

Balancing Global Exploration and Local-connectivity Exploitation with Rapidly-exploring Random disjointed-Trees

Tin Lai, Fabio Ramos and Gilad Francis*

Abstract—Sampling efficiency in a highly constrained environment has long been a major challenge for sampling-based planners. In this work, we propose Rapidly-exploring Random disjointed-Trees* (RRdT*), an incremental optimal multi-query planner. RRdT* uses multiple disjointed-trees to exploit local-connectivity of spaces via Markov Chain random sampling, which utilises neighbourhood information derived from previous successful and failed samples. To balance local exploitation, RRdT* actively explore unseen global spaces when local-connectivity exploitation is unsuccessful. The active trade-off between local exploitation and global exploration is formulated as a multi-armed bandit problem. We argue that the active balancing of global exploration and local exploitation is the key to improving sample efficient in sampling-based motion planners. We provide rigorous proofs of completeness and optimal convergence for this novel approach. Furthermore, we demonstrate experimentally the effectiveness of RRdT*'s locally exploring trees in granting improved visibility for planning. Consequently, RRdT* outperforms existing state-of-the-art incremental planners, especially in highly constrained environments.

I. INTRODUCTION

Sampling-based path planners (SBPs) provide a robust approach to robotic motion planning, where the objective is to produce a sequence of actions for the system to transits from an initial point to a goal point under a set of constraints. These planners are favourable as they offer robustness in high-dimensional configuration spaces (*C-space*), since the sampling procedure replaces explicit construction of the—often intractable—*C-space*. An SBP samples configuration points randomly and connects valid points in a graph or tree-like structure. This structure is then searched for a possible solution, and given sufficient time the planner is guaranteed to find a solution if one exists. This guarantee is often called *probabilistic completeness* [1].

An SBP can be categorically classified as either single-query or multi-query. A single-query planner such as Rapidly-exploring Random Tree (RRT) [2] returns a feasible path that connects a pair of initial and goal points. Multi-query planners such as Probabilistic Roadmap (PRM) [3], on the other hand, constructs a topological graph that allows the planner to perform different instances of queries efficiently. Star variants of these planners, such as RRT* [4] and PRM* [5], are considered *asymptotically optimal* [1], as they further guarantee that the solution will converge, in the limit, to the optimal solution, given a user-defined path associated cost.

Although SBPs are probabilistically complete, their runtime performance is significantly influenced by *C-space*'s complex-

ity. Intuitively, tightly constrained regions are problematic as they limit the connectivity of free space. Indeed, this is a widely recognised issue of SBPs, which is often framed as planning in narrow passages [6]. Limited free space within narrow passages implies a low probability of picking a point inside. Narrow passages severely restrict the performance of single-query algorithms since low-occurring samples that do happen to fall inside narrow passages are merely thrown away if the tree fails to expand. Meaning, the growth of the tree is restricted by the surrounding obstacles, resulting in limited visibility and high failure rate in extending new connections. Consequently, the growth of tree will be bottlenecked by narrow passages until a series of successful points within narrowed passages had been sampled.

In this paper, we propose Rapidly-exploring Random disjointed-Trees* (RRdT*) as a novel anytime SBP [7] based on RRT*, which inherit RRT*'s *probabilistic completeness* and *asymptotic optimality* guarantees. RRdT* explores *C-space* with multiple locally exploring disjointed trees (d-trees). The initial roots of the d-trees are scattered randomly in *C-space*, which is followed by expansion of the d-trees, exploiting local connectivity. This approach overcomes standard SBPs' failure in creating connections in highly restricted space, as RRdT* will create a new d-tree at the sampled point if it fails to connect to an existing d-trees. Consequently, the RRdT* path planner can be re-framed as a multi-armed bandit (MAB), which balances global exploration and local-connectivity exploitation of *C-space*.

Our contribution is an incremental SBP that exploits local-connectivity to improve performance while balancing exploration. This is achieved by formulating the balance of the competing choices as an MAB with infinite mortal arms and non-stationary reward sequence. We provide rigorous proofs on completeness and optimal convergence guarantees. We show in simulations that such formulation yields superior results with high sampling efficiency in highly constrained *C-space*, tackling the limited visibility issue that many SBPs faces.

II. RELATED WORK

Research in SBPs focused on: (i) runtime performance for finding an initial solution, and (ii) convergence rate on obtaining an optimal solution. Improvements to convergence rates are achieved by focusing sampling to certain regions in *C-space*, e.g., restricting search space to a prolate hypersphere [8], or adaptively bias toward regions with limited visibility [9]. Several SBP algorithms use bridge tests to discover narrow passages [10], followed by dense re-sampling

*Tin Lai, Fabio Ramos and Gilad Francis are with The School of Information Technologies, University of Sydney, Australia tin.lai@sydney.edu.au

at those regions [11]–[13]. Another approach is retraction-based planners that optimise to generate samples close to the boundary of obstacles [14], [15]. However, all methods require a user-defined heuristic for finding narrow spaces that naturally has no explicit representation.

Planners using multiple exploring trees were presented previously too, for example, growing bidirectional trees [16], growing multiple local trees with a heuristic tree selection scheme [17], and combining bridge test with a learning technique to model the probability of tree selection [18]. While the concept of utilising multiple trees had improved sampling efficiency, these algorithms are restrictive in tree locations or require explicit computations for finding narrow passages, which are hard to generalise to high-dimensional C -space.

Markov Chain Monte Carlo (MCMC) has also been used by SBPs, as it utilises information observed from previous samples, instead of merely discarding failed attempts [19]. Al-Blawi *et al.* showed that a PRM is in fact the result of a set of MCMC explorations being run simultaneously [20]. A Monte Carlo Random Walk path planner was proposed in [21], where a stochastic planner explores the neighbourhood of a search state by constructing a Markov Chain to propose spaces with high contributions.

Several authors proposed SBPs that balance between exploration and exploitation behaviours during planning, borrowing from standard practice in reinforcement learning. Heuristic biasing was used to guide tree search greedily while retaining probability for random exploration [22]. However, this method does not incorporate valuable information from invalid samples. Rickert *et al.* [23] proposed a 2-step SBP that exploits C -space's structure after an initial workspace exploration step. While this approach significantly improved performance while avoiding being trapped globally, it lacks *probabilistic completeness* guarantee.

Our method, RRdT*, actively balances explorative and exploitative behaviour to achieve high sampling efficiency. As an incremental multi-query SBP it uses multiple trees to exploit local spaces via chained sampling at unvisited spaces. The active balances between the two enhance performance while maintaining completeness and convergence guarantees.

III. RAPIDLY-EXPLORING RANDOM DISJOINTED-TREES*

In this section, we formalise the path planning problem and describe the RRdT* algorithm.

A. Problem formulation

Let $C \subseteq \mathbb{R}^d$ denotes the C -space, where $d \geq 2$; and $C_{obs} \subseteq C$ denotes the set of invalid states. The obstacle-free space C_{free} is the closure set $\text{cl}(C \setminus C_{obs})$. Let $q \in C$ denotes a state in C -space. A feasible path is formally defined as follows:

Definition 1 (Feasible path): A path is a sequence of consecutively connected configurations $\sigma : [0, 1] \rightarrow C_{free}$. A path is said to be feasible if $\sigma(0) = q_{init}$, $\sigma(1) = q_{goal}$, and it is collision free: $\sigma(\tau) \in C_{free} \forall \tau \in [0, 1]$.

The objective of a path planning algorithm is to construct a feasible path σ from an initial point q_{init} to a goal point

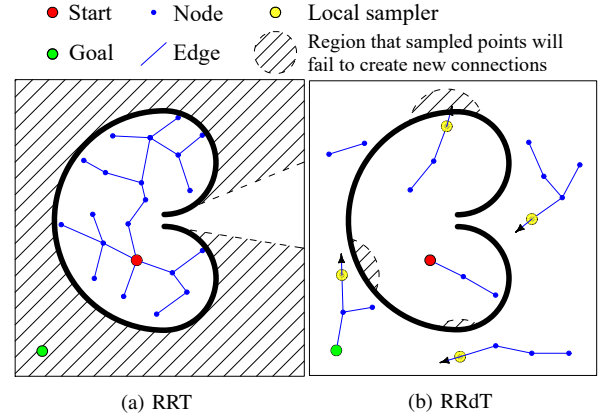


Fig. 1. Sampling scenarios in the bug-trap problem [9] with same number of nodes. (a) RRT: Expansions of tree being restricted by surrounding obstacles, has very limited visibility (white regions) to utilise sampled points. (b) RRdT: Large visible region with high probability of creating new connections. Local samplers can exploit local-connectivity in parallel with less failures.

q_{goal} . Therefore, path planning can be considered as one of the following problems:

Problem 1 (Feasibility planning): Given a configuration space C , a free space C_{free} , an obstacle space C_{obs} , an initial configuration $q_{init} \in C_{free}$, and a goal configuration $q_{goal} \in C_{free}$. Find a path that satisfy definition 1.

Problem 2 (Optimal planning): Let there be a *cost function* $c(\sigma)$ that assigns non-negative cost to all non-trivial feasible paths defined by definition 1. For all possible paths from problem 1, find $\sigma^* : [0, 1] \rightarrow \text{cl}(C_{free})$, such that $\sigma^*(0) = q_{init}$, $\sigma^*(t) = q_{goal}$, and $c(\sigma^*) = \min_{\sigma \in \text{cl}(C_{free})} c(\sigma)$.

B. High-level description

RRdT* aims to strike a balance between global exploration and local exploitation with effective utilisation of information obtained from sampled points. RRT¹ tree expansion is limited to a local scope bounded by the neighbourhood visibility of the tree nodes as illustrated in Fig. 1a. It continuously tries to grow the root tree outwards by sampling random points and creating new connections to the closest nodes. Therefore, RRT will often reject updates from valid samples when there exists no free route from the closest existing node towards a sampled point. For example, in Fig. 1a, any sampled point that falls inside the hatched region will be discarded due to failure in tree expansion. This problem is exacerbated in a tightly constrained C -space, where there are not many readily available free routes. RRdT* maintains high visibility by using d-trees to explore C -space. It exploits local connectivity by employing local samplers (yellow circle in Fig. 1b) that serve as density estimators. D-trees explore within C_{free} by estimating which portions of its surrounding have the highest probability of being free space. Thus, maintaining local connectivity information without wasting samples. As illustrated in Fig. 1b, RRdT* mitigates the limited visibility by creating new local samplers in regions, which are not visible from the root tree. Local sampling in RRdT* is performed by moving an ϵ -distance step proposed by the local

¹Although RRdT* is based on RRT*, the concept is interchangeable with RRT, which will be used for ease of illustration.

sampler. By employing local sampling, RRdT* will only fail to create new connections at non-visible locations within the ϵ -ball, which has a smaller volume compared to the entire C -space. Sampling locally avoids explicit computation for identifying narrow passages, which often lacks theoretical performance guarantees [18], or trades-off *completeness* with computational speed gains [23].

RRdT* can be used for re-planning, and it returns a feasible solution faster than general multi-query planners like PRM. PRM performs shortest path search only after exhausting its sampling budget. In contrast, RRdT* is incremental and maintains an exploring tree from a root that is continuously adding new samples. The balance between exploration and exploitation fuses the advantageous properties from the two main SBPs—tree connectivity from RRT and high visibility from PRM. Effective balancing can ensure faster convergence rate while maintaining asymptotic *completeness*. The proposal of RRdT* aims to fill in this gap with provable guarantees.

C. Balancing exploration and exploitation

Exploiting local connectivity comes naturally in path planning, where the objective is to construct a directed path that connects two given points. Let σ_{sol} denotes a potential solution path, $\sigma_{\overline{q_i q_j}}$ denotes a path with $\sigma(0) = q_i, \sigma(1) = q_j$, and $\mathcal{V}(q)$ denotes the visibility set of q , which represents the region of C_{free} visible from some $q \in C_{free}$. Intuitively, each point q_i that lies on σ_{sol} is connected to at least two other points (parent and child node) that lie on the same solution. That is, each point on the solution path has at least two exploitable locally connected nodes; only exceptions are q_{init} and q_{goal} , which does not has a parent and child node respectively.

$$\begin{aligned} \exists \geq 2 q_j, q_j \in \mathcal{V}(q_i) \wedge q_i \neq q_j \wedge \sigma_{\overline{q_i q_j}} \in C_{free} \\ \forall q_i \in \sigma_{sol} \setminus \{q_{init}, q_{goal}\} \end{aligned} \quad (1)$$

This statement of at least two local connections is a lower bound that remains true for all path planning settings. Higher visibility implies more available local connections to exploit. However, most SBPs do not utilise—or exploit—this fact to avoid the pitfall of being stuck in dead-ends; for example, random sampling by PRM ignores valuable local connectivity information. On the other hand, heuristic-based algorithms such as potential field method [24] tries to exploit obstacle information, but often will not find a feasible path even if one exists [25]. The proposed RRdT* strikes a balance between exploiting the locality of C -space and avoid being trapped within obstacles by actively balancing the exploration of unvisited space, as we describe below:

1) *Exploitation on local-connectivity*: Local exploitation in RRdT* is performed by a *local sampler*—a moving density estimator that defines a proposal distribution based on previous successes, which is utilised by MCMC random walk. This attempts to exploit the local connectivity between each point as a chained sampling procedure built from the previous movement’s direction. Fig. 2 demonstrates a local sampler navigating within a narrow passage, utilising previous successful observations to adapt its proposal distribution.

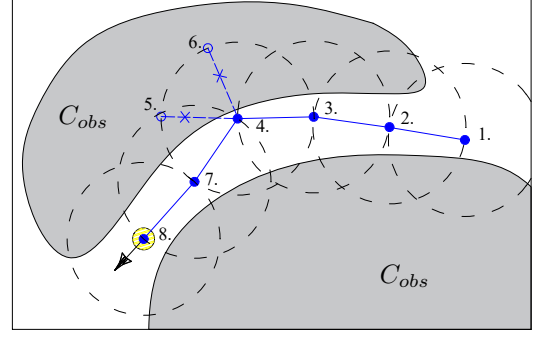


Fig. 2. Local connectivity exploitation using MCMC random walk inside narrow passage. With a local sampler first initialised at location 1, 2-4 are successive and successful samples. The 5th sample fails the collision check. Local sampler then retries to re-sample sampler’s proposal distribution (6th sample) and fails again. The 7th sample finds a free route to continue the chained sampling. If local sampler get stuck during the process (e.g. at a dead-end), MAB scheduler will restart arm at a new location.

2) *Exploration on global-space*: Exploration is implicitly performed by restarting a local sampler at a new location. Global exploration in RRdT* utilise generic uniformly random sampling of C -space, such that it inherits *probabilistic completeness* and asymptotic guarantees of optimal convergence from RRT*. RRdT* as a framework does not use a global dynamic sampling schedule. Replacing uniform schedule with a dynamic schedule, such as the one used in Informed RRT* [8], is trivial, but is beyond the scope of this paper.

3) *Balancing scheduler*: Balancing global exploration and local exploitation in RRdT* can be formulated as an MAB. The objective of an MAB scheduler is to allocate resources effectively based on the complexity of C -space. Spaces with a complex environment can utilise local-connectivity information to exploits obstacles’ structure; whilst spaces with high visibility can allocate more resources on rapid exploring.

We employ Mortal MAB [26] with non-stationary reward sequences [27] and infinitely many arms [28] as RRdT*’s MAB scheduler. An arm a_i samples and moves within C_{free} ; its state at time t is denoted as $a_{i,t}$ where it will makes an observation o_t and receives a reward $R(a_{i,t})$. We consider discrete time and Bernoulli arms, with success probability as:

$$\mathbb{P}(a_{i,t} \mid a_{i,t-1}, o_{t-1}) \forall i \in \{1, \dots, k\} \quad (2)$$

We consider the arms as independent random variables with $\mathbb{P}(a_{i,t})$ having non-stationary distribution (as each arm represents a moving density estimator of connections in C_{free}), which depends on its previous successes. The reward sequence of a_i is a stochastic sequence with unknown payoff distributions and changes rapidly according to the complexity of C -space. The reward $R(a_{i,t})$ is a mapping of C -space, where $a_{i,t}$ is mapped to a pay-off of 1 if $a_i \in C_{free}$ at t , and 0 otherwise. An arm is ‘mortal’ in a sense that it is used to model a local sampler that moves around in C -space; thus it has a stochastic lifetime after which they expire due to events such as running into a dead-end or stuck in a narrow passage. The action space is infinite ($k = \infty$) because a new arm can be created on demand by randomly sampling a new point, especially after an arm is expired.

Algorithm 1: MAB arm restart scheduler

```
1 function RestartArm( $A, D, \eta$ )
2   foreach  $a_i$  in  $\{a_j \in A \mid \mathbb{P}(a_j) < \eta\}$  do
3      $q_{rand} \leftarrow \text{SampleFree}$ 
4     Join  $q_{rand}$  to all  $d_i \in D$  within  $\epsilon$ -distance
5     if Not joined to existing tree then
6       Create new arm  $a_{new}$  at  $q_{rand}$ 
7        $A \leftarrow A \cup \{a_{new}\} \setminus a_i$ 
8     return True  $\triangleright$  new node is added
9   return False  $\triangleright$  No sampler restarted
```

D. Implementation

Algorithm 1 shows RRdT*'s MAB restart scheduler. The scheduler checks whether the probability of each arm a_i in the set of arms A is lower than η , an arbitrarily small constant for $\eta \in \mathbb{R}, \eta > 0$. A low probability indicates successive failures (e.g. stuck in a dead-end), which results in the arm randomly relocated to a new location in C_{free} . The new location, q_{rand} , is assigned by global exploration sampling schedule, `SampleFree`. However, if the new sampled point, q_{rand} , is within ϵ -distance of an existing d-tree d_i from the set of trees D , the new point will be added to the appropriate d-tree instead (lines 4 to 7). The resulting behaviour is identical to RRT* (extending existing node ϵ -distance towards sampled point), which implies RRdT*'s asymptotic behaviour converges to RRT* after C -space is covered by d-trees.

Algorithm 2 defines the main RRdT* algorithm. RRdT* exploration starts by restarting k arms randomly. Then, it draws an arm from a multinomial distribution in `PickArm`, with each arm $a_i \in A$ having probability formulated by (2). Once an arm had been picked, RRdT* exploits arm's neighbourhood by sampling using a local proposal distribution (line 7). To incorporate past sampling successes, we employed an MCMC random walker with a von Mises-Fisher (vMF) [29] distribution, which is a wrapped Gaussian distribution, commonly used for directional statistics. It enables us to construct a chained directed sampling using a random walker that produces a sample q_{new} . If q_{new} is valid, the arm's position $a_i.pos$ is updated accordingly (line 11). A schematic example of a directed local-connectivity search of C_{free} is shown in Fig. 2. A practical illustration of local sampling at different locations simultaneously is shown in Fig. 4a.

Asymptotic optimality of RRdT* is maintained by `Rewire` (line 10), a procedure that rewires edges based on their neighbours' accumulated cost [4]. Global exploration in RRdT* is implicitly performed when the MAB scheduler restarts an arm at a randomly chosen location in C_{free} (line 3). Thus, exploration is performed when local exploitation is unsuccessful, which actively adjust the planner's immediate objective based on C -space's complexity.

IV. ANALYSIS

We will first define the notion of *expansiveness* [30].

Definition 2 (expansiveness): Let $\mu(X)$ denotes the Lebesgue measure of X , which represents its volume. The free

Algorithm 2: RRdT* Algorithm

```
Input:  $q_{init}, q_{goal}, N, k, \epsilon, \eta$ 
Initialise:  $Root \leftarrow G(V = \{q_{init}\}, E = \emptyset)$ 
               $T \leftarrow \{Root\}; A \leftarrow \emptyset; n \leftarrow 1$ 
1 Initialise  $k$  arms into  $A$  and d-trees into  $D$ 
2 while  $n \leq N$  do
3   if RestartArm( $A, D, \eta$ ) then  $\triangleright$  Alg. 1
4      $n \leftarrow n + 1$ 
5   else
6      $a_i \leftarrow \text{PickArm}(A)$ 
7      $q_{new} \leftarrow a_i$  samples locally via MCMC
8     if  $\sigma_{a_i.pos, q_{new}} \in C_{free}$  then
9       Join  $q_{new}$  to all  $d_i \in D$  within  $\epsilon$ -distance
10      if  $q_{new}$  joined to  $Root$  then Rewire( $Root$ )
11      Update  $a_i.pos$  to  $q_{new}$ 
12       $n \leftarrow n + 1$ 
13    Updates  $a_i$  probability
```

space C_{free} is said to be $(\alpha, \beta, \epsilon)$ -expansive [30] if each of its connected components $C'_{free} \subseteq C_{free}$ satisfy the following conditions:

- (i) for every point $q \in C'_{free}$, $\mu(\mathcal{V}(q)) \geq \epsilon$
- (ii) for any connected subset $S \subseteq C'_{free}$, the set $\beta\text{-LOOKOUT}(S) = \{q \in S \mid \mu(\mathcal{V}(q) \setminus S) \geq \beta \mu(C'_{free} \setminus S)\}$ has volume $\mu(\beta\text{-LOOKOUT}(S)) \geq \alpha \mu(S)$

This notation guarantees C_{free} is ϵ -good [31]; and measures complexity of C with the quantity of visibility by $q \in S$.

A. Feasibility planning

Let n_{dtree} denotes the total number of d-trees restart in the time-span of the algorithm. Then, although a new d-tree is restarted whenever a new arm is added, n_{dtree} will be bounded by a real value constant. Consider the followings:

Assumption 1: The free space $C_{free} \subseteq C$ is a finite set.

Lemma 1 (Termination of d-tree restarting): Let Assumption 1 hold. Then, the total number of d-trees restarted for any given C -space is always finite. That is, there exists a constant $\phi \in \mathbb{R}$ such that $n_{dtree} < \phi$ for any given C .

Proof: A new d-tree will be restarted by sampling a new point in \mathbb{R}^d . From Alg. 1, if there exists a d-tree within ϵ -distance, the new point will be added as a node to that d-tree instead of creating a new d-tree. Hence, every new d-tree must be at least ϵ -distance away from each other. We can formulate each d-tree as an ϵ -ball centred at the d-tree's origin. In the limiting case, C_{free} will be filled by the volume of ϵ -balls. Let $\epsilon(q)$ denotes the ϵ -ball of q , and V be the set of all nodes. Then, the above statement is formally defined as:

$$\lim_{t \rightarrow \infty} C_{free} \setminus \bigcup_{q \in V_t} \epsilon(q) = \emptyset \quad (3)$$

Therefore, n_{dtree} is upper bounded by how many ϵ -balls can C_{free} fits. It is immediate that if C_{free} is a finite set, n_{dtree} will always be bounded by a constant. This guarantee ensures the tractability of using d-trees to explore C_{free} will always terminate. \blacksquare

Assumption 2: The MAB scheduler has a discounting factor such that the probability of all arms will eventually decay to zero.

Lemma 2 (Infinite random sampling): Let Assumption 2 holds, and $n_{i,t}$ denotes the number of uniformly random points added to a d-tree $d_i \in D$ at time t . Then, $n_{i,t}$ always increases without bound, i.e., as $t \rightarrow \infty$, $n_{i,t} \rightarrow \infty$.

Proof: Assumption 2 ensures that the probability of all arms in A will eventually decay to zero, no matter how successful they are. Therefore, the MAB restart scheduler, as defined in Alg. 1, will perform infinitely many times. Therefore, there exist infinitely-many uniformly random samples, and $n_{i,t}$ always increases without bound. ■

Let $n_t^{RRdT^*}$ and $n_t^{RRT^*}$ be the number of uniformly random points sampled at time t for RRdT* and RRT* respectively. The differences of the asymptotic sampling scheme between RRdT* and RRT* is at most a constant.

Theorem 1: There exists a constant $\phi \in \mathbb{R}$ such that

$$\lim_{t \rightarrow \infty} \mathbb{E} \left[\frac{n_t^{RRdT^*}}{n_t^{RRT^*}} \right] \leq \phi. \quad (4)$$

Proof: There exists two different sampling schemes being employed in RRdT*.

- (i) **Global:** *Random sampling* when an arm restarts (Alg. 1)
- (ii) **Local:** *MCMC random walk sampling* when an arm exploits local neighbourhood (Alg. 2 line 7).

The total number of uniformly random points sampled for RRdT* is the summation of each individual d-tree's randomly sampled points, i.e., $n_t^{RRdT^*} = \sum_{i=1}^{|D|} n_{i,t}$. Deriving from Lemma 2, it guarantees each d-tree $d_i \in D$ will always has infinite uniformly sampled points $n_{i,t}$ to improve the tree structure. Hence, $\lim_{t \rightarrow \infty} n_t^{RRdT^*} = \infty$. Therefore, no matter how successful (ii) is, it will be dominated by (i) in the limiting case. Therefore, the ratio of the total number of uniformly random points between RRdT* and RRT* will be bounded by a constant as the behaviour of RRdT* will converge to RRT* as $t \rightarrow \infty$. ■

With Lemma 2 and theorem 1, the *probabilistic completeness* of RRdT* is immediate.

Theorem 2 (Probabilistic Completeness): RRdT* inherits the same probabilistic completeness of RRT*. That is, if there exists a feasible solution to problem 1, then $\lim_{i \rightarrow \infty} \mathbb{P}(\{\sigma_{sol}(i) \cap q_{goal} \neq \emptyset\}) = 1$

B. Asymptotic optimality planning

Theorem 3 (see [18]): Select k points q_1, \dots, q_k randomly from C'_{free} including q_{init} and q_{goal} . Set q_1, \dots, q_k as root nodes and extend k trees from these points. Let n be the total number of nodes that all these trees extended, and $\gamma \in \mathbb{R}$ be a real number in $(0, 1]$. If n satisfies:

$$n \geq k(\alpha\beta\epsilon)^{-1} \ln[4(1-\epsilon)] \ln\{3 \ln[2k^2(1-\epsilon)]/\gamma\beta\} + k\mu(C_{free})/\mu(C) \ln(3k^2/2\gamma) \quad (5)$$

then the probability that each pair of these k trees can be attached successfully is at least $1 - \gamma$.

Due to space constraints, please refers to [18] for full proof of Theorem 3. Then, the joining of d-trees in RRdT* can be shown to conform to the same bound.

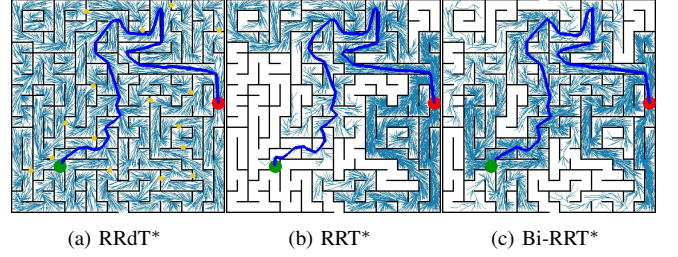


Fig. 3. Maze environment. Incremental SBPs with 10,000 nodes. (a) edges evenly distributed, and visible to all C_{free} . (b) uneven edges distribution with limit visibility. (c) same as RRT* but edges can grow from two points.

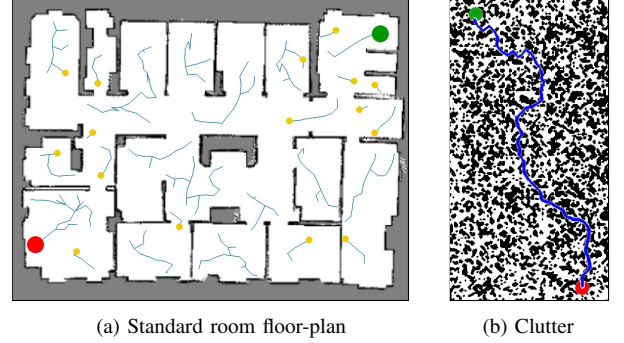


Fig. 4. Environments. (a) Room: typical map for baseline (least constrained); also showcasing RRdT* local samplers exploit multiple local-spaces simultaneously. (b) Clutter: randomly generated, extremely limited visibility.

Theorem 4 (Joining of d-trees): Let $C'_{free} \subseteq C_{free}$ be a connected free space; $d_i, d_j \in D$ be an instance of d-tree in C'_{free} where $d_i \neq d_j$. If there exists a feasible path to connect d_i to d_j , both trees will eventually join as $t \rightarrow \infty$.

Proof: Lemma 1 states that the restarting of new d-trees will terminate; hence, the number of d-trees is finite. Lemma 2 ensures the number of uniformly random points sampled for all $d_i \in D$ is infinite as $t \rightarrow \infty$. Hence, in the limiting case, all d-trees satisfy the real-valued bound given by Theorem 3. Therefore, all d-trees in C'_{free} will eventually join to a single tree. ■

Theorem 5 (Asymptotic optimality): Let $\sigma_t^{RRdT^*}$ be the solution returned by RRdT* at time t , and c^* is the minimal path cost for problem 2. If a solution exists, then the cost of $\sigma_t^{RRdT^*}$ will converges to optimal cost almost-surely. That is:

$$\mathbb{P} \left(\left\{ \lim_{t \rightarrow \infty} c(\sigma_t^{RRdT^*}) = c^* \right\} \right) = 1 \quad (6)$$

Proof: From Theorem 4, all d-trees in the same C'_{free} will converge to a single tree. According to Lemma 2 there will be infinite sampling available to improve that tree; and together with adequate rewiring procedure [4] by Alg. 2 line 10, it is guaranteed that the solution will converge to the optimal solution as $t \rightarrow \infty$. ■

V. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of our method in simulated environments with various degrees of complexity: a standard room floor-plan² (Fig. 4a), a maze³ (Fig. 3), and

²Available at <http://robotang.co.nz/projects/robotics/custom-player-plugins>

³Available at <http://www.mazegenerator.net/Examples.aspx>

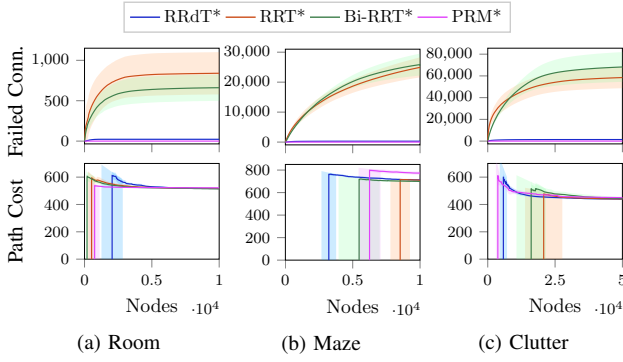


Fig. 5. Comparison of failed connections (top) and path cost (bottom) as a function of tree nodes in various planning scenarios, shaded region indicates the standard deviation of the 20 repetitions. RRdT* maintains a low failed connections rate in all environments, whereas PRM* does not creates any connections while sampling. (a) Room: simple environment, RRdT* requires more nodes to return initial solution, however, converges to the same solution. (b) Maze: more complex, RRdT* reached initial solution faster than other planners. (c) Clutter: highly complex environment, RRdT* consistently requires similar amount of nodes and outperforms other incremental SBPs.

TABLE I
SAMPLED POINTS FOR 10,000 NODES. FAILURE COMPARISON.
($\mu \pm 2\sigma$) WITH TWO SIGNIFICANT FIGURES

Env.	Method	Sampled Points ($\cdot 10^3$)		
		Total	Fail to connect	Points in C_{obs}
Room	RRdT*	21 \pm 0.32	0.023\pm0.010	11 \pm 0.32
	RRT*	22 \pm 0.88	0.84 \pm 0.52	11 \pm 0.70
	Bi-RRT*	21 \pm 0.64	0.66 \pm 0.32	11 \pm 0.54
	PRM*	21 \pm 0.30	N/A	11 \pm 0.30
Maze	RRdT*	12\pm0.11	0.29\pm0.052	1.7 \pm 0.10
	RRT*	40 \pm 6.8	25 \pm 6.6	5.3 \pm 1.0
	Bi-RRT*	41 \pm 6.8	26 \pm 6.8	5.3 \pm 1.0
	PRM*	12\pm0.090	N/A	1.5\pm0.088
Clutter	RRdT*	22 \pm 0.29	1.3\pm0.11	10.93 \pm 0.26
	RRT*	100 \pm 22	42 \pm 16	48 \pm 15
	Bi-RRT*	102 \pm 15	44 \pm 12	47 \pm 10
	PRM*	19\pm0.34	N/A	9.1\pm0.34

clutter (Fig. 4b); listed in decreasing order of their $(\alpha, \beta, \epsilon)$ -expansiveness [30]. Experiments were performed with 20 different pairs of start and goal locations, each repeated 20 times. Planning was performed with a fixed budget of nodes, and planners could run until the budget is exhausted. The pairs of locations with the highest cost were plotted as Fig. 5. RRdT*, RRT* [4], Bi-RRT* [16], and PRM* [5] were implemented under the same planning framework in Python. Note that since PRM* is not an anytime algorithm [7], a solution is only available after the entire budget is exhausted. Therefore, for comparing RRdT* against an equivalent multi-query planner, the result of PRM* presented here reconstructs its entire graph every 250 iterations (and discarded afterwards).

We used the number of sampled points as a metric to measure sampling efficiency. Table I shows results of sampled points by the various SBPs. Because PRM* does not attempt to create connections during sampling, the *fail to connect* measurements is not applicable. All planners behave simi-

arly in the simple *Room* environment. In more complex environments like the *Maze* and *Clutter*, RRdT* obtained the same number of nodes with fewer samples compared to other incremental SBPs. In fact, sampling efficiency of RRT* and Bi-RRT* in Table I is identical. However, as depicted in Fig. 5, Bi-RRT* outperforms RRT* in time-to-solution, which is indicative of the benefits of growing multiple trees from different roots. Indeed, RRdT* significantly outperforms RRT* and Bi-RRT* in finding an initial solution in environments with limited visibility. RRdT*'s high visibility allows it to exploits multiple local spaces simultaneously, without the restriction of having only a single root. In the event that such space brings no new connections, RRdT* can restarts the arm at other locations while retaining the previous d-tree for re-connection later. Hence, it retains information that other incremental SBPs throw away during failed tree expansion.

Fig. 3 highlights the different behaviour between each incremental SBPs. Both RRT* and Bi-RRT* are restricted by surrounding obstacles which limits its visibility to explores the spaces thoroughly. Whereas RRdT* has a faster time-to-solution (Fig. 5) in the *Maze*, explore spaces incrementally and evenly, whilst having the ability to perform re-planning. The MCMC local sampling employed by RRdT* also performs better than PRM* (Fig. 5), since PRM*'s visibility between milestones is blocked by the straight walls, rendering inferior result in creating linkages. Besides, PRM* is not an anytime algorithm; a solution is only available after the budget is exhausted. Thus, in a normal setting, it would never outperform RRdT*. Moreover, RRdT* fuses the advantageous properties of RRT* and PRM*. RRdT* maintains similar sample counts as PRM*, but being an incremental SBP, it can return an initial solution faster.

VI. CONCLUSION

We presented RRdT*, an incremental multi-query SBP that actively balances global exploration and local-connectivity exploitation, formulated as an MAB problem. By exploiting local connection, RRdT* maintains sample efficiency and high visibility with robustness against highly complex spaces.

Limited visibility, especially within narrow passages, has a restricting effect on tree expansions in incremental planners. RRdT* mitigates it by exploring *C-space* with multiple exploring disjointed-trees, which on their own exploits local-connectivity. High visibility of RRdT* saves computational resources, allowing it to exploits local structures and behave consistently even in a highly constrained environment. Active exploration combined with MCMC Random Walk exploitation is a novel approach that brings sampling efficiency to incremental SBPs, whilst keeping theoretical guarantees.

We believe the exploration-exploitation balances makes RRdT* highly adaptive and robust towards different complexity of *C-space*. In future, a dynamic sampling procedure can be used in place of the generic uniform sampling for further enhancement. Exploring with multiple trees in RRdT* is highly suitable to be implemented as parallel programming. The concept of an active balance can also be applied in a wide range of SBPs, to utilise resources available efficiently.

REFERENCES

- [1] M. Elbanhawi and M. Simic, 'Sampling-based robot motion planning: A review', *Ieee access*, vol. 2, pp. 56–77, 2014.
- [2] S.M. LaValle, 'Rapidly-exploring random trees: A new tool for path planning', 1998.
- [3] L. E. Kavraki, P. Svestka, J. C. Latombe *et al.*, 'Probabilistic roadmaps for path planning in high-dimensional configuration spaces', *IEEE transactions on Robotics and Automation*, vol. 12, no. 4, pp. 566–580, Aug. 1996.
- [4] S. Karaman and E. Frazzoli, 'Incremental Sampling-based Algorithms for Optimal Motion Planning', *Robotics: Science and Systems Foundation*, 27 Jun. 2010.
- [5] S. Karaman and E. Frazzoli, 'Sampling-based algorithms for optimal motion planning', *The international journal of robotics research*, vol. 30, no. 7, pp. 846–894, 2011.
- [6] D. Hsu, L. E. Kavraki, J.-C. Latombe *et al.*, 'On finding narrow passages with probabilistic roadmap planners', in *Robotics: THE ALGORITHMIC PERSPECTIVE: 1998 Workshop on the Algorithmic Foundations of Robotics*, 1998, pp. 141–154.
- [7] S. Karaman, M. R. Walter, A. Perez *et al.*, 'Anytime Motion Planning using the RRT*', in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 1478–1483.
- [8] J. D. Gammell, T. D. Barfoot and S. S. Srinivasa, 'Informed sampling for asymptotically optimal path planning', *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 966–984, 2018.
- [9] A. Yershova, L. Jaillet, T. Siméon *et al.*, 'Dynamic-domain RRTs: Efficient exploration by controlling the sampling domain', in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference On*, IEEE, 2005, pp. 3856–3861.
- [10] S. A. Wilmarth, N. M. Amato and P. F. Stiller, 'MAPRM: A probabilistic roadmap planner with sampling on the medial axis of the free space', in *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference On*, vol. 2, IEEE, 1999, pp. 1024–1031.
- [11] D. Hsu, T. Jiang, J. Reif *et al.*, 'The bridge test for sampling narrow passages with probabilistic roadmap planners', in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference On*, vol. 3, IEEE, 2003, pp. 4420–4426.
- [12] Z. Sun, D. Hsu, T. Jiang *et al.*, 'Narrow passage sampling for probabilistic roadmap planning', *IEEE Transactions on Robotics*, vol. 21, no. 6, pp. 1105–1115, 1 Dec. 2005.
- [13] W. Wang, X. Xu, Y. Li *et al.*, 'Triple RRTs: An Effective Method for Path Planning in Narrow Passages', *Advanced Robotics*, vol. 24, no. 7, p. 943, 1 Jan. 2010.
- [14] L. Zhang and D. Manocha, 'An efficient retraction-based RRT planner', in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference On*, IEEE, 2008, pp. 3743–3750.
- [15] J. Lee, O. Kwon, L. Zhang *et al.*, 'SR-RRT: Selective retraction-based RRT planner', in *Robotics and Automation (ICRA), 2012 IEEE International Conference On*, IEEE, 2012, pp. 2543–2550.
- [16] J. J. Kuffner and S. M. LaValle, 'RRT-connect: An efficient approach to single-query path planning', in *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference On*, vol. 2, IEEE, 2000, pp. 995–1001.
- [17] M. Strandberg, 'Augmenting RRT-planners with local trees', in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference On*, vol. 4, IEEE, 2004, pp. 3258–3262.
- [18] W. Wang, L. Zuo and X. Xu, 'A Learning-based Multi-RRT Approach for Robot Path Planning in Narrow Passages', *Journal of Intelligent & Robotic Systems*, vol. 90, no. 1-2, pp. 81–100, 2018.
- [19] W. Chen, 'Motion Planning with Monte Carlo Random Walks', PhD Thesis, University of Alberta, 2015.
- [20] I. Al-Bluwai, T. Siméon and J. Cortés, 'Motion planning algorithms for molecular simulations: A survey', *Computer Science Review*, vol. 6, no. 4, pp. 125–143, 2012.
- [21] H. Nakhost, J. Hoffmann and M. Müller, 'Resource-Constrained Planning: A Monte Carlo Random Walk Approach.', in *ICAPS, 2012*.
- [22] C. Urmson and R. Simmons, 'Approaches for heuristically biasing RRT growth', in *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference On*, vol. 2, IEEE, 2003, pp. 1178–1183.
- [23] M. Rickert, O. Brock and A. Knoll, 'Balancing exploration and exploitation in motion planning', in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference On*, IEEE, 2008, pp. 2812–2817.
- [24] O. Khatib, *Real-Time Obstacle Avoidance for Manipulators and Mobile Robots*. 1986.
- [25] Y. Koren and J. Borenstein, 'Potential field methods and their inherent limitations for mobile robot navigation', in *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference On*, IEEE, 1991, pp. 1398–1404.
- [26] D. Chakrabarti, R. Kumar, F. Radlinski *et al.*, 'Mortal multi-armed bandits', in *Advances in Neural Information Processing Systems*, 2009, pp. 273–280.
- [27] O. Besbes, Y. Gur and A. Zeevi, 'Stochastic multi-armed-bandit problem with non-stationary rewards', in *Advances in Neural Information Processing Systems*, 2014, pp. 199–207.
- [28] D. A. Berry, R. W. Chen, A. Zame *et al.*, 'Bandit problems with infinitely many arms', *The Annals of Statistics*, pp. 2103–2116, 1997.
- [29] N. I. Fisher, T. Lewis and B. J. Embleton, *Statistical analysis of spherical data*. Cambridge university press, 1993.
- [30] D. Hsu, J.-C. Latombe and R. Motwani, 'Path planning in expansive configuration spaces', in *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference On*, vol. 3, IEEE, 1997, pp. 2719–2726.
- [31] S. M. LaValle and J. J. Kuffner Jr, 'Randomized kinodynamic planning', *The international journal of robotics research*, vol. 20, no. 5, pp. 378–400, 2001.