Metacognitive Radar: Masking Cognition From an Inverse Reinforcement Learner

KUNAL PATTANAYAK, Student Member, IEEE
VIKRAM KRISHNAMURTHY, Fellow, IEEE
Cornell University, Ithaca, NY USA

CHRISTOPHER M. BERRY, Student Member, IEEE
Lockheed Martin Advanced Technology Laboratories, Cherry Hill, NJ USA

A metacognitive radar switches between two modes of cognition—one mode to achieve a high-quality estimate of targets, and the other mode to hide its utility function (plan). To achieve high-quality estimates of targets, a cognitive radar performs a constrained utility maximization to adapt its sensing mode in response to a changing target environment. If an adversary can estimate the utility function of a cognitive radar, it can determine the radar's sensing strategy and mitigate the radar performance via electronic countermeasures (ECM). This article discusses a metacognitive radar that switches between two modes of cognition: achieving satisfactory estimates of a target while hiding its strategy from an adversary that detects cognition. The radar does so by transmitting purposefully designed suboptimal responses to spoof the adversary's Neyman–Pearson detector. We provide theoretical guarantees by ensuring that the Type-I error probability of the adversary's detector exceeds a predefined level for a specified tolerance on the radar's performance loss. We illustrate our cognition-masking scheme via numerical examples involving waveform adaptation and beam allocation. We show that small purposeful deviations from the optimal emission confuse the adversary by significant amounts, thereby masking the radar's cognition. Our approach uses ideas from revealed preference in microeconomics and adversarial inverse reinforcement learning. Our proposed algorithms provide a principled approach for system-level electronic counter-countermeasures to hide the radar's strategy from an adversary. We also provide performance bounds for our cognition-masking scheme when the adversary has misspecified measurements of the radar's response.

Manuscript received 13 October 2022; revised 22 March 2023 and 22 June 2023; accepted 27 August 2023. Date of publication 12 September 2023; date of current version 8 December 2023.

DOI. No. 10.1109/TAES.2023.3314406

Refereeing of this contribution was handled by R. Romero.

This work was supported in part by a Research Contract from Lockheed Martin, the Army Research Office under Grant W911NF-21-1-0093, in part by the Air Force Office of Scientific Research under Grant FA9550-22-1-0016, and in part by National Science Foundation under Grant CCF-2312198.

Authors' addresses: Kunal Pattanayak and Vikram Krishnamurthy are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA, E-mail: (kp487@cornell.edu; vikramk@cornell.edu); Christopher M. Berry is with Lockheed Martin Advanced Technology Laboratories, Cherry Hill, NJ 08002 USA, E-mail: (christopher.m.berry@lmco.com). *(Corresponding author: Kunal Pattanayak.)*

This article has supplementary downloadable material available at https://doi.org/10.1109/TAES.2023.3314406, provided by the authors.

GLOSSARY OF SYMBOLS

Abbreviations

IRL	Inverse reinforcement learning.
I-IRL	Inverse–inverse reinforcement learning.

IRL for Identifying Radar Cognition (see Section II)

$k = 1, 2, \ldots, K$	Time index.
$\alpha_k \in \mathbb{R}_+^d$	Target probe.
$\beta_k \in \mathbb{R}_+^d$	Radar action.
x_k	Target state.
$y_k \sim p_{\beta_k}(y\|x_k)$	Radar observation.
$\pi_k = \mathcal{T}(\pi_{k-1}, y_k)$	Radar tracker.
$\hat{\beta}_k = \beta_k + \omega_k$	Observed radar action.
$\omega_k \sim f_\omega$	Measurement noise.
R	Radar tracker obs. noise covariance.
Q	Radar tracker state noise covariance.
C	Radar sensor gain.
Σ	Radar tracker covariance.
$\mathbf{u}$	Radar utility function.
$\mathbf{g}(\cdot) \leq 0$	Radar resource constraint.
$\mathcal{D}$	Adversary's IRL dataset.
$\mathcal{D} \equiv \mathcal{D}_g$	(a) When constraint is known.
$\mathcal{D} \equiv \mathcal{D}_u$	(b) When utility is known.
$\mathcal{A}(\cdot, \mathcal{D}_{u/g}) \leq 0$	Adversary's IRL feasibility test.
θ	Variable for IRL feasibility test.
u_{IRL}	Reconstructed utility function.
$g_{\text{IRL}} - \gamma \leq 0$	Reconstructed resource constraint.

Masking Radar Cognition (see Section III)

	Margin of IRL feasibility test.
$\mathcal{M}_u(\mathcal{D}_g)$	(a) When constraint is known.
$\mathcal{M}_g(\mathcal{D}_u)$	(b) When utility is known.
η	Extent of cognition masking.
$\{\beta_k^*\}_{k=1}^K$	Radar's naive utility-masking response.
$\{\tilde{\beta}_k^*\}_{k=1}^K$	Radar's cognition-masking response.

Masking Radar Cognition in Noise (see Section IV)

	IRL detector for noisy radar responses.
$\phi^*(\widehat{\mathcal{D}})$	Statistical test $\lessgtr_{H_1}^{H_0} h(\gamma)$.
$\widehat{\mathcal{D}}$	Adversary's noisy IRL dataset.
$\phi^*(\widehat{\mathcal{D}})$	Test statistic.
γ	Significance level.
λ	Extent of IRL detector mitigation.

I. INTRODUCTION

In abstract terms, a cognitive radar is a constrained utility maximizer with multiple sets of utility functions and constraints that allow the radar to deploy different strategies depending on changing environments. Cognitive radars adapt their waveform scheduling and beam allocation by optimizing their utility functions in different situations. If a smart adversary can estimate the utility function or constraints of the cognitive radar, then it can exploit this information to mitigate the radar's performance (e.g., jam the radar with purposefully designed interference). A natural question is: *How can a cognitive radar hide its cognition from an adversary?* Put simply, how can a smart sensor hide its strategy by acting dumb? We term this cognition-masking functionality as metacognition.[1] A metacognitive radar [1] switches between two modes of cognition; one mode to achieve a high-quality estimate of a target, and the other mode to hide its utility function (plan).

A metacognitive radar pays a penalty for stealth—it deliberately transmits suboptimal responses to keep its strategy hidden from the adversary resulting in performance degradation. This article investigates how a cognitive radar hides its strategy when the adversary observes the radar's responses. Our metacognition results are inspired by privacy-preserving mechanisms in differential privacy and adversarial obfuscation in deep learning with related works discussed in the following text. Although this article is radar-centric, we emphasize that the problem formulation and algorithms also apply to adversarial inverse reinforcement learning (IRL) in general machine learning applications, namely, how to purposefully choose suboptimal actions to hide a strategy.

Related Works

Cognitive radars are widely studied in the literature [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34];[2] see [2], [3], [4], and [22] for comprehensive discussions on the cognitive radar literature. More recently, our articles [35] and [36] deal with IRL algorithms for cognitive radars, namely, how can an adversary estimate the utility function of a cognitive radar by observing its decisions. Reconstructing a decision maker's utility function by observing its actions is the main focus of IRL [37], [38], [39] in machine learning and revealed preference [40], [41] in the microeconomics literature. In the radar literature, such IRL-based adversarial actions to mitigate the radar's operations are called electronic countermeasures (ECM) [35], [42], [43]. This article builds on [35], [36], [44] and develops electronic counter-countermeasures (ECCM) [45], [46], [47] to mitigate ECM. This article assumes that the adversary's ECM is unaware if the radar has ECCM capability, which is consistent with state-of-the-art ECCM literature. The central theme of this article is to apply results from revealed preference in microeconomics theory [40], [48]. To the best of our knowledge, this approach for ECCM to hide cognition is not explored in the literature.

Several works in the literature [49], [50], [51] highlight how an adversary benefits from learning the radar's utility function. In [49], the adversary optimizes its probes to increase the power of its statistical hypothesis test for utility maximization. The authors in [50] and [51] show how revealed preference-based IRL techniques can be used to manipulate consumer behavior.

In the radar context, Sakuma et al. [52] use the Laplacian mechanism for metacognition; the cognitive radar anonymizes its trajectories via additive Laplacian noise. Differential privacy-based adversarial obfuscation has seen success in applications, such as ML [53], user data sharing [54], and recommendation systems [55]. In our cognition-masking approach, the radar mitigates adversarial IRL via purposeful perturbations from an optimal strategy, where the perturbations are computed via stochastic gradient algorithms (see Algorithm 2 in Section IV-B).

Outline and Organization of Results

1) *Background—IRL:* In Section II, we formulate the interaction between a cognitive radar and an adversary target. We first discuss several cognitive radar models studied in optimal waveform design and sensor management in Section II-B. We then review the main idea of revealed preference-based adversarial IRL algorithms, namely, Theorems 1 and 5 in Section II-C, that the adversary uses to reconstruct the radar's strategy from its actions. Then, we outline two examples of cognitive radar functionalities, namely, waveform adaptation and beam allocation. Theorem 6 stated that in Appendix F, the supplementary document extends adversarial IRL to the case where the cognitive radar faces multiple constraints. Theorem 6 is omitted from the main text for readability.
2) *Masking Radar's Strategy From Adversarial IRL:* Section III contains our main metacognition results, namely, Theorem 2 for mitigating adversarial IRL by masking the radar's strategy. The key idea is for the radar to deliberately deviate from its optimal (naive) response to ensure the following.
   a) Its true strategy almost fails to rationalize its perturbed responses (masked from adversarial IRL).
   b) Its performance degradation due to suboptimal responses does not exceed a particular threshold. Theorem 7 in Appendix F extends Theorem 2 to the case where the cognitive radar has multiple constraints. Theorem 8 provides performance bounds on the cognition-masking scheme of Theorem 2 when

[1] "Metacognition" [1] is used to describe a sensing platform that switches between multiple objectives (constrained utility functions).

[2] We discuss cognitive radars in more detail in Section II-B and contextualize the conventional models of radar cognition to the abstract constrained utility maximization framework assumed throughout this article.

the adversary has misspecified measurements of the radar's response.

3) *Masking Radar's Strategy From Adversarial IRL Detectors in Noise:* Section IV extends our IRL and cognition-masking results to the case where the adversary has noisy measurements of the radar's response. First, we define IRL detectors (Definition 4) that *detect* radar's cognition in noise. Then, we enhance our cognition-masking scheme of Theorem 2 to mitigate the IRL detectors. The radar's cognition-masking objective is now used to maximize the detectors' conditional Type-I error probability, subject to a bound on its deliberate performance degradation.
4) *Numerical Illustration of Masking Cognition by Metacognitive Radars:* Section V illustrates our metacognition results on two target tracking functionalities, namely, waveform adaptation and beam allocation. Our numerical experiments show that the metacognition algorithms in this article can effectively mask both the radar's utility function and resource constraint when the cognitive radar is probed by the adversarial target. Our main finding is that a small deliberate performance loss of the metacognitive radar suffices to mask the radar's strategy from the adversary to a large extent. For conciseness, we include the appendix in an online document separate from the main text as supplementary material.

Running Example: Since the concept of ECCM via cognition masking is somewhat abstract, for the reader's convenience, we relate each assumption, definition, and theorem introduced in this article at an implementation level to a real-world cognitive radar example. Specifically, we consider a cognitive radar [20] tracking an adversarial target.

II. BACKGROUND: IRL TO ESTIMATE COGNITIVE RADAR

Since this article investigates how to construct a cognitive radar that hides its utility from an adversarial IRL system, this section gives the background on how an adversarial system can use IRL to estimate the radar's utility. An important aspect of the IRL framework below is that it is a necessary and sufficient condition for identifying cognition (utility maximization behavior); hence, it can be considered an optimal IRL scheme. Appendixes H and G discuss cognition masking when the adversary performs suboptimal IRL.

A. Radar–Adversary Dynamics

MODEL 1 (RADAR–TARGET INTERACTION) The cognitive radar–adversary interaction has the following dynamics:

$$\begin{aligned} &\textit{target probe}: \alpha_k \in \mathbb{R}_+^d \\ &\textit{radar action}: \beta_k \in \mathbb{R}_+^d \\ &\textit{target state}: x_k = \{x_k(t),\ t = 1, 2, \ldots\}, \\ &\qquad x_k(t+1) \sim p_{\alpha_k}(x|x_k(t)),\ x_0 \sim \pi_0 \\ &\textit{radar observation}: y_k \sim p_{\beta_k}(y|x_k) \\ &\textit{radar tracker}: \pi_k = \mathcal{T}(\pi_{k-1}, y_k) \\ &\textit{observed radar action}: \hat{\beta}_k = \beta_k + \omega_k,\ \omega_k \sim f_\omega \end{aligned} \tag{1}$$

REMARKS We now give examples for the abstract model (1).

1) A widely used example [56], [57] for the radar–adversary dynamics model (1) is that of linear Gaussian dynamics for target kinematics and linear Gaussian measurements

$$\begin{aligned} x_k(t+1) &= Ax_k(t) + w_t(\alpha_k),\ x_k(0) \sim \pi_0 = \mathcal{N}(\hat{x}_0, \Sigma_0) \\ y_k(t) &= Cx_k(t) + v_t(\beta_k), \quad k = 1, 2, \ldots, K. \end{aligned} \tag{2}$$

Here, $x_k(t) \in \mathcal{X} = \mathbb{R}^X$ and $y_k(t) \in \mathcal{Y} = \mathbb{R}^Y$. A is a block diagonal matrix [58] when the target state represents its position and velocity in Euclidean space. The variables $w_t \sim \mathcal{N}(0, Q(\alpha_k))$ and $v_t \sim \mathcal{N}(0, R(\beta_k))$ are mutually independent Gaussian noise processes.
2) In this article, we are only concerned with the asymptotic statistics of the radar tracker $\mathcal{T}$ (1) for our cognition-masking algorithms. One example is that of a Bayesian tracker (Kalman filter) where the asymptotic covariance of the state estimate is the unique positive semidefinite solution of the algebraic Riccati equation (ARE). Other tracker examples include the particle filter, interacting multiple-model filter, etc.

We now proceed to define a cognitive radar, which we assume in this article to be a constrained utility maximizer.

DEFINITION 1 (COGNITIVE RADAR) Consider the radar–adversary interaction dynamics of Model 1. The cognitive radar chooses its response β_k^* (1) at time k by maximizing a utility function $\mathbf{u}(\alpha_k, \cdot)$ subject to constraint $\mathbf{g}(\alpha_k, \cdot) \leq 0$

$$\begin{aligned} \beta_k^* \in \operatorname{argmax}\ & \mathbf{u}(\alpha_k, \beta) \\ & \mathbf{g}(\alpha_k, \beta) \leq 0. \end{aligned} \tag{3}$$

We assume that $\mathbf{g}(\cdot)$ is an increasing function of β.

From a radar practitioner's perspective, let us briefly relate the parameters in Definition 1 to a cognitive radar–adversary interaction. Consider a cognitive radar as modeled in [20, Sec. 4B] tracking an adversarial target. The response β parameterizes the radar's transmitted waveform, and the probe α_k parameterizes the state noise covariance matrix due to the adversary's maneuvers. In the cognitive radar context of [20], the utility function $u(\cdot)$ is equivalent to the inverse of the transmitted signal power (radar minimizes its transmission power); the constraints $g(\alpha_k, \cdot) \leq 0$

can be interpreted as posterior Crámer–Rao bound (PCRB) constraints on the radar's estimate of the target's state[3]

REMARKS

1) In the main text of this article, we consider a single constraint. This is consistent with most works in cognitive radar literature, which also assume a single operating constraint. For example, in [59], the cognitive radar is constrained by a bound on the target dwell time (monotone in the time the radar spends tracking each target). In [22], the radar's constraint is a bound on the receiver sensor processing cost (monotone in the radar's choice of sensor accuracy for target tracking). Hence, we only consider the operating cost of the radar in the main text, which is reflected in the radar's scalar-valued constraint $\mathbf{g}$ in (3).
2) *Multiple Resource Constraints:* Our IRL methodology discussed in the following text can be extended to multiple resource constraints ($\mathbf{g}$ is vector valued). However, for readability, we only consider a scalar-valued constraint $\mathbf{g}$ in the main text of this article. We consider multiple resource constraints in Appendix F. The notation for IRL and cognition-masking results is complicated for vector-valued cost $\mathbf{g}(\cdot)$ and, hence, omitted from the main text and discussed in the supplementary document.

B. Radar Cognition as Constrained Utility Maximization

Cognitive radars have been studied extensively in the literature [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34]. In this section, we discuss relevant works from the cognitive radar literature and contextualize widely used models of radar cognition to the abstract constrained utility maximization framework proposed in Definition 1.

Cognitive Radars: The term "cognition" in cognitive radars is used to describe a number of functionalities, such as optimal waveform design, knowledge-aided radar detection and tracking for minimizing response times, and sensor management. A cognitive radar [22], [59], [60] uses the perception–action cycle of cognition to sense the environment and learn from it relevant information about the target and the environment. Cognitive radars have also been modeled as reinforcement learners in the literature that maximize their utility [35], [36], [61], [62], [63] and tune their sensing resources to optimally satisfy mission objectives.[4]

Table I displays works in the cognitive radar literature related to the constrained utility maximizer framework of Definition 1. For brevity, we limit our discussion of cognitive radars to waveform design, sensor management, and joint waveform–receiver filter design.

1) *Radar cognition for optimal waveform design:* The signal-to-interference noise ratio (SINR) is a widely used objective maximized by cognitive radars for waveform adaptation [5], [6], [7], [8], [9], [10]. In [5] and [6], the radar is constrained by the maximum peak-to-average ratio (PAR) of the transmission code that controls the variation of the code about its mean value, and hence, controls the transmission bandwidth. In [7], [8], and [9], the radar is constrained by the total contiguous bandwidth available for transmission, and the resulting optimization problem results in the well-studied sense–react–notch paradigm. The cognitive radar in [10] faces multiple constraints, namely, bounds on the total transmission power, Hamming/Manhattan distance with respect to a reference code, and the interference power spilled over in undesirable frequency bands. We extend our cognition-masking result of Theorem 2 to vector-valued constraints in Theorem 7 in the supplementary document.

The cognitive radar discussed in [11] minimizes a convex combination of two metrics, namely, the spectral-integrated level ratio (SILR), a variable that is inversely proportional to the SINR, and the integrated cross-correlation level (ICCL) that measures the cross correlation of the transmitted waveforms across multiple antennas. The transmitted waveform is constrained to be either constant modulus or discrete phase (equivalent to M-ary phase-shift keying with a prespecified alphabet size). In [12], the radar minimizes the $\mathcal{L}_2$-norm between the ambiguity function of the transmitted waveform and that of a reference waveform constrained by the total transmission power. The waveform design scheme in [16] resembles that of [12] in which the cognitive radar minimizes a convex combination of the interference power and the side lobe correlation, subject to a bound on the transmission power. In [13], [14], and [15], the radar maximizes the M.I. (based on differential entropy) between the received signal and the impulse response of the target subject to a bound on the transmission power. In [21], the cognitive radar minimizes the posterior Crámer–Rao lower bound (CRLB) of the target estimate subject to a bound on the transmission power. The CRLB for the target estimate is also widely used in cognitive radars performing optimal sensor management as discussed in the following text.

2) *Radar cognition for optimal sensor management:* Analogous to optimal waveform design, SINR is also a widely used objective for optimal sensor management in cognitive radars [17], [18], where the radar is constrained by sensing constraints, such as the cost of changing the tracked cell in Euclidean space [17] and bound on downlink interference power [19]. The posterior and predicted CRLBs for the target estimate are also widely used optimization metrics for cognitive radar performing optimal sensor deployment [21], [22], where the radar faces constraints, such as bounds on the communication cost with the central processing unit [21] and bounds on the sensing and processing cost [22]. A similar model is proposed in [23], where the radar optimizes its sensor deployment locations and the number of active

[3]It is straightforward to show that PCRB is inversely proportional to the radar sensor's SNR that depends on the target's maneuvers; hence, $g(\alpha_k)$ can be viewed as SNR constraints with explicit dependence on the adversarial probe α_k.

[4]In the context of Data Fusion Information Group (DFIG) process model [64], sensor adaptation by the radar can be viewed as Level 4-process refinement in the DFIG model.

TABLE I
Cognitive Radars as Constrained Utility Maximizers

Works	Category	Utility	Constraint
[5], [6]	Waveform	Minimum SINR of finitely many users	Bound on Peak-to-Average Ratio (PAR)
[7], [8], [9]	Waveform	SINR (Sense–React–Notch Paradigm)	Bound on contiguous transmission bandwidth
[10]	Waveform	SINR	Bounds on interference power in restricted frequency bands, total transmission power, and lower bound on similarity wrt a reference code
[11]	Waveform	Negative of convex combination of SILR (interference), ICCL (waveform correlation)	Baseband transmission codes are constrained to be either constant modulus or discrete phase
[12]	Waveform	Negative of $\mathcal{L}_2$-deviation from a desired ambiguity function	Bound on transmission power
[13], [14], [15]	Waveform	M.I. the between measured signal and target impulse response	Bound on transmission power
[16]	Waveform	Negative of convex combination of interference power in restricted frequency bands and side-lobe correlation	Bound on transmission power
[17], [18], [19]	Tracked cell	SINR	Bounds on Euclidean distance between current and next tracked cell (cost of shifting target cell), downlink interference power
[20], [21]	Sensor, Waveform	Negative of predicted posterior CRLB for target estimate	Bounds on transmission power, communication cost
[22]	Sensor	Negative of predicted conditional CRLB of target estimate	Bounds on sensing cost, processing cost
[23]	Sensor	Negative of root-mean-squared error between target state and estimate	(i) *For sensor deployment locations:* unconstrained (global optimum achieved in finitely many steps), (ii) *For number of sensors to be deployed:* Bound on deployed sensors
[24], [25], [26]	Beamsteering	Target position entropy (as a function of target cell)	Bound on target tracking entropy
[27], [28]	Sensor	Likelihood of emission on a 2-D grid under Gaussian measurement model	Emission activity norm (block-sparsity constraint)
[29]	Joint waveform–receiver filter	Negative of interference and clutter power at the receiver	Normalization constraint on the received signal power (Capon constraint), Bound on transmission power
[30]	Joint waveform–receiver filter	Negative of interference power at the receiver	Capon constraint, code constraint from [12]
[31], [32], [33]	Joint waveform–receiver filter	SINR, signal power at the receiver	Orthogonality constraint between waveforms, bounds on transmission power, bounds on $\mathcal{L}_2$-distance from a set of reference waveforms
[34]	Joint waveform–receiver filter	SCNR	Bound on transmission power

In the table above, we contextualize several notable works in the cognitive radar literature according to the abstract constrained utility maximization setup of Definition 1. For every cognitive radar model, as discussed in Section II-B, we list the optimization type or "category," the equivalent utility being maximized by the radar and the resource constraint faced by the radar. The metacognition algorithms in this article provide a principled approach to spoof an adversary that can identify the radar's plan and mitigate the radar's operations.

sensors. The radar minimizes the mean squared tracking error subject to constraints on the number of sensors deployed. The authors in [24], [25], and [26] address optimal beamsteering for cognitive radars. To choose the optimal cell for focusing its transmit beam, the radar maximizes the entropy of the target's location. For target tracking, the optimal sensor parameters minimize the target's tracking entropy (based on the location and velocity of the target). Finally, Aubry et al. [27], [28] design cognitive radars for adaptive target detection that maximize the likelihood of target emission on a two-dimensional (2-D) grid, subject to block sparsity constraints on the target location.

3) *Radar cognition for joint waveform–receiver filter optimization:* Joint optimization of waveform and receiver filter design is well explored in the cognitive radar literature; we discuss a few notable works in the following text. Note that the radar optimizes over two variables, namely, the receiver filter and the transmitted waveform. In [29] and [30], the radar minimizes the clutter/interference power at the receiver subject to the well-known Capon [65] constraint, namely, a normalization constraint on the received signal power. In addition, the radar in [29] is subject to an equality (normalization) constraint on the received signal power, and a bound on the transmission power in [29]. The radar in [30] faces an additional waveform constraint (identical to [12]), namely, the transmission waveform is constrained to be either constant modulus or discrete phase. On a related note, robust constrained Capon beamforming is investigated in [66], [67], and [68]. Rossetti and Lambotharan [31] consider a bistatic cognitive radar transmitting two waveforms. The joint waveform–receiver filter optimization is done in two steps: First, the optimum receiver filters are computed that maximize the receiver SINR. Then, the optimal waveforms are computed that maximize the signal power at the radar receiver subject to orthogonality constraints on the two waveforms, transmission power constraints, and bounds on $\mathcal{L}_2$-deviation from a set of reference waveforms. Rossetti and Lambotharan [31] generalize their work to multistatic radars in [32] and to cognitive radar networks in [33]. Finally, Guerci et al. [34] maximize the signal-to-clutter noise ratio (SCNR) at the receiver subject to a bound on the transmission power. The key idea is that the introduction of a physics-based scattering model for the clutter environment makes the maximization of SCNR tractable unlike the traditional approaches.

Metacognitive Radars: A metacognitive radar [2], [69], [70], [71] transcends conventional notions of "cognition" in radars. In this article, we view metacognition

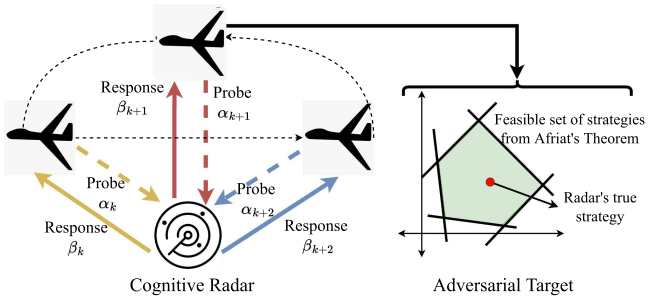


Fig. 1. Schematic of adversarial IRL against cognitive radars. The adversary observes a sequence of decisions of the cognitive radar in response to a sequence of adversarial probes. Revealed preference-based adversarial IRL (Afriat's Theorem) [40], [48] is equivalent to checking the existence of a feasibility polytope for a set of inequalities (Afriat's Theorem [40], [48]). Our aim in this article is to make adversarial IRL cumbersome—how to purposefully distort radar responses metacognition objective in this article is to spoof adversarial IRL, namely, how to make checking linear feasibility difficult.

as the radar's sensing ability to identify an adversary in its environment and strategic ability to spoof the adversary using inverse–inverse reinforcement learning (I-IRL) techniques to "mask" its cognition. The working assumption of the article is that an adversary can identify the cognitive ability of a radar and mitigate the radar's operations based on this information. Recent works address how to identify cognitive radars by analyzing a finite time series of emission exchanges with the radar [35], [36], [44]; a summary of the strategy identification results is presented in Section II-C in the following text.

Radar functionalities that mitigate adversarial systems are termed ECCM in the radar literature; see [72] for a comprehensive discussion. Low-probability-of-intercept (LPI) transmission design [45], [73], [74] achieves stealth for cognitive radars and avoids cognition detection. Waveform adaptation schemes to counter barrage jamming are studied in [45] and [46]. Frequency diversity for stealth-based ECCM in multitarget and moving target tracking applications is studied in [75], [76], and [77].

While the works discussed above mitigate an adversary, the ECCM measures do not necessarily mask the radar's cognition. The metacognitive radar's aim in this article is to *confuse* the adversary's detector and *hide* its cognition, i.e., ensuring the adversary incorrectly reconstructs the radar's strategy with high probability, by deliberately transmitting suboptimal responses. Specifically, this article contributes to antistealth and anti-ARM ECCM [78] by ensuring that adversarial mitigation is ineffective with a large probability.

C. Adversarial IRL for Identifying Strategy of Cognitive Radar

We now review the main results for adversarial IRL, namely, how an adversary can identify and reconstruct the radar's strategy by observing the radar's responses. The adversarial IRL system is schematically shown in Fig. 1. The key idea is to formulate the adversary's task of identifying the radar's strategy as a linear feasibility problem in terms of the radar's responses. This article considers two distinct scenarios in terms of the dependence of the adversary's probe α_k on the radar's utility $\mathbf{u}$ and resource constraint $\mathbf{g}$ in (3). The two scenarios are formalized in Assumptions 1 and 2 below in our IRL results, Theorems 1 and 5, and justified in Section II-D in the tracking examples of waveform adaptation and beam allocation.

IRL for Identifying Radar's Utility Function

In works, such as [11] and [16], the adversary can mitigate the cognitive radar if the adversary knows the utility weights. Theorem 1 below provides a set-valued reconstruction algorithm to estimate the radar's utility function when the adversary controls the radar's resource constraint. Such scenarios where the adversary knows the radar's resource constraint is formalized below in Assumption 1.

ASSUMPTION 1 The radar's resource constraint $\mathbf{g}(\cdot)$ in (3) is linear in the adversary's probe α_k and the radar's utility $\mathbf{u}(\cdot)$ is independent of α_k

$$\mathbf{g}(\alpha_k, \beta) = \alpha_k'\beta - 1, \ \ \mathbf{u}(\alpha_k, \beta) \equiv \mathbf{u}(\beta) \tag{4}$$

IRL objective: The adversary aims to reconstruct the radar's utility $\mathbf{u}(\cdot)$ using the dataset $\mathcal{D}_g$, where $\mathcal{D}_g$ is defined as follows:

$$\mathcal{D}_g = \{\mathbf{g}(\alpha_k, \cdot), \beta_k\}_{k=1}^K \tag{5}$$

where $\mathbf{g}(\alpha_k, \cdot)$ is defined in (4).

In spite of its linear structure, the constraint in (4) can model nonlinear radar constraints via a suitable definition of the radar's response β and the adversary's probe α. For example, an upper bound on the asymptotic precision of the radar's state estimate (inverse of the solution of the ARE) can be expressed as a linear constraint in terms of the eigenvalues of the state and noise covariance matrix; see [35, Lemma 3] for a detailed exposition. Let us now state Theorem 1 for achieving IRL when assumption 1 holds.

THEOREM 1 (IRL FOR IDENTIFYING RADAR'S UTILITY FUNCTION) Consider the cognitive radar described in Model 1. Suppose assumption 1 holds. Then

1) The adversary checks for the existence of a feasible utility function that satisfies (3) by checking the feasibility of a set of linear inequalities

$$\begin{aligned} &There\ exists\ a\ feasible\ \theta \in \mathbb{R}_+^{2K}\ s.t.\ \mathcal{A}(\theta, \mathcal{D}_g) \leq \mathbf{0} \\ \Leftrightarrow\ &\exists\ u\ s.t. \beta_k \in \operatorname{argmax} u(\beta),\ \alpha_k'\beta \leq 1\ \forall k \end{aligned} \tag{6}$$

where dataset $\mathcal{D}_g$ is defined in (5) and the set of inequalities $\mathcal{A} \leq \mathbf{0}$ is defined in Appendix A.

2) If $\mathcal{A}(\cdot, \mathcal{D}_g) \leq \mathbf{0}$ has a feasible solution, the set-valued IRL estimate of the radar's utility $\mathbf{u}$ is given by

$$\begin{aligned} u_{\text{IRL}}(\beta) &\equiv \{u_{\text{IRL}}(\beta; \theta) : \mathcal{A}(\theta, \mathcal{D}_g) \leq \mathbf{0}\} \\ u_{\text{IRL}}(\beta; \theta) &= \min_{k \in \{1,2,\ldots,K\}} \{\theta_k + \theta_{k+K}\ \alpha_k'(\beta - \beta_k)\}. \end{aligned} \tag{7}$$

Theorem 1 is well known in microeconomics as Afriat's theorem [40], [48] and widely used for set-valued estimation of consumer utilities from the offline data. In complete analogy, the adversary also performs IRL on a batch of

probe–response exchanges with the cognitive radar to reconstruct the radar's utility.[5] Abstractly, Theorem 1 says that given a finite dataset, the adversary can at best construct a polytope of feasible strategies that rationalize the adversary's dataset. Theorem 1 achieves IRL when the radar faces a single operating constraint. We discuss adversarial IRL for multiple resource constraints in Theorem 6 in Appendix F. Then, the linear feasibility test of (6) generalizes to a mixed-integer linear feasibility test, linear in the real-valued feasible variables in the multiconstraint case.

The important aspects of Theorem 1 to a practitioner are the following: Unlike typical *reactive* ECM systems, the adversarial target in this article is assumed to be a *cognitive* entity [80]. The cognitive ECM entity has the capability to estimate the radar's strategy encoded in its utility function $\mathbf{u}$, and then perform adversarial maneuvers $(\alpha_{1:K})_{\text{Adv}}$ that minimize the radar's utility

$$(\alpha_{1:K})_{\text{Adv}} \in \underset{\alpha_{1:K}}{\operatorname{argmin}} \sum_{k=1}^{K} \max_{\beta_{1:K}} \mathbf{u}(\beta_k),\ \mathbf{g}(\alpha_k, \beta_k) \leq 0 \quad (8)$$

In the context of the cognitive radar modeled in [20], the utility function could be a Quality-of-Service (QoS) metric [81] the radar maximizes to yield the optimal waveform parameter (instead of simply minimizing the transmission power). The ECM objective in this scenario would be to identify the radar's QoS function for mitigating its operations. Through the reconstructive procedure of (47) in Theorem 1, the adversary can estimate the radar's utility, and then use (8) to design optimal maneuvers that minimize the radar's QoS.[6]

IRL for Identifying Radar's Resource Constraints

In certain scenarios, the utility of the radar is well known [e.g., signal-to-noise ratio (SNR)], but the operational constraints of the radar are not known, for example, bound on the PAR [5], [6]. We formalize such scenarios where the adversary knows the radar's utility function below as Assumption 2:

ASSUMPTION 2 The radar's utility function $\mathbf{u}(\cdot)$ (3) is controlled by the adversary's probe α_k, the radar's resource constraint $\mathbf{g}$ is independent of α_k and has the following form:

$$\mathbf{g}(\alpha_k, \beta) \equiv \mathbf{g}(\beta) - \gamma_k,\ \gamma_k > 0 \quad (9)$$

where γ_k, g are independent of α_k.

IRL objective: The adversary aims to reconstruct $\mathbf{g}(\cdot)$ using the dataset $\mathcal{D}_u$, where $\mathcal{D}_u$ is defined as follows:

$$\mathcal{D}_u = \{\mathbf{u}(\alpha_k, \cdot), \beta_k\}_{k=1}^{K}. \quad (10)$$

IRL for estimating the radar resource constraints has the same structure as that of Theorem 1 and is discussed in the online supplementary document. IRL for Assumption 2 is formally stated in Theorem 5 in Appendix B and summarized as follows:

$$g_{\text{IRL}}(\beta) \equiv \{g_{\text{IRL}}(\beta; \theta) : \mathcal{A}(\theta, \mathcal{D}_u) \geq \mathbf{0}\}$$
$$g_{\text{IRL}}(\beta; \theta) = \max_{k \in \{1,2,\ldots,K\}} \{\theta_k + \theta_{K+k}(\mathbf{u}(\alpha_k, \beta) - \mathbf{u}(\alpha_k, \beta_k))\} \quad (11)$$

where g_{IRL} is the adversary's set-valued estimate of the radar's constraint $\mathbf{g}$, dataset $\mathcal{D}_u$ is defined in (10) and $\theta \in \mathbb{R}_+^{2K}$ is a feasible vector w.r.t. the feasibility test $\mathcal{A}(\cdot, \mathcal{D}_u) \geq \mathbf{0}$. Note how the IRL feasibility inequalities in (11) are identical to that of (6) in Theorem 1 but with the inequality direction reversed.

Theorem 5 is useful when the adversary is interested in evaluating the radar's constraints. Consider the cognitive radar in [20]. The adversary knows the radar's utility, for example, the SNR. The adversary's aim instead is to estimate the radar's constraints on the cost of communication [20, eq. (40)] with the central processing unit. Knowledge of the radar's communication cost facilitates adversarial maneuver selection as follows:

$$(\alpha_{1:K})_{\text{Adv}} \in \underset{\alpha_{1:K}}{\operatorname{argmin}} \sum_{k=1}^{K} \max_{\beta_{1:K}} \mathbf{u}(\alpha_k, \beta_k),\ \mathbf{g}(\beta_k) \leq \gamma_k \quad (12)$$

where the utility function is simply the radar sensor's SNR that indeed depends on the adversary's maneuvers (parametrized by probe α_k), $g(\cdot)$ is the radar's communication cost, and γ_k is the cost threshold at time step k.

D. Examples of IRL for Identifying Radar Cognition

Below, we discuss two examples of cognitive radar functionalities, namely, waveform adaptation and beam allocation. Throughout this article, we will use the two examples below for contextualizing our cognition-masking results.[7]

1) *Example 1—Waveform Adaptation for Cognitive Radar:* Waveform adaptation [84], [85], [86], [87], [88], [89] is a crucial functionality of a cognitive radar. Consider a cognitive radar with linear Gaussian dynamics and measurements (2). The cognitive radar's aim is to choose the optimal sensor mode (observation noise covariance) based on the target's maneuvers. The more accurate sensor results in more precise observations but is also costlier to deploy. Appendix D formalizes the optimal waveform adaptation and abstracts the problem as the constrained utility maximization problem of (3). In simple terms, the cognitive radar maximizes its observation noise covariance (least accurate sensing mode) subject to a lower bound on the radar's SNR. The key idea is to assume that the adversary's probe α_k and radar's response β_k are the eigenvalues of covariance matrices Q and R^{-1}, respectively, and hence, parameterize

[5] Afriat's theorem with linear constraints (4) has been generalized to nonlinear monotone constraints in the literature [79]. For the radar context in this article, it suffices to assume a linear constraint when the adversary is trying to estimate the radar's utility.

[6] Popular framework to study radar–adversary interactions of the form in (8) is the principal agent problem (PAP). We refer the reader to [82] and [83], where Krishnamurthy et al. design ECCM strategies using a PAP framework for adversarial mitigation.

[7] In Appendixes C and D, we formally relate the variables in (13) and (14) to tracker-level parameters of the cognitive radar.

the state and observation noise covariance in the state-space model of (2). Appendix D then shows the equivalence between an upper bound on the radar's asymptotic covariance $(\Sigma^*(\alpha_k, \beta_k))^{-1}$ and the linear constraint $\alpha_k'\beta \leq 1$. In summary, the cognitive radar's optimal waveform adaptation strategy can be abstracted as follows:

$$\beta_k \in \text{argmax } \mathbf{u}(\beta),\ \alpha_k'\beta \leq 1 \tag{13}$$

where u is the radar's utility, and the linear constraint $\alpha_k'\beta \leq 1$ equivalently bounds the *asymptotic precision* of the radar.

Let us briefly discuss the state-of-the-art in waveform design in the radar literature and show how optimal waveform design can be embedded in the abstract constrained utility maximization setup of (13). In [84], the constraint in (13) is a bound on the waveform power; the utility function is either the conditional M.I. between the target impulse response and the reflected waveforms, given the knowledge of transmitted waveform, or simply the negative of the mean squared error between the true and estimated location of the target being tracked, with both choices of utility function yielding the same optimal waveform choice. Liu et al. [86] study waveform design in omnidirectional radars where the radar's utility function (13) is the negative of the downlink multiuser interference and the resource constraint is simply a bound on the transmitted power. In [88], the radar's utility is the negative of the Crámer–Rao bound on the variance of the radar's state estimate; the radar's resource constraint is a bound on its transmission power. Wei et al. [87] design optimal waveforms with an added ECCM functionality to mitigate ECM. The key idea is to first send a pilot waveform to estimate the parameters of the adversary's ECM, followed by intrapulse frequency coding with appropriate parameters to deceive the adversary's ECM. Our ECCM approach is similar to that of [87] with the only difference that, instead of increasing the bandwidth of our transmitted signal to combat smart noise jamming, the cognitive radar transmits suboptimal waveforms to avoid its strategy from being reconstructed by the adversary.

IRL for optimal waveform adaptation: The adversary's aim is to identify the radar's utility function u. Also, the setup of (13) falls under Assumption 1. Hence, the adversary uses the IRL test of (6) in Theorem 1 for identifying u.

2) *Example 2—Beam Allocation for Cognitive Radar:* Appendix C discusses optimal beam allocation [90], [91], [92], [93], [94]. The cognitive radar's aim is to allocate its beam intensity optimally between multiple targets. Compared to a target with less jerky maneuvers, a target with unpredictable maneuvers requires a more focused beam for the SNR to lie above a certain threshold. Appendix C formalizes the beam allocation problem and abstracts the problem as a constrained utility maximization problem (3). The key idea is to relate the adversary's probe α_k to the asymptotic *predicted* precision of the radar tracker. In summary, the cognitive radar's optimal waveform adaptation problem can be abstracted as follows:

$$\beta_k \in \text{argmax } \mathbf{u}(\alpha_k, \beta) \equiv \prod_{i=1}^{m} \beta(i)^{\alpha_k(i)},\ \|\beta\|_\kappa \leq \gamma_k \tag{14}$$

where the radar maximizes a Cobb–Douglas utility subject to a bound γ_k on the total transmit beam intensity (κ-norm of intensity vector) for all k.

IRL for optimal beam allocation: Since the adversary knows the radar's utility (Assumption 2), its aim is to identify the radar's constraint $\mathbf{g}(\cdot) - \gamma_k \leq 0$ using the IRL test (50) in Theorem 5.

Summary: This section discussed how an adversary can deploy IRL to estimate a cognitive radar's utility and constraint. While IRL with a single operational constraint is discussed in [35], the IRL algorithm for multiple constraints (in Appendix F) is new. This section also related Theorem 1 for identifying radar cognition to the parameters of a cognitive radar [20].

III. I-IRL: MASKING RADAR UTILITY AND CONSTRAINTS FROM ADVERSARIAL IRL

Having discussed how an IRL system can detect a cognitive radar, we are now ready to design a cognitive radar that is aware of the adversary's IRL motives and hides its strategy (utility function and resource constraints) from the IRL system. In radar terminology, IRL for mitigating a radar system falls under the field of ECM. Since metacognition deals with spoofing adversarial IRL, it can be viewed as a form of ECCM against ECM, see schematic outlined in Fig. 2.

Rationale: How to hide cognition? Recall that the feasibility of (6) and (50) is both necessary and sufficient for identifying utility maximization behavior (3); see [40] and [48] for the proof. Hence, a cognitive radar's true strategy lies within the polytope of feasible strategies computed by the adversary (see Fig. 1). The cognition-masking rationale in this article is to transmit purposefully the designed perturbed responses that ensure that the radar's true strategy lies close to the edge of the polytope of feasible strategies. The distance from the edge of the feasibility polytope is a measure of goodness-of-fit of the strategy to the radar's responses; see Definitions 2 and 3 in the following text. In other words, the radar deliberately sacrifices performance to ensure that its strategy poorly rationalizes its perturbed responses, hence hiding its strategy from adversarial IRL.

Main Result: How a radar can mask its utility/constraints

Theorem 2 below is our main result for cognition masking. Theorem 2 uses the concept of feasibility margin—how far is a strategy from failing the IRL feasibility tests (6) or (50). We define two margins—$\mathcal{M}_u$ and $\mathcal{M}_g$—for the feasibility margins of feasible utilities u and constraints g, respectively.

DEFINITION 2 (FEASIBILITY MARGIN FOR RECONSTRUCTED UTILITY (6)) Consider the dataset $\mathcal{D}_g$ defined in (5). The feasibility margin $\mathcal{M}_u(\mathcal{D}_g)$ defined below measures how

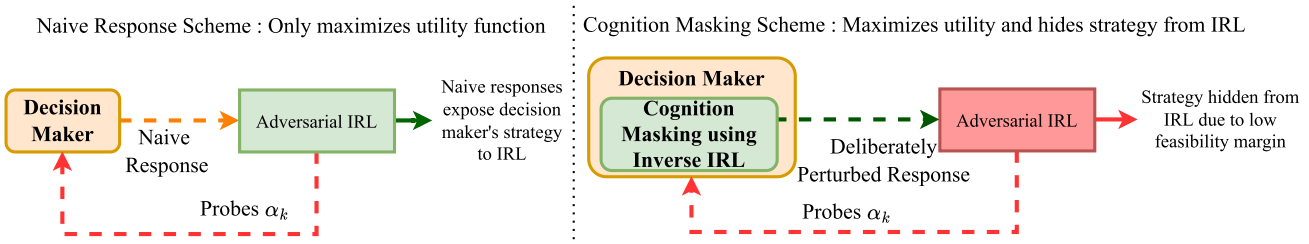


Fig. 2. Schematic of the cognitive radar masking its strategy from adversarial IRL (via Theorem 2).
Naive response scheme (Left): The adversary sends a sequence of probe signals to the radar and records its responses to the adversary's probes. The radar's strategy passes the IRL feasibility test of Theorem 1 with a large margin if the radar transmits naive responses (3) and can be reconstructed by IRL.
Cognition-masking scheme (Right): If the radar is aware of adversarial IRL, the radar deliberately perturbs its responses according to Theorem 2 to hide its strategy from the adversary at the cost of performance degradation. In Section V, we illustrate via numerical examples how small deliberate perturbations in the radar's naive responses mask the radar's strategy from adversarial IRL to a large extent.

far is the utility u from failing the IRL feasibility test (6)

$$\mathcal{M}_u(\mathcal{D}_g) = \min_{\epsilon \geq 0} \epsilon, \ \mathcal{A}(u, \mathcal{D}_g) + \epsilon \mathbf{1} \geq \mathbf{0} \tag{15}$$

where $\mathbf{1}$ is the column vector of all ones.

DEFINITION 3 (FEASIBILITY MARGIN FOR RECONSTRUCTED CONSTRAINTS (50)) Consider the dataset $\mathcal{D}_u$ defined in (10). The feasibility margin $\mathcal{M}_g(\mathcal{D}_u)$ defined below measures how far the constraint g is from failing the IRL feasibility test (50)

$$\mathcal{M}_g(\mathcal{D}_u) = \min_{\epsilon \geq 0} \epsilon, \ \mathcal{A}(g, \mathcal{D}_u) - \epsilon \mathbf{1} \leq \mathbf{0} \tag{16}$$

where $\mathbf{1}$ is the column vector of all ones.

The margins (15) and (16) are measure of goodness-of-fit for the IRL feasibility inequalities (6) and (50), respectively, for any feasible strategy.[8] If u is a feasible utility that rationalizes $\mathcal{D}_g$ (5), we have $\mathcal{A}(u, \mathcal{D}_g) \leq \mathbf{0}$ from (6). Hence, the margin for u is the minimum *nonnegative* perturbation so that the IRL test of (6) fails, that is, $\mathcal{A}(\cdot, \mathcal{D}_g) + \epsilon \mathbf{1} \geq \mathbf{0}$. Similarly, if g is a feasible resource constraint that rationalizes $\mathcal{D}_u$ (10), we have $\mathcal{A}(u, \mathcal{D}_g) \geq \mathbf{0}$ from (50). Hence, the margin for u is the minimum *nonpositive* perturbation so that the IRL test of (6) fails, that is, $\mathcal{A}(\cdot, \mathcal{D}_g) - \epsilon \mathbf{1} \geq \mathbf{0}$. Equivalently, the margin measures how far a strategy lies from the edge of the polytope of feasible strategies.[9] The concept of margins arises in many prominent areas of machine learning, for example, in support vector machines [99] for classification tasks and also max-margin IRL [100]. In the radar context, a strategy with a large feasible margin is a high-confidence point estimate of the radar's strategy and, hence, at higher risk of getting exposed.

We are now ready to state our first cognition-masking result, Theorem 2. Theorem 2 ensures that the radar's true strategy has a low feasibility margin w.r.t. the IRL tests of Theorems 1 and 5 by deliberately perturbing the radar's naive responses (3). In a sense, the radar optimally switches between maximizing its performance and maximizing the privacy of its plan.

THEOREM 2 (MASKING COGNITION FROM ADVERSARIAL IRL FEASIBILITY TESTS.) Consider the cognitive radar (3) in Definition 1. Let $\{\beta_k^*\}_{k=1}^K$ denote the naive response sequence (3) that maximizes the cognitive radar's utility. Then:

1) *Masking Utility Function From IRL:* Suppose Assumption 1 holds. The response sequence $\{\tilde{\beta}_{1:K}^*\}$ defined below masks the radar's utility $\mathbf{u}$ from the adversary by ensuring that $\mathbf{u}$ passes the IRL feasibility test (6) with a sufficiently low margin (15) parametrized by $\eta \in [0, 1]$

$$\{\tilde{\beta}_{1:K}^*\} = \operatorname*{argmin}_{\{\beta_k \geq \mathbf{0}, \ \alpha_k' \beta_k \leq 1\}} \sum_{k=1}^K \mathbf{u}(\beta_k^*) - \mathbf{u}(\beta_k) \tag{17}$$

$$\mathcal{M}_{\mathbf{u}}(\mathcal{D}_g) \leq (1 - \eta) \ \mathcal{M}_{\mathbf{u}}(\mathcal{D}_g^*) \tag{18}$$

where dataset $\mathcal{D}_g^* = \{\alpha_k'(\cdot) - 1, \beta_k^*\}_{k=1}^K$ is the adversary's dataset when the radar transmits naive responses $\{\beta_k^*\}_{k=1}^K$, and $\mathcal{D}_g$ is defined in (5).

2) *Masking Resource Constraint From IRL:* Suppose Assumption 2 holds. The response sequence $\{\tilde{\beta}_{1:K}^*\}$ defined below masks the radar's resource constraint $\mathbf{g}$ from the adversary by ensuring that $\mathbf{g}$ passes the IRL feasibility test (50) with a sufficiently low margin (16) parametrized by $\eta \in [0, 1]$

$$\{\tilde{\beta}_{1:K}^*\} = \operatorname*{argmin}_{\{\beta_k \geq \mathbf{0}, \ g(\beta_k) \leq \gamma_k\}} \sum_{k=1}^K \mathbf{u}(\beta_k^*) - \mathbf{u}(\beta_k) \tag{19}$$

$$\mathcal{M}_{\mathbf{g}}(\mathcal{D}_u) \leq (1 - \eta) \, \mathcal{M}_{\mathbf{g}}(\mathcal{D}_u^*) \tag{20}$$

[8] Strictly speaking, the margin (15) is the minimum perturbation so that $\mathcal{A}(u_{\mathcal{A}}, \mathcal{D}_u)$ is infeasible, where $u_{\mathcal{A}}$ is the finite-dimensional projection of u for the IRL feasibility test defined in (48) in Appendix A. However, we abuse notation and express the feasibility test as $\mathcal{A}(u, \mathcal{D}_u)$ for the sake of simplicity of exposition. We abuse notation in a similar way for (16).

[9] There exist several robustness measures in the literature [95], [96], [96], [97], [98] that check how well a *dataset* satisfies economic-based rationality. Our cognition-masking aim is more subtle—our aim is to ensure that a particular *strategy* rationalizes a dataset poorly by minimizing its feasibility margin (15), (16).

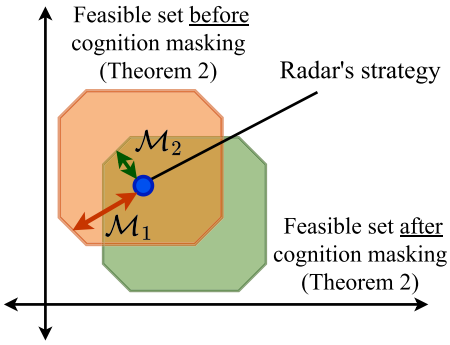


Fig. 3. Cognition masking for mitigating adversarial IRL. The radar's naive responses pass the IRL feasibility tests in Theorems 1 and 5 with a large feasibility margin $\mathcal{M}_1$. Cognition masking distorts the feasibility polytope so that the radar's true strategy is almost infeasible (low margin $\mathcal{M}_2$) w.r.t. the IRL feasibility inequalities (close to the edge of feasibility polytope). Hence, the true strategy is a low-confidence estimate for IRL and successfully hidden from the adversary.

where dataset $\mathcal{D}_u^* = \{\mathbf{u}(\alpha_k, \cdot), \beta_k^*\}_{k=1}^K$ is the adversary's dataset when the radar transmits naive responses $\{\beta_k^*\}_{k=1}^K$, and $\mathcal{D}_u$ is defined in (10).

Theorem 2 is our first result for masking cognition; see Algorithm 1 for a stepwise procedure for masking the radar's utility (17). This is schematically illustrated in Fig. 3. Theorem 2 computes the *optimal* suboptimal response of the radar that sufficiently mitigates adversarial IRL. The radar minimizes its performance degradation (maximizes *QoS*) due to suboptimal responses subject to a bound (18) and (20) on the feasibility margin of the radar's strategy (maximizes *adversarial confusion*). Theorem 2 can be viewed as an I-IRL scheme that mitigates an IRL system and is a critical feature of a *metacognitive* radar that switches between different plans. For completeness, Appendix F extends cognition masking to the case where the cognitive radar faces multiple constraints. Theorem 7 generalizes the cognition-masking scheme of Theorem 2 to the multiconstraint case where the adversary uses Theorem 6 for optimal IRL. Also, Appendix G discusses cognition masking when the adversary has *misspecified* measurements of the radar's responses. Our key result is Theorem 8 that provides a performance bound on the cognition-masking scheme of Theorem 2 in terms of the misspecification error magnitude.

Extent of cognition masking η in Theorem 2: A smaller value of η implies a larger extent of cognition masking from adversarial IRL and also a greater degradation in the radar's performance. One extreme case is setting $\eta = 0$. This results in maximal masking of the radar's strategy. That is, the IRL feasibility inequalities (6) and (50) are no more feasible and there exists *no feasible strategy* that rationalizes the radar's responses. Setting $\eta = 0$ also causes the radar to deviate maximally from its naive responses (3), and hence results in a large performance degradation. The other extreme case is setting $\eta = 1$. In this case, the radar simply transmits its naive response (3) and its strategy is not hidden from the adversary.

Algorithm 1: Masking Radar's Utility Via Theorem 2 From IRL Feasibility Test (6).

Step 1. Compute radar's naive response sequence $\beta_{1:K}^*$ by solving the convex optimization problem (3)

$$\beta_k^* = \operatorname{argmin} \mathbf{u}(\beta), \mathbf{g}(\alpha_k, \beta) \leq 0, \beta \geq \mathbf{0} \forall k \in \{1, 2, \ldots, K\}$$

where $\mathbf{u}$ is the concave monotone in β and $\mathbf{g}(\alpha_k, \beta)$ is the convex monotone in β.

Step 2. Choose $\eta \in [0, 1]$ (extent of cognition masking from IRL feasibility test).

Step 3. Compute upper bound $\mathcal{M}_{\text{thresh}}$ on the desired margin (15) after cognition masking: $\mathcal{M}_{\text{thresh}} = (1 - \eta)\ \mathcal{M}_{\mathbf{u}}(\{\alpha_k, \beta_k^*\}_{k=1}^K)$, where $\mathcal{M}_{\mathbf{u}}$ is defined in (15).

Step 4. Compute the cognition-masking responses by solving the following optimization problem:

$$\{\tilde{\beta}_{1:K}^*\}_{\text{MASK-U}} = \operatorname{argmin} \sum_{k=1}^{K} \mathbf{u}(\beta_k^*) - \mathbf{u}(\beta_k)$$
$$\beta_k \geq \mathbf{0},\ \alpha_k' \beta_k \leq 1\ \forall k \in \{1, 2, \ldots, K\}$$
$$\mathcal{M}_{\mathbf{u}}(\{\alpha_k, \beta_k\}_{k=1}^K) \leq \mathcal{M}_{\text{thresh}}\,. \tag{21}$$

Due to the nonlinear margin constraint in (21), the optimization problem can be solved using a general purpose nonlinear programming solver, for example, `fmincon` in MATLAB, to obtain a local minimum.

Let us briefly explain the essence of the cognition-masking algorithm in Theorem 2 through our running cognitive radar example from [20]. We first assume that the naive cognitive radar maximizes its QoS subject to constraints on its PCRB. The adversary exploits the ECM scheme of Theorem 1 to estimate the radar's QoS function and generates malicious probes (8). As an ECCM measure, the radar intentionally chooses a suboptimal waveform that trades off between maximizing the radar's QoS (17) and ensuring a poor reconstruction of the radar's strategy by the adversary [margin constraint (18)]. Let us consider the second scenario where the cognitive radar's utility is the inverse of the PCRB, that is, the radar minimizes its PCRB [20, eq. (40)] subject to a constraint on its communication cost with the central processing unit. The adversary can use Theorem 5 to estimate the radar's communication cost and can then use (12) to generate malicious probes. As an ECCM measure, the radar intentionally violates the communication cost constraint that trades off between minimizing the radar's transmission power (19) and ensuring a poor reconstruction of the radar's communication cost by the adversary [margin constraint (20)].

Summary

In this section, we introduced our key cognition-masking result, namely, Theorem 2 that mitigates the ECM attempts of the adversary (Theorems 1 and 5) to estimate the radar's strategy (utility function/resource constraint). From

a practitioner's perspective, we also related the cognition-masking scheme to a formal model of a cognitive radar [20] that chooses its waveform by solving a constrained optimization problem. This section sets the stage to address cognition masking from an adversary under noisy measurements. In the rest of this article, we motivate our cognition-masking results using two radar functionalities, namely, optimal waveform adaptation and optimal beam allocation, instead of the cognitive radar model of [20].

IV. HOW TO MASK COGNITION FROM DETECTOR?

The framework considered in Theorem 2 was deterministic; we assumed that the adversary had accurate measurements of the radar's responses. In this section, we generalize Theorem 2 to the case where the adversary has *noisy* measurements of the radar's decisions. That is, the noise term ω_k in the radar's response measurement $\hat{\beta}_k$ in (1) of Model 1 is a nonzero random variable with pdf f_ω. If the adversary deploys a Neyman–Pearson[10] type detector, how can we design our cognition-masking strategy to spoof this detector so that the radar can hide its utility and constraints? Before generalizing Theorem 2 to the noisy case, we first address the following question: *How do the adversary's IRL algorithms, Theorems 1 and 5, adapt to noisy measurements?*

A. Noisy Adversarial IRL Detectors Against Cognitive Radars

Our key IRL results for noisy radar measurements are outlined in Definition 4 in the following text. Recall from Section II that the adversary's IRL algorithm in Theorem 1 comprises a linear feasibility test to identify a feasible strategy that rationalizes the radar's responses. When the adversary has noisy measurements of the radar's response, the deterministic feasibility test generalizes to a *feasibility hypothesis test* to detect the existence of feasible strategies (utilities and constraints) so that the radar responses satisfy utility maximization (3).

For our hypothesis tests below, let H_0 and H_1 denote the null and alternate hypotheses that the adversary's noiseless datasets defined in (5) and (10) pass, and not pass, respectively, the IRL feasibility tests (6) and (50), respectively.

$$\begin{aligned} H_0 &: \text{ Radar is a constrained utility maximizer (3)} \\ H_1 &: \text{ Radar is NOT a constrained utility maximizer (3).} \end{aligned} \tag{22}$$

The two types of error that arise in hypothesis testing are Type-I and Type-II errors. In the radar context, the Type-I and Type-II errors have the following interpretation:

Type−I: Classify a cognitive radar as noncognitive

Type−II: Classify a noncognitive radar as cognitive. (23)

In analogy to Theorems 1 and 5, our IRL detectors defined below assume two scenarios, namely, Assumptions 3 and 4 that generalize Assumptions 1 and 2, respectively, to the case where the adversary has noisy response measurements.

ASSUMPTION 3 Consider the radar–adversary interaction scenario of Assumption 1. The adversary has access to the noisy dataset $\widehat{\mathcal{D}}_g$ defined as follows:

$$\widehat{\mathcal{D}}_g = \{\mathbf{g}(\alpha_k, \cdot), \hat{\beta}_k\}_{k=1}^K,\ \hat{\beta}_k = \beta_k + \omega_k,\ \omega_g \sim f_\omega \tag{24}$$

where $\mathbf{g}(\alpha_k, \cdot)$ is defined in (4), β_k is the radar's response, and ω_k is the adversary sensor's measurement noise (1) with pdf f_ω known to the radar.

IRL objective: The adversary uses the IRL detector (27) in Definition 4 to detect if the noise-free dataset $\mathcal{D}_g$ (5) passes the IRL test (6) of Theorem 1

ASSUMPTION 4 Consider the radar–adversary interaction scenario of Assumption 2. The adversary has access to the noisy dataset $\widehat{\mathcal{D}}_u$ defined as follows:

$$\widehat{\mathcal{D}}_u = \{\mathbf{u}(\alpha_k, \cdot), \hat{\beta}_k\}_{k=1}^K,\ \hat{\beta}_k = \beta_k + \omega_k,\ \omega_g \sim f_\omega \tag{25}$$

where β_k is the radar's response, and ω_k is the adversary sensor's measurement noise (1) with pdf f_ω known to the radar.

IRL objective: The adversary uses the IRL detector (27) in Definition 4 to detect if the noise-free dataset $\mathcal{D}_u$ (10) passes the IRL test (50) of Theorem 5.

Our IRL hypothesis tests for detecting radar's cognition (feasible utilities and resource constraints) for noisy radar response measurements are stated in Definition 4.

DEFINITION 4 (IRL DETECTORS FOR NOISY RESPONSE MEASUREMENTS) Consider the cognitive radar (3) from Definition 1 and the radar–adversary interaction from Model 1.

1) *IRL for detecting feasible utilities:* Suppose Assumption 3 holds. Then, the statistical test below detects if the radar's responses satisfy utility maximization behavior (3)

$$\boxed{\mathbb{P}(\phi_u^*(\widehat{\mathcal{D}}_g) \le L_g) \lesseqgtr_{H_0}^{H_1} \gamma}. \tag{26}$$

2) *IRL for detecting feasible resource constraints:* Suppose Assumption 4 holds. Then, the statistical test below detects if the radar's responses satisfy utility maximization behavior (3)

$$\boxed{\mathbb{P}(\phi_g^*(\widehat{\mathcal{D}}_u) \le L_u) \lesseqgtr_{H_0}^{H_1} \gamma}. \tag{27}$$

In the statistical tests (26) and (27) $\gamma \in [0, 1]$ is the "significance level" of the test. L_g and L_u are the random variables defined as follows:

$$L_g \equiv \max_{s,k}\ \alpha_k'(\omega_k - \omega_s) \tag{28}$$

[10]By Neyman–Pearson's lemma [101], it is impossible to maximize the Type-I and Type-II error of a detector simultaneously. In this article, we focus on mitigating the detector by maximizing its *conditional* Type-I error probability.

$$L_u \equiv \max_{s,k} \ (\mathbf{u}(\alpha_k, \hat{\beta}_k) - \mathbf{u}(\alpha_k, \hat{\beta}_s)) - (\mathbf{u}(\alpha_k, \hat{\beta}_k - \omega_k)) - \mathbf{u}(\alpha_k, \hat{\beta}_s - \omega_s)) \quad (29)$$

where $\omega_k \sim f_\omega$ is the measurement noise in the adversary's measurement of the radar's response (1). The test statistics $\phi_g^*(\cdot)$ and $\phi_u^*(\cdot)$ are the minimum perturbations required for the noisy datasets $\widehat{\mathcal{D}}_g$ and $\widehat{\mathcal{D}}_u$, respectively, to pass the IRL feasibility tests (6) and (50)

$$\phi_u^*(\widehat{\mathcal{D}}_g) = \min_{\epsilon,\theta>0} \epsilon, \ \mathcal{A}(\theta, \widehat{\mathcal{D}}_g + \epsilon) \leq 0 \quad (30)$$

$$\phi_g^*(\widehat{\mathcal{D}}_u) = \max_{\epsilon,\theta>0} \epsilon, \ \mathcal{A}(\theta, \widehat{\mathcal{D}}_u - \epsilon) \geq 0 \quad (31)$$

REMARKS

1) The random variable L_g (28) bounds the perturbation needed for $\widehat{\mathcal{D}}_g$ to pass the IRL test (6), if H_0 holds

$$H_0: \ \exists \theta > 0 \text{ s.t. } \mathcal{A}(\theta, \mathcal{D}_g) \leq 0 \Rightarrow \mathcal{A}(\theta, \widehat{\mathcal{D}}_g + L_g) \leq 0$$

where $\mathcal{D}_g$ is the noise-free version of the noisy dataset $\widehat{\mathcal{D}}_g$. Similarly, the random variable L_u (29) bounds the perturbation needed for $\widehat{\mathcal{D}}_u$ to pass the IRL test (50), if H_0 holds

$$H_0: \ \exists \theta > 0 \text{ s.t. } \mathcal{A}(\theta, \mathcal{D}_u) \geq 0 \Rightarrow \mathcal{A}(\theta, \widehat{\mathcal{D}}_u + L_u) \geq 0$$

where $\mathcal{D}_u$ is the noise-free version of the noisy dataset $\widehat{\mathcal{D}}_u$.

2) The IRL detectors (26) and (27) classify the radar as a utility maximizer if the perturbation needed for the feasibility of the IRL inequalities lies under a particular threshold, and vice-versa. Consider the statistical test of (26). Equation (26) can be expressed differently as follows:

$$\phi_u^*(\widehat{\mathcal{D}}_g) \underset{H_1}{\overset{H_0}{\lessgtr}} F_{L_\alpha}^{-1}(1-\gamma) \quad (32)$$

where the RHS term in (32) is the test threshold for test statistic $\phi_u^*(\cdot)$. Intuitively, the larger the perturbation needed for the feasibility of the IRL inequalities, the less confidence the adversary has to classify the radar as a utility maximizer.

Computational Complexity of IRL Detectors: The constrained optimization problems (30) and (31) are nonconvex since the RHS of the constraint is bilinear in the feasible variable. However, since the objective function depends only on a scalar, a 1-D line search algorithm can be used to solve for $\phi_u^*(\cdot)$ in (30) and $\phi_g^*(\cdot)$ in (31). That is, for any fixed value of ϵ, the constraints in (30) and (31) specialize to a set of linear inequalities for which feasibility is straightforward to check.

We now discuss a key feature of the statistical tests (26) and (27) in Theorem 3 that bounds the Type-I error probability $\mathbb{P}(H_1|H_0)$ of the IRL detectors. Recall that the Type-I error probability is the probability of incorrectly classifying the radar as noncognitive, when the radar's response is the solution of a constrained utility maximization problem (3).

THEOREM 3 (PERFORMANCE OF IRL DETECTORS (DEFINITION 4)) Consider the statistical tests (26) and (27) in Definition 4. The Type-I error probability of the tests is bounded by the significance level of the tests γ

$$\mathbb{P}(H_1|H_0) \leq \gamma \ \textit{for both detectors (26) and (27).} \quad (33)$$

The proof of Theorem 3 is in Appendix E. The key idea in the proof is to show that, given that the null hypothesis H_0 holds, the random variables L_g and L_u dominate the test statistics $\phi_g^*(\widehat{\mathcal{D}}_u)$ and $\phi_u^*(\widehat{\mathcal{D}}_g)$, respectively. Since the IRL detectors have a bounded Type-I probability, our cognition-masking rationale for the noisy case discussed in the following text is to maximize their conditional Type-I error probability.

B. Masking Cognition From IRL Detectors

In Section IV-A, we generalize the IRL results of Theorems 1 and 5 in Section II to the case where the adversary has noisy measurements of the radar's responses. The key idea is that the IRL feasibility tests (6) and (50) generalize to IRL detectors (26) and (27) in Definition 4, respectively, that detect utility maximization behavior. This section addresses cognition masking when the adversary uses the IRL detectors of Definition 4: *How to mitigate the statistical tests of (26) and (27) and make utility maximization detection difficult?*

Intuition for hiding cognition from IRL detectors: Suppose the radar follows the cognition-masking scheme of Theorem 2 for the noisy case. Indeed, the radar's strategy is hidden from the IRL feasibility tests of Theorems 1 and 5 but does not affect the performance of the IRL detectors of Definition 4. To do so, the radar maximizes the *conditional Type-I error probability*[11] of the IRL detectors by deliberately deviating from its naive responses (3). The conditional Type-I error probability can be viewed as the noisy analog of the inverse of the feasibility margin in the noiseless case.

DEFINITION 5 (CONDITIONAL TYPE-I ERROR PROBABILITY FOR IRL DETECTORS (DEFINITION 4)) Consider the datasets $\mathcal{D}_g$ and $\mathcal{D}_u$ defined in (5) and (10), and their corresponding noisy versions $\widehat{\mathcal{D}}_g$ and $\widehat{\mathcal{D}}_u$ defined in (24) and (25), respectively. Let $\phi_u(\widehat{\mathcal{D}}_g, \mathbf{u})$ and $\phi_g(\widehat{\mathcal{D}}_u, \mathbf{g})$ denote the minimum perturbations required for the tuples $(\widehat{\mathcal{D}}_g, \mathbf{u})$ and $(\widehat{\mathcal{D}}_u, \mathbf{g})$, respectively, to pass the IRL feasibility tests (6), (50)

$$\phi_u^*(\widehat{\mathcal{D}}_g, \mathbf{u}) = \min_{\epsilon \geq 0} \epsilon, \ \mathcal{A}(\mathbf{u}, \widehat{\mathcal{D}}_g + \epsilon) \leq 0$$
$$\phi_g^*(\widehat{\mathcal{D}}_u, \mathbf{g}) = \min_{\epsilon \geq 0} \epsilon, \ \mathcal{A}(\mathbf{g}, \widehat{\mathcal{D}}_u - \epsilon) \geq 0 \quad (34)$$

where $\mathbf{u}$ and $\mathbf{g}$ are the radar's utility and resource constraint, respectively. Then:

1) For IRL detector (26), the conditional Type-I error probability, conditioned on $\widehat{\mathcal{D}}_g$ (24) and radar's utility

[11] Radar can at best maximize the conditional Type-I error probability to mitigate the IRL detectors as the Type-I error probability is bounded by the detectors' significance level γ due to Theorem 3.

$\mathbf{u}$, is given by $\mathbb{P}(H_1|\mathcal{D}_g, \mathbf{u})$ and defined as follows:

$$\mathbb{P}(H_1|\mathcal{D}_g, \mathbf{u}) = \mathbb{P}(\ \phi_u^*(\widehat{\mathcal{D}}_g, \mathbf{u}) > F_{L_g}^{-1}(1-\gamma)\ ). \tag{35}$$

2) For IRL detector (27), the conditional Type-I error probability conditioned on $\widehat{\mathcal{D}}_u$ (25) and radar's constraint $\mathbf{g}$ is given by $\mathbb{P}(H_1|\widehat{\mathcal{D}}_u, \mathbf{g})$, and defined as follows:

$$\mathbb{P}(H_1|\mathcal{D}_u, \mathbf{g}) = \mathbb{P}(\ \phi_g^*(\widehat{\mathcal{D}}_u, \mathbf{g}) > F_{L_u}^{-1}(1-\gamma)\ ) \tag{36}$$

In (35) and (36), the alternate hypothesis event H_1 is expressed differently in the equivalent representation form of (32), and the random variables L_u and L_g are defined in (28) and (29)

REMARKS

1) The test statistics of the IRL detectors defined in (35) and (36) are computed via an optimization over $\mathbb{R}_+^{2K+1}$, whereas the optimization in (36) is over $\mathbb{R}_+$. Hence, $\phi_u^*(\widehat{\mathcal{D}}_g, \mathbf{u})$ and $\phi_g^*(\widehat{\mathcal{D}}_u, \mathbf{g})$ (36) are cheaper to compute than the test statistics $\phi_u^*(\widehat{\mathcal{D}}_g)$ (35) and $\phi_g^*(\widehat{\mathcal{D}}_u)$ (36), respectively.
2) The IRL detector performance is already constrained due to Theorem 3 (bounded Type-I error probability). Hence, to mitigate the IRL detector, the best the radar can do is to maximize its conditional Type-I error probability using the statistics defined in (34).

We are now ready to state our cognition-masking result, Theorem 4, that mitigates IRL detectors (Definition 4). In analogy to Theorem 2 for mitigating the IRL feasibility tests of Theorems 1 and 5, the radar deliberately degrades its performance to maximize the IRL detectors' conditional Type-I error probability defined in (35) and (36).

THEOREM 4 (MASKING COGNITION FROM ADVERSARIAL IRL DETECTORS) Consider the cognitive radar (3) from Definition 1. Let $\{\beta_k^*\}_{k=1}^K$ denote the naive response sequence (3) that maximizes the cognitive radar's utility. Then:

1) *Masking Utility Function From Detector:* Suppose Assumption 3 holds. Then, the response sequence defined below makes cognition detection difficult by ensuring that the detector (26) has a sufficiently large conditional Type-I error probability

$$\{\tilde{\beta}_{1:K}^*\} = \underset{\{\beta_k \geq \mathbf{0},\ \alpha_k'\beta_k \leq 1\}}{\operatorname{argmin}} \sum_{k=1}^{K} \mathbf{u}(\beta_k^*) - \mathbf{u}(\beta_k) - \lambda\ \mathbb{P}(H_1|\mathcal{D}_g, \mathbf{u}). \tag{37}$$

2) *Masking Resource Constraint From Detector:* Suppose Assumption 4 holds. Then, the response sequence below makes cognition detection difficult by ensuring that the detector (27) has a sufficiently large conditional Type-I error probability

$$\{\tilde{\beta}_{1:K}^*\} = \underset{\{\beta_k \geq \mathbf{0},\ \mathbf{g}(\beta_k) \leq \gamma_k\}}{\operatorname{argmin}} \sum_{k=1}^{K} \mathbf{u}(\beta_k^*) - \mathbf{u}(\beta_k) - \lambda\ \mathbb{P}(H_1|\mathcal{D}_u, \mathbf{g}). \tag{38}$$

In (37) and (38), the positive scalar λ parameterizes the extent of mitigation of the IRL detector.

Theorem 4 is our second result for cognition masking; see Algorithm 2 for a stepwise procedure for masking cognition in noise (37) when the adversary knows the radar's constraints. Equations (37) and (38) compute the *optimal* suboptimal radar response that sufficiently hides the radar's cognition from being detected by the IRL hypothesis tests of Definition 4. The parameter λ in Theorem 4 is analogous to parameter η in Theorem 2. A larger value of λ (37) results in a larger conditional Type-I error probability for the IRL detector (larger adversarial confusion) while increasing the radar's deviation from its optimal response (greater performance degradation), and vice-versa.

The optimization problems (37) and (38) can be solved by stochastic gradient algorithms. Algorithm 2 outlines a constrained simultaneous perturbation stochastic approximation (SPSA) [102], [103] implementation for computing the cognition-masking scheme of Theorem 4. The objective function is nonconvex in the radar's responses; hence, SPSA converges to a local optimum. SPSA is a generalization of adaptive algorithms where the gradient computation in (37) requires only two empirical estimates of the objective function per iteration, i.e., the number of evaluations is independent of the dimension of the radar's response. For decreasing step size $\eta = 1/i$ (42), the SPSA algorithm converges with probability one to a local stationary point. For constant step size η, it converges weakly (in probability).

Summary: In this section, we generalized our cognition-masking results of Theorem 2 to the case where the adversary has noisy measurements of the radar's responses. We first generalized our adversarial IRL feasibility tests of Theorems 1 and 5 to IRL hypothesis tests (26) and (27) in Definition 4 to detect utility maximization behavior, given noisy radar response measurements. We then present Theorem 4 that masks the radar's cognition by making utility maximization detection erroneous by maximizing the conditional Type-I error probability of the IRL detectors via purposeful suboptimal responses. Our cognition-masking results can be extended without loss of generality (WLOG) to any suboptimal IRL algorithm, as discussed in Appendix H.

V. NUMERICAL RESULTS FOR I-IRL

In this section, we illustrate how our cognition-masking results of Theorems 2 and 4 successfully confuse adversarial IRL via the two radar tracking functionalities, namely, waveform adaptation and beam allocation, as discussed in Section II.

TABLE II
Parameters for Numerical Experiments

Masking smart waveform adaptation	
Time horizon	$K = 50$
Probe/response dimension	$m = 4$
Probe	$\alpha_k(i) \overset{\text{i.i.d}}{\sim} \text{Unif}(0.2, 2.5)$, $i = 1, 2, \ldots, m$
Utility function	$\mathbf{u}_1(\beta) = \sum_{i=1}^{m} \sqrt{\beta(i)}$ $\mathbf{u}_2(\beta) = \sum_{i=1}^{m} \beta(i)^2$
Resource constraint	$\mathbf{g}(\alpha_k, \beta) = \alpha_k' \beta - 1$
Masking smart beam allocation	
Time horizon	$K = 50$
Probe/response dimension	$m = 4$
Probe	$\alpha_k(i) \overset{\text{i.i.d}}{\sim} \text{Unif}(0.1, 0.7)$, $i = 1, 2, \ldots, m$
Utility function	$\mathbf{u}(\alpha_k, \beta) = \prod_{i=1,2,\ldots,m} \beta(i)^{\alpha_k(i)}$
Resource constraint	$\mathbf{g}(\alpha_k, \beta) = \|\beta\|_\kappa - \gamma_k$, $\gamma_k \overset{\text{i.i.d}}{\sim} \text{Unif}(0.5, 2)$.

A. Cognition Masking Via Theorem 2 for Noiseless Adversary Measurements

Consider the scenario where the adversary has accurate measurements of the radar's responses. Recall from Section II-D that the adversary knows the radar's constraints for waveform adaptation and the radar's utility function for beam allocation. For waveform adaptation, the probe signal parameterizes the state covariance matrix of the radar's Kalman filter due to the adversary's maneuvers, and the response signal parameterizes the sensory accuracy chosen by the radar. Recall that the probe signal α_k is the diagonal of the state noise covariance matrix: $Q_k = \text{diag}[\alpha_k(1), \alpha_k(2)]$. For beam allocation, the ith component of the probe signal $\alpha_k(i)$ is the asymptotic predicted precision of the radar tracker for target i. The probe α_k parameterizes the radar's Cobb–Douglas utility for beam allocation. Our simulation parameters for our numerical experiments are listed in Table II.

Parameters for Numerical Experiments

In Table II, $\text{Unif}(a, b)$ denotes the uniform pdf with support (a, b). The elements of the probes α_k (3) and intensity thresholds γ_k (14) are generated randomly and independently over time $k = 1, 2, \ldots, K$. For waveform adaptation, we conduct our numerical experiments for two distinct utility functions $\mathbf{u}_1$ and $\mathbf{u}_2$. Given the probe sequence $\{\alpha_k\}_{k=1}^K$, the cognitive radar chooses its smart response sequence via (17) for masking optimal waveform adaptation and via (19) for masking optimal beam allocation. Recall from Section II that response β_k is the diagonal of the inverse of radar's observation noise covariance matrix for waveform adaptation. For beam allocation, $\beta_k(i)$ is the beam intensity directed toward target i at time k.

Figs. 4 and 5 show the performance loss (minimum perturbation from optimal response computed via (17) and (19) in Theorem 2) of the cognitive radar as a function of η (extent of cognition masking) when the cognitive radar performs waveform adaptation and beam allocation, respectively. We see that for both functionalities, both the radar's performance loss and adversarial IRL mitigation increase with η. This is expected since larger η implies a larger shift of the set of feasible strategies computed via IRL to ensure that the radar's strategy is sufficiently close to the edge of the feasible set at the cost of greater deviation from the radar's optimal strategy.

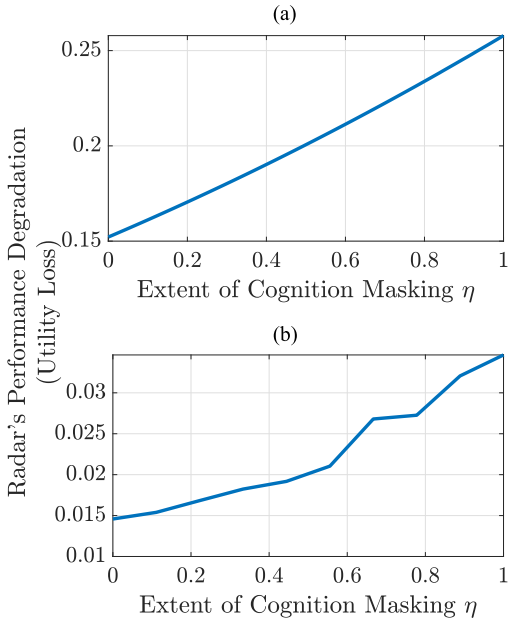


Fig. 4. Masking waveform adaptation strategy from adversarial IRL: Small deliberate performance loss (vertical axis) of the cognitive radar results in large performance mitigation of the adversary (horizontal axis). The figure illustrates a cognitive radar operating with two distinct utility functions.
1) $\eta = 1$ corresponds to maximum cognition masking and, hence, results in maximum performance loss. 2) For a fixed value of η, the quadratic utility (b) requires smaller perturbation (≈ 10 times) from the optimal response compared with the sublinear utility of subfigure (a).
(a) $\mathbf{u}(\beta) = \sum_{i=1}^{m} \sqrt{\beta(i)}$. (b) $\mathbf{u}(\beta) = \sum_{i=1}^{m} \beta(i)^2$.

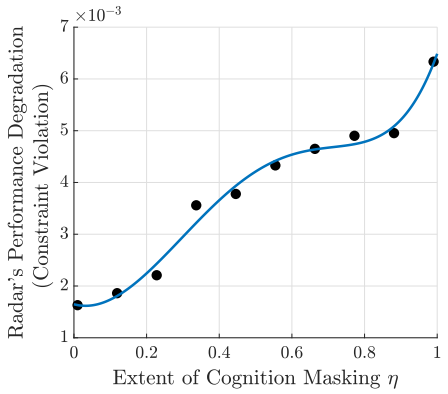


Fig. 5. Masking beam allocation strategy from adversarial IRL: Small deliberate utility loss of the radar (vertical axis) results in large performance loss (extent of strategy masking η) of the adversarial IRL algorithm (horizontal axis). $\eta = 0$ corresponds to zero strategy masking, and $\eta = 1$ corresponds to complete strategy masking by the radar. As expected, the deliberate utility loss of the radar increases with η.

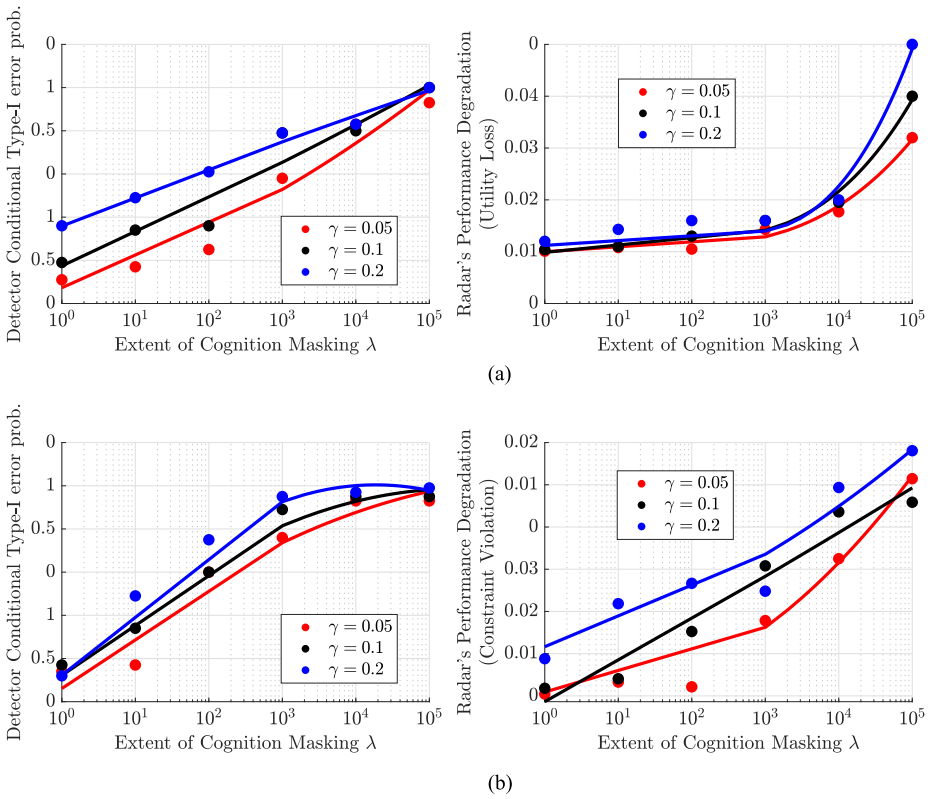


Fig. 6. Masking cognition from IRL detectors: Performance of metacognitive radar for waveform adaptation (a) and beam allocation (b) when the adversary deploys an IRL detector (26) and (27) for cognition detection. The key takeaway is that a small sacrifice in performance of the radar results in a large performance loss of adversary's IRL detector. The performance loss of both the radar and the adversary due to metacognition increases with scaling factor λ (37) and significance level γ of the adversary's IRL detectors (26) and (27).

B. Cognition Masking Via Theorem 4 for Noisy Adversary Measurements

We now consider the scenario where the adversary has noisy measurements of the radar's response. Consider the simulation parameters of Table II. For our second set of numerical experiments for both waveform adaptation and beam allocation, we set the noise pdf f_ω (1) to $\mathcal{N}(0, 0.3I)$, where $\mathcal{N}(\mu, \Sigma)$ denotes the multivariate normal distribution with mean μ and covariance Σ, and I denotes the identity matrix in Theorem 4.

For the noisy case, we consider only a single utility function for waveform adaption, namely, $\mathbf{u}(\beta) = \sum_{i=1}^{m} \sqrt{\beta(i)}$. We performed our numerical experiments for three values of $\gamma = \{0.05, 0.1, 0.2\}$ for both waveform adaptation and beam allocation. Recall from Section IV that γ is the significance level of the adversary's IRL detectors (26) and (27) in Definition 4.

Given the probe sequence $\{\alpha_k\}_{k=1}^{K}$, we generated the cognition-masking response sequence via (37) for waveform adaption and (38) for beam allocation by varying the parameter λ (37) over the interval $[10^0, 10^5]$. Recall from Theorem 4 that the radar minimizes the detectors' conditional Type-I error probabilities (35) and (36) to mitigate adversarial IRL while deliberately compromising on its performance (utility).

Our SPSA algorithm [102], [103] (Algorithm 2) for stochastic gradient descent was executed over 10^4 iterations for all pairs of (λ, γ), $\lambda \in \{10^0, 10^1, 10^2, 10^3, 10^4, 10^5\}$ and $\gamma \in \{0.05, 0.1, 0.2\}$. Fig. 6 shows the conditional Type-I error probability (adversarial IRL mitigation) of the detector and performance loss of the radar as the parameter λ is varied for three different values of the significance level α of the adversary's detector. Recall from Theorem 4 that the parameter λ controls the extent of cognition masking for noisy I-IRL. From Fig. 6, we see that both the conditional Type-I error probability of the IRL detectors and radar's performance loss increase with λ as well as γ.

If $\lambda = 0$, the radar simply transmits its naive response that maximizes its utility (no performance loss) and also results in zero adversarial mitigation. For the limiting case of $\lambda \to \infty$, the radar's cognition-masking response computed via Theorem 4 degenerates to a constant for all time k, hence maximizing the conditional Type-I error probability of the detector at the cost of maximal performance loss for the radar.

Let us briefly discuss the variation of the radar performance and adversarial mitigation as the parameter γ is varied. γ (26) can be viewed as the risk-aversion tendency of the adversary's IRL system since it bounds the detector's Type-I error probability. Recall from (22) that the Type-I

error is the probability of detecting a cognitive radar as noncognitive. Higher γ implies that the detector is *risk seeking* and a lower γ implies that the detector is *risk averse*. Naturally, a larger deviation from the optimal strategy is required to mitigate a risk-averse detector compared with a risk-seeking detector to the same extent.

VI. CONCLUSION AND EXTENSIONS

This article investigated how a cognitive radar can hide its cognition from an adversary when the adversary performs IRL to estimate the radar's utility function by observing its actions. The adversary's IRL estimate of the radar's strategy is a polytope of feasible solutions to a set of convex inequalities. Our first cognition-masking result is Theorem 2. When the adversary has accurate measurements of the radar's response, cognition masking via Theorem 2 ensures that the radar's true strategy lies close to the edge of the feasibility polytope computed via adversarial IRL (true strategy poorly rationalizes adversary's dataset). When the adversary has noisy measurements of the radar's response, adversarial IRL generalizes to a cognition detector defined in Definition 4. Our second cognition-masking result is Theorem 4. The key idea is to maximize the probability of the radar being classified as noncognitive by the detector subject to a bound on the radar's performance loss. Finally, in Section V, we illustrate our cognition-masking results on a cognitive radar that performs waveform adaptation and beam allocation for target tracking. We show that small purposeful deviations from the optimal strategy of the radar suffice to significantly confuse the adversarial IRL system.

This article builds significantly on our previous work [35] on ECM for identifying cognitive radars, and [104], [105], and [106] on ECCM for masking radar cognition. Theorem 6 extends IRL for cognitive radars [35] when the radar faces multiple resource constraints. The linear IRL feasibility test for a single constraint case generalizes to a mixed-integer feasibility test. Theorem 7 generalizes the cognition-masking result of [104] to multiple constraints. Our previous works [104], [105], [106] assume optimal adversarial IRL via Afriat's theorem. This article generalizes cognition masking to suboptimal adversarial IRL algorithms. Algorithm 3 outlines a cognition scheme when the adversary uses an arbitrary IRL algorithm to estimate the radar's strategy. Theorem 8 provides performance bounds for our cognition-masking scheme when the adversary has misspecified measurements of the radar's response. Although this article is radar-centric, we emphasize that the problem formulation and algorithms developed also apply to adversarial IRL in general machine learning applications.

Finally, a useful extension of this article would be to study cognition masking in a dynamic radar–adversary interaction environment in comparison with the batchwise probe–response exchange considered in this article. Also, how to mask cognition when the adversary knows of the radar's ECCM capability? Such an approach warrants a game-theoretic discussion in terms of a Stackelberg game where the adversary moves first and the radar responds to the adversary's probes. It is also worthwhile exploring state-of-the-art concepts in chance constrained optimization [107] and robust optimization [108], [109] to achieve cognition masking under uncertainty—when the radar has noisy measurements of the adversary's probes.

Algorithm 2: SPSA for Mitigating Utility Maximization Detection for Adversarial IRL Detector (26) [(37) in Theorem 4].

Step 1. Set $\boldsymbol{\beta}_0 = \{\beta^*_{1:K}\}$, the naive response sequence (13) that maximizes the radar's utility (3).
Step 2. Choose $\lambda > 0$ (extent of cognition masking).
Step 3. For iterations $i = 0, 1, 2, \ldots$, (i) Compute $\widehat{\mathbb{P}}(H_1|\{\alpha_k\}_{k=1}^K, \boldsymbol{\beta}_i, \mathbf{u})$, the empirical probability estimate of the conditional Type-I error probability of the detector (26) defined in (35) using $R \times K$ fixed realizations $\{\omega_{r,k}\}_{r,k=1}^{R,K}$ of adversary's measurement noise $\omega_k \sim f_\omega$ (1)

$$\frac{1}{R}\sum_{r=1}^{R} \mathbb{1}\left\{\phi^*_u(\{\alpha_k, \beta_{i,k} + \omega_{r,k}\}_{k=1}^K, \mathbf{u}) > F^{-1}_{L_g}(1-\gamma)\right\} \tag{39}$$

In (39)
- $\boldsymbol{\beta}_i \equiv \{\beta_{i,1:K}\} \geq \mathbf{0}$ is a vector of responses.
- $\mathbb{1}\{\cdot\}$ denotes the indicator function.
- R controls the accuracy of the empirical probability estimate.
- $F_{L_g}(\cdot)$ is the distribution function of the r.v. L_g (26).
- The statistic $\phi^*_u(\cdot, \mathbf{u})$ is defined in (34). Let $J(\boldsymbol{\beta}_i)$ denote the objective being maximized in (37)

$$J(\boldsymbol{\beta}_i) = \sum_{k=1}^{K} \mathbf{u}(\beta_{i,k}) - \mathbf{u}(\beta_{i,k}) - \lambda \mathbb{P}(H_1|\{\alpha_k\}_{k=1}^K, \boldsymbol{\beta}_i, \mathbf{u}) \tag{40}$$

Then: (ii) Compute empirical estimate $\widehat{J}(\boldsymbol{\beta}_i)$

$$\widehat{J}(\boldsymbol{\beta}_i) = \sum_{k=1}^{K} \mathbf{u}(\beta^*_k) - \mathbf{u}(\beta_{i,k}) - \lambda\, \widehat{\mathbb{P}}(H_1|\{\alpha_k\}_{k=1}^K, \boldsymbol{\beta}_i, \mathbf{u}) \tag{41}$$

where $\widehat{\mathbb{P}}(H_1|\{\alpha_k\}_{k=1}^K, \boldsymbol{\beta}_i, \mathbf{u})$ is computed in (39).
(ii) Compute the estimate of the gradient $\nabla_{\boldsymbol{\beta}} J(\boldsymbol{\beta}_i)$ as follows:

$$\widehat{\nabla}_{\boldsymbol{\beta}}(\widehat{J}(\boldsymbol{\beta}_i)) = \frac{\Delta_i}{\omega \|\Delta_i\|_F^2}\, \widehat{J}(\boldsymbol{\beta}_i + \delta\ \Delta_i) - \widehat{J}(\boldsymbol{\beta}_i - \delta\ \Delta_i)$$

where δ is the gradient step size, $||\cdot||_2$ denotes the Frobenius norm, and $\Delta_i \in \{-1, +1\}^{m\times K}$ is a random perturbation vector whose each element is ± 1 with probability $1/2$. (iii) Update the radar's response as follows:

$$\boldsymbol{\beta}_{i+1} = \text{Proj}_{S_\alpha}\left(\boldsymbol{\beta}_i + \eta\, \frac{\Delta_i}{\|\Delta_i\|_F} \widehat{\nabla}_{\boldsymbol{\beta}} \widehat{J}(\boldsymbol{\beta}_i)\right) \tag{42}$$

where η is the response update step size and Proj_{S_α} is the projection operator to the hyperplane $S_\alpha = \{\beta_{1:K} : \alpha_k' \beta_k = 1, \beta_k \geq \mathbf{0}\}$.
Step 4. Set $i \leftarrow i + 1$ and go to Step 3.

REFERENCES

[1] K. V. Mishra, M. B. Shankar, and B. Ottersten, "Toward metacognitive radars: Concept and applications," in *Proc. IEEE Int. Radar Conf.*, 2020, pp. 77–82.

[2] X. Wang, Z. Fei, J. A. Zhang, J. Huang, and J. Yuan, "Constrained utility maximization in dual-functional radar-communication multi-UAV networks," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2660–2672, Apr. 2021.

[3] A. F. Martone et al., "Closing the loop on cognitive radar for spectrum sharing," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 9, pp. 44–55, Sep. 2021.

[4] A. F. Martone, "Cognitive radar demystified," *URSI Radio Sci. Bull.*, vol. 2014, no. 350, pp. 10–22, Sep. 2014.

[5] L. Zhao and D. P. Palomar, "Maximin joint optimization of transmitting code and receiving filter in radar and communications," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 850–863, Feb. 2017.

[6] L. Wu, P. Babu, and D. P. Palomar, "Cognitive radar-based sequence design via SINR maximization," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 779–793, Feb. 2017.

[7] A. F. Martone et al., "Practical aspects of cognitive radar," in *Proc. IEEE Radar Conf.*, 2020, pp. 1–6.

[8] S. D. Blunt et al., "Principles and applications of random FM radar waveform design," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 35, no. 10, pp. 20–28, Oct. 2020.

[9] B. Ravenscroft, J. W. Owen, J. Jakabosky, S. D. Blunt, A. F. Martone, and K. D. Sherbondy, "Experimental demonstration and analysis of cognitive spectrum sensing and notching for radar," *IET Radar, Sonar Navig.*, vol. 12, pp. 1466–1475, 2018.

[10] A. Aubry, A. De Maio, M. Piezzo, M. M. Naghsh, M. Soltanalian, and P. Stoica, "Cognitive radar waveform design for spectral coexistence in signal-dependent interference," in *Proc. IEEE Radar Conf.*, 2014, pp. 474–478.

[11] M. Alaee-Kerahroodi, E. Raei, S. Kumar, and B. S. M. R. Rao, "Cognitive radar waveform design and prototype for coexistence with communications," *IEEE Sensors J.*, vol. 22, no. 10, pp. 9787–9802, May 2022.

[12] H. Esmaeili-Najafabadi, H. Leung, and P. W. Moo, "Unimodular waveform design with desired ambiguity function for cognitive radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 3, pp. 2489–2496, Jun. 2020.

[13] M. R. Bell, "Information theory and radar waveform design," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1578–1597, Sep. 1993.

[14] R. Romero and N. A. Goodman, "Information-theoretic matched waveform in signal dependent interference," in *Proc. IEEE Radar Conf.*, 2008, pp. 1–6.

[15] N. A. Goodman, P. R. Venkata, and M. A. Neifeld, "Adaptive waveform design and sequential hypothesis testing for target recognition with active sensors," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 105–113, Jun. 2007.

[16] H. He, P. Stoica, and J. Li, "Waveform design with stopband and correlation constraints for cognitive radar," in *Proc. 2nd Int. Workshop Cogn. Inf. Process.*, 2010, pp. 344–349.

[17] W. Melvin, M. Wicks, P. Antonik, Y. Salama, P. Li, and H. Schuman, "Knowledge-based space-time adaptive processing for airborne early warning radar," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 13, no. 4, pp. 37–42, Apr. 1998.

[18] W. W. Howard, A. F. Martone, and R. M. Buehrer, "Distributed online learning for coexistence in cognitive radar networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 2, pp. 1202–1216, Apr. 2022.

[19] S. Maleki et al., "Cognitive spectrum utilization in Ka band multibeam satellite communications," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 24–29, Mar. 2015.

[20] P. Chavali and A. Nehorai, "Scheduling and power allocation in a cognitive radar network for multiple-target tracking," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 715–729, Feb. 2012.

[21] J. Li, L. Xu, P. Stoica, K. W. Forsythe, and D. W. Bliss, "Range compression and waveform optimization for MIMO radar: A Cramér–Rao bound based study," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 218–232, Jan. 2008.

[22] K. L. Bell, C. J. Baker, G. E. Smith, J. T. Johnson, and M. Rangaswamy, "Cognitive radar framework for target detection and tracking," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1427–1439, Dec. 2015.

[23] M. Hernandez, T. Kirubarajan, and Y. Bar-Shalom, "Multisensor resource deployment using posterior Cramér-Rao bounds," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 2, pp. 399–416, Apr. 2004.

[24] V. C. Vannicola and J. A. Mineo, "Applications of knowledge based systems to surveillance," in *Proc. IEEE Nat. Radar Conf.*, 1988, pp. 157–164.

[25] R. A. Romero and N. A. Goodman, "Cognitive radar network: Cooperative adaptive beamsteering for integrated search-and-track application," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 2, pp. 915–931, Apr. 2013.

[26] R. A. Romero and N. A. Goodman, "Adaptive beamsteering for search-and-track application with cognitive radar network," in *Proc. IEEE RadarCon*, 2011, pp. 1091–1095.

[27] A. Aubry, A. De Maio, and M. Govoni, "Two-dimensional spectrum sensing for cognitive radar," in *Proc. IEEE Radar Conf.*, 2018, pp. 815–820.

[28] A. Aubry, V. Carotenuto, A. De Maio, and M. A. Govoni, "Multi-snapshot spectrum sensing for cognitive radar via block-sparsity exploitation," *IEEE Trans. Signal Process.*, vol. 67, no. 6, pp. 1396–1406, Mar. 2018.

[29] P. Setlur and M. Rangaswamy, "Joint filter and waveform design for radar stap in signal dependent interference," *IEEE Trans. Sig. Process.*, vol. 64, no. 1, pp. 19–34, 2015.

[30] E. Raei, M. Alaee-Kerahroodi, and M. B. Shankar, "ADMM based transmit waveform and receive filter design in cognitive radar systems," in *Proc. IEEE Radar Conf.*, 2020, pp. 1–6.

[31] G. Rossetti and S. Lambotharan, "Waveform optimization techniques for bi-static cognitive radars," in *Proc. IEEE 12th Int. Colloquium Signal Process. Appl.*, 2016, pp. 115–118.

[32] G. Rossetti, A. Deligiannis, and S. Lambotharan, "Waveform design and receiver filter optimization for multistatic cognitive radar," in *Proc. IEEE Radar Conf.*, 2016, pp. 1–5.

[33] G. Rossetti and S. Lambotharan, "Coordinated waveform design and receiver filter optimization for cognitive radar networks," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, 2016, pp. 1–5.

[34] J. Guerci, J. Bergin, R. Guerci, M. Khanin, and M. Rangaswamy, "A new MIMO clutter model for cognitive radar," in *Proc. IEEE Radar Conf.*, 2016, pp. 1–6.

[35] V. Krishnamurthy, D. Angley, R. Evans, and B. Moran, "Identifying cognitive radars—Inverse reinforcement learning using revealed preferences," *IEEE Trans. Sig. Process.*, vol. 68, pp. 4529–4542, 2020, doi: 10.1109/TSP.2020.3013516.

[36] V. Krishnamurthy, K. Pattanayak, S. Gogineni, B. Kang, and M. Rangaswamy, "Adversarial radar inference: Inverse tracking, identifying cognition, and designing smart interference," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 4, pp. 2067–2081, Aug. 2021.

[37] A. Y. Ng et al., "Algorithms for inverse reinforcement learning," in *Proc. 17th Int. Conf. Mach. Learn.*, 2000, pp. 663–670.

[38] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proc. 21st Int. Conf. Mach. Learn.*, 2004, pp. 1–8.

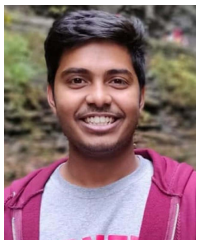
[39] B. D. Ziebart et al., "Maximum entropy inverse reinforcement learning," in *Proc. 23rd Nat. Conf. Artif. Intell.*, 2008, pp. 1433–1438.

[40] S. Afriat, "The construction of utility functions from expenditure data," *Int. Econ. Rev.*, vol. 8, no. 1, pp. 67–77, 1967.

[41] A. Caplin and M. Dean, "Revealed preference, rational inattention, and costly information acquisition," *Amer. Econ. Rev.*, vol. 105, no. 7, pp. 2183–2203, 2015.

[42] J. Boyd, D. B. Harris, D. D. King, and H. W. Welch Jr., "Electronic countermeasures," *Electron. Countermeasures*, 1978. [Online]. Available: https://ui.adsabs.harvard.edu/abs/1978elcm.book.....B
[43] A. Kuptel, "Counter unmanned autonomous systems (CUAXS): Priorities, policy, future capabilities," *Policy. Future Capabilities. Multinational Capability Develop. Campaign*, pp. 15–16, May 5, 2017.
[44] L. Snow, V. Krishnamurthy, and B. M. Sadler, "Identifying coordination in a cognitive radar network—A multi-objective inverse reinforcement learning approach," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2023, pp. 1–5.
[45] C. Shi, F. Wang, M. Sellathurai, and J. Zhou, "Low probability of intercept-based distributed MIMO radar waveform design against barrage jamming in signal-dependent clutter and coloured noise," *IET Signal Process.*, vol. 13, no. 4, pp. 415–423, 2019.
[46] F. A. Butt, I. H. Naqvi, and U. Riaz, "Hybrid phased-MIMO radar: A novel approach with optimal performance under electronic countermeasures," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1184–1187, Jun. 2018.
[47] S. Gong, X. Wei, and X. Li, "ECCM scheme against interrupted sampling repeater jammer based on time-frequency analysis," *J. Syst. Eng. Electron.*, vol. 25, no. 6, pp. 996–1003, 2014.
[48] H. Varian, "Revealed preference and its applications," *Econ. J.*, vol. 122, no. 560, pp. 332–338, 2012.
[49] V. Krishnamurthy and W. Hoiles, "Afriat's test for detecting malicious agents," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 801–804, Dec. 2012.
[50] K. Amin, R. Cummings, L. Dworkin, M. Kearns, and A. Roth, "Online learning and profit maximization from revealed preferences," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 770–776.
[51] A. Roth, J. Ullman, and Z. S. Wu, "Watch and learn: Optimizing from revealed preferences feedback," in *Proc. 28th Annu. ACM Symp. Theory Comput.*, 2016, pp. 949–962.
[52] Y. Sakuma, T. P. Tran, T. Iwai, A. Nishikawa, and H. Nishi, "Trajectory anonymization through Laplace noise addition in latent space," in *Proc. 9th Int. Symp. Comput. Netw.*, 2021, pp. 65–73.
[53] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate, "Differentially private empirical risk minimization," *J. Mach. Learn. Res.*, vol. 12, no. 3, pp. 1069–1109, 2011.
[54] R. Shokri, "Privacy games: Optimal user-centric data obfuscation," in *Proc. Privacy Enhancing Technol.*, vol. 2015, no. 2, 2015, pp. 1–17.
[55] G. Beigi, A. Mosallanezhad, R. Guo, H. Alvari, A. Nou, and H. Liu, "Privacy-aware recommendation with private-attribute protection using adversarial learning," in *Proc. 13th Int. Conf. Web Search Data Mining*, 2020, pp. 34–42.
[56] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation*. Hoboken, NJ, USA: Wiley, 2008.
[57] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking—Part I: Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.
[58] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems.* Norwood, MA, USA: Artech House, 1999.
[59] S. Haykin, "Cognitive radar: A way of the future," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 30–40, Jan. 2006.
[60] S. Haykin, "Cognitive dynamic systems: Radar, control, and radio [point of view]," *Proc. IEEE*, vol. 100, no. 7, pp. 2095–2103, Jul. 2012.
[61] F. Smits, A. Huizing, W. van Rossum, and P. Hiemstra, "A cognitive radar network: Architecture and application to multiplatform radar management," in *Proc. Eur. Radar Conf.*, 2008, pp. 312–315.
[62] M. Kozy, J. Yu, R. M. Buehrer, A. Martone, and K. Sherbondy, "Applying deep-Q networks to target tracking to improve cognitive radar," in *Proc. IEEE Radar Conf.*, 2019, pp. 1–6.
[63] C. E. Thornton, R. M. Buehrer, A. F. Martone, and K. D. Sherbondy, "Experimental analysis of reinforcement learning techniques for spectrum sharing radar," in *Proc. IEEE Int. Radar Conf.*, 2020, pp. 67–72.
[64] E. Blasch et al., "Issues and challenges of knowledge representation and reasoning methods in situation assessment (level 2 fusion)," *Proc. SPIE*, vol. 6235, 2006, Art. no. 623510.
[65] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.
[66] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
[67] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2407–2423, Sep. 2004.
[68] P. Stoica, Z. Wang, and J. Li, "Robust capon beamforming," in *Proc. Conf. Rec. 36th Asilomar Conf. Signals, Syst. Comput.*, 2002, pp. 876–880.
[69] A. Stringer, G. Dolinger, D. Hogue, L. Schley, and J. G. Metcalf, "A meta-cognitive approach to adaptive radar detection," *IEEE Trans. Aerosp. Electron. Syst.*, early access, May 8, 2023, doi: 10.1109/TAES.2023.3274101.
[70] G. T. Capraro and M. C. Wicks, "Metacognition for waveform diverse radar," in *Proc. Int. Waveform Diversity Des. Conf.*, 2012, pp. 348–351.
[71] A. F. Martone et al., "Metacognition for radar coexistence," in *Proc. IEEE Int. Radar Conf.*, 2020, pp. 55–60.
[72] L. Neng-Jing and Z. Yi-Ting, "A survey of radar ECM and ECCM," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1110–1120, Jul. 1995.
[73] D. C. Schleher, "LPI radar: Fact or fiction," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 21, no. 5, pp. 3–6, May 2006.
[74] P. E. Pace, *Detecting and Classifying Low Probability of Intercept Radar*. Norwood, MA, USA: Artech House, 2009.
[75] W.-Q. Wang, "Moving-target tracking by cognitive RF stealth radar using frequency diverse array antenna," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3764–3773, Jul. 2016.
[76] W.-Q. Wang, "Adaptive RF stealth beamforming for frequency diverse array radar," in *Proc. 23rd Eur. Signal Process. Conf.*, 2015, pp. 1158–1161.
[77] Z. Zhang, S. Salous, H. Li, and Y. Tian, "Optimal coordination method of opportunistic array radars for multi-target-tracking-based radio frequency stealth in clutter," *Radio Sci.*, vol. 50, no. 11, pp. 1187–1196, 2015.
[78] L. Neng-Jing, "Radar ECCMs new area: Anti-stealth and anti-ARM," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1120–1127, Jul. 1995.
[79] F. Forges and E. Minelli, "Afriat's theorem for general budget sets," *J. Econ. Theory*, vol. 144, no. 1, pp. 135–145, 2009.
[80] M. Arik and O. B. Akan, "Enabling cognition on electronic countermeasure systems against next-generation radars," in *Proc. MILCOM IEEE Mil. Commun. Conf.*, 2015, pp. 1103–1108.
[81] A. Charlish, F. Hoffmann, C. Degen, and I. Schlangen, "The development from adaptive to cognitive radar resource management," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 35, no. 6, pp. 8–19, Jun. 2020.
[82] A. Gupta and V. Krishnamurthy, "Principal agent problem as a principled approach to electronic counter-countermeasures in radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 4, pp. 3223–3235, 2022.
[83] S. Jain, K. Pattanayak, V. Krishnamurthy, and C. Berry, "Adaptive ECCM for mitigating smart jammers," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2023, pp. 1–5.
[84] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 330–343, Jan. 2007.
[85] F. Gini, A. De Maio, and L. Patton, *Waveform Design and Diversity for Advanced Radar Systems*. London, U.K.: IET, 2012.
[86] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4264–4279, Aug. 2018.

[87] Z. Wei, Z. Liu, B. Peng, and R. Shen, "ECCM scheme against interrupted sampling repeater jammer based on parameter-adjusted waveform design," *Sensors*, vol. 18, no. 4, 2018, Art. no. 1141.
[88] Y. Liu, G. Liao, Z. Yang, and J. Xu, "Multiobjective optimal waveform design for OFDM integrated radar and communication systems," *Signal Process.*, vol. 141, pp. 331–342, 2017.
[89] E. Grossi and M. Lops, "MIMO radar waveform design: A divergence-based approach for sequential and fixed-sample size tests," in *Proc. IEEE 3rd Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process.*, 2009, pp. 165–168.
[90] V. Krishnamurthy and R. Evans, "Hidden Markov model multi-arm bandits: A methodology for beam scheduling in multi-target tracking," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 2893–2908, Dec. 2001.
[91] V. Krishnamurthy and D. Djonin, "Optimal threshold policies for multivariate POMDPs in radar resource management," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3954–3969, Oct. 2009.
[92] M. Xie, W. Yi, L. Kong, and T. Kirubarajan, "Receive-beam resource allocation for multiple target tracking with distributed MIMO radars," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 5, pp. 2421–2436, Oct. 2018.
[93] J. Wang and H. Zhu, "Beam allocation and performance evaluation in switched-beam based massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, 2015, pp. 2387–2392.
[94] J. Wang, H. Zhu, L. Dai, N. J. Gomes, and J. Wang, "Low-complexity beam allocation for switched-beam based multiuser massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8236–8248, Dec. 2016.
[95] H. R. Varian, *Goodness-of-Fit for Revealed Preference Tests*. Ann Arbor, MI, USA: Dept. Econ., Univ. Michigan, 1991.
[96] M. Dean and D. Martin, "Measuring rationality with the minimum cost of revealed preference violations," *Rev. Econ. Statist.*, vol. 98, no. 3, pp. 524–534, 2016.
[97] F. Echenique, S. Lee, and M. Shum, "The money pump as a measure of revealed preference violations," *J. Political Econ.*, vol. 119, no. 6, pp. 1201–1223, 2011.
[98] B. Smeulders, F. C. Spieksma, L. Cherchye, and B. De Rock, "Goodness-of-fit measures for revealed preference tests: Complexity results and algorithms," *ACM Trans. Econ. Comput.*, vol. 2, no. 1, pp. 1–16, 2014.
[99] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory*, 1992, pp. 144–152.
[100] N. D. Ratliff, J. A. Bagnell, and M. A. Zinkevich, "Maximum margin planning," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 729–736.
[101] H. V. Trees, *Detection, Estimation and Modulation Theory*. Hoboken, NJ, USA: Wiley, 1968.
[102] J. Spall, *Introduction to Stochastic Search and Optimization*. Hoboken, NJ, USA: Wiley, 2003.
[103] I.-J. Wang and J. C. Spall, "A constrained simultaneous perturbation stochastic approximation algorithm based on penalty functions," in *Proc. Amer. Control Conf.*, 1999, pp. 393–399.
[104] K. Pattanayak, V. Krishnamurthy, and C. Berry, "How can a cognitive radar mask its cognition?," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2022, pp. 5897–5901.
[105] K. Pattanayak, V. Krishnamurthy, and C. Berry, "Meta-cognition: An inverse-inverse reinforcement learning approach for cognitive radars," in *Proc. 25th Int. Conf. Inf. Fusion*, 2022, pp. 1–8.
[106] K. Pattanayak, V. Krishnamurthy, and C. Berry, "Inverse-inverse reinforcement learning: How to hide strategy from an adversarial inverse reinforcement learner," in *Proc. IEEE 61st Conf. Decis. Control*, 2022, pp. 3631–3636.
[107] A. Nemirovski and A. Shapiro, "Scenario approximations of chance constraints," in *Probabilistic and Randomized Methods for Design Under Uncertainty*. Berlin, Germany: Springer, 2006, pp. 3–47.
[108] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*, vol. 28. Princeton, NJ, USA: Princeton Univ. Press, 2009.
[109] H.-G. Beyer and B. Sendhoff, "Robust optimization—A comprehensive survey," *Comput. Methods Appl. Mech. Eng.*, vol. 196, pp. 3190–3218, 2007.
[110] K. Pattanayak, V. Krishnamurthy, and E. Blasch, "Inverse sequential hypothesis testing," in *Proc. IEEE 23rd Int. Conf. Inf. Fusion*, 2020, pp. 1–7.
[111] K. Pattanayak and V. Krishnamurthy, "Unifying classical and Bayesian revealed preference," 2023, *arXiv:2106.14486*.
[112] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice Hall, 1979.
[113] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[114] J. Akhtar, "Orthogonal block coded ECCM schemes against repeat radar jammers," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 3, pp. 1218–1226, Jul. 2009.
[115] M. Soumekh, "SAR-ECCM using phase-perturbed LFM chirp signals and DRFM repeat jammer penalization," in *Proc. IEEE Int. Radar Conf.*, 2005, pp. 507–512.
[116] D. S. Garmatyuk and R. M. Narayanan, "ECCM capabilities of an ultrawideband bandlimited random noise imaging radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1243–1255, Oct. 2002.



Kunal Pattanayak (Student Member, IEEE) received the integrated B.tech. and M.tech. degrees in electronics and electrical communication from the Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA, in 2018.

His research interests broadly include statistical signal processing, focusing primarily on inverse sensing and information economics.

Mr. Pattanayak is a recipient of the McMullen graduate fellowship from Cornell University, and has been a speaker at the 2020 Sloan-NOMIS Conference on Attention and Applied Economics.

Vikram Krishnamurthy (Fellow, IEEE) received the Ph.D. degree in department of systems engineering from Australian National University, Canberra, ACT, Australia, in 1992.

From 2002 to 2016, he was a Professor and Canada Research Chair with the University of British Columbia, Canada. He is a Professor with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. His research interests include statistical signal processing and stochastic control in social networks and adaptive sensing. He served as a Distinguished Lecturer of the IEEE Signal Processing Society and an Editor-in-Chief for IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING. In 2013, he was awarded an Honorary Doctorate from KTH (Royal Institute of Technology), Sweden. He is the author of two books *Partially Observed Markov Decision Processes* and *Dynamics of Engineered Artificial Membranes and Biosensors* published by Cambridge University Press in 2016 and 2018, respectively.

Christopher M. Berry (Student Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Drexel University, Philadelphia, PA, USA, in 2010 and 2013. He is currently working toward the Ph.D. degree in electrical engineering with the New Jersey Institute of Technology, Newark, NJ, USA.

He is currently a Senior Member of the Engineering Staff with Lockheed Martin Advanced Technology Laboratories, Cherry Hill, NJ, USA. His research interests include statistical signal processing, machine learning, multitarget tracking, and inverse sensing.