_K\}$.

THEOREM 3.2 *Suppose that $K \subseteq C_K^\epsilon + B(0, R_\epsilon)$ is an unbounded convex set, and f_t is defined as in (1). Let $\delta \in (0, 1)$ and suppose*

$$
t \leq \left(\frac{\delta^2 \tau_{C_K}}{96enR_\epsilon}\right)^p,
$$

and let X_t be a random variable distributed according to f_t. Then

$$
P\left(X_t \in C_K \setminus \{0\}\right) \geq (1 - \delta)\left(1 - \frac{2\varepsilon}{\tau_{C_K}}\right)^n.
$$

PROOF. All that is needed is to combine Lemmas 3.2 and 3.3. Clearly,

$$P\left(X_t \in C_K\right) \geq P\left(\|X_t\| \geq \rho\right) P\left(X_t \in C_K \mid \|X_t\| \geq \rho\right).$$

By Lemmas 3.2 and 3.3, we have

$$P\left(\|X_t\| \geq \rho\right) \geq 1 - 4e\rho t^{1/p} \text{ and } P\left(X_t \in C_K \mid \|X_t\| \geq \rho\right) \geq \left(1 - \frac{6R}{\tau_{C_K} \cdot \rho}\right)^n \left(1 - \frac{2\varepsilon}{\tau_{C_K}}\right)^n.$$

It is sufficient to ensure that

$$1 - 4et\rho > 1 - \frac{\delta}{2} \quad \text{and} \quad \left(1 - \frac{6R}{\tau_{C_K} \cdot \rho}\right)^n > 1 - \frac{\delta}{2}$$

for some $\rho \geq 6R/\tau_{C_K}$. Noting that $(1-x)^{1/n} \leq 1 - \frac{x}{n}$ for all $x \in [0,1]$, the second relation also holds since if

$$1 - \frac{6R}{\tau_{C_K} \cdot \rho} > 1 - \frac{\delta}{2n}.$$

It suffices to choose $\rho = \frac{12nR}{\delta\tau_{C_K}}$ for the second relation to hold, and the first relation holds since

$$4e\rho t^{1/p} \leq \frac{4\delta^2 \tau_{C_K} e 12nR}{96enR\delta\tau_{C_K}} = \frac{\delta}{2}.$$

$\square$

In the case of K being unbounded, Theorem 3.2 characterizes the behavior of random variables $X_t \sim f_t$ for values of t that are "relatively small" with respect to the unbounded support K. It is natural to ask what kind of behavior one should expect for values of t that are "relatively large" with respect to the support K. The next lemma, which holds regardless of K being unbounded, gives a partial answer to this question. When t is "relatively large", there is a fixed probability for which there exist a direction s for which $|\langle s, X_t \rangle|$ is also reasonably large.

LEMMA 3.4 *Assume that there exists $\bar{v} \in K$ such that $\|\bar{v}\| = D$. If $t^{1/p} \geq 1/D$, then*

$$\max_{s \in S_*^{n-1}} P\left(|\langle s, X_t \rangle| > \frac{1}{4ent^{1/p}}\right) > \frac{1}{9}.$$

PROOF. Let $\gamma = \frac{1}{4ent^{1/p}}$. Since $D = \|\bar{v}\| = \max\{\langle s, \bar{v}\rangle : s \in B_*(0,1)\}$, there exists $v \in S_*^{n-1}$ such that $\langle v, \bar{v} \rangle = D$.

By obtaining a lower bound for v we automatically obtain a lower bound for the maximum over the dual sphere. So,

$$\max_{s \in S_*^{n-1}} P\left(|\langle s, X_t \rangle| > \gamma\right) \geq P\left(|\langle v, X_t \rangle| > \gamma\right).$$

It will be convenient to define the following sets

$$A = \{x \in K : |\langle v, x\rangle| < \gamma\} \quad \text{and} \quad B = \left\{(1-\alpha)x + \alpha\bar{v} \; : \; x \in A\right\},$$

where $\alpha = \gamma \frac{2}{D} < \frac{1}{2en}$.

Note that for any $y \in B$ we have $y \in K$ since x and $\bar{v} \in K$ and $\alpha \in (0,1)$. Moreover,

$$|\langle v, y \rangle| \geq \langle v, y \rangle = (1-\alpha)\langle v, x\rangle + \alpha\langle v, \bar{v}\rangle > -\gamma + \gamma\frac{2}{D}D = \gamma.$$

Thus, $y \notin A$.

Since $\alpha\|v\| = 2\gamma$, note that for any $M > 0$ we have

$$\|(1-\alpha)x + \alpha\bar{v}\|^p \leq (\|x\| + 2\gamma)^p \leq \|x\|^p\left(1 + \frac{2\gamma}{M}\right)^p + (M + 2\gamma)^p. \tag{14}$$

where each term in the right hand side follows by considering that $\|x\| \geq M$ or $\|x\| \leq M$. By choosing $M = 2e\gamma n$, we have

$$
\begin{aligned}
P\left(X_t \in K \setminus A\right) \;\geq\;& P\left(X_t \in B\right) = \int_B \frac{e^{-t\|y\|^p} dy}{\int_K e^{-t\|z\|^p} dz} \\[2mm]
=\;& \frac{1}{\int_K e^{-t\|z\|^p} dz} \int_A e^{-t\|(1-\alpha)x + \alpha\bar{v}\|^p} (1-\alpha)^n dx \\[2mm]
\geq\;& \frac{e^{-t(M+2\gamma)^p}(1-\alpha)^n}{\int_K e^{-t\|z\|^p} dz} \int_A e^{-t\left\|\left(1+\frac{2\gamma}{M}\right)x\right\|^p} dx \\[2mm]
\geq\;& \frac{e^{-t(2\gamma)^p(en+1)^p}(1-\alpha)^n \left(1+\frac{1}{en}\right)^{-n}}{\int_K e^{-t\|z\|^p} dz} \int_{\left(1+\frac{2\gamma}{M}\right)A} e^{-t\|x\|^p} dx \\[2mm]
\geq\;& e^{-t(2\gamma)^p 4^p n^p} e^{-1/(2e-1)} e^{-1/e} P(X_t \in A) \\[2mm]
\geq\;& \tfrac{1}{e^2} P(X_t \in A) = \tfrac{1}{e^2}\left(1 - P(X_t \in K \setminus A)\right)
\end{aligned}
$$

and we have $P(X_t \in K \setminus A) \geq 1/9$. $\qquad\square$

This result allows us to construct several useful bounds on the moments of X_t. Moreover, in the case of K being unbounded, the condition on t is always satisfied.

In the case of $p = 1$, we can improve on Lemma 3.4 as follows.

COROLLARY 3.2 *If in addition we have $p = 1$, for $t \geq \sqrt{n}/D$ we have*

$$
\max_{s \in S_*^{n-1}} P\left(|\langle s, X_t\rangle| > \frac{1}{4et\sqrt{n}}\right) > \frac{1}{3}.
$$

PROOF. The proof is similar to Lemma 3.4. We define $A := \left\{x \in K : |\langle v, x\rangle| < \frac{1}{4et\sqrt{n}}\right\}$, and $\alpha := \frac{1}{2etD\sqrt{n}} < \frac{1}{2en}$. Next, note that the inequality (14) becomes

$$
\|(1-\alpha)x + \alpha\bar{v}\| \leq \|x\| + \alpha\|\bar{v}\|,
$$

and there is no need to use M. In this case, it follows that

$$
P(X_t \in K \setminus A) \geq e^{-t\alpha\|\bar{v}\|}(1-\alpha)^n P(X_t \in A) \geq \frac{1}{2}P(X_t \in A) = \frac{1}{2}\left(1 - P(X_t \in K \setminus A)\right)
$$

and we have $P(X_t \in K \setminus A) \geq 1/3$. This holds since we have

$$
\begin{aligned}
e^{-t\alpha\|\bar{v}\|}(1-\alpha)^n \;=\;& e^{-1/(2e\sqrt{n})}\left(1 - \frac{1}{2etD\sqrt{n}}\right)^n \geq e^{-1/2e}\left(1 - \frac{1}{2en}\right)^n \\[2mm]
\geq\;& e^{-1/2e} \cdot e^{-1/(2e-1)} \geq 0.664 \geq 1/2.
\end{aligned}
$$

$\qquad\square$

For the reader convenience, the next corollary is specialized for the case of $\|\cdot\| = \|\cdot\|_2$ and $p = 1$; it will be used in the following section. Recall that given a random variable X_t, its covariance matrix is given by $V_t = E\left[(X_t - E[X_t])(X_t - E[X_t])'\right]$ and it relates with the second moment matrix of X_t by $E[X_t X_t'] = V_t + E[X_t]E[X_t]'$. The corollary shows how one can relate the eigenvalues of the second moment matrix, the parameter t, and the diameter of K (if the latter is bounded).

COROLLARY 3.3 *Assume that $p = 1$ and $\|\cdot\| = \|\cdot\|_2$. If K is unbounded, then for every $t > 0$*

$$
\lambda_{\max}\left(E[X_t X_t']\right) \geq \frac{1}{n(7et)^2}.
$$

Otherwise, suppose that for some $t > 0$ we have

$$
\lambda_{\max}\left(E[X_t X_t']\right) < \frac{1}{n(7et)^2}.
$$

Then K is bounded and $K \subset B(0, \overline{R})$ for $\overline{R} = \frac{\sqrt{n}}{t}$.

PROOF. If one is using the Euclidean norm, we have $\|\cdot\|_2 = \|\cdot\| = \|\cdot\|_*$. Moreover, we have that

$$\lambda_{\max}\left(E\left[X_t X_t'\right]\right) = \max_{s \in S^{n-1}} \langle s, E\left[X_t X_t'\right] s \rangle = \max_{s \in S^{n-1}} E\left[\langle s, X_t \rangle^2\right].$$

If K is unbounded, there exists $v \in K$ such that $\|v\| \geq \sqrt{n}/t$. Thus, we have

$$\max_{s \in S^{n-1}} E\left[\langle s, X_t \rangle^2\right] \geq \frac{1}{n(4et)^2} \max_{s \in S^{n-1}} P\left(\left|\langle s, X_t \rangle\right| > \frac{1}{\sqrt{n}(4et)}\right) \geq \frac{1}{n(4et)^2}\frac{1}{3} \geq \frac{1}{n(7et)^2}$$

where the second inequality follows from Corollary 3.2.

For the second part, suppose there exists $\bar{v} \in K$, with $\|\bar{v}\| \geq \frac{\sqrt{n}}{t}$. Let $\overline{R} := \|\bar{v}\| \geq \sqrt{n}/t$. Then $t \geq \sqrt{n}/\overline{R}$ and $\|\bar{v}\| = \overline{R}$, so by Corollary 3.2 we have a contradiction. $\square$

A similar result can be established for small values at t. Intuitively we will recover results known for the uniform density as we let t goes to zero.

LEMMA 3.5 *Assume that there exists $\bar{v} \in K$ such that $\|\bar{v}\| = D$. If $t^{1/p} < 1/(2D)$,*

$$\max_{s \in S_*^{n-1}} P\left(\left|\langle s, X_t \rangle\right| > \frac{D}{4en}\right) > \frac{1}{9}.$$

PROOF. The proof is similar to the proof of Lemma 3.4 if one defines

$$A = \left\{x \in K : \left|\langle v, x \rangle\right| < \frac{D}{4en}\right\}.$$

$\square$

COROLLARY 3.4 *Assume that K has diameter D and $t < 1/(2D)$. Then,*

$$E\left[\|X_t\|\right] \geq \frac{D}{72en}.$$

PROOF. If K has diameter D, there exist two points $\bar{v}, \bar{w} \in K$ such that $\|\bar{v} - \bar{w}\| = D$. We can assume that $\|\bar{v}\| \geq D/2$. The result follows from Lemma 3.5. $\square$

We close this section with a simple observation with respect to the *entropy* of a density function

$$Ent(f) = -\int_{\mathbb{R}^n} f(x)\ln f(x)dx.$$

COROLLARY 3.5 *If f_t is a norm induced density function, then*

$$Ent(f_t) = tE[\|X_t\|^p].$$

Thus, Lemma 3.4, Lemma 3.5, and Corollary 3.2 can be used to bound the entropy.

PROOF. The result follows by noting that $-f(x)\ln f(x) = t\|x\|^p e^{-t\|x\|^p}$. $\square$

4. Testing the Boundedness of a Convex Set: a Density Homotopy. The algorithm we propose is a homotopy procedure to simulate a random variable which has desirable properties with respect to K. Motivated by the geometry of unbounded convex sets, the uniform density over K would be an interesting candidate. Unfortunately, as opposed to most frameworks in the literature, a random variable which is uniformly distributed over K will not be *proper* if K is unbounded and cannot be used. Instead, we will work with a parameterized family of densities, $\mathcal{F} = \{f_t : t \in (0, t_0]\}$, such that f_t is a proper density for every t. In addition, for any fixed compact subset of K the parameterized density uniformly converges to the uniform density over that compact set as $t \to 0$. As mentioned earlier, the algorithm must provide us with a certificate of boundedness or unboundedness. Any nonzero element of the recession cone of K is a valid certificate of unboundedness. We will assume that a membership oracle for the recession cone of K itself is available.

On the other hand, the certificate of boundedness is more thought-provoking. If K is described by a set of linear inequalities, $K = \{x \in \mathbb{R}^n : Ax \leq b\}$, K will be bounded if and only if positive combinations of the rows of A span $\mathbb{R}^n$. More generally, if K is represented by a separation oracle, a valid certificate of boundedness would be a set of normal vectors associated with hyperplanes returned by the separation oracle whose positive combinations span $\mathbb{R}^n$. Note that a membership oracle provides much less information and we cannot sensibly extend the previous concept to our framework. Instead, our certificate of boundedness will be given by the eigenvalues of the second moment matrix associated with the random variables induced by the family $\mathcal{F}$. In contrast with the previous certificates, it will be a "probabilistic certificate of boundedness" since the true second moment matrix is unknown and must be estimated via a probabilistic method.

4.1 Assumptions and condition measures. In addition to Assumption 3.1, we make the following assumptions on the set K:

ASSUMPTION 4.1 *K is a closed convex set given by a membership oracle.*

ASSUMPTION 4.2 *There exists $r > 0$ such that $B(0, r) \subseteq K$.*

ASSUMPTION 4.3 *A membership oracle is available for C_K, the recession cone of K.*

The closedness of K could be relaxed with minor adjustments on the implementation of the random walk. Assumption 4.3 specifies how K is represented.

Assumption 4.2 is stronger than Assumption 3.1. It requires that we are given a point in the interior of K, which is assumed to be the origin without loss of generality. That is standard in the membership oracle framework, since the problem of finding a point in a convex set given only by a membership oracle is hard in general. Finally, we emphasize that only a lower bound on r is required to implement our algorithm. Section 5.3 gives a simple procedure to obtain an approximation of r within a factor of $\sqrt{n}$.

In our analysis, besides the dimension of K, there are three geometric quantities that naturally arise: r, R_ϵ, and τ_{C_K}. Not surprisingly, the dependence of the computational complexity of our algorithm on these geometric quantities differs if K is bounded or unbounded (recall that the case of $\tau_{C_K} = 0$ is fundamentally different if K is bounded or unbounded). Nonetheless, in either case the dependence on these quantities will be only logarithmic. An instance of the problem is said to be *ill-conditioned* if $\tau_{C_K} = 0$ and K is unbounded, otherwise the instance is said to be *well-conditioned*.

4.2 The algorithm. In order to define the algorithm, let f_t be defined as (1) with $\|\cdot\| = \|\cdot\|_2$ (the Euclidean norm). Let $\alpha \in (0, 1)$ and let *hit-and-run* be a geometric random walk which will simulate the next random variable (see Section 6 for details). This yields the following "exact" method to test the boundedness of K:

Density Homotopy Algorithm (Exact):
Input: r such that $B(0, r) \subset K$, define $t_0 = t_{\text{initial}}(r)$, $\alpha \in (0, 1)$, and set $k \leftarrow 0$.
Step 1. (Initialization) Draw $X_{t_0} \sim f_{t_0}(x)$.
Step 2. (Testing Unboundedness) If $X_{t_k} \in C_K \setminus \{0\}$, **stop**.
Step 3. (Variance and Mean) Compute the mean and covariance of X_{t_k}: z_k and V_k.
Step 4. (Testing Boundedness) If $\lambda_{\max}\left(V_k + z_k z_k^T\right) < \frac{1}{n(7et_k)^2}$, **stop**.
Step 5. (Update Density) Update the parameter: $t_{k+1} = (1 - \alpha) \cdot t_k$.
Step 6. (Random Point) Draw $X_{t_{k+1}} \sim$ *hit-and-run*$(f_{t_{k+1}}, X_{t_k}, V_k, m)$
Step 7. (Loop) Set $k \leftarrow k + 1$, goto **Step 2**.

This (exact) method requires r, the exact draw of $X_{t_{k+1}}$, and the exact computation of the mean z_{k+1} and covariance matrix V_{k+1} of the random variable $X_{t_{k+1}}$. In order to obtain an implementable method, we can use only approximations of these objects.

Detailed bounds on the computational complexity of the hit-and-run procedure, on the estimation of $\hat{z}_{k+1}$ and $\hat{V}_{k+1}$, and $\hat{r}$ are provided in Sections 5.3, 6, and 7. Moreover, the use of an approximate mean and approximate covariance matrix must be taken into account in the test of boundedness (Step 4), which is presented in Theorem 4.1.

Each loop of the algorithm (Steps 2-7) is called an *iteration* of the algorithm. Thus, the work per iteration consists of (i) performing the hit-and-run random walk, (ii) computing an approximation of the covariance matrix, (iii) testing if the current point belongs to the recession cone, and (iv) computing the largest eigenvalue of a positive definite matrix. Although a highly accurate approximation of the covariance matrix is not needed, the probabilistic method used to estimate such matrix requires at least $O^*(n)$ samples. Such estimation will dominate the computational complexity per iteration.

Letting t_{final} denote the final value of the parameter t when the algorithm stops, the total number of iterations of the algorithm will be

$$\left\lceil \frac{1}{\alpha} \cdot \ln\left(\frac{t_{\text{initial}}}{t_{\text{final}}}\right) \right\rceil.$$

The following theorem characterizes the complexity of the homotopy algorithm.

THEOREM 4.1 *Let K be a convex set satisfying Assumptions 4.2, 4.3, and 4.1 and consider the homotopy algorithm using a family of densities $\mathcal{F} = \{f_t(x) \sim 1_K(x) \cdot e^{-t\|x\|_2} : t \in (0, t_0]\}$. Then:*

(i) If $K \subseteq C_K^\epsilon + B(0, R_\epsilon)$, where $\epsilon = \tau_{C_K}/4n$, is unbounded, the algorithm will compute a valid certificate of unboundedness in at most

$$O\left(\sqrt{n}\ln\left(\frac{n}{\delta}\frac{1}{\tau_{C_K}}\frac{R_\epsilon}{r}\right)\right) \quad \text{iterations with probability } 1 - \delta,$$

where each iteration makes at most $O^*\left(n^4\ln\left(\frac{1}{\delta}\ln\left(\frac{1}{\delta}\frac{1}{\tau_{C_K}}\frac{R_\epsilon}{r}\right)\right)\right)$ calls to the membership oracle.

(ii) If $K \subseteq B(0, R)$ is bounded, the algorithm will compute a valid certificate of boundedness in at most

$$O\left(\sqrt{n}\ln\left(n\frac{R}{r}\right)\right) \quad \text{iterations with probability } 1 - \delta,$$

where each iteration makes at most $O^*\left(n^4\ln\left(\frac{1}{\delta}\ln\left(\frac{R}{r}\right)\right)\right)$ calls to the membership oracle.

The proof of Theorem 4.1, which is provided in Section 10, is built upon the analysis of the next six sections.

5. Analysis of the homothopy algorithm

5.1 Stopping criteria: unbounded case. An appropriate certificate of unboundedness for a convex set is to exhibit a non-zero element of the recession cone of K. Assumption 4.3 allows us to correctly verify if a point is an element of C_K. For example, in the case of linear conic systems (3), any membership oracle for C itself can be used to construct a membership oracle for K and C_K.

If the algorithm terminates indicating that K is unbounded, a nonzero element of the recession cone was found (a certificate of unboundedness). Thus, the algorithm always terminates correctly in this case. The following corollary of Theorem 3.2 ensures that we can find such certificate, which provides a desirable stopping criteria in the case of K being unbounded.

COROLLARY 5.1 *Suppose $K \subseteq C_K^\epsilon + B(0, R_\epsilon)$, where $\epsilon = \tau_{C_K}/4n$, is unbounded. After at most*

$$T = 6\ln\frac{1}{\delta} + \frac{1}{\alpha}\ln\left(t_0\frac{96enR_\epsilon}{(1/2)^2\tau_{C_K}}\right)$$

iterations of the exact algorithm, we will have found an nonzero element of the recession cone with probability at least $1 - \delta$.

PROOF. Let $\delta_1 = 1/2$, $T_1 = T - 6\ln\frac{1}{\delta}$, and $\epsilon = \tau_{C_K}/4n$. We start the algorithm with $t = t_0$, and after T_1 iterations, we obtain $t_{T_1} = \frac{\delta_1^2\tau_{C_K}}{96enR_\epsilon}$. For our choice of ϵ, applying Theorem 3.2 to $X_{t_{T_1}}$ yields

$$P(X_{t_{T_1}} \in C_K \setminus \{0\}) \geq \frac{1}{2e}.$$

For every following iteration, since t is decreasing, we have at least that same probability. Therefore, the probability of $X_t \in C_K$ in any of the next $6\ln\frac{1}{\delta}$ iterations is at least $1 - \delta$. $\quad\square$

5.2 Stopping criteria: bounded case. In contrast with the unbounded case, we lack a straightforward certificate for the case of K being bounded. In addition, an unbounded set whose recession cone has zero width should not be wrongly classified as bounded. That is, our analysis should cover such an ill-conditioned case as well.

In the search for an appropriate certificate, the mean of the random variable X_t appears as a natural candidate. Assuming that the set K is unbounded and line-free, its norm should increase as the parameter t decreases. On the other hand, if K is bounded, the mean will remain bounded no matter how much t decreases. Unfortunately, that analysis breaks down for sets that contain lines. For example, if K is symmetric the mean of X_t is zero for every $t > 0$, whether K is bounded or not.

In order to overcome this we consider the second moment matrix Ω_t of the random variable X_t. The matrix Ω_t will be large, in the positive definite sense, if either the covariance matrix or the norm of the mean of X_t is large. Again, if K is unbounded the maximum eigenvalue of Ω_t increases as the parameter t decreases. Otherwise, K being bounded, the maximum eigenvalue will eventually be bounded. This provides a nice criterion which is robust to instances where K contains lines and/or $\tau_{C_K} = 0$. We emphasize that the second order information is readily available, since we are required to compute the covariance matrix and the mean of X_t to implement the hit-and-run random walk for f_t (see Section 6 for details on the sampling theory and the reasons why we need to compute the covariance matrix to keep f_t in near-isotropic position). The next corollary provides the desirable stopping criteria.

COROLLARY 5.2 *Suppose K is bounded. The exact algorithm will detect boundedness for all $t < \frac{1}{7eR\sqrt{n}}$, and will bound R by $\frac{\sqrt{n}}{t}$. Moreover, this will happen in at most $T = \frac{1}{\alpha}\ln\left(t_0\, 7eR\sqrt{n}\right)$ iterations.*

PROOF. If $t < \frac{1}{7eR\sqrt{n}}$, then $\lambda_{\max}\left(E[X_t X_t']\right) < \frac{1}{(7et)^2 n}$ from Corollary 3.3. From Corollary 3.3, this means that K is bounded and $R < \frac{\sqrt{n}}{t}$.

Moreover, starting with $t = t_0$, after $T = \frac{1}{\alpha}\ln\left(t_0\, 7eR\sqrt{n}\right)$ iterations we obtain $t_T \leq \frac{1}{7eR\sqrt{n}}$. $\qquad\square$

To generate a certificate of boundedness, the maximum eigenvalue of the second moment matrix must be estimated (Section 7 covers the necessary theory for that). Since it will be estimated via a probabilistic method, there is a probability of failure on each iteration which must be properly controlled (see Section 9 for details). Thus, in contrast to the unbounded case, if the algorithm terminates indicating that K is bounded, there is a probability of failure associated with that decision which can be made smaller at the expense of computational time.

5.3 Initialization of the algorithm: unknown r. This section clarifies how to start the algorithm, that is, how should we choose the initial value t_0 based on r. As is usual under the membership oracle framework, it is assumed that we know a point in the interior of K, which is taken to be the origin for convenience (Assumption 4.2). In some applications of interest such points are readily available, for example $0 \in \text{int } K$ in the conic system (3).

The implementation of Step 1 will be done by a simple accept-reject method, see [3]. Note that it is simple to draw a random variable X_{t_0} whose density is proportional to $f_{t_0}(x) \sim e^{-t_0\|x\|_2}$ on $\mathbb{R}^n$ instead of only on K (pick a point uniformly on S_2^{n-1} and then scale using a $\Gamma(n, t_0)$ random variable[1]). If it is the case that $X_{t_0} \in K$, we accept that point; otherwise, we draw another point according to f_{t_0} and repeat the process until a point in K is obtained.

Now, we need a constructive approach to bound r from below. Fortunately, a simple procedure is available for estimating r up to a factor of $\sqrt{n}$. That will be satisfactory since the final dependence on r is only logarithmic. Consider the following quantity:

$$\hat{r} = \min_{i=1,\ldots,n} \max\{\, t \;:\; t\mathbf{e}_i \in K, \;\; -t\mathbf{e}_i \in K\,\}, \tag{15}$$

where $\mathbf{e}_i \in \mathbb{R}^n$ denotes the ith unit vector of $\mathbb{R}^n$. It is clear that $\hat{r}$ can be approximated in $O\left(n|\ln\frac{1}{r}|\right)$ operations (via a simple binary search) and will not increase the order of the computational complexity of the algorithm. The next lemma provides a guarantee of the quality of the approximation provided by $\hat{r}$.

---

[1] A $\Gamma(\alpha, \beta)$ random variable is characterized by the density $f(x) = (\beta^\alpha/\Gamma(\alpha))x^{\alpha-1}e^{-\beta x}$ for $x \geq 0$, and zero otherwise.

LEMMA 5.1 *Let r be the radius of the largest ball centered at the origin contained in K and let $\hat{r}$ be as defined in (15). Then $\hat{r} \geq r \geq \hat{r}/\sqrt{n}$.*

PROOF. Clearly, $\hat{r} \geq r$. Note that $\hat{r}\mathbf{e}_i \in K$ for every i. Thus, the convex hull of these points contains a ball of radius $\hat{r}/\sqrt{n}$ which is contained in K. Therefore, $\hat{r} \geq r \geq \hat{r}/\sqrt{n}$. □

We also need to ensure that the probability of the event $\{X_{t_0} \in K\}$ is reasonably large. The next lemma achieves this by a suitable choice of the initial parameter t_0 based on the radius r of the largest ball centered at the origin contained in K.

LEMMA 5.2 *Assume that the ball centered at the origin with radius r is contained in K. Let X_{t_0} be a random variable whose density is proportional to $e^{-t_0\|x\|_2}$ for any $x \in \mathbb{R}^n$. Then if $t_0 \geq 2\frac{(n-1)}{r}$, we have*

$$P(X_{t_0} \in K) \geq 1 - e^{n-(t_0 r/2)}.$$

PROOF. Using Lemma 5.16 of [12] (since $t_0 r/(n-1) \geq 2$) for the second inequality, we have

$$
\begin{aligned}
P(X_{t_0} \notin K) &\leq P(X_{t_0} \notin B(0,r)) = P\big(f_{t_0}(X_{t_0}) \leq f_{t_0}(0)e^{-t_0 r}\big) \leq \left(e^{1-\frac{t_0 r}{n-1}}\frac{t_0 r}{n-1}\right)^{n-1} \\
&\leq e^{n-1}\left(e^{-\frac{t_0 r}{n-1}}\frac{t_0 r}{n-1}\right)^{n-1} \leq e^{n-(t_0 r/2)}
\end{aligned}
$$

since $e^{-c}c \leq e^{-c/2}$ for every $c \geq 0$.

□

This allows us to efficiently implement the accept-reject method in Step 1 of the algorithm.

COROLLARY 5.3 *After computing $\hat{r}$ as in (15), it suffices to initialize the algorithm with*

$$t_0 = \frac{8n^{3/2}}{\hat{r}}$$

to obtain that $P(X_{t_0} \in K) \geq 1 - e^{-3n}$.

PROOF. Recall that Lemma 5.1 implies that $\hat{r} \geq r/\sqrt{n}$. By Lemma 5.2, $P(X_{t_0} \in K) \geq 1 - e^{n-4n^{3/2}\frac{r}{\hat{r}}} \geq 1 - e^{-3n}$. □

6. Sampling f_t via a geometric random walk. The ability to sample according to any density in the family $\mathcal{F}$ is the driving force of our algorithm. Although a variety of densities can be perfectly simulated with negligible computational effort, that is no longer the case if we restrict the support of the density to be an arbitrary convex set K given by a membership oracle. In fact, even to generate a random point distributed uniformly over a convex set is an interesting problem with many remarkable applications (linear programming, computing the volume, etc., see [7],[14],[11]).

Important tools to generate random points proportional to a density function restricted to a high dimensional convex set K are the so-called geometric random walks. Starting with a point in K, at each step the random walk moves to a point according to some distribution that depends only (i) on the current point, and (ii) on the desired density f to be simulated. Thus, the sequence of points of the random walk is a Markov chain whose state space is K. Moreover, there are simple choices of transition kernels (which is the continuous state space analog for the transition matrix for a finite state Markov chain) that make f the unique stationary distribution of this Markov chain (for example, the celebrated Metropolis filter), which automatically ensures several asymptotic results for arbitrary Markov chains [3]. Going one step further, we are interested in the convergence rate to the stationary distribution, which is a much more challenging question (which could be arbitrarily slow in general). So we can bound the necessary number of steps required by the random walk to generate a random point whose density is approximately f.

By choosing $\mathcal{F}$ to be a family of logconcave densities (in our case the parameter $p = 1$), we will be able to invoke several results from a recent literature which demonstrate the efficiency of one particular random walk called hit-and-run, see [9, 12, 13]. Roughly speaking, these results show that if (i) a relatively good

approximation of the covariance matrix is available, and (ii) the density of the current point is close to the desired density, then only $O^*(n^3)$ steps of the random walk are necessary to generate a random point whose distribution is a good approximation of the desired (target) distribution.

In our context, recall that the distribution of interest f_{t_k} is changing at each iteration. The current approximation of the covariance matrix $\hat{V}_k$ will be used as the approximation to the true covariance matrix of the next iteration V_{k+1}, which in turn will be estimated by $\hat{V}_{k+1}$. In a similar way, the current point X_{t_k}, distributed approximately according to f_{t_k}, will be used as the starting point for the random walk to approximate a random variable distributed according to $f_{t_{k+1}}$. The parameter α, which dictates the factor by which t is decreased at each iteration, will be the largest value such that these approximations are valid from a theoretical perspective.

6.1 A geometric random walk: hit-and-run. There are several possible choices of geometric random walks. We refer to [18] for a recent survey. Here we use the so-called hit-and-run random walk. The implementation of this random walk requires the following as input: a density function f (known up to a multiplicative constant), a starting point X^0, a covariance matrix V, and a number of steps m.

Subroutine: *hit-and-run* $\left(f,\ X^0,\ V,\ m\right)$

Step 0 Set $k \leftarrow 0$.
Step 1 Pick a random vector $d \sim N(0, V)$.
Step 2 Define the line $\ell(X^k, d) = \{X^k + td :\ t \in \mathbb{R}\}$.
Step 3 Move to a point X^{k+1} chosen according to f restricted to $\ell(X^k, d)$.
Step 4 Set $k \leftarrow k + 1$. If $k \leq m$, goto **Step 1**.
Step 5 Report X^m.

Although hit-and-run can be implemented for arbitrary densities, we will restrict ourselves to the case of logconcave densities. In such case, the implementation of hit-and-run can be done efficiently, and we refer to the Appendix for a complete description for the case in which f_t is defined as in (1) with $\|\cdot\| = \|\cdot\|_2$, the Euclidean norm, which is the proposed density for the algorithm.

6.2 Sampling complexity. Here we state without proof results in the literature of sampling random points according to logconcave densities. We start with a complexity result for the mixing time of the hit-and-run random walk. Recall that the distribution associated with a density f is given by $\pi_f(A) = \int_A f(x)dx$, and the mean associated with f is $z_f = E_f[X] = \int x f(x)dx$.

THEOREM 6.1 *([11] Theorem 1.1) Let f be a logconcave density and let X denote the associated random variable. Moreover, f is such that (i) the level set of probability $1/8$ contains a ball of radius s, (ii) $E_f\left[\|X - z_f\|_2^2\right] \leq S^2$, and (iii) the L_2 norm of the starting distribution σ with respect to the stationary distribution π_f is at most M. Let σ^m be the distribution of the current point after m steps of hit-and-run applied to π_f with $V = I$. Then, after*

$$m = O\left(\frac{n^2 S^2}{s^2} \ln^5\left(\frac{nM}{\varepsilon}\right)\right) \quad steps,$$

the total variation distance of σ^m and f is at most ε.

Theorem 6.1 bounds the rate of convergence of the geometric random walk not only by the dimension but also by the L^2 norm of the starting density with respect to the stationary density f_t, and by how "well-rounded" is f_t via the ratio S/s. The notion of "well-rounded" is deeply connected with the concept of near isotropic position. The next lemma quantifies this connection.

LEMMA 6.1 *([12] Lemma 5.13) Let f be a density in C-isotropic position. Then $E_f\left[\|X - z_f\|_2^2\right] \leq Cn$, and any upper level set U of f contains a ball of radius $\pi_f(U)/e\sqrt{C}$.*

Any (full-dimensional) density can be put in near-isotropic position by a suitable linear transformation. By using an approximation of the covariance matrix $\hat{V}$ to implement the hit-and-run random walk such that all eigenvalues of $\hat{V}^{-1}V$ are between $1/C$ and C, f_t is in C-isotropic position. Thus, the ratio

S/s can be bounded by $8eC\sqrt{n}$ (in this case, note that only $m = O^*(n^3)$ steps of the random walk will be necessary to generate one random sample). The next section summarizes how to compute an approximation $\hat{V}_k$ which puts f_{t_k} in 2-isotropic position. Moreover, it will be shown that all densities simulated by the algorithm will be at most C-isotropic for a constant C (independent of the dimension) as we decrease the homotopy parameter t.

REMARK 6.1 *In our analysis we will be assuming independence among different samples for simplicity (recall that they are separated by $m = O^*(n^3)$ steps of the random walk). Although this is not the case, independence can be approximated at the cost of an additional constant factor in the number m of steps of the random walk. Here we have chosen exposition over formalism since no additional insight is gained if we work out all the details.*

7. Estimating the covariance matrix, the mean, and the second moment matrix. In this section, we recall estimation results for the mean and covariance matrix of a logconcave random variable. Moreover, we show that these estimates can be used to approximate the second moment matrix with a desired relative accuracy. Herein it will be assumed that independent identically distributed samples $\{X^i\}$ are available. We emphasize that these estimations depend only on the samples and not on the isotropic position of the density function. As stated before, the isotropic position plays an important role to bound the number of steps of the chain required to obtain each sample.

First we recall a result for estimating the mean and covariance matrix.

LEMMA 7.1 *([12] Lemma 5.17) Let z and V denote respectively the mean and covariance matrix of a logconcave random variable. For any $\xi > 0$ and $\delta \in (0, 1)$, using*

$$N > \frac{4}{\xi^2} n \ln^2 \frac{1}{\delta} \quad \text{samples and} \quad \hat{X} = \frac{1}{N} \sum_{i=1}^{N} X^i,$$

we have that

$$P\left(\left\| V^{-1/2} \left(\hat{X} - z\right) \right\|_2 > \xi \right) \leq \delta.$$

LEMMA 7.2 *([7] Corollary A.2) Let V denote the covariance matrix of a logconcave random variable. Using $N > O\left(n \ln^2 n \ \ln^3\left(\frac{1}{\delta}\right)\right)$ samples, where $\delta < 1/n$, define*

$$\hat{V} = \frac{1}{N} \sum_{i=1}^{N} (X^i - \hat{X})(X^i - \hat{X})^T$$

we have that all eigenvalues of the matrix $\hat{V}^{-1}V$ are in the interval $[1/2,\ 2]$ with probability at least $1 - \delta$.

These results yield a useful estimation procedure for the second moment matrix of a logconcave random variable . In particular the maximum eigenvalue of such matrix, which is used in Step 4 in the homotopy algorithm to test for boundedness, will be estimated up to a (known) constant factor.

LEMMA 7.3 *Let $\Omega := E\left[XX'\right]$ denote the second moment matrix associated with a logconcave random variable X. Then for $\xi < 1/4$ and using $N > O(\ln^3 \frac{1}{\delta} n \ln^2 n + \frac{1}{\xi^2} n \ln^2 \frac{1}{\delta})$, for $\delta < 1/n$, the matrix*

$$\hat{\Omega} = \hat{V} + \hat{X}\hat{X}^T$$

is such that all eigenvalues of $\hat{\Omega}^{-1}\Omega$ are in he interval $[(1/2 - \xi),\ (2 + \xi + \xi^2)]$ with probability at least $1 - \delta$.

PROOF. In this proof let $\|\cdot\| = \|\cdot\|_2$. Lemma 7.1 yields that $\|V^{-1/2}(\hat{X} - z)\| \leq \xi$ with probability greater than $1 - \delta/2$. In this event, there exists $d \in \mathbb{R}^n$ with $\|d\| \leq 1$ satisfying

$$\hat{X} = z + \xi V^{1/2} d.$$

For any $w \in \mathbb{R}^n$, we have that

$$
\begin{aligned}
\left\langle w, \hat{\Omega}w \right\rangle &= \left\langle w, \hat{V}w \right\rangle + \left\langle \hat{X}, w \right\rangle^2 = \left\langle w, \hat{V}w \right\rangle + \left\langle z + \xi V^{1/2}d, w \right\rangle^2 \\
&= \left\langle w, \hat{V}w \right\rangle + \langle z, w \rangle^2 + 2\xi \langle z, w \rangle \langle d, V^{1/2}w \rangle + \xi^2 \langle d, V^{1/2}w \rangle^2 \\
&\geq \tfrac{1}{2} \langle w, Vw \rangle + \tfrac{1}{2} \langle z, w \rangle^2 - \xi \langle w, \Omega w \rangle = (\tfrac{1}{2} - \xi) \langle w, \Omega w \rangle .
\end{aligned}
\tag{16}
$$

where we used the geometric/arithmetic mean inequality that yields $|\langle z, w \rangle| \|V^{1/2}w\| \leq \frac{\langle z, w \rangle^2 + \langle w, Vw \rangle}{2}$. The upper bounds follows from the same argument. $\square$

We will use the following corollary in our algorithm.

COROLLARY 7.1 *Using $N = O\left(n \ln^2 n \ln^3 \left(\frac{1}{\delta}\right)\right)$ samples to estimate the second moment matrix $\hat{\Omega}$, with probability at least $1 - \delta$ we have that all eigenvalues of $\hat{\Omega}^{-1}\Omega$ are between $1/3$ and 3.*

PROOF. Set $\xi = 1/24$ and apply Lemma 7.3. $\square$

Thus, with $O^*(n)$ samples per iteration one can properly estimate the mean, covariance matrix, and second moment matrix to conduct the algorithm.

8. Updating the parameter t: warm-start. Since the parameter α controls how fast the sequence $\{t_k\}$ decreases, its value is critical to the computational complexity of the algorithm. Although we are tempted to decrease t as fast as possible, we still need to use the current density as a "warm-start" to approximate the next density. That is, the L^2 norm of f_{t_k} with respect to $f_{t_{k+1}}$ needs to be controlled. Moreover, the covariance matrix associated with f_{t_k} should also be close to the covariance matrix of the next iterate $f_{t_{k+1}}$. The trade-off among these quantities will decide how fast one can decrease the parameter t. Kalai and Vempala show how to relate the L^2 norm of two logconcave densities and their covariance matrices. For the reader's convenience we re-state their results here. Recall that given two densities f and g, the L^2 norm of f with respect to g is defined by $\|f/g\| = E_f[f(X)/g(X)]$ (see Section 1.2 for details).

LEMMA 8.1 *([7] Lemma 3.8) Let f and g be logconcave densities over K with mean $z_f = E_f[X]$ and $z_g = E_g[X]$, respectively. Then, for any $c \in \mathbb{R}^n$*

$$E_f\left[\langle c, X - z_f \rangle^2\right] \leq 16 \|f/g\| E_g\left[\langle c, X - z_g \rangle^2\right]$$

This result implies that it suffices to control the L^2 norm of $f_{t_{k+1}}$ with respect to f_{t_k} in order to guarantee that the covariance matrices of these densities are relatively close. A simple extension of Lemma 3.11 of [7] yields a bound this L^2 norm.

COROLLARY 8.1 *Consider the densities f_{t_k} and $f_{t_{k+1}}$ such that $t_{k+1} = (1 - \alpha)t_k = (1 - \frac{1}{\sqrt{n}})t_k$, i.e., $\alpha = 1/\sqrt{n}$. For $n \geq 5$,*

$$\|f_{t_k}/f_{t_{k+1}}\| \leq e^{n/(n-1)} \quad \text{and} \quad \|f_{t_{k+1}}/f_{t_k}\| \leq e^{n/(n-2\sqrt{n})}.$$

The next lemma combines the previous results to ensure that all the densities used in the algorithm will be in 2^8-isotropic position.

LEMMA 8.2 *Assuming that $n \geq 16$, and using $N = O^*(\kappa^3 n)$ samples at any iteration, the distribution encountered by the sampling algorithm in the next iteration is 2^8-isotropic with probability $1 - 2^{-\kappa}$.*

PROOF. By using Lemma 7.2, we have that f_{t_k} is 2-isotropic after we estimated its covariance matrix in Step 2 of the algorithm with probability $1 - 2^{-\kappa}$. Using Lemma 8.1 and Corollary 8.1, for any $v \in S^{n-1}$,

$$E_{f_{t_{k+1}}}\left[\left\langle v, x - z_{f_{t_{k+1}}} \right\rangle^2\right] \leq 16 \, e^{n/(n-2\sqrt{n})} E_{f_{t_k}}\left[\left\langle v, x - z_{f_{t_k}} \right\rangle^2\right] \leq 32 \, e^{n/(n-2\sqrt{n})} \leq 2^8,$$

since $n \geq 16$. $\square$

Lemma 8.2 ensures that $O^*(n)$ samples suffice to estimate the current covariance matrix accurately enough to be used as an approximation to the covariance matrix associated with the next iteration.

9. Controlling the overall probability of failure. Before we proceed to the proof of Theorem 4.1, we prove a technical lemma which allows us to properly control the overall probability of failure of the algorithm. First note that if the algorithm finds an element of C_K it will always stop correctly. Thus, the algorithm can fail only if it (wrongly) declares that K is a bounded set or if it does not terminate. The latter concerns with the stopping criteria of the algorithm and it was already analyzed in Sections 5.1 and 5.2. The first issue can occur only if the estimated second moment matrix $\hat{\Omega}_k$ differs significantly from the true matrix Ω_{t_k}. In turn their difference is controlled by the number of samples used to estimate $\hat{\Omega}_{t_k}$, which depends on the probability of failure used for the current iteration. Recall that we do not know the total number of iterations T a priori (since we do not have any estimates for R or τ_{C_K}), and so we cannot simply set the probability of failure to δ/T for each iteration. Instead of using a constant probability of failure for all iterations, we slightly increase the number of samples used to estimate $\hat{\Omega}_{t_k}$ at each iteration to ensure that the probability of failure will not accumulate indefinitely.

LEMMA 9.1 *Let T be the total number of iterations of the algorithm and $1 - \delta_f$ the desired probability of success. Initializing the algorithm with $\delta = \delta_f/4$ and setting the probability of failure of the ith iteration to $\delta_i = \frac{1}{i^2}\delta_f/4$, we obtain:*
(i) the probability of success is at least $1 - \delta_f/2$;
(ii) The smallest δ used by the algorithm satisfies $\delta \geq \frac{\delta_f}{4}\frac{1}{T^2}$.

PROOF. (i) By setting a probability of failure of the ith iteration of the algorithm to $(1/i^2)\delta_f/4$, we have that the total probability of failure is bounded by

$$\sum_{i=1}^{T} \delta_i = \sum_{i=1}^{T} \frac{\delta_f}{4i^2} \leq \frac{\delta_f}{4}\sum_{i=1}^{\infty} \frac{1}{i^2} \leq \delta_f/2$$

since $\sum_{i=1}^{\infty} \frac{1}{i^2} = \pi^2/6 < 2$.

(ii) follows since the algorithm terminates after T iterations.

$\square$

10. Proof of Theorem 4.1. Let $\alpha = \frac{1}{\sqrt{n}}$ and let $1 - \delta_f$ be the desired probability of success. The algorithm terminates in $T = \sqrt{n}\ln\left(\frac{t_{initial}}{t_{final}}\right)$ iterations. Next we need to: *(i)* properly bound $t_{initial}$, *(ii)* ensure that the algorithm terminates, i.e., properly bound t_{final} (which also bounds T), and *(iii)* control the probability that it wrongly terminates.

(i) Corollary 5.3 yields $t_{initial} = t_0 = 8n^{3/2}/\hat{r}$ which is at least $8n/r$ and at most $8n^{3/2}/r$.

(ii) We will use Corollaries 5.2 and 5.1 with $\delta = \delta_f/2$. Thus, the algorithm will correctly terminate with probability at least $1 - \delta_f/2$ after at most

$$\overline{T} = 6\ln\frac{1}{\delta_f/2} + \sqrt{n}\ln\left((8n^{3/2}/r)(96enR_\epsilon)/((1/2)^2\tau_{C_K})\right)$$

iterations if K is unbounded, and after at most

$$\overline{T} = \sqrt{n}\ln\left((8n^{3/2}/r)21eR\sqrt{n}\right)$$

iterations if K is bounded (note that we use 21 instead of 7 to take into account a factor of 3 in the approximation of the maximum eigenvalue of V_k).

(iii) Finally, we ensure that the probability of (wrongly) terminating before reaching iteration $\overline{T}$ is at most $\delta_f/2$. This is achieved by slightly reducing the probability of failure as described in Lemma 9.1.

Now we analyze the impact of reducing the probability of failure on the computational complexity of each iteration. Lemma 9.1(*ii*) ensures that it is sufficient to use

$$\delta = \frac{\delta_f/2}{4\overline{T}^2} = \frac{\delta_f}{32n\ln^2\left(t_{initial}/t_{final}\right)}.$$

This will dictate the number of samples needed to estimate the covariance matrix. In Lemma 8.2 we will need at most

$$\kappa = \frac{1}{\ln 2}\ln\frac{32\overline{T}^2}{\delta_f}$$

at any given iteration (given that we terminate correctly). That leads to a total of at most

$$O\left(n\ln^3\left(\frac{\overline{T}}{\delta_f}\right)\right)$$

samples per iteration of the algorithm where each sample requires $O^*(n^3)$ steps of the geometric random walk.

Defining a condition measure $\mu := \frac{R}{r}$ if K is bounded, or $\mu := \frac{R_\epsilon}{r}\frac{1}{\tau_{C_K}}$ if K is unbounded, the overall complexity is

$$O^*\left(\overline{T}\,n^4\ln^3\left(\frac{\overline{T}}{\delta_f}\right)\right) = O^*\left(n^{4.5}\ln\left(\frac{\mu}{\delta_f}\right)\,\ln^3\left(\frac{n\ln(n\mu/\delta)}{\delta_f}\right)\right).$$

The results of Theorem 4.1 follow by using the appropriate $\overline{T}$ on each case as defined in (ii).

11. Conclusions. In this work we study probability densities f_t induced by an arbitrary norm with a convex support K as defined by (1). Our goal is to relate geometric properties of K with analytical properties of f_t.

Using these properties, we also develop an algorithm to test whether a convex set, given only by a membership oracle, is bounded or unbounded. The computational complexity of the probabilistic method proposed is polynomial in the dimension of the set and only logarithmic in other condition measures (such as the width of the recession cone and the diameter of the set, respectively, for the unbounded and bounded cases).

Clearly, one can derive a similar method based on uniform densities, instead of norm-induced densities, by defining $f_t(x) \sim 1_K(x)1_{B(0,1/t)}(x)$. These densities would be well-defined and many results of Section 3 can be adapted for this case as well. Although this could lead to a simpler analysis, the parameter α that is used to decrease t on each iteration must be set to $1/n$ instead of $1/\sqrt{n}$ for the current density be a good "starting point" for the next density. This leads to an algorithm of the order $O^*(n^5)$.

Exploiting the geometry of unbounded convex sets was key in developing such probabilistic method. A concentration phenomenon guarantees that most points of an unbounded convex set will be in its recession cone if that cone has positive width. In such a case, random points with large norms are likely to belong to the recession cone which yields a certificate of unboundedness. In the case of the set being bounded, the second moment matrix associated with the random variable must also be bounded, in opposition to the unbounded case, which yields a certificate of boundedness.

In contrast with probabilistic methods over convex sets in the literature, we need to explicitly deal with unbounded sets, and additional effort is needed to ensure that all the densities are well defined. Moreover, if K is unbounded our analysis shows that an element of C_K^ϵ can be computed in

$$O^*\left(\sqrt{n}\ln\left(\frac{R_\epsilon}{r\epsilon}\right)\right) \quad \text{iterations.}$$

That is particularly relevant for the ill-conditioned case of $\tau_{C_K} = 0$. Although we cannot find an element of C_K (which has zero volume), the algorithm will generate a direction d, $\|d\| = 1$, such that $\mathbf{dist}(d, C_K) < \epsilon$.

References

[1] K. Ball, *Logarithmically concave functions and sections of convex sets in* IR^n, Studia Math. **88** (1988), no. 1, 69–84.

[2] D. Bertsimas and S. Vempala, *Solving convex programs by random walks*, Journal of the ACM **51** (2004), no. 4, 540–556.

[3] G. Casella and C. P. Robert, *Monte carlo statistical methods*, Springer, 1999.

[4] G. de Ghellink and J.-P. Vial, *A polynomial Newtow method for linear programming*, Algorithmica (1986), no. 1, 425–453.

[5] M. Gromov and V. Milman, *Brunn Theorem and a concentration of volume of convex bodies*, GAFA Seminar Notes, Tel Aviv University **V** (1984), 12pp.

[6] A. Iusem and A. Seeger, *Axiomatization of the index of pointedness for closed convex cones*, Mat. apl. comput. **24** (2005), no. 2, 245–283.

[7] A. Kalai and S. Vempala, *Simulating annealing for convex optimization*, to appear on Mathematic of Operations Research (2006).

[8] L. G. Khachiyan, *Rounding of polytopes in the real number model of computation*, Mathematics of Operations Research **21** (1996), no. 2, 307–320.

[9] L. Lovász, *Hit-and-run mixes fast*, Mathematical Programming **86** (1998), 443–461.

[10] L. Lovász and M. Simonovits, *On the randomized complexity of volume and diameter*, Proc. 33rd IEEE FOCS (1992), 482–491.

[11] L. Lovász and S. Vempala, *Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization*, Submitted. (2006).

[12] _____, *The geometry of logconcave functions and sampling algorithms*, to appear in RSA (2006).

[13] _____, *Hit-and-run from a corner*, SIAM Journal of Computing **35** (2006), no. 4, 985–100.

[14] _____, *Simulated annealing in convex bodies and an $o^*(n^4)$ volume algorithm*, Journal of Computer and System Sciences **72** (2006), 392–417.

[15] A. Naor and D. Romik, *Projecting the surface measure of the sphere of ℓ_p^n*, Annales de l'Institut Henri Poincare' (B), Probability and Statistics **39** (2003), 241–261.

[16] Y. Nesterov, M. Todd, and Y. Ye, *Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems*, Mathematical Programming **84** (1999), 227–267.

[17] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[18] S. Vempala, *Geometric random walks: A survey*, Combinatorial and Computational Geometry **52** (2005), 573–612.

Appendix A. Hit-and-run for $f(x) \sim e^{-\|x\|_2}$. Here we specialize the method for implementing the hit-and-run method for logconcave densities proposed in [12] to the densities defined as in (1) with the Euclidean norm.

Given $x_0 \in K$, $d \in S^{n-1}$, define the line segment $\ell(x_0, d) = K \cap \{x_0 + td : t \in \mathbb{R}\}$ (note that if it is not a line segment, either d or $-d$ belongs to C_K). The next iterate should be chosen from $\ell(x_0, d)$ according to the density f. Let $M_\ell = \max_{y \in \ell(x_0,d)} f(y)$ denote the maximum density on $\ell(x_0, d)$ and, for $v \in (0, M_\ell)$, let $L_\ell(v)$ denote the corresponding level set restricted to $\ell(x_0, d)$.

It is a two step method:

Step 1. Choose v uniformly on $(0, M_\ell)$;

Step 2. Choose y uniformly on $L_\ell(v)$.

To implement this, we need to compute three quantities: M_ℓ and the two endpoints of $L_\ell(v)$. It is possible to speed up the process by using closed-form formulas since we know the functional form of f (although additional effort is necessary to adjust it for the support of f).

On $\ell(x_0, d)$, f can be written as $f(x_0 + td) \sim e^{-\sqrt{\|x_0\|^2 + 2t\langle x_0, d\rangle + t^2\|d\|^2}}$. It is convenient to note that

$$t^* = -\frac{\langle x_0, d\rangle}{\|d\|^2} \in \arg\max\{f(x_0 + td) : t \in \mathbb{R}\},$$

that is, for $x^* = x_0 + t^* d$, $f(x^*) \geq M_L$. If $x^* \in K$, we have found M_ℓ, otherwise, we need to make a binary search on $[x_0, x^*]$ to find it (note that in the second case we already have one of the endpoints of $L_\ell(v)$).

After drawing $v \in (0, M_\ell)$, again we can compute the explicit "unrestricted" point of where the endpoints should be

$$\hat{t} \in \left\{t : v = e^{-\sqrt{\|x_0\|^2 + 2t\langle x_0, d\rangle + t^2\|d\|^2}}\right\}$$

or equivalently, the solution for the following quadratic equation

$$\hat{t}^2\|d\|^2 + 2\hat{t}\langle x_0, d\rangle + \|x_0\|^2 + \ln^2 v = 0.$$

Again, if the solutions $\hat{t}$ lie in K we found an endpoint; otherwise, we need to perform binary search on $[x_0, x_0 + \hat{t}d]$.