

On the Trade-off Between Angle of Arrival and Symbol Estimation in Bistatic ISAC Systems Using Unitary Signaling

Sebastian Fodor^b, Gábor Fodor^{*†}, Miklós Telek^{‡#}

^bStockholm University, Stockholm, Sweden. E-mail: sebbifodor@fastmail.com

^{*}Ericsson Research, Stockholm, Sweden. E-mail: Gabor.Fodor@ericsson.com

[†]KTH Royal Institute of Technology, Stockholm, Sweden. E-mail: gaborf@kth.se

[‡]Budapest University of Technology and Economics, Budapest, Hungary. E-mail: telek@hit.bme.hu

[#]HUNREN-BME Information Systems Research Group, Budapest, Hungary. E-mail: telek@hit.bme.hu

Abstract—Previous works in array processing have proposed two types of snapshot models for the angle of arrival (AoA) estimation problem in multi-antenna systems. The deterministic model assumes that the source waveforms are non-random, while the random sensor noise is white Gaussian with a known covariance matrix. The stochastic model assumes that both the waveforms and the noise are zero-mean Gaussian. Interestingly, the performance of these two models have rarely been compared in integrated sensing and communication (ISAC) systems. Therefore, in this paper, we consider the uplink of a bistatic ISAC system that uses unitary constant envelope signaling and pilot-based channel estimation while transmitting a sensing signal simultaneously with the communication signals. The base station uses both the pilot and data signals to estimate the angle of a passive source and the transmitted data symbol by an active (connected) user equipment device. For this system, we derive the classical Cramér-Rao bound for unbiased estimators of the AoA and the transmitted symbol, along with the Bayesian Cramér-Rao bound, which bounds the error of all estimators. We also derive the ISAC-aware minimum mean squared error receiver for both the deterministic and stochastic models. We study the trade-off between sensing and communication under the deterministic and stochastic waveform assumptions. Specifically, we show that the fundamental trade-off between sensing and communication power allocations is expressed differently in the deterministic and stochastic models and argue that the results serve as basic considerations when designing pilot and sensing signals for ISAC systems.

Index terms— angle of arrival estimation, Cramér-Rao bound, integrated sensing and communication.

I. INTRODUCTION

The emerging sixth generation (6G) of mobile networks is expected to be a multi-functional network that integrates communications, localisation, sensing, computation and security to offer a wide range of services in a spectral efficient fashion [1]–[3]. Integrated sensing and communication (ISAC) has recently emerged as one of the key 6G enabling technologies, which extends the capabilities of traditional localisation schemes by the ability of detecting, tracking and estimating the parameters of both active (connected) user equipment (UE) devices and passive objects in the environment [4]–[6]. In fact,

recently, several research projects have demonstrated that 5G pilot signals, designed for synchronisation and channel state information (CSI) acquisition – such as the synchronisation signal block (SSB) and demodulation reference signal (DMRS) of 5G New Radio systems – can be used as sensing signals in integrated passive radar and communication systems [7], [8].

Recognizing the advantages of reusing pilot and communication signals as illuminators of opportunity in passive radars [9], recent works have studied the fundamental trade-offs between communication and sensing in terms of the achievable spectral and energy efficiency, detection probability and accuracy of parameter estimation of target objects [5], [10]–[12]. In particular, establishing Cramér-Rao Bounds (CRBs) for parameter estimation has provided valuable insights in some of the inherent trade-offs in ISAC systems operating as either monostatic, bistatic or multistatic radars [10], [13], [14]. Specifically, when the signals of UE devices are used for sensing purposes (as illustrated in Figure 1), there is a trade-off not only between allocating transmit power to pilot and data symbols – which has been studied for long [15], [16], albeit not in ISAC systems – but also between the energy dedicated for CSI acquisition, and maintaining appropriate levels of signal-to-noise ratio (SNR) for both communications and sensing [8], [12], [17].

Several papers have proposed mechanisms whereby the total available power at an ISAC transmitter can be either used jointly for sensing and communications or split between generating the sensing and communication signals [10], [19]–[23]. Such a power split can be achieved in the frequency domain [10], [19] or in the spatial domain using multiple antennas [4], [20]–[23]. Specifically, in analog beamforming architectures, the multibeam technique to generate suitable communication and sensing beams was proposed in, for example, [14], [20], [21]. When hybrid and fully digital architectures are available, suitable precoding techniques for generating ISAC signals was studied in [4], [22]–[24]. When such techniques are employed at a UE device, the UE can allocate a certain portion of its transmit power resources to generating a sensing signal, while using the remaining power to transmit suitable uplink pilot (reference) signals and information-carrying data signals

G. Fodor was supported by the 6G-MUSICAL EU Horizon 2023 project, Project ID: 101139176. M. Telek was supported by the OTKA K-138208 project of the Hungarian Scientific Research Fund.

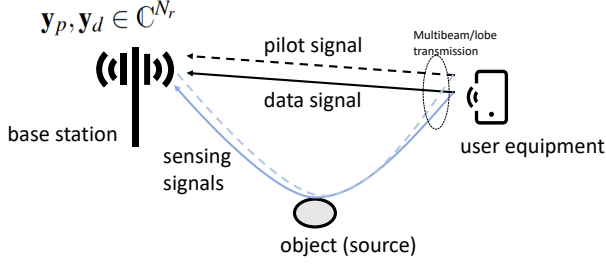


Figure 1. A simple bistatic ISAC system consisting of a multi-antenna user equipment (UE) that transmits a pilot and a unitary constant envelope data signal in two subsequent time instances, a passive object (referred to as the source of the sensing signal, see [18, Chapter 12], and a multi-antenna base station equipped with N_r receive antennas. We assume that the UE utilizes multi-antenna transmission to transmit the pilot/data symbols and simultaneously the sensing symbols, as proposed in, for example, [4].

(Figure 1).

Previous works in array processing have proposed two types of snapshot models of the angle of arrival (AoA) estimation problem in multi-antenna systems. The deterministic model assumes that the source (sensing) waveforms are nonrandom, while the stochastic model assumes that both the waveforms and the noise are zero-mean Gaussian, as discussed in details in [18, Chapter 12], and illustrated in Figure 1. It is interesting to note that the rigorous analysis, including establishing CRBs for deterministic (also called "conditional") and stochastic (also called "unconditional") waveforms dates back to the 90's, see [25], [26], and briefly discussed in [18, Chapter 12], while more recent works on sensing with random signals include [27], [28].

It is important to distinguish these two models, since they result in different AoA CRBs, and different maximum likelihood (ML) direction finding techniques [26], [29]. As discussed also in [18], the stochastic model is applicable for cases of Gaussian waveforms and a great number of snapshots, while the deterministic model is more suitable in scenarios with unknown statistics of the source waveforms and small sample size, where it is natural to treat the waveforms as unknown deterministic variables. While both models are insightful, and have been widely used for both establishing CRBs and developing practical AoA estimation methods – including multiple signal classification (MUSIC) or estimation of signal parameters via rotational invariance (ESPRIT)-based algorithms – the performance of these two models have not been compared in bistatic ISAC systems, where both the AoA and the transmitted communication symbol must be estimated and the transmit power resources must be split between pilot signaling, data transmission and sensing. The most important related works are summarized in Table I.

In this paper, we argue that – using a high level of abstraction, as in [18, Chapter 12] and assuming a single radar target, as illustrated in Figure 1 – studying and deriving CRBs

for estimating the AoA and the transmitted communication symbol in the uplink of a simple ISAC system are conceptually insightful and practically useful. Since the classical CRB provides a bound for unbiased estimators only, it has limited use when studying e.g. minimum mean squared error (MMSE) estimators, which may be biased. For this reason we also derive the Bayesian Cramér-Rao Bound (BCRB), which bounds the error of all estimators [33]. In particular, we assume unitary constant envelope signaling – that is the data symbols are zero-mean non-correlated unitary normalized symbols, as in e.g. [34], [35], [36], [37] – for communication and aim to derive and compare the sensing and communication CRBs in the following four cases, which capture the aspects of the sensing waveform being deterministic or stochastic and whether the bound is the classical CRB or the BCRB:

- **Case 1 (CRB) and Case 2 (BCRB):** In these cases, we use the deterministic waveform assumption, which is applicable for cases, in which the source signal is predefined and known by the receiver [18].
- **Case 3 (CRB) and Case 4 (BCRB):** In these cases, the sensing waveform is a zero-mean complex Gaussian random variable, which is a suitable model for a priori unknown sensing signals with known second-order statistics.

Arguably, these CRBs remain valid for any M -ary phase-shift keying (PSK) signaling, where the possible values of the ϕ are selected from a predetermined set. In practice, the estimation error of ϕ incurs symbol errors, which depend on the signal constellation used in the system, see for example see for example [38], [39], and more recently [40], [41]. However, the connection between the CRBs established in this paper and the symbol error is out of the scope of this paper.

The scenario, in which we study the above four cases differ from that studied in [12] to establish some fundamental trade-offs of ISAC systems in the following aspects:

- Reference [12] considers the scenario in which the communication and sensing receivers are separate, and therefore, there is no interference between the communication and sensing signals.
- Additionally, reference [12] assumes that the sensing receiver knows the transmitted communications symbol, while the ISAC receiver studied in this paper estimates the transmitted symbol and the AoA from a passive object simultaneously.
- Our paper uses a separate communication and sensing symbol, which will be later denoted by $\mathbf{p}$ and x respectively.

For these cases, we derive CRBs for both the transmitted communication symbol and the AoA. Note that under the unitary constant-envelope communication assumption, estimating the symbol is equivalent to estimating the phase of the complex symbol (whose magnitude is assumed to be 1) [35], [36], [37]. Interestingly, in all four cases, closed forms of the associated CRBs can be derived. As the symbolic expressions for the CRBs and the numerical examples illustrate, various engineering insights can be derived. One of the important

Table I
OVERVIEW OF RELATED LITERATURE

Reference	Main scope/contribution	Addressed ISAC trade-offs	Are CRB or Bayesian CRB (BCRB) derived?	Deterministic or stochastic sensing signal	Key performance indicators	Comment
Chiriyath et al., [10]	Bounds on the perf. of joint radar and comm. co-existence	Multiple access rate region vs. Fisher information when dividing the total BW and power	CRB for time-delay estimation, data rate and estimation rate bounds	Stochastic sensing signal	Achievable rate region, Fisher information, estimation rate	AoA is out of scope due to single-antenna model.
Liu et al., [19]	Adaptive OFDM integrated waveform	Trade-off between data inf. rate and mutual information for radar target identification and classification	No	Stochastic sensing signal	SNR, data rate, mutual information	Transmit power and waveform are adaptive, AoA is out of scope.
Liu et al., [4]	Joint radar and communication (comprehensive) design, Integrating radar sensing in MIMO UL/DL communication systems	UL/DL Spectral efficiency vs. parameter (e.g. delay/Doppler) estimation	No	Stochastic sensing signal	SNR, data rate, spectral efficiency, parameter estimation MSE	Pilot/data/sensing trade-off in bi-static sensing is not modelled.
Huang et al., [30]	Multiple ISAC transmitters coordinated by a central controller	Communication SNR vs. CRB for target location estimation	CRB for location estimation is used as a constraint in various opt. problems	Stochastic sensing signal	CRB for location estimation, SNR for communication, total transmit power	AoA and pilot modeling are out of scope.
Behdad et al., [31]	Comm. and multistatic sensing in a cloud radio access network	Detection probability vs. used power via power allocation for ISAC using a maximum a posteriori ratio test detector	No	Stochastic sensing signal	Detection probability, total power consumption	AoA and CRB derivations are out of scope.
Liu et al., [32]	Several performance trade-offs and theoretical limits, signal processing aspects and relation of ISAC to overall 6G network design	Various trade-offs at the PHY layer and cross-layers	CRB for delay estimation is discussed/derived	Deterministic and stochastic sensing signals are discussed (not systematically compared)	SNR, parameter estimation, detection probability, data rate, etc.	The three-way trade-off between pilot/data/sensing power and comparing CRB/BCRB are out of scope.
Baig et al., [14]	Bistatic BS-to-BS ISAC system, SNR and CRB are derived	Communication rate vs. AoA estimation performance in the time and power domains	Classical CRB for AoA estimation is derived	Stochastic sensing signal	Data rate, AoA estimation error	Pilot signals are not modelled. Simulation study: No analytical derivations of performance metrics.
Xu et al., [23]	Cellular network-based ISAC, where several cells are coordinated in a coordinated multipoint fashion to minimize the beam pattern mismatch error.	SINR achieved for communication vs. reliability of detection under power budget constraint	No	Stochastic sensing signal	Achieved quality of service for communication users and target sensing/detection	AoA and CRB calculations are out of scope.
Xiong et al., [12]	Fundamental trade-offs in point-to-point ISAC systems under Gaussian channels	Achievable communication rate vs. CRB for sensing	CRBs for various estimated parameters based on sensing	Both stochastic and deterministic sensing signals are considered	Achievable data rate, CRB for various sensing parameters	Pilot/data signals are not modelled. ISAC-aware MIMO receiver is out of scope.
Present paper	CRB and BCRB for both the deterministic and the stochastic sensing models	CRB for symbol estimation vs CRB for AoA estimation	CRB and BCRB with varying degree of available prior information	Deterministic and stochastic sensing signals	CRBs and variances of estimates	CRB, BCRB and estimate variances are investigated as the function of the allocated power levels to sensing vs. communication

insights is that the inherent ISAC trade-off is much more articulated in the random waveform case, while the deterministic waveform allows for an almost independent allocation of the sensing and communication resources, since the impact of the communication and sensing signals on each other is much less in the deterministic model than when using random waveforms for sensing.

The paper is structured as follows. The next section describes the system and signal models. Next, Section III and Section IV derive CRBs for the deterministic and stochastic models respectively. Section V develops the maximum likelihood symbol and AoA estimators, while Section VI describes the MMSE channel and symbol estimators. Section VII discusses numerical results, and Section VIII summarizes the most important insights, draws conclusions and proposes open research questions.

II. SYSTEM MODEL

We consider a single-user multiple input multiple output (SU-MIMO) bistatic ISAC system, in which the UE device transmits both sensing and communication (i.e. pilot or data) signals [3], [4], [20], see Figure 1. In general, there are P objects that reflect the sensing signals.

In this paper, we assume unitary constant envelope signaling, that is the communication symbol x is uniformly distributed over the complex unit circle, $x = e^{i\phi}$, where ϕ is a random variable with support $[-\pi, +\pi]$. Furthermore, we will denote the transmitted sensing symbol by η , where both x and η are scalars.

As proposed by e.g. [20], [14] and [42], the transmitter node can use proper precoding to spatially separate the communication and sensing symbols and avoid or minimize the effect that the communication signal has on the sensing process. The combined transmitted symbol then becomes (see e.g. [42] for a combined transmitted symbol model):

$$\mathbf{w}x + \mathbf{w}_s\eta \in \mathbb{C}^{N_t}, \quad (1)$$

where $\mathbf{w}$ and $\mathbf{w}_s$ denote the $\mathbb{C}^{N_t}$ precoding vectors for the communication and sensing symbols respectively, and the scalars x and η denote the transmitted communication symbol and sensing symbol respectively and N_t denotes the number of transmit antennas.

The received pilot and data signals at the BS then become:

$$\begin{aligned} \mathbf{y}_p &= \underbrace{\mathbf{A}(\theta)}_{N_r \times P} \underbrace{\mathbf{D}}_{P \times P} \underbrace{\mathbf{H}_t(\mathbf{w}s^* + \mathbf{w}_s\eta)}_{P \times 1} \\ &\quad + \alpha\sqrt{P_p} \underbrace{\mathbf{H}(\mathbf{w}s^* + \mathbf{w}_s\eta)}_{N_r \times 1} + \mathbf{n}_p \in \mathbb{C}^{N_r}, \text{ and} \\ \mathbf{y}_d &= \mathbf{A}(\theta)\mathbf{D}\mathbf{H}_t(\mathbf{w}x + \mathbf{w}_s\eta) \\ &\quad + \alpha\sqrt{P_d}\mathbf{H}(\mathbf{w}x + \mathbf{w}_s\eta) + \mathbf{n}_d \in \mathbb{C}^{N_r}, \end{aligned} \quad (2)$$

respectively, where the diagonal sensing matrix, which collects the sensing path loss $\alpha_{s,i}$ (which includes the radar cross section of the object), and the sensing power P_s for each

Table II
SYSTEM PARAMETERS

Notation	Meaning
N_r	Number of receive antennas at BS
$P_p, P_d, P_s; P_{\text{TOT}} = \max(P_p, P_d) + P_s$	Transmit power of pilot, data and sensing signals respectively, and their total sum.
P	Number of objects (targets).
$\mathbf{p} \in \mathbb{C}^P$	Sensing signals from the P objects (referred to as sources in [18, Chapter 12]).
$\theta_p \in [-\pi/2, \pi/2]$	Angle of arrival of object p , $p = 1 \dots P$
$\mathbf{D} \in \mathbb{C}^{P \times P}$	Diagonal sensing matrix.
$\mathbf{a}(\theta_p) \triangleq [\dots e^{i2\pi j \ell \sin(\theta_p)} \dots]^T \in \mathbb{C}^{N_r}$	Steering vector associated with θ_p , $j = 0 \dots N_r - 1$, where ℓ denotes the antenna spacing per wavelength ratio.
$\boldsymbol{\theta} = [\theta_1, \dots, \theta_P]^T$	Vector of AoA-s from the P objects.
$\mathbf{A}(\boldsymbol{\theta}) \triangleq [\mathbf{a}(\theta_1) \dots \mathbf{a}(\theta_P)] \in \mathbb{C}^{N_r \times P}$	Steering matrix [18].
$s; s^*$	Transmitted uplink pilot symbol and its complex conjugate.
$x = e^{i\phi}; x = 1; \phi \in [-\pi, +\pi]$	Transmitted unitary normalized uplink data symbol.
$\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$; where $\text{vec}(\mathbf{H}) \sim \mathcal{CN}(\mathbf{0}, \text{Cov}(\text{vec}(\mathbf{H})))$	Communication channel matrix between the UE equipped with N_t transmit antennas and the BS equipped with N_r receive antennas.
$\mathbf{w} \in \mathbb{C}^{N_t}$	Transmit precoder at the UE.
$\mathbf{h} \triangleq \mathbf{H}\mathbf{w} \in \mathbb{C}^{N_r}$; $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$	Complex (effective) channel between UE-BS.
α	Path loss between the UE and the BS.
α_s	Aggregate path loss between (1) the UE and the object and (2) the object and the BS, including the (complex) radar cross section (RCS) of the object.
$\mathbf{n}_p, \mathbf{n}_d \in \mathbb{C}^{N_r}$	Additive white Gaussian noise (AWGN) at the receiver when receiving the pilot and data signals.
$\mathbf{C} \in \mathbb{C}^{N_r \times N_r}$	Stationary covariance matrix of the effective (block-fading) communication channel.
$\mathbf{y}_p; \mathbf{y}_d \in \mathbb{C}^{N_r}$	Received uplink pilot and data signals at the BS.
σ^2	Variance of the AWGN.
$\boldsymbol{\Omega} \in \mathbb{C}^{P \times P}$	Covariance matrix of the P -dimensional sensing signal in the stochastic model.
$\mathbf{I}(\phi, \boldsymbol{\theta}); \tilde{\mathbf{I}}(\phi, \boldsymbol{\theta})$	Fisher information matrix in the deterministic/stochastic model.
$Z(\mathbf{u}, \mathbf{v}, \phi, \boldsymbol{\theta}); \tilde{Z}(\mathbf{u}, \mathbf{v}, \phi, \boldsymbol{\theta})$	Loglikelihood function in the deterministic/stochastic model.

object, is defined as:

$$\mathbf{D} \triangleq \begin{bmatrix} \alpha_{s,1}\sqrt{P_s} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_{s,P}\sqrt{P_s} \end{bmatrix} \in \mathbb{R}^{P \times P}, \quad (3)$$

$\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{C}^{N_r \times P}$ is the steering matrix, and s is the uplink

pilot symbol (see also Table II). Furthermore, $\mathbf{H}_t$ is the $P \times N_t$ transmitter-objects channel matrix, whose p^{th} row contains the N_t dimensional channel between the transmitter and the p^{th} object and $\mathbf{H}$ is the $N_r \times N_t$ dimensional communication channel matrix between the transmitter (UE) and the BS, and α , P_p and P_d denote the communication path loss, pilot and data transmit powers respectively.

In this paper, similarly to [20], [14] and [42], we assume that $N_t > P$, that is the number of transmit antennas must be greater than the number of objects that we intend to sense and that the transmitter has perfect CSI so that it can apply proper precoding such that $\mathbf{w}$ is in the null-space of $\mathbf{H}_t$ (meaning that the transmitter zero-forces the communication symbol away from the objects):

$$\mathbf{H}_t \mathbf{w} = \mathbf{0}, \text{ and } \mathbf{H} \mathbf{w}_s \approx \mathbf{0}. \quad (4)$$

Indeed, when such transmit precoding is employed, substituting (4) into (2), we have:

$$\mathbf{y}_p = \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \underbrace{(\mathbf{H}_t \mathbf{w}_s \eta)}_{\triangleq \mathbf{p}} + \alpha \sqrt{P_d} \underbrace{(\mathbf{H} \mathbf{w}_s)}_{\triangleq \mathbf{h}} + \mathbf{n}_p \in \mathbb{C}^{N_r},$$

$$\mathbf{y}_d = \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} (\mathbf{H}_t \mathbf{w}_s \eta) + \alpha \sqrt{P_d} (\mathbf{H} \mathbf{w}_s) + \mathbf{n}_d \in \mathbb{C}^{N_r}, \quad (5)$$

where $\mathbf{p} \triangleq \mathbf{H}_t \mathbf{w}_s \eta \in \mathbb{C}^P$ is the sensing signal, and $\mathbf{h} \triangleq \mathbf{H} \mathbf{w}_s \in \mathbb{C}^{N_r}$ is the effective communication channel. The parameters characterizing this system are summarized in Table II. As discussed in [18], given the above definition of the sensing signal $\mathbf{p}$, the sensing signal can be advantageously modelled as a zero-mean complex normal random variable if $\mathbf{H}_t$ contains complex normal distributed elements. However, $\mathbf{p}$ can also be modeled as a deterministic signal under the assumption that a good estimate of $\mathbf{H}_t$ is available at the receiver.

Note that this signal model is an extension of the radar signal models used in [3], [5], [14], [18] by explicitly distinguishing the received pilot and data signals ($\mathbf{y}_p$ and $\mathbf{y}_d$). We will assume that the UE is equipped with multiple transmit antennas that enable it to divide its total transmit power between the communication (i.e. pilot and data) signals and the sensing signal, such that their sum $P_{\text{TOT}} = \max(P_p, P_d) + P_s$ remains under a power budget dictated by physical limitations and regulatory constraints on the uplink of cellular systems [43].

III. CRAMÉR-RAO BOUND IN THE DETERMINISTIC MODEL

To proceed, notice that in a SU-MIMO system, the pilot symbol can be set, without loss of generality, to $s = 1$. Also, recall that the transmitted symbol is of the form of $x = e^{i\phi}$. It will be convenient to rewrite (5) as:

$$\mathbf{y}_p = \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \mathbf{p} + \alpha \sqrt{P_p} \mathbf{h} + \mathbf{n}_p, \quad (6)$$

$$\mathbf{y}_d = \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \mathbf{p} + \alpha \sqrt{P_d} \mathbf{h} e^{i\phi} + \mathbf{n}_d. \quad (7)$$

Recalling the system model and Table II in the previous section, notice that in these expressions $\mathbf{y}_p$, $\mathbf{y}_d$ are observed random variables, $\mathbf{h}$, $\mathbf{n}_p$, $\mathbf{n}_d$ are non-observed random variables, α , $\sqrt{P_p}$, $\sqrt{P_d}$, $\mathbf{D}$, $\mathbf{p}$, $\mathbf{C}$ and σ^2 are known parameters, and $\mathbf{A}(\cdot)$ is a known function. In practice, the serving BS maintains an estimate of the large scale parameters (path loss) and channel covariance matrices for mobility (handover), and

quality of service (QoS) management [44], [45], link failure prediction [46] and power control [47] purposes based on measurement reports on reference signals (RSRP) delivered by the UE. Based on such measurement reports, the BS continuously can maintain an estimate of the channel covariance matrix with sophisticated techniques [48], [49], [50], [51], [52], [53]. Furthermore, in practice, the BS continuously estimates the thermal noise variance as well as the total received interference power, which are often expressed in the form of reference signal received quality (RSRQ), based on UE measurement reports and measurements at the BS [54], [55].

We intend to estimate $\{\phi, \boldsymbol{\theta}\}$ based on $\mathbf{y}_p$, $\mathbf{y}_d$. In the deterministic model [18, Chapter 12], for a given value of $\{\phi, \boldsymbol{\theta}\}$, $\mathbf{y}_p$ and $\mathbf{y}_d$ are complex normally distributed random vectors with expected value:

$$\boldsymbol{\mu}(\boldsymbol{\theta}) \triangleq \mathbb{E} \left[\begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix} \middle| \{\phi, \boldsymbol{\theta}\} \right] = \begin{pmatrix} \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \mathbf{p} \\ \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \mathbf{p} \end{pmatrix}, \quad (8)$$

where $\mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \mathbf{p} = \sum_{i=1}^P \mathbf{a}(\theta_i) \alpha_{s,i} \sqrt{P_s} \mathbf{p}_i$. Using the above notation, we can first state a theorem, that will be useful for proving a proposition regarding the Fisher information matrix (FIM) in the deterministic model.

Let the scalar-matrix function $\Psi(\phi)$ be defined as follows:

$$\begin{aligned} \Psi(\phi) &\triangleq \mathbb{E} \left[\left(\begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \left(\begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \middle| \{\phi, \boldsymbol{\theta}\} \right] \\ &= \begin{pmatrix} \alpha^2 P_p \mathbf{C} + \sigma^2 \mathbf{I}_{N_r} & \alpha^2 \sqrt{P_p P_d} \mathbf{C} e^{-i\phi} \\ \alpha^2 \sqrt{P_p P_d} \mathbf{C} e^{i\phi} & \alpha^2 P_d \mathbf{C} + \sigma^2 \mathbf{I}_{N_r} \end{pmatrix}. \end{aligned} \quad (9)$$

That is,

$$\begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix} \sim \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\theta}), \Psi(\phi)). \quad (10)$$

Theorem 1. For $\Psi^{-1}(\phi)$, the following holds:

$$\Psi^{-1}(\phi) = \begin{pmatrix} \mathcal{F}_1(\mathbf{C}) & \mathcal{F}_2(\mathbf{C}) e^{-i\phi} \\ \mathcal{F}_2(\mathbf{C}) e^{i\phi} & \mathcal{F}_4(\mathbf{C}) \end{pmatrix}, \quad (11)$$

where

$$\mathcal{F}_1(u) = \frac{\sigma^2 + \alpha^2 P_d u}{\sigma^2 (\sigma^2 + \alpha^2 (P_p + P_d) u)},$$

$$\mathcal{F}_2(u) = \frac{-\alpha^2 \sqrt{P_p P_d} u}{\sigma^2 (\sigma^2 + \alpha^2 (P_p + P_d) u)},$$

$$\mathcal{F}_4(u) = \frac{\sigma^2 + \alpha^2 P_p u}{\sigma^2 (\sigma^2 + \alpha^2 (P_p + P_d) u)},$$

and the matrix function $\mathcal{F}(\mathbf{C})$ of Hermitian matrix $\mathbf{C}$ is defined based on the spectral decomposition of $\mathbf{C}$. If $\mathbf{C} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^H$ is the spectral decomposition of $\mathbf{C}$, where $\boldsymbol{\Lambda} = \text{diag}(\lambda_i)$ then $\mathcal{F}(\mathbf{C}) = \mathbf{U} \mathcal{F}(\boldsymbol{\Lambda}) \mathbf{U}^H = \mathbf{U} \text{diag}(\mathcal{F}(\lambda_i)) \mathbf{U}^H$.

The proof of Theorem 1 is in Appendix A.

While Theorem 1 applies generally to any covariance matrix $\mathbf{C}$, a case of interest is when $\mathbf{C}$ is diagonal or proportional to the identity matrix. Lemma 1 below states when these special cases occur.

Lemma 1. If $\mathbf{H}$ follows a complex multivariate normal distribution such that the covariance of the i 'th and j 'th rows of $\mathbf{H}$ is

$$\mathbb{E}[(\mathbf{H}_{i*})^H \mathbf{H}_{j*}] = \mathbf{C}_{\mathbf{H}}^{(i,j)} \in \mathbb{C}^{N_t \times N_t},$$

then $\mathbf{C}$ is diagonal if the structure of $\mathbf{C}_{\mathbf{H}}^{(i,j)}$ is as follows:

$$\mathbf{C}_{\mathbf{H}}^{(i,j)} = \mathbf{0} \text{ for } i \neq j \implies \mathbf{C} = \text{diag}(\lambda_i), \quad (12)$$

where the λ_i -s were introduced in Theorem 1. Additionally, if $\mathbf{C}_{\mathbf{H}}^{(i,j)}$ has the following structure:

$$\mathbf{C}_{\mathbf{H}}^{(i,j)} = \begin{cases} \mathbf{0} & \text{if } i \neq j \\ \mathbf{C}_{\mathbf{H}}^{(1,1)} & \text{otherwise} \end{cases} \quad (13)$$

then $\mathbf{C} = c\mathbf{I}_{N_r}$ is proportional to the identity matrix with $c = \mathbf{w}^H \mathbf{C}_{\mathbf{H}}^{(1,1)} \mathbf{w}$.

Proof. Note that the (i, j) element of $\mathbf{C}$, denoted by $\mathbf{C}_{ij}$, is obtained as:

$$\begin{aligned} \mathbf{C}_{ij} &= \mathbb{E}[(\mathbf{h})_i (\mathbf{h})_j^*] = \mathbb{E}[(\mathbf{H}\mathbf{w})_i (\mathbf{H}\mathbf{w})_j^*] \\ &= \mathbf{w}^H \mathbb{E}[(\mathbf{H}_{j*})^H \mathbf{H}_{i*}] \mathbf{w} = \mathbf{w}^H \mathbf{C}_{\mathbf{H}}^{(j,i)} \mathbf{w}, \end{aligned} \quad (14)$$

from which both parts of the Lemma follow. $\square$

Now we are in the position of stating the following proposition about the FIM when using the deterministic model.

For $i, j = 1, \dots, P$, the elements of the FIM are defined as

$$\mathbf{I}(\phi, \theta)_{1,1} = -\mathbb{E} \frac{\partial^2}{\partial \phi^2} \log f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v}),$$

$$\mathbf{I}(\phi, \theta)_{1,i+1} = \mathbf{I}(\phi, \theta)_{i+1,1} = -\mathbb{E} \frac{\partial^2}{\partial \phi \partial \theta_i} \log f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v}),$$

$$\mathbf{I}(\phi, \theta)_{i+1,j+1} = -\mathbb{E} \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v}),$$

where $\frac{\partial^2}{\partial \theta_i \partial \theta_j} = \frac{\partial^2}{\partial \theta_j^2}$ for $i = j$.

Proposition 1. The elements of the FIM are as follows:

$$\mathbf{I}(\phi, \theta)_{1,1} = \text{tr} \left(\left(\frac{\partial^2}{\partial \phi^2} \Psi^{-1}(\phi) \right) \Psi(\phi) \right), \quad (15)$$

$$\mathbf{I}(\phi, \theta)_{1,i+1} = \mathbf{I}(\phi, \theta)_{i+1,1} = 0, \quad (16)$$

$$\mathbf{I}(\phi, \theta)_{i+1,j+1} = 2\text{Re} \left[\left(\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\theta) \right)^H \Psi^{-1}(\phi) \left(\frac{\partial}{\partial \theta_j} \boldsymbol{\mu}(\theta) \right) \right], \quad (17)$$

where $\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\theta)$ is defined in (49). Calculating the FIM, and hence the CRB has time complexity $\mathcal{O}(N_r^{2.4})$.

The proof of Proposition 1 is in Appendix B. Note that since Proposition 1 derives the classical CRB, ϕ is assumed to be a deterministic unknown parameter. Hence, the CRB holds for any signaling where the data symbol lies on the unit circle, such as the PSK signaling. The following corollaries will be useful to obtain numerical results without extensive simulation experiments.

Corollary 1. Consider the example of $P = 1$ (i.e. the receiver estimates the AoA from a single object), $N_r = 2$ and $\mathbf{C} = c\mathbf{I}$. In this simple example, – using the notation $\alpha_s = \alpha_{s,1}$ – the

diagonal elements of the FIM become:

$$\begin{aligned} \mathbf{I}(\phi, \theta_1)_{11} &= \frac{4c^2 P_p P_d \alpha^4}{c(P_p + P_d) \alpha^2 \sigma^2 + \sigma^4} \text{ and} \\ \mathbf{I}(\phi, \theta_1)_{22} &= \frac{8\pi^2 \ell^2 P_s \alpha_s^2 \cos^2(\theta_1) (c(P_p + P_d) \alpha^2 + 2\sigma^2 - 2c\sqrt{P_p P_d} \alpha^2 \cos(\phi))}{c(P_p + P_d) \alpha^2 \sigma^2 + \sigma^4}, \end{aligned} \quad (18)$$

where ℓ denotes the antenna spacing per wavelength ratio (see Table II).

It is interesting to note that the $\mathbf{I}(\phi, \theta_1)_{11}$ element is not a function of either θ_1 or ϕ , while $\mathbf{I}(\phi, \theta_1)_{22}$ is a function of both. In particular, as intuitively expected, $\mathbf{I}(\phi, \theta_1)_{22}$ becomes zero when the AoA of the single object, θ_1 , becomes $\frac{\pi}{2}$. On the other hand, from the point of view of estimating θ_1 , $\mathbf{I}(\phi, \theta_1)_{22}$ is maximal in ϕ , when $\phi = \pi$ and $\cos(\phi) = -1$. In contrast, the Fisher information is minimal, when $\phi = 0$ and $\cos(\phi) = 1$. The terms $\alpha\sqrt{P_p}\mathbf{h}$ and $\alpha\sqrt{P_d}\mathbf{h}e^{j\phi}$ appear as noise in (6) and (7) respectively when estimating θ_1 . Hence, it is intuitive that the Fisher information with respect to θ_1 is maximal when the pilot and data signals in (6) and (7) have a perfect negative correlation.

As (18) suggests, $\mathbf{I}(\phi, \theta_1)_{22}$ may benefit from a priori available information on the AoA. This is illustrated by the following example.

Corollary 2. Consider the example defined in Corollary 1, assume that $P_p = P_d$, and that the AoA estimator at the BS has the a priori information that θ_1 is uniformly distributed in the interval $[-\frac{\pi}{4}, \frac{\pi}{4}]$, and ϕ is uniformly distributed over $[-\pi, \pi]$. Then, the $\mathbf{I}_{22}$ element of the a posteriori FIM is

$$\mathbf{I}_{22} = \frac{8\ell^2 \pi (\pi + 2) P_s \alpha_s^2 (cP_1 \alpha^2 + \sigma^2)}{2cP_1 \alpha^2 \sigma^2 + \sigma^4}, \quad (19)$$

where we used the notation $P_1 = P_p = P_d$.

(19) is obtained from (18) as

$$\mathbf{I}_{22} = \frac{1}{2\pi} \int_{\phi=-\pi}^{\pi} \frac{2}{\pi} \int_{\theta_1=-\pi/4}^{\pi/4} \mathbf{I}(\phi, \theta_1)_{22} d\theta_1 d\phi.$$

When the total power budget is not fully utilized ($P_s < P_{\text{TOT}} - P_1$) in (19) then increasing P_s and keeping P_1 fix, increases $\mathbf{I}_{22}$. When the total power budget is utilized, i.e. $P_s = P_{\text{TOT}} - P_1$, the derivative of (19) with respect to P_1 is negative for $0 < P_1 < P_{\text{TOT}}$ which indicates that $\mathbf{I}_{22}$ is maximized at $P_1 = 0$ (i.e. the total power is devoted to sensing), which is align with intuition.

IV. CRAMÉR-RAO BOUND IN THE STOCHASTIC MODEL

According to the discussion in [18, Chapter 12], the stochastic model assumes that both the noise and the source waveforms are zero-mean Gaussian. In this stochastic case, in addition to the known parameters discussed for the deterministic model in the previous section, the covariance matrix of the source waveform is assumed to be known, and here it is denoted as $\boldsymbol{\Omega} \in \mathbb{C}^{P \times P}$. That is:

$$\mathbf{p} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega}). \quad (20)$$

It is reasonable to assume that the BS knows the distribution of the source waveform. This is because the BS is in full control

of configuring the pilot (reference) signals, the UE transmit power and other transmission parameters used by the UE [56, Chapter 5].

The received pilot and data signals remain the same as in the deterministic model discussed in the previous section, given in (6) and (7). For the joint distribution of $\mathbf{y}_p$ and $\mathbf{y}_d$, we now have:

$$\begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix} \sim \mathcal{CN}(\mathbf{0}, \tilde{\Psi}(\phi, \theta)), \quad (21)$$

where

$$\tilde{\Psi}(\phi, \theta) \triangleq \mathbb{E} \left[\begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix} \begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix}^H \middle| \{\phi, \theta\} \right].$$

To determine $\tilde{\Psi}(\phi, \theta)$, notice that:

$$\begin{aligned} \tilde{\Psi}(\phi, \theta) &= \Psi(\phi) + \mathbb{E} \left[\begin{pmatrix} \mathbf{A}(\theta)\mathbf{D}\mathbf{p} \\ \mathbf{A}(\theta)\mathbf{D}\mathbf{p} \end{pmatrix} \begin{pmatrix} \mathbf{A}(\theta)\mathbf{D}\mathbf{p} \\ \mathbf{A}(\theta)\mathbf{D}\mathbf{p} \end{pmatrix}^H \right] \\ &= \Psi(\phi) + \begin{pmatrix} \mathbf{A}(\theta)\mathbf{D} \\ \mathbf{A}(\theta)\mathbf{D} \end{pmatrix} \Omega \begin{pmatrix} \mathbf{A}(\theta)\mathbf{D} \\ \mathbf{A}(\theta)\mathbf{D} \end{pmatrix}^H \\ &= \begin{pmatrix} \alpha^2 P_p \mathbf{C} + \sigma^2 \mathbf{I}_{N_r} + \mathbf{M}(\theta) & \alpha^2 \sqrt{P_p P_d} \mathbf{C} e^{-i\phi} + \mathbf{M}(\theta) \\ \alpha^2 \sqrt{P_p P_d} \mathbf{C} e^{i\phi} + \mathbf{M}(\theta) & \alpha^2 P_d \mathbf{C} + \sigma^2 \mathbf{I}_{N_r} + \mathbf{M}(\theta) \end{pmatrix}, \end{aligned} \quad (22)$$

where $\mathbf{M}(\theta) = \mathbf{A}(\theta)\mathbf{D}\Omega\mathbf{D}^H\mathbf{A}^H(\theta)$. When $\Omega = \text{Diag}(\omega_1, \dots, \omega_P)$, then $\mathbf{M}(\theta) = \sum_{i=1}^P \omega_i \alpha_s^2 P_s \mathbf{a}(\theta_i) \mathbf{a}(\theta_i)^H$ is a dyadic decomposition of $\mathbf{M}(\theta)$.

The following proposition defines the elements of the FIM for the stochastic model.

Proposition 2. For $i, j = 1, \dots, P$, the elements of the FIM $\tilde{\mathbf{I}}(\phi, \theta)$ associated with the deterministic unknown parameters $\{\phi, \theta\}$ become as follows:

$$\begin{aligned} \tilde{\mathbf{I}}(\phi, \theta)_{1,1} &= \frac{\partial^2}{\partial \phi^2} \log \det \tilde{\Psi}(\phi, \theta) \\ &+ \text{tr} \left(\tilde{\Psi}(\phi, \theta) \frac{\partial^2}{\partial \phi^2} \tilde{\Psi}^{-1}(\phi, \theta) \right), \end{aligned} \quad (23)$$

$$\begin{aligned} \tilde{\mathbf{I}}(\phi, \theta)_{1,i+1} &= \tilde{\mathbf{I}}(\phi, \theta)_{i+1,1} = \\ &= \frac{\partial^2}{\partial \phi \partial \theta_i} \log \det \tilde{\Psi}(\phi, \theta) + \text{tr} \left(\tilde{\Psi}(\phi, \theta) \frac{\partial^2}{\partial \phi \partial \theta_i} \tilde{\Psi}^{-1}(\phi, \theta) \right), \end{aligned} \quad (24)$$

$$\begin{aligned} \tilde{\mathbf{I}}(\phi, \theta)_{i+1,j+1} &= \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log \det \tilde{\Psi}(\phi, \theta) \\ &+ \text{tr} \left(\tilde{\Psi}(\phi, \theta) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \tilde{\Psi}^{-1}(\phi, \theta) \right). \end{aligned} \quad (25)$$

Calculating the FIM, and hence the CRB, has time complexity $\mathcal{O}(N_r^{2.4})$.

The proof of Proposition 2 is in Appendix C. Similarly to Proposition 1, Proposition 2 applies to any signaling, where the data symbol lies on the unit circle, such as the PSK signaling.

The following corollary will be useful in the numerical section.

Corollary 3. Consider the example of $P = 1$ (i.e. single object), $N_r = 2$, $\mathbf{C} = c\mathbf{I}$, $\Omega = 1$. Under the assumption $P_p = P_d$ the $\tilde{\mathbf{I}}(\phi, \theta_1)_{11}$ and the $\tilde{\mathbf{I}}(\phi, \theta_1)_{22}$ elements of the FIM are as follows:

$$\begin{aligned} \tilde{\mathbf{I}}(\phi, \theta_1)_{11} &= \frac{4\alpha^4 c^2 P_1^2}{(2\alpha^2 c P_1 + \sigma^2)(2\alpha^2 c P_1 \sigma^3 + 4\sigma P_s \alpha_s^2 (c P_1 v(\phi) + \sigma^2) + \sigma^5)^2} \\ &\cdot \left[\sigma^4 (2\alpha^2 c P_1 + \sigma^2)^2 + 6\sigma^2 P_s \alpha_s^2 (2\alpha^2 c P_1 + \sigma^2) (c P_1 v(\phi) + \sigma^2) \right. \\ &\quad \left. + 2P_s^2 \alpha_s^4 \left(4c P_1 v(\phi) (c P_1 v(\phi) + \sigma^2 (\cos(\phi) + 3)) \right. \right. \\ &\quad \left. \left. - \sigma^4 (\cos(2\phi) - 5) \right) \right], \end{aligned} \quad (26)$$

$$\begin{aligned} \tilde{\mathbf{I}}(\phi, \theta_1)_{22} &= \frac{32\pi^2 l^2 \cos^2(\theta_1) P_s^2 \alpha_s^4 (c P_1 v(\phi) + \sigma^2)^2}{\sigma^2 (2\alpha^2 c P_1 + \sigma^2) (2\alpha^2 c P_1 \sigma^2 + 4P_s \alpha_s^2 (c P_1 v(\phi) + \sigma^2) + \sigma^4)}, \end{aligned} \quad (27)$$

where $P_1 = P_p = P_d$, and $v(\phi) = \alpha^2(1 - \cos(\phi))$.

Corollary 3 allows a similar qualitative analysis as Corollary 2. When the total power budget is utilized, i.e. $P_s = P_{\text{TOT}} - P_1$, then the derivative of (26) with respect to P_1 is positive for $0 < P_1 < P_{\text{TOT}}$, which indicates that $\tilde{\mathbf{I}}(\phi, \theta_1)_{11}$ is maximized at $P_1 = P_{\text{TOT}}$; and the derivative of (27) with respect to P_1 is negative for $0 < P_1 < P_{\text{TOT}}$, from which $\tilde{\mathbf{I}}(\phi, \theta_1)_{22}$ is maximized at $P_1 = 0$.

The relative simplicity of $\tilde{\mathbf{I}}(\phi, \theta_1)_{22}$ compared to $\tilde{\mathbf{I}}(\phi, \theta_1)_{11}$ in Corollary 3, is due to the following property.

Corollary 4. $\det \tilde{\Psi}(\phi, \theta_1)$ is independent of θ_1 when $P = 1$ and consequently $\frac{\partial^2}{\partial \theta_1^2} \log \det \tilde{\Psi}(\phi, \theta_1) = 0$

We prove Corollary 4 in Appendix D.

As we will see in the next section – where we derive the derivatives of the loglikelihood functions – unbiased estimators that attain the CRB cannot be found either in the deterministic or in the stochastic model. Therefore, we seek alternative estimators, such as the maximum likelihood and MMSE estimators of ϕ and θ_1 .

V. DETERMINING THE MAXIMUM LIKELIHOOD CHANNEL AND SYMBOL ESTIMATORS

In this section we are interested in formulating the maximum likelihood estimators (MLEs) for the deterministic and stochastic models. The logarithm of the respective likelihood functions are available in (51) and (60) respectively.

A. Maximum Likelihood Estimation of the Deterministic Model

To compute the maximum likelihood estimator, we need to find the proper solution of the set of equations $\frac{\partial}{\partial \phi} Z(\mathbf{u}, \mathbf{v}, \phi, \theta) = 0$, $\frac{\partial}{\partial \theta_i} Z(\mathbf{u}, \mathbf{v}, \phi, \theta) = 0$, for $i = 1, \dots, P$, where (see (51)):

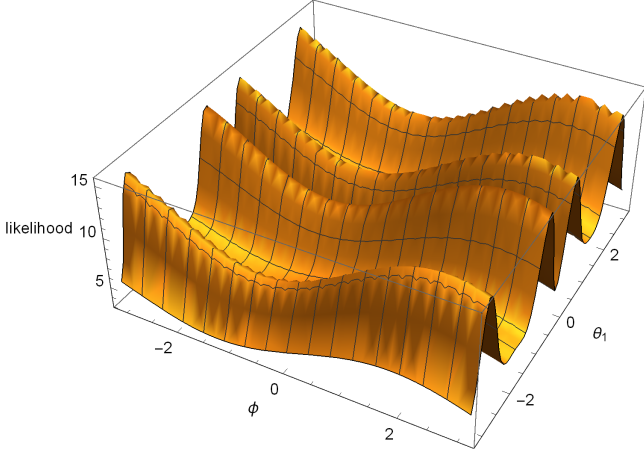


Figure 2. An example of the likelihood function as a function of ϕ and θ_1 in the deterministic case, which illustrates why it is difficult to find its global maximum.

$$Z(\mathbf{u}, \mathbf{v}, \phi, \theta) \triangleq \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\theta) \right)^H \boldsymbol{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\theta) \right) \quad (28)$$

is the ϕ and θ_i -dependent term of the loglikelihood function. For the derivative with respect to ϕ we have:

$$\begin{aligned} \frac{\partial}{\partial \phi} Z(\mathbf{u}, \mathbf{v}, \phi, \theta) = \\ 2 \operatorname{Re} \left[(\mathbf{v} - \mathbf{A}(\theta) \mathbf{Dp})^H \mathcal{F}_2(\mathbf{C}) e^{i(\phi + \pi/2)} (\mathbf{u} - \mathbf{A}(\theta) \mathbf{Dp}) \right], \end{aligned} \quad (29)$$

where $\mathcal{F}_2(\mathbf{C})$ is defined in Theorem 1. For the derivative with respect to θ_i , utilizing that $\boldsymbol{\Psi}^{-1}$ is Hermitian and does not depend on θ_i , we have:

$$\begin{aligned} \frac{\partial}{\partial \theta_i} Z(\mathbf{u}, \mathbf{v}, \phi, \theta) \\ = 2 \operatorname{Re} \left[\left(-\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\theta) \right)^H \boldsymbol{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\theta) \right) \right]. \end{aligned} \quad (30)$$

Due to (29) and (30) the derivative of the loglikelihood functions with respect to ϕ and θ_i is a non-linear function of ϕ and θ_i . Therefore, an unbiased estimator that attains the CRB for ϕ or θ_i cannot be found [57, Chapter 3.4].

As the likelihood is a non-linear function of ϕ and θ_1 , it may exhibit several local extreme values. Figure 2 plots $Z(\mathbf{u}, \mathbf{v}, \phi, \theta)$ from (28) (the likelihood function apart from a constant) as a function of θ_1 and ϕ for a simple example with $P = 1$ and $N_r = 2$.

This motivates the use of the maximum likelihood estimators derived in this section and the MMSE estimators derived in the sequel.

B. Maximum Likelihood Estimation in the Stochastic Model

Similarly to the deterministic case, to compute the maximum likelihood estimator in the stochastic case, we need to find the

proper solution of the set of equations $\frac{\partial}{\partial \phi} \tilde{Z}(\mathbf{u}, \mathbf{v}, \phi, \theta) = 0$ and $\frac{\partial}{\partial \theta_i} \tilde{Z}(\mathbf{u}, \mathbf{v}, \phi, \theta) = 0$ for $i = 1, \dots, P$, where

$$\tilde{Z}(\mathbf{u}, \mathbf{v}, \phi, \theta) \triangleq \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^H \tilde{\boldsymbol{\Psi}}^{-1}(\phi, \theta) \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}, \quad (31)$$

and its derivative according to ϕ and θ_i can be computed similarly:

$$\frac{\partial}{\partial \phi} \tilde{Z}(\mathbf{u}, \mathbf{v}, \phi, \theta) = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^H \frac{\partial}{\partial \phi} \tilde{\boldsymbol{\Psi}}^{-1}(\phi, \theta) \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}, \quad (32)$$

$$\frac{\partial}{\partial \theta_i} \tilde{Z}(\mathbf{u}, \mathbf{v}, \phi, \theta) = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^H \frac{\partial}{\partial \theta_i} \tilde{\boldsymbol{\Psi}}^{-1}(\phi, \theta) \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}, \quad (33)$$

where $\tilde{\boldsymbol{\Psi}}(\phi, \theta)$ is defined in (22).

The derivatives of the loglikelihood functions with respect to ϕ and θ_i in (32) and (33) are non-linear functions of ϕ and θ_i . As a consequence, an unbiased estimator of the stochastic model, that attains the CRB for ϕ or θ_i cannot be found.

C. Brute Force Search for Maximum Likelihood

While symbolically finding the ML estimate of ϕ and θ_i is hard, one can use a brute force grid search algorithm to find the maximum of $Z(\mathbf{u}, \mathbf{v}, \phi, \theta)$, as given by Algorithm 1 for $P = 1$.

Algorithm 1 Brute Force Grid Search to Maximize $Z(\mathbf{u}, \mathbf{v}, \phi, \theta)$

Input : Function $Z(\mathbf{u}, \mathbf{v}, \phi, \theta)$, observed $\mathbf{u}$ and $\mathbf{v}$, search parameters $[\theta_{\min}, \theta_{\max}]$, $[\phi_{\min}, \phi_{\max}]$, N_θ , N_ϕ

Output : $\hat{\phi}, \hat{\theta}$

Initialize : $\max_value \leftarrow -\infty$
 $\hat{\theta} \leftarrow \theta_{\min}$
 $\hat{\phi} \leftarrow \phi_{\min}$

```

1 for  $i \leftarrow 0$  to  $N_\theta - 1$  do
2   for  $j \leftarrow 0$  to  $N_\phi - 1$  do
3      $\theta \leftarrow \theta_{\min} + i \cdot \Delta_\theta$ 
4      $\phi \leftarrow \phi_{\min} + j \cdot \Delta_\phi$ 
5      $\text{current\_value} \leftarrow Z(\mathbf{u}, \mathbf{v}, \phi, \theta)$ 
6     if  $\text{current\_value} > \max\_value$  then
7        $\max\_value \leftarrow \text{current\_value}$ 
8        $\hat{\theta} \leftarrow \theta$ 
9        $\hat{\phi} \leftarrow \phi$ 
10    else
11      end if
12  end for
13 end for
14 return  $\hat{\phi}, \hat{\theta}$ 
```

This algorithm finds the ML estimates of ϕ and θ_1 in ranges $[\phi_{\min}, \phi_{\max}]$ and $[\theta_{\min}, \theta_{\max}]$, respectively, with precision Δ_ϕ and Δ_θ , respectively. The time complexity of the algorithm is $\mathcal{O}(N_r^{2.4} N_\theta N_\phi)$. A similar algorithm with the same time complexity can be used to find the ML estimates in the stochastic model by replacing Z with $\tilde{Z}$.

VI. DETERMINING THE MMSE CHANNEL AND SYMBOL ESTIMATORS

A. MMSE Channel and Symbol Estimation in the Deterministic Model

Recall that the MMSE channel estimator – which we will denote as $\mathbf{H}_{\text{ISAC}}$ to emphasize that the channel estimator takes into account the effect of the sensing signal – aims to minimize the mean squared error (MSE) between the estimate $\hat{\mathbf{h}} = \mathbf{H}_{\text{ISAC}}\mathbf{y}_p$ and the channel $\mathbf{h}$, where [16], [58]:

$$\mathbf{H}_{\text{ISAC}} \triangleq \arg \min_{\mathbf{H}} \mathbb{E}\{\|\mathbf{H}\mathbf{y}_p - \mathbf{h}\|_F^2\}. \quad (34)$$

For $\mathbf{H}_{\text{ISAC}}$ and the associated channel estimate $\hat{\mathbf{h}}$, we can state the following.

Proposition 3. *In the deterministic model, the MMSE channel estimator and the associated MMSE channel estimate are expressed as $\text{vec}(\mathbf{H}_{\text{ISAC}}) = \mathbf{T}^{-1}\mathbf{b}$, where*

$$\mathbf{T}^{-1} = \tilde{\mathbf{M}}^{-1} \otimes \mathbf{I}, \quad (35)$$

$$\mathbf{b} = \alpha\sqrt{P_p}\text{vec}(\mathbf{C}). \quad (36)$$

Furthermore, the estimated channel when using the MMSE receiver is expressed as:

$$\hat{\mathbf{h}}_{\text{ISAC}} = \alpha\sqrt{P_p}\mathbf{C}(\tilde{\mathbf{M}}^{-1})^T\mathbf{y}_p, \quad (37)$$

where $\tilde{\mathbf{M}} \triangleq \sigma^2\mathbf{I}_{N_r} + \alpha^2P_p\mathbf{C} + \mathbf{A}(\boldsymbol{\theta})\mathbf{D}\mathbf{p}\mathbf{p}^H\mathbf{D}^H\mathbf{A}^H(\boldsymbol{\theta})$.

The proof of Proposition 3 is in Appendix E.

The MMSE symbol estimator that utilizes an MMSE channel estimation has been derived in several papers, see for example, [59], [16]. Similarly to the derivation of the ISAC-aware channel estimator, the ISAC-aware receiver can be derived by considering the sensing signal as additional Gaussian noise at the communication receiver:

$$\mathbf{G}_{\text{ISAC}} = \alpha\sqrt{P_d}\hat{\mathbf{h}}_{\text{ISAC}}^H \cdot \left(\alpha^2P_d \left(\hat{\mathbf{h}}_{\text{ISAC}}\hat{\mathbf{h}}_{\text{ISAC}}^H + \mathbf{Q} \right) + \mathbf{A}(\boldsymbol{\theta})\mathbf{D}\mathbf{p}\mathbf{p}^H\mathbf{D}^H\mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2\mathbf{I}_{N_r} \right)^{-1}, \quad (38)$$

where the $\mathbf{Q}$ regularization matrix is the covariance matrix of the conditional distribution of $\mathbf{h}$ [59], [58]:

$$(\mathbf{h}|\hat{\mathbf{h}}) \sim \mathcal{CN}(\hat{\mathbf{h}}, \mathbf{Q}). \quad (39)$$

Note that as it was proven in [58], regularizing the MMSE receiver with the $\mathbf{D}$ and $\mathbf{Q}$ matrices as shown in (38) minimizes the expectation of the squared symbol error in the presence of channel estimation errors. Thus, the intuition behind the ISAC-aware receiver in (38) is that since the sensing signal appears as noise at the communication receiver, the regularization must include the sensing signal accordingly.

B. MMSE Channel and Symbol Estimation in the Stochastic Model

The analysis of the stochastic model follows the same pattern as the one of the deterministic model and here we only summarize the main results.

Table III
SETTING OF THE SYSTEM PARAMETERS

Parameter	Value
N_r	4
$\mathbf{C} = c\mathbf{I}_{N_r}$, with $c = 1$	Covariance matrix of the effective channel $\mathbf{h} = \mathbf{H}\mathbf{w}$.
P_p, P_d, P_s	Total power budget $P_p + P_s = 250$ mW; $P_d + P_s = 250$ mW.
P	1 (single object)
$\mathbf{p} \in \mathbb{C}^P$	$p = 1$ and $p \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega})$, where $\boldsymbol{\Omega} = 1$ (scalar).
θ_p	Angle of arrival of object p , $p = 1 \dots P$
s	$s = 1$
x	$x = e^{i\phi}$, where $\phi \in [-\pi, \pi]$
α, α_s	60 dB ("low path loss, (PL)") or 80 dB ("high path loss, (PL)")

The estimated channel when using the MMSE receiver is expressed as:

$$\check{\mathbf{h}}_{\text{ISAC}} = \alpha\sqrt{P_p}\mathbf{C}(\tilde{\mathbf{M}}^{-1})^T\mathbf{y}_p, \quad (40)$$

where $\tilde{\mathbf{M}} \triangleq \sigma^2\mathbf{I}_{N_r} + \alpha^2P_p\mathbf{C} + \mathbf{A}(\theta_1)\mathbf{D}\boldsymbol{\Omega}\mathbf{D}^H\mathbf{A}^H(\theta_1)$. The ISAC-aware receiver is:

$$\check{\mathbf{G}}_{\text{ISAC}} = \alpha\sqrt{P_d}\check{\mathbf{h}}_{\text{ISAC}}^H \cdot \left(\alpha^2P_d \left(\check{\mathbf{h}}_{\text{ISAC}}\check{\mathbf{h}}_{\text{ISAC}}^H + \mathbf{Q} \right) + \mathbf{A}(\theta_1)\mathbf{D}\boldsymbol{\Omega}\mathbf{D}^H\mathbf{A}^H(\theta_1) + \sigma^2\mathbf{I}_{N_r} \right)^{-1}. \quad (41)$$

C. Summary

Section V has derived analytical formulas for the first derivatives of the likelihood functions in the deterministic and stochastic models, which form the basis of finding the maximum likelihood estimations of the AoA and the transmitted communication symbol (θ_1 and ϕ) for each observation of $(\mathbf{y}_p, \mathbf{y}_d)$. While the maximum likelihood estimation of θ_1 and ϕ is appealing, maximizing the likelihood function is difficult in practice, and therefore in this section we have noted that a viable alternative is to use MMSE estimation for both the communication channel $\mathbf{h}$ and the transmitted symbol x , while employing the well-known MUSIC algorithm for estimating θ_1 . To this end, we have established Proposition 3, that, together with (38), defines the ISAC-aware MMSE receiver, which can be used as an alternative to maximum likelihood estimation.

VII. NUMERICAL RESULTS

Recall that Cases 1-2 refer to the deterministic model, in which the source signal is predefined and known by the receiver, while in Cases 3-4 the sensing waveform is zero-mean Gaussian. Cases 1 and 3 concern the case without a priori information about the AoA, while in Cases 2 and 4 some a priori information about the AoA is available at the receiver, and we are interested in establishing the BCRB.

To gain some basic insights, in this section, we consider the uplink of a SU-MIMO ISAC system that consists of a single UE, a serving BS equipped with $N_r = 4$ receive antennas and a single passive object (i.e. $P = 1$). The UE transmits

Table IV
SUMMARY OF FIGURES IN THE NUMERICAL SECTION

Figure	Description and Take-Away
Figure 3	Comparing the CRBs in the deterministic and stochastic models.
Figure 4	Variances of the $\hat{\phi}$ and $\hat{\theta}_1$ estimates – using maximum likelihood estimation – in the deterministic and stochastic models. The figure also shows the associated deterministic model CRBs for comparison.
Figure 5	CRBs for $\hat{\phi}$ and $\hat{\theta}_1$ in the deterministic model as functions of ϕ .
Figure 6	CRB for $\hat{\theta}_1$ in the deterministic and stochastic models at low and high path loss values as functions of θ_1 .
Figure 7	CRBs for $\hat{\phi}$ and $\hat{\theta}_1$ in the deterministic and stochastic models as functions of the communication power (i.e. $P_p = P_d$)
Figure 8	CRB for $\hat{\theta}_1$ in the stochastic model as a function of the communication power (i.e. $P_p = P_d$) when the a priori information of θ_1 is available (i.e. Case 4) that it is uniformly distributed in the interval $[0, \frac{\pi}{2a_x}]$; $a_x \geq 1$. The figure examines two cases: $a_x = 1$ and $a_x \rightarrow \infty$.
Figure 9	CRB for $\hat{\theta}_1$ in the stochastic model as a function of the sensing power P_s when a certain power budget is allocated to the total communication power (i.e. $P_p + P_d = 100$ mW).
Figure 10	Pilot and data power trade-off in Case 3 in terms of the CRB for $\hat{\theta}_1$ when the pilot and data power are set equally/unequally under a fixed communication power budget.
Figure 11	CDF of the squared communication symbol estimation error under different communication power settings.
Figure 12	CDF of the AoA estimation error under different communication power (and thereby sensing power) settings.

a pilot symbol s and subsequently a data symbol $x = e^{i\phi}$ over the effective communication channel $\mathbf{h}$, which yields the received pilot and data signals as described by (6) and (7). Recall that in this example we assume that there is a single slot used for transmitting the pilot symbol and a single subsequent slot for transmitting the data symbol according to the snapshot model assumption (discussed in [18]). Note that all angles are measured in radian, and the units are indicated in the figures, except for the unit-less quantities such as radian or probability.

In line with the assumption on dividing the total available power P_{TOT} discussed in Section II, the total power budget of the UE is divided between the communication power (P_p in the pilot slot and P_d in the data slot) and the sensing power P_s , where we set the total power budget to 250 mW. The main parameters of this system are summarized in Table III. $\mathbf{C} =$

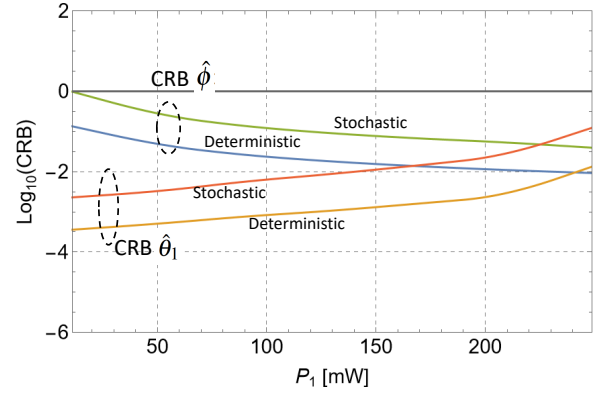


Figure 3. Comparing the CRBs for $\hat{\phi}$ and $\hat{\theta}_1$ in the deterministic (Case 1) and stochastic (Case 3) models. As p_1 , that is the transmit power level used for both pilot and data symbols increases, the CRB for $\hat{\phi}$ decreases, while the CRB for $\hat{\theta}_1$ increases. Notice that the CRBs in the deterministic model are somewhat lower due to the fact that the applied sensing signal is deterministic.

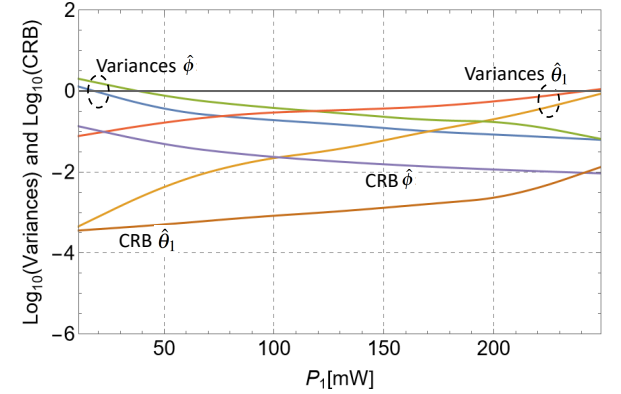


Figure 4. Comparing the variances of $\hat{\phi}$ and $\hat{\theta}_1$ in the deterministic (Case 1) and stochastic (Case 3) models and relating them to the CRB (obtained with the deterministic model). The figure shows the empirically obtained variances for both $\hat{\phi}$ and $\hat{\theta}_1$ (as indicated with the dashed ellipsoids) when using maximum likelihood estimation as well as the respective CRBs. Note that the variances obtained in the deterministic sensing model ($p = 1$) are lower for all P_1 (communication power) values than when using the stochastic (Gaussian) sensing model.

$c\mathbf{I}_{N_r}$ means that the rows of $\mathbf{H}$ are assumed to be independent with identical covariance matrices. It is important to note that the CRB results presented in this section are obtained using the closed form expressions in Propositions 1 and 2, while the variances of the estimated parameters $\hat{\phi}$ and $\hat{\theta}_1$ (shown only in Figure 4) are obtained by Monte Carlo simulations. For convenience, the figures in this section show the 10-base logarithm of the CRBs and variances.

Figure 3 shows the CRBs as a function of $P_1 \triangleq P_p = P_d$ when using the deterministic and stochastic sensing signal models. As P_1 increases, the sensing power decreases, which explains why the CRB for $\hat{\theta}_1$ increases, while the CRB for $\hat{\phi}$ decreases. These opposing trends are due to the intrinsic trade-off in the power domain between sensing and communications. Also, as expected, both CRBs are lower when using the deterministic model due to the fact that when the sensing signal is stochastic, it increases the variance of the received pilot and data signals.

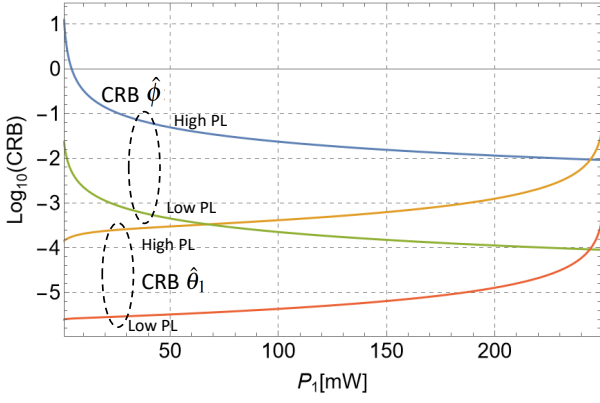


Figure 5. CRB trade-off in the deterministic model (Case 1) when the path loss between the UE and the BS is high (80 dB) and low (60 dB), denoted as “High PL” and “Low PL”, respectively. Both estimates are sensitive to both the power setting (i.e. the value of P_1) and the path loss.

Figure 4 shows the variance of $\hat{\phi}$ and $\hat{\theta}_1$ as a function of $P_1 \triangleq P_p = P_d$ in the deterministic and stochastic models, that is for the cases when the sensing signal p is deterministic and when it is a random signal drawn from a Gaussian distribution according to the parameter setting in Table III. For comparison, this figure also plots the respective (deterministic) CRBs. The variances are obtained by Monte Carlo simulations using the ML estimation method to obtain the $\hat{\phi}$ and $\hat{\theta}_1$ estimates based on the received signal $\mathbf{y}_p$ and $\mathbf{y}_d$. Notice that when maximizing the loglikelihood function in $\hat{\phi}$ and $\hat{\theta}_1$, the receiver (i.e. the serving BS) does not need to estimate $\mathbf{h}$ and in this sense the ML is blind.

Figure 5 focuses on Case 1 (deterministic model without a priori information on the parameters) and shows the CRBs for $\hat{\phi}$ and $\hat{\theta}_1$ as a function of the communication power P_1 for the high and low path loss ($\alpha = 80$ dB and $\alpha = 60$ dB, respectively). Similarly to the trends observed in Figure 3, the CRB for $\hat{\phi}$ decreases, while the CRB for $\hat{\theta}_1$ increases. Low path loss between the UE and the serving BS facilitates higher quality estimates characterized by much lower CRBs for both symbol and AoA estimation due to the higher SNR from the perspectives of both communication and sensing.

Figure 6 shows the CRBs for $\hat{\theta}_1$ and $\hat{\phi}$ as a function of the actual value (ground truth) of ϕ in Case 1. Note as we discussed in conjunction with Proposition 1, the CRB for $\hat{\phi}$ does not depend on ϕ , while the phase of the transmitted symbol (ϕ) affects heavily the CRB for $\hat{\theta}_1$. This is because when $\phi = 0$ (i.e. $x = 1$), the communication signal represents the highest interference in the observation at the BS and makes the angle estimation problematic. Notice in the figure that the CRB at this value of ϕ is the same in the low and high path loss scenarios. This observation may be relevant in a future work aiming to design not only the sensing signal, but also to design “sensing-friendly” communication signals.

Figure 7 shows the CRB for $\hat{\theta}_1$ in Case 1 and Case 3 (deterministic and stochastic model respectively), as a function of θ_1 (recall that the CRB for $\hat{\phi}$ is not a function of θ_1). The CRB is low when $\theta_1 = 0$ or when $\theta_1 = \pm\pi$ and high when

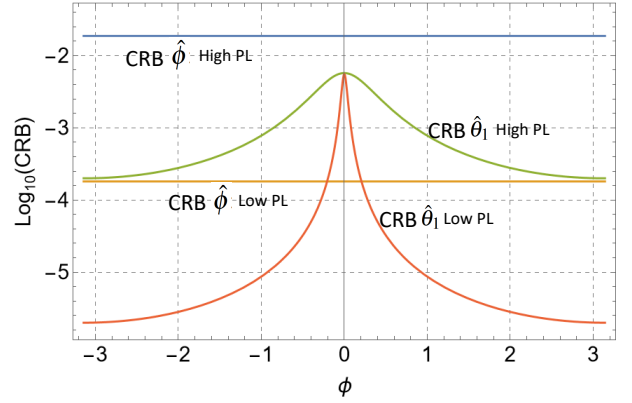


Figure 6. CRBs for $\hat{\theta}_1$ and $\hat{\phi}$ as the function of ϕ in Case 1. Interestingly, the CRB for $\hat{\phi}$ is not a function of ϕ , whereas the CRB for $\hat{\theta}_1$ is quite sensitive to the phase of the transmitted symbol. This result indicates that the AoA estimation is most problematic (high CRB) when $\phi = 0$, i.e. $x = 1$. (The value of θ_1 is assumed to be the ground truth $\theta_1 = \frac{\pi}{5}$.)

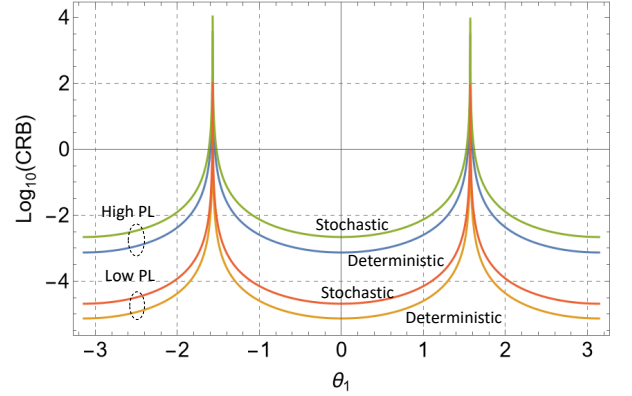


Figure 7. CRB for $\hat{\theta}_1$ as a function of θ_1 in the deterministic and stochastic models (Case 1 and Case 3). When $\theta_1 = \pm\frac{\pi}{2}$, the quality of the AoA estimation is poor (high CRB), because the time difference of arrival and phase difference at the different antenna elements is less than when, e.g. $\theta_1 = 0$.

$\theta_1 = \pm\frac{\pi}{2}$.

The CRB is low when the AoA of the impinging signals is $\theta_1 = 0$ (i.e. arriving from the boresight direction), because in this case the change in the phase difference at the antenna elements of the antenna array due to a small movement of the reflecting object is the largest over all possible values of θ_1 . This can be seen by considering that the steering vector depends on the sine of θ_1 , whose rate of change is greatest at $\theta_1 = 0$.

Figure 8 compares Case 2 and Case 4 when a priori knowledge about the distribution of the AoA is available, and examines how this a priori knowledge affects the BCRB. This figure assumes that θ_1 is uniformly distributed in the interval $[0, \frac{\pi}{2a_x}]$, which is applicable in cases, where the sensed passive objects are restricted to be in a certain angular domain. As a_x increases along the abscissa, the support of the AoA gets smaller in the interval $[0, \frac{\pi}{2a_x}]$, which lowers the BCRB. Notice that the BCRB is lower in the deterministic model (Case 2), and that this difference does not get smaller as the support of the AoA decreases (i.e. as a_x increases).

Figure 9 shows the BCRB in Case 4 (i.e. stochastic model)

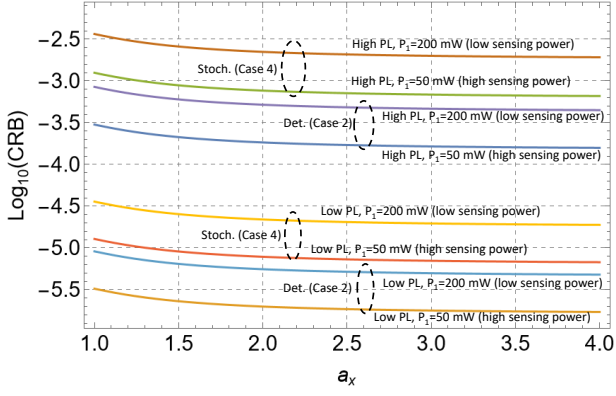


Figure 8. The impact of a priori knowledge about the AoA distribution in Case 2 (deterministic model) and Case 4 (stochastic model). As a_x increases along the abscissa, the support of the θ_1 gets smaller in the interval $[0, \frac{\pi}{2a_x}]$, which lowers the BCRB in both cases.

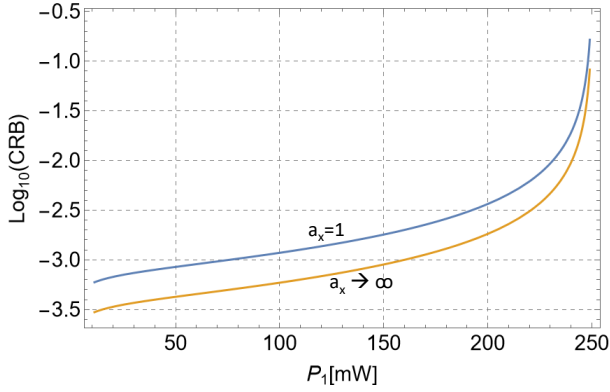


Figure 9. BCRB for $\hat{\theta}_1$ as a_x goes to ∞ in the stochastic model (Case 4). In this figure, $a_x = 1$ corresponds to the a priori information that θ_1 is uniformly distributed in the interval $[0, \frac{\pi}{2}]$, whereas increasing a_x narrows down the support of θ_1 to $[0, \frac{\pi}{2a_x}]$. The BCRB decreases due to narrowing down the possible values of θ_1 .

as a function of P_1 when a priori information is available about $\hat{\theta}_1$. The a priori information about θ_1 is represented by the information that θ_1 is uniformly distributed over the closed interval $[0, \frac{\pi}{2a_x}]$ when $a_x = 1$ and $a_x \rightarrow \infty$. More specifically, the BCRB plotted in this figure is defined as:

$$\tilde{\mathbf{I}}(a_x)_{22} \triangleq \frac{1}{2\pi} \cdot \frac{2a_x}{\pi} \cdot \int_{\phi=-\pi}^{\pi} \int_{\theta_1=0}^{\frac{\pi/2}{a_x}} \tilde{\mathbf{I}}(\phi, \theta_1)_{22} d\phi d\theta_1. \quad (42)$$

As Figure 9 shows, the BCRB decreases as the interval within which θ_1 lies decreases (as a_x tends to infinity), although the CRB remains finite. This fact can also be seen by evaluating the integral in (42) (not shown here).

Figure 10 illustrates the pilot and data transmit power trade-off in Case 3 (stochastic model). This figure shows the CRB for $\hat{\theta}_1$ as a function of the sensing power P_s assuming a fixed communication power budget $P_p + P_d = 100$ mW. As the sensing power increases, the CRB for $\hat{\theta}_1$ decreases. The CRB is quite insensitive to how the communication power budget is divided between the transmit power level of the pilot (P_p) and data (P_d) signals. Note that two curves corresponding to how

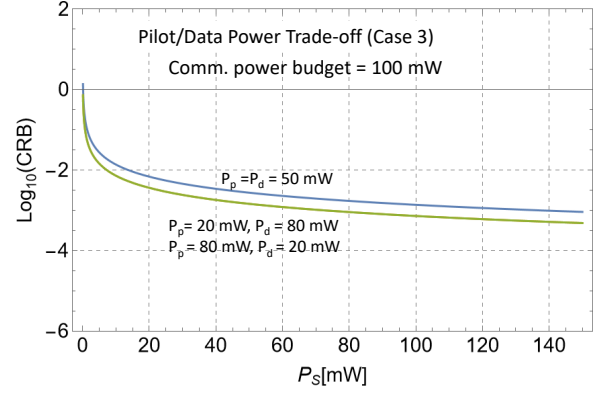


Figure 10. Pilot and data transmit power trade-off in Case 3 (stochastic sensing signal) illustrated by the CRB for $\hat{\theta}_1$ as a function of the sensing power P_s assuming a fixed communication power $P_p + P_d = 100$ mW. As the sensing power increases, the CRB for $\hat{\theta}_1$ decreases. The CRB is quite insensitive to how the communication power budget is divided between the transmit power level of the pilot (P_p) and data (P_d) signals, which explains why the two curves ($P_p = 20$ mW and $P_p = 80$ mW) overlap.

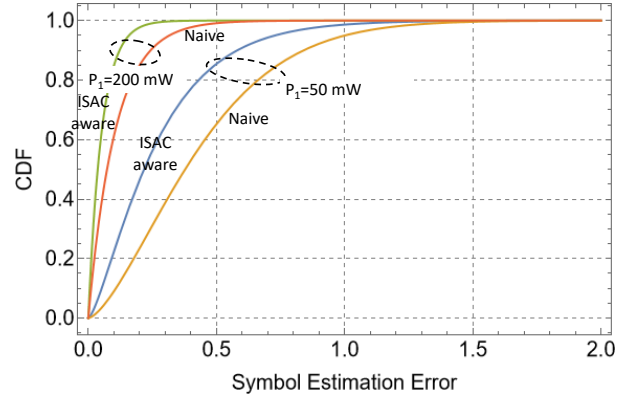


Figure 11. The distribution of the squared communication symbol error in the cases of using the naive and the ISAC-aware MMSE receivers and a stochastic sensing signal when the communication power is set to P_1 to 50 mW and 200 mW. As the communication power increases, the symbol estimation error decreases, and the gap between the naive and ISAC-aware receivers vanishes.

the communication budget is divided between the pilot and data power overlap. We observe a similar behavior in Case 1 (deterministic model, not shown here).

Finally, we study the impact of power allocation on the symbol and AoA estimation performance when using the proposed ISAC-aware communication receiver and the naive (legacy) MMSE receiver. Recall that the naive receiver is derived under the assumption that perfect channel state information at the receiver (CSIR) is available, and uses the estimated channel $\hat{\mathbf{h}}$ as if it was the actual channel $\mathbf{h}$ [16], [60]:

$$\mathbf{G}_{\text{naive}} = \alpha \sqrt{P_d} \hat{\mathbf{h}}^H \cdot \left(\alpha^2 P_d \hat{\mathbf{h}} \hat{\mathbf{h}}^H + \sigma^2 \mathbf{I}_{N_r} \right)^{-1}. \quad (43)$$

For ease of illustration, we again set the pilot and data power levels equally (denoted by $P_1 \triangleq P_p = P_d$) in Figures 11 and 12 and assume that $P_1 + P_s = 250$ mW. Figure 11 shows the distribution of the squared error of communication symbol estimation (i.e. $|x - \hat{x}|^2$). When the pilot and data power levels are low $P_1 = 50$ mW, (and even when $P_1 = 200$ mW),

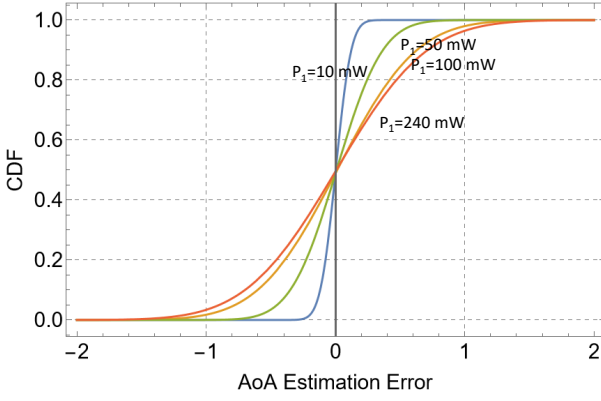


Figure 12. The distribution of the AoA estimation error as the function of the power level that is allocated to communication, P_1 . Recall that the sensing power increases, since we assume a fix total communication and sensing power budget (in this case 250 mW). As P_1 decreases, the sensing power increases and the distribution of the AoA estimation concentrates around zero with a lower variance than when the sensing power is low.

the squared symbol error depends critically on employing the ISAC-aware communication receiver. For example, when $P_1 = 50$ mW, in 80% of the cases the squared symbol error is less than 0.42 when using the ISAC-aware receiver, while it is less than 0.68 when using the naive receiver. However, when P_1 is set to 200 mW, the difference between the naive and ISAC-aware receivers becomes negligible.

As expected, Figure 12 shows an opposite trend for the AoA-estimation error, which further illustrates the inherent trade-off between the communication and sensing performance. When P_1 is low, the sensing power can be set to higher values, which significantly improves the AoA estimation quality. This is clearly visible in the cumulative distribution function (CDF) of the AoA estimation error, where the CDF has a much narrower spread around zero when $P_1 = 10$ mW than when $P_1 = 240$ mW.

VIII. CONCLUSIONS

In this paper, we argued that both the deterministic and the stochastic models provide meaningful CRBs for symbol and AoA estimation in ISAC systems. However, the dependencies of the CRBs in the two models are different. In the deterministic model, the CRB for the symbol phase (ϕ) is not a function of the symbol phase and the angle, while the CRB for the AoA is a function of both the symbol phase and the AoA. In contrast, in the stochastic model, the CRB on the symbol phase is a function of the phase (but not of the AoA), while the CRB for the AoA is a function of both. We have also shown that a priori information on the distribution of the symbol phase and the AoA enables to derive the Bayesian CRB, which in turn helps to lower the CRB. ML estimation is challenging in terms of finding the maximum of the likelihood functions and does not achieve the CRB either in the deterministic or in the stochastic models. These results can serve as basic considerations when designing pilot and sensing signals for ISAC systems. Our future work includes extending

these models to cases, in which the wireless channels undergo channel aging.

APPENDIX A PROOF OF THEOREM 1

Proof. According to (9)

$$\Psi(\phi) = \begin{pmatrix} \mathbf{K}_1(\mathbf{C}) & \mathbf{K}_2(\mathbf{C})e^{-i\phi} \\ \mathbf{K}_2(\mathbf{C})e^{i\phi} & \mathbf{K}_4(\mathbf{C}) \end{pmatrix}, \quad (44)$$

where $\mathbf{K}_1(\mathbf{C}) \triangleq \alpha^2 P_p \mathbf{C} + \sigma^2 \mathbf{I}_{N_r}$, $\mathbf{K}_2(\mathbf{C}) \triangleq \alpha^2 \sqrt{P_p P_d} \mathbf{C}$ and $\mathbf{K}_4(\mathbf{C}) \triangleq \alpha^2 P_d \mathbf{C} + \sigma^2 \mathbf{I}_{N_r}$. We note that $\mathbf{C}$, $\mathbf{K}_1(\mathbf{C})$, $\mathbf{K}_2(\mathbf{C})$ and $\mathbf{K}_4(\mathbf{C})$ are Hermitian and commute.

For the inverse of $\Psi(\phi)$, utilizing that $e^{i\phi}e^{-i\phi} = 1$, we have:

$$\Psi^{-1}(\phi) \triangleq \begin{pmatrix} \hat{\mathbf{K}}_1(\mathbf{C}) & \hat{\mathbf{K}}_2^H(\mathbf{C})e^{-i\phi} \\ \hat{\mathbf{K}}_2(\mathbf{C})e^{i\phi} & \hat{\mathbf{K}}_4(\mathbf{C}) \end{pmatrix}, \quad (45)$$

where

$$\begin{aligned} \hat{\mathbf{K}}_1(\mathbf{C}) &\triangleq \left(\mathbf{K}_1(\mathbf{C}) - \mathbf{K}_2^H(\mathbf{C})\mathbf{K}_4^{-1}(\mathbf{C})\mathbf{K}_2(\mathbf{C}) \right)^{-1}, \\ \hat{\mathbf{K}}_4(\mathbf{C}) &\triangleq \left(\mathbf{K}_4(\mathbf{C}) - \mathbf{K}_2(\mathbf{C})\mathbf{K}_1^{-1}(\mathbf{C})\mathbf{K}_2^H(\mathbf{C}) \right)^{-1}, \\ \hat{\mathbf{K}}_2(\mathbf{C}) &\triangleq -\mathbf{K}_1^{-1}(\mathbf{C})\mathbf{K}_2^H(\mathbf{C})\hat{\mathbf{K}}_4(\mathbf{C}). \end{aligned}$$

Utilizing the commutativity of the matrices, we have

$$\begin{aligned} \hat{\mathbf{K}}_1(\mathbf{C}) &\triangleq \left(\mathbf{K}_1(\mathbf{C}) - \mathbf{K}_2^H(\mathbf{C})\mathbf{K}_4^{-1}(\mathbf{C})\mathbf{K}_2(\mathbf{C}) \right)^{-1} \\ &= \left(\alpha^2 P_p \mathbf{C} + \sigma^2 \mathbf{I}_{N_r} - \alpha^4 P_p P_d \mathbf{C}^2 (\alpha^2 P_d \mathbf{C} + \sigma^2 \mathbf{I}_{N_r})^{-1} \right)^{-1}. \end{aligned}$$

That is, $\hat{\mathbf{K}}_1$ is such that

$$\hat{\mathbf{K}}_1(u) = \frac{1}{\alpha^2 P_p u + \sigma^2 - \frac{\alpha^4 P_p P_d u^2}{\alpha^2 P_d u + \sigma^2}} = \frac{\sigma^2 + \alpha^2 P_d u}{\sigma^4 + \alpha^2 \sigma^2 (P_p + P_d) u}.$$

The rest of the proof comes similarly from $\hat{\mathbf{K}}_2(\mathbf{C})$ and $\hat{\mathbf{K}}_4(\mathbf{C})$. $\square$

APPENDIX B PROOF OF PROPOSITION 1

Using $e^{i\phi}e^{-i\phi} = 1$, for the determinant of $\Psi(\phi)$ we have:

$$\det \Psi(\phi) = \det \mathbf{K}_1(\mathbf{C}) \times \det \left(\mathbf{K}_4(\mathbf{C}) - \mathbf{K}_2(\mathbf{C})\mathbf{K}_1^{-1}(\mathbf{C})\mathbf{K}_2^H(\mathbf{C}) \right) \quad (46)$$

That is, $\det \Psi(\phi)$ is independent of ϕ .

For $\frac{\partial}{\partial \phi} \Psi^{-1}(\phi)$, utilizing Theorem 1, we have:

$$\frac{\partial}{\partial \phi} \Psi^{-1}(\phi) \triangleq \begin{pmatrix} \mathbf{0} & \hat{\mathbf{K}}_2^H(\mathbf{C})e^{-i(\phi+\pi/2)} \\ \hat{\mathbf{K}}_2(\mathbf{C})e^{i(\phi+\pi/2)} & \mathbf{0} \end{pmatrix}. \quad (47)$$

The derivative of $\mathbf{a}_j(\theta)$ with respect to θ is

$$\frac{\partial}{\partial \theta} \mathbf{a}_j(\theta) = \frac{\partial}{\partial \theta} e^{i2\pi j \ell \sin(\theta)} = i2\pi j \ell \cos(\theta) e^{i2\pi j \ell \sin(\theta)} \quad (48)$$

for $j \in \{0, \dots, N_r - 1\}$ and

$$\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\boldsymbol{\theta}) = \left(\frac{\partial}{\partial \theta_i} \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \mathbf{p} \right) = \left(\mathbf{a}'(\theta_i) \alpha_{s,i} \sqrt{P_s} \mathbf{p}_i \right), \quad (49)$$

where $\mathbf{a}'(\theta_i) \triangleq \left[\dots i2\pi j \ell \cos(\theta_i) e^{i2\pi j \ell \sin(\theta_i)} \dots \right]^T$.

The joint probability density function of $(\mathbf{y}_p, \mathbf{y}_d)$ is

$$f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v}) = \frac{1}{\pi^{2Nr} \det \mathbf{\Psi}(\phi)} \cdot \exp \left(- \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \cdot \mathbf{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \right). \quad (50)$$

Noting that $\det \mathbf{\Psi}(\phi)$ is neither a function of ϕ nor $\boldsymbol{\theta}$ (see (46) above), we define $d \triangleq \det \mathbf{\Psi}(\phi)$, and for the logarithm of the density function we write

$$\log f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v}) = -\log(\pi^{2Nr} d) - \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \mathbf{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right). \quad (51)$$

To determine the elements of the FIM, we need the second order partial derivatives of the logarithm of the density function with respect to θ_i and ϕ .

For the first-order derivative, we have:

$$\begin{aligned} \frac{\partial \log f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v})}{\partial \theta_i} &= \left(\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \mathbf{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \\ &+ \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \mathbf{\Psi}^{-1}(\phi) \left(\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\boldsymbol{\theta}) \right). \end{aligned} \quad (52)$$

The second-order derivatives can be written as:

$$\begin{aligned} \frac{\partial^2 \log f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v})}{\partial \phi^2} &= \\ &= - \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \frac{\partial^2}{\partial \phi^2} \mathbf{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right), \end{aligned} \quad (53)$$

$$\begin{aligned} \frac{\partial^2 \log f_{\mathbf{y}_p(t), \mathbf{y}_d(t)}(\mathbf{u}, \mathbf{v})}{\partial \phi \partial \theta_i} &= \\ &= \left(\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \frac{\partial}{\partial \phi} \mathbf{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \\ &+ \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \frac{\partial}{\partial \phi} \mathbf{\Psi}^{-1}(\phi) \left(\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \\ &= 2 \operatorname{Re} \left[\left(\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \frac{\partial}{\partial \phi} \mathbf{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \right], \end{aligned} \quad (54)$$

and

$$\begin{aligned} \frac{\partial^2 \log f_{\mathbf{y}_p(t), \mathbf{y}_d(t)}(\mathbf{u}, \mathbf{v})}{\partial \theta_i \partial \theta_j} &= \\ &= 2 \operatorname{Re} \left[\left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \mathbf{\Psi}^{-1}(\phi) \left(\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \right] \\ &- 2 \operatorname{Re} \left[\left(\frac{\partial}{\partial \theta_i} \boldsymbol{\mu}(\boldsymbol{\theta}) \right)^H \mathbf{\Psi}^{-1}(\phi) \left(\frac{\partial}{\partial \theta_j} \boldsymbol{\mu}(\boldsymbol{\theta}) \right) \right]. \end{aligned} \quad (55)$$

Next, to get the FIM, we substitute $(\mathbf{y}_p, \mathbf{y}_d)$ into $(\mathbf{u}, \mathbf{v})$, and take the expected value. Notice that since

$$\mathbb{E} \left[\begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_d \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \right] = \mathbf{0}, \quad (56)$$

any term that is a linear function of this expression has an expected value of $\mathbf{0}$. Finally, noticing that

$$\mathbb{E}[\mathbf{z}^H \mathbf{A} \mathbf{z}] = \mathbb{E}[\operatorname{tr}(\mathbf{z}^H \mathbf{A} \mathbf{z})] = \mathbb{E}[\operatorname{tr}(\mathbf{A} \mathbf{z} \mathbf{z}^H)] = \operatorname{tr}(\mathbf{A} \operatorname{Cov}(\mathbf{z})), \quad (57)$$

for any zero mean random vector $\mathbf{z}$ and non-random matrix $\mathbf{A}$, and substituting

$$\mathbf{z} = \begin{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \boldsymbol{\mu}(\boldsymbol{\theta}) \end{pmatrix}, \quad (58)$$

and

$$\mathbf{A} = \frac{\partial^2}{\partial \phi^2} \mathbf{\Psi}^{-1}(\phi), \quad (59)$$

into (57), the first part of the proposition follows. Regarding the time complexity of calculating the FIM, note that the operations in the calculation are:

- Inversion of a matrix of dimension $2N_r \times 2N_r$;
- Multiplication of two matrices of dimensions $2N_r \times 2N_r$;
- Partial derivation of a matrix of dimension $2N_r \times 2N_r$;
- Multiplication of a matrix of dimensions $2N_r \times 2N_r$ and a vector of dimension $2N_r$.

The highest time complexity steps are the matrix multiplication and inversion steps, for which many $\mathcal{O}(N_r^{2.4})$ algorithms exist, e.g. [61].

APPENDIX C PROOF OF PROPOSITION 2

The logarithm of the density function in the stochastic case becomes:

$$\begin{aligned} \log f_{\mathbf{y}_p, \mathbf{y}_d}(\mathbf{u}, \mathbf{v}) &= \\ &= -\log(\pi^{2Nr}) - \log \det \tilde{\mathbf{\Psi}}(\phi, \boldsymbol{\theta}) - \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^H \tilde{\mathbf{\Psi}}^{-1}(\phi, \boldsymbol{\theta}) \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}. \end{aligned} \quad (60)$$

The derivatives of the first term of the likelihood function in (60) are zero. The second order partial derivatives of the third term of (60) with respect to ϕ is:

$$\frac{\partial^2}{\partial \phi^2} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^H \tilde{\mathbf{\Psi}}^{-1}(\phi, \boldsymbol{\theta}) \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}^H \frac{\partial^2}{\partial \phi^2} \tilde{\mathbf{\Psi}}^{-1}(\phi, \boldsymbol{\theta}) \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}. \quad (61)$$

The other second order derivative (i.e. with respect to θ_i and θ_j) are obtained similarly. To obtain the elements of the FIM, we need to take the expectations of the second order partial derivatives of (60). Using the identity in (57), the FIM as stated in the proposition follows. The time complexity is the same as in Proposition 1.

APPENDIX D
PROOF OF COROLLARY 4

Proof. Notice that due to (11), when $P=1$ (single object), we have:

$$\det \tilde{\Psi}(\phi, \theta) = \det \Psi(\phi) \cdot \left(1 + \Omega \left(\mathcal{F}_1(m(\theta)) + \mathcal{F}_2(m(\theta))e^{-i\phi} + \mathcal{F}_2(m(\theta))e^{i\phi} + \mathcal{F}_4(m(\theta)) \right) \right)$$

where Ω and $m(\theta) = \mathbf{D}^H \mathbf{A}(\theta)^H \mathbf{C} \mathbf{A}(\theta) \mathbf{D}$ are scalars. For $m(\theta)$ we have

$$m(\theta) = \mathbf{D}^H \mathbf{A}(\theta)^H \mathbf{C} \mathbf{A}(\theta) \mathbf{D} = \alpha_{s,1}^2 P_s \left(\sum_{i=1}^{N_r} \mathbf{C}_{ii} + \sum_{i=1}^{N_r} \sum_{j=1, j \neq i}^{N_r} \mathbf{C}_{ij} e^{i2\pi \ell(i-j) \sin(\theta)} \right),$$

which implies that when $\mathbf{C}$ is diagonal, that is $\mathbf{C}_{ij} = 0$ for $i \neq j$, $\det \tilde{\Psi}(\phi, \theta)$ is not a function of θ . $\square$

APPENDIX E
PROOF OF PROPOSITION 3

Proof. We first derive the MMSE channel estimator $\mathbf{H}_{\text{ISAC}}$. Using that $\mathbf{H}\mathbf{y}_p = \text{vec}(\mathbf{H}\mathbf{y}_p) = (\mathbf{y}_p^T \otimes \mathbf{I})\text{vec}(\mathbf{H})$, for $\mathbb{E}\{\|\mathbf{H}\mathbf{y}_p - \mathbf{h}\|_F^2\}$ we have:

$$\begin{aligned} \mathbb{E}\{\|\mathbf{H}\mathbf{y}_p - \mathbf{h}\|_F^2\} &= \mathbb{E}\{\text{vec}(\mathbf{H})^H (\mathbf{y}_p^T \otimes \mathbf{I})^H (\mathbf{y}_p^T \otimes \mathbf{I}) \text{vec}(\mathbf{H})\} + \mathbb{E}\{\mathbf{h}^H \mathbf{h}\} \\ &\quad - \mathbb{E}\{\text{vec}(\mathbf{H})^H (\mathbf{y}_p^T \otimes \mathbf{I})^H \mathbf{h}\} - \mathbb{E}\{\mathbf{h}^H (\mathbf{y}_p^T \otimes \mathbf{I}) \text{vec}(\mathbf{H})\}, \end{aligned}$$

which is a quadratic optimization problem for $\mathbf{z} = \text{vec}(\mathbf{H})$, that is of the form $\mathbf{z}^H \mathbf{T} \mathbf{z} - \mathbf{z}^H \mathbf{b} - \mathbf{b}^H \mathbf{z} + \text{constant}$, where the optimal solution is $\mathbf{z}_{\text{opt}} = \text{vec}(\mathbf{H}_{\text{ISAC}})$. The solution of the quadratic optimization problem is $\mathbf{z}_{\text{opt}} = \mathbf{T}^{-1} \mathbf{b}$, where:

$$\begin{aligned} \mathbf{T} &= \mathbb{E}\{(\mathbf{y}_p^T \otimes \mathbf{I})^H (\mathbf{y}_p^T \otimes \mathbf{I})\} = \mathbb{E}\{\mathbf{y}_p^* \mathbf{y}_p^T\} \otimes \mathbf{I} \\ &= (\sigma^2 \mathbf{I}_{N_r} + \alpha^2 P_p \mathbf{C} + \mathbf{A}(\theta) \mathbf{D} \mathbf{p} \mathbf{p}^H \mathbf{D}^H \mathbf{A}^H(\theta)) \otimes \mathbf{I} = \bar{\mathbf{M}} \otimes \mathbf{I}, \\ \mathbf{b} &= \mathbb{E}\{(\mathbf{y}_p^T \otimes \mathbf{I})^H \mathbf{h}\} = \mathbb{E}\{\mathbf{y}_p^* \otimes \mathbf{h}\} = \alpha \sqrt{P_p} \text{vec}(\mathbf{C}), \end{aligned}$$

and

$$\mathbf{T}^{-1} = \bar{\mathbf{M}}^{-1} \otimes \mathbf{I}.$$

The next step is to compute $\hat{\mathbf{h}}_{\text{ISAC}}$.

$$\begin{aligned} \hat{\mathbf{h}}_{\text{ISAC}} &= \mathbf{H}_{\text{ISAC}} \mathbf{y}_p = (\mathbf{y}_p^T \otimes \mathbf{I}) \text{vec}(\mathbf{H}_{\text{ISAC}}) \\ &= \alpha \sqrt{P_p} (\mathbf{y}_p^T \otimes \mathbf{I}) (\bar{\mathbf{M}}^{-1} \otimes \mathbf{I}) \text{vec}(\mathbf{C}) \\ &= \alpha \sqrt{P_p} (\mathbf{y}_p^T \bar{\mathbf{M}}^{-1} \otimes \mathbf{I}) \text{vec}(\mathbf{C}) \\ &= \alpha \sqrt{P_p} \mathbf{C} (\bar{\mathbf{M}}^{-1})^T \mathbf{y}_p, \end{aligned}$$

which is identical with (37). $\square$

REFERENCES

- [1] C.-X. Wang, X. You, X. Gao, X. Zhu, Z. Li, C. Zhang, H. Wang, Y. Huang, Y. Chen, H. Haas, J. S. Thompson, E. G. Larsson, M. D. Renzo, W. Tong, P. Zhu, X. Shen, H. V. Poor, and L. Hanzo, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 905–974, 2023.
- [2] Z. Wei, H. Qu, Y. Wang, X. Yuan, H. Wu, Y. Du, K. Han, N. Zhang, and Z. Feng, "Integrated sensing and communication signals towards 5G-A and 6G: A survey," *IEEE Internet of Things Journal*, pp. 1–1, 2023.
- [3] G. Caire, C. R. C. M. da Silva, T. Gu, and W. Yuan, "Integrating sensing into communications in multi-functional networks," *IEEE Communications Magazine*, vol. 61, no. 5, pp. 24–25, 2023.
- [4] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint Radar and Communication Design: Applications, State-of-the-Art, and the Road Ahead," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, 2020.
- [5] C. Yuanhao, *Integrated Sensing and Communications*, F. Liu, C. Masouros, and Y. Eldar, Eds. Springer, 2023.
- [6] C. Mollén, G. Fodor, R. Baldemair, J. Huschke, and J. Vinogradova, "Joint multistatic sensing of transmitter and target in OFDM-based JCAS systems," in *2023 Joint European Conference on Networks and Comm. & 6G Summit (EuCNC/6G Summit)*, 2023, pp. 144–149.
- [7] P. Samczynski, K. Abratkiewicz, M. Plotka, T. P. Zielinski, J. Wszolek, S. Hausman, P. Korbel, and A. Ksiezyk, "5G network-based passive radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–9, 2022.
- [8] K. Abratkiewicz, A. Ksiezyk, M. Plotka, P. Samczynski, J. Wszolek, and T. P. Zielinski, "SSB-based signal processing for passive radar using a 5G network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 3469–3484, 2023.
- [9] M. Malanowski, M. Zywek, M. Plotka, and K. Kulpa, "Passive bistatic radar detection performance prediction considering antenna patterns and propagation effects," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.
- [10] A. R. Chiriyath, B. Paul, G. M. Jacyna, and D. W. Bliss, "Inner bounds on performance of radar and communications co-existence," *IEEE Trans. on Signal Processing*, vol. 64, no. 2, pp. 464–474, 2016.
- [11] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, and K. Chetty, "A survey on fundamental limits of integrated sensing and communication," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 994–1034, 2022.
- [12] 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*, vol. 69, no. 9, pp. 5723–5751, 2023.
- [13] J. Tong, H. Gaoming, T. Wei, and P. Huaufu, "Cramér-Rao lower bound analysis for stochastic model based target parameter estimation in multistatic passive radar with direct-path interference," *IEEE Access*, vol. 7, pp. 106 761–106 772, 2019.
- [14] M. U. Baig, J. Vinogradova, G. Fodor, and C. Mollén, "Joint communication and sensing beamforming for passive object localization," in *WSA and SCC 2023; 26th Int. ITG Workshop on Smart Antennas and 13th Conf. on Systems, Communications, and Coding*, 2023, pp. 1–6.
- [15] N. Jindal and A. Lozano, "A unified treatment of optimum pilot overhead in multipath fading channels," *IEEE Transactions on Communications*, vol. 58, no. 10, pp. 2939–2948, 2010.
- [16] G. Fodor, P. D. Marco, and M. Telek, "On the impact of antenna correlation and CSI errors on the pilot-to-data power ratio," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2622–2633, 2016.
- [17] H. Liu and E. Alsusa, "A novel ISAC approach for uplink NOMA system," *IEEE Communications Letters*, vol. 27, no. 9, pp. 2333–2337, 2023.
- [18] A. B. Gershman, *Array Signal Processing*, H. Bölcskei, D. Gesbert, C. B. Papadias, and A.-J. van der Veen, Eds. Cambridge University Press, 2008, vol. Space-Time Wireless Systems: From Array Processing to MIMO Communications.
- [19] Y. Liu, G. Liao, J. Xu, Z. Yang, and Y. Zhang, "Adaptive OFDM integrated radar and communications waveform design based on information theory," *IEEE Comm. Letters*, vol. 21, no. 10, pp. 2174–2177, 2017.
- [20] J. A. Zhang, X. Huang, Y. J. Guo, J. Yuan, and R. W. Heath, "Multibeam for joint communication and radar sensing using steerable analog antenna arrays," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 671–685, 2019.
- [21] C. B. Barneto, S. D. Liyanaarachchi, T. Riihonen, L. Anttila, and M. Valkama, "Multibeam design for joint communication and sensing in 5G new radio networks," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [22] 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.
- [23] D. Xu, C. Liu, S. Song, and D. W. Kwan Ng, "Integrated sensing and communication in coordinated cellular networks," in *2023 IEEE Statistical Signal Processing Workshop (SSP)*, 2023, pp. 90–94.

- [24] A. Zhang, M. L. Rahman, X. Huang, Y. J. Guo, S. Chen, and R. W. Heath, "Perceptive mobile networks: Cellular networks with radio vision via joint communication and radar sensing," *IEEE Vehicular Technology Magazine*, vol. 16, no. 2, pp. 20–30, 2021.
- [25] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, 1990.
- [26] —, "MUSIC, maximum likelihood, and Cramér-Rao bound: Further results and comparisons," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 12, pp. 2140–2150, 1990.
- [27] S. Lu, F. Liu, F. Dong, Y. Xiong, J. Xu, and Y.-F. Liu, "Sensing with random signals," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024, pp. 12 961–12 965.
- [28] L. Xie, F. Liu, Z. Xie, Z. Jiang, and S. Song, "Sensing mutual information with random signals in Gaussian channels," in *ICC 2024 - IEEE International Conference on Communications*, 2024, pp. 2228–2233.
- [29] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.
- [30] Y. Huang, Y. Fang, X. Li, and J. Xu, "Coordinated power control for network integrated sensing and communication," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 12, pp. 13 361–13 365, 2022.
- [31] Z. Behdad, O. T. Demir, K. W. Sung, E. Björnson, and C. Cavdar, "Power allocation for joint communication and sensing in cell-free massive mimo," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 4081–4086.
- [32] 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.
- [33] H. van Trees, "Bounds on the accuracy attainable in the estimation of continuous random processes," *IEEE Transactions on Information Theory*, vol. 12, no. 3, pp. 298–305, 1966.
- [34] L. C. Tran, A. Mertins, E. Dutkiewicz, and X. Huang, "Unitary differential space-time-frequency codes for MB-OFDM UWB," in *2009 9th International Symposium on Communications and Information Technology*, 2009, pp. 1161–1166.
- [35] J. Duplcy, J. Louveaux, and L. Vandendorpe, "Utility-based MIMO uplink beamforming," in *Fourth IEEE Workshop on Sensor Array and Multichannel Processing*, 2006., 2006, pp. 254–257.
- [36] L. C. Tran, A. Mertins, and T. A. Wysocki, "Unitary differential space-time-frequency codes for MB-OFDM UWB wireless communications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 862–876, 2013.
- [37] P. Singh, S. Srivastava, A. K. Jagannatham, and L. Hanzo, "Second-order statistics-based semi-blind techniques for channel estimation in millimeter-wave MIMO analog and hybrid beamforming," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6886–6901, 2020.
- [38] J. Komo and K. Barnett, "Improved bounds for coherent M-ary PSK symbol error probability," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 2, pp. 396–399, 1997.
- [39] N. C. Beaulieu and C. Jiang, "A new solution for the SEP of MPSK," *IEEE Communications Letters*, vol. 17, no. 1, pp. 12–14, 2013.
- [40] L. Rugini, "SEP bounds for MPSK with low SNR," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2473–2477, 2020.
- [41] H. Jafari, H. Miar-Naimi, and J. Kazemitabar, "Generalised expression for the symbol error floor of M-ary phase shift keying in the presence of phase noise, I/Q imbalance and DC-offset," *IET Communications*, vol. 14, no. 14, pp. 2319–2325, 2020. [Online]. Available: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-com.2020.0001>
- [42] N. Babu, C. Masouros, C. B. Papadias, and Y. C. Eldar, "Precoding for multi-cell ISAC: From coordinated beamforming to coordinated multipoint and bi-static sensing," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 14 637–14 651, 2024.
- [43] "New Radio (NR): UE radio access capabilities," 3GPP, Technical Specification (TS) 38.306, 09 2023, version 17.6.0.
- [44] H. Khzaali, J. Zec, and I. Kostanic, "Repeatability of reference signal received power measurements in lte networks," in *2018 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2018, pp. 1–6.
- [45] I. Ullah, H. E. Sayed, S. Malik, and M. A. Khan, "Performance evaluation of quality of experience aware mobility management in heterogeneous cellular networks," *IEEE Access*, pp. 1–1, 2024.
- [46] K. Boutiba, M. Bagaa, and A. Ksentini, "Radio link failure prediction in 5G networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.
- [47] A. Simonsson and A. Furuskar, "Uplink power control in lte - overview and performance, subtitle: Principles and benefits of utilizing rather than compensating for sinr variations," in *2008 IEEE 68th Vehicular Technology Conference*, 2008, pp. 1–5.
- [48] M. Mahot, F. Pascal, P. Forster, and J.-P. Ovarlez, "Asymptotic properties of robust complex covariance matrix estimates," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3348–3356, 2013.
- [49] W. Rao, "Low rank regularized ML estimation of structured covariance matrix," in *2016 IEEE International Conference on Digital Signal Processing (DSP)*, 2016, pp. 54–57.
- [50] S. Wu and X. Zhang, "A low-complexity antenna-la-aware spatial covariance matrix estimation method," in *2019 IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*, 2019, pp. 1–6.
- [51] T. L. Marzetta, G. H. Tucci, and S. H. Simon, "A random matrix-theoretic approach to handling singular covariance estimates," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6256–6271, 2011.
- [52] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 863–867, 2018.
- [53] S. Zou and Z. Zhao, "Large covariance matrix estimation based on factor models via nonconvex optimization," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024, pp. 9656–9660.
- [54] V. Raida, M. Lerch, P. Svoboda, and M. Rupp, "Deriving cell load from RSRQ measurements," in *2018 Network Traffic Measurement and Analysis Conference (TMA)*, 2018, pp. 1–6.
- [55] M.-W. Wu, Y. Jin, Y. Li, T. Song, and P.-Y. Kam, "Maximum-likelihood, magnitude-based, amplitude and noise variance estimation," *IEEE Signal Processing Letters*, vol. 28, pp. 414–418, 2021.
- [56] S. Ahmadi, *LTE Advanced: A Practical Systems Approach to Understanding 3GPP LTE Releases 10 and 11 Radio Access Technologies*. Academic Press, 2014, <https://doi.org/10.1016/B978-0-12-405162-1.00005-8>.
- [57] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory*, ser. Prentice Hall Signal Processing. Prentice Hall, 1993, no. ISBN: 013-345711-7.
- [58] G. Fodor, P. Di Marco, and M. Telek, "Performance analysis of block and comb type channel estimation for massive MIMO systems," in *1st International Conference on 5G for Ubiquitous Connectivity*, 2014, pp. 62–69.
- [59] J. Hoydis, S. T. Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need ?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [60] G. Fodor, P. D. Marco, and M. Telek, "On minimizing the MSE in the presence of channel state information errors," *IEEE Communications Letters*, vol. 19, no. 9, pp. 1604–1607, 2015.
- [61] R. Duan, H. Wu, and R. Zhou, "Faster matrix multiplication via asymmetric hashing," 2023. [Online]. Available: <https://arxiv.org/abs/2210.10173>