Cognitive beamspace algorithm for integrated sensing and communications

Petteri Pulkkinen*†, Majdoddin Esfandiari†, and Visa Koivunen†
†Information and Communications Engineering, Aalto University, Espoo, Finland
*Saab Finland Oy, Helsinki, Finland
Email: petteri.pulkkinen@aalto.fi, majdoddin.esfandiari@aalto.fi, visa.koivunen@aalto.fi

***Abstract*—This paper presents a novel cognitive beamspace algorithm for integrated sensing and communications (ISAC) systems, focusing on the optimization of spatial resources. The proposed method operates in the beamspace domain that enables an orthogonal design between the sensing and communication functions. Additionally, the approach leverages the principles of Thompson sampling, known for effectively balancing exploration and exploitation in uncertain environments. It enables the ISAC system to dynamically adjust radar target search strategies based on environmental feedback while maintaining acceptable communication rates with the user equipment. We demonstrate the algorithm's effectiveness over traditional methods through numerical simulations. These simulations reveal significant improvements in radar search functionality while meeting the communications quality of service constraints.**

***Index Terms*—Integrated sensing and communications, cognitive radar, beamspace algorithm, Thompson sampling, spatial resource optimization**

I. Introduction

Integration of radar and communication functionalities into unified systems, known as ISAC systems, is one of the key new technologies envisioned for the emerging 6G wireless communications [1], [2]. By merging these traditionally distinct functionalities, ISAC systems offer a promising avenue toward more efficient use of the hardware and spectrum resources. However, ISAC systems introduce new challenges, particularly in resource management and simultaneously optimizing the performance of the two functionalities [3].

This paper considers the optimization of spatial resources in an ISAC system. Transmit beamformer designs for ISAC systems have been considered in several works, which can be classified into three main categories of *radar-centric*, *communications-centric*, and *joint* designs [4]. The objective of the radar-centric designs is to add communications functionality to radar systems [5]–[7], whereas the communications-centric designs utilize communications signals not only for transmitting data through some communications links but also for obtaining information about targets present in environment [8]. The joint designs view both communications and sensing tasks as one functionality, leading to more general designs [9]. For instance, two beamforming algorithms were proposed in [10]. One allocates separate sub-arrays of the base station (BS) antennas to sensing and communications functionalities while the other uses all antennas for both functions. It was shown in [10] that the latter configuration outperforms the former. In [11], the transmit beamforming has been designed by optimizing the desired beam pattern while constraining the signal-to-interference-plus-noise ratio (SINR) at the communications user side. Utilizing backscattered radar and communications signals for detecting targets enhances the power efficiency of the developed ISAC system in [11]. The joint optimization of radar spatial spectrum matching error and the spectral efficiency for communications has been considered in [12] for beamforming design in wideband ISAC systems.

In this paper, we develop a new cognitive resource allocation method which enhances the radar search functionality of an ISAC system while providing an acceptable communications rate to the user equipments (UEs). The concept of cognitive radar [13], [14] refers to fully adaptive radar systems capable of learning and making decisions based on environmental closed-loop feedback. The proposed method is based on a design in a beamspace domain [15] that facilitates a convenient orthogonal model of the sensing and communications tasks. It also enables us to develop reinforcement learning methods based on multi-armed bandit (MAB) framework to address the exploration-exploitation tradeoff [16]. Particularly, the proposed beam allocation algorithm stems from Thompson sampling [17]. Thompson sampling is known for its efficacy in balancing exploration and exploitation and can be adapted to the unique requirements of ISAC systems. Combining ISAC with cognitive radar principles and Thompson sampling results in dramatically boosting the radar search functionality.

II. Problem Formulation

Consider an ISAC BS operating in an environment containing Q radar targets and, for the sake of clarity of the derivation, a single UE. The model and method can be extended to multi-user scenarios. The BS is equipped with an uniform linear array (ULA) of N antenna elements, and the UE is equipped with a single omnidirectional antenna. The steering vector $\mathbf{u}(\theta)$ of the ULA is written as

$$[\mathbf{u}(\theta)]_n = e^{-j(n-1)\pi \sin(\theta)} \; \forall \; n \in [N] \tag{1}$$

where $[N] := \{1, \ldots, N\}$.

The task is to guarantee a desired downlink (DL) communications rate to the user while simultaneously searching for radar targets. For the communications task, we assume perfect channel state information (CSI) estimated either from channel feedback or utilizing the channel reciprocity. For the target search task, it is assumed that the number of targets and their directions are initially unknown. We propose a closed-loop cognitive algorithm [13] for the radar search task based on an active Sense-Learn-Adapt (SLA) cycle.

A. Signal Model

We consider the Swerling model for the target scattering coefficient such that the target complex gain $\alpha_q \, \forall \, q \in [Q]$, including all the losses such as propagation loss and the radar cross section (RCS), obeys zero-mean circular Gaussian distribution $\alpha_q \sim \mathcal{CN}(0, \sigma^2_{\alpha,q})$. Thus, the monostatic sensing channel can be written as follows

$$\mathbf{H}_r(t,\tau) = \sum_{q=1}^{Q} \alpha_q^r \mathbf{u}(\theta_q^r)\mathbf{u}^{\mathrm{H}}(\theta_q^r) e^{j2\pi\nu_q t}\delta(\tau - \tau_q) \quad (2)$$

where θ_q^r is the target angle with respect to the boresight, ν_q is the target radial velocity, and τ_q is the two-way propagation delay to the target.

Let us consider that the radar signal s_r and the DL communications symbol s_c are transmitted simultaneously. Thus, the overall transmit signal of the ISAC system can be written as

$$\mathbf{x}(t) = a(t)\sqrt{P_{\mathrm{tot}}}(s_r \mathbf{w}_r + s_c \mathbf{w}_c) \quad (3)$$

where $a(t)$ is the sub-pulse shape and P_{tot} is the total power as well as $\mathbf{w}_r$ and $\mathbf{w}_c$ are the radar and communications beamformers, respectively. The received signal is filtered with a matched filter in the analog domain with the sub-pulse $a(t)$. We assume that either there is no clutter or the impact of clutter has been suppressed through a pre-processing step. Thus, the backscattered signal at the ISAC receiver can be written as

$$\mathbf{y}_r(t) = \sum_{q=1}^{Q} \alpha_q \sqrt{P_{\mathrm{tot}}} e^{j2\pi\nu_q t} \mathcal{A}_0(\tau_q - t) \times \mathbf{u}(\theta_q^r)\mathbf{u}^{\mathrm{H}}(\theta_q^r)(s_r\mathbf{w}_r + s_c\mathbf{w}_c) + \mathbf{v}_r(t) \quad (4)$$

where $\mathcal{A}_0(t)$ is the zero Doppler cut of the (delay) ambiguity function of the sub-pulse $a(t)$, and the DL communications signal $s_c\mathbf{w}_c$ is considered as mutual interference. We sample the received signal $\mathbf{y}_r(t)$ at the time instants $t_m = \frac{m-1}{T} + T_0 \, \forall \, m \in [M]$ to obtain observation matrix $\mathbf{Y}_r = [\mathbf{y}_1^r, \ldots, \mathbf{y}_M^r] \in \mathbb{C}^{N\times M}$ where M is the number of samples and T_0 is a guard time to prevent receiving backscattered signals while transmitting. For full-duplex configurations, the guard time is not necessary. The noise $\mathbf{v}_r(t_m) \coloneqq \mathbf{v}_m^r \sim \mathcal{CN}(0, \sigma_r^2\mathbf{I})$ obeys independent and identically distributed (i.i.d.) complex Gaussian distribution. With a minor abuse of notation, we can write $\mathbf{y}_r \in \{\mathbf{y}_m^r\}_{m=1}^M$ as

$$\mathbf{y}_r = \mathbf{H}_r\mathbf{x} + \mathbf{v}_r \quad (5)$$

where $\mathbf{x} = \sqrt{P_{\mathrm{tot}}}(s_r\mathbf{w}_r + s_c\mathbf{w}_c)$, and $\mathbf{v}_r \in \{\mathbf{v}_m^r\}_{m=1}^M$. Furthermore, the sensing channel matrix is

$$\mathbf{H}_r = \sum_{q=1}^{Q} \alpha_q e^{j2\pi\nu_q t} \mathcal{A}_0(\tau_q - t)\mathbf{u}(\theta_q^r)\mathbf{u}^{\mathrm{H}}(\theta_q^r) \quad (6)$$

for $t \in \{t_m\}_{m=1}^M$.

Considering the extended Saleh-Valenzuela model [18], the DL communication channel to the UE can be expressed as

$$\mathbf{h}_c = \sum_{l=1}^{L} \alpha_l^c \mathbf{u}(\theta_l^c) \quad (7)$$

where L is the number of paths, and α_l^c and θ_l^c are the complex gain and the direction of the path $l \in [L]$, respectively. The received signal at the UE receiver can be written as

$$y_c = \mathbf{h}_c^{\mathrm{H}}\mathbf{x} + v_c \quad (8)$$

where $v_c \sim \mathcal{CN}(0, \sigma_c^2)$ is the receiver noise of the UE.

B. Beamspace Transform

The beamspace refers to the transformation domain where channels and signals are analyzed in the angle or beam domain, as opposed to the element domain. The beamspace model defines N discrete Fourier transform (DFT) beams of the ULA to directions

$$\phi_n = \begin{cases} \sin^{-1}\left(\frac{2(n-1)}{N}\right), & \forall n = 1, \ldots, \frac{N}{2} \\ \pm\frac{\pi}{2}, & n = \frac{N}{2}+1 \\ \sin^{-1}\left(\frac{2(n-1)}{N} - 2\right), & \forall n = \frac{N}{2}+2, \ldots, N \end{cases} \quad (9)$$

where all the beams are orthogonal. We use matrix $\mathbf{U} \in \mathbb{C}^{N\times N}$ to denote the unitary DFT matrix. Pre-multiplying (5) by $\mathbf{U}^{\mathrm{H}}$ and using the unitary property of the DFT matrix, we can write

$$\bar{\mathbf{y}}_r \coloneqq \mathbf{U}^{\mathrm{H}}\mathbf{y}_r = \bar{\mathbf{H}}_r\bar{\mathbf{x}} + \bar{\mathbf{v}}_r \quad (10)$$

where $\bar{\mathbf{H}}_r = \mathbf{U}^{\mathrm{H}}\mathbf{H}\mathbf{U}$ is the beamspace channel matrix, $\bar{\mathbf{x}} = \mathbf{U}^{\mathrm{H}}\mathbf{x}$ is the beamspace transmitted signal, and $\bar{\mathbf{v}}_r = \mathbf{U}^{\mathrm{H}}\mathbf{v}_r$ is the noise in beamspace domain. Since $\mathbf{U}$ is a unitary matrix, the distribution of noise $\bar{\mathbf{v}}_r$ is the same as in the element domain. Furthermore, the beamspace transmitted signal can be expressed as

$$\bar{\mathbf{x}} = \sqrt{P_{\mathrm{tot}}}(s_r\bar{\mathbf{w}}_r + s_c\bar{\mathbf{w}}_c) \quad (11)$$

where $\bar{\mathbf{w}}_r = \mathbf{U}^{\mathrm{H}}\mathbf{w}_r$ and $\bar{\mathbf{w}}_c = \mathbf{U}^{\mathrm{H}}\mathbf{w}_c$ are the beamformers in the beamspace domain. Similarly, we can write the received communications signal at the UE as

$$y_c = \bar{\mathbf{h}}_c^{\mathrm{H}}\bar{\mathbf{x}} + v_c \quad (12)$$

where $\bar{\mathbf{h}}_c = \mathbf{U}^{\mathrm{H}}\mathbf{h}_c$.

III. Optimization in the Beamspace Domain

According to (6) and (10), the beamspace channel matrix is diagonal when assuming that each target is perfectly at the center of one of the DFT beams with direction $\phi_n, \forall n \in [N]$ defined in (9). Therefore, we can approximate

$$\bar{\mathbf{H}}_r \approx \mathrm{diag}(\boldsymbol{\beta}) \quad (13)$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_N]^{\mathrm{T}}$, and $\beta_n, \forall n \in [N]$ correspond to the complex target gains in the directions $\phi_n, \forall n \in [N]$. Let us also define a binary random variable $\mathbf{z} \in \{0,1\}^N$ where z_n indicates the target presence in the beam $n \in [N]$. Let us assume that $\mathbf{z}$ is constant during the considered time period. The variables $\mathbf{z}$ and $\boldsymbol{\beta}$ are coupled as follows:

- $z_n = 0 \Leftrightarrow \beta_n \sim \delta(\beta_n)$ where $\delta(\cdot)$ is the Dirac delta, and
- $z_n = 1 \Leftrightarrow \beta_n \sim \mathcal{CN}(0, g_n)$ which is a zero mean circular complex Gaussian distribution with variance $g_n > 0$.

The vector $\mathbf{g} = [g_1, g_2, \ldots, g_N]^{\mathrm{T}}$ contains the variances of the target complex gains (see eq.(2)) for hypothetical targets existing in the different beams. We assume that elements of $\mathbf{z} = [z_1, z_2, \ldots, z_N]^{\mathrm{T}}$ obey independent Bernoulli distributions with parameters $\boldsymbol{\rho} = [\rho_1, \rho_2, \ldots, \rho_N]^{\mathrm{T}}$.

The radar search function aims to find the (coarse) directions of unknown targets.[1] This can be interpreted as finding the elements of $\mathbf{z}$ equal to 1. We address this problem by employing the mutual information (MI) criterion $I(\bar{\mathbf{y}}; \boldsymbol{\beta})$ that is the information we receive about $\boldsymbol{\beta}$ by observing $\bar{\mathbf{y}}$. If β_n contains information, a target is in the nth DFT beam. Consequently, maximizing $I(\bar{\mathbf{y}}_r; \boldsymbol{\beta})$ leads to identifying the DFT beams that contain targets. This criterion is written as

$$I(\bar{\mathbf{y}}_r; \boldsymbol{\beta}) = \sum_{n=1}^{N} \xi_n \log\left(1 + \frac{g_n |\bar{w}_n^r|^2}{g_n |\bar{w}_n^c|^2 + \sigma^2 / P_{\mathrm{tot}}}\right) \tag{14}$$

where we consider the back-scattered communications signals as interference, and the variables $\xi_n \in \{0, 1\}, \forall\, n \in [N]$ are introduced to indicate the beams that have been already declared to have a target present, i.e., tracking task has been initialized.[2] Thus, $\xi_n = 0$ gives zero utility if the track has been created and $\xi_n = 1$ otherwise.

For communications, we consider communications MI as the optimization criterion. It can be written as

$$C = \log\left(1 + \frac{|\bar{\mathbf{h}}_c^{\mathrm{H}} \bar{\mathbf{w}}_c|^2}{|\bar{\mathbf{h}}_c^{\mathrm{H}} \bar{\mathbf{w}}_r|^2 + \sigma_c^2 / P_{\mathrm{tot}}}\right). \tag{15}$$

We formulate an optimization problem that is solved sequentially to facilitate active target search while communicating with the UE. The optimization problem can be written as

$$\underset{\bar{\mathbf{w}}_c, \bar{\mathbf{w}}_r}{\arg\max}\, I(\bar{\mathbf{y}}_r; \boldsymbol{\beta}) \tag{16a}$$
$$\text{s.t. } \|\bar{\mathbf{w}}_c\|_2^2 + \|\bar{\mathbf{w}}_r\|_2^2 \leq 1 \tag{16b}$$
$$C \geq C_{\min} \tag{16c}$$

where $C_{\min}$ is the minimum required communications rate. However, the optimization problem of (16) is not convex, and the gain $\mathbf{g}$ is not known in advance in a more realistic scenario. We address these issues in our proposed algorithm.

IV. Proposed method based on Thompson sampling

Thompson sampling [17] is a method that balances exploration (gathering more information about unknown parameters) and exploitation (maximizing current performance using existing knowledge of these parameters). It efficiently addresses various problems compared to other algorithms developed for exploration and exploitation problems, such as MAB algorithms [16] that work only in certain types of problems. The immediate optimization objective this paper considers is (16), and the unknown parameters are the gains $\mathbf{g}$. Thompson sampling is a Bayesian approach where $\mathbf{g}$ has a prior distribution, and the posterior distribution of $\mathbf{g}$ is updated sequentially based on the new observations. Let us define time slot index $k \in \mathbb{N}$ to refer to a specific coherent processing interval (CPI) of the sensing functionality. We use the notation $\bar{\mathbf{y}}_k^r$ to refer to the beamspace observation of equation (10) at specific time index k. In the decision-making phase, we sample from the posterior $\Pr\{\mathbf{g}|\bar{\mathbf{y}}_1^r, \ldots \bar{\mathbf{y}}_k^r\}$ and use the sampled $\mathbf{g}$ in the optimization problem (16).

A. Optimization method

We consider an orthogonal design between the communications and sensing for the ISAC system, which can be achieved by utilizing the orthogonality of the DFT beams in Section II-B. Setting $\bar{\mathbf{h}}_c^{\mathrm{H}} \bar{\mathbf{w}}_r = 0$ in (15), the UE receiver does not suffer interference from the sensing functionality. Furthermore, we require that $g_i |\bar{w}_i^c|^2 = 0$ if $|\bar{w}_i^r|^2 > 0$ for all $i \in [N]$ to remove the interference of the back-scattered communications signals to the sensing receiver (see (14)). The latter constraint means we cannot simultaneously transmit communications and sensing signals along the same beams. We propose a two-stage optimization algorithm similar to the one proposed for allocating sub-carriers of multicarrier ISAC system in [19].

Let us first optimize the performance of the communications functionality by allowing transmission to all beams. Thus, we can optimize

$$\underset{\bar{\mathbf{w}}_c}{\arg\max}\ \log\left(1 + \frac{|\bar{\mathbf{h}}_c^{\mathrm{H}} \bar{\mathbf{w}}_c|^2}{\sigma^2} P_{\mathrm{tot}}\right) \tag{17a}$$
$$\text{s.t. } \|\bar{\mathbf{w}}_c\|_2^2 \leq 1 \tag{17b}$$

which has the optimal solution $\bar{\mathbf{w}}_c^* = \bar{\mathbf{h}}_c / \|\bar{\mathbf{h}}_c\|_2$. The desired communications beamformer is obtained by keeping the minimum number of the elements of $\bar{\mathbf{w}}_c^*$ with the largest magnitudes that satisfy (16c) and setting the remaining elements to zero. Furthermore, to remove the excess power, the resulting beamformer is scaled to meet (16c) strictly. After the communications beamformer optimization, we can optimize the sensing beamformer as follows

$$\underset{\bar{\mathbf{w}}_r}{\arg\max}\, I(\bar{\mathbf{y}}_r; \boldsymbol{\beta}) \tag{18a}$$
$$\text{s.t. } \|\bar{\mathbf{w}}_r\|_2^2 \leq 1 - \|\bar{\mathbf{w}}_c\|_2^2 \tag{18b}$$
$$\bar{\mathbf{h}}_c^{\mathrm{H}} \bar{\mathbf{w}}_r = 0 \tag{18c}$$
$$\mathrm{diag}\left(\mathbb{1}_{\{|\mathbf{w}_c| > 0\}}\right) \bar{\mathbf{w}}_r = \mathbf{0} \tag{18d}$$

where $|\cdot|$ is the element-wise absolute value. The optimization problem of (18) is not convex since $I(\bar{\mathbf{y}}; \boldsymbol{\beta})$ is not concave with

[1] The method is straightforward to extend to range and Doppler dimensions. These additional dimensions are not included in the model to simplify the presentation.

[2] To consider the range and Doppler dimension, the MI is the sum over range, Doppler, and beam dimensions.

respect to $\bar{\mathbf{w}}_r$. However, by the change of variables $p_n^r = |\bar{w}_n^r|^2$ and $\mathbf{p}_r = [p_1^r, \dots, p_N^r]^{\mathrm{T}}$, we can first optimize

$$\arg\max_{\mathbf{p}_r} I(\bar{\mathbf{y}}; \boldsymbol{\beta}) \tag{19a}$$

$$\text{s.t. } \mathbf{1}^T \mathbf{p}_r \leq 1 - \|\bar{\mathbf{w}}_c\|_2^2 \tag{19b}$$

$$\mathbf{p}_r \geq 0 \tag{19c}$$

which is a well-known water-filling problem that can be solved efficiently [20]. To satisfy constraints in (18c) and (18d) we use projection, i.e, minimize L2-norm $\left\|\sqrt{\mathbf{p}_r} - \bar{\mathbf{w}}_r\right\|_2^2$ subject to the constaints. This projection can be implemented efficiently by computing the null space $\mathbf{A}_0$ of the linear constraints using singular value decomposition (SVD). Then, the optimal power allocation of (19) along with the projection can be used to construct $\bar{\mathbf{w}}_r = \mathbf{A}_0(\mathbf{A}_0^{\mathrm{H}}\mathbf{A}_0)^{-1}\mathbf{A}_0^{\mathrm{H}}\sqrt{\mathbf{p}_r}$.

B. Posterior update equations

The proposed algorithm utilizes the posterior distributions of the variables $\mathbf{z}$ and $\mathbf{g}$. Since the signal model is assumed to be independent among the beams, we can derive the update expressions independently. Thus, to simplify the notation, we can derive the posterior update equations using the scalar signal model

$$y_{k,n}^r = \beta_{k,n} x_{k,n} + v_{k,n}^r \tag{20}$$

where $\beta_{k,n} \sim \mathcal{CN}(0, g_n)$ is the target complex gain, $x_{k,n}$ is the transmitted signal and $v_{k,n} \in \mathcal{CN}(0, \sigma_k^2)$ is the noise at specific time index k and beam $n \in [N]$. We will use the shorthand $p_{k,n} = |x_{k,n}|^2$ for the power.

First, let us consider an update rule for the distribution $\eta_k(g_n) \coloneqq \Pr(g_n | y_{1:k,n}^r, z_n = 1)$ that is the distribution of gain g_n if the target is present given all the observations up to and including time instance k. We can write the distribution as follows

$$\eta_k(g_n) \propto \mathcal{CN}(y_{k,n}^r; 0, p_{k,n} g_n + \sigma_k^2)\eta_{k-1}(g_n) \tag{21}$$

using the Bayes rule where $\eta_0(g_n) = \Pr(g_n | z_n = 1)$ is the prior distribution. However, since $p_{k,n}$ and σ_k^2 vary as a function of k, there is no conjugate prior distribution to implement the update (21) in a closed form. Moreover, the computational and memory requirements for the update in (21) must be reasonable, and we need to be able to sample from η_k to use Thompson sampling. We address this problem using the Laplace approximation and Newton–Raphson method to update the posterior parameters efficiently [17].

Laplace approximation approximates a distribution by fitting the normal distribution to the (local) mode of the posterior. We use it to approximate the distribution of $u_n = \log g_n$ such that g_n is log-normal distributed. Thus, we set $\Pr(u_n | y_{1:k,n}^r, z_n = 1) \approx \mathcal{N}(u_n; \mu_{k,n}, s_{k,n}^2)$ where $\mu_{k,n}$ and $s_{k,n}^2$ are the mean and the variance of the Laplace approximation. By approximating $\Pr(u | y_{1:k',n}, z_n = 1) \approx \mathcal{N}(u; \mu_{k',n}, s_{k',n}^2)\ \forall\ k' \in [k-1]$ the memory complexity of the Laplace approximation can be reduced from linear (as a function of k) to constant. Thus, the proposed method computes the mean as

$$\mu_{k,n} = \arg\max_{u_n}\ \log \mathcal{CN}(y_{k,n}; p_{k,n} e^{u_n} + \sigma_k^2) + \log \mathcal{N}(u_n; \mu_{k-1,n}, s_{k-1,n}) \tag{22}$$

and the variance as

$$s_{k,n} = -\left(-\frac{1}{s_{k-1,n}} + \frac{\partial^2}{\partial u_n^2} \log \mathcal{CN}(y_{k,n}^r; p_{k,n} e^{u_n} + \sigma_k^2) \right)^{-1} \Bigg|_{u_n = \mu_{k,n}} \tag{23}$$

where the maximization in (22) can be done using the Newton-Raphson method initialized with $\mu_{k-1,n}$. It is reasonable to assume that $\mu_{k-1,n}$ is a good starting point since the posterior modes are typically close in subsequent time steps. However, a robust version of the Newton-Raphson method is required to address the rare cases where the objective is not locally concave. We employ an approach that forces the second derivative to be negative via the absolute value.

Next, we consider the posterior $\rho_{k,n} = \Pr(z_n = 1 | y_{1:k,n}^r)$. It can be written as follows

$$\rho_{k,n} \propto \Pr(y_{k,n}^r | y_{1:k-1,n}^r, z_n = 1)\rho_{k-1,n} \tag{24}$$

where

$$\Pr(y_{k,n}^r | y_{1:k-1,n}^r, z_n = 1) = \int_0^\infty \Pr(y_{k,n}^r | z_n = 1, g_n)\eta_{k-1}(g_n) dg \tag{25}$$

Equation (25) is approximated by a Gaussian distribution $\mathcal{CN}(y_{k,n}^r; 0, p_{k,n}\, \mathbb{E}_{\eta_{k-1}}[g_n] + \sigma_k^2)$ by matching the first and second moments. Note that (24) can be normalized to probability by using

$$\Pr(z_n = 0 | y_{1:k,n}^r) \propto \mathcal{CN}(y_{k,n}^r; 0, \sigma_k^2)(1 - \rho_{k-1,n}). \tag{26}$$

Algorithm 1 summarizes the overall method using the posterior updates and the proposed optimization method.

Algorithm 1: Thompson Sampling Based Optimization for ISAC System

Initialize priors η_0 and $\boldsymbol{\rho}_0$
for *each time index* $k \in \mathbb{N}$ **do**
    For all n sample $z_{k,n} \sim \text{Bernoulli}(\rho_{k-1,n})$
    If $z_{k,n} = 0$ set $g_n = 0$, else sample $g_n \sim \eta_{k-1}(g_n)$
    Optimize $\bar{\mathbf{w}}_c$ and $\bar{\mathbf{w}}_r$ according to Section IV-A
    Transmit the signal $\mathbf{x}_k$ and observe $\bar{\mathbf{y}}_k^r$
    Update $\eta_k(\mathbf{g}) \propto \eta_{k-1}(\mathbf{g})\, \mathcal{CN}\left(\bar{\mathbf{y}}_k^r; \mathbf{0}, \text{diag}\left(\mathbf{g} \odot |\mathbf{x}_k|^2\right) + \sigma_k^2 \mathbf{I}\right)$ using Laplace approximation and Newton-Raphson method
    Update $\boldsymbol{\rho}_k$ by using the equation (24)
    Set $\xi_n = 1$ for all n that had track initialized
endfor

V. Numerical examples

In this section, we evaluate the developed algorithm through numerical simulations. To focus on the core functionalities of the algorithm, we limit our simulations only to the beam domain. Thus, the processing in Doppler and range domains is not implemented. We evaluate the method over a sequence of time steps to understand its dynamic behavior. The number of correct tracks is used to assess the performance as a function of time. A track is correctly initialized if the target is within 3dB beamwidth of the initialized beam. We also evaluate the number of incorrect tracks. Incorrect track means that a track is initialized to a beam where a target does not exist within the 3dB beamwidth. Maximizing the number of correct tracks subject to keeping the number of incorrect tracks manageable is central to the performance of the search functionality.

Two variants of the proposed algorithms are evaluated; one as written in Section IV, and the other ignoring the sampling step and instead using the learned expected sensing gain $\mathbb{E}[\mathbf{g}|\mathbf{y}_1^r,\ldots,\mathbf{y}_k^r]$ in the optimization, referred to as the *greedy* method. Both variants initialize a track to beam $n \in [N]$ if ρ_n exceeds the threshold 0.98. We compare the proposed algorithm to three other baseline algorithms. All baseline algorithms ensure the desired communications performance as in Section IV-A. Then, the sensing performance is optimized as follows.

1) An algorithm we refer to as *conventional* optimizes the radiated power to one DFT direction at a time, and sequentially scans through all beams. In case of detection, the same beam is used in the next time slot until the criterion for track initialization is met. In the simulations, two subsequent detections are required to initialize a track. It uses a constant false alarm detector with a false alarm rate of 0.1 in order to detect low observable targets.
2) An algorithm referred to as *random* is a simplified variant of the proposed algorithm that samples beam powers randomly and uses the projection to satisfy the orthogonality constraints.
3) An algorithm we refer to as *single beam Thompson* is another simplified variant of the proposed algorithm that selects one beam for sensing at a time by choosing the beam with maximum signal-to-noise ratio (SNR) from the sampled SNRs. Furthermore, it uses an inverse-gamma distribution as the prior. This choice of prior, along with the fact that $p_{k,n}$ is either $P_{\text{tot}} - \|\bar{\mathbf{w}}_c\|^2$ or 0, facilitates closed-form the posterior updates.

The algorithms are compared with Monte Carlo with 1000 in two settings. In the first setting, the targets are precisely at one of the DFT directions, so the assumption about the sensing channel employed in this paper is satisfied. In the second setting, the target directions can be between the DFT directions. The number of beams we consider is 32, the number of propagation paths to the UE is $L = 4$, and the communications requirement is $C_{\min} = 3$. The SNRs to $K = 10$ targets are from 0dB to 20dB when a DFT beam is steered to the correct direction with maximum power P_{tot}. The target RCSs are selected to fill this interval uniformly in the dB scale. The communications rate requirement $C_{\min}$ is always satisfied for all the scenarios and algorithms. Thus, our results focus on the sensing performance.

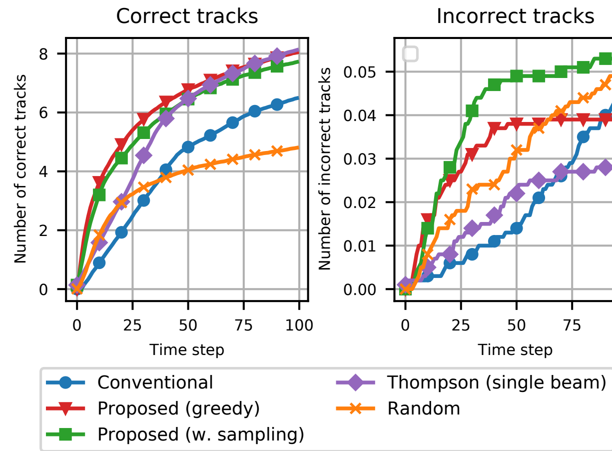

Fig. 1: The proposed method acquires a large number of correct tracks the quickest while keeping the number of incorrect tracks comparable to the other methods.

A. Targets precisely at the DFT directions

Fig. 1 shows the performance in the setting where targets are precisely located at one of the DFT directions. We first observe that the proposed method finds targets faster than the baseline methods. Furthermore, the proposed *greedy* method finds approximately as many targets as the best baseline, namely *single beam Thompson*, in the simulated time frame. The simulations surprisingly show that the performance of the proposed *sampling* method is worse compared to the *greedy* method. This may be caused by the prior η_0 and $\boldsymbol{\rho}_0$ implicitly encouraging the necessary exploration, and thus the sampling is not required. In fact, due to the prior, the learned expected gains tend to overestimate the true gain values, thus resembling upper confidence bound algorithms in MAB literature [16].

The proposed methods achieve quick track acquisition in the initial phases by distributing the power to multiple beams to gain more information than using a single beam. However, it can be seen from the performance of the *random* method that more than distributing the power randomly is required for optimal performance. The *random* method initially (the first 25 time instances) finds targets faster than the other baseline methods but is still slower than the proposed methods. Moreover, the number of tracks the *random* method generates in the simulation period is the worst because it cannot find targets with low SNR. The results also show that limiting the beamforming to a single beam, as the *single beam Thompson* does, significantly slows the track acquisition even if the method utilizes the acquired posterior of the target gains and presence. However, the performance of the *single beam Thompson* is still better compared to the conventional method.

B. Target directions on a continuous scale

Fig 2 shows the performance in the setting where the target directions are randomly sampled from the uniform distribution

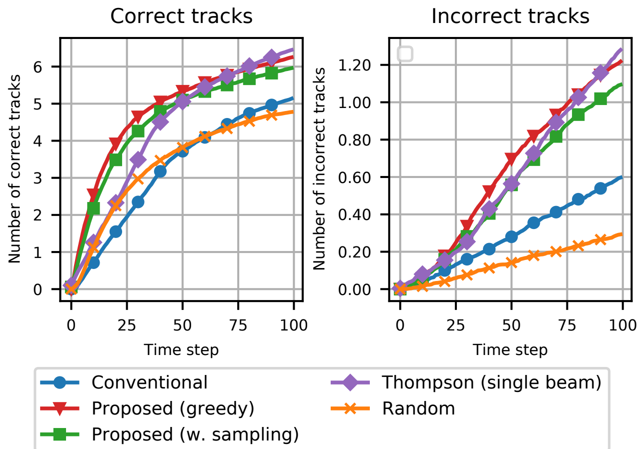

Fig. 2: The number of false tracks is increased significantly in the scenario where targets can be between the DFT directions due to the power leakage problem in the beamspace domain. The proposed method acquires correct tracks the fastest.

defined on the interval of $(-\pi/2, \pi/2)$. The main finding is that the number of incorrect tracks has increased. This occurs because of power leakage in the beamspace domain, which means that several beams can illuminate a single target simultaneously and receive the signals that are scattered back from it. Thus, typically, an incorrect track is created to the adjacent beam of the correct beam. More complicated receiver processing would be required to eliminate those incorrect tracks. The *random* method obtains fewer incorrect tracks than others due to its inability to steer the power toward the incorrect targets caused by the power leakage problem. The results in the correct tracks align mostly with the findings we made in the ideal scenario. Namely, the proposed methods find the targets quickest, and the number of correct tracks acquired in the simulation period is almost as good as with the *single beam Thompson* method.

VI. Conclusions

This paper considered optimizing spatial resources of integrated sensing and communications (ISAC) systems to manage the dual demands of sensing and communication tasks. We proposed a novel cognitive beamspace algorithm that markedly improved the overall performance of the ISAC system. It achieved this by intelligently optimizing in the beamspace domain to achieve cognitive target search behavior while communicating with a user equipment (UE) with an acceptable rate. A vital feature of this algorithm was its adaptive learning capability, enabling it to adjust its radar target search strategy dynamically in real-time according to environmental feedback. Future research directions include evaluating and improving the cognitive beamspace algorithm to be more robust in realistic operational environments.

References

[1] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, “Joint radar and communication design: Applications, state-of-the-art, and the road ahead,” *IEEE Trans. on Commun.*, vol. 68, no. 6, pp. 3834–3862, 2020.

[2] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, “Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond,” *IEEE J. on Sel. Areas in Comms.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[3] N. C. Luong, X. Lu, D. T. Hoang, D. Niyato, and D. I. Kim, “Radio resource management in joint radar and communication: A comprehensive survey,” *IEEE Comms. Surveys & Tut.*, vol. 23, no. 2, pp. 780–814, 2021.

[4] D. Ma, N. Shlezinger, T. Huang, Y. Liu, and Y. C. Eldar, “Joint radar-communication strategies for autonomous vehicles: Combining two key automotive technologies,” *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 85–97, 2020.

[5] A. Hassanien, M. G. Amin, Y. D. Zhang, and F. Ahmad, “Dual-function radar-communications: Information embedding using sidelobe control and waveform diversity,” *IEEE Trans. on Signal Process.*, vol. 64, no. 8, pp. 2168–2181, 2015.

[6] T. Huang, N. Shlezinger, X. Xu, Y. Liu, and Y. C. Eldar, “MAJoRCom: A dual-function radar communication system using index modulation,” *IEEE Trans. on Signal Process.*, vol. 68, pp. 3423–3438, 2020.

[7] D. Ma, N. Shlezinger, T. Huang, Y. Shavit, M. Namer, Y. Liu, and Y. C. Eldar, “Spatial modulation for joint radar-communications systems: Design, analysis, and hardware prototype,” *IEEE Trans. on Veh. Technol.*, vol. 70, no. 3, pp. 2283–2298, 2021.

[8] P. Kumari, J. Choi, N. González-Prelcic, and R. W. Heath, “IEEE 802.11 ad-based radar: An approach to joint vehicular communication-radar system,” *IEEE Trans. on Veh. Technol.*, vol. 67, no. 4, pp. 3012–3027, 2017.

[9] S. P. Chepuri, N. Shlezinger, F. Liu, G. C. Alexandropoulos, S. Buzzi, and Y. C. Eldar, “Integrated sensing and communications with reconfigurable intelligent surfaces: From signal modeling to processing,” *IEEE Signal Process. Mag.*, vol. 40, no. 6, pp. 41–62, 2023.

[10] F. Liu, C. Masouros, A. Li, H. Sun, and L. Hanzo, “MU-MIMO communications with MIMO radar: From co-existence to joint transmission,” *IEEE Trans. on Wireless Commun.*, vol. 17, no. 4, pp. 2755–2770, 2018.

[11] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, “Joint transmit beamforming for multiuser MIMO communications and MIMO radar,” *IEEE Trans. on Signal Process.*, vol. 68, pp. 3929–3944, 2020.

[12] Z. Cheng, Z. He, and B. Liao, “Hybrid beamforming design for ofdm dual-function radar-communication system,” *IEEE J. of Sel. Topics in Signal Process.*, vol. 15, no. 6, pp. 1455–1467, 2021.

[13] S. Haykin, “Cognitive radar: a way of the future,” *IEEE Signal Process. Mag.*, vol. 23, pp. 30–40, Jan. 2006.

[14] J. R. Guerci, “Cognitive radar: A knowledge-aided fully adaptive approach,” in *2010 IEEE Radar Conference (RadarConf10)*, pp. 1365–1370, 2010.

[15] J. Brady, N. Behdad, and A. M. Sayeed, “Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements,” *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 7, pp. 3814–3827, 2013.

[16] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Cambridge, MA, USA: Cambridge University Press, 2020.

[17] D. J. Russo, B. Van Roy, A. Kazerouni, I. Osband, and Z. Wen, “A tutorial on Thompson sampling,” *Found. Trends Mach. Learn.*, vol. 11, p. 1–96, July 2018.

[18] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, “Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems,” *IEEE J. of Sel. Topics in Signal Process.*, vol. 10, no. 3, pp. 485–500, 2016.

[19] M. Bică and V. Koivunen, “Multicarrier radar-communications waveform design for RF convergence and coexistence,” in *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP19)*, pp. 7780–7784, 2019.

[20] P. He, L. Zhao, S. Zhou, and Z. Niu, “Water-filling: A geometric approach and its application to solve generalized radio resource allocation problems,” *IEEE Trans. on Wireless Commun.*, vol. 12, no. 7, pp. 3637–3647, 2013.