Abstract—The optimal control for mobile agents is an important and challenging issue. Recent work shows that using randomized mechanism in agents’ control can make the state unpredictable, and thus improve the security of agents. However, the unpredictable design is only considered in single period, which can lead to intolerable control performance in long time horizon. This paper aims at the trade-off between the control performance and state unpredictability of mobile agents in long time horizon. Utilizing random perturbations consistent with uniform distributions to maximize the attackers’ prediction errors of future states, we formulate the problem as a multi-period convex stochastic optimization problem and solve it through dynamic programming. Specifically, we design the optimal control strategy considering both unconstrained and input constrained systems. The analytical iterative expressions of the control are further provided. Simulation illustrates that the algorithm increases the prediction errors under Kalman filter while achieving the control performance requirements successfully.

I. INTRODUCTION

Nowadays, as technologies of perception, localization and motion planning gradually mature, mobile agents such as autonomous vehicles have been widely applied in various fields [1], [2]. However, traditional researches on optimal control and planning usually pursue the objectives of optimal time, shortest path or minimum energy consumption [3]. Since autonomous vehicles become more and more pervasive in the industry and military fields, how to protect their security while controlling have raised great concerns these years [4], but still remains an unsolved problem [5].

In this paper, we focus on the security breach where an external attacker can predict the states of a mobile agent deployed in the physical environment. This security issue is of practical importance, since the historical trajectory of a vehicle is inevitably public to the attacker, who can leverage numerous estimation methods for interception or attack [6]. For instance, [7] uses extended Kalman filter to estimate the states of a moving object and predict its trajectory by a UAV and [8] presents a multiple model unscented Kalman filter to predict multi-agent trajectory. Notice that for an attacker, an accurate prediction of the system state is essential to the latter attack. Therefore, it is necessary for a control system to protect its history trajectory for being unpredictable to attackers. Some studies have already been carried out in these years. [9] introduces a novel coding scheme to protect the secrecy of the robot motion planning. [10] considers an entropy maximization problem on a partially observable Markov decision process to decrease the predictability of the decision-maker’s trajectory. A secure control method is shown in [11] to ensure the resilience to attacks.

Nevertheless, to the best of our knowledge, few studies have yielded the unpredictable control of the mobile agents. In [12], the author proves the unpredictability of the system is greatest when the inputs satisfy a uniform distribution. However, the proposed stochastic control method is a single-period optimization, which can only guarantee an unpredictable trajectory but omits the overall system performance in a long time horizon. Due to the addition of random disturbances, the final state of the system may not reach the expected target state.

Inspired by the above discussion, we present a novel control strategy for mobile agents with linear systems to ensure both unpredictability and control performance. A stochastic perturbation is added to control inputs at each step in order to make the system states unpredictable and we translate the problem into a multi-period convex optimization problem. Moreover, simple input constraints are also considered in the problem. The main contributions are summarized as follows:

- Aiming at the trade-off between state unpredictability and the control performance in long time horizon, we formulate a multi-period stochastic optimization problem for linear systems and solve it through stochastic dynamic programming.
- We design an optimal control strategy for both unconstrained and input-constrained mobile agents to ensure the security and be unpredictable for external attackers while achieving the system performance requirements successfully. Performance showed by simulation demonstrates the effectiveness of the algorithm.

The remainder of the paper is organized as follows. Section II describes the problem of interest. Section III solves the optimal solution of the unconstrained problem with dynamic programming, while Sec. IV studies the problem with simple input constraints. Simulation results are shown in Sec. V, followed by conclusions and future directions in Sec. VI.

II. PROBLEM FORMULATION AND PRELIMINARIES

A. Model Description

- System model of mobile agent

Consider a mobile agent with a linear dynamic model whose discrete form is

\[ x_{k+1} = A_k x_k + B_k u_k, \]  

where \( x_k \) is the state vector of the mobile agent, \( u_k \) is the control input and \( A_k, B_k \) are \( n \times n \) and \( n \times m \) matrices for...
any $k = 0, 1, \cdots, N - 1$. The output model of the mobile agent is

$$ y_k = C x_k, $$

(2)

where $y_k$ is the output information such as agent’s positions, and we have $C = [I_q \times q \ 0_{q \times (n-q)}]$. 

- **Prediction model of attacker**

Suppose there is an attacker having exact knowledge of the dynamic model $A_k$, $B_k$ and $C$ of the mobile agent. Through observing the output information, the attacker is able to predict the future trajectory of the agent with some data fusion methods. The prediction model is described as:

$$ \hat{y}_{k+1|k} = C \hat{x}_{k+1|k} = C (A_k \hat{x}_{k|k} + B_k \hat{u}_{k|k}), $$

(3)

where $\hat{y}_{k+1|k}$ and $\hat{u}_{k|k}$ are predictions to the system output and control input at time $k$ separately. $\hat{x}_{k|k}$ is the posterior estimate of $x_k$. In our assumption, $\hat{x}_{k|k}$ is an unbiased estimation. Denote the attacker’s prediction error at time $k$ as:

$$ e_{k+1|k} = y_{k+1} - \hat{y}_{k+1|k}. $$

(4)

**B. Problem of Interest**

Consider a task of controlling the mobile agent from an initial state $x_0$ to a target state $x_N^\ast$. An external attacker is trying to predict the future state during this process and carry out an attack or interception. Our objective is to design a control policy $u_k$ to maximize the state unpredictability while achieving the control performance requirements in the long time horizon.

In order to increase the unpredictability, a random disturbance term $\delta_k$ is added to the control input at each step. The random variable $\delta_k$ satisfies a distribution $f_k(y)$ with the expectation $E(\delta_k) = 0$ and variance $\mathbb{D}(\delta_k) = \sigma_k^2$, where $\delta_{k,i}$ and $\sigma_{k,i}$ are $i$ th component of $\delta_k$ and $\sigma_k$ separately. Note that each component of $\delta_k$ is independent. We have

$$ u_k = \mu_k + \delta_k, $$

(5)

which means $u_k$ is also a random variable and for all $k = 0, 1, \cdots, N - 1$, $i = 1, \cdots, m$

$$ \mathbb{E}(u_k) = \mu_k, \quad \mathbb{D}(u_{k,i}) = \sigma_{k,i}^2. $$

(6)

In this paper, we define $\sigma_k^2 = [\sigma_{k,1}^2, \cdots, \sigma_{k,m}^2]^T$ as a vector.

We now express the control objective in mathematical forms. Giving a fixed terminal time $N$ and target state $x_N^\ast$, we use a linear quadratic objective function to represent the system performance, denoted as $J_c\{u_{0:N-1}\}$. The optimization problem is written as

$$ \min J_c\{u_{0:N-1}\} = \mathbb{E} \left[ (x_N - x_N^\ast)^T H (x_N - x_N^\ast) \right] + \sum_{k=0}^{N-1} \mathbb{E} (x_k^T Q_k x_k + u_k^T R_k u_k), $$

(7)

where $u_{0:N-1} = \{u_0, u_1, \cdots, u_{N-1}\}$ is the set of control inputs, $H, Q_k$ are positive semi-definite matrices, and each $R_k$ is a positive definite matrix. Note that $J_c$ reflects both the deviation to the target state and the cost of control during the process. Since each control input $u_k$ in (5) is a random variable, the above optimization function is expressed in an expectation form.

On the other hand, we use the attacker’s prediction error $e_{k+1|k}$ at each step $k$ to represent the performance of the unpredictability, which is denoted as $J_p\{k\}$. Notice that $e_{k+1|k}$ is also a random variable, so we define $J_p\{k\}$ with the expectation form:

$$ J_p\{k\} = \mathbb{E}(\|e_{k+1|k}\|_2^2), k = 0, 1, \cdots, N - 1. $$

(8)

Regard (8) as a measurement of the state unpredictability. According to [12], we formulate the state unpredictability objective into a max-min optimization problem:

$$ \max \min J_p\{k\}, $$

where $J_p$ is first minimized with the attacker’s optimal estimation $\hat{u}_{k|k}$ to the control input, and then maximized with $\delta_k$ generated from the optimal distribution $f_k(y)$.

Therefore, considering both of the control performance and the state unpredictability, our objective is to optimize both (7) and (9). In the next section, we will introduce one basic lemma and transform this multi-objective problem into a solvable single-objective problem.

**III. PROBLEM REFORMULATION AND THE OPTIMAL SOLUTION**

**A. Problem Reformulation**

In this subsection, we will introduce a lemma from [12] and formulate an optimization problem combining (7) and (9).

**Lemma 1.** Consider the problem (9). Suppose the maximum variance of the random perturbation $\delta_k$ is $\sigma_k^2$. Then, $f_k(y)$ is the optimal distribution in the sense of probability iff

$$ \mathbb{D}(\delta_k) = \sigma_k^2, $$

and

$$ f_k^\ast(y) = \left\{ \begin{array}{ll} \frac{1}{(2\sqrt{3})^m \prod_{i=1}^{m} \sigma_{k,i}^2}, & y_i \in [-\sqrt{3}\sigma_{k,i}^2, \sqrt{3}\sigma_{k,i}^2], \\ 0, & \text{otherwise}, \end{array} \right. $$

where $\sigma_{k,i}$ is the component of $\sigma_k$, $i = 1, \cdots, m$.

The detailed proof process is in [12].

Lemma 1 shows that when the variance of the perturbation $\sigma_k^2$ reaches the maximum value, the expectation of the attacker’s prediction error $J_p\{k\}$ is the largest. This conclusion is consistent with our intuitions, which means the larger the control variances are, the more difficult it is for an attacker to predict the future states accurately. Therefore, to maximize the state unpredictability is to maximize the variance $\sigma_k^2$, and using the multivariate uniform distribution as $f_k^\ast(y)$ can optimize the problem (9). Thus, we define an unpredictable utility function as follows to help formulate our optimization problem.

**Unpredictable Utility Function:** Based on the above discussion, we design a utility function, denoted by $J_u\{k\}$, to represent the unpredictability at each period $k$, and combine (7) and (9) into a single-objective optimization problem. There are two requirements for designing this utility function:

- **The utility function is negatively correlated with all the variances $\sigma_{k,i}^2$, i.e., $J_u\{k\}$ is a monotonically decreasing linear function of $\sigma_{k,i}^2$.** When we minimize the value of
the function $J_u \{k \}$, it needs to enlarge the variance, and then ensure the higher unpredictability. Thus, the trade-off between the system performance and the unpredictability is well formulated.

- To maintain the convexity of the optimization objective, each $J_u \{k \}$ needs to be convex with $\sigma_{k,i}^2$ for $k = 0, \cdots, N - 1$. Clearly, if the function is non-convex, the global optimal solution is hard to be obtained and not unique. Meanwhile, it may be taken at the infinite boundary when there are no constraints on the variable.

Considering these two requirements, without loss of generality, the unpredictable utility function is defined as

$$J_u \{k \} = \sum_{i=1}^{m} \frac{1}{\sigma_{k,i}^2}$$

Then, we formulate a single-objective optimization problem:

$$\mathbf{P_1 :} \quad \min_{\mu_k, \sigma_k} \left\{ \begin{array}{c} \mu_k = -G_k x_k + M_k \\ \sigma_{k,i}^2 = (\frac{\lambda_{k,i}}{P_{k,i}^2})^\frac{1}{2} \end{array} \right\}$$

s.t. 

$$J_u \{k \} = \sum_{i=1}^{m} \frac{1}{\sigma_{k,i}^2}$$

where $\lambda_1 > 0, \lambda_2 > 0, \lambda_{3,k} \geq 0$ are weights of each term respectively and $\sigma_{k,i}$ is the $i$th component of $\sigma_k$. The optimization variables are $\mu_k$ and $\sigma_k$ at each step. We can see that the third term in $J$ is $\sum_{k=0}^{N-1} \lambda_{3,k} J_u \{k \}$, representing the unpredictability utility, and when the variance of control is larger, the function value is smaller. Moreover, the Bellman functions of this problem for all $k$ are convex, which will be shown in the next subsection.

In this way, we obtain a multi-period and single-object convex optimization problem. We will solve the Problem $\mathbf{P_1}$ in the next subsection.

### B. The Optimal Control Policy

Since Problem $\mathbf{P_1}$ is multi-period and convex, we solve the optimal solution with dynamic programming [13].

**Theorem 1.** The optimal solution of Problem $\mathbf{P_1}$ is given by

$$\left\{ \begin{array}{c} \mu_k = -G_k x_k + M_k \\ \sigma_{k,i}^2 = (\frac{\lambda_{k,i}}{P_{k,i}^2})^\frac{1}{2} \end{array} \right\} \tag{11}$$

where for $k = 0, 1, \cdots, N - 1$

$$G_k = P_k^{-1} B_k^T J_{k+1}^T A_k$$

$$M_k = \frac{1}{2} P_k^{-1} B_k^T J_{k+1}^T$$

$$P_k = \lambda_k R_k + B_k^T J_{k+1} B_k$$

and

$$W_k = \lambda_k Q_k + A_k^T J_{1,k+1} A_k, \quad Z_k = J_{2,k+1} A_k$$

$$J_{1,k} = W_k - A_k^T J_{1,k+1} B_k G_k, \quad J_{1,N} = \lambda_1 H$$

$$J_{2,k} = Z_k - J_{2,k+1} B_k G_k, \quad J_{2,N} = 2 \lambda_1 x_N^2 T H.$$}

**Proof.** See the proof in Appendix.

Once the parameters $G_k, M_k, P_k$ are obtained offline, we are able to calculate a sequence of $\sigma_k^2$ with Theorem 1. According to Lemma 1, the optimal distribution of the perturbation $\delta_k$ is the multivariate uniform distribution. Therefore, we generate the $\delta_k$ at each step from a uniform distribution $f_k(y)$ whose expectation is 0 and variance is $\sigma_k^2$.

$$\delta_{k,i} \sim \mathcal{U}[-(3\sigma_{k,i}^2)^\frac{1}{2}, (3\sigma_{k,i}^2)^\frac{1}{2}].$$

Then, together with formula (1) and (5), we can get the expectation sequence $\mu_k$ and control inputs $u_k$.

The time complexity of the dynamic programming algorithm is $O(Nn^3)$, where $n$ is the dimension of state $x_k$ and $N$ is total control steps.

Note that the expectation sequence $\mu_k$ decides the direction of convergence of the system, while the variance $\sigma_k^2$ determines the unpredictability of the state system. In a traditional LQR problem without uncertainty, the control law is usually given by $u_k = -K_k x_k$. As for our algorithm, an additional parameter $M_k$ is used to adjust $\mu_k$ (and $u_k$) since the disturbance term $\delta_k$ is added. We can observe from the formula (16) that $\sigma_k^2$ is positively correlated with $\lambda_{3,k}$, which means the unpredictability of the system can be increased by enlarging the value of $\lambda_{3,k}$. We will detail the effect of weight parameters on control performance in the simulation section.

### IV. DEALING WITH SIMPLEx INPUT-CONSTRAINTS

In most practical situations, the systems contain multiple constraints. However, it is not easy to solve a multi-period linear quadratic optimal control problem with complex constraints. Inspired by [14], we give an algorithm to solve the problem $\mathbf{P_3}$ with simple input-constraints through dynamic programming.

Consider a constrained problem

$$\mathbf{P_2 :} \quad \min_{\mu_k, \sigma_k} \left\{ \begin{array}{c} \mu_k = -G_k x_k + M_k \\ \sigma_{k,i}^2 = (\frac{\lambda_{k,i}}{P_{k,i}^2})^\frac{1}{2} \end{array} \right\}$$

s.t. 

$$J_u \{k \} = \sum_{i=1}^{m} \frac{1}{\sigma_{k,i}^2}$$

$$N^{-1} \sum_{k=0}^{N-1} \left\{ \begin{array}{c} J_u \{k \} \\ \sigma_{k,i}^2 = (\frac{\lambda_{k,i}}{P_{k,i}^2})^\frac{1}{2} \end{array} \right\}$$

where $\pi > 0$ is the upper bound of $|u_k|$ at each step. From Theorem 1, we obtain

$$\delta_{k,i} \sim \mathcal{U}[-\tilde{\delta}_k, \tilde{\delta}_k], \quad \tilde{\delta}_{k,i} = (3\sigma_{k,i}^2)^\frac{1}{2}. \tag{16}$$

Then we can set a conservative bound for each $\mu_k$ as

$$-\bar{\mu}_k \leq \mu_k \leq \bar{\mu}_k, \quad \bar{\mu}_k = \pi - \tau \tilde{\delta}_k, \tag{17}$$

where $\tau$ can determine how conservative the control is. In this way, if we set $\tau = 1$, the control inputs $u_k$ will definitely satisfy the constraints (15b).

The following theorems provide a solution for $\mathbf{P_2}$.

**Theorem 2.** The solution of Problem $\mathbf{P_2}$ is given by

$$\left\{ \begin{array}{c} \mu_k = -G_k x_k + \tilde{M}_k \\ \sigma_{k,i}^2 = (\frac{\lambda_{k,i}}{P_{k,i}^2})^\frac{1}{2} \end{array} \right\} \tag{18}$$

where for $k = 0, 1, \cdots, N - 1$ $G_k$ and $\tilde{M}_k$ are defined as:

$$\tilde{G}_k = \left\{ \begin{array}{c} G_k & |f_k - G_k x_k + \tilde{M}_k| \leq \bar{\mu}_k, \\ 0 & otherwise \end{array} \right\}$$

$$\tilde{M}_k = \left\{ \begin{array}{c} M_k & |f_k - G_k x_k + \tilde{M}_k| \leq \bar{\mu}_k, \\ -\bar{\mu}_k & f_k - G_k x_k + \tilde{M}_k \leq -\bar{\mu}_k, \\ \bar{\mu}_k & f_k - G_k x_k + \tilde{M}_k > \bar{\mu}_k. \end{array} \right\} \tag{19}$$
Remark 1. The process of using dynamic programming to solve $P_2$ is similar to $P_1$. The difference is that due to the constraints on $\mu_k$, at each step the new parameter $G_k$ and $M_k$ need to satisfy the equation (19) according to three different situations, and $G_k, M_k$ are substituted into $G_k, M_k$ during the continue calculation.

Remark 2. The problem of Theorem 2 is that we cannot get $x_k$ when computing the parameters $G_k$ and $M_k$ off-line. Therefore, we need to traverse all the possibilities of the parameter pair $(G_k, M_k)$. Note that there are 3 possible values for $(G_k, M_k)$ at each step and $3^N$ possible values for $(G_0, M_0)$. But for a certain initial state $x_0$, only one pair of $(G_0, M_0)$ satisfies all constraints in the problem. We propose Theorem 3 to help find the feasible parameters.

Theorem 3. Denote $\tilde{G}_{0:N-1} = [\tilde{G}_0, \cdots, \tilde{G}_{N-1}], \tilde{M}_{0:N-1} = [\tilde{M}_0, \cdots, \tilde{M}_{N-1}]$. For an initial state $x_0$, whether the parameter pair $(\tilde{G}_{0:N-1}, \tilde{M}_{0:N-1})$ are feasible with the constraints is determined by following $N$ inequalities.

$$-\pi_k \leq E_k x_0 + F_k \leq \pi_k, \quad k = 0, 1, \cdots, N - 1$$

where

$$\begin{cases} E_k = -\tilde{G}_k \Pi_{i=0}^{k-1} (A_k - B_k \tilde{G}_i) \\ F_k = -\tilde{G}_k \sum_{i=0}^{k-1} \Pi_{j=0}^{k-1} (A_j - B_j \tilde{G}_j) B_k \tilde{M}_i + \tilde{M}_k. \end{cases}$$

V. SIMULATION

In this section, we conduct multiple simulations on our algorithm to show the performance and the unpredictability of the system control.

In order to simplify the problem, we choose a single-input-single-out system. The dynamic equation is

$$x_{k+1} = x_k + (x_k + u_k) \Delta t,$$

where $\Delta t = T/N$ means to divide the total duration $T$ evenly into $N$ steps. Then we have $A = 1 + \Delta t, B = 1$ according to formula (1). Our control object is to maximize the unpredictability of the state system while meeting the last performance requirement. Let $x_0^N = 0$ and $H = 1, Q = 0, R = \Delta t$. The optimization function is described as $\min_{\mu_k, \sigma_k} J_1 = \lambda_1 E(x_k^2) + \lambda_2 \sum_{k=0}^{N-1} E(\sigma_k^2) + \sum_{k=0}^{N-1} \lambda_3 \sigma_k^2.

The upper bound of control inputs $|u_k|$ is $\pi = 4$. We use Kalman Filter to do the one-step prediction. The observation noise is set to be $N(0, 0.5)$ in the algorithm. At each step, the prediction of the next state is computed by $\hat{x}_{k+1|k} = A\hat{x}_{k|k} + B\hat{u}_{k|k}$. With two kinds of control methods, the prediction results and errors are illustrated in Fig.3. We can see that the prediction errors increase significantly after adding perturbation to the control inputs. The average and maximum prediction errors under different $\lambda_3$ are shown in Table I. As $\lambda_3$ grows larger, the system state is more difficult to be predicted accurately.

The control results with and without simple input-constraints are shown in Fig.4. Set the parameters $N = 15$ and $\lambda_1 = 5, \lambda_2 = 1, \lambda_3 = 0.5$. The upper bound for control inputs $|u_k|$ is $\pi = 4$. We can find that our control policy with constraints has $|u_k| \leq 4$ for all $k$, while the original policy without constraints exceeds the bound at $k = 4$.

VI. CONCLUSION

In this paper, an optimal control method considering state unpredictability for mobile agents is proposed. We add uni-

| $\lambda_3$ | 0 | 0.2 | 0.5 | 1 |
|---|---|---|---|---|
| Ave. Error | 0.401 | 0.850 | 1.037 | 1.251 |
| Max Error | 1.389 | 2.440 | 2.795 | 3.124 |
The Bellman function at time $N$ where $J$, where $q$ states and guarantees the security of mobile agents, but also with the input-constrained systems. Our control method not through dynamic programming. The algorithm can also deal with formly distributed random perturbations to the control inputs and formulate a multi-period convex optimization problem. The expectation and variance of the control inputs are solved through dynamic programming. The algorithm can also deal with the input-constrained systems. Our control method not only maximizes the attackers’ prediction errors to the future states and guarantees the security of mobile agents, but also satisfies the system performance requirements.

**APPENDIX**

Theorem 1 is proved by mathematical induction. According to the boundary conditions of the Bellman function, the value function at $N$ is written as:

$$V(x_N) = \lambda_1 (x_N^T H x_N - 2q_N x_N),$$

where $q_N = x_N^T H$. Let $J_{1,N} = \lambda_1 H, J_{2,N} = 2\lambda_1 q_N$ and $J_{3,N} = 0$. We have:

$$V(x_N) = x_N^T J_{1,N} x_N - J_{2,N} x_N + J_{3,N}.$$  \hspace{1cm} (21)

The Bellman function at time $N-1$ is:

$$V(x_{N-1}) = \min \{ \lambda_2 \mathbb{E}(x_{N-1}^T Q_{N-1} x_{N-1} + u_{N-1}^T R_{N-1} u_{N-1}) + \sum_{i=1}^{m} \frac{\lambda_{3,N-1}}{\sigma_{N-1,i}^2} + \mathbb{E}(V(x_N)) \}.$$ 

Substituting the dynamic model (1) into (21), we obtain:

$$\mathbb{E}(V(x_N)) = \mathbb{E}(x_N^T J_{1,N} x_N) - J_{2,N} \mathbb{E}(x_N) + J_{3,N}.$$ 

The function $V(x_{N-1})$ is simplified as:

$$V(x_{N-1}) = x_{N-1}^T W_{N-1} x_{N-1} - Z_{N-1} x_{N-1} + J_{3,N} + \min\{ K_{N-1} \mu_{N-1} + \mu_{N-1}^T P_{N-1} \mu_{N-1} + \text{Tr}(\Sigma_{N-1} P_{N-1}) + \sum_{i=1}^{m} \frac{\lambda_{3,N-1}}{\sigma_{N-1,i}^2} \},$$

where

$$W_k = \lambda_2 Q_k + A_k^T J_{1,k+1} A_k,$$

$$Z_k = J_{2,k+1} A_k,$$

$$P_k = \lambda_2 R_k + B_k^T J_{1,k+1} B_k,$$

$$K_k = 2x_k^T A_k^T J_{2,k+1} B_k - J_{2,k+1} B_k,$$

and $\Sigma_{N-1}$ is the covariance matrix of $\delta_{N-1}$. By differentiating $\mu_{N-1}$ and $\sigma_{N-1,i}$ separately, we have

$$\frac{\partial V(x_{N-1})}{\partial \mu_{N-1}} = 2P_{N-1} x_{N-1} + K_{N-1} = 0$$

$$\frac{\partial V(x_{N-1})}{\partial \sigma_{N-1,i}} = 2P_{N-1,i} \sigma_{N-1,i} - 2 \frac{\lambda_3 \mu_{N-1}}{\sigma_{N-1,i}^2} = 0.$$ 

Therefore the global optimal solution is:

$$\begin{align*}
\mu_{N-1} &= -G_{N-1} x_{N-1} + M_{N-1} \\
\sigma_{N-1,i} &= \left( \frac{\lambda_3 \sigma_{N-1,i}}{\mu_{N-1}^2} \right)^{\frac{1}{2}}.
\end{align*}$$

Substitute $\mu_{N-1}$ and $\sigma_{N-1}$ into $V(x_{N-1})$. Let

$$J_{1,k} = W_k - A_k^T J_{1,k+1} A_k,$$

$$J_{2,k} = Z_k - J_{2,k+1} B_k G_k,$$

$$J_{3,k} = -\frac{1}{2} J_{2,k+1} B_k M_k + 2 \sum_{i=1}^{m} (\lambda_{3,k} P_{k,i})^{\frac{1}{2}},$$

then we have

$$V(x_{N-1}) = x_{N-1}^T J_{1,N-1} x_{N-1} - J_{2,N-1} x_{N-1} + J_{3,N-1}.$$ 

Continue the above process for $k = N-2, \ldots, 0$, then the mathematical induction is done.

**REFERENCES**

[1] T. Raunholt, I. Rodriguez, P. Mogensen, and M. Larsen, “Towards a 5G mobile edge cloud planner for autonomous mobile robots,” in 2021 IEEE 94th Vehicular Technology Conference.

[2] Q. Jiao, Y. Li, X. Duan, J. He, and Q.-G. Wang, “Multi-robot target search under multi-peak distribution: A dynamic approach based on high confidence area,” in 2021 IEEE 94th Vehicular Technology Conference.

[3] J. Connors and G. Elkaim, “Analysis of a spline based, obstacle avoiding path planning algorithm,” in 2007 IEEE 65th Vehicular Technology Conference.

[4] F. Pasqualetti, F. Dörfler, and F. Bullo, “Attack detection and identification in cyber-physical systems,” IEEE Transactions on Automatic Control, vol. 58, no. 11, pp. 2715–2729, 2013.

[5] S. Bijani and D. Robertson, “A review of attacks and security approaches in open multi-agent systems,” Artificial Intelligence Review, vol. 42, no. 4, pp. 607–636, 2014.

[6] C. Qiu, J. He, J. Li, C. Fang, and Y. Mo, “Moving target interception considering dynamic environment,” in Proceedings of IEEE ACC, 2022.

[7] C. G. Prevost, A. Desbiens, and E. Gagnon, “Extended Kalman filter for state estimation and trajectory prediction of a moving object detected by an unmanned aerial vehicle,” in Proceedings of IEEE ACC, 2007.

[8] J. Schulz, C. Houben, J. Lüchner, and D. Burschka, “Multiple model unscented Kalman filtering in dynamic Bayesian networks for intention estimation and trajectory prediction,” in 2018 21st International Conference on Intelligent Transportation Systems.

[9] A. Tsiamis, A. B. Alexandr, and G. J. Pappas, “Motion planning with secrecy,” in Proceedings of IEEE ACC, 2019.

[10] M. Hibbard, Y. Savas, B. Wu, T. Tanaka, and U. Topcu, “Unpredictable planning under partial observability,” in Proceedings of IEEE CDC, 2019.

[11] G. Bianchin, Y.-C. Liu, and F. Pasqualetti, “Secure navigation of robots in adversarial environments,” IEEE Control Systems Letters, vol. 4, 2019.

[12] J. Li, J. He, Y. Li, and X. Guan, “Unpredictable trajectory design for mobile agents,” in Proceedings of IEEE ACC, 2020.

[13] R. Bellman, “Dynamic programming,” Science, vol. 153, 1966.

[14] J. B. Mare and J. A. De Doná, “Solution of the input-constrained LQR problem using dynamic programming,” Systems & Control Letters, vol. 56, 2007.