Dynamics-compatible potential fields using stochastic perturbations

Shridhar K. Shah and Herbert G. Tanner

Abstract— **This paper suggests a method for numerically constructing almost globally converging artificial potential fields for motion planning, in a way that ensures that the resulting gradient field is compatible with the dynamics of the navigating robot. Convergence to an arbitrarily small destination set can be guaranteed, and the size of the destination set can be reduced at the expense of additional off-line computational time. The construction is based on the solution of the Hamilton-Jacobi-Bellman (HJB) equation associated with a related stochastic optimal control problem. This partial differential equation (PDE) is solved numerically by simulating paths of the system with Gaussian random perturbation applied to the input. The resulting control laws are optimal in terms of the magnitude of control actuation. The method is applied to the case of a Dubin's car navigating amongst obstacles.**

Keywords - stochastic optimal control, potential field, motion planning, navigation functions

I. Introduction

This paper describes a stochastic simulation method that builds a potential field for robot navigation, which not only is free of local minima, but also provides directly implementable state feedback laws ensuring convergence to goal sets with minimum input effort. The control laws constructed are compatible with the dynamics and differential constraints of the system, since the latter are taken directly into account during the construction process. The result is a numerical method for the construction of *input-optimal* feedback motion plans with guaranteed convergence properties, which are customized to the system's dynamics and have no need for manual parameter tuning.

The connection between the hitting probabilities of random processes and artificial potential fields has long been recognized [1]. While it is possible to generate optimal feedback motion plans by solving the HJB PDE [2], [3], analytical solutions of this partial differential equation cannot be obtained in general. Numerical solution methods on a discrete grid space can be constructed using dynamic programing [4]; the latter reference contains a comprehensive literature review on the issue, which cannot be matched here due to space limitations. The approach of [4] allows a few algorithmic improvement options to speed up the computation of the solution to the HJB, and results in a numerical potential field. More recent approaches [3] exploit advances in the field of numerical solutions of PDEs to scale up the dimension of the state space of the underlying (generally, nonlinear) dynamical system. Kinodynamic motion planning [5], [6] is another approach that incorporates part of the dynamics of the system in the resulting open-loop reference trajectories. These motion references have to be tracked using a separate trajectory tracking controller. If one prefers a one-stage approach like the one in [3], [4], an available alternative is to *learn* a *navigation function* (a particular type of potential function with guaranteed absence of local minima [7]) based on experimentally measured feasible system trajectories [8]. In contrast to the preceding approaches, the latter makes no assumption regarding any knowledge of obstacle locations, and instead approximates an obstacle function component of the navigation function, through the solution of an inverse problem. The desirable feature of absence of local minima is also present in *harmonic* potential fields [1], obtained through the solution of a Laplace PDE [9]. This class of methods are inspired by vector fields associated with physical phenomena such as fluid flow [10] and heat conduction [11], giving local-minima-free potential fields with slightly different properties each time.

Regarding the latter approach, one of the common limitations in harmonic potential field methods is that they do not take into account the system's dynamics during the construction process; rather they focus on the topology of the obstacle-free workspace. As a result, the resulting potential field is not compatible with the system dynamics in general (cf. [3] for exceptions). The learning approach of [8] on the other hand, which does incorporate some dynamic effects through the experimental data that it uses, requires a sufficient number of successful system runs in order for this body of data to become available. Collecting this data may be difficult or labor-intensive, albeit possible. Still, the method cannot guarantee that the resulting potential field is completely compatible with the system dynamics, and there is no apparent guarantee of optimality of the derived control law. Kinodynamic motion planning methods are particularly capable of handling large-dimensional problems, but they are point-to-point methods; that is, they solve the problem of navigating from point A to point B. If a feedback strategy is to be generated using this method, the process needs to be repeated from all practically possible initial conditions. In addition, while the reference trajectories themselves might be time-optimal, the performance of the required trajectory tracking controller may not be. These methods have a history of application for motion planning in different type of robotic systems [4], [6]. However, these tools do not offer a one-stage optimal motion planning feedback control strategy, which is directly implementable on the system dynamics without the need for prior numerical or experimental testing.

Earlier numerical solution approaches to the HJB PDE [4] fit the stated problem specifications; on a closer look,

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

Shridhar K. Shah is with Mathworks Inc (shridhar.shah@mathworks.com). Herbert Tanner is with Department of Mechanical Engineering, University of Delaware, (btanner@udel.edu)

though, it is not clear how dynamic effects such as system inertia is sufficiently captured in the one-step solution computation (between neighboring grid nodes). Furthermore, and although it might be possible in principle to specify goal sets yielding full-state convergence for the system at hand, this has not been explicitly done—unicycle dynamics are steered to a ball around the goal point with arbitrary orientation. Finally, the cost in computation time imposed by executing the dynamic programming algorithm is a fixed overhead, which might not be suited for real-time implementation on vehicles with limited computational capabilities. More recent developments [3] directly address the issue of incorporating system dynamics in the resulting potential field, offering one of the most advanced solutions to date, with the only potential limitation being that sophisticated numerical PDE solvers are brought to bear.

The method proposed in this paper offers a straightforward alternative to direclty involving PDE solvers, by approximating the solution of the HJB through a process of stochastic simulation on the basis of the Feynman-Kac lemma [12]. Just like [3], the motion planning problem is being mapped to a stochastic optimal control problem, but from that point on the solution approach diverges. The stochastic optimal control formulation, together with the specific structure of the resulting stochastic HJB equation, affords a linearization of the PDE through a logarithmic transformation. Then, application of the Feynman-Kac lemma [12] permits an approximation of the solution to the HJB PDE through numerical simulation of an associated stochastic differential equation (SDE). It appears that this specific formulation permits numerical solutions in dimensions slightly higher than alternative constructions based on the deterministic version of the HJB.

The logarithm transform used to linearize the HJB equation associated with the stochastic optimal control problem appears in [13]. Based on this transform, path integral methods have been developed to solve the PDE [14]. Different variations of this path-integral method for the solution of stochastic optimal control problems have been explored, for reinforcement learning [15], variable stiffness control (equivalent to automatic tuning of PD gains) [16], and risk sensitive control idea [17]. To a great extend, path integral tools [14], [15] are used to learn control strategies to follow desired trajectories, and do not naturally lend themselves to applications in feedback motion planning problems. To achieve this, we deviate from the classic path integral formulation, and instead use the notion of exit time. A stochastic optimal control problem with exit time was considered in [18]. The introduction of the exit time allows us to simulate an SDE, which is truncated at the exit time and derive a solution of the related PDE on which to base our potential function. These ideas have been explored for enabling *stochastic* nonholonomic systems to navigate in constrained environments by transitioning between waypoints in a receding horizon fashion [19]. In that later work, however, the model of the vehicle was stochastic, orientation was ignored, and local navigation between obstacles did not have to consider collisions to obstacles. Here, in contrast, the system is deterministic (its stochastic extension is used only for design) and obstacles are interspersed between initial and desired poses.

The method proposed in this paper is associated with computation effort that is tunable; obviously, the price of reducing computation is less accuracy and sparser potential field representation. In addition, in the implementation example reported in this paper, a nonholonomic system (Dubin's car) is shown to converge to goal sets which restrict final orientation within a pre-specified range. The proposed method is based on forward stochastic simulation of the system dynamics, which offers two advantages: explicit consideration of long-range inertial effects, and robustness to noise and perturbations up to thresholds determined during the design phase. Finally, the potential function produced as the solution of the HJB can be interpreted directly as the logarithm of the probability of achieving convergence to the goal when using the suggested control law, in the presence of noise contaminating the input.

The paper is organized in the following way. Section II states the problem considered. Section III contains the main technical result, describes the construction of the potential field, and establishes its convergence guarantees. Section IV illustrates the application of the method to the case of steering Dubin's car amongst obstacles, and the paper concludes with Section VI, which summarizes the results and hints at future extensions.

II. Convergence with obstacle avoidance

Let $\mathcal{D} \subseteq \mathbb{R}^n$ be a bounded domain, and let the obstacles within this space be represented as a set $\mathcal{O}$. The collision-free workspace where the system can move is expressed as $\mathcal{P} \triangleq \mathcal{D} \setminus \mathcal{O}$. Assume that the boundary of $\mathcal{P}$, denoted $\partial\mathcal{P}$ is described by a function which is twice differentiable (i.e., in $\mathcal{C}^2$), and consider a time invariant nonlinear system with state $\mathbf{q}$ and control input $u : \mathbb{R}^n \to \mathbb{R}^n$ having dynamics of the form

$$\dot{q} = f(\mathbf{q}) + g(\mathbf{q})u(\mathbf{q}) \ . \tag{1}$$

The the nominal drift vector field $f : \mathbb{R}^n \to \mathbb{R}^n$ and the control matrix $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are assumed to be Lipschitz continuous—fairly standard assumptions for the drift and diffusion terms in the stochastic process [12] which will be associated with (1); local relaxations of these continuity assumptions are possible [20].

The objective is to design a static feedback control $u(\mathbf{q})$ such that the system converges to an arbitrarily small set around origin $\mathbf{q} = \mathbf{0}$.

III. Learning to navigate optimally

Consider the free workspace $\mathcal{P}$ as a bounded domain with $\mathcal{C}^2$ boundary $\partial\mathcal{P}$ and the closure $\overline{\mathcal{P}}$. The goal set for the system is denoted $\mathcal{G}$ and without loss of generality is assumed to be an ε-neighborhood of the origin:

$$\mathcal{G} = \{\mathbf{q} \in \mathcal{P} \mid \|\mathbf{q}\| \leq \varepsilon\} \ .$$

Assume that the boundary of $\mathcal{G}$, $\partial\mathcal{G}$, is also $\mathcal{C}^2$ and is disjoint from $\partial\mathcal{P}$.

Let $W = \{W(t), \mathcal{F}_t : 0 \le t < \infty\}$ be an m-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 the probability measure, and $\{\mathcal{F}_t : t \ge 0\}$ 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 [12]. Taking now

$$V(\mathbf{q}) \triangleq -\log\Big(\mathbb{P}\left[\mathbf{q}(\tau) \in \partial\mathcal{G} \mid \mathbf{q}(0) = \mathbf{q}\right]\Big) \ ,$$

and defining $\Sigma : \mathbb{R}^n \to \mathbb{R}^{m\times m}$ as a bounded matrix with bounded inverse which is also Lipschitz continuous on $\mathcal{P}\setminus\mathcal{G}$, and after setting $a(\mathbf{q}) \triangleq \Sigma(\mathbf{q})\Sigma^\intercal(\mathbf{q})$, we expect that the feedback law we seek to derive is of the form

$$u(\mathbf{q}) = -a(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} V(\mathbf{q}) \ . \tag{2}$$

Here, the function $V(\mathbf{q})$ will be computed by numerically simulating the stochastic system

$$\mathrm{d}\mathbf{q} = f(\mathbf{q})\,\mathrm{d}t + g(\mathbf{q})(\Sigma(\mathbf{q})\,\mathrm{d}W); \ \ \mathbf{q}(0) = \mathbf{q} \tag{3}$$

The continuity conditions on $f(\mathbf{q})$ and $g(\mathbf{q})$ are still in force.

With respect to (3) define τ to be the first (stopping) time at which the system hits either the boundary $\partial\mathcal{P}$ (which represents the boundary of obstacles) or that of $\partial\mathcal{G}$. In this case, function V represents the negative logarithm of the *probability* of sample paths hitting the goal set boundary *before* any other boundary surface. This probability can be estimated empirically by simulating a number sample paths of (3) for different initial conditions.

A. Derivation of the potential field

Let $\mathbb{E}^{\mathbf{q}}$ denote expectation over all trajectories starting from initial condition $\mathbf{q}$, and define $t \wedge \tau \triangleq \min(t, \tau)$. Consider a stochastic perturbation applied to (1) in the form of

$$\mathrm{d}\mathbf{q} = f(\mathbf{q})\,\mathrm{d}t + g(\mathbf{q})\big[u(\mathbf{q})\,\mathrm{d}t + \Sigma(\mathbf{q})\,\mathrm{d}W\big]; \ \ \mathbf{q}(0) = \mathbf{q} \ . \tag{4}$$

This section presents the control law $u^*(\mathbf{q})$ that minimizes the cost functional

$$V(t,\mathbf{q}) = \min_{u(\mathbf{q})} \mathbb{E}^{\mathbf{q}}\left[\int_0^{t\wedge\tau} \frac{1}{2} u^\intercal(s)\, a(\mathbf{q})^{-1}\, u(s)\,\mathrm{d}s + \Phi\big(\mathbf{q}(t\wedge\tau)\big)\right] \tag{5}$$

recalling that τ is the first exit time from $\mathcal{P}\setminus\mathcal{G}$.

Letting $t \to \infty$, the value function takes the form

$$V(\mathbf{q}) = \min_{u(\mathbf{q})} \mathbb{E}^{\mathbf{q}}\Big[\int_0^{\tau} \frac{1}{2} u^\intercal(s)\, a(\mathbf{q})^{-1}\, u(s)\,\mathrm{d}s + \Phi\big(\mathbf{q}(\tau)\big)\Big] \ . \tag{6}$$

The HJB equation associated with the optimal control problem (5) is expressed in terms of the generator (second-order partial differential operator) $\mathcal{A}$ corresponding to the diffusion (4),

$$\mathcal{A} \triangleq f^T(\mathbf{q})\partial_{\mathbf{q}} + u^\intercal(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} + \frac{1}{2}\mathrm{tr}\big\{\Sigma^\intercal(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{qq}}\, g(\mathbf{q})\, \Sigma(\mathbf{q})\big\} \ , \tag{7}$$

in the form

$$\min_{u(x)} \big\{\mathcal{A}V(\mathbf{q}) + \frac{1}{2} u^\intercal(\mathbf{q})\, a(\mathbf{q})^{-1}\, u(\mathbf{q})\big\} = 0 \ .$$

This minimizer is [18]

$$u^*(\mathbf{q}) = -a(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} V(\mathbf{q}) \tag{8}$$

and substituting in the HJB equation yields

$$f^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} V(\mathbf{q}) - \frac{1}{2}\partial_{\mathbf{q}} V^\intercal(\mathbf{q})\, g(\mathbf{q})\, a(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} V(\mathbf{q}) + \frac{1}{2}\mathrm{tr}\big\{\Sigma^\intercal(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{qq}} V(\mathbf{q}) g(\mathbf{q}) \Sigma(\mathbf{q})\big\} = 0 \ . \tag{9}$$

Suppose now that the solution to the PDE is of the form [18]

$$V(\mathbf{q}) = -\log\psi(\mathbf{q})$$

and substitute to (9) to obtain the PDE

$$\partial_{\mathbf{q}}\psi^\intercal(\mathbf{q})\, f(\mathbf{q}) + \frac{1}{2}\mathrm{tr}\big\{\Sigma^\intercal(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{qq}}\psi(\mathbf{q})\, g(\mathbf{q})\, \Sigma(\mathbf{q})\big\} = 0 \tag{10}$$

with boundary condition

$$\psi(\mathbf{q}) = \exp\big[-\Phi\big(\mathbf{q}(\tau)\big)\big]; \ \ \mathbf{q} \in \partial(\mathcal{P}\setminus\mathcal{G}) \ .$$

An analytic solution of (10) is generally not possible for complex nonlinear systems. However, the Feynman-Kac formula relates such a PDE to a SDE, and allows one to determine the solution of the PDE by simulating the SDE. The associated SDE is constructed as follows. Recalling that τ has been defined as the first exit time from domain $\mathcal{P}\setminus\mathcal{G}$, let $\zeta(t)$ be the Markov process on $\mathcal{P}\setminus\mathcal{G} \subset \mathbb{R}^n$ obeying [12]

$$\mathrm{d}\zeta(t) = b\big(\zeta(t)\big)\,\mathrm{d}t + g\big(\zeta(t)\big)\,\Sigma\big(\zeta(t)\big)\,\mathrm{d}W(t) \ . \tag{11}$$

Then the solution of (10) satisfies

$$\psi(\mathbf{q}) = \mathbb{E}\left[\exp\big[-\Phi\big(\zeta(\tau)\big)\big] \mid \zeta(0) = \mathbf{q}\right] \ .$$

It is possible to guarantee that the system does not exit through the part of the boundary of $\partial\mathcal{P}$ adjacent to obstacles by imposing an infinite cost on hitting that part of the boundary. Defining $\mathcal{X}_{\partial\mathcal{P}}$ to be the indicator function on $\partial\mathcal{P}$,

$$\mathcal{X}_{\partial\mathcal{P}} = \begin{cases} 0 & \text{on } \partial\mathcal{G} \\ 1 & \text{on } \partial\mathcal{P} \end{cases} \ ,$$

the penalty for hitting the wrong boundary is introduced with

$$\Phi(\mathbf{q}) = +\infty \cdot \mathcal{X}_{\partial\mathcal{P}} \ .$$

Then, function $\psi(\mathbf{q})$ becomes the probability of the state hitting the goal boundary $\partial\mathcal{G}$ before any other:

$$\psi(\mathbf{q}) = \mathbb{P}\left[\zeta(\tau) \in \mathcal{G} \mid \zeta(0) = \mathbf{q}\right] \ . \tag{12}$$

With this insight, the optimal control law (8) is rewritten as

$$u^* = -a(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} \{ -\log \mathbb{P}[\zeta(\tau) \in \mathcal{G} \mid \zeta(0) = \mathbf{q}] \} \quad (13)$$

which is essentially the analytic expression of the right-hand side of (2).

Remark 1: Evaluating (8) requires $\psi(\mathbf{q})$, which is found either by solving (10), or by numerically simulating (11) and computing (12). Since Φ imposes an infinite penalty on states exiting the obstacle boundary $\mathcal{P}$, the above construction offers a guarantee that system exits first at $\partial\mathcal{G}$ and not on any other portion of the boundary. As states approach $\partial\mathcal{P}$, inputs pushing the system away from it increase. This means that in the case where the control input of (1) is constrained, there can be initial conditions from which no feasible control input exists to prevent the system from hitting the wrong boundary, $\partial\mathcal{P}$, or to force it enter $\mathcal{G}$.

Here we assume that $u(\mathbf{q})$ is unconstrained, and we state the following result.

Theorem 2: Given that $u \in \mathbb{R}^n$ defined in (2) exists, the system (1) steered by (8) converges to $\mathcal{G}$ from all initial conditions in the interior of $\mathcal{P}$.

Proof: We treat the deterministic system (1) as a special case of (4) where $\mathrm{d}W \equiv 0$.

First we show that the system will not hit $\partial\mathcal{P}$. This part of the proof is established by contradiction. Consider the closed-loop system (4)–(8)

$$\mathrm{d}\mathbf{q} = f(\mathbf{q})\,\mathrm{d}t + g(\mathbf{q}) \left(-a(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} V(\mathbf{q})\,\mathrm{d}t + \Sigma(\mathbf{q})\,\mathrm{d}W \right) ,$$

and assume that there exists an initial condition $\mathbf{q}_0$ from which at least one sample path exits through $\partial\mathcal{P}$ instead of $\partial\mathcal{G}$. As the terminal cost $\Phi \to \infty$ on $\mathbf{q} \in \partial\mathcal{P}$, the value function

$$V(\mathbf{q}_0) = \min_{u(\mathbf{q})} \mathbb{E}^{\mathbf{q}_0} \left[\int_0^\tau \frac{1}{2} u^\intercal(s)\, a(\mathbf{q})^{-1}\, u(s)\,\mathrm{d}s + \Phi(\mathbf{q}(\tau)) \right]$$

becomes unbounded with $\mathbf{q} \to \mathbf{q}(\tau) \in \partial\mathcal{P}$. We know that $\tau < \infty$, so $V(\mathbf{q}_0)$ is to escape to infinity in finite time. However, the mere existence of an optimal control law implies by default that $V(\mathbf{q}_0) < \infty$, and thus an unbounded value function directly contradicts the feasibility of the optimal control problem. It thus follows, that since almost no sample paths hit $\partial\mathcal{P}$ and the first exit time is finite, they should necessarily exit from $\partial\mathcal{G}$:

$$\mathbb{P}\left[\mathbf{q}(\tau) \in \partial\mathcal{G} \mid \mathbf{q}(0) = \mathbf{q}_0 \right] = 1 \ . \quad (14)$$

Now recall (1); this closed loop system has dynamics of the form

$$\dot{\mathbf{q}} = f(\mathbf{q})\,\mathrm{d}t - g(\mathbf{q})\, a(\mathbf{q})\, g^\intercal(\mathbf{q})\, \partial_{\mathbf{q}} V(\mathbf{q})$$

where the value function $V(\mathbf{q})$ is

$$V(\mathbf{q}) = \min_{u(\mathbf{q})} \left[\int_0^\tau \frac{1}{2} u^\intercal(s)\, a(\mathbf{q})^{-1}\, u(s)\,\mathrm{d}s + \Phi(\mathbf{q}(\tau)) \right] \ .$$

Since in this case $dW \equiv 0$, all sample paths are identical and (14) reduces to

$$\mathbf{q}(\tau) \in \partial\mathcal{G}, \ \ \forall \mathbf{q}(0) = \mathbf{q} \ .$$

■

IV. Dubin's Car

We demonstrate the proposed method using the kinematics of Dubin's car [21] as an example. This kinematics is expressed as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

where $u \in [-1, 1]$. The Dubin's car has a constant linear speed of 1 m/s; because of the constraint on the angular speed, there is a minimum turning radius. We identify the state as $\mathbf{q} = (x, y, \theta)^\intercal$, the drift vector field as $f(\mathbf{q}) = (\cos\theta, \sin\theta, 0)^\intercal$ and the control vector field as $g(\mathbf{q}) = (0, 0, 1)^\intercal$.

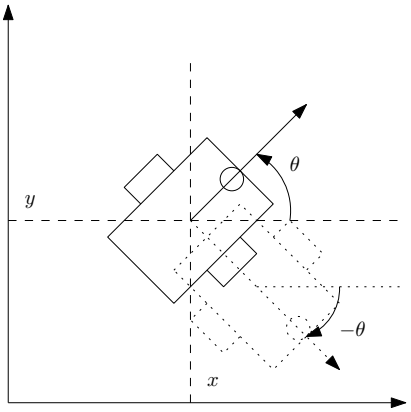

Fig. 1. A graphical representation of Dubin's car

To construct the potential field, we simulate the following stochastically perturbed model parameterized by Σ

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}\theta \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} (\Sigma\,\mathrm{d}W) \ . \quad (15)$$

For simulation purposes we take $\Sigma = 0.5$ and consider a spherical workspace of radius $R = 10$ m containing two obstacles of radius 1 m at coordinates $(-6, -2)$ and $(-2, -6)$. System (15) is a stochastic differential equation, which we simulate using Euler-Maruyama method [22]. We divide the state space into a grid of $41 \times 41 \times 41$ grid, and treat each grid point as a potential initial condition. The discretization time is chosen to be $\mathrm{d}t = 0.1$ sec, and the simulation time is continuously increased until 95% of the paths computed exit either at the goal set or at the obstacle boundary. The Brownian noise term can be intuitively thought of as Gaussian noise with zero mean and variance equal to $\sqrt{\mathrm{d}t}$.

In the simulation results depicted in Fig. 2, The free workspace is enclosed in a circle of radius 5 punctured by two disk-shaped obstacles marked in black. The final orientation of the vehicle is not regulated. The goal set is defined as

$$\mathcal{G} = \{ \mathbf{q} \in \mathcal{P} \mid \sqrt{x^2 + y^2} \leq 1, \theta \in [-\pi, \pi] \}$$

, and is marked with a blue circle. Several initial conditions are tested, each marked with a small blue circle, while the initial orientation of the vehicle is represented by a smaller red dot on the small blue circle showing the initial position. Because the cost function penalizes control input, the resulting trajectories use minimum control effort to reach the goal set, although the path traversed may not necessarily the shortest possible. The control inputs used to produce the sample path annotated with numeral 1, are shown in Fig. 2.

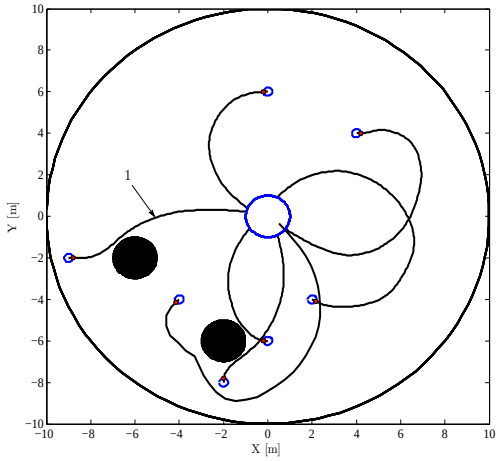

Fig. 2. Simulated paths for different initial conditions for a Dubin's car, ignoring final orientation. Initial positions are marked with a small circle

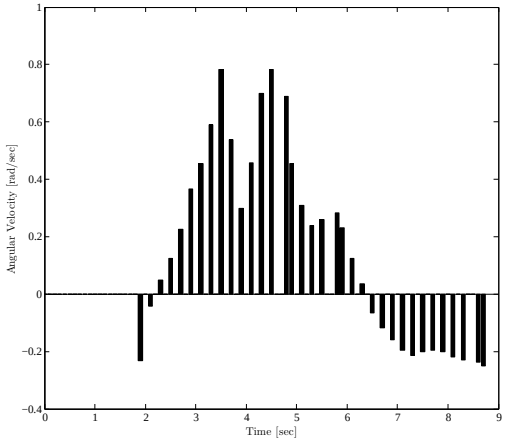

Fig. 3. Control inputs for the trajectory annotated with 1 in Fig. 2

Function ψ is the probability that sample paths reach the goal boundary before hitting any other:

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

This function can be estimated as the ratio of sample paths hitting the goal over all sample paths generated. In these simulations, function $\psi(\mathbf{q})$ is approximated by simulating 500 sample paths of (15) for each initial condition. A plot of $\psi(\mathbf{q})$ estimated in this way is shown in Fig. 4 for $\theta = \frac{\pi}{2}$. Once ψ becomes available, the control law is expressed as

$$u(\mathbf{q}) = -0.25 \cdot \partial_{\mathbf{q}}(-\log \psi(\mathbf{q})) \ . \tag{16}$$

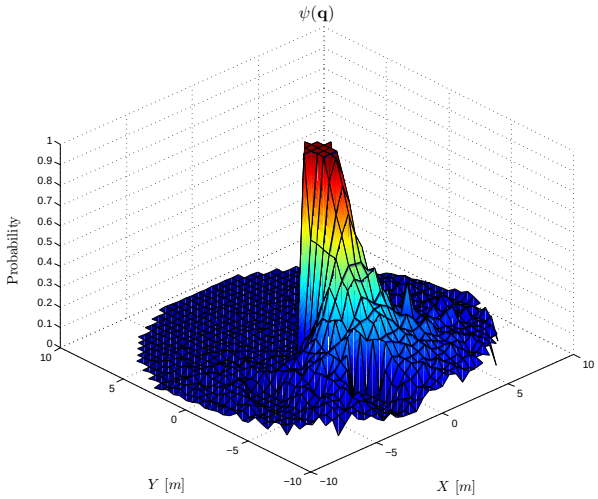

Fig. 4. Function $\psi(\mathbf{q})$ for $\theta = \frac{\pi}{2}$ and $\Sigma = 0.5$ for the case of convergence without orientation regulation

To avoid numerical overflows, all zero values of $\psi(\mathbf{q})$ were mapped to an arbitrarily small value (e.g., 10^{-10}). Equation (16) prescribes large values near the obstacle boundary, but Dubin's vehicle input constraints saturate the values obtained within the interval $[-1, 1]$.

In a second set of simulation trials, we regulate final orientation too by by defining the goal set as

$$\mathcal{G} = \left\{ \mathbf{q} \in \mathcal{P} \mid \sqrt{x^2 + y^2} \leq 1, \theta \in \left[\tfrac{\pi}{8}, \tfrac{\pi}{4}\right] \right\} \ .$$

The resulting sample paths are shown Fig. 5, where it is evident that all generated paths hit the target set with an orientation between $\frac{\pi}{8}$ and $\frac{\pi}{4}$. It is possible to reduce the acceptable range of final orientation angles, but with the constraint on angular velocity this would limit the set of initial conditions from which convergence with regulated orientation can be achieved without collision to obstacles.

V. Discussion

The method presented in this paper is a numerical approach and hence the resolution of the discretization influences accuracy. It should be noted, however, that state and time spaces generally admit discretizations with different resolution. The question of what is the optimal resolution is beyond the scope of this paper; the resolution used in the examples of the preceding section is chosen empirically. Another important parameter 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 then requires increasingly fine resolution in the discretization of the state space.

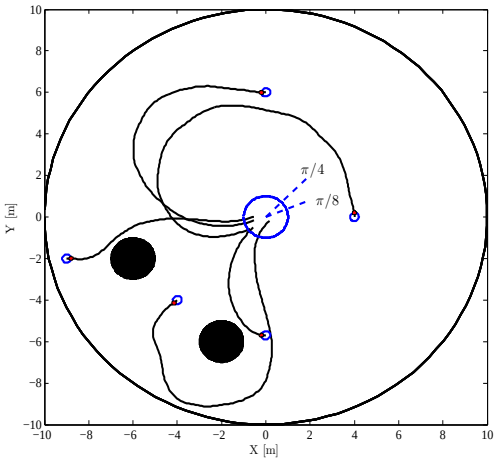

Fig. 5. Simulated paths for different initial conditions for a Dubin's car with orientation convergence between $\frac{\pi}{8}$ and $\frac{\pi}{4}$

The selection of the stochastic perturbation is yet another design parameter that affects the convergence of sample paths simulated. For the case of Dubin's car, a relatively large value for Σ (in the order of 1.0) results in a function $\psi(\cdot)$ with gradient too small to yield a uniform average convergence rate. Too small values for Σ (in the order of 0.1 and below) result in $\psi(\cdot)$ varying too abruptly, suggesting that a finer discretization of the state space might be necessary. Ultimately, what is an appropriate value for Σ depends on the magnitude of the control authority given; in this example, that would be the interval $[-1, 1]$. Reasonable choices for Σ are those that produce perturbations in the order of, and of size comparable to available control inputs. (These perturbations are samples taken from a distribution $\mathcal{N}(0, \Sigma \cdot \sqrt{dt})$.

The presented approach is practical for systems of up to five or six dimensions; given sufficient computational power, eight dimensions can also be handled. Alternative motion planning methods applicable to high-dimensional spaces typically do not produce *feedback* control policies as do the ones presented here. They are effective for steering a vehicle from (a given) point A to (a given) point B, but may not be used to obtain motion planning strategies for all possible initial conditions.

VI. Conclusions

The problem of deriving feedback control policies to optimally steer vehicles with nontrivial kinematics in obstacle environments with guaranteed convergence properties to a goal set, is ultimately related to the solution of a HJB PDE. In the specific case of control laws which are optimal in terms of actuation effort, the solution of this PDE can be interpreted as the the negated logarithm of the probability of a stochastically perturbed version of the unforced original dynamics to hit the goal set. This insight allows the calculation of optimal control laws based on the empirical exit probability, obtained through Monte Carlo simulations of the associated stochastic dynamics.

References

[1] C. I. Connolly, "Harmonic functions and collision probabilities," *The International Journal of Robotics Research*, vol. 16, no. 4, pp. 497–507, 1997.

[2] S. Sundar and Z. Shiller, "Optimal obstacle avoidance based on the Hamilton-Jacobi-Bellman equation," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, may 1994, pp. 2424–2429.

[3] 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.

[4] S. M. Lavalle and P. Konkimalla, "Algorithms for computing numerical optimal feedback motion strategies," *The International Journal of Robotics Research*, vol. 20, no. 9, pp. 729–752, 2001.

[5] B. Donald, P. Xavier, J. Canny, and J. Reif, "Kinodynamic motion planning," *Journal of the Association for Computing Machinery*, vol. 40, no. 5, pp. 1048–1066, Nov. 1993.

[6] S. M. LaValle and J. J. Kuffner, "Randomized kinodynamic planning," *The International Journal of Robotics Research*, vol. 20, no. 5, pp. 378–400, 2001.

[7] D. E. Koditschek and E. Rimon, "Robot navigation functions on manifolds with boundary," *Advances in Applied Mathematics*, vol. 11, no. 4, pp. 412–442, 1990.

[8] I. Filippidis, K. Kyriakopoulos, and P. Artemiadis, "Navigation functions learning from experiments: Application to anthropomorphic grasping," in *Proceedings of IEEE International Conference on Robotics and Automation*, 2012, pp. 570–575.

[9] C. Connolly, J. Burns, and R. Weiss, "Path planning using Laplace's equation," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, 1990, pp. 2102–2106.

[10] C. Louste and A. Lieacute;geois, "Path planning for non-holonomic vehicles: a potential viscous fluid field method," *Robotica*, vol. 20, no. 3, pp. 291–298, 2002.

[11] Y. Wang and G. Chirikjian, "A new potential field method for robot path planning," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 2, 2000, pp. 977–982.

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

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

[14] 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.

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

[16] 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.

[17] 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.

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

[19] S. Shah and H. G. Tanner, "Stochastic receding horizon control: Application to an octapedal robot," in *Proceedings of the SPIE Defense, Security, and Sensing*, 2013, pp. 87 251D–87 251D–9.

[20] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. Springer, 2010.

[21] L. E. Dubins, "On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents," *American Journal of Mathematics*, vol. 79, pp. 497–516, 1957.

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