Online Bayesian Meta-Learning for Cognitive Tracking Radar

CHARLES E. THORNTON, Graduate Student Member, IEEE
RICHARD M. BUEHRER, Fellow, IEEE
Virginia Tech, Blacksburg, VA USA

ANTHONY F. MARTONE, Fellow, IEEE
U.S. Army Research Laboratory, Adelphi, MD USA

A key component of cognitive radar is the ability to generalize or achieve consistent performance across a range of sensing environments, since aspects of the physical scene may vary over time. This presents a challenge for learning-based waveform selection approaches, since transmission policies that are effective in one scene may be highly suboptimal in another. We address this problem by strategically biasing a learning algorithm by exploiting high-level structure across tracking instances, referred to as Meta-Learning. In this work, we develop an online Meta-Learning approach for waveform-agile tracking. This approach uses information gained from previous target tracks to speed up and enhance learning in new tracking instances. This results in sample-efficient learning across a class of finite-state target channels by exploiting inherent similarity across tracking scenes, attributed to common physical elements such as target type or clutter statistics. We formulate the online waveform selection problem within the framework of Bayesian learning and provide prior-dependent performance bounds for the Meta-Learning problem using probability approximately correct Bayes theory. We present a computationally feasible metaposterior sampling algorithm and study the performance in a simulation study consisting of diverse scenes. Finally, we examine the potential performance benefits and practical challenges associated with online Meta-Learning for waveform-agile tracking.

Manuscript received 7 July 2022; revised 14 December 2022 and 26 April 2023; accepted 7 May 2023. Date of publication 12 May 2023; date of current version 1 October 2023.

DOIs. No. 10.1109/TAES.2023.3275552

Refereeing of this contribution was handled by A. Charlish.

This work was supported by the U.S. Army Research Office (ARO). An earlier version of this paper was presented in part at the 2022 IEEE Radar Conference [DOI: 10.1109/RadarConf2248738.2022.9763914].

Authors’ addresses: Charles E. Thornton and Richard M. Buehrer are with the Department of ECE, Virginia Tech, Blacksburg, VA 24061 USA, E-mail: (thorntoc@vt.edu; buehrer@vt.edu); Anthony F. Martone is with U.S. Army Research Laboratory, Adelphi, MD 20783 USA, E-mail: (anthony.f.martone.civ@army.mil). (Corresponding author: Charles E. Thornton.)

1By sample-efficient, we refer to algorithms with bounded regret, which captures the deficit between a given decision-making strategy and an optimal decision rule, given a fixed objective [13, Sec. 4.4].
that the learner uses to make predictions about new data, where both time and observations are limited. In the present work, we study an online or sequential approach to Meta-Learning where the inductive bias is gradually inferred over time after each task is completed. Here, the base learning process occurs within a track while the Meta-Learning process occurs on a track-to-track basis.

Deducing the value of unseen future data using the past is impossible without further assumptions [19]. Thus, all data-driven approaches are inherently biased by said assumptions. Selecting an appropriate bias by hand is often cumbersome, as encoding domain knowledge into complex model classes is notoriously difficult. It is important to note that conventional learning approaches involve learning a new model from scratch for each task. Such approaches do not exploit task similarity and can be data inefficient. An approach alternative to Meta-Learning is learning a single model, which performs well on all tasks within a given class. Unfortunately, such a model may not exist or sufficient data and learning time may not be available. This is especially true for the current application of target tracking.

Here, we consider a scenario in which a stationary monostatic radar engages in a sequence of target tracks. The radar may track multiple targets at a time and wishes to select waveforms, which allow for accurate estimates of the target’s position in the presence of clutter, interference, and noise. This waveform-agile tracking scenario is posed as an online Meta-Learning problem. We cast the repeated waveform selection process as a linear contextual bandit problem, and use a posterior (Thompson) sampling algorithm [20], [21] to select waveforms. Posterior sampling approaches are known to provide both good theoretical and practical performance but are sensitive to the choice of the prior distribution, which is difficult to select by hand. Thus, we aim to gradually learn a prior distribution, which allows for efficient learning across a class of tracking instances. For a stationary radar platform, we expect the target and clutter types the radar experiences from track-to-track to exhibit learnable structure, due to similarity across target and clutter impulse responses. Thus, we expect the experience to carry over from one track to another, allowing the radar to learn an effective bias for any tracking instance.

In this formulation, the waveform selection process for each time-limited target track corresponds to a new learning instance or “task.” We consider a very general signal model, which incorporates spatially extended target and clutter responses, meaning that returns may occupy multiple range cells. This matched-illumination setup is commonly studied in the waveform design literature [22], [23], [24] and is appropriate for modern radar systems, which are capable of transmitting increasingly wideband pulses. Each learning task is then specified by the target’s physical trajectory, the target impulse response, clutter impulse response, and additive noise model.

We argue that Meta-Learning has significant potential for performance improvements in tracking radar systems due to several aspects of the problem. First, the radar is assumed to be embedded in a physical environment, which remains relatively stationary from track-to-track. We expect there to be a high degree of similarity between tracking instances, inducing some learnable structure. However, each track is not identical, due to the new targets’ trajectory and changes in the physical scene. Furthermore, the number of observations during each target track is limited, which requires the radar to learn quickly and sequentially. Finally, the first few measurements of a track are of particular importance so that the track is not lost. Thus, learning a reasonable waveform selection strategy quickly will result in significantly fewer lost tracks. Here, we propose a practical online Meta-Learning scheme that drastically reduces the probability of losing a track across a variety of physical scenes.

A. Related Work

This work addresses the problem of learning optimal adaptive waveform selection strategies for a broad class of radar environments. Therefore, some notes on the various notions of optimality for radar waveforms are required. There is a rich history of the literature on radar waveform design and providing a comprehensive survey is beyond the scope of the present work. The interested reader can consult several excellent surveys on waveform diversity and design [25], [26], [27], [28], [29] and a good work on radar signals by Levanon [30]. Quoting Levanon, “the work (or art) of designing radar waveforms is based mostly on experience and expertise obtained through successive designs” [30, Sec. 1.5]. Here, we seek to automate the process of gaining experience by way of Meta-Learning.

The general principle of waveform design is to utilize a transmit signal that maximizes information gained about the target while accounting for channel effects due to propagation loss, noise, and interference. The following papers have been influential, but the list is by no means exhaustive.

Bell [31] proposed mutual information between the target impulse response and the received signal as a waveform design criterion for estimating the stochastic parameters of extended targets. Kay [22] addressed the problem of optimal signal design and detection for point targets in signal-dependent Gaussian distributed clutter. Pillai et al. [24] consider an SINR-based design of pulsed waveforms for the detection of extended targets in signal-dependent noise, importantly showing that chirplike signals are often insufficient in the presence of low-pass dominant noise. However, an issue with both SINR and mutual-information-based waveform design procedures is that closed-form solutions are often not possible and on-the-fly waveform design can be computationally burdensome. These and other issues are addressed by Romero et al. [23], which extend the mutual-information-based design to the case of signal-dependent clutter. Some of the information-theoretic tradeoffs in waveform design for extended target detection have been examined by Zhu et al. [32], [33]. Other works have focused on optimizing the transmitted code sequence and receive filter simultaneously using sophisticated optimization techniques [34], [35], [36].
In a real-time target tracking system, it is generally preferable to forgo waveform design and instead select from a catalog of known waveforms. This is due to various resource constraints that make on-the-fly waveform design difficult [37]. The problem of waveform-agile tracking is surveyed in [38]. Numerous approaches have followed the Fisher information perspective, introduced by Kershaw and Evans [39], [40], [41], [42]. Other approaches have focused on waveform optimization by estimating clutter statistics on-the-fly [43], [44]. However, the aforementioned waveform selection techniques become computationally burdensome when longer-term scheduling is considered.

It has been known for some time that nonmyopic scheduling techniques are effective for a diverse range of sensor management problems, which often require planning [45], [46]. However, computational concerns historically limited the practical applications of stochastic control processes in large-scale waveform selection problems. In recent years, computational and algorithmic advances have allowed for real-time adaptive transmission schemes based on dynamic programming [47], deep reinforcement learning [48], universal source coding [49], and more computationally feasible schemes based on multiarmed bandit and contextual bandit learning [14], [50]. While these learning schemes have proven to be effective in environments with stationary dynamics, the major issues of generalization and sample-efficiency remained understudied. In Table I, we outline the relative advantages and drawbacks of the learning methods described in such works briefly.

This article proposes online Meta-Learning as a method for real-time waveform selection. Meta-Learning is broadly defined as a method of learning from examples that leverages experience gained from previous interactions. This principle is based on the idea that learning multiple related tasks should improve performance [17], [18], [53], [54], [55]. The metalearner’s goal is to generalize from a finite set of observed tasks to arbitrary new tasks. In this present context of radar tracking, we consider each track to be a new task. Our assumption at the outset is that learning a waveform selection strategy for a given target track should give the radar prior knowledge about which waveforms may be appropriate to use for future target tracks. This is because of similarities in the target impulse response and clutter impulse response from track-to-track.

Recently, Meta-Learning has been proposed for several wireless communication and networking applications [56], [57], [58], [59], [60], with the goal of reducing training overhead and complexity. This application is reasonable, as domain knowledge is available in many wireless applications, but is often difficult to encode this knowledge by hand. Meta-Learning circumvents this challenge by learning an inductive bias automatically. Several contributions have focused on using Meta-Learning to train detectors with a limited number of training examples [57], [60]. Here, we focus exclusively on the problem of transmit waveform selection. The number of observations available for learning is inherently limited by the application of target tracking. Additionally, we assume that the radar is physically stationary, from which we may expect a great deal of similarity across tracking instances. Thus, we argue that the waveform-agile tracking problem is well-posed to benefit from Meta-Learning techniques.

### Table I

| Formulation                  | Contextual Bandit       | MDP/POMDP               | Context/Tree Weighting | Meta-Learning + Contextual Bandit |
|-----------------------------|-------------------------|-------------------------|------------------------|----------------------------------|
| **Advantages**              | Low-Complexity/Online/Bounded Regret | Accounts for impact of radar’s actions/Bellman Optimality/Scales to large action spaces via deep learning | Bellman Optimality/Few a priori assumptions/Online | Low-Complexity/Bounded Regret/Generalization/Online |
| **Drawbacks**               | No context awareness/Converges to single action | Does not account for impact of radar’s actions/Unknown generalization performance | May require offline training/Sample inefficient/Generalization unknown | Limited to low dimension problems/Poor generalization | Assumes task similarity |

B. Our Contributions

This work is the first known investigation of online Meta-Learning to a waveform selection problem, other than our preliminary research [1]. In the current article, we extend the work in [1] by providing additional details regarding the learning framework and present general signal and channel models, incorporating both extended targets and signal-dependent clutter. We also present an extended simulation study.

More specifically, we make the following contributions to the known literature.

C1. We develop a general framework for the Meta-Learning problem by defining class of finite-state representations for the waveform-agile target tracking problem. To develop a model for this class, we extend the notion of a finite-state target channel (FSTC), which describes the radar propagation environment for a specific tracking instance [61, Ch. 3.2.3] to an ensemble, or class of such channels that can be used to specify a Meta-Learning problem.

C2. We identify the benefits and limitations of Bayesian learning for the application of radar tracking in more detail than the previous work and discuss the significant impact of prior misspecification on tracking.
performance. In particular, we note that prior misspecification can result in a significant probability of low SNR measurements early in a track, when a target is most likely to be lost.

C3. We demonstrate that a meta-Thompson sampling algorithm can be applied to effectively learn a prior for a class of FSTCs satisfying sufficient structure, namely that the class of channels can be parameterized by a vector, and the FSTC for each tracking instance can be thought of as independent samples coming from a fixed distribution. We argue that this is due to inherent similarities between radar tracking instances, such as common target material and size.

C4. We discuss some limitations of the proposed Meta-Learning approach and discuss practical considerations. We identify how a Meta-Learning algorithm may interact with other processes in a multifunction radar system.

C. Organization

The rest of this article is organized as follows. Section II describes the radar signal model and propagation characteristics and proceeds to a discussion of the FSTC considered. Section III presents the Bayesian learning problem of interest and describes the process of Meta-Learning by which the radar learns an inductive bias. Section IV presents the Meta-Learning algorithm used and describes associated considerations. Section V examines the performance of the proposed approach in simulation. Finally, Section VI concludes this article.

II. SYSTEM MODEL

We consider a physically stationary and monostatic radar system. The radar engages in an ongoing sequence of \( m \) target tracking intervals. The number of targets tracked, target impulse responses, and object trajectories may vary from one interval to another. The radar may track multiple targets concurrently. The targets are assumed to be spatially extended over multiple range cells, against a cluttered background, which is also extended in range. Each target tracking interval lasts for a fixed number \( 1 < n < \infty \) of coherent processing intervals (CPIs). During each tracking interval, the radar monitors a single target that is assumed to have been detected during an earlier scan period. The radar wishes to maintain adequate detection performance while accurately estimating target parameters of interest.

Let the area covered by the radar’s beam be represented by a discrete sampling grid in delay-Doppler space, with delay cells indexed by \( \tau = 1, \ldots, \tau_{\text{max}} \) and Doppler cells indexed by \( v = 1, \ldots, v_{\text{max}} \). During each CPI, the radar transmits a finite-duration bandpass signal expressed by the model

\[
s_{\tau}(t) = \sqrt{2}\text{Re}\left[\sqrt{E_{\tau}}\hat{s}(t)\exp(j2\pi f_{c}t)\right]
\]

where \( f_{c} \) is the carrier frequency, \( E_{\tau} \) is the energy of the transmitted pulse, and \( \hat{s}(t) \) is the complex envelope of the signal [62, Appendix], which is assumed to be normalized, i.e., \( \int |\hat{s}(t)|^{2}dt = 1 \). We note that the delay-Doppler ambiguity properties of the transmitted waveform are completely determined by the complex envelope of the signal [30, Ch. 3]. We assume that \( E_{\tau} \) and \( f_{c} \) are fixed, and wish only to vary the complex envelope on a CPI-to-CPI basis. This is a standard problem considered in the waveform design literature [22, 23, 44].

Here, we assume the transmitted signal during each CPI is a coherent train of \( N_{p} \) identical pulses. Thus, the complex envelope of the transmitted signal is expressed by the sum

\[
\hat{s}(t) = \frac{1}{\sqrt{N_{p}}} \sum_{i=1}^{N_{p}} p_{i}(t - (i - 1)T_{r})
\]

where \( p_{i} \) is the complex envelope of the \( i \)th pulse in the CPI and \( T_{r} \) is the pulse repetition period.

In this formulation, we constrain the choice of envelope to a finite set of known waveforms. However, we wish to diversify the envelope enough such that a good waveform can be selected for a variety of tracking instances. Thus, we allow \( p(t) \) to be either a generalized frequency modulated (GFM) chirp or a phase-coded pulse. The complex envelope of a GFM chirp pulse is given by

\[
p_{i}(t) = a(t) \exp\left(j2\pi b\xi(t/t')\right)
\]

where \( a(t) \) is an envelope function, \( b \) is a scalar FM rate parameter, \( \xi(t) \) is the chirp phase function [41], and \( t' > 0 \) is a reference time point. The complex envelope of the phase-coded pulse is given by

\[
p_{i}(t) = \frac{1}{\sqrt{T}} \sum_{m=1}^{M} u_{m} \text{rect}\left[t - (m - 1)t_{b}\right]
\]

where \( M \) is the number of subpulses, \( T \) is the pulse duration, \( u_{m} = \exp(j\phi_{m}) \), and the set of \( M \) phases \( \{\phi_{m}\}_{m=1}^{M} \) is the specific phase code associated with \( \hat{s}(t) \). A detailed description of commonly used phase codes and their desirable properties can be found in [30, Ch. 6].

Each CPI, we assume that the transmitted pulse train illuminates both moving targets and surrounding clutter. The received signal is then given by

\[
r(t) = \chi(t) * [s_{R}(t) + n(t)]
\]

where \( \chi(t) \) is the complex-valued receive filter impulse response, which we take to be a matched filter (i.e., \( \chi(t) = \hat{s}^{*}(-t) \)) and \( n(t) \) is a complex-valued signal-independent additive noise term, due to receiver noise, interference, or jamming. Finally, \( s_{R}(t) \) is the reflected signal. Assuming \( L \) targets and \( H \) clutter components, the received signal during the CPI can be expressed as

\[
s_{R}(t) = \sum_{g=1}^{G} e^{j2\pi f_{g}t} \int_{0}^{t_{\text{max}}} \frac{1}{R_{g}} K_{\xi}(\tau) \hat{s}\left(t - \left(\frac{R_{g}}{c} + \tau\right)\right) d\tau
\]
\[
+ \sum_{h=1}^{H} e^{j2\pi f^h t} \int_{0}^{\tau^{\text{max}}} \frac{1}{R_h} \lambda_h(t) s \left( t - \frac{R_h}{c} + \tau \right) d\tau
\]  

(6)

where \( \kappa_g(t) \) is the complex impulse response of target \( g \), which may be modeled as a finite-energy stochastic process that varies on a CPI-to-CPI basis, \( R_g \) and \( f^h_g \) are the range-Doppler frequency of target \( g \), respectively, and \( \lambda(t) \) is the clutter’s complex impulse response, which also may be modeled as a finite-energy random process in general. We assume that the number of targets \( L \) is known a priori from a detection process.

Analogously, it is often more convenient to specify the target and clutter models by their frequency responses, given by

\[
h_g(f) = \int_{0}^{\tau^{\text{max}}} \kappa_g(\tau) \exp(-j2\pi ft)d\tau
\]  

(7)

\[
c_h(f) = \int_{0}^{\tau^{\text{max}}} \lambda_h(\tau) \exp(-j2\pi ft)d\tau. 
\]  

(8)

This system model assumes that both the target and clutter are extended in range and, thus, exhibit frequency selectivity [62, Ch. 12]. This is a reasonable assumption when the target and clutter occupy multiple range cells, which is relevant for many high-resolution applications. For example, extended target modeling for missile- and airplane-shaped targets has been conducted in [63], [64], and [65].

From (6), we observe that the complex envelope should in some sense be “matched” to a composite channel response that incorporates both the target and clutter impulse responses. This is the problem investigated in [23], [66], and [67]. However, in many cases, it is difficult or impossible to solve for the optimal waveform in a closed form. Furthermore, the approach in [23] and [66] assumes knowledge of the target and clutter impulse response a priori, which is not practical for real-time tracking systems.

In our model (6), target illumination, clutter illumination, and various signal-independent noise components all affect the received signal and may change significantly over the course of a track. Thus, solving for the optimal waveform in a closed form will be intractable in general. A helpful model for such time-varying channels is the finite-state representation [68, Sec. 4.6]. This model has been widely applied in communication theory to model channels with intersymbol interference and fading and has also been suggested for radar [61].

To motivate the Meta-Learning problem, we consider a scenario in which each tracking instance is described by a unique FSTC with memory. We assume that the radar engages in an ongoing sequence of target tracks, and that there is an underlying class of FSTCs that compose the physical scene, accounting for variability due to changes in the target and clutter impulse responses. In the next section, we introduce the FSTC model and relevant assumptions.

A. FSTC Model

Each discrete time index \( k \) denoted by \( k = 1, 2, \ldots, n \) corresponds to the present radar CPI within a target track. During CPI \( k \), the radar scene is said to be in a state denoted by \( s_k \), which takes values in finite set \( S \). The state describes relative losses due to the scattering effects of the propagation environment, as well as from the target, representing a general notion of the channel’s effect. Thus, the state captures the relevant effects present in (6), namely due to the targets’ positions, Doppler shifts, and impulse responses, as well as the analogous scatterers modeled as clutter.

The state varies over time according to a discrete-time stochastic process \( \{s_k\}_{k \in \mathbb{N}_+} \), with a fixed memory length \( L < \infty \). More specifically, this means that knowledge of the entire state sequence does not provide more information about the current state than knowledge of the past \( L \) states. In mathematical terms

\[
P(s_k | s_{k-1}^{k-1}) = P(s_k | s_{k-L}^{k-1}) 
\]  

(9)

for every possible state sequence. The memory length is unknown to the radar. Thus, the state-generating process is said to have transition probabilities \( P(s_k | s_{k-L}^{k-1}) \), which are unknown to the radar a priori. In this formulation, we assume the scene’s state transitions occur independently of the radar’s transmissions. This is true of most radar tracking scenarios but notably excludes the cases of an adversarial target response or adversarial interference.

Since the radar is a measurement system, it does not observe the underlying state directly. The radar instead has access to an observation \( o_k \in O \) at each time step. This observation is a hypothesis regarding the target’s position and the scattering effects of the channel. The observation process is governed by the probability kernel \( P(o_k | s_k) \), defined for each \( o_k \in O \) and \( s_k \in S \).

Based on the sequence of observations \( o_t \), the radar must select a waveform \( w_k \) from a finite alphabet of waveforms \( \mathcal{W} \). The radar wishes to measure a time-varying vector of physical target parameters, such as range, velocity, and target size, expressed by \( z_k \in \mathbb{Z} \). This vector is assumed to be generated by a fixed source distribution \( P(z_k) \), which is to be independent of the radar’s choice of waveform. We have now established the tools necessary to define our FSTC, which the radar will repeatedly interact with during a target track.

**DEFINITION 1 (FSTC)** An FSTC is specified by the state transition probabilities defined in (9), the source distribution

\[ ^2 \text{For convenience, we consider discrete time with finite observation and waveform sets. The extension to continuous domains comes with a loss of generality, and requires more sophisticated assumptions.} \]

\[ ^3 \text{We use the notation } s_t \text{ to denote the sequence of states from time } 1 \text{ to } k. \]
where \( \mathcal{S} \) is a finite alphabet of states, \( \mathcal{W} \) is a finite waveform alphabet, \( \mathcal{Z} \) is a finite set of possible target parameter vectors, \( \mathcal{Y} \) is a not necessarily finite set of received signals, and \( P: \mathcal{W} \times \mathcal{S} \times \mathcal{Z} \mapsto \mathcal{Y} \) is a matrix of probabilities, which expresses the probability of receiving a particular signal given a waveform, state sequence, and target parameter.

**Remark 1** Our channel model is a generalization of the FSTC proposed by Bell [61], as the state sequence of our FSTC is not assumed to be a first-order Markov process and may depend on the entire past state sequence. Our FSTC also differs from that of Bell in that we do not allow the target parameter vector to depend on the radar’s choice of the waveform.

**Remark 2** A conventional learning approach would seek to optimize the choice of waveforms for a fixed but a priori unknown, FSTC. We will refer to the problem of learning a waveform selection strategy for a particular FSTC as the task, to remain consistent with the Meta-Learning literature. Meta-Learning aims to find an inductive bias, which allows for efficient learning across a class of FSTCs with parameters drawn from a common distribution.

**Definition 2 (FSTC Class)** Given a common state, waveform, and target parameter alphabet \((\mathcal{W}, \mathcal{S}, \mathcal{Z})\), we can consider a class \( \Theta \) of FSTCs, where each channel is specified by the FSTC parameter vector \( \theta \in \Theta \) and is characterized by the conditional probability assignment

\[
P(y|w, s, z, \theta)
\]

for every \( w \in \mathcal{W}, s \in \mathcal{S}, \) and \( z \in \mathcal{Z} \).

The above definition is analogous to a class of finite-state fading channels [69]. The FSTC parameter vector \( \theta \) specifies the unique propagation characteristics of each channel not captured by the state. These effects are due to any variations in the target and clutter impulse responses in (6), which are not explicitly captured by the state. It is this parameter vector, which we seek to learn in Section III.

**Assumption 1 (Task Structure):** Each target track corresponds to a new FSTC, drawn independently from a fixed class \( \Theta \). For a sequence of tracks, we assume the existence of a distribution \( P(\theta) \) on \( \Theta \). This assumption is reasonable since a new target track corresponds to a new trajectory, target scattering behavior, and possible changes in clutter and noise distribution. However, since the physical scene is expected to remain relatively stationary over time, the scene should have enough structure to define \( \Theta \) and \( P(\theta) \).

Since capacity-achieving schemes are impractical without explicit assumptions and a priori knowledge of the target channel’s behavior, we frame the waveform selection process for a fixed FSTC as a statistical learning problem.

During each track instance, the radar wishes to select the waveform that minimizes an unknown loss function during each CPI, which is a sequential decision problem over a finite time horizon. The loss represents a performance characteristic, which is calculable without knowledge of the underlying channel state. Instead of learning a fixed waveform selection strategy that performs well across an entire class of FSTCs, we seek to learn an inductive bias that allows for efficient learning of any new channel within the class.

### III. Learning Framework

The radar engages in a sequence of \( m \) target tracking periods. During each tracking instance, \( s = \{1, 2, \ldots, m\}, \) the radar maintains an estimate of the target parameter vector \( \{z_k\}_{k=1}^n \) over a fixed time horizon of \( n \) CPIs. Based on the selected waveform at CPI \( k \), namely \( w_k \) and underlying state \( s_k \), the radar receives a loss, defined by the sequence of mappings \( \epsilon_k: \mathcal{W} \times \mathcal{S} \mapsto \mathbb{R} \). The loss must be a quantity, which is directly observable by the radar. For example, features from the range-Doppler response, such as estimated SINR, or outputs from the radar’s tracking filter, such as the normalized innovation, are natural choices for the loss function [70, Sec. III]. Since the state is not fully observable by the radar, the loss appears stochastic. The goal of the learning problem at each time step is to select the waveform

\[
\omega_k^* = \min_{\omega_k \in \mathcal{W}} \mathbb{E}[\epsilon_k(w_k, s_k)]
\]

where the expectation is taken over the variability in the loss mapping. Unfortunately, this expectation is not computable as both the sequences of loss mappings \( \{\epsilon_k\} \) and state values \( \{s_k\} \) are unknown to the radar. Thus, the radar must select waveforms using only the history of observations, transmitted waveforms, and received losses up to step \( k - 1 \), given by the set

\[
\mathcal{F}_{k-1} = \{(a_i, w_i, \ell_i)\}_{i=1}^{k-1}.
\]

In general, this problem is practically challenging for the following two reasons.

1. Each state value \( s_k \) is only partially observed through \( a_k \). Thus, the radar must select waveforms with only partial knowledge of the underlying FSTC.
2. The structure of \( \ell_k \) is unknown to the radar a priori and must be inferred in the small-sample regime. Thus, the radar must gather information about the losses associated with each waveform while trying to minimize the total number of waveforms transmitted, which do not achieve close to the lowest possible loss.

Thus, the radar must gather information about the current FSTC instantiation by gradually building \( \mathcal{F}_{k-1} \) through

---

As mentioned in Section II-A, the true value of the state is unobserved by the radar.

We describe our specific choice of the loss function in Section V. An overview of possible choices can be found in [70, Sec. III].
repeated experience. The information gathered in the history must be sufficiently rich, such that the radar has enough knowledge to select near-optimal waveforms, in terms of our loss function $\ell_k$. However, this information must be gathered quickly, such that the total number of highly suboptimal waveforms transmitted during the tracking interval is kept low. Thus, the canonical dilemma of exploration and exploitation is present. This challenge is particularly relevant for the inherently time-sensitive tracking problem, as the number of highly suboptimal waveforms transmitted must be kept low throughout the entirety of the tracking instance to avoid losing the target.

We now introduce a pseudo-Bayesian learning formulation, in which waveforms are selected based on the estimated posterior probability that they yield the lowest loss. We develop a scheme to estimate the FSTC parameter $\theta$, which can be used to map radar observations to expected losses. We note that in this scheme, the estimated posterior distribution does not model any inherent randomness in the value of $\theta$ itself, but rather reflects the radar’s uncertainty about the true value of $\theta$.

In essence, we parameterize each target tracking instance by a vector $\theta \in \mathbb{R}^d$, which is unknown a priori. This vector can be used to map the radar’s observations to expected losses by weighting over context features. We express uncertainty about the true, unknown, value of $\theta$ by introducing a prior distribution over the range of possible values, expressed by $P(\theta)$. As the radar gains “experience,” we expect the posterior distribution $P(\theta|F)$ to begin to concentrate around a certain value. We have thus reduced the sequential decision problem to a density estimation problem. This approach, known as posterior sampling, or Thompson sampling, is widely used for sequential decision making due to a unique combination of theoretical performance and practical efficiency.

A. Thompson Sampling Approach

A common framework used for sequential decision making is the Bayesian approach. Under this framework, decisions are made with respect to the posterior probability that they will yield the lowest loss [18, Sec. 2.2]. This approach assumes the learning task can be described by a set of probability distributions $\{P_i\}$ parameterized by $\theta \in \Theta \subset \mathbb{R}^d$.

A computationally feasible example of Bayesian learning is posterior sampling in the well-known stochastic linear bandit problem [13, Ch. 19], which is adopted here. In this framework, we define $\varphi: \mathcal{O} \times \mathcal{W} \mapsto \mathbb{R}^d$ to be a feature mapping, which describes aspects of the physical scene. The stochastic linear bandit model is then specified by assuming the relation

$$\ell_k = \langle \theta^*, \varphi(o_i, w_i) \rangle + \eta_k$$

holds for each pair $(o_i, w_i) \in \mathcal{O} \times \mathcal{W}$, where $\eta_k$ is a sub-Gaussian random variable. Under this assumption, knowledge of the vector $\theta$ is sufficient to predict $E[\ell_k|$ for each pair of waveform and observation $(o_i, w_i)$. Taking a Bayesian perspective, the radar must use observed data $\mathcal{F}$ to learn a posterior distribution over the parameter $P(\theta|\mathcal{F})$. This is the base learning problem we will study in the remainder of this article.

Given the stochastic linear bandit assumption (13), the ability to parameterize the FSTC by $\theta$ as discussed in Definition 1 is of great importance, and we now discuss it briefly. The inner product relation (13) dictates the mapping of context features onto expected losses with a reasonable degree of certainty. In order for such a relation to be satisfied for a realistic tracking scenario, the FSTC must be adequately described by the set of features $\{\varphi(o_i, w_j)\}$ $\forall i \in \mathcal{O}, j \in \mathcal{W}$. The function $\varphi$ maps observable quantities, such as estimated target position and SINR, to a vector describing the utility of waveform $w_i$. Additionally, a suitable weighting over these features must reliably predict the associated loss for each pair $(o, w)$.

We now describe the process for estimating a posterior distribution describing the radar’s uncertainty about the value of $\theta$ using the radar’s observations directly. The radar starts with a prior over possible FSTC parameter values $P(\theta)$. Using the observed data $\mathcal{F}_{k-1}$, the radar then updates its prior distribution to a posterior distribution by applying the Bayes rule

$$P(\theta|\mathcal{F}_{k-1}) = \frac{P(\mathcal{F}_{k-1}|\theta)P(\theta)}{P(\mathcal{F}_{k-1})} \propto P(\mathcal{F}_{k-1}|\theta)P(\theta).$$

Since $\theta$ is a fixed, continuous-valued parameter vector, we may reflect the radar’s uncertainty about its value by choosing a Gaussian prior distribution. Assuming a Gaussian likelihood function, computation of the posterior in (14) is straightforward and computationally efficient. A detailed computation of (14) may be found in Appendix A.

DEFINITION 3 (THOMPSON SAMPLING): Thompson sampling proceeds by selecting the waveform $w_{TS} \in \mathcal{W}$ such that

$$\int_{\Theta} \mathbb{I} \left[ E[\ell|w_{TS}, o, \theta] = \min_{w'} E[\ell|w', o, \theta] \right] P(\theta|\mathcal{F}_{k-1}) d\theta$$

where $\mathbb{I}[-]$ is the indicator function. Details of a practical algorithm to achieve this policy in the linear contextual bandit setting can be found in [20]. Details of a context-aware waveform selection strategy based on the Thompson Sampling can be found in [14].
DEFINITION 4 (BAYESIAN REGRET) The expected, or Bayesian, n-round regret of a decision strategy is given by

\[ BR^*_n = E \left[ \sum_{k=1}^{n} \ell_k - \sum_{k=1}^{n} \min_{w \in W} \langle \theta, w^* \rangle \right] \] (16)

which is an indication of the average performance of a decision algorithm. However, an algorithm with low Bayesian regret may perform poorly in particular scenarios.

We note that many Bayesian and frequentist bounds on the sample-efficiency of TS exist [20], [21], [72]. However, these bounds generally focus on the asymptotic case, where the effects of the prior distribution “wash out.” In practice, performance can be highly suboptimal if the prior is misspecified and the number of observations is limited, as in a target tracking problem.

THEOREM 1 (PRIOR DEPENDENT UPPER/LOWER BOUND FOR BAYESIAN BANDIT [13]) For any prior distribution \( P \) over the parameter space \( \Theta \), the Bayesian regret of an A-armed bandit satisfies

\[ BR^*_n(P) \leq C \sqrt{A \times n} \] (17)

where \( C > 0 \) is a universal constant. Furthermore, there exists a prior \( P \) such that

\[ BR^*_n(P) \geq c \sqrt{A \times n} \] (18)

where \( c > 0 \) is a universal constant. Thus, we see that \( \sup_P BR^*_n(P) = \Theta(\sqrt{A \times n}) \), where \( f(x) = \Theta(g(x)) \) denotes that \( f(x) \) is bounded both above and below by \( g(x) \) asymptotically.

Thus, we see that for the case of small time-horizon \( n \), the choice of prior can make a relatively large difference. For example, in [73], a lower bound on the Bayesian regret of TS is established to be \( \frac{1}{2\sqrt{A \times n}} \), along with a prior independent upper-bound of \( 14\sqrt{A \times n} \).

Since the sample-efficiency of TS over a limited time horizon is highly dependent on the choice of the prior distribution [73], [74], we discuss a Meta-Learning procedure for sequentially learning a prior over time, when the radar is embedded in an environment of tasks coming from a common task distribution. We will see that in many scenarios, Meta-Learning an effective prior for Thompson Sampling can provide major benefits in tracking performance. We now describe a Meta-Learning formulation in which uncertainty about the prior distribution is captured by modeling a distribution over instance priors.

THEOREM 2 (SIMPLE-TASK PAC-BAYES BOUND [75]) Let \( P \in \mathcal{M} \) be a prior distribution over the hypothesis space \( \Theta \). Then, for any \( \delta \in (0, 1] \), the following inequality holds uniformly for all posterior distributions \( \hat{Q} \in \mathcal{M} \) with probability at least \( 1 - \delta \)

\[ \text{er}(Q, D) \leq \widehat{\text{er}}(Q, S) + \sqrt{\frac{D_{\text{KL}}(Q||P) + \log m_s}{2(m_s - 1)}} \] (19)

where \( \text{er}(Q, D) \) is the expected loss when acting according to posterior distribution \( Q \) when the true underlying distribution is \( D \), \( \widehat{\text{er}}(Q, S) \) is the empirically observed loss when data \( S \sim D^{m_s} \) are observed, and \( m_s \) is the sample size.

REM 3 While appearing abstract, the two terms in the probability approximately correct (PAC)-Bayes bound illustrate a tradeoff between fitting the data and complexity. The first term encourages picking a \( Q \) that provides a low empirical error, whereas the second term emphasizes \( Q \) with a small distance from the prior \( P \). The choice of prior affects the tightness of the bound. In general, we want to select a prior, which is close to posteriors that give us a low empirical error.

B. Meta-Learning Formulation

In this section, we describe a Meta-Learning procedure, which aims to improve radar tracking performance using information gained via repeated interaction over \( m \) target tracking instances. Each individual tracking instance corresponds to a new instance of a \( n \)-step waveform-agile tracking problem on an FSTC described in Section II-A, where the radar has access to a catalog of \(|\mathcal{W}| = K\) waveforms.

The radar ultimately wishes to learn the true parameter vector for each tracking instance \( s \), denoted by \( \theta_s \in \mathbb{R}^d \). Given knowledge of \( \theta_s \), contexts can be directly mapped to average losses using the inner product relationship \( \ell_k = \langle \theta, \phi(s, w_s) \rangle + \eta_k \) for stochastic linear bandits. As candidates, we consider the class of distributions \( \mathcal{P} = \{ P_\theta : \theta \in \Theta \} \), where \( \Theta \) is a compact subset of \( \mathbb{R}^d \). Just as in the conventional learning setting, the radar is equipped with a prior distribution \( P(\theta) \), which expresses a subjective belief about the value of \( \theta \) based on the current sequence of observations \( \mathcal{F} \).

At the beginning of each tracking instance \( v \in \{1, 2, \ldots, n_{\text{track}}\} \), an instance of the learning problem, specified by \( \theta_{v,s} \), is sampled from a task distribution \( P_s \), which...
Algorithm 2: Contextual Meta-TS for Waveform Agile Tracking.

**Input** meta-prior distribution $Q$, Loss function $\ell$

Set $Q_1 \leftarrow Q$

for Each target track $v = 1, \ldots, N_{\text{track}}$ do

1. Sample $P_v \sim Q_v$;
2. Apply Thompson Sampling waveform selection policy (15) with prior $P_v$ to problem parameterized by $\theta_{v,*} \sim P_\star$ for $n$ CPIs;
3. Update Meta-Posterior $Q_{v+1}$ according to the update rules (22) and (23);
end

is fixed but unknown to the radar. Since $P_\star$ corresponds to the true distribution of tasks, each task being parameterized by $\theta_{v,*}$, it can also be interpreted as the best possible choice of prior for a Bayesian learning algorithm. Thus, the Meta-Learning process consists of estimating the true prior $P_\star$ sequentially interacting with problem instances that are assumed to be sampled i.i.d. from a common distribution $Q$.

The radar’s uncertainty about the true value of $P_\star$ is reflected by assuming it is sampled from a fixed distribution over instance priors $P_\star \sim Q$. We refer to $Q$ as a **metaprior**.

The Meta-Learning environment is then characterized by the pair $(P_\star, Q)$. We denote by $Q_v$, the **metaposterior**, which is the radar’s current estimate of the instance prior $P_v$ using information gained up to track $v$. Good performance will occur when $D_{KL}(Q_v||P_\star)$ is small, which corresponds to a low degree of uncertainty regarding the true prior $P_\star$. The metaposterior is estimated sequentially using the standard Bayesian update rule

$$Q_{v+1}(\hat{P}) \propto P(H_v|P_\star = \hat{P})Q_v(\hat{P}) = Q_v(\hat{P}) \int_0 P(H_v|\theta_{v,*} = \theta)P(\theta_{v,*} = \theta|P_\star = \hat{P})d\theta. \quad (20)$$

The Meta-Learning formulation is an example of a hierarchical Bayes model.

**IV. ALGORITHM AND PRACTICE**

At the heart of the waveform selection approach is posterior sampling [14, 20], which is also described for the waveform-agile target tracking problem in Algorithm 1. The Meta-Learning algorithm used in this article is described in Algorithm 2, which seeks to learn an effective prior distribution for the linear contextual bandit described in Algorithm 1.

To ensure that the metaposterior update (20) is computable in a closed form, and the existence of an efficient algorithm for online Meta-Learning, we consider a normal–normal conjugacy scenario, in which we study a Gaussian bandit with a Gaussian metaprior. The Gaussian–Gaussian bandit model is commonly used in the literature on sequential decision processes over continuous parameter spaces. We note that the parameter $\theta$ is assumed to be a fixed value over each target track, and the Gaussian model is used here to reflect the radar’s uncertainty about the true value of $\theta$. Thus, explicit assumptions about the environment are avoided. We assume a normal distribution over instance priors, expressed by $P(\theta) = \mathcal{N}(\mu, \sigma^2_0 I_d)$, where the noise level $\sigma^2_0$ is fixed.

The metaprior is then a distribution over instance prior means, given by $Q_v(\mu) = \mathcal{N}(0, \sigma^2_0 I_d)$. We assume the noise level $\sigma^2_0$ is fixed and known to the radar. The Meta-Learning process then maintains a metaposterior $Q_v(\mu) = \mathcal{N}(\mu_{0,v}, \Sigma_v)$. Due to the conjugacy properties of the normal distribution, the metaposterior $Q_v$ can be simply updated in a closed form, even for a contextual bandit algorithm. Following the standard computations for Bayesian multitask regression (reviewed in [76, Appendix D]), we see that the metaposterior can be simply expressed by

$$Q_v \sim \mathcal{N}(\mu_v, \Lambda^{-1}_v) \quad \text{(21)}$$

where the parameters are updated for each track $v$ by

$$\Lambda_v = \Lambda_{v-1} + X^T_v (\sigma^2 I + X_v \Sigma X^T_v)^{-1}X_v \quad \text{(22)}$$

$$\mu_v = \Lambda_v^{-1}(\Lambda_{v-1}\mu_{v-1} + X^T_v (\sigma I_n) + X_v \Sigma X^T_v)^{-1}L_v \quad \text{(23)}$$

where $\Sigma = \sigma^2_0 I_d$, $X_v = \{\varphi_{k=1}, \varphi_2, \ldots, \varphi_n\}$ is the vector of observed contexts over each time index $k = 1, 2, \ldots, n$ of the present track $v$, and $L_v = \{\ell_{k=1}, \ell_2, \ldots, \ell_n\}$ is a vector containing the observed losses during track $v$. The updates (22) and (23) are easily calculated for low-dimension cases, such as the waveform selection problem examined here. For high-dimensional problems, the Woodbury matrix identity can be applied to speed up computations. A description of the Meta-Learning process can be seen in Algorithm 2.

We note that an analytical characterization of the sample-efficiency of this algorithm is carried out in [76, Sec. 4.1].

We expect Meta-Learning to provide major benefits for cases where $\sigma^2_v \gg \sigma^2_0$, since uncertainty about the true prior $P_\star$ is significant. In cases where $\sigma^2_v \ll \sigma^2_0$, Meta-Learning is very close to the standard single-task learning problem, and we expect little benefit. Thus, the amount of relatedness between tasks is an important indicator of Meta-Learning performance. The issue of task similarity has been studied from an information-theoretic perspective in [77]. Future work in this domain could focus on a rigorous justification of task similarity for radar tasks of practical interest, such as tracking under interference.

**Theorem 3 (Extended PAC-Bayes Bound [55])** Let $Q$ be a learning algorithm, which maps observations and a prior to a posterior distribution on $\mathcal{M}$, and let $P$ be some predefined hyperprior distribution. Let $\tau_v$ be a new learning instance [55, Sec. 3.1]. Then, for any $\delta \in (0, 1)$, the following holds uniformly for the set of all hyperposterior

---

10This quantity is sometimes referred to as the **hyperprior** in the field of Bayesian statistics.
distributions \( Q \) with probability at least \( 1 - \delta \):

\[
\text{er}(Q, r_e) \leq \frac{1}{n_{\text{obs}}} \sum_{i=1}^{n_{\text{obs}}} \mathbb{E} \hat{r}_i(Q, S_i)
\]

Empirical multitask error

\[+
\frac{1}{n_{\text{obs}}} \sum_{i=1}^{n_{\text{obs}}} \sqrt{D_{\text{KL}}(Q \| P) + \mathbb{E}_{P \sim Q} D_{\text{KL}}(Q_i \| P) + \log \frac{2n_{\text{obs}} \delta}{2(m_i - 1)}}
\]

Avg. Task Complexity

\[+
\sqrt{D_{\text{KL}}(Q \| P) + \log \frac{2n_{\text{obs}} \delta}{2(n_{\text{obs}} - 1)}}
\]

Environment Complexity

(24)

where \( Q_i \triangleq Q(S_i, P) \).

Remark 4

The extended PAC-Bayes bound consists of the empirical multitask error plus two complexity terms. The first is the average of task complexity terms for the observed tasks. This term tends to zero as the number of samples in each task \( m_i \to \infty \).

V. SIMULATION STUDY

In this section, we examine the performance of the contextual meta-TS waveform selection approach presented in Algorithm 2. Meta-TS is compared with three baseline waveform selection strategies. The first is an “oracle,” which performs idealized Thompson Sampling. The oracle begins with knowledge of the underlying prior distribution. Additionally, we consider a Thompson sampling algorithm, which is initialized with an uninformative prior, namely \( P(\theta) \sim \mathcal{N}(0, I_d) \). We further consider a random waveform selection strategy, in which waveforms from the radar’s library are transmitted with uniform probability, and a rule-based adaptive strategy which modifies a linear FM waveform depending on the targets’ estimated positions. If the average of the estimated target positions is less than 3 km, the radar transmits a linear FM waveform with a bandwidth of 80 MHz. If the average estimated position is 3–5 km, the bandwidth of the LFM waveform is 40 MHz. If the average estimated position is greater than 5 km, the LFM bandwidth is 20 MHz.

We simulate a waveform-agile tracking scenario in which the radar engages in a temporal sequence of target tracks and adapts its waveform each CPI. Three targets are present in the scene during each tracking instance. Each target tracking instance is defined by a new FSTC parameter vector \( \theta \in \mathbb{R}^d \). We note that during each track, the channel is nonstationary, as the target and clutter impulse responses vary on a CPI-to-CPI basis. However, the mean values of these impulse responses remain stationary. Thus, the weighting vector \( \theta \in \mathbb{R}^d \) is constant during the track.

In the simulations, the radar undergoes a sequence of \( N_{\text{track}} = 80 \) target tracks. Each target track lasts for \( n = 500 \) CPIs, and each CPI consists of 256 coherent pulses. Assuming a pulse repetition frequency of 10 kHz, each track lasts for approximately 12.8 s. The radar has access to a waveform library with \( A = 50 \) waveforms. The waveform library consists of 10 LFM chirp pulses with a varying chirp rate, 10 Zadoff-Chu phase-coded pulses, 10 Frank-coded pulses, 10 Barker-coded pulses, and 10 exponential FM pulses with a variable parameter \( \alpha \). The radar measurements of each target are filtered using a Kalman filter after each CPI.

At the beginning of each track, the initial position, velocity, and trajectory of each of the three targets is randomized. The target frequency responses \( h_i(f) \) and the clutter frequency responses \( c_i(f) \) are Gaussian processes, and vary from CPI-to-CPI. The mean frequency response for each target is Gaussian shaped, and is expressed by

\[
\mathbb{E}[h_i(f)] = \exp\left(-\frac{(f - f_0)^2}{2\gamma_0^2}\right)
\]

where \( f_0 \) is a target specific parameter that defines the spatial extent of the target. The mean clutter frequency responses are similarly defined by

\[
\mathbb{E}[c_i(f)] = \exp\left(-\frac{(f - f_0)^2}{2\gamma_0^2}\right)
\]

where \( \gamma_0 \) determines the spatial extent of clutter component \( h \).

These simulation conditions, along with the loss function and context representation, determine the parameter vector \( \theta \in \mathbb{R}^d \), which the radar wishes to learn. The vector \( \theta \) represents a weighting over context features in order to predict the loss associated with each waveform. This underlying structure from track-to-track is what the Meta-Learning algorithm exploits to effectively bias the algorithm. Over the course of the track, the radar uses its current estimate of the target’s position, as well as an estimate of the channel quality, in order to select the best waveform for the next CPI. The loss function used by the radar is

\[
\ell_k = \min \left\{ \max \left[ \frac{\text{SINR}_{\text{mean}}}{\text{SINR}_{\text{target}}}, 0 \right], 1 \right\}
\]

(27)

where \( \text{SINR}_{\text{mean}} \) is the SINR, averaged across each of the targets measured from the range-Doppler response. This is computed by performing CFAR detection, clustering the detections using the DBSCAN clustering algorithm [78], and averaging the power over each of the target clusters. Let \( RD_{\text{DB}} \) be the range-Doppler map in dB units, indexed by range and Doppler cells, respectively. Then, \( \text{SINR}_{\text{mean}} \) is computed as follows:

\[
\text{SINR}_{\text{mean}} = \frac{1}{|T|} \sum_{i,j \in T} RD_{\text{DB}}(i,j) - \frac{1}{|RD|} \sum_{i,j} RD_{\text{DB}}(i,j)
\]

(28)

where \( T \) is the set of target bins, established using a prior estimate on the targets’ positions, followed by constant-false alarm rate detection and clustering, which is performed using the DBSCAN algorithm [78].

The term \( \text{SINR}_{\text{target}} \) is a target SINR value, which roughly corresponds to the highest SINR the radar can expect to achieve during a given CPI. In these simulations, the target SINR is 22 dB. This loss function is effective since it is bounded from above and below and directly incorporates measurement quality. Furthermore, it is reasonable
to assume that a target SINR value will be approximately known for most tracking applications.

The set of context vectors \( \varphi(w_i, o_j) \) is defined by computing

\[
\varphi(w_i, o_j) = \left[ \mathbb{E}[\ell_i(w_i)|o_j], \text{Var}[\ell_i(w_i)|o_j], \max_{m < k} \mathbb{E}[\ell_m(w_i)|o_j] \right]
\]

for each \( w_i \in \mathcal{W} \) and \( o_j \in \mathcal{O} \). We note that these context features will vary from one tracking interval to another, depending on the target and clutter frequency responses and trajectories. These changes induce the variation over the parameter vector \( \theta \).

In Fig. 1, we examine the outage behavior of each of the waveform selection strategies over the sequence of \( m = 80 \) tracks. We observe that the meta-TS algorithm quickly approximates the true prior and performs nearly as effectively as the oracle both terms of average SINR, average tracking error, weakest target SINR, and maximum tracking error. Thus, it can be maintained that knowledge of the underlying prior drastically improves outage behavior, especially over the limited time horizon of 500 CPIs each track. We also observe that random waveform selection serves the worst performing baseline, with outage events occurring in 30%–45% of all transmissions. The meta-TS approach generally results in a 30%–40% reduction outage probability as compared to random waveform selection. We note that the uninformed TS strategy and the rule-based adaptive LFM strategy perform similarly, due to the radar’s inability to quickly infer \( \theta \) in a short time horizon using an uninformative prior. However, we note that the adaptive LFM approach performs noticeably worse in terms of weakest target SINR and maximum tracking error.

Fig. 2 shows the outage performance, in terms of the previously considered performance statistics, over the entire sequence of 80 tracks. We observe once more that all posterior sampling techniques enjoy a large performance benefit over random waveform selection, which consistently selects waveforms that are poorly matched to the composite target/clutter response of the channel. Additionally, we see that the adaptive LFM strategy performs comparatively poorly in terms of tracking error for the weakest target response.
Fig. 2. **Cumulative Outage Behavior** of each waveform selection algorithm over a sequence of $N_{\text{track}} = 80$ target tracks, each lasting for $n = 500$ CPIs. Posterior sampling results in a drastic cumulative performance improvement over random waveform selection. We observe that the most substantial performance improvements as a result of learning the best prior occur in the regime of $5 \text{ dB} \leq \text{SINR} \leq 15 \text{ dB}$ and $60 \text{ m} \leq \text{tracking error} \leq 40 \text{ m}$. Learning approaches provide the most substantial benefits over rule-based LFM selection in terms of weakest target performance.

Fig. 3. **Average KL Divergence** between the metaposterior $Q_v$ and the true prior $P^*$. After 20 tracks, the Meta-Learning algorithm has closely approximated the oracle.

In Fig. 3, we observe the KL divergence between the metaposterior $Q_v$ and the best prior $P_v$. After just a few tracks, the metaposterior concentrates toward the true prior in terms of KL divergence, and very closely approximates the oracle after 20 tracks. Thus, we see that the metalearner
is converging toward a distribution that is in line with the true data. We note that the KL divergence does not necessarily need to go to zero in order to achieve acceptable radar performance. The metalearner’s aim is to learn a prior that is effective over each of the possible FSTCs the radar may experience, which is generally feasible even with a small number of examples. Explicitly estimating the true prior, on the other hand, may require far more samples.

In Fig. 4, we observe the cumulative strong regret experienced by each waveform selection algorithm. This performance measure is widely used in the online learning literature and gives a sense of how closely a decision rule follows the best possible strategy under the given problem formulation. We see that when the true prior is known, the radar incurs regret at an approximately constant rate, which is expected given that the true prior is suitable for all FSTCs in the class. Similarly, when an uninformative prior is used, the radar also incurs regret at a nearly constant rate, but much faster given that more exploration is needed to learn the dynamics of the new FSTC during each track. When Meta-Learning is applied, the radar incurs regret at a decreasing rate over time as the true prior is approximated by the meta-TS algorithm. We see that each of the contextual bandit algorithms significantly outperforms the naive random selection strategy, and more narrowly the adaptive LFM strategy, as expected.

In Fig. 5, behavior over a single track is isolated. We examine performance during the 80th and final target track for each waveform selection algorithm. In particular, we see the cumulative SINR averaged across the three targets for each algorithm on the left side of the figure. We observe that the meta-TS algorithm performs close to the oracle throughout the entire track duration. The uninformed TS algorithm performs significantly worse over the entirety the track, due to the short time horizon, which makes learning without a strong prior difficult. In order to learn effectively with an uninformative prior, the number of waveforms in the catalog would need to be reduced. Thus, we find that Meta-Learning has greatly improved the learning efficiency in the cases tested. The same conclusion can also be drawn from the empirical distribution function of the average target SINR seen on the right side of the figure, from which it is observed that Meta-Learning provides substantial gains in the $5 \leq \text{SINR} \leq 15$ dB regime once again.

VI. CONCLUSION

This work has presented an online Meta-Learning approach for waveform-agile radar tracking. This approach allows the radar to learn transmission policies from a limited number of interactions by exploiting inherent similarities in the physical scene from track-to-track, when an underlying structure is present, such as similarity in target type and trajectory. We have herein defined the waveform selection problem as a temporal sequence of Bayesian learning tasks, where each learning instance corresponds to a new target tracking interval. Each tracking interval, the radar sequentially chooses waveforms based on an estimate of the posterior distribution of an FSTC coming from a fixed class. We have observed that misspecification of the prior distribution can cause significant performance degradation, motivating the study of algorithms, which can automatically learn a prior that works well across the entire class of channels.

In essence, the radar learns about the underlying structure of the target and clutter frequency responses and target trajectories, which carry over across tracking intervals. This information is used to bias the learning process toward certain model parameter values so that less exploration steps or highly suboptimal transmissions are required to develop a transmission policy. Some broader potential implications for target tracking are as follows.

1) When exploration early in a tracking interval is reduced, there is a significantly lower occurrence of “outage” events, which may result in a lost track.

2) An initial bias may allow for the use of larger waveform catalogs, which can provide more diversity for...
challenging target identification and tracking applications.

3) High-level bias information could potentially be communicated and used to inform other radars or processes in a distributed network.

Thus, the general problem of bias learning has the potential for direct practical applications, especially in terms of reducing worst case performance when the scene has fundamental characteristics, which remain relatively stationary over time. However, we note that Meta-Learning may not be appropriate for all radar tracking applications, and robust systems are expected to use an array of rule-based and learning approaches to reduce worst case performance in unexpected conditions. Future work could focus on developing more explicit connections between the physical signal model and probabilistic FSTC model to identify cases in which Meta-Learning would be most beneficial.

Herein, we have shown that the problem can be effectively formulated in terms of Bayesian learning, for which a computationally feasible contextual bandit algorithm is developed. This approach is satisfying as the online waveform selection naturally fits well within the contextual bandit framework. We have shown that when the target and clutter impulse responses have underlying similarities in structure, learning a bias over the model parameter can drastically reduce the time required to learn a waveform selection strategy, which is near-optimal with respect to the chosen loss function.

However, the Bayesian interpretation of Meta-Learning is not exclusive, and the problem can also be approached from a functional viewpoint [18]. Future research comparing the Bayesian and functional interpretations would be of value, especially the possible identification of structural properties that may simplify the learning problem. Additionally, an investigation into the sharing of bias information in a distributed network is an avenue for future work. Knowledge of spatial correlation structure could significantly speed up bias learning.

APPENDIX A

COMPUTATION OF THE POSTERIOR \( P(\theta_k | F_{k-1}) \)

Let the prior distribution be \( \mathcal{N}(\delta_{\theta_k}, B_{\theta_k}^{-1}) \), where \( \delta_{\theta_k} \in \mathbb{R}^d \) and \( B_{\theta_k} \in \mathbb{R}^{d \times d} \) are the mean vector and covariance matrix of a multivariate normal distribution. Furthermore, let \( \ell_k(w_i) \) be the loss associated with transmitting waveform \( w_i \) at step \( k \). Then, we may compute the posterior distribution as

\[
P(\theta_k \mid \ell_k(w_i)) \propto P(\ell_k(w_i) \mid \theta) P(\theta) \]

\[
\propto \exp \left\{ -\frac{1}{2} \left[ (\ell_k(w_i) - \theta^T \phi_{i,k})^2 \right] \right\} 
\]

\[
\exp \left\{ -\frac{1}{2} \left[ \ell_k(w_i)^2 + \theta^T \phi_{i,k} \phi_{i,k}^T \theta \right] \right\} 
\]

\[
\propto \exp \left\{ -\frac{1}{2} (\theta - \hat{\theta}_{k+1})^T B_{k+1} (\theta - \hat{\theta}_{k+1}) \right\}
\]

\[ (33) \]

\[ (34) \]

\[ (35) \]

REFERENCES

[1] C. E. Thornton, R. M. Buehrer, and A. F. Martone, “Online Meta-Learning for scene-diverse waveform-agile radar target tracking,” in *Proc. IEEE Radar Conf.*, 2022, pp. 1–6.

[2] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, “The rise of radar for autonomous vehicles: Signal processing solutions and future research directions,” *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 20–31, Sep. 2019.

[3] K. V. Mishra, M. R. B. Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, “Toward millimeter-wave joint radar communications: A signal processing perspective,” *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 100–114, Sep. 2019.

[4] A. C. Aydogdu et al., “Radar interference mitigation for automated driving: Exploring proactive strategies,” *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 72–84, Jul. 2020.

[5] A. Hassanien, M. G. Amin, E. Aboutanios, and B. Himed, “Dual-function radar communication systems: A solution to the spectrum congestion problem,” *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 115–126, Sep. 2019.

[6] S. Haykin, Y. Xue, and P. Setoodeh, “Cognitive radar: Step toward bridging the gap between neuroscience and engineering,” *Proc. IEEE*, vol. 100, no. 11, pp. 3102–3110, Nov. 2012.

[7] 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.

[8] M. S. Greco, F. Gini, P. Stinco, and K. Bell, “Cognitive radars: On the road to reality: Progress thus far and possibilities for the future,” *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 112–125, Jul. 2018.

[9] A. F. Martone, “Cognitive radar demystified,” *URSI Radio Sci. Bull.*, vol. 2014, no. 350, pp. 10–22, 2014.

[10] 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.

[11] S. Z. Gurbuz, H. D. Griffiths, A. Charlish, M. Rangaswamy, M. S. Greco, and K. Bell, “An overview of cognitive radar: Past, present, and future,” *IEEE Aerosp. Electron. Syst. Mag.*, vol. 34, no. 12, pp. 6–18, Dec. 2019.

[12] B. Hajek and M. Rangaswamy, “Course notes, ECE 543: Statistical learning theory,” Univ. Illinois Urbana-Champaign, Champaign, IL, USA, 2021. [Online]. Available: http://maxim.ece.illinois.edu/teaching/SLT/SLT.pdf

[13] C. Szepesvari and T. Lattimore, *Bandit Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2020.

[14] C. E. Thornton, R. M. Buehrer, and A. Martone, “Constrained contextual bandit learning for adaptive radar waveform selection,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 2, pp. 1133–1148, Apr. 2022.

[15] O. Bousquet and A. Elisseeff, “Stability and generalization,” *J. Mach. Learn. Res.*, vol. 2, pp. 499–526, 2002.

[16] X. Ying, “An overview of overfitting and its solutions,” *J. Phys.: Conf. Series*, vol. 1168, no. 2, 2019, Art. no. 022022.

[17] C. Finn, A. Rajeswaran, S. Kakade, and S. Levine, “Online meta-learning,” in Proc. Int. Conf. Mach. Learn., 2019, pp. 1920–1930.

[18] J. Baxter, “Theoretical models of learning to learn,” in *Learning to Learn*. Berlin, Germany: Springer, 1998, pp. 71–94.

[19] V. Vapnik, *Estimation of Dependencies Based on Empirical Data*. Berlin, Germany: Springer, 2006.
[20] S. Agrawal and N. Goyal, “Thompson sampling for contextual bandits with linear payoffs,” in Proc. Int. Conf. Mach. Learn., 2013, pp. 127–135.

[21] D. J. Russo et al., “A tutorial on Thompson sampling,” Found. Trends Mach. Learn., vol. 11, no. 1, pp. 1–96, 2018.

[22] S. Kay, “Optimal signal design for detection of Gaussian point targets in stationary Gaussian clutter/reverberation,” IEEE J. Sel. Topics Signal Process., vol. 1, no. 1, pp. 31–41, Jun. 2007.

[23] R. A. Romero, J. Bae, and N. A. Goodman, “Theory and application of SNR and mutual information matched illumination waveforms,” IEEE Trans. Aerosp. Electron. Syst., vol. 47, no. 2, pp. 912–927, Apr. 2011.

[24] S. Pillai, H. Oh, D. Youla, and J. Guerci, “Optimal transmit-receiver design in the presence of signal-dependent interference and channel noise,” IEEE Trans. Inf. Theory, vol. 46, no. 2, pp. 577–584, Mar. 2000.

[25] R. Calderbank, S. D. Howard, and B. Moran, “Waveform diversity in radar signal processing,” IEEE Signal Process. Mag., vol. 26, no. 1, pp. 32–41, Jan. 2009.

[26] S. D. Blunt and E. L. Mokole, “Overview of radar waveform diversity,” IEEE Aerosp. Electron. Syst. Mag., vol. 31, no. 11, pp. 2–42, Nov. 2016.

[27] W. Roberts, H. He, J. Li, and P. Stoica, “Probing waveform synthesis and receiver filter design,” IEEE Signal Process. Mag., vol. 27, no. 4, pp. 99–110, Jul. 2010.

[28] J. J. Benedetto, I. Konstantinidis, and M. Rangaswamy, “Phase-coded waveforms and their design,” IEEE Signal Process. Mag., vol. 26, no. 1, pp. 22–31, Jan. 2009.

[29] F. Giin, A. De Maio, and L. Patton, Waveform Design and Diversity for Advanced Radar Systems. London, U.K.: IET, 2012.

[30] N. Levanon and E. Mozeson, Radar Signals. Hoboken, NJ, USA: Wiley, 2004.

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

[32] Z. Zhu, S. Kay, and R. S. Raghavan, “Information-theoretic optimal radar waveform design,” IEEE Signal Process. Lett., vol. 24, no. 3, pp. 274–278, Mar. 2017.

[33] Z. Zhu, S. Kay, and R. S. Raghavan, “Locally optimal radar waveform design for detecting doubly spread targets in colored noise,” IEEE Signal Process. Lett., vol. 25, no. 6, pp. 833–837, Jun. 2018.

[34] M. M. Naghsh, M. Soltanalian, P. Stoica, M. Modarres-Hashemi, A. D. Maio, and A. Aubry, “A Doppler robust design of transmit sequence and receive filter in the presence of signal-dependent interference,” IEEE Trans. Signal Process., vol. 62, no. 4, pp. 772–785, Feb. 2014.

[35] S. M. Karbasi, A. Aubry, A. D. Maio, and M. H. Bastani, “Robust transmit code and receive filter design for extended targets in clutter,” IEEE Trans. Signal Process., vol. 63, no. 8, pp. 1965–1976, Apr. 2015.

[36] L. Wu, P. Babu, and D. P. Palomar, “Transmit waveform/receive filter design for MIMO radar with multiple waveform constraints,” IEEE Trans. Signal Process., vol. 66, no. 6, pp. 1526–1540, Mar. 2018.

[37] D. Cochran, S. Suvorova, S. D. Howard, and B. Moran, “Waveform libraries,” IEEE Signal Process. Mag., vol. 26, no. 1, pp. 12–21, Jan. 2009.

[38] S. P. Sira, Y. Li, A. Papandreou-Suppappola, D. Morrell, D. Cochran, and M. Rangaswamy, “Waveform-agile sensing for tracking,” IEEE Signal Process. Mag., vol. 26, no. 1, pp. 53–64, Jan. 2009.

[39] D. J. Kershaw and R. J. Evans, “Optimal waveform selection for tracking systems,” IEEE Trans. Inf. Theory, vol. 40, no. 5, pp. 1536–1550, Sep. 1994.

[40] D. J. Kershaw and R. J. Evans, “Waveform selective probabilistic data association,” IEEE Trans. Aerosp. Electron. Syst., vol. 33, no. 4, pp. 1180–1188, Oct. 1997.

[41] S. P. Sira, A. Papandreou-Suppappola, and D. Morrell, “Dynamic configuration of time-varying waveforms for agile sensing and tracking in clutter,” IEEE Trans. Signal Process., vol. 55, no. 7, pp. 3207–3217, Jul. 2007.

[42] C. O. Savage and B. Moran, “Waveform selection for maneuvering targets within an IMM framework,” IEEE Trans. Aerosp. Electron. Syst., vol. 43, no. 3, pp. 1205–1214, Jul. 2007.

[43] T. Zhang, G. Cui, L. Kong, W. Yi, and X. Yang, “Phase-modulated waveform evaluation and selection strategy in compound-Gaussian clutter,” IEEE Trans. Signal Process., vol. 61, no. 5, pp. 1143–1148, Mar. 2013.

[44] A. Aubry, A. DeMaio, A. Farina, and M. Wicks, “Knowledge-aided (potentially cognitive) transmit signal and receive filter design in signal-dependent clutter,” IEEE Trans. Aerosp. Electron. Syst., vol. 49, no. 1, pp. 93–117, Jan. 2013.

[45] B. Scala, M. Rezaeian, and B. Moran, “Optimal adaptive waveform selection for target tracking,” in Proc. Int. Conf. Inf. Fusion, 2005, p. 6.

[46] A. Charlish and F. Hoffmann, “Anticipation in cognitive radar using stochastic control,” in Proc. IEEE Radar Conf., 2015, pp. 1692–1697.

[47] E. Selvi, R. M. Buehrer, A. Martone, and K. Sherbondy, “Reinforcement learning for adaptable bandwidth tracking radars,” IEEE Trans. Aerosp. Electron. Syst., vol. 56, no. 5, pp. 3904–3921, Oct. 2020.

[48] C. E. Thornton et al., “Deep reinforcement learning control for radar detection and tracking in congested spectral environments,” IEEE Trans. Cogn. Commun. Netw., vol. 6, no. 4, pp. 1335–1349, Dec. 2020.

[49] C. E. Thornton, R. M. Buehrer, H. S. Dhillon, and A. F. Martone, “Universal learning waveform selection strategies for adaptive target tracking,” IEEE Trans. Aerosp. Electron. Syst., vol. 58, no. 6, pp. 5798–5814, Dec. 2022.

[50] W. W. Howard, C. E. Thornton, A. F. Martone, and R. M. Buehrer, “Multi-player bandits for distributed cognitive radar,” in Proc. IEEE Radar Conf., 2021, pp. 1–6.

[51] A. O. Hero, D. Castañón, D. Cochran, and K. Kastella, Foundations and Applications of Sensor Management. Berlin, Germany: Springer, 2007.

[52] V. Krishnamurthy and R. J. Evans, “Hidden Markov model multiarm bandits: A methodology for beam scheduling in multitaget tracking,” IEEE Trans. Signal Process., vol. 49, no. 12, pp. 2893–2908, Dec. 2001.

[53] S. Thrun, “Is learning the n-th thing any easier than learning the first?,” in Proc. Adv. Neural Inf. Process. Syst., 1996, pp. 640–646.

[54] J. Baxter, “A Bayesian/information theoretic model of learning to learn via multiple task sampling,” Mach. Learn., vol. 28, pp. 7–39, 1997.

[55] R. Amit and R. Meir, “Meta-learning by adjusting priors based on extended PAC-Bayes theory,” Proc. Int. Conf. Mach. Learn., 2018, pp. 205–214.

[56] O. Simeone, S. Park, and J. Kang, “From learning to meta-learning: Reduced training overhead and complexity for communication systems,” in Proc. 6G Wireless Summit, 2020, pp. 1–5.

[57] S. Park, H. Jang, O. Simeone, and J. Kang, “Learning to demodulate from few pilots via offline and online meta-learning,” IEEE Trans. Signal Process., vol. 69, pp. 226–239, 2021.

[58] A. E. Kalor, O. Simeone, and P. Popovski, “Prediction of mmWave/Terahertz link blockages through meta-learning and recurrent neural networks,” IEEE Wireless Commun. Lett., vol. 10, no. 12, pp. 2815–2819, Dec. 2021.

[59] Y. Hu, M. Chen, W. Saad, H. V. Poor, and S. Cui, “Distributed multi-agent meta learning for trajectory design in wireless drone networks,” IEEE J. Sel. Areas Commun., vol. 39, no. 10, pp. 3177–3192, Oct. 2021.

[60] W. Jiang, A. M. Haimovich, M. Govoni, T. Garner, and O. Simeone, “Fast data-driven adaptation of radar detection via meta-learning,” in Proc. 55th Asilomar Conf. Signals, Syst., Comput., 2021, pp. 1618–1622.

[61] M. R. Bell, “Information theory and radar: Mutual information and the design and analysis of radar waveforms and systems,” Ph.D. dissertation, Elect. Eng., California Inst. Technol., Pasadena, CA, USA, 1988.
[62] H. L. Van Trees, Detection, Estimation, and Modulation Theory: Part III. Radar-Sonar Signal Processing and Gaussian Signal in Noise. Hoboken, NJ, USA: Wiley, 2001.

[63] S. Kashyap, J. Stanier, G. Painchaud, and A. Louie, “Radar response of missile-shaped targets,” in Proc. IEEE Antennas Propag. Soc. Int. Symp., 1995, pp. 1910–1913.

[64] K. M. Cuomo, J. E. Pion, and J. T. Mayhan, “Ultrawide-band coherent processing,” IEEE Trans. Antennas Propag., vol. 47, no. 6, pp. 1094–1107, 1999.

[65] N. Jiang, R. Wu, and J. Li, “Super resolution feature extraction of moving targets,” IEEE Trans. Aerosp. Electron. Syst., vol. 37, no. 3, pp. 781–793, Jul. 2001.

[66] A. Gopalan, S. Mannor, and Y. Mansour, “Thompson sampling for complex online problems,” in Proc. Int. Conf. Mach. Learn., 2014, pp. 100–108.

[67] W. R. Thompson, “On the likelihood that one unknown probability exceeds another in view of the evidence of two samples,” Biometrika, vol. 25, no. 3/4, pp. 285–294, 1933.

[68] S. Bubeck and C.-Y. Liu, “Prior-free and prior-dependent regret minimization,” Information- theoretic bounds and information meta-risk minimization,” in Proc. IEEE Radar Conf., 2020, pp. 1–6.

[69] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, “DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN,” ACM Trans. Database Syst., vol. 42, no. 3, pp. 1–21, 2017.

[70] D. A. McAllister, “PAC-Bayesian model averaging,” in Proc. Adv. Neural Inf. Process. Syst., 2021, pp. 26382–26394.

[71] B. Kveton et al., “Meta-Thompson sampling,” in Proc. Int. Conf. Mach. Learn., 2021, pp. 5884–5893.

[72] G. Durisi, “Transfer meta-learning: Information-theoretic bounds and information meta-risk minimization,” IEEE Trans. Inf. Theory, vol. 68, no. 1, pp. 474–501, Jan. 2022.

[73] Charles E. Thornton, Antonio, TX, USA), with The Pennsylvania State University, State College, PA, USA, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, and Bowie State University, Bowie, MD, USA. Since joining ARL, he has authored more than 140 journal and conference publications, two book chapters, ten patents, drafted five new spectrum sharing standards for the IEEE 686 Radar Standards document, and is also the Director of Wireless@Virginia Tech, a comprehensive research group focusing on wireless communications, radar, and localization. During 2009, he was a visiting Researcher with the Laboratory for Telecommunication Sciences (LTS), a federal research lab, which focuses on telecommunication challenges for national defense. While with LTS, his research focus was in the area of cognitive radio with a particular emphasis on statistical learning techniques. He has authored or coauthored more than 80 journals and approximately 250 conference papers and holds 18 patents in the area of wireless communications. His current research interests include machine learning for wireless communications and radar, geolocation, position location networks, cognitive radio, cognitive radar, electronic warfare, dynamic spectrum sharing, communication theory, multi-input multi-output radar systems, spread spectrum, interference avoidance, and propagation modeling.

Richard M. Buehler (Fellow, IEEE) received Virginia Tech from Bell Labs as an Assistant Professor with the Bradley Department of Electrical and Computer Engineering in 2001. He is currently a Professor of Electrical Engineering and is also the Director of Wireless@Virginia Tech, a comprehensive research group focusing on wireless communications, radar, and localization. During 2009, he was a visiting Researcher with the Laboratory for Telecommunication Sciences (LTS), a federal research lab, which focuses on telecommunication challenges for national defense. While with LTS, his research focus was in the area of cognitive radio with a particular emphasis on statistical learning techniques. He has authored or coauthored more than 80 journals and approximately 250 conference papers and holds 18 patents in the area of wireless communications. His current research interests include machine learning for wireless communications and radar, geolocation, position location networks, cognitive radio, cognitive radar, electronic warfare, dynamic spectrum sharing, communication theory, multi-input multi-output radar systems, spread spectrum, interference avoidance, and propagation modeling.

Anthony F. Martone (Fellow, IEEE) received the B.S. (summa cum laude) degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, USA, in 2001, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 2007. In 2007, he joined the U.S. Army Research Laboratory (ARL), Adelphi, MD, USA, as a Researcher in the RF Signal Processing and Modeling branch, where his research interests include radar, cognitive radar, sensing through the wall technology, spectrum sharing, and radar signal processing. He established himself as a subject matter expert (SME) for radar spectrum sharing and nonlinear radar research within ARL and academic communities. He is currently leading multiple Cognitive Radar research investigations at ARL, where he is overseeing, directing, and collaborating with multiple universities to address spectrum sharing for radar and communication systems, software-defined transceiver control, and adaptive processing techniques. He was a Ph.D. Committee Member for eight graduate students with The Pennsylvania State University, State College, PA, USA, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, and Bowie State University, Bowie, MD, USA. Since joining ARL, he has authored more than 140 journal and conference publications, two book chapters, ten patents, drafted five new spectrum sharing standards for the IEEE 686 Radar Standards document, provided the Plenary Presentation at the 2022 IEEE Radar Conference (New York, NY, USA).

Dr. Martone was elevated to IEEE Fellow in 2023 for contributions to the development and validation of cognitive radar systems. Since 2017, he has been an Associate Editor for IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS and has been on the Spectral Innovation, Standards, and Publications Committees of IEEE Aerospace and Electronic Systems Society Radar Systems Panel since 2019. He was the recipient of the Commanders Award for Civilian Service in 2011 for his research and development of sensing through the wall signal processing techniques. He was the General Co-Chair of the 2023 IEEE Radar Conference (San Antonio, TX, USA).