

# MOTION PLANNING FOR CONTINUOUS TIME STOCHASTIC PROCESSES: A DYNAMIC PROGRAMMING APPROACH

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 a control enabling the process to execute the desired maneuver with probability no less than some pre-specified value. The proposed DPP comprises auxiliary value functions defined in terms of discontinuous payoff functions. A concrete instance of the use of this novel DPP in the case of diffusion processes is also presented. In this case, we establish 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 on an example involving biological switches.

## 1. INTRODUCTION

Motion planning of dynamical systems aims to steer the state of the system through certain given sets in a specific order and pre-assigned time schedule. This problem finds a wide spectrum of applications ranging from air traffic management [LTS00] and security of power networks [MVM<sup>+</sup>10] to navigation of unmanned air vehicles [MB00, BMGA02].

1.1. **Context.** The two fields of robotics and control have contributed much to this problem. Particular problems of interest in this context are the control synthesis for a given initial condition to accomplish the motion planning task, and the determination of the set of initial conditions for which such a controller exists.

The focus in the robotics community has mainly been on the first objective (i.e., control synthesis for a given initial condition) with particular emphasis on computational issues. To this end, the motion planning problem is typically *approximated* in an optimal control framework wherein the cost function rewards the target set while penalizing the obstacles. In this vein, there is a rich literature to tackle the corresponding approximation problems. Examples include [JM70] that capitalizes on the synergy between local and global methods, [TL05] that builds on [JM70] to propose an iterative LQG scheme, and [Kap07] extended later by [TBS10] that

---

*Date:* January 22, 2016.

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.

proposes a reinforcement learning based approach to generate parametrized control policy via path integrals.

Research toward the second objective (i.e. characterizing the desired set of initial conditions) is traditionally conducted in the control community, and is the focus of this article. A technical difficulty to *accurately* characterize this set is a potential non-smooth behavior commonly arising in the optimal control context, which consequently renders the analysis of the value function more involved. This issue is a computational concern for control synthesis purposes as well; see, for example, [AM11] that opts to circumvent this non-smoothness by introducing additional noise.

**1.2. Literature on Set Characterization.** From a set characterization viewpoint, motion planning problems in the deterministic setting have been studied extensively from different perspectives; see, for example, [Aub91] in the language of viability theory, [SM86] for a dynamic programming approach, and also [Sus91, CS98] for a differential geometric perspective. In the stochastic setting, stochastic viability and controlled invariance are studied in an almost-sure treatment in [APF00, BJ02]; see also the references therein. Following the same objective, methods involving stochastic contingent sets [AP98], viscosity solutions to second-order partial differential equations [BG99], and also the equivalence for the invariance problem between a stochastic differential equation (SDE) and a certain deterministic control system were developed in this context [DF04].

Although almost-sure motion planning maneuvers are interesting in their own right, they maybe be too conservative in applications with less stringent safety requirements where a common specification involves only bounding the probability that undesirable events take place. This line of research in the stochastic setting has received relatively little attention, in particular for systems governed by SDEs. In fact, it was not until recently that the basic motion planning problem involving one target and one obstacle set, the so-called reach-avoid problem, was investigated in the context of finite probability spaces for a class of continuous-time Markov decision processes [BHKH05], and in the discrete-time stochastic hybrid systems context [CCL11, SL10].

In the continuous time and continuous space settings, one may tackle the dynamic programming formulation of the reach-avoid problem from two perspectives: a direct technique based on the theory of stochastic target problems, and an indirect approach via an exit-time stochastic optimal control formulation. For the former, we refer the reader to [BT11]; see also the recent book [Tou13]. In our earlier works [MCL11, MCL15] we focused on the latter perspective for reachability of controlled diffusion processes. Here we continue in the same spirit by going beyond the reach-avoid problem to more complex motion planning problems for a larger class of stochastic processes with possibly discontinuous sample paths. Preliminary results in this direction were reported in [MMAC13] with an emphasis on the application side and without covering the technical details.

**1.3. Our Contributions.** In this context the main contributions of the present article are summarized below:

- (i) Based on a formal definition of the different stochastic motion planning maneuvers (Section 2), we establish a connection between the desired task and a class of stochastic optimal control problems (Section 3);
- (ii) we propose a weak dynamic programming principle (DPP) under mild assumptions on the admissible controls and the stochastic process (Section 4);

(iii) in the context of controlled SDEs, based on the proposed DPP we derive a partial differential equation (PDE) characterization of the desired set of initial conditions (Section 5).

In more detail, we start with the formal definition of a motion planning objective comprising of two fundamental reachability maneuvers. To the best of our knowledge, this is new in the literature. We address the following natural question: for which initial states do there exist an admissible control such that the stochastic processes satisfy the motion planning objective with a probability greater than a given value  $p$ ? We then characterize this set of initial states by establishing a connection between the motion planning specifications and a class of stochastic optimal control problems involving discontinuous payoff functions and a sequence of successive exit-times.

Due to the discontinuity of the payoff functions, the classical results in stochastic optimal control problems and its connection to Hamilton-Jacobi-Bellman PDEs are not applicable here, see for instance [FS06, Section IV.7]. Under mild technical assumptions, we propose a weak DPP involving auxiliary value functions which only requires semicontinuity of the value functions. It is worth noting that the DPP is developed in a rather general setting, emphasizing particular features of the underlying process that lead to such a characterization of the motion planning. Closest in spirit to our work is [BT11] in which the optimal control problem is studied in a fixed time horizon as well as an optimal stopping setting. Neither of these settings is applicable to our exit-time framework. From a technical standpoint, this article extends the technology developed by [BT11] to deal with the class of exit-time problems suitable to our motion planning tasks.

Finally, we focus on a class of controlled SDEs in which the required assumptions of the proposed DPP are investigated. Indeed, it turns out that the standard uniform non-degeneracy and exterior cone conditions of the involved sets suffice to fulfill the DPP requirements. Subsequently, we demonstrate how the DPP leads to a new framework for characterizing the desired set of initial conditions based on tools from PDEs. Due to the discontinuities of the value functions involved, all the PDEs are understood in the generalized notion of the so-called discontinuous viscosity solutions. In this context, we show how the value functions can be solved by a means of a sequence of PDEs, in which the preceding PDE provides the boundary condition of the current one.

**1.4. Computational Complexity and Existing Numerical Tools.** On the computational side, it is well-known that PDE techniques suffer from the curse of dimensionality. In the literature a class of suboptimal control methods referred to as Approximate Dynamic Programming (ADP) have been developed for dealing with this difficulty; for a sampling of recent works see [dFVR03] for a linear programming approach, [KT03] for actor-critic algorithms, and [Ber05] for a comprehensive survey on the entire area. Besides the ADP literature, very recent progress on numerical methods based on tensor train decompositions holds the potential of substantially ameliorating this curse of dimensionality; see a representative article [KS11] and the references therein. In this light, taken in its entirety, the results in this article can be viewed as a theoretical bridge between the motion planning objective formalized in Section 2 and sophisticated numerical methods that can be used to address real problem instances. Here we demonstrate the practical use of this bridge by addressing a stochastic motion planning problem for biological switches.

The organization of the article is as follows: In Section 2 we formally introduce the stochastic motion planning problems. In Section 3 we establish a connection between the motion planning

objectives and a class of stochastic optimal control problems, for which a weak DPP is proposed in Section 4. A concrete instance of the use of the novel DPP in the case of controlled diffusion processes is presented in Section V, leading to characterization of the motion planning objective with the help of a sequence of PDE's in an iterative fashion. To validate the performance of the proposed methodology, in Section 6 the theoretical results are applied to a biological two-gene network. For better readability, the technical proofs along with required preliminaries are provided in the appendices.

Notation. Given  $a, b \in \mathbb{R}$ , we define  $a \wedge b := \min\{a, b\}$  and  $a \vee b := \max\{a, b\}$ . We denote by  $A^c$  (resp.  $A^\circ$ ) the complement (resp. interior) of the set  $A$ . We also denote by  $\bar{A}$  (resp.  $\partial A$ ) the closure (resp. boundary) of  $A$ . We let  $B_r(x)$  be an open Euclidean ball centered at  $x$  with radius  $r$ . The Borel  $\sigma$ -algebra on a topological space  $\mathbb{A}$  is denoted by  $\mathfrak{B}(\mathbb{A})$ , and measurability on  $\mathbb{R}^d$  will always refer to Borel-measurability. Given function  $f : \mathbb{A} \rightarrow \mathbb{R}$ , the lower and upper semicontinuous envelopes of  $f$  are defined, respectively, by  $f_*(x) := \liminf_{x' \rightarrow x} f(x')$  and  $f^*(x) := \limsup_{x' \rightarrow x} f(x')$ . Throughout this article all (in)equalities between random variables are understood in almost sure sense. For the ease of the reader, we also provide here a partial notation list which will be also explained in more details later throughout the article:

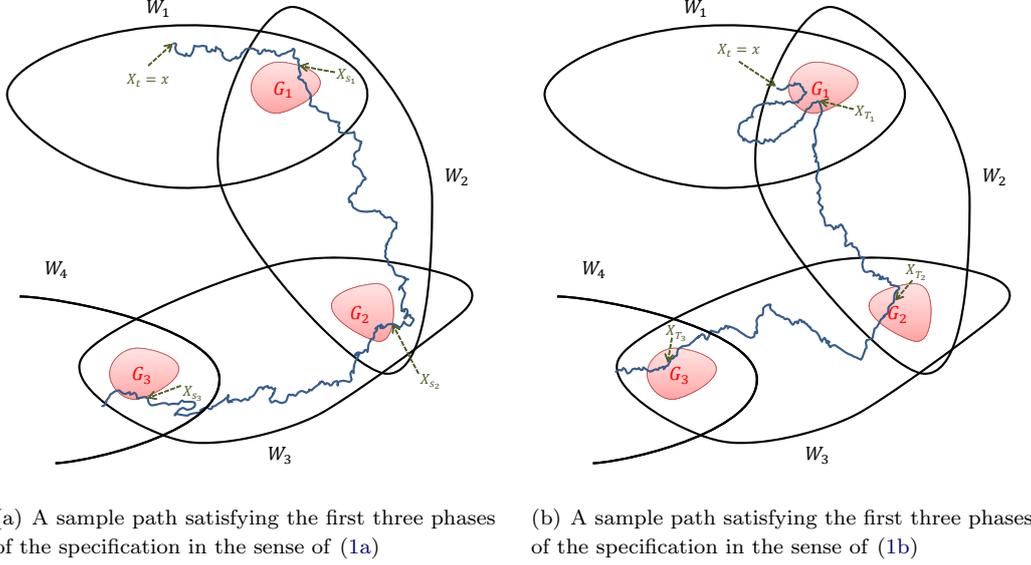
- $\mathcal{U}_t$ : set of admissible controls at time  $t$ ;
- $(X_s^{t,x;\mathbf{u}})_{s \geq 0}$ : stochastic process under the control  $\mathbf{u}$  and convention  $X_s^{t,x;\mathbf{u}} := x$  for all  $s \leq t$ ;
- $(W_i \rightsquigarrow G_i)_{\leq T_i}$  (resp.  $W_i \xrightarrow{T_i} G_i$ ) : motion planning event of reaching  $G_i$  sometime before time  $T_i$  (resp. at time  $T_i$ ) while staying in  $W_i$ , see Definition 2.1;
- $(\Theta_i^{A_i})_{i=k}^n$ : sequential exit-times from the sets  $(A_i)_{i=k}^n$  in order, see Definition 3.1;
- $\mathcal{L}^u$ : Dynkin operator, see Definition 5.5.

## 2. 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  $\mathbf{z} := (z_s)_{s \geq 0}$  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 [KS91, p. 48]. Consider also an auxiliary subfiltration  $\mathbb{F}_t := (\mathcal{F}_{t,s})_{s \geq 0}$ , where  $\mathcal{F}_{t,s}$  is the  $\mathbb{P}$ -completion of  $\sigma(z_{r \vee t} - z_t, r \in [0, s])$ . It is obvious to observe that any  $\mathcal{F}_{t,s}$ -random variable is independent of  $\mathcal{F}_t$ ,  $\mathcal{F}_{t,s} \subseteq \mathcal{F}_s$  with equality in case of  $t = 0$ , and for  $s \leq t$ ,  $\mathcal{F}_{t,s}$  is the trivial  $\sigma$ -algebra.

The object of our study is an  $\mathbb{R}^d$ -valued controlled random process  $(X_s^{t,x;\mathbf{u}})_{s \geq t}$ , initialized at  $(t, x)$  under the control  $\mathbf{u} \in \mathcal{U}_t$ , where  $\mathcal{U}_t$  is the set of admissible controls at time  $t$ . Since the precise class of admissible controls does not play a role until Section IV we defer the formal definition of these until then. Let  $T > 0$  be a fixed time horizon, and let  $\mathbb{S} := [0, T] \times \mathbb{R}^d$ . We assume that for every  $(t, x) \in \mathbb{S}$  and  $\mathbf{u} \in \mathcal{U}_t$ , the process  $(X_s^{t,x;\mathbf{u}})_{s \geq t}$  is  $\mathbb{F}_t$ -adapted process whose sample paths are right continuous with left limits. 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  $\tau$  such that  $\tau_1 \leq \tau \leq \tau_2$   $\mathbb{P}$ -a.s.

Given sets  $(W_i, G_i) \in \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$  for  $i \in \{1, \dots, n\}$ , we are interested in a set of initial conditions  $(t, x) \in \mathbb{S}$  such that there exists an admissible control  $\mathbf{u} \in \mathcal{U}_t$  steering the process  $X_s^{t,x;\mathbf{u}}$  through  $(W_i)_{i=1}^n$  ("way point" sets) while visiting  $(G_i)_{i=1}^n$  ("goal" sets) in a pre-assigned order. One may pose this objective from different perspectives based on different time scheduling


 FIGURE 1. Sample paths of the process  $X^{t,x;\mathbf{u}}$  for a fix control  $\mathbf{u} \in \mathcal{U}_t$ 

for the excursions between the sets. We formally introduce some of these notions which will be addressed throughout this article.

**Definition 2.1** (Motion Planning Events). *Consider a fixed initial condition  $(t, x) \in \mathbb{S}$  and admissible control  $\mathbf{u} \in \mathcal{U}_t$ . 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:*

$$(1a) \quad \left\{ X^{t,x;\mathbf{u}} \models [(W_1 \rightsquigarrow G_1) \circ \dots \circ (W_n \rightsquigarrow G_n)]_{\leq T} \right\} :=$$

$$\left\{ \exists (s_i)_{i=1}^n \subset [t, T] \mid X_{s_i}^{t,x;\mathbf{u}} \in G_i \quad \text{and} \quad X_r^{t,x;\mathbf{u}} \in W_i \setminus G_i, \quad \forall r \in [s_{i-1}, s_i], \quad \forall i \leq n \right\},$$

$$(1b) \quad \left\{ X^{t,x;\mathbf{u}} \models (W_1 \xrightarrow{T_1} G_1) \circ \dots \circ (W_n \xrightarrow{T_n} G_n) \right\} :=$$

$$\left\{ X_{T_i}^{t,x;\mathbf{u}} \in G_i \quad \text{and} \quad X_r^{t,x;\mathbf{u}} \in W_i, \quad \forall r \in [T_{i-1}, T_i], \quad \forall i \leq n \right\},$$

where  $s_0 = T_0 := t$ .

The set in (1a), roughly speaking, contains the events in the underlying probability space that the trajectory  $X^{t,x;\mathbf{u}}$ , initialized at  $(t, x) \in \mathbb{S}$  and controlled via  $\mathbf{u} \in \mathcal{U}_t$ , succeeds in visiting  $(G_i)_{i=1}^n$  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  $W_i$  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 usually more restricted in comparison to (1a). Indeed, not only is the trajectory confined to  $W_i$  on the way between  $G_{i-1}$  and  $G_i$ , 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  $T_i$  for each  $i$ . Figure 1(b) demonstrates one sample path in which the first three phases of the excursion are successfully fulfilled.

Note that once a trajectory belonging to the set in (1a) visits  $G_i$  for the first time, it is required to remain in  $W_{i+1}$  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  $G_i$  at the time  $T_i$ . As an illustration, one can easily inspect that the sample path in Figure 1(b) indeed violates the requirements of the definition (1a) as it leaves  $W_2$  after it visits  $G_1$  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  $G_i$ , while the definition (1b) only changes the admissible way set *only after* the intermediate time  $T_i$  irrespective of whether the trajectory visits  $G_i$  prior to  $T_i$ .

For simplicity we may impose the following assumptions:

**Assumption 2.2.** *We stipulate that*

- a. *The sets  $(G_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$  are closed.*
- b. *The sets  $(W_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$  are open.*

Concerning Assumption 2.2.a., if  $G_i$  is not closed, then it is not difficult to see that there could be some continuous transitions through the boundary of  $G_i$  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  $G_i$  at the time  $s_i$ . Notice that this is not the case for the definition (1b) since the trajectory only visits the sets  $G_i$  at the specific times  $T_i$  while any continuous transition and maneuver inside  $G_i$  are allowed. Assumption 2.2.b. is rather technical and required for the analysis employed in subsequent sections.

The events introduced in Definition 2.1 depend, of course, on the control  $\mathbf{u} \in \mathcal{U}$  and initial condition  $(t, x) \in \mathbb{S}$ . The main objective of this article is to determine the set of initial conditions  $x \in \mathbb{R}^d$  such that there exists an admissible control  $\mathbf{u}$  where the probability of the motion planning events is higher than a certain threshold. Let us formally introduce these sets as follows:

**Definition 2.3** (Motion Planning Initial Condition 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 condition sets:*

$$(2a) \quad \text{MP}(t, p; (W_i, G_i)_{i=1}^n, T) := \left\{ x \in \mathbb{R}^d \mid \exists \mathbf{u} \in \mathcal{U}_t : \right. \\ \left. \mathbb{P}\{X^{t,x;\mathbf{u}} \models [(W_1 \rightsquigarrow G_1) \circ \dots \circ (W_n \rightsquigarrow G_n)]_{\leq T}\} > p \right\},$$

$$(2b) \quad \widetilde{\text{MP}}(t, p; (W_i, G_i)_{i=1}^n, (T_i)_{i=1}^n) := \left\{ x \in \mathbb{R}^d \mid \exists \mathbf{u} \in \mathcal{U}_t : \right. \\ \left. \mathbb{P}\{X^{t,x;\mathbf{u}} \models (W_1 \xrightarrow{T_1} G_1) \circ \dots \circ (W_n \xrightarrow{T_n} G_n)\} > p \right\}.$$

**Remark 2.4** (Stochastic Reach-Avoid Problem). *The motion planning scenarios for only two sets  $(W_1, G_1)$  basically reduce to the basic reach-avoid maneuver studied in our earlier work [] by setting the reach set to  $G_1$  and the avoid set to  $\mathbb{R}^d \setminus W_1$ . See also [GLQ06] for the corresponding deterministic and [SL10] for the corresponding discrete time stochastic reach-avoid problems.*

**Remark 2.5** (Mixed Motion Planning Events). *One may also consider an event that consists of a mixture of the events in (1), e.g.,  $(W_1 \rightsquigarrow G_1)_{\leq T_1} \circ (W_2 \xrightarrow{T_2} G_2)$ . Following essentially*

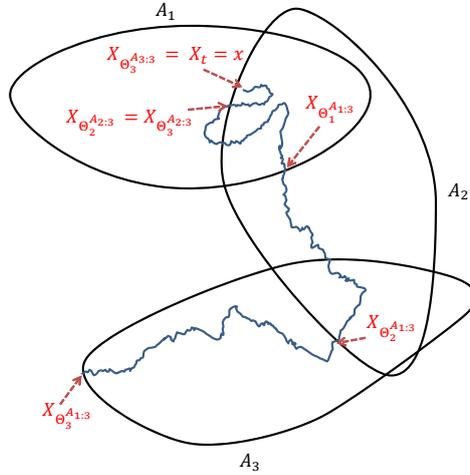


FIGURE 2. Sequential exit-times of a sample path through the sets  $(A_i)_{i=k}^3$  for different values of  $k$

the same analytical techniques as the ones proposed in subsequent sections, one can also address these mixed motion planning objectives; see [MMAC13] for an example of this nature.

**Remark 2.6** (Time-varying Goal and Way Point Sets). *In Definition 2.3 the motion planning objective is introduced in terms of stationary (time-independent) goal and way point sets. However, note that one can always augment the state space with time, and introduce a new stochastic process  $Y_t := [X_t^\top, t]^\top$ . Therefore, a motion planning concerning moving sets for  $X_t$  can be viewed as a motion planning with stationary sets for the process  $Y_t$ .*

### 3. CONNECTION TO STOCHASTIC OPTIMAL CONTROL PROBLEMS

In this section we establish a connection between stochastic motion planning initial condition sets MP and  $\widetilde{\text{MP}}$  of Definition 2.3 and a class of stochastic optimal control problems involving stopping times. First, given a sequence of sets we introduce a sequence of random times that corresponds to the times that the process  $X^{t,x;\mathbf{u}}$  exits each set in the sequence for the first time.

**Definition 3.1** (Sequential Exit-Time). *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  $(\Theta_i^{A_k:n})_{i=k}^n$  defined<sup>1</sup> by*

$$\begin{cases} \Theta_i^{A_k:n}(t, x) := \inf \{ r \geq \Theta_{i-1}^{A_k:n}(t, x) : X_r^{t,x;\mathbf{u}} \notin A_i \}, \\ \Theta_{k-1}^{A_k:n}(t, x) := t, \end{cases}$$

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  $\mathbf{u}$  in addition to the initial condition  $(t, x)$ , but here and later in the sequel we shall suppress this dependence. For notational simplicity, we may also drop  $(t, x)$  in subsequent sections. We refer the interested reader to Lemma I.1 showing that these sequential stopping times are indeed well-defined.

<sup>1</sup>By convention,  $\inf \emptyset = \infty$ .

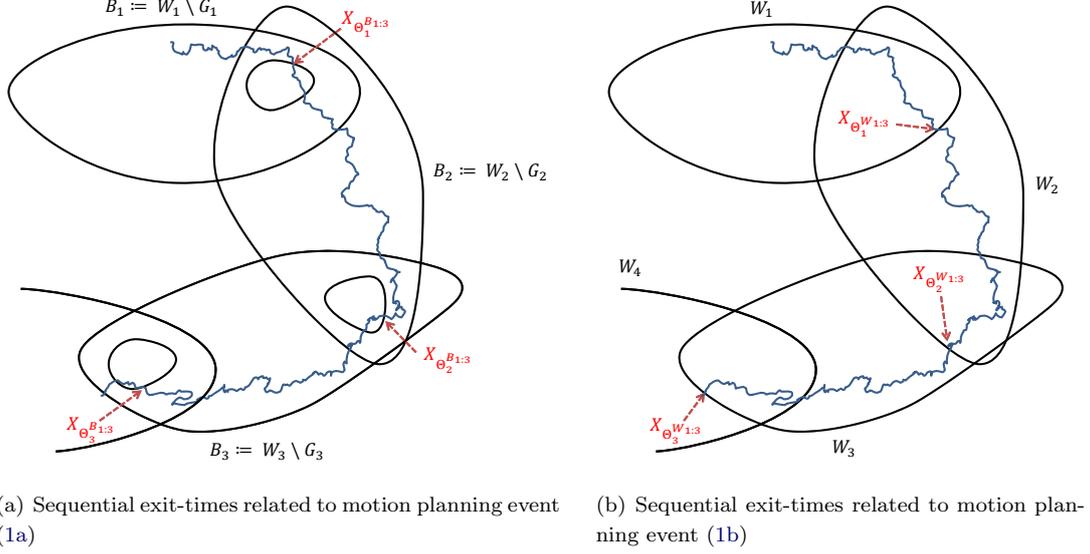


FIGURE 3. Sequential exit-times corresponding to different motion planning events as introduced in (1)

In Figure 2 a sample path of the process  $X^{t,x;\mathbf{u}}$  along with the sequential exit-times  $(\Theta_i^{A_{k:3}})_{i=k}^n$  is depicted for different  $k \in \{1, 2, 3\}$ . Note that since the initial condition  $x$  does not belong 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. In Section 5 we shall see that these differences will lead to different definitions of value functions in order to derive a dynamic programming argument.

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  $V, \tilde{V} : \mathbb{S} \rightarrow [0, 1]$  defined by

$$(3a) \quad \begin{cases} V(t, x) := \sup_{\mathbf{u} \in \mathcal{U}_t} \mathbb{E} \left[ \prod_{i=1}^n \mathbf{1}_{G_i} (X_{\eta_i}^{t,x;\mathbf{u}}) \right], \\ \eta_i := \Theta_i^{B_{1:n}} \wedge T, \quad B_i := W_i \setminus G_i, \end{cases}$$

$$(3b) \quad \begin{cases} \tilde{V}(t, x) := \sup_{\mathbf{u} \in \mathcal{U}_t} \mathbb{E} \left[ \prod_{i=1}^n \mathbf{1}_{G_i \cap W_i} (X_{\tilde{\eta}_i}^{t,x;\mathbf{u}}) \right], \\ \tilde{\eta}_i := \Theta_i^{W_{1:n}} \wedge T_i, \end{cases}$$

where  $\Theta_i^{W_{1:n}}, \Theta_i^{B_{1:n}}$  are the sequential exit-times in the sense of Definition 3.1. Figure 3(a) and 3(b) illustrate the sequential exit-times corresponding to the sets  $B_i$  and  $W_i$ , respectively. The main result of this section, Theorem 3.2 below, establishes a connection from the sets  $\text{MP}, \widetilde{\text{MP}}$  and superlevel sets of the value functions  $V$  and  $\tilde{V}$ .

**Theorem 3.2** (Optimal Control Characterization). *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,

$$(4) \quad \text{MP}(t, p; (W_i, G_i)_{i=0}^n, T) = \{x \in \mathbb{R}^d \mid V(t, x) > p\}.$$

Moreover, suppose Assumption 2.2.b. holds. Then,

$$(5) \quad \widetilde{\text{MP}}(t, p; (W_i, G_i)_{i=0}^n, (T_i)_{i=1}^n) = \{x \in \mathbb{R}^d \mid \widetilde{V}(t, x) > p\},$$

where the value functions  $V$  and  $\widetilde{V}$  are as defined in (3).

*Proof.* See Appendix I. □

Intuitively speaking, observe that the value functions (3) consist of a sequence of indicator functions, where the reward is 1 when the corresponding phase (i.e., reaching  $G_i$  while staying in  $W_i$ ) of motion planning is fulfilled, while the reward is 0 if it fails. Let us also highlight that the difference between the time schedule between the two motion planning problems in (3) is captured via the stopping times  $\eta_i$  and  $\tilde{\eta}_i$ : the former refers to the first time to leave  $W_i$  or hit  $G_i$  before  $T$ , and the latter only considers the exit time from  $W_i$  prior to  $T_i$ . Hence, the product of the indicators evaluates to 1 if and only if the entire journey comprising  $n$  phases is successfully accomplished. In this light, taking expectations yields the probability of the desired event, and the supremum over admissible controls leads to the assertion that there exists a control for which the desired properties hold.

#### 4. DYNAMIC PROGRAMMING PRINCIPLE

The objective of this section is to derive a DPP for the value functions  $V$  and  $\widetilde{V}$  introduced in (3). The DPP provides a bridge between the theoretical characterization of the solution to our motion planning problem through value functions (Section 3) and explicit characterizations of these value functions using, for example, PDEs (Section 5), which can then be used to solve the original problem numerically.

Let  $(T_i)_{i=1}^n \subset [0, T]$  be a sequence of times,  $(A_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$  be a sequence of open sets, and  $\ell_i : \mathbb{R}^r \rightarrow \mathbb{R}$  for  $i = 1, \dots, n$  be a sequence of measurable and bounded payoff functions. We define the sequence of value functions  $V_k : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  for each  $k \in \{1, \dots, n\}$  as

$$(6) \quad \begin{cases} V_k(t, x) := \sup_{u \in \mathcal{U}_t} \mathbb{E} \left[ \prod_{i=k}^n \ell_i(X_{\tau_i^k}^{t,x;u}) \right], \\ \tau_i^k(t, x) := \Theta_i^{A_k:n}(t, x) \wedge T_i, \quad i \in \{k, \dots, n\}, \end{cases}$$

where the stopping times  $(\Theta_i^{W_k:n})_{i=k}^n$  are sequential exit-times in the sense of Definition 3.1. Recall that the sequential exit-times of  $V_k$  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 function  $V$  (resp.  $\widetilde{V}$ ) in (3) is a particular case of the value function  $V_1$  defined as in (6) when  $A_i := W_i \setminus G_i$  (resp.  $A_i := W_i$ ),  $\ell_i := \mathbb{1}_{G_i}$  (resp.  $\ell_i := \mathbb{1}_{G_i \cap W_i}$ ), and  $T_i := T$ .

To state the main result of this section, Theorem 4.3 below, some technical definitions and assumptions concerning the stochastic processes  $X^{t,x;u}$ , the admissible controls  $\mathcal{U}_t$ , and the payoff functions  $\ell_i$  are needed:

**Assumption 4.1.** For all  $(t, x) \in \mathbb{S}$ ,  $\theta \in \mathcal{T}_{[t,T]}$ , and  $\mathbf{u}, \mathbf{v} \in \mathcal{U}_t$ , we stipulate the following assumptions on

**a. Admissible controls:**

$\mathcal{U}_t$  is the set of  $\mathbb{F}_t$ -progressively measurable processes taking values in a given control set. That is, the value of  $\mathbf{u} := (u_s)_{s \in [0, T]}$  at time  $s$  is a measurable mapping  $(z_{r \vee t} - z_t)_{r \in [0, s]} \mapsto u_s$  for all  $s \in [0, T]$ , see [KS91, Def. 1.11, p. 4] for more details.

**b. Stochastic process:**

i. Causality: If  $\mathbb{1}_{[t, \theta]} \mathbf{u} = \mathbb{1}_{[t, \theta]} \mathbf{v}$ , then we have  $\mathbb{1}_{[t, \theta]} X^{\cdot, x; \mathbf{u}} = \mathbb{1}_{[t, \theta]} X^{\cdot, x; \mathbf{v}}$ .

ii. Strong Markov property: For each  $\omega \in \Omega$  and the sample path  $(z_r)_{r \in [0, \theta(\omega)]}$  up to the stopping time  $\theta$ , let the random control  $\mathbf{u}_\theta \in \mathcal{U}_{\theta(\omega)}$  be the mapping  $(z_{\cdot \vee \theta(\omega)} - z_{\theta(\omega)}) \mapsto \mathbf{u}(z_{\cdot \wedge \theta(\omega)} + z_{\cdot \vee \theta(\omega)} - z_{\theta(\omega)}) =: \mathbf{u}_\theta$ .<sup>2</sup> Then, it holds with probability one that

$$\mathbb{E} \left[ \ell(X_{\theta+s}^{t, x; \mathbf{u}}) \mid \mathcal{F}_\theta \right] = \mathbb{E} \left[ \ell(X_{\theta+s}^{\bar{t}, \bar{x}; \bar{\mathbf{u}}}) \mid \bar{t} = \theta, \bar{x} = X_\theta^{t, x; \mathbf{u}}, \bar{\mathbf{u}} = \mathbf{u}_\theta \right],$$

for all bounded measurable functions  $\ell : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $s \geq 0$ .

iii. Continuity of the exit-times: Given initial condition  $(t_0, x_0) \in \mathbb{S}$ , for all  $k \in \{1, \dots, n\}$  and  $i \in \{k, \dots, n\}$  the stochastic mapping  $(t, x) \mapsto X_{\tau_i^k(t, x)}^{t, x; \mathbf{u}}$  is  $\mathbb{P}$ -a.s. continuous at  $(t_0, x_0)$  where the stopping times  $\tau_i^k$  are defined as in (6).

**c. Payoff functions:**

$(\ell_i)_{i=1}^n$  are lower semicontinuous for all  $i \leq n$ .

**Remark 4.2.** Some remarks on the above assumptions are in order:

- Assumption 4.1.a. implies that the admissible controls  $\mathbf{u} \in \mathcal{U}_t$  take action at time  $t$  independent of future information arriving at  $s > t$ . This is known as the non-anticipative strategy and is a standard assumption in stochastic optimal control [Bor05].
- Assumption 4.1.b. imposes three constraints on the process  $X^{\cdot, x; \mathbf{u}}$  defined on the prescribed probability space: i) causality of the solution processes for a given admissible control ii) strong Markov property iii) continuity of exit-time. The causality property is always satisfied in practical applications; uniqueness of the solution process  $X^{\cdot, x; \mathbf{u}}$  under any admissible control process  $\mathbf{u}$  guarantees it. The class of Strong Markov processes is fairly large; for instance, it contains the solution of SDEs under some mild assumptions on the drift and diffusion terms [Kry09, Thm. 2.9.4]. The almost sure continuity of the exit-time with respect to the initial condition of the process is the only potentially restrictive part of the assumptions. 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  $(t, x) \mapsto X_{\tau_i^k(t, x)}^{t, x; \mathbf{u}}$  at the given initial condition with probability one. We shall elaborate on this issue and its ramifications for a class of diffusion processes in Section 5.
- Assumption 4.1.c. imposes a fairly standard assumption on the payoff functions. In case  $\ell_i$  is the indicator function of a given set, for example in (3), this assumption requires the set to be open. This issue will be addressed in more details in Subsection 5.3, in particular to bring a reconciliation with Assumption 2.2.a..

<sup>2</sup>Notice that  $z_{\cdot} \equiv z_{\cdot \wedge \theta(\omega)} + z_{\cdot \vee \theta(\omega)} - z_{\theta(\omega)}$ . Thus, the randomness of  $\mathbf{u}_\theta$  is referred to the term  $z_{\cdot \wedge \theta(\omega)}$ .

Let function  $J_k : \mathbb{S} \times \mathcal{U}_0 \rightarrow \mathbb{R}$  be

$$(7) \quad J_k(t, x; \mathbf{u}) := \mathbb{E} \left[ \prod_{i=k}^n \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}}) \right],$$

where  $(\tau_i^k)_{i=k}^n$  are as defined in (6).

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 (6). 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 4.3** (Dynamic Programming Principle). *Consider the value functions  $(V_j)_{j=1}^n$  and the sequential stopping times  $(\tau_j^k)_{j=k}^n$  introduced in (6) where  $k \in \{1, \dots, n\}$ . Under Assumptions 4.1, for all  $(t, x) \in \mathbb{S}$  and family of stopping times  $\{\theta^{\mathbf{u}}, \mathbf{u} \in \mathcal{U}_t\} \subset \mathcal{T}_{[t, T]}$ , we have*

$$(8a) \quad V_k(t, x) \leq \sup_{\mathbf{u} \in \mathcal{U}_t} \mathbb{E} \left[ \sum_{j=k}^{n+1} \mathbf{1}_{\{\tau_{j-1}^k \leq \theta^{\mathbf{u}} < \tau_j^k\}} V_j^*(\theta^{\mathbf{u}}, X_{\theta^{\mathbf{u}}}^{t,x;\mathbf{u}}) \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}}) \right],$$

$$(8b) \quad V_k(t, x) \geq \sup_{\mathbf{u} \in \mathcal{U}_t} \mathbb{E} \left[ \sum_{j=k}^{n+1} \mathbf{1}_{\{\tau_{j-1}^k \leq \theta^{\mathbf{u}} < \tau_j^k\}} V_{j*}(\theta^{\mathbf{u}}, X_{\theta^{\mathbf{u}}}^{t,x;\mathbf{u}}) \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}}) \right],$$

where  $V_j^*$  and  $V_{j*}$  are, respectively, the upper and the lower semicontinuous envelope of  $V_j$ ,  $\tau_{k-1}^k := t$ ,  $V_{n+1} \equiv 1$ , and  $\tau_{n+1}^k$  is any constant time strictly greater than  $T$ , say  $\tau_{n+1}^k := T + 1$ .

*Proof.* See Appendix I. □

In our context the DPP proposed in Theorem 4.3 allows us to characterize the value function (6) through a sequence of value functions  $(V_j)_{j=k+1}^n$ . That is, (8a) and (8b) impose mutual constraints on a value function and subsequent value functions in the sequence. Moreover, the last function in the sequence is fixed to a constant by construction. Therefore, assuming that an algorithm to sequentially solve for these mutual constraints can be established, one could in principle use it to compute all value functions in the sequence and solve the original motion planning problem. In Section 5 we show how, for a class of controlled diffusion processes, the constraints imposed by (8) reduce to PDEs that the value functions need to satisfy. This enables the use of numerical PDE solution algorithms for this purpose.

**Remark 4.4** (Semicontinuity and Measurability). *Theorem 4.3 introduces DPP’s in the spirit of [BT11] but in an exit-time framework, which is in a weaker sense than the standard DPP in stochastic optimal control problems [FS06, Section IV.7]. Namely, the DPP only requires a semicontinuity of the payoff functions while it avoids the verification of the measurability of the value functions  $V_k$  in (3). Notice that in general this measurability issue is non-trivial due to the supremum operation running over possibly uncountably many controls.*

## 5. THE CASE OF CONTROLLED DIFFUSIONS

In this section we come to the last step in our construction. We demonstrate how the DPP derived in Section 4, in the context of controlled diffusion processes, gives rise to a series of PDE’s. Each PDE is understood in the discontinuous viscosity sense with boundary conditions in both Dirichlet (pointwise) and viscosity senses. This paves the way for using PDE numerical

solvers to numerically approximate the solution of our original motion planning problem for specific examples. We demonstrate an instance of such an example in Section 6.

We first introduce formally the standard probability space setup for SDEs, then proceed with some preliminaries to ensure that the requirements of the proposed DPP, Assumptions 4.1, hold. The section consists of subsections concerning PDE derivation and boundary conditions along with further discussions on how to deploy existing PDE solvers to numerically compute our PDE characterization.

Let  $\Omega$  be  $\mathcal{C}([0, T], \mathbb{R}^{z_d})$ , the set of continuous functions from  $[0, T]$  to  $\mathbb{R}^{z_d}$ , 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  $\Omega$  to which the process  $(z_t)_{t \geq 0}$  is adapted. Let us recall that  $\mathbb{F}_t := (\mathcal{F}_{t,s})_{s \geq 0}$  is the auxiliary subfiltration defined as  $\mathcal{F}_{t,s} := \sigma(z_{r \vee t} - z_t, r \in [0, s])$ . 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  $\mathbf{u} = (u_t)_{t \geq 0}$  we consider the  $\mathbb{R}^d$ -valued SDE<sup>3</sup>

$$(9) \quad dX_s = f(X_s, u_s) ds + \sigma(X_s, u_s) dW_s, \quad X_t = x,$$

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

**Assumption 5.1.** *We stipulate that*

- a.  $\mathbb{U} \subset \mathbb{R}^m$  is compact;
- b.  $f$  and  $\sigma$  are continuous and Lipschitz in first argument uniformly with respect to the second;
- c. The diffusion term  $\sigma$  of the SDE (9) is uniformly non-degenerate, i.e., there exists  $\delta > 0$  such that for all  $x \in \mathbb{R}^d$  and  $u \in \mathbb{U}$ ,  $\|\sigma(x, u)\sigma^\top(x, u)\| > \delta$ .

It is well-known that under Assumptions 5.1.a. and 5.1.b. there exists a unique strong solution to the SDE (9) [Bor05]; let us denote it by  $(X_s^{t,x;\mathbf{u}})_{s \geq t}$ . For future notational simplicity, we slightly modify the definition of  $X_s^{t,x;\mathbf{u}}$ , and extend it to the whole interval  $[0, T]$  where  $X_s^{t,x;\mathbf{u}} := x$  for all  $s$  in  $[0, t]$ .

In addition to Assumptions 5.1 on the SDE (9), we impose the following assumption on the motion planning sets that allows us to guarantee the continuity of sequential exit-times, as required for the DPP obtained in the preceding section.

**Assumption 5.2.** *The open sets  $(A_i)_{i=1}^n$  satisfy the exterior cone condition, i.e., for every  $i \in \{1, \dots, n\}$ , there are positive constants  $h, r$  and  $\mathbb{R}^d$ -value bounded map  $\eta : A_i^c \rightarrow \mathbb{R}^d$  such that*

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

where  $B_r(x)$  denotes an open ball centered at  $x$  and radius  $r$  and  $A_i^c$  stands for the complement of the set  $A_i$ .

**Remark 5.3** (Smooth Boundary). *If the set  $A_i$  is bounded and its boundary  $\partial A_i$  is smooth, then Assumption 5.2 holds. Furthermore, boundaries with corners may also satisfy Assumption 5.2; Figure 4 depicts two different examples.*

<sup>3</sup>We slightly abuse notation and earlier used  $\sigma$  for the sigma algebra as well. However, it will be always clear from the context to which  $\sigma$  we refer.

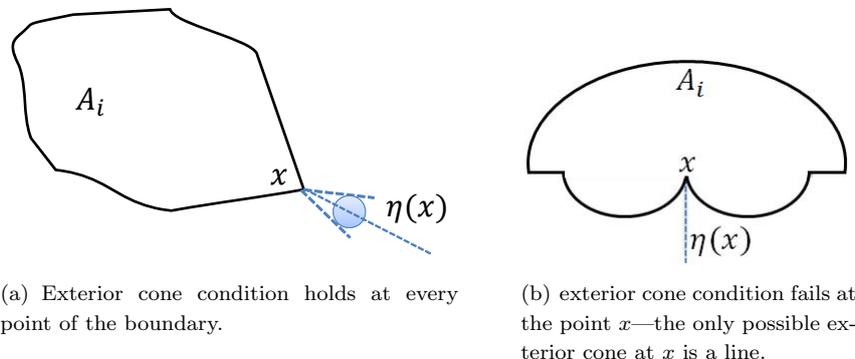


FIGURE 4. Exterior cone condition of the boundary

**5.1. Sequential PDEs.** In the context of SDEs, we show how the abstract DPP of Theorem 4.3 results in a sequence of PDEs, to be interpreted in the sense of discontinuous viscosity solutions; for the general theory of viscosity solutions we refer to [CIL92] and [FS06]. For numerical solutions to these PDEs, one also needs appropriate boundary conditions which is addressed in the next subsection.

To apply the proposed DPP, one has to make sure that Assumptions 4.1 are satisfied. As pointed out in Remark 4.2, the only nontrivial assumption in the context of SDEs is Assumption 4.1.b.iii. The following proposition addresses this issue, and allows us to employ the DPP of Theorem 4.3 for the main result of this subsection.

**Proposition 5.4** (Stopped-Process Continuity). *Consider the SDE (9) where Assumptions 5.1 hold. Suppose the open sets  $(A_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$  satisfy the exterior cone condition in Assumption 5.2. Let  $(\Theta_i^{A_{1:n}})_{i=1}^n$  be the respective sequential exit-times as defined in Definition 3.1. Given intermediate times  $(T_i)_{i=1}^n$ , and control  $\mathbf{u} \in \mathcal{U}_t$ , consider the stopping time  $\tau_i := \Theta_i^{A_{1:n}} \wedge T_i$ . Then, for any  $i \in \{1, \dots, n\}$ , initial condition  $(t, x) \in \mathbb{S}$ , and sequence of initial conditions  $(t_m, x_m) \rightarrow (t, x)$ , we have*

$$\lim_{m \rightarrow \infty} \tau_i(t_m, x_m) = \tau_i(t, x) \quad \mathbb{P}\text{-a.s.},$$

As a consequence, the stochastic mapping  $(t, x) \mapsto X_{\tau_i(t, x)}^{t, x; \mathbf{u}}$  is continuous with probability one, i.e.,  $\lim_{m \rightarrow \infty} X_{\tau_i(t_m, x_m)}^{t_m, x_m; \mathbf{u}} = X_{\tau_i(t, x)}^{t, x; \mathbf{u}}$   $\mathbb{P}$ -a.s. for all  $i$ .

*Proof.* See Appendix II. □

**Definition 5.5** (Dynkin Operator). *Given  $u \in \mathbb{U}$ , we denote by  $\mathcal{L}^u$  the Dynkin operator (also known as the infinitesimal generator [KS91, Chap. 5]) associated to the SDE (9) as*

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

where  $\Phi$  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  $x$ .

Theorem 5.6 is the main result of this subsection, which provides a characterization of the value functions  $V_k$  in terms of Dynkin operator in Definition 5.5 in the interior of the set of

interest, i.e.,  $[0, T_k[ \times A_k$ . We refer to [Kal97, Thm. 17.23] for details on the above differential operator.

**Theorem 5.6** (Dynamic Programming Equation). *Consider the system (9), and suppose that Assumptions 5.1 hold. Let the value functions  $V_k : \mathbb{S} \rightarrow \mathbb{R}^d$  be as defined in (6), where the sets  $(A_i)_{i=1}^n$  satisfy Assumption 5.2, and the payoff functions  $(\ell_i)_{i=1}^n$  are all lower semicontinuous. Then,*

◦  $V_{k*}$  is a viscosity supersolution of

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

◦  $V_k^*$  is a viscosity subsolution of

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

*Proof.* See Appendix II. □

**5.2. Boundary Conditions.** To numerically solve the PDE of Theorem 5.6, one needs boundary conditions on the complement of the set where the PDE is defined. This requirement is addressed in the following proposition.

**Proposition 5.7** (Boundary Conditions). *Suppose that the hypotheses of Theorem 5.6 hold. Then the value functions  $V_k$  introduced in (6) satisfy the following boundary value conditions:*

$$(10a) \quad \text{Dirichlet: } \begin{cases} V_k(t, x) = V_{k+1}(t, x) \ell_k(x) \\ \forall (t, x) \in [0, T_k] \times A_k^c \cup \{T_k\} \times \mathbb{R}^d \end{cases}$$

$$(10b) \quad \text{Viscosity: } \begin{cases} \limsup_{\substack{A_k \ni x' \rightarrow x \\ t' \uparrow t}} V_k(t', x') \leq V_{k+1}^*(t, x) \ell_k^*(x) \\ \liminf_{\substack{A_k \ni x' \rightarrow x \\ t' \uparrow t}} V_k(t', x') \geq V_{k+1*}(t, x) \ell_k(x) \\ \forall (t, x) \in [0, T_k] \times \partial A_k \cup \{T_k\} \times \bar{A}_k \end{cases}$$

*Proof.* See Appendix II. □

Proposition 5.7 provides boundary condition for  $V_k$  in both Dirichlet (pointwise) and viscosity senses. The Dirichlet boundary condition (10a) is the one usually employed to numerically compute the solution via PDE solvers, whereas the viscosity boundary condition (10b) is required for theoretical support of the numerical schemes and comparison results.

**Remark 5.8** (Lower Semicontinuity). *In the SDE setting, one can, without loss of generality, extend the class of admissible controls in the definition of  $V_k$  to  $\mathcal{U}_0$ , i.e.,  $V_k(t, x) = \sup_{\mathbf{u} \in \mathcal{U}_0} J_k(t, x; \mathbf{u})$ ; for a rigorous technology to prove this assertion see [BT11, Remark 5.2]. Thus,  $V_k$  is lower semicontinuous as it is a supremum over a fixed family of lower semicontinuous functions, see Lemma I.3 in Appendix I. In this light, one may argue that in the viscosity boundary condition (10b), the second assertion is subsumed by the Dirichlet boundary condition (10a).*

**5.3. Discussion on Numerical Issues.** For the class of controlled diffusion processes (9), Subsection 5.1 developed a PDE characterization of the value function  $V_k$  within the set  $[0, T_k[ \times A_k$  along with boundary conditions in terms of the successor value function  $V_{k+1}$  provided in Subsection 5.2. Since  $V_{n+1} \equiv 1$ , one can infer that Theorem 5.6 and Proposition 5.7 provide a series of PDE where the last one has known boundary condition, while the boundary conditions of earlier in the sequence are determined by the solution of subsequent PDE, 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 [MCL15].

Before proceeding with numerical solutions, we need to properly address two technical 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 2.2.a. and the following discussion. On the other hand, in order to invoke the DPP argument of Section 4 and its consequent PDE in Subsection 5.1, we need to impose that the payoff functions  $(\ell_i)_{i=1}^n$  are all lower semicontinuous; see Assumption 4.1.c. In the case of the value function  $V$  in (3a), this constraint results in  $(G_i)_{i=1}^n$  all being open, which in general contradicts Assumption 2.2.a..
- (ii) Most of the existing PDE solvers provide theoretical guarantees for *continuous* viscosity solutions, e.g., [Mit05]. Theorem 5.6, on the other hand, characterizes the solution to the motion planning problem in terms of *discontinuous* viscosity solutions. Therefore, it is a natural question whether we could employ any of available numerical methods to approximate the solution of 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 (3) are in general discontinuous, and any PDE approach has to rely on discontinuous versions.

To address the above concerns, we propose an  $\varepsilon$ -conservative but precise way of characterizing the motion planning initial set. Given  $(W_i, G_i) \in \mathfrak{B}(\mathbb{R}^d) \times \mathfrak{B}(\mathbb{R}^d)$ , let us construct a smaller goal set  $G_i^\varepsilon \subset G_i$  such that  $G_i^\varepsilon := \{x \in G_i \mid \text{dist}(x, G_i^c) \geq \varepsilon\}$ .<sup>4</sup> For sufficiently small  $\varepsilon > 0$  one may observe that  $W_i \setminus G_i^\varepsilon$  satisfies Assumption 5.2. Note that this is always possible if  $W_i \setminus G_i$  satisfies Assumption 5.2 since one can simply take  $\varepsilon < h/2$ , where  $h$  is as defined in Assumption 5.2. Figure 5 depicts this situation.

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

$$\ell_i^\varepsilon(x) := \left(1 - \frac{\text{dist}(x, G_i^c)}{\varepsilon}\right) \vee 0.$$

Replacing the goal sets  $G_i^\varepsilon$  and payoff functions  $\ell_i^\varepsilon$  in (3a), we arrive at the value function

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

<sup>4</sup> $\text{dist}(x, A) := \inf_{y \in A} \|x - y\|$ , where  $\|\cdot\|$  stands for the Euclidean norm.

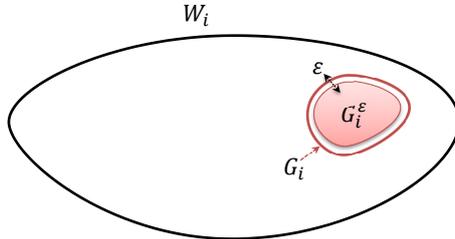


FIGURE 5. Construction of the sets  $G_i^\varepsilon$  from  $G_i$  as described in Subsection 5.3

It is straightforward to inspect that  $V^\varepsilon \leq V$  since  $G_i^\varepsilon \subset G_i$ . Moreover, with a similar technique as in [MCL15, Thm. 5.1], one may show that  $V(t, x) = \lim_{\varepsilon \downarrow 0} V^\varepsilon(t, x)$  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  $\ell_i^\varepsilon$  are, by construction, Lipschitz continuous that in light of uniform continuity of the process, Lemma II.2 in Appendix II, leads to the continuity of the value function  $V^\varepsilon$ .<sup>5</sup> Hence, the discontinuous PDE characterization of Subsection 5.1 can be approximated arbitrarily closely in the continuous regime.

Let us recall that having reduced the motion planning problems to PDEs, numerical methods and computational algorithms exist to approximate its solution [Mit05]. In Section 6 we demonstrate how to use such methods to address practically relevant problems. In practice, such methods are effective for systems of relatively small dimension due to the curse of dimensionality. To alleviate this difficulty and extend the method to large problems, we can leverage on ADP [dFVR03, CRVRL06] or other advances in numerical mathematics, such as tensor trains [KS11]. The link between motion planning and the PDEs through DPP is precisely what allows us to capitalize on any such developments in the numerics.

## 6. NUMERICAL EXAMPLE: CHEMICAL LANGEVIN EQUATION FOR A BIOLOGICAL SWITCH

When modeling uncertainty in biochemical reactions, one often resorts to countable Markov chain models [Wil06] 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 ordinary differential equations called the *chemical master equation* (CME) [AGA09, ESKPG05], also known as the forward Kolmogorov equation.

Though close to the physical reality, the CME is particularly difficult to work with analytically. One therefore typically employs different approximate solution methods, for example the Finite State Projection method [Kha10] or the moment closure method [SH07]. Such approximation method resorts to approximating discrete molecule numbers by a continuum and capturing the stochasticity in their evolution through a stochastic differential equation. This stochastic continuous-time approximation is called the *chemical Langevin equation* or the *diffusion approximation*, see for example [Kha10] and the references therein. The Langevin approximation can be inaccurate for chemical species with low copy numbers; it may even assign a negative number to some molecular species. To circumvent this issue we assume here that the species of interest come in sufficiently high copy numbers to make the Langevin approximation reasonable.

<sup>5</sup>This continuity result can, alternatively, be deduced via the comparison result of the viscosity characterization of Theorem 5.6 together with boundary conditions (10b) [CIL92].

Multistable biological systems are often encountered in nature [BSS01]. In this section we consider the following chemical Langevin formulation of a bistable two gene network:

$$(11) \quad \begin{cases} dX_t = (f(Y_t, \mathbf{u}_x) - \mu_x X_t)dt + \sqrt{f(Y_t, \mathbf{u}_x)}dW_t^1 + \sqrt{\mu_x X_t}dW_t^2, \\ dY_t = (g(X_t, \mathbf{u}_y) - \mu_y Y_t)dt + \sqrt{g(X_t, \mathbf{u}_y)}dW_t^3 + \sqrt{\mu_y Y_t}dW_t^4, \end{cases}$$

where  $X_t$  and  $Y_t$  are the concentration of the two repressor proteins with the respective degradation rates  $\mu_x$  and  $\mu_y$ ;  $(W_t^i)_{t \geq 0}$  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  $\mathbf{u}_x$  and  $\mathbf{u}_y$ .

In the absence of exogenous control signals, the authors of [Che00] study sufficient conditions on the drifts  $f$  and  $g$  under which the system dynamic (11) without the diffusion term has two (or more) stable equilibria. In this case, system (11) can be viewed as a biological switch network. The theoretical results of [Che00] are also experimentally investigated in [GCC00] for a genetic toggle switch in *Escherichia coli*.

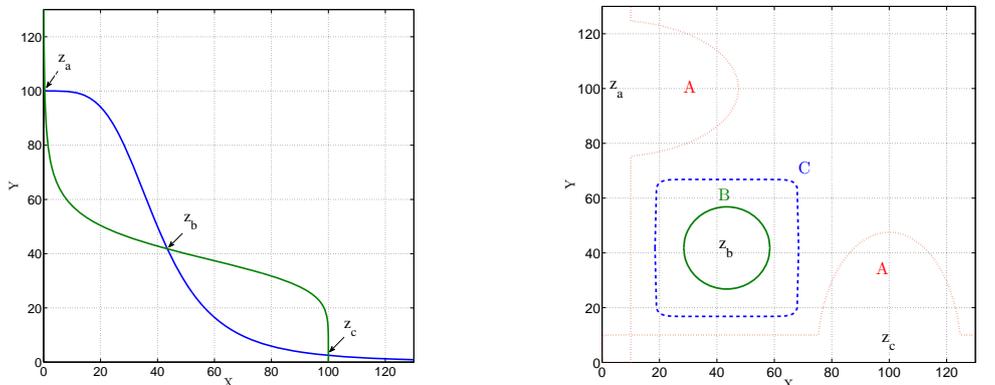
Here we consider the biological switch dynamics where the production rates of proteins are influenced by external control signals; experimental constructs that can be used to provide such inputs have recently been reported in the literature [MSSO<sup>+</sup>11]. The level of repression is described by a *Hill* function, which models cooperativity of binding as follows:

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

where  $\theta_i$  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  $\mathbb{U}_x := [\underline{u}_x, \bar{u}_x]$  and  $\mathbb{U}_y := [\underline{u}_y, \bar{u}_y]$ . In this example we consider system (11) 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 6(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 (11) is understood in the absence of the diffusion term as the noise may very well push the states from one stable equilibrium to another.

Driving bistable systems away from the commonly observed states may provide useful insights into their functional organization. Motivated by this, in this example we consider a motion planning task comprising two phases. First, we aim to steer the number of proteins to be at a target set around the unstable equilibrium at a certain time horizon, say  $T_1$ , by synthesizing appropriate input signals  $\mathbf{u}_x$  and  $\mathbf{u}_y$ . During this phase we opt to avoid the region of attraction of the stable equilibria as well as low numbers for each protein; the latter justifies our Langevin model being well-posed in the region of interest. These target and avoid sets are denoted, respectively, by the closed sets  $B$  and  $A$  in Figure 6(b). In the second phase of the task, it is required to keep the molecular populations within a larger margin around the unstable equilibrium till sometime after  $T_1$ , say  $T_2 > T_1$ ; Figure 6(b) depicts this maintenance margin by the open set  $C$ . In the context of reachability, the second phase is known as *viability* [AP98].

In view of motion planning events introduced in Definition 2.1, in particular (1b), the above motion planning task can be expressed by  $(A^c \xrightarrow{T_1} B) \circ (C \xrightarrow{T_2} C)$ . Let  $Z^{t,z;\mathbf{u}} := [X^{t,x;\mathbf{u}}, Y^{t,y;\mathbf{u}}]$  be the joint process with initial condition  $z := [x, y]$  and controller  $\mathbf{u} := [\mathbf{u}_x, \mathbf{u}_y]$ . Following the



(a) Nullclines and equilibria of the drift of the SDE (11);  $x_a$  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.

FIGURE 6. State space of the biological switch (11) with desired motion planning sets.

framework of Section 4, we consider the following value functions:

$$(12a) \quad V_1(t, z) := \sup_{\mathbf{u} \in \mathcal{U}_t} \mathbb{E} \left[ \mathbf{1}_B(Z_{\tau_1^1}^{t,z;\mathbf{u}}) \mathbf{1}_C(Z_{\tau_2^1}^{t,z;\mathbf{u}}) \right],$$

$$(12b) \quad V_2(t, z) := \sup_{\mathbf{u} \in \mathcal{U}_t} \mathbb{E} \left[ \mathbf{1}_C(Z_{\tau_2^2}^{t,z;\mathbf{u}}) \right],$$

where the stopping times  $\tau_1^1, \tau_2^1, \tau_2^2$  are defined as in (6) with sets  $A_1 := A^c$  and  $A_2 := C$ . Let us recall that by Theorem 3.2 the desired set of initial conditions in Definition 2.3 is indeed the  $p$  superlevel set of  $V_1$ . The solution of the motion planning objective is the value function  $V_1$  in (12a), which in view of Theorem 5.6 is characterized by the Dynkin PDE operator in the interior of  $[0, T_1] \times A^c$ . However, due to the boundary condition (10a), we first need to compute  $V_2$  in (12b) to provide boundary conditions for  $V_1$  according to

$$(13) \quad \begin{cases} V_1(t, z) = V_2(t, z) \mathbf{1}_B(z), \\ \forall(t, z) \in [0, T_1] \times A \cup \{T_1\} \times \mathbb{R}^2. \end{cases}$$

Observe that the boundary condition for the value function  $V_2$  is

$$V_2(t, z) = \mathbf{1}_C(z), \quad \forall(t, z) \in [0, T_2] \times C^c \cup \{T_2\} \times \mathbb{R}^2.$$

Therefore, we need to solve the PDE of  $V_2$  backward from the terminal time  $T_2$  to  $T_1$  together with the above boundary condition. Then, the value function  $V_1$  can be computed via solving the same PDE from  $T_1$  to 0 with the boundary condition (13). The Dynkin operator  $\mathcal{L}^u$  reduces to

$$\begin{aligned} \sup_{\mathbf{u} \in \mathbb{U}} \mathcal{L}^u \phi(t, x, y) &= \max_{\mathbf{u} \in \mathbb{U}} \left[ \partial_t \phi + \partial_x \phi (f(y, u_x) - \mu_x x) + \partial_y \phi (g(x, u_y) - \mu_y y) \right. \\ &\quad \left. + \frac{1}{2} \partial_x^2 \phi (f(y, u_x) + \mu_x x) + \frac{1}{2} \partial_y^2 \phi (g(x, u_y) + \mu_y y) \right] \end{aligned}$$

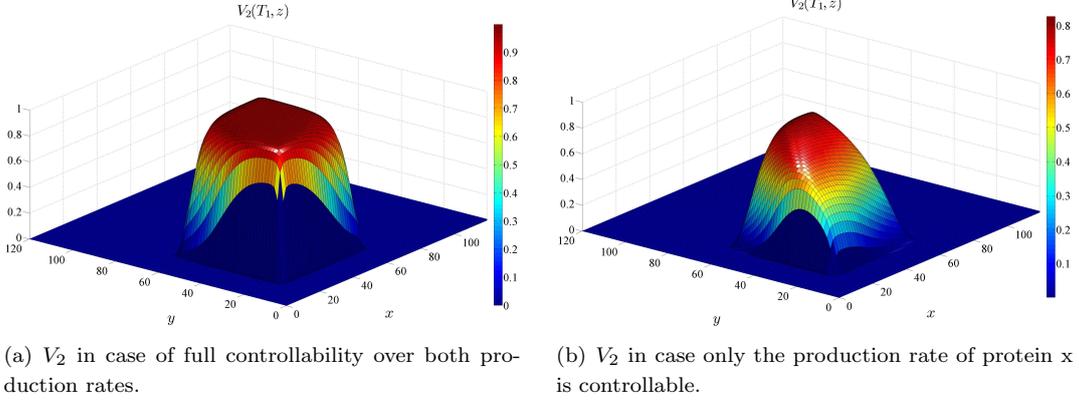


FIGURE 7. The value function  $V_2$  as defined in (12b) corresponding to the probability of staying in  $C$  for 60 time units.

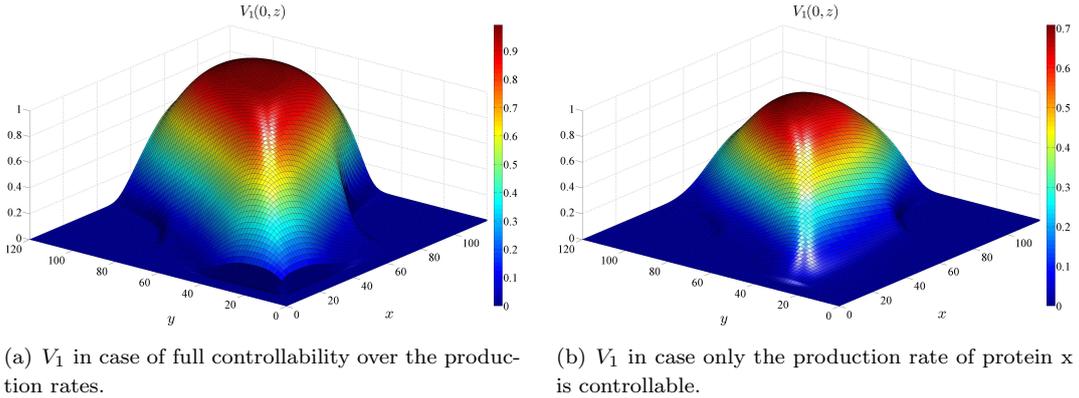


FIGURE 8. The value function  $V_1$  as defined in (12a) corresponding to the probability of being at  $B$  at 20 time units and staying in  $C$  afterward till 80 time units, while avoiding  $A$  throughout the motion.

$$\begin{aligned}
 &= \partial_t \phi - \left( \partial_x \phi - \frac{1}{2} \partial_x^2 \phi \right) \mu_x x - \left( \partial_y \phi - \frac{1}{2} \partial_y^2 \phi \right) \mu_y y \\
 &\quad + \max_{u_x \in [\underline{u}_x, \bar{u}_x]} \left[ f(y, u_x) \left( \partial_x \phi + \frac{1}{2} \partial_x^2 \phi \right) \right] + \max_{u_y \in [\underline{u}_y, \bar{u}_y]} \left[ g(x, u_y) \left( \partial_y \phi + \frac{1}{2} \partial_y^2 \phi \right) \right].
 \end{aligned}$$

Thanks to the linearity of the drift term in  $\mathbf{u}$ , an optimal control can be expressed in terms of derivatives of the value functions  $V_1$  and  $V_2$  as

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

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

For this system we investigate two scenarios: One where full control over both production rates is possible, and one where only the production rate of protein  $x$  can be controlled. Accordingly, in the first scenario we set  $\underline{u}_x = \underline{u}_y = 0$  and  $\bar{u}_x = \bar{u}_y = 2$  while in the second we set  $\underline{u}_x = 0$ ,  $\bar{u}_x = 2$  and  $\bar{u}_y = \underline{u}_y = 1$ . Figure 7 depicts the probability distribution of staying in set  $C$  within the time horizon  $[T_1, T_2] = [20, 80]$  for 60 time units<sup>6</sup> in terms of the initial conditions  $(x, y) \in \mathbb{R}^2$ . Value function  $V_2$  is zero outside set  $C$ , as the process has obviously left  $C$  if it starts outside it. Figures 7(a) and 7(b) demonstrate the first and second scenarios, respectively. Note that in the second case the probability of success dramatically decreases in comparison to the first. This result indicates the importance of full controllability of the production rates for the achievement of the desired control objective.

Figure 8 depicts the probability of successively being at set  $B$  at  $T_1 = 60$  time units while avoiding  $A$ , and staying in set  $C$  till  $T_2 = 80$  time units thereafter. Since the objective is to avoid  $A$  throughout the motion, the value function  $V_1$  takes zero value on  $A$ . Figures 8(a) and 8(b) demonstrate the first and second control scenarios, respectively. It is easy to observe the non-smooth behavior of the value function  $V_1$  on the boundary of set  $A$  in Figure 8(b). All simulations in this subsection were obtained using the Level Set Method Toolbox [Mit05] (version 1.1), with a grid  $121 \times 121$  in the region of interest.

## 7. CONCLUSION AND FUTURE DIRECTIONS

We introduced different notions of stochastic motion planning problems. Based on a class of stochastic optimal control problems, we characterized the set of initial conditions from which there exists an admissible control to execute the desired maneuver with probability no less than some pre-specified value. We then established a weak DPP in terms of auxiliary value functions. Subsequently, we focused on a case of diffusions as the solution of a controlled SDE, and investigated the required conditions to apply the proposed DPP. It turned out that invoking the DPP one can solve a series of PDEs in a recursive fashion to numerically approximate the desired initial set as well as the admissible control for the motion planning specifications. Finally, the performance of the proposed stochastic motion planning notions was illustrated for a biological switch network.

For future work, as Theorem 4.3 holds for the broad class of stochastic processes whose sample paths are right continuous with left limits, we aim to study the required conditions of the proposed DPP (Assumptions 4.1) for a larger class of stochastic processes, e.g., controlled Markov jump-diffusions. Furthermore, motivated by the fact that full state measurements may not be available in practice, an interesting question is to address the motion planning objective with imperfect information, i.e., an admissible control would be only allowed to utilize the information of the process  $Y_s := h(X_s)$  where  $h : \mathbb{R}^d \rightarrow \mathbb{R}^{d_y}$  is a given measurable mapping.

### I. APPENDIX: TECHNICAL PROOFS OF SECTIONS 3 & 4

We first start with a rather technical lemma showing that the sequential stopping times in Definition 3.1 are indeed well defined.

---

<sup>6</sup>Notice that the half-life of each protein is assumed to be 17.32 time units

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

*Proof.* Let  $\tau_A$  be the first exit-time from the set  $A_i$ :

$$(I.1) \quad \tau_{A_i}(t, x) := \inf\{s \geq 0 : X_{t+s}^{t,x;\mathbf{u}} \notin A_i\}.$$

We know that  $\tau_A$  is an  $\mathbb{F}_t$ -stopping time [EK86, Thm. 1.6, Chapter 2]. Let  $\omega(\cdot) \mapsto \vartheta_s(\omega(\cdot)) := \omega(s + \cdot)$  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  $\vartheta$  and right continuity of the filtration  $\mathbb{F}_t$ ; see [EK86, Prop. 1.4, Chapter 2] for more details in this regard.  $\square$

Before proceeding with the proof of Theorem 3.2, we start with a fact which is an immediate consequence of right continuity of the process  $X^{t,x;\mathbf{u}}$ :

**Fact I.2.** *Fix a control  $\mathbf{u} \in \mathcal{U}_t$  and an initial condition  $(t, x) \in \mathbb{S}$ . Let  $(A_i)_{i=1}^n \subset \mathfrak{B}(\mathbb{R}^d)$  be a sequence of open sets. Then, for all  $i \in \{1, \dots, n\}$*

$$X_{\Theta_i^{A_{1:n}}}^{t,x;\mathbf{u}} \notin A_i, \quad \text{on } \{\Theta_i^{A_{1:n}} < \infty\},$$

where  $(\Theta_i^{A_{1:n}})_{i=1}^n$  are the sequential exit-times in the sense of Definition 3.1.

*Proof of Theorem 3.2.* We first show (4). Observe that it suffices to prove that

$$(I.2) \quad \left\{ X^{t,x;\mathbf{u}} \models [(W_1 \rightsquigarrow G_1) \circ \dots \circ (W_n \rightsquigarrow G_n)]_{\leq T} \right\} = \bigcap_{i=1}^n \left\{ X_{\eta_i}^{t,x;\mathbf{u}} \in G_i \right\}$$

for all initial condition  $(t, x)$  and control  $\mathbf{u}$ , where the stopping time  $\eta_i$  is as defined in (3a). Let  $\omega$  belong to the left-hand side of (I.2). In view of the definition (1a), there exists a set of instants  $(s_i)_{i=1}^n \subset [t, T]$  such that for all  $i$ ,  $X_{s_i}^{t,x;\mathbf{u}}(\omega) \in G_i$  while  $X_r^{t,x;\mathbf{u}}(\omega) \in W_i \setminus G_i =: B_i$  for all  $r \in [s_{i-1}, s_i[$ , where we set  $s_0 = t$ . It then follows by an induction argument that  $\eta_i(\omega) = \Theta_i^{B_{1:n}} = s_i$ , which immediately leads to  $X_{\eta_i(\omega)}^{t,x;\mathbf{u}}(\omega) \in G_i$  for all  $i \leq n$ . This proves the relation “ $\subset$ ” between the left- and right-hand sides of (I.2). Now suppose that  $\omega$  belongs to the right-hand side of (I.2). Then, we have  $X_{\eta_i(\omega)}^{t,x;\mathbf{u}}(\omega) \in G_i$  for all  $i \leq n$ . In view of the definition of stopping times  $\eta_i$  in (3a), it follows that  $X_r^{t,x;\mathbf{u}}(\omega) \in B_i := W_i \setminus G_i$  for all  $r \in [\eta_{i-1}(\omega), \eta_i(\omega)[$ . Introducing the time sequence  $s_i := \eta_i(\omega)$  implies the relation “ $\supset$ ” between the left- and right-hand sides of (I.2). Together with preceding argument, this implies (I.2).

To prove (5) we only need to show that

$$(I.3) \quad \left\{ X^{t,x;\mathbf{u}} \models (W_1 \xrightarrow{T_1} G_1) \circ \dots \circ (W_n \xrightarrow{T_n} G_n) \right\} = \bigcap_{i=1}^n \left\{ X_{\tilde{\eta}_i}^{t,x;\mathbf{u}} \in G_i \cap W_i \right\}$$

for all initial condition  $(t, x)$  and controls  $\mathbf{u}$ , where the stopping time  $\tilde{\eta}_i$  is introduced in (3b). To this end, let us fix  $(t, x) \in \mathbb{S}$  and  $\mathbf{u} \in \mathcal{U}_t$ , and assume that  $\omega$  belongs to the left-hand side of (I.3). By definition (1b), for all  $i \leq n$  we have  $X_{T_i}^{t,x;\mathbf{u}}(\omega) \in G_i$  and  $X_r^{t,x;\mathbf{u}}(\omega) \in W_i$  for all  $r \in [T_{i-1}, T_i]$ . By a straightforward induction, we see that  $\tilde{\eta}_i(\omega) = T_i$ , and consequently  $X_{\tilde{\eta}_i(\omega)}^{t,x;\mathbf{u}}(\omega) \in G_i \cap W_i$  for all  $i \leq n$ . This establishes the relation “ $\subset$ ” between the left- and right-hand sides of (I.3). Now suppose  $\omega$  belongs to the right-hand side of (I.3). Then, for all  $i \leq n$  we have  $X_{\tilde{\eta}_i(\omega)}^{t,x;\mathbf{u}}(\omega) \in G_i \cap W_i$ . By virtue of Fact I.2 and an induction argument once again, it is

guaranteed that  $\tilde{\eta}_i(\omega) = T_i$ , and consequently it follows that  $X_{T_i}^{t,x;\mathbf{u}}(\omega) \in G_i$  and  $X_r^{t,x;\mathbf{u}}(\omega) \in W_i$  for all  $r \in [T_{i-1}, T_i]$ . This establishes the relation “ $\supset$ ” in (I.3), and the assertion follows.  $\square$

We now continue with the missing proof of Section 4. Before proceeding with the proof of Theorem 4.3, we need a preparatory lemma.

**Lemma I.3.** *Under Assumptions 4.1.b.iii. and 4.1.c., the function  $\mathbb{S} \ni (t, x) \mapsto J_k(t, x; \mathbf{u}) \in \mathbb{R}$  is lower semicontinuous for all  $k \in \{1, \dots, n\}$  and control  $\mathbf{u} \in \mathcal{U}_0$ .*

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

$$(I.4) \quad \begin{aligned} \liminf_{(s,y) \rightarrow (t,x)} J_k(s, y; \mathbf{u}) &= \liminf_{(s,y) \rightarrow (t,x)} \mathbb{E} \left[ \prod_{i=k}^n \ell_i(X_{\tau_i^k}^{s,y;\mathbf{u}}) \right] \\ &\geq \mathbb{E} \left[ \liminf_{(s,y) \rightarrow (t,x)} \prod_{i=k}^n \ell_i(X_{\tau_i^k}^{s,y;\mathbf{u}}) \right] \end{aligned}$$

$$(I.5) \quad \begin{aligned} &\geq \mathbb{E} \left[ \prod_{i=k}^n \liminf_{(s,y) \rightarrow (t,x)} \ell_i(X_{\tau_i^k}^{s,y;\mathbf{u}}) \right] \\ &\geq \mathbb{E} \left[ \prod_{i=k}^n \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}}) \right] = J_k(t, x; \mathbf{u}), \end{aligned}$$

where the inequality in (I.4) follows from the Fatou’s lemma, and (I.5) is a direct consequence of Assumptions 4.1.b.iii. and 4.1.c.  $\square$

*Proof of Theorem 4.3.* The proof extends the main result of our earlier work [MCL15, Thm. 4.7] on the so-called reach-avoid maneuver. Let  $u \in \mathcal{U}_t$ ,  $\theta := \theta^{\mathbf{u}} \in \mathcal{T}_{[t,T]}$ , and  $\mathbf{u}_\theta$  be the random control as introduced in Assumption 4.1.b.ii. Then we have

$$(I.6a) \quad \begin{aligned} &\mathbb{E} \left[ \prod_{i=k}^n \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}}) \mid \mathcal{F}_\theta \right] \\ &= \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \mathbb{E} \left[ \prod_{i=j}^n \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}}) \mid \mathcal{F}_\theta \right] \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}}) \end{aligned}$$

$$(I.6b) \quad = \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} J_j(\theta, X_\theta^{t,x;\mathbf{u}}; \mathbf{u}_\theta) \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}})$$

$$(I.6c) \quad \leq \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} V_j(\theta, X_\theta^{t,x;\mathbf{u}}) \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t,x;\mathbf{u}})$$

where (I.6b) follows from Assumption 4.1.b.ii. and right continuity of the process, and (I.6c) is due to the fact that  $\mathbf{u}_\theta \in \mathcal{U}_{\theta(\omega)}$  for each realization  $\omega \in \Omega$ . In light of the tower property of conditional expectation [Kal97, Thm. 5.1], arbitrariness of  $\mathbf{u} \in \mathcal{U}_t$ , and obvious inequality  $V_j \leq V_j^*$ , we arrive at (8a).

To prove (8b), consider uniformly bounded upper semicontinuous functions  $(\phi_j)_{j=k}^n$  such that  $\phi_j \leq V_{j*}$  on  $\mathbb{S}$ . Mimicking the ideas in the proof of our earlier work [MCL15, Thm. 4.7] and due to Lemma I.3, one can construct an admissible control  $\mathbf{u}_j^\varepsilon$  for any  $\varepsilon > 0$  and  $j \in \{k, \dots, n\}$  such

that

$$(I.7) \quad \phi_j(t, x) - 3\varepsilon \leq J_j(t, x; \mathbf{u}_j^\varepsilon) \quad \forall (t, x) \in \mathbb{S}.$$

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

$$(I.8) \quad \mathbf{v}^\varepsilon := \mathbb{1}_{[t, \theta]} \mathbf{u} + \mathbb{1}_{] \theta, T]} \sum_{j=k}^n \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \mathbf{u}_j^\varepsilon,$$

where  $\mathbf{u}_j^\varepsilon$  satisfies (I.7). Notice that Assumption 4.1.a. ensures  $\mathbf{v}^\varepsilon \in \mathcal{U}_t$ . By virtue of the tower property, Assumptions 4.1.b.i. and 4.1.b.ii., and the assertions in (I.7) and (I.8), it follows

$$\begin{aligned} V_k(t, x) &\geq J_k(t, x; \mathbf{v}^\varepsilon) = \mathbb{E} \left[ \mathbb{E} \left[ \prod_{i=k}^n \ell_i(X_{\tau_i^k}^{t, x; \mathbf{v}^\varepsilon}) \mid \mathcal{F}_\theta \right] \right] \\ &= \mathbb{E} \left[ \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} J_j(\theta, X_\theta^{t, x; \mathbf{u}}; \mathbf{u}_j^\varepsilon) \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t, x; \mathbf{u}}) \right] \\ &= \mathbb{E} \left[ \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \left( \phi_j(\theta, X_\theta^{t, x; \mathbf{u}}) - 3\varepsilon \right) \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t, x; \mathbf{u}}) \right]. \end{aligned}$$

Now, consider a sequence of increasing continuous functions  $(\phi_j^m)_{m \in \mathbb{N}}$  that converges point-wise to  $V_{j*}$ . The existence of such sequence is ensured by Lemma 1.3, see [Ren99, Lemma 3.5]. By boundedness of  $(\ell_j)_{i=1}^n$  and the dominated convergence Theorem, we get

$$\begin{aligned} V_k(t, x) &\geq \\ &\mathbb{E} \left[ \sum_{j=k}^{n+1} \mathbb{1}_{\{\tau_{j-1}^k \leq \theta < \tau_j^k\}} \left( V_{j*}(\theta, X_\theta^{t, x; \mathbf{u}}) - 3\varepsilon \right) \prod_{i=k}^{j-1} \ell_i(X_{\tau_i^k}^{t, x; \mathbf{u}}) \right] \end{aligned}$$

Since  $\mathbf{u} \in \mathcal{U}_t$  and  $\varepsilon > 0$  are arbitrary, this leads to (8b).  $\square$

## II. APPENDIX: TECHNICAL PROOFS OF SECTION 5

*Proof of Proposition 5.4.* The key step in the proof relies on the two Assumptions 5.1.c. and 5.2. There is a classical result on non-degenerate diffusion processes indicating that if the process starts from the tip of a cone, then it enters the cone with probability one [RB98, Corollary 3.2, p. 65]. This hints at the possibility that the aforementioned Assumptions together with almost sure continuity of the strong solution of the SDE (9) result in the continuity of sequential exit-times  $\Theta_i^{A_1:n}$  and consequently  $\tau_i$ . In the following we shall formally work around this idea.

Let us assume that  $t_m \leq t$  for notational simplicity, but one can effectively follow similar arguments for  $t_m > t$ . By the definition of the SDE (9),

$$\begin{aligned} X_r^{t_m, x_m; \mathbf{u}} &= X_t^{t_m, x_m; \mathbf{u}} + \int_t^r f(X_s^{t_m, x_m; \mathbf{u}}, u_s) ds \\ &\quad + \int_t^r \sigma(X_s^{t_m, x_m; \mathbf{u}}, u_s) dW_s, \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

By virtue of [Kry09, Thm. 2.5.9, p. 83], for all  $q \geq 1$  we have

$$\begin{aligned} &\mathbb{E} \left[ \sup_{r \in [t, T]} \|X_r^{t, x; \mathbf{u}} - X_r^{t_m, x_m; \mathbf{u}}\|^{2q} \right] \\ &\leq C_1(q, T, K) \mathbb{E} \left[ \|x - X_t^{t_m, x_m; \mathbf{u}}\|^{2q} \right] \end{aligned}$$

$$\leq 2^{2q-1}C_1(q, T, K)\mathbb{E}\left[\|x - x_m\|^{2q} + \|x_m - X_t^{t_m, x_m; \mathbf{u}}\|^{2q}\right],$$

whence, in light of [Kry09, Corollary 2.5.12, p. 86], we get

$$(II.1) \quad \mathbb{E}\left[\sup_{r \in [t, T]} \|X_r^{t, x; \mathbf{u}} - X_r^{t_n, x_n; \mathbf{u}}\|^{2q}\right] \leq C_2(q, T, K, \|x\|)(\|x - x_n\|^{2q} + |t - t_n|^q).$$

In the above inequalities,  $K$  is the Lipschitz constant of  $f$  and  $\sigma$  mentioned in Assumption 5.1.b.;  $C_1$  and  $C_2$  are constant depending on the indicated parameters. Hence, in view of Kolmogorov's continuity criterion [Pro05, Corollary 1 Chap. IV, p. 220], one may consider a version of the stochastic process  $X^{t, x; \mathbf{u}}$  which is continuous in  $(t, x)$  in the topology of uniform convergence on compacts. This leads to the fact that  $\mathbb{P}$ -a.s, for any  $\varepsilon > 0$ , for all sufficiently large  $m$ ,

$$(II.2) \quad X_r^{t_m, x_m; \mathbf{u}} \in B_\varepsilon(X_r^{t_0, x_0; \mathbf{u}}), \quad \forall r \in [t_m, T],$$

where  $B_\varepsilon(y)$  denotes the ball centered at  $y$  and radius  $\varepsilon$ . For simplicity, let us define the shorthand  $\tau_i^m := \tau_i(t_m, x_m)$ .<sup>7</sup> By the definition of  $\tau_i$  and Definition 3.1, since the set  $A_i$  is open, we conclude that

$$(II.3) \quad \exists \varepsilon > 0, \quad \bigcup_{s \in [\tau_{i-1}^0, \tau_i^0[} B_\varepsilon(X_s^{t_0, x_0; \mathbf{u}}) \cap A_i^c = \emptyset \quad \mathbb{P}\text{-a.s.}$$

By definition  $\tau_0^0 := \tau_0(t_0, x_0) = t_0$ . As an induction hypothesis, let us assume  $\tau_{i-1}^0$  is  $\mathbb{P}$ -a.s. continuous, and we proceed with the induction step. One can deduce that (II.3) together with (II.2) implies that  $\mathbb{P}$ -a.s. for all sufficiently large  $m$ ,

$$X_r^{t_m, x_m; \mathbf{u}} \in A_i, \quad \forall r \in [t_m, \tau_i^0[.$$

In conjunction with  $\mathbb{P}$ -a.s. continuity of sample paths, this immediately leads to

$$(II.4) \quad \liminf_{m \rightarrow \infty} \tau_i^m := \liminf_{m \rightarrow \infty} \tau_i(t_m, x_m) \geq \tau_i(t_0, x_0) \quad \mathbb{P}\text{-a.s.}$$

On the other hand, as mentioned earlier, the Assumptions 5.1.c. and 5.2 imply that the set of sample paths that hit the boundary of  $A_i$  and do not enter the set is negligible [RB98, Corollary 3.2, p. 65]. Thus, with probability one we have

$$\forall \delta > 0, \quad \exists s \in [\Theta_i^{A_1:n}(t_0, x_0), \Theta_i^{A_1:n}(t_0, x_0) + \delta[ : \\ X_s^{t_0, x_0; \mathbf{u}} \in A_i$$

Hence, in light of (II.2),  $\mathbb{P}$ -a.s. there exists  $\varepsilon > 0$ , possibly depending on  $\delta$ , such that for all sufficiently large  $m$  we have

$$X_s^{t_m, x_m; \mathbf{u}} \in B_\varepsilon(X_s^{t_0, x_0; \mathbf{u}}) \subset A_i^c$$

Recalling the induction hypothesis, we note that in accordance with the definition of sequential stopping times  $\Theta_i^{A_1:n}$ , one can infer that  $\Theta_i^{A_1:n}(t_m, x_m) \leq s < \Theta_i^{A_1:n}(t_0, x_0) + \delta$ . From arbitrariness of  $\delta$  and the definition of  $\tau_i$ , this leads to

$$\limsup_{m \rightarrow \infty} \tau_i(t_m, x_m) := \limsup_{m \rightarrow \infty} (\Theta_i^{A_1:n}(t_m, x_m) \wedge T_i) \leq \tau_i(t_0, x_0) \quad \mathbb{P}\text{-a.s.},$$

<sup>7</sup>This notation is only employed in this proof.

where in conjunction with (II.4), P-a.s. continuity of the map  $(t, x) \mapsto \tau_i(t, x)$  at  $(t_0, x_0)$  for any  $i \in \{1, \dots, n\}$  follows. The assertion follows by induction.

The continuity of the mapping  $(t, x) \mapsto X_{\tau_i(t, x)}^{t, x; \mathbf{u}}$  follows immediately from the almost sure continuity of the stopping time  $\tau_i(t, x)$  in conjunction with the almost sure continuity of the version of the stochastic process  $X^{t, x; \mathbf{u}}$  in  $(t, x)$ ; for the latter let us note again that Kolmogorov's continuity criterion guarantees the existence of such a version in light of (II.1).  $\square$

*Proof of Theorem 5.6.* Here we briefly sketch the proof of the first assertion of the theorem (supersolution property), and refer the reader to [MCL15, Thm. 4.10] for details concerning the same technology to prove the theorem.

Note that any  $\mathbb{F}_t$ -progressively measurable  $\mathbf{u} \in \mathcal{U}_t$  satisfies Assumptions 4.1.a.. It is a classical result [Øks03, Chap. 7] that the strong solution  $X^{t, x; \mathbf{u}}$  satisfies Assumptions 4.1.b.i. and 4.1.b.ii. Furthermore, Proposition 5.4 together with almost sure path-continuity of the strong solution guarantees Assumption 4.1.b.iii. Hence, having fulfilled all the required assumptions of Theorem 4.3, we can employ the DPP (8). For the sake of contraction to the first assertion, suppose there exists  $(t_0, x_0) \in [0, T_k] \times A_k$  such that  $-\sup_{u \in \mathbb{U}} \mathcal{L}^u V_{k*}(t_0, x_0) < 0$  in the viscosity sense. That is, there exist a smooth function  $\phi$  and  $\delta > 0$  such that

$$\begin{cases} \min_{(t, x) \in \mathbb{S}} (V_{k*} - \phi)(t, x) = (V_{k*} - \phi)(t_0, x_0) = 0, \\ -\sup_{u \in \mathbb{U}} \mathcal{L}^u \phi(t_0, x_0) < -2\delta. \end{cases}$$

Since  $\phi$  is smooth, the map  $(t, x) \mapsto \mathcal{L}^u \phi(t, x)$  is continuous for each  $u \in \mathbb{U}$ . Therefore, there exist  $u_0 \in \mathbb{U}$  and  $r > 0$  such that

$$-\mathcal{L}^{u_0} \phi(t, x) < -\delta, \quad \forall (t, x) \in B_r(t_0, x_0).$$

Let us define the stopping time  $\theta(t, x)$  as the first exit time of trajectory  $X^{t, x; u_0}$  from the ball  $B_r(t_0, x_0)$ . Note that by continuity of the solution process of the SDE (9), it holds that  $t < \theta(t, x)$  with probability 1 for all  $(t, x) \in B_r(t_0, x_0)$ . Therefore, selecting  $r > 0$  sufficiently small so that  $B_r(t_0, x_0) \subset [0, T_k] \times A_k$  and applying Itô's formula, we see that for all  $(t, x) \in B_r(t_0, x_0)$ , we have

$$(II.5) \quad \phi(t, x) < \mathbb{E}[\phi(\theta(t, x), X_{\theta(t, x)}^{t, x; u_0})].$$

Let us take a sequence  $(t_m, x_m, V_k(t_m, x_m))_{m \in \mathbb{N}}$  converging to  $(t_0, x_0, V_{k*}(t_0, x_0))$ , i.e.,  $\phi(t_m, x_m) \rightarrow \phi(t_0, x_0) = V_{k*}(t_0, x_0)$ . Due to the inequality (II.5), for sufficiently large  $m$  we have

$$V(t_m, x_m) < \mathbb{E}[V_{k*}(\theta(t_m, x_m), X_{\theta(t_m, x_m)}^{t_m, x_m; u_0})],$$

which, in view of the fact that  $\theta(t_m, x_m) < \tau_k \wedge T_k$ , contradicts the DPP in (8a). The *subsolution* property is proved effectively in a similar fashion.  $\square$

To provide boundary conditions, in particular in the viscosity sense (10b), we need some preliminaries as follows:

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

$$\Theta_j^{A_{k:n}}(t, x) = \Theta_j^{A_{i:n}}(\theta, X_{\theta}^{t, x; \mathbf{u}}),$$

pointwise on the set  $\{\Theta_{i-1}^{A_{k:n}}(t, x) \leq \theta < \Theta_i^{A_{k:n}}(t, x)\}$ .

**Lemma II.2.** *Suppose that the conditions of Proposition 5.4 hold. Given a sequence of controls  $(\mathbf{u}_m)_{m \in \mathbb{N}} \subset \mathcal{U}$  and initial conditions  $(t_m, x_m) \rightarrow (t, x)$ , we have*

$$\lim_{m \rightarrow \infty} \left\| X_{\tau_i(t, x)}^{t, x; \mathbf{u}_m} - X_{\tau_i(t_m, x_m)}^{t_m, x_m; \mathbf{u}_m} \right\| = 0 \quad \mathbb{P}\text{-a.s.},$$

where  $\tau_i(t, x) := \Theta_i^{A_{1:n}}(t, x) \wedge T_i$ .

Note that Lemma II.2 is indeed a stronger statement than Proposition 5.4 as the desired continuity is required uniformly with respect to the control. Let us highlight that the stopping times  $\tau_i(t, x)$  and  $\tau_i(t_m, x_m)$  are both effected by the control  $\mathbf{u}_m$ . But nonetheless, the mapping  $(t, x) \mapsto X_{\tau_i}^{t, x; \mathbf{u}_m}$  is almost surely continuous irrespective of the controls  $(\mathbf{u}_m)_{m \in \mathbb{N}}$ . For the proof we refer to an identical technique used in [MCL15, Lemma 4.11]

*Proof of Proposition 5.7.* The boundary condition in (10a) is an immediate consequence of the definition of the sequential exit-times introduced in Definition 3.1. Namely, for any initial state  $x \in A_k^c$  we have  $\Theta_k^{A_{k:n}}(t, x) = t$ , and in light of Fact II.1 for all  $i \in \{k, \dots, n\}$

$$\begin{aligned} \Theta_i^{A_{k:n}}(t, x) &= \Theta_i^{A_{k+1:n}}(t, x), \\ \forall(t, x) \in [0, T_k] \times A_k^c \bigcup \{T_k\} \times \mathbb{R}^d. \end{aligned}$$

Since  $\tau_k^k = t$  for the above initial conditions, then  $X_{\tau_k^k}^{t, x; \mathbf{u}} = x$  which yields to (10a).

For the boundary conditions (10b), we show the first assertion; the second follows similarly. Let  $(t_m, x_m) \rightarrow (t, x)$  where  $t_m < T_k$  and  $x_m \in A_k$ . Invoking the DPP in Theorem 4.3 and introducing  $\theta := \tau_{k+1}^k$  in (8a), we reach

$$V_k(t_m, x_m) \leq \sup_{\mathbf{u} \in \mathcal{U}_t} \mathbb{E} \left[ V_{k+1}^* (\tau_k^k, X_{\tau_k^k}^{t_m, x_m; \mathbf{u}}) \ell_k (X_{\tau_k^k}^{t_m, x_m; \mathbf{u}}) \right].$$

Note that one can replace a sequence of controls in the above inequalities to attain the supremum. This sequence, of course, depends on the initial condition  $(t_m, x_m)$ . Hence, let us denote it via two indices  $(\mathbf{u}_{m,j})_{j \in \mathbb{N}}$ . One can deduce that there exists a subsequence of  $(\mathbf{u}_{m,j})_{j \in \mathbb{N}}$  such that

$$\begin{aligned} & \lim_{m \uparrow \infty} V_k(t_m, x_m) \\ & \leq \lim_{m \uparrow \infty} \lim_{j \uparrow \infty} \mathbb{E} \left[ V_{k+1}^* (\tau_k^k, X_{\tau_k^k}^{t_m, x_m; \mathbf{u}_{m,j}}) \ell_k (X_{\tau_k^k}^{t_m, x_m; \mathbf{u}_{m,j}}) \right] \\ & \leq \lim_{j \uparrow \infty} \mathbb{E} \left[ V_{k+1}^* (\tau_k^k, X_{\tau_k^k}^{t_j, x_j; \mathbf{u}_{m_j}}) \ell_k (X_{\tau_k^k}^{t_j, x_j; \mathbf{u}_{m_j}}) \right] \\ \text{(II.6)} \quad & \leq \mathbb{E} \left[ \lim_{j \uparrow \infty} V_{k+1}^* (\tau_k^k, X_{\tau_k^k}^{t_j, x_j; \mathbf{u}_{m_j}}) \ell_k^* (X_{\tau_k^k}^{t_j, x_j; \mathbf{u}_{m_j}}) \right] \\ \text{(II.7)} \quad & = V_{k+1}^*(t, x) \ell_k^*(x) \end{aligned}$$

where (II.6) and (II.7) follow, respectively, from Fatou's lemma and the uniform continuity assertion in Lemma II.2. Let us recall that by Lemma II.2 we know  $\tau_k^k(t_j, x_j) \rightarrow \tau_k^k(t, x) = t$  as  $j \rightarrow \infty$  uniformly with respect to the controls  $(\mathbf{u}_{m_j})_{j \in \mathbb{N}}$ . Similar analysis would follow for the second part of (10b) by using the other side of DPP in (8b).  $\square$

## REFERENCES

- [AGA09] Steven Andrews, Tuan Dinh, and Adam Arkin, *Stochastic models of biological processes*, Meyers RA, ed. Encyclopedia of Complexity and System Science (2009), 87308749.

- [AM11] Ross Anderson and Dejan Milutinovic, *A stochastic approach to dubins feedback control for target tracking*, IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2011, pp. 3917–3922.
- [AP98] J.P. Aubin and G. Da Prato, *The viability theorem for stochastic differential inclusions*, Stochastic Analysis and Applications **16** (1998), no. 1, 1–15.
- [APF00] J.P. Aubin, G. Da Prato, and H. Frankowska, *Stochastic invariance for differential inclusions*, Set-Valued Analysis. An International Journal Devoted to the Theory of Multifunctions and its Applications **8** (2000), no. 1-2, 181–201.
- [Aub91] J.P. Aubin, *Viability Theory*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 1991.
- [Ber05] D. Bertsekas, *Dynamic programming and suboptimal control: A survey from {ADP} to mpc*, European Journal of Control **11** (2005), no. 45, 310 – 334.
- [BG99] M. Bardi and P. Goatin, *Invariant sets for controlled degenerate diffusions: a viscosity solutions approach*, Stochastic analysis, control, optimization and applications, Systems Control Found. Appl., Birkhäuser Boston, Boston, MA, 1999, pp. 191–208.
- [BHKH05] Christel Baier, Holger Hermanns, Joost-Pieter Katoen, and Boudewijn R. Haverkort, *Efficient computation of time-bounded reachability probabilities in uniform continuous-time markov decision processes*, Theor. Comput. Sci. **345** (2005), no. 1, 2–26.
- [BJ02] M. Bardi and R. Jensen, *A geometric characterization of viable sets for controlled degenerate diffusions*, Set-Valued Analysis **10** (2002), no. 2-3, 129–141.
- [BMGA02] R.W. Beard, T.W. McLain, M.A. Goodrich, and E.P. Anderson, *Coordinated target assignment and intercept for unmanned air vehicles*, IEEE Transactions on Robotics and Automation **18** (2002), no. 6, 911–922.
- [Bor05] V. S. Borkar, *Controlled diffusion processes*, Probability Surveys **2** (2005), 213–244 (electronic).
- [BSS01] A. Becskei, B. Seraphin, and L. Serrano, *Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion*, EMBO J **20** (2001), no. 10, 2528–2535.
- [BT11] B. Bouchard and N. Touzi, *Weak dynamic programming principle for viscosity solutions*, SIAM Journal on Control and Optimization **49** (2011), no. 3, 948–962.
- [CCL11] Debasish Chatterjee, Eugenio Cinquemani, and John 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>.
- [Che00] J. Cherry, *How to make a Biological Switch*, Journal of Theoretical Biology **203** (2000), no. 2, 117–133.
- [CIL92] M. G. Crandall, H. Ishii, and P. L. Lions, *User’s guide to viscosity solutions of second order partial differential equations*, American Mathematical Society **27** (1992), 1–67.
- [CRVRL06] Randy Cogill, Michael Rotkowitz, Benjamin Van Roy, and Sanjay Lall, *An approximate dynamic programming approach to decentralized control of stochastic systems*, Control of Uncertain Systems: Modelling, Approximation, and Design, Lecture Notes in Control and Information Science, Springer, 2006, pp. 243–256.
- [CS98] Y. Chitour and H. J. Sussmann, *Line-integral estimates and motion planning using the continuation method*, Essays on Mathematical Robotics, Springer-Verlag, New York, 1998, pp. 91–126.
- [DF04] G. Da Prato and H. Frankowska, *Invariance of stochastic control systems with deterministic arguments*, Journal of Differential Equations **200** (2004), no. 1, 18–52.
- [dFVR03] D. P. de Farias and B. Van Roy, *The linear programming approach to approximate dynamic programming*, Operations Research **51** (2003), no. 6, 850–865.
- [EK86] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Ltd., New York, 1986.
- [ESKPG05] Hana El Samad, Mustafa Khammash, Linda Petzold, and Dan Gillespie, *Stochastic modelling of gene regulatory networks*, Internat. J. Robust Nonlinear Control **15** (2005), no. 15, 691–711.
- [FS06] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solution*, 3 ed., Springer-Verlag, 2006.
- [GCC00] T. S. Gardner, C. R. Cantor, and J. J. Collins, *Construction of a genetic toggle switch in Escherichia coli*, Nature **403** (2000), no. 6767, 339–42.

- [GLQ06] Y. Gao, J. Lygeros, and M. Quincampoix, *The Reachability Problem for Uncertain Hybrid Systems Revisited: a Viability Theory Perspective*, Hybrid systems: computation and control, Lecture Notes in Comput. Sci., vol. 3927, Springer, Berlin, 2006, pp. 242–256.
- [JM70] David H. Jacobson and David Q. Mayne, *Differential Dynamic Programming*, American Elsevier Publishing Co., Inc., New York, 1970. MR 0263418 (41 #8023)
- [Kal97] Olav Kallenberg, *Foundations of Modern Probability*, Probability and its Applications (New York), Springer-Verlag, New York, 1997.
- [Kap07] Hilbert J. Kappen, *An introduction to stochastic control theory, path integrals and reinforcement learning*, Cooperative behavior in neural systems, vol. 887, Amer. Inst. Phys., Melville, NY, 2007, pp. 149–181. MR 2441307 (2009i:93140)
- [Kha10] Mustafa Khammash, *Modeling and analysis of stochastic biochemical networks*, Control Theory and System Biology (Pablo A. Iglesias and Brian P. Ingalls, eds.), MIT Press, Cambridge, MA, 2010, pp. xii+345.
- [Kry09] N.V. Krylov, *Controlled Diffusion Processes*, Stochastic Modelling and Applied Probability, vol. 14, Springer-Verlag, Berlin Heidelberg, 2009, Reprint of the 1980 Edition.
- [KS91] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, 2 ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.
- [KS11] Boris N. Khoromskij and Christoph Schwab, *Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs*, SIAM Journal on Scientific Computing **33** (2011), no. 1, 364–385. MR 2783199 (2012e:65273)
- [KT03] Vijay R. Konda and John N. Tsitsiklis, *On actor-critic algorithms*, SIAM Journal on Control and Optimization **42** (2003), no. 4, 1143–1166 (electronic). MR 2044789 (2004m:93151)
- [LTS00] J. Lygeros, C. Tomlin, and S.S. Sastry, *A game theoretic approach to controller design for hybrid systems*, Proceedings of IEEE **88** (2000), no. 7, 949–969.
- [MB00] Timothy W. McClain and Randal W. Beard, *Trajectory planning for coordinated rendezvous of unmanned air vehicles*, Proc. GNC’2000, 2000, pp. 1247–1254.
- [MCL11] Peyman Mohajerin Esfahani, Debasish Chatterjee, and John Lygeros, *On a problem of stochastic reach-avoid set characterization*, 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), Dec 2011, pp. 7069–7074.
- [MCL15] ———, *On a stochastic reach-avoid problem and set characterization*, Automatica, accepted (2015), Available at <http://arxiv.org/abs/1202.4375>.
- [Mit05] I. Mitchell, *A toolbox of hamilton-jacobi solvers for analysis of nondeterministic continuous and hybrid systems*, Hybrid systems: computation and control (M. Morari and L. Thiele, eds.), Lecture Notes in Comput. Sci., no. 3414, Springer-Verlag, 2005, pp. 480–494.
- [MMAC13] P. Mohajerin Esfahani, A. Miliadis-Argeitis, and D. Chatterjee, *Analysis of controlled biological switches via stochastic motion planning*, European Control Conference (ECC), July 2013, pp. 93–98.
- [MSSO<sup>+</sup>11] A. Miliadis, 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 **29** (2011), no. 12, 1114–1116.
- [MVM<sup>+</sup>10] Peyman Mohajerin Esfahani, Maria Vrakopoulou, Kostas Margellos, John Lygeros, and Goran Andersson, *Cyber attack in a two-area power system: Impact identification using reachability*, American Control Conference, 2010, pp. 962–967.
- [Øks03] Bernt Øksendal, *Stochastic differential equations*, sixth ed., Universitext, Springer-Verlag, Berlin, 2003. MR 2001996 (2004e:60102)
- [Pro05] Philip E. Protter, *Stochastic Integration and Differential Equations*, Stochastic Modelling and Applied Probability, vol. 21, Springer-Verlag, Berlin, 2005, Second edition. Version 2.1, Corrected third printing.
- [RB98] Richard Bass, *Diffusions and Elliptic Operators*, Probability and its Applications (New York), Springer-Verlag, New York, 1998.
- [Ren99] P. J. Reny, *On the existence of pure and mixed strategy nash equilibria in discontinuous games*, Econometrica **67** (1999), 1029–1056.
- [SH07] Abhyudai Singh and JooPedro Hespanha, *A derivative matching approach to moment closure for the stochastic logistic model*, Bulletin of Mathematical Biology **69** (2007), no. 6, 1909–1925.
- [SL10] Sean Summers and John Lygeros, *Verification of discrete time stochastic hybrid systems: a stochastic reach-avoid decision problem*, Automatica **46** (2010), no. 12, 1951–1961. MR 2878218

- [SM86] Kang G. Shin and Neil D. McKay, *A dynamic programming approach to trajectory planning of robotic manipulators*, IEEE Transactions on Automatic Control **31** (1986), no. 6, 491–500.
- [Sus91] H. J. Sussmann, *Two new methods for motion planning for controllable systems without drift*, Proceedings of the First European Control Conference (ECC91) (Grenoble, France), 1991, pp. 1501–1506.
- [TBS10] Evangelos A. Theodorou, Jonas Buchli, and Stefan Schaal, *A generalized path integral control approach to reinforcement learning*, J. Mach. Learn. Res. **11** (2010), 3137–3181. MR 2746549 (2011h:62256)
- [TL05] Emanuel Todorov and Weiwei Li, *A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems*, American Control Conference, June 2005, pp. 300–306.
- [Tou13] Nizar Touzi, *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*, Fields Institute Monographs, vol. 29, Springer, New York, 2013. MR 2976505
- [Wil06] Darren James Wilkinson, *Stochastic modelling for systems biology*, Chapman & Hall/CRC, Boca Raton, FL, 2006. MR 2222876 (2006k:92040)