

Smooth Information Gathering in Two-Player Noncooperative Games

Fernando Palafox
University of Texas at Austin
Austin, TX, United States
fernandopalafox@utexas.edu

Jesse Milzman
DEVCOM Army Research Laboratory
Adelphi, MD, United States
jesse.m.milzman.civ@army.mil

Dong Ho Lee
University of Texas at Austin
Austin, TX, United States

Ryan Park
University of Texas at Austin
Austin, TX, United States

David Fridovich-Keil
University of Texas at Austin
Austin, TX, United States

ABSTRACT

We present a mathematical framework for modeling two-player noncooperative games in which one player is uncertain of the other player's costs but can preemptively allocate information-gathering resources to reduce this uncertainty. We refer to the players as the uncertain player (UP) and the certain player (CP), respectively. We obtain UP's decisions by solving a two-stage problem where, in Stage 1, UP allocates information-gathering resources that smoothly transform the information structure in the second stage. Then, in Stage 2, a signal (that is, a function of the Stage 1 allocation) informs UP about CP's costs, and both players execute strategies which depend upon the signal's value. This framework allows for a smooth resource allocation, in contrast to existing literature on the topic. We also identify conditions under which the gradient of UP's overall cost with respect to the information-gathering resources is well-defined. We then provide a gradient-based algorithm to solve the two-stage game. Finally, we apply our framework to a tower-defense game which can be interpreted as a variant of a Colonel Blotto game with smooth payoff functions and uncertainty over battlefield valuations. We include an analysis of how optimal decisions shift with changes in information-gathering allocations and perturbations in the cost functions.

CCS CONCEPTS

• **Mathematics of computing** → **Stochastic optimization and control**; • **Theory of computation** → **Solution concepts in game theory**.

KEYWORDS

Noncooperative game theory; incomplete information; information asymmetry; information-gathering

ACM Reference Format:

Fernando Palafox, Jesse Milzman, Dong Ho Lee, Ryan Park, and David Fridovich-Keil. 2025. Smooth Information Gathering in Two-Player Noncooperative Games. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025, IFAAMAS*, 9 pages.



This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

1 INTRODUCTION

Incomplete information games provide a mathematical formalism for understanding the behavior of rational agents who lack perfect knowledge of one another's objectives or other aspects of the game [7]. In real-world scenarios, agents often attempt to preemptively gather information before such interactions in order to reduce their uncertainty and gain a strategic advantage. However, any player gaining new information may have to contend with other players shifting their strategies in response. Traditionally, such rational information-gathering activities are framed as a discrete choice, e.g., whether or not to pay for costly information with a pre-defined structure [22]. In contrast, real-world information-gathering decisions, e.g., distributed sensor placement [12], are often continuously parametrized.

In this paper, we address this question of continuously parametrized, preemptive information-gathering. To this end, we develop a model for two-player non-cooperative games of incomplete information with the following features:

- The uncertain player (UP) does not know the certain player (CP)'s cost.
- UP can smoothly allocate information-gathering resources to reduce uncertainty, and CP is aware of this allocation.

In such games, there is a coupling between the optimal allocation of information-gathering resources and both players' strategic reactions.

Our contributions are as follows: (1) A game-theoretic model for two-player noncooperative games with one-sided uncertainty and two stages, as shown in Figure 1. In Stage 1, UP selects how to allocate information-gathering resources from a continuous decision landscape. This allocation parameterizes the relationship between a world unknown to UP, and a signal that provides UP with limited information about the world. In Stage 2, UP receives the value of the signal which both players then use to play a non-cooperative game. (2) A local descent algorithm to solve both Stage 1 and Stage 2 for each player's decisions. (3) Conditions under which gradients of costs and solutions with respect to decision variables are well-defined. (4) An application of this model to a tower defense scenario—akin to a Colonel Blotto game with smooth payoff functions—and an analysis of the solutions.

2 RELATED WORK

Interactions in which some players are uncertain about others' objectives were first formally modeled in Harsanyi's seminal work on

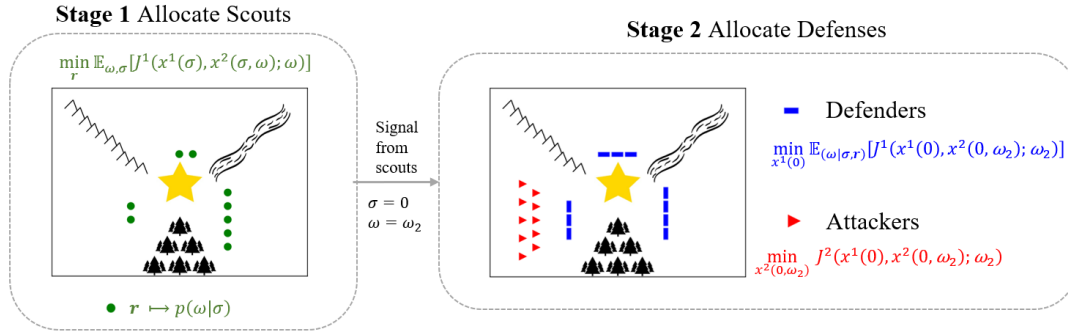


Figure 1: Overall schematic of the two-stage game applied to a tower defense scenario. In this scenario, UP must defend a tower (denoted by a star) that can be attacked from three directions. In Stage 1, UP seeks to minimize its expected cost by allocating information-gathering resources r , and thereby defining posterior distribution $p(\omega|\sigma)$. Here, “world” $\omega \in \{\omega_1, \omega_2, \omega_3\}$ reflects the direction which CP wishes to attack, and “signal” $\sigma \in \{0, 1, 2, 3\}$ is the output of a noisy sensor. In Stage 2, both players seek to minimize their costs J^1, J^2 ; Their allocations (x^1, x^2) are functions of the signal σ (UP, player 1) and/or the world ω (CP, player 2).

Bayesian games [7]. Since then, Bayesian games have been applied to a variety of domains including, but not limited to, cybersecurity of nuclear plants [17], intrusion detection in wireless networks [16], and decision-making in military command and control settings [3, 11]. Much of the existing literature focuses primarily on optimal decision-making in these uncertain environments, typically without examining the capacity of agents to transform the information landscape to their advantage. A notable exception is the literature on deception, e.g., [8, 21, 22].

In single-agent settings, it is common to study this concept of altering the information landscape via the value of information (VoI), which quantifies the expected reduction in cost attributable to a given source of information—i.e., a given variable available to the decision-maker. VoI can be understood as the amount a decision-maker would be willing to pay for information before making a decision [10]. VoI was traditionally related to the comparison of experiments [2] from statistical decision theory (SDT) [19], which has game-theoretic roots [1]. In recent decades, VoI in Bayesian games has been studied as the comparison of information structures [9, 13, 14, 18], which essentially adapts the SDT framework for experiments to the more complex information-cost interactions that emerge from multiple decision-makers.

In this work, we focus on optimizing preemptive actions that selectively gather information to strategically minimize uncertainty in a noncooperative, Bayesian, two-player game. Incorporating VoI-based information gathering into a game where one player has the opportunity to rationally manipulate their information structure *ex ante* is a less-studied phenomenon. To the best of our knowledge, it has not been studied for smooth manipulations of the information structure, as we do in this work.

We highlight a few similar efforts. In earlier work, Hespanha et al. [8] introduced a partial information attacker/defender game in which the defender acts first and may selectively reveal some information about their resource allocation, in order to deceive their opponent. Fuchs and Khargonekar [6] similarly proposed a two-stage, Stackelberg-like partial information Colonel Blotto game. In this game, one player allocates their resources, and then the second player receives a signal (with a given structure) from a

sensor system which alerts them to whether the resources allocated to each battlefield are above a fixed threshold. The second player then uses this information for their allocation. However, they do not consider sensor placement (as a resource allocation problem) or individual sensor tuning, i.e., VoI-based decisions on what information structures would be optimal. More directly comparable to the problem we investigate, Xu and Zhuang [22] set up an incomplete information attacker-defender game, in which the attacker has a costly choice of whether or not to try to learn the vulnerability of the defender—a binary variable chosen by nature. This leads to a characterization of the conditions under which the attacker benefits from trying to learn the defender’s vulnerability.

To the best of our knowledge, no existing literature has explicitly modeled the pre-emptive acquisition of information as a continuously-parametrized decision on the part of one of the players, prior to strategic engagement. Insofar as many real-world informational decisions are smoothly parametrized—e.g. the placement, orientation, and calibration of stationary sensors, or the utilization of drones to conduct reconnaissance along spatially and temporally continuous paths—there is a need for the development of theory and methods for optimizing smooth information structures for decision-making in non-cooperative settings, adversarial or otherwise. This work is an initial effort in that direction.

3 FORMULATION

We begin by defining a noncooperative complete information game and build on it to formulate a Bayesian game with asymmetric uncertainty about CP’s objective. Then, we introduce the concept of signals and signal structure in order to model UP’s information-gathering measures and define a smooth, two-player game with asymmetric and incomplete information. Finally, we formulate the problem of optimally allocating information-gathering as a two-stage problem, in which the continuous decision of the first stage smoothly transforms the information structure in the second stage.

Throughout the paper, superscripts denote indexing by a player, where 1 and 2 denote the uncertain (UP) and certain (CP) players,

respectively. Subscripts denote indexing elements in a vector. We use the following notation to refer to sets $[n] \equiv \{1, \dots, n\}$.

3.1 Complete Information Game

We begin by defining two-player, static games with complete information. Each player i seeks to minimize a cost function $J^i : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ subject to constraints $G^i(x^i) = 0$ and $H^i(x^i) \geq 0$, where $G^i : \mathbb{R}^n \rightarrow \mathbb{R}$, $H^i : \mathbb{R}^n \rightarrow \mathbb{R}$, and n is the dimension of each player's action space.

Mathematically, each player's problem is given by

$$\min_{x^1} J^1(x^1, x^2) \text{ s.t. } x^1 \in \mathcal{X}^1 \quad (1a)$$

$$\min_{x^2} J^2(x^1, x^2) \text{ s.t. } x^2 \in \mathcal{X}^2 \quad (1b)$$

where $\mathcal{X}^i = \{x^i \mid G^i(x^i) = 0, H^i(x^i) \geq 0\}$ is player i 's feasible set. Note that the players' decisions x^i are coupled via $J^i(\cdot)$.

3.2 Bayesian Game With Asymmetric Uncertainty

Now we extend the game to include UP's uncertainty about CP's cost, where UP must make a decision without any new information but its prior knowledge. We introduce uncertainty by parameterizing the costs with an unknown "world" denoted $\omega \in \Omega = [\omega_m]$, where m is the number of worlds. We assume UP has a prior belief about what the world is in the form of a discrete distribution $p : \Omega \rightarrow [0, 1]$, $\sum_{\omega} p(\omega) = 1$. Since UP is uncertain about the value of ω , it now seeks to minimize its expected cost over the prior distribution $p(\omega)$. However, computing this expectation requires knowledge of CP's decisions for every possible world, i.e., $x^2(\omega)$ for every $\omega \in \Omega$. Therefore, the new game is given by

$$\min_{x^1} \mathbb{E}_{\omega} [J^1(x^1, x^2(\omega); \omega)] \text{ s.t. } x^1 \in \mathcal{X}^1 \quad (2a)$$

$$\min_{x^2(\omega_1)} J^2(x^1, x^2(\omega_1); \omega_1) \text{ s.t. } x^2(\omega_1) \in \mathcal{X}^2 \quad (2b)$$

$$\vdots$$

$$\min_{x^2(\omega_m)} J^2(x^1, x^2(\omega_m); \omega_m) \text{ s.t. } x^2(\omega_m) \in \mathcal{X}^2. \quad (2c)$$

Note that the information structure of (2) implies that CP is aware of UP's uncertainty since all of CP's types $x^2(\omega)$, $\omega \in \Omega$ play against a single UP x^1 , a decision made on the basis of the common prior $p(\omega)$.

3.3 Signals and Signal Structures

Our next step is to provide a mathematical framework that describes UP's capacity to deploy information-gathering measures, e.g., by deploying surveillance resources. To that end, we introduce the concept of a signal and a signal structure. Before deciding on the value of x^1 , suppose that UP receives a signal $\sigma \in \mathcal{S} = \{0, 1, \dots, m\}$ with information about the true value of ω . We associate one signal value k for each world ω_k , with a signal of 0 signifying an information-gathering failure. The relationship between signal σ_i and world ω_j is determined by the signal structure: the conditional probability $p(\sigma_i | \omega_j)$, which will be determined by UP's information-gathering decision.

We make the following assumptions regarding the signal structure:

ASSUMPTION 1 (NO FALSE POSITIVES). $p(i | \omega_j) = 0 \forall i \neq j, i > 0$. This implies $p(0 | \omega_i) = 1 - p(i | \omega_i) \forall \omega_i \in \Omega$.

ASSUMPTION 2 (CP'S AWARENESS). CP is aware of both the signal value and signal structure.

From Assumption 1 it follows that signal $i > 0$ always implies that $\omega = \omega_i$. Assumption 2 adds information asymmetry to the interaction: not only is UP uncertain about CP's true intentions, but CP is also aware of the signal structure (e.g., surveillance allocation) and the received signal value. This models a worst-case scenario where CP is fully aware of the information available to the UP. Neither assumption is essential to our framework, but they simplify our presentation and analysis, particularly for Proposition 1.

3.4 Smooth, Two-Player Game with Asymmetric and Incomplete Information

We now introduce a two-player game that encodes the knowledge gained from the signal structure defined in the previous section. UP's prior distribution remains unchanged. In this game, UP makes their decision using the signal value, i.e., $x^1(\sigma)$, and CP makes a decision using both the signal and the world value, i.e., $x^2(\sigma, \omega)$. Thus, we are solving the Bayesian game given by

$$\min_{x^1} \mathbb{E}_{\omega, \sigma} [J^1(x^1(\sigma), x^2(\sigma, \omega); \omega)] \quad (3a)$$

$$\min_{x^2} \mathbb{E}_{\omega, \sigma} [J^2(x^1(\sigma), x^2(\sigma, \omega); \omega)] \quad (3b)$$

where the player strategies are maps of the form

$$x^1 : \mathcal{S} \rightarrow \mathcal{X}^1 \quad (3c)$$

$$x^2 : \mathcal{S} \times \Omega \rightarrow \mathcal{X}^2. \quad (3d)$$

Assumption 1 implies that UP's decision when $\sigma \neq 0$ does not depend on its prior $p(\omega)$, since both players know ω . However, if $\sigma = 0$, ω could still be any ω_i for which $p(0 | \omega_i) > 0$. Therefore, to select $x^1(0)$, UP must minimize the expectation of its cost over the conditional probability $p(\omega_i | 0)$ for every ω_i such that $p(0 | \omega_i) > 0$. Using this information, we may break up the game given by (3) into its component decisions, given by:

$$\min_{x^1(0)} \mathbb{E}_{\omega|0} [J^1(x^1(0), x^2(0, \omega); \omega)] \quad (4a)$$

$$\min_{x^1(1)} J^1(x^1(1), x^2(1, \omega_1); \omega_1) \quad (4b)$$

$$\vdots$$

$$\min_{x^1(m)} J^1(x^1(m), x^2(m, \omega_m); \omega_m) \quad (4c)$$

$$\min_{x^2(0, \omega_1)} J^2(x^1(0), x^2(0, \omega_1); \omega_1) \quad (4d)$$

$$\vdots$$

$$\min_{x^2(0, \omega_m)} J^2(x^1(0), x^2(0, \omega_m); \omega_m) \quad (4e)$$

$$\min_{x^2(1, \omega_1)} J^2(x^1(1), x^2(1, \omega_1); \omega_1) \quad (4f)$$

$$\vdots$$

$$\min_{x^2(m, \omega_m)} J^2(x^1(m), x^2(m, \omega_m); \omega_m), \quad (4g)$$

subject to $x^1(\sigma) \in \mathcal{X}^1, \forall \sigma \in \mathcal{S}$ and $x^2(\sigma, \omega) \in \mathcal{X}^2, \forall \sigma \in \mathcal{S}, \omega \in \Omega$. Note that (4a) depends on the signal structure $p(0 | \omega_i)$ via Bayes' rule. The term $x^1(0)$ in (4a) can be interpreted as UP's decision given no warning and accounting for its knowledge of how it allocated information-gathering resources. For example, consider a situation

where UP allocated enough information-gathering resources to ensure that it will always be warned when $\omega = \omega_1$, i.e., $p(1|\omega_1) = 1$. Then, it need not account for $x^2(0, \omega_1)$ when minimizing the expected cost in (4a) because it knows that $p(0|\omega_1) = 0$.

By contrast, the solutions to the complete information subgames given by (4b)-(4c) and (4f)-(4g) are completely independent of the signal structure.

We remark that this game is a generalization of the Bayesian game described in Section 3.2. To see why, note that setting $p(0|\omega) = 1, \forall \omega \in \Omega$ exactly reduces (4) into (2) since the only relevant decisions are those for which $\sigma = 0$ (as $p(\sigma|\omega) = 0, \forall \sigma > 0$). This corresponds to the case of a UP without the capability to gather new information.

3.5 Signal Structure Selection

UP seeks a signal structure that will strategically minimize their expected cost (3a). To that end, we parametrize the signal structure from Sec. 3.3 with the decision variable $\mathbf{r} \in \mathbb{R}_{\geq 0}^m, \sum_i r_i = 1$ such that $r_i = p(\sigma_i|\omega_i)$. We may then formulate the signal structure selection problem as

$$\min_{\mathbf{r}} \tilde{J} \quad (5a)$$

$$\text{s.t. } 0 \leq r_i \leq 1 \quad (5b)$$

$$\sum_{i=1}^m r_i = 1 \quad (5c)$$

where $\tilde{J} = \mathbb{E}_{\omega, \sigma}[J^1(x^1(\sigma), x^2(\sigma, \omega))]$ for brevity, and (5c) can be interpreted as encoding UP's limited uncertainty-reducing resources, e.g., a limited number of security cameras.

3.6 Two-Stage Problem

Our goal in this work is to develop a decision-making algorithm for a UP aiming to optimally allocate information-gathering assets, and then optimally play a two-player non-cooperative game where the other player's costs are unknown. We now have all the parts to model this scenario as a two-stage problem composed of a signal structure selection problem (5) and the smooth game described in Section 3.4. Given a prior distribution over the worlds $p(\omega)$, UP solves two stages:

- Stage 1: Solve (5) to obtain a signal structure $\mathbf{r}$.
- Stage 2: Given a signal structure, assemble a policy that maps signals to UP decisions, i.e., $x^1 : \mathcal{S} \rightarrow \mathcal{X}^1$ using the solution to (4).

For both stages, we seek locally optimal solutions. In general, the Stage 1 decision landscape will not be convex (see e.g. Fig. 2 in Sec. 5), and thus we cannot guarantee global optimality.

4 SOLVING THE TWO-STAGE PROBLEM

Algorithm 1 summarizes our approach for solving this two-stage problem. At a high level: (i) we solve the two-stage problem by making an initial guess for $\mathbf{r}$, (ii) then, we use this guess to solve Stage 2 in (4), and (iii) finally, we descend the gradient $\frac{d\tilde{J}}{d\mathbf{r}}$ and project the resulting value of $\mathbf{r}$ onto the simplex constraints (5b) and (5c).

As Algorithm 1 depends upon the gradient $\frac{d\tilde{J}}{d\mathbf{r}}$, we first discuss the existence and computation of this derivative. The Stage 1 objective $\tilde{J}$ from (5a) is a function of the decision variables for Stage 2, i.e., $x^1(\sigma), x^2(\sigma, \omega)$. The value of these variables depends on the Stage 1 signal structure selection via (4a). Therefore, when computing $\frac{d\tilde{J}}{d\mathbf{r}}$ we must consider the relationship between the solution of the Stage 2 problem (4) and changes in $\mathbf{r}$. Proceeding formally, we compute the total derivative of $\tilde{J}$ with respect to $\mathbf{r}$ as

$$\frac{d\tilde{J}}{d\mathbf{r}} = \nabla_{\mathbf{r}} \tilde{J} + \nabla_{\mathbf{x}} \tilde{J} \nabla_{\mathbf{r}} \mathbf{x} \quad (6)$$

$$= \nabla_{\mathbf{r}} \tilde{J} + \sum_{j=0}^m \nabla_{x^1(j)} \tilde{J} \nabla_{\mathbf{r}} x^1(j) + \sum_{j,i=0}^m \nabla_{x^2(j, \omega_i)} \tilde{J} \nabla_{\mathbf{r}} x^2(j, \omega_i) \quad (7)$$

Thus, computing $\frac{d\tilde{J}}{d\mathbf{r}}$ requires us to compute $\nabla_{\mathbf{r}} \mathbf{x}$, which is the derivative of a Nash equilibrium solution with respect to parameters of players' objectives, which may not be well-defined, in general. Thus, we offer the following proposition that provides sufficient conditions for the existence of $\frac{d\tilde{J}}{d\mathbf{r}}$ in unconstrained games with $\mathcal{X}^i = \mathbb{R}^n, \forall i \in \{1, 2\}$. In the experimental section we describe how our solver is extended to the constrained setting.

PROPOSITION 1. *Let $\mathbf{r}$ be a point in the relative interior of the simplex, and let $(\mathbf{x}^1, \mathbf{x}^2)$ be a Nash equilibrium solution for an associated Stage 2 game with no constraints ($\mathcal{X} = \mathbb{R}^n$). Then, the gradient $\frac{d\tilde{J}}{d\mathbf{r}}$ exists at $\mathbf{r}$ if the following conditions hold:*

- (1) $\mathbb{E}_{\omega|0}[J^1(\underline{x}^1, \underline{x}^2; \omega)]$ and $J^2(\underline{x}^1, \underline{x}^2_k; \omega_k), k \in [m]$ are twice-differentiable with respect to $\underline{x}^1$ and $\underline{x}^2_k$, respectively.
- (2) $\nabla_{\underline{x}^1}^2 \mathbb{E}_{\omega|0}[J^1(\underline{x}^1, \underline{x}^2; \omega)]$ and each $\nabla_{\underline{x}^2_k}^2 J^2(\underline{x}^1, \underline{x}^2_k; \omega_k), k \in [m]$, are invertible.
- (3) The matrix $\mathbf{E}$ given by

$$\mathbf{E} = \nabla_{\underline{x}^1}^2 \mathbb{E}[J^1] - \sum_{i=1}^m \nabla_{\underline{x}^2_i, \underline{x}^1} \mathbb{E}[J^1] (\nabla_{\underline{x}^2_i}^2 J^2)^{-1} \nabla_{\underline{x}^1, \underline{x}^2_i} J^2$$

is invertible.

where we employ the notation $\underline{x}^1 = x^{1*}(0)$ and $\underline{x}^2_i = x^{2*}(0, \omega_i)$ for brevity. All the matrices are evaluated at $(\mathbf{r}, \mathbf{x}^{1*}, \mathbf{x}^{2*})$

Algorithm 1: Solve Two-Stage Problem

```

1 Input:  $p(\omega)$ , step size  $\alpha \in \mathbb{R}$ , initial guess  $\mathbf{r}_0$ .
2  $\mathbf{r} \leftarrow \mathbf{r}_0$ 
3 while !converged do
4    $x^1, x^2 \leftarrow \text{solveStage2}(\mathbf{r}, p(\omega))$ 
5    $\frac{d\tilde{J}}{d\mathbf{r}} \leftarrow \text{composeDerivative}(\mathbf{r}, x^1, x^2)$ 
6    $\mathbf{r}_0 \leftarrow \mathbf{r} - \alpha \frac{d\tilde{J}}{d\mathbf{r}}$ 
7    $\mathbf{r} \leftarrow \text{project}(\mathbf{r}_0)$  // Onto constraints
8 end
9  $x^1, x^2 \leftarrow \text{solveStage2}(\mathbf{r}, p(\omega))$ 
10 return  $\mathbf{r}, x^1, x^2$ 
```

PROOF. Stage 2 can be decoupled into a set of perfect information games with signals $\sigma > 0$ —i.e., Equations (4b), (4c), (4f) and (4g)—and a single imperfect information game, where $\sigma = 0$ —i.e., Equations (4a), (4d) and (4e). Due to Assumption 1, only the imperfect information game depends on the signal structure selection $\mathbf{r}$, and only these solutions need to be considered when computing $\nabla_{\mathbf{r}}\tilde{J}$. That is, $\nabla_{\mathbf{r}}x^1(j) = 0$ and $\nabla_{\mathbf{r}}x^2(j, \omega_j) = 0$ for all $j > 0$, and thus Eq. 6 reduces to

$$\frac{d\tilde{J}}{d\mathbf{r}} = \nabla_{\mathbf{r}}\tilde{J} + \underbrace{\nabla_{\mathbf{z}}\tilde{J}\nabla_{\mathbf{r}}\mathbf{z}}_{\nabla_{\mathbf{z}}\tilde{J}\nabla_{\mathbf{r}}\mathbf{z}} + \sum_{j=1}^m \nabla_{\mathbf{x}_j^2}\tilde{J}\nabla_{\mathbf{r}}\mathbf{x}_j^2 \quad (8)$$

where $\mathbf{z} \triangleq (\mathbf{x}_1^1, \mathbf{x}_1^2, \dots, \mathbf{x}_m^2)$. The terms $\nabla_{\mathbf{r}}\tilde{J}$ and $\nabla_{\mathbf{z}}\tilde{J}$ are straightforward to compute due to Assumption 1. However, $\nabla_{\mathbf{r}}\mathbf{z}$ requires more work since it is not immediately clear how to differentiate the solution of the Stage 2 subgame (4).

Fortunately, we can find an expression for $\nabla_{\mathbf{r}}\mathbf{z}$ by analyzing the structure of the game solutions and applying the implicit function theorem [4]. First, we define a solution to the Stage 2 game as a Nash equilibrium. Then, using the fact that a Nash equilibrium implies first-order stationarity for all players, we leverage the implicit function theorem and derive sufficient conditions for the existence of $\nabla_{\mathbf{r}}\mathbf{z}$.

Solutions to the Stage 2 game (4) are points that satisfy Nash equilibrium conditions where no player has a unilateral incentive to deviate from their chosen strategy. Concretely, in a game with N players, strategy $\mathbf{x}^* = [x^{1*}, \dots, x^{N*}]$ is a Nash equilibrium if it satisfies

$$J^i(\mathbf{x}^*) \leq J^i([x^i, x^{-i*}]) \quad \forall i \in [N], \quad (9)$$

where $[x^i, x^{-i*}]$ denotes a strategy where only player i deviates from $\mathbf{x}^*$.

Inequality (9) implies that first-order stationarity must hold for all players, i.e.,

$$\mathbf{F} = \begin{bmatrix} \nabla_{\mathbf{x}_1^1}\mathbb{E}_{\omega|0}[J^1(\mathbf{x}_1^{1*}, \mathbf{x}_1^{2*}; \omega)] \\ \nabla_{\mathbf{x}_1^2}J^2(\mathbf{x}_1^{1*}, \mathbf{x}_1^{2*}; \omega_1) \\ \vdots \\ \nabla_{\mathbf{x}_m^2}J^2(\mathbf{x}_1^{1*}, \mathbf{x}_m^{2*}; \omega_m) \end{bmatrix} = \mathbf{0}. \quad (10)$$

Then, assuming certain regularity conditions are met, by the implicit function theorem [4, Thm. 1B.1], there exists a localized solution map $\mathbf{z} : \mathbf{r}' \mapsto \mathbf{z}(\mathbf{r}')$ for any $\mathbf{r}' \in U$, a neighborhood of $\mathbf{r}$ on which $\mathbf{F}(\mathbf{z}(\mathbf{r}'), \mathbf{r}') = \mathbf{0}$. Moreover, for this localized solution map, $\nabla_{\mathbf{r}}\mathbf{z}$ exists and is given by

$$\nabla_{\mathbf{r}}\mathbf{z} = \nabla_{\mathbf{z}}\mathbf{F}^{-1}\nabla_{\mathbf{r}}\mathbf{F}. \quad (11)$$

The following are the sufficient conditions for the existence and differentiability of this solution map in U :

- (1) $\mathbf{F} = \mathbf{0}$
- (2) $\mathbf{F}$ is continuously differentiable
- (3) $\nabla_{\mathbf{z}}\mathbf{F}$ is invertible.

The first condition is already given by (10). The second condition follows if we assume costs are twice-differentiable with respect to their respective decision variables $\mathbf{x}_1^1$ and $\mathbf{x}_i^2$ (Condition 1 in Proposition 1).

To show that the third condition is met we compute $\nabla_{\mathbf{z}}\mathbf{F}$ and analyze its structure. We begin by noting that

$$\nabla_{\mathbf{z}}\mathbf{F} = \begin{bmatrix} \nabla_{\mathbf{x}_1^1}^2\mathbb{E}[J^1] & \nabla_{\mathbf{x}_1^2\mathbf{x}_1^1}\mathbb{E}[J^1] & \dots & \nabla_{\mathbf{x}_m^2\mathbf{x}_1^1}\mathbb{E}[J^1] \\ \nabla_{\mathbf{x}_1^1\mathbf{x}_1^2}J_1^2 & \nabla_{\mathbf{x}_1^2}^2J_1^2 & \dots & \nabla_{\mathbf{x}_0^1\mathbf{x}_m^2}J_1^2 \\ \vdots & \vdots & \ddots & \vdots \\ \nabla_{\mathbf{x}_1^1\mathbf{x}_m^2}J_m^2 & \nabla_{\mathbf{x}_1^2\mathbf{x}_m^2}J_m^2 & \dots & \nabla_{\mathbf{x}_m^2}^2J_m^2 \end{bmatrix}. \quad (12)$$

where $\nabla_{u,v} = \nabla_u\nabla_v$, $J_k^i = J^i(\cdot; \omega_k)$ and we omit the other arguments for brevity. Decisions made by Player 2 are independent for all worlds. Therefore, $\nabla_{\mathbf{x}_i^2\mathbf{x}_j^2}J_j^2 = \mathbf{0}$, $i \neq j$ and all off-diagonal $n \times n$ blocks in the bottom-right $nm \times nm$ block of $\nabla_{\mathbf{z}}\mathbf{F}$ are zero.

$$\Rightarrow \nabla_{\mathbf{z}}\mathbf{F} = \begin{bmatrix} \nabla_{\mathbf{x}_1^1}^2\mathbb{E}[J^1] & \nabla_{\mathbf{x}_1^2\mathbf{x}_1^1}\mathbb{E}[J^1] & \dots & \nabla_{\mathbf{x}_m^2\mathbf{x}_1^1}\mathbb{E}[J^1] \\ \nabla_{\mathbf{x}_1^1\mathbf{x}_1^2}J_1^2 & \nabla_{\mathbf{x}_1^2}^2J_1^2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \nabla_{\mathbf{x}_1^1\mathbf{x}_m^2}J_m^2 & \mathbf{0} & \dots & \nabla_{\mathbf{x}_m^2}^2J_m^2 \end{bmatrix} \quad (13)$$

We can re-write $\nabla_{\mathbf{z}}\mathbf{F}$ as the block matrix

$$\nabla_{\mathbf{z}}\mathbf{F} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \quad (14)$$

where

$$\mathbf{A} = \nabla_{\mathbf{x}_1^1}^2\mathbb{E}[J^1] \quad (15a)$$

$$\mathbf{B} = \left[\nabla_{\mathbf{x}_1^2\mathbf{x}_1^1}\mathbb{E}[J^1] \dots \nabla_{\mathbf{x}_m^2\mathbf{x}_1^1}\mathbb{E}[J^1] \right] \quad (15b)$$

$$\mathbf{C} = \begin{bmatrix} \nabla_{\mathbf{x}_1^1\mathbf{x}_1^2}J_1^2 \\ \vdots \\ \nabla_{\mathbf{x}_1^1\mathbf{x}_m^2}J_m^2 \end{bmatrix} \quad (15c)$$

$$\mathbf{D} = \begin{bmatrix} \nabla_{\mathbf{x}_1^2}^2J_1^2 & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \nabla_{\mathbf{x}_m^2}^2J_m^2 \end{bmatrix}. \quad (15d)$$

If Player 1 and Player 2's cost Hessians are invertible (Condition 2), so are $\mathbf{A}$ and $\mathbf{D}$, and we can express $\det(\nabla_{\mathbf{z}}\mathbf{F})$ in terms of the Schur complement $\nabla_{\mathbf{z}}\mathbf{F}/\mathbf{D}$:

$$\det(\nabla_{\mathbf{z}}\mathbf{F}) = \det(\mathbf{A}) \det(\nabla_{\mathbf{z}}\mathbf{F}/\mathbf{D}), \quad (16)$$

where $\nabla_{\mathbf{z}}\mathbf{F}/\mathbf{D} = \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$. Then, $\nabla_{\mathbf{z}}\mathbf{F}$ is invertible if and only if $\det(\nabla_{\mathbf{z}}\mathbf{F}/\mathbf{D}) \neq 0$ (since $\det(\mathbf{A}) \neq 0$).

If we let $\mathbf{E} = \nabla_{\mathbf{z}}\mathbf{F}/\mathbf{D}$ we have

$$\mathbf{E} = \nabla_{\mathbf{x}_1^1}^2\mathbb{E}[J^1] - \sum_{j=1}^m \nabla_{\mathbf{x}_j^2\mathbf{x}_1^1}\mathbb{E}[J^1](\nabla_{\mathbf{x}_j^2}^2J_j^2)^{-1}\nabla_{\mathbf{x}_1^1\mathbf{x}_j^2}J_j^2. \quad (17)$$

Thus, $\nabla_{\mathbf{z}}\mathbf{F}$ is invertible if and only if $\mathbf{E}$ is invertible (Condition 3).

Therefore, by the implicit function theorem, $\nabla_{\mathbf{r}}\mathbf{z}$ exists if Conditions 1-3 hold. Moreover, using (8), it follows that $\frac{d\tilde{J}}{d\mathbf{r}}$ exists as well. $\square$

5 EXPERIMENTAL RESULTS

In this section we demonstrate our proposed formulation with a zero-sum tower-defense game. We include a variety of experiments meant to give intuition about the Stage 1 cost landscape and the

optimal selection given a selected signal structure. All code is found in a public repository.¹

We define a zero-sum tower-defense game in which the tower, referred to as the defender, may be attacked from one of three directions by an attacker. In Stage 1, the defender (UP) allocates scouts in each direction, corresponding to the selection of a signal structure. These scouts are meant to warn the defender of the preferred attack direction. In Stage 2, the defender receives a signal from its scouts, represented as an integer $\sigma \in \{0, 1, 2, 3\}$. Integers $1 \leq \sigma \leq 3$ inform the defender what the world is, and in every world $\omega \in \Omega$, the attacker has a preferred attack direction. A zero signal corresponds to the case where the defender gets no warning about the world value. After receiving a signal, the defender plays a zero-sum game with the defender using a policy $x^1 : \mathcal{S} \rightarrow \mathcal{X}^1$ that maps signals to defense allocations. Similarly, the attacker plays the game with a policy $x^2 : \mathcal{S} \times \Omega \rightarrow \mathcal{X}^2$, that maps the defender signal (which we assume the attacker knows) and the world (known to the attacker), to attack allocations.

Stage 1 consists of the signal structure selection problem defined in (5) with defender cost function J^1 given by

$$J^1(x^1(i), x^2(i, \omega); \omega) = -J^2(x^1(i), x^2(i, \omega); \omega) \quad (18a)$$

where J^2 will be defined shortly. Each world ω corresponds to an attacker cost function with a preference towards a particular direction of attack.

In Stage 2, the policies x^1, x^2 are assembled from the solution of the Stage 1 game (4) with J^1 and J^2 as defined in this section. They are also constrained such that $\mathcal{X}^i = \{x^i \mid \sum_{j=0}^n x_j^i = 1, x^i \geq \mathbf{0}\}$.

The attacker cost function is given by

$$J^2(x^1(i), x^2(i, \omega), \omega) = - \sum_{j=1}^m \zeta(\delta_j(\omega)) \delta_j^2(\omega) \quad (19a)$$

$$\text{where } \delta_j(\omega) = \beta_j(\omega)x_j^2(i) - x_j^1(i, \omega) \quad (19b)$$

$$\zeta(\delta) = \frac{1}{1 + e^{-2k\delta}} \quad (19c)$$

$$0 < \beta_k(\omega_j) < \beta_j(\omega_j), \forall k \neq j. \quad (19d)$$

The term $\delta_j(\omega)$ in (19b) is the difference between a scaled attacker's allocation $x_j^2(i)$ and the defender allocation $x_j^1(i)$ in direction j . Informally, a large $\delta_j(\omega)$ corresponds to a mismatch between defense and attack allocations in direction j , a win for the attacker. The scaling effect of $\beta_j(\omega)$ can be interpreted as a force multiplier in each direction of attack. Inequality (19d) implies that $\beta_k(\omega_j)$ is larger when $k = j$. Therefore, it is easier for the attacker to gain a numerical advantage in that direction. This is how we encode a direction preference, since the cost in (19a) is decreasing in each δ_j . The cost function in (19c) uses a logistic factor with sharpness parameter k that "activates" only if $\delta > 0$. This ensures that the attacker is not penalized for situations where the defender allocates resources in a direction where the attacker is not present.

We gather the preferences for each world and direction in the matrix $\mathbf{B}$ as follows:

$$\mathbf{B} = \begin{bmatrix} \beta^T(\omega_1) \\ \beta^T(\omega_2) \\ \beta^T(\omega_3) \end{bmatrix} = \begin{bmatrix} \beta_1(\omega_1) & \beta_2(\omega_1) & \beta_3(\omega_1) \\ \beta_1(\omega_2) & \beta_2(\omega_2) & \beta_3(\omega_2) \\ \beta_1(\omega_3) & \beta_2(\omega_3) & \beta_3(\omega_3) \end{bmatrix}. \quad (20)$$

¹github.com/CLeARoboticsLab/GamesVol.jl

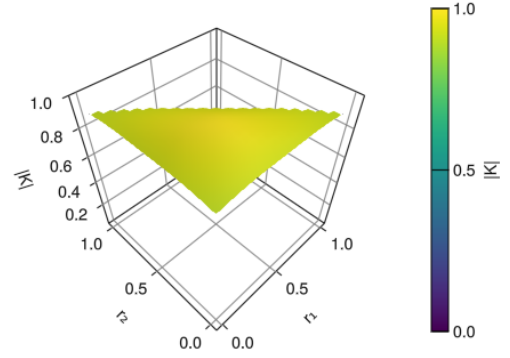


Figure 2: Normalized Stage 1 cost $|J|$ as a function of scout allocation $\mathbf{r}$.

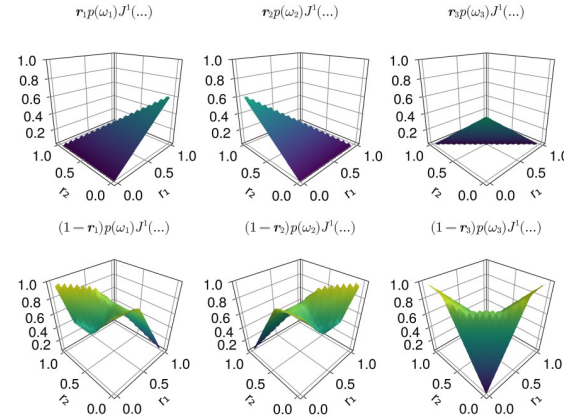


Figure 3: Values for each term in (22c) as a function of $\mathbf{r}$. Normalized by the highest value observed across all six terms.

In this case, (19d) implies that the diagonal elements are the largest for every row.

This experiment can be understood as a version of the Colonel Blotto game with smooth payoff functions. In the standard formulation of a Colonel Blotto game, two players must simultaneously allocate forces across a set of battlefields [20]. At every battlefield, the player with the largest number of forces wins, and the payoff for all players depends on the number of battlefields they win. By contrast, ours is a game of degree, as there is a continuously varying quantity that may be won from each battlefield.

Computing the derivative for this experiment requires considering the effects of the constraints on the solution of Stage 2 as these constraints were not considered in Proposition 1. As shown in Liu et al. [15], when strict complementarity holds, and under mild assumptions, the derivative $\nabla_{\mathbf{r}} \mathbf{z}$ can be uniquely computed. For the case of weak complementarity, the derivative is not well-defined, but subgradients can still be computed. Our implementation of Algorithm 1 accounts for both cases.

5.1 Understanding the Cost Landscape

We begin by plotting the Stage 1 cost as a function of scout allocation. To do so, we first select a valid scout allocation $\mathbf{r}$ and solve (4)

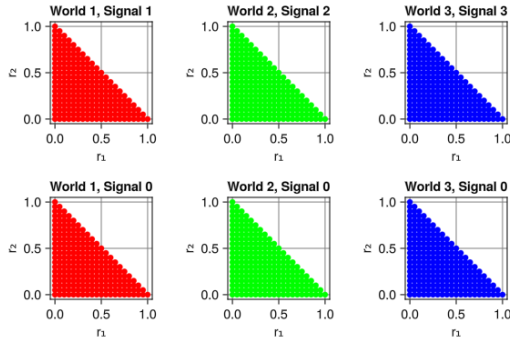


Figure 4: Attacker policy x^2 as a function of $\mathbf{r}$ and signal/world.

to obtain $x^1(\sigma)$ and $x^2(\sigma, \omega)$, $\forall \omega \in \Omega, \sigma \in \mathcal{S}$ for which $p(\sigma, \omega) > 0$. Then, we use these variables to compute the Stage 1 cost (5a). Note that $\mathbf{r}$ belongs to the 2-simplex, and therefore, has only two degrees of freedom. This means we can plot the Stage 1 cost (5a) as a 3-D surface where the (x, y) -axes are r_1 and r_2 , and the z -axis is the normalized expected cost. We normalize the costs by dividing with the highest observed cost among all $\mathbf{r}$ in the simplex.

For this experiment, we start with a uniform prior $p(\omega_i) = \frac{1}{3}$, $\forall i \in \{1, 2, 3\}$, a sharpness parameter $k = 10.0$, and the following attacker preference matrix:

$$\mathbf{B} = \begin{bmatrix} 3.0 & 2.0 & 2.0 \\ 2.0 & 3.0 & 2.0 \\ 2.0 & 2.0 & 3.0 \end{bmatrix}. \quad (21)$$

The resulting plot in Figure 2 reveals the surprising fact that, for this parameter regime, the Stage 1 cost landscape is relatively flat. This implies that scout allocation has only a small effect on the expected cost for the defender at Stage 1.

To understand why, we begin by separating the Stage 1 cost into the six terms from which it is composed:

$$\mathbb{E}_{\omega, \sigma} [J^1(x^1(\sigma), x^2(\sigma, \omega); \omega)] = \quad (22a)$$

$$\sum_{i=1}^3 \mathbf{r}_i p(\omega_i) J^1(x^1(i), x^2(i, \omega_i)) + \quad (22b)$$

$$\sum_{i=1}^3 (1 - \mathbf{r}_i) p(\omega_i) J^1(x^1(0), x^2(0, \omega_i)). \quad (22c)$$

Then, we plot the separate contribution of each term in Figure 3, where we readily see how the structure of the individual terms results in a cancellation effect that leads to the relatively flat Stage 1 cost in Figure 2.

The result is that the optimal scout allocations are found at the vertices of the simplex. This result has an interesting interpretation: it is in the defender's best interest to completely remove uncertainty about one world, instead of distributing its information-gathering resources across many directions.

5.2 Corresponding Attack and Defense Decisions

We now present how players' decisions change as a function of signal structure selection $\mathbf{r}$. To this end, in Figure 5 we color each

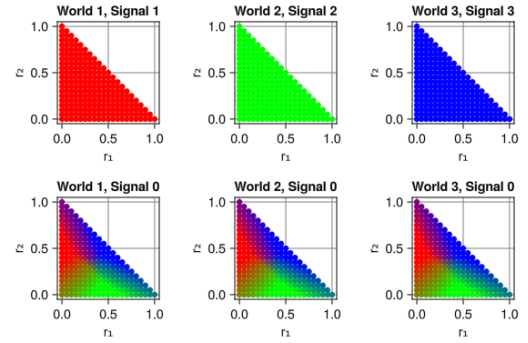


Figure 5: Defender policy x^1 as a function of $\mathbf{r}$ and signal/world.

point in the $\mathbf{r}$ simplex with an RGB color whose component intensity is given by a Stage 2 decision variable. For example, given an $\mathbf{r}$, if the defender's optimal Stage 2 decision for $\sigma = 0$ is $x^1(0) = [0.5, 0.5, 0.0]^T$, then point $\mathbf{r}$ is colored by the (normalized) RGB triplet $(0.5, 0.5, 0.0)$. This figure illustrates what the defense policy x^1 is for every possible combination of signal and signal structure selection.

We note that in the case of detection ($\sigma \neq 0$, top three plots in Figure 5), the defense policy is to allocate all of the resources in the direction of the highest attack preference. For example, consider the top-middle plot in Figure 5 which displays the defender's decisions when $(\omega, \sigma) = (\omega_2, 2)$. The preference matrix (20) implies that in ω_2 , the attacker prefers direction 2. The resulting optimal defensive action is to allocate all resources in that direction, as shown by the green surface with RGB triplet $(0.0, 1.0, 0.0)$. To help understand why, we examine the attack policy, shown in Figure 4, for the same world and signal combination. Note that the surface in the top-middle plot is all green, which means that the attacker allocates all of its resources in direction 2. The interpretation is that even though a signal $\sigma = 2$ informs the defender what the world is, and therefore, what the preferred attack direction is, the optimal choice for the attacker is still to attack in its preferred direction. In response, the defender allocates all of its resources in that same direction. This pattern holds for any of the complete information cases, i.e., $(\omega, \sigma) = (\omega_i, i)$. This situation arises because the benefits conferred by a larger attack multiplier in the preferred direction are large enough to offset the actions taken by a defender who is aware of this preference.

When the defender has no information about what the world is ($\sigma = 0$), its decision will be based only on its knowledge of how it allocated information-gathering resources in Stage 1. This may result in a mixed allocation in different directions. For example, if the defender's Stage 1 decision is $\mathbf{r} = [1, 0, 0]^T$ and it received a signal of 0, then $p(0|\omega_1) = 0$, implying that the value of the world must be ω_2 or ω_3 . Therefore, given a uniform prior and no other information, the optimal decision must be to distribute its resources between the preferred directions in ω_2 and ω_3 . This is seen in Figure 5 for $\sigma = 0$, with the selected decision for the bottom right corner ($\mathbf{r} = [1, 0, 0]^T$) represented in a teal color, which is a mixed allocation between directions 2 and 3 (green and blue, respectively).

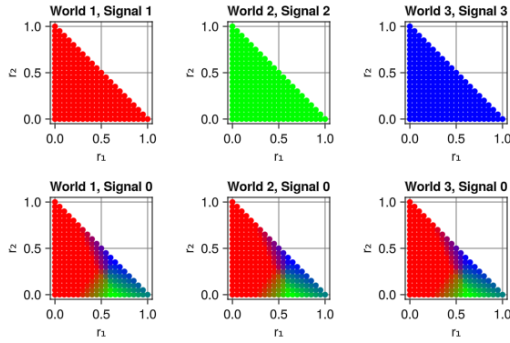


Figure 6: Defender policy x^1 given a perturbed preference matrix B_θ (23) with $\theta = 2$. Red “spreads out” to the right side of the simplex, because the lower probabilities of the risky world (as one moves toward the right vertex of the simplex) are outweighed by the additional expected cost.

5.3 Making One World Riskier

We now examine how the decision landscape changes as we perturb the cost functions. Let us consider the scenario in which one world is much riskier than the others—that is, for one of the worlds, the preferred direction has a larger force multiplier than the force multipliers of the preferred directions of the other worlds. Mathematically, we can model this by defining a perturbed preference matrix B_θ as follows:

$$B_\theta = \begin{bmatrix} 3.0 + \theta & 2.0 & 2.0 \\ 2.0 & 3.0 & 2.0 \\ 2.0 & 2.0 & 3.0 \end{bmatrix}. \quad (23)$$

In Figure 6 we present the resulting defender policy for the case where $\theta = 2$. Compared to the original defender policy in Figure 5, the defender policy remains the same for the complete information cases in the top three plots, i.e., when the world is detected and $(\omega, \sigma) = (\omega_i, i)$. This makes sense because, within each world, the preferred attack direction has not changed, so once the world has been detected the best defense action remains unchanged whether $\theta = 0$ or not.

On the other hand, in the incomplete information case ($\sigma = 0$), the defense policy becomes more biased towards allocating defense resources in the risky direction. The expected cost associated with the risky world remains high enough to bias the minimizer of (4a) towards allocating resources in that direction—even if there is a low probability of that world given $\sigma = 0$ (as given by the scout allocation r).

This explains the results shown in the bottom three plots of Figure 6, where a larger portion of the simplex is red when compared to Figure 5. Moreover, as one increases θ , the red region expands over more of the r -simplex. This is because the additional risk of ω_1 results in a situation where the defender’s best decision is $x^1(0) = [1, 0, 0]^T$, even when $p(\omega_1|0)$ is low (which happens for r closer to the right side of the simplex). We note that the attacker policy remains unchanged and identical to the one seen in Figure 4.

6 CONCLUSION

We provide a two-stage, game-theoretic model for two-player non-cooperative games where an uncertain player can pre-emptively gather information. In the first stage, an uncertain player (UP) chooses how to allocate their information-gathering resources. Critically, this model allows for a smooth resource allocation, in contrast to existing literature on the topic. In the second stage, UP receives information—also available to the certain player (CP)—in the form of a signal, which UP and CP use to play a non-cooperative game. We develop an algorithm that solves this two-stage problem, returning UP’s optimal allocation of information-gathering resources and UP’s policy that maps signals to actions in the noncooperative game.

This work can be extended in several directions. Within Blotto games, one could characterize how the cost landscapes change with different priors, cost parameters, and/or cost functions. One specific cost function of interest is the arctan payoff function introduced by Ferdowsi et al. [5] as part of the Generalized Blotto Game. Other extensions include: (1) relaxing the assumptions on the signal structure, e.g., a game without the “no-false-positive” Assumption 1, (2) applying this to a game with higher dimensional decision variables and/or more worlds, and (3) playing stage 2 many times (perhaps infinitely), allowing for the UP to iteratively update their prior and improve outcomes. (4) A deeper analysis of conditions in Proposition 1, and how they intuitively relate to the player objectives.

ACKNOWLEDGMENTS

We would like to thank Michael Dorothy at the DEVCOM Army Research Laboratory and Yue Yu at the University of Minnesota at Austin for their valuable insights during the research and writing process.

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-23-2-0011.

REFERENCES

- [1] David Blackwell. 1949. *Comparison of reconnaissances*. Rand Corporation, Santa Monica, CA. https://www.rand.org/pubs/research_memoranda/RM241.html
- [2] David Blackwell. 1953. Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics* 24, 2 (1953), 265 – 272. <https://doi.org/10.1214/aoms/1177729032>
- [3] Joel Brynielsson and Stefan Arnborg. 2004. Bayesian Games for Threat Prediction and Situation Analysis. In *Proceedings of the Seventh International Conference on Information Fusion (FUSION 2004)*. International Society of Information Fusion (ISIF), Stockholm, Sweden, 1125–1132.
- [4] Asen L Dontchev and R Tyrrell Rockafellar. 2009. *Implicit functions and solution mappings*. Vol. 543. Springer, New York, NY, USA.
- [5] Aidin Ferdowsi, Anibal Sanjab, Walid Saad, and Tamer Basar. 2018. Generalized colonel Blotto game. In *2018 Annual American Control Conference (ACC)*. IEEE, Piscataway, NJ, USA, 5744–5749.
- [6] Zachariah E. Fuchs and Pramod P. Khargonekar. 2012. A sequential Colonel Blotto game with a sensor network. In *2012 American Control Conference (ACC)*. IEEE, Piscataway, NJ, USA, 1851–1857. <https://doi.org/10.1109/ACC.2012.6315589>
- [7] John C. Harsanyi. 1968. Games with Incomplete Information Played by “Bayesian” Players, I–III. Part III. The Basic Probability Distribution of the Game. *Management Science* 14, 7 (1968), 486–502.
- [8] Joao P Hespanha, Yusuf S Ateskan, H Kizilcok, et al. 2000. Deception in non-cooperative games with partial information. In *Proceedings of the 2nd DARPA-JFACC Symposium on Advances in Enterprise Control*. Citeseer, DARPA, Arlington, VA, 1–9.
- [9] Ian Hogeboom-Burr and Serdar Yuksel. 2021. Comparison of information structures for zero-sum games and a partial converse to blackwell ordering in standard

- borel spaces. *SIAM Journal on Control and Optimization* 59, 3 (2021), 1781–1803.
- [10] Ronald A. Howard. 1966. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics* 2, 1 (Aug. 1966), 22–26. <https://doi.org/10.1109/TSSC.1966.300074>
- [11] Kyle Hunt and Jun Zhuang. 2024. A review of attacker-defender games: Current state and paths forward. *European Journal of Operational Research* 313, 2 (2024), 401–417. <https://doi.org/10.1016/j.ejor.2023.04.009>
- [12] Lance M Kaplan. 2006. Global node selection for localization in a distributed sensor network. *IEEE Transactions on Aerospace and Electronic systems* 42, 1 (2006), 113–135.
- [13] Ehud Lehrer and Dinah Rosenberg. 2006. What restrictions do Bayesian games impose on the value of information? *Journal of Mathematical Economics* 42, 3 (June 2006), 343–357. <https://doi.org/10.1016/j.jmateco.2005.09.002>
- [14] Ehud Lehrer, Dinah Rosenberg, and Eran Shmaya. 2010. Signaling and mediation in games with common interests. *Games and Economic Behavior* 68, 2 (2010), 670–682.
- [15] Xinjie Liu, Lasse Peters, and Javier Alonso-Mora. 2023. Learning to play trajectory games against opponents with unknown objectives. *IEEE Robotics and Automation Letters* 8, 7 (2023), 4139–4146.
- [16] Yu Liu, Cristina Comaniciu, and Hong Man. 2006. A Bayesian game approach for intrusion detection in wireless ad hoc networks. In *Proceeding from the 2006 workshop on Game theory for communications and networks (GameNets '06)*. Association for Computing Machinery, New York, NY, USA, 4–es. <https://doi.org/10.1145/1190195.1190198>
- [17] Lee T. Maccarone and Daniel G. Cole. 2022. Bayesian games for the cybersecurity of nuclear power plants. *International Journal of Critical Infrastructure Protection* 37 (July 2022), 100493. <https://doi.org/10.1016/j.ijcip.2021.100493>
- [18] Marcin Peški. 2008. Comparison of information structures in zero-sum games. *Games and Economic Behavior* 62, 2 (2008), 732–735.
- [19] Howard Raiffa and Robert Schlaifer. 1968. *Applied Statistical Decision Theory*. MIT Press, Cambridge, MA.
- [20] Brian Roberson. 2006. The Colonel Blotto game. *Economic Theory* 29, 1 (Sept. 2006), 1–24. <https://doi.org/10.1007/s00199-005-0071-5>
- [21] Violetta Rostobaya, Yue Guan, James Berneburg, Michael Dorothy, and Daigo Shishika. 2023. Deception by Motion: The Eater and the Mover Game. *IEEE Control Systems Letters* 7 (2023), 3157–3162. <https://doi.org/10.1109/LCSYS.2023.3291385>
- [22] Jie Xu and Jun Zhuang. 2016. Modeling costly learning and counter-learning in a defender-attacker game with private defender information. *Annals of Operations Research* 236, 1 (2016), 271–289.