Conditional Density Growth (CDG) Model: A Simplified Model of RRT Coverage for Kinematic Systems

Joel M. Esposito

Abstract—It has been shown before that the Rapidly Exploring Random Tree (RRT) algorithm is probabilistically complete; and that the probability of finding a particular solution path can be related to the number of nodes. However, little analysis exists on the rate at which the tree covers the configuration space. In this paper, we present a stochastic difference equation which models how the tree covers the configuration space as a function of the number of nodes in the tree. Using two simplifying assumptions, appropriate for holonomic, kinematic, systems in expansive configuration spaces, we derive a closed form solutions for the expected value and variance of configuration space coverage, which only depend on two easily computable parameters. Using a grid-based coverage measurement, we present experimental evidence supporting this model across a range of dimensions, obstacle densities and parameter choices. Collecting data from 1,000 RRTs, we provide evidence that configuration space coverage concentrates tightly around the expected coverage predicted by the model; and the results of a Chi-Squared test suggest that the distribution of coverage across these runs is highly Gaussian. Together these results enable one to predict the expected coverage, along with a confidence interval, after a certain number of nodes have been added to the tree. We also applied the model to an example with extremely narrow passages and to a system with non-holonomic kinematics. The expected value prediction is still qualitatively accurate; but the rate constant is reduced and the variance is higher. Overall, in addition to its theoretical value, the model may find future applications as an online measure of search-progress and problem difficulty, useful for adaptive variants of the basic RRT algorithm.

Index Terms—Path Planning for Manipulators, Motion Control, Motion Planning, Randomized Methods.

I. INTRODUCTION

Randomized algorithms, such as Probabilistic Road Maps (PRM) [2], [17], Rapidly Exploring Random Trees (RRTs) [24], [25] and their variants [4], [21], [3], are a powerful and practically important class of motion planning methods that have been used to address a variety of real world problems successfully. Their appeal lies in their ability to address large or complex problems in an incremental fashion [22]. In contrast, deterministic, computational geometry-type approaches to motion planning (for example [6]) suffer from well known complexity limitations because they attempt to explicitly construct and manipulate the free configuration space. However, the price of this incremental approach is a reduction in completeness. Most computational geometry-type algorithms are algorithmically complete – meaning that they are guaranteed to find a solution to a problem if one exists; or report none exists upon termination. They are also guaranteed to terminate in finite time. Randomized methods sacrifice algorithmic completeness for weaker Probabilistic Completeness [23] - meaning that if a solution exists, the probability that it is found approaches one as the number of iterations approaches infinity. While establishing an algorithm’s probabilistic completeness is of great theoretical importance, it tells us nothing about the rate at which the configuration space is covered. Little work has been done to enhance the theoretical foundation of RRTs, in this regard.

From a practical point of view, understanding the expected rate of coverage would be useful for planning problems in which it is not known a priori if a solution exists. Because, as long as no solution is found, the algorithm has no standard termination criteria. Therefore the roboticist must use experience and intuition to decide when to stop the search. It would be nice to be able to say, for example, that after 100,000 nodes have been added, we are 95% confident that 99.99% of the space has been explored.

This paper presents an idealized model of RRT growth, called the Conditional Density Growth (CDG) Model, primarily suited to holonomic systems operating in expansive configuration spaces. From this model we are able to derive various statistical properties of the RRT’s configuration space coverage such as the expected value, variance and distribution properties1. Section II provides a problem statement, reviews the RRT algorithm, and surveys related work. Section III presents the main analysis. First, coverage rate (a.k.a. growth) is defined here as the incremental change in the volume fraction of the configuration space that could be connected to the tree in the next iteration. Based on two simplifying assumptions, we develop a stochastic difference equation which models the coverage of holonomic systems operating in expansive spaces. Our main result is that the expected value of coverage is exponential in the number of nodes; and that the variance is a cubic function of coverage, peaking when two-thirds of the space has been explored. The model and assumptions are then validated against experimental computational data in Section IV. A grid-based method is used to approximately measure the configuration space coverage; and the model is fit with good agreement across a variety of dimensions and parameter values. Evidence suggests that, over many

1An earlier version of this work was presented at the 2012 IEEE/RSJ International Conference on Intelligent Robot and Systems, during the invited session 50 Years of Robot Motion Planning: Achievements and Emerging Approaches. The current version has been significantly expanded to include analysis of coverage variance and concentration, as well as more examples.
trials, configuration space coverage concentrates tightly around the expected value predicted by the model in a Gaussian-like fashion. This is significant because it suggests that the expected value and the variance give sufficient information to place a confidence interval on the expected coverage after a given number of nodes have been added to the tree. Section V examines the application of the model to nonholonomic systems and poorly expansive workspaces. Section VI discusses the significance of the model, as well as ways in which it could be extended. While this class of systems is admittedly limited, the CDG model introduced in this paper enhances our theoretical understanding of the algorithm and can be used as a base line for comparing algorithmic enhancements.

II. BACKGROUND

Problem 1. Given: the free configuration space \( X \); a compact time interval, \( T = [t^{\text{init}}, t^{\text{final}}] \); a compact set of possible initial states \( X^{\text{init}} \subset X \); a compact set of input values \( U \); dynamics \( \dot{x} = f(x,u) \), Lipschitz in \( x \in X \) and \( u \in U \); and a compact goal set \( S \subset X \). Find a function \( u : T \rightarrow U \) (if one exists) such that, when the state evolves from \( x \in X^{\text{init}} \), and is restricted to lie in \( X \), it produces a trajectory such that \( \exists^* \in T | x(t^*) \in S \).

Note that, in the first part of this paper we strictly focus on versions of this problem in which the dynamics \( f(x,u) \) are holonomic; and \( X \), while not necessarily simply connected, is expansive as defined in [12].

A Rapidly-exploring Random Tree (RRT) is an incremental randomized search algorithm that is designed for a broad class of motion planning problems [24]. It utilizes a tree-like data structure, \( T \), whose vertices (a.k.a. nodes) are points in the free configuration space \( x \in X \), and whose edges are constant inputs \( u \in U \) connecting two vertices in time step \( \Delta t \). Many variants of the algorithm have been introduced. While we consider kinematic systems in this paper, our ultimate goal is to extend the analysis to a richer class of problems. Among other things this often precludes bi-directional searches and requires the use of a local planner, numerical integration method and limited time step size. To that end, we have chosen the forward search version of the Rapidly Exploring Random Tree (RRT) method presented in [23] (Chapter 14, page 834), to implement and analyze. This extension of the original RRT is capable of solving problems with differential constraints.

Consider the outlines provided in Algorithm 1 and 2. Algorithm 1 simply calls Algorithm 2 Extend, repeatedly adding vertices and edges to the tree until a solution is found. As long as no solution has been found there is no accepted termination criteria. Algorithm 2 is the heart of the RRT method. It generates a sample, \( x^{\text{rand}} \), from a random variable \( X^{\text{rand}} \), with sample space \( X \) and density function \( p(\cdot) \). We assume the uniform distribution is used, since it is by far the most common choice in the literature. It then uses \( p(\cdot, \cdot) \), a distance function defined according to some suitable (pseudo) metric, to select \( x^{\text{near}} \), the closest existing vertex to the sample \( x^{\text{rand}} \). A local motion planner is then used to select a new input \( u^{\text{new}} \) which can steer the system as close as possible to \( x^{\text{rand}} \) in time \( \Delta t \), beginning from \( x^{\text{near}} \). A local planner is a quick, but possibly incomplete, algorithm that attempts to connect two configurations in close proximity to each other (small \( \Delta t \)). In the case of holonomic, kinematic, systems, the local planner may simply use line segments to connect \( x^{\text{rand}} \) and \( x^{\text{near}} \). For systems with more complex dynamics possible local planning techniques include steering controllers which exploit the structure of the dynamics, heuristic approaches such as potential fields, or discretization techniques which approximate the set of admissible input values \( U \) with a countable set of inputs \( \hat{U} \). Once an input is selected, any numerical integration technique that can simulate the system’s response to piecewise constant inputs can then be used, with \( u^{\text{new}} \) as an input and \( x^{\text{near}} \) as an initial condition, in order to generate \( x^{\text{new}} \). If the new state is not already a vertex in the tree, it will be added along with \( u^{\text{new}} \) as an edge. It is possible for the local planner to fail to generate a new tree node, for example if all possible values in \( \hat{U} \) have been explored from that particular \( x^{\text{near}} \), but we do not consider such resolution issues in this paper.

Algorithm 1 Generate RRT: \( T \) (Based on [23], Chapt. 14)

\[
\text{Initialize RRT: } T. \text{addVertex}(x^1 \in X^{\text{init}}) \\
\text{while } \exists x \in T \cap S \text{ do} \\
\quad \text{Extend}(T) \\
\text{end while}
\]

Algorithm 2 Extend(\( T \)) (Based on [23], Chapt. 14)

\[
x^{\text{rand}} \leftarrow X^{\text{rand}} \\
x^{\text{near}} \leftarrow \arg \min_{x \in T} \rho(x, x^{\text{rand}}) \\
u^{\text{new}} = \arg \min_{u \in U} \{ \rho(x^{\text{near}} + f(x, u)\Delta t, x^{\text{rand}}) \} \\
x^{\text{new}} = x^{\text{near}} + f(\cdot, u^{\text{new}})\Delta t \\
\text{if } x^{\text{new}} \notin T \text{ then} \\
\quad T.\text{addVertex}(x^{\text{new}}) \\
\quad T.\text{addEdge}(u^{\text{new}}, x^{\text{near}} \rightarrow x^{\text{new}}) \\
\text{end if}
\]

The advantage of the RRT algorithm is that it works directly with the set of admissible inputs and is therefore applicable to systems with complex dynamics. It is well suited to the problem of quickly searching high-dimensional spaces that have both algebraic and differential constraints. A second advantage is that it biases the exploration toward unexplored portions of the space by randomly sampling points in the state space and incrementally pulling the search tree toward them, reducing the size of largest Voronoi regions as the tree grows [25], [26]. The RRT algorithm was shown to be probabilistically complete in [24], with some discussion of the odds of discovering particular solution paths in [25]. In [7] the RRT algorithm, with some technical modifications, was shown to be resolution complete – meaning that accounting for discretization effects in time, input and the state space the algorithm is guaranteed to find goal sets of a certain minimum size. Recently the quality of the solutions emerging from the RRT have been shown to be asymptotically sub-optimal; and a new algorithm, RRT*, has been introduced which is asymptotically optimal [15].
From a theoretical point of view, the solution time (vice coverage) has been analyzed before – both for the RRT [25] and other randomized planning algorithms [16]. Typically in such “path finding” approaches, a particular path of interest is identified (ex. a path connecting a specific initial state to a specific goal state). Given some constants capturing the properties of this path, such as length and obstacle clearance within the configuration space, one determines the odds of discovering a homotopically equivalent path as a function of the number of sample nodes. A similar result, bounding the probability of failure was reported in [8], in terms of the number of milestones and lookout sets. The approach was generalized in [20] to include more abstract metric spaces. The approach is of great theoretical importance since it explains known challenges in finding paths through narrow passageways in the free-space. In [28] several interesting observations regarding the variance in coverage rate were reported. However, surprisingly little attention has been directed at developing predictive models for the rate at which the free space is covered by the tree. Of course this is partially because, in practice, the goal of such methods is not necessarily to explore the entire state space but rather to find a sequence of inputs from \( x^1 \) to \( S \) as quickly as possible. Indeed many adaptive methods and enhancements to the RRT (such as [5], [1] and [19]) improve performance without necessarily improving the coverage of the state space.

Despite its natural connection to pragmatic goal of finding a solution quickly, “path finding” analysis has several limitations. First, in realistic applications, one can rarely identify such paths of interest in advance. Second, assuming a path is identified, many of the constants used in establishing convergence bounds may be impractical to compute. Finally, the approach is extremely sensitive to the problem definition. In particular, a small perturbation in the location of the initial state or goal state can dramatically alter which areas of the state space are considered “interesting”.

We believe that a space covering approach to analysis overcomes some of these challenges. Space covering approaches attempt to determine the probability of finding all possible paths [13], [12]. Therefore, this type of analysis finds application in so-called verification or testing problems, such as those considered in [19], [18]. Because a “path of interest” does not need to be identified, limited a priori knowledge is required. More importantly, the approach is insensitive to perturbations in the initial and goal configurations. Analyzing coverage is also akin to considering worst case solution times.

III. Analysis

In this section, we define what we mean by “exploring” the state space, and present a stochastic model of this process.

A. Coverage Definition

We would like to determine what fraction of the reachable set \( a, T \), has explored. However, because a tree node \( x^k \) is a point and an edge is a line, the tree itself has measure zero and, therefore, can never fill the state space. Extending some concepts and notation from [19] and [23], define the set of states \( R \) that could be connected to the tree from a node \( x^k \) in a single iteration, using time step \( \Delta t \), as

\[
R(x^k, \Delta t) = \{ x \in X | \exists u \in U, x = x^k + \int_{0}^{\Delta t} f(x,u)dt \}.
\]

Note that this set has a non-zero measure, allowing us to discuss its volume which can be related to the fraction of the space that can be explored from that node. This notation has three principle uses, as illustrated in Figure 1.

1) The Connection Set \( C = R(x^{near}, \Delta t) \) is set of states that can be connected to \( x^{near} \) in a single iteration.
2) The Explored Set, is defined by \( E_K = R(T_K, \Delta t) = \bigcup_{k=1}^{K} R(x^k, \Delta t) \). It is the set of states that can be connected to, from any existing node in the tree, in a single iteration. It is used to define coverage. Note that if the intersection of the explored set and the goal set \( S \) is non-empty, the algorithm can terminate.
3) The Reachable Set, is defined as \( \mathcal{R} = R(x^1, T) \). It is is the set of all states that can possibly be explored from initial condition \( x^1 \) within the given time interval \( T = [t_{init}, t_{final}] \). It is a super set of the Explored Set.

Using this notation we define the Covered Volume Fraction, or coverage for short, of a tree with \( K \) nodes as

\[
C_K = \frac{\mu(E_K)}{\mu(\mathcal{R})}, \quad (III.1)
\]

the ratio of the volume of the explored set to that of the reachable set, where \( \mu(\cdot) \) is used to denote the volume (measure) of a set.

When a new node, \( x^K \) is added to the tree the change in the covered volume fraction is the growth

\[
G_K = \frac{\mu(E_K - E_{K-1})}{\mu(\mathcal{R})} = C_K - C_{K-1}. \quad (III.2)
\]
Remark 1. Note that obstacles in the configuration space are explicitly accounted for in the coverage measures in eqs. III.1 and III.2 because they reduce the volume of $R$ in the denominator.

B. Coverage Model: Holonomic Systems in Expansive Spaces

In the remainder of this paper we use the following notation.

If $X$ is a random variable with probability density function $p(x)$, $Pr[|x| < c]$ is the probability the absolute value of $x$ is less than some positive number $c$. $E(X)$ and $V(X)$ denote the expected value and variance of $X$. The notation $|$ is used to denote probabilities, expected value or variances conditioned on some other event.

Since Algorithm 2 is driven by a random variable $X^{\text{rand}}$, the coverage of a tree with $K$ nodes, $C_K \in [0, 1]$, is the state of a random, discrete time process

$$C_K = C_{K-1} + G_K,$$  

(III.3)

with the exception of $C_1$ which is determined solely by the initial condition. We choose to model the growth as a function of two random variables and the parameter $\Delta t$, $G_K(X^{\text{rand}}, C_{K-1}; \Delta t)$.

Assumption 1. Resolution factors such as the choice of integration method do not significantly influence the expected value of $G_K$. For a more complete discussion of resolution, the reader is referred to [7].

![Counterexamples. Assumption 2 (left), the tree can grow even if $x^{\text{rand}} \in \mathcal{E}$, especially if $x^{\text{near}}$ is on the boundary of the explored region.](image1)

![Counterexamples. Assumption 2 (right), it is possible to have very small growth if $x^{\text{rand}} \in \mathcal{X} - \mathcal{E}$, especially when the metric poorly reflects the “cost to go”.](image2)

Remark 2. Note that the assumption does not state that the growth is zero every time $x^{\text{rand}} \in \mathcal{E}$. In particular we expect it to be non-zero when $x^{\text{near}}$ supports the boundary of $\mathcal{E}$ (see Fig. 2 for a counter-example). Also, it implicitly accounts for the role of the current tree configuration and density function, because it is conditional on $x^{\text{rand}}$ and $\mathcal{E}$. It simplifies the analysis, by making the conditional expected value $E(G_K|x^{\text{rand}} \in \mathcal{E}_K)$ independent of $C_{K-1}$, $x^{\text{rand}}$, or $\Delta t$ (if one assumes $g = 0$). The experimental results tend to validate the accuracy of this assumption (see Sect. IV).

Assumption 3. If $x^{\text{rand}}$ is generated outside the explored set, $\mathcal{X} - \mathcal{E}$, we assume the conditional expected value of growth is larger and solely a function of the step size, $E(G_K|x^{\text{rand}} \in \mathcal{X} - \mathcal{E}_K) = \hat{G}(\Delta t) \gg g(\Delta t)$. See Figure 3 for an example.

Remark 3. While the growth at a given iteration is determined by a complex relationship between the location of $x^{\text{rand}}$, $x^{\text{near}}$, the configuration of the tree, the Lipschitz constant, and the time step $\Delta t$, the assumption allows one, on average, to replace them with a simplified function $\hat{G}$ when $x^{\text{rand}} \in \mathcal{X} - \mathcal{E}$. Obviously the assumption is not uniformly true, especially for systems that are not small time locally controllable (see Fig. 2 for counter-examples). In general, $\hat{G}$ is directly proportional to $\Delta t$ and the local Lipschitz constant of the dynamics $f(x, u)$ under the set of admissible inputs $U$. Increasing either has the effect of accelerating coverage.

It follows directly from eq. III.4 that

$$E(G_K|x^{\text{rand}}) \approx \begin{cases} g(\Delta t), & \text{if } x^{\text{rand}} \in \mathcal{E}, \\ \hat{G}(\Delta t), & \text{if } x^{\text{rand}} \in \mathcal{X} - \mathcal{E}. \end{cases}$$  

(III.5)

Therefore, our first main result is that the expected value of the growth at each iteration can be modeled as follows.

$$E(G_K|\mathcal{E}) \approx g \cdot Pr[x^{\text{rand}} \in \mathcal{E}] + \hat{G} \cdot (1 - Pr[x^{\text{rand}} \in \mathcal{E}]).$$

Assuming $x^{\text{rand}}$ is generated using a uniform distribution (the most common choice in the literature by far), $Pr[x^{\text{rand}} \in \mathcal{E}] = C_{K-1}$ and

$$E(G_K|C_{K-1}) \approx g \cdot C_{K-1} + (1 - C_{K-1}) \cdot \hat{G}. \quad (III.6)$$
Note that this formula predicts the tree grows quickly at first (small \( C_{K-1} \)) and more slowly as \( C_{K-1} \) approaches one, which supports Observation 3.

Our second main result concerns the variance of the growth. Using the fact that the total variance of two sub-populations can be computed as the variance of the sub-population means, plus the mean of the sub-population variances,

\[
V(G_K | C_{K-1}) \approx (\hat{G} - g)^2 C_{K-1} (1 - C_{K-1}) + \frac{1}{2} C_{K-1} V(G | x^{\text{rand}} \in \mathcal{E}) + (1 - C_{K-1}) V(G | x^{\text{rand}} \in \mathcal{X} - \mathcal{E}) \quad (\text{III.7})
\]

The first term is the variance of the conditional means. Our experiments suggest this term is dominant. The second and third terms are mean of the conditional variances. We neglect their effects in the sequel. Under this simplification, the variance vanishes when \( C_{K-1} = 0 \) or \( C_{K-1} = 1 \), and attains its maximum value of \( (\hat{G} - g)^2 / 4 \) at \( C_{K-1} = 0.5 \). It predicts the variance to be higher for systems with large values of \( \hat{G} \), either due to a large time step or Lipschitz constant. Another observation is that, the ratio \( V(G_K | C_{K-1}) / C_{K-1} \) monotonically decreases as new nodes are added to the tree.

Our third result provides an estimate of the expected value of coverage. Assume \( g = 0 \). Substituting eq. \text{III.6} into eq. \text{III.3}, the expected value of \( C_K \), is

\[
E(C_K) \approx E(C_{K-1}) + (1 - E(C_{K-1})) \cdot \hat{G}. \quad (\text{III.8})
\]

Whose solution is

\[
E(C_K) \approx C_1 (1 - \hat{G})^K + \hat{G} \sum_{k=0}^{K-1} (1 - \hat{G})^k. \quad (\text{III.9})
\]

Note that \( C_1 \) is completely determined by the initial condition and is not a random variable; however, further simplification is possible by assuming \( C_1 \approx \hat{G} \), allowing the two terms to be condensed into a single geometric series. Simplifying yields the third result

\[
E(C_K) \approx 1 - (1 - \hat{G})^K. \quad (\text{III.10})
\]

Observe that in accordance with previous results on probabilistic completeness [25], \( E(C_K) \to 1 \) as \( K \to \infty \). Obviously, increasing \( \hat{G} \) has the effect of hastening coverage. We remark that a similar analysis, relaxing the assumption that \( g = 0 \), yields a solution

\[
E(C_K) = \frac{\hat{G}}{\hat{G} - g} \left[ 1 - (1 - \hat{G} - g)^K \right], \quad (\text{III.11})
\]

which only converges to 1, if \( g \to 0 \) as \( C \to 1 \). While this is true in practice, we prefer the simplicity of eq. \text{III.10}.

Our fourth result predicts the expected value of the number of nodes, \( K \), required to achieve a particular coverage level, \( C \),

\[
E(K) \approx \frac{\ln \left( \frac{\hat{G}}{\hat{G} - g} - C \right)}{\ln \left( \frac{\hat{G}}{\hat{G} - g} \left( 1 - \hat{G} - g \right) \right)}, \quad (\text{III.12})
\]

Example: to achieve 95% coverage (\( C = 0.95 \)), with \( \hat{G} = 0.01 \) and \( g = 0 \), the expected number of nodes is \( K = 298 \).

Our last result concerns the variance of coverage. Using the definition of the covariance on can show the relationship between the growth during iterations \( i \) and \( j \) to be

\[
\text{Cov}(G_i, G_j) = -(\hat{G} - g)^2 C_{\min(i,j)} (1 - C_{\min(i,j)}).\]

Then, using the fact that \( C_K = \sum_{i=1}^{N} G_i \), and that the variance of a sum is equal to the sum of the entries in the covariance matrix,

\[
V(C_K) \approx \frac{(\hat{G} - g)^2}{2} C_K \cdot (1 - C_K), \quad (\text{III.13})
\]

suggesting the maximum variance is \( 2/27 \cdot (\hat{G} - g) \), which occurs at \( C_K = 2/3 \). These results are summarized in Table IV.

IV. Computational Experiments

A. Holonomic Systems in Expansive Spaces

The results in this section were obtained for an \( N \) dimensional, holonomic system with no obstacles. \( \mathcal{X} = [-0.5, 0.5]^{\mathbb{N}} \subset \mathbb{R}^N \) is the simply connected unit hypercube, the initial condition \( x^1 \) is the zero vector, \( U = \{ u \in \mathbb{R}^N \| \| u \|_2 \leq u_{\max} \} \), and \( J = u \).

Fig. 3. For holonomic systems with no obstacles, the connection sets are balls. The maximum growth \( G_{\max} \) is the volume of the gray area formed by \( R(x^{\text{new}}, \Delta t) - R(x^{\text{near}}, \Delta t) \), which is the hypersphere minus the vesica piscis.

B. Parameter Computation

For holonomic, kinematic systems, \( R(x^k, \Delta t) \) is simply a ball centered at \( x^k \) of radius \( u_{\max} \cdot \Delta t \). Regarding eq. \text{III.6}, in order to generate an easily computable approximation, we begin by assuming \( g = 0 \) and \( \hat{G} \) is equal to the upper bound of \( \hat{G}_K \), which is the fractional volume of the shaded piece of the ball shown in Figure 3. In arbitrary dimension \( N \) the fractional volume of the shaded set can be computed via numerical integration using

\[
G_{\max} = \frac{1}{\mu(K)} \left\{ \frac{\Delta t^{N/2}}{\tau_{N/2}} (u_{\max} \cdot \Delta t)^{N-1} \right\}^{\mu(K)} 
\]

where \( \Gamma \) is the generalization of the factorial function. The first term is \( \mu(B^N(u_{\max} \cdot \Delta t)) \)– the volume of a hypersphere with radius \( u_{\max} \cdot \Delta t \). The second term is twice the volume of the
Table I

<table>
<thead>
<tr>
<th>N</th>
<th>( \alpha_N )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.5</td>
<td>0.609</td>
<td>0.688</td>
<td>0.747</td>
<td>0.793</td>
<td>0.830</td>
</tr>
<tr>
<td>7</td>
<td>( \alpha_N )</td>
<td>0.859</td>
<td>0.883</td>
<td>0.902</td>
<td>0.918</td>
<td>0.985</td>
<td>0.997</td>
</tr>
</tbody>
</table>

Results of non-dimensionalized numerical integration used with Eq. IV.2 to compute \( G^\text{max} \) in dimensions 1 through 30.

hypersphere cap formed by an intersecting hyperplane whose distance from the sphere’s center is half the radius – sometimes referred to as the \( \text{vesica piscis} \). We compute this quantity by numerical integration. Readers interested in reproducing these results can quickly compute \( G^\text{max} \) using:

\[
\mu(\mathbb{B}^N(u^\text{max} \cdot \Delta t)) = \frac{\pi^{N/2}}{\Gamma(N/2 + 1)} (u^\text{max} \cdot \Delta t)^N\text{IV.2)}
\]

\[
G^\text{max} = \alpha_N \cdot \frac{\mu(\mathbb{B}^N(u^\text{max} \cdot \Delta t))}{\mu(\mathbb{R})}.
\]

Where, \( \alpha_N \) can be found in Table I, which summarizes the results of our unit-less numerical integration of eq. IV.1 in dimensions \( N = 1, \ldots, 30 \). Note as \( N \to \infty, \alpha_N \to 1 \).

**C. RRT Implementation**

A uniform distribution was used to select \( x^\text{rand} \). In all cases the tree was allowed to grow until \( C_K = 0.995 \). A 4\textsuperscript{th} order Runge-Kutta integration method was used, with variable step size and a tolerance of \( 10^{-8} \). The Euclidean metric was used, and a closed-form optimal planner generated \( u^\text{new} \).

There are no subsequent references to run-times or memory usage statistics, making the results machine-independent. All subsequent reported results are averaged over 1,000 trials unless otherwise stated.

**D. Discretized Coverage Measure**

While the notions of growth and coverage in Sect. III-A are useful for analysis, computing them exactly for a given tree is intractable, especially in the case of high dimensional systems with complex dynamics. Instead an approximation is used (Fig. 4, top). Given a free state space, define a uniform, axis-aligned grid with spacing \( \delta x \). Each grid cell is a closed set \( X_{i,j,...} \) indexed by an integer in each dimension. As each new node is added to the tree, a cell \( X_{i,j,...} \) is considered occupied if \( \exists x \in \mathbb{R}(T_{K_i} \cup T_{K_j}) \bigcap X_{i,j,...} \). The coverage \( \tilde{C} \) is approximated as the fraction of grid cells which are occupied. For example, Figure 4 (bottom) is a snapshot of the approximated \( \tilde{C} \) when \( C_{220} = 0.68 \). This is much less conservative, and easier to compute, than dispersion-based measures [10], [11].

**E. Experimental Results and Model Validity**

Our initial experiments used \( \Delta t \cdot u^\text{max} = 0.05, N = 3 \), and a grid size of \( \delta x = 0.005 \). Figures 5 and 6 compare the experimental mean and variance of the tree growth at each iteration, to the values predicted by eqs. III.6 and III.7. The value of \( \hat{G} \) was assumed to be \( G^\text{max} \), and computed via eq. IV.2 and Fig. 1, while \( g \) was assumed to be zero.

Regarding Assumption 2, \( E(G_K | x^\text{rand} \in \mathcal{E}) \approx 0 \), consider Fig. 7 (top), which shows a histogram of experimental \( G_K \)'s when \( x^\text{rand} \) was generated inside the explored set. Each bin in the histogram includes a range of volumes equal to 10 grid cells. More than 1/3 of the time \( G_K = 0 \), making it the mode by a large margin; and the mean is \( 0.52 \times 10^{-5} \) (variance \( 3.21 \times 10^{-9} \)). Since the growth can never be negative, it is unlikely the expected value would be zero in practice; however it is indeed very small considering that \( G^\text{max} \) is \( 3.56 \times 10^{-4} \).

Regarding Assumption 3, \( E(G_K | x^\text{rand} \in \mathcal{X} \setminus \mathcal{E}) \approx \hat{G} \), consider Fig. 7 (bottom) which shows a histogram of actual \( G_K \)'s when \( x^\text{rand} \) was generated outside the explored set. The experimental mode is \( 3.60 \times 10^{-4} \) and the mean
One of our original goals was to identify the underlying distributions seen in Fig. 7 for the purpose of refining the model. Motivated by the fact that both the selection of distributions seen in Fig. 7 for the purpose of refining the assumptions in Section III are reasonable.

For each of these candidates, we applied the Chi-squared test to determine the likelihood the data in Fig. 7 could have come from one of the distributions. To date none of the results have been statistically significant at the 0.05 level.

One observation worth mentioning is that, while the overall shape of the histograms in Fig. 7 is fairly typical, the best-fit distribution parameters vary significantly as coverage evolves. For example, the ratio of the number of nodes in the “peak” of the histogram of \( G_K(x_{\text{rand}} \in \mathcal{E}) \) (Fig. 7 -top) as compared to the “tail” is much lower for the first few thousand nodes (small \( C_K \)); while the peak is significantly higher for the final few thousand nodes (\( C_K \) close to 1). This suggests that \( g \to 0 \) as \( C_K \to 1 \), as required by eq. III.11. Therefore we do not expect that the distributions parameters are static, but rather a function of coverage.

Figures 8 and 10 show the experimental mean and variance of coverage as blue dots; while, the red dashed lines show the values predicted by eq. III.10 and eq. III.12, using \( G_{\text{max}} \).

An interesting way to visualize the agreement of the model with the experimental data is to note that according to eq. III.10, the function \(-\ln(1 - E(C_K))\) should vary linearly with \( K \) (solid line in Fig. 10). The dotted red line shows a linear fit to the log-transformed data, with \( r^2 = 0.989 \), indicating an excellent agreement. While intuitive to visualize this way, we caution that \(-\ln(1 - C_K)\) is ill-conditioned as \( C_K \to 1 \) since \( \ln 0 \) is undefined. Instead, we use the fit of the raw data to eq. III.10 as discussed in the next paragraph as our primary accuracy measure.

To quantitatively test the accuracy of the coverage model, we used a trust-region-based nonlinear least-squares algorithm to numerically curve fit eq. III.10 to the experimental data with a relative tolerance of \( 10^{-10} \), leaving \( G \) as a free parameter.
The green dashed line in Fig. 8 was created with a computed fit-coefficient of \( \hat{G} = (4.02 \pm 0.002) \times 10^{-4} \) (95% confidence). The “goodness of fit” was \( r^2 = 0.9991 \), indicating the form of eq. III.10 described the shape of the coverage curve extremely well. For comparison, according to Fig. 3, \( G_{\text{max}} = 3.56 \times 10^{-4} \).

Table II shows the result of similar analysis across examples with dimensions 2 - 8, holding \( \Delta t \cdot u_{\text{max}} = 0.2 \) constant. Note that \( N = 7, 8 \) use a coarser grid size due to memory limitations. First note that \( r^2 \approx 0.999 \) for all examples, indicating that the form of eq. III.10 holds well. Columns 4-6 show some dimensionless quantities that also speak to model fidelity. For a perfect model \( G_{\text{ext}} / \hat{G} \) would be 1 and \( G_{\text{int}} / \hat{G} \) would be \( \approx 0 \). An interesting trend is that \( \hat{G} / G_{\text{max}} \) (column 4) gets smaller, as \( N \) increases (\( N = 7, 8 \) are exceptions due to coarse grid size). One reason for this is that when \( x_{\text{rand}} \) is generated near the boundary of \( \mathcal{E} \), the growth can be smaller than expected (Fig. 2 - right). Note that in high

dimensional state spaces the volume of the boundary layer (thickness \( \Delta t \cdot u_{\text{max}} \)) becomes larger relative to the volume of the interior of \( \mathcal{E} \).

![Logarithmic curve fit. – \( \ln(1 - C_K) \) should be linear in \( K \). Note the log is ill-conditioned as \( C_K \to 1 \).](image)

**Table II**

<table>
<thead>
<tr>
<th>N</th>
<th>Nodes</th>
<th>( r^2 )</th>
<th>( G_{\text{max}} )</th>
<th>( G_{\text{ext}} )</th>
<th>( G_{\text{int}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3.10 \times 10^3</td>
<td>0.9991</td>
<td>1.2744</td>
<td>0.5859</td>
<td>0.1098</td>
</tr>
<tr>
<td>3</td>
<td>1.38 \times 10^2</td>
<td>0.9995</td>
<td>1.0743</td>
<td>0.6035</td>
<td>0.0968</td>
</tr>
<tr>
<td>4</td>
<td>6.67 \times 10^2</td>
<td>0.9997</td>
<td>0.8903</td>
<td>0.60732</td>
<td>0.1202</td>
</tr>
<tr>
<td>5</td>
<td>3.37 \times 10^3</td>
<td>0.9993</td>
<td>0.8932</td>
<td>0.5748</td>
<td>0.1257</td>
</tr>
<tr>
<td>6</td>
<td>1.81 \times 10^4</td>
<td>0.9983</td>
<td>0.7609</td>
<td>0.5622</td>
<td>0.1218</td>
</tr>
<tr>
<td>7</td>
<td>2.76 \times 10^5</td>
<td>0.9999</td>
<td>0.9831</td>
<td>0.9139</td>
<td>0.1232</td>
</tr>
<tr>
<td>8</td>
<td>1.28 \times 10^6</td>
<td>0.9995</td>
<td>0.8823</td>
<td>0.9948</td>
<td>0.0727</td>
</tr>
</tbody>
</table>

**Table III**

<table>
<thead>
<tr>
<th>( \Delta t \cdot u_{\text{max}} )</th>
<th>Nodes</th>
<th>( r^2 )</th>
<th>( G_{\text{max}} )</th>
<th>( G_{\text{ext}} )</th>
<th>( G_{\text{int}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.200</td>
<td>151</td>
<td>0.999</td>
<td>1.097</td>
<td>0.6067</td>
<td>0.0884</td>
</tr>
<tr>
<td>0.100</td>
<td>1258</td>
<td>0.998</td>
<td>1.166</td>
<td>0.6380</td>
<td>0.1062</td>
</tr>
<tr>
<td>0.050</td>
<td>9127</td>
<td>0.996</td>
<td>1.236</td>
<td>0.6343</td>
<td>0.1209</td>
</tr>
<tr>
<td>0.025</td>
<td>6918</td>
<td>0.994</td>
<td>1.266</td>
<td>0.6355</td>
<td>0.1283</td>
</tr>
</tbody>
</table>

Another issue we investigated was the validity of eq. III.10 as \( \Delta t \cdot u_{\text{max}} \) varies which directly alters \( \hat{G} \) and influences the rate at which the tree grows. Table III summarizes the results. First note that for all trials \( r^2 > 0.99 \) indicating good agreement with the model for a wide range of step sizes. Second, as this parameter decreases, \( G_{\text{ext}} = E(G_K | x_{\text{rand}} \notin \mathcal{E}) \) increases relative to \( \hat{G} \), indicating Assumption 3 becomes more accurate. However the ratio in the final column increases slightly, implying that assumption 2 (no growth when \( x_{\text{rand}} \in \mathcal{E} \)) becomes weaker. Again we feel this trend is related to the ratio of the boundary layer to the interior volume, which increases as \( \Delta t \cdot u_{\text{max}} \) becomes smaller. The surface area in Figure 4
is much larger when this parameter is decreased. For large values the individual branches begin to merge and the surface area is smaller.

**F. Concentration Properties**

While the mean and variance of coverage are important parameters, they are not sufficient to determine the shape of its probability distribution. One could use Chebyshev’s inequality to produce a loose bound on the deviation from the expected value after a certain number of nodes have been added,

\[ Pr\left(|C_K - E(C_K)| \leq m \cdot \sqrt{V(C_K)} \right) \geq 1 - \frac{1}{m^2}, \]  

where \( m > 0 \). However, considering that the coverage is a sum of many random variables, one would expect the distribution of coverage to be more tightly concentrated around its expected value; yet, most concentration inequalities are not applicable. For example, Hoeffding’s inequality only applies to sums of independent random variables; while Azuma’s only applies to martingales.

We grew 1,000 RRTs using \( \Delta t \cdot u^{\text{max}} = 0.05 \) in \( \mathcal{X} \subset \mathbb{R}^3 \), and a grid size of \( \delta x = 0.005 \). We observed that, once \( C_K > 0.03 \), the distribution of coverage looked quite Gaussian (Fig. 11), and that approximately 68%, 95% and 99% of the trees were within 1, 2, and 3 standard deviations, respectively, from the mean (Fig. 12) – suggesting much tighter concentration than that predicted by Chebyshev’s inequality. In light of this we applied the Chi-squared goodness of fit test at each iteration to determine the probability the 1,000 samples were drawn from a truncated Gaussian distribution (i.e. \( 0 \leq C_K \leq 1 \)). For 84% of the iterations the test was significant at the 0.05 level, suggesting the distribution of coverage is likely Gaussian. Most of the iterations that did not meet this criteria occurred when \( C_K < 0.03 \) or \( C_K > 0.99 \). We obtained similar results for problems with 2-6 dimensional state spaces dimensions; and across a variety of step sizes. These finding lead to the conjecture that inequality IV.3 can be tightened to the well know relation for Gaussian distributions

\[ Pr(|C_K - E(C_K)| \leq m \cdot \sqrt{V(C_K)}) = erf(m/\sqrt{2}). \]

In addition, consider that the ratio of the standard deviation to the mean decreases as coverage approaches 1. Blue solid line is experimental measurement; red dashed line is predicted value.

**V. ADDITIONAL EXAMPLES**

While the coverage model is designed for holonomic systems operating in obstacle free workspaces, few problems of interest fall into that class. Here we examine the model’s applicability to other classes of systems.

**A. Example 1: Expansive Spaces With Obstacles**

Figure 14 depicts a unit cube state space (N=3), whose middle section (width 0.2) is populated by randomly placed spherical obstacles of radius 0.025 (inspired by [12]). But the percent of the midsection covered by obstacles was varied between 0% and 50%. As the density increases, passing from the left half of \( \mathcal{X} \) to the right may require traveling through
passages as narrow as 0.025. However, because there are many such passages the workspace remains expansive.

Figure 14 shows experimental coverage averaged over 50 trials vs. the predicted coverage (red dashed line). We used \( \Delta t \cdot u_{\text{max}} = 0.05 \) and \( \delta x = 0.005 \). Note that to a certain extent the model is able account for the presence of obstacles, by reducing the denominator in eq. III.1. Because \( \Delta t \cdot u_{\text{max}} \) is held constant but the volume of the free space is reduced, \( \hat{G} \) is effectively increased when more obstacles are present – hastening coverage. Apparently the predictive power of the model is not significantly diminished compared with the 0% (no obstacles) example. Also note that the histograms (not shown) are virtually identical to the ones shown in Fig. 7 – implying Assumptions 2 and 3 are not adversely affected.

\[ \text{Fig. 14. (top) An expansive workspace filled with randomly generated spherical obstacles. (bottom) Actual coverage (blue line) vs predicted coverage (red dashed) for obstacle densities of 0%, 25% and 50%).} \]

B. Example 2: Poorly Expansive Spaces

Problems with a single long narrow passages are much more likely to weaken Assumption 3. Consider the problem in Figure 15 (top), where two large open regions within the unit cube are connected by a very long narrow tube of diameter \( \text{GapWidth} \) and length \( \text{GapLength} = 0.6 \) (also inspired by [12]).

Figure 15 (bottom) shows the coverage as we vary \( \text{GapWidth} \). First note that we saw no appreciable change in expected coverage when \( 0.16 \leq \text{GapWidth} \leq 1.0 \). When \( 0.04 \leq \text{GapWidth} \leq 0.16 \) the model fit is still very high, with \( r^2 \approx 0.998 \), however the growth rate, \( \hat{G} \), is effectively reduced by up to 15%. This is explained by examining the histogram of \( G_k|\text{rand} \in X - \mathcal{E} \) (Fig. 16), which is now bi-modal. It resembles a mixture of Figs. 7 (top) and (bottom), with an effectively reduced mean. Clearly the last term in eq. III.7 is no longer negligible, suggesting the variance is much higher than originally predicted.

When \( \text{GapWidth} \leq 0.04 \) the fit to the model is qualitatively poor (\( r^2 \leq 0.8 \)). Rather than look at the expected coverage, Fig. 20 shows coverage for 20 individual trials, for the extreme example of \( \text{GapWidth} = \delta x = 0.005 \). The growth actually takes place in two separate phases. The first phase corresponds to the tree filling the free space to the left of the passage, where the initial condition is located (see Fig. 18 - top). The number of nodes required to find the passage varies considerably among the trials – accounting for the variation in the length of the “tail” of the first phase. The second growth phase corresponds to the filling of the free space to the right of the passage. In the best case scenarios, when the passage is found quickly, the original coverage model is still remarkably accurate (upper red dashed line, Fig. 20). However in the worst case scenarios, an extended model is needed.

\[ \text{Fig. 15. (Top) A non-expansive workspace. (Bottom) Actual mean coverage (solid) for a variety of GapWidths vs predicted coverage (red dashed).} \]
Fig. 16. A histogram of $G_k(x^{rand}) \in X - E$ for a GapWidth = 0.005.

Fig. 17. Blue lines are coverage as a function of number of nodes for 20 individual RRTs, applied to the problem in Fig. 15, with GapWidth=0.005. Mean coverage shown in black. Red dashed line is expected value predicted by eq. III.10. Green dashed line is coverage predicted by sequential application of eq. III.10 in each expansive component of the workspace.

$X_L$ denote the expansive portion of the workspace to the left of the passage, where the tree is rooted; and $X_R$ denote the expansive portion to the right. Using eq. III.12, compute $K_L$ – the number of nodes required to achieve sufficient coverage of $X_L$ so that the expected number of nodes in the lookout set of $X_L$ is 1. Then, ignoring the volume of the narrow passage, $E(C_{K_L}) = \mu(X_L) \cdot \left(1 - \left(1 - \frac{G}{\mu(X_L)}\right)^K\right)$ if $K < K_L$,

$\mu(X_L) \cdot \left(1 - \left(1 - \frac{G}{\mu(X_L)}\right)^K\right) + \mu(X_R)$,

$\left(1 - \left(1 - \frac{G}{\mu(X_R)}\right)^{K-K_L}\right)$ if $K \geq K_L$.

This extension of the model is naive in many respects. For example, it ignores the shape and length of the passage; and it is not clear how to extend it to the case where there are more than two expansive components connected in a cyclic fashion. However, Fig. 20 shows that this model does qualitatively reproduce the twice exponential coverage trend in many of the trials. We chose not to develop this further because models that require in depth knowledge of solution parameters, such as passage length, diameter or connectivity, are antithetical to our goal of finding a simple coverage formula of practical significance. However, it is a topic of ongoing work.

C. Example 3: Non-Holonomic

Consider the classic unicycle model of a nonholonomic mobile robot where $x'_1 = u_1 \cos x_3$, $x'_2 = u_1 \sin x_3$, $x'_3 = u_2$ and $u_1 \in [-1, 1]$ and $u_2 \in [-\bar{\omega}, \bar{\omega}]$. For this example we used the (imperfect) metric

$$\rho(x^A, x^B) = \sqrt{(x^A_1 - x^B_1)^2 + (x^A_2 - x^B_2)^2 + \frac{1}{\pi} \min(|x^A_3 - x^B_3|, 2\pi - |x^A_3 - x^B_3|)},$$

and use an arc search algorithm [27] as the local planner to connect $x^{near}$ and $x^{rand}$.

We applied our coverage model to the nonholonomic system and compared it to a holonomic system whose connection sets have identical volume. On a positive note, we found that the coverage model remains qualitatively valid for a wide range of $\omega$ values. Consider Fig. 18- bottom, leaving $G$ as a free parameter, the fit of eq. III.10 to the non-holonomic coverage data (red lines) is quite good ($r^2 > 0.99$).

However, the growth rate is reduced from the equivalent holonomic system (green lines) by about 25%. Note that, the connection set is bowtie-shaped as shown in Fig. 18-top (vs circular in Fig. 3). A theoretical result from the study of stochastic coverage processes [9] suggests that the shape and orientation of the connection sets should not impact the expected value of coverage; however our initial results indicate this is not true for the RRT. An investigation of the histogram of growth reveals a situation similar to the moderately narrow passage examples in Fig. 16 – a bimodal histogram.

VI. CONCLUSIONS AND FUTURE WORK

In this paper we provided a stochastic difference equation which models the fractional volume of the configuration space covered by the forward search RRT algorithm, as a function of the number of nodes in the tree. We model the growth at each iteration as a random variable drawn from one of two distinct distributions – termed the Conditional Density Growth (CDG) Model. Under two simplifying assumptions, the model permits simple closed form approximations of the expected value and variance of the coverage. The expected value of coverage follows an exponential convergence pattern, while the variance peaks when 2/3 of the state space has been explored then decays to zero. The model only depends on a single rate constant; and we provided a formula and table to compute it for systems with single integrator dynamics, using the time step and a Lipschitz constant – two easily obtainable parameters. The model recovers many of the properties observed by users of the RRT such as probabilistic completeness, and rapid initial exploration followed by a slower exploration phase. Our results are summarized in Table IV.

When applied to holonomic systems operating in expansive configuration spaces, the computational experiments and model exhibit strong agreement across a variety of dimensions and parameter choices. The presence of obstacles alone does not appear to invalidate the model. Furthermore, we presented experimental evidence that the probability distribution of coverage is highly concentrated around its expected value, for
First, it addresses some previously open foundational questions. Second, we feel that one of the strengths of the model lies in its computational and informational simplicity. Estimating the coverage only requires a handful of floating point operations. In addition, all the parameters of the model, when applicable, are easily computable, and do not require a priori knowledge of the solution. Third, from a practitioner’s perspective, the formula can be used to provide estimates of the number of nodes required to fill a specified percent of the space (ex. 1,500 nodes for 99.9% coverage with 95% confidence). To our knowledge this is the first RRT termination criteria developed using a formal approach. Fourth, many adaptive variants of the RRT algorithm have been introduced in the literature (e.g. [1], [19], [5], or [21]), for the purpose of addressing difficult problems that may contain narrow passages. Our simple formula could be used as a baseline metric of just how difficult a problem is, by examining the lag in predicted vs. actual coverage. For example, if after 1,500 nodes the model predicts 99.9% coverage, but only 50% of the space has been covered, it is highly unlikely the configuration space is expansive.

Possible directions of future work include formalizing the extended model of coverage (eq. V-B) for systems with narrow passages – in particular to model the variance; estimating the nature of the distributions seen in Fig. 7, for the purpose of developing a mixture model for the bi-modal distribution in Fig. 16; and developing an analogous model for Probabilistic Road Map planners.

### References


From Assumptions 2 and 3

Assuming conditional variances are negligible

Assuming conditional variances are negligible

Assume $g = 0$, if $C \approx 1$.

Chi-squared test suggests coverage is Gaussian

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Function</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(G_K)$</td>
<td>$g \cdot C_K + G(1 - C_K)$</td>
<td>From Assumptions 2 and 3</td>
</tr>
<tr>
<td>$V(G_K)$</td>
<td>$(G - g)C_K \cdot (1 - C_K)$</td>
<td>Assuming conditional variances are negligible</td>
</tr>
<tr>
<td>$E(C_K)$</td>
<td>$\frac{G}{G - g} \cdot (1 - (G - g)K)$</td>
<td>where $g \to 0$ as $C \to 1$, simplify by assuming $g \equiv 0$</td>
</tr>
<tr>
<td>$V(C_K)$</td>
<td>$\left( \frac{G}{G - g} \right)^2 C_K \cdot (1 - C_K)$</td>
<td>Assuming conditional variances are negligible</td>
</tr>
<tr>
<td>$E(K)$</td>
<td>$\ln \left( \frac{\hat{G} - g}{G - g} \right)$</td>
<td>Assume $g = 0$, if $C \approx 1$.</td>
</tr>
</tbody>
</table>

Pr $|C_K - E(C_K)| \leq m \cdot \sqrt{V(C_K)}$ erfc $\left( \frac{m}{\sqrt{2}} \right)$

TABLE IV
SUMMARY OF RESULTS.


Joel M. Esposito earned the BS degree in Mechanical Engineering from Rutgers University, New Brunswick, NJ, USA and the MS and PhD, in 1999 and 2002 respectively, from the University of Pennsylvania, Philadelphia, PA, USA. He is currently an Associate Professor in the Systems Engineering Department at the United States Naval Academy, Annapolis, MD, USA; and serves as the Editor in Chief of the American Society of Engineering Education’s Computers in Education Journal. His research interests include multi-robot coordination, motion planning, computational tools for motion control and robotics pedagogy. Prof. Esposito is the recipient of the Naval Academy’s 2010 R. A. Raouff Teaching Award.