Sampling s-Concave Functions: The Limit of Convexity Based Isoperimetry

Karthekeyan Chandrasekaran[1], Amit Deshpande[2], and Santosh Vempala[1]

[1] School of Computer Science, Georgia Institute of Technology
karthe@cc.gatech.edu, vempala@cc.gatech.edu
[2] Microsoft Research India
amitdesh@microsoft.edu

Abstract. Efficient sampling, integration and optimization algorithms for logconcave functions [BV04, KV06, LV06a] rely on the good isoperimetry of these functions. We extend this to show that $-1/(n-1)$-concave functions have good isoperimetry, and moreover, using a characterization of functions based on their values along every line, we prove that this is *the largest class* of functions with good isoperimetry in the spectrum from concave to quasi-concave. We give an efficient sampling algorithm based on a random walk for $-1/(n-1)$-concave probability densities satisfying a smoothness criterion, which includes heavy-tailed densities such as the Cauchy density. In addition, the mixing time of this random walk for Cauchy density matches the corresponding best known bounds for logconcave densities.

1 Introduction

Given a function $f : \mathbb{R}^n \to \mathbb{R}_+$, accessible by querying the function value at any point $x \in \mathbb{R}^n$, and an error parameter $\epsilon > 0$, three fundamental problems are: (i) Integration: estimate $\int f$ to within $1 \pm \epsilon$, (ii) Maximization: find x that approximately maximizes f, i.e., $f(x) \geq (1-\epsilon)\max f$, and (iii) Sampling: generate x from density π with $d_{tv}(\pi, \pi_f) \leq \epsilon$ where d_{tv} is the total variation distance and π_f is the density proportional to f. The complexity of an algorithm is measured by the number of queries for the function values.

The most general class of functions for which these problems are known to have compexity polynomial in the dimension, is the class of logconcave functions. A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is logconcave if its logarithm is concave on its support, i.e., for any two points $x, y \in \mathbb{R}^n$ and any $\lambda \in (0,1)$,

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

Logconcave functions generalize indicator functions of convex bodies (and hence the problems subsume convex optimization and volume computation) as well as Gaussians. Following the polynomial time algorithm of Dyer, Frieze and Kannan [DFK91] for estimating the volume of a convex body, a long line of work

I. Dinur et al. (Eds.): APPROX and RANDOM 2009, LNCS 5687, pp. 420–433, 2009.

[AK91, Lov90, DF91, LS92, LS93, KLS97, LV07, LV06c, LV06b] culminated in the results that both sampling and integration have polynomial complexity for any logconcave density. Integration is done by a reduction to sampling and sampling also provides an alternative to the Ellipsoid method for optimization [BV04, KV06, LV06a]. Sampling itself is achieved by a random walk whose stationary distribution has density proportional to the given function. The key question is thus the rate of convergence of the walk, which depends (among other things) on the isoperimetry of the target function.

Informally, a function has good isoperimetry if one cannot remove a set of small measure from its domain and partition it into two disjoint sets of large measure. Logconcave functions satisfy the following isoperimetric inequality:

Theorem 1. *[DF91, LS93] Let $f : \mathbb{R}^n \to \mathbb{R}_+$ be a logconcave function with a convex support K of diameter D, $\int_{\mathbb{R}^n} f < \infty$, and S_1, S_2, S_3 be any partition of K into three measurable sets. Then, for a distribution π_f with density proportional to f,*

$$\pi_f(S_3) \geq \frac{2d(S_1, S_2)}{D} \min\{\pi_f(S_1), \pi_f(S_2)\},$$

where $d(S_1, S_2)$ refers to the minimum distance between any two points in S_1 and S_2.

Although the class of logconcave functions is fairly large, it does not capture all the functions with good isoperimetry. The definition of logconcavity says that, for every line segment in the domain, the value at its midpoint is at least the geometric mean of the values at its endpoints. This is a generalization of concavity where, for every line segment in the domain, the value at its midpoint is at least the arithmetic mean of the values at its endpoints. This motivates the following question: What condition should a function satisfy along every line segment to have good isoperimetry?

In this paper, using a characterization of functions based on generalized means, we present a class of functions with good isoperimetry that is the largest under this particular characterization. We also give an efficient algorithm to sample from these functions; a well-known example among these is the Cauchy density (which is not logconcave and is heavy-tailed).

To motivate and state our results, we begin with a discussion of one-dimensional conditions.

1.1 From Concave to Quasi-concave

Definition 1. *(s-concavity of probability density) A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is said to be s-*concave*, for $-\infty \leq s \leq 1$, if*

$$f\left(\lambda x + (1-\lambda)y\right) \geq \left(\lambda f(x)^s + (1-\lambda) f(y)^s\right)^{1/s},$$

for all $\lambda \in [0,1], \forall x, y \in \mathbb{R}^n$.

The following are some special cases: A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is said to be

$$\begin{cases} \textit{concave} \text{ if,} & f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda) f(y) \\ \textit{logconcave} \text{ if,} & f(\lambda x + (1-\lambda)y) \geq f(x)^{\lambda} f(y)^{1-\lambda} \\ \textit{harmonic-concave} \text{ if,} & f(\lambda x + (1-\lambda)y) \geq \left(\frac{\lambda}{f(x)} + \frac{(1-\lambda)}{f(y)}\right)^{-1} \\ \textit{quasi-concave} \text{ if,} & f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\} \end{cases}$$

for all $\lambda \in [0,1], \forall x, y \in \mathbb{R}^n$.

These conditions are progressively weaker, restricting the function value at a convex combination of x and y to be at least the arithmetic average, geometric average, harmonic average and minimum, respectively. Note that s_1-concave functions are also s_2-concave if $s_1 > s_2$. It is thus easy to verify that:

concave $\subsetneq$ s-concave ($s > 0$) $\subsetneq$ logconcave $\subsetneq$ s-concave ($s < 0$) $\subsetneq$ quasi-concave.

Relaxing beyond quasi-concave would violate unimodality, i.e., there could be two distinct local maxima, which appears problematic for all of the fundamental problems. Also, it is well-known that quasi-concave functions have poor isoperimetry.

There is a different characterization of probability measures based on a generalization of the Brunn-Minkowski inequality. The Brunn-Minkowski inequality states that the Euclidean volume (or Lebesgue measure) μ satisfies

$$\mu(\lambda A + (1-\lambda)B)^{1/n} \geq \lambda \mu(A)^{1/n} + (1-\lambda)\mu(B)^{1/n},$$

for $\lambda \in [0,1]$ and compact subsets $A, B \subseteq \mathbb{R}^n$, where $\lambda A + (1-\lambda)B = \{\lambda a + (1-\lambda)b \,:\, a \in A, b \in B\}$ is the Minkowski sum.

Definition 2. *(κ-concavity of probability measure) A probability measure μ over $\mathbb{R}^n$ is κ-concave if*

$$\mu(\lambda A + (1-\lambda)B)^{\kappa} \geq \lambda \mu(A)^{\kappa} + (1-\lambda)\mu(B)^{\kappa},$$

$\forall A, B \subseteq \mathbb{R}^n$, $\forall \lambda \in [0,1]$.

Note that the Euclidean volume (or Lebesgue measure) is quasi-concave according to Definition 1 but $1/n$-concave according to Definition 2. Borell [Bor74, Bor75] showed an equivalence between these two definitions as follows.

Lemma 1. *An absolutely continuous probability measure μ on a convex set $K \subseteq \mathbb{R}^n$ is κ-concave, for $-\infty < \kappa \leq 1/n$, if and only if there is a density function $p : \mathbb{R}^n \to \mathbb{R}_+$, which is s-concave for $s = \dfrac{\kappa}{1-\kappa n}$.*

Thus, if the density function is s-concave for $s \in [-1/n, 0]$, then the corresponding probability measure is κ-concave for $\kappa = \frac{s}{1+ns}$. Bobkov [Bob07] proves the following isoperimetric inequality for κ-concave probability measures for $-\infty < \kappa \le 1$.

Theorem 2. *Given a κ-concave probability measure μ, for any measurable subset $A \subseteq \mathbb{R}^n$,*

$$\mu(\delta A) \ge \frac{c(\kappa)}{m} \min\{\mu(A), 1 - \mu(A)\}^{1-\kappa}$$

where m is the μ-median of the Euclidean norm $x \mapsto \|x\|$, for some constant $c(\kappa)$ depending on κ.

Therefore, by Lemma 1, we get an isoperimetric inequality for any s-concave function $f : \mathbb{R}^n \to \mathbb{R}_+$, for $s \in [-1/n, 0]$, as

$$\pi_f(\delta A) \ge \frac{c(s)}{m} \min\{\pi_f(A), 1 - \pi_f(A)\}^{1-\frac{s}{1+ns}},$$

for any measurable set $A \subseteq \mathbb{R}^n$.

In comparison, we prove a stronger isoperimetric inequality for the class of $-1/(n-1)$-concave functions (which subsumes $-1/n$-concave functions) and we remove the dependence on s in the inequality completely.

1.2 The Cauchy Density

The generalized Cauchy probability density $f : \mathbb{R}^n \to \mathbb{R}_+$ parameterized by a positive definite matrix $A \in \mathbb{R}^{n \times n}$ and a vector $m \in \mathbb{R}^n$, is given by

$$f(x) \propto \frac{\det(A)^{-1}}{\left(1 + \|A(x-m)\|^2\right)^{(n+1)/2}}.$$

For simplicity, we assume $m = \bar{0}$ using a translation. It is easy to sample this distribution in full space (by an affine transformation it becomes spherically symmetric and therefore a one-dimensional problem) [Joh87]. We consider the problem of sampling according to the Cauchy density restricted to a convex set. This is reminiscent of the work of Kannan and Li who considered the problem of sampling a Gaussian distribution restricted to a convex set [KL96].

1.3 Our Results

Our first result establishes good isoperimetry for $-1/(n-1)$-concave functions in $\mathbb{R}^n$.

Theorem 3. *Let $f : \mathbb{R}^n \to \mathbb{R}_+$ be a $-1/(n-1)$-concave function with a convex support $K \subseteq \mathbb{R}^n$ of diameter D, and let $\mathbb{R}^n = S_1 \cup S_2 \cup S_3$ be a measurable partition of $\mathbb{R}^n$ into three non-empty subsets. Then*

$$\pi_f(S_3) \ge \frac{d(S_1, S_2)}{D} \min\{\pi_f(S_1), \pi_f(S_2)\}.$$

It is worth noting that the isoperimetric coefficient above is only smaller by a factor of 2 when compared to that of logconcave functions (Theorem 1).

Next, we prove that beyond the class of $-1/(n-1)$-concave functions, there exist functions with exponentially small isoperimetric coefficient.

Theorem 4. *For any $\epsilon > 0$, there exists a $-1/(n-1-\epsilon)$-concave function $f : \mathbb{R}^n \to \mathbb{R}_+$ with a convex support K of finite diameter and a partition $\mathbb{R}^n = S \cup T$ such that*

$$\frac{\pi_f(\partial S)}{\min\{\pi_f(S), \pi_f(T)\}} \leq Cn(1+\epsilon)^{-\epsilon n}$$

for some constant $C > 0$.

Theorems 3 and 4 can be summarized by the following figure.

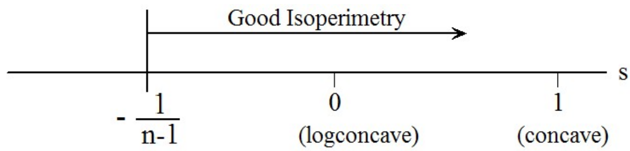

Fig. 1. Limit of isoperimetry for s-concave functions

We prove that the ball walk with a Metropolis filter can be used to sample efficiently according to $-1/(n-1)$-concave densities which satisfy a certain Lipschitz condition. In each step, the ball walk picks a new point y, uniformly at random from a small ball around the current point x, and moves to y with probability $\min\{1, f(y)/f(x)\}$. A distribution σ_0 is said to be an H-warm start ($H > 0$) for the distribution π_f if for all $S \subseteq \mathbb{R}^n$, $\sigma_0(S) \leq H\pi_f(S)$. Let σ_m denote the distribution after m steps of the ball walk with a Metropolis filter.

Definition 3. *We call a function $f : \mathbb{R}^n \to \mathbb{R}_+$ to be (α, δ)-smooth if*

$$\max\left\{\frac{f(x)}{f(y)}, \frac{f(y)}{f(x)}\right\} \leq \alpha,$$

for all x, y in the support of f with $\|x - y\| \leq \delta$.

Theorem 5. *Let $f : \mathbb{R}^n \to \mathbb{R}_+$ be proportional to an s-concave (α, δ)-smooth function, restricted to a convex body $K \subseteq \mathbb{R}^n$ of diameter D, where $s \geq -1/(n-1)$. Let K contain a ball of radius δ and σ_0 be an H-warm start. Then, after*

$$m \geq \left(\frac{CnD^2}{\delta^2} \log \frac{2H}{\epsilon}\right) \cdot \max\left\{\frac{nH^2}{\epsilon^2}, \frac{(\alpha^{-s}-1)^2}{s^2}\right\}$$

steps of the ball walk with radius $r \leq \min\left\{\frac{\epsilon\delta}{16H\sqrt{n}}, \frac{|2s\delta|}{\alpha^{-s}-1}\right\}$, *we have that*

$$d_{tv}(\sigma_m, \pi_f) \leq \epsilon,$$

for some absolute constant C, where $d_{tv}(\cdot,\cdot)$ *is the total variation distance.*

Applying the above theorem directly to sample according to the Cauchy density, we get a mixing time of $O\left(\left(\frac{n^3H^2}{\epsilon^2}\log\frac{2H}{\epsilon}\right)\cdot\max\left\{\frac{H^2}{\epsilon^2}, n\right\}\right)$ using parameters $\delta = 1$, $\alpha = e^{\frac{n+1}{2}}$ and, $D = \frac{8\sqrt{2n}H}{\epsilon}$ (one can prove that the probability measure outside the ball of radius D around the origin is at most $\epsilon/2H$ for the chosen value of D). Using a more careful analysis (comparison of 1-step distributions), this bound can be improved to match the current best bounds for sampling logconcave functions.

Theorem 6. *Consider the Cauchy probability density f defined in Section 1.2, restricted to a convex set* $K \subseteq \mathbb{R}^n$ *containing a ball of radius* $\|A^{-1}\|_2$ *and let* σ_0 *be an H-warm starting distribution. Then after*

$$m \geq O\left(\frac{n^3H^4}{\epsilon^4}\log\frac{2H}{\epsilon}\right)$$

steps with ball-walk radius $r = \epsilon/8\sqrt{n}$, *we have*

$$d_{tv}(\sigma_m, \pi_f) \leq \epsilon,$$

where $d_{tv}(.,.)$ *is the total variation distance.*

The proof of this theorem departs from its earlier counterparts in a significant way. In addition to isoperimetry, and the closeness of one-step distributions of nearby points, we have to prove that most of the measure is contained in a ball of not-too-large radius. For logconcave densities, this large-ball probability decays exponentially with the radius. For the Cauchy density it only decays linearly (Proposition 3).

All missing proofs are available in the full version of the paper[1].

2 Preliminaries

Let $r\mathbb{B}_x$ denote a ball of radius r around point x. One step of the ball walk at a point x defines a probability distribution P_x over $\mathbb{R}^n$ as follows.

$$P_x(S) = \int_{S\cap r\mathbb{B}_x} \min\left\{1, \frac{f(y)}{f(x)}\right\} dy.$$

For every measurable set $S \subseteq \mathbb{R}^n$ the ergodic flow from S is defined as

$$\Phi(S) = \int_S P_x(\mathbb{R}^n \setminus S) f(x) dx,$$

[1] http://arxiv.org/abs/0906.2448

and the measure of S according to π_f is defined as $\pi_f(S) = \int_S f(x)dx / \int_{\mathbb{R}^n} f(x)dx$. The s-conductance ϕ_s of the Markov chain defined by ball walk is

$$\phi_s = \inf_{s \le \pi_f(S) \le 1/2} \frac{\Phi(S)}{\pi_f(S) - s}.$$

To compare two distributions Q_1, Q_2 we use the total variation distance between Q_1 and Q_2, defined by $d_{tv}(Q_1, Q_2) = \sup_A |Q_1(A) - Q_2(A)|$. When we refer to the distance between two sets, we mean the minimum distance between any two points in the two sets. That is, for any two subsets $S_1, S_2 \subseteq \mathbb{R}^n$, $d(S_1, S_2) := \min\{|u - v| : u \in S_1, v \in S_2\}$. Next we quote a lemma from [LS93] which relates the s-conductance to the mixing time.

Lemma 2. *Let $0 < s \le 1/2$ and $H_s = \sup_{\pi_f(S) \le s} |\sigma_0(S) - \pi_f(S)|$. Then for every measurable $S \subseteq \mathbb{R}^n$ and every $m \ge 0$,*

$$|\sigma_m(S) - \pi_f(S)| \le H_s + \frac{H_s}{s}\left(1 - \frac{\phi_s^2}{2}\right)^m.$$

Finally, the following localization lemma [LS93, KLS95] is a useful tool in the proofs of isoperimetric inequalities.

Lemma 3. *Let $g : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}$ be two lower semi-continuous integrable functions such that*

$$\int_{\mathbb{R}^n} g(x)dx > 0 \quad \textit{and} \quad \int_{\mathbb{R}^n} h(x)dx > 0.$$

Then there exist two points $a, b \in \mathbb{R}^n$ and a linear function $l : [0,1] \to \mathbb{R}_+$ such that

$$\int_0^1 g((1-t)a + tb)l(t)^{n-1}dt > 0 \quad \textit{and} \quad \int_0^1 h((1-t)a + tb)l(t)^{n-1}dt > 0.$$

3 Isoperimetry

Here we prove an isoperimetric inequality for functions satisfying a certain unimodality criterion. We further show that $-1/(n-1)$-concave functions satisfy this unimodality criterion and hence have good isoperimetry.

We begin with a simple lemma that will be used in the proof of the isoperimetric inequality.

Lemma 4. *Let $p : [0,1] \to \mathbb{R}_+$ be a unimodal function, and let $0 \le \alpha < \beta \le 1$. Then*

$$\int_\alpha^\beta p(t)dt \ge |\alpha - \beta| \min\left\{\int_0^\alpha p(t)dt, \int_\beta^1 p(t)dt\right\}.$$

Now we are ready to prove an isoperimetric inequality for functions satisfying a certain unimodality criterion.

Theorem 7. *Let $f : \mathbb{R}^n \to \mathbb{R}_+$ be a function whose support has diameter D, and f satisfies the following unimodality criterion: For any affine line $L \subseteq \mathbb{R}^n$ and any linear function $l : K \cap L \to \mathbb{R}_+$, $h(x) = f(x)l(x)^{n-1}$ is unimodal. Let $\mathbb{R}^n = S_1 \cup S_2 \cup S_3$ be a partition of $\mathbb{R}^n$ into three non-empty subsets. Then*

$$\pi_f(S_3) \geq \frac{d(S_1, S_2)}{D} \min\left\{\pi_f(S_1), \pi_f(S_2)\right\}.$$

Proof. Suppose not. Define $g : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}$ as follows.

$$g(x) = \begin{cases} \frac{d(S_1, S_2)}{D} f(x) & \text{if } x \in S_1 \\ 0 & \text{if } x \in S_2 \\ -f(x) & \text{if } x \in S_3 \end{cases} \quad \text{and} \quad h(x) = \begin{cases} 0 & \text{if } x \in S_1 \\ \frac{d(S_1, S_2)}{D} f(x) & \text{if } x \in S_2 \\ -f(x) & \text{if } x \in S_3. \end{cases}$$

Thus

$$\int_{\mathbb{R}^n} g(x)dx > 0 \quad \text{and} \quad \int_{\mathbb{R}^n} h(x)dx > 0,$$

Lemma 3 implies that there exist two points $a, b \in \mathbb{R}^n$ and a linear function $l : [0,1] \to \mathbb{R}_+$ such that

$$\int_0^1 g((1-t)a + tb)l(t)^{n-1}dt > 0 \quad \text{and} \quad \int_0^1 h((1-t)a + tb)l(t)^{n-1}dt > 0. \quad (2)$$

Moreover, w.l.o.g. we can assume that the points a and b are within the support of f, and hence $\|a - b\| \leq D$. We may also assume that $a \in S_1$ and $b \in S_2$. Consider a partition of the interval $[0,1] = Z_1 \cup Z_2 \cup Z_3$, where

$$Z_i = \left\{z \in [0,1] \;:\; (1-z)a + zb \in S_i\right\}.$$

For $z_1 \in Z_1$ and $z_2 \in Z_2$, we have

$$d(S_1, S_2) \leq d\left((1-z_1)\,a + z_1 b, (1-z_2)\,a + z_2 b\right) \leq |z_1 - z_2| \cdot \|a - b\| \leq |z_1 - z_2|\, D,$$

and therefore $d(S_1, S_2) \leq d(Z_1, Z_2)D$. Now we can rewrite Equation (2) as

$$\begin{aligned} \int_{Z_3} f((1-t)a + tb)l(t)^{n-1}dt &< \frac{d(S_1, S_2)}{D} \int_{Z_1} f((1-t)a + tb)l(t)^{n-1}dt \\ &\leq d(Z_1, Z_2) \int_{Z_1} f((1-t)a + tb)l(t)^{n-1}dt \end{aligned}$$

and similarly

$$\int_{Z_3} f((1-t)a + tb)l(t)^{n-1}dt \leq d(Z_1, Z_2) \int_{Z_2} f((1-t)a + tb)l(t)^{n-1}dt$$

Define $p : [0,1] \to \mathbb{R}_+$ as $p(t) = f((1-t)a + tb)l(t)^{n-1}$. From the unimodality assumption in our theorem, we know that p is unimodal. Rewriting the above equations, we have

$$\int_{Z_3} p(t)dt < d(Z_1, Z_2) \int_{Z_1} p(t)dt \quad \text{and} \quad \int_{Z_3} p(t)dt < d(Z_1, Z_2) \int_{Z_2} p(t)dt. \quad (3)$$

Now suppose Z_3 is a union of disjoint intervals, i.e., $Z_3 = \bigcup_i (\alpha_i, \beta_i)$, $0 \leq \alpha_1 < \beta_1 < \alpha_2 < \beta_2 < \cdots \leq 1$. By Lemma 4 we have

$$\int_{\alpha_i}^{\beta_i} p(t)dt \geq |\alpha_i - \beta_i| \cdot \min\left\{\int_0^{\alpha_i} p(t)dt, \int_{\beta_i}^1 p(t)dt\right\}.$$

Therefore, adding these up we get

$$\begin{aligned}\int_{Z_3} p(t)dt &= \sum_i \int_{\alpha_i}^{\beta_i} p(t)dt \\ &\geq |\alpha_i - \beta_i| \cdot \sum_i \min\left\{\int_0^{\alpha_i} p(t)dt, \int_{\beta_i}^1 p(t)dt\right\} \\ &\geq d(Z_1, Z_2) \cdot \min\left\{\int_{Z_1} p(t)dt, \int_{Z_2} p(t)dt\right\}.\end{aligned}$$

The last inequality follows from the fact that either every interval in Z_1 or every interval in Z_2 is accounted for in the summation. Indeed, suppose some interval in Z_2 is not accounted for in the summation. Then, that interval has to be either the first or the last interval in $[0, 1]$ in which case all intervals in Z_1 are accounted for. But this is a contradiction to Inequality (3). This completes the proof of Theorem 7.

3.1 Isoperimetry of $-1/(n-1)$-Concave Functions

We show that $-1/(n-1)$-concave functions satisfy the unimodality criterion used in the proof of Theorem 7. Therefore, as a corollary, we get an isoperimetric inequality for $-1/(n-1)$-concave functions.

Proposition 1. *Let $f : \mathbb{R}^n \to \mathbb{R}_+$ be a smooth $-1/(n-1)$-concave function and $l : [0, 1] \to \mathbb{R}_+$ be a linear function. Now let $a, b \in \mathbb{R}^n$ and define $h : [0, 1] \to \mathbb{R}_+$ as $h(t) = f((1-t)a + tb)l(t)^{n-1}$. Then h is a unimodal function.*

We get Theorem 3 as a corollary of Theorem 7 and Proposition 1.

3.2 Lower Bound for Isoperimetry

In this section, we show that $-1/(n-1)$-concave functions are the limit of isoperimetry by showing a $-1/(n-1-\epsilon)$-concave function with poor isoperimetry for $0 < \epsilon \leq 1$.

Proof (Proof of Theorem 4). The proof is based on the following construction. Consider $K \subseteq \mathbb{R}^n$ defined as follows.

$$K = \left\{x \ : \ 0 \leq x_1 < \frac{1}{1+\delta} \text{ and } x_2^2 + x_3^2 + \ldots + x_n^2 \leq (1 - x_1)^2\right\},$$

where $\delta > 0$. K is a parallel section of a cone symmetric around the X_1-axis and is therefore convex. Now we define a function $f : \mathbb{R}^n \to \mathbb{R}_+$ whose support is K.

$$f(x) = \begin{cases} \dfrac{C}{(1-(1+\delta)x_1)^{n-1-\epsilon}} & \text{if } x \in K, \\ 0 & \text{if } x \notin K, \end{cases}$$

where C is the appropriate constant so as to make $\pi_f(K) = 1$. By definition, f is a $-1/(n-1-\epsilon)$-concave function.

Define a partition $\mathbb{R}^n = S \cup T$ as $S = \{x \in K \ : \ 0 \le x_1 \le t\}$ and $T = \mathbb{R}^n \setminus S$. It can be shown that the theorem holds for a suitable choice of t.

4 Sampling s-Concave Functions

Throughout this section, let $f : \mathbb{R}^n \to \mathbb{R}_+$ be an s-concave (α, δ)-smooth function given by an oracle such that $s \ge -1/(n-1)$. Let K be the convex body over which we want to sample points according to f. We also assume that K contains a ball of radius δ and is contained in a ball of radius D. We state a technical lemma related to the smoothness and the concavity of the function.

Lemma 5. *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a s-concave (α, δ)-smooth function. For any constant c such that $1 < c < \alpha$, if $\|x - z\| \le \frac{|cs\delta|}{\alpha^{-s}-1}$, then $\frac{f(x)}{f(z)} \le c$.*

The above lemma states that every s-concave (α, δ)-smooth function, is also $(c, \frac{|cs\delta|}{(\alpha^{-s}-1)})$-smooth for any constant c such that $1 < c < \alpha$. In particular, if $\alpha > 2$, this suggests that we may use the smoothness parameters to be $(\alpha' = 2, \delta' = \frac{|2s\delta|}{\alpha^{-s}-1})$ and if $\alpha \le 2$, then we may use $(\alpha' = 2, \delta' = \delta)$ as the parameters. Thus, the function can be assumed to be $(2, \min\{\delta, \frac{|2s\delta|}{\alpha^{-s}-1}\})$-smooth.

In order to sample, we need to show that K contains points of good local conductance. For this, define

$$K_r = \left\{ x \in K \ : \ \frac{\text{vol}\,(r\mathbb{B}_x \cap K)}{\text{vol}\,(r\mathbb{B}_x)} \ge \frac{3}{4} \right\}.$$

The idea is that, for appropriately chosen r, the log-lipschitz-like constraint will enforce that the points in K_r have good local conductance. Further, we have that the measure in K_r is close to the measure of f in K based on the radius r.

Lemma 6. *For any $r > 0$, the set K_r is convex and*

$$\pi_f(K_r) \ge 1 - \frac{4r\sqrt{n}}{\delta}.$$

4.1 Coupling

In order to prove conductance, we need to prove that when two points are geometrically close, then their one-step distributions overlap. We will need the following technical lemma about spherical caps to prove this.

Lemma 7. *Let H be a halfspace in $\mathbb{R}^n$ and B_x be a ball whose center is at a distance at most $tr/\sqrt{n}$ from H. Then*

$$e^{-\frac{t^2}{4}} > \frac{2\,\mathrm{vol}\,(H \cap r\mathbb{B})}{\mathrm{vol}\,(r\mathbb{B})} > 1 - t$$

Lemma 8. *For $r \leq \min\{\delta, \frac{|2s\delta|}{\alpha^{-s}-1}\}$, if $u, v \in K_r$, $\|u - v\| < r/16\sqrt{n}$, then*

$$d_{tv}(P_u, P_v) \leq 1 - \frac{7}{16}$$

Proof. We may assume that $f(v) \geq f(u)$. Then,

$$d_{tv}(P_u, P_v) \leq 1 - \frac{1}{\mathrm{vol}\,(r\mathbb{B})} \int_{r\mathbb{B}_v \cap r\mathbb{B}_u \cap K} \min\left\{1, \frac{f(y)}{f(v)}\right\} dy$$

Let us lower bound the second term in the right hand side.

$$\begin{aligned}
\int_{r\mathbb{B}_v \cap r\mathbb{B}_u \cap K} \min\left\{1, \frac{f(y)}{f(v)}\right\} dy &\geq \int_{r\mathbb{B}_v \cap r\mathbb{B}_u \cap K} \min\left\{1, \frac{f(y)}{f(v)}\right\} dy \\
&\geq \left(\frac{1}{2}\right) \mathrm{vol}\,(r\mathbb{B}_v \cap r\mathbb{B}_u \cap K) \qquad \text{(By Lemma 5)} \\
&\geq \left(\frac{1}{2}\right) (\mathrm{vol}\,(r\mathbb{B}_v) - \mathrm{vol}\,(r\mathbb{B}_v \setminus r\mathbb{B}_u) - \mathrm{vol}\,(r\mathbb{B}_v \setminus K)) \\
&\geq \left(\frac{1}{2}\right) \left(\mathrm{vol}\,(r\mathbb{B}_v) - \frac{1}{16}\mathrm{vol}\,(r\mathbb{B}) - \frac{1}{16}\mathrm{vol}\,(r\mathbb{B})\right) \\
&\geq \left(\frac{7}{16}\right) \mathrm{vol}\,(r\mathbb{B})
\end{aligned}$$

where the bound on $\mathrm{vol}\,(r\mathbb{B}_v \setminus r\mathbb{B}_u)$ is derived from Lemma 7 and $\mathrm{vol}\,(r\mathbb{B}_v \setminus K)$ is bounded using the fact that $v \in K_r$. Hence,

$$d_{tv}(P_u, P_v) \leq 1 - \frac{7}{16}$$

4.2 Conductance and Mixing Time

Consider the ball walk with metropolis filter using the s-concave (α, δ)-smooth density function oracle with ball steps of radius r.

Lemma 9. *Let $S \subseteq \mathbb{R}^n$ be such that $\pi_f(S) \geq \epsilon_1$ and $\pi_f(\mathbb{R}^n \setminus S) \geq \epsilon_1$. Then, for ball walk radius $r \leq \min\left\{\frac{\epsilon_1\delta}{8\sqrt{n}}, \frac{|2s\delta|}{\alpha^{-s}-1}\right\}$, we have that*

$$\Phi(S) \geq \frac{r}{2^9\sqrt{n}D} \min\{\pi_f(S) - \epsilon_1, \pi_f(\mathbb{R}^n \setminus S) - \epsilon_1\}$$

Using the above lemma, we prove Theorem 5.

Proof (Proof of Theorem 5). On setting $\epsilon_1 = \epsilon/2H$ in Lemma 9, we have that for ball-walk radius $r = \min\{\frac{\epsilon\delta}{16H\sqrt{n}}, \frac{|2s\delta|}{(\alpha^{-s}-1)}\}$,

$$\phi_{\epsilon_1} \geq \frac{r}{2^9\sqrt{n}D}.$$

By definition $H_s \leq H \cdot s$ and hence by Lemma 2,

$$|\sigma_m(S) - \pi_f(S)| \leq H \cdot s + H \cdot exp\left\{-\frac{mr^2}{2^{19}nD^2}\right\}$$

which gives us that beyond

$$m \geq \frac{2^{19}nD^2}{r^2}\log\frac{2H}{\epsilon}$$

steps, $|\sigma_m(S) - \pi_f(S)| \leq \epsilon$. Substituting for r, we get the theorem.

4.3 Sampling the Cauchy Density

In this section, we prove certain properties of the Cauchy density along with the crucial coupling lemma leading to Theorem 6. Without loss of generality, we may assume that the distribution given by the oracle is,

$$f(x) \propto \begin{cases} 1/(1+||x||^2)^{\frac{n+1}{2}} & \text{if } x \in K, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

This is because, either we are explicitly given the matrix A of a general Cauchy density, or we can compute it using the function f at a small number of points and apply a linear transformation. Further, note that by the hypothesis of Theorem 6, we may assume that K contains a unit ball.

Proposition 2. *The Cauchy density function is $-1/(n-1)$-concave.*

Proposition 3 says that we can find a ball of radius $O(\sqrt{n}/\epsilon_1)$ outside which the Cauchy density has at most ϵ_1 probability mass.

Proposition 3

$$\mathsf{Pr}\left(\|x\| \geq \frac{2\sqrt{2n}}{\epsilon_1}\right) \leq \epsilon_1.$$

Proposition 4 shows the smoothness property of the Cauchy density. This is the crucial ingredient used in the stronger coupling lemma. Define K_r as before. Then,

Proposition 4. *For $x \in K_r$, let*

$$C_x = \{y \in r\mathbb{B}_x : |x \cdot (x-y)| \leq \frac{4r||x||}{\sqrt{n}}\}$$

and $y \in C_x$. Then,

$$\frac{f(x)}{f(y)} \geq 1 - 4r\sqrt{n}$$

Finally, we have the following coupling lemma.

Lemma 10. *For $r \leq 1/\sqrt{n}$, if $u, v \in K_r$, $\|u - v\| < r/16\sqrt{n}$, then*

$$d_{tv}(P_u, P_v) < \frac{1}{2}.$$

The proof of conductance and mixing bound follow the proof of mixing bound for s-concave functions closely. Comparing the above coupling lemma with that of s-concave functions (Lemma 8), we observe that the improvement is obtained due to the constraint on the radius of the ball walk in the coupling lemma. In the case of Cauchy, a slightly relaxed radius suffices for points close to each other to have a considerable overlap in their one-step distribution.

4.4 Discussion

There are two aspects of our algorithm and analysis that merit improvement. The first is the dependence on the diameter, which could perhaps be made logarithmic by applying an appropriate affine transformation as in the case of logconcave densities. The second is eliminating the dependence on the smoothness parameter entirely, by allowing for sharp changes locally and considering a smoother version of the original function. Both these aspects seem to be tied closely to proving a tail bound on a 1-dimensional marginal of an s-concave function.

References

[AK91] Applegate, D., Kannan, R.: Sampling and integration of near log-concave functions. In: STOC 1991: Proceedings of the twenty-third annual ACM symposium on Theory of computing, pp. 156–163. ACM, New York (1991)

[Bob07] Bobkov, S.G.: Large deviations and isoperimetry over convex probability measures with heavy tails. Electronic Journal of Probability 12, 1072–1100 (2007)

[Bor74] Borell, C.: Convex measures on locally convex spaces. Ark. Math. 12, 239–252 (1974)

[Bor75] Borell, C.: Convex set functions in d-space. Periodica Mathematica Hungarica 6, 111–136 (1975)

[BV04] Bertsimas, D., Vempala, S.: Solving convex programs by random walks. J. ACM 51(4), 540–556 (2004)

[DF91] Dyer, M.E., Frieze, A.M.: Computing the volume of a convex body: a case where randomness provably helps. In: Proc. of AMS Symposium on Probabilistic Combinatorics and Its Applications, pp. 123–170 (1991)

[DFK91] Dyer, M.E., Frieze, A.M., Kannan, R.: A random polynomial-time algorithm for approximating the volume of convex bodies. J. ACM 38(1), 1–17 (1991)

[Joh87] Johnson, M.E.: Multivariate statistical simulation. Wiley, Chichester (1987)

[KL96] Kannan, R., Li, G.: Sampling according to the multivariate normal density. In: FOCS 1996: Proceedings of the 37th Annual Symposium on Foundations of Computer Science, Washington, DC, USA, p. 204. IEEE Computer Society Press, Los Alamitos (1996)

[KLS95] Kannan, R., Lovász, L., Simonovits, M.: Isoperimetric problems for convex bodies and a localization lemma. J. Discr. Comput. Geom. 13, 541–559 (1995)

[KLS97] Kannan, R., Lovász, L., Simonovits, M.: Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. Random Structures and Algorithms 11, 1–50 (1997)

[KV06] Kalai, A., Vempala, S.: Simulated annealing for convex optimization. Math. Oper. Res. 31(2), 253–266 (2006)

[Lov90] Lovász, L.: How to compute the volume? Jber. d. Dt. Math.-Verein, Jubiläumstagung 1990, 138–151 (1990)

[LS92] Lovász, L., Simonovits, M.: On the randomized complexity of volume and diameter. In: Proc. 33rd IEEE Annual Symp. on Found. of Comp. Sci., pp. 482–491 (1992)

[LS93] Lovász, L., Simonovits, M.: Random walks in a convex body and an improved volume algorithm. Random Structures and Alg 4, 359–412 (1993)

[LV06a] Lovász, L., Vempala, S.: Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization. In: FOCS 2006: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, Washington, DC, USA, pp. 57–68. IEEE Computer Society Press, Los Alamitos (2006)

[LV06b] Lovász, L., Vempala, S.: Hit-and-run from a corner. SIAM J. Computing 35, 985–1005 (2006)

[LV06c] Lovász, L., Vempala, S.: Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm. J. Comput. Syst. Sci. 72(2), 392–417 (2006)

[LV07] Lovász, L., Vempala, S.: The geometry of logconcave functions and sampling algorithms. Random Struct. Algorithms 30(3), 307–358 (2007)