
Hit-and-Run for Sampling and Planning in Non-Convex Spaces

Yasin Abbasi-Yadkori
Adobe Research

Peter L. Bartlett
UC Berkeley & QUT

Victor Gabillon
Queensland University of Tech.

Alan Malek
UC Berkeley

Abstract

We propose the HIT-AND-RUN algorithm for planning and sampling problems in non-convex spaces. For sampling, we show the first analysis of the HIT-AND-RUN algorithm in non-convex spaces and show that it mixes fast as long as certain smoothness conditions are satisfied. In particular, our analysis reveals an intriguing connection between fast mixing and the existence of smooth measure-preserving mappings from a convex space to the non-convex space. For planning, we show advantages of HIT-AND-RUN compared to state-of-the-art planning methods such as Rapidly-Exploring Random Trees.

1 Introduction

Rapidly-Exploring Random Trees (RRT) (LaValle, 1998, LaValle and Kuffner, 2001) is one of the most popular planning algorithms, especially when the search space is high-dimensional and finding the optimal path is computationally expensive. RRT performs well on many problems where classical dynamic programming based algorithms, such as A*, perform poorly. RRT is essentially an exploration algorithm, and in the most basic implementation, the algorithm even ignores the goal information, which seems to be a major reason for its success. Planning problems, especially those in robotics, often feature narrow pathways connecting large explorable regions; combined with high dimensionality, this means that finding the optimal path is usually intractable. However, RRT often provides a *feasible* path quickly.

Although many attempts have been made to improve the basic algorithm (Abbasi-Yadkori et al., 2010, Karaman and Frazzoli, 2011), RRT has proven difficult

to improve upon. In fact, given extra computation, repeatedly running RRT often produces competitive solutions. In this paper, we show that a simple alternative greatly improves upon RRT. We propose using the HIT-AND-RUN algorithm for feasible path search. Arguably simpler than RRT, the HIT-AND-RUN is a rapidly mixing MCMC sampling algorithm for producing a point uniformly at random from a convex space (Smith, 1984). Not only HIT-AND-RUN finds a feasible path faster than RRT, it is also more robust with respect to the geometry of the space.

Before giving more details, we define the *planning* and *sampling* problems that we consider. Let Σ be a bounded connected subset of \mathbb{R}^n . For points $a, b \in \Sigma$, we use $[a, b]$ to denote their (one-dimensional) convex hull. Given a starting point a_1 and a goal region $\mathcal{G} \subset \Sigma$, the *planning problem* is to find a sequence of points $\{a_1, a_2, \dots, a_\tau\}$ for $\tau \geq 1$ such that all points are in Σ , a_τ is in \mathcal{G} , and for $t = 2, \dots, \tau$, $[a_{t-1}, a_t] \subset \Sigma$.

The *sampling problem* is to generate points uniformly at random from Σ . Sampling is often difficult, but Markov Chain Monte Carlo (MCMC) algorithms have seen empirical and theoretical success (Lovász and Vempala, 2007). MCMC algorithms, such as HIT-AND-RUN and BALL-WALK (Vempala, 2005), sample a Markov Chain on Σ that has a stationary distribution equal to the uniform distribution on Σ ; then, if we run the Markov Chain long enough, the marginal distribution of the sample is guaranteed to come from a distribution exponentially close to the target distribution. Solving the sampling problem yields a solution to the planning problem; one can generate samples and terminate when a_t hits \mathcal{G} .

Let us define HIT-AND-RUN and the RRT algorithms (see also Figure 1 for an illustration). HIT-AND-RUN defines a Markov chain on Σ where the transition dynamics are as follows. A direction is chosen uniformly at random, and a_{t+1} is chosen uniformly from the largest chord contained in Σ in this direction passing through a_t . This Markov Chain has a uniform stationary distribution on Σ (Smith, 1984). As a planning algorithm, this chain continues until it hits the goal region. Let τ be the stopping time. The solution path

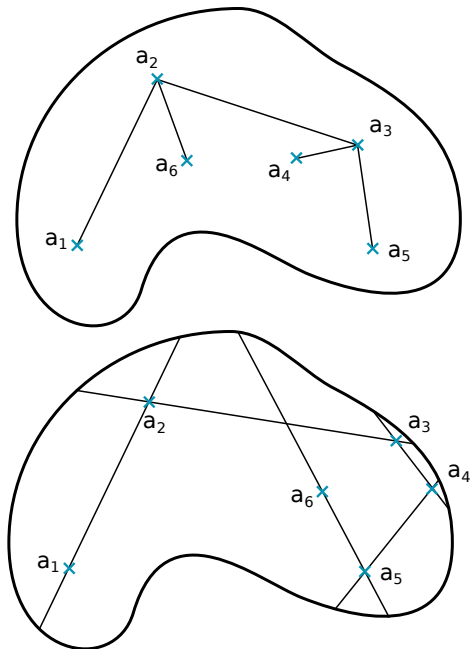


Figure 1: RRT (top) and HIT-AND-RUN (bottom)

is $\{a_1, a_2, \dots, a_7\}$.

On the other hand, the RRT algorithm iteratively builds a tree T with a_1 as a root and nodes labeled as $a^n \in \Sigma$ and edges $\{a^m, a^n\}$ that satisfy $[a^m, a^n] \subseteq \Sigma$. To add a point to the tree, a^r is uniformly sampled from Σ and its nearest neighbor $a^n \in T$ is computed. If $[a^n, a^r] \subseteq \Sigma$, then node a^r and edge $[a^n, a^r]$ are added to T . Otherwise, we search for the point $a^e \in [a^n, a^r]$ farthest from a^n such that $[a^n, a^e] \subseteq \Sigma$. Then a^e and $[a^n, a^e]$ are added to the tree. This process is continued until we add an edge terminating in \mathcal{G} and the sequence of points on that branch is returned as the solution path. In the presence of dynamic constraints, a different version of RRT that makes only small local steps is used. These versions will be discussed in the experiments section.

There are two main contributions on this paper. First, we analyze the HIT-AND-RUN algorithm in a non-convex space and show that the mixing time is polynomial in dimensionality as long as certain smoothness conditions are satisfied. The mixing time of HIT-AND-RUN for convex spaces is known to be polynomial (Lovász, 1999). However, to accommodate planning problems, we focus on non-convex spaces. Our analysis reveals an intriguing connection between fast mixing and the existence of smooth measure-preserving mappings. The only existing analysis of random walk algorithms in non-convex spaces is due to Chandrasekaran et al. (2010) who analyzed BALL-

WALK in star-shaped bodies.¹ We show that the mixing time of HIT-AND-RUN is $O(n^6)$ in a class of nonconvex spaces. It is possible that the n^6 rate could be improved: the analogous best result for HIT-AND-RUN on a convex set is n^2 . However, the n^2 complexity in the convex case required decades of refinement. Second, we propose HIT-AND-RUN for planning problems as an alternative to RRT and show that it finds a feasible path quickly. From the mixing rate, we obtain a bound on the expected length of the solution path in the planning problem: If the mixing time is m , diameter of the space is D_Σ , and the volume of the goal space is v , then, roughly, this expected length is $O(mD_\Sigma/v)$. Although this might not be tight, this is much stronger than the available theory for RRT (that show only that a path is found in the limit).

The current proof techniques in the analysis of HIT-AND-RUN heavily rely on the convexity of the space. It turns out that non-convexity is specially troubling when points are close to the boundary. We overcome these difficulties as follows. First, Lovász and Vempala (2006) show a tight isoperimetric inequality in terms of average distances instead of minimum distances. This enables us to ignore points that are sufficiently close to the boundary. Next we show that as long as points are sufficiently far from the boundary, the cross-ratio distances in the convex and non-convex spaces are closely related. Finally we show that, given a curvature assumption, if two points are close geometrically and are sufficiently far from the boundary, then their proposal distributions must be close as well.

HIT-AND-RUN has a number of advantages compared to RRT; it does not require random points sampled from the space (which is itself a hard problem), and it is guaranteed to reach the goal region with high probability in a polynomial number of rounds. In contrast, there are cases where RRT growth can be very slow. For example, consider a spiral shaped space. RRT can be slow in this space because in many rounds the tree does not grow in the *right* direction. As HIT-AND-RUN only considers the space that is visible to the current point, it is less sensitive to the geometry of the free space. We can make this problem arbitrarily hard for RRT by making certain parts of the spiral fatter. HIT-AND-RUN, on the other hand, is insensitive to such changes. Additionally, the growth of the RRT tree can become very slow towards the end. This is because the rest of the tree absorbs most samples, and the tree grows only if the random point falls in the vicinity of the goal. (see the experiments sections for more details). Moreover, HIT-AND-RUN provides *safer* solutions, as its paths are more likely to stay away

¹We say S is star-shaped if the kernel of S , define by $K_S = \{x \in S : \forall y \in S [x, y] \subseteq S\}$, is nonempty.

from the boundary. In contrast, a common issue with RRT solutions is that they tend to be close to the boundary. Because of this, further post-processing steps are needed to smooth the path.

1.1 Notation

For a set K , we will denote the n -dimensional volume by $\text{VOL}(K)$, the $(n - 1)$ -dimensional surface volume by $S_K = \text{VOL}_{n-1}(\partial K)$, and the boundary by ∂K . The diameter of K is $D_K = \max_{x, x' \in K} |x - x'|$, where $|\cdot|$ will be used for absolute value and Euclidean norm, and the distance between sets K_1 and K_2 is defined as $d(K_1, K_2) = \min_{x \in K_1, y \in K_2} |x - y|$. Similarly, $d(x, K) = d(\{x\}, K)$. For a set K , we use K^ϵ to denote $\{x \in K : d(x, \partial K) \geq \epsilon\}$. Finally, for distributions P and Q , we use $d_{tv}(P, Q)$ to denote the total variation distance between P and Q .

We will also need some geometric quantities. We will denote lines (i.e., 1-dimensional affine spaces) by ℓ . For $x, y \in K$, we denote their convex hull, that is, the line segment between them, by $[x, y]$ and $\ell(x, y)$ the line that passes through x and y (which contains $[x, y]$). We also write $[x_1, \dots, x_k]$ to denote that x_1, \dots, x_k are collinear.

We also use $\ell_K(x, y)$ to denote the longest connected chord through $x \in K$ and $y \in K$ contained in K and $|\ell_K(x, y)|$ its length. We use $a(x, y)$ and $b(x, y)$ to denote the endpoints of $\ell_K(x, y)$ that are closer to x and y , respectively, so that $\ell_K(x, y) = [a(x, y), b(x, y)] = [b(y, x), a(y, x)]$. The Euclidean ball of unit radius centered at the origin, $B_n(0, 1) \subset \mathbb{R}^n$, has volume π_n . We use $x_{1:m}$ to denote the sequence x_1, \dots, x_m . Finally we use $a \wedge b$ to denote $\min(a, b)$.

2 Sampling from Non-Convex Spaces

Most of the known results for the sampling times of the HIT-AND-RUN exist for convex sets only. We will think of Σ as the image of some convex set Ω under a measure preserving, bilipschitz function g . The goal is to understand the relevant geometric quantities of Σ through properties of g and geometric properties of Ω . We emphasize that the existence of the map g and its properties are necessary for the analysis, but the actual algorithm does not need to know g . We formalize this assumption below as well as describe how we interact with Σ and present a few more technical assumptions required for our analysis. We then present our main result, and follow that with some conductance results before moving on to the proof of the theorem in the next section.

Assumption 1 (Oracle Access). *Given a point u and a line ℓ that passes through u , the oracle returns whether*

$u \in \Sigma$, and, if so, the largest connected interval in $\ell \cap \Sigma$ containing u .

The complexity of computing the longest chord in an arbitrary nonconvex space depends on the specific problem. For example, if an agent controls a robot and the robot has some distance sensors, then the complexity is low. On the other hand, if we want to estimate the length by probing the space and there is a thin wall, this operation can be expensive.

Assumption 2 (Bilipschitz Measure-Preserving Embeddings). *There exist a convex set $\Omega \subset \mathbb{R}^n$ and a bilipschitz, measure-preserving map g such that Σ is the image of Ω under g . That is, there exists a function g with $|D_g(x)| = 1$ (i.e. the Jacobian has unit determinant) with constants L_Σ and L_Ω such that, for any $x, y \in \Omega$,*

$$\frac{1}{L_\Omega} |x - y| \leq |g(x) - g(y)| \leq L_\Sigma |x - y|.$$

In words, g is measure-preserving, g is L_Σ -Lipschitz, and g^{-1} is L_Ω -Lipschitz.

As an example, Fonseca and Parry (1992) shows that for any star-shaped space, a smooth measure-preserving embedding exists. Also the spiral shaped spaces that we use in our experiments satisfy this assumption. The basic intuition is that we need to disallow sets with very narrow bridges connecting large masses (e.g. a dumbbell) and very deep corners; there are examples of both sets which can be made to mix arbitrarily slowly.

One interesting consequence of Assumption 2 is that because the mapping is measure-preserving, there must exist a pair $x, y \in \Omega$ such that $|g(x) - g(y)| \geq |x - y|$. Otherwise, $\int_\Omega g \leq 1$, a contradiction. Similarly, there must exist a pair $u, v \in \Sigma$ such that $|g^{-1}(u) - g^{-1}(v)| \geq |u - v|$. Thus,

$$L_\Omega, L_\Sigma \geq 1. \tag{1}$$

To simplify the analysis, we will assume that Ω is a ball with radius r . In what follows, we use x, y, z to denote points in Ω , and u, v, w to denote points in Σ . We will also assume that Σ has no sharp corners and has a smooth boundary:

Assumption 3 (Low Curvature). *For any two dimensional plane $\mathcal{H} \subset \mathbb{R}^n$, let $\kappa_{\mathcal{H}}$ be the curvature of $\partial\Sigma \cap \mathcal{H}$ and $\mathcal{R}_{\mathcal{H}}$ be the perimeter of $\partial\Sigma \cap \mathcal{H}$. We assume that Σ has low curvature, i.e. that $\kappa = \sup_{\mathcal{H}} \kappa_{\mathcal{H}} \mathcal{R}_{\mathcal{H}}$ is finite.*

We require this assumption to show that if two points are close geometrically, then their proposal distributions must be close as well. Assumption 2 does not imply low curvature, as there exist smooth measure-preserving mappings from the unit ball to a cube (Griepentrog et al., 2008).

Assumption 4. We assume that the volume of Σ is equal to one. We also assume that Σ contains a Euclidean ball of radius one.

Note that the unit ball has volume less than 1 for $n > 12$, so for small dimensional problems, we will need to relax this assumption.

We motivate the forthcoming technical machinery by demonstrating what it can accomplish. The following theorem is the main result of the paper, and the proof makes up most of Section 3.

Theorem 5. Consider the HIT-AND-RUN algorithm. Let σ_0 be the distribution of the initial point given to HIT-AND-RUN, σ_t be the distribution after t steps of HIT-AND-RUN, and σ be the stationary distribution (which is uniform). Let $M = \sup_A \sigma_0(A)/\sigma(A)$. Let ϵ be a positive scalar. After

$$t \geq C'n^6 \log \frac{M}{\epsilon}$$

steps, we have $d_{tv}(\sigma_t, \sigma) \leq \epsilon$. Here C' is a low order polynomial of $L_\Omega, L_\Sigma, \kappa$.

3 Analysis

This section proves Theorem 5. We begin by stating a number of useful geometrical results, which allow us to prove the two main components: an isoperimetric inequality in Section 3.2 and a total variation inequality in Section 3.3. We then combine everything in Appendix 3.4.

3.1 Fast Mixing Markov Chains

We rely on the notion of conductance as our main technical tool. This section recalls the relevant results.

We say that points $u, v \in \Sigma$ see each other if $[u, v] \subseteq \Sigma$. We use $\text{VIEW}(u)$ to denote all points in Σ visible from u . Let $\ell_\Sigma(u, v)$ denote the chord through u and v inside Σ and $|\ell_\Sigma(u, v)|$ its length. Let $P_u(A)$ be the probability of being in set $A \subset \Sigma$ after one step of HIT-AND-RUN from u and f_u its density function. By an argument similar to the argument in Lemma 3 of Lovász (1999), we can show that

$$f_u(v) = 2 \frac{\mathbf{1}\{v \in \text{VIEW}(u)\}}{n\pi_n |\ell_\Sigma(u, v)| \cdot |u - v|^{n-1}}. \quad (2)$$

The conductance of the Markov process is defined as

$$\Phi = \inf_{A \subset \Sigma} \frac{\int_A P_u(\Sigma \setminus A) du}{\min(\text{VOL}(A), \text{VOL}(\Sigma \setminus A))}.$$

We begin with a useful conductance result that applies to general Markov Chains.

Lemma 6 (Corollary 1.5 of Lovász and Simonovits (1993)). Let $M = \sup_A \sigma_0(A)/\sigma(A)$. Then for every $A \subset \Sigma$,

$$|\sigma_t(A) - \sigma(A)| \leq \sqrt{M} \left(1 - \frac{\Phi^2}{2}\right)^t.$$

Proving a lower bound on the conductance is therefore a key step in the mixing time analysis. Previous literature has shown such lower bounds for convex spaces. Our objective in the following is to obtain such bounds for more general non-convex spaces that satisfy bilipschitz measure-preserving embedding and low curvature assumptions.

As in previous literature, we shall find that the following *cross-ratio distance* is very useful in deriving an *isoperimetric inequality* and a *total variation inequality*.

Definition 1. Let $[a, u, v, b]$ be collinear and inside Σ , such that $a, b \in \partial\Sigma$. Define

$$d_\Sigma(u, v) = \frac{|a - b| |u - v|}{|a - u| |v - b|}.$$

It is easy to see that $d_\Sigma(u, v) \geq 4|u - v|/D_\Sigma$. We define the following distance measure for non-convex spaces.

Definition 2. A set Σ will be called τ -best if, for any $u, v \in \Sigma$, there exist points $z_1, \dots, z_{\tau-1}$ such that $[u, z_1], [z_{\tau-1}, v]$, and $[z_i, z_{i+1}]$ for $i = 1, \dots, \tau - 2$ are all in Σ ; i.e., any two points in Σ can be connected by τ line segments that are all inside Σ . We define the distance

$$\tilde{d}_\Sigma(u, v) = \inf_{z_1, \dots, z_{\tau-1} \in \Sigma} (d_\Sigma(u, z_1) + d_\Sigma(z_1, z_2) + \dots + d_\Sigma(z_{\tau-1}, v)),$$

and, by extension, the distance between two subsets $\Sigma_1, \Sigma_2 \subset \Sigma$ as $\tilde{d}_\Sigma(\Sigma_1, \Sigma_2) = \inf_{u \in \Sigma_1, v \in \Sigma_2} \tilde{d}_\Sigma(u, v)$.

The analysis of the conductance is often derived via an *isoperimetric inequality*.

Theorem 7 (Theorem 4.5 of Vempala (2005)). Let Ω be a convex body in \mathbb{R}^n . Let $h : \Omega \rightarrow \mathbb{R}^+$ be an arbitrary function. Let $(\Omega_1, \Omega_2, \Omega_3)$ be any partition of Ω into measurable sets. Suppose that for any pair of points $x \in \Omega_1$ and $y \in \Omega_2$ and any point z on the chord of Ω through x and y , $h(z) \leq (1/3) \min(1, d_\Omega(x, y))$. Then

$$\text{VOL}(\Omega_3) \geq \mathbf{E}_\Omega(h) \min(\text{VOL}(\Omega_1), \text{VOL}(\Omega_2)),$$

where the expectation is defined with respect to the uniform distribution on Ω .

Given an isoperimetric inequality, a total variation inequality is typically used in a mixing time analysis

to lower bound cross-ratio distances and then lower bound the conductance. Our approach is similar. We use the embedding assumption to derive an isoperimetric inequality in the non-convex space Σ . Then we relate cross-ratio distance d_Ω to distance \tilde{d}_Σ . This approximation is good when the points are sufficiently far from the boundary. We incur a small error in the mixing bound by ignoring points that are too close to the boundary. Finally we use the curvature condition to derive a total variation inequality and to lower bound the conductance.

3.2 Cross-Ratio Distances

The first step is to show the relationship between cross-ratio distances in the convex and non-convex spaces. We show that these distances are close as long as points are far from the boundary. These results will be used in the proof of the main theorem in Section 3.4 to obtain an isoperimetric inequality in the non-convex space. First we define a useful quantity.

Definition 3. Consider a convex set Ω with some subset Ω' and collinear points $\{a, x, b\}$ with $a, b \in \partial\Omega$, $x \in \Omega'$, and $|x - b| \leq |x - a|$. Let c be a point on $\partial\Omega$. Let $R(a, x, b, c) = |x - b| / |x - c|$. We use $R(\Omega, \Omega')$ to denote the maximum of $R(a, x, b, c)$ over all such points. We use R_ϵ to denote $R(\Omega, \Omega^\epsilon)$.

The following lemma is the main technical lemma, and we use it to express \tilde{d}_Σ in terms of d_Ω .

Lemma 8. Let ϵ be a positive scalar such that $R_\epsilon(1 + 8R_\epsilon) \geq 2/3$. Let $\{a, x_1, x_2, b\}$ be collinear such that a and b are on the boundary of Ω , $x_1, x_2 \in \Omega^\epsilon$, and $|x_1 - a| < |x_2 - a|$. Let c and d be two points on the boundary of Ω . Then

$$\frac{|b - a|}{|a - x_1|} \cdot \frac{|x_1 - c|}{|c - d|} \cdot \frac{|d - x_2|}{|x_2 - b|} \geq \frac{1}{4R_\epsilon(1 + 2R_\epsilon)}.$$

Proof. Let

$$A = \frac{|b - a| |x_1 - x_2|}{|a - x_1| |x_2 - b|}, \quad B = \frac{|c - d| |x_1 - x_2|}{|c - x_1| |x_2 - d|}.$$

We prove the claim by proving that $A/B \geq 1/(4R_\epsilon(1 + 2R_\epsilon))$. We show this inequality by considering different configurations of $\{a, x_1, x_2, b, c, d\}$. To save space, we present the proof for only a few cases. The complete proof is in Appendix A.

Case 1, $|x_1 - b| \leq |x_1 - a|$: In this case, x_1 and x_2 are both on the line segment $[(a + b)/2, b]$. We consider two cases.

Case 1.1, $|x_2 - d| \leq |x_2 - b|$: We have that

$$\begin{aligned} |c - d| &= |c - x_1| + |x_1 - x_2| + |x_2 - d| \\ &\leq |c - x_1| + |x_1 - b| + |x_2 - b|. \end{aligned} \quad (3)$$

Because $|x_1 - a| < |x_2 - a|$ by the assumption of the lemma, we have $|x_2 - b| < |x_1 - b|$. Also because $|x_1 - b| \leq |x_1 - a|$ in Case 1, we have $|x_1 - b| / |c - x_1| \leq R_\epsilon$. Thus

$$\frac{|x_2 - b|}{|c - x_1|} \leq \frac{|x_1 - b|}{|c - x_1|} \leq R_\epsilon. \quad (4)$$

By (3) and (4),

$$\begin{aligned} \frac{|c - d|}{|c - x_1|} &\leq 1 + \frac{|x_1 - b|}{|c - x_1|} + \frac{|x_2 - b|}{|c - x_1|} \\ &\leq 1 + R_\epsilon + \frac{|x_2 - b|}{|c - x_1|} \leq 1 + 2R_\epsilon. \end{aligned}$$

We use also that by definition of R_ϵ $|x_2 - b| \leq R_\epsilon |x_2 - d|$. This and the previous result lets us bound

$$\begin{aligned} B &\leq (1 + 2R_\epsilon) \frac{|x_1 - x_2|}{|x_2 - d|}, \\ A &\geq \frac{|b - a| |x_1 - x_2|}{R_\epsilon |a - x_1| |x_2 - d|} \geq \frac{|x_1 - x_2|}{R_\epsilon |x_2 - d|}, \end{aligned}$$

and conclude

$$\frac{A}{B} \geq \frac{1}{R_\epsilon(1 + 2R_\epsilon)} \geq \frac{1}{4R_\epsilon(1 + 2R_\epsilon)}.$$

The proof for other cases is in Appendix A. \square

The following lemma states that the distance d_Ω does not increase by adding more steps.

Lemma 9. Let $a, y_1, y_2, \dots, y_m, b$ be in the convex body Ω such that the points $\{a, y_1, y_2, \dots, y_m, b\}$ are collinear. Further assume that $a, b \in \partial\Omega$. We have that

$$d_\Omega(y_1, y_2) + \dots + d_\Omega(y_{m-1}, y_m) \leq d_\Omega(y_1, y_m).$$

Proof. We write

$$\begin{aligned} d_\Omega(y_1, y_m) &= \frac{|a - b| |y_1 - y_m|}{|a - y_1| |y_m - b|} \\ &= \frac{|a - b| |y_1 - y_2|}{|a - y_1| |y_m - b|} + \frac{|a - b| |y_2 - y_3|}{|a - y_1| |y_m - b|} \\ &\quad + \dots + \frac{|a - b| |y_{m-1} - y_m|}{|a - y_1| |y_m - b|} \\ &\geq \frac{|a - b| |y_1 - y_2|}{|a - y_1| |y_2 - b|} + \frac{|a - b| |y_2 - y_3|}{|a - y_2| |y_3 - b|} \\ &\quad + \dots + \frac{|a - b| |y_{m-1} - y_m|}{|a - y_{m-1}| |y_m - b|} \\ &= d_\Omega(y_1, y_2) + \dots + d_\Omega(y_{m-1}, y_m). \end{aligned}$$

\square

The next lemma upper bounds \tilde{d}_Σ in terms of d_Ω .

Lemma 10. *Let $x_1, x_2 \in \Omega^\epsilon$. We have that*

$$\tilde{d}_\Sigma(g(x_1), g(x_2)) \leq 4L_\Sigma^2 L_\Omega^2 R_\epsilon (1 + 2R_\epsilon) d_\Omega(x_1, x_2).$$

Proof. First we prove the inequality for the case that $g(x_1) \in \text{VIEW}(g(x_2))$. Let $a, b \in \partial\Omega$ be such that the points $\{a, x_1, x_2, b\}$ are collinear. Let $c, d \in \Omega$ be points such that the points $\{g(c), g(x_1), g(x_2), g(d)\}$ are collinear and the line connecting $g(c)$ and $g(d)$ is inside Σ . By the Lipschitzity of g and g^{-1} and Lemma 8,

$$\begin{aligned} \tilde{d}_\Sigma(g(x_1), g(x_2)) &= \frac{|g(c) - g(d)| |g(x_1) - g(x_2)|}{|g(c) - g(x_1)| |g(x_2) - g(d)|} \\ &\leq \frac{L_\Sigma^2 L_\Omega^2 |c - d| |x_1 - x_2|}{|c - x_1| |x_2 - d|} \\ &\leq L_\Sigma^2 L_\Omega^2 4R_\epsilon (1 + 2R_\epsilon) \frac{|a - b| |x_1 - x_2|}{|a - x_1| |x_2 - b|} \\ &= 4L_\Sigma^2 L_\Omega^2 R_\epsilon (1 + 2R_\epsilon) d_\Omega(x_1, x_2). \end{aligned} \quad (5)$$

Now consider the more general case where $g(x_1) \notin \text{VIEW}(g(x_2))$. Find a set of points y_1, \dots, y_τ such that the line segments $[g(x_1), g(y_1)], [g(y_1), g(y_2)], \dots, [g(y_\tau), g(x_2)]$ are all inside Σ . By definition of \tilde{d}_Σ , (5), and Lemma 9, $\tilde{d}_\Sigma(g(x_1), g(x_2))$ can be upper bounded by

$$\begin{aligned} &\inf_{u_1, \tau-1 \in \Sigma} d_\Sigma(g(x_1), u_1) + d_\Sigma(u_1, u_2) + \dots \\ &\quad + d_\Sigma(u_{\tau-1}, g(x_2)) \\ &\leq d_\Sigma(g(x_1), g(y_1)) + d_\Sigma(g(y_1), g(y_2)) \\ &\quad + \dots + d_\Sigma(g(y_{\tau-1}), g(x_2)) \\ &\leq 4L_\Sigma^2 L_\Omega^2 R_\epsilon (1 + 2R_\epsilon) (d_\Omega(x_1, y_1) \\ &\quad + d_\Omega(y_1, y_2) + \dots + d_\Omega(y_{\tau-1}, x_2)) \\ &\leq 4L_\Sigma^2 L_\Omega^2 R_\epsilon (1 + 2R_\epsilon) d_\Omega(x_1, x_2) \end{aligned}$$

□

3.3 Total Variation Inequality

In this section, we show that if two points $u, v \in \Sigma$ are close to each other, then P_u and P_v are also close. First we show that if the two points are close to each other, then they have similar views.

Lemma 11 (Overlapping Views). *Given the curvature κ defined in Assumption 3, for any $u, v \in \Sigma^\epsilon$ such that $|u - v| \leq \epsilon' \leq \epsilon$,*

$$P_u(\{x : x \notin \text{VIEW}(v)\}) \leq \max\left(\frac{4}{\pi}, \frac{\kappa}{\sin(\pi/8)}\right) \frac{\epsilon'}{\epsilon}.$$

The proof is in Appendix A. Next we define some notation and show some useful inequalities. For $u \in \Sigma$, let w be a random point obtained by making one step of HIT-AND-RUN from u . Define $F(u)$ by

$\mathbb{P}(|w - u| \leq F(u)) = 1/8$. If $d(u, \partial\Sigma) \geq h$, less than $1/8$ of any chord passing through u is inside $B(u, h/16)$. Thus $\mathbb{P}(|u - w| \leq h/16) \leq 1/8$, which implies

$$F(u) \geq \frac{h}{16}. \quad (6)$$

Intuitively, the total variation inequality implies that if u and v are close geometrically, then their proposal distributions must be close as well.

Lemma 12. *Let $u, v \in \Sigma^\epsilon$ be two points that see each other. Let $\epsilon' = \frac{\epsilon}{6} \min\left(\frac{\pi}{4}, \frac{\sin(\pi/8)}{\kappa}\right)$. Suppose that*

$$d_\Sigma(u, v) < \frac{\epsilon}{24D_\Sigma} \text{ and } |u - v| < \min\left(\frac{2F(u)}{\sqrt{n}}, \epsilon'\right).$$

Then,

$$|P_u - P_v| < 1 - \frac{\epsilon}{8e^4 D_\Sigma}.$$

The proof is in Appendix A. The proof uses ideas from proof of Lemma 9 of Lovász (1999). The proof of Lovász (1999) heavily relies on the convexity of the space, which does not hold in our case. We overcome the difficulties using the low curvature assumption and the fact that u and v are sufficiently far from the boundary.

3.4 Putting Everything Together

Now we are ready to prove Theorem 5. We proceed as follows. First, Theorem 7 shows an isoperimetric inequality in the convex case in terms of average distances. As we have an integration over Ω , we can ignore points that are very close to the boundary at a moderate cost in terms of the tightness of the inequality. Second, Lemma 10 shows that as long as points are sufficiently far from the boundary, the cross-ratio distances in the convex and non-convex spaces are closely related. This lemma and the isoperimetric inequality for the convex space gives an isoperimetric inequality for Σ . The next step is to show that if two points are close geometrically, then their proposal distributions must be close as well. We achieve this by using the fact that points are sufficiently far from the boundary and the boundary has low curvature. This argument bounds the conductance of HIT-AND-RUN. The complete proof is in Appendix A.

Lemma 13. *Let $\delta = \frac{9r}{320e^4 n L_\Omega D_\Sigma}$, $G = \frac{1}{6} \min\left(\frac{\pi}{4}, \frac{\sin(\pi/8)}{\kappa}\right)$, $\epsilon' = \frac{9r}{20n}$, where r is the radius of ball Ω (so $r^n \pi_n = 1$). Let $N = \frac{9r}{80n L_\Sigma^2 L_\Omega^2 R_{\epsilon'} (1+2R_{\epsilon'})}$. The conductance Φ of HIT-AND-RUN is at least*

$$\frac{\delta}{4} \left(\frac{2}{5nD_\Omega} \wedge N \left(\frac{1}{24D_\Sigma} \wedge \frac{2}{\sqrt{n}} \left(\frac{1}{8\sqrt{n}} \wedge G \right) \right) \right).$$

If we treat $L_\Sigma, L_\Omega, \kappa$ as constants and collect all constants in C , we have a $\Phi \geq C/n^3$ lower bound for the conductance.

Using Lemma 6 and Lemma 13,

$$d_{tv}(\sigma_t, \sigma) \leq \sqrt{M} \left(1 - \frac{C^2}{2n^6}\right)^t,$$

which gives the bound in Theorem 5 after rearrangement.

4 Planning

This section makes an empirical argument for use of the HIT-AND-RUN in trajectory planning. In the first of two experiments, the state space is a position vector constrained to some map illustrated by the bottom plots of Figure 2. The second experiment also includes two dimensions of velocity in the state and limits state transitions to those that respect the map as well as kinematics and requires the planning to control the system explicitly (by specifying an acceleration vector for every time step). We use spiral shaped spaces in both experiments. It can be shown that the embedding assumption (Assumption 2) holds for these examples: first, using results of Griepentrog et al. (2008), we construct a smooth measure-preserving mapping from the unit ball to a cube. Then we construct a map from the unit cube to a long tube. Finally, we obtain the desired shape by bending the tube in appropriate locations. We will show that HIT-AND-RUN outperforms RRT in both cases by requiring fewer transitions to reach the goal state across a wide variety of map difficulties.

4.1 Position only

The state starts at the bottom left of the spiral and the goal is the top right. Both algorithms are implemented as described in the introduction. The number of transitions needed to reach the goal of both algorithms is plotted as a function of the width of the spiral arms; the larger the width, the easier the problem.

The results are presented in Figure 2. The top plot show the number of transitions needed by both algorithms as the width of the arms changes, averaged over 500 independent runs. We see that the HIT-AND-RUN outperforms RRT for all but the hardest problems, usually by a large margin. The two lower plots show the sample points produced from one run with width equal to 1.2; we see that RRT has more uniform coverage, but that HIT-AND-RUN has a large speedup over linear sections, therefore justifying its faster exploration.

RRT is slow in this problem because in many rounds the tree does not grow in the *right* direction. For example

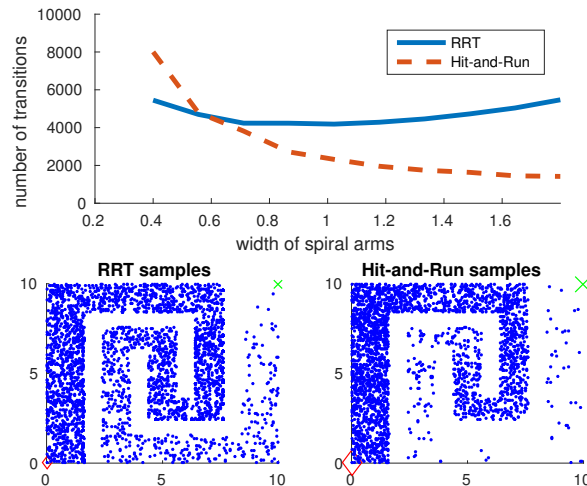


Figure 2: Position only planning example

at the beginning the tree needs to grow upwards, but most random samples will bias the growth to right. As HIT-AND-RUN only considers the space that is visible to the current point, it is less sensitive to the geometry of the free space. We can make this problem arbitrarily hard for RRT by making the middle part of the spiral fatter. HIT-AND-RUN, on the other hand, is insensitive to such changes. Additionally, the growth of the RRT tree can become very slow towards the end. This is because the rest of the tree absorbs most samples, and the tree grows only if the random point falls in the vicinity of the goal.

4.2 Kinematic Planning

In this set of simulations, we constrain the state transitions to adhere to the laws of physics: the state propagates forward under kinematics until it exits the permissible map, in which case it stops inelastically at the boundary. The position map is the two-turn corridor, illustrated in the bottom plots of Figure 3. Both algorithms propose points in the analogous manner to the previous section (where a desired speed is sampled in addition to a desired position); then, the best acceleration vector in the unit ball is calculated and the sample is propagated forward by the kinematics. If the sample point encounters the boundary, the velocity is zeroed. Both RRT and HIT-AND-RUN are constrained to use the same controller and the only difference is what points are proposed.

We see that HIT-AND-RUN again outperforms RRT across a large gamut of path widths by as much as a factor of three. The bottom two plots are of a typical sample path, and we see that HIT-AND-RUN has two advantages: it accelerates down straight hallways, and it samples more uniformly from the state space. In

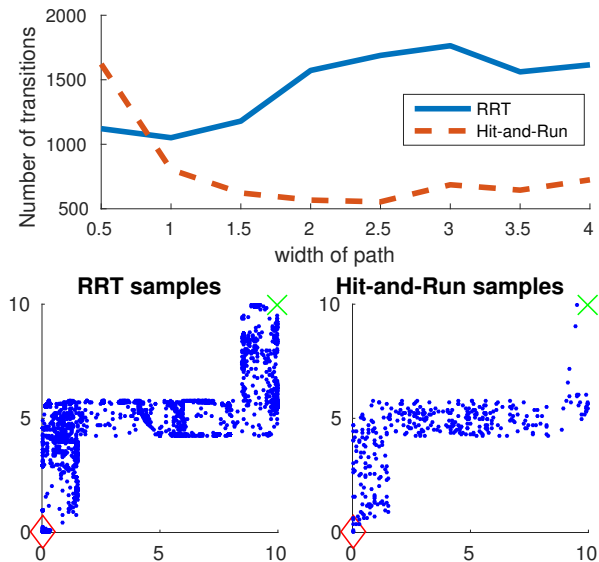


Figure 3: Performance under kinematic constraints

contrast, RRT wastes many more samples along the boundaries.

5 Conclusions and Future Work

This paper has two main contributions. First, we use a measure-preserving biLipschitz map to extend the analysis of the HIT-AND-RUN random walk to non-convex sets. Mixing time bounds for non-convex sets open up many applications, for example non-convex optimization via simulated annealing and similar methods. The second contribution of this paper has been to study one such application: the planning problem.

In contrast to RRT, using HIT-AND-RUN for planning has stronger guarantees on the number of samples needed and faster convergence in some cases. It also avoids the need for a sampling oracle for Σ , since it combines the search with an approximate sampling oracle. One drawback is that the sample paths for HIT-AND-RUN have no pruning and are therefore longer than the RRT paths. Hybrid approaches that yield short paths but also explore quickly are a promising future direction.

Acknowledgements

We gratefully acknowledge the support of the Australian Research Council through an Australian Laureate Fellowship (FL110100281) and through the Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS), and of the NSF through grant IIS-1619362.

References

- Y. Abbasi-Yadkori, J. Modayil, and Cs. Szepesvari. Extending rapidly-exploring random trees for asymptotically optimal anytime motion planning. In *International Conference on Intelligent Robots and Systems (IROS)*, 2010.
- K. Chandrasekaran, D. Dadush, and S. Vempala. Thin partitions: Isoperimetric inequalities and sampling algorithms for some nonconvex families. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2010.
- I. Fonseca and G. Parry. Equilibrium configurations of defective crystals. *Arch. Rational Mech. Anal.*, 120: 245–283, 1992.
- J. A. Griepentrog, W. Höppner, H-C. Kaiser, and J. Rehberg. A bi-Lipschitz continuous, volume preserving map from the unit ball onto a cube. *Note di Matematica*, 28:177–193, 2008.
- S. Karaman and E. Frazzoli. Sampling-based algorithms for optimal motion planning. *International Journal of Robotics Research.*, 30(7):846–894, 2011.
- S. M. LaValle. Rapidly-exploring random trees: A new tool for path planning. Technical Report TR 98-11, Computer Science Department, Iowa State University, 1998.
- S. M. LaValle and J. J. Kuffner. Randomized kinodynamic planning. *The International Journal of Robotics Research*, 20(5), 2001.
- L. Lovász. Hit-and-run mixes fast. *Mathematical Programming*, 86(3):443–461, 1999.
- L. Lovász and M. Simonovits. Random walks in a convex body and an improved volume algorithm. *Random Structures and Algorithms*, 4(4):359–412, 1993.
- L. Lovász and S. Vempala. Hit-and-run from a corner. *SIAM Journal on Computing*, 35(4):985–1005, 2006.
- L. Lovász and S. Vempala. The geometry of logconcave functions and sampling algorithms. *Random Structures & Algorithms*, 30(3):307–358, 2007.
- R. L. Smith. Efficient Monte-Carlo procedures for generating points uniformly distributed over bounded regions. *Operation Research*, 32:1296–1308, 1984.
- S. Vempala. Geometric random walks: A survey. *Combinatorial and Computational Geometry*, 52, 2005.