Motion Planning via Optimal Control for Stochastic Processes

Peyman Mohajerin Esfahani, Debasish Chatterjee, and John Lygeros

Abstract—**We study stochastic motion planning problems which involve a controlled process, with possibly discontinuous sample paths, visiting certain subsets of the state-space while avoiding others in a sequential fashion. For this purpose, we first introduce two basic notions of motion planning, and then establish a connection to a class of stochastic optimal control problems concerned with sequential stopping-times. A weak dynamic programming principle (DPP) is then proposed, which characterizes the set of initial states that admit the existence of a policy enabling the process to execute the desired maneuver with probability at least as much as some pre-specified value. The proposed DPP consists of some auxiliary value functions defined in terms of discontinuous payoff functions. An application of the DPP is demonstrated in the context of controlled diffusion processes thereafter. It turns out that the aforementioned set of initial states can be characterized as the level set of a discontinuous viscosity solution to a sequence of partial differential equations, for which the first one has a known boundary condition, while the boundary conditions of the subsequent ones are determined by the solutions to the preceding steps. Finally, the generality and flexibility of the theoretical results are illustrated with the aid of an example involving biological switches.**

I. INTRODUCTION

Motion planning can be viewed as a scheme of excursions to visiting certain specific sets in a specific order according to a specified time schedule. In the context of motion planning for controlled dynamical systems, the central issue is to determine whether there exits an admissible policy to drive the process through some sets while visiting certain targets in a pre-assigned order and scheduled times. In the deterministic setting, motion planning problems have been studied extensively from different perspectives; here we cite two representative articles [**?**], [**?**] and refer to the references therein for further details of the literature. In this article we focus on the stochastic counterpart of the motion planning. The basic motion planning problem involving two targets and obstacle sets has been investigated in different contexts, e.g., from computational standpoint in finite probability spaces for a class of continuous-time Markov decision processes (CTMDPs) [**?**], or in discrete-time stochastic hybrid systems based on a dynamic programming approach in [1], [**?**]. In our earlier works [2] and [3] we focused on reachability of controlled diffusion processes; the reachable sets in these works

PME and JL are with the Automatic Control Laboratory, ETH Zürich, 8092 Zürich, Switzerland; DC is with the Systems & Control Engineering, IIT-Bombay, Powai, Mumbai 400076, India. Emails: {mohajerin,lygeros}@control.ee.ethz.ch, chatterjee@sc.iitb.ac.in

were characterized as superlevel sets of value functions which are given by the discontinuous viscosity solutions to certain partial differential equations with some associated boundary conditions. Here we continue our study by moving beyond reachability to more complex motion planning specifications for a larger class of stochastic processes with possibly discontinuous sample paths.

We first introduce different motion planning scenarios in the context of piecewise continuous processes. We address the following natural question: for which set of initial states does there exist an admissible policy such that the controlled stochastic processes satisfy the motion planning specifications with a probability greater than a given value p? To characterize this set of initial states, we establish a connection between the motion planning specifications to a class of stochastic optimal control problems involved sequential stopping times. We shall be concerned with a stochastic process started and stopped when it hits certain subsets of the state-space, or equivalently, the process obtained by concatenating segments of the original process between consecutive stopping times.

Under certain mild assumptions on the admissible polices, the stochastic process, and the sets concerned the motion planning, we propose a Dynamic Programming Principle (DPP) in which some auxiliary value functions are required. The DPP is introduced in a weak version in the spirit of [4]; this formulation does not require measurability of the value functions. In the following sections we will focus on a class of diffusion processes as the strong solution of a stochastic differential equation (SDE), in which the required assumptions of the DPP are investigated. In the light of the proposed DPP, we develop a new framework to characterize the desired initial sets based on tools from partial differential equations (PDE's). Due to the discontinuities in the value functions corresponding to these problems, all the PDE's are understood in the generalized notion of the so called viscosity solutions. It turns out to solve the value functions a series of PDEs are considered, in which the preceding PDE provides the boundary condition of the proceeding PDE, i.e., the PDEs are solved in a recursive fashion. In order to numerically compute the desired initial sets by means of off-the-shelf PDE solvers, some numerical issues are discussed thereafter.

Mention may be made of the fact that the techniques proposed here suffer from the curse of dimensionality. Given a continuous-time controlled stochastic process, e.g., a controlled diffusion, one approach to solving motion planning problems may proceed with a discretization of the state-space, constructing a controlled Markov chain on the discretized space that approximates the original process in a certain way, and then to deal with this Markov chain insofar as motion plan-

ning is concerned. While it may appear that this discretized setting is conceptually simpler, it does not lead to dimensionality reduction, and moreover, introduces the non-trivial issue of the quality of the approximation and the associated errors due to the discretization involved. Furthermore, to our knowledge, there appears to be no off-the-shelf software that algorithmically leads to the aforementioned discretization. In contrast, our techniques deal directly with the given controlled stochastic process. The implementation of our techniques require employment of off-the-shelf PDE solvers, which is a well-studied topic. The errors in our numerical solutions occur only due to the employment of PDE solvers.

The article is organized as follows: in §II we formally introduce the stochastic motion planning problems on the prescribed probability space. In §III we construct a connection between the motion planning problems to a class of stochastic optimal control problems, for which a weak DPP is proposed in §IV in terms of some auxiliary value functions. An application of the proposed DPP is illustrated in §V, which leads to an alternative characterization of the motion planning objective by a series of PDE's in a recursive fashion. To validate the performance of the proposed methodology, in §VI the theoretical results are applied to a biological two-gene network, where the quality of the biological switch is investigated. For better readability, some of the technical proofs of §III and §V are removed, and we refer interested readers to [5, Appendix].

NOTATION

For the ease of readers, we provide here a partial notation list which will be also explained in more details later throughout the article:

- $\wedge$ (resp. $\vee$): minimum (resp. maximum) operator;
- A^c (resp. A°): complement (resp. interior) of the set A;
- $B_r(x)$: open Euclidian ball centered at x and radius r;
- $\mathfrak{B}(\mathbb{A})$: Borel σ-algebra on a topological space $\mathbb{A}$;
- $\mathcal{U}_t$: set of admissible policies at time t;
- $(X_s^{t,x;\boldsymbol{u}})_{s\geq 0}$: stochastic process under the control policy $\boldsymbol{u}$ and convention $X_s^{t,x;\boldsymbol{u}} := x$ for all $s \leq t$;
- $(W_i \rightsquigarrow G_i)_{\leq T_i}$ (resp. $W_i \xrightarrow{T_i} G_i$) : motion-planning events of reaching G_i sometime before time T_i (resp. at time T_i) while staying in W_i, see Definition II.1;
- $\left(\Theta_i^{A_{k:n}}\right)_{i=k}^n$: sequential exit-times from the sets $(A_i)_{i=k}^n$ in order, see Definition III.1;
- V^* (resp. V_*): upper (resp. lower) semicontinuous envelope of the function V;
- $\mathcal{L}^u$: Dynkin operator, see Definition V.5.

II. GENERAL SETTING AND PROBLEM DESCRIPTION

Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ whose filtration $\mathbb{F} := (\mathcal{F}_s)_{s\geq 0}$ is generated by an $\mathbb{R}^{d_z}$-valued process $\boldsymbol{z} := (z_s)_{s\geq 0}$ with independent increments. Let this natural filtration be enlarged by its right-continuous completion, i.e. it satisfies the usual conditions of completeness and right continuity [6, p. 48]. For future purposes, we introduce an auxiliary subfiltration $\mathbb{F}_t := (\mathcal{F}_{t,s})_{s\geq 0}$, where $\mathcal{F}_{t,s}$ is the $\mathbb{P}$-completion of $\sigma(z_r - z_t, t \leq r \leq t \vee s)$. Note that for $s \leq t$, $\mathcal{F}_{t,s}$ is

the trivial σ-algebra, and any $\mathcal{F}_{t,s}$-random variable is independent of $\mathcal{F}_t$. By definitions, it is obvious that $\mathcal{F}_{t,s} \subseteq \mathcal{F}_s$ with equality in case of $t = 0$.

The object of our study is an $\mathbb{R}^d$-valued controlled random process $\left(X_s^{t,x;\boldsymbol{u}}\right)_{s\geq t}$, initialized at (t,x) under the control policy $\boldsymbol{u} \in \mathcal{U}_t$, where $\mathcal{U}_t$ is the set of admissible policies at time t. Let $T > 0$ be a fixed time horizon, and let $\mathbb{S} := [0,T] \times \mathbb{R}^d$. Throughout this work we assume that for every $(t,x) \in \mathbb{S}$ and $\boldsymbol{u} \in \mathcal{U}_t$, the process $\left(X_s^{t,x;\boldsymbol{u}}\right)_{s\geq t}$ is $\mathbb{F}$-adapted process with RCLL sample paths.[1] We denote by $\mathcal{T}$ the collection of all $\mathbb{F}$-stopping times; for $\tau_1, \tau_2 \in \mathcal{T}$ with $\tau_1 \leq \tau_2$ $\mathbb{P}$-a.s. we let the subset $\mathcal{T}_{[\tau_1,\tau_2]}$ denote the collection of all $\mathbb{F}_{\tau_1}$-stopping times τ such that $\tau_1 \leq \tau \leq \tau_2$ $\mathbb{P}$-a.s. Measurability on $\mathbb{R}^d$ will always refer to Borel-measurability, and $\mathfrak{B}(\mathbb{A})$ stands for the Borel σ-algebra on a topological space $\mathbb{A}$. Throughout this article all the (in)equalities between random variables are understood in almost sure sense.

Given sets $(W_i, G_i) \in \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$ for $i \in \{1, \cdots, n\}$, we are interested in a set of initial conditions $(t,x) \in \mathbb{S}$ such that there exists an admissible strategy $\boldsymbol{u} \in \mathcal{U}_t$ steering the process $X_\cdot^{t,x;\boldsymbol{u}}$ through the sets $(W_i)_{i=1}^n$ while visiting $(G_i)_{i=1}^n$ in a pre-assigned order. In fact, W_i and G_i stand for *"Way"* and *"Goal"* respectively. One may pose this objective from different perspectives based on different time scheduling for the excursions between the sets. We formally introduce some of these notions which will be addressed throughout this article.

Definition II.1 (Motion-Planning Events)**.** *Consider a fixed initial condition $(t,x) \in \mathbb{S}$ and admissible policy $\boldsymbol{u} \in \mathcal{U}_t$. Given a sequence of pairs $(W_i, G_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$ and horizon times $(T_i)_{i=1}^n \subset [t, T]$, we introduce the following **motion-planning events**:*

$$\left\{ X_\cdot^{t,x;\boldsymbol{u}} \models \left[(W_1 \rightsquigarrow G_1) \circ \cdots \circ (W_n \rightsquigarrow G_n) \right]_{\leq T} \right\} := \tag{1a}$$
$$\left\{ \exists (s_i)_{i=1}^n \subset [t,T] \mid X_{s_i}^{t,x;\boldsymbol{u}} \in G_i \quad and \right.$$
$$\left. X_r^{t,x;\boldsymbol{u}} \in W_i \setminus G_i, \ \forall r \in [s_{i-1}, s_i[, \quad \forall i \leq n \right\},$$

$$\left\{ X_\cdot^{t,x;\boldsymbol{u}} \models (W_1 \xrightarrow{T_1} G_1) \circ \cdots \circ (W_n \xrightarrow{T_n} G_n) \right\} := \tag{1b}$$
$$\left\{ X_{T_i}^{t,x;\boldsymbol{u}} \in G_i \ and \ X_r^{t,x;\boldsymbol{u}} \in W_i, \quad \forall r \in [T_{i-1}, T_i], \quad \forall i \leq n \right\},$$

where in the above definitions $s_0 = T_0 := t$.

The set in (1a), roughly speaking, contains those events that the trajectory $X_\cdot^{t,x;\boldsymbol{u}}$, initialized at $(t,x) \in \mathbb{S}$ and controlled via $\boldsymbol{u} \in \mathcal{U}_t$, succeeds in visiting the sets $(G_i)_{i=1}^n$ in a certain order, while the entire duration between the two visits to G_{i-1} and G_i is spent in W_i, all within the time horizon T. In other words, the journey from G_{i-1} to the next destination G_i must belong to the way W_i for all i. Figure 1(a) depicts a sample path that successfully contributes to the first three phases of the excursion in the sense of (1a).

In the case of (1b), the set of paths is more restricted in comparison to (1a). Indeed, not only is the trajectory confined to the ways W_i, but also there is a time schedule $(T_i)_{i=1}^n$ that a priori forces the process to be at the goal sets G_i at the specific times $(T_i)_{i=1}^n$. Figure 1(b) demonstrates one sample path in which the first three phases of the

---

[1]That is, processes with paths that are right continuous with left limits.
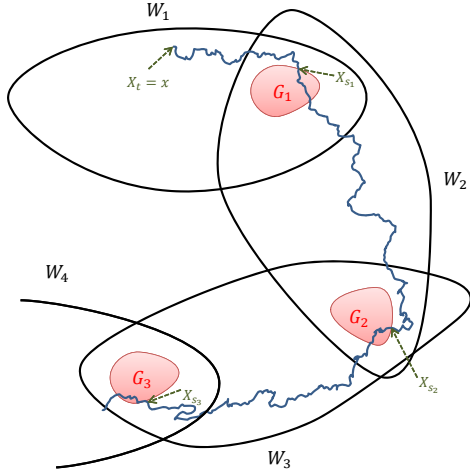
(a) A sample path satisfying the first three phases of the specification in the sense of (1a)
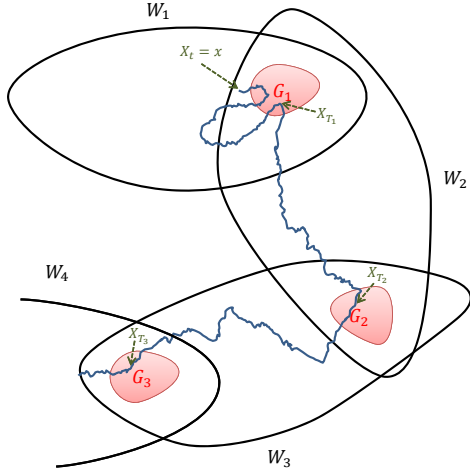
(b) A sample path satisfying the first three phases of the specification in the sense of (1b)

Fig. 1. Sample paths of the process $X^{t,x;\boldsymbol{u}}_{\cdot}$ for a fix policy $\boldsymbol{u} \in \mathcal{U}_t$

excursion are successfully fulfilled.

Note that once a trajectory belonging ot the set in (1a) visits G_i for the first time, it is required to remain in the way W_{i+1} until the next goal G_{i+1} is reached, whereas a trajectory belonging to the set in definition (1b) may visit the destination G_i several times, while staying in W_i, until the intermediate time schedule T_i. The only requirement, in contrast to (1a), is to confine the trajectory to be at the goal G_i at the time T_i.

As an illustration, one can easily inspect that the successful sample path in Figure 1(b) indeed violates the requirements of the definition (1a) as it leaves W_2 after it visits the goal set G_1 for the first time. In other words, the definition (1a) changes the admissible way set W_i to W_{i+1} *immediately after* the trajectory visits the goal set G_i, while the definition (1b) only changes the admissible way set *only after* the intermediate time T_i irrespective of whether the trajectory

visits the goal set G_i prior to T_i.

From the technical standpoint, if the target set G_i is not closed, then it is not difficult to see that there could be some continuous transitions through the boundary of the goal G_i that are not admissible in view of the definition (1a) since the trajectory must reside in $W_i \setminus G_i$ for the whole interval $[s_{i-1}, s_i[$ and just hit the set G_i at the time s_i. Notice that we do not need to consider this issue for the set in definition (1b) since in this case the trajectory only visits the sets G_i at the specific times T_i while any continuous transition and maneuver inside the target sets G_i are allowed. In order to address the aforementioned issue, we may impose the following:

Assumption II.2. *The sets* $(G_i)^n_{i=1} \subset \mathfrak{B}(\mathbb{R}^d)$ *are closed.*

Having discussed about the properties of motion planning as detailed in Definition II.1, one may conclude that in general none of the definitions of (1) is more restrictive than the other. However, in the particular motion planning case where the family of the way sets $(W_i)^n_{i=1}$ is nested, i.e., $W_i \subseteq W_{i+1}$, one can see that the motion planning definition of (1b) actually imposes more constraints on the trajectories of the random process in light of the right continuity of $X^{t,x;\boldsymbol{u}}_{\cdot}$ and Assumption II.2. The following Fact formally addresses this issue.

Fact II.3. *Consider a family of set pairs* $(W_i, G_i)^n_{i=1} \subset \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$ *where* $(G_i)^n_{i=1}$ *satisfies Assumption II.2 and* $(W_i)^n_{i=1}$ *is nested, i.e.* $W_i \subseteq W_{i+1}$. *Then, for all initial condition* $(t, x) \in \mathbb{S}$, *intermediate times* $(T_i)^n_{i=1} \subset [t, T]$, *and control policy* $\boldsymbol{u} \in \mathcal{U}_t$, *it holds that*

$$\left\{ X^{t,x;\boldsymbol{u}}_{\cdot} \models (W_1 \xrightarrow{T_1} G_1) \circ \cdots \circ (W_n \xrightarrow{T_n} G_n) \right\}$$
$$\subseteq \left\{ X^{t,x;\boldsymbol{u}}_{\cdot} \models \left[(W_1 \rightsquigarrow G_1) \circ \cdots \circ (W_n \rightsquigarrow G_n) \right]_{\leq T} \right\}$$

Proof: Let ω be contained in the set defined in (1b). This means that $X^{t,x;\boldsymbol{u}}_{T_i}(\omega) \in G_i$, and for all $r \in [T_{i-1}, T_i]$ we have $X^{t,x;\boldsymbol{u}}_r(\omega) \in W_i$. Due to the right continuity of the sample paths and Assumption II.2, one can see that for all $i \in \{1, \cdots, n\}$ there exists $s_i \in [T_{i-1}, T_i]$ so that $X^{t,x;\boldsymbol{u}}_{s_i}(\omega) \in G_i$ while $X^{t,x;\boldsymbol{u}}_r(\omega) \in W_i \setminus G_i$ for all $r \in [T_{i-1}, s_i[$, where $T_0 := t$. Since the sets $(W_i)^n_{i=1}$ are nested, $X^{t,x;\boldsymbol{u}}_r(\omega) \in W_{i-1} \subseteq W_i$ for all $r \in [s_{i-1}, T_{i-1}]$. An induction argument quickly leads to

$$X^{t,x;\boldsymbol{u}}_{s_i}(\omega) \in G_i,$$
$$X^{t,x;\boldsymbol{u}}_r(\omega) \in W_i \setminus G_i, \quad \forall r \in [s_{i-1}, s_i[,$$

where $s_0 := t$. In fact one may introduce s_i as the first hitting time of the set G_i after time T_{i-1}.[2] This implies that ω is also contained in the set defined in (1a) and proves the assertion. ∎

Let us note that to satisfy the order of visiting the goal/target sets $(G_i)^n_{i=1}$, it only suffices to exclude the target sets $(G_i)^n_{i=k+1}$ from the set W_k. Therefore, the hypothesis in Fact II.3 concerning the nested way sets W_i does not really impose any restriction on the motion planning objectives.

---

[2]The formal definition of this stopping time and its application are considered later for Fact III.3.

Remark II.4. The motion planning scenarios for only two sets (W_1, G_1) are essentially the basic reachability maneuver that was studied in our earlier work [3]. The definition (1a) suggests the same Reach-Avoid problem as in [3, Definition 2.4], where the target and obstacle sets are G_1 and $\mathbb{R}^d \setminus W_1$ respectively. In this special case, the definition (1b) also follows the same concept as Reach-Avoid problem in [3, Definition 3.4] with the same target and obstacle sets.

Remark II.5. A particular case of the Definition II.1 is the following: $G_i := A_{i+1} \setminus A_i$ and $G_i := A_{i+1} \cap A_i$ for the definitions (1a) and (1b) respectively. Here the motion planning objective is to pass through n given sets $(A_i)_{i=1}^n$ in a certain order.

The events introduced in Definition II.1 depend, of course, on the control policy $\boldsymbol{u} \in \mathcal{U}$ and initial condition $(t, x) \in \mathbb{S}$. The central objective of this work is to determine the set of initial conditions $x \in \mathbb{R}^d$ such that there exists an admissible policy $\boldsymbol{u}$ where the probability of the above path-planning events is higher than a certain threshold. To this end, we formally introduce these sets as follows:

Definition II.6 (Motion-Planning Initial Set). *Consider a fixed initial time $t \in [0, T]$. Given a sequence of set pairs $(W_i, G_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$ and horizon times $(T_i)_{i=1}^n \subset [t, T]$, we define the following motion-planning initial sets:*

$$\mathrm{PP}\big(t, p; (W_i, G_i)_{i=1}^n, T\big) := \Big\{ x \in \mathbb{R}^d \mid \exists \boldsymbol{u} \in \mathcal{U}_t \ : \tag{2a}$$

$$\mathbb{P}\big\{ X_{\cdot}^{t,x;\boldsymbol{u}} \models \big[(W_1 \rightsquigarrow G_1) \circ \cdots \circ (W_n \rightsquigarrow G_n) \big]_{\leq T} \big\} > p \Big\},$$

$$\widetilde{\mathrm{PP}}\big(t, p; (W_i, G_i)_{i=1}^n, (T_i)_{i=1}^n\big) := \Big\{ x \in \mathbb{R}^d \mid \exists \boldsymbol{u} \in \mathcal{U}_t \ : \tag{2b}$$

$$\mathbb{P}\big\{ X_{\cdot}^{t,x;\boldsymbol{u}} \models (W_1 \xrightarrow{T_1} G_1) \circ \cdots \circ (W_n \xrightarrow{T_n} G_n) \big\} > p \Big\}.$$

III. CONNECTION TO STOCHASTIC OPTIMAL CONTROL

In this section we establish a connection from stochastic motion-planning initial sets PP and $\widetilde{\mathrm{PP}}$, defined in Definition II.6, and a class of stochastic optimal control problems involving stopping times. We introduce a sequence of random times that corresponds to the times that the process $X_{\cdot}^{t,x;\boldsymbol{u}}$ for the first time exits from the sequence of sets one after another in a certain order:

Definition III.1. *Given an initial condition $(t, x) \in \mathbb{S}$ and a sequence of measurable sets $(A_i)_{i=k}^n \subset \mathfrak{B}(\mathbb{R}^d)$, the sequence of random times $\big(\Theta_i^{A_{k:n}}\big)_{i=k}^n$ defined[3] by*

$$\Theta_i^{A_{k:n}}(t, x) := \inf \big\{ r \geq \Theta_{i-1}^{A_{k:n}}(t, x) \ : \ X_r^{t,x;\boldsymbol{u}} \notin A_i \big\}, \quad \Theta_{k-1}^{A_{k:n}}(t, x) := t,$$

*is called the **sequential exit-time** through the set A_k to A_n.*

Note that the sequential exit-time $\Theta_i^{A_{k:n}}$ depends on the control policy $\boldsymbol{u}$ in addition to the initial condition (t, x), but here and later in the sequel we shall suppress this dependence. For notational simplicity, we also drop (t, x) in the subsequent sections.

In Figure 2 a sample path of the process $X_{\cdot}^{t,x;\boldsymbol{u}}$ along with the sequential exit-times $(\Theta_i^{A_{k:3}})_{i=k}^n$ is depicted for different $k \in \{1, \cdots, 3\}$. Note that since the initial condition x does not belongs

to A_3, the first exit-time of the set A_3 is indeed the start time t, i.e., $\Theta_3^{A_{3:3}} = t$. Let us highlight the difference between stopping times $\Theta_2^{A_{1:3}}$ and $\Theta_2^{A_{2:3}}$. The former is the first exit-time of the set A_2 after the time that the process leaves A_1, whereas the latter is the first exit-time of the set A_2 from the very beginning.[4]
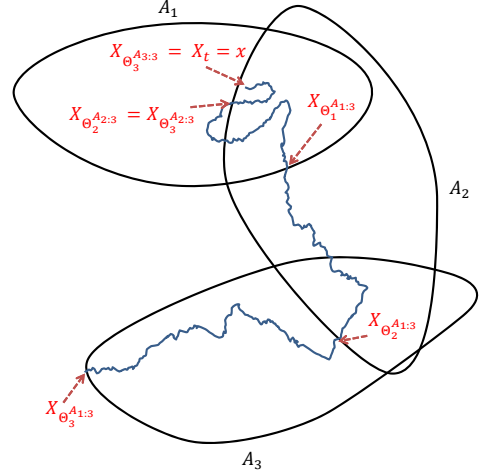

Fig. 2. Sequential exit-times of a sample path through the sets $(A_i)_{i=k}^3$ for different values of k
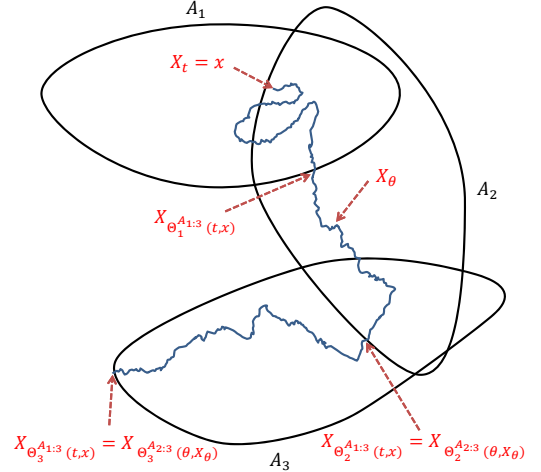

Fig. 3. Sequential exit-times corresponding to different initial conditions

Given a stopping time $\theta \in \mathcal{T}_{[t, T]}$, let $\omega \in \Omega$ be a sample realization such that the process $\big(X_s^{t,x;\boldsymbol{u}}(\omega) \big)_{t \leq s \leq \theta(\omega)}$ successfully visits the first sets $(A_j)_{j=1}^{i-1}$. For this ω, it then follows from Definition III.1 that the sequential stopping times $\Theta_j^{A_{1:n}}$ of visiting the sets $(A_j)_{j=i}^n$ starting at the initial condition (t, x) is the same as the sequential exit-times $\Theta_j^{A_{i:n}}$ of visiting the sets $(A_j)_{j=i}^n$ starting at $(\theta(\omega), X_{\theta(\omega)}^{t,x;\boldsymbol{u}})$. Figure 3 depicts one of these sample paths where $\Theta_1^{A_{1:3}}(\omega) \leq \theta(\omega) < \Theta_1^{A_{1:3}}(\omega)$. It is obvious from the figure that starting from an initial condition (t, x), the process views the exit-times of A_2 and A_3 after leaving A_1 as if it starts from the initial

---

[3]By convention, $\inf \emptyset = \infty$.

[4]In §V we shall see that these differences will lead to different definitions of value functions in order to derive a dynamic programming argument.

condition $(\theta, X_\theta^{t,x;\boldsymbol{u}})$ and proceeds to exit from the set A_2 and A_3 irrespective of the history before θ. Lemma III.2 formally presents this result and will be employed later to prove one of the main result of this article, i.e., Theorem IV.5.

Lemma III.2. *Consider a control policy $\boldsymbol{u} \in \mathcal{U}_t$ and initial condition $(t,x) \in \mathbb{S}$. Given a sequence of measurable sets $(A_i)_{i=k}^n \subset \mathfrak{B}(\mathbb{R}^d)$ and stopping time $\theta \in \mathcal{T}_{[t,T]}$, for all $k \in \{1, \cdots, n\}$ and $j \geq i \geq k$ we have*

$$\Theta_j^{A_{k:n}}(t,x) = \Theta_j^{A_{i:n}}\big(\theta, X_\theta^{t,x;\boldsymbol{u}}\big) \quad on \quad \big\{\Theta_{i-1}^{A_{k:n}}(t,x) \leq \theta < \Theta_i^{A_{k:n}}(t,x)\big\}$$

We need to establish that the sequential stopping-times are well-defined, hence:

Fact III.3 (Measurability)**.** *Consider a sequence of measurable sets $(A_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$ and initial condition $(t,x) \in \mathbb{S}$. The sequential exit-time $\Theta_i^{A_{1:n}}(t,x)$ is an $\mathbb{F}_t$-stopping time for all $i \in \{1, \cdots, n\}$, i.e., $\big\{\Theta_i^{A_{1:n}}(t,x) \leq s\big\} \in \mathcal{F}_{t,s}$ for all $s \geq 0$.*

Proof: Let τ_A be the first exit-time from the set A_i:

$$\tau_{A_i}(t,x) := \inf\{s \geq 0 \ : \ X_{t+s}^{t,x;\boldsymbol{u}} \notin A_i\}. \tag{3}$$

We know that τ_A is an $\mathbb{F}_t$-stopping time [7, Theorem 1.6, Chapter 2]. Let $\omega(\cdot) \mapsto \vartheta_s\big(\omega(\cdot)\big) := \omega(s+\cdot)$ be the time-shift operator. From the definition it follows that for all $i \geq 0$

$$\Theta_{i+1}^{A_{1:n}} = \Theta_i^{A_{1:n}} + \tau_{A_i} \circ \vartheta_{\Theta_i^{A_{1:n}}}.$$

Now the assertion follows directly in light of the measurability of the mapping ϑ and right continuity of the filtration $\mathbb{F}_t$.[5] ∎

For technical reasons we stipulate that the way sets satisfy the following:

Assumption III.4. *The sets $(W_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$ are open.*

Under Assumption III.4, the following fact is an immediate consequence of right continuity of the stochastic process $X_\cdot^{t,x;\boldsymbol{u}}$:

Fact III.5. *Fix a control policy $\boldsymbol{u} \in \mathcal{U}_t$ and an initial condition $(t,x) \in \mathbb{S}$. Let measurable sets $(W_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$ be given, and suppose that Assumption III.4 holds for $(W_i)_{i=1}^n$. Then, for all $i \in \{1, \cdots, n\}$*

$$X_{\Theta_i^{W_{1:n}}}^{t,x;\boldsymbol{u}} \notin W_i, \qquad on \quad \big\{\Theta_i^{W_{1:n}} < \infty\big\},$$

*where $\big(\Theta_i^{W_{1:n}}\big)_{i=1}^n$ are the sequential exit-times introduced as in Definition III.1.*

Given $(W_i, G_i, T_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d) \times [t,T]$, we introduce two value functions $V, \widetilde{V} : \mathbb{S} \to [0,1]$ defined by

$$V(t,x) := \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\left[\prod_{i=1}^n \mathbb{1}_{G_i}\big(X_{\eta_i}^{t,x;\boldsymbol{u}}\big)\right], \quad \begin{cases} \eta_i := \Theta_i^{B_{1:n}} \wedge T, \\ B_i := W_i \setminus G_i, \end{cases} \tag{4a}$$

$$\widetilde{V}(t,x) := \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\left[\prod_{i=1}^n \mathbb{1}_{G_i \cap W_i}\big(X_{\widetilde{\eta}_i}^{t,x;\boldsymbol{u}}\big)\right], \quad \widetilde{\eta}_i := \Theta_i^{W_{1:n}} \wedge T_i, \tag{4b}$$

[5]See, for instance, [7, Proposition 1.4, Chapter 2] for more details in this regard.

where $\Theta_i^{W_{1:n}}, \Theta_i^{B_{1:n}}$ are the sequential exit-times in the sense of Definition III.1. Figure 4(a) and 4(b) illustrate the sequential exit-times corresponding to the sets B_i and W_i, respectively; the sample trajectories are the same as Figure 1. The main result of this section, Proposition III.6 below, establishes a connection from the sets PP, $\widetilde{\text{PP}}$ and superlevel sets of the value functions V and $\widetilde{V}$.
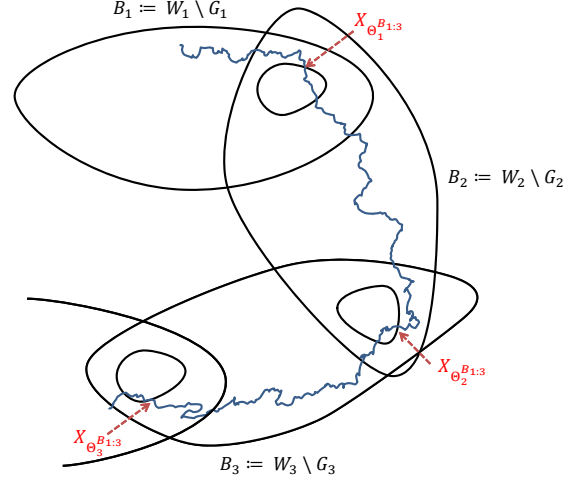
(a) Sequential exit-times related to motion-planning event (1a)
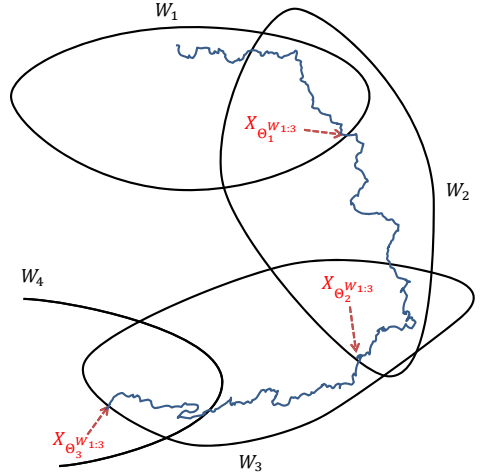
(b) Sequential exit-times related to motion-planning event (1b)

Fig. 4. Sequential exit-times corresponding to different motion-planning events as introduced in (1)

Proposition III.6. *Fix a probability level $p \in [0,1]$, a sequence of set pairs $(W_i, G_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$, an initial time $t \in [0,T]$, and intermediate times $(T_i)_{i=1}^n \subset [t,T]$. Then*

$$\text{PP}\big(t,p;(W_i,G_i)_{i=0}^n, T\big) = \big\{x \in \mathbb{R}^d \mid V(t,x) > p\big\}. \tag{5}$$

Moreover, suppose Assumption III.4 holds. Then,

$$\widetilde{\text{PP}}\big(t,p;(W_i,G_i)_{i=0}^n, (T_i)_{i=1}^n\big) = \big\{x \in \mathbb{R}^d \mid \widetilde{V}(t,x) > p\big\}, \tag{6}$$

where the value functions V and $\widetilde{V}$ are as defined in (4).

Proof: See [5, Appendix A]. ∎

Remark III.7 (Mixed Motion-Planning Events). In this section we focus on two sets of events as introduced in Definition II.1, and will continue doing so for our subsequent results. However, it is of interest to consider an event that consists of a mixture of the events in (1), e.g., $(W_1 \rightsquigarrow G_1)_{\leq T_1} \circ (W_2 \xrightarrow{T_2} G_2)$. One can observe that essentially the same analytical techniques as the ones proposed here can be employed to address these mixed motion planning objectives, and establish a connection to a class of optimal control problems with some appropriate sequential stopping times. We shall provide an example of this nature in §VI.

IV. DYNAMIC PROGRAMMING PRINCIPLE

The objective of this section is to derive a DPP for the value functions V and $\widetilde{V}$ introduced in (4). We proceed with a more abstract setting for transparent elucidation in the following fashion: Let $(T_i)_{i=1}^n \subset [0, T]$ be a sequence of times, $(A_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^n)$ be a family of open sets, and payoff functions $\ell_i : \mathbb{R}^r \to \mathbb{R}$ that are measurable and bounded, $i = 1, \ldots, n$. We define the sequence of value functions $V_k : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $k = 1, \ldots, n$,

$$V_k(t, x) := \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\Big[\prod_{i=k}^n \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big], \tag{7}$$

$$\tau_i^k(t, x) := \Theta_i^{A_{k:n}}(t, x) \wedge T_i, \quad i \in \{k, \cdots, n\},$$

where the stopping times $(\Theta_i^{W_{k:n}})_{i=k}^n$ are sequential exit-times in the sense of Definition III.1. Notice that the sequential exit-times of the value function V_k correspond to an excursion through the sets $(A_i)_{i=k}^n$ irrespective of the first $(k-1)$ sets. It is straightforward to observe that the value functions V and $\widetilde{V}$ in (4) are particular cases of the value function V_1 defined as in (7) for an appropriate selection of the sets $(A_i)_{i=1}^n$, functions $(\ell_i)_{i=1}^n$, and intermediate times $(T_i)_{i=1}^n$.

To state the main result of this section, Theorem IV.5 below, some technical definitions and assumptions concerning the stochastic processes $X^{t,x;\boldsymbol{u}}$, admissible strategies $\mathcal{U}_t$, and the payoff functions ℓ_i, are needed: Let $\mathbb{A}$ be a metric space, and let $f : \mathbb{A} \longrightarrow \mathbb{R}$ be a function. The lower and upper semicontinuous envelopes, respectively, of f are defined as:

$$f_*(x) := \liminf_{x' \to x} f(x') \qquad f^*(x) := \limsup_{x' \to x} f(x').$$

We denote by $\mathrm{USC}(\mathbb{A})$ and $\mathrm{LSC}(\mathbb{A})$ the collection of all upper-semicontinuous and lower-semicontinuous functions from $\mathbb{A}$ to $\mathbb{R}$, respectively.

Assumption IV.1. *We stipulate the following assumptions on*

a. ***Admissible control policies:***

   i. Non-anticipative policy: *Given a control set $\mathbb{U} \subset \mathbb{R}^{d_u}$, any $\boldsymbol{u} \in \mathcal{U}_t$ is $\mathbb{U}$-valued $\mathbb{F}_t$-adapted stochastic process;*

   ii. Stability under concatenation: *The set of admissible control policies at time t, $\mathcal{U}_t$, is closed under concatenation. That is, for all $\boldsymbol{u_1}, \boldsymbol{u_2} \in \mathcal{U}_t$ and stopping time $\theta \in \mathcal{T}_{[t,T]}$, it holds that*

$$\mathbb{1}_{[t,\theta]}\boldsymbol{u_1} + \mathbb{1}_{]\theta,T]}\boldsymbol{u_2} \in \mathcal{U}_t;$$

b. ***Stochastic process*** $X_{\cdot}^{t,x;\boldsymbol{u}}$***:***

   i. Causality: *For all initial conditions $(t, x) \in \mathbb{S}$, any control policies $\boldsymbol{u_1}, \boldsymbol{u_2} \in \mathcal{U}_t$, and stopping time $\theta \in \mathcal{T}_{[t,T]}$, it holds that*

$$\mathbb{1}_{[t,\theta]}\boldsymbol{u_1} = \mathbb{1}_{[t,\theta]}\boldsymbol{u_2} \implies \mathbb{1}_{[t,\theta]}X_{\cdot}^{t,x;\boldsymbol{u_1}} = \mathbb{1}_{[t,\theta]}X_{\cdot}^{t,x;\boldsymbol{u_2}}$$

   ii. Strong Markov property: *Given initial condition $(t, x) \in \mathbb{S}$, control policy $\boldsymbol{u} \in \mathcal{U}_t$, and stopping time $\theta \in \mathcal{T}_{[t,T]}$, for all bounded measurable functions $\ell : \mathbb{R}^n \to \mathbb{R}$ and $s \geq 0$ it holds that*

$$\mathbb{E}\Big[\ell\big(X_{\theta+s}^{t,x;\boldsymbol{u}}\big)\Big|\mathcal{F}_\theta\Big] = \mathbb{E}\Big[\ell\big(X_{\theta+s}^{t,x;\boldsymbol{u}}\big)\Big|X_\theta^{t,x;\boldsymbol{u}}\Big] \qquad \mathbb{P}\text{-}a.s.;$$

   iii. Continuity of exit-time: *Given initial condition $(t_0, x_0) \in \mathbb{S}$ and control policy $\boldsymbol{u} \in \mathcal{U}_t$, for all $k \in \{1, \cdots, n\}$ and $i \in \{k, \cdots, n\}$ the stochastic mapping $(t, x) \mapsto X_{\tau_i^k(t,x)}^{t,x;\boldsymbol{u}}$ is $\mathbb{P}$-a.s. continuous at the point (t_0, x_0) where the stopping time τ_i^k is defined as in (7);*

c. ***Payoff functions:***
   $(\ell_i)_{i=1}^n$ *are lower semicontinuous, i.e., $\ell_i \in \mathrm{LSC}(\mathbb{R}^d)$ for all $i \leq n$.*

Remark IV.2. Some remarks on the above assumptions are in order:

○ Assumption IV.1.a.i. implies that admissible strategies $\boldsymbol{u} \in \mathcal{U}_t$ take action at time t independent of future information arriving at $s > t$. This is known as a *non-anticipative* strategy [8], and is a standard assumption.

○ Assumption IV.1.a.ii. is also standard for dynamic programming arguments and it holds for a large class of admissible strategies, e.g., progressively measurable processes.

○ Assumption IV.1.b. imposes three constraints on the process $X_{\cdot}^{t,x;\boldsymbol{u}}$ defined on the prescribed probability space: i) causality of the solution processes for a given admissible policy ii) strong Markov property iii) continuity of exit-time. The causality property is always satisfied in practical applications; uniqueness of the solution process $X_{\cdot}^{t,x;\boldsymbol{u}}$ under any admissible control process $\boldsymbol{u}$ guarantees it. The class of Markovian processes is fairly large; for instance, it contains the solution of SDEs under some mild assumptions on the drift and diffusion terms [9, Theorem 2.9.4]. The almost sure continuity of the exit-time with respect to the initial condition of the process is the only restrictive. Note that this condition does not always hold even for deterministic processes with continuous trajectories. One may need to impose conditions on the process and possibly the sets involved in motion planning in order to satisfy continuity of the mapping $(t, x) \mapsto X_{\tau_i^k(t,x)}^{t,x;\boldsymbol{u}}$ at the given initial condition with probability one. We shall elaborate on this issue and its ramifications to a class of diffusions in §V.

We present two Lemmas preparatory to our main result. For $k = 1, \ldots, n$, we define the function $J_k : \mathbb{S} \times \mathcal{U} \to \mathbb{R}$ as:

$$J_k(t, x; \boldsymbol{u}) := \mathbb{E}\Big[\prod_{i=k}^n \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big], \tag{8}$$

where $\left(\tau_i^k\right)_{i=k}^n$ are as defined in (7).

Lemma IV.3. *Consider the value function V_k and J_k as defined in (7) and (8) respectively. Then, for any $k \in \{1, \cdots, n\}$*

$$V_k(t,x) = \sup_{\boldsymbol{u} \in \mathcal{U}} J_k(t,x;\boldsymbol{u}), \qquad \forall(t,x) \in \mathbb{S}.$$

Proof: Note that all the admissible policies $\boldsymbol{u} \in \mathcal{U}_t$ are also contained in $\mathcal{U}$. Therefore, the inequality " $\geq$ " between the left- and right-hand sides is immediate. In order to show the reverse inequality, observe that any control policy $\boldsymbol{u} \in \mathcal{U}$ is a measurable function of the process $(z_s)_{s \geq 0}$. For any fixed $(z_s)_{0 \leq s \leq t}$ it holds that $\boldsymbol{u} := \boldsymbol{u}\big((z_s)_{0 \leq s \leq T}\big) = \widetilde{\boldsymbol{u}}\big((z_s)_{0 \leq s \leq t}, (z_s - z_t)_{t \geq s \geq T}\big)$, where $\widetilde{\boldsymbol{u}}$ can be viewed as a policy independent of $\mathcal{F}_t$ due to independent increments of z. By [4, Remark 5.2] the inequality " $\leq$ " also follows, and the assertion of the lemma follows.[6] ∎

Lemma IV.4. *Under Assumptions IV.1.b.iii. and IV.1.c., the function $\mathbb{S} \ni (t,x) \mapsto J_k(t,x;\boldsymbol{u}) \in \mathbb{R}$ is lower semicontinuous for all $k \in \{1, \cdots, n\}$ and control policy $\boldsymbol{u} \in \mathcal{U}$.*

Proof: Fix $k \in \{1, \ldots, n\}$. It is obvious that the function J_k is uniformly bounded since ℓ_k are. Therefore,

$$\liminf_{(s,y) \to (t,x)} J_k(s,y;\boldsymbol{u})$$
$$= \liminf_{(s,y) \to (t,x)} \mathbb{E}\Big[\prod_{i=k}^n \ell_i\big(X_{\tau_i^k(s,y)}^{s,y;\boldsymbol{u}}\big)\Big] \geq \mathbb{E}\Big[\liminf_{(s,y) \to (t,x)} \prod_{i=k}^n \ell_i\big(X_{\tau_i^k(s,y)}^{s,y;\boldsymbol{u}}\big)\Big]$$
$$\geq \mathbb{E}\Big[\prod_{i=k}^n \liminf_{(s,y) \to (t,x)} \ell_i\big(X_{\tau_i^k(s,y)}^{s,y;\boldsymbol{u}}\big)\Big] \geq \mathbb{E}\Big[\prod_{i=k}^n \ell_i\big(X_{\tau_i^k(s,y)}^{t,x;\boldsymbol{u}}\big)\Big]$$
$$= J_k(t,x;\boldsymbol{u}),$$

where the inequality in the first line follows from the Fatou's lemma, and the second inequality in the second line is a direct consequence of Assumptions IV.1.b.iii. and IV.1.c. ∎

The following Theorem, the main result of this section, establishes a dynamic programming argument for the value function V_k in terms of the "successor" value functions $(V_j)_{j=k+1}^n$, all defined as in (7). For ease of notation we shall introduce deterministic times $\tau_{k-1}^k, \tau_{n+1}^k$, and a trivial constant value function V_{n+1}.

Theorem IV.5. *Consider the value functions $(V_j)_{j=1}^n$ and the sequential stopping times $(\tau_j^k)_{j=k}^n$ introduced in (7). Under Assumption IV.1, for all initial conditions $(t,x) \in \mathbb{S}$, stopping times $\theta \in \mathcal{T}_{[t,T]}$, $k \in \{1, \cdots, n\}$ we have*

$$V_k(t,x) \leq \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\Big[\sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} V_j^*\big(\theta, X_\theta^{t,x;\boldsymbol{u}}\big) \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big], \tag{9a}$$

$$V_k(t,x) \geq \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\Big[\sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} V_{j_*}\big(\theta, X_\theta^{t,x;\boldsymbol{u}}\big) \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big], \tag{9b}$$

where V_j^* and V_{j_*} are upper and lower semicontinuous envelopes of the value function V_j, respectively, $\tau_{k-1}^k := t$, $V_{n+1} \equiv 1$, and

---

[6]For a similar result in the context of SDEs, see [9, Theorem 3.1.7, p. 132].

τ_{n+1}^k *can be chosen any constant time strictly greater than T, say $\tau_{n+1}^k := T + 1$.*

Proof: The proof extends the main result of our earlier work [3, Theorem 4.7] on the so-called "reach-avoid" motion planning maneuver. Based on the tower property of conditional expectation [10, Theorem 5.1], we have

$$\mathbb{E}\Big[\prod_{i=k}^n \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big|\mathcal{F}_\theta\Big] \tag{10a}$$
$$= \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \mathbb{E}\Big[\prod_{i=j}^n \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big|\mathcal{F}_\theta\Big] \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)$$
$$= \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} J_j\big(\theta, X_\theta^{t,x;\boldsymbol{u}};\boldsymbol{u}\big) \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big) \tag{10b}$$
$$\leq \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} V_j\big(\theta, X_\theta^{t,x;\boldsymbol{u}}\big) \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big) \tag{10c}$$

where (10b) and (10c) follow from Assumption IV.1.b.ii., and Lemma III.2, respectively. In the light of the tower property of conditional expectation [10, Theorem 5.1], arbitrariness of $\boldsymbol{u} \in \mathcal{U}_t$, and obvious inequality $V_j \leq V_j^*$, we arrive at (9a).

To prove (9b), we define a sequence of uniformly bounded upper semicontinuous functions $(\phi_j)_{j=k}^n \subset \text{USC}(\mathbb{S})$ such that $\phi_j \leq V_{j_*}$ on $\mathbb{S}$. Mimicking the ideas in the proof of our earlier work [3, Theorem 4.7], it is possible to establish that given $\epsilon > 0$ for all $j \in \{k, \cdots, n\}$ there exists an admissible control policy $\boldsymbol{u}_j^\epsilon$ such that

$$\phi_j(t,x) - 3\epsilon \leq J_j(t,x;\boldsymbol{u}_j^\epsilon) \qquad \forall(t,x) \in \mathbb{S}. \tag{11}$$

Let us fix $\boldsymbol{u} \in \mathcal{U}_t$ and $\epsilon > 0$, and define

$$\boldsymbol{v}^\epsilon := \mathbb{1}_{[t,\theta]}\boldsymbol{u} + \mathbb{1}_{]\theta,T]} \sum_{j=k}^n \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \boldsymbol{u}_j^\epsilon, \tag{12}$$

where $\boldsymbol{u}_j^\epsilon$ satisfies (11). Notice that Assumption IV.1.a. guarantees $\boldsymbol{v}^\epsilon \in \mathcal{U}_t$. In accordance with (11), (12) and in light of Assumption IV.1.i., one obtains the tower property from (10b) as follows:

$$V_k(t,x) \geq J_k(t,x;\boldsymbol{v}^\epsilon) = \mathbb{E}\Big[\mathbb{E}\Big[\prod_{i=k}^n \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{v}^\epsilon}\big)\Big|\mathcal{F}_\theta\Big]\Big]$$
$$= \mathbb{E}\Big[\sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} J_j\big(\theta, X_\theta^{t,x;\boldsymbol{u}};\boldsymbol{u}_j^\epsilon\big) \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big]$$
$$= \mathbb{E}\Big[\sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \big(\phi_j\big(\theta, X_\theta^{t,x;\boldsymbol{u}}\big) - 3\epsilon\big) \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big].$$

Since V_{j_*} is lower semicontinuous, in view of [11, Lemma 3.5] one may pick a sequence of increasing continuous functions $(\phi_j^m)_{m \in \mathbb{N}}$ that converges point-wise to V_j. By boundedness of $(\ell_j)_{i=1}^n$ and the dominated convergence Theorem, we get

$$V_k(t,x) \geq \mathbb{E}\Big[\sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \big(V_{j_*}\big(\theta, X_\theta^{t,x;\boldsymbol{u}}\big) - 3\epsilon\big) \prod_{i=k}^{j-1} \ell_i\big(X_{\tau_i^k}^{t,x;\boldsymbol{u}}\big)\Big]$$

Since $\boldsymbol{u} \in \mathcal{U}_t$ and $\epsilon > 0$ are arbitrary, this implies the assertion (9b). ∎

Remark IV.6. Theorem IV.5 introduces DPP's in a weaker sense than the standard DPP in stochastic optimal control problems [12]. Namely, one does not need to verify the measurability of the value functions V_k in (4) so as to apply the DPP's. Notice that in general this measurability issue is non-trivial due to the supremum operation running over possibly uncountably many policies.

V. APPLICATIONS TO CONTROLLED DIFFUSIONS

The objective of this section is to demonstrate how the weak DPP derived in §IV adapts to the context of controlled diffusion processes. This application results in a series of Hamilton-Javobi-Bellman PDE's, where each PDE is understood in the discontinuous viscosity sense with some boundary conditions both in viscosity and Dirichlet (pointwise) senses. To this end, we shall first introduce formally the standard probability space setup for SDEs, then proceed with some preliminaries so as to pave the ground for the required Assumptions IV.1 to hold. The section consists of subsections concerning PDE derivation and boundary conditions along with further discussions so as to deploy the existing PDE solvers to numerically compute our PDE characterization.

Let Ω be $\mathcal{C}\left([0,T],\mathbb{R}^{z_d}\right)$, the set of continuous functions from $[0,T]$ into $\mathbb{R}^{z_d}$, and let $(z_t)_{t\geq 0}$ be the canonical process, i.e., $z_t(\omega) := \omega_t$. We consider $\mathbb{P}$ as the Wiener measure on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F})$, where $\mathbb{F}$ is the smallest right continuous filtration on Ω such that the process $(z_t)_{t\geq 0}$ is adapted to. Let us recall that $\mathbb{F}_t := (\mathcal{F}_{t,s})_{s\geq 0}$ is the auxiliary subfiltration defined as $\mathcal{F}_{t,s} := \sigma(z_r - z_t, t \leq r \leq t \vee s)$

Let $\mathbb{U} \subset \mathbb{R}^{d_u}$ be a control set, and $\mathcal{U}_t$ denote the set of all $\mathbb{F}_t$-progressively measurable mappings into $\mathbb{U}$. For every $\boldsymbol{u} = (u_t)_{t\geq 0}$ we consider the following $\mathbb{R}^d$-valued SDE[7]:

$$\mathrm{d}X_s = f(X_s, u_s)\,\mathrm{d}s + \sigma(X_s, u_s)\,\mathrm{d}W_s, \qquad X_t = x, \quad s \geq t, \quad (13)$$

where $f : \mathbb{R}^d \times \mathbb{U} \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^d \times \mathbb{U} \to \mathbb{R}^{d\times d_z}$. are measurable maps, and $W_s := z_s$ is the canonical process.

Assumption V.1. *We stipulate that*

a. $\mathbb{U} \subset \mathbb{R}^m$ *is compact;*

b. *f and σ are continuous and Lipschitz in its first argument uniformly with respect to the second;*

c. *The diffusion term σ of the SDE (13) is uniformly non-degenerate, i.e., there exists $\delta > 0$ such that for all $x \in \mathbb{R}^d$ and $u \in \mathbb{U}$, $\|\sigma\sigma^\top\| > \delta$.*

It is well-known [8] that under Assumptions V.1.a. and V.1.b. there exits a unique strong solution to the SDE (13); let us denote it by $\left(X_s^{t,x;\boldsymbol{u}}\right)_{s\geq t}$. For future notational simplicity, we slightly modify the definition of $X_s^{t,x;\boldsymbol{u}}$, and extend it to the whole interval $[0,T]$ where $X_s^{t,x;\boldsymbol{u}} := x$ for all s in $[0,t]$.

In addition to Assumptions V.1 on the SDE (13), we impose the following assumption on the motion planning sets that allows us to

[7]We slightly abuse notation and earlier used σ for the sigma algebra as well. However, it will be always clear from the context to which σ we refer.

(a) Exterior cone condition holds at every point of the boundary.

(b) exterior cone condition fails at the point x—the only possible exterior cone at x is a line.
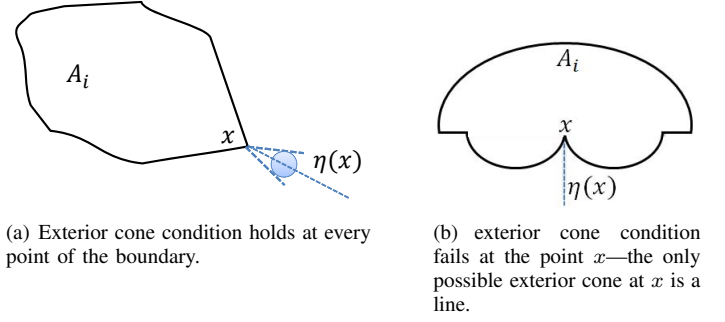
Fig. 5.   Exterior cone condition of the boundary

guarantee the continuity of sequential exit-times as required for the DPP obtained in the preceding section.

Assumption V.2 (Exterior Cone Condition)**.** *The open sets $(A_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$ satisfy the following condition: for every $i \in \{1, \cdots, n\}$, there are positive constants h, r an $\mathbb{R}^d$-value bounded map $\eta : A_i^c \to \mathbb{R}^n$ such that*

$$\mathrm{B}_{rt}(x + \eta(x)t) \subset A_i^c \qquad \text{for all } x \in A_i^c \text{ and } t \in (0, h]$$

where $\mathrm{B}_r(x)$ denotes an open ball centered at x and radius r and A_i^c stands for the complement of the set A_i.

Remark V.3. If the set A_i is bounded and its boundary ∂A_i is smooth, then Assumption V.2 holds. Furthermore, boundaries with corners may also satisfy Assumption V.2; Figure 5 depicts two different examples.

A. Sequential Partial Differential Equations

This subsection establishes a connection between the DPP introduced in Theorem IV.5 and a sequence of PDEs, all of the latter are meant in the sense of discontinuous viscosity solutions; for the general theory of viscosity solutions we refer to [13] and [12]. For numerical solutions to these PDEs, one also needs some boundary conditions, and that will be the objective of the next subsection.

To apply the proposed DPP, one has to make sure that Assumptions IV.1 are satisfied. As pointed out in Remark IV.2, the only nontrivial assumption in the context of SDEs is Assumption IV.1.b.iii. The following proposition addresses this issue, and allows us to employ the DPP of Theorem IV.5 for the main result of this subsection.

Proposition V.4. *Consider the SDE (13) where Assumptions V.1 holds. Suppose that the open sets $(A_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$ satisfy the exterior cone condition in Assumption V.2. Let $\left(\Theta_i^{A_{1:n}}\right)_{i=1}^n$ be the respective sequential exit-times as defined in Definition III.1. Given intermediate times $(T_i)_{i=1}^n$ and control policy $\boldsymbol{u} \in \mathcal{U}_t$, for any $i \in \{1,\cdots, n\}$, initial condition $(t, x) \in \mathbb{S}$, and sequence of initial conditions $(t_m, x_m) \to (t, x)$, we have*

$$\lim_{m\to\infty} \tau_i(t_m, x_m) = \tau_i(t, x) \quad \mathbb{P}\text{-a.s.}, \qquad \tau_i(t,x) := \Theta_i^{A_{1:n}}(t,x) \wedge T_i.$$

Moreover, one can show that the above result readily leads to the continuity of the stochastic mapping $(t,x) \mapsto X_{\tau_i(t,x)}^{t,x;\boldsymbol{u}}$ with probability one, i.e., $\lim_{m\to\infty} X_{\tau_i(t_m,x_m)}^{t_m,x_m;\boldsymbol{u}} = X_{\tau_i(t,x)}^{t,x;\boldsymbol{u}}$ $\mathbb{P}$-a.s. for all i.

Proof: See [5, Appendix B]. ∎

Definition V.5 (Dynkin Operator)**.** *Given* $u \in \mathbb{U}$*, we denote by* $\mathcal{L}^u$ *the Dynkin operator (also known as the infinitesimal generator) associated to the controlled diffusion* (13) *as*

$$\mathcal{L}^u \Phi(t,x) := \partial_t \Phi(t,x) + f(x,u).\partial_x \Phi(t,x) + \frac{1}{2} Tr[\sigma\sigma^\top(x,u)\partial_x^2 \Phi(t,x)],$$

where Φ *is a real-valued function smooth on the interior of* $\mathbb{S}$*, with* $\partial_t \Phi$ *and* $\partial_x \Phi$ *denoting the partial derivatives with respect to* t *and* x *respectively, and* $\partial_x^2 \Phi$ *denoting the Hessian matrix with respect to* x*. We refer to [10, Theorem 17.23] for more details on the above differential operator.*

Theorem V.6 is the main result of this subsection, which provides a characterization of the value functions V_k in terms of Dynkin operator in Definition V.5 in the interior of the set of interest, i.e., $[0, T_k[\times A_k$. This result is a direct consequence of the DPP in Theorem IV.5, and Itô's formula; for similar technique we refer to the proof of [3, Theorem 4.10].

Theorem V.6. *Consider the system* (13)*, and suppose that Assumptions V.1 hold. Let the value functions* $V_k : \mathbb{S} \to \mathbb{R}^d$*,* $k = 1, \ldots, n$ *be as defined in* (7)*, where the sets* $(A_i)_{i=1}^n$ *satisfy the Assumption V.2, and the payoff functions* $(\ell_i)_{i=1}^n$ *are all lower semicontinuous. Then,*

○ *the lower semicontinuous envelope of* V_k *is a viscosity supersolution of*

$$-\sup_{u \in \mathbb{U}} \mathcal{L}^u V_{k*}(t,x) \geq 0 \qquad on \quad [0, T_k[\times A_k;$$

○ *the upper semicontinuous envelope of* V *is a viscosity subsolution of*

$$-\sup_{u \in \mathbb{U}} \mathcal{L}^u V_k^*(t,x) \leq 0 \qquad on \quad [0, T_k[\times A_k.$$

Proof: We refer to [5, Appendix B] for a sketch of proof and [3, Theorem 4.10] for a detailed analysis of the same technique. ∎

B. Boundary Conditions

To numerically solve the PDE characterization of the previous part, one needs some boundary value conditions on the complement set of the PDE, which is addressed in the following:

Proposition V.7. *Suppose that the hypotheses of Theorem V.6 hold. Then the value functions* V_k *introduced in* (7) *satisfy the following boundary value conditions:*

$$V_k(t,x) = V_{k+1}(t,x)\ell_k(x) \qquad on \quad [0,T_k] \times A_k^c \bigcup \{T_k\} \times \mathbb{R}^d \quad (14a)$$

$$V_k^*(t,x) \leq V_{k+1}^*(t,x)\ell_k^*(x) \qquad on \quad [0,T_k] \times A_k^c \bigcup \{T_k\} \times \mathbb{R}^d \quad (14b)$$

Proof: See [5, Appendix B], along with a preparatory lemma. ∎

Proposition V.7 provides boundary conditions for the value function V_k not only in Dirichlet (pointwise) (14a) sense but also in the discontinuous viscosity sense (14b). Observe that the value functions

V_k are all lower semicontinuous since each of them is the supremum over family of lower semicontinuous functions J_k, see [5, Lemma B.1]. It then follows that the pointwise boundary condition (14a) indicates the other side of viscosity boundary condition of (14b), i.e., $V_{k*}(t,x) \geq V_{k+1*}(t,x)\ell_{k*}(x)$.

C. Discussion on Numerical Issues

The objective of the current section was to propose an alternative characterization of the value functions $V_k : \mathbb{S} \to \mathbb{R}$ in (7) for the case of controlled diffusion processes governed via the SDE (13). To this end, §V-A developed a PDE characterization of the value function V_k within the set $[0, T_k[\times A_k$ along with some boundary conditions in terms of the successor value function V_{k+1} provided in §V-B. Given value function V_{k+1}, to obtain the value function V_k, the only non-trivial set of initial conditions becomes $[0, T_k[\times A_k$ for which it is required to solve a certain PDE with some boundary (possibly both terminal and lateral) conditions. Since $V_{n+1} \equiv 1$, one can infer that Theorem V.6 and Proposition V.7 suggest a series of PDE equations for which the first one has known boundary condition ℓ_n, while the boundary conditions of the subsequent steps are determined by the solution of the preceding PDE step, i.e., V_{k+1} provides boundary conditions for the PDE corresponding to the value function V_k. Let us highlight once again that the basic motion planning maneuver involving only two sets is effectively the same as the first step of this series of PDEs and was studied in our earlier work [3].

Before proceeding to apply the PDE characterization of this section to obtain the desired motion-planning initial sets introduced in Definition II.6, we need to properly justify the following two concerns:

(i) On the one hand, for the definition (2a) we need to assume that the goal set G_i is closed so as to allow continuous transition into G_i; see Assumption II.2 and the discussion preceding it. On the other hand, in order to invoke the DPP argument of §IV and its consequent PDE in §V-A, we need to impose that the payoff functions $(\ell_i)_{i=1}^n$ are all lower semicontinuous; see Assumption IV.1.c. In the case of the value function V in (4a), this constraint results in $(G_i)_{i=1}^n$ all being open which is obviously in contradiction to our earlier assumption in accordance to (2a).

(ii) Due to the numerical issues, most of existing PDE solvers provide theoretical guarantees for *continuous* viscosity solutions, e.g., [14], whereas our characterization is indeed in discontinuous form. Therefore, it is a natural question whether we could employ the existing off-the-shelf toolbox to numerically calculate our desired value function.

Let us initially highlight the following points: Concerning (i) it should be mentioned that this contradiction is not applicable for the motion-planning initial set (2b) since the goal set G_i can be simply chosen to be open without confining the continuous transitions. Concerning (ii), we would like to stress that this discontinuous formulation is inevitable since the value functions defined in (4) are in general discontinuous, and any PDE approach has to rely on discontinuous versions.

Moreover, we next propose an ϵ-conservative but precise way of characterizing the motion-planning initial set of (2a) as well as justifying the numerical concern in (ii) to employ the existing off-the-shelf PDE solvers. Given $(W_i, G_i) \in \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$, let us construct a smaller goal set $G_i^\epsilon \subset G_i$ such that $G_i^\epsilon := \{x \in G_i \,|\, \mathrm{dist}(x, G_i^c) \geq \epsilon\}$.[8] For sufficiently small $\epsilon > 0$ one may observe that $W_i \setminus G_i^\epsilon$ satisfies Assumption V.2. Note that this is always possible if $W_i \setminus G_i$ satisfies Assumption V.2 since one can simply take $\epsilon < h/2$, where h is as defined in Assumption V.2. Figure 6 depicts this situation.
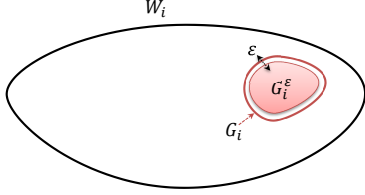

Fig. 6. Construction of the sets G_i^ϵ from G_i as described in §V-C

Formally we define the payoff function $\ell_i^\epsilon : \mathbb{R}^d \to \mathbb{R}$ as follows:

$$\ell_i^\epsilon(x) := \left(1 - \frac{\mathrm{dist}(x, G_\epsilon)}{\epsilon}\right) \vee 0.$$

Replacing the goal sets G_i^ϵ and payoff functions ℓ_i^ϵ in (4a), we arrive at the value functions

$$V^\epsilon(t,x) := \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\left[\prod_{i=1}^n \ell_i^\epsilon\left(X_{\eta_i^\epsilon}^{t,x;\boldsymbol{u}}\right)\right], \quad \begin{cases} \eta_i^\epsilon & := \Theta_i^{B_{1:n}^\epsilon} \wedge T, \\ B_i^\epsilon & := W_i \setminus G_i^\epsilon. \end{cases}$$

It is straightforward to inspect that $V^\epsilon \leq V$ since the goal sets are smaller with respect to the actual goal sets G_i. Moreover, one can show that $V(t,x) = \lim_{\epsilon \downarrow 0} V^\epsilon(t,x)$ on the set $(t,x) \in [t, T[\times \mathbb{R}^d$, which indicates that the approximation scheme can be arbitrarily precise. Note that the approximated payoff functions ℓ_i^ϵ are, by construction, Lipschitz continuous that in light of uniform continuity of the process, see [5, Lemma B.1], yields to the continuity of the value function V^ϵ. Hence, the discontinuous PDE characterization of §V-A is simplified to continuous regime, for which there exist PDE solvers for numerical computations. We refer to [3, Section 5] for a detailed analysis of the proposed approximation scheme.

VI. NUMERICAL EXAMPLE: CHEMICAL LANGEVIN EQUATION FOR A BIOLOGICAL SWITCH

In this section we apply the theoretical results in the preceding sections to a biological switch network. Multistable biological systems are often seen in nature [15]. The inherent stochasticity in these systems can be substantial—they influence convergence, or may even lead to switching behavior from one equilibrium to another. The pioneer works on modeling reactions in biochemical network are based on countable state Markov chains [16], which describe the evolution of molecular numbers. Due to the Markov property of chemical reactions, one can track the time evolution of the probability distribution for molecular populations as a family of

[8] $\mathrm{dist}(x, A) := \inf_{y \in A} \|x - y\|$, where $\|\cdot\|$ stands for the Euclidean norm.

ordinary differential equations called the *chemical master equation* (CME) [17], [18]; it is also known as the forward Kolmogorov equation.

An SDE approach toward modeling the stochastic molecular numbers of species has been proposed in the literature. It is, of course, assumed that a molecular species number may take non-integer numbers. This assumption is usually reasonable for large molecular populations, whereas for low copy numbers, it may not be reliable—a small number, say, 0.1 protein copies can lead to an entirely different dynamic behavior than one would observe from exactly 0 copies. The time-continuity of stochastic processes makes the analysis significantly more tractable than the CME. In this method, one can approximate the molecular numbers via continuous time Markov process, where the latter is an approximation of the jump Markov process that underlines the CME. This stochastic continuous-time approximation is called the *chemical Langevin equation* or the *diffusion approximation*, see [19] and the reference therein for further details on modeling and analysis of stochastic biochemical networks.

Another difficulty with the diffusion approximation is that the model is typically not well-posed since it may assign a negative number to a molecular species. Nevertheless, this issue can be neglected when the focus of observation is away from low numbers of each species. In this section we consider the following chemical Langevin formulation of a two gene network:

$$\begin{cases} \mathrm{d}X_t = \left(f(Y_t, \boldsymbol{u}_x) - \mu_x X_t\right)\mathrm{d}t + \sqrt{f(Y_t, \boldsymbol{u}_x)}\mathrm{d}W_t^1 + \sqrt{\mu_x X_t}\mathrm{d}W_t^2, \\ \mathrm{d}Y_t = \left(g(X_t, \boldsymbol{u}_y) - \mu_y Y_t\right)\mathrm{d}t + \sqrt{g(X_t, \boldsymbol{u}_y)}\mathrm{d}W_t^3 + \sqrt{\mu_y Y_t}\mathrm{d}W_t^4, \end{cases} \tag{15}$$

where X_t and Y_t are the concentration of the two repressor proteins with the respective degradation rates μ_x and μ_y; $(W_t^i)_{t \geq}$ are independent standard Brownian motion processes. Functions f and g are repression functions that describe the impact of each protein on the other's rate of synthesis controlled via some external inputs $\boldsymbol{u}_x$ and $\boldsymbol{u}_y$.

In the absence of exogenous control signals, the authors of [20] study sufficient conditions on the drifts f and g under which the system dynamic (15) without the diffusion term has two (or more) stable equilibria. In this case, system (15) can be viewed as a biological switch network. The aforementioned theoretical results are also experimentally investigated in [21] for a genetic toggle switch in *Escherichia coli*.

In this article we consider the biological switch dynamics such that the degradation rates of proteins are influenced by some external control signals. The practical feasibility of this assumption has successfully been experimented in [22]. The level of repression is described by a *Hill* function, which models cooperativity of binding as follows:

$$f(y,u) := \frac{\theta_1^{n_1} k_1}{y^{n_1} + \theta_1^{n_1}} u, \qquad g(x,u) := \frac{\theta_2^{n_2} k_2}{x^{n_2} + \theta_2^{n_2}} u,$$

where θ_i are the threshold of the production rate with respective exponents n_i, and k_i are the production scaling factors. The parameter u represents the role of external signals that affect the production rates, for which the control sets are $\mathbb{U}_x := [\underline{u}_x, \overline{u}_x]$ and $\mathbb{U}_y := [\underline{u}_y, \overline{u}_y]$

with nominal value $\hat{u} := 1$. As explained in [20], the *nullclines* of the system are $\frac{f(y,\hat{u})}{\mu_1}$ and $\frac{g(x,\hat{u})}{\mu_2}$, which determine whether the system has multiple stable equilibria. In this example we consider system (15) with the following parameters: $\theta_i = 40$, $\mu_i = 0.04$, $k_i = 4$ for both $i \in \{1,2\}$, and exponents $n_1 = 4$, $n_2 = 6$. Figure 7(a) depicts the drift nullclines and the equilibria of the system. The equilibria z_a and z_c are stable, while z_b is the unstable one. We should remark that the "stable equilibrium" of SDE (15) is understood in the absence of the diffusion term, as the noise may very well push the states from one stable equilibrium to another.
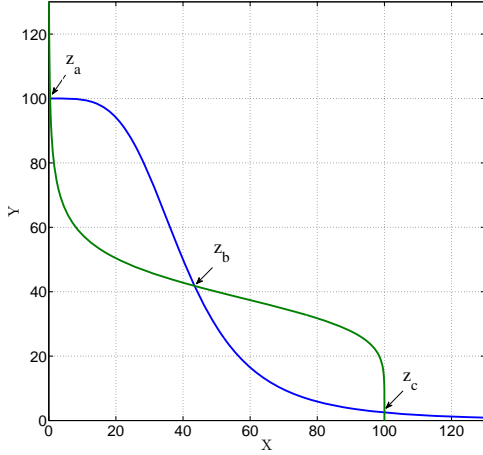
(a) Nullclines and equilibria of the drift of the SDE (15); x_a and x_b are stable, and x_c is unstable.
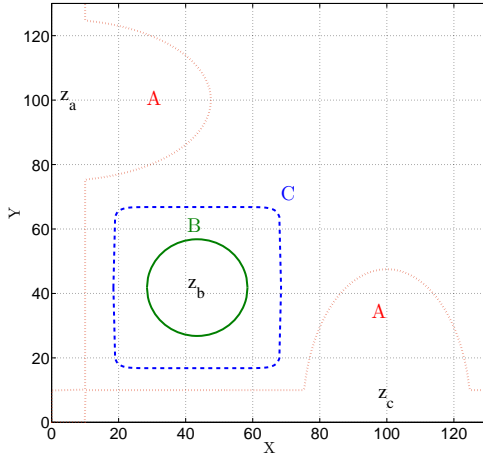
(b) The set A is an avoidance region contained in the region of attraction of the stable equilibria x_a and x_c, B is the target set around the unstable equilibrium x_b, and C is the maintenance margin.

Fig. 7. State space of the biological switch (15) with desired motion planning sets.

In this simulation, we first aim to steer the number of proteins towards a target set around the unstable equilibrium by synthesizing appropriate input signals $\boldsymbol{u}_x$ and $\boldsymbol{u}_y$ within a certain time horizon, say T_1. During this task we opt to avoid the region of attraction of the stable equilibria as well as low numbers for each protein; the latter justifies our model being well-posed in the region of interest. The aforementioned target and avoid sets are denoted, respectively,

by the closed sets B and A in Figure 7(b). In the second phase of the task, once the trajectory visits the target set B, it is required to keep the molecular populations within a slightly larger margin around equilibria for some time, say T_2; Figure 7(b) depicts this maintenance margin by the open set C. In the context of reachability, the second phase is known as *viability* [23]; for an almost-sure stochastic counterpart see for instance [24], [25].

Let us highlight two technical points here: (i) set C must be chosen strictly larger than B in order to allow the process to maneuver inside the interior of the maintenance set C once it hits the target set B. From a theoretical standpoint, this is necessary since the probability of hitting set B and remaining inside B for all future has zero probability; this is due to the non-degenerate property of SDE (15), see the proof of Proposition V.4 for a rigorous analysis of this issue. (ii) Since the process is non-degenerate, in principle the noise can push the process anywhere in the long-run with positive probability. This fact indicates that the probability of success in the second phase of the motion planning task decreases with the time horizon T_2, and it tends to zero as T_2 goes to ∞.

Therefore, the motion planning consists of two parts: reaching the target set B while avoiding the set A within the certain horizon T_1, and staying in the set C for the certain time T_2 after visiting the set B for the first time. In view of motion-planning events introduced in Definition II.1, the first phase of the path can be expressed as $(A^c \rightsquigarrow B)_{\leq T_1}$, and the second phase as $(C \xrightarrow{T_2} C)$; see (1) for detailed definitions of these symbols. By defining the joint process $Z^{t,z;\boldsymbol{u}}_\cdot := \left[X^{t,x;\boldsymbol{u}}_\cdot, Y^{t,y;\boldsymbol{u}}_\cdot\right]$, with the initial condition $z := (x,y)$, the desired excursion is a combination of the events studied in the preceding sections and, with a slight abuse of notation, can be expressed by

$$\left\{ Z^{t,z;\boldsymbol{u}}_\cdot \models (A^c \rightsquigarrow B)_{\leq T_1} \circ (C \xrightarrow{T_2} C) \right\}.$$

The above event depends, of course, on the initial condition (t,x) and control policy $\boldsymbol{u}$, and the objective is to maximize its probability over all admissible policies $\boldsymbol{u} := [\boldsymbol{u}_x, \boldsymbol{u}_y]$. The desired path is not exactly in the framework of Definition II.1 but, nonetheless, one can invoke the same ideas as in §III and introduce the following value functions:

$$V_1(t,z) := \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\left[\mathbb{1}_B\big(Z^{t,z;\boldsymbol{u}}_{\tau_1^1}\big) \mathbb{1}_C\big(Z^{t,z;\boldsymbol{u}}_{\tau_2^1}\big) \right], \quad (16a)$$

$$V_2(t,z) := \sup_{\boldsymbol{u} \in \mathcal{U}_t} \mathbb{E}\left[\mathbb{1}_C\big(Z^{t,z;\boldsymbol{u}}_{\tau_2^2}\big) \right], \quad (16b)$$

where τ_1^1 and τ_2^2 are defined in a same spirit as (7) with given sets $A_1 := (A \cup B)^c$ and $A_2 := C$. However, the stopping time τ_2^1 requires a slight modification so as to address the combination of both motion-planning events introduced in Definition II.1: $\tau_2^1 := \Theta_2^{A_{1:2}} \wedge (\tau_1^1 + T_2)$.

The solution of our motion planning objective is the value function V_1 in (16a), which in view of Theorem V.6 is characterized by the Dynkin differential operator in the interior of $[0,T_1[\times(A \cup B)^c$. However, we need first to solve numerically for the auxiliary value function V_2 in (16b) in order to provide boundary conditions for the PDE corresponding to V_1 by

$$V_1(t,z) = \mathbb{1}_B(z)V_2(t,z), \ \forall (t,x) \in [0,T_1] \times (A \cup B)\bigcup\{T_1\} \times \mathbb{R}^n. \quad (17)$$

It is straightforward to observe that the boundary condition for the value function V_2 is

$$V_2(t,z) = \mathbb{1}_C(z), \ \forall (t,x) \in [0, T_1 + T_2] \times C^c \bigcup \{T_1 + T_2\} \times \mathbb{R}^n.$$

Therefore, we need to solve the PDE of V_2 with the above boundary condition backward from the time $T_1 + T_2$ to the time T_1, and then at time T_1 restrict the value function V_2 onto the set B to provide boundary conditions for the value function V_1. Thus, the value function V_1 can be computed via solving the same PDE from T_1 to 0 but along with different boundary conditions provided by the preceding step. According to Definition V.5 for any smooth function $\phi := \phi(t,x,y)$ the Dynkin operator $\mathcal{L}^u$ can be simplified to

$$\sup_{u \in \mathbb{U}} \mathcal{L}^u \phi(t,x,y)$$

$$= \max_{u \in \mathbb{U}} \Big[\partial_t \phi + \partial_x \phi \big(f(y, u_x) - \mu_x x \big) + \partial_y \phi \big(g(x, u_y) - \mu_y y \big)$$

$$+ \frac{1}{2} \partial_x^2 \phi \big(f(y, u_x) + \mu_x x \big) + \frac{1}{2} \partial_y^2 \phi \big(g(x, u_y) + \mu_y y \big) \Big]$$

$$= \partial_t \phi - \big(\partial_x \phi - \frac{1}{2} \partial_x^2 \phi \big) \mu_x x - \big(\partial_y \phi - \frac{1}{2} \partial_y^2 \phi \big) \mu_y y$$

$$+ \max_{u_x \in [\underline{u}_x, \overline{u}_x]} \Big[f(y, u_x) \big(\partial_x \phi + \frac{1}{2} \partial_x^2 \phi \big) \Big]$$

$$+ \max_{u_y \in [\underline{u}_y, \overline{u}_y]} \Big[g(x, u_y) \big(\partial_y \phi + \frac{1}{2} \partial_y^2 \phi \big) \Big].$$

On account of Theorem V.6 and linearity of the drift terms in u, one can propose an optimal policy in terms of derivatives of the value functions V_1 and V_2 in (16), respectively, for the first and second phase of the motion:

$$\boldsymbol{u}_x^*(t,x,y) = \begin{cases} \overline{u}_x(t,x,y) & \text{if } \partial_x V_i(t,x,y) + \frac{1}{2} \partial_x^2 V_i(t,x,y) \geq 0, \\ \underline{u}_x(t,x,y) & \text{if } \partial_x V_i(t,x,y) + \frac{1}{2} \partial_x^2 V_i(t,x,y) < 0, \end{cases}$$

$$\boldsymbol{u}_y^*(t,x,y) = \begin{cases} \overline{u}_y(t,x,y) & \text{if } \partial_y V_i(t,x,y) + \frac{1}{2} \partial_y^2 V_i(t,x,y) \geq 0, \\ \underline{u}_y(t,x,y) & \text{if } \partial_y V_i(t,x,y) + \frac{1}{2} \partial_y^2 V_i(t,x,y) < 0, \end{cases}$$

where $i \in \{1, 2\}$ corresponds to the phase of the path.

In the sequel we investigate two scenarios: first, when full control over both production rates is possible, i.e., $\underline{u}_x = \underline{u}_y = 0$ and $\overline{u}_x = \overline{u}_y = 2$; second, when we only have access to the production rate of protein x, i.e., $\overline{u}_y = \underline{u}_y = \hat{u}$. Figure 8 depicts the probability distribution of staying in set C within the time horizon $T_2 = 120$ time units [9] in terms of the initial conditions $(x,y) \in \mathbb{R}^2$. V_2 is zero outside set C, as the process has obviously left C if it starts outside it. Figures 8(a) and 8(b) demonstrate the first and second control scenarios, respectively. Note that in the second case the probability of success dramatically decreases in comparison with the first. This result indicates the importance of full controllability of the production rates for the achievement of the given control objective.

Figure 9 depicts the probability of successively reaching set B within the time horizon $T_1 = 60$ time units and staying in set C for $T_2 = 120$ time units thereafter. Since the objective is to avoid set A, the value function V_1 takes zero value on A. Figures 9(a) and 9(b) demonstrate the first and second control scenarios, respectively. It is

[9]Notice that the half-life of each protein is assumed to be 17.32 time units
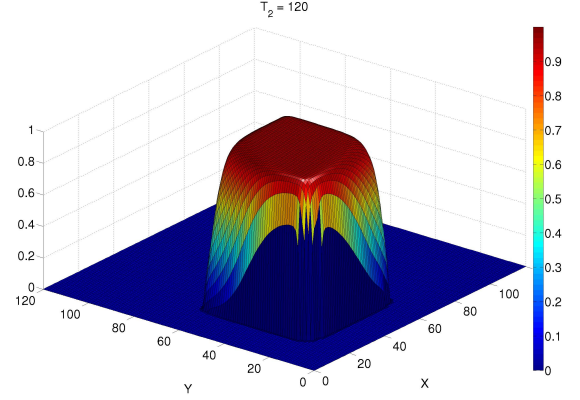
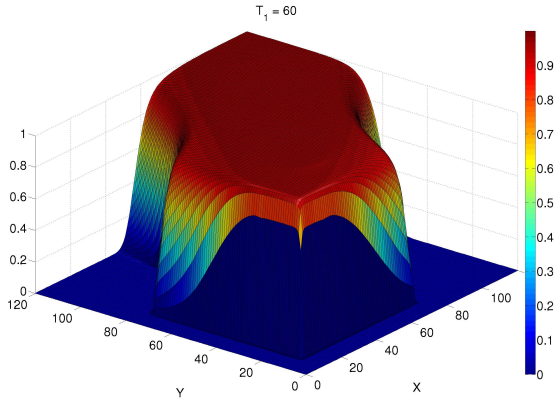(a) V_2 in case of full controllability over both production rates.
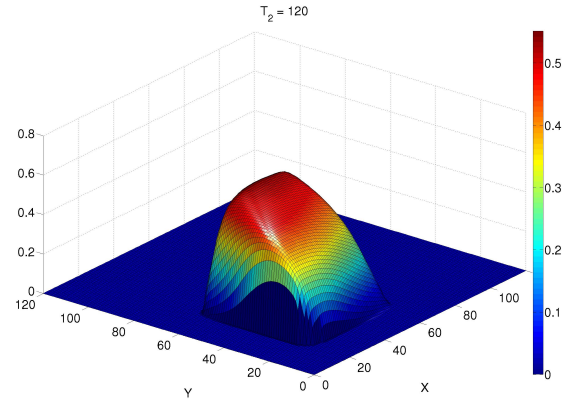
(b) V_2 in case of half controllability over only the production rate of protein x.

Fig. 8. The value function V_2 as defined in (16b) corresponding to probability of staying in C for 120 time units.

easy to observe the non-smooth behavior of the value function V_1 on the boundary of set B in Figure 9(b). This is indeed a consequence of the boundary condition explained in (17). All simulations in this subsection were obtained using the Level Set Method Toolbox [14] (version 1.1), with a grid 121×121 in the region of simulation.

VII. CONCLUSION AND FUTURE DIRECTIONS

In this article we introduced different notions of stochastic motion planning problems, with RCLL sample-paths realizations. Based on a class of stochastic optimal control problems, we proposed an alternative characterization of the set of initial conditions from which there exists an admissible policy to execute the desired maneuver with probability at least as much as some pre-specified value. We then established a weak DPP, which does not need to verify the measurability of the value functions, in terms of some auxiliary value functions. Subsequently, we focused on a case of diffusions as the solution of an SDE, and investigated the required conditions to apply the proposed DPP in the preceding section. It turned out that invoking the DPP, one can solve a series of PDEs in a recursive fashion so as to numerically compute the desired initial set as well as admissible policy for the motion planning specifications.

(a) V_1 in case of full controllability over the production rates.
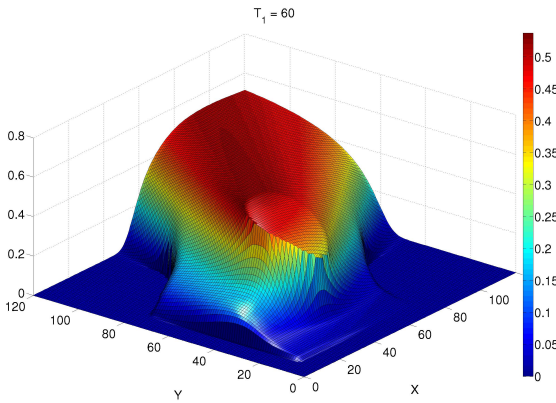
(b) V_1 in case of inability to increase the production rate of protein y.

Fig. 9. The value function V_1 as defined in (16a) corresponding to probability of staying in C for 120 time units, once it reaches B while avoiding A within 60 time units.

For future work, in light of Theorem IV.5 which is in fact developed for RCLL processes, we aim to study the required conditions of the proposed DPP, Assumptions IV.1, for a larger class of stochastic processes, e.g., controlled Markov jump-diffusions.

REFERENCES

[1] D. Chatterjee, E. Cinquemani, and J. Lygeros, "Maximizing the probability of attaining a target prior to extinction," *Nonlinear Analysis: Hybrid Systems*, 2011, http://dx.doi.org/10.1016/j.nahs.2010.12.003.

[2] P. Mohajerin Esfahani, D. Chatterjee, and J. Lygeros, "On a problem of stochastic reach-avoid set characterization for diffusions," in *IEEE Conference on Decision and Control*, dec. 2011, pp. 7069 –7074.

[3] ——, "On A Stochastic Reach-Avoid Problem and Set Characterization," Tech. Rep., Oct. 2012, [Online]. Available: http://arxiv.org/abs/1202.4375.

[4] B. Bouchard and N. Touzi, "Weak dynamic programming principle for viscosity solutions," *SIAM J. Control Optim.*, vol. 49, no. 3, pp. 948–962, 2011.

[5] P. Mohajerin Esfahani, D. Chatterjee, and J. Lygeros, "Motion Planning via Optimal Control for Stochastic Processes," Tech. Rep., Nov. 2012. [Online]. Available: http://arxiv.org/abs/1211.1138

[6] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed., ser. Graduate Texts in Mathematics. New York: Springer-Verlag, 1991, vol. 113.

[7] S. Ethier and T. Kurtz, *Markov Processes: Characterization and Convergence*, ser. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons, Ltd., 1986.

[8] V. S. Borkar, "Controlled diffusion processes," *Probability Surveys*, vol. 2, pp. 213–244 (electronic), 2005. [Online]. Available: http://dx.doi.org/10.1214/154957805100000131

[9] N. Krylov, *Controlled Diffusion Processes*, ser. Stochastic Modelling and Applied Probability. Berlin Heidelberg: Springer-Verlag, 2009, vol. 14, reprint of the 1980 Edition.

[10] O. Kallenberg, *Foundations of Modern Probability*, ser. Probability and its Applications (New York). New York: Springer-Verlag, 1997.

[11] P. J. Reny, "On the existence of pure and mixed strategy nash equilibria in discontinuous games," *Econometrica*, vol. 67, pp. 1029–1056, 1999. [Online]. Available: http://dx.doi.org/10.1214/154957805100000131

[12] W. Fleming and H. Soner, *Controlled Markov Processes and Viscosity Solution*, 3rd ed. Springer-Verlag, 2006.

[13] M. G. Crandall, H. Ishii, and P. L. Lions, "User's guide to viscosity solutions of second order partial differential equations," *American Mathematical Society*, vol. 27, pp. 1–67, 1992. [Online]. Available: http://dx.doi.org/10.1214/154957805100000131

[14] I. Mitchell, "A toolbox of hamilton-jacobi solvers for analysis of nondeterministic continuous and hybrid systems," in *Hybrid systems: computation and control*, ser. Lecture Notes in Comput. Sci., M. Morari and L. Thiele, Eds. Springer-Verlag, 2005, no. 3414, pp. 480–494.

[15] A. Becskei, B. Seraphin, and L. Serrano, "Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion," *EMBO J*, vol. 20, no. 10, pp. 2528–2535, May 2001.

[16] D. J. Wilkinson, *Stochastic modelling for systems biology*. Chapman & Hall/CRC, Boca Raton, FL, 2006.

[17] S. Andrews, T. Ginh, and A. Arkin, "Stochastic models of biological processes," *Meyers RA, ed. Encyclopedia of Complexity and System Science*, p. 87308749, 2009.

[18] H. El Samad, M. Khammash, L. Petzold, and D. Gillespie, "Stochastic modelling of gene regulatory networks," *Internat. J. Robust Nonlinear Control*, vol. 15, no. 15, pp. 691–711, 2005.

[19] M. Khammash, "Modeling and analysis of stochastic biochemical networks," in *Control Theory and System Biology*.

[20] J. Cherry, "How to make a Biological Switch," *Journal of Theoretical Biology*, vol. 203, no. 2, pp. 117–133, Mar. 2000.

[21] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in Escherichia coli," *Nature*, vol. 403, no. 6767, pp. 339–42, Jan. 2000.

[22] A. Milias, S. Summers, J. Stewart-Ornstein, I. Zuleta, D. Pincus, H. El-Samad, M. Khammash, and J. Lygeros, "In silico feedback for in vivo regulation of a gene expression circuit," *Nature Biotechnology*, vol. 29, no. 12, p. 11141116, Dec. 2011. [Online]. Available: http://control.ee.ethz.ch/index.cgi?page=publications;action=details;id=3903

[23] J. Aubin, *Viability Theory*, ser. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston Inc., 1991.

[24] J. Aubin and G. Da Prato, "Stochastic viability and invariance," *Annali della Scuola Normale Superiore di Pisa. Classe di Scienze. Serie IV*, vol. 17, no. 4, pp. 595–613, 1990. [Online]. Available: http://www.numdam.org/item?id=ASNSP_1990_4_17_4_595_0

[25] J. Aubin and G. D. Prato, "The viability theorem for stochastic differential inclusions," *Stochastic Analysis and Applications*, vol. 16, no. 1, pp. 1–15, 1998. [Online]. Available: http://dx.doi.org/10.1080/07362999808809512