

Sample Complexity of Probabilistic Roadmaps via ϵ -nets

Matthew Tsao, Kiril Solovey, and Marco Pavone

Abstract— We study fundamental theoretical aspects of probabilistic roadmaps (PRM) in the finite time (non-asymptotic) regime. In particular, we investigate how completeness and optimality guarantees of the approach are influenced by the underlying deterministic sampling distribution \mathcal{X} and connection radius $r > 0$. We develop the notion of (δ, ϵ) -completeness of the parameters \mathcal{X}, r , which indicates that for every motion-planning problem of clearance at least $\delta > 0$, PRM using \mathcal{X}, r returns a solution no longer than $1 + \epsilon$ times the shortest δ -clear path. Leveraging the concept of ϵ -nets, we characterize in terms of lower and upper bounds the number of samples needed to guarantee (δ, ϵ) -completeness. This is in contrast with previous work which mostly considered the asymptotic regime in which the number of samples tends to infinity. In practice, we propose a sampling distribution inspired by ϵ -nets that achieves nearly the same coverage as grids while using fewer samples.

I. INTRODUCTION

The Probabilistic Roadmap Method (PRM) [1] is one of the most widely used sampling-based technique for motion planning. PRM generates a graph approximation of the full free space of the problem, by generating a set of configuration samples and connecting nearby samples when it is possible to move between configurations without collision using straight-line paths. PRM is particularly suitable in multi-query settings, where the workspace environment needs to be preprocessed to answer multiple queries consisting of different start and goal points. Recently PRM has been applied to challenging robotic settings, including manipulation planning [2], inspection planning and coverage [3], task planning [4], and multi-robot motion planning [5], [6]. PRM is instrumental in many modern single-query planners, which implicitly maintain a PRM graph and return a solution that minimizes the path's cost [7], [8], [9], [10], [11].

Extensive study of PRM's theoretical properties quickly followed its inception. The first question that occupied the research community was whether PRM guarantees to find a solution if one exists [12], [13], [14], and later on the quality of the returned solution. Several works have established the magnitude of the connection radius r sufficient to guarantee the convergence of the solution returned by PRM to an optimal solution [7], [15], [16]. However, the majority of works addressing those two questions consider in their analysis the (somewhat unrealistic) asymptotic regime, where the number of samples n tends to infinity. *The question of what are the smallest values of n, r to guarantee a high-quality solution in practice, i.e., when n is fixed, remains open.*

Statement of Contributions: In this work we make progress toward addressing the aforementioned question. In particular, we study how the sample set \mathcal{X} , and its cardinality n , as well as the size of the connection radius r , affect completeness

and optimality guarantees of PRM. We develop the notion of (δ, ϵ) -completeness of the parameters \mathcal{X}, r , which indicates that for every motion-planning problem of clearance $\delta > 0$, PRM using \mathcal{X}, r returns a solution no longer than $1 + \epsilon$ times the shortest path with at least δ clearance from the obstacles.

The concept of ϵ -nets [17] plays a key role in our contributions in both the theory and application. From a theoretical perspective, we leverage properties of ϵ -nets to characterize in terms of lower and upper bounds the sample size and connection radius needed to guarantee (δ, ϵ) -completeness. This is in contrast with previous work which mostly considered the asymptotic regime in which the number of samples tends to infinity. From an application perspective, we leverage properties of ϵ -nets via a template method to produce sample sets that efficiently cover the workspace. We observe empirically that these sample sets offer nearly the same coverage as grids while using fewer samples. Grids are an important baseline because they are used widely in practice and offer better coverage (i.e., dispersion) than uniform random sampling [18]. The increased efficiency provided by the template method over grids can improve the runtime of PRM and related algorithms. An extended version the paper provides additional information and missing proofs [19].

This paper is organized as follows. Related work is surveyed in Section II. Preliminaries are discussed in Section III. Section IV presents the main theoretical contributions and Section V presents proof sketches for some of the results. Numerical experiments comparing the efficiency of our ϵ -net based template method to grids are presented in Section VI. We summarize our work and discuss future directions in Section VII.

II. RELATED WORK

We provide a literature review of results concerning the theoretical properties of PRM. The majority of results apply to the setting of a Euclidean configuration spaces, and samples that are generated in a uniform and random fashion.

The study of asymptotic optimality in sampling-based planning was initiated in [15]. This paper proves that if $r > \gamma \left(\frac{\log n}{n}\right)^{1/d}$, for some constant $\gamma > 0$, then the length of the solution returned by PRM converges asymptotically almost surely (a.a.s.), as the number of samples $n \rightarrow \infty$, to length cost of the robust optimal solution. Such a connection radius leads to a graph of size $\Theta(n \log n)$, in contrast to a size of $\Theta(n^2)$ induced by a constant radius (as in [1]). Subsequent work managed to further reduce the constant γ [7], [20]. A recent paper [16] establishes the existence of a critical connection radius $r^* = \gamma' \left(\frac{1}{n}\right)^{1/d}$, where γ' is constant: if $r < r^*$ then PRM is guaranteed to fail (even when $n \rightarrow \infty$), and if $r > r^*$ then it is guaranteed to converge

The authors are with the Autonomous Systems Laboratory, Stanford University, CA 94305, USA. This work was supported in part by NSF, Award Number: 1931815.

a.a.s. to a near-optimal solution. A finite-time analysis of PRM, providing probabilistic bounds for achieving a given stretch factor $1 + \epsilon$ for a fixed number of samples n and a specific connection radius of the aforementioned form was established in [21], [22]. An algorithm has stretch factor β if it produces a solution whose length is no more than β times the length of the optimal solution.

The aforementioned results assume a uniform random sampling scheme. A recent work [18] establishes that using a low-dispersion deterministic sampling scheme (e.g., Halton and Sukharev sequences), asymptotic optimality is achieved with a radius as small as $f(n) \left(\frac{1}{n}\right)^{1/d}$, for any $\lim_{n \rightarrow \infty} f(n) = \infty$.

Two recent works [23], [24] consider non-Euclidean systems and develop sufficient conditions for asymptotic optimality with uniform random sampling.

III. PRELIMINARIES

We provide several basic definitions. Given two points $x, y \in \mathbb{R}^d$, denote by $\|x - y\|_p := (\sum_{i=1}^d |x_i - y_i|^p)^{1/p}$ the ℓ_p distance between them. When $p = 2$ we obtain the standard Euclidean distance. We denote the d -dimensional ℓ_p ball with radius $r > 0$ centered at $x \in \mathbb{R}^d$ as $B_p(x, r) := \{y : \|x - y\|_p \leq r\}$. For a Euclidean set \mathcal{X} , $\text{co}(\mathcal{X})$ denotes its convex hull and $\text{vol}(\mathcal{X})$ to denote its volume.

A. Motion planning

Denote by \mathcal{C} the configuration space of the robot, which, by rescaling we will assume is $[0, 1]^d$ throughout this paper. The free space $\mathcal{F} \subset \mathcal{C}$ denotes all collision-free configurations. A motion-planning problem is then specified by the tuple $(\mathcal{F}, x_{\text{start}}, x_{\text{goal}})$. The objective is to find a (continuous) path $p : [0, 1] \rightarrow [0, 1]^d$ that a) moves the robot from the start to goal location, i.e. $p(0) = x_{\text{start}}, p(1) = x_{\text{goal}}$ and b) avoids collisions with obstacles, i.e. $p(t) \in \mathcal{F}$ for all $t \in [0, 1]$.

We measure the quality of a path p by its length $\ell(p)$. A crucial property of paths in sampling-based planning is the notion of clearance. A motion-planning problem $(\mathcal{F}, x_{\text{start}}, x_{\text{goal}})$ has δ -clearance if there exists a path $p : [0, 1] \rightarrow [0, 1]^d$ with $p(0) = x_{\text{start}}, p(1) = x_{\text{goal}}$, and $\bigcup_{t \in [0, 1]} B_2(p(t), \delta) \subset \mathcal{F}$.

B. Probabilistic Roadmaps

We provide a formal definition of the Probabilistic Roadmaps Method (PRM). For a given motion-planning problem $\mathcal{M} := (\mathcal{F}, x_{\text{start}}, x_{\text{goal}})$, PRM generates a graph $G_{\mathcal{M}}$, whose vertices and edges represent collision-free configurations and straight-line paths connecting configurations, respectively.

The PRM graph induced by $\mathcal{M}, \mathcal{X}, r$ is denoted by $G_{\mathcal{M}}(\mathcal{X}, r) = (V, E)$. The vertices V are all collision-free configurations in $\mathcal{X} \cup \{x_{\text{start}}, x_{\text{goal}}\}$. The (undirected) edges E connect between every pair of vertices $u, v \in V$ such that (i) the distance between them is at most r , and (ii) the straight-line path between them is entirely collision free. Formally,

$$\begin{aligned} V &:= (\mathcal{X} \cup \{x_{\text{start}}, x_{\text{goal}}\}) \cap \mathcal{F}, \\ E &:= \{(u, v) \in V \times V : \|u - v\|_2 \leq r, \text{co}(\{u, v\}) \subset \mathcal{F}\}. \end{aligned}$$

C. Sampling distributions

The choice of sample set \mathcal{X} has significant implications on the properties of $G_{\mathcal{M}}(\mathcal{X}, r)$. We discuss two types of sample sets, namely ϵ -nets and grids, that are particularly useful for PRM. We use *sample set* and *sample distribution* interchangeably when referring to \mathcal{X} .

Definition 1 (ϵ -nets). A set $\mathcal{B} \subset \mathbb{R}^d$ is a ϵ -net for a set $\mathcal{A} \subset \mathbb{R}^d$ if for every $a \in \mathcal{A}$, there exists a $b \in \mathcal{B}$ so that $\|a - b\|_2 \leq \epsilon$.

As far as motion planning is concerned, ϵ -nets are good candidates for deterministic sampling schemes since by definition they have uniformly dense coverage over the entire space. The low dispersion sequences mentioned in [18] (i.e. Sukharev Grids) are special cases of ϵ -nets, and we will show that studying these objects in full generality leads to improved covering efficiency. *Minimal* ϵ -nets, i.e. nets with the smallest possible cardinality, are then good candidates for *efficiently* covering the entire space. We will make these intuitions formal in the coming sections. Grids have long seen use as sample sets in motion planning [25], so we use grids as benchmarks against which we compare the performance of sample sets inspired by ϵ -nets. We define a Sukharev grid of $[0, 1]^d$ with spacing w to be the following collection of points:

$$\mathcal{X}_{\text{grid}}(w) := \left\{ x \in \mathbb{R}^d : \frac{x_i}{w} + \frac{1}{2} \in \left\{ 1, 2, \dots, \frac{1}{w} \right\} \forall i \right\}.$$

See Section IV-C for a detailed comparison between grid sampling and ϵ -net sampling.

D. Completeness as a Benchmark

We will use the following benchmark to measure the quality of samples set \mathcal{X} and a connection radius.

Definition 2 ((δ, ϵ) -completeness). Given a samples set \mathcal{X} and a connection radius r , we say that the pair (\mathcal{X}, r) is (δ, ϵ) -complete if for every δ -clear \mathcal{M} it holds that

$$d_{G_{\mathcal{M}}(\mathcal{X}, r)}(x_{\text{start}}, x_{\text{goal}}) < (1 + \epsilon) \cdot \text{OPT}_{\delta},$$

where $d_{G_{\mathcal{M}}(\mathcal{X}, r)}(x_{\text{start}}, x_{\text{goal}})$ denotes the length of the shortest path from start to goal in the graph $G_{\mathcal{M}}(\mathcal{X}, r)$, and OPT_{δ} is the length of the shortest δ -clear solution to \mathcal{M} .

Note that $\epsilon = \infty$ corresponds to the case where finding a feasible path is the only objective.

IV. SAMPLE COMPLEXITY OF PRM

Our main objective is to study the sample complexity of PRM. For a clearance level $\delta > 0$ and a stretch tolerance $\epsilon > 0$, we wish to find comprehensive conditions on when a sample set \mathcal{X} and connection radius r are (δ, ϵ) -complete. The properties of ϵ -nets are central to our contributions to this goal, as well as our experimental results in Section VI.

We focus on deterministic sampling but mention that our techniques can be used to derive results for i.i.d. sampling procedures as well, which show a poorer performance of i.i.d. sampling, in line with the findings of [18]. As required by our main objective, all of our results are presented in the finite-sample, non-asymptotic setting.

Algorithm 1: Build-Net

1 **Input:** Set \mathcal{A} , cover radius $\epsilon > 0$;
2 **Output:** An ϵ -net \mathcal{B} for \mathcal{A} ;
3 $\mathcal{B} \leftarrow \emptyset$;
4 **while** $\mathcal{A} \setminus \cup_{b \in \mathcal{B}} B_2(y, \epsilon) \neq \emptyset$ **do**
5 | Find $y \in \mathcal{A} \setminus \cup_{b \in \mathcal{B}} B_2(b, \epsilon)$;
6 | $\mathcal{B} \leftarrow \mathcal{B} \cup \{y\}$;
7 **return** \mathcal{B} . ;

Algorithm 2: Epsilon Net Sampling (ENS)

1 **Input:** Sample size n , workspace dimension d ,
stretch tolerance $\epsilon > 0$, δ clearance parameter;
2 **Output:** Sample set \mathcal{X} and connection radius r ;
3 $\alpha \leftarrow \frac{\epsilon}{\sqrt{1+\epsilon^2}}$;
4 $n_\delta \leftarrow \min \left\{ \sqrt{\pi d} \left(\sqrt{\frac{2d}{\pi e}} \cdot \frac{1-(2-\alpha)\delta}{\alpha\delta} \right)^d, n \right\}$;
5 $\delta_{\min} \leftarrow \sqrt{\frac{2d}{\pi e}} \left(\frac{\sqrt{\pi d}}{n_\delta} \right)^{\frac{1}{d}} \left[\alpha - (2-\alpha) \sqrt{\frac{2d}{\pi e}} \left(\frac{\sqrt{\pi d}}{n_\delta} \right)^{\frac{1}{d}} \right]^{-1}$;
6 $r \leftarrow 2 \left(\alpha + \sqrt{1-\alpha^2} \right) \delta_{\min}$;
7 $\mathcal{X} \leftarrow \text{Build-Net}([\delta_{\min}, 1 - \delta_{\min}]^d, \delta_{\min})$;
8 **return** (\mathcal{X}, r) .

We will prove sufficient conditions on sample complexity in a constructive way by exhibiting algorithms that achieve the conditions. To this end, the **Epsilon Net Sampling** (ENS) procedure presented in Algorithm 2 will be used to certify the sufficient conditions. ENS leverages properties of ϵ -nets to determine a connection radius r and net resolution δ_{\min} . It then calls as a subroutine the **Build-Net** algorithm, which constructs a δ_{\min} -net of the configuration space by including points until a δ_{\min} -net is obtained, while ensuring that all points included are at pairwise distance at least δ_{\min} . In the next section, we will prove the efficacy of ENS which in turn gives sufficient conditions on the complexity of PRM.

A. Main Results

The following theorem provides a necessary condition on the size of a sample set \mathcal{X} and radius r to be (δ, ϵ) -complete.

Theorem 1 (Necessary Conditions). *Let $\mathcal{X} \subset [0, 1]^d$ be a set of n points and let $\delta > 0$. If (\mathcal{X}, r) is (δ, ∞) -complete then*

$$n \geq \sqrt{\frac{e}{2}} \left(1 - \frac{2\delta}{1-2\delta} \right)^2 \left(\sqrt{\frac{d-1}{2\pi e}} \cdot \frac{(1-2\delta)}{\delta} \right)^d$$
$$\text{and } r \geq (1-2\delta) \left(\sqrt{\pi d} \right)^{1/d} \sqrt{\frac{d}{2\pi e}} \left(\frac{1}{n} \right)^{1/d}.$$

In particular, any sample set smaller than the lower bound, regardless of r and how the points are chosen, cannot be (δ, ∞) -complete. Next, we focus our efforts in finding (\mathcal{X}, r) with size and radius comparable to the lower bound from Theorem 1, for which (δ, ϵ) -completeness is guaranteed. The following theorem leverages ENS to achieve this goal.

Theorem 2 (Sufficient Conditions). *Let $\epsilon > 0, \delta > 0$. If*

$$n \geq \sqrt{\pi d} \left(\sqrt{\frac{2d}{\pi e}} \cdot \frac{1-(2-\alpha)\delta}{\alpha\delta} \right)^d \text{ and}$$
$$r \geq 2 \left(1 + \frac{1}{\epsilon} \right) \left(\sqrt{\pi d} \right)^{1/d} \sqrt{\frac{d}{2\pi e}} \left(\frac{1}{n} \right)^{1/d}$$

where $\alpha = \frac{\epsilon}{\sqrt{1+\epsilon^2}}$, then $\text{ENS}(n, d, \delta)$ is (δ, ϵ) -complete.

See Section V for proof sketches of the sample size results from Theorems 1 and 2. See [19] for the full proofs.

B. Discussion

A few comments are in order. The connection radius condition in Theorem 2 generalizes the works of [15] and [18] to finite sample settings. Concretely, choosing $\epsilon = (\log n)^{-1/d}$ achieves asymptotic optimality with a radius $r = \gamma \left(\frac{\log n}{n} \right)^{1/d}$, recovering the result of [15]. This is because $\epsilon \rightarrow 0$ as $n \rightarrow \infty$, and the minimum δ satisfying the sample size condition in Theorem 2 goes to zero as $n \rightarrow \infty$. More generally, for any diverging function $f(n)$, choosing $\epsilon = \frac{1}{f(n)}$ recovers the asymptotic optimality condition $r = \gamma f(n) n^{-1/d}$ from [18].

Theorem 2 provides several other additions to the motion planning literature. It provides guarantees on the stretch factor achievable with finite n , which are not specified in the aforementioned works. This result is actionable in the sense that for a given δ and ϵ , it provides a sample size and distribution guaranteed to find a solution with the desired quality. Conversely, Theorem 2 can be used to certify of hardness (or as a non-existence proof [26], [27]): for any $n, \delta > 0$ satisfying the condition in Theorem 2, if there does not exist a feasible path in $\mathcal{G}_{\mathcal{M}}(\mathcal{X}, r)$ where $(\mathcal{X}, r) = \text{ENS}(n, d, \delta)$, then \mathcal{M} has clearance strictly less than δ .

We discuss the implications of our results on the sample complexity of PRM. For small dimensions, i.e., $d \leq 4$, Theorem 2 shows that $O(\delta^{-d})$ samples are sufficient for (δ, ∞) -completeness, since $\epsilon = \infty$ corresponds to $\alpha = 1$. Conversely, Theorem 1 shows that this sample complexity is essentially optimal (up to a multiplicative constant factor) in the sense that every (δ, ∞) -complete (\mathcal{X}, r) must have $|\mathcal{X}| = \Omega(\delta^{-d})$. In high dimensions, the ratio between the sufficient and necessary conditions on sample size is $\Omega(2^d \sqrt{d})$. Thus when d is no longer a small constant, there is a significant gap between the achievability result of Theorem 2 to the lower bound on sample complexity given by Theorem 1.

Evaluating the Bounds: To give a sense of the sample size conditions specified by Theorems 1 and 2, Table I evaluates bounds for $n = |\mathcal{X}|$ with respect various values of δ, d, ϵ .

From this we observe a poor scaling with clearance level δ , implying that PRM with classical sampling methods is likely not the right tool for high dimensional, low clearance motion-planning problems with low error tolerance (i.e., where success must be guaranteed). For example, if \mathcal{M} is a maze in $d = 5$ dimensions with path width 0.01 (i.e. clearance $\delta = 0.005$), then by Theorem 1, at least $9.2 \cdot 10^9$ samples are needed to ensure that PRM will find a solution to the maze. A graph this size is beyond what can be stored by modern computers.

TABLE I: Sample Complexity Examples

| δ | d | Thm. 1 LB | Thm. 2 UB | | |
|----------|-----|-------------------|---------------------|-------------------|---------------------|
| | | | $\epsilon = \infty$ | $\epsilon = 1$ | $\epsilon = 0.25$ |
| 0.25 | 4 | 0 | 252 | 669 | 22737 |
| 0.25 | 5 | 0 | 1430 | 4837 | $3.9 \cdot 10^5$ |
| 0.25 | 6 | 0 | 8781 | 37930 | $7.5 \cdot 10^6$ |
| 0.1 | 4 | 82 | 20411 | $7.15 \cdot 10^4$ | $4.2 \cdot 10^6$ |
| 0.1 | 5 | 570 | $3.48 \cdot 10^5$ | $1.66 \cdot 10^6$ | $2.6 \cdot 10^8$ |
| 0.1 | 6 | 4313 | $6.41 \cdot 10^6$ | $4.19 \cdot 10^7$ | $1.8 \cdot 10^{10}$ |
| 0.05 | 4 | 2983 | $4.1 \cdot 10^5$ | $1.52 \cdot 10^6$ | $9.9 \cdot 10^7$ |
| 0.05 | 5 | 46201 | $1.46 \cdot 10^7$ | $7.62 \cdot 10^7$ | $1.4 \cdot 10^{10}$ |
| 0.05 | 6 | $7.86 \cdot 10^5$ | $5.67 \cdot 10^8$ | $4.13 \cdot 10^9$ | $2.2 \cdot 10^{12}$ |

The story is more optimistic for larger values of δ (i.e., problems with higher clearance) and lower dimensions. When $\delta = 0.25$, a feasible solution can be guaranteed (see $\epsilon = \infty$) with 8000 samples. For $d = 4, \delta = 0.1$, a feasible solution can be guaranteed with 20,000 samples. In fact, a stretch factor of 2 (see $\epsilon = 1$) can be guaranteed in $\delta = 0.25$ clearance with a similar sample size on the order of 10^4 . Theorem 2 reveals that, with the right sampling scheme, PRM is guaranteed to find a solution efficiently and can do so in real time. For smaller clearance levels like $\delta = 0.1$, a solution can be guaranteed with somewhere between 10^5 and 10^6 samples, depending on the dimension. A solution with stretch factor 2 can be guaranteed by inflating the sample size by an additional factor of 10. This is no longer acceptable for real-time applications, but is still manageable for offline preprocessing if the environment will be queried many times, allowing for a cheap amortized cost.

While many practical robotic systems are high dimensional, i.e., $d \geq 6$, some application instances naturally admit decoupling of the degrees of freedom of the system, which induces lower-dimensional configuration subspaces. For instance, manipulation problems (see, e.g., [2]) can be typically decomposed into a sequence of tasks where the manipulator is driving toward an object (while fixing its arms), then moves an arm towards the object, and finally grasps it by actuating its fingers. Additionally, in some settings prior knowledge about the structure of the environment or a lower-dimensional space can inform sampling in the full configuration space, which can lead to more informative sampling distributions [28], [29], [30], [31]. Therefore, while PRM with classical sampling methods may not be the right tool for solving an entire motion planning problem, it can still be useful for solving subtasks in hierarchical task models or covering latent spaces efficiently.

C. Implications for Grid Sampling

The achievability result from Theorem 2 is obtained by using ENS. In this section we provide sufficient conditions on sample size for PRM to be (δ, ϵ) -complete when using grid sampling. Such a result serves as a benchmark for the proposed sampling algorithm ENS, and may be of independent interest as grids are commonly used in motion planning.

Corollary 1 (Sufficient Conditions for Grid Sampling). *Take r to satisfy the connection radius condition given in Theorem 2. Then $(\mathcal{X}_{grid}(\frac{2\alpha\delta}{\sqrt{d}}), r)$ is (δ, ϵ) -complete and*

$$\left| \mathcal{X}_{grid} \left(\frac{2\alpha\delta}{\sqrt{d}} \right) \right| = \left(\frac{\sqrt{d}}{2} \cdot \frac{1-2\delta}{\alpha\delta} \right)^d, \text{ where } \alpha = \frac{\epsilon}{\sqrt{1+\epsilon^2}}.$$

See [19] for a proof of corollary 1.

One natural comparison to make then is between the quality of coverage offered by grids versus ϵ -net approaches like Build-Net. A high resolution grid is in fact an example of an ϵ -net. There is, however, one key distinction between a grid and general ϵ -nets. Technically, a grid is a covering of the space using ℓ_∞ balls, whereas ϵ -nets provide coverings via ℓ_2 balls. Grids obtain their status as ϵ -nets through norm equivalence, i.e. $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{d}\|x\|_\infty \forall x \in \mathbb{R}^d$. The ℓ_2 norm is more prevalent than ℓ_∞ in the motion planning literature as δ -clearance and connection radius conditions are both stated with respect to Euclidean distance. It stands to reason that general ϵ -nets would provide a more efficient sampling strategy for motion planning problems and algorithms that are specified by ℓ_2 distance. Corollary 1 providing a worse (i.e. larger) sufficient condition than Theorem 2 corroborates this intuition. Furthermore, in this section we show that, under the condition $\epsilon < \sqrt{\frac{\pi e}{8}} - 1 \approx 0.033$, there exist ϵ -nets that are asymptotically (as $\epsilon \rightarrow 0, d \rightarrow \infty$) more efficient than grids when it comes to covering the space with ℓ_2 balls.

Lemma 1 (Net size and Grid size). *$\mathcal{X}_{grid}(w)$ is an ϵ -net if and only if $w \leq \frac{2\epsilon}{\sqrt{d}}$, thus the smallest grid that is also an ϵ -net is $\mathcal{X}_{grid}(\frac{2\epsilon}{\sqrt{d}})$. If \mathcal{X} is an ϵ -net obtained from Build-Net($[0, 1]^d, \epsilon$), then we have*

$$\frac{|\mathcal{X}|}{\left| \mathcal{X}_{grid} \left(\frac{2\epsilon}{\sqrt{d}} \right) \right|} \leq \sqrt{\pi d} \left(\frac{\sqrt{8}(1+\epsilon)}{\sqrt{\pi e}} \right)^d.$$

See [19] for a proof of Lemma 1. When $\epsilon < \sqrt{\frac{\pi e}{8}} - 1$ we have $\frac{\sqrt{8}(1+\epsilon)}{\sqrt{\pi e}} < 1$, hence the ratio given by Lemma 1 goes to zero as $d \rightarrow \infty$. We can then conclude that there exist ϵ -nets that are more efficient than grids for high dimensional problems. For lower dimensions, the upper bound from Lemma 1 is larger than 1. We suspect however, that the bound is loose in this regime and that ϵ -nets remain competitive with grids even for small dimensions. To test our hypothesis, we conduct an empirical comparison between ϵ -net inspired sample sets and grids in Section VI.

For some intuition as to why ϵ -nets provide more efficient coverage than grids in high dimensions, we note that Algorithm 1 returns an ϵ -net \mathcal{X} whose points are pairwise at distance at least ϵ . The grid on the other hand, has spacing at most $\frac{2\epsilon}{\sqrt{d}}$. The shrinkage factor of $O\left(\frac{1}{\sqrt{d}}\right)$ is precisely from the error in approximating ℓ_2 by ℓ_∞ . Since the minimum separation in the grid is smaller, the balls in $\cup_{x \in \mathcal{X}_{grid}(2\epsilon/\sqrt{d})} B_2(x, \epsilon)$ overlap more than the balls in $\cup_{x \in \mathcal{X}} B_2(x, \epsilon)$ and thus lead to less efficient coverage.

V. PROOF SKETCHES

In this section we sketch the proofs of the sample size conditions in Theorems 1 and 2. To set the stage, we first discuss the cardinality of minimal ϵ -nets, which play a key role in the proofs. The following result provides upper and lower bounds on the size of minimal nets.

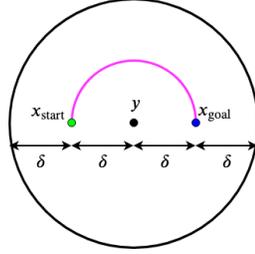
Theorem 3 (Cardinality of ϵ -nets). *Let $\mathcal{A} \subset \mathbb{R}^d$ be a set.*

- 1) Every ϵ -net of \mathcal{A} must have cardinality at least $\text{vol}(\mathcal{A})\sqrt{\pi d} \left(\sqrt{\frac{d}{2\pi e}} \frac{1}{\epsilon} \right)^d$.
- 2) There exists a ϵ -net of \mathcal{A} with cardinality at most $\text{vol}(\mathcal{A} \oplus B_2(y, \frac{\epsilon}{2})) \sqrt{\pi d} \left(\sqrt{\frac{2d}{\pi e}} \frac{1}{\epsilon} \right)^d$.

Where $\mathcal{A} \oplus \mathcal{B} := \{a + b : a \in \mathcal{A}, b \in \mathcal{B}\}$ denotes the Minkowski sum of the sets \mathcal{A} and \mathcal{B} . See [32] for a proof of Theorem 3. The **Build-Net** procedure outlined by Algorithm 1 is one particular way to generate an ϵ -net with property 2 of Theorem 3.

A. Proof Sketch for Theorem 1

The outline for the proof sketch is as follows. (1) Suppose the sample set $\{x_i\}_{i=1}^n$ has size $n < \sqrt{\pi d} \left(\sqrt{\frac{d}{2\pi e}} \frac{1-4\delta}{2\delta} \right)^d$. (2) By part 1 of Theorem 3, $\{x_i\}_{i=1}^n$ cannot be a 2δ -net of $[2\delta, 1-2\delta]^d$. (3) This means that there exists some y so that $\{x_i\}_{i=1}^n \cap B_2(y, 2\delta) = \emptyset$. Furthermore, since $y \in [2\delta, 1-2\delta]^d$, we have $B_2(y, 2\delta) \subset [0, 1]^d$. (4) Consider the motion planning problem $\mathcal{M} := (\mathcal{F}, x_{\text{start}}, x_{\text{goal}})$: The black dot is y , $x_{\text{start}} = y - \delta e_1$, $x_{\text{goal}} = y + \delta e_1$ ($\{e_i\}_{i=1}^d$ are the standard basis vectors for \mathbb{R}^d .) are depicted by the green and blue dots respectively, and the obstacle set $\bar{\mathcal{F}} := \{y\} \cup \{x : \|x - y\|_2 = 2\delta\}$ is shown in black. This problem is δ -clear, and a solution trajectory is shown in pink. There is, however, no way to reach x_{goal} from x_{start} in $G_{\mathcal{M}}$. This is because the obstacle $\{y\}$ blocks the edge between x_{start} and the x_{goal} , and the shell $\{x : \|x - y\|_2 = 2\delta\}$ blocks any edges from $\{x_{\text{start}}, x_{\text{goal}}\}$ to $\{x_i\}_{i=1}^n$, since the points $\{x_i\}_{i=1}^n$ are outside of $B_2(y, 2\delta)$. Thus, the nodes $x_{\text{start}}, x_{\text{goal}}$ will have no neighbors in $G_{\mathcal{M}}$. We conclude that, regardless of the value of r , no sample set of size $n < \sqrt{\pi d} \left(\sqrt{\frac{d}{2\pi e}} \frac{1-4\delta}{2\delta} \right)^d$ can be (δ, ∞) -complete in $[0, 1]^d$. The proof of Theorem 1 uses this idea but packs the space using a torus instead of a ball. Since a similar argument can be done using the torus, and the torus used has smaller volume than the ball, more tori can be packed, leading to a larger lower bound.



B. Proof Sketch for Theorem 2

Suppose \mathcal{M} is a δ_{\min} -clearance motion planning problem for some $\delta_{\min} > 0$. It then has a shortest δ_{\min} -clear solution path p . Consider points $\{p_j\}_{j=1}^m$ on the path p so that consecutive points are at $2\sqrt{1-\alpha^2}\delta_{\min}$ euclidean distance apart. See Fig. 1 for an illustration. For each $1 \leq j \leq m-1$, by δ_{\min} -clearance the balls of radius δ_{\min} centered at p_j, p_{j+1} are collision free. This is illustrated by the blue region in Fig. 1. If we have a collection of samples \mathcal{X} which forms a $\alpha\delta_{\min}$ -net of the space, then we are guaranteed that there are $z_j, z_{j+1} \in \mathcal{X}$ that are within $\alpha\delta_{\min}$ distance of p_j and p_{j+1} respectively. This is illustrated by the green regions. The condition on n given in the statement of Theorem 2 in conjunction with Theorem 3 guarantees the existence of

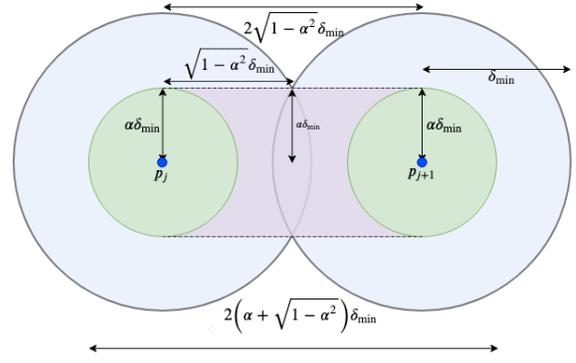


Fig. 1: A visualization of the proof technique of Theorem 2 in 2 dimensions.

such a \mathcal{X} . The spacing $2\sqrt{1-\alpha^2}\delta_{\min}$ was chosen so that the convex hull of the green set, depicted as the union of the green and violet regions, is entirely contained in the blue, collision free region. This means that any line segment joining a point in the left green ball to a point in the right green ball is collision-free. Therefore, for each j , the line segment joining z_j to z_{j+1} is collision-free for every $1 \leq j \leq m-1$. Furthermore, by triangle inequality, the length of these segments can be at most $2(\alpha + \sqrt{1-\alpha^2})\delta_{\min}$. Thus, if we choose r to be this value, then the edges $\{(z_j, z_{j+1})\}_{j=1}^{m-1}$ will all be in the graph $G_{\mathcal{M}}(\mathcal{X}, r)$. Since p is a solution path, we can choose the first and last samples $p_1 = x_{\text{start}}$ and $p_m = x_{\text{goal}}$. Therefore the path that concatenates these edges will be a collision-free path from the start to the goal in the graph $G_{\mathcal{M}}(\mathcal{X}, r)$. Finally, to bound the length of this path, note that the length of the path from p_j to p_{j+1} is $\|p_j - p_{j+1}\|_2 = 2\sqrt{1-\alpha^2}\delta_{\min}$. Thus,

$$\frac{\|z_j - z_{j+1}\|_2}{\|p_j - p_{j+1}\|_2} \leq \frac{2(\alpha + \sqrt{1-\alpha^2})\delta_{\min}}{2\sqrt{1-\alpha^2}\delta_{\min}} = 1 + \frac{\alpha}{\sqrt{1-\alpha^2}} = 1 + \epsilon$$

where the last equality is due to the definition of α . Since each piece of the path defined by $\{z_j\}_{j=1}^m$ is not more than $1 + \epsilon$ times its corresponding piece in p , this gives a path whose total length is at most $1 + \epsilon$ times the length of p .

VI. EXPERIMENTS

In this section we demonstrate the improved efficiency of sampling methods that are based on ϵ -nets, compared to grids. While Lemma 1 shows that ϵ -nets are provably more efficient than grids asymptotically as $d \rightarrow \infty$, it does not address the comparison in low dimensions when d is small. To make an empirical comparison in the non-asymptotic regime, we numerically construct sample sets based on ϵ -nets and compare their size and coverage quality to grids.

A. A Sampling Procedure via Templating

When ϵ is small, it is computationally expensive to find a cover of the space with balls of radius ϵ . To alleviate this computational burden, we make the following observation. Suppose \mathcal{T} is a ϵ -net of $[0, m^{-1}]^d$ for some $m \in \mathbb{N}_+$. Since $[0, 1]^d$ can be written as a union of m^d translations of $[0, m^{-1}]^d$, the union of the analogous m^d translates of \mathcal{T} gives an ϵ -net of $[0, 1]^d$. Grids are in fact one example of ϵ -nets with this periodic structure. See Fig. 2 for an example in $d = 2$ dimensions. If instead we want an $\epsilon/2$ net, we can replicate $\frac{1}{2}\mathcal{T}$ a total of $(2m)^d$ times. We call \mathcal{T} a *template*,

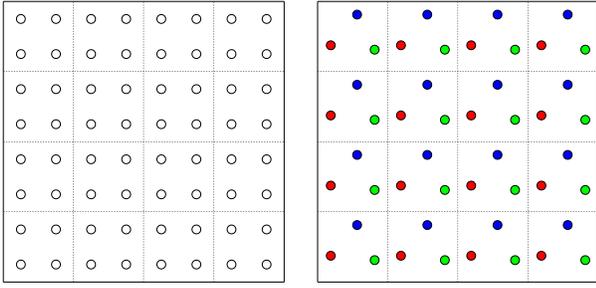


Fig. 2: **Left:** A grid of width 8 in $d = 2$ dimensions can be represented by repeating a grid of width 2 a total of 16 times. **Right:** More generally, any $\mathcal{T} := (x_1, x_2, x_3)$ which forms an ϵ -net of $[0, \frac{1}{4}]^2$ can be replicated 16 times to attain an ϵ -net of $[0, 1]^2$.

since scaling it and replicating it appropriately can create ϵ -nets of arbitrary resolution and size. Since rescaling and replication are computationally cheap, the sample sets can be computed online so long as the templates are precomputed. Due to this versatility, it is acceptable for the computation of these templates to be done offline.

With this in mind our goal is to find templates of small cardinality for a large range of problem dimensions. Our benchmark will be a grid with spacing $w := \frac{2\epsilon}{\sqrt{d}}$, since we showed in Lemma 1 that a grid is an ϵ -net if and only if its spacing is at most w . The cube $[0, kw]^d$ is covered by a grid of spacing w with exactly k^d points. See Fig. 2 for a 2D example. Thus our goal is to find a ϵ -net of $[0, kw]^d$ with fewer than k^d points. This problem is homogeneous in ϵ , so by re-scaling, this is equivalent to finding a $\frac{\sqrt{d}}{2k}$ -net of $[0, 1]^d$ using fewer than k^d points. Moreover, since grids themselves are periodic, the ratio of sizes of the sample set obtained by repeating \mathcal{T} to the grid will also be $|\mathcal{T}|/k^d$.

Certifying that \mathcal{T} is a $\frac{\sqrt{d}}{2k}$ -net of $[0, 1]^d$ is nontrivial because there are infinitely many conditions to satisfy. Instead we sample a large collection of points \mathcal{V} uniformly at random from $[0, 1]^d$, and obtain a $\frac{\sqrt{d}}{2k}$ -net of those points via the output of Build-Net($\mathcal{V}, \frac{\sqrt{d}}{2k}$). The resulting set of points \mathcal{T} may not be a $\frac{\sqrt{d}}{2k}$ -net of $[0, 1]^d$ since it is only certified to be a ϵ -net of the set of densely sampled points. We say a point is uncovered by \mathcal{T} if its nearest neighbor in \mathcal{T} is at distance more than $\frac{\sqrt{d}}{2k}$. We verify using Monte Carlo simulation that only a negligible fraction of $[0, 1]^d$ is uncovered.

B. Numerical Results

We created templates \mathcal{T} using the procedure described in the previous section for various values of dimension d and k . We recorded the size of the resulting template and the empirical estimate \hat{p} on the volume of uncovered space. For each trial, \hat{p} was computed via Monte Carlo with 10 million samples. The ratio between $|\mathcal{T}|$ and the size of the grid benchmark, k^d , is denoted as ρ . An efficient sample set \mathcal{T} will lead to a small value of ρ . Table II summarizes the results of this experiment.

C. Discussion

From Table II we see that the relative efficiency compared to the grid, ρ , improves as the dimension grows. This observation is consistent with the asymptotic implications of Lemma 1. For all dimensions, the value of \hat{p} is on the order of

TABLE II: Template sizes for various dimensions

| d | $k = 2$ | | | $k = 3$ | | |
|-----|-----------------|--------|---------------------|-----------------|--------|---------------------|
| | $ \mathcal{T} $ | ρ | \hat{p} | $ \mathcal{T} $ | ρ | \hat{p} |
| 4 | 15 | 0.93 | $3.7 \cdot 10^{-2}$ | 77 | 0.95 | $1.4 \cdot 10^{-2}$ |
| 5 | 27 | 0.84 | $3.3 \cdot 10^{-2}$ | 189 | 0.77 | $5.7 \cdot 10^{-3}$ |
| 6 | 57 | 0.89 | $5.5 \cdot 10^{-3}$ | 457 | 0.63 | $2.4 \cdot 10^{-3}$ |
| 7 | 105 | 0.82 | $3.1 \cdot 10^{-3}$ | 1078 | 0.50 | $1.4 \cdot 10^{-3}$ |
| 8 | 173 | 0.68 | $1.6 \cdot 10^{-3}$ | 2477 | 0.38 | $6.4 \cdot 10^{-4}$ |
| 9 | 291 | 0.58 | $1.3 \cdot 10^{-3}$ | 5650 | 0.29 | $2.6 \cdot 10^{-4}$ |

10^{-3} , which means that the templates are $\frac{\sqrt{d}}{2k}$ -nets of almost the entire space. Notice that the achieved value of ρ is better when $k = 3$ than when $k = 2$ across all tested dimensions. This is because the Build-Net procedure does not exploit the fact that \mathcal{T} will be used in a periodic manner. Simply put, there are covering inefficiencies at the boundaries when templates are used in a periodic manner. When covering $[0, 1]^d$ with larger templates, there are fewer boundaries between individual template replicas, and thus fewer instances where this boundary inefficiency exists. This reveals a natural trade-off between computation and sample efficiency. Indeed, the cover radius $\frac{\sqrt{d}}{2k}$ is a decreasing function of k , meaning it is more difficult computationally to find templates for larger k . Thus using small k is computationally cheap, but as Table II shows, the quality of the template is worse than what would be obtained for larger k . Looking at the extremes, $k = 1$ corresponds to the grid, and $k = \frac{\sqrt{d}}{2\epsilon}$ corresponds to finding an ϵ -net of $[0, 1]^d$ without exploiting templates or periodic translation in any way.

VII. CONCLUSION AND FUTURE WORK

We made progress in the characterization of sample complexity of PRM. We provided lower bounds on the sample size that is necessary for (δ, ϵ) -complete sampling algorithms. We then complemented the lower bound with achievability results by analyzing ϵ -net and grid based sampling schemes. These sampling schemes are then showed to attain, up to constant factors, the optimal sample complexity for lower dimensional problems. Through numerical experiments we exhibited an ϵ -net inspired sampling strategy, termed templating, that offers nearly the same coverage quality as grids while using significantly fewer samples.

There are several interesting directions for future research. First, the gap between the sufficient and necessary conditions for (δ, ϵ) -completeness is dimension dependent. In high dimensions, the characterization is no longer tight, and the precise dependence of sample complexity on dimension is not yet known. In fact, the gap in our results is due to the gap in characterization of ϵ -net sizes in Theorem 3, thus closing that gap would have implications for our results. Second, while the templates proposed in our experiments show nearly the same coverage quality as grids, they may not be true ϵ -nets since they are only certified to be an ϵ -net of a large subset of the space. While we showed empirically that the volume of uncovered points is typically on the order of 10^{-3} , we wish to build templates that are certifiably ϵ -nets for the whole space while retaining reduced size observed in our experiments. Since Build-Net is used in both our theoretical and experimental results, a better algorithm for constructing ϵ -nets would improve results both in theory and practice.

REFERENCES

- [1] L. E. Kavraki, P. Svestka, J. Latombe, and M. H. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Trans. Robotics and Automation*, vol. 12, no. 4, pp. 566–580, 1996. [1](#)
- [2] A. Kimmel, R. Shome, Z. Littlefield, and K. E. Bekris, "Fast, anytime motion planning for prehensile manipulation in clutter," in *18th IEEE-RAS International Conference on Humanoid Robots, Humanoids 2018, Beijing, China, November 6-9, 2018*, 2018, pp. 1–9. [1](#), [4](#)
- [3] M. Fu, A. Kuntz, O. Salzman, and R. Alterovitz, "Toward asymptotically-optimal inspection planning via efficient near-optimal graph search," in *Robotics: Science and Systems*, Freiburg im Breisgau, Germany, 2019. [1](#)
- [4] C. R. Garrett, T. Lozano-Pérez, and L. P. Kaelbling, "Ffrob: Leveraging symbolic planning for efficient task and motion planning," *International Journal of Robotics Research*, vol. 37, no. 1, pp. 104–136, 2018. [1](#)
- [5] R. Shome, K. Solovey, A. Dobson, D. Halperin, and K. E. Bekris, "dRRT*: Scalable and informed asymptotically-optimal multi-robot motion planning," *Autonomous Robots*, Jan 2019. [1](#)
- [6] W. Hönig, J. A. Preiss, T. K. S. Kumar, G. S. Sukhatme, and N. Ayanian, "Trajectory planning for quadrotor swarms," *IEEE Trans. Robotics*, vol. 34, no. 4, pp. 856–869, 2018. [1](#)
- [7] L. Janson, E. Schmerling, A. A. Clark, and M. Pavone, "Fast marching tree: A fast marching sampling-based method for optimal motion planning in many dimensions," *International Journal of Robotics Research*, vol. 34, no. 7, pp. 883–921, 2015. [1](#)
- [8] J. A. Starek, J. V. Gómez, E. Schmerling, L. Janson, L. Moreno, and M. Pavone, "An asymptotically-optimal sampling-based algorithm for Bi-directional motion planning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2015, pp. 2072–2078. [1](#)
- [9] O. Salzman and D. Halperin, "Asymptotically-optimal motion planning using lower bounds on cost," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 4167–4172. [1](#)
- [10] J. D. Gammell, S. S. Srinivasa, and T. D. Barfoot, "Batch informed trees (BIT*): Sampling-based optimal planning via the heuristically guided search of implicit random geometric graphs," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 3067–3074. [1](#)
- [11] K. Solovey and D. Halperin, "Sampling-based bottleneck pathfinding with applications to Fréchet matching," in *European Symposium on Algorithms*, 2016, pp. 76:1–76:16. [1](#)
- [12] L. E. Kavraki, M. N. Kolountzakis, and J. Latombe, "Analysis of probabilistic roadmaps for path planning," *IEEE Trans. Robotics and Automation*, vol. 14, no. 1, pp. 166–171, 1998. [1](#)
- [13] A. M. Ladd and L. E. Kavraki, "Fast tree-based exploration of state space for robots with dynamics," in *Algorithmic Foundations of Robotics*, 2004, pp. 297–312. [1](#)
- [14] S. Chaudhuri and V. Koltun, "Smoothed analysis of probabilistic roadmaps," *Comput. Geom.*, vol. 42, no. 8, pp. 731–747, 2009. [1](#)
- [15] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *International Journal of Robotics Research*, vol. 30, no. 7, pp. 846–894, 2011. [1](#), [3](#)
- [16] K. Solovey and M. Kleinbort, "The critical radius in sampling-based motion planning," *International Journal of Robotics Research*, 2019. [1](#)
- [17] N. H. Mustafa and K. Varadarajan, "Epsilon-nets and epsilon-approximations," in *Handbook of Discrete and Computational Geometry*, 3rd ed., J. E. Goodman, J. O'Rourke, and C. D. Toth, Eds. CRC Press LLC, 2016, ch. 47. [Online]. Available: <http://www.csun.edu/~ctoth/Handbook/HDCG3.html> [1](#)
- [18] L. Janson, B. Ichter, and M. Pavone, "Deterministic sampling-based motion planning: Optimality, complexity, and performance," *International Journal of Robotics Research*, 2017. [1](#), [2](#), [3](#)
- [19] M. Tsao, K. Solovey, and M. Pavone, "Sample complexity of probabilistic roadmaps via epsilon nets, extended version," 2019. [Online]. Available: <https://arxiv.org/pdf/1909.06363.pdf>
- [20] K. Solovey, O. Salzman, and D. Halperin, "New perspective on sampling-based motion planning via random geometric graphs," *International Journal of Robotics Research*, vol. 37, no. 10, 2018. [1](#)
- [21] A. Dobson and K. E. Bekris, "A study on the finite-time near-optimality properties of sampling-based motion planners," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, November 3-7, 2013*, 2013, pp. 1236–1241. [2](#)
- [22] A. Dobson, G. V. Moustakides, and K. E. Bekris, "Geometric probability results for bounding path quality in sampling-based roadmaps after finite computation," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 4180–4186. [2](#)
- [23] E. Schmerling, L. Janson, and M. Pavone, "Optimal sampling-based motion planning under differential constraints: The drift case with linear affine dynamics," in *IEEE Conference on Decision and Control*, 2015, pp. 2574–2581. [2](#)
- [24] —, "Optimal sampling-based motion planning under differential constraints: The driftless case," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 2368–2375. [2](#)
- [25] S. M. LaValle, *Planning Algorithms*. Cambridge, U.K.: Cambridge University Press, 2006. [2](#)
- [26] J. Basch, L. J. Guibas, D. Hsu, and An Thai Nguyen, "Disconnection proofs for motion planning," in *IEEE International Conference on Robotics and Automation*, vol. 2, 2001, pp. 1765–1772. [3](#)
- [27] Z. McCarthy, T. Bretl, and S. Hutchinson, "Proving path non-existence using sampling and alpha shapes," in *IEEE International Conference on Robotics and Automation*, 2012, pp. 2563–2569. [3](#)
- [28] W. Reid, R. Fitch, A. Gktoan, and S. Sukkarieh, "Motion planning for reconfigurable mobile robots using hierarchical fast marching trees," in *Workshop on the Algorithmic Foundations of Robotics*, 2016. [4](#)
- [29] B. Ichter, J. Harrison, and M. Pavone, "Learning sampling distributions for robot motion planning," in *IEEE International Conference on Robotics and Automation*, Brisbane, Australia, May 2018. [4](#)
- [30] B. Ichter and M. Pavone, "Robot motion planning in learned latent spaces," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2407–2414, 2019. [4](#)
- [31] S. Choudhury, M. Bhardwaj, S. Arora, A. Kapoor, G. Ranade, S. Scherer, and D. Dey, "Data-driven planning via imitation learning," *International Journal of Robotics Research*, vol. 37, no. 13-14, 2018. [4](#)
- [32] Y. Wu, "Packing, covering, and consequences on minimax risk," in *Course Lecture notes for ECE598: Information-theoretic methods for high-dimensional statistics*, 2016, ch. 14. [Online]. Available: <http://www.stat.yale.edu/~yw562/teaching/it-stats.pdf>