This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

Faster-Than-Nyquist Symbol-Level Precoding for Wideband Integrated Sensing and Communications

Zihan Liao, Fan Liu, *Senior Member, IEEE*, Ang Li, *Senior Member, IEEE*, Christos Masouros, *Senior Member, IEEE*

***Abstract*—In this paper, we present an innovative symbol-level precoding (SLP) approach for a wideband multi-user multi-input multi-output (MU-MIMO) downlink integrated sensing and communications (ISAC) system employing faster-than-Nyquist (FTN) signaling. Our proposed technique minimizes the minimum mean squared error (MMSE) for the sensed parameter estimation while ensuring the communication per-user quality-of-service through the utilization of constructive interference (CI) methodologies. While the formulated problem is non-convex in general, we tackle this issue using proficient minorization and successive convex approximation (SCA) strategies. Numerical results substantiate that our FTN-ISAC-SLP framework can increase communication throughput by up to 20% while reducing sensing MMSE by about 1 dB.**

***Index Terms*—ISAC, dual-functional radar-communication, faster-than-nyquist, constructive interference, symbol-level precoding**

I. Introduction

INTEGRATED sensing and communications (ISAC) has emerged as a pivotal enabling technology for next-generation wireless networks, such as 5G-advanced and 6G. This technology seeks profound integration between wireless sensing and communication (S&C) to facilitate the co-design of both functionalities, thereby enhancing hardware, spectral, and energy efficiency while obtaining mutual performance gains [1]. As a result, ISAC has found applications in numerous emerging areas, including vehicular networks, industrial IoT, and smart homes [2].

Various signaling schemes have been developed for ISAC, which can be broadly classified into two primary methodologies: orthogonal resource allocation and fully unified waveform design. The former aims to allocate orthogonal or orthogonal wireless resources to S&C, thus to prevent interference between them, namely, time-, spectral-, spatial-, and code-division methods. However, this approach suffers from poor resource efficiency. Consequently, it is more advantageous to create a fully unified ISAC waveform through the shared utilization of wireless resources between S&C. This strategy is generally referred to as dual-functional radar-communication (DFRC) design.

Z. Liao and F. Liu (corresponding author) are with the Department of Electrical and Electronic Engineering and the School of System Design and Intelligent Manufacturing, Southern University of Science and Technology, Shenzhen, China. (email: liaozh2020@mail.sustech.edu.cn, liuf6@sustech.edu.cn).

A. Li is with the School of Information and Communications Engineering, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China. (email: ang.li.2020@xjtu.edu.cn).

C. Masouros is with the Department of Electronic and Electrical Engineering, University College London, WC1E 7JE London, U.K. (e-mail: c.masouros@ucl.ac.uk).

DFRC systems inherently exhibit conflicting requirements between radar and communication functionalities concerning aspects such as antenna placement, power amplifier operation regions, and signal formats, due to their different and often contradictory performance metrics and constraints. Thus, the transmit waveform must be meticulously designed to balance these requirements and enhance system performance. In general, DFRC designs can follow one of three schemes: sensing-centric design (SCD), communication-centric design (CCD), and joint design (JD) [3]. The former two schemes prioritize the sensing or communication capabilities of the ISAC system, considering the other functionality as ancillary. In contrast, JD schemes strive to create an ISAC signal from scratch instead of relying on pre-existing S&C waveforms, resulting in a scalable tradeoff between S&C [4].

Recently, multi-input multi-output (MIMO) architectures are extensively employed in DFRC systems to offer waveform diversity for radar target detection [5] and beamforming gains and spatial multiplexing for multi-user communications. Numerous researchers have focused on transmit precoding designs in MIMO DFRC systems [6]–[17], where a precoding matrix is conceived to optimize radar sensing and communication metrics. Notable radar metrics include radar receiver's signal-to-interference-plus-noise ratio (SINR) [7], beampattern mean squared error (MSE) [8], Cramér-Rao bound [9], and the similarity between the designed DFRC beamformer and the its reference radar-only counterpart [10]–[13]. Prevalent communication metrics encompass achievable rate [14], [15], communication user's SINR [10], [12], [16], and multi-user interference (MUI) [8], [11]. The amalgamation of radar sensing and communication metrics furnishes a comprehensive criterion for designing and evaluating DFRC systems.

Although existing DFRC schemes are ingeniously conceived through sophisticated approaches, they generally assume Nyquist pulse shaping implicitly. Our goal in this paper is to achieve a reasonable communication rate while not spoiling the sensing performance, thus providing a better communication-sensing tradeoff. The most intuitive way to improve data rate under limited bandwidth would be increasing constellation size or utilizing FTN-signaling. However, increasing the constellation size amplifies the communication system's vulnerability to noise, and may also incur performance loss in sensing as per the recent research results on the deterministic-random tradeoff of the ISAC system [18]. This leaves FTN signaling a more feasible solution. FTN signaling's core concept is to enhance the data rate by accelerating pulse transmission in the temporal dimension. Moreover, by implementing FTN signaling, the power of the transmitted waveform at some time slot not only includes

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

the power from the transmitted symbols at that time slot but also the power from the symbols at other time slots. In this way, the transmitted waveform is amplified, resulting in a potential to improve sensing performance. However, FTN signaling violates the Nyquist criterion and brings inter-symbol interference (ISI) [19]. In a multi-user MIMO (MU-MIMO) system employing FTN signaling, interference exists in both spatial and temporal domains, specifically, MUI and ISI due to non-orthogonality in users' channels and pulse shaping, which needs to be carefully coped with through multi-antenna precoding designs.

In the aforementioned DFRC waveform designs [6]–[17], conventional linear block-level precoding (BLP) embeds communication symbols into the dual-function waveform. However, these approaches' available degrees of freedom (DoFs) are proven to be limited by the number of users [20]–[22]. More importantly, these methods adopted block-level beampattern, a function of the signal sample covariance matrix, as a design objective, where the radar sensing performance may be guaranteed only when the number of transmitted symbols is sufficiently large. Consequently, instantaneous transmit beampatterns in different time slots might exhibit significant distortions, leading to severe performance degradation in target detection and parameter estimation if only a limited number of samples are collected. Additionally, conventional BLP designs mitigate MUI and ISI via channel equalization techniques, such as zero forcing, which overlook the fact that known interference can be harnessed to enhance useful signal power [23].

To tackle the issues above, symbol-level precoding (SLP) has been proposed as a means of exploiting, rather than merely eliminating, interference in multi-user communication systems [24]–[29]. Particularly well-suited for ISAC applications, FTN signaling and SLP form an ideal pairing, as both spatial and temporal interference can be harnessed to enhance communication performance without compromising sensing performance. Unlike conventional BLP, SLP is a non-linear and symbol-dependent approach, optimizing each instantaneous transmitted vector based on specific symbols to be transmitted. From a radar perspective, this method enables meticulous design of the instantaneous transmit beampattern in each time slot in a symbol-by-symbol manner, providing more DoFs for the sensing functionality. From a communication perspective, SLP can exploit transmitted symbol information to convert interference into constructive components, thereby enhancing the quality-of-service (QoS) of multi-user communications. Given the flexibility of JD-based DFRC waveform design, SLP can fully exploit constructive interference (CI). By incorporating FTN signaling and SLP in the ISAC system, we aim to harness the benefits of both techniques, facilitating the exploitation of both temporal and spatial interference. This combination is particularly suitable for ISAC applications and results in performance augmentation for sensing and communication from both temporal and spatial dimensions.

Previous research on ISAC has predominantly adopted a narrowband model, with complex wideband communication tasks typically addressed through orthogonal frequency division multiplexing (OFDM) [30]. However, as our work seeks to enhance data rates in the time domain as opposed to the frequency domain, the applicability of OFDM becomes less evident. This challenge is compounded by the expansion of the baseband signal's bandwidth under FTN signaling due to reduced symbol duration, thereby necessitating the consideration of a wideband communication model in our paper. While certain basic sensing tasks, such as angle detection, do not demand a wideband radar system, an increasing range of sensing requirements, including the detection of fast-moving objects and high-range resolution detection, can only be adequately addressed using wideband radar systems [31]. Therefore, to ensure comprehensive applicability and cater to the evolving requirements of our proposed ISAC system, we have opted to incorporate a wideband signal model in our research.

In this paper, we propose a novel DFRC precoding technique referred to as FTN-ISAC-SLP for a MIMO ISAC system, wherein a multi-antenna BS simultaneously serves multiple single-antenna communication users and detects target response matrices for radar sensing. This approach amalgamates the strategies discussed above, thus actualizing performance enhancement for S&C from both temporal and spatial dimensions. The existing literature on ISAC predominantly focuses on narrowband conditions, rendering the devised systems inapplicable in certain scenarios. In this paper, we extend the discussion to encompass wideband conditions, and develop ISAC signaling strategies for wideband systems. The primary contributions of this work are summarized as follows:

- We develop the system model for wideband DFRC transmission using FTN signaling, and formulate the FTN-ISAC-SLP waveform design as an optimization problem.
- To efficiently solve the non-convex waveform design problem, we devise a pair of algorithm frameworks that employ minorization or SCA methods, which transform the problem into two solvable second-quadratically constrained quadratic programming (QCQP) sub-problems. The minorization approach demonstrates rapid convergence, while the SCA approach excels in minimizing the objective function.
- We further propose a more computationally efficient method, termed binary penalty search (BPS), to solve the sub-problems in minorization and SCA methods. The BPS method converts the QCQP into sequential quadratic programming (QP) problems, which can be readily solved with significantly reduced computational overheads.
- We provide extensive numerical examples to illustrate the superiority of the proposed wideband FTN-ISAC-SLP designs over its Nyquist BLP counterparts, which demonstrate considerable performance gains in both radar sensing and multi-user communications.

The remainder of this paper is structured as follows. Section II introduces the system model, the performance metrics for multi-user communications, and radar sensing, as well as the problem formulation. The proposed minorization and SCA algorithms, in addition to the BPS method, are developed in Section III. Simulation results are presented in Section IV, and conclusions are provided in Section V. Lastly, some proofs of the propositions in the paper are appended in Section VI.

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

TABLE I
NOTATION TABLE

Notation	**Description**
Boldface lower-case letter	column vector
Boldface upper-case letter	matrix
$(\cdot)^\top$	Transpose operation
$(\cdot)^H$	Transpose-conjugate operation
$(\cdot)^*$	Conjugate of a complex number or matrix
$\mathbb{R}$	Set of real numbers
$\mathbb{C}$	Set of complex numbers
$\|a\|$	Absolute value of real scalar a
$\|\|b\|\|$	Magnitude of complex scalar b
$\|\| \cdot \|\|_F$	Frobenius norm of a matrix
$\Re\{\cdot\}$	Real part of a complex number or matrix
$\Im\{\cdot\}$	Imaginary part of a complex number or matrix
$*$	Convolution operation
$\otimes$	Kronecker product
$\circ$	Hadamard product
$\mathbf{A} \succeq \mathbf{0}$	Matrix $\mathbf{A}$ is positive semi-definite
$\mathbf{I}_M$	$M \times M$ identity matrix
$\mathbf{0}_{M,N}$	$M \times N$ matrix with all entries being 0
$\text{Diag}(\mathbf{a})$	Diagonal matrix with diagonal entries from vector $\mathbf{a}$
$\text{diag}(\mathbf{A})$	Vector with entries being diagonal entries of matrix $\mathbf{A}$
$\mathbb{E}[\mathbf{A}]$	Expectation matrix of random matrix $\mathbf{A}$
$\text{tr}(\cdot)$	Trace of a square matrix
$\text{vec}(\mathbf{A})$	Vector obtained by column-wise stacking of entries of matrix $\mathbf{A}$
$\mathbf{a} \sim \mathcal{CN}(\mathbf{m}, \mathbf{R})$	$\mathbf{a}$ obeys a complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$

II. SYSTEM MODEL

We consider a wideband MIMO ISAC BS equipped with N_t transmit antennas and N_r receive antennas, which is serving K downlink single-antenna users while detecting targets as a monostatic radar. Without loss of generality, we assume $K < N_t$. Before formulating the FTN-ISAC-SLP problem, we first elaborate on the system model and performance metrics of both radar sensing and communications.

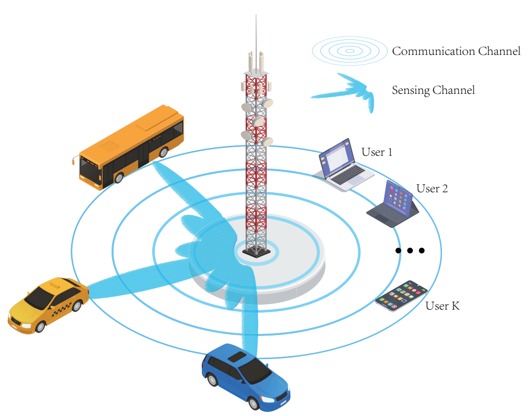

Fig. 1. ISAC Downlink System.

A. General Signal Model

Let $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_K]^\top \in \mathbb{C}^{K \times L}$ denote the symbol matrix to be transmitted, with $\mathbf{d}_k \in \mathbb{C}^{L \times 1}$ being the symbol stream intended for the k-th user with a block length L, and each entry being drawn from a given constellation. Unless otherwise specified, in this paper we consider a PSK constellation, since the extension to QAM constellations is straightforward using approaches from the literature, for example, those in [27]. Moreover, let $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_{N_t}]^\top \in \mathbb{C}^{N_t \times L}$ be the precoded signal matrix, with its entry $s_{n,i}$ at n-th row and i-th column representing the precoded symbols to be transmitted from n-th antenna at the i-th time slot. Suppose that the precoded symbols $\mathbf{S}$ are passed through a root-raised-cosine (RRC) shaping filter $\varphi(t)$ with a roll-off factor α and a duration T_0. The band-limited signal is transmitted with an FTN-specific symbol interval $T = \tau T_0$ where $\tau \in [0, 1]$. Under such a setting, the transmitted FTN signal $x_n(t)$ at n-th antenna can be expressed as

$$x_n(t) = \sum_{i=1}^{L} \varphi(t - (i-1)T) s_{n,i}, \tag{1}$$

where $s_{n,i}$ is i-th element of $\mathbf{s}_n$.

1) Communication Model: Suppose $h^{ij}(t)$ denotes the impulse response from the j-th transmitting antenna to the i-th receiving antenna. In the context of narrowband signaling, $h^{ij}(t)$ may be approximated as an impulse function, due to the frequency response of the channel being nearly constant within a narrow frequency range. However, in the case of wideband signaling, the same assumption becomes untenable, necessitating the consideration of a more general form for $h^{ij}(t)$. Consequently, the signal received at the i-th receiving antenna can be expressed as $y_i = \sum_j h^{ij}(t) * x(t) + n_i(t)$, where $*$ represents convolution. Given that $h^{ij}(t)$ can no longer be approximated as an impulse function, the convolution operation cannot be simplified to multiplication by a constant. Therefore, rather than directly using $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, we propose $\mathbf{y}(t) = \mathbf{H}(t) * \mathbf{x}(t) + \mathbf{n}(t) = \int \mathbf{H}(\tau)\mathbf{x}(t-\tau)d\tau + \mathbf{n}(t)$.

Therefore the wideband MIMO input-output relationship in the communication model is given by

$$\mathbf{r}_c(t) = \mathbf{H}_c(t) * \mathbf{x}(t) + \mathbf{n}_c(t). \tag{2}$$

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

where $\mathbf{n}_c(t)$ is the complex-valued AWGN with zero mean and variance σ_C^2, and the element of the channel impulse response matrix $\mathbf{H}_c(t)$, namely, $h_c^{ij}(t)$, is the impulse response from the j-th transmit antenna to the i-th receive antenna.

The received FTN signal after passing through a matched filter $\varphi^*(-t)$ is given by

$$\begin{aligned}\mathbf{y}_c(t) &= (\varphi^*(-t)\mathbf{I}) * \mathbf{r}_c(t) \\ &= (\varphi^*(-t)\mathbf{I}) * \mathbf{H}_c(t) * \mathbf{x}(t) + (\varphi^*(-t)\mathbf{I}) * \mathbf{n}_c(t) \\ &= \mathbf{H}_c(t) * (\phi(t)\mathbf{I}) * \mathbf{s}(t) + \boldsymbol{\eta}(t),\end{aligned} \tag{3}$$

where

$$\begin{aligned}\phi(t) &= \int_{-\infty}^{\infty} \varphi(\zeta)\varphi^*(\zeta - t)d\zeta, \\ \boldsymbol{\eta}(t) &= \int_{-\infty}^{\infty} \mathbf{n}_c(\zeta)\varphi^*(\zeta - t)d\zeta.\end{aligned} \tag{4}$$

Let $\mathbf{X}_C = [\mathbf{x}_{C,1}, \mathbf{x}_{C,2}, \cdots, \mathbf{x}_{C,L}]$ and $\mathbf{x}_{C,i}$ be the sample of $(\phi(t)\mathbf{I}) * \mathbf{s}(t)$ at $t = (i-1)T$.

Then we can discretize $(\phi(t)\mathbf{I}) * \mathbf{s}(t)$ with respect to time as

$$\mathbf{X}_C = \mathbf{S}\boldsymbol{\Omega}_\phi^\top. \tag{5}$$

where

$$\boldsymbol{\Omega}_\phi = \begin{bmatrix} \phi(-QT) & 0 & \cdots & 0 \\ \vdots & \phi(-QT) & \ddots & \vdots \\ \phi(0) & \vdots & \ddots & 0 \\ \vdots & \phi(0) & \ddots & \phi(-QT) \\ \phi(QT) & \vdots & \ddots & \vdots \\ \vdots & \phi(QT) & \ddots & \phi(0) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \phi(QT) \end{bmatrix} \in \mathbb{R}^{(L+2Q)\times L}. \tag{6}$$

By letting $\mathbf{H}_C = [\mathbf{H}_{C,1}, \mathbf{H}_{C,2}, \cdots, \mathbf{H}_{C,P}]$ represent the sampled impulse response of $\mathbf{H}_c(t)$ at time $t = iT, i = 0, 1, \cdots, P-1$, $L_0 = L + 2Q$ and $L_1 = L + 2Q + P - 1$, the equation (3) can be discretized as

$$\widetilde{\mathbf{Y}}_C = \mathbf{H}_C * \mathbf{X}_C + \widetilde{\mathbf{N}}_C, \tag{7}$$

where the i-th column of $\mathbf{H}_C * \mathbf{X}_C$ can be expressed as

$$\begin{aligned}&\mathbf{z}_{C,i} = \sum_j \mathbf{H}_{C,j}\mathbf{x}_{C,i-j}, \\ &1 \le i \le L_1, 1 \le j \le P, 1 \le i - j \le L_0.\end{aligned} \tag{8}$$

which is the result of matrix discrete convolution.

Notice that $\widetilde{\mathbf{Y}}_C \in \mathbb{C}^{K\times L_1}$. However, we need to recover $\mathbf{D} \in \mathbb{C}^{K\times L}$ from $\widetilde{\mathbf{Y}}_C$. To accomplish that, we right multiply a matrix $\mathbf{G} \in \mathbb{C}^{L_1\times L}$ to reduce the size of $\widetilde{\mathbf{Y}}_C$ to obtain the received symbol stream $\mathbf{Y}_C \in \mathbb{C}^{K\times L}$:

$$\widetilde{\mathbf{Y}}_C\mathbf{G} = (\mathbf{H}_C * \mathbf{X}_C)\mathbf{G} + \widetilde{\mathbf{N}}_C\mathbf{G}. \tag{9}$$

One possible choice of $\mathbf{G}$ is $\mathbf{G}_1 = [\mathbf{0}_{P+Q,L}^\top, \mathbf{I}_L^\top, \mathbf{0}_{Q-1,L}^\top]^\top$, which means we only take the first L sample points and discard the rest ones. However, that would result in the incomplete use of the received signal energy. Another choice of $\mathbf{G}$ is

$$\mathbf{G}_2 = \begin{bmatrix} \mathbf{I}_{2L-L_1} & \mathbf{0}_{2L-L_1,L_1-L} \\ \mathbf{0}_{L_1-L,2L-L_1} & \mathbf{I}_{L_1-L} \\ \mathbf{0}_{L_1-L,2L-L_1} & \mathbf{I}_{L_1-L} \end{bmatrix}, \tag{10}$$

which means we add the last $L_1 - L$ sample points to the previous points. However, that would result in the interference between consecutive symbols. The design of $\mathbf{G}$ may be adjusted according to different scenarios and in this paper, we take the first choice.

Moreover, we note that the noise $\widetilde{\mathbf{N}}_C\mathbf{G} = [\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_K]^\top\mathbf{G}$, with $\boldsymbol{\eta}_k = [\eta_k(0), \eta_k(T), \cdots, \eta_k((L-1)T)]^\top$ being the corresponding received noise vector at the k-th user, is not independent at each time slot.

Proposition 1: For the noise $\boldsymbol{\eta}_k$ received at k-th user, we have

$$\mathbb{E}[\boldsymbol{\eta}_k\boldsymbol{\eta}_k^H] = \sigma_C^2\boldsymbol{\Phi}_1, \tag{11}$$

where

$$\boldsymbol{\Phi}_1 = \begin{bmatrix} \phi(0) & \phi(-T) & \cdots & \phi(-(L_1-1)T) \\ \phi(T) & \phi(0) & \cdots & \phi(-(L_1-2)T) \\ \vdots & \vdots & \ddots & \vdots \\ \phi((L_1-1)T) & \phi((L_1-2)T) & \cdots & \phi(0) \end{bmatrix}. \tag{12}$$

Proof: See section Appendix A.

Thereby $\mathbb{E}[\mathbf{G}^\top\boldsymbol{\eta}_k\boldsymbol{\eta}_k^H\mathbf{G}] = \sigma_C^2\mathbf{G}^\top\boldsymbol{\Phi}_1\mathbf{G}$. To decorrelate the noise, let the eigenvalue decomposition of $\mathbf{G}^\top\boldsymbol{\Phi}_1\mathbf{G}$ be $\mathbf{U}_\phi\boldsymbol{\Lambda}_\phi\mathbf{U}_\phi^H$ where $\mathbf{U}_\phi$ is a unitary matrix containing eigenvectors and $\boldsymbol{\Lambda}_\phi$ is a diagonal matrix composed by eigenvalues. Right-multiplying $\mathbf{U}_\phi$ at both sides of (9) yields

$$\mathbf{Y}_C = (\mathbf{H}_C * \mathbf{X}_C)\mathbf{G}\mathbf{U}_\phi + \mathbf{N}_C, \tag{13}$$

where $\mathbf{Y}_C = \widetilde{\mathbf{Y}}_C\mathbf{G}\mathbf{U}_\phi$ and $\mathbf{N}_C = \widetilde{\mathbf{N}}_C\mathbf{G}\mathbf{U}_\phi$. By doing so, the covariance matrix for row vectors of $\mathbf{N}_C$ becomes $\sigma_C^2\boldsymbol{\Lambda}_\phi$, i.e., a diagonal matrix.

2) Radar Sensing Model: Consider the target response matrix (TRM) $\mathbf{H}_r(t) \in \mathbb{C}^{N_r\times N_t}$ that models the sensing channel. Depending on the sensing scenarios, $\mathbf{H}_r(t)$ can be of different forms.

Consider the target response matrix (TRM) $\mathbf{H}_r(t) \in \mathbb{C}^{N_r\times N_t}$ that models the sensing channel. Depending on the sensing scenarios, $\mathbf{H}_r(t)$ can be of different forms. In practical scenarios, the sensing channel $\mathbf{H}_R(\boldsymbol{\eta})$ may be viewed as a nonlinear function of the parameters $\boldsymbol{\eta}$, and the goal is to estimate the desired parameters $\boldsymbol{\eta}$ from the received $\mathbf{Y}_R$. However, characterizing the MMSE for estimating certain parameters in a potentially non-linear observation model can be intractable. In fact, a closed-form MMSE is attainable only when $\mathbf{Y}_R$ and $\boldsymbol{\eta}$ are jointly Gaussian, which requires $\mathbf{H}_R(\boldsymbol{\eta})$ to be a linear transformation of $\boldsymbol{\eta}$, which holds only for a few cases. Moreover, the prior knowledge of targets, e.g., the number of scatterers and angle parameters of the to-be-sensed targets may not be available at the BS. As a result, there is typically no definitive structure for $\mathbf{H}_R(\boldsymbol{\eta})$, resulting in significant difficulties in directly optimizing the MMSE of $\boldsymbol{\eta}$.

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

An alternative, more practical, general, and robust approach, in line with the insights provided in [32], is to use a two-stage process. To ensure the generality of the proposed method, following [3], we consider a generic Target Response Matrix (TRM) $\mathbf{H}_R$ instead of a specific model. This allows us to initially estimate the overall TRM by minimizing MMSE($\mathbf{H}$) to obtain $\hat{\mathbf{H}}_R$, which is a more achievable task. Subsequently, $\boldsymbol{\eta}$ can be determined by minimizing $\|\hat{\mathbf{H}}_R - \mathbf{H}_R(\boldsymbol{\eta})\|_F^2$. The main objective of this paper is to estimate the overall channel, which is the first stage of the proposed two-stage approach.

At the sensing receiver, we directly sample the received signal without passing it through the pulse-shaping filter, yielding the following radar-receiving signal model. Similar to the equation (3), the received echo signal at the receive antennas can be written as

$$\mathbf{y}_r(t) = \mathbf{H}_r(t) * (\varphi(t)\mathbf{I}) * \mathbf{s}(t) + \mathbf{n}_r(t). \tag{14}$$

where $\mathbf{n}_r(t)$ is the complex-valued AWGN at the receive antennas with zero mean and variance σ_R^2. We can then discretize $(\varphi(t)\mathbf{I}) * \mathbf{s}(t)$ in the same way with equation (5) as

$$\mathbf{X}_R = \mathbf{S}\boldsymbol{\Omega}_\varphi^\top. \tag{15}$$

where

$$\boldsymbol{\Omega}_\varphi = \begin{bmatrix} \varphi(-QT) & 0 & \cdots & 0 \\ \vdots & \varphi(-QT) & \ddots & \vdots \\ \varphi(0) & \vdots & \ddots & 0 \\ \vdots & \varphi(0) & \ddots & \varphi(-QT) \\ \varphi(QT) & \vdots & \ddots & \vdots \\ \vdots & \varphi(QT) & \ddots & \varphi(0) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \varphi(QT) \end{bmatrix}. \tag{16}$$

Then similar to equation (7) we can rewrite equation (14) as

$$\mathbf{Y}_R = \mathbf{H}_R * \mathbf{X}_R + \mathbf{N}_R, \tag{17}$$

where $\mathbf{H}_R = [\mathbf{H}_{R,1}, \mathbf{H}_{R,2}, \cdots, \mathbf{H}_{R,P}]$ and $\mathbf{H}_{R,i}$ is the sampled TRM of $\mathbf{H}_r(t)$ at $t = (i-1)T$ and $\mathbf{N}_R$ denotes an AWGN matrix, with zero mean and the variance of each entry being σ_R^2. Here we assume every entry of $\mathbf{H}_R$ follows complex Gaussion distribution $\mathcal{CN}(0, \sigma_H^2)$.

Remark: In the communication model we attempt to detect the signal $\mathbf{D}$ from $\mathbf{r}_c(t)$ in the receiver side, thus we pass the received signal to RRC matched filter to maximize the received SINR for each precoded symbol. In the sensing model, our aim is to recover the TRM $\mathbf{H}_R$ from the raw observation (14), rather than to recover $\mathbf{D}$. Therefore, we treat $\mathbf{X}_R$ as an equivalent transmitted waveform and regard (14) as the sufficient statistics for estimating $\mathbf{H}_R$, which needs not to be match-filtered by the RRC pulse.

B. Constraints and Objective Function for MIMO Model

1) MIMO Communication Model and CI constraint: The original form of communication model that contains matrix convolution is not easy to handle. To that end, we convert matrix convolution to matrix multiplication by rewriting equation (9) as

$$\mathbf{Y}_C{}^\top = (\mathbf{G}\mathbf{U}_\phi)^\top \overline{\mathbf{X}}_C \mathbf{H}_C{}^\top + \mathbf{N}_C{}^\top, \tag{18}$$

where

$$\overline{\mathbf{X}}_C = \begin{bmatrix} \mathbf{x}_{C,1}^\top & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{x}_{C,2}^\top & \mathbf{x}_{C,1}^\top & \ddots & \vdots \\ \vdots & \mathbf{x}_{C,2}^\top & \ddots & \mathbf{0} \\ \mathbf{x}_{C,L_0}^\top & \vdots & \ddots & \mathbf{x}_{C,1}^\top \\ \mathbf{0} & \mathbf{x}_{C,L_0}^\top & \ddots & \mathbf{x}_{C,2}^\top \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x}_{C,L_0}^\top \end{bmatrix} \in \mathbb{C}^{L_1 \times NP}. \tag{19}$$

Proposition 2: By defining $\mathbf{E}_p = [\mathbf{0}_{L_0\times(p-1)}, \mathbf{I}_{L_0}, \mathbf{0}_{L_0\times(P-p)}]^\top$, we are able to rewrite equation (18) as

$$\mathrm{vec}(\mathbf{Y}_C^\top) = \overline{\mathbf{H}}_C \mathrm{vec}(\mathbf{S}^\top) + \mathrm{vec}(\mathbf{N}_C^\top), \tag{20}$$

where

$$\overline{\mathbf{H}}_C = (\mathbf{H}_C \otimes (\mathbf{G}\mathbf{U}_\phi)^\top) \begin{bmatrix} \mathbf{I}_{N_t} \otimes (\mathbf{E}_1\boldsymbol{\Omega}_\phi) \\ \mathbf{I}_{N_t} \otimes (\mathbf{E}_2\boldsymbol{\Omega}_\phi) \\ \vdots \\ \mathbf{I}_{N_t} \otimes (\mathbf{E}_P\boldsymbol{\Omega}_\phi) \end{bmatrix}, \tag{21}$$

and $\mathrm{vec}(\mathbf{N}_C^\top) \sim \mathcal{CN}(\mathbf{0}, \sigma_C^2 \mathbf{I}_K \otimes \boldsymbol{\Lambda}_\phi)$.

Proof: See section Appendix B.

This system has two types of interference: temporal interference caused by FTN signaling and wideband signaling, and spatial interference in MIMO systems between users. By expressing the communication model as in equation (20), it is viable to consider both interferences in a unified manner.

Subsequently, we are now able to present our CI constraint for communication performance. Note that such a constraint is different from the SINR constraint used in BLP, which cancels the interference so that the received symbol lies in the correct region. CI constraint, on the other hand, pushes the received symbols away from their corresponding detection thresholds within the modulated-symbol constellation, instead of canceling the interference. In this way, the symbol is kept in the desired region, and the power of the interference is also used to help maintain the symbol in the correct region, thus contributing positively to the overall use of signal power. In this paper, we will use PSK modulation for simplicity, as the CI constraint for a PSK symbol can be represented by two real linear constraints. Other modulations are also possible in our system, which may be characterized by more linear constraints.

According to [25], for any transmitted PSK symbol d and its corresponding noise-free received symbol y, the CI constraint guarantees that

$$|\Im\{d^* y\}| - \Re\{d^* y\} \tan\theta \le (-\sqrt{\Gamma}\tan\theta)\sigma. \tag{22}$$

where Γ represents the requisite SINR at the receiver end, θ is related to the type of constellation and is $\pi/4$ for QPSK, and

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

σ denotes the standard deviation of the corresponding noise imposed on y.

Suppose $\overline{\mathbf{H}}_C \text{vec}(\mathbf{S}^\top) = [\mathbf{y}_{C,1}^\top, \mathbf{y}_{C,2}^\top, \cdots, \mathbf{y}_{C,K}^\top]^\top$, where $\mathbf{y}_{C,k}$ and $\mathbf{d}_k$ represent the symbol stream received by the k-th user in the absence of noise and the symbol stream intended for transmission to the k-th user, respectively. Define $\boldsymbol{\varsigma} = \sqrt{\text{diag}(\sigma_C^2 \mathbf{I}_K \otimes \boldsymbol{\Lambda}_\phi)}$, following the inequality (22), the CI constraint imposed on the k-th user can be expressed as

$$|\Im\{\mathbf{d}_k^* \circ \mathbf{y}_{C,k}\}| - \Re\{\mathbf{d}_k^* \circ \mathbf{y}_{C,k}\} \tan\theta \leq (-\sqrt{\Gamma_k}\tan\theta)\boldsymbol{\varsigma}, \forall k \tag{23}$$

where $\circ$ denotes the Hadamard product, and Γ_k is the required SNR for the k-th user. Subsequently, we aim to consolidate the k CI constraints and recast them into a single matrix inequality.

Proposition 3: Define $\boldsymbol{\Gamma} = \text{Diag}([\Gamma_1, \Gamma_2, \cdots, \Gamma_K]^\top)$ and $\overline{\mathbf{D}} = \text{Diag}(\text{vec}(\mathbf{D}^\top))$; then, the CI constraint for k users can be formulated as

$$\begin{aligned}\left|\Im\left\{\overline{\mathbf{D}}^* \overline{\mathbf{H}}_C \text{vec}(\mathbf{S}^\top)\right\}\right| - \Re\left\{\overline{\mathbf{D}}^* \overline{\mathbf{H}}_C \text{vec}(\mathbf{S}^\top)\right\}\tan\theta \\ \leq (-\sqrt{\boldsymbol{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}.\end{aligned} \tag{24}$$

Proof: See section Appendix C.

2) MIMO Radar Model and MMSE for Sensing: Similar to the communication model, the radar model can also be expressed as

$$\mathbf{Y}_R{}^\top = \overline{\mathbf{X}}_R \mathbf{H}_R{}^\top + \mathbf{N}_R{}^\top, \tag{25}$$

where $\overline{\mathbf{X}}_R$ is defined in the same fashion with equation (19) as

$$\overline{\mathbf{X}}_R = [\mathbf{E}_1 \mathbf{X}_R^\top, \mathbf{E}_2 \mathbf{X}_R^\top, \cdots, \mathbf{E}_P \mathbf{X}_R^\top]. \tag{26}$$

Different from the communication model, we extract $\text{vec}(\mathbf{H}_R^\top)$ for the sake of derivation of MMSE. Thus we have

$$\text{vec}(\mathbf{Y}_R^\top) = (\mathbf{I}_{N_r} \otimes \overline{\mathbf{X}}_R)\text{vec}(\mathbf{H}_R^\top) + \text{vec}(\mathbf{N}_R^\top). \tag{27}$$

Let us assume $\text{vec}(\mathbf{H}_R^\top) = \mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \sigma_H^2 \mathbf{I})$, then according to [32], [33], MMSE with respect to $\mathbf{h}$ can be written as

$$\begin{aligned}\text{MMSE} &= \sigma_R^2 \text{tr}\left(\left(\frac{\sigma_R^2}{\sigma_H^2}\mathbf{I} + (\mathbf{I}_{N_r}\otimes\overline{\mathbf{X}}_R)^H(\mathbf{I}_{N_r}\otimes\overline{\mathbf{X}}_R)\right)^{-1}\right) \\ &= \sigma_R^2 N_r \text{tr}\left(\left(\frac{\sigma_R^2}{\sigma_H^2}\mathbf{I} + \overline{\mathbf{X}}_R^H \overline{\mathbf{X}}_R\right)^{-1}\right) \\ &= N_r \text{tr}(\sigma_H^2 \mathbf{I} - \sigma_H^4 \overline{\mathbf{X}}^H (\sigma_H^2 \overline{\mathbf{X}}_R \overline{\mathbf{X}}_R^H + \sigma_R^2 \mathbf{I})^{-1} \overline{\mathbf{X}}_R),\end{aligned} \tag{28}$$

where the MMSE estimator is expressed as

$$\hat{\mathbf{h}}_{\text{MMSE}} = \sigma_H^2 \overline{\mathbf{X}}_R^H (\sigma_H^2 \overline{\mathbf{X}}_R \overline{\mathbf{X}}_R^H + \sigma_R^2 \mathbf{I})^{-1} \mathbf{y}. \tag{29}$$

Upon successful estimation of the TRM, it becomes feasible to determine the parameters of the sensing channel, utilizing the resultant TRM matrices.

3) MIMO Energy Constraint: Due to the fact that our transmitted impulses are no longer orthogonal to each other, using $\|\mathbf{S}\|_F^2$ as the energy might not be suitable.

Proposition 4: The energy of the transmitted waveform $x_n(t)$ at n-th antenna is given by

$$\int \|x_n(t)\|^2 dt = \mathbf{s}_n^H \boldsymbol{\Phi} \mathbf{s}_n. \tag{30}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi(0) & \phi(-T) & \cdots & \phi(-(L-1)T) \\ \phi(T) & \phi(0) & \cdots & \phi(-(L-2)T) \\ \vdots & \vdots & \ddots & \vdots \\ \phi((L-1)T) & \phi((L-2)T) & \cdots & \phi(0) \end{bmatrix}. \tag{31}$$

Proof: See section Appendix D.

Therefore, the energy constraint under a given budget E may be written as

$$\sum_{n=1}^{N_t} \mathbf{s}_n^H \boldsymbol{\Phi} \mathbf{s}_n = \text{tr}(\mathbf{S}\boldsymbol{\Phi}\mathbf{S}^H) \leq E. \tag{32}$$

or in the following quadratic form with respect to $\text{vec}(\mathbf{S}^\top)$

$$\text{vec}(\mathbf{S}^\top)^H (\mathbf{I}_{N_t} \otimes \boldsymbol{\Phi}) \text{vec}(\mathbf{S}^\top) \leq E. \tag{33}$$

Alternatively, one may also impose a per-antenna energy constraint, namely

$$\mathbf{s}_n^H \boldsymbol{\Phi} \mathbf{s}_n = \text{tr}(\mathbf{S}\boldsymbol{\Phi}\mathbf{S}^H) \leq E_n, \ \forall n. \tag{34}$$

We will compare the resultant MMSE performance under both energy constraints in numerical results.

III. FTN-ISAC Symbol-Level Precoding Design

A. Problem Formulation

Based on the discussion above, the precoding optimization problem for the MIMO model can be expressed as

$$\begin{aligned}\min_{\mathbf{S}} \ & f(\mathbf{S}) = \text{tr}\left(\left(\frac{\sigma_R^2}{\sigma_H^2}\mathbf{I} + \overline{\mathbf{X}}_R^H \overline{\mathbf{X}}_R\right)^{-1}\right) \\ s.t. \ & \left|\Im\left\{\overline{\mathbf{D}}^* \overline{\mathbf{H}}_C \text{vec}(\mathbf{S}^\top)\right\}\right| - \Re\left\{\overline{\mathbf{D}}^* \overline{\mathbf{H}}_C \text{vec}(\mathbf{S}^\top)\right\}\tan\theta \\ & \qquad \leq (-\sqrt{\boldsymbol{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}, \\ & \text{vec}(\mathbf{S}^\top)^H (\mathbf{I}_{N_t}\otimes\boldsymbol{\Phi})\text{vec}(\mathbf{S}^\top) \leq E, \\ & \mathbf{X}_R = \mathbf{S}\boldsymbol{\Omega}_\varphi^\top, \overline{\mathbf{X}}_R = [\mathbf{E}_1 \mathbf{X}_R^\top, \mathbf{E}_2 \mathbf{X}_R^\top, \cdots, \mathbf{E}_P \mathbf{X}_R^\top].\end{aligned} \tag{35}$$

By formulating the problem above, we aim to construct the pre-encoded symbols $\mathbf{S}$ corresponding to the given data matrix $\mathbf{D}$ intended for transmission, in such a manner that the MMSE pertinent to radar detection is reduced to its lowest possible value, whilst concurrently ensuring the CI constraint is satisfied in the context of a energy budget E.

B. Minorization Approach for FTN-ISAC-SLP

It is important to acknowledge that the optimization problem (35) is non-convex in nature. To tackle this challenge, we devise an optimization framework based on the minorization approach in this section.

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

According to (28), solving problem (35) is equivalent to solving problem (36).

$$\begin{aligned}\max_{\mathbf{S}}\ & f_m(\mathbf{S}) = \mathrm{tr}(\overline{\mathbf{X}}_R^H(\sigma_H^2\overline{\mathbf{X}}_R\overline{\mathbf{X}}_R^H + \sigma_R^2\mathbf{I})^{-1}\overline{\mathbf{X}}_R)\\ s.t.\ & \left|\Im\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\right| - \Re\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\tan\theta\\ & \qquad \le (-\sqrt{\boldsymbol{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma},\\ & \mathrm{vec}(\mathbf{S}^\top)^H(\mathbf{I}_{N_t}\otimes\boldsymbol{\Phi})\mathrm{vec}(\mathbf{S}^\top) \le E,\\ & \mathbf{X}_R = \mathbf{S}\boldsymbol{\Omega}_\varphi^\top, \overline{\mathbf{X}}_R = [\mathbf{E}_1\mathbf{X}_R^\top, \mathbf{E}_2\mathbf{X}_R^\top, \cdots, \mathbf{E}_P\mathbf{X}_R^\top].\end{aligned} \tag{36}$$

The fundamental concept of the proposed framework revolves around deriving a minorizer for $f_m(\mathbf{S})$. More specifically, the derived minorizers (denoted by $g_m(\mathbf{S};\mathbf{S}_k)$) ought to satisfy the following conditions:

$$g_m(\mathbf{S};\mathbf{S}_k) \le f_m(\mathbf{S}),\ g_m(\mathbf{S}_k;\mathbf{S}_k) = f_m(\mathbf{S}_k), \tag{37}$$

According to the work of [33], we can construct a minorizer using the following inequality

$$\begin{aligned}f_m(\mathbf{S}) &= \mathrm{tr}\left(\overline{\mathbf{X}}_R^H\left(\sigma_H^2\overline{\mathbf{X}}_R\overline{\mathbf{X}}_R^H + \sigma_R^2\mathbf{I}\right)^{-1}\overline{\mathbf{X}}_R\right)\\ &\ge 2\Re\left\{\mathrm{tr}(\mathbf{Q}_k^H\overline{\mathbf{X}}_R)\right\} - \mathrm{tr}(\mathbf{T}_k(\sigma_H^2\overline{\mathbf{X}}_R\overline{\mathbf{X}}_R^H + \sigma_R^2\mathbf{I}))\end{aligned} \tag{38}$$

where

$$\mathbf{Q}_k = \sigma_H^4(\sigma_H^2\overline{\mathbf{X}}_{R,k}\overline{\mathbf{X}}_{R,k}^H + \sigma_R^2\mathbf{I})^{-1}\overline{\mathbf{X}}_{R,k}, \mathbf{T}_k = \mathbf{Q}_k\mathbf{Q}_k^H/\sigma_H^4. \tag{39}$$

Proposition 5: By defining

$$\mathbf{E}_R = \begin{bmatrix}\mathbf{I}_N\otimes(\mathbf{E}_1\boldsymbol{\Omega}_\varphi)\\ \mathbf{I}_N\otimes(\mathbf{E}_2\boldsymbol{\Omega}_\varphi)\\ \vdots\\ \mathbf{I}_N\otimes(\mathbf{E}_P\boldsymbol{\Omega}_\varphi)\end{bmatrix}, \tag{40}$$

we are able to minorize $f_m(\overline{\mathbf{X}})$ by

$$g_m(\mathbf{S};\mathbf{S}_k) = c_k - 2\Re\left\{\mathrm{vec}(\mathbf{S}^\top)^H\mathbf{b}_k\right\} - \mathrm{vec}(\mathbf{S}^\top)^H\mathbf{B}_k\mathrm{vec}(\mathbf{S}^\top). \tag{41}$$

where $c_k = -\mathrm{tr}(\sigma_R^2\mathbf{T}_k)$, and

$$\mathbf{b}_k = -\mathbf{E}_R^H\mathrm{vec}(\mathbf{Q}_k), \tag{42}$$

$$\mathbf{B}_k = \mathbf{E}_R^H(\sigma_H^2\mathbf{I}\otimes\mathbf{T}_k)\mathbf{E}_R \succeq \mathbf{0}. \tag{43}$$

Proof: See section Appendix E.

Then we can express the minorizing problem at $k+1$-th iteration as

$$\begin{aligned}\min_{\mathbf{S}}\ & 2\Re\left\{\mathrm{vec}(\mathbf{S}^\top)^H\mathbf{b}_k\right\} + \mathrm{vec}(\mathbf{S}^\top)^H\mathbf{B}_k\mathrm{vec}(\mathbf{S}^\top)\\ s.t.\ & \left|\Im\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\right| - \Re\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\tan\theta\\ & \qquad \le (-\sqrt{\boldsymbol{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma},\\ & \mathrm{vec}(\mathbf{S}^\top)^H(\mathbf{I}_{N_t}\otimes\boldsymbol{\Phi})\mathrm{vec}(\mathbf{S}^\top) \le E.\end{aligned} \tag{44}$$

Assume that a solution $\mathbf{S}_k$ has been obtained upon the completion of the k-th iteration. Subsequently, by solving problem (44) at the $(k+1)$-th iteration, an optimal solution $\mathbf{S}^\star$ is acquired. We can confidently assert that $f_m(\mathbf{S}^\star) \ge g_m(\mathbf{S}^\star;\mathbf{S}_k) \ge g_m(\mathbf{S}_k;\mathbf{S}_k) = f_m(\mathbf{S}_k)$, which implies that a superior solution for minimizing MMSE can be attained at the $(k+1)$-th iteration. This process can be iteratively repeated to continuously optimize the solution until the convergence is achieved.

We are now ready to present Algorithm 1 to solve the problem (35) based on the discussion above. The detail concerning solving the problem (44) is elaborated in section III-D.

Algorithm 1 Minorization Method for Solving (35)

Require: N_t, L, σ_C^2, σ_R^2, σ_H^2, $\mathbf{D}$, $\mathbf{H}_C$, $\boldsymbol{\Gamma}$, the execution threshold ϵ and the maximum iteration number $i_{\max}$.

Ensure: $\mathbf{S}^\star$

1: initialize $\mathbf{S}_0 \in \mathcal{Q}$ by picking up $\mathbf{S}_{-1}$ randomly and solving problem (44), $k=0$.
2: **repeat**
3: Calculate the $\mathbf{b}_k$ and $\mathbf{B}_k$ by equation (42) and (43).
4: Solve problem (44) to obtain $\mathbf{S}^\star$.
5: $k = k+1$.
6: **until** $\|\mathbf{S}_k - \mathbf{S}_{k-1}\|_F^2 \le \epsilon$ or $i = i_{\max}$.
7: $\mathbf{S}^\star = \mathbf{S}_k$

C. SCA Approach for FTN-ISAC-SLP

In this section, we will develop another optimization scheme to solve the problem (35) by the idea of SCA. To proceed with the SCA algorithm, we approximate the objective function by its first-order Taylor expansion near a given point $\mathbf{S}_k$ (and hence $\overline{\mathbf{X}}_{R,k}$) as

$$f(\mathbf{S}) \approx f(\mathbf{S}_k) + \Re\left\{\mathrm{vec}\left(\frac{\partial f}{\partial\overline{\mathbf{X}}_R}\right)^H\mathrm{vec}\left(\overline{\mathbf{X}}_R - \overline{\mathbf{X}}_{R,k}\right)\right\} \tag{45}$$

where

$$\frac{\partial f}{\partial\overline{\mathbf{X}}_R} = -2\overline{\mathbf{X}}_{R,k}\left(\overline{\mathbf{X}}_{R,k}^H\overline{\mathbf{X}}_{R,k}\right)^{-1}\left(\overline{\mathbf{X}}_{R,k}^H\overline{\mathbf{X}}_{R,k}\right)^{-1} \tag{46}$$

stands for the gradient at the point $\mathbf{S}_k$. By using the fact that $\mathrm{vec}(\overline{\mathbf{X}}_R) = \mathbf{E}_R\mathrm{vec}(\mathbf{S}^\top)$ and define

$$\mathbf{t}_{R,k}^H = \mathrm{vec}\left(\frac{\partial f}{\partial\overline{\mathbf{X}}_R}\right)^H\mathbf{E}_R, \tag{47}$$

we are able to approximate $f(\mathbf{S})$ around $\mathbf{S}_k$ by

$$\begin{aligned}g_l(\mathbf{S};\mathbf{S}_k) \approx\ & \Re\left\{\mathbf{t}_{R,k}^H\mathrm{vec}(\mathbf{S}^\top)\right\}\\ & + f(\mathbf{S}_k) - \Re\left\{\mathrm{vec}\left(\frac{\partial f}{\partial\overline{\mathbf{X}}_R}\right)^H\mathrm{vec}\left(\overline{\mathbf{X}}_{R,k}\right)\right\}.\end{aligned} \tag{48}$$

Then we proceed to solve the following sub-problem (49) in $(k+1)$-th iteration.

$$\begin{aligned}\min_{\mathbf{S}}\ & \Re\left\{\mathbf{t}_{R,i}^H\mathrm{vec}(\mathbf{S}^\top)\right\}\\ s.t.\ & \left|\Im\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\right| - \Re\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\tan\theta\\ & \qquad \le (-\sqrt{\boldsymbol{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma},\\ & \mathrm{vec}(\mathbf{S}^\top)^H(\mathbf{I}_{N_t}\otimes\boldsymbol{\Phi})\mathrm{vec}(\mathbf{S}^\top) \le E.\end{aligned} \tag{49}$$

Suppose that a solution $\mathbf{S}_k$ has been procured at the k-th iteration. By subsequently solving problem (49) during the $(k+1)$-th iteration, an optimal solution $\mathbf{S}^\star$ is obtained. When $\mathbf{S}^\star$ is

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

in close proximity to $\mathbf{S}_k$ and the SCA approximation holds, it follows that $f(\mathbf{S}^\star) \le g_l(\mathbf{S}^\star;\mathbf{S}_k) \le g_l(\mathbf{S}_k;\mathbf{S}_k) = f(\mathbf{S}_k)$. Albeit $\mathbf{S}^\star$ is not necessarily adjacent to $\mathbf{S}_k$, the difference $\mathbf{S}^\star - \mathbf{S}_k$ provides a decent direction for the optimization of $f(\mathbf{S})$. By iteratively taking small steps along the direction of $\mathbf{S}^\star - \mathbf{S}_k$, it is possible to successively obtain superior solutions that minimize MMSE prior to reaching convergence.

With a properly chosen step size $t \in [0,1]$, one may get the $(k+1)$-th iteration point as

$$\mathbf{S}_{k+1} = \mathbf{S}_k + t(\mathbf{S}^\star - \mathbf{S}_k) = (1-t)\mathbf{S}_k + t\mathbf{S}^\star. \tag{50}$$

Since $\mathbf{S}_k, \mathbf{S}^\star \in \mathcal{Q}$ by the definition of convexity, we have $\mathbf{S}_{k+1} \in \mathcal{Q}$, which is a feasible solution to problem (35).

We are now ready to present Algorithm 2 to solve the problem (35) based on the discussion above. The detail concerning solving the problem (44) is elaborated in section III-D.

Algorithm 2 SCA Method for Solving (35)

Require: N_t, L, σ_C^2, σ_R^2, σ_H^2, $\mathbf{D}$, $\mathbf{H}_C$, $\mathbf{\Gamma}$, the execution threshold ϵ and the maximum iteration number $i_{\max}$.
Ensure: $\mathbf{S}^\star$

1: Initialize $\mathbf{S}_0 \in \mathcal{Q}$ by picking up $\mathbf{S}_{-1}$ randomly and solving problem (49), $k=0$.
2: **repeat**
3: Calculate the $\mathbf{t}_{R,k}$ by equation (46) and (47).
4: Solve problem (49) to obtain $\mathbf{S}^\star$.
5: Update the solution by $\mathbf{S}_{k+1} = \mathbf{S}_k + t\left(\mathbf{S}^\star - \mathbf{S}_k\right)$, where t is determined by using the exact line search.
6: $k = k+1$.
7: **until** $\|\mathbf{S}_k - \mathbf{S}_{k-1}\|_F^2 \le \epsilon$ or $i = i_{\max}$.
8: $\mathbf{S}^\star = \mathbf{S}_k$

D. *Efficient Algorithm for Solving Sub-problems (44) and (49)*

Notice that both problems (44) and (49) can be written in the form of

$$\begin{aligned}\min_{\mathbf{S}}\ & 2\Re\left\{\mathrm{vec}(\mathbf{S}^\top)^H\mathbf{a}_k\right\} + \mathrm{vec}(\mathbf{S}^\top)^H\mathbf{A}_k\mathrm{vec}(\mathbf{S}^\top)\\ s.t.\ & \left|\Im\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\right| - \Re\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\tan\theta\\ & \qquad\le (-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma},\\ & \mathrm{vec}(\mathbf{S}^\top)^H(\mathbf{I}_{N_t}\otimes\mathbf{\Phi})\mathrm{vec}(\mathbf{S}^\top) \le E.\end{aligned} \tag{51}$$

which is a QCQP with linear inequality constraint and quadratic energy constraint. Specifically, for problem (49), $\mathbf{A}_k = \mathbf{0}$. This problem is convex and can be solved using the popular PDIP algorithm. However, we are looking for a more efficient way to solve this type of QCQP. There have been many efficient techniques developed to solve real QP with linear inequality constraints, such as problem (52), for both small and large problems, such as active set methods and the Mehrotra predictor-corrector algorithm.

$$\min_{\mathbf{x}}\ \mathbf{x}^\top\mathbf{A}\mathbf{x} + 2\mathbf{x}^\top\mathbf{a}\ s.t.\ \mathbf{B}\mathbf{x} \le \mathbf{b}. \tag{52}$$

Therefore, in order to enhance the efficiency of our algorithm, which involves solving QCQP problems of the form (51), we aim to develop an algorithm that can transform the process of solving QCQP into solving QP problems with only linear inequality constraints. Following this idea, we develop the BPS algorithm, which can convert solving (51) into iteratively solving a small amount of (52).

By letting $\mathbf{P}_\Re = \Re\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\right\}$, $\mathbf{P}_\Im = \Im\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\right\}$, $\mathbf{s}_\Re = \Re\left\{\mathrm{vec}(\mathbf{S}^\top)\right\}$ and $\mathbf{s}_\Im = \Im\left\{\mathrm{vec}(\mathbf{S}^\top)\right\}$, the inequalities (62) can be decomposed to

$$\begin{aligned}&(\mathbf{P}_\Im - \mathbf{P}_\Re\tan\theta)\mathbf{s}_\Re + (\mathbf{P}_\Re + \mathbf{P}_\Im\tan\theta)\mathbf{s}_\Im\\ &\qquad\le (-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma},\\ &(-\mathbf{P}_\Im - \mathbf{P}_\Re\tan\theta)\mathbf{s}_\Re + (-\mathbf{P}_\Re + \mathbf{P}_\Im\tan\theta)\mathbf{s}_\Im\\ &\qquad\le (-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma},\end{aligned} \tag{53}$$

or in single matrix inequality form

$$\begin{bmatrix}\mathbf{P}_\Im - \mathbf{P}_\Re\tan\theta & \mathbf{P}_\Re + \mathbf{P}_\Im\tan\theta\\ -\mathbf{P}_\Im - \mathbf{P}_\Re\tan\theta & -\mathbf{P}_\Re + \mathbf{P}_\Im\tan\theta\end{bmatrix}\begin{bmatrix}\mathbf{s}_\Re\\ \mathbf{s}_\Im\end{bmatrix} \le \begin{bmatrix}(-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}\\ (-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}\end{bmatrix}. \tag{54}$$

Taking

$$\begin{aligned}&\mathbf{\Psi} = \begin{bmatrix}\mathbf{P}_\Im - \mathbf{P}_\Re\tan\theta & \mathbf{P}_\Re + \mathbf{P}_\Im\tan\theta\\ -\mathbf{P}_\Im - \mathbf{P}_\Re\tan\theta & -\mathbf{P}_\Re + \mathbf{P}_\Im\tan\theta\end{bmatrix},\\ &\hat{\mathbf{s}} = \begin{bmatrix}\mathbf{s}_\Re\\ \mathbf{s}_\Im\end{bmatrix}, \boldsymbol{\gamma} = \begin{bmatrix}(-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}\\ (-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}\end{bmatrix}\\ &\hat{\mathbf{A}}_k = \begin{bmatrix}\Re\{\mathbf{B}_k\} & -\Im\{\mathbf{B}_k\}\\ \Im\{\mathbf{B}_k\} & \Re\{\mathbf{B}_k\}\end{bmatrix}, \hat{\mathbf{a}}_k = \begin{bmatrix}\Re\{\mathbf{b}_k\}\\ -\Im\{\mathbf{b}_k\}\end{bmatrix},\\ &\mathbf{\Upsilon} = \begin{bmatrix}\mathbf{I}_{N_t}\otimes\mathbf{\Phi} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{N_t}\otimes\mathbf{\Phi}\end{bmatrix},\end{aligned} \tag{55}$$

we can rewrite problem (44) as

$$\begin{aligned}&\min_{\hat{\mathbf{s}}}\ \hat{\mathbf{s}}^\top\hat{\mathbf{A}}_k\hat{\mathbf{s}} + 2\hat{\mathbf{s}}^\top\hat{\mathbf{a}}_k\\ &s.t.\ \hat{\mathbf{s}}^\top\mathbf{\Upsilon}\hat{\mathbf{s}} \le E,\ \mathbf{\Psi}\hat{\mathbf{s}} \le \boldsymbol{\gamma}.\end{aligned} \tag{56}$$

We then try to remove the energy constraint in (51) by introducing the penalty factor $\rho \ge 0$ and the penalty problem (57).

$$\begin{aligned}\mathcal{K}(\rho) : &\min_{\hat{\mathbf{s}}}\ \hat{\mathbf{s}}^\top(\hat{\mathbf{A}}_k + \rho\mathbf{\Upsilon})\hat{\mathbf{s}} + 2\hat{\mathbf{s}}^\top\hat{\mathbf{a}}_k\\ &s.t.\ \mathbf{\Psi}\hat{\mathbf{s}} \le \boldsymbol{\gamma}\end{aligned} \tag{57}$$

The key idea here is to introduce a regularization term, $\rho\hat{\mathbf{s}}^\top\mathbf{\Upsilon}\hat{\mathbf{s}}$, into the objective function to ensure that the energy term $\hat{\mathbf{s}}^\top\mathbf{\Upsilon}\hat{\mathbf{s}}$ does not become excessively large. However, it is important to note that if ρ is too large, the optimal solution of problem (57) would primarily minimize $\hat{\mathbf{s}}^\top\mathbf{\Upsilon}\hat{\mathbf{s}}$, consequently leading to inadequate reduction of the original objective function $\hat{\mathbf{s}}^\top\hat{\mathbf{B}}_k\hat{\mathbf{s}} + 2\hat{\mathbf{s}}^\top\hat{\mathbf{b}}_k$. Conversely, if ρ is too small, the optimal solution would not sufficiently consider the term $\hat{\mathbf{s}}^\top\mathbf{\Upsilon}\hat{\mathbf{s}}$, potentially resulting in violation of the constraint $\hat{\mathbf{s}}^\top\mathbf{\Upsilon}\hat{\mathbf{s}} \le E$. To address this issue, we propose a binary penalty search (BPS) algorithm, which aims to identify an appropriate value for ρ and effectively solve the problem (56).

Suppose we now have two penalty factors, $\rho_r > \rho_l > 0$, such that solving $\mathcal{K}(\rho_r)$ results in a solution $\hat{\mathbf{s}_r}$ that satisfies the energy constraint $\hat{\mathbf{s}_r}^\top\mathbf{\Upsilon}\hat{\mathbf{s}_r} \le E$, while solving $\mathcal{K}(\rho_l)$ yields a solution $\hat{\mathbf{s}_l}$ that violates the energy constraint. We can infer that there may exist a ρ in the interval (ρ_l, ρ_r), for which solving

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

$\mathcal{K}(\rho)$ provides a solution that more effectively minimizes the objective function than solving $\mathcal{K}(\rho_r)$, while still adhering to the energy constraint.

By realizing the fact above, we set the penalty factor $\rho = (\rho_r + \rho_l)/2$, which is the mid-point of the interval (ρ_l, ρ_r). If solving $\mathcal{K}(\rho)$ yields a new solution that violates the energy constraint, we set $\rho_l = \rho$. Otherwise, we set $\rho_r = \rho$. This approach progressively narrows the search range. By iteratively applying this process, we eventually reach a sufficiently small search range where $\rho_r \approx \rho_l$. The most appropriate penalty factor would then be ρ_r. Consequently, the Binary Penalty Search (BPS) algorithm can be presented in Algorithm 3.

Proposition 6: The solution derived by this algorithm is the optimal solution of the problem (56).

Proof: See section Appendix F.

Algorithm 3 Binary Penalty Search

Require: $\hat{\mathbf{A}}_k$, $\hat{\mathbf{a}}_k$, E, $\mathbf{\Psi}$, $\mathbf{\Upsilon}$, $\boldsymbol{\gamma}$, execution threshold ϵ.
Ensure: optimal solution $\hat{\mathbf{s}}_\star$

```
Initialize ρ_l = 0, ρ_r = ρ_max.
repeat
    ρ = (ρ_l + ρ_r)/2;
    Solve problem (57) by active-set method to obtain ŝ_⋆.
    if ŝ_⋆^T Υ ŝ_⋆ ≤ E then
        ρ_r = ρ;
    else
        ρ_l = ρ;
    end if
until ρ_r − ρ_l ≤ ε
```

Comparison of Time Complexity between Primal-Dual Interior Point and BPS Algorithm: The computational complexity of the Primal-Dual Interior Point Method (PDIP) algorithm for addressing problem (51) typically falls within the bounds of $\mathrm{O}((4NL+1)^{3.5}\log(1/\epsilon))$ to $\mathrm{O}((4NL+1)^{4}\log(1/\epsilon))$. This signifies a marginal elevation in computational complexity in contrast to the problem (57), which exhibits a time complexity of $\mathrm{O}((4NL+1)^{3.5}\log(1/\epsilon))$.

When employing the BPS approach to solve the problem (51), the procedure necessitates the resolution of $\log(\rho_{max})$ instances of the problem (57). Consequently, the algorithmic time complexity for the BPS approach escalates to $\mathrm{O}(\log(\rho_{max}/\epsilon)(4NL)^{3.5}\log(1/\epsilon))$. In essence, this indicates that the BPS algorithm's computational complexity surpasses that of the PDIP algorithm when the constraints imposed by the energy limit are less stringent, signified by a comparably diminished ρ_{max}.

Regarding problem (57), various efficacious algorithms for QP can be adapted to accommodate a diversity of situations. For example, when operating within a relatively small-scale problem space, the active-set strategy can be implemented to lower the practical computational complexity effectively.

IV. Numerical Results

In this section, we provide numerical results to verify the superiority of the proposed FTN-ISAC-SLP approaches. Without loss of generality, we consider an ISAC BS that is equipped with $N_r = 8$ antennas for its receiver. The noise variances are set as $\sigma_C^2 = \sigma_R^2 = 0$ dBm. The quantity of $\boldsymbol{\mu}_{\mathbf{h}}$ has minimal impact on the optimization discussed in this paper and is therefore set to $\mathbf{0}$. The variance of TRM fluctuations is set as $\sigma_H^2 = 20$ dBm, with each element of $\mathbf{H}_R$ drawn from $\mathcal{CN}(0, \sigma_H^2)$. Symbol duration T_0 is set to 1 ms. Each element of $\mathbf{H}_C$ independently is drawn from $\mathcal{CN}(0, \sigma_C^2)$. Without loss of generality, all the communication users are imposed with the same worst-case QoS, i.e., $\Gamma_k = \Gamma, \forall k$.

Fig. 2 presents the constellation plot of the received symbols of the FTN-ISAC-SLP system before noise imposition. Generally, the symbols are distanced from the detection thresholds, namely, the x and y axes. However, some symbols are observed to be closer to the thresholds. This occurrence can be attributed to the noise $\mathrm{vec}(\mathbf{N}_C^\top) \sim \mathcal{CN}(\mathbf{0}, \sigma_C^2 \mathbf{I}K \otimes \mathbf{\Lambda}\phi)$, which is imposed on the symbols and possesses varying variances. Consequently, certain symbols are more susceptible to noise interference, while others remain unaffected. The system adapts by allocating more power to the received symbols subjected to noise with higher variance.

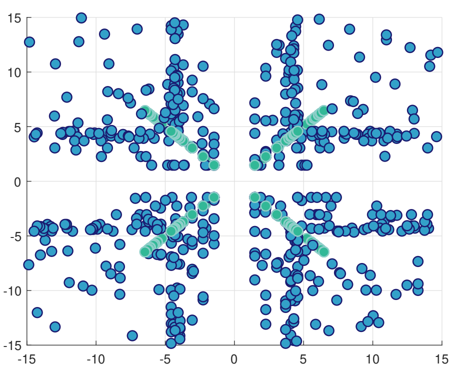

Fig. 2. Constellation plot of the received symbols without noise. Here we take $\tau = 0.9$. The green dots are the nominal constellation points.

In Fig. 3, we compare the convergence behavior of Algorithm 1 (minorization approach) and Algorithm 2 (SCA approach) by illustrating how MMSE changes with the number of iterations. It is observed that the minorization approach reaches a better sensing performance in the early iterations, while the SCA approach ultimately achieves better results after more iterations. The slower convergence of the SCA approach can be due to the fact that the SCA approximation requires the solution to take only small steps along the optimization direction to ensure that the approximation remains valid. This leads to more iterations being needed for the SCA approach to converge to the ultimate solution. However, it is worth noting that despite its slower convergence, the SCA approach eventually results in better sensing performance, which could be advantageous in scenarios where the final performance is of higher importance than the speed of convergence.

In Fig. 4, we compare the performance of the minorization and SCA approaches under different energy budgets E and fixed communication and sensing conditions, using 1000 initial data points for each method. Subsequently, we compute the

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

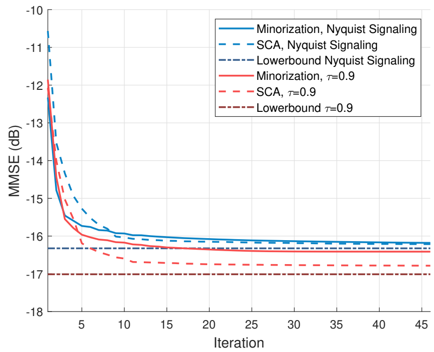

Fig. 3. MMSE versus minorization/SCA iteration in case of $N_t = 3$, $K = 2$, $L = 15$, $P = 3$, $Q = 3$, $\Gamma = 15$ dBm, $E = 30$ dBm.

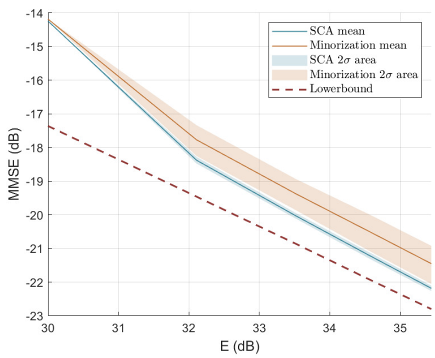

Fig. 4. MMSE versus E in case of $N_t = 3$, $K = 3$, $L = 20$, $P = 3$, $Q = 3$, $\Gamma = 15$ dBm, $\tau = 0.9$.

mean and standard deviation of the MMSE generated by distinct initial points. The 2σ region denotes the area encompassing points situated within a distance of twice the standard deviation from the mean. As the energy budget E increases, both methods exhibit better sensing performance and approach the lower bound of MMSE. This is expected, as a higher energy budget allows for more flexibility in satisfying both the communication and sensing requirements. It is observed that the minorization approach exhibits a larger fluctuation in the results compared to the SCA approach, as indicated by the wider 2σ region. This suggests that the SCA approach might be more robust and consistent in terms of its performance across different initial data points.

In Fig. 5, we compare the results of solving the minorization problems using the BPS algorithm and the Primal-Dual Interior Point (PDIP) algorithm directly. The purpose of this comparison is to evaluate the effectiveness of the BPS algorithm in solving minorization problems, as compared to a well-established optimization algorithm like PDIP. The figure shows that the BPS algorithm and the PDIP algorithm provide essentially the same results at each iteration of the algorithms. Moreover, as proved in the later section of the study, the BPS algorithm yields the optimal solution for the energy-constrained QP problem akin to the PDIP algorithm.

Fig. 6 delineates the relationship between the Central Processing Unit (CPU) execution time and frame length L. The iterations are curtailed once the MMSE reaches a threshold of $-15;$ dBm. As anticipated, an augmentation in L corresponds

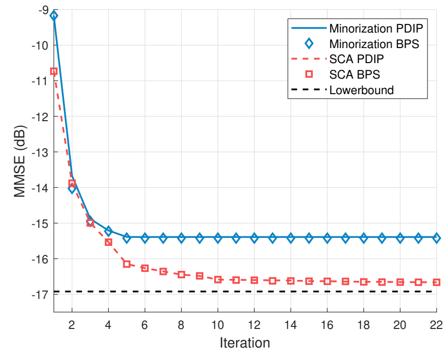

Fig. 5. MMSE versus minorization/SCA iteration in case of $N = 3$, $K = 2$, $L = 15$, $P = 3$, $Q = 3$, $\Gamma = 15$ dBm, $E = 30$ dBm, $\tau = 0.9$.

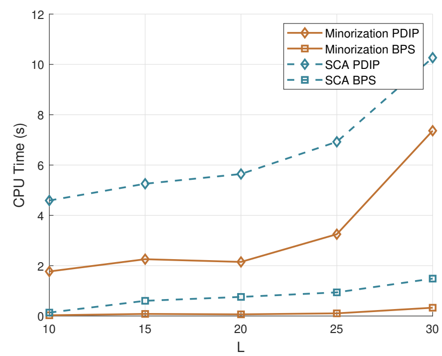

Fig. 6. CPU Time versus L in case of $N_t = 3$, $K = 2$, $P = 3$, $Q = 3$, $\Gamma = 15$ dBm, $E = 35$ dBm.

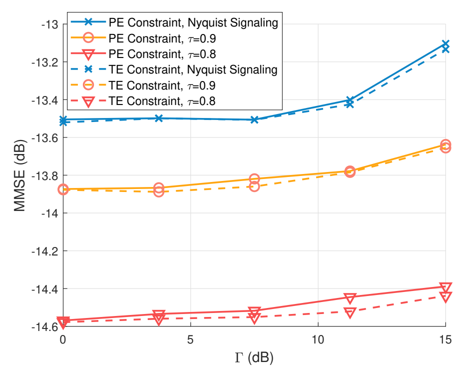

Fig. 7. MMSE versus Γ comparison between total energy (TE) constraint and per-antenna (PE) constraint in case of $N_t = 4$, $K = 3$, $L = 15$, $P = 3$, $Q = 3$, $E = 30$ dBm.

to an expansion in the problem size, leading to a heightened demand for CPU time. The BPS exhibits a more efficient performance in terms of CPU time when juxtaposed with the PDIP method. Moreover, the minorization approach manifests superior convergence speed towards a reasonably low MMSE threshold compared to the SCA approach.

Fig. 7 delineates the comparative analysis between the per-antenna and total energy constraints. For the sake of simplicity, we set each antenna's energy budget as $E_n = E/N_t$. The results indicate a negligible discrepancy between the two constraints. Generally, the MMSE, when applied to the per-antenna energy constraint, is marginally higher than its

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

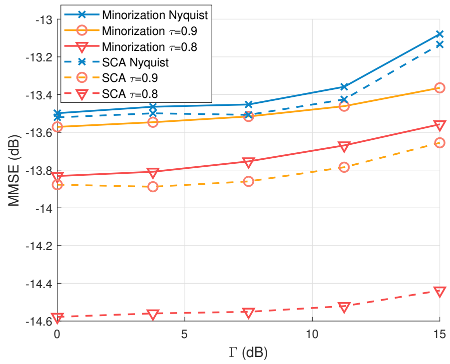

Fig. 8. MMSE versus Γ in case of $N_t = 4$, $K = 3$, $L = 15$, $P = 3$, $Q = 3$, $E = 30$ dBm.

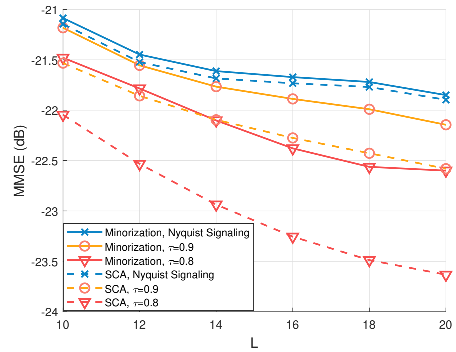

Fig. 11. MMSE versus L in case of $N_t = 4$, $K = 3$, $\Gamma = 15$ dBm, $P = 3$, $Q = 3$, $E = 35$ dBm.

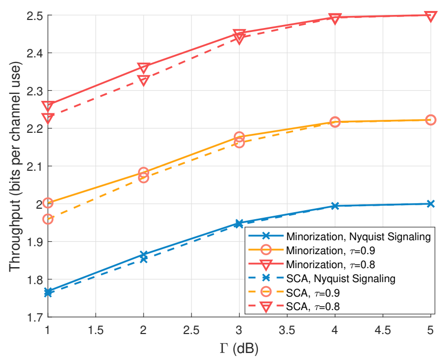

Fig. 9. Throughput versus Γ in case of $N_t = 4$, $K = 3$, $L = 15$, $P = 3$, $Q = 3$, $E = 30$ dBm.

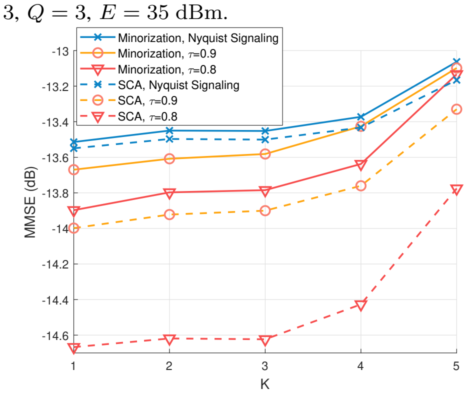

Fig. 12. MMSE versus K in case of $N_t = 5$, $L = 20$, $\Gamma = 15$ dBm, $P = 3$, $Q = 3$, $E = 32$ dBm.

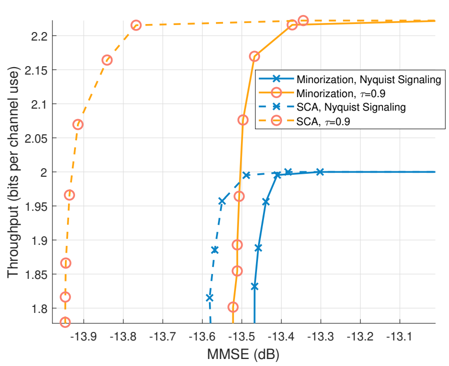

Fig. 10. Throughput versus MMSE in case of $N_t = 4$, $K = 3$, $L = 15$, $P = 3$, $Q = 3$, $E = 30$ dBm.

counterpart under the total energy constraint.

In Fig. 8, the impact of SNR threshold for communication users on the radar estimation MMSE is depicted. It can be observed that with an elevated communications SNR, the estimation performance deteriorates, signifying an intrinsic tradeoff between communication and sensing performance. Concurrently, it is observed that as τ increases, the MMSE exhibits a decline. This phenomenon can be ascribed to the fact that interference in FTN signaling is harnessed to generate a positive impact on sensing performance. The smaller the value of τ, the greater the extent to which one pulse can contribute its energy to adjacent pulses. In communication, such interference may lead to challenges in recovering the correct constellation. However, in sensing, a larger energy contribution often corresponds to enhanced sensing performance.

In Fig. 9, we demonstrate the impact of the SNR threshold for communication on communication throughput. Inspired by the computation of the bit error rate (BER) in [34], we develop a way to calculate throughput from the constellation that guarantees fair comparison under FTN signaling. Let's denote the symbol stream received sans noise as $\mathbf{y}$. In the context of QPSK, the probability of an individual symbol $y \in \mathbf{y}$ being received with an incorrect in-phase sign, under Gaussian noise, is represented as $p_1 = Q(|\Re(y)|/\sqrt{\sigma_n^2/2})$. Concurrently, the probability for an incorrect quadrature sign is given by $p_2 = Q(|\Im(y)|/\sqrt{\sigma_n^2/2})$, where σ_n^2 signifies the noise variance imposed on the symbol y and $Q(\cdot)$ denotes the Q-function. Accordingly, the probability for the correct transmission of the two bits encapsulated within this symbol is calculated as $(1-p_1)(1-p_2)$, and the probability for the correct transmission of a single bit is derived as $(1-p_1)p_2 + p_1(1-p_2)$. The expected number of successfully transmitted bits is quantified as $2(1-p_1)(1-p_2) + (1-p_1)p_2 + p_1(1-p_2)$, which is also $2(1-P_e)$ where P_e is the calculation of BER in [34]. It is important to note that under FTN signaling, this symbol occupies only τT_0 time. To ensure a fair comparison, we normalize the expected number of successfully transmitted bits by τ, which furnishes us with the formula for computing the throughput from a given constellation:

$$\frac{1}{KL}\sum_{y \in \mathbf{y}} \left(2(1-p_1)(1-p_2) + (1-p_1)p_2 + p_1(1-p_2)\right)/\tau.$$

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

It is observed that as the SNR threshold for communication increases, the communication throughput increases, which is essentially a result of received symbols being pushed further from the detection thresholds in the constellation plot. At the same time, it is observed that as τ increases, the throughput increases, which is due to the fact that the same amount of symbols are transmitted in less time. This highlights the effectiveness of FTN signaling in achieving higher throughput for communication while maintaining a balance between communication and sensing performance.

Fig. 10 illustrates the relationship between radar estimation MMSE and throughput. It can be observed that as MMSE increases, the throughput also increases. This indicates an inherent tradeoff between communication and sensing performance. When MMSE is sufficiently large, the throughput converges to the maximum transmission rate, signifying a bit error rate of zero. In this scenario, the system primarily emphasizes communication performance. Furthermore, it is noted that when throughput diminishes, the MMSE converges to a specific value, representing the lower bound of MMSE when the system is solely focused on sensing performance.

In Fig. 11, the influence of frame length L on radar estimation performance is illustrated. It is observed that as the frame length increases, the estimation performance improves, even though communication performance constraints become more stringent. Additionally, in the case of SCA with $\tau = 0.8$, the MMSE displays a convergence trend as L increases, indicating that the sensing performance is limited by the energy constraint.

In Fig. 12, we demonstrate the impact of the communication user number K on radar estimation performance. Unlike increasing the frame length L, which also increases the number of symbols to be transmitted, the estimation performance worsens as K increases. This is because when we increase L by 1, we only tighten the communication constraint on K more symbols. However, when we increase K by 1, we tighten the communication constraint on L more symbols. In the two cases illustrated by Fig. 11 and 12, L in Fig. 12 is significantly larger than K in Fig. 11, resulting in different trends of MMSE.

V. Conclusion

In conclusion, this paper introduces a novel wideband FTN-ISAC-SLP precoding technique for MIMO DFRC systems, adeptly combining FTN signaling and SLP to enhance sensing and communication performance across temporal and spatial dimensions. To tackle the complex non-convex waveform design problem, we develop two algorithm frameworks based on minorization and SCA methods, transforming the problem into solvable QCQP sub-problems. Additionally, we propose a computationally efficient BPS method to solve these sub-problems. Extensive simulation results validate the effectiveness of the proposed FTN-ISAC-SLP design in both radar sensing and multi-user communication performance. The future research may consider to look into not only utilizing the ISI but also the inter-carrier interference (ICI) in the context of multi-carrier transmission, thus to further enhance the performance of the FTN-ISAC-SLP system. Additionally, one may also attempt to develop a more computationally efficient algorithm using the concept of sequential FTN SLP to improve the system's effectiveness.

Appendix A: Proof of Proposition 1

For any i, j that $1 \leq i, j \leq L$, we have

$$\begin{aligned}
&\mathbb{E}[\eta_k(iT)\eta(jT)_k^*] \\
&= \mathbb{E}\left[\int \mathbf{n}_c(t)\varphi^*(t-iT)dt \int \mathbf{n}_c^*(t)\varphi(t-jT)dt\right] \\
&= \mathbb{E}\left[\int\int \mathbf{n}_c(t_1)\mathbf{n}_c(t_2)^*\varphi^*(t_1-iT)\varphi(t_2-jT)dt_1dt_2\right] \\
&= \int\int \mathbb{E}[\mathbf{n}_c(t_1)\mathbf{n}_c(t_2)^*]\varphi^*(t_1-iT)\varphi(t_2-jT)dt_1dt_2 \\
&= \int\int \sigma_C^2\delta(t_1-t_2)\varphi^*(t_1-iT)\varphi(t_2-jT)dt_1dt_2 \\
&= \int \sigma_C^2\varphi^*(t-iT)\varphi(t-jT)dt \\
&= \int \sigma_C^2\varphi^*(t)\varphi(t-(i-j)T)dt \\
&= \sigma_C^2\phi((i-j)T) = \sigma_C^2\mathbf{\Phi}_{i,j}.
\end{aligned} \tag{58}$$

Thus $\mathbb{E}[\boldsymbol{\eta}_k\boldsymbol{\eta}_k^H] = \sigma_C^2\mathbf{\Phi}$.

Appendix B: Proof of Proposition 2

We can express $\overline{\mathbf{X}}_C$ as

$$\overline{\mathbf{X}}_C = [\mathbf{E}_1\mathbf{X}_C^\top, \mathbf{E}_2\mathbf{X}_C^\top, \cdots, \mathbf{E}_P\mathbf{X}_C^\top], \tag{59}$$

which yields a way to vectorize $\overline{\mathbf{X}}_C$ with respect to $\text{vec}(\mathbf{S}^\top)$

$$\begin{aligned}
\text{vec}(\overline{\mathbf{X}}_C) &= \begin{bmatrix} \text{vec}(\mathbf{E}_1\mathbf{X}_C^\top) \\ \text{vec}(\mathbf{E}_2\mathbf{X}_C^\top) \\ \vdots \\ \text{vec}(\mathbf{E}_P\mathbf{X}_C^\top) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{N_t}\otimes\mathbf{E}_1 \\ \mathbf{I}_{N_t}\otimes\mathbf{E}_2 \\ \vdots \\ \mathbf{I}_{N_t}\otimes\mathbf{E}_P \end{bmatrix} \text{vec}(\mathbf{X}_C^\top) \\
&= \begin{bmatrix} \mathbf{I}_{N_t}\otimes(\mathbf{E}_1\mathbf{\Omega}_\phi) \\ \mathbf{I}_{N_t}\otimes(\mathbf{E}_2\mathbf{\Omega}_\phi) \\ \vdots \\ \mathbf{I}_{N_t}\otimes(\mathbf{E}_P\mathbf{\Omega}_\phi) \end{bmatrix} \text{vec}(\mathbf{S}^\top).
\end{aligned} \tag{60}$$

By using the fact that $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^\top \otimes \mathbf{A})\text{vec}(\mathbf{X})$, we have

$$\begin{aligned}
\text{vec}(\mathbf{Y}_C^\top) =& (\mathbf{H}_C \otimes (\mathbf{GU}_\phi)^\top)\text{vec}(\overline{\mathbf{X}}_C) + \text{vec}(\mathbf{N}_C^\top) \\
=& \overline{\mathbf{H}}_C\text{vec}(\mathbf{S}^\top) + \text{vec}(\mathbf{N}_C^\top),
\end{aligned} \tag{61}$$

Appendix C: Proof of Proposition 3

Define $\mathbf{\Gamma} = \text{Diag}([\Gamma_1, \Gamma_2, \cdots, \Gamma_K])$, $\overline{\mathbf{D}} = \text{Diag}(\text{vec}(\mathbf{D}^\top))$ and $\boldsymbol{\varsigma} = \sqrt{\text{diag}(\sigma_C^2\mathbf{I}_K \otimes \mathbf{\Lambda}_\phi)}$; then, the CI constraint for k users can be formulated as

$$\begin{aligned}
&\left|\Im\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\text{vec}(\mathbf{S}^\top)\right\}\right| - \Re\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\text{vec}(\mathbf{S}^\top)\right\}\tan\theta \\
&\qquad\qquad \leq (-\sqrt{\mathbf{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}.
\end{aligned} \tag{62}$$

We first stack the k linear inequalities

$$|\Im\{\mathbf{d}_k^* \circ \mathbf{y}_{C,k}\}| - \Re\{\mathbf{d}_k^* \circ \mathbf{y}_{C,k}\}\tan\theta \leq (-\sqrt{\Gamma_k}\tan\theta)\boldsymbol{\sigma} \tag{63}$$

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

in a column to form a united equality:

$$\left|\Im\left\{\begin{bmatrix}\mathbf{d}_1\\\mathbf{d}_2\\\vdots\\\mathbf{d}_k\end{bmatrix}^*\circ\begin{bmatrix}\mathbf{y}_{C,1}\\\mathbf{y}_{C,2}\\\vdots\\\mathbf{y}_{C,k}\end{bmatrix}\right\}\right|-\Re\left\{\begin{bmatrix}\mathbf{d}_1\\\mathbf{d}_2\\\vdots\\\mathbf{d}_k\end{bmatrix}^*\circ\begin{bmatrix}\mathbf{y}_{C,1}\\\mathbf{y}_{C,2}\\\vdots\\\mathbf{y}_{C,k}\end{bmatrix}\right\}\tan\theta$$
$$=\left|\Im\left\{\mathrm{vec}(\mathbf{D}^\top)\circ\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\right|$$
$$-\Re\left\{\mathrm{vec}(\mathbf{D}^\top)\circ\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\tan\theta\le\begin{bmatrix}(-\sqrt{\Gamma_1}\tan\theta)\boldsymbol{\varsigma}\\(-\sqrt{\Gamma_2}\tan\theta)\boldsymbol{\varsigma}\\\vdots\\(-\sqrt{\Gamma_k}\tan\theta)\boldsymbol{\varsigma}\end{bmatrix}. \tag{64}$$

Notice that for Hardmard product of two vectors we have $\mathbf{a}\circ\mathbf{b}=\mathrm{Diag}(\mathbf{a})\mathbf{b}$. Thus we can rewrite (64) as

$$\left|\Im\left\{\mathrm{Diag}(\mathrm{vec}(\mathbf{D}^\top))\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\right|$$
$$-\Re\left\{\mathrm{Diag}(\mathrm{vec}(\mathbf{D}^\top))\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\tan\theta$$
$$=\left|\Im\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\right|-\Re\left\{\overline{\mathbf{D}}^*\overline{\mathbf{H}}_C\mathrm{vec}(\mathbf{S}^\top)\right\}\tan\theta$$
$$\le\begin{bmatrix}(-\sqrt{\Gamma_1}\tan\theta)\boldsymbol{\varsigma}\\(-\sqrt{\Gamma_2}\tan\theta)\boldsymbol{\varsigma}\\\vdots\\(-\sqrt{\Gamma_k}\tan\theta)\boldsymbol{\varsigma}\end{bmatrix}=(-\sqrt{\boldsymbol{\Gamma}\otimes\mathbf{I}_L}\tan\theta)\boldsymbol{\varsigma}. \tag{65}$$

Appendix D: Proof of Proposition 4

$$\int\|x_n(t)\|^2dt=\int\left\|\sum_{i=1}^{L}\varphi(t-(i-1)T)s_{n,i}\right\|^2dt$$
$$=\int\left(\sum_{i=0}^{L-1}\varphi(t-iT)s_{n,i+1}\right)\left(\sum_{i=0}^{L-1}\varphi(t-iT)^*s_{n,i+1}^*\right)dt$$
$$=\int\sum_{i=1}^{L}\sum_{j=1}^{L}\left(\varphi(t-(i-1)T)\varphi(t-(j-1)T)^*s_{n,i}s_{n,j}^*\right)dt$$
$$=\sum_{i=1}^{L}\sum_{j=1}^{L}\left(\int\varphi(t-(i-1)T)\varphi(t-(j-1)T)^*dt\right)s_{n,i}s_{n,j}^*dt$$
$$=\sum_{i=1}^{L}\sum_{j=1}^{L}\left(\int\varphi(t)\varphi(t-(i-j)T)^*dt\right)s_{n,i}s_{n,j}^*dt$$
$$=\sum_{i=1}^{L}\sum_{j=1}^{L}\phi((i-j)T)s_{n,i}s_{n,j}^*=\mathbf{s}_n^H\boldsymbol{\Phi}\mathbf{s}_n. \tag{66}$$

Appendix E: Proof of Proposition 5

Similar to the vectorization of $\overline{\mathbf{X}}_C$, we are able to vectorize $\overline{\mathbf{X}}$ as

$$\mathrm{vec}(\overline{\mathbf{X}}_R)=\mathrm{vec}([\mathbf{E}_1\mathbf{X}_R^\top,\mathbf{E}_2\mathbf{X}_R^\top,\cdots,\mathbf{E}_P\mathbf{X}_R^\top])$$
$$=\begin{bmatrix}\mathbf{I}_N\otimes(\mathbf{E}_1\boldsymbol{\Omega}_\varphi)\\\mathbf{I}_N\otimes(\mathbf{E}_2\boldsymbol{\Omega}_\varphi)\\\vdots\\\mathbf{I}_N\otimes(\mathbf{E}_P\boldsymbol{\Omega}_\varphi)\end{bmatrix}\mathrm{vec}(\mathbf{S}^\top)=\mathbf{E}_R\mathrm{vec}(\mathbf{S}^\top). \tag{67}$$

Using the fact that $\mathrm{tr}(\mathbf{A}^H\mathbf{B})=\mathrm{vec}^H(\mathbf{A})\mathrm{vec}(\mathbf{B})$ and $\mathrm{tr}(\mathbf{ABCD})=\mathrm{vec}^\top(\mathbf{D})(\mathbf{A}\otimes\mathbf{C}^\top)\mathrm{vec}(\mathbf{B}^\top)$, we have

$$\begin{aligned}\mathrm{tr}(\mathbf{Q}_k^H\overline{\mathbf{X}}_R)&=\mathrm{vec}^H(\mathbf{Q}_k)\mathrm{vec}(\overline{\mathbf{X}}_R)\\&=\mathrm{vec}^H(\mathbf{Q}_k)\mathbf{E}_R\mathrm{vec}(\mathbf{S}^\top)=-\mathbf{b}_k^H\mathrm{vec}(\mathbf{S}^\top),\end{aligned} \tag{68}$$

$$\begin{aligned}\mathrm{tr}(\mathbf{T}_k\overline{\mathbf{X}}_R\overline{\mathbf{X}}_R^H)&=\mathrm{tr}(\mathbf{I}_L\overline{\mathbf{X}}_R^H\mathbf{T}_k\overline{\mathbf{X}}_R)\\&=\mathrm{vec}^\top(\overline{\mathbf{X}}_R)(\mathbf{I}_L\otimes\mathbf{T}_k^*)\mathrm{vec}(\overline{\mathbf{X}}_R^*)\\&=\mathrm{vec}(\mathbf{S}^\top)^\top\mathbf{E}_R^\top(\mathbf{I}_L\otimes\mathbf{T}_k^*)\mathbf{E}_R^*\mathrm{vec}(\mathbf{S}^\top)^*\\&=\mathrm{vec}(\mathbf{S}^\top)^H\mathbf{E}_R^H(\mathbf{I}_L\otimes\mathbf{T}_k)\mathbf{E}_R\mathrm{vec}(\mathbf{S}^\top)\\&=\mathrm{vec}(\mathbf{S}^\top)^H\mathbf{B}_k\mathrm{vec}(\mathbf{S}^\top)/\sigma_H^2.\end{aligned} \tag{69}$$

Thus we have

$$\begin{aligned}&2\Re\left\{\mathrm{tr}(\mathbf{Q}_k^H\overline{\mathbf{X}}_R)\right\}-\mathrm{tr}(\mathbf{T}_k(\sigma_H^2\overline{\mathbf{X}}_R\overline{\mathbf{X}}_R^H+\sigma_R^2\mathbf{I}))\\&=\mathrm{tr}(\sigma_R^2\mathbf{T}_k)+2\Re\left\{\mathrm{tr}(\mathbf{Q}_k^H\overline{\mathbf{X}}_R)\right\}-\sigma_H^2\mathrm{tr}(\mathbf{T}_k\overline{\mathbf{X}}_R\overline{\mathbf{X}}_R^H)\\&=c_k-2\Re\left\{\mathrm{vec}(\mathbf{S}^\top)^H\mathbf{b}_k\right\}-\mathrm{vec}(\mathbf{S}^\top)^H\mathbf{B}_k\mathrm{vec}(\mathbf{S}^\top).\end{aligned} \tag{70}$$

Appendix F: Proof of Proposition 6

The proof of the convergence of the BPS algorithm to the optimal solution is presented below. Consider the following convex problem:

$$\begin{aligned}&\min_{\mathbf{x}}\ f_o(\mathbf{x})\\&s.t.\ \mathbf{x}\in\mathcal{S},\ f_p(\mathbf{x})\le\mathcal{E}.\end{aligned} \tag{71}$$

where f_o and f_p are convex functions and $\mathcal{S}$ is a convex region. Also consider the penalty problem, which eliminates the energy constraint

$$\begin{aligned}\mathcal{P}(\rho):&\min_{\mathbf{x}}\ f_o(\mathbf{x})+\rho f_p(\mathbf{x})\\&s.t.\ \mathbf{x}\in\mathcal{S}.\end{aligned} \tag{72}$$

Initially, we will prove that as ρ increases, the optimal solution of $\mathcal{P}(\rho)$, $\mathbf{x}^\star$, exhibits a larger $f_o(\mathbf{x}^\star)$ and a smaller $f_p(\mathbf{x}^\star)$. Subsequently, we will establish that the optimal solution of problem (71) is also the optimal solution of problem $\mathcal{P}(\rho)$ for a certain ρ.

$$f_o(\mathbf{x}_1^\star)+\rho_2f_p(\mathbf{x}_1^\star)\ge f_o(\mathbf{x}_2^\star)+\rho_2f_p(\mathbf{x}_2^\star), \tag{73}$$
$$f_o(\mathbf{x}_1^\star)+\rho_1f_p(\mathbf{x}_1^\star)\le f_o(\mathbf{x}_2^\star)+\rho_1f_p(\mathbf{x}_2^\star). \tag{74}$$

Then we rearrange the inequality (74) as

$$\begin{aligned}&f_o(\mathbf{x}_1^\star)+\rho_2f_p(\mathbf{x}_1^\star)+(\rho_1-\rho_2)f_p(\mathbf{x}_1^\star)\\&\le f_o(\mathbf{x}_2^\star)+\rho_2f_p(\mathbf{x}_2^\star)+(\rho_1-\rho_2)f_p(\mathbf{x}_2^\star).\end{aligned} \tag{75}$$

According to inequality (73) and (75) we have

$$\begin{aligned}0&\le(f_o(\mathbf{x}_1^\star)+\rho_2f_p(\mathbf{x}_1^\star))-(f_o(\mathbf{x}_2^\star)+\rho_2f_p(\mathbf{x}_2^\star))\\&\le(\rho_1-\rho_2)(f_p(\mathbf{x}_2^\star)-f_p(\mathbf{x}_1^\star)),\end{aligned} \tag{76}$$

which yields

$$f_p(\mathbf{x}_1^\star)\le f_p(\mathbf{x}_2^\star). \tag{77}$$

Then combining inequality (73) and (77) we have

$$f_o(\mathbf{x}_1^\star)\ge f_o(\mathbf{x}_2^\star) \tag{78}$$

Inequality (77) and (78) reveal that an increase in the penalty factor results in an increment of the objective function and a reduction in the penalty function. The objective of the BPS

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3372186

algorithm is to identify the minimal penalty factor such that the penalty function f_p does not surpass $\mathcal{E}$.

We now prove that the optimal solution $\mathbf{x}^\star$ of problem (71) corresponds to the optimal solution of problem $\mathcal{P}(\rho)$ for a particular ρ. Assuming $\mathbf{x}^\star$ is not constrained by $f_p(\mathbf{x}) \leq \mathcal{E}$, it is evident that $\mathbf{x}^\star$ represents the optimal solution of $\mathcal{P}(0)$. If it is constrained, in accordance with the KKT complementary slackness condition, we ascertain that for certain values of μ and $\boldsymbol{\nu}$:

$$\frac{\partial f_o}{\partial \mathbf{x}}(\mathbf{x}^\star) + \mu \frac{\partial f_p}{\partial \mathbf{x}}(\mathbf{x}^\star) + \frac{\partial \mathcal{L}_{\mathcal{S}}(\boldsymbol{\nu})}{\partial \mathbf{x}}(\mathbf{x}^\star) = 0, \tag{79}$$

where $\mathcal{L}_{\mathcal{S}}$ and $\boldsymbol{\nu}$ denote the Lagrange augmentation function and dual variables for the constraint $\mathbf{x} \in \mathcal{S}$. Consequently, we deduce that this condition also represents the KKT complementary slackness condition for problem $\mathcal{P}(\mu)$. Therefore, $\mathbf{x}^\star$ serves as the optimal solution of $\mathcal{P}(\mu)$.

References

[1] D. K. Pin Tan, J. He, Y. Li, A. Bayesteh, Y. Chen, P. Zhu, and W. Tong, "Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions," in *2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S)*, 2021, pp. 1–6.

[2] L. Zheng, M. Lops, Y. C. Eldar, and X. Wang, "Radar and communication coexistence: An overview: A review of recent methods," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 85–99, 2019.

[3] 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 Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[4] Z. Feng, Z. Fang, Z. Wei, X. Chen, Z. Quan, and D. Ji, "Joint radar and communication: A survey," *China Communications*, vol. 17, no. 1, pp. 1–27, 2020.

[5] J. Li and P. Stoica, "MIMO radar with colocated antennas," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, 2007.

[6] P. M. McCormick, S. D. Blunt, and J. G. Metcalf, "Simultaneous radar and communications emissions from a common aperture, part i: Theory," in *2017 IEEE Radar Conference (RadarConf)*. IEEE, 2017, pp. 1685–1690.

[7] J. Qian, M. Lops, L. Zheng, X. Wang, and Z. He, "Joint system design for coexistence of MIMO radar and MIMO communication," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3504–3519, 2018.

[8] B. Tang, H. Wang, L. Qin, and L. Li, "Waveform design for dual-function MIMO radar-communication systems," in *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE, 2020, pp. 1–5.

[9] P. Kumari, S. A. Vorobyov, and R. W. Heath, "Adaptive virtual waveform design for millimeter-wave joint communication–radar," *IEEE Transactions on Signal Processing*, vol. 68, pp. 715–730, 2019.

[10] F. Liu, C. Masouros, T. Ratnarajah, and A. Petropulu, "On range sidelobe reduction for dual-functional radar-communication waveforms," *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1572–1576, 2020.

[11] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Transactions on Signal Processing*, vol. 66, no. 16, pp. 4264–4279, 2018.

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

[13] Z. Cheng, B. Liao, and Z. He, "Hybrid transceiver design for dual-functional radar-communication system," in *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE, 2020, pp. 1–5.

[14] W. Yuan, F. Liu, C. Masouros, J. Yuan, D. W. K. Ng, and N. González-Prelcic, "Bayesian predictive beamforming for vehicular networks: A low-overhead joint radar-communication approach," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 1442–1456, 2020.

[15] C. Xu, B. Clerckx, and J. Zhang, "Multi-antenna joint radar and communications: Precoder optimization and weighted sum-rate vs probing power tradeoff," *IEEE Access*, vol. 8, pp. 173 974–173 982, 2020.

[16] N. Su, F. Liu, and C. Masouros, "Secure radar-communication systems with malicious targets: Integrating radar, communications and jamming functionalities," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 83–95, 2020.

[17] A. Bazzi and M. Chafii, "On outage-based beamforming design for dual-functional radar-communication 6g systems," *IEEE Transactions on Wireless Communications*, 2023.

[18] Y. Xiong, F. Liu, Y. Cui, W. Yuan, T. X. Han, and G. Caire, "On the fundamental tradeoff of integrated sensing and communications under gaussian channels," *IEEE Transactions on Information Theory*, 2023.

[19] J. B. Anderson, F. Rusek, and V. Öwall, "Faster-than-Nyquist signaling," *Proceedings of the IEEE*, vol. 101, no. 8, pp. 1817–1830, 2013.

[20] 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 Transactions on Signal Processing*, vol. 68, pp. 3929–3944, 2020.

[21] L. Chen, F. Liu, J. Liu, and C. Masouros, "Composite signalling for DFRC: Dedicated probing signal or not?" *arXiv preprint arXiv:2009.03528*, 2020.

[22] F. Liu and C. Masouros, "Joint beamforming design for extended target estimation and multiuser communication," in *2020 IEEE Radar Conference (RadarConf20)*. IEEE, 2020, pp. 1–6.

[23] D. Spano, M. Alodeh, S. Chatzinotas, and B. Ottersten, "Faster-than-Nyquist signaling through spatio-temporal symbol-level precoding for the multiuser miso downlink channel," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 5915–5928, 2018.

[24] C. Masouros and E. Alsusa, "Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1396–1404, 2009.

[25] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3628–3640, 2015.

[26] M. Alodeh, D. Spano, A. Kalantari, C. G. Tsinos, D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Symbol-level and multicast precoding for multiuser multiantenna downlink: A state-of-the-art, classification, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1733–1757, 2018.

[27] A. Li, D. Spano, J. Krivochiza, S. Domouchtsidis, C. G. Tsinos, C. Masouros, S. Chatzinotas, Y. Li, B. Vucetic, and B. Ottersten, "A tutorial on interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 796–839, 2020.

[28] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "Joint symbol-level precoding and reflecting designs for irs-enhanced mu-miso systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 798–811, 2020.

[29] R. Liu, M. Li, Q. Liu, A. L. Swindlehurst, and Q. Wu, "Intelligent reflecting surface based passive information transmission: A symbol-level precoding approach," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 7, pp. 6735–6749, 2021.

[30] D. Garmatyuk, J. Schuerger, K. Kauffman, and S. Spalding, "Wideband ofdm system for radar and communications," in *2009 IEEE radar conference*. IEEE, 2009, pp. 1–6.

[31] I. I. Immoreev and P. D. V. Fedotov, "Ultra wideband radar systems: advantages and disadvantages," in *2002 IEEE Conference on Ultra Wideband Systems and Technologies (IEEE Cat. No. 02EX580)*. IEEE, 2002, pp. 201–205.

[32] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., 1993.

[33] B. Tang, J. Liu, H. Wang, and Y. Hu, "Constrained radar waveform design for range profiling," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1924–1937, 2021.

[34] P. K. Vitthaladevuni and M.-S. Alouini, "Ber computation of 4/m-qam hierarchical constellations," *IEEE Transactions on Broadcasting*, vol. 47, no. 3, pp. 228–239, 2001.

Authorized licensed use limited to: University College London. Downloaded on March 21,2024 at 17:20:35 UTC from IEEE Xplore. Restrictions apply.
© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.