Synthesizing Attack-Aware Control and Active Sensing Strategies Under Reactive Sensor Attacks

Sumukha Udupa, Abhishek N. Kulkarni, *Graduate Student Member, IEEE*, Shuo Han, *Member, IEEE*, Nandi O. Leslie, Charles A. Kamhoua, *Senior Member, IEEE*, and Jie Fu, *Member, IEEE*

***Abstract*—We consider the probabilistic planning problem for a defender (P1) who can jointly query the sensors and take control actions to reach a set of goal states while being aware of possible sensor attacks by an adversary (P2) who has perfect observations. To synthesize a provably-correct, attack-aware joint control and active sensing strategy for P1, we construct a stochastic game on graph with augmented states that include the actual game state (known only to the attacker), the belief of the defender about the game state (constructed by the attacker based on his knowledge of the defender's observations). We present an algorithm to compute a belief-based, randomized strategy for P1 to satisfy the reachability objective with probability one, under the worst-case sensor attacks carried out by an informed P2. We prove the correctness of the algorithm and illustrate it using an example.**

***Index Terms*—Discrete event systems, sensor attacks, cyber-physical system, stochastic games on graphs.**

I. Introduction

IN THIS letter, we develop a formal methods based approach to synthesize provably correct attack-aware cyber-physical systems (CPSs), featured by strategic interactions between a controller/defender and an attacker who carries out sensor attacks on the system. We address the following question: Given the objective of reaching a subset of states in the system, how does one plan the defender's control actions and active information acquisition to satisfy the objective with probability one, under the worst-case sensor attack strategy?

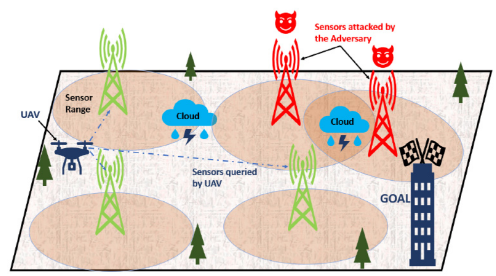

Fig. 1. The strategic interaction between a controller and an adversary.

As a motivating example, consider Fig. 1, where a UAV must reach the flag before its battery is depleted. When the UAV encounters a cloud, it stops moving forward until the cloud moves away. The cloud moves randomly. To complete the task, the UAV deploys a network of sensors to detect the cloud's location. An adversary may attack the sensors to prevent the UAV from accomplishing its mission. Such adversarial interactions include security patrolling robots or search and rescue in a contested environment.

We model the interaction between the defender and the attacker as a partially observable *stochastic* system where the defender's observation is partially controlled by the attacker: At each time step, the defender can choose what sensors to query and control actions to take, and the attacker can choose what sensors to attack. The defender receives observations from the uncompromised sensors and aims to reach a set of *goal states*. In our previous work [1], we analyzed the cost of attack-unaware control where the defender mistakes the compromised sensors as probabilistic sensor failures. This letter investigates the synthesis of *attack-aware controllers with active perception and control*. The key insight is that by knowing which sensors are susceptible to attacks, the defender can selectively choose which sensors to query *in anticipation of the attacker's best response*. Our solution assumes the

Manuscript received 21 March 2022; revised 22 May 2022; accepted 13 June 2022. Date of publication 29 June 2022; date of current version 15 July 2022. This work was supported in part by the Army Research Office under Grant W911NF-22-1-0034, and in part by NSF under Grant CNS-0257512. Recommended by Senior Editor C. Seatzu. *(Corresponding author: Jie Fu.)*

Sumukha Udupa, Abhishek N. Kulkarni, and Jie Fu are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: sudupa@ufl.edu; a.kulkarni2@ufl.edu; fujie@ufl.edu).

Shuo Han is with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: hanshuo@uic.edu).

Nandi O. Leslie is with Raytheon Technologies, Waltham, MA 02451 USA (e-mail: nandi.o.leslie@raytheon.com).

Charles A. Kamhoua is with the Network Security Branch, U.S. Army Research Laboratory, Adelphi, MD 20783 USA (e-mail: charles.a.kamhoua.civ@mail.mil).

Digital Object Identifier 10.1109/LCSYS.2022.3187313

worst-case scenario of asymmetric information: The attacker observes the state and the defender's action before deciding which sensors to attack. Under such a worst-case attacker's information, the attack-aware controller, if it exists, can provide a strong guarantee of the correctness of the closed-loop system.

Our Approach and Contributions: First, we model the adversarial sensor attacks with a new class of turn-based, one-sided, partially observable stochastic games (POSGs), in which the observation function is dynamically changing and jointly determined by the defender and the attacker. Second, we construct an augmented game in which a state includes the actual game state (known by the attacker) and, the belief of the defender about the game state (constructed by the attacker who knows the defender's observation and the actual state). We develop an algorithm to solve a belief-based Almost-Sure Winning (ASW) strategy for the defender in the augmented game and prove that this strategy ensures that the control objective in the original system can be satisfied with probability one, regardless of any sensor attack strategy. The problem is EXPTIME-complete, which matches the lower-bound complexity for one-sided POSG with a fixed observation function [2] (see also the survey on stochastic games on graphs [3]).

Related Work: Our work closely relates to the supervisory control of discrete event systems under sensor/actuator attacks. In the literature on supervisory control, the authors [4] studied various sensor and actuator attacks, including replacement, injection, deletion, and replay of observable and controllable events to a discrete event system (DES) and investigated the controllability of the system and the design of attack-resilient supervisory controllers. Sensor deception attacks have been studied in [5], [6] from the attacker's perspective. The goal is to synthesize a sensor deception attack strategy that misleads the system to reach unsafe states. Given that the system is modeled as a probabilistic automaton, the authors in [5] proposed to construct a $1\frac{1}{2}$ stochastic graph game, also known as a Markov decision process (MDP), and employ a linear program solution of MDPs to design the optimal strategy that maximizes the attack success probability. Covert attacks are investigated in [7] for DES and [8] for networked DES, where the attacker's goal is to remain hidden while compromising the system via stealthy sensor/actuator attacks.

Our game model is different from both deterministic and stochastic DES models. First, deterministic DESs capture the adversarial interactions by a deterministic transition system with controllable and uncontrollable events and observable and unobservable events. In our model, the system dynamics are stochastic, and the observation is partially controlled due to sensor attacks. In works on stochastic DESs [5], [9], the system is modeled as a probabilistic automaton that specifies, for each state, the probability distribution over possible events. This model also differs from our game in which the defender decides an action, and the outcome of that action is determined by a probability distribution. Our model can reduce to a stochastic DES if the defender's policy is fixed. However, we aim to synthesize an attack-aware control strategy for the defender and thereby use the two-player stochastic game formulation.

II. Preliminaries and Problem Formulation

Given a set X, $\mathcal{D}(X)$ is the set of all probability distributions over X, and for a distribution $d \in \mathcal{D}(X)$, $\mathsf{Supp}(d) = \{x \in X \mid d(x) > 0\}$ denotes the support of d.

We introduce a new class of partially observable stochastic games played under asymmetric information. In this game, an autonomous agent (Player 1, P1) actively queries sensors to obtain task-relevant information. Meanwhile, an attacker (Player 2, P2) seeks to compromise P1's mission by reactively blocking the sensor information requested by P1.

Definition 1 (Zero-Sum Stochastic Reachability Game With Partially Controllable Observation Function): A two-player stochastic game with a partially controllable observation function in which P1 has a reachability objective is a tuple

$$G = \langle S, A, P, s_0, \Gamma, \Sigma \times \mathcal{B}, \mathcal{O}, \mathsf{Obs}, o_0, F\rangle,$$

where 1) $\langle S, A, P, s_0\rangle$ is an MDP where S is a finite set of states; A is a finite set of actions; s_0 is an initial state; $P : S\times A \to \mathcal{D}(S)$ is a probabilistic transition function such that $P(s, a, s')$ is the probability of reaching state s' given action a taken at state s. 2) $\Gamma = \{0, 1, \ldots, N\}$ is a set of indexed sensors. 3) $\Sigma \subseteq 2^\Gamma$ is a set of sensor query actions of P1, each of which acquires sensing information from a subset of sensors from Γ; 4) $\mathcal{B} \subseteq 2^\Gamma$ is a set of sensor attack actions of P2, each of which blocks sensing information of a subset of sensors from Γ, similar to jamming attacks [10], [11]. 5) $\mathcal{O} \subseteq 2^S$ is a set of observations. 6) $\mathsf{Obs} : S \times \Sigma \times \mathcal{B} \to \mathcal{O}$ is a deterministic observation function of P1, which maps a state $s \in S$, a sensor query action σ, and a sensor attack action β into an observation $o = \mathsf{Obs}(s, \sigma, \beta) \in \mathcal{O}$. 7) $o_0 \in \mathcal{O}$ is an initial observation and $s_0 \in o_0$. 8) $F \subseteq S$ is a set of final states. P1 must enforce a visit to F to satisfy the *reachability objective*.

In contrast to the standard POSG models [12] where the observation functions are fixed, in our game, the observation function of P1 is determined dynamically by P1's active sensing and P2's reactive sensor attacks. In particular, the observation generated due to P1's sensor query and P2's sensor attack is understood as follows: Each sensor $i \in \Gamma$ covers a subset S_i of states S. Assuming s to be the current state, sensor i returns a Boolean value v_i : $v_i = \mathsf{True}$ if $s \in S_i$, otherwise $v_i = \mathsf{False}$. Given a state $s \in S$, a sensing action $\sigma \in \Sigma$ and a sensor attack action $\beta \in \mathcal{B}$, the observation $\mathsf{Obs}(s, \sigma, \beta)$ of state s is given by

$$\mathsf{Obs}(s, \sigma, \beta) = S - \left(\bigcup_{i\in\sigma\setminus\beta, v_i=\mathsf{True}} S \setminus S_i \cup \bigcup_{i\in\sigma\setminus\beta, v_i=\mathsf{False}} S_i \right).$$

Two states $s, s' \in S$ are said to be observation equivalent given the sensing action and sensor attack action σ, β if $\mathsf{Obs}(s, \sigma, \beta) = \mathsf{Obs}(s', \sigma, \beta)$.

Information structure: In this game, we assume that 1) P2 has perfect observation of states and actions, i.e., P2 can directly observe the current state and P1's control and sensing actions. 2) P1 knows which sensors are attacked by P2—this assumption holds for jamming attacks.

Game Play: The game play in G is constructed as follows. From the initial state s_0, P1 obtains the initial observation o_0. Based on the observation, P1 selects a control action $a_0 \in A$

and a sensor query action $\sigma_0 \in \Sigma$. The system moves to state s_1 with probability $P(s_0, a_0, s_1)$. At state s_1, P2 selects an attack action $\beta_0 \in \mathcal{B}$. The system generates a new observation $o_1 = \mathsf{Obs}(s_1, \sigma_0, \beta_0)$ determined by the state, P1's sensing action and P2's sensor attack action. We denote the resulting play as $\rho = s_0(a_0, \sigma_0)s_1\beta_0(a_1, \sigma_1)s_2\beta_1 \ldots$ Note that P2's attack action is taken after P2 observes the current state. The set of plays in G is denoted by $\mathsf{Plays}(G)$ and the set of finite prefixes of plays is denoted by $\mathsf{Prefs}(G)$.

Observation Equivalent Plays to P1: Given a play $\rho = s_0(a_0, \sigma_0)s_1\beta_0(a_1, \sigma_1)s_2\beta_1 \ldots$, P1's *observation* of ρ is $\rho^o = o_0(a_0, \sigma_0, \beta_0)o_1(a_1, \sigma_1, \beta_1) \ldots$ where $o_{i+1} = \mathsf{Obs}(s_{i+1}, \sigma_i, \beta_i)$ for all $i \geq 0$ and o_0 is the initial observation. For notation convenience, we denote the observation of play ρ as $\mathsf{Obs}(\rho)$. Two plays (or play prefixes) ρ, ρ' are said to be observation-equivalent to P1, denoted by $\rho \sim \rho'$, if and only if $\mathsf{Obs}(\rho) = \mathsf{Obs}(\rho')$.

P1's Reachability Objective and Strategy: A play $\rho = s_0(a_0, \sigma_0)s_1\beta_0(a_1, \sigma_1)s_2\beta_1 \ldots$ is *winning* for P1 if $s_k \in F$ for some $k \geq 0$. Otherwise, it is winning for P2. During interaction, P1 must determine, simultaneously, a control action $a \in A$, and a sensor query action $\sigma \in \Sigma$. We denote P1's set of actions by $\mathcal{A}_1 = A \times \Sigma$, and that of P2 by $\mathcal{A}_2 = \mathcal{B}$. A finite-memory, randomized (resp., deterministic) strategy for player $j \in \{1, 2\}$ is a function $\pi_j : \mathsf{Prefs}(G) \to \mathcal{D}(\mathcal{A}_j)$ (resp., $\pi_j : \mathsf{Prefs}(G) \to \mathcal{A}_j$). A player j is said to follow strategy π_j if for any prefix $\rho \in \mathsf{Prefs}(G)$ at which π_j is defined, player j takes the action $\pi_j(\rho)$ if π_j is deterministic, or an action $a \in \mathsf{Supp}(\pi_j(\rho))$ with probability $\pi_j(\rho, a)$ if π_j is randomized. A strategy is said to be *observation-based* if $\pi_j(\rho) = \pi_j(\rho')$ whenever $\rho \sim \rho'$.

Problem 1: Given a game G in Def. 1, determine if there exists an observation-based strategy using which P1 can satisfy the reachability objective with probability one, for any sensor attack strategy played by P2 with perfect observations.

III. Synthesizing Attack-Aware Strategies With Active Perception

Given P2's perfect observation, P2 can construct P1's belief given his information and higher-order information (what P2 knows P1 observes). To solve Problem 1, we reduce the original game in Def. 1 into a two-player stochastic game whose states include *P2's belief of P1's belief.*

To ease the construction, we introduce a function $\mathsf{Post}_G : 2^S \times A \to 2^S$ that maps a given set of states $B \subseteq S$ and an action $a \in A$ to the possible reachable states $\mathsf{Post}_G(B, a) = \{s' \in S \mid \exists s \in B : P(s, a, s') > 0\}$. We then denote $\mathsf{Post}_G(\{s\}, a)$ as $\mathsf{Post}_G(s, a)$.

Definition 2: Given the zero-sum stochastic game with partially controllable observations G (Def. 1), the stochastic two-player reachability game augmented with P2's belief and P2's belief of P1's belief is a tuple

$$\mathcal{G} = \langle Q \cup \{q_F\}, \mathcal{A}_1 \cup \mathcal{A}_2, \delta, q_0, q_F \rangle,$$

where

- $Q = Q_1 \cup Q_N \cup Q_2$ is the state set consisting of P1, P2 and nature states (c.f. [3]). $Q_1 = \{(s, B) \mid s \in S, B \subseteq S\}$ is the set of states where P1 selects a (control and sensing) action (a, σ). $Q_N = \{(s, B, a, \sigma) \mid s \in S, B \in 2^S, (a, \sigma) \in \mathcal{A}_1\}$ is the set of nature's states. $Q_2 = \{(s, B, \sigma) \mid s \in S, B \in 2^S, \sigma \in \Sigma\}$ is a set of states where P2 selects a sensor attack action. For any $q \in Q$, the first component s is the state in the original game G and the second component B is the belief state that P1 constructs given P1's partial observations.
- q_F is a single final state. It is also a sink state.
- $\mathcal{A}_1 = A \times \Sigma$ is a set of P1's actions and $\mathcal{A}_2 = \mathcal{B}$ is a set of P2's actions.
- $q_0 = (s_0, o_0)$ is the initial state.
- $\delta : (Q_1 \times \mathcal{A}_1) \cup Q_N \cup (Q_2 \times \mathcal{A}_2) \to \mathcal{D}(Q \cup \{q_F\})$ is the probabilistic transition function defined as follows: For a P1's state $(s, B) \in Q_1$ and action $(a, \sigma) \in \mathcal{A}_1$, $\delta((s, B), (a, \sigma), (s, B', a, \sigma)) = 1$, where $B' = \mathsf{Post}_G(B, a)$. That is, with probability one, a nature's state (s, B', a, σ) is reached. For a nature's state $(s, B', a, \sigma) \in Q_N$, we distinguish three cases: 1) If $\mathsf{Post}_G(s, a) \subseteq F$ then $\delta((s, B', a, \sigma), q_F) = 1$. 2) If $\mathsf{Post}_G(s, a) \cap F = \emptyset$, then $\delta((s, B', a, \sigma), (s', B', \sigma)) = P(s, a, s')$. 3) If $\mathsf{Post}_G(s, a) \cap F \neq \emptyset$ and $\mathsf{Post}_G(s, a) \setminus F \neq \emptyset$, then, for some $\epsilon \in (0, 1)$, $\delta((s, B', a, \sigma), q_F) = \epsilon$ and $\delta((s, B', a, \sigma), (s', B', \sigma)) = (1 - \epsilon) \cdot P(s, a, s')$. That is, with some positive probability ϵ, the final, sink state q_F is reached. For a P2's state $(s', B', \sigma) \in Q_2$, and an attack action $\beta \in \mathcal{A}_2$, $\delta((s', B', \sigma), \beta, (s', B'')) = 1$ where $B'' = B' \cap \mathsf{Obs}(s', \sigma, \beta)$.

A sequence of transitions $(s, B) \xrightarrow{(a,\sigma)} (s, B', a, \sigma) \dashrightarrow (s', B', \sigma) \xrightarrow{\beta} (s', B'')$ is understood as follows: At the state (s, B), the true state is s and P1 believes any state in B is possibly the true state. P1 selects a pair of control and sensing actions (a, σ) and updates B to B', which includes a set of states that may be reached if action a is taken at some state in B. Then, the nature player makes a probabilistic transition (represented by the dash arrow) to a new state s' according to the stochastic system dynamics. P2 observes the true state s' and, then, chooses a sensor attack action β. With this sensor attack, P1 observes $\mathsf{Obs}(s', \sigma, \beta)$ and updates P1's belief to eliminate states that are not consistent with the observation.

Definition 2 makes it explicit that while P2 cannot directly control the true state of the game, P2 can affect the augmented state of game $\mathcal{G}$ by influencing the belief of P1. Our belief structure is inspired by stochastic games with signals [13], where a player constructs a belief of his own and the belief of his opponent's belief. However, our modeling and solutions are different from [13].

Next, we describe how to use the game $\mathcal{G}$ augmented with beliefs to solve an attack-aware strategy in the original game G. First, we show that when P2 is limited to blocking sensor readings, regardless of P2's attack, P1 is sure that one of the state in P1's belief is the true state.

Lemma 1: If a state (s, B) is reachable from the initial state q_0, then $s \in B$.

Proof: By induction. The initial state $q_0 = (s_0, o_0)$ satisfies the condition (See Def. 1). Consider a play in the game $\mathcal{G}$ such that q^1_k is the k-th state reached by a sequence of players' actions (P1, P2's actions and the nature's stochastic choices). Suppose $q^1_k = (s_k, B_k)$ that satisfies $s_k \in B_k$. For any action $(a, \sigma) \in \mathcal{A}_1$ of P1, the next state reached is $(s_k, \mathsf{Post}_G(B_k, a), a, \sigma)$. From that state, the nature's probabilistic action will determine the next state $(s_{k+1}, \mathsf{Post}_G(B_k, a), \sigma)$. Note that because $s_k \in B_k$, then $s_{k+1} \in \mathsf{Post}_G(B_k, a)$ by construction.

Then, the attacker P2 takes an action β to generate an observation for P1, $o = \mathsf{Obs}(s_{k+1}, \sigma, \beta)$, which is the set of observation equivalent states. As the attacker can only hide sensor readings, it holds that $\mathsf{Obs}(s_{k+1}, \sigma, \lambda) \subseteq \mathsf{Obs}(s_{k+1}, \sigma, \beta)$ where λ means no attack. And $s_{k+1} \in \mathsf{Obs}(s_{k+1}, \sigma, \lambda)$ implies $s_{k+1} \in \mathsf{Obs}(s_{k+1}, \sigma, \beta)$. The new belief for P1 is $B_{k+1} = o \cap \mathsf{Post}_G(B_k, a)$ and since $s_{k+1} \in o$ and $s_{k+1} \in \mathsf{Post}_G(B_k, a)$, it holds that $s_{k+1} \in B_{k+1}$. ■

This property is critical to construct P1's observation-based strategy to reach F, even if P1 may not know when F is reached. Consider a transition $(s, B', a, \sigma) \dashrightarrow q_F$ where $\mathsf{Post}_G(s, a) \cap F \neq \emptyset$. Since $\mathsf{Post}_G(s, a) \subseteq B'$ implies $B' \cap F \neq \emptyset$, P1 *knows*, without observing, the probability that F is reached is greater than 0.

Definition 3 (Belief-Based Almost-Sure Winning Strategy/Region): Given the two-player game $\mathcal{G}$, a strategy π_1 is *almost-sure winning* for P1 if by following π_1, regardless of P2's strategy, P1 ensures to reach q_F with probability one. A strategy π_1 is *belief-based* provided that for two states $(s, B), (s', B') \in Q_1$, if $B = B'$ then $\pi_1((s, B)) = \pi_1((s', B'))$. A set of states from which P1 has a *belief-based, almost-sure winning strategy* is called P1's almost-sure winning region with partial observation, denoted $\mathsf{Win}_1^{=1}$.

Note that any belief-based strategy is observation-based because the belief is constructed from P1's observations. Next, we prove that by solving the game $\mathcal{G}$ in Def. 2, we can obtain a joint control and active sensing strategy to satisfy the objective against sensor attacks in the game G.

Theorem 1: A belief-based almost-sure winning strategy to reach $\{q_F\}$ in P1's belief-based game $\mathcal{G}$ is also almost-surely winning for P1 to visit F in the game with partially controllable observation function, G, regardless of the sensor attack strategy carried out by P2.

Proof: By the construction of the game $\mathcal{G}$, the event of reaching q_F is conditioned on the event that a nature state $(s, B, a, \sigma) \in Q_N$ where $\mathsf{Post}_G(s, a) \cap F \neq \emptyset$ or $\mathsf{Post}_G(s, a) \subseteq F$ is visited. Let $Y \subseteq Q_N$ be all nature states that can be reached prior to visiting q_F given the almost-sure winning strategy π. If q_F is visited with probability one from any state in the almost-sure winning region, then the set Y must be visited with probability one from any state in $\mathsf{Win}_1^{=1}$. Let $p = \min_{(s,B,a,\sigma)\in Y} \Pr(F \mid s, a)$ be the minimal probability of reaching F from a state in Y. The probability of not reaching F in k visits to Y is smaller than $(1 - p)^k$. In addition, if F is not reached, the almost-sure winning strategy will reach some state $q' \in \mathsf{Win}_1^{=1}$ from which Y is revisited with probability one. Hence, the probability of eventually reaching F is $\lim_{k\to\infty} 1 - (1 - p)^k = 1$. That is, π is also almost-surely winning to visit F in game G. ■

Next, we introduce Alg. 1 to compute a belief-based, ASW *randomized* strategy for P1. The algorithm includes the following steps: In the first step, we use the solution of two-player stochastic games with two-sided perfect observations [14], to solve the positive winning region for P2, denoted $\mathsf{Win}_2^{>0} \subsetneq Q$, which includes a set of states from which P2 can ensure a winning play with a positive probability, when both players have perfect observations. Starting from any state $q \in \mathsf{Win}_2^{>0}$, if P1 cannot reach q_F with probability one even if P1 has perfect observation, then P1 cannot reach q_F with probability one given partial observations.

In the second step, we initialize a set $Y_0 = Q \setminus \mathsf{Win}_2^{>0}$ and iteratively refine the set Y_i to obtain Y_{i+1}, for $i \geq 0$. At iteration i, Alg. 1 computes a set of states, from which P1 can ensure to stay within Y_i with probability one, and with a positive probability, to reach q_F in finite steps. The following functions are defined: For each P1's state $q \in Q_1$, $Y \subseteq Q$, let

$$\mathsf{Allow}(q, Y) = \{(a, \sigma) \in \mathcal{A}_1 \mid \mathsf{Post}_{\mathcal{G}}(q, (a, \sigma)) \subseteq Y\},$$

where $\mathsf{Post}_{\mathcal{G}}(q, (a, \sigma)) = \{q' \mid \delta(q, (a, \sigma), q') > 0\}$ is the set of states that can be reached given P1 applies (a, σ) at state q. By definition, P1 ensures that the next state stays within Y by taking an allowed action in $\mathsf{Allow}(q, Y)$.

Given $q = (s, B) \in Q_1$, let $[q]_\sim = [(s, B)]_\sim = \{(s', B)\} \mid B' = B\}$ be the set of states in which P1 has the same belief as q. We define

$$\mathsf{Allow}([q]_\sim, Y) = \bigcap_{q'\in[q]_\sim} \mathsf{Allow}(q', Y),$$

That is, an action is allowed at q if and only if it is allowed at any other state q' that shares the same belief as q,

Given a set Y and a set $R \subseteq Y$, we define three functions:

$$\mathsf{Prog}_1(R, Y) = \{q \mid \exists (a, \sigma) \in \mathsf{Allow}([q]_\sim, Y), \mathsf{Post}_{\mathcal{G}}(q, (a, \sigma)) \subseteq R\},$$

which outputs a set of states from which P1 has an allowed action to reach R in one step.

$$\mathsf{Prog}_2(R, Y) = \{q \mid \forall \beta \in \mathcal{B}, \mathsf{Post}_{\mathcal{G}}(q, \beta) \subseteq R\},$$

which outputs a set of states from which P2 cannot prevent reaching R in the next step.

$$\mathsf{Prog}_N(R, Y) = \{q \mid \mathsf{Supp}(\delta(q)) \cap R \neq \emptyset \wedge \mathsf{Supp}(\delta(q)) \subseteq Y\},$$

which outputs a set of states from which the nature player can ensure to reach R with a positive probability, while staying within Y with probability one.

In the inner loop of Alg. 1, given a set Y, after initializing $R_0 = \{q_F\}$, Alg. 1 iteratively computes R_{k+1} given R_k for all $k > 0$ until a fixed point is reached. At iteration $k+1$, R_{k+1} is obtained as the union of R_k, $\mathsf{Prog}_1(R_k, Y)$, $\mathsf{Prog}_2(R_k, Y)$ and $\mathsf{Prog}_N(R_k, Y)$. By definition, from any state in $R_{k+1} \setminus R_k$, P1 can ensure to reach R_k with a positive probability. The iteration terminates when $R_{k+1} = R_k$. Then, this new fixed point is the new set Y as the outer fixed point computation. Alg. 1 terminates when $Y_{j+1} = Y_j$ and the fixed point is P1's ASW region $\mathsf{Win}_1^{=1}$.

We establish the correctness and completeness of Alg. 1 by showing that $\mathsf{Win}_1^{=1}$ is indeed the ASW region of P1. And, at any state in $\mathsf{Win}_1^{=1}$, P1 has an ASW strategy to visit F.

Lemma 2: The set $\mathsf{Win}_1^{=1}$ obtained from Alg. 1 is the almost-sure winning region for P1.

Proof: Let N be the index where $Y_N = Y_{N+1}$. To prove that $Y_N = \mathsf{Win}_1^{=1}$, we prove the following: 1) $\mathsf{Win}_1^{=1} \subseteq Y_j$, for all $0 \leq j \leq N$. By induction, first, given that $Y_0 = Q \setminus \mathsf{Win}_2^{>0}$, it holds that $\mathsf{Win}_1^{=1} \subseteq Y_0$. Assume that $\mathsf{Win}_1^{=1} \subseteq Y_i$ for some

Algorithm 1 Belief-Based Almost-Sure Winning Region

Inputs: Two-player reachability game with augmented states $\mathcal{G}$ and P2's positive winning region $\mathsf{Win}_2^{>0}$ in $\mathcal{G}$.

Outputs: P1's ASW region $\mathsf{Win}_1^{=1}$.

```
j ← 0; Y_j ← Q \ Win_2^{>0}
repeat
    k ← 0; R_k ← {q_F}
    repeat
        R_{k+1} ← R_k ∪ Prog_1(R_k, Y_j) ∪ Prog_2(R_k, Y_j) ∪ Prog_N(R_k, Y_j)
        k ← k + 1
    until R_{k+1} = R_k
    Y_{j+1} ← R_k; j ← j + 1
until Y_{j+1} = Y_j
return Win_1^{=1} ← Y_j.
```

$i > 0$, we show that $\mathsf{Win}_1^{=1} \subseteq Y_{i+1}$ as follows: Note that $Y_{i+1} = R_k \cup \mathsf{Prog}_1(R_k, Y_i) \cup \mathsf{Prog}_2(R_k, Y_i) \cup \mathsf{Prog}_N(R_k, Y_i)$. By construction, Y_{i+1} includes any state from which P1 has a strategy to reach q_F with a positive probability, while staying in Y_i. Thus, for any state $q \in Y_i \setminus Y_{i+1}$, either 1) P1 cannot ensure to stay within Y_i with probability one, or 2) P2 has a strategy to ensure q_F is not reached with probability 1. For case 1), if from a state, P1 cannot ensure to stay in Y_i, then that state is not in $\mathsf{Win}_1^{=1}$. This is because for any state in $\mathsf{Win}_1^{=1}$, P1 has a strategy to ensure staying within $\mathsf{Win}_1^{=1}$ and thereby Y_i given $\mathsf{Win}_1^{=1} \subseteq Y_i$. A state satisfying the condition in case 2) is P2's ASW region and thus not in $\mathsf{Win}_1^{=1}$. Therefore, Y_{i+1} only removes states that are not in $\mathsf{Win}_1^{=1}$ from Y_i and thus $\mathsf{Win}_1^{=1} \subseteq Y_{i+1}$.

2) $Y_N \setminus \mathsf{Win}_1^{=1} = \emptyset$, By contradiction, assume there exists a state $q \in Y_N \setminus \mathsf{Win}_1^{=1}$. By construction, for any $q \in R_k \cup \mathsf{Prog}_1(R_k, Y_N) \cup \mathsf{Prog}_2(R_k, Y_N) \cup \mathsf{Prog}_N(R_k, Y_N)$, P1 has a strategy to reach q_F with a positive probability in finitely many steps, regardless of the strategy of P2. Let E_T be the event that "starting from a state in Y_N, a run reaches the final state q_F in T steps." and let $\gamma > 0$ be the minimal probability for the event E_T to occur for any state $q \in Y_N$. Then, the probability of not reaching q_F in infinitely many steps can be upper bounded by $\lim_{k\to\infty} (1-\gamma)^k = 0$. Therefore, for any $q \in Y_N$, P1 has a strategy to ensure q_F is reached with probability one and thus ensures F is reached with probability one (Theorem 1). This contradicts with the assumption that $q \notin \mathsf{Win}_1^{=1}$. Combining 1) and 2), we show that $\mathsf{Win}_1^{=1} = Y_N$. ■

Given $\mathsf{Win}_1^{=1}$, P1's belief-based, ASW strategy is defined by a set-based function $\pi_1^* : \mathsf{Win}_1^{=1} \to 2^{\mathcal{A}_1}$ such that

$$\pi_1^*(q) = \{(a, \sigma) \mid (a, \sigma) \in \mathsf{Allow}([q]_\sim, \mathsf{Win}_1^{=1})\}. \quad (1)$$

At each state $q \in \mathsf{Win}_1^{=1}$, P1 must take any action in $\pi_1^*(q)$ with a nonzero probability. By definition, $\pi_1(q) = \pi_1(q')$ if q, q' share the same belief.

Theorem 2: By following π_1^* defined in Eq. 1, P1 ensures that the game eventually reaches the state q_F.

Proof: Let $R_0, R_1, \ldots, R_K$ be the set of level sets constructed using Alg. 1 given input $\mathsf{Win}_1^{=1}$. For level $0 < j \le K$ and a state $q \in R_j$, let $(a, \sigma) \in \pi_1^*(q)$ such that $\mathsf{Post}_{\mathcal{G}}(q, (a, \sigma)) \in R_{j-1}$. Because taking the action (a, σ) has a nonzero probability, then the level will strictly decrease with a *positive* probability. In addition, with probability one, the game stays within $\mathsf{Win}_1^{=1}$ for any action in $\pi_1^*(q)$ and its probabilistic outcomes. Then, let E_n be the event that "Reaching R_{j-1} from a state in R_j in n steps." It holds that $\lim_{n\to\infty} P(E_n) = 1$. Thus, by repeating the same argument for $j = K, K-1, \ldots, 1$, $R_0 = \{q_F\}$ will be reached with probability one. ■

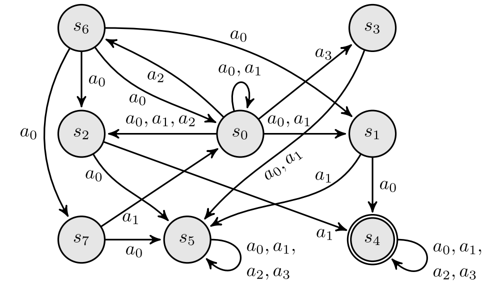

Fig. 2. An example for attack-aware planning. We omit exact transition probabilities and indicate the possible outcomes for each state-action pair.

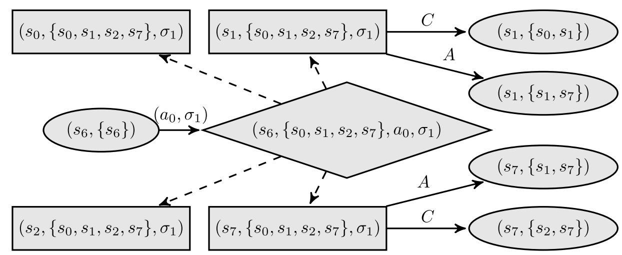

Fig. 3. A fragment of the augmented game $\mathcal{G}$. P1's states are ellipses, P2's states are rectangles, and the nature player's states are diamonds.

Remark 1: Note that in computation, P2's strategy is not restricted to be belief-based. Therefore, for any state $q \in \mathsf{Win}_1^{=1}$, P1 can ensure almost-sure winning regardless of P2's strategy given P2's perfect observation.

Complexity analysis: The time complexity for solving P1's ASW belief-based strategy in $\mathcal{G}$ is $\mathcal{O}(|Q|(|Q_1| \cdot (|A| \cdot |\Sigma|) + |Q_2| \cdot |\mathcal{B}| + |Q_N|))$. In terms of the original game, we have the complexity to be $\mathcal{O}(2^{|S|} \cdot |A| \cdot (|\Sigma| + |\mathcal{B}|))$ due to the subset construction for beliefs. The complexity matches the lower bound for one-sided partial information games [2].

IV. An Illustrative Example

In this section, we present an example to illustrate the proposed algorithm. Consider the MDP shown in Fig. 2, P1 has 5 sensors, A, B, C, D, and E covering the states $\{s_0, s_1\}$, $\{s_1, s_2\}$, $\{s_0, s_2, s_3\}$, $\{s_4, s_5\}$, and $\{s_2, s_6, s_7\}$ respectively and four control actions $\{a_0, a_1, a_2, a_3\}$ with probabilistic outcomes. P1 has four sensor query actions σ_0, σ_1, σ_2, and σ_3 which query the sensors $\{A, B\}$, $\{A, C\}$, $\{B\}$, and $\{B, E\}$ respectively. P1's goal is to reach s_4.

Fig. 3 is a fragment of game $\mathcal{G}$ augmented with beliefs. Starting with P1's state $(s_6, \{s_6\})$, P1 takes action a_0 and queries sensor $\{A, C\}$ with action σ_1. The next state is a nature state $(s_6, \{s_0, s_1, s_2, s_7\}, a_0, \sigma_1)$ where the belief is updated to $\mathsf{Post}_G(\{s_6\}, a_0) = \{s_0, s_1, s_2, s_7\}$ which are possible next states given the action a_0 taken at s_6. Then, the nature player

TABLE I
SUMMARY OF THREE CASES

Case	Almost-sure winning initial states
1. No Attack & 2. Restricted Attack	s_0, s_1, s_2, s_6, s_7
3. Unrestricted Attack	s_0, s_1, s_2, s_7

randomly selects one of the states, say s_1, and arrives at $(s_1, \{s_0, s_1, s_2, s_7\}, \sigma_1)$. If there is no sensor attack, P1 should obtain True for sensor A and False for sensor C and deduce the current state is s_1. However, P2 attacks sensor C so that P1 only receives reading from A and deduces the current state can either be s_0 or s_1—resulting in P1's state $(s_1, \{s_0, s_1\})$.

We use three variations of the example to highlight the system performance given attackers with different capabilities: **Case 1:** No attack: P2 has no sensor attack actions. In this case, P1 planning with joint control and sensing actions; **Case 2:** Restricted Attack: each time P2 can attack one of the sensors from $\{B, E\}$. **Case 3:** Unrestricted Attack: each time P2 can attack any one of the sensors. Assuming P1 knows the initial state, under three cases, the sets of initial states from which P1 has an ASW strategy are shown in Table I. The strategies for three cases were computed in 4.3 s, 6.5 s and 14.1 s on a laptop with AMD RYZEN 9 processor and 16 GB of RAM.

Note that starting from s_6, P1 has ASW strategies for reaching $\{s_4\}$ in Cases 1 and 2, but not in Case 3. Consider Case 1 (no attack), from state s_6, P1 only has action a_0 and thus, can reach one of s_0, s_1, s_2 or s_7 with some positive probability. P1's ASW strategy assigns the following actions with nonzero probabilities: $\{(a_0, \sigma_0), (a_0, \sigma_1), (a_0, \sigma_3)\}$. Action (a_0, σ_2) is not allowed because, with some positive probability the next state is s_1 and P1 refines her belief given the sensor information to $\{s_1, s_7\}$ where P1 has no actions to ensure reaching s_4: In the state s_1, a_0 reaches s_4 and a_1 reaches a sink state s_5. However, at s_7, a_0 leads to the sink state s_5. Similar statement holds if the state s_7 is reached. As P1 must choose between a_0 and a_1 at $(s_1, \{s_1, s_7\})$ and $(s_7, \{s_1, s_7\})$ and there is no belief-based ASW strategy to reach s_4 from these two states.

Though the set of winning initial states are the same for Case 1 and 2. The winning strategies for P1 are different. In Case 2, P1's winning actions at $(s_6, \{s_6\})$ are (a_0, σ_0) and (a_0, σ_1), excluding action (a_0, σ_3) which was a winning action in case 1. The reason is as follows: Suppose (s_0, σ_3) is taken, with a positive probability, the game reaches a P2's state $(s_2, \{s_0, s_1, s_2, s_7\}, \sigma_3)$. P2 attacks sensor E and results in P1's state $(s_2, \{s_1, s_2\})$. Given P1's belief $\{s_1, s_2\}$, the action a_0 is winning for state s_1 but losing for state s_2. Thus, P1 does not have an action at $(s_2, \{s_1, s_2\})$ and $(s_1, \{s_1, s_2\})$ to ensure reaching s_4 with probability one. Thus, action (a_0, σ_3) is not an action from P1's almost-sure winning strategy under P2's restricted attack.

Finally, in Case 3, state s_6 is no longer in $\mathsf{Win}_1^{=1}$. Consider the two actions (a_0, σ_0) and (a_0, σ_1) allowed by the winning strategy at $(s_6, \{s_6\})$ for Case 2. With the action (a_0, σ_0), the game transitions, with a positive probability, to P2 state $(s_1, \{s_0, s_1, s_2, s_7\}, \sigma_0)$, which is not in $\mathsf{Win}_1^{=1}$ as P2 can drive the game into the P1 state $(s_1, \{s_1, s_2\})$ by attacking sensor A. Consider the action (a_0, σ_1), with a positive probability, the game reaches P2 state $(s_7, \{s_0, s_1, s_2, s_7\}, \sigma_1)$, which is not in $\mathsf{Win}_1^{=1}$ as P2 can drive the game to reach $(s_7, \{s_1, s_7\})$ by attacking sensor A. Starting from s_6, P1 has no strategy to reach s_4 if P2 can attack any sensor.

V. CONCLUSION AND FUTURE WORK

In this letter, we studied qualitative planning of control and active information acquisition given adversarial sensor attacks and presented a method to synthesize an observation-based strategy, for the attack-aware defender, to satisfy a reachability objective with probability one, under the worst case sensor attacks on its observations. With the formal-method based modeling framework and solution approach, several future directions are considered: 1) Strategic sensor placement can be studied to ensure the existence of an attack-aware almost-sure winning strategy. 2) The synthesis of P1's strategies can be analyzed given other asymmetric information structures between P1 and P2, including concurrent game interactions, and two-sided partial observations. 3) Symbolic approaches (c.f. [2]) for solving POSGs may be investigated to avoid explicit exponential subset construction.

REFERENCES

[1] A. N. Kulkarni, S. Han, N. O. Leslie, C. A. Kamhoua, and J. Fu, "Qualitative planning in imperfect information games with active sensing and reactive sensor attacks: Cost of unawareness," in *Proc. IEEE Conf. Decis. Control (CDC)*, 2021, pp. 586–592.

[2] K. Chatterjee and L. Doyen, "Partial-observation stochastic games: How to win when belief fails," in *Proc. 27th Annu. IEEE Symp. Logic Comput. Sci.*, 2012, pp. 175–184.

[3] K. Chatterjee and T. A. Henzinger, "A survey of stochastic ω-regular games," *J. Comput. Syst. Sci.*, vol. 78, no. 2, pp. 394–413, 2012.

[4] Y. Wang and M. Pajic, "Supervisory control of discrete event systems in the presence of sensor and actuator attacks," in *Proc. IEEE Conf. Decis. Control (CDC)*, 2019, pp. 5350–5355.

[5] R. Meira-Góes, R. Kwong, and S. Lafortune, "Synthesis of sensor deception attacks for systems modeled as probabilistic automata," in *Proc. Amer. Control Conf. (ACC)*, 2019, pp. 5620–5626.

[6] R. Meira-Góes, E. Kang, R. H. Kwong, and S. Lafortune, "Synthesis of sensor deception attacks at the supervisory layer of cyber–physical systems," *Automatica*, vol. 121, Nov. 2020, Art. no. 109172.

[7] L. Lin and R. Su, "Synthesis of covert actuator and sensor attackers as supervisor synthesis," *IFAC-PapersOnLine*, vol. 53, no. 4, pp. 1–6, 2020.

[8] R. Tai, L. Lin, Y. Zhu, and R. Su, "Synthesis of covert sensor attacks in networked discrete-event systems with non-FIFO channels," 2021, *arXiv:2103.07132*.

[9] J. Chen, M. Ibrahim, and R. Kumar, "Quantification of secrecy in partially observed stochastic discrete event systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 1, pp. 185–195, Jan. 2017.

[10] A. Tayebi, S. M. Berber, and A. Swain, "Wireless sensor network attacks: an overview and critical analysis with detailed investigation on jamming attack effects," in *Sensing Technology: Current Status and Future Trends III*, A. Mason, S. C. Mukhopadhyay, and K. P. Jayasundera, Eds. Cham, Switzerland: Springer Int., 2015.

[11] K. Grover, A. Lim, and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: A survey," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 17, no. 4, pp. 197–215, 2014.

[12] J.-F. Raskin, T. A. Henzinger, L. Doyen, and K. Chatterjee, "Algorithms for omega-regular games with imperfect information," *Logical Methods Comput. Sci.*, vol. 3, no. 3, pp. 1–23, 2007. [Online]. Available: https://lmcs.episciences.org/1094

[13] N. Bertrand, B. Genest, and H. Gimbert, "Qualitative determinacy and decidability of stochastic games with signals," *J. ACM*, vol. 64, no. 5, pp. 1–48, 2017.

[14] R. Bloem, K. Chatterjee, and B. Jobstmann, "Graph games and reactive synthesis," in *Handbook of Model Checking*, E. M. Clarke, T. A. Henzinger, H. Veith, and R. Bloem, Eds. Cham, Switzerland: Springer Int., 2018.