Probabilistic Safety Certificate for Multi-agent Systems

Haoming Jing[1] and Yorie Nakahira[1]

***Abstract*— This paper focuses on the multi-agent safe control problem for stochastic systems. We propose a probabilistic certificate for safety and performance specifications and use it to construct a distributed algorithm. The certificate integrates the reachability- and invariance-based (barrier-function-based) approaches via a new notion of forward invariance defined on the long-term probability. The proposed method has two features. First, it can guarantee a long-term probability of safety and performance satisfaction using myopic evaluation. Second, each agent can collaboratively ensure system-wide specifications even if each does not have sufficient information to evaluate the specifications. The effectiveness of the proposed method is tested using numerical experiments.**

I. Introduction

Multi-agent autonomous systems must balance safety and performance specifications in uncertain environments with distributed control in real-time. There can be information sharing constraints between agents due to limited communication or the scale of the network. In such systems, an agent may only have access to the information of a small subset of the whole network or its neighboring agents. Despite the information sharing constraints, the safety and performance specifications are often given as global specifications that need to be ensured in the long-term. For example, swarms of autonomous agents must collaboratively achieve some common goal (*e.g.,* when a swarm of surveillance drones need to collectively cover the search areas, and when at least one robot should reach a target area to perform some tasks). Safety (*e.g.,* collision avoidance, stability) must also be satisfied at all times. Nonlinear systems can have an unsafe (unstable) region of attractions, which often cannot be avoided by myopically moving away from unsafe regions. Moreover, the environments in which these agents operate can be highly uncertain and dynamic. These uncertainties can come from a multitude of factors, arising from human and other agents' behaviors, disturbance and noise, limited communications, and unmodeled dynamics. Due to the highly dynamic nature, agents must have a fast feedback loop and respond quickly. Such latency requirements may prohibit the use of cloud computation and delayed communication, and the agents' control actions often need to be computed using onboard hardware. To tackle these challenges, this paper studies how to ensure long-term global objectives with information sharing constraints and limited online computation in uncertain environments.

Related work. There has been great advancement in safe control techniques for uncertain or multi-agent systems in the past decade. When the control objective is to avoid obstacles (one agent crashes into another), the existing literature has proposed to use a distance-based barrier function that can be evaluated using local information [1], [2]. This approach is based on the idea that one only needs to know the distance between two agents to control their possibility of crashes. More generally, this approach works when a control objective can be translated into a local condition whose safe set is defined by a level set of decomposable Barrier or Lyapunov functions [3], [4], [5]. Here, decomposable functions refer to the ones that can be evaluated using only the information available to each agent. Although global objectives are quite common in many multi-agent systems (as stated above), this approach cannot account for global objectives that cannot be represented by non-decomposable Barrier/Lyapunov functions.

There often exists stringent tradeoffs between assuring long-term behaviors vs. computational efficiency. On one hand, there exist model prediction control and reachability-based techniques that account for future trajectories of long time horizon to ensure long-term safety [6], [7], [8]. These techniques are often computationally expensive because the space of possible trajectories exponentially increases with the horizon. To reduce the computation burden, various techniques based on barrier function approaches have been proposed to ensure short-term safety conditions myopically [9], [10], [11]. On the other hand, approaches such as stochastic control barrier functions achieve a significant reduction in computational cost due to their use of myopic controllers, but can result in unsafe behaviors in a longer time horizon due to the compounding probabilities of unsafe events [5], [12], [13], [14]. These approaches cannot control the accumulation of tail distribution and may result in small long-term safe probability. To better account for tradeoffs, we have proposed a framework to ensure long-term safe probability using myopic evaluation for fully observable centralized systems [15]. In this paper, we will generalize our prior work to distributed systems with information sharing constraints.

Contribution of this paper. In this paper, we propose a stochastic safe control technique that can ensure multiple global objectives for multi-agent systems in uncertain environments. We first define a new notion of probabilistic forward invariance and forward convergence which can represent the satisfaction of safety and operational specifications with high probability. The specifications can be given global specifications in the form of unions and intersections of forward invariance and forward convergence conditions. Then, we show a sufficient condition for the probability of the

[1]Haoming Jing and Yorie Nakahira are with the Department of Electrical and Computer Engineering, Carnegie Mellon Universty, {haomingj, ynakahir}@andrew.cmu.edu.

control objectives to be within a desired range. This sufficient condition has two features. First, it can achieve all global objectives using local computation. The global objectives can be something that cannot be represented by decomposable barrier functions. At the same time, the condition can be used by each agent with only local information to certify the safety of an existing action or modify it to satisfy the safety and operational specifications. Second, it can achieve long-term safety or performance specifications using myopic evaluation. The specification can be defined as satisfying forward invariance condition (safety) or forward convergence condition (operational) through an outlook time horizon, while its condition can be evaluated using future evolution of an immediate next step. When the sampling frequency is sufficiently high, the certification and modification scheme can be done using a linear constraint and be integrated into a convex/quadratic program. Using this condition, we propose a distributed control algorithm for each agent: Each agent solves an optimization problem with the linear constraints; the information sharing structure or the decision of control actions has a tree structure that can accommodate the relative priority (power) between agents.

The proposed methods have the following advantages.

Advantage 1: The proposed methods can use local information to ensure global safety and performance specifications. The proposed method can be implemented in a decentralized manner: Each agent can use its local information to certify or modify its control actions based on the sufficient condition described above. If all agents can find a feasible action, the global safety or operational specifications will be satisfied with desired probabilities, even global specifications which are represented using non-decomposable Barrier or Lyapunov functions.

Advantage 2: The proposed methods can ensure long-term safety using myopic evaluation. The proposed method embeds the probability of long-term safety or performance into a Barrier-like function. This embedding allows a new notion of conditional forward invariance to be applied on the long-term probability. This new notion allows each agent to typically evaluate the outcome of the immediate future horizon, only using its local information, to ensure long term probability.

To achieve advantages 1 and 2, our novel definition of conditional probabilistic forward invariance and forward convergence condition (see section III-C for detail) is critical to achieve this property. To the best of our knowledge, there does not exist any existing methods that can achieve advantages 1 and 2 simultaneously.

II. Preliminary

Let $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}^n$, and $\mathbb{R}^{m\times n}$ be the set of real numbers, the set of non-negative real numbers, the set of n-dimensional real vectors, and the set of $m \times n$ real matrices, respectively. Let $x[k]$ be the k-th element of vector x. Let $f : \mathcal{X} \to \mathcal{Y}$ represent that f is a mapping from space $\mathcal{X}$ to space $\mathcal{Y}$. Let $\mathbb{1}\{\mathcal{E}\}$ be an indicator function, which takes 1 when condition $\mathcal{E}$ holds and 0 otherwise. Let $\mathbf{0}_{m\times n}$ be an $m \times n$ matrix with all entries 0. Let $\mathbf{1}_n^{[m]}$ be a length n column vector with the m-th entry 1 and other entries 0. Let $\mathbb{I}_m$ be the $m \times m$ identity matrix. Let $\bigtriangledown_x f$ be the gradient of a real valued function f with respect to x. Let $\mathbb{H}_x f$ be the hessian of a real valued function f with respect to x. Given events $\mathcal{E}$ and $\mathcal{E}_c$, let $\mathbb{P}(\mathcal{E})$ be the probability of $\mathcal{E}$ and $\mathbb{P}(\mathcal{E}|\mathcal{E}_c)$ be the conditional probability of $\mathcal{E}$ given the occurrence of $\mathcal{E}_c$. Given random variables X and Y, let $\mathbb{E}[X]$ be the expectation of X and $\mathbb{E}[X|Y = y]$ be the conditional expectation of X given $Y = y$. We use upper-case letters (*e.g.,* Y) to denote random variables and lower-case letters (*e.g.,* y) to denote their specific realizations.

III. Problem Statement

A. System Description

We consider a multi-agent time-invariant stochastic dynamical system with M agents. The dynamics of agent i, $i \in \{1, 2, \cdots, M\}$, is given by the stochastic differential equation (SDE):

$$dX^i = (F^i(X^i) + G^i(X^i)U^i)dt + \Sigma^i(X^i)dW^i, \quad (1)$$

where $X^i \in \mathbb{R}^{m^i}$ is the system state of agent i, $U^i \in \mathbb{R}^{n^i}$ is the control input of agent i, and $W^i \in \mathbb{R}^{\omega^i}$ captures the system uncertainties of agent i. We assume that W^i is the independent standard Brownian motions with 0 initial value. The value of $\Sigma^i(X^i)$ is determined based on the size of uncertainty in agent i. We assume that the dynamics of agent i does not depend on other agents. Thus, the dynamics of the entire multi-agent system can be written as

$$dX = (F(X) + G(X)U)dt + \Sigma(X)dW, \quad (2)$$

where

$$X = \begin{bmatrix} X^1 \\ X^2 \\ \vdots \\ X^N \end{bmatrix}, U = \begin{bmatrix} U^1 \\ U^2 \\ \vdots \\ U^N \end{bmatrix}, W = \begin{bmatrix} W^1 \\ W^2 \\ \vdots \\ W^N \end{bmatrix}, \quad (3)$$

and

$$F = \text{diag}(F^1, F^2, \cdots, F^M) \quad (4)$$

$$G = \text{diag}(G^1, G^2, \cdots, G^M) \quad (5)$$

$$\Sigma = \text{diag}(\Sigma^1, \Sigma^2, \cdots, \Sigma^M). \quad (6)$$

Let

$$m = \sum_{i=1}^{M} m^i \quad (7)$$

denote the dimension of the state, *i.e.,* $X \in \mathbb{R}^m$. To implement the controller in digital system, we discretize the time into sampled points of equal interval Δt, *i.e.,* $t_k = \Delta t k, \forall k \in \mathbb{Z}_+$. Accordingly, system (1) and (2) can be written in discrete-time as

$$X^i_{k+1} = \mathcal{F}^i(X^i_k, U^i_k, W^i_k) \quad (8)$$

and

$$X_{k+1} = \mathcal{F}(X_k, U_k, W_k), \quad (9)$$

respectively. With slight abuse of notation, we use X_k to denote X evaluated at time $k\Delta t$.

We assume that agent i can access the information of its own states and the states and control inputs of a few other agents. Let $\mathcal{A}^i$ be the set of agents whose states and information can be accessed by agent i. Then, the information available to agent i at time k is given by

$$Q_k^i = \{X_k^j, U_k^l : j \in \mathcal{A}^i, l \in \mathcal{A}^i \backslash \{i\}\}. \tag{10}$$

B. Nominal Controller

We assume the existence of a nominal controller

$$U_k^i = N^i(Q_k^i) \tag{11}$$

for each agent i. The nominal controller is assumed to satisfy some performance specifications, but not necessarily all safety and operational specifications, as we will introduce in section III-C. The proposed framework does not restrict the choice of nominal controllers, and each agent can have different forms of nominal controllers.

C. Design Goal

Our goal is to ensure long term safety of all agents as well as satisfaction of operational specifications. We assume that there are B such specifications, indexed by $j = 1, 2, \cdots, B$, and each specification is represented as follows: at time k, specification j is defined by the event

$$\mathcal{C}_k^j = \{x \in \mathbb{R}^m : \phi_k^j(x) \geq 0\}, \tag{12}$$

where $\phi_k^j(x) : \mathbb{R}^m \to \mathbb{R}$ is a continuous mapping. Here, B is the number of safety/operational specifications. We consider two forms of conditions: forward invariance and forward convergence, formally defined below.

1) Forward Invariance: The forward invariance specifications require the condition to continuously hold. If the j-th condition is given as a forward invariance condition, its satisfaction during time horizon T_k^j is given by

$$S_k^j = \{x_\tau \in \mathcal{C}_k^j, \forall \tau \in \{k, k+1, \cdots, k+T_k^j\}\}. \tag{13}$$

2) Forward Convergence: The forward convergence specifications require the system to satisfy the condition eventually. If the j-th condition is given as a forward convergence condition, its satisfaction before time horizon T_k^j is given by

$$S_k^j = \{\exists \tau \in \{k, k+1, \cdots, k+T_k^j\} \text{ s.t. } x_\tau \in \mathcal{C}_k^j\}. \tag{14}$$

The overall performance specification can be represented by the intersections and/or unions of condition $S^j, j \in \{1, 2, \cdots, B\}$, denoted by S. The design goal is to satisfy the S with probability above $1-\epsilon$ at each time k, *i.e.,*

$$\mathbb{P}(S_k) \geq 1 - \epsilon, \ \forall k \geq 0. \tag{15}$$

The forward invariance specifications and forward convergence specifications are combined in (15). This is different from existing techniques that use two separate processes. In a separate design, the control input calculated based on one specification may compromise other specifications. The advantage of combining them into one condition is to jointly account for multiple specifications of both types and not compromising any specification.

IV. Proposed Method

Here, we present a sufficient condition to achieve the design goal in section IV-A and prove its performance guarantee in section IV-C. Based on this condition, we propose a distributed controller in section IV-B.

A. Conditions to Assure Safety and Operational Specifications

In this subsection, we present a sufficient condition to satisfy the performance and safety specifications. Let

$$\Psi_k(I) := \mathbb{P}(S_k|I) \in \mathbb{R} \tag{16}$$

be the sequence of probability of event S_k conditioned on the information I. We define a new notion of conditional discrete-time generator as below.

Definition 1 (Conditional discrete-time generator)**.** The conditional discrete-time generator A of a discrete-time stochastic process $\{x_k\}_{k\in\mathbb{Z}_+}$ conditioned on another process $\{y_k\}_{k\in\mathbb{Z}_+}$ with sampling interval Δt evaluated at time k is given by

$$A\phi(x_k|y_k) = \frac{\mathbb{E}[\phi(x_{k+1})|y_k] - \mathbb{E}[\phi(x_k)|y_k]}{\Delta t} \tag{17}$$

whose domain is the set of all functions $\phi : \mathbb{R}^n \to \mathbb{R}$ of the stochastic process.

When $x_k = y_k$, this generator can be considered as the discrete-time counterpart of the continuous-time infinitesimal generator. We additionally add the conditioning of y_k in order to capture the information-sharing constraints. Although the value of $A\phi(y_k)$ depends on both x_k and y_k, with slight abuse of notation, for the rest of the paper, we will use $A\phi(y_k)$ where the discrete-time stochastic process x_k in Definition 1 is the full state of the system, *i.e.,* X_k in (9).

We consider the following condition at all time k:

$$A\Psi_k(Q_k^i) \geq -\gamma(\Psi_k(Q_k^i) - (1-\epsilon)), \ \forall k \geq 0. \tag{18}$$

Here, $\gamma : \mathbb{R} \to \mathbb{R}$ is a function that satisfies the following 2 design requirements:

Requirement 1: $\gamma(h)$ is linear in h.

Requirement 2: $\gamma(h) \leq h$ for any $h \in \mathbb{R}$.

The probability measure of $\mathbb{P}(S_k|I)$ is taken over X, the global state, conditioned on Q^i, the information that can be accessed by agent i. Therefore, the values on both sides of (18) can be computed using Q^i. Thus, the form of (18) is advantageous in distributed networks without centralized information or computing (see section IV-B). Interestingly, this localizable property does not require the global safety and operational specifications S to be decomposable (*i.e.,* the design specifications S can depend on the value of all states). This is in stark contrast with the existing literature for deterministic and standard barrier functions: agent i can only evaluate the safety constraint S only depending on the information of Q^i.

Theorem 1. *Consider system* (8) *and* (9)*. We assume the initial condition* $X_0 = x$ *satisfies* $\mathbb{P}(S_0|X_0 = x) \geq 1 - \epsilon$*. If*

at each time k, each agent i generates a control policy that satisfies (18)*, then the following condition holds:*

$$\mathbb{E}[\mathbb{P}(S_k|X_k)] \geq 1-\epsilon,\ \forall k \geq 0. \tag{19}$$

Interestingly, although the conditions in (18) can be imposed by each agent i using its local information Q_k^i, the behavior can be guaranteed for global safety and operational specifications. The proof of theorem 1 is given in section IV-C.

B. Proposed Controller

To efficiently implement condition (18) in real time, we first show that (18) can be implemented as a linear function of U_k^i. We define $\Gamma_{\mathcal{D}}^i : \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{\sum_{j\in\mathcal{A}^i} m^j}$ to be[1]

$$\begin{aligned}&\Gamma_{\mathcal{D}}^i(X_k, a, \delta)\\ =&\{X_k^j + \delta \mathbb{1}\{1 \leq a - \sum_{l=1}^{j-1} m^l \leq m^j\}\mathbf{1}_{m^j}^{[a-\sum_{l=1}^{j-1} m^l]} : j \in \mathcal{A}^i\}.\end{aligned} \tag{20}$$

Note that although $\Gamma_{\mathcal{D}}^i$ takes input of the state of the whole system, only the information available to agent i is required for evaluation. Let $\mathcal{D}^i(X_k)$ be defined as

$$\mathcal{D}^i(X_k) = [\mathcal{D}_{(1)}^i(X_k), \mathcal{D}_{(2)}^i(X_k), \cdots, \mathcal{D}_{(m)}^i(X_k)]^\intercal \in \mathbb{R}^m, \tag{21}$$

where

$$\mathcal{D}_{(a)}^i(X_k) = \frac{\Psi_k(\Gamma_{\mathcal{D}}^i(X_k, a, \Delta)) - \Psi_k(\Gamma_{\mathcal{D}}^i(X_k, a, -\Delta))}{2\Delta}. \tag{22}$$

Here, Δ is the step size to calculate the finite difference of Ψ_k. We additionally define $\Gamma_{\mathcal{H}}^i : \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{\sum_{j\in\mathcal{A}^i} m^j}$ to be

$$\begin{aligned}&\Gamma_{\mathcal{H}}^i(X_k, a, b, \delta_a, \delta_b)\\ =&\{X_k^j + \delta_a \mathbb{1}\{1 \leq a - \sum_{l=1}^{j-1} m^l \leq m^j\}\mathbf{1}_{m^j}^{[a-\sum_{l=1}^{j-1} m^l]}\\ &+\delta_b \mathbb{1}\{1 \leq b - \sum_{l=1}^{j-1} m^l \leq m^j\}\mathbf{1}_{m^j}^{[b-\sum_{l=1}^{j-1} m^l]} : j \in \mathcal{A}^i\}.\end{aligned} \tag{23}$$

Note that although $\Gamma_{\mathcal{H}}^i$ takes input of the state of the whole system, only the information available to agent i is required for evaluation. Let

$$\begin{aligned}&\mathcal{H}^i(X_k)\\ =&\begin{bmatrix} \mathcal{H}_{(1,1)}^i(X_k) & \mathcal{H}_{(1,2)}^i(X_k) & \cdots & \mathcal{H}_{(1,m)}^i(X_k) \\ \mathcal{H}_{(2,1)}^i(X_k) & \mathcal{H}_{(2,2)}^i(X_k) & \cdots & \mathcal{H}_{(2,m)}^i(X_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{H}_{(m,1)}^i(X_k) & \mathcal{H}_{(m,2)}^i(X_k) & \cdots & \mathcal{H}_{(m,m)}^i(X_k) \end{bmatrix},\end{aligned} \tag{24}$$

[1]Here, we use $\mathbb{R}^{\sum_{j\in\mathcal{A}^i} m^j}$ to denote the set that has $\sum_{j\in\mathcal{A}^i} m^j$ real elements.

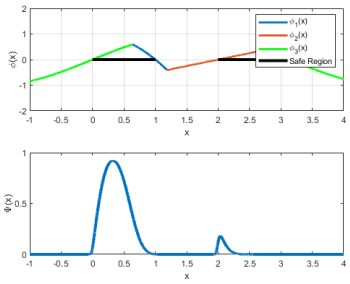

Fig. 1: Min and max composition of 3 barrier functions. Note that although $\phi(x)$ (top plot) is not differentiable, the probablistic formulation $\Psi(x)$ (bottom plot) is smooth.

where

$$\begin{aligned}\mathcal{H}_{(a,b)}^i(X_k) =& \frac{1}{\Delta^2}(\Psi_k(\Gamma_{\mathcal{H}}^i(X_k, a, b, \Delta, \Delta))\\ &- \Psi_k(\Gamma_{\mathcal{H}}^i(X_k, a, b, -\Delta, \Delta))\\ &- \Psi_k(\Gamma_{\mathcal{H}}^i(X_k, a, b, \Delta, -\Delta))\\ &+ \Psi_k(\Gamma_{\mathcal{H}}^i(X_k, a, b, -\Delta, -\Delta))).\end{aligned} \tag{25}$$

Lemma 1. *If the limit of $\lim_{\Delta t\to 0} A\Psi_k(Q_k^i)$ exists, the following condition holds:*

$$\begin{aligned}\lim_{\Delta t\to 0} A\Psi_k(Q_k^i) =& \lim_{\Delta\to 0}(\mathcal{D}^i(X_k)\cdot(F(X_k) + G(X_k)U_k)\\ &+ \frac{1}{2}tr(\Sigma^\intercal(X_k)\mathcal{H}^i(X_k)\Sigma(X_k))).\end{aligned} \tag{26}$$

Remark 1. The function $\Psi_k(x)$ can be smooth even when $\phi(x)$ is not differentiable. For example, consider the case with system (2) with $F = -\frac{1}{2}X$, $G = 0$ and $\Sigma = 2$. The system is discretized with $\Delta t = 0.1$. As an example, we define 3 barrier functions:

$$\phi_1(x) = -x^2 - 1 \tag{27}$$

$$\phi_2(x) = \frac{1}{2}x - 1 \tag{28}$$

$$\phi_3(x) = \sin(x), \tag{29}$$

and specify a composition of $\phi_1(x)$, $\phi_2(x)$, and $\phi_3(x)$, given by

$$\phi(x) = \min(\max(\phi_1(x), \phi_2(x)), \phi_3(x)). \tag{30}$$

Observe that $\phi(x)$ is not differentiable. However,

$$\Psi(x) = \mathbb{P}(\phi(X_\tau) \geq 0, \forall \tau \in \{1, 2, \cdots, 10\}|X_0 = x) \tag{31}$$

is smooth, as shown in fig. 1.

For each agent i, if the information of the a-th entry of X is not accessible, then the a-th entry of $\mathcal{D}^i$ will be 0. Similarly, if either information of the a-th or b-th entry of X is not accessible, then the (a, b)-th entry of $\mathcal{H}^i$ will be 0. Therefore, (26) can be evaluated using only the local information available to agent i and the current control action of other agents whose states are available to agent i, *i.e.*, $\{U_k^j : j \in \mathcal{A}^i\}$. Since the control action of agent i is computed based on the control actions of other agents, the control actions of agents must be available in an order such

that later agents can compute their control actions based on previously available control actions. To calculate (16) using the state information of the agents whose control actions are not available yet and ensure that there exist feasible control actions for these agent, we assume the existence of another controller:

$$U_k^i = R^i(Q_k^i). \tag{32}$$

This controller can be considered to be a controller that is conservative in terms of performance. With this controller, we propose a cascading architecture. We assume that there exists a way to rank all agents such that agent i computes its control action using its own state measurement, and the state measurements and control actions of a subset of agents j, $1 \leq j < i$. Based on theorem 1 and (26), we propose the following constrained optimization problem to find the control action. Each agent i uses Q_k^i to solve

$$\begin{aligned} U_k^i = \arg\min_{u^i} \; & J^i(N(Q_k^i), u^i) \\ \text{s.t. } & \mathcal{D}^i(X_k) \cdot (F(X_k) + G(X_k)u) \\ & + \frac{1}{2}\text{tr}(\Sigma^\intercal(X_k)\mathcal{H}^i(X_k)\Sigma(X_k)) \\ & \geq -\gamma(\Psi_k(Q_k^i) - (1-\epsilon)) \\ & u^j = 0, \; \forall j \notin \mathcal{A}^i. \end{aligned} \tag{33}$$

Here, u contains the control actions of all agents and u^i is the control action of agent i. When computing (16), the agents with index $j < i$ are assumed to use the nominal control action N^j, while the agents with index $j > i$ are assumed to use R^j. The mapping $J^i : \mathbb{R}^{n^i} \times \mathbb{R}^{n^i} \to \mathbb{R}$ is an objective function that penalizes the derivation from the nominal controller policy for agent i. Additional constraints can be added to the optimization problem (33) to account other constraints, such as actuation limits. The proposed algorithm is shown in algorithm 1.

Remark 2. Although the input of $\mathcal{D}^i$ and $\mathcal{H}^i$ is the full state X_k, they can be evaluated using Q_k^i only, as defined in (20) to (25). Therefore, the constraint of (33) can be evaluated using local information Q_k^i only.

Algorithm 1 Proposed control algorithm

```
k ← 0
while k < K_max do
    for i = 1 : M do
        Obtain Q_k^i
        Receive U_k^l, l ∈ A^i\{i}
        Find U_k^i ← solve {u^i in (33)}
    end for
    Execute control actions U_k^i, 1 ≤ i ≤ M
    k ← k + 1
end while
```

Remark 3. In algorithm 1, agents with larger indexes make decisions based on the actions of agents with smaller indexes, so agents with smaller index gets more priority in decision making. Apart from this priority hierarchy, the information sharing structure can also take forms of general tree structures, where the agents on the child nodes make decisions based on the actions of all the nodes on the path to the root node. There exists multiple ways to structure the information sharing structure and choose priorities for agents. One example is based on the physics of the system (*e.g.,* in a truck platooning system, the vehicles in behind make control decisions based on the vehicles before them). Another example is based on pre-defined priority (*e.g.,* in an intersection, emergency vehicles such as ambulance have higher priority in making control decisions compared to other vehicles).

C. Proof of Theorem 1

Lemma 2. *Let S be an event with marginal probability $\mathbb{P}(S)$ and conditional probability $\mathbb{P}(S|Y)$, where Y is a random variable with probability density function $f_Y(y)$. Then, we have the following condition.*

$$\mathbb{E}[\mathbb{P}(S|Y)] = \mathbb{P}(S). \tag{34}$$

Proof (lemma 2). We have

$$\begin{aligned} \mathbb{E}[\mathbb{P}(S|Y)] &= \int_{-\infty}^{\infty} \mathbb{P}(S|Y=y) f_Y(y) dy && (35) \\ &= \mathbb{P}(S) && (36) \end{aligned}$$

due to the law of total probability. ■

Proof (theorem 1). We first show that

$$\mathbb{E}[\Psi_k(X_k)] = \mathbb{E}[\Psi_k(Q_k^i)], \; \forall i \in \{1, 2, \cdots, M\}. \tag{37}$$

We have

$$\begin{aligned} & \mathbb{E}[\Psi_k(X_k)] \\ =& \mathbb{E}[\mathbb{P}(S_k|X_k)] && (38) \\ =& \mathbb{P}(S_k) && (39) \\ =& \mathbb{E}[\mathbb{P}(S_k|Q_k^i)], \; \forall i \in \{1, 2, \cdots, M\} && (40) \\ =& \mathbb{E}[\Psi_k(Q_k^i)], \; \forall i \in \{1, 2, \cdots, M\}. && (41) \end{aligned}$$

Here, (38) and (41) is due to definition (16), and (39) and (40) is due to lemma 2. Therefore, for all $i \in \{1, 2, \cdots, M\}$, we have

$$\begin{aligned} & \mathbb{E}[-\gamma(\Psi_k(Q_k^i) - (1-\epsilon))] \\ =& -\gamma(\mathbb{E}[\Psi_k(Q_k^i)] - (1-\epsilon)) && (42) \\ =& -\gamma(\mathbb{E}[\Psi_k(X_k)] - (1-\epsilon)) && (43) \\ =& \mathbb{E}[-\gamma(\Psi_k(X_k) - (1-\epsilon))]. && (44) \end{aligned}$$

Here, (42) and (44) is due to design requirement 1. In

addition, for all $i \in \{1, 2, \cdots, M\}$, we have

$$\begin{aligned}
&\mathbb{E}[A\Psi_k(Q_k^i)] \\
=&\mathbb{E}[\frac{\mathbb{E}[\Psi_{k+1}(X_{k+1})|Q_k^i] - \mathbb{E}[\Psi_k(X_k)|Q_k^i]}{\Delta t}] && (45)\\
=&\frac{\mathbb{E}[\mathbb{E}[\Psi_{k+1}(X_k^i)|Q_k^i]]}{\Delta t} - \frac{\mathbb{E}[\mathbb{E}[\Psi_k(X_k^i)|Q_k^i]]}{\Delta t} && (46)\\
=&\frac{\mathbb{E}[\Psi_{k+1}(X_{k+1})]}{\Delta t} - \frac{\mathbb{E}[\Psi_k(X_k)]}{\Delta t} && (47)\\
=&\frac{\mathbb{E}[\mathbb{E}[\Psi_{k+1}(X_{k+1})|X_k]]}{\Delta t} - \frac{\mathbb{E}[\Psi_k(X_k)]}{\Delta t} && (48)\\
=&\mathbb{E}[\frac{\mathbb{E}[\Psi_{k+1}(X_{k+1})|X_k] - \Psi_k(X_k)}{\Delta t}] && (49)\\
=&\mathbb{E}[A\Psi_k(X_k)]. && (50)
\end{aligned}$$

Here, (45) and (50) is due to (17), and (47) and (48) is due to the law of total expectation. From (44) and (50), we know that

$$\mathbb{E}[A\Psi_k(Q_k^i)] \geq \mathbb{E}[-\gamma(\Psi_k(Q_k^i) - (1-\epsilon))] \tag{51}$$

implies

$$\mathbb{E}[A\Psi_k(X_k)] \geq \mathbb{E}[-\gamma(\Psi_k(X_k) - (1-\epsilon))]. \tag{52}$$

Here, (51) holds because of safety condition (18).

Next, we use mathematical induction to prove (19). Condition (19) holds for $k = 0$ due to the assumption on initial condition. We suppose (19) holds at time $k > 0$, and show (19) holds at time $k+1$. Let

$$\mathbb{E}[\Psi_k(X_k)] = \mathbb{E}[\mathbb{P}(S_k|X_k)] = 1 - \epsilon + h \tag{53}$$

for some $b > 0$. We define the set of events V_i and variables $v_i, h_i,$ and δ_i, $i \in \{0, 1\}$, as follows:

$$\begin{aligned}
V_0 &= \{\Psi_k(X_k) < 1 - \epsilon\}, && (54)\\
V_1 &= \{\Psi_k(X_k) \geq 1 - \epsilon\}, && (55)\\
v_0 &= \mathbb{E}[\Psi_k(X_k) \mid V_0] = 1 - \epsilon - \delta_0, && (56)\\
v_1 &= \mathbb{E}[\Psi_k(X_k) \mid V_1] = 1 - \epsilon + \delta_1, && (57)\\
h_0 &= \mathbb{P}(V_0), && (58)\\
h_1 &= \mathbb{P}(V_1). && (59)
\end{aligned}$$

The left hand side of (53) can then be written as

$$\begin{aligned}
&\mathbb{E}[\Psi_k(X_k)] \\
=&\mathbb{E}[\Psi_k(X_k) \mid V_0]\,\mathbb{P}(V_0) + \mathbb{E}[\Psi_k(X_k) \mid V_1]\,\mathbb{P}(V_1) \\
=&v_0 h_0 + v_1 h_1. && (60)
\end{aligned}$$

From

$$\begin{aligned}
\mathbb{E}[\Psi_k(X_k) \mid V_0] &< 1 - \epsilon, \\
\mathbb{E}[\Psi_k(X_k) \mid V_1] &\geq 1 - \epsilon,
\end{aligned} \tag{61}$$

we obtain

$$\delta_0 \geq 0 \tag{62}$$

and

$$\delta_1 \geq 0. \tag{63}$$

Moreover, $\{b_i\}_{i\in\{0,1\}}$ satisfies

$$\mathbb{P}(V_0) + \mathbb{P}(V_1) = h_0 + h_1 = 1. \tag{64}$$

Combining (53) and (60) gives

$$1 - \epsilon + h = v_0 h_0 + v_1 h_1. \tag{65}$$

Applying (56) and (57) to (65) gives

$$1 - \epsilon + h = (1 - \epsilon - \delta_0)\, h_0 + (1 - \epsilon + \delta_1)\, h_1, \tag{66}$$

which, combined with (64), yields

$$h = \delta_1 h_1 - \delta_0 h_0. \tag{67}$$

On the other hand, we have

$$\begin{aligned}
&\mathbb{E}\left[\gamma\left(\Psi_k(X_k) - (1-\epsilon)\right)\right] \\
=&\mathbb{P}(V_0)\left(\mathbb{E}\left[\gamma\left(\Psi_k(X_k) - (1-\epsilon)\right) \mid V_0\right]\right) \\
&+ \mathbb{P}(V_1)\left(\mathbb{E}\left[\gamma\left(\Psi_k(X_k) - (1-\epsilon)\right) \mid V_1\right]\right) && (69)\\
=&h_0\left(\mathbb{E}\left[\gamma\left(\Psi_k(X_k) - (1-\epsilon)\right) \mid V_0\right]\right) \\
&+ h_1\left(\mathbb{E}\left[\gamma\left(\Psi_k(X_k) - (1-\epsilon)\right) \mid V_1\right]\right) && (70)\\
=&h_0\left(\gamma\left(\mathbb{E}\left[\Psi_k(X_k) - (1-\epsilon) \mid V_0\right]\right)\right) \\
&+ h_1\left(\gamma\left(\mathbb{E}\left[\Psi_k(X_k) - (1-\epsilon) \mid V_1\right]\right)\right) && (71)\\
=&h_0\left(\gamma\left(-\delta_0\right)\right) + h_1\left(\gamma\left(\delta_1\right)\right) && (72)\\
=&\gamma\left(-h_0\delta_0 + h_1\delta_1\right) && (73)\\
=&\gamma(h) && (74)\\
\leq&h. && (75)
\end{aligned}$$

Here, (70) is due to (58) and (59); (71) is obtained from design requirement 1; (72) is based on (56) and (57); (73) is due to design requirement 1 and (64); (74) is due to (67); and (75) is due to design requirement 2. From (69) to (75), we have

$$\mathbb{E}\left[-\gamma\left(\Psi_k(X_k) - (1-\epsilon)\right)\right] \geq -h. \tag{76}$$

Recall that the control action is chosen to satisfy (18). Now, we take the expectation over both side of (18) to obtain

$$\mathbb{E}[A\Psi_k(X_k)] \geq \mathbb{E}[-\gamma(\Psi_k(X_k) - (1-\epsilon))]. \tag{77}$$

From (17), we have

$$A\Psi_k(X_k) = \frac{\mathbb{E}[\Psi_{k+1}(X_{k+1}) - \Psi_k(X_k)|X_k]}{\Delta t}. \tag{78}$$

Therefore, (77) can be written as

$$\begin{aligned}
&\frac{\mathbb{E}[\mathbb{E}[\Psi_{k+1}(X_{k+1}) - \Psi_k(X_k)|X_k]]}{\Delta t} \\
\geq&\mathbb{E}[-\gamma(\Psi_k(X_k) - (1-\epsilon))].
\end{aligned} \tag{79}$$

Using the law of total expectation, we have

$$\begin{aligned}
&\frac{\mathbb{E}[\Psi_{k+1}(X_{k+1}) - \Psi_k(X_k)]}{\Delta t} \\
\geq&\mathbb{E}[-\gamma(\Psi_k(X_k) - (1-\epsilon))].
\end{aligned} \tag{80}$$

Combining (53), (76), (80) and design requirement 2 yields

$$\begin{aligned}&\mathbb{E}[\Psi_{k+1}(X_{k+1})]\\ \geq&\mathbb{E}[\Psi_k(X_k)] + \mathbb{E}[-\gamma(\Psi_k(X_k) - (1-\epsilon))]\Delta t &(81)\\ \geq&1 - \epsilon + h - h\Delta t &(82)\\ =&1 - \epsilon + h(1-\Delta t). &(83)\end{aligned}$$

Since $\Delta t \ll 1$ and $h \geq 0$, we have

$$\mathbb{E}[\Psi_{k+1}(X_{k+1})] \geq 1 - \epsilon. \tag{84}$$

V. Numerical Simulation

In this section, we test the empirical performance of the proposed method using numerical simulation. We consider a multi-agent system whose setting resembles the group robot operations. Examples of such operations include warehouse robots operations and swarm vehicle operations. The simulation runs for a total time of t_{max}. The system consists a total of M autonomous agents. Let superscript i denote the i-th agent. All agents are governed by the following nonlinear dynamical system:

$$\begin{aligned}dp_t^{xi} &= v_t^i \cos(\theta_t^i)dt &(85)\\ dp_t^{yi} &= v_t^i \sin(\theta_t^i)dt &(86)\\ dv_t^i &= a_t^i dt + \sigma^{vi} dW^{vi} &(87)\\ d\theta_t^i &= \phi_t^i dt + \sigma^{\phi i} dW^{\phi i}, &(88)\end{aligned}$$

where p^{xi} and p^{yi} are the position, v^i is the speed, ϕ^i is the heading angle, a^i is the acceleration, and ϕ^i is the steering rate. The amount of uncertainty is characterized by W^{vi} and $W^{\phi i}$, which we assume are the independent Brownian motions with 0 initial value. For all $i \in \{1, 2, \cdots, M\}$, let

$$\begin{aligned}X_t^i &:= \begin{bmatrix} v_t^i \\ \theta_t^i \end{bmatrix}, U_t^i := \begin{bmatrix} a_t^i \\ \phi_t^i \end{bmatrix}, W^i := \begin{bmatrix} W^{vi} \\ W^{\phi i} \end{bmatrix}\\ G^i &:= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Sigma^i := \begin{bmatrix} \sigma^{vi} & 0 \\ 0 & \sigma^{\phi i} \end{bmatrix}. &(89)\end{aligned}$$

Thus, (87) and (88) can be written as

$$dX_t^i = G^i U_t^i dt + \Sigma^i dW^i. \tag{90}$$

To implement the controller in digital system, we discretize the time into sampled points of equal interval Δt, *i.e.,* $t_k = k\Delta t, \forall k \in \mathbb{Z}_+$, such that (90) can be written in discrete time as

$$X_{k+1}^i = \mathcal{F}^i(U_k^i, W_k^i). \tag{91}$$

Let p_0^{xi} and p_0^{yi} be the starting point of agent i, and p_{goal}^{xi} and p_{goal}^{yi} be the goal of agent i. The set of agents whose information is available to agent i at time k is given by

$$\mathcal{A}_k^i = \{j : \sqrt{(p_k^{xi} - p_k^{xj})^2 + (p_k^{yi} - p_k^{yj})^2} \leq r\}, \tag{92}$$

where r is the maximum range that an agent can broadcast its state and control action information. The operational goal for each agent i is to follow a reference trajectory X^{ri} that enables them to reach the goal, *i.e.,*

$$X_k^{ri} = \begin{bmatrix} v_{max} \\ \text{atan2}(\frac{p_{goal}^{yi} - p_k^{yi}}{p_{goal}^{xi} - p_k^{xi}}) \end{bmatrix}, \tag{93}$$

where v_{max} is the maximum speed. The nominal controller aims to follow this reference using a proportional controller, *i.e.,*

$$N^i(Q_k^i) = K(X_k^i - X_k^{ri}),\ \forall i \in \{1, 2, \cdots, M\}, \tag{94}$$

where K is the controller gain. In addition to the nominal controller, which is considered to give the most aggressive control action, we also assume there exists a controller that gives the most conservative control action. One example is a controller that makes the vehicle decelerate in the maximum rate, *i.e.,*

$$R^i(X_k^i) = \begin{bmatrix} -\text{sign}(v_k^i) a_{max} \\ 0 \end{bmatrix},\ \forall i \in \{1, 2, \cdots, M\}, \tag{95}$$

where a_{max} is the maximum acceleration rate. In addition to the aforementioned agent, we also add an agent, labeled $M+1$, who does not execute safe control policies and whose state is completely unobservable to other agents. However, the other agents knows the initial location, the system dynamics, and the control policy of this agent. Specifically, the system dynamics of this agent is identical to other agents except for having larger uncertainties, and the control policy is identical to the nominal control policy for the other agents, given in (94). The safety specification for all agent is given by

$$\begin{aligned}S_k = \{&\sqrt{(p_\tau^{xi} - p_\tau^{xj})^2 + (p_\tau^{yi} - p_\tau^{yj})^2} \geq l,\\ &\forall \tau \in \{k, k+1, \cdots, k+T\}, i, j \in \{1, 2, \cdots, M+1\}, i \neq j\},\end{aligned} \tag{96}$$

where l is the lowest safe distance, and T is the outlook time horizon. We implement the controller based on Algorithm 1. The objective function in (33) is given by

$$J^i(N(Q_k^i), u^i) = ||N(Q_k^i) - u^i||_2,\ \forall i \in \{1, 2, \cdots, M\}. \tag{97}$$

The key simulation parameters are shown in Table I. In the simulation, we use

$$\gamma(h) = h - 10. \tag{98}$$

Using the same randomly generated starting points and goals, we run simulation with the nominal controller and the proposed control policy. The results are illustrated in Figure 2, and Figure 3. A video of the simulation showing the evolution of the trajectories is available at https://github.com/haomingj/Probabilistic-Safety-Certificate-for-Multi-agent-Systems.

Analysis. The proposed controller is able to ensure safety while preserving the performance of the system. This is shown in Figure 2, where all vehicles reach their goals in the simulation time. In addition, in Figure 3, the successful achievement of the safety objective at all times shows the proposed controller's ability to maintain safety under a few

Parameter	Value	Parameter	Value
Δt	0.01	T	100
K	1	l	0.5
r	10	v_{max}	10
a_{max}	20	ϵ	0.15
M	15	t_{max}	20
$\sigma^{vi},\ \forall i$	2	$\sigma^{\phi i},\ \forall i$	2

TABLE I: Key parameters in the simulation.

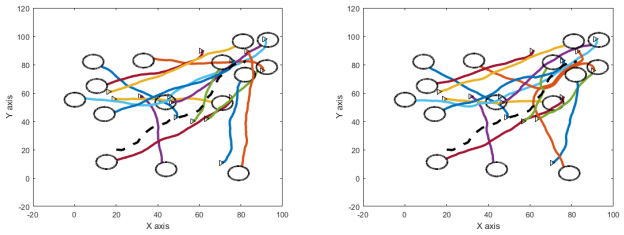

Fig. 2: Agent trajectories with nominal controller (left) and proposed controller (right). The triangles show the starting point and direction for the agents and the circles show the goal regions. The dashed line shows the trajectory of the unobservable agent. All agents reach their goals within simulation time.

challenging conditions. Firstly, the system (85) to (88) is nonlinear. Secondly, the agents only have partial system state information specified in $\mathcal{A}^i$ and cannot evaluate the full safety condition. The state information of one of the agent is completely inaccessible to other agents, which breaks of the assumptions of many existing method since the barrier function cannot be explicitly evaluated. Thirdly, the safety condition is composed of multiple barrier functions representing the distance between all agent pairs. This barrier function can be non-differentiable with respect to state since the closest agent to ego agent may change any time. Since most existing methods are not designed to ensure long term

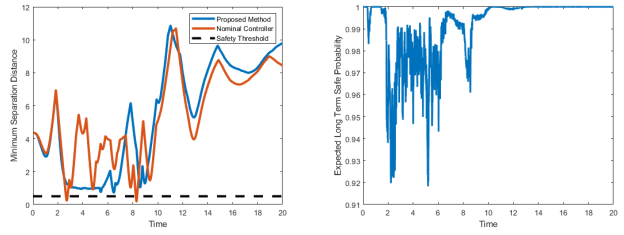

Fig. 3: The minimum distance between any 2 agents (including the unobservable agent) for the nominal controller and the proposed controller (left) and the expected long term safe probability of the proposed method (right). For the nominal controller, the safety specification is violated several times. For the proposed controller, the safety specification is never violated, and the expected safe probability is maintained over $1-\epsilon$.

safety and performance under uncertainty as well as incorporating a binary composition of multiple barrier functions that can only be evaluated locally based on partial information or cannot be explicitly evaluated, comparison with existing methods is not included in the simulation.

VI. Conclusion

In this paper, we propose a safety certificate for multi-agent systems that ensures long term safety and performance using myopic controllers, achieves safety without overly compromising performance, and provides global guarantee on safety and performance conditions that can not be sufficiently evaluated with the information locally available to each agent. We verify the effectiveness of the proposed method in a simulation setting concerning group robot operation.

■

References

[1] H. Zhu, B. Brito, and J. Alonso-Mora, "Decentralized probabilistic multi-robot collision avoidance using buffered uncertainty-aware voronoi cells," *Autonomous Robots*, pp. 1–20, 2022.

[2] D. Claes, D. Hennes, K. Tuyls, and W. Meeussen, "Collision avoidance under bounded localization uncertainty," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 1192–1198.

[3] L. Lindemann and D. V. Dimarogonas, "Control barrier functions for signal temporal logic tasks," *IEEE control systems letters*, vol. 3, no. 1, pp. 96–101, 2018.

[4] R. Cheng, M. J. Khojasteh, A. D. Ames, and J. W. Burdick, "Safe multi-agent interaction through robust control barrier functions with learned uncertainties," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 777–783.

[5] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 3420–3431.

[6] M. Farina, L. Giulioni, and R. Scattolini, "Stochastic linear model predictive control with chance constraints–a review," *Journal of Process Control*, vol. 44, pp. 53–67, 2016.

[7] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, "Learning-based model predictive control: Toward safe learning in control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 269–296, 2020.

[8] M. Chen and C. J. Tomlin, "Hamilton–jacobi reachability: Some recent theoretical advances and applications in unmanned airspace management," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 333–358, 2018.

[9] A. Clark, "Control barrier functions for complete and incomplete information stochastic systems," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 2928–2935.

[10] W. Luo, W. Sun, and A. Kapoor, "Multi-robot collision avoidance under uncertainty with probabilistic safety barrier certificates," *arXiv preprint arXiv:1912.09957*, 2019.

[11] M. Ahmadi, X. Xiong, and A. D. Ames, "Risk-sensitive path planning via cvar barrier functions: Application to bipedal locomotion," *arXiv preprint arXiv:2011.01578*, 2020.

[12] C. Santoyo, M. Dutreix, and S. Coogan, "A barrier function approach to finite-time stochastic system verification and control," *Automatica*, p. 109439, 2021.

[13] V. Dhiman, M. J. Khojasteh, M. Franceschetti, and N. Atanasov, "Control barriers in bayesian learning of system dynamics," *IEEE Transactions on Automatic Control*, 2021.

[14] K. P. Wabersich, L. Hewing, A. Carron, and M. N. Zeilinger, "Probabilistic model predictive safety certification for learning-based control," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 176–188, 2021.

[15] Z. Wang, H. Jing, C. Kurniawan, A. Chern, and Y. Nakahira, "Myopically verifiable probabilistic certificate for long-term safety," *arXiv preprint arXiv:2110.13380*, 2021.