Control of Stochastic Unicycle-type Robots

Shridhar K. Shah and Herbert G. Tanner

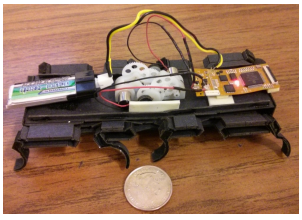
***Abstract*— This paper addresses the problem of optimal control of a unicycle-type robot perturbed with stochastic noise in an environment with sparsely populated obstacles. The objective is that the robot pose converges to a neighborhood of a desired position and orientation. A feedback control law is constructed such that it is compatible with the differential constraints of the unicycle. The construction is based on numerical solution of the Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE) associated with a stochastic optimal control problem. The control law is optimal in terms of control effort and comes with probabilistic guarantees of convergence to the goal set.**

Keywords — wheeled mobile robots, stochastic optimal control, unicycle, obstacle avoidance

I. Introduction

The approach in this paper is motivated by recent advancements in manufacturing of miniature bio-inspired crawling robots (Fig. 1(a)). For systems of the scale of the robot Fig. 1(a), one it is controllers that require minimal computation and actuation effort are desirable because of the severe limitations in payload and available energy storage capacity.

The motion behavior of these robots is inherently stochastic due to the pronounced scale-dependent effects of manufacturing process variations, compliance, ground interaction, and battery charge fluctuations [1], [2]. Under certain conditions, the kinematics of multi-legged robots with differential drive mechanisms can be reasonably approximated by a unicycle model [3], [4]. Driven by these realizations, the paper develops input-optimal controllers for stochastic unicycles, which can extend to cases of motion in constrained environments.



(a) OctoRoACH platform [1]

(b) A wheeled surrogate platform

Fig. 1. Motion of legged platforms at the miniature scale can be effectively controlled using stochastic versions of well-studied macroscopic wheeled vehicle models

Control of unicycles in a *deterministic* setting has been extensively studied (see [5]–[7] and the references within for a sparse sample of available literature) Optimal, with respect to control effort, nonholonomic control designs exist [8], [9], some incorporating obstacle avoidance constraints [10]. It is unclear to what extent available solutions are robust with respect to stochastic perturbations and noise. Although feedback can be extremely effective at rejecting perturbations (c.f. [5]) this typically comes at the expense of control effort. In addition, most existing nonholonomic control solutions for obstacle-free environments do not naturally carry over to cluttered environments, mainly due to the use of non-invertible state transformations [5], [11]. Finally, when stochastic noise perturbs deterministic closed loop systems, the probability that the latter can still reach a neighborhood of their goal configuration without violating obstacle avoidance constraints has only be assessed *a posteriori* [12]; stochastic noise being theoretically unbounded prevents one from establishing deterministic performance bounds. This paper fills the gap in the middle with a control-effort optimal stochastic control design, which can be also adjusted to probabilistically guarantee obstacle avoidance.

Nonholonomic convergence to goal and avoiding obstacles under stochastic perturbations has been treated as a stochastic reach-avoid control problem [13], but without being concerned about the final orientation. Final orientation is regulated in stochastic unicycle systems [14] using switched feedback controllers, but here the environment is obstacle-free. There is some somewhat related work in the context of formation control for stochastic unicycles [15], where obstacle avoidance is approached deterministically and the control objective is to achieve a distance-based formation. For stochastic Dubins-type vehicles, there are minimum-expected-time feedback controllers [16] computed using the HJB PDE, but neither obstacle avoidance, nor orientation convergence is addressed. From an algorithmic viewpoint, there exist adaptations of rapidly exploring random trees [17] that explicitly take into account model uncertainty. However, the latter approaches are fundamentally open-loop. There is recent work, evolving parallel to the one reported here, which involves the solution of a PDE for the purpose of constructing navigation functions for robots exhibiting stochastic nonlinear dynamics [18]. However, it is unclear if those PDEs, which use more general incremental cost functions, admit the probabilistic approximations to their solutions as the ones in this paper do.

For more general stochastic systems with nonlinear dynamics, the stochastic optimal control problem has a longer history [19]–[22]. Path integrals have been used for control design in applications ranging from reinforcement learning [22], to variable impedance control [23], risk sensitive control [24], and receding horizon-based robot motion planning [25], [26]. The theoretical foundation [20] of the latter work

This work was supported by ARL MAST CTA # W911NF-08-2-0004

Shridhar Shah is with the MathWorks Inc. [shridhar.shah@mathworks.com]. Herbert Tanner is with the Department of Mechanical Engineering, University of Delaware [btanner@udel.edu].

allows an exit-time formulation of the robot convergence problem. Whereas stochastic optimal control methods of this nature have been developed for holonomic systems moving in obstacle-free *local* neighborhoods [25], [26], this paper particularizes the analysis to stochastic unicycle systems, allows for the possibility of (small) obstacles inside neighborhoods, and targets the regulation of both position and orientation. This is done through off-line, Feynman-Kac based approximations of solutions to HJB PDE (c.f. [20], [22], [27]). The solution of this PDE yields a potential field that steers the stochastic system away from obstacles and toward the target set.

The paper is organized in the following way. Section II states the problem. It is followed by Section III, which gives technical details of the adopted stochastic optimal control formulation with exit constraints. Section IV presents simulations and controller analysis, as well as experimental results, for an application in an environment without obstacles. Section V provides simulations for environment with obstacles, Section VI discusses implementation issues, and Section VII concludes.

II. Problem Statement

Let $\mathcal{D} \subseteq \mathbb{R}^3$ be a bounded domain, and let the obstacles within $\mathcal{D}$ be represented in the form of a closed set $\mathcal{O}$. The free robot workspace is thus $\mathcal{P} \triangleq \mathcal{D} \setminus \mathcal{O}$. Assume that the boundary of $\mathcal{P}$, denoted $\partial\mathcal{P}$, is expressed by a twice differentiable function (i.e., $\mathcal{C}^2$). Figure 2 shows the coordinate system assigned on $\mathcal{D}$. Consider now a unicycle perturbed with stochastic noise, with $\mathrm{q} = (x, y, \theta)$ denoting its state (x-y position and orientation) and v, ω its linear and angular speed (inputs). Noise here affects the inputs. Let $W = \{W(t), \mathcal{F}_t : 0 \leq t < \infty\}$ model a 2-dimensional Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, $\mathcal{F}$ is a σ-algebra on Ω, $\mathbb{P}$ is a probability measure, and $\{\mathcal{F}_t : t \geq 0\}$ is a filtration (i.e., an increasing family of sub-σ-algebras of $\mathcal{F}$), assumed right-continuous and such that $\mathcal{F}_0$ contains all $\mathbb{P}$-null (of measure zero) sets [27]. The diffusion terms, introduced by W, are assumed to have state-dependent diffusion coefficients $\Sigma_1(\mathrm{q})$, $\Sigma_2(\mathrm{q})$, respectively, giving rise to a two-dimensional diffusion matrix $\Sigma(\mathrm{q}) = \mathrm{diag}\{\Sigma_1(\mathrm{q}), \Sigma_2(\mathrm{q})\}$. With these in place, we express the dynamics of the system as

$$\underbrace{\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}\theta \end{bmatrix}}_{\mathrm{dq}} = \underbrace{\begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix}}_{g(\mathrm{q})} \Big(\begin{bmatrix} v \\ \omega \end{bmatrix} \mathrm{d}t + \Sigma(\mathrm{q})\, \mathrm{d}W \Big) \tag{1}$$

An input-optimal feedback control law is sought, such that the system converges to a goal set $\mathcal{E} \triangleq \{(x, y, \theta) \in \mathbb{R}^2 \times \mathbb{S} \mid \|(x, y, \theta)\| \leq \varepsilon\}$ where ε is a small constant and[1] $\|(x, y, \theta)\| \triangleq \sqrt{x^2 + y^2 + (\cos\theta - 1)^2 + \sin^2\theta}$.

[1]Formally, $\mathrm{q} = [x, y, \theta]^\intercal$ belongs in the two-dimensional special Euclidean group SE(2); it can, however, be embedded in $\mathbb{R}^4$ [28], where the usual metrics can be used. Here, the metric $\|[x\, y\, \theta]^\intercal\| = \sqrt{x^2 + y^2 + (\cos\theta - 1)^2 + (\sin\theta)^2}$ is used.

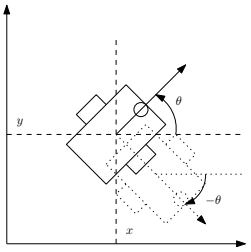


Fig. 2. Description of the coordinates for system (1)

Assume that the goal set $\mathcal{E}$ is within the free workspace $\mathcal{P}$, and that the domain $\mathcal{P} \setminus \mathcal{E} \subset \mathbb{R}^3$ is bounded with a $\mathcal{C}^2$ boundary $\partial(\mathcal{P} \setminus \mathcal{E})$. Its closure is denoted $\overline{(\mathcal{P} \setminus \mathcal{E})}$.

III. Stochastic Optimal Control

Consider a class of stochastic systems evolving on the open set $\mathcal{P} \subset \mathbb{R}^n$,

$$\mathrm{dq} = b(\mathrm{q})\, \mathrm{d}t + G(\mathrm{q})[u(\mathrm{q})\, \mathrm{d}t + \Sigma(\mathrm{q})\, \mathrm{d}W] \ . \tag{2}$$

Assume that $b(\mathrm{q})$, $G(\mathrm{q})$, $\Sigma(\mathrm{q})$, and $\Sigma^{-1}(\mathrm{q})$ are bounded and Lipschitz continuous on $\mathcal{P} \setminus \mathcal{E}$—standard assumptions [27]. Taking τ to be the first time $\mathrm{q}(t)$ hits the boundary of $\mathcal{P} \setminus \mathcal{E}$, $a(\mathrm{q}) \triangleq \Sigma(\mathrm{q})\Sigma^\intercal(\mathrm{q})$, and choosing a terminal cost Φ

$$\Phi(\mathrm{q}(\tau)) = \begin{cases} 0 & \text{on } \partial\mathcal{E} \\ \infty & \text{on } \partial\mathcal{P} \end{cases} ,$$

we seek a feedback control law that minimizes the following cost functional

$$V(\mathrm{q}) = \min_{u(\mathrm{q})} \mathbb{E}\Big[\int_0^\tau \frac{1}{2} u^\intercal\big(\mathrm{q}(s)\big)\, a^{-1}\big(\mathrm{q}(s)\big)\, u\big(\mathrm{q}(s)\big)\, \mathrm{d}s + \Phi\big(\mathrm{q}(\tau)\big) \Big]. \tag{3}$$

Denoting ∂_q, ∂_{qq} the gradient row vector $\frac{\partial}{\partial \mathrm{q}}$ and the Jacobian matrix $\frac{\partial^2}{\partial \mathrm{q}^2}$, respectively, we consider the generator $\mathcal{A}$ of (2) (see [29])

$$\mathcal{A}V = \partial_\mathrm{q} V \cdot (\mathrm{b}(t) + G(\mathrm{q})\mathrm{u}(\mathrm{q})) + \frac{1}{2}\mathrm{tr}\{\partial_{\mathrm{qq}} V \cdot \mathrm{G}(\mathrm{q})\Sigma(\mathrm{q})\Sigma^\intercal(\mathrm{q})\mathrm{G}^\intercal(\mathrm{q})\}\,,$$

and express the HJB equation associated with (3) as

$$\min_{u(\mathrm{q})} \Big\{ \mathcal{A}V(\mathrm{q}) + \frac{1}{2} u^\intercal(\mathrm{q})\, a^{-1}(\mathrm{q})\, u(\mathrm{q}) \Big\} = 0\,. \tag{4}$$

The optimal control law u^* that is the solution of (4) is analytically expressed as (see [20])

$$u^* = -a\, G^\intercal\, \partial_\mathrm{q} V^\intercal \ . \tag{5}$$

(Dependence of terms on q dropped for brevity.) Substituting (5) in (4) and applying the logarithmic transformation $V(\mathrm{q}) = -\log g(\mathrm{q})$ [19], yields

$$\partial_\mathrm{q} g\, \mathrm{b} + \frac{1}{2}\mathrm{tr}\big\{ \partial_{\mathrm{qq}} g\, G\, \Sigma\, \Sigma^\intercal\, G^\intercal \big\} = 0 \tag{6}$$

with boundary conditions

$$g\big(\mathrm{q}(\tau)\big) = \begin{cases} 1 & \text{on} \quad \partial\mathcal{E} \\ 0 & \text{on} \quad \partial\mathcal{P} \end{cases} \quad .$$

Using the Feynman-Kac formula [27], the solution of (6) is

$$g(\mathrm{q}) = \mathbb{P}\left[\zeta(\tau) \in \partial\mathcal{E} \mid \zeta(0) = \mathrm{q}\right] \tag{7}$$

where $\zeta(t)$ is the Markov process

$$\mathrm{d}\zeta(t) = b(\zeta)\ \mathrm{d}t + G(\mathrm{q})\,\Sigma(\zeta)\ \mathrm{d}W \tag{8}$$

evolving on $\mathcal{P} \setminus \mathcal{E}$. The optimal control (5) can thus be expressed as

$$u^* = -a\ G^{\intercal}\left(\partial_{\mathrm{q}} \log \mathbb{P}\big[\zeta(\tau) \in \partial\mathcal{E} \mid \zeta(0) = \mathrm{q}\big]\right)^{\intercal} . \tag{9}$$

With the infinite penalty on $\partial\mathcal{P} \setminus \mathcal{E}$, the optimal control ensures that the system exits from $\partial\mathcal{E}$ with probability one at time τ, which is finite if for $1 \le l \le m$ [27, Lemma 7.4]

$$\min_{\mathrm{q}\in\mathcal{P}\setminus\mathcal{E}} a_{ll}(\mathrm{q}) > 0 \implies \mathbb{E}[\tau \mid \mathrm{q}(0) \in \mathcal{P} \setminus \mathcal{E}] < \infty \ . \tag{10}$$

If Σ and Σ^{-1} are bounded, then (10) is satisfied. For the first exit time τ, the constraint that $\mathrm{q}(\tau) \in \mathcal{E}$ is equivalent to [20]

$$\mathbb{P}\big[\mathrm{q}(\tau) \in \partial\mathcal{P} \mid \mathrm{q}(0) = \mathrm{q}\big] = 0 \ .$$

If this constraint is satisfied, then the robot avoids obstacles and converges to the goal set with probability one. The caveat is that infinitely large inputs may be required arbitrarily close to the obstacle surface, and realistically, control inputs are always bounded. As a result, depending on the magnitude of noise and control input bounds, the actual probability of success of convergence to goal can be smaller than one. This issue is discussed in Section IV-B.

IV. Bounded Inputs Without Obstacles

This section presents simulation and experimental results that investigate the behavior of optimal feedback control laws described in Section III, for the case where the control input is bounded. For computational purposes we assume that $\mathcal{D}$ is free of internal obstacles. The workspace $\mathcal{P}$ is circular with radius 1.0 $\mathtt{m}$, $\mathcal{P} = \{(x, y, \theta) \in \mathbb{R}^2 \times \mathbb{S} \mid \sqrt{(x^2 + y^2} < 1.0,\ \forall\theta\}$. The goal set is given as $\mathcal{E} = \{(x, y, \theta) \in \mathbb{R}^2 \times \mathbb{S} \mid \|(x, y, \theta)\| \le 0.1\}$, and the system evolves within the domain $\mathcal{P} \setminus \mathcal{E}$. The unicycle is assumed to be a point.

Let us particularize the general dynamics (2) to the case of unicycle kinematics (1), where, in addition, $b\big(\mathrm{q}(t)\big) = 0$. Compared to the general case (2), system (1) does not have a drift term. For the numerical implementation we choose $\Sigma_1 = 0.1$ and $\Sigma_2 = 0.5$ to be constant variances. We constrain the linear and angular velocity in the intervals $v \in [-0.3, 0.3]$ $\mathtt{m/s}$ and $\omega \in [-2, 2]$ $\mathtt{rad/s}$, respectively.

The optimal control input (9) can be obtained by simulating (8)—in most practical cases the solution of the PDE (6) cannot be given analytically. System (8) is written as

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1 \\ \mathrm{d}W_2 \end{bmatrix} \ ,$$

which is a stochastic differential equation that can be simulated using the Euler-Maruyama method [30]. The Brownian noise term is generated as a Gaussian noise with zero mean and variance equal to $\mathrm{d}t$ [30]. With the function $g(\mathrm{q})$ estimated of line, as the empirical probability of sample paths of (8) first hitting the goal boundary, the control law (9), now takes the form

$$u^*(\mathrm{q}) = -\Sigma\Sigma^T \cdot \partial_{\mathrm{q}}\big(-\log g(\mathrm{q})\big) \ .$$

A. Simulation results

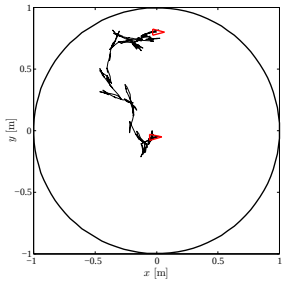
The state space $\mathcal{P}\setminus\mathcal{E}$ is discretized into a $41\times41\times41$ grid, and each grid point is treated as a potential initial condition. Five hundred sample paths are simulated for *each* initial condition. The discretization time is chosen to be $\mathrm{d}t = 0.05$ seconds, and the total simulation time is kept large enough such that at least 95% of the sample paths, from each initial condition, exit either at the goal boundary $\partial\mathcal{E}$ or at the obstacle boundary $\partial\mathcal{P} \setminus \partial\mathcal{E}$. A visual representation of the function $g(\mathrm{q})$ computed for unicycle kinematics of system (11) is shown in Figure 4(b).

There are three factors that affect computation speed when estimating $g(\mathrm{q})$: grid size, number of samples per cell, and whether a drift term is included in the stochastic simulations at each cell. The following computation times are reported for the purpose of assessing the effect of changing these problem parameters, and are conservative since no effort was made to optimize the algorithm itself for speed. That algorithm, running on a common laptop with a four-core i5-3210M CPU clocked at 2.5 GHz and 3.2 GB of RAM, computing $g(\mathrm{q})$ for a configuration of a 21×21 grid with 50 samples out of each cell and a nudging drift term of the form $(0.2\cos\theta\ \ 0.2\sin\theta\ \ 0)^{\intercal}$ required 1 minute and 5 seconds. The same grid and sample configuration without the drift term required 9 minutes and 45 seconds. Increasing the resolution with a 41×41 grid, and taking 10 times more samples out of each initial condition results in computation times of 85 minutes (with drift) and 794 minutes (without the drift term).

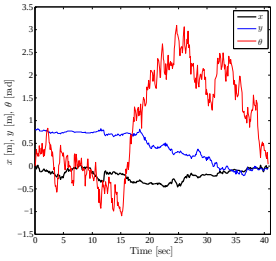
Applying optimal control u^* of (9) results in sample paths like the one shown in Figure 3(a). The trajectories of individual state components are depicted in Figure 3(b) and the optimal control inputs applied to the system are contained in Figure 3(c). With this implementation, control inputs can attain large values near the obstacle boundary and are subsequently saturated so that v and ω remain within their designated intervals. As a result, some configurations near the outer boundary have no feasible control solution. In practice, restricting control authority impacts the probability of reaching the goal set. Section IV-B elaborates.

B. Probability of success

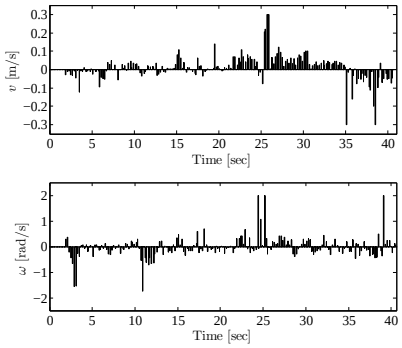
When input constraints are imposed on a system, the probability of convergence to the goal set is no longer one. The empirical probability of success is defined as the ratio of sample paths that exit at the goal set boundary, to the total number of sample paths which exit either at the goal or at the outer (or obstacle) boundary. To assess the impact of input



(a) A sample path of (1)–(9)



(b) State trajectory for the sample path of Figure 3(a)



(c) Control inputs given by controller (9)

Fig. 3. Simulation of (9) applied to (1). A sample path is shown in part (a). The outer boundary is shown as a circle and the initial and final orientation of unicycle are shown as triangles, with the goal being at the origin. Individual state trajectories are shown in part (b), indicating convergence to a neighborhood of origin. The control inputs applied during this simulation are shown in part (c).

saturation on the probability of success, we simulate 500 sample paths from initial position $(x, y, \theta) = (0.0, 0.8, 0.0)$ (cf. Figure 4(a)). We find that for $v \in [-0.3, 0.3]$, and without constraints on angular velocity, the empirical probability of success is still 1.0. But when an additional constraint of the form $\omega \in [-2, 2]$ is imposed on the angular velocity, the estimated probability of success drops to 0.66.

C. Comparison to a deterministic controller

Control law (9) lets the system 'sail' with noise for as long as it can, and exercises corrective action at 'full throttle' only when necessary to avoid collisions. Instead, a deterministic feedback controller would constantly use control effort to steer the system toward its goal region. In some cases, (e.g. [5]) the combination of feedback with the absence of singular state transformations yields closed loop systems that are extremely robust to perturbations and noise. The difference in philosophy on control effort application between (9) and deterministic alternatives (specifically [5]), is illustrated in Table I. In three simulation trials, both controllers succeed in driving the simulated robot to its goal, but (9) uses consistently significantly less power.

D. Experimental testing

Intuitively, controller (9) works by applying minimal effort in general, and use full force to avoid the obstacle boundary.

TABLE I

COMPARISON OF CONTROL EFFORTS BETWEEN THE STOCHASTIC OPTIMAL CONTROLLER AND THE FEEDBACK CONTROLLER OF [5].

	Stochastic optimal controller		Feedback controller of [5]	
	$\sum v$	$\sum \omega$	$\sum v$	$\sum \omega$
Trial 1	8.4561	49.0434	21.0152	138.6997
Trial 2	8.6472	27.6933	31.8662	406.6131
Trial 3	9.0893	51.5052	23.6704	174.2929

As a result, convergence time can be long, but input effort is small. One way to reduce convergence-time is to push the system gently toward the goal by means of a small controlled drift term (which is missing from (1)). Hence, consider (1) with a constant drift velocity V_n, compactly written as

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} V_n + v \\ \omega \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1 \\ \mathrm{d}W_2 \end{bmatrix} \right) \quad (11)$$

where $\Sigma_1 = 0.1$ and $\Sigma_2 = 0.5$ are constant variances as before, and V_n is taken[2] to be 0.2 $\mathtt{m/sec}$. The workspace $\mathcal{P}$ is obstacle free, and of radius 1.0 $\mathtt{m}$.

A three-wheeled, differential-steering robot (Figure 1(b)) is used as an experimental platform, with kinematics described by those of a unicycle. The control law is computed on a dual core 2.0 GHz processor, and is sent over a wireless network to the robot. Although this robot exhibits uncertain behavior on its own due to wheel slippage, ground friction, random network delays and inaccurate velocity command realization, we additionally contaminate the control inputs by Gaussian noise. Full state (x, y and θ) feedback data is obtained using a motion capture system, which also introduces some noise of small magnitude. Linear velocity inputs are saturated at ± 0.3 $\mathtt{m/s}$ and angular velocity commands at ± 1.0 $\mathtt{rad/s}$. Hardware limitations do not allow the platform to realize linear velocities of magnitude smaller than 0.1 $\mathtt{m/s}$. Thus control inputs are constrained in the intervals $v \in [-0.3, -0.1] \cup \{0\} \cup [0.1, 0.3]$ $\mathtt{m/s}$, and $\omega \in [-1, 1]$ $\mathtt{rad/s}$.

Velocity commands are sent to the robot at a frequency of 5 $\mathtt{Hz}$, in the form

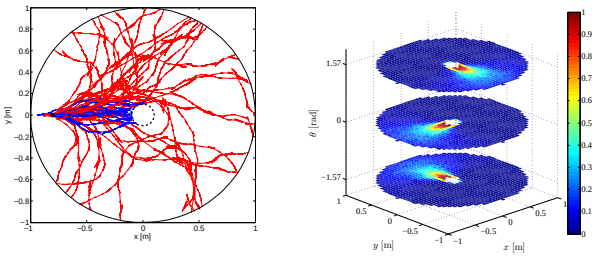
$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \mathrm{sat}^{0.3}_{1.0} \left(\begin{bmatrix} V_n + v^* \\ \omega^* \end{bmatrix} + \begin{bmatrix} \Sigma_1 \, \mathrm{d}W_1 / \Delta t \\ \Sigma_2 \, \mathrm{d}W_2 / \Delta t \end{bmatrix} \right) \quad (12)$$

where $\mathrm{sat}_{(0.3,1.0)^\intercal}$ denotes the saturation function, applied element-wise, capping the magnitude of the first component at 0.3 and of the second at 1.0. The noise components were divided with $\Delta t = 0.2$ seconds to make the implementation compatible with (11) in the form

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} \mathrm{sat}^{0.3}_{1.0} \left(\begin{bmatrix} V_n + v^* \\ \omega^* \end{bmatrix} + \begin{bmatrix} \Sigma_1 \, \mathrm{d}W_1 / \Delta t \\ \Sigma_2 \, \mathrm{d}W_2 / \Delta t \end{bmatrix} \right) \Delta t$$

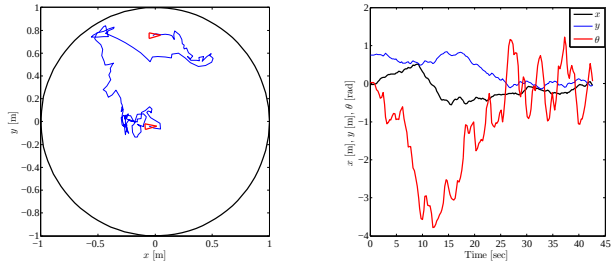
Function $g(\mathrm{q})$ is obtained by simulating (13) and estimating

[2]The physical system's absolute linear velocity is lower bounded at 0.1 m/s.



(a) 50 sample paths for (13) (b) $g(\mathrm{q})$

Fig. 4. Computation of function $g(\mathrm{q})$. Part (a) shows 50 sample paths for (13) used for computation of $g(\mathrm{q})$. Sample paths are color coded; red paths exit at outer boundary (circle), while blue sample paths exit at goal boundary (dotted circle) for an initial condition $(-0.8,\ 0.0,\ 0.0)$. Part (b) gives three slices of the function $g(\mathrm{q})$ for system (11) computed using 500 sample paths.



(a) A sample path with control (9) applied to the system (1) (b) Trajectory of state components for sample path shown in Figure 5(a)

Fig. 5. Experimental result of controller (9) applied to a unicycle-type robot. The observed path is shown in part (a). The outer boundary is circular, the initial configuration is the one on top, and the desired one is at the center, with zero orientation. Initial and final unicycle configurations are marked with triangles. Trajectories of individual state components are shown in part (b), indicating convergence to a neighborhood of the desired configuration. The curve with the dip around 10 sec represents the unicycle's orientation.

$g(\mathrm{q})$ from the sample paths. The control inputs can be computed as per (9).

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}\theta \end{bmatrix} = \begin{bmatrix} V_n \cos\theta \\ V_n \sin\theta \\ 0 \end{bmatrix} + \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1 \\ \mathrm{d}W_2 \end{bmatrix} \quad (13)$$

Figure 4(a) shows 50 sample paths for this system where 28% of the paths exited at the goal boundary, implying that $g([-0.8,\ 0.0,\ 0.0]^\intercal) = 0.28$. Figure 4(b) shows three slices for $\theta = -\pi/2$, 0 and $\pi/2$ from the 4-dimensional plot of the function $g(\mathrm{q})$, estimated using 500 sample paths.

Control law (5)–(12) is applied to the robot of Figure 1(b). The observed path taken by the robot is shown in Figure 5(a) for an initial condition $(x, y, \theta) = (0.0, 0.8, 0.0)$, exiting at a neighborhood of the origin. Figure 5(b) shows the time profile of the path. Figures 6(a)–(b) give the resulting optimal control inputs in a single sample path realization, and Figures 6(c)–(d) reveal the noise components that contaminated the nominal optimal control inputs during the experiment. Figures 6(e)–(f) show the saturated control inputs (12) sent to the robot.

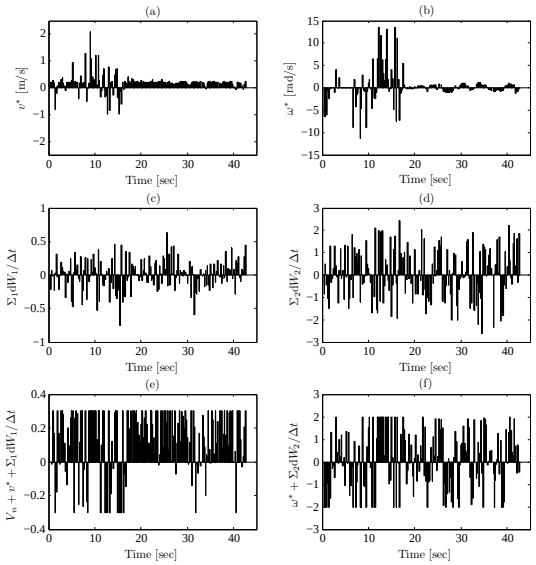


Fig. 6. Experimental data for the trajectories of Figure 5. Plots (a)–(b) depict the optimal linear and angular velocity control, respectively, from (9), Plots (c)–(d) record the noise that disturbed the system, while (e)–(f) show the applied linear and angular velocity inputs (after saturation at control limits).

V. Bounded Inputs And Obstacles

In this section we demonstrate the extension of the control design (9) to cases where the workspace is cluttered with obstacles. The setup is otherwise identical to that of Section IV-D, but the system here is simulated in Player/Stage [31] with control loop frequency set at 5 Hz; hence disturbances such as wheel slip and surface friction are not present, and inputs velocities are assumed to be accurately realized. The system's velocities, however, are constrained in the intervals $[-0.3, 0.3]$ and $[-2, 2]$, for v and ω, respectively. The workspace $\mathcal{P}$ is of radius 1.5 m, containing four obstacles of radius 0.1 m at coordinates $(-0.7, 0.0)$, $(0.7, 0.0)$, $(0.0, -0.7)$ and $(0.0, 0.7)$.

In this scenario, the domain is

$$\mathcal{D} \triangleq \{\mathrm{q} \in \mathbb{R}^2 \times \mathbb{S} \mid \sqrt{x^2 + y^2} < 1.5, \forall\theta\} \ ,$$

with the obstacle region defined as

$$\mathcal{O} \triangleq \{\mathrm{q} \in \mathcal{D} \mid [(x-O_{x_i})^2+(y-O_{y_i})^2]^{\frac{1}{2}} \leq 0.1, \forall\theta, i = 1, \ldots, 4\} \ ,$$

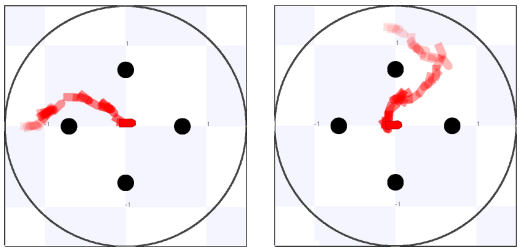
for obstacles centered at (O_{x_i}, O_{y_i}), $i = 1, \ldots, 4$. The free workspace is $\mathcal{P} = \mathcal{D} \setminus \mathcal{O}$, and the goal set is defined as

$$\mathcal{E} = \{\mathrm{q} \in \mathcal{P} \mid \sqrt{x^2 + y^2 + (\cos\theta - 1)^2 + (\sin\theta)^2} \leq 0.1\} \ .$$

Figure 7 shows Stage simulation results for two initial conditions. The robot—although modeled as a point—is shown in the figures as a rectangle, rounded in the front, and the obstacles are depicted as solid disks.

VI. Discussion

The control design method presented is input-optimal and feedback-based, and uses the solution of a PDE precomputed numerically off-line. Due to the numerical solution, the



(a) $q_0 = (-1.2, 0.0, 0.0)$ (b) $q_0 = (0.0, 1.2, 0.0)$

Fig. 7. Simulated paths of a stochastic differential drive robot

resolution of the discretization influences accuracy. State and time can be discretized at different resolution. The question of optimal resolution is beyond the scope of this paper, and the one used in the examples of this paper is chosen empirically; so is the size of the goal set. As the goal set becomes smaller, the set of initial conditions producing trajectories hitting the target becomes smaller as well, and sufficiently accurate estimation of the probability function requires increasingly fine resolution for state-space discretization. Finally, the obstacles considered in Section V are implicitly assumed sparse. Heavily cluttered environments are not in the context of the scenario discussed at the beginning of that section, and are more appropriately treated using alternative techniques (see, for instance, [25], [26]).

VII. Conclusions and Future Work

Unicycle-type robots perturbed by stochastic noise can be steered to a neighborhood of a desired position and orientation using the control law presented in this paper. This method naturally extends to cases where the workspace is sparsely populated by obstacles, and yields controllers that utilize minimum control effort while steering the system to the designated goal set. The design has been validated both in simulations and in experiments.

References

[1] A. Pullin, N. Kohut, D. Zarrouk, and R. Fearing, "Dynamic turning of 13 cm robot comparing tail and differential drive," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2012, pp. 5086–5093.

[2] K. Karydis, I. Poulakakis, and H. G. Tanner, "Probabilistic validation of a stochastic kinematic model for an eight-legged robot," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2013, pp. 2562–2567.

[3] D. Panagou and H. Tanner, "Modeling of a Hexapod Robot; Kinematic Equivalence to a Unicycle," University of Delaware, Department of Mechanical Engineering, Tech. Rep., 04 2009.

[4] K. Karydis, D. Zarouk, I. Poulakakis, R. S. Fearing, and H. G. Tanner., "Planning with the STARs," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014, pp. 3033–3038.

[5] C. Canudas de Wit and O. Sørdalen, "Exponential stabilization of mobile robots with nonholonomic constraints," *IEEE Transactions on Automatic Control*, vol. 37, no. 11, pp. 1791–1797, 1992.

[6] G. Oriolo, A. De Luca, and M. Vendittelli, "WMR control via dynamic feedback linearization: design, implementation, and experimental validation," *IEEE Transactions on Control Systems Technology*, vol. 10, no. 6, pp. 835 – 852, 2002.

[7] D. Panagou, H. G. Tanner, and K. J. Kyriakopoulos, "Dipole-like fields for stabilization of systems with pfaffian constraints." in *Proceedings of IEEE International Conference on Robotics and Automation*, 2010, pp. 4499–4504.

[8] K. A. Morgansen and R. W. Brockett, "Nonholonomic control based on approximate inversion," in *Proceedings of the IEEE American Control Conference*, 1999, pp. 3515–3519.

[9] C.-C. Yih and P. I. Ro, "Near-optimal motion planning for nonholonomic systems using a multi-point shooting method," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 1996, pp. 2943–2948.

[10] M. Kobilarov and G. Sukhatme, "Optimal control using nonholonomic integrators," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2007, pp. 1832–1837.

[11] A. Astolfi, "Discontinuous control of nonholonomic systems," *Systems and Control Letters*, vol. 27, no. 1, pp. 37–45, Jan. 1996.

[12] S. Shah, C. Pahlajani, and H. Tanner, "Probability of success in stochastic robot navigation with state feedback," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 3911–3916.

[13] S. Summers, M. Kamgarpour, C. J. Tomlin, and J. Lygeros, "A Stochastic Reach-Avoid Problem with Random Obstacles," in *Hybrid Systems: Computation and Control*. ACM, 2011, pp. 251–260.

[14] W. Zhao-Jing and L. Yong-Hui, "Stochastic stabilization of nonholonomic mobile robot," in *Proceedings of 30th Chinese Control Conference*, 2011, pp. 1290–1295.

[15] R. Anderson and D. Milutinovic, "A stochastic optimal enhancement of feedback control for unicycle formations," in *International Symposium on Distributed Autonomous Robotic Systems*, 2012.

[16] R. P. Anderson, E. Bakolas, D. Milutinovic, and P. Tsiotras, "The markov-dubins problem in the presence of a stochastic drift field," in *Proceedings of the IEEE Conference on Decision and Control*, 2012, pp. 130–135.

[17] G. Kewlani, G. Ishigami, and K. Iagnemma, "Stochastic mobility-based path planning in uncertain environments," in *Proceedings of the IEEE/RSJ international conference on Intelligent robots and systems*, 2009, pp. 1183–1189.

[18] M. B. Horowitz and J. W. Burdick, "Optimal navigation functions for nonlinear stochastic systems," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 224–231.

[19] W. H. Fleming, "Exit probabilities and optimal stochastic control," *Applied Mathematics and Optimization*, vol. 4, pp. 329–346, 1977.

[20] M. Day, "On a stochastic control problem with exit constraints," *Applied Mathematics and Optimization*, vol. 6, pp. 181–188, 1980.

[21] H. J. Kappen, "Path integrals and symmetry breaking for optimal control theory," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, no. 11, p. P11011, 2005.

[22] E. Theodorou, F. Stulp, J. Buchli, and S. Schaal, "Iterative path integral stochastic optimal control for learning robotic tasks," in *The 18th World Congress of The International Federation of Automatic Control, Milan, Italy*, 2011.

[23] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 820–833, 2011.

[24] B. van den Broek, W. Wiegerinck, and B. Kappen, "Stochastic optimal control of state constrained systems," *International Journal of Control*, vol. 84, no. 3, pp. 597–615, 2011.

[25] S. Shah, C. Pahlajani, N. Lacock, and H. Tanner, "Stochastic receding horizon control for robots with probabilistic state constraints," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2012, pp. 2893–2898.

[26] S. K. Shah, H. G. Tanner, and C. D. Pahlajani, "Optimal navigation for vehicles with stochastic dynamics," *IEEE Transactions on Control Systems Technology*, (to appear) 2015.

[27] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus (Graduate Texts in Mathematics)*, 2nd ed. Springer, 1991.

[28] S. M. LaValle, *Planning Algorithms*. Cambridge, U.K.: Cambridge University Press, 2006.

[29] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed. Springer, 2005.

[30] D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations," *SIAM Review*, vol. 43, no. 3, pp. 525–546, 2001.

[31] B. P. Gerkey, R. T. Vaughan, and A. Howard, "The player/stage project: Tools for multi-robot and distributed sensor systems," in *Proceedings of the 11th International Conference on Advanced Robotics*, 2003, pp. 317–323.