Automatica 87 (2018) 274–280

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief Paper

Markov inequality rule for switching among time optimal controllers in a multiple vehicle intercept problem☆

Dejan Milutinović [a], David W. Casbeer [b], Meir Pachter [c]

[a] Computer Engineering Department, University of California, Santa Cruz, CA 95064, United States
[b] Control Science Center of Excellence, Air Force Research Laboratory, Wright-Patterson AFB, OH 45433, United States
[c] Department of Electrical Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433, United States

ARTICLE INFO

Article history:
Received 24 April 2016
Received in revised form 4 June 2017
Accepted 25 August 2017
Available online 6 November 2017

ABSTRACT

In this paper, a Markov inequality based switching rule is proposed to switch among numerically computed, time optimal controllers in a multiple vehicle intercept problem. Each controller is optimal for the intercept of a single vehicle, i.e., for the *segment* of the complete time varying multiple vehicle target set. The switching rule guarantees that after every switch the time to the target set is shorter with a certain predefined probability. Furthermore, the rule guarantees that the target set is reached after a finite number of switches and the rule scales well with the number of vehicles, i.e., the segments covering the target set. The problem and results are illustrated by a numerical example.

1. Introduction

Numerical methods of stochastic optimal control (Kushner & Dupuis, 2001) are attractive for designing state feedback controllers for nonlinear dynamical systems described by stochastic differential equations (Øksendal, 2010). Generally, these computed controllers are in the form of a lookup table and can be executed fast in real-time applications for reaching a target set in a minimum expected time (Anderson, Efstathios, Milutinović, & Panagiotis, 2012). However, the computed controller is specific to the target set, and in the case of a time varying target set, the controller must be recomputed, which in most cases is too slow for real time implementations.

This paper is motivated by a multiple vehicle intercept problem presented in Section 2. While we know how to compute a minimum time optimal control to intercept a single vehicle by using its relative position coordinates, in the case of multiple vehicles, the number of necessary relative coordinates increases linearly, which quickly exhausts the computational power for computing an optimal solution due to the so-called *curse of dimensionality*. In addition, let us assume that we can compute the controller for N vehicles. That controller will be valid only for the specific number N. Adding or removing one vehicle would require that the optimal control be changed. It would mean that we need to compute not one, but multiple optimal controllers, one for each possible number of vehicles up to a maximum number of vehicles, and select or switch to the controller that matches the current number of vehicles. Since it is known (Branicky, 1998) that switching among multiple satisfactory controllers may lead to unsatisfactory outcomes, this is unacceptable and an additional level of analysis is necessary.

For the case of multiple vehicles, we propose to use the minimum time optimal control to intercept a single vehicle and a switching rule to select which vehicle to intercept. The proposed rule guarantees that after every switch, the time to intercept is reduced with a certain predefined probability. The rule is based on the computed state dependent expected time to intercept and the optimal control; therefore, the switching is also state dependent. The resulting switching rule based navigation scales well with the number of vehicles.

The dynamics of the motivating problem (Section 2) with switchings among the vehicles can be modeled with *deterministic transitions* over a finite number of discrete states (one per vehicle) and stochastic differential equation dynamics in each state, the so-called stochastic switched system (see Table 1 in Teel, Subbaramana, and Sferlazza (2014)). The stability of deterministic versions of such systems has been studied using Lyapunov function and multiple Lyapunov function approaches (Branicky, 1998) and a stability analysis of stochastic versions used similar tools applied to statistical estimates of Lyapunov functions (Chatterjee & Liberzon, 2004) and a comparison principle (Chatterjee & Liberzon, 2006). The moment stability for such systems has also been analyzed in Boukas (2006), Feng, Jieand, and Ping (2011), Feng and Zhang

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michael V. Basin under the direction of Editor Ian R. Petersen.

E-mail addresses: dmilutin@ucsc.edu (D. Milutinović), david.casbeer@us.af.mil (D.W. Casbeer), meir.pachter@afit.edu (M. Pachter).

https://doi.org/10.1016/j.automatica.2017.09.009
0005-1098/

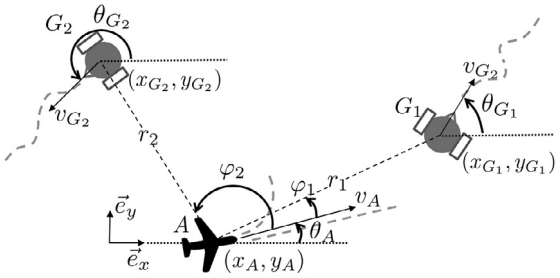

Fig. 1. Geometry and state-space representation of an aerial vehicle A at position (x_A, y_A) and two ground vehicles (G_i), at (x_{G_i}, y_{G_i}), $i = 1, 2$; θ_{EGi}, θ_A are heading angles; φ_i and r_i are bearing angles and ranges from A to G_i; v_A is the aerial vehicle velocity and v_{G_i} are ground vehicle velocities.

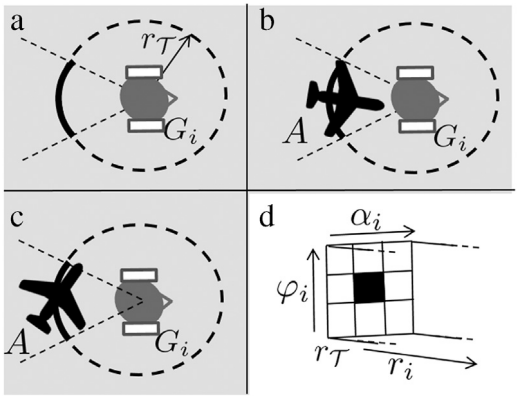

Fig. 2. (a) The target segment set $\mathcal{T}_i(t)$, which is the circular sector behind the moving G_i . (b) Agent A is inside the target set, while in (c) agent A is not in the target set, because its heading is not aligned with G_i. (d) The time invariant set $\mathcal{S}_i^{\mathcal{T}}$ is shown in a 3D-space, r_i, φ_i, α_i.

(2006) and Filipović (2009). However, in all these references, including (Chatterjee & Liberzon, 2004, 2006), the analysis is focused on a time-dependent switching signal and not on a state-dependent switching (Zhang, Wu, & Xia, 2014). A hysteresis-based switching in the context of supervisory control of uncertain systems has been proposed in Hespanha, Liberzon, and Morse (2003), but only recently has the state-dependent switching law for the stochastic switched systems been considered in Wu, Cui, Shi, and Karimi (2013) and Zhang et al. (2014). Furthermore, the stability of the sliding mode control for semi-Markovian jump systems has been considered in Li, Wua, Shi, and Lim (2015). To the best of our knowledge, a *state-dependent* switching policy that reduces the time to reach a target set and prevents an infinite number of switchings has not been considered so far. Another novelty of the work presented here is that the policy is scalable, which makes it suitable for the navigation of an autonomous vehicle surrounded by a number of other vehicles.

While the problem motivation in Section 2 suggests the application of the switching rule to navigation in multi-agent systems (Dimarogonas & Kyriakopoulos, 2004), in Section 3 we discuss the switching in the context of a *target set* which is a union of *target segment sets*, where each segment set corresponds to a single vehicle. Therefore, the switching rule is potentially relevant to other robotics applications that appear in the similar form, for example, those in which segments can be associated with sets corresponding to robot end effector grasping configurations. Section 4 illustrates our results using the multiple vehicle problem and statistical analysis, and Section 5 gives conclusions.

2. Problem motivation

Let us consider a scenario with three agents depicted in Fig. 1. Two of the agents are ground vehicles G_1 and G_2 with equal speeds $v_{G_1} = v_{G_2} = v_G$. The third agent is an aerial vehicle (A) flying at a constant altitude. The kinematic model of A is a deterministic Dubins vehicle model describing the vehicle's position x_A, y_A and heading angle θ_A as

$$dx_A = v_A \cos(\theta_A)dt \tag{1}$$
$$dy_A = v_A \sin(\theta_A)dt \tag{2}$$
$$d\theta_A = u_A dt \tag{3}$$

where the A's velocity, $v_A > v_G$, is a known constant, and its control input is the bounded heading rate $u_A \in [-1, 1]$. The agent A's goal is to navigate in a *minimum time* into the vulnerable tail sector $\mathcal{T}_i(t)$, $i = 1, 2$ of one or the other ground vehicle for inspection purposes. Therefore, the target *set for* A is $\mathcal{T}(t) = \mathcal{T}_1(t) \cup \mathcal{T}_2(t)$ and the time dependence is the consequence of ground vehicle motion. However, *A has no knowledge of ground vehicles' navigation strategy*. To anticipate that uncertainty, the kinematics of each ground vehicle is modeled by the stochastic dynamics

$$dx_{G_i} = v_G \cos(\theta_{G_i})dt, \quad i = 1, 2 \tag{4}$$
$$dy_{G_i} = v_G \sin(\theta_{G_i})dt \tag{5}$$
$$d\theta_{G_i} = \sigma_G dw_i, \tag{6}$$

where the vehicle positions are given by x_{G_i}, y_{G_i} and the heading angles are $\theta_{G_i}(t) = \int_0^t \sigma_E dw_i$, which are continuous time random walks since dw_i denotes the Wiener process increments. The scaling parameter σ_G is identical for both vehicles.

The vulnerable tail $\mathcal{T}_i$(t) sector is a circular sector attached to the back of the ground vehicle G_i, see Fig. 2, and we also refer to it as a target segment. The relative position between A and G_i is uniquely defined based on the triple of relative coordinates $(r_i, \varphi_i, \alpha_i)$, where r_i is the distance between A and G_i, φ_i is the bearing angle from A to G_i, and α_i is the difference between the A's and G_i's heading angles. Therefore, the time varying target segment $\mathcal{T}_i(t)$ in the Cartesian space can be represented as a time invariant set $\mathcal{S}_i^{\mathcal{T}} \in \mathcal{R}$ in the space of relative coordinates $\mathcal{R} \subset \mathbb{R}^3$

$$\mathcal{S}_i^{\mathcal{T}} = \left\{[\underline{r}, \overline{r}] \times [-\overline{\varphi}, \overline{\varphi}] \times [-\overline{\alpha}, \overline{\alpha}]\right\}, \tag{7}$$

with $0 < \underline{r} < \overline{r}$, $\overline{\varphi} > 0$, and $\overline{\alpha} > 0$ (see Fig. 2). The time invariant $\mathcal{S}_i^{\mathcal{T}}$ is a box in the state space defined by r_i, φ_i, and α_i, and the facet of this cube for $r_i = \overline{r}$ is depicted in Fig. 2d. For the time invariant set $\mathcal{S}_i^{\mathcal{T}}$, we can formulate and compute the Hamilton–Jacobi–Bellman (HJB) equation solution for the minimum time optimal controller to reach $\mathcal{T}_i(t)$. The result is the optimal control $u_i(r_i, \varphi_i, \alpha_i)$ which is state dependent and defines the value of control variable for a given relative position $(r_i, \varphi_i, \alpha_i)$ between A and G_i, and for every i. If the aerial vehicle A *is tasked to* reach the tail sector $\mathcal{T}_i$ of a specific G_i, then the computed optimal control u_i defines the optimal way to do it and a further analysis of the problem is unnecessary.

However, computations for a minimum time optimal controller for the target set $\mathcal{T}(t) = \mathcal{T}_1(t) \cup \mathcal{T}_2(t)$ require the time invariant set $\mathcal{S}^{\mathcal{T}} \subset \mathbb{R}^6$ since the configuration of agents requires 6 relative coordinates, i.e., three per the target segment. The same number of dimensions is also necessary to describe the state feedback control. If we consider reaching the segmented target set $\mathcal{T}(t) = \cup_{i=1}^{N} \mathcal{T}_i(t)$, then the number of state variables is $3N$. For example, in Section 4, we present the multiple vehicle problem with 5 vehicles, which requires dealing with a 15-dimensional state space. This large number of dimensions quickly exhausts the computational power for computing the HJB equation solution and the optimal control.

3. Switching rule

Let us consider a nonempty and not necessarily connected time-varying target set $\mathcal{T}(t) \subset \mathbb{R}^n$. The first assumption (A1) is that the set $\mathcal{T}(t)$ is a union of sets, $\mathcal{T}(t) = \cup_{i=1}^{N} \mathcal{T}_i(t)$, where $\mathcal{T}_i(t) \subset \mathbb{R}^n$ are not necessarily disjoint sets. In the sequel, these sets are called *target segments* $\mathcal{T}_i(t)$ and the set $\mathcal{T}(t)$ is a *segmented target set*. The term *segment* is borrowed from image processing literature (Klette, 2014) where image segments with certain features cover collectively the entire image, which is similar to our assumption that the segments cover collectively the entire target set.

The second assumption (A2) is that for each target segment $\mathcal{T}_i(t)$, there exist a time invariant set $\mathcal{S}_i^{\mathcal{T}} \subset \mathcal{R}$ in the space $\mathcal{R}$ of relative coordinates $\mathbf{r}_i \in \mathcal{R}$, $i = 1, \ldots, N$, and a state feedback control law $u_i(\mathbf{r}_i) : \mathcal{R} \to \mathcal{U}$, which takes values in the control variable set $\mathcal{U}$, that can take the controlled stochastic system $d\mathbf{r}_i = f(\mathbf{r}_i, u_i(\mathbf{r}_i))dt + g(\mathbf{r}_i, u_i(\mathbf{r}_i))d\mathbf{w}_i$ from any initial condition $\mathbf{r}_i(0)$ at $t = 0$ to the state $\mathbf{r}(\tau_F^i) \in \mathcal{S}_i^{\mathcal{T}}$ with the expected time $U_i(\mathbf{r}_i(0)) = E\left\{\tau_F^i\right\}$, where $\tau_F^i > 0$ is the random arrival time. In the stochastic system description, $\mathbf{w}_i$ is the vector valued Wiener process (Øksendal, 2010), and following the work on state-dependent switching (Wu et al., 2013), $f(\cdot, \cdot)$ and $g(\cdot, \cdot)$ are locally Lipschitz in $\mathbf{r}_i$ and u_i, and the state dependent $u_i(t) = u_i(\mathbf{r}_i(t))$ is an ultimately uniformly bounded and piecewise continuous process. Naturally, the state-dependent function $U_i(\mathbf{r}_i) : \mathcal{R} \to \mathbb{R}$ and $\mathbf{r}_i \in \mathcal{S}_i^{\mathcal{T}} \Rightarrow U_i(\mathbf{r}_i) = 0$. For $\mathbf{r}_i(t)$ which varies in time, we define $U_i(t)$ as $U_i(t) = U_i(\mathbf{r}_i(t))$ and use it in the following theorem which defines the switching rule.

Theorem 1. *Assume that at time $\tau^s \geq 0$, the control $u_i(\mathbf{r}_i)$, $i = 1, \ldots, N$ is active and that up to τ^s, the number of switches is s, $s = 0, 1, 2..$, in which case $s = 0$ corresponds to the initial time $\tau^0 = 0$. Let the switch from active $u_i(\mathbf{r}_i)$ to $u_j(\mathbf{r}_j)$, $i \neq j$, $j = 1, \ldots, N$ take place at time τ^{s+1}, which is the first time $t > \tau^s$ when the following inequality is satisfied*

$$\frac{\min\{U_i(t), U_i(\tau^s)\} - U_j(t)}{\min\{U_i(t), U_i(\tau^s)\}} \geq p \tag{8}$$

where $U_i(t) > 0$, $U_i(\tau^s) > 0$, $U_j(t) \geq 0$ and $p \in (0, 1)$. Then, (a) the sequence $U(\tau^s)$, defined as $U(\tau^s) \triangleq U_i(\tau^s)$ with i indicating currently active control, is bounded from above as $U(\tau^s) \leq (1-p)^s U(0)$ for all $s = 1, 2, 3, 4....$ (b) If we define a threshold $\underline{U} > 0$ so that the switching stops once $U(\tau^s) \leq \underline{U}$, then the total number of switches is bounded by s^b defined as

$$s^b = 1 + \left\lceil \frac{\ln(\underline{U}/U(0))}{\ln(1-p)} \right\rceil \tag{9}$$

where the brackets $\lceil \cdot \rceil$ denote the closest integer number greater than the expression in them. (c) The switch from active $u_i(\mathbf{r}_i)$ to $u_j(\mathbf{r}_j)$ based on the condition (8) guarantees that

$$Pr\left\{\tau_F^j < \min\left\{U_i(t), U_i(\tau^s)\right\}\right\} \geq p \tag{10}$$

or, in other words, that the time of reaching the target set is shorter than $\min\{U_i(t), U(\tau^s)\}$ with a probability greater than p.

Proof. To prove (a), we start from (8) to obtain

$$U_j(t) \leq (1-p)\min\{U_i(t), U_i(\tau^s)\}. \tag{11}$$

At the time point $t = \tau^{s+1}$ of the switch from i to j, this inequality is $U_j(\tau^{s+1}) \leq (1-p)\min\{U_i(\tau^{s+1}), U_i(\tau^s)\}$. If $U_i(\tau^{s+1}) < U_i(\tau^s)$, we have

$$U_j(\tau^{s+1}) < (1-p)U_i(\tau^s). \tag{12}$$

Otherwise, we have $U_i(\tau^{s+1}) \geq U_i(\tau^s)$ and

$$U_j(\tau^{s+1}) \leq (1-p)U_i(\tau^s). \tag{13}$$

Based on the definition, $U(\tau^{s+1}) = U_j(\tau^{s+1})$ and $U(\tau^s) = U_i(\tau^s)$; therefore, the last two inequalities can be summarized as $U(\tau^{s+1}) \leq (1-p)U(\tau^s)$. Iterations of this inequality backward result in

$$U(\tau^{s+1}) \leq (1-p)U(\tau^s) \leq (1-p)^2 U(\tau^{s-1}) \tag{14}$$

and once we reach $s = 0$, we can conclude that $U(\tau^{s+1}) \leq (1-p)^{s+1}U(\tau^0)$, i.e.,

$$U(\tau^s) \leq (1-p)^s U(\tau^0) \tag{15}$$

which finalizes the proof of (a) since $\tau^0 = 0$.

To prove (b), we start with the condition that the switching stops $U(\tau^s) \leq \underline{U}$ and (15), to obtain

$$U(\tau^s) \leq (1-p)^{s-1}U(0) \leq \underline{U} \tag{16}$$

and conclude that there exists a finite s, such that the switching stop condition is satisfied $U(\tau^s) \leq \underline{U}$. Having in mind that $\ln(1-p) < 0$, this yields the following inequality $s \geq 1 + \frac{\ln(\underline{U}/U(0))}{\ln(1-p)}$ and the smallest integer s when $U(\tau^s) \leq \underline{U}$ is s^b defined by expression (9). The bound s^b value defines the total number of switches after which $U(\tau^s)$ has to be smaller than or equal to the threshold value $\underline{U}$. In other words, it defines the maximal numbers of switches until the switching stops.

The proof of (c) is based on the Markov inequality. First note that $U_i(t) = U_i(\mathbf{r}_i(t)) > 0$, for all $i = 1, 2, \ldots, N$, is equivalent to the statement $\mathbf{r}_i(t) \notin \mathcal{S}_i^{\mathcal{T}}$, i.e., that the target set $\mathcal{T}(t)$ is not reached. Also to shorten the notation, let us define $U_i^{min}(t) = \min\{U_i(t), U_i(\tau^s)\}$. The time τ_F^j to reach the target segment $\mathcal{T}_j(t)$ after the switch is a nonnegative random variable with the expected value $U_j(t) = U_j(\mathbf{r}_j(t))$; therefore, it satisfies the Markov inequality

$$Pr\left\{\tau_F^j \geq U_i^{min}(t)\right\} \leq \frac{E\{\tau_F^j\}}{U_i^{min}(t)} = \frac{U_j(t)}{U_i^{min}(t)}. \tag{17}$$

Since $Pr\{\tau_F^j < U_i^{min}(t)\} = 1 - Pr\{\tau_F^j \geq U_i^{min}(t)\}$, the elementary algebraic transformations of the inequality yield

$$Pr\left\{\tau_F^j < U_i^{min}(t)\right\} \geq \frac{U_i^{min}(t) - U_j(t)}{U_i^{min}(t)}. \tag{18}$$

The right side of the inequality is identical to the left side of (8); therefore, $Pr\left\{\tau_F^j < U_i^{min}(t)\right\} \geq p$, which finalizes the proof. □

4. Example

In the example, we use the scenario described in Section 2 and illustrated in Fig. 1, but the number of ground vehicles G_i is $N = 5$ and $i = 1, 2, \ldots, 5$. The relative position between A and G_i is uniquely defined based on the relative coordinates $\mathbf{r}_i = (r_i, \varphi_i, \alpha_i)$

$$r_i(t) = \sqrt{(x_A - x_{G_i})^2 + (y_A - y_{G_i})^2}, i = 1, 2...5 \tag{19}$$

$$\varphi_i(t) = -\theta_P + atan\left(\frac{y_{G_i} - y_A}{x_{G_i} - x_A}\right), \varphi_i \in [-\pi, \pi] \tag{20}$$

$$\alpha_i(t) = \theta_{G_i} - \theta_A, \alpha_i \in [-\pi, \pi] \tag{21}$$

where r_i is the distance between A and G_i, φ_i is the bearing angle from A to G_i, and α_i is the difference between the A's and G_i's heading angles. The target segment $\mathcal{T}_i(t)$, i.e., the corresponding invariant set $\mathcal{S}_i^{\mathcal{T}} \subset \mathbb{R}^3$ is defined by (7). Using Îto calculus, the relative coordinates (19)–(21), the kinematic model for A (1)–(3), and the stochastic model for G_i (4)–(6), we can derive the following

evolution for the relative coordinates between A and G_i

$$dr_i = (v_G\cos(\varphi-\alpha) - v_A\cos\varphi)\,dt = b_{r_i}dt \tag{22}$$
$$d\varphi_i = \left(-u_A + \frac{-v_G\sin(\varphi-\alpha)+v_A\sin\varphi}{r_i}\right)dt = b_{\varphi_i}dt$$
$$d\alpha_i = -u_A dt + \sigma_G dw = b_{\alpha_i}dt + \sigma_G dw.$$

The minimum expected time control problem to reach the target segment $\mathcal{T}_i(t)$ is to find a control u_A that minimizes the cost $\mathcal{J}(u_A) = E\left\{g(\tau_F^i) + \int_0^{\tau_F^i} 1dt\right\}$ where the terminal cost $g(\tau_F^i) = g(r_i(\tau_F^i), \varphi_i(\tau_F^i), \alpha_i(\tau_F^i))$ is

$$g(\tau_F^i) = \begin{cases} 0, & \text{if } (r_i(\tau_F^i), \varphi_i(\tau_F^i), \alpha_i(\tau_F^i)) \in \mathcal{S}_i^{\mathcal{T}} \\ M, & \text{if } (r_i(\tau_F^i), \varphi_i(\tau_F^i), \alpha_i(\tau_F^i)) \in \mathcal{A}_i \end{cases} \tag{23}$$

with the set $\mathcal{A}_i$ defined as $\mathcal{A}_i = \left\{(r_i, \varphi_i, \alpha_i) | r_i \le \underline{r}\right\} \setminus \mathcal{S}_i^{\mathcal{T}}$. The terminal cost for the set $\mathcal{A}_i$ involves a large positive penalty $M \gg 0$ and yields the optimal control that avoids the set $\mathcal{A}_i$. There is no penalty for reaching the invariant set $\mathcal{S}_i^{\mathcal{T}}$. In other words, the cost function is constructed to yield the optimal control u_A that minimizes the time for A to reach the segment $\mathcal{T}_i(t)$ and avoids configurations in which A is in the proximity of G_i, but not in the tail sector set $\mathcal{T}_i(t)$.

The cost function $\mathcal{J}(u_A)$ gives rise to the HJB equation defining the evolution of the cost-to-go function V_i for the optimal control

$$0 = \min_{u_A(t)}\left\{b_{r_i}\frac{\partial V_i}{\partial r_i} + b_{\varphi_i}\frac{\partial V_i}{\partial \varphi_i} + b_{\alpha_i}\frac{\partial V_i}{\partial \alpha_i} + \sigma_E^2\frac{\partial^2 V_i}{\partial \alpha_i^2} + 1\right\} \tag{24}$$

with the two boundary conditions $V_i = 0$ for all $(r_i, \varphi_i, \alpha_i) \in \mathcal{S}_i^{\mathcal{T}}$ and $V_i = M$ for all $(r_i, \varphi_i, \alpha_i) \in \mathcal{A}_i$.[1] The solution of the HJB equation is $V_i = V(r_i, \varphi_i, \alpha_i)$, and the solution also yields the optimal state feedback control $u_i(\mathbf{r}_i) = u_A(r_i, \varphi_i, \alpha_i)$. Once the optimal control is computed, it can be used to compute the expected time U_i from the steady-state solution of the backward Kolmogorov (BK) equation (Gardiner, 2009),

$$0 = b_{r_i}\frac{\partial U_i}{\partial r_i} + b_{\varphi_i}\frac{\partial U_i}{\partial \varphi_i} + b_{\alpha_i}\frac{\partial U_i}{\partial \alpha_i} + \sigma_E^2\frac{\partial^2 U_i}{\partial \alpha_i^2} + 1 \tag{25}$$

with the boundary condition $U_i = 0$ for all $(r_i, \varphi_i, \alpha_i) \in \mathcal{S}_i^{\mathcal{T}}$. The BK equation solution is $U_i = U_i(r_i, \varphi_i, \alpha_i)$. Naturally, as $(r_i(t), \varphi_i(t), \alpha_i(t))$ varies in time, there is a corresponding evolution of the expected time $U_i(t) = U_i(r_i(t), \varphi_i(t), \alpha_i(t))$, where U_i denotes the expected time to reach the segment $\mathcal{T}_i(t)$.

Numerical method: The solution of the optimal control problem is based on the so-called locally consistent Markov chain discretization of the HJB equation. The discretization yields a Markov chain with control u_A-dependent transition probabilities while the problem of solving the HJB equation is converted into a discrete state space dynamic programming problem. Then, we use value iterations (Thrun, Burgard, & Fox, 2005) to find the optimal feedback control $u_i^h = u_A^h(r_i^h, \varphi_i^h, \alpha_i^h)$ of the discrete Markov chain. The superscript h indicates that the control is computed for the discretized state space. Once we compute the optimal control, we can compute the expected time U_i, i.e., its discrete approximation U_i^h, exploiting the similarity of the Kolmogorov equation (25) with (24). To do that, we use the same discretization scheme and value iterations in which the control is already computed u_i^h and the *min* operator is excluded. Further details about the numerical method can be found in Section III of Munishkin, Milutinović, and Casbeer (2016). In the same paper, the controller has been implemented and tested with ground robots as a part of a different control problem. Similar controllers were used for the navigation of a small UAV in the presence of stochastic winds (Anderson, Bakolas, Milutinović & Tsiotras, 2012). A general explanation of the method for the control of nonholonomic vehicles is in Choi and Milutinović (2015). The method was also used for target tracking problems (Anderson & Milutinović, 2014) and was flight tested with UAVs (Milutinović, Casbeer, & Kingston, 2017).

[1] Note that because of the second boundary condition, which follows from the terminal costs in the set $\mathcal{A}_i$, the cost-to-go function V_i cannot be interpreted as the expected time to reach $\mathcal{T}_i$.

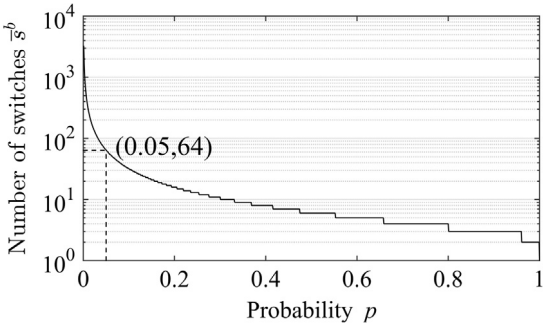

Fig. 3. Conservative upper bound for the number of switches, which is independent of the initial position of A relative to the ground vehicles, $\bar{s}^b = 1 + \left\lceil \frac{\log(U/\underline{U})}{\log(1-p)} \right\rceil$.

The units in (22), which we used to compute the numerical optimal control in the example of this section, are normalized so that all the angles are in radians, and the velocities $v_A = 1$ and $v_G = 0.5$. The noise scaling parameter $\sigma_E = 10\pi/180$ and the maximum turning rate of A is $u_{max} = 0.05$. The tail sectors (7) to be reached by A are defined by $\underline{r} = 5, \bar{r} = 15, \overline{\varphi} = -\underline{\varphi} = 10\pi/180$ and $\overline{\alpha} = -\underline{\alpha} = 20\pi/180$. The computational domain is

$$\mathcal{K} = \{[R_{min}, R_{max}] \times [-\pi, \pi - \Delta\varphi) \times [-\pi, \pi - \Delta\alpha)\}$$

with $R_{min} = \underline{r} = 5$, $R_{max} = 204$ and discretized with the steps $\Delta r = (R_{max} - R_{min})/99 \approx 2.01$ and $\Delta\varphi_i = \Delta\alpha = 5\pi/180$ in the direction of r_i, φ_i and α_i state space variables. Since in our problem formulation, the angles φ and α have full ranges, in the discretized computational domain, the pairs of points $(r_i^h, -\pi, \alpha_i^h)$ and $(r_i^h, \pi - \Delta\varphi, \alpha_i^h)$, as well as $(r_i^h, \varphi_i^h, -\pi)$ and $(r_i^h, \varphi_i^h, \pi - \Delta\alpha)$ are next to each other, i.e., we use periodic boundary conditions along the φ_i and α_i state space variables. Other boundary conditions of the discrete approximation based on the locally consistent Markov chain approximation method take into account (23) and the boundary conditions of (24) and (25).

Note that the target sets $\mathcal{S}_i^{\mathcal{T}}$ and kinematic parameters are equal for all five vehicles, which means that for the same relative coordinates $\mathbf{r}_i = \mathbf{r}_j \Rightarrow u_i(\mathbf{r}_i) = u_j(\mathbf{r}_j)$ and $U_i(\mathbf{r}_i) = U_j(\mathbf{r}_j)$. Therefore, we need to run value iterations to compute the control and the expected times only once and use their values for all vehicles based on their relative positions.

Results: In the results presented here we use the fact that the expected times are computed based on the discrete approximation. Because of that we can say that the largest possible value for $U(0)$ (Theorem 1(b)) is $\overline{U} = \max_{r_i,\varphi_i,\alpha_i} U_i^h$. Similarly, the smallest possible threshold that we should consider is $\underline{U} = \min_{r_i,\varphi_i,\alpha_i} U_i^h$. With these two values, we can define a conservative upper bound $\bar{s}^b$ for the number of switches in our example, which is plotted in Fig. 3 as a function of p. This figure illustrates well that the number of switches is larger for a smaller p (Theorem 1(c)). As customary in statistics, we use $p = 0.05$, which corresponds to $\bar{s}^b = 64$ switches. The upper bound guarantees a finite switching sequence, and, as expected from the conservative nature of this bound, the simulation results show that the number of switches is much smaller than 64, see Fig. 6.

Fig. 4 shows one of many simulation results we obtained. Even if we use the same initial position in our simulation, the results

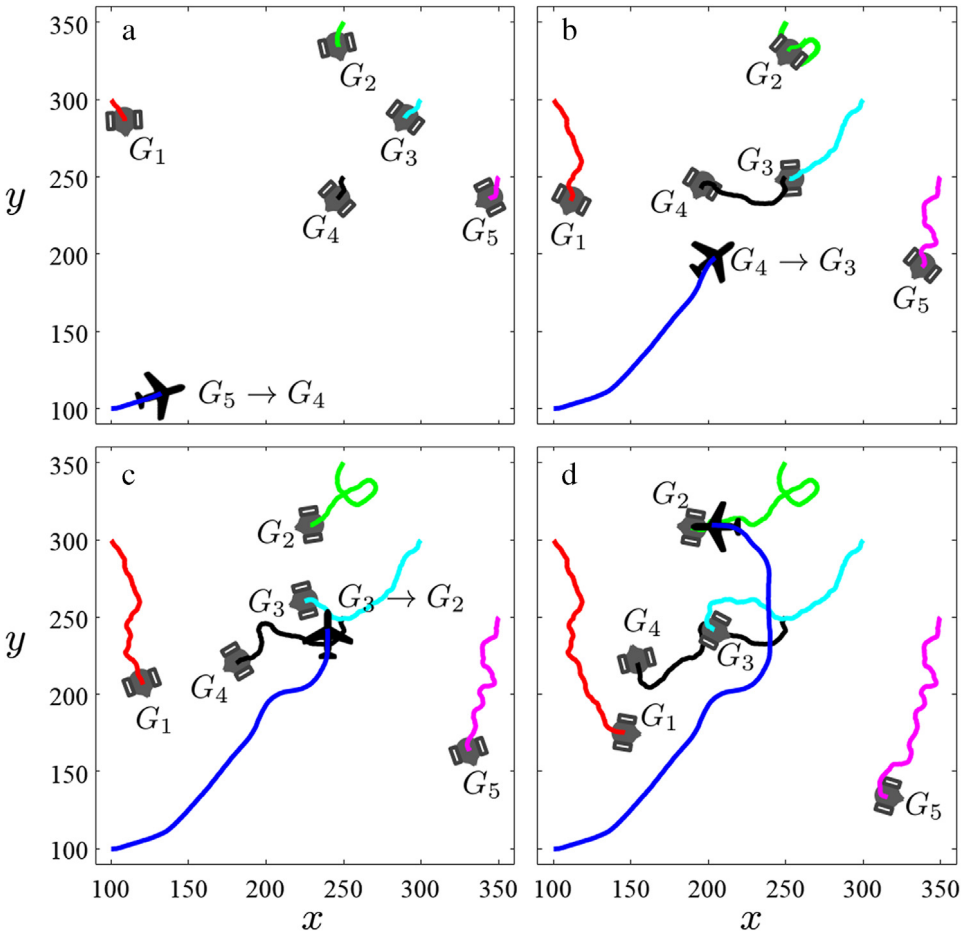

Fig. 4. Numerical simulation example with five ground vehicles G_i, $i = 1, 2...5$. Trajectories up to the point of the switch from (a) G_5 to G_4; (b) G_4 to G_3; (c) G_3 to G_2. (d) Trajectories up to the point in which the pursuer enters the tail of G_2.

are always different due to the stochastic nature of the ground vehicle dynamics. We selected the realization in Fig. 4 because there are three switches with more or less straight trajectories for the agents. More curved trajectories, including multiple loops, are not unusual, however they are difficult to interpret due to the trajectories cluttering the figure.

Another characteristic of our example is that out of five ground vehicles, only G_1 is initially in the computational domain of the numerical optimal control. In the simulation, whenever there is the chance that the distance between G_i and A exceeds R_{max}, we set $r_i = R_{max}$ and use the resulting control, as well as the corresponding expected time in our switching rule. Because of that, A initially follows G_5 with the heading angle and expected time evaluated at the distance $r_5 = R_{max}$. G_5, i.e., the target segment $\mathcal{T}_5$ is more favorable to reach than any other segment, including $\mathcal{T}_1$ from the closest vehicle G_1. In the simulation, the switching rule is based on the discrete approximation $U_i^h(t)$ resulting from the discrete approximation of the Kolmogorov equation U_i^h and evaluated at discrete points corresponding to the relative position $(r_i(t), \varphi_i(t), \alpha_i(t))$ between A and G_i at time t. The simulation is performed in discrete steps of $\Delta t = 1$.

Fig. 4a shows trajectories up to the point in which A switches from reaching $\mathcal{T}_5$ to reaching $\mathcal{T}_4$. The next time, A switches to $\mathcal{T}_3$ (Fig. 4b), and finally to $\mathcal{T}_2$ (Fig. 4c), which is reached as illustrated in Fig. 4d. The switching times are represented by Fig. 5a with the plot of the subscript i of the segment to be reached against the time.

Fig. 5b shows the evolution of $U_i^h(t)$ in time for each segment, $i = 1, 2...5$ and of the value that corresponds to the currently pursued segment $\mathcal{T}_i$. The diagram shows that before the first switch ($\mathcal{T}_5 \to \mathcal{T}_4$), the expected times to reach segments $\mathcal{T}_4$ and $\mathcal{T}_1$ are shorter, but the switch happens at $t = 35$ once the switching rule is satisfied. The next switch at $t = 150$ happens after the expected time to reach $\mathcal{T}_4$ starts to increase. The increase is due to G_4 changing its heading, when A misses the opportunity to enter quickly the segment $\mathcal{T}_4$. The switch takes place once the condition for the switch to $\mathcal{T}_3$ is satisfied, otherwise A would continue pursuing $\mathcal{T}_4$. Similarly, before the next switch, A misses the opportunity to enter quickly $\mathcal{T}_3$ and, at $t = 215$, it switches to $\mathcal{T}_2$. After this switch, A reaches $\mathcal{T}_2$ at $\tau_F = 304$ and at that point of time $U_i^h(t) = 0$.

The numerical simulation is repeated 1000 times with the same initial condition. The number of switches before A enters the target set is counted and represented in Fig. 6 histogram, which is normalized to represent its empirical probability distribution. Note that for these 1000 cases, the number of switches is rarely 0 and never more than 9, which indicates how conservative the upper bound $\bar{s}^b$ is. From these simulation runs, we also computed the average time of $\overline{\tau_F} = 420.41$ with a 95% confidence interval [404.41, 436.41] until the target set $\mathcal{T}$, i.e., one of the segments $\mathcal{T}_i$, was reached. Without the switching, i.e., when A did not switch from reaching $\mathcal{T}_5$, the average time to reach the target set was $\overline{\tau_F} = 513.95$ with a 95% confidence interval [497.32, 530.57]. Obviously, the average performance of reaching the target set using switching is better.

5. Conclusions

We considered in this paper the problem of reaching a target set, which is composed of individual segments. Each individual

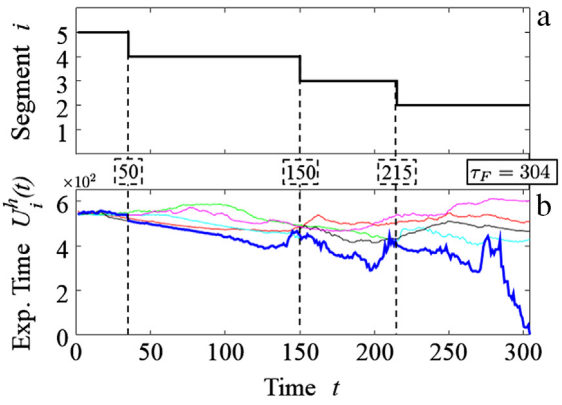

Fig. 5. Numerical simulation example with the aerial vehicle A and five ground vehicles G_i. The target set is $\mathcal{T} = \cup_{i=1}^{5} \mathcal{T}_i$, $i = 1, 2, \ldots, 5$: (a) Index i of the target segment to be reached. The switching times are indicated with the dashed lines. The time τ_F at which A enters the segment $\mathcal{T}_2$ is $\tau_F = 304$. ; (b) The expected times to segments $\mathcal{T}_1$-red, $\mathcal{T}_2$-green, $\mathcal{T}_3$-cyan, $\mathcal{T}_4$-black, $\mathcal{T}_5$-magenta. The thick blue line corresponds to the current $\mathcal{T}_i$ (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.).

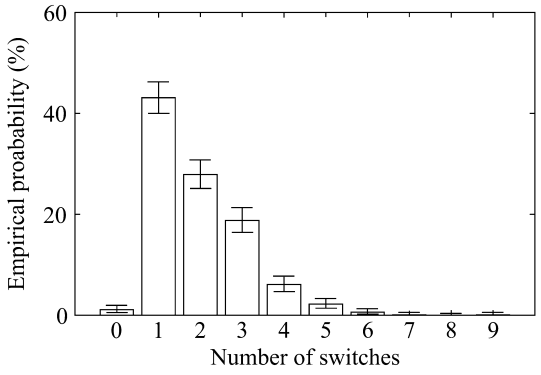

Fig. 6. Histogram of the number of switches from 1000 simulations with 95% confidence intervals. The histogram is normalized to represent the empirical probability distribution.

segment can be reached by a corresponding controller and, in order to reach the target set faster, the policy for switching among the controllers was proposed. The policy is based on expected times to reach individual segments and it reduces, with a specified probability, the time to reach the target set. The policy is state dependent, has a bounded number of switches, and scales well with the number of target segments.

The proposed policy's properties were proven and the policy was illustrated by the multiple vehicle interception problem. Our statistical analysis shows the reduction of the average time of reaching the target and the distribution of the number of switches, which is skewed to numbers that are significantly smaller than a conservative bound for the number of switches.

References

Anderson, R. P., Bakolas, E., Milutinović, D., & Tsiotras, P. (2012). Optimal feedback guidance of a small aerial vehicle in the presence of stochastic wind. *AIAA Journal of Guidance, Control and Dynamics*, *36*(4), 975–985.

Anderson, R. P., & Milutinović, D. (2014). A stochastic approach to dubins vehicle tracking problems. *IEEE Transactions on Automatic Control*, *59*(10), 2801–2806.

Boukas, E.-K. (2006). *Stochastic switching systems: Analysis and design*. Birkhäuser.

Branicky, M. S. (1998). Multiple Lyapunov functions and other analysis tools for switched and hybrid systems. *IEEE Transactions on Automatic Control*, *43*(4), 475–482.

Chatterjee, D., & Liberzon, D. (2004). On stability of stochastic switched systems. In *the proceedings of the 43rd IEEE conference on decision and control*, Vol. 4, (pp. 4125–4127).

Chatterjee, D., & Liberzon, D. (2006). Stability analysis of deterministic and stochastic switched systems via a comparison principle and multiple Lyapunov functions. *SIAM Journal on Control and Optimization*, *45*(1), 174–206.

Choi, J., & Milutinović, D. (2015). Tips on stochastic optimal feedback control and bayesian spatio-temporal models: applications to robotics. *ASME Journal of Dynamic Systems, Measurement, and Control*, *137*(3), 030801.

Dimarogonas, D. V., & Kyriakopoulos, K. J. (2004). Lyapunov-like stability of switched stochastic systems. In *the proceedings of the 2004 American control conference*, Vol. 2, (pp. 1868–1872).

Feng, W., Jieand, T., & Ping, Z. (2011). Stability analysis of switched stochastic systems. *Automatica*, *47*, 148–157.

Feng, W., & Zhang, J. F. (2006). Stability analysis and stabilization control of multivariable switched stochastic systems. *Automatica*, *42*, 169–176.

Filipović, V. (2009). Exponential stability of stochastic switched systems. *Transactions of the Institute of Measurement and Control*, *31*(2), 205–212.

Gardiner, C. (2009). *Stochastic methods: A handbook for the natural and social sciences*. Springer.

Hespanha, J., Liberzon, D., & Morse, A. S. (2003). Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, *39*, 263–272.

Klette, R. (2014). *Concise computer vision: An introduction into theory and algorithms*. Springer.

Kushner, H. J., & Dupuis, P. (2001). *Numerical methods for stochastic control problems in continuous time*. Berlin: Springer.

Li, F., Wua, G., Shi, P., & Lim, C-C. (2015). state estimation and sliding mode control for semi-markovian jump systems with mismatched uncertainties. *Automatica*, *51*, 385–393.

Milutinović, D., Casbeer, D. W., Rasmussen, S., & Kingston, D. (2017). A stochastic approach to small uav feedback control for target tracking and blind spot avoidance. In *the proceedings of the 2017 IEEE conference on control technology and applications*.

Munishkin, A. A., Milutinović, D., & Casbeer, D. W. (2016). Stochastic optimal control navigation with the avoidance of unsafe configurations. In *the proceedings of the 2016 international conference on unmanned aircraft systems*.

Øksendal, B. (2010). *Stochastic differential equations: An introduction with applications*. Berlin: Springer.

Teel, A. R., Subbaramana, A., & Sferlazza, A. (2014). Stability analysis for stochastic hybrid systems: A survey. *Automatica*, *50*, 2435–2456.

Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic robotics*. MIT Press.

Wu, Z., Cui, M., Shi, P., & Karimi, H. R (2013). Stability of stochastic nonlinear systems with state-dependent switching. *IEEE Transactions on Automatic Control*, *58*(9), 1904–1918.

Zhang, H., Wu, Z., & Xia, Y. (2014). Stability of stochastic nonlinear systems with state-dependent switching. *Automatica*, *50*, 599–606.

Dejan Milutinović earned Dipl.-Ing (1995) and Magister's (1999) degrees in Electrical Engineering from the University of Belgrade, Serbia and a doctoral degree in Electrical and Computer Engineering (2004) from Instituto Superior Técnico, Lisbon, Portugal. From 1995 to 2000, he worked as a research engineer in the Automation and Control Division of Mihajlo Pupin Institute, Belgrade, Serbia.

His doctoral thesis was the first runner-up for the best Ph.D. thesis of European Robotics in 2004 by EURON. He won the NRC award of the US Academies in 2008 and Hellman Fellowship in 2012.

He is currently an Associate Professor in the Department of Computer Engineering, UC Santa Cruz. His research interests are in the area of modeling and control of stochastic dynamical systems applied to robotics. He currently serves as an Editor for the Journal of Intelligent and Robotic Systems, Springer.

David W. Casbeer is the Technical Area Lead for the UAV Cooperative and Intelligent Control Team within the Control Science Center of Excellence in the Air Force Research Laboratory's Aerospace Systems Directorate. The UAV team focuses on decision making, planning, and coordination for multiple autonomous UAVs acting and reacting in uncertain and adversarial environments. Dr. Casbeer received the B.S. and Ph.D. degrees from Brigham Young University in 2003 and 2009, respectively, where he advanced theory describing the statistics of decentralized estimation techniques. In 2016, Dr. Casbeer was awarded AFRL's Early Career Award for distinguished foundational research in multi-agent control. He currently serves as the Chair for the AIAA Intelligent Systems Technical Committee and as a Senior Editor for the Journal of Intelligent and Robotic Systems.

Meir Pachter is a Professor of Electrical Engineering at the Air Force Institute of Technology, Wright-Patterson AFB.

Dr. Pachter received the B.S. and M.S. degrees in Aerospace Engineering in 1967 and 1969, respectively, and the Ph.D. degree in Applied Mathematics in 1975, all from the Israel Institute of Technology.

Dr. Pachter held research and teaching positions at the Israel Institute of Technology, the Council for Scientific and Industrial Research in South Africa, Virginia Polytechnic Institute, Harvard University and Integrated Systems, Inc.

Dr. Pachter is interested in the application of mathematics to the solution of engineering and scientific problems. His current areas of interest include military operations optimization, dynamic games, cooperative control, estimation and optimization, statistical signal processing, adaptive optics, inertial navigation, and GPS navigation. For his work on adaptive and reconfigurable flight control he received the AF Air Vehicle's Directorate Foulois award for 1994, together with Phil Chandler and Mark Mears.

Dr. Pachter is a Fellow of the IEEE.