From Torch to Projector: Fundamental Tradeoff of Integrated Sensing and Communications

Yifeng Xiong, *Member, IEEE*, Fan Liu, *Member, IEEE*, Kai Wan, *Member, IEEE*, Weijie Yuan, *Member, IEEE*, Yuanhao Cui, *Member, IEEE*, and Giuseppe Caire, *Fellow, IEEE*

***Abstract*—Sensing and communications (S&C) were historically developed in parallel. In the last decade, they have been evolving from separation to integration, giving rise to the integrated sensing and communication (ISAC) paradigm, that has been recognized as one of 6G key usage scenarios. Despite the plethora of research works dedicated to ISAC signal processing, the fundamental performance limits of S&C remain widely unexplored in ISAC systems. In this tutorial paper, we attempt to summarize the recent research findings in characterizing the performance boundary of ISAC systems and the resulting S&C tradeoff from an information-theoretical viewpoint. We begin with a folkloric "torch metaphor" that depicts the subspace tradeoff between S&C. Then, we elaborate on the fundamental capacity-distortion (C-D) theory, indicating the incompleteness of this metaphor. Towards that end, we further elaborate on the S&C tradeoff by discussing a special case within the C-D framework, namely the Cramér-Rao bound (CRB)-rate region. In particular, S&C have preference discrepancies over both the signal subspace and the adopted codebook, leading to a "projector metaphor" complementary to the torch analogy. We also present two practical design examples by leveraging the lessons learned from fundamental theories. Finally, we conclude the paper by identifying a number of open challenges.**

***Index Terms*—Integrated sensing and communications, capacity-distortion theory, fundamental limits, CRB-rate region.**

I. Introduction

ISAC has recently been recognized by the international telecommunication union (ITU) as one of the vertices supporting the 6G hexagon of usage scenarios [1]. By relying on unified hardware platforms and radio waveforms, such integration enables resource-efficient cooperation between S&C subsystems, supporting various emerging applications including vehicle-to-everything (V2X) networking, extended reality (XR), and digital twins.

This work was supported in part by National Natural Science Foundation of China under Grants 62301060, 62331023, 62101232 and 12141107, in part by the Guangdong Province "Pearl River" Young Talent Support Program under Grant 2021QN02X128, and in part by the Shenzhen Science and Technology Program under Grants 20220815100308002, RCBS20210609103227018 and JCYJ20220530114412029. (*Corresponding author: Fan Liu.*)

Y. Xiong is with the School of Information and Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China (e-mail: yifengxiong@bupt.edu.cn).

F. Liu, Y. Cui and W. Yuan are with the School of System Design and Intelligent Manufacturing, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: liuf6@sustech.edu.cn, cuiyuanhao@bupt.edu.cn, yuanwj@sustech.edu.cn).

K. Wan is with School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: kai_wan@hust.edu.cn).

G. Caire is with the Chair of Communications and Information Theory, Technical University of Berlin, 10623 Berlin, Germany (e-mail: caire@tu-berlin.de).

Early endeavours on ISAC system design aim to extend the capability of existing infrastructures, hence developed into *sensing-centric* and *communication-centric* paradigms. Representative sensing-centric schemes include communicating using radar sidelobes and the permutation degree of freedom (DoF) in the waveform-antenna mapping of multiple-input multiple-output (MIMO) radars [2]. Communication-centric schemes are exemplified by sensing relying on orthogonal frequency-division multiplexing (OFDM) waveforms [3]. These schemes do not target the optimal S&C performance. To push the benefit of ISAC (termed as the "integration gain" [4]) to its limit, schemes of *joint design* emerged more recently [5], [6], which aim at conceiving novel joint signaling strategies from the ground-up, capable of accomplishing both tasks simultaneously. A natural, but easily overlooked question that arises is: What is the fundamental limit of the integration gain?

If we focus back on conventional individual S&C systems, we see that their fundamental performance limits originate from the resource budget. For example, the celebrated Shannon capacity formula for the scalar Gaussian channel exactly expresses the dependency of the communication performance on the available power and bandwidth. However, the performance limit of ISAC systems is vastly different. In its essence, the problem of ISAC system design is a multi-objective optimization problem. To elaborate, separated S&C design can be viewed as a special case of ISAC design, namely employing a time-sharing strategy. By utilizing the synergies between S&C, more favorable performance can be achieved. However, besides some rare occasions, it is highly unlikely that S&C would achieve their optimal performance at the same time, suggesting the existence of a fundamental S&C tradeoff in an ISAC system. Such a tradeoff may be framed by the Pareto boundary of the multi-objective ISAC system design problem.

For decades, S&C have been regarded as an information-theoretical "odd couple" that are mutually intertwined in profound ways [7]. At large, the S&C tradeoff may be understood from the perspective of a *signal-and-system* duality. For instance, let us consider a simple linear Gaussian model given by

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z} \tag{1}$$

with $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{H}$ and $\mathbf{Z}$ being the transmitted ISAC signal, the received signal, the channel, and the noise, respectively. For communication, the task is to decode the message encoded in $\mathbf{X}$, whereas for sensing, the task is to extract information from the channel $\mathbf{H}$. In light of this, one may view $\mathbf{H}$ as the "transmitted signal" from the environment to be sensed, while viewing $\mathbf{X}$ as the "channel' that this "signal" would

This article has been accepted for publication in IEEE BITS the Information Theory Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MBITS.2024.3376638

pass through. In this paper, we will encounter several ISAC channel models widely applied in the literature, most of which are variants of (1).

While the connection between estimation and information theories has been well-studied in the context of, e.g., I-MMSE equation [8], the fundamental tradeoff still remains widely open in general. To this end, this article will summarize current understandings of different components in the S&C tradeoff in a coherent manner, through the prism of information theory [9]–[11]. We shall commence with a folkloric "torch metaphor" depicting the resource competition between S&C subsystems, followed by the general capacity-distortion theory suggesting the incompleteness of this metaphor. We then introduce the Cramér-Rao bound (CRB)-rate region which clearly indicates that the S&C tradeoff is two-fold: Apart from resources, S&C subsystems would also have different preferences on the input distribution. For each component in this tradeoff, we provide design examples illustrating its practical implications. Finally, we conclude this paper with some open challenges.

II. The Torch Metaphor

The fundamental S&C tradeoff in ISAC has been exemplified (implicitly or explicitly) by the "torch metaphor" in some works [5], [12], as illustrated in Fig. 1. In this picture, the role played by the ISAC base station (BS) resembles a child holding a torch: If he points the torch towards the communication user, the user would receive the message, while the sensing target is left in the dark and hence can hardly be seen. On the other hand, if the target is maximally illuminated, the user would receive very weak signals, resulting in highly noise-corrupted messages. This metaphor is intuitive, and it provides us with the following basic understandings of the S&C tradeoff:

- Power allocation across orthogonal or quasi-orthogonal dimensions (in the case of Fig. 1 it is space, or angle, but it could be time, or frequency) offers an immediate way to tradeoff communication and sensing performance;
- Since we are using a unified signal, when the communication user and the sensing target are "close to each other", the S&C tradeoff becomes less prominent.

Inspired by these intuitions, most existing research contributions on ISAC system design focus on power allocation in a generalized sense, including beamforming in MIMO systems [5] and subcarrier power allocation in OFDM systems [13]. Despite the effectiveness of these techniques, we would still wonder, whether the torch metaphor depicts the full picture of the S&C tradeoff.

Recently, it has been found that the torch metaphor does cover the full picture, when the sensing task is to detect the presence of a potential target, under some specific choice of sensing performance metrics [11]. Specifically, in this scenario, during the n-th channel use, the ISAC BS would determine the absence/presence of the target (represented by a state $\eta \in \{0,1\}$) based on an echo $\mathbf{Y}_{\mathrm{s},n}$, while the user aims for decoding the information conveyed in the transmitted ISAC signal $\mathbf{X}_n$, based on its received signal $\mathbf{Y}_{\mathrm{c},n}$. Both

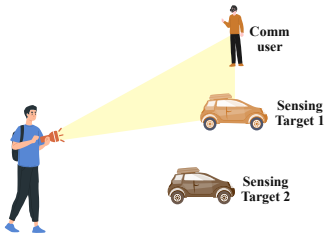


Figure 1. Graphical illustration of the torch metaphor which illustrates that the S&C tradeoff contains ST.

the echo and the user-received signal are contaminated by circularly symmetric Gaussian noises, denoted by $\mathbf{Z}_{\mathrm{s},n}$ and $\mathbf{Z}_{\mathrm{c},n}$, respectively. Such a scenario may be characterized as

$$\mathbf{Y}_{\mathrm{c},n} = \boldsymbol{H}_{\mathrm{c}}\mathbf{X}_n + \mathbf{Z}_{\mathrm{c},n}, \tag{2a}$$

$$\mathbf{Y}_{\mathrm{s},n} = \eta\boldsymbol{H}_{\mathrm{s}}\mathbf{X}_n + \mathbf{Z}_{\mathrm{s},n}, \tag{2b}$$

for $n = 1, 2, \ldots, N$, where N is the coding block length, $\boldsymbol{H}_{\mathrm{c}}$ denotes the communication channel, and $\boldsymbol{H}_{\mathrm{s}}$ represents the target response matrix (also referred to as the sensing channel [9]–[11]). In this problem, both $\boldsymbol{H}_{\mathrm{c}}$ and $\boldsymbol{H}_{\mathrm{s}}$ are fixed over time. For communication, a natural performance metric is the communication rate given by

$$R = \lim_{N\to\infty} \frac{1}{N} \log M_N,$$

under the assumption that the decoding error probability vanishes as N tends to infinity, and M_N denotes the size of the communication codebook. For sensing, the performance metric chosen in [11] is the error exponent[1] defined as

$$E = \lim_{N\to\infty} \frac{1}{N} \log \frac{1}{\delta_N}, \tag{3}$$

where δ_N denotes the maximum detection error probability over all codewords and target states, namely

$$\delta_N = \max_{w\in[M_N]} \max_{s\in\{0,1\}} \mathbb{P}\left\{\hat{S}\left(\mathbf{Y}_{\mathrm{s}}^N, \mathbf{X}^N(\mathsf{W})\right) \neq \eta | \eta = s, \mathsf{W} = w\right\},$$

with $\mathsf{W} \in [M_N]$ being the encoded message, $\mathbf{Y}_{\mathrm{s}}^N = (\mathbf{Y}_{\mathrm{s},1}, \mathbf{Y}_{\mathrm{s},2}, \ldots, \mathbf{Y}_{\mathrm{s},N})$, and $\mathbf{X}^N(\mathsf{W}) = (\mathbf{X}_1(\mathsf{W}), \mathbf{X}_2(\mathsf{W}), \ldots, \mathbf{X}_N(\mathsf{W}))$. Under the aforementioned setting, the set constituted by all achievable pairs (R, E) (termed as the rate-exponent region) is shown to be characterized by [11]

$$R \leqslant \log\left|\boldsymbol{I} + \sigma_{\mathrm{c}}^{-2}\boldsymbol{H}_{\mathrm{c}}\widetilde{\boldsymbol{R}}_{\mathbf{X}}\boldsymbol{H}_{\mathrm{c}}^{\mathrm{H}}\right|, \tag{4a}$$

$$E \leqslant \frac{1}{4}\mathrm{Tr}\left\{\sigma_{\mathrm{s}}^{-2}\boldsymbol{H}_{\mathrm{s}}\widetilde{\boldsymbol{R}}_{\mathbf{X}}\boldsymbol{H}_{\mathrm{s}}^{\mathrm{H}}\right\}, \tag{4b}$$

where $\widetilde{\boldsymbol{R}}_{\mathbf{X}} = \mathbb{E}\left\{\frac{1}{N}\sum_{n=1}^{N}\mathbf{X}_n\mathbf{X}_n^{\mathrm{H}}\right\}$ is the statistical covariance matrix of the transmitted ISAC signal, satisfying a power budget constraint $\mathrm{Tr}\{\widetilde{\boldsymbol{R}}_{\mathbf{X}}\} \leqslant P$. Note that $\mathrm{Tr}\{\boldsymbol{A}\}$ represents the trace of matrix $\boldsymbol{A}$ and $\mathbb{E}\{\cdot\}$ represents the expectation operation.

From (4) we observe a concrete version of the torch metaphor. Apparently, the communication-optimal $\widetilde{\boldsymbol{R}}_{\mathbf{X}}$ would

[1]The error exponent is related to a class of more frequently used sensing performance metrics, i.e., statistical divergences, via the large deviation theory [14].

This article has been accepted for publication in IEEE BITS the Information Theory Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MBITS.2024.3376638

have its eigenvectors matching those of $\boldsymbol{H}_{\rm c}$, while the eigenvalues are determined by the water-filling power allocation [14]. By contrast, the sensing-optimal $\widetilde{\boldsymbol{R}}_{\mathbf{X}}$ would be aligned with the dominant eigenvector (possibly having multiplicity larger than 1) of $\boldsymbol{H}_{\rm s}$ [11]. Therefore, as long as the eigenspaces of $\boldsymbol{H}_{\rm s}$ and $\boldsymbol{H}_{\rm c}$ do not match, or the water-filling strategy does not concentrate the power on the dominant eigenvector, there would be a S&C tradeoff. Furthermore, (4) also gives us an explicit sense in which the communication user and the sensing target are "close to each other": their distance may be characterized by a measure of discrepancy between the "communication subspace" $\mathrm{span}(\boldsymbol{H}_{\rm c})$ and the "sensing subspace" – the dominant eigenvector of $\boldsymbol{H}_{\rm s}$. In light of this, such a tradeoff is termed as the ST in [9], and thus the torch metaphor may be represented concisely as the statement of

$$\text{S\&C tradeoff} = \text{ST}. \tag{5}$$

To further reveal the nature of the ST, let us consider a tangible example, in which the BS is equipped with collocated transmit and receive uniform linear antenna arrays (half-wavelength spacing) of the size $N_{\rm s} = M = 10$, the user has a single antenna (i.e., $N_{\rm c} = 1$), while the sensing target is point-like. In this scenario, the communication and sensing channels are given by

$$\boldsymbol{H}_{\rm c} = \alpha_{\rm c}\boldsymbol{a}^{\rm H}(\theta_{\rm c}), \quad \boldsymbol{H}_{\rm s} = \alpha_{\rm s}\boldsymbol{a}(\theta_{\rm s})\boldsymbol{a}^{\rm H}(\theta_{\rm s}),$$

where $\alpha_{\rm c}$ and $\alpha_{\rm s}$ denote the scalar channel coefficients, $\theta_{\rm c}$ and $\theta_{\rm s}$ denote the bearing angle of the user and the target relative to the BS, respectively, and $\boldsymbol{a}(\theta)$ is the array steering vector given by

$$\boldsymbol{a}(\theta) = [1,\ e^{j\pi\sin(\theta)},\ e^{j2\pi\sin(\theta)}, \ldots, e^{j\pi(N_{\rm s}-1)\sin(\theta)}]^{\rm T}.$$

Correspondingly, the characterization of the rate-exponent region (4) can then be simplified as

$$\boldsymbol{r}(\lambda) = \underset{\boldsymbol{v}:\|\boldsymbol{v}\|=1}{\mathrm{argmax}}\ (1-\lambda)|\boldsymbol{v}^{\rm H}\boldsymbol{a}(\theta_{\rm c})|^2 + \lambda|\boldsymbol{v}^{\rm H}\boldsymbol{a}(\theta_{\rm s})|^2,\ \lambda\in[0,1], \tag{6a}$$

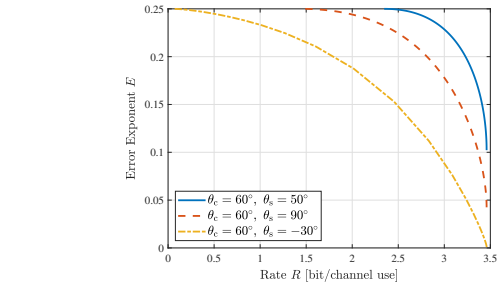
$$R(\lambda) \leqslant \log(1 + P\sigma_{\rm c}^{-2}|\alpha_{\rm c}|^2|\boldsymbol{r}^{\rm H}(\lambda)\boldsymbol{a}(\theta_{\rm c})|^2), \tag{6b}$$

$$E(\lambda) \leqslant \frac{1}{4}\mathrm{Tr}\left\{P\sigma_{\rm s}^{-2}|\alpha_{\rm s}|^2|\boldsymbol{r}^{\rm H}(\lambda)\boldsymbol{a}(\theta_{\rm s})|^2\right\}, \tag{6c}$$

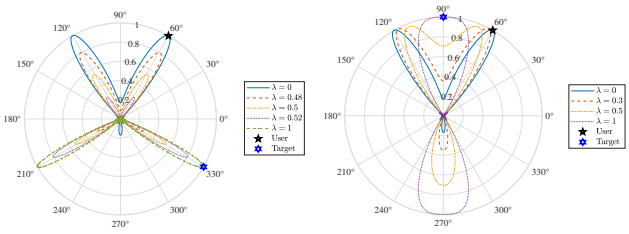
where λ is a parameter controlling the S&C tradeoff. The communication and sensing signal-to-noise ratios (SNRs) may be expressed as

$$\mathrm{SNR}_{\rm c} = P\|\boldsymbol{H}_{\rm c}\|_{\rm F}^2\sigma_{\rm c}^{-2}, \quad \mathrm{SNR}_{\rm s} = P\|\boldsymbol{H}_{\rm s}\|_{\rm F}^2\sigma_{\rm s}^{-2}.$$

Using (6), we may readily obtain the rate-exponent regions for given configurations of $(\theta_{\rm c}, \theta_{\rm s})$, as portrayed in Fig. 2a. Note that the closeness between the S&C subspaces in this scenario is characterized by the intersection angle between $\boldsymbol{a}(\theta_{\rm c})$ and $\boldsymbol{a}(\theta_{\rm s})$. We may observe from Fig. 2a that the S&C tradeoff becomes more prominent as the angular separation between the target and the user increases. More intuitively, as can be seen from Fig. 2a, in the $\theta_{\rm c} = 60°$, $\theta_{\rm s} = 90°$ scenario, the sensing-optimal and communication-optimal beam patterns (corresponding to $\lambda = 0$ and $\lambda = 1$, respectively) have a large overlap. By contrast, in the $\theta_{\rm c} = 60°$, $\theta_{\rm s} = -30°$ scenario, the sensing- and communication-optimal beam patterns are almost



(a) Rate-exponent region



(b) $\theta_{\rm c} = 60°,\ \theta_{\rm s} = 90°$ (c) $\theta_{\rm c} = 60°,\ \theta_{\rm s} = -30°$

Figure 2. The rate-exponent region (as well as its beamspace illustrations) of the single-antenna user, point-like target scenario, with communication and sensing SNRs equal to 10dB and 0dB, respectively. The "ISAC torch" can simultaneously illuminate both the user and the target in case (b), while it has to apply a power-splitting strategy in case (c).

orthogonal to each other. This corroborates our intuition about the ST, that more synergies between S&C tasks would be witnessed as the corresponding subspaces become closer to each other. In the language of the torch metaphor, we may say that the "ISAC torch" in the case of Fig. 2b can simultaneously illuminate both the user and the target, while it has to apply a power-splitting strategy in the case of Fig. 2c.

Now we see that (5) holds true for the target presence detection problem under the sensing metric of error exponent. We would naturally ask in the sequel that:

- Does (5) hold in general?
- If not, what is the key condition for (5) to hold? Is it the metric chosen (error exponent), the specific type of the sensing task (target detection), or both?

To answer these questions, one has to rely on a more general analytical framework, that applies to both estimation and detection tasks and can handle favorably all reasonable sensing performance metrics, besided error exponent. As we shall see later in Sections III and IV, such a general framework gives rise to a more complete picture compared to the torch metaphor (termed as the "projector metaphor"), which takes the empirical distribution (type) of $\mathbf{X}$ into account.

III. Capacity-Distortion Theory

The capacity-distortion theory is conceived in the hope of building the universal framework mentioned at the end of Section II.

A. Rate-Distortion Region

As the name suggests, the capacity-distortion theory investigates the tradeoff between communication capacity and sensing distortion. Originally proposed by Shannon in the

context of rate-distortion theory for lossy data compression [15], distortion refers to a wide range of functions taking the form of $d(\boldsymbol{\eta}, \hat{\boldsymbol{\eta}})$, whose inputs are the true value of some quantity (in sensing problems, the sensing parameter) $\boldsymbol{\eta}$ and its estimate $\hat{\boldsymbol{\eta}}$. Due to the randomness of the communication message, the ISAC waveform (codeword) $\mathbf{X}^N$ is also random. To reflect the sensing performance over a relatively long period of time, a common practice is to use the expectation of the distortion, instead of its instantaneous values, as the performance metric. For example, in estimation tasks, the squared Euclidean distance $d(\boldsymbol{\eta}, \hat{\boldsymbol{\eta}}) = \|\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}\|^2$ is a widely applied distortion function, whose expectation is the mean-squared error (MSE). In binary (e.g., target presence) detection problems with the task of determining the value of a binary variable $\eta \in \{0, 1\}$, a valid distortion function is the Hamming distance given by $d(\eta, \hat{\eta}) = \eta \oplus \hat{\eta}$,[2] whose expectation is related to commonly used detection metrics, including the detection probability P_{D} and the false alarm rate P_{FA}, as follows:

$$\begin{aligned}\mathbb{E}\{\eta \oplus \hat{\eta}\} &= (1 \oplus 1)\mathbb{P}\{\hat{\eta} = 1 | \eta = 1\} &(7)\\ &\quad + (0 \oplus 0)\mathbb{P}\{\hat{\eta} = 0 | \eta = 0\}\\ &\quad + (1 \oplus 0)\mathbb{P}\{\hat{\eta} = 1 | \eta = 0\} &(8)\\ &\quad + (0 \oplus 1)\mathbb{P}\{\hat{\eta} = 0 | \eta = 1\}\\ &= 1 - P_{\mathrm{D}} + P_{\mathrm{FA}}. &(9)\end{aligned}$$

Note that for constant false-alarm rate (CFAR) detectors based on the Neyman-Pearson criterion with fixed P_{FA} [16], minimizing the Hamming distance is equivalent to maximizing the detection probability P_{D}.

Besides capacity and distortion, there is yet another important ingredient in the capacity-distortion theory, namely the transmission cost. To elaborate, not all ISAC waveforms (or codewords) cost equally in terms of wireless resources. For example, the points in quadrature amplitude modulation (QAM) constellations having different amplitudes would yield different power consumptions. Apparently, as the overall resource budget increases, S&C performances can be simultaneously enhanced. Therefore, one has to discuss the capacity-distortion tradeoff under a specific resource budget. Similar to the sensing distortion, in order to take into account the randomness of the codewords, we typically use the expectation of the resource cost over all possible codewords as the measure of transmission cost.

Once the resource budget is given and the sensing distortion metric is chosen, the S&C tradeoff is expressed in terms of the largest achievable rate-distortion region. Formally speaking, given an expected resource budget B, the rate-distortion-cost triple (R, D, B) is achievable (in the infinite block length regime), if there exists a sequence of $(2^{NR}, R)$ codes $\{\mathbf{X}^N | N \in \mathbb{N}\}$ encoding the message $\mathsf{W} \in \{1, \ldots, 2^{NR}\}$, and a state estimator function $\hat{\boldsymbol{\eta}} : (\mathbf{Y}_{\mathrm{s}}, \mathbf{X}^N) \mapsto \hat{\boldsymbol{\eta}}$, such that the

[2]The symbol $\oplus$ represents the "exclusive OR" operation.

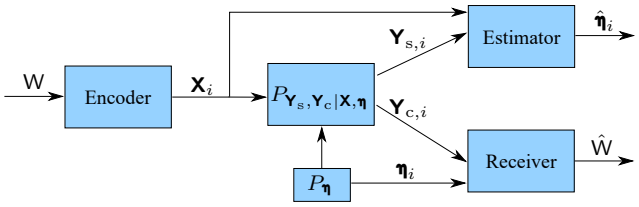


Figure 3. The structure of a generic ISAC system with block-fading state models, often considered in the capacity-distortion theory.

following holds [9]

$$\mathbb{E}\{d(\boldsymbol{\eta}^N, \hat{\boldsymbol{\eta}}^N)\} \leqslant D, \quad (10a)$$

$$\mathbb{E}\{b(\mathbf{X}^N)\} \leqslant B, \quad (10b)$$

$$P_{\mathrm{e}}^{(N)} := \frac{1}{2^{NR}} \sum_{i=1}^{2^{NR}} \mathbb{P}\{\hat{\mathsf{W}} \neq i | \mathsf{W} = i\} \to 0, \quad (10c)$$

as $N \to \infty$, where $b(\cdot)$ denotes the instantaneous cost of single codeword. In addition, for generality, the values of the sensing parameters are allowed to vary across time, denoted by $\boldsymbol{\eta}_i$ at the i-th channel use, constituting a parameter sequence $\boldsymbol{\eta}^N$. The capacity-distortion function given a specific resource budget B is then defined as

$$C_B(D) = \sup\{R | (R, D, B) \text{ is achievable}\}. \quad (11)$$

For the simplicity of discussion, let us consider block-fading state models (also known as in-block memory channels [17]), exemplified by (2), and illustrated in Fig. 3. The effect of such channels can be expressed as

$$\begin{aligned}&p_{\mathsf{W},\mathbf{Y}_{\mathrm{c}}^N,\mathbf{Y}_{\mathrm{s}}^N,\mathbf{X}^N,\boldsymbol{\eta}^N}(W, \boldsymbol{Y}_{\mathrm{s}}, \boldsymbol{Y}_{\mathrm{c}}, \boldsymbol{X}, \boldsymbol{\eta}) = p_{\mathsf{W}}(W)\\ &\quad \times \prod_{i=1}^{N} p_{\boldsymbol{\eta}}(\boldsymbol{\eta}_i) p_{\mathbf{X}|\mathsf{W}}(\boldsymbol{X}_i | \mathsf{W}) p_{\mathbf{Y}_{\mathrm{s}},\mathbf{Y}_{\mathrm{c}}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_{\mathrm{s},i}, \boldsymbol{Y}_{\mathrm{c},i} | \boldsymbol{X}_i, \boldsymbol{\eta}_i).\end{aligned} \quad (12)$$

In particular, the channel $p_{\mathsf{W},\mathbf{Y}_{\mathrm{c}}^N,\mathbf{Y}_{\mathrm{s}}^N,\mathbf{X}^N,\boldsymbol{\eta}^N}(W, \mathbf{Y}_{\mathrm{s}}, \boldsymbol{Y}_{\mathrm{c}}, \boldsymbol{X}, \boldsymbol{\eta})$ produces a communication output $\mathbf{Y}_{\mathrm{c}}$ and a sensing output $\mathbf{Y}_{\mathrm{s}}$, whereas their relationship with the channel input $\mathbf{X}$ can be different from the linear model (2). Moreover, the way that the sensing parameter $\boldsymbol{\eta}$ couples with the channel can also be different. For such channels, it is natural to consider per-block resource budgets and distortion metrics, given by

$$\mathbb{E}\{b(\mathbf{X}^N)\} = \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{b(\mathbf{X}_i)\},$$

$$\mathbb{E}\{d(\boldsymbol{\eta}^N, \hat{\boldsymbol{\eta}}^N)\} = \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{d(\boldsymbol{\eta}_i, \hat{\boldsymbol{\eta}}_i)\}.$$

The conditions (10) can then be concretized as

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{d(\boldsymbol{\eta}_i, \hat{\boldsymbol{\eta}}_i)\} \leqslant D, \quad (13a)$$

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{b(\mathbf{X}_i)\} \leqslant B, \quad (13b)$$

$$\limsup_{N\to\infty} P_{\mathrm{e}}^{(N)} = 0. \quad (13c)$$

In order to gain insights from (13), observe that when the sensing distortion constraint (13a) is absent, the capacity is given by the classical result of "capacity with cost" [18]

$$C_B = \max_{p_{\mathbf{X}}(\boldsymbol{X})\in\mathcal{P}_B} I(\mathbf{X};\mathbf{Y}_{\mathrm{c}}), \tag{14}$$

where

$$\mathcal{P}_B = \{p_{\mathbf{X}}(\boldsymbol{X})|\mathbb{E}\{b(\mathbf{X})\} \leqslant B\}.$$

The result (14) implies that the capacity in the presence of cost budget has a single-letter representation, which would extremely simplify further analysis. Furthermore, since the sensing channel is independent across blocks, the optimal estimator does not rely on historical observations, and hence the expected distortion can be written as a function of the channel input, i.e.,

$$\mathbb{E}\{d(\boldsymbol{\eta}_i,\hat{\boldsymbol{\eta}}_i)|\mathbf{X}=\boldsymbol{X}\} = c(\boldsymbol{X}).$$

In light of this, the capacity-distortion function can now be viewed as a capacity with two costs, namely [9]

$$C_B(D) = \sup_{p_{\mathbf{X}}(\boldsymbol{X})\in\mathcal{P}_B\cap\mathcal{P}_D} I(\mathbf{X};\mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}}), \tag{15a}$$

$$\mathcal{P}_B = \{p_{\mathbf{X}}(\boldsymbol{X})|\mathbb{E}\{b(\mathbf{X})\} \leqslant B\}, \tag{15b}$$

$$\mathcal{P}_D = \{p_{\mathbf{X}}(\boldsymbol{X})|\mathbb{E}\{c(\mathbf{X})\} \leqslant D\}. \tag{15c}$$

Remarkably, the result (15) suggests that the expected sensing distortion may be alternatively viewed as a kind of "sensing-induced communication cost". This understanding enables us to extend even further the range that the capacity-distortion theory is applicable, to scenarios in which the sensing performance cannot be described by a proper distortion function. For example, the CRB widely used as the performance metric of estimation problems is not dependent on any specific estimator $\hat{\boldsymbol{\eta}}$. Sometimes it does not even depend on the true value of the parameter $\boldsymbol{\eta}$. Therefore, CRB is not a distortion function, but it is related to the transmitted codeword $\mathbf{X}$ (as will be discussed later), and hence the capacity-distortion theory can still be applied in the generalized sense.

In this aforementioned scenario, we did not consider feedback. One may wonder whether designing the ISAC codebook by relying on feedback can further enhance the S&C performance. To this end, the state-dependent memoryless channel with delayed feedback (SDMC-DF) model has been investigated [9]. The effect of an SDMC-DF can be expressed as

$$\begin{aligned} &p_{\mathsf{W},\mathbf{Y}^N,\mathbf{Z}^N,\mathbf{X}^N,\boldsymbol{\eta}^N}(W,\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{X},\boldsymbol{\eta}) = p_{\mathsf{W}}(W) \\ &\times\prod_{i=1}^{N} p_{\boldsymbol{\eta}}(\boldsymbol{\eta}_i)p_{\mathbf{X}|\mathsf{W},\mathbf{Z}}(\boldsymbol{X}_i|\mathsf{W},\boldsymbol{Z}_{i-1})p_{\mathbf{Y},\mathbf{Z}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_i,\boldsymbol{Z}_i|\boldsymbol{X}_i,\boldsymbol{\eta}_i) \end{aligned} \tag{16}$$

where both sensing and communication rely on the same channel, which produces the output $\mathbf{Y}^N$.[3] For sensing tasks, the channel input $\mathbf{X}_i$ at the i-th channel use cannot be designed based on the real-time channel output $\mathbf{Y}_i$ or the parameter value $\boldsymbol{\eta}_i$, since the design process has to be causal. Rather, one can only rely on a *delayed feedback* $\mathbf{Z}_{i-1}$, which may be a function of the channel output $\mathbf{Y}_i$.

It turns out that, for SDMC-DF models, (15) also applies, with the slight modification that $\mathbf{Y}_{\mathrm{c}}$ is replaced with $\mathbf{Y}$. To see this, first note that feedback does not improve the capacity of memoryless channels. For the sensing performance, it has been shown that the optimal sensing distortion is achievable by the simple letter-wise minimum expected posterior distortion estimator [9] taking the form of

$$\gamma(\mathbf{X},\mathbf{Z}) = \left(\hat{\boldsymbol{\eta}}^*(\mathbf{X}_1,\mathbf{Z}_1),\ldots,\hat{\boldsymbol{\eta}}^*(\mathbf{X}_N,\mathbf{Z}_N)\right),$$

where

$$\hat{\boldsymbol{\eta}}^*(\boldsymbol{X},\boldsymbol{Z}) = \arg\min_{\boldsymbol{\eta}'} \int d(\boldsymbol{\eta},\boldsymbol{\eta}')p_{\boldsymbol{\eta}|\mathbf{X},\mathbf{Z}}(\boldsymbol{\eta}|\boldsymbol{X},\boldsymbol{Z})\mathrm{d}\boldsymbol{\eta}.$$

We may now see that the expected distortion also admits a single-letter representation

$$c(\boldsymbol{X}) = \mathbb{E}\big\{d\big(\boldsymbol{\eta},\hat{\boldsymbol{\eta}}^*(\mathbf{X},\mathbf{Z})\big)|\mathbf{X}=\boldsymbol{X}\big\}, \tag{17}$$

and thus (15) still applies.

From the aforementioned discussion, we may conclude that a key condition for (15) to hold is that the channel states are independent across blocks. For general channels with memory, even the capacity itself remains open. Furthermore, for such channels, online and offline estimators may have very different performances. These problems deserve further investigation.

B. Computing the Capacity-Distortion Boundary

From a pure theorist's perspective, the result (15) fully characterizes the capacity-distortion region. However, recall that our expectation on the capacity-distortion theory in the first place was to help understand the nature of the S&C tradeoff, hopefully beyond the ST. Unfortunately, we cannot even dig ST itself out of (15), not to mention any further insight.

To serve our purpose, a possible approach is to compute explicitly the capacity-distortion functions for some representative scenarios, in the hope of obtaining useful intuitions. It turns out that the renowned Blahut-Arimoto (B-A) algorithm [19] originally proposed for computing the unconstrained capacity can be applied to compute capacity-distortion functions with some modifications. To elaborate, given an initialization of the trail distribution $r(\boldsymbol{X})$ for $p_{\mathbf{X}}(\boldsymbol{X})$, the original B-A algorithm solving $C = \max_{p_{\mathbf{X}}(\boldsymbol{X})} I(\mathbf{X};\mathbf{Y})$ is a fixed-point iteration repeating the following two steps in each round:

1) Update the trail distribution $q(\boldsymbol{X}|\boldsymbol{Y})$ for the *a posteriori* distribution $p_{\mathbf{X}|\mathbf{Y}}(\boldsymbol{X}|\boldsymbol{Y})$ according to

$$q(\boldsymbol{X}|\boldsymbol{Y}) \leftarrow \frac{p_{\mathbf{X}}(\boldsymbol{X})p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})}{\int r(\boldsymbol{X})p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})\mathrm{d}\boldsymbol{X}}; \tag{18}$$

2) Update $r(\boldsymbol{X})$ by

$$r(\boldsymbol{X}) \leftarrow \frac{e^{\int p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})\log q(\boldsymbol{X}|\boldsymbol{Y})\mathrm{d}\boldsymbol{Y}}}{\int e^{\int p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})\log q(\boldsymbol{X}|\boldsymbol{Y})\mathrm{d}\boldsymbol{Y}}\mathrm{d}\boldsymbol{X}}, \tag{19}$$

[3]Note that (12) and (16) are the most general ISAC channel models appeared in this paper. Especially, (1), (2) and (20) can all be viewed as specific instances of (12). However, our discussion will focus on linear models, since the subspace tradeoff would otherwise be ill-defined.

which is equivalent to solving the following constrained entropy maximization problem:

$$r(\boldsymbol{X}) \leftarrow \arg\max_{p(\boldsymbol{X})} -\int p(\boldsymbol{X}) \log p(\boldsymbol{X}) \mathrm{d}\boldsymbol{X},$$
$$\text{s.t. } p(\boldsymbol{X}) \geqslant 0 \ \forall \boldsymbol{X}, \ \int p(\boldsymbol{X}) \mathrm{d}\boldsymbol{X} = 1,$$
$$\mathbb{E}_{p(\boldsymbol{X})}\left\{ \int p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X}) \log q(\boldsymbol{X}|\boldsymbol{Y}) \mathrm{d}\boldsymbol{Y} \right\} \geqslant t$$

for some constant t.

Let us take the model (12) as an example. Now that our objective function is the conditional mutual information $I(\mathbf{X}; \mathbf{Y}_\mathrm{c}|\mathbf{H}_\mathrm{c})$, we shall replace (18) with

$$q(\boldsymbol{X}|\boldsymbol{Y}_\mathrm{c}, \boldsymbol{\eta}) \leftarrow \frac{p_{\mathbf{X}}(\boldsymbol{X}) p_{\mathbf{Y}_\mathrm{c}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_\mathrm{c}|\boldsymbol{X}, \boldsymbol{\eta})}{\int p_{\mathbf{X}}(\boldsymbol{X}) p_{\mathbf{Y}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_\mathrm{c}|\boldsymbol{X}, \boldsymbol{\eta}) \mathrm{d}\boldsymbol{X}}.$$

In addition, since two extra constraints (15b) and (15c) are now enforced, we should replace (19) with

$$r(\boldsymbol{X}) \leftarrow \frac{e^{\int [p_{\boldsymbol{\eta}}(\boldsymbol{\eta}) p_{\mathbf{Y}_\mathrm{c}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_\mathrm{c}|\boldsymbol{X},\boldsymbol{\eta}) \log q(\boldsymbol{X}|\boldsymbol{Y}_\mathrm{c},\boldsymbol{\eta}) - \lambda b(\boldsymbol{X}) - \mu c(\boldsymbol{X})] \mathrm{d}\boldsymbol{\eta} \mathrm{d}\boldsymbol{Y}_\mathrm{c}}}{\int e^{\int [p_{\boldsymbol{\eta}}(\boldsymbol{\eta}) p_{\mathbf{Y}_\mathrm{c}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_\mathrm{c}|\boldsymbol{X},\boldsymbol{\eta}) \log q(\boldsymbol{X}|\boldsymbol{Y}_\mathrm{c},\boldsymbol{\eta}) - \lambda b(\boldsymbol{X}) - \mu c(\boldsymbol{X})] \mathrm{d}\boldsymbol{\eta} \mathrm{d}\boldsymbol{Y}_\mathrm{c}} \mathrm{d}\boldsymbol{X}},$$

where λ and μ are the Lagrangian multipliers corresponding to the constraints (15b) and (15c), respectively. Similar results can also be obtained under the model (16).

From the above discussions, we may get a vague sense that the sensing distortion requirements render the resulting ISAC signal $\mathbf{X}$ "less random", as they impose constraints on the entropy maximization problem. To validate our intuition, let us consider the toy example of a real-valued, single-input single-output (SISO) Gaussian SDMC-DF channel model (with Rayleigh fading)

$$\mathsf{Y}_i = \eta_i \mathsf{X}_i + \mathsf{N}_i,$$

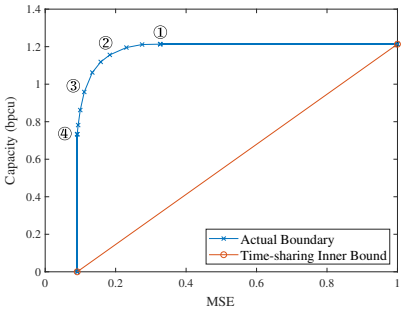
where η_i and N_i are mutually independent zero-mean Gaussian variables with unit variance, while X_i is the ISAC waveform satisfying the power constraint $\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{|\mathsf{X}_i|^2\} \leqslant B = 10\text{dB}$. We consider the perfect feedback $\mathsf{Z}_i = \mathsf{Y}_i$ and the quadratic sensing distortion $d(\eta, \hat{\eta}) = (\eta - \hat{\eta})^2$. In this scenario, the sensing-optimal estimator is the letter-wise minimum mean squared error (MMSE) estimator given by $\hat{\eta}_{i,\mathrm{MMSE}} = \mathbb{E}\{\eta_i|\mathsf{X}_i, \mathsf{Y}_i\}$, whose MSE can be calculated as

$$\frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\{(\eta - \hat{\eta}_{i,\mathrm{MMSE}})^2\} = \mathbb{E}\left\{ \frac{1}{1 + |\mathsf{X}|^2} \right\},$$

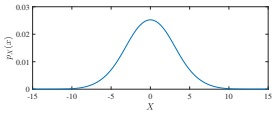
where X is the single-letter representation of the channel input following the probability distribution $p_{\mathsf{X}}(X)$.

Under the aforementioned formulation, the capacity-distortion boundary may be numerically computed using the modified B-A algorithm, as portrayed in Fig. 4. At the communication-optimal point ①, the input distribution is Gaussian, and we have

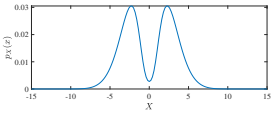
$$C_B(D) = \frac{1}{2} \mathbb{E}\{\log_2(1 + B|\eta|^2)\} \approx 1.213,$$
$$\mathbb{E}\{c(\mathsf{X})\} = \mathbb{E}\left\{ \frac{1}{1 + |\mathsf{X}|^2} \right\} \approx 0.327.$$



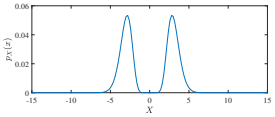
(a) The capacity-distortion boundary



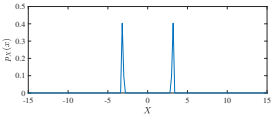
(b) $p_{\mathsf{X}}(X)$ at ①



(c) $p_{\mathsf{X}}(X)$ at ②



(d) $p_{\mathsf{X}}(X)$ at ③



(e) $p_{\mathsf{X}}(X)$ at ④

Figure 4. The capacity-distortion boundary of the real-valued SISO Gaussian channel scenario with $B = 10\text{dB}$, as well as the Pareto-optimal input distributions $p_{\mathsf{X}}(X)$ along the boundary.

By contrast, at the sensing-optimal point ④, the input distribution corresponds to binary phase-shift keying (BPSK) modulation, for which we have

$$C_B(D) \approx 0.733,$$
$$\mathbb{E}\{c(\mathsf{X})\} = \frac{1}{1 + B} \approx 0.091.$$

As we move along the capacity-distortion boundary from the communication-optimal point ① to the sensing-optimal point ④, the corresponding input distribution $p_{\mathsf{X}}(X)$ exhibits a smooth transition from the Gaussian distribution to the BPSK modulation, which agrees with our intuition that the ISAC signal becomes less random as the sensing distortion requirement becomes more stringent.

Note that there is not enough room to accommodate ST in the SISO scenario. Therefore, the S&C tradeoff in the aforementioned example follows solely from the difference in signal preference between communication and sensing tasks. Intuitively, communication tasks would favor signals with a higher degree of randomness, in order to pack more information into the signal. For example, the capacity for AWGN channels is achieved by using Gaussian-distributed signals, where the Gaussian distribution has the maximum entropy under power constraint. By contrast, sensing tasks favor signals that are deterministic in some sense, in order to better distinguish the received signals (echoes) coupled with sensing parameters taking different values. Such a tradeoff is termed as deterministic-random tradeoff (DRT) [10].

Naturally, we would then wonder whether the DRT exists in more general scenarios, what effect it would exhibit, and how it would interact with the ST when it exists. Unfortunately, although the modified B-A algorithm is universal in principle, it is hardly applicable to the general settings in a practical

sense, due to its enormous computational complexity. To elaborate, the modified B-A algorithm relies on numerical integration, which could be computationally prohibitive when the dimensionality of the signals or the sensing parameters is high. For example, if the number of samples per dimension is K, the total number of samples N_{MC} that the widely used Monte Carlo integration method would require will be on the order of $N_{\mathrm{MC}} = O\big(K^{N_{\mathbf{Y}}+N_{\mathbf{X}}+N_{\boldsymbol{\eta}}}\big)$, where $N_{\mathbf{Y}}$, $N_{\mathbf{X}}$ and $N_{\boldsymbol{\eta}}$ are the dimensionalities of $\mathbf{Y}$, $\mathbf{X}$, and $\boldsymbol{\eta}$, respectively. The exponentially increasing sample complexity would easily surpass the capability of most computing devices, even for small-scale MIMO systems. Furthermore, the modified B-A algorithm cannot provide analytical solutions, which can be useful for practical system design.

Considering these difficulties, to push further our understanding of the S&C tradeoff, we might need to sacrifice the generality of the capacity-distortion theory to some degree. For example, we may try and find specific distortion functions that are both easy to analyze and sufficiently general to reflect ST and DRT. In what follows, we shall consider one such example in detail.

IV. CRB-Rate Region

We now focus our attention from the distortion measure $d(\boldsymbol{\eta}, \hat{\boldsymbol{\eta}})$ to a specific sensing metric, namely, CRB for target parameter estimation. Unlike the MSE that specifically relies on the employed estimator, CRB serves as a globally lower bound for all unbiased estimators (satisfying the regularity condition), which usually leads to more tractable analytical expressions, and is achievable at high-SNR regimes [16]. Recalling the single-letterization of the expected distortion in (17), one may treat the CRB as a cost function of the transmitted codeword $\mathbf{X}$, and consider the interplay between the CRB and communication rate as a special case of the C-D tradeoff.

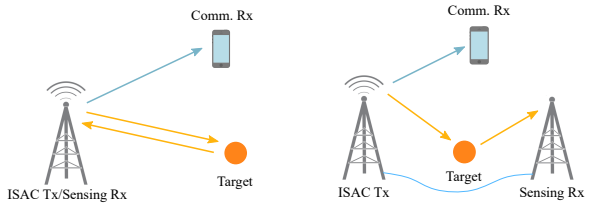
A. Vector Gaussian System Model

We commence by re-examining the model in (2), which may be extended to a more generic form as

$$\mathbf{Y}_{\mathrm{c}} = \mathbf{H}_{\mathrm{c}}\mathbf{X} + \mathbf{Z}_{\mathrm{c}}, \tag{20a}$$

$$\mathbf{Y}_{\mathrm{s}} = \mathbf{H}_{\mathrm{s}}(\boldsymbol{\eta})\mathbf{X} + \mathbf{Z}_{\mathrm{s}}, \tag{20b}$$

where the sensing channel $\mathbf{H}_{\mathrm{s}} \in \mathbb{C}^{N_{\mathrm{s}} \times M}$ is now defined as a deterministic, possibly nonlinear function of the sensing parameter $\boldsymbol{\eta} \in \mathbb{R}^K$, e.g., angular MIMO radar channel [20]. If not otherwise specified, the channel input $\mathbf{X} \in \mathbb{C}^{M \times T}$ will be referred to as an ISAC signal matrix in this section. This model may be viewed as a special case of the generic model shown in Fig. 3. Following the preceding block-fading state model, both the sensing parameter $\boldsymbol{\eta}$ and communication channel $\mathbf{H}_{\mathrm{c}} \in \mathbb{C}^{N_{\mathrm{c}} \times M}$ vary every T symbols in an i.i.d. manner. The discussion on other channels with memory (e.g., Markov channels) is designated as our future works. For convenience, we also assume that the ISAC Tx has perfect knowledge on $\mathbf{H}_{\mathrm{c}}$. Finally, $\mathbf{Z}_{\mathrm{c}}$ and $\mathbf{Z}_{\mathrm{s}}$ are zero-mean white Gaussian noise matrices with variances of σ_{c}^2 and σ_{s}^2, respectively.



(a) Monostatic sensing (b) Bistatic sensing

Figure 5. The ISAC scenarios described in (20), where the dual-functional waveform $\mathbf{X}$ is known to both the ISAC Tx and sensing receiver (Rx).

At this point, it would be worthwhile to concretize the separate channel models in (20) by illustrating the specific scenarios concerned in this section. As shown in Fig. 5, the ISAC Tx emits a dual-functional signal $\mathbf{X}$ to simultaneously communicate with a single communication Rx and sense one or more targets. The sensing Rx is either collocated with the Tx (monostatic sensing), or connected with the Tx with a wired link (bistatic sensing). In both cases, both the Tx and sensing Rx have perfect knowledge about the ISAC signal $\mathbf{X}$. Nevertheless, $\mathbf{X}$ is unknown to the communication Rx as it conveys useful information intended for the communication user. We are thus temped to model $\mathbf{X} \sim p_{\mathbf{X}}(\boldsymbol{X})$ as a random matrix, whose realization is perfectly known at both the ISAC Tx and sensing Rx, but is unknown at the communication Rx.[4]

Under the model (20), the phenomenon that sensing distortion requirements reduce the randomness of the ISAC signal can be characterized with better clarity. To elaborate, the capacity maximization problem amounts to

$$\begin{aligned}\max_{p(\mathbf{X})}\ I(\mathbf{X};\mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}}) &\Longleftrightarrow \max_{p(\mathbf{X})}\ h(\mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}}) - h(\mathbf{Y}_{\mathrm{c}}|\mathbf{X},\mathbf{H}_{\mathrm{c}}) \\ &\Longleftrightarrow \max_{p(\mathbf{X})}\ h(\mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}}).\end{aligned}$$

The sensing distortion requirement would impose constraints on the maximization of the differential entropy $h(\mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}})$, making the resulting $\mathbf{Y}_{\mathrm{c}}$ less random. According to the entropy power inequality and the Brascamp-Lieb inequality, the differential entropy $h(\mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}})$ is both upper- and lower-bounded by the entropies of linear transformations of $\mathbf{H}_{\mathrm{c}}\mathbf{X}$ and $\mathbf{Z}_{\mathrm{c}}$ (conditioned on $\mathbf{H}_{\mathrm{c}}$) [22]. In light of this, sensing distortion requirements might also reduce the entropy $h(\mathbf{X})$. Similar arguments are also applicable to linear Gaussian models belonging to the class of (16) (i.e., with delayed feedback). In the following subsections, we shall further quantize the phenomenon under the framework of CRB-rate region.

B. CRB with Random but Known Nuisance Parameters

Since the communication performance of (20a) can be directly measured by the mutual information $I(\mathbf{X};\mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}})$, we are now in a position to rethink the sensing performance evaluation in ISAC systems. Indeed, in conventional radar systems, probing signals are typically deterministic, well-designed with good ambiguity properties. In ISAC systems, on the other hand, the transmit signal randomly varies from block to block given the communication data embedded in the

[4]For bistatic sensing scenarios where $\mathbf{X}$ is not known to the sensing Rx, interested readers are referred to [21], in which several achievable strategies are proposed.

This article has been accepted for publication in IEEE BITS the Information Theory Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MBITS.2024.3376638

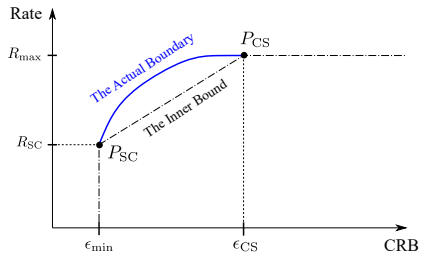


Figure 6. CRB-rate region, Pareto boundary, and time-sharing inner bound.

signal. This imposes a unique challenge in defining the CRB, which now becomes a function of the random signal $\mathbf{X}$.

One possible approach would be to treat $\mathbf{X}$ as a nuisance parameter, and either consider it as a part of the unknown sensing parameter, or integrate it out of the likelihood function describing the observations, leading to classical hybrid or marginal CRB expressions, respectively. Nonetheless, neither of these methods grasps the fundamental feature of the ISAC system, that the random signal $\mathbf{X}$ is *known* to the sensing Rx, and is thus loose in general. To that end, we resort to a Miller-Chang type CRB [23] by computing the CRB for a given instance of $\mathbf{X}$, and take the expectation over $\mathbf{X}$. For any weakly unbiased estimator of $\boldsymbol{\eta}$, the MSE is lower bounded by the Miller-Chang Bayesian CRB in the form of

$$\mathrm{MSE}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\eta}}) \geqslant \mathbb{E}\left\{\mathrm{tr}\left\{\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}}^{-1}\right\}\right\}, \tag{21}$$

where the expectation is taken with respect to $\mathbf{X}$, and $\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}}$ denotes the Bayesian Fisher Information Matrix (BFIM) of $\boldsymbol{\eta}$ given by

$$\begin{aligned}\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}} :=& \mathbb{E}\left\{\frac{\partial \ln p_{\mathbf{Y}_{\rm s}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_{\rm s}|\boldsymbol{X},\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \frac{\partial \ln p_{\mathbf{Y}_{\rm s}|\mathbf{X},\boldsymbol{\eta}}(\boldsymbol{Y}_{\rm s}|\boldsymbol{X},\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^{\rm T}}\middle|\mathbf{X}\right\} \\ &+ \mathbb{E}\left\{\frac{\partial \ln p_{\boldsymbol{\eta}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \frac{\partial \ln p_{\boldsymbol{\eta}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^{\rm T}}\right\}.\end{aligned} \tag{22}$$

More precisely, the BFIM $\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}}$ can be expressed as an affine map of the sample covariance matrix $\mathbf{R}_{\mathbf{X}} = T^{-1}\mathbf{X}\mathbf{X}^{\rm H}$ [10]

$$\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}} = \frac{T}{\sigma_{\rm s}^2}\boldsymbol{\Phi}(\mathbf{R}_{\mathbf{X}}), \tag{23}$$

where

$$\boldsymbol{\Phi}(\boldsymbol{A}) = \sum_{i=1}^{r_1} \widetilde{\boldsymbol{F}}_i \boldsymbol{A}^{\rm T} \widetilde{\boldsymbol{F}}_i^{\rm H} + \sum_{j=1}^{r_2} \widetilde{\boldsymbol{G}}_j \boldsymbol{A} \widetilde{\boldsymbol{G}}_j^{\rm H} + \widetilde{\boldsymbol{J}}_{\rm P}, \tag{24}$$

and $\widetilde{\boldsymbol{J}}_{\rm P} = \sigma_{\rm s}^2 T^{-1} \boldsymbol{J}_{\rm P}$, with the term $\boldsymbol{J}_{\rm P}$ being contributed by the prior distribution $p_{\boldsymbol{\eta}}(\boldsymbol{\eta})$, i.e., the second term on the right hand side of (22). In particular, the matrices $\widetilde{\boldsymbol{F}}_i$ and $\widetilde{\boldsymbol{G}}_j$ are partitioned from the Jacobian matrix $\mathbf{F} := \frac{\partial \mathrm{vec}(\mathbf{H}_{\rm s}^*)}{\partial \boldsymbol{\eta}}$.

By noting (21), it turns out that the Miller-Chang CRB is nothing but an equivalent "expected sensing distortion" discussed in Sec. III-A (though it is not a real distortion measure), and may hence be viewed as a sensing-induced cost imposed on signaling resources. Although with high dimensionality, one may still deduce useful results on the CRB-rate tradeoff by exploiting the affine structure of the BFIM, as discussed in the sequel.

C. CRB-Rate Tradeoff

The CRB-rate tradeoff can be characterized as the following Pareto optimization problem

$$\min_{p_{\mathbf{X}}(\boldsymbol{X})} \quad \alpha\mathbb{E}\left(\mathrm{tr}\left\{[\boldsymbol{\Phi}(\mathbf{R}_{\mathbf{X}})]^{-1}\right\}\right) - (1-\alpha) I(\mathbf{X};\mathbf{Y}_{\rm c}|\mathbf{H}_{\rm c}), \tag{25a}$$

$$\text{s.t.} \quad \mathbb{E}\left(\mathrm{tr}\left\{\mathbf{R}_{\mathbf{X}}\right\}\right) = P_{\rm T}, \tag{25b}$$

where $\alpha \in [0,1]$ is a weighting factor controlling the priority of S&C performance. We highlight here that by moving the CRB from (25a) to (25b), (25) may be equivalently recast as a constrained capacity characterization problem with two cost functions as in (15).

Needless to say, fully depicting the Pareto boundary of the CRB-rate region would incur unaffordable computational overheads, as one has to numerically seek for optimal $p_{\mathbf{X}}(\boldsymbol{X})$ by leveraging the modified B-A algorithm discussed in Sec. III-B. To reveal fundamental insights into the CRB-rate tradeoff, we are interested in the two corner points over the Pareto frontier as shown in Fig. 6, namely, $P_{\rm CS}$, the minimum achievable sensing CRB constrained by the maximum communication capacity, and $P_{\rm SC}$, the maximum achievable rate constrained by the minimized CRB. It is obvious that the line segment linked to two points forms a time-sharing inner-bound. In what follows, we briefly characterize the S&C performance at the two points.

D. $P_{\rm CS}$ Performance Characterization

Let us first examine the point $P_{\rm CS}$. For the point-to-point Gaussian channel, it is well-known that Gaussian distribution is the unique capacity-achieving input distribution (CAID) under an average power cost [24], in which case (25) reduces to a capacity characterization problem with only power constraint by simply letting $\alpha = 0$. More specifically, at $P_{\rm CS}$, each column of $\mathbf{X}$ follows a circularly symmetric complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\rm CS})$ in an i.i.d. manner, where the statistical covariance matrix $\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\rm CS}$ is obtained by solving the following rate maximization problem

$$\begin{aligned} R_{\max} =& \max_{\widetilde{\boldsymbol{R}}\succcurlyeq\mathbf{0}, \widetilde{\boldsymbol{R}}=\widetilde{\boldsymbol{R}}^{\rm H}} \mathbb{E}\left\{\log\left|\boldsymbol{I} + \sigma_c^{-2}\mathbf{H}_c\widetilde{\boldsymbol{R}}\mathbf{H}_c^{\rm H}\right|\right\} \\ &\text{s.t.} \quad \mathrm{tr}\left\{\widetilde{\boldsymbol{R}}\right\} \leqslant P_{\rm T} \\ =& \mathbb{E}\left\{\log\left|\boldsymbol{I} + \sigma_{\rm c}^{-2}\mathbf{H}_{\rm c}\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\rm CS}\mathbf{H}_c^{\rm H}\right|\right\}. \end{aligned} \tag{26}$$

It is readily observed that the optimal solution of (26) has the following eigenvalue decomposition structure

$$\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\rm CS} = \boldsymbol{U}_{\rm c}\boldsymbol{\Lambda}_{\rm c}(\mathbf{H}_{\rm c})\boldsymbol{U}_{\rm c}^{\rm H}, \tag{27}$$

where $\boldsymbol{U}_{\rm c}$ contains the right singular vectors of $\mathbf{H}_{\rm c}$, and $\boldsymbol{\Lambda}_{\rm c}(\mathbf{H}_{\rm c})$ contains the optimal eigenvalues that can be attained from the water-filling method. Accordingly, the optimal ISAC signal structure at $P_{\rm CS}$ is

$$\mathbf{X}^{\rm CS} = \boldsymbol{U}_{\rm c}\boldsymbol{\Lambda}_{\rm c}^{\frac{1}{2}}(\mathbf{H}_{\rm c})\mathbf{D}, \tag{28}$$

where the entries of $\mathbf{D}$ i.i.d. subject to $\mathcal{CN}(0,1)$.

One would then raise the natural question: What is the sensing performance if a Gaussian signal is transmitted? From

(21) to (24), it is obvious that the CRB is determined by the sample covariance matrix $\mathbf{R}_{\mathbf{X}} = T^{-1}\mathbf{X}\mathbf{X}^{\mathrm{H}}$ rather than the statistical covariance matrix $\widetilde{\boldsymbol{R}}_{\mathbf{X}} = \mathbb{E}(\mathbf{R}_{\mathbf{X}})$. At P_{CS}, since each column of $\mathbf{X}$ is i.i.d. Gaussian distributed, $\mathbf{R}_{\mathbf{X}}$ follows a complex Wishart distribution. Note the fact that $\mathrm{Tr}\left([\boldsymbol{\Phi}(\mathbf{R}_{\mathbf{X}})]^{-1}\right)$ is a convex function in $\mathbf{R}_{\mathbf{X}}$. Upon denoting the CRB at P_{CS} as ϵ_{CS}, and based on Jensen's inequality, we have

$$\begin{aligned}\epsilon_{\mathrm{CS}} &\triangleq \frac{T}{\sigma_{\mathrm{s}}^2}\mathbb{E}\left\{\mathrm{Tr}\left([\boldsymbol{\Phi}(\mathbf{R}_{\mathbf{X}})]^{-1}\right)\right\} \geqslant \frac{T}{\sigma_{\mathrm{s}}^2}\mathrm{Tr}\left\{\left(\boldsymbol{\Phi}\left[\mathbb{E}(\mathbf{R}_{\mathbf{X}})\right]\right)^{-1}\right\} \\ &= \frac{T}{\sigma_{\mathrm{s}}^2}\mathrm{Tr}\left\{\left[\boldsymbol{\Phi}\left(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{CS}}\right)\right]^{-1}\right\},\end{aligned} \tag{29}$$

which holds for arbitrarily distributed $\mathbf{R}_{\mathbf{X}}$. This suggests there is certain performance loss for sensing due to the Wishart distributed $\mathbf{R}_{\mathbf{X}}$, since the Jensen lower bound is attained when $\mathbf{R}_{\mathbf{X}} = \mathbb{E}(\mathbf{R}_{\mathbf{X}}) = \widetilde{\boldsymbol{R}}_{\mathbf{X}}$, which holds for Wishart matrix only when $T \to \infty$.

A more non-trivial result in [10] depicts the upper bound of ϵ_{CS}, given by

$$\mathbb{E}\left\{\mathrm{Tr}\left([\boldsymbol{\Phi}(\mathbf{R}_{\mathbf{X}})]^{-1}\right)\right\} \leqslant \frac{T\cdot\mathrm{tr}\left\{\boldsymbol{\Phi}(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{CS}})^{-1}\right\}}{T-\min\{K,\mathsf{M}_{\mathrm{CS}}\}}, \tag{30}$$

with $\mathsf{M}_{\mathrm{CS}} = \mathrm{rank}(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{CS}})$, which clearly indicates that the maximum sensing performance loss at P_{CS} is jointly determined by the number of sensing parameters K and the rank[5] of $\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{CS}}$. Note again that when $T \to \infty$, the sensing performance is lossless since the upper bound converges to its lower counterpart.

E. P_{SC} Performance Characterization

The performance characterization at P_{SC} becomes more challenging compared to that of P_{CS}, as the achieving strategy remains unknown in general. By denoting the CRB at P_{SC} as $\epsilon_{\min}$, and using the Jensen's inequality again, we have

$$\begin{aligned}&\frac{T}{\sigma_{\mathrm{s}}^2}\mathbb{E}\left\{\mathrm{Tr}\left([\boldsymbol{\Phi}(\mathbf{R}_{\mathbf{X}})]^{-1}\right)\right\} \geqslant \frac{T}{\sigma_{\mathrm{s}}^2}\mathrm{Tr}\left\{\left(\boldsymbol{\Phi}\left[\mathbb{E}(\mathbf{R}_{\mathbf{X}})\right]\right)^{-1}\right\} \\ &= \frac{T}{\sigma_{\mathrm{s}}^2}\mathrm{Tr}\left\{\left[\boldsymbol{\Phi}\left(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}}\right)\right]^{-1}\right\} \triangleq \epsilon_{\min}\end{aligned} \tag{31}$$

holds for any complex semidefinite matrix $\mathbf{R}_{\mathbf{X}}$ satisfying the average power constraint, where $\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}}$ is the solution of the deterministic CRB minimization problem

$$\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}} = \underset{\widetilde{\boldsymbol{R}}\succcurlyeq\mathbf{0},\,\widetilde{\boldsymbol{R}}=\widetilde{\boldsymbol{R}}^{\mathrm{H}}}{\arg\min}\ \mathrm{Tr}\left\{\left[\boldsymbol{\Phi}\left(\widetilde{\boldsymbol{R}}\right)\right]^{-1}\right\}\ \ \mathrm{s.t.}\ \mathrm{Tr}\left(\widetilde{\boldsymbol{R}}\right) \leqslant P_{\mathrm{T}}. \tag{32}$$

While problem (32) is a convex semidefinite programming (SDP), it is not strictly convex. As a result, the optimal solution is not necessarily unique. In fact, all the solutions of an SDP belongs to the subspace spanned by the maximum-rank solution, hence may be parameterized as

$$\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}} = \boldsymbol{U}_{\mathrm{s}}\boldsymbol{\Lambda}_{\mathrm{s}}\boldsymbol{U}_{\mathrm{s}}^{\mathrm{H}}, \tag{33}$$

[5]In the high-SNR regime we have $\mathrm{rank}(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{CS}}) = \mathrm{rank}(\mathbf{H}_{\mathrm{c}})$.

where $\boldsymbol{U}_{\mathrm{s}}$ consists of the eigenvectors of the maximum-rank sensing-optimal solution corresponding to the non-zero eigenvalues, while $\boldsymbol{\Lambda}_{\mathrm{s}}$ is a positive semidefinite Hermitian matrix.

Provably, (32) admits a unique solution in many situations [10]. When the solution is unique, the equality in (31) holds iff

$$\mathbf{R}_{\mathbf{X}} = \mathbb{E}(\mathbf{R}_{\mathbf{X}}) = \widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}}. \tag{34}$$

That is, the sample covariance matrix $\mathbf{R}_{\mathbf{X}}$ becomes a deterministic matrix when the global minimum of the CRB $\epsilon_{\min}$ is attained. One may then wonder whether there is any communication DoF left in the ISAC signal at P_{SC}. The answer is non-trivially yes. This is because a deterministic $\mathbf{R}_{\mathbf{X}}$ does not necessarily imply a deterministic $\mathbf{X}$. The latter may still be a random signal conveying information, given by

$$\mathbf{X}^{\mathrm{SC}} = \sqrt{T}(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}})^{\frac{1}{2}}\mathbf{Q} = \sqrt{T}\boldsymbol{U}_{\mathrm{s}}\boldsymbol{\Lambda}_{\mathrm{s}}^{\frac{1}{2}}\mathbf{Q}, \tag{35}$$

where $\mathbf{Q} \in \mathbb{C}^{\mathsf{M}_{\mathrm{SC}}\times T}$ is a random semi-unitary matrix such that $\mathbf{Q}\mathbf{Q}^{\mathrm{H}} = \boldsymbol{I}$, where $\mathsf{M}_{\mathrm{SC}} = \mathrm{rank}(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}})$. Since $(\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}})^{\frac{1}{2}}$ is deterministic, the communication DoFs at P_{SC} are only contributed by the randomness of $\mathbf{Q}$.

We are now ready to characterize the achievable communication rate at P_{SC}. That is, seeking for the optimal distribution $p_{\mathbf{Q}}(\boldsymbol{Q})$ over the set of all $\mathsf{M}_{\mathrm{SC}} \times T$ semi-unitary matrices, namely, the Stiefel manifold $\mathcal{S}(T, \mathsf{M}_{\mathrm{SC}})$, such that the mutual information $I(\mathbf{Q}; \mathbf{Y}_{\mathrm{c}}|\mathbf{H}_{\mathrm{c}})$ is maximized. In the high-SNR regime, this is equivalent to solving the sphere packing problem over the Stiefel manifold, where the optimal $p_{\mathbf{X}}(\boldsymbol{X})$ is uniform distribution, leading to the following asymptotic achievable rate [10]

$$R_{\mathrm{SC}} = \mathbb{E}\Big\{\Big(1-\frac{\mathsf{M}_{\mathrm{SC}}}{2T}\Big)\log|\sigma_{\mathrm{c}}^{-2}\mathbf{H}_{\mathrm{c}}\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\mathrm{SC}}\mathbf{H}_{\mathrm{c}}^{\mathrm{H}}| + \mathsf{c}_0\Big\} + O(\sigma_{\mathrm{c}}^2), \tag{36}$$

where

$$\mathsf{c}_0 = \frac{\mathsf{M}_{\mathrm{SC}}}{T}\Big[\Big(T-\frac{\mathsf{M}_{\mathrm{SC}}}{2}\Big)\log\frac{T}{e} - \log\Gamma(T) + \log(2\sqrt{\pi})\Big] \tag{37}$$

converges to zero as $T \to \infty$.

Observe immediately that when $T \to \infty$, the communication DoFs are lossless, since even a Gaussian matrix would have asymptotically orthogonal rows with the increase of T, making it asymptotically equivalent to a semi-unitary matrix.

F. The Two-fold S&C Tradeoff: A Projector Metaphor

The above results clearly demonstrate the effect of both ST and DRT in an ISAC system. By comparing the left-most parts of $\mathbf{X}^{\mathrm{CS}}$ and $\mathbf{X}^{\mathrm{SC}}$, we see that the communication- and sensing-optimal ISAC signals should be aligned to $\boldsymbol{U}_{\mathrm{c}}$ and $\boldsymbol{U}_{\mathrm{s}}$, respectively, which may be regarded as the orthogonal bases of communication and sensing subspaces. Then the ST is nothing but to flow the signal power from $\mathrm{span}(\boldsymbol{U}_{\mathrm{c}})$ to $\mathrm{span}(\boldsymbol{U}_{\mathrm{s}})$, perfectly fitting the picture of the "ISAC torch" metaphor. More interestingly, by comparing the right-most parts of $\mathbf{X}^{\mathrm{CS}}$ and $\mathbf{X}^{\mathrm{SC}}$, e.g., $\mathbf{D}$ and $\mathbf{Q}$, we see that communication- and sensing-optimal signals adopt Gaussian and semi-unitary codebooks, respectively, which again reflects the DRT. That is, when the ISAC system moves along the Pareto frontier

This article has been accepted for publication in IEEE BITS the Information Theory Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MBITS.2024.3376638

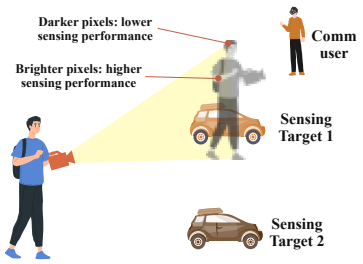


Figure 7. Graphical illustration of the projector metaphor which illustrates that the S&C tradeoff contains ST and DRT simultaneously.

from $P_{\rm CS}$ to $P_{\rm SC}$, $p_{\mathbf{X}}(\boldsymbol{X})$ gradually changes from Gaussian to less random distribution, and eventually becomes uniform distribution over semi-unitary matrices. In that sense, the DRT considered in Fig. 4 is simply a one-dimensional special case, since the semi-unitary matrix reduces to constant-modulus signals in its scalar form, namely, the BPSK modulation in Fig. 4(e).

In addition to the tradeoff between input distributions, the DRT may also be observed in the attainable communication and sensing DoF at the two corner points. At $P_{\rm CS}$, the communication subsystem apparently acquires the full DoF of the Gaussian channel, namely, $\mathsf{M}_{\rm CS}$, which is the rank of $\widetilde{\boldsymbol{R}}_{\mathbf{X}}^{\rm CS}$. As shown in (30), due to the Gaussian signaling, the sensing subsystem suffers from the DoF loss, or, equivalently, the reduction in the number of individual observations of the target, which is up to $\min\{K, \mathsf{M}_{\rm CS}\}$. At $P_{\rm SC}$, the sensing subsystem attains the full DoF of T thanks to the deterministic signal sample covariance matrix. In contrast, the semi-unitary signaling incurs a communication DoF loss of $\frac{\mathsf{M}_{\rm SC}^2}{2T}$, as indicated by (36). Both ST and DRT can be negligible in certain practical scenarios. For example, the sensing and communication subspaces may coincide in free-space optical systems [25], rendering the ST negligible, while the DRT would be negligible when $\min\{K, \mathsf{M}_{\rm CS}\}/T$ is small.

To achieve the S&C performance tradeoff, it is critical to determine the steering direction of the ISAC signal. More importantly, what kind of codebook is transmitted through that direction also matters. With the above understanding, we may now correct the "ISAC torch" metaphor with a more comprehensive picture, i.e., the "ISAC projector". As shown in Fig. 7: A child (the ISAC Tx) holds a projector, and wishes to simultaneously illuminate a target (sensing) while sending an image to a receiver (communication). To form an image, the brightness of each pixel may be used to convey information. Nevertheless, those dark pixels result in imperfect illumination of the target.

V. DRT and ST in Practical ISAC Systems

A. DRT: Sensing with Random Signals

The lessons learned from the DRT inspired us to rethink the practical design philosophy of ISAC systems. In particular, one has to take the randomness of communication data into account while conceiving a sensing strategy, which is a unique challenge emerged in the context of ISAC. In this subsection, we investigate a novel precoding design for sensing with random signals.

Let us consider again the sensing model (20b), with the sensing parameters being the entries of the channel matrix $\mathbf{H}_{\rm s}$, i.e., $\boldsymbol{\eta} = \mathrm{vec}(\mathbf{H}_{\rm s})$. For a given instance of $\mathbf{X}$, the linear minimum MSE (LMMSE) estimator reads

$$\hat{\mathbf{H}}_{\rm LMMSE} = \mathbf{Y}_{\rm s}\left(\mathbf{X}^{\rm H}\boldsymbol{R}_{\mathbf{H}}\mathbf{X} + \sigma_{\rm s}^2 N_{\rm s}\boldsymbol{I}\right)^{-1}\mathbf{X}^{\rm H}\boldsymbol{R}_{\mathbf{H}}, \tag{38}$$

where $\boldsymbol{R}_{\mathbf{H}} = \mathbb{E}(\mathbf{H}_{\rm s}^{\rm H}\mathbf{H}_{\rm s})$ represents the channel correlation matrix. The resulting estimation error is given as

$$\xi_{\mathbf{H}_{\rm s}|\mathbf{X}} = \mathrm{Tr}\left[\left(\boldsymbol{R}_{\mathbf{H}}^{-1} + \frac{1}{\sigma_{\rm s}^2 N_{\rm s}}\mathbf{X}\mathbf{X}^{\rm H}\right)^{-1}\right]. \tag{39}$$

Once again, the estimation error relies on the instantaneous realization of the random ISAC signal $\mathbf{X}$. To depict the average sensing performance, we take the expectation of (39) over $\mathbf{X}$, yielding

$$\xi_{\mathbf{H}_{\rm s}} = \mathbb{E}\left\{\mathrm{Tr}\left[\left(\boldsymbol{R}_{\mathbf{H}}^{-1} + \frac{1}{\sigma_{\rm s}^2 N_{\rm s}}\mathbf{X}\mathbf{X}^{\rm H}\right)^{-1}\right]\right\}. \tag{40}$$

We refer the term (40) to as the ergodic LMMSE (E-LMMSE) [26], as it may be understood as a time average over different realizations of $\mathbf{X}$. Obviously, it is lower-bounded by the Miller-Chang CRB in (21).

We now investigate a specific form of the ISAC signal by letting $\mathbf{X} = \boldsymbol{W}\mathbf{S}$, where $\boldsymbol{W} \in \mathbb{C}^{M\times M}$ is a precoding matrix, and $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_T] \in \mathbb{C}^{M\times T}$ contains column-wise i.i.d. data symbols, satisfying $\mathbb{E}(\mathbf{s}_i) = \mathbf{0}$ and $\mathbb{E}(\mathbf{s}_i\mathbf{s}_i^{\rm H}) = \boldsymbol{I}$. A fundamental question to ask is: What is the optimal precoder $\boldsymbol{W}$ that minimizes the E-LMMSE? In classical MIMO radar waveform design, strictly orthogonal signals are typically employed, namely, $\frac{1}{T}\mathbf{S}\mathbf{S}^{\rm H} = \boldsymbol{I}$, where $\mathbf{S}$ is a semi-unitary matrix, corresponding to the $P_{\rm SC}$ point discussed in Section IV. In such a case, it is known that the LMMSE-optimal precoder has a water-filling structure given by [27]

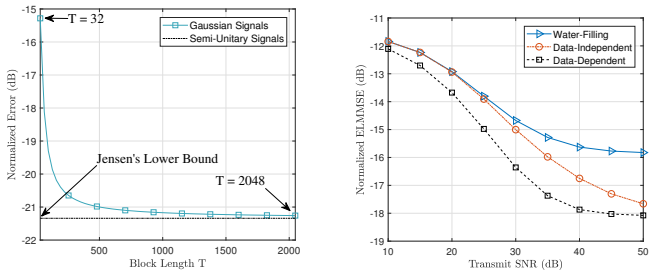
$$\boldsymbol{W}_{\rm WF} = \sqrt{\frac{\sigma_{\rm s}^2 N_{\rm s}}{T}}\boldsymbol{Q}\left[\left(\mu_0\boldsymbol{I} - \boldsymbol{\Lambda}^{-1}\right)^{+}\right]^{\frac{1}{2}}, \tag{41}$$

where $\boldsymbol{Q}$ and $\boldsymbol{\Lambda}$ contain the eigenvectors and eigenvalues of $\boldsymbol{R}_{\mathbf{H}}$, respectively, and μ_0 is a constant meeting the power constraint $\|\boldsymbol{W}_{\rm WF}\|_F^2 = P_{\rm T}$. However, in the ISAC scenario, the water-filling solution (41) may not be optimal due to the randomness in $\mathbf{S}$. Evidently, applying the Jensen's inequality to (40) yields

$$\begin{aligned}\xi_{\mathbf{H}_{\rm s}} &= \mathbb{E}\left\{\mathrm{Tr}\left[\left(\boldsymbol{R}_{\mathbf{H}}^{-1} + \frac{1}{\sigma_{\rm s}^2 N_{\rm s}}\boldsymbol{W}\mathbf{S}\mathbf{S}^{\rm H}\boldsymbol{W}^{\rm H}\right)^{-1}\right]\right\}\\ &\overset{(a)}{\geqslant} \mathrm{Tr}\left[\left(\boldsymbol{R}_{H}^{-1} + \frac{1}{\sigma_{\rm s}^2 N_{\rm s}}\mathbb{E}_{\boldsymbol{S}}\left\{\boldsymbol{W}\mathbf{S}\mathbf{S}^{\rm H}\boldsymbol{W}^{\rm H}\right\}\right)^{-1}\right]\\ &\overset{(b)}{=} \mathrm{Tr}\left[\left(\boldsymbol{R}_{H}^{-1} + \frac{T}{\sigma_{\rm s}^2 N_{\rm s}}\boldsymbol{W}\boldsymbol{W}^{H}\right)^{-1}\right],\end{aligned} \tag{42}$$

where (a) is due to the convexity of $\xi_{\mathbf{H}_{\rm s}|\mathbf{X}}$ with respect to $\mathbf{S}\mathbf{S}^{\rm H}$, and (b) holds from the fact $\mathbb{E}(\mathbf{S}\mathbf{S}^{\rm H}) = T\boldsymbol{I}$. The equality in (a) holds only asymptotically[6] for $\frac{T}{M} \to \infty$, since $\frac{1}{T}\mathbf{S}\mathbf{S}^{\rm H} \neq \boldsymbol{I}$ due to the i.i.d. assumption in the columns of $\mathbf{S}$.

[6]To the best of our knowledge, most existing ISAC precoding literature overlooked the data randomness by assuming $\frac{1}{T}\mathbf{S}\mathbf{S}^{\rm H} \approx \boldsymbol{I}$.



(a) $M = 64, N_{\rm s} = 32$ (b) $M = 64, N_{\rm s} = 32, T = 32$

Figure 8. Sensing with random signals. (a). Tightness of the Jensen bound with an increasing block length. (b). Estimation performance of the DRT-aware precoding designs.

It turns out that the water-filling solution (41) minimizes the Jensen lower bound in (42), rather than $\xi_{\mathbf{H}_{\rm s}}$ itself. Specifically, when T is comparable to M, and when non-unitary codebooks, e.g., Gaussian codebook (where $\mathbf{S}\mathbf{S}^{\rm H}$ follows Wishart distribution), are employed, the orthogonality of $\mathbf{S}$ breaks and the Jensen bound is no longer tight. To see this, we show an example in Fig. 8a with $M = 64, N_{\rm s} = 32$, where the water-filling precoder (41) is applied to both Gaussian distributed and semi-unitary data matrix. The Jensen bound (semi-unitary sensing performance) is attained when $T \geqslant 2048$. To account for the random nature of the ISAC signal, one needs to develop novel precoders that directly minimize the E-LMMSE, rather than its Jensen lower bound. Here we briefly introduce two possible methodologies [26], namely, data-dependent and data-independent designs, both of which yield better performance than that of the water-filling precoder.

1) Data-Dependent Precoding: Despite its randomness, the fact that $\mathbf{S}$ is known at the ISAC Tx enables a precoding design based on given instances of $\mathbf{S}$. Let us denote the nth realization of $\mathbf{S}$ as $\boldsymbol{S}_n$, and the to-be-designed precoder as $\boldsymbol{W}_n$. One may then directly minimize $\xi_{\mathbf{H}_{\rm s}|\mathbf{X}}$ based on the known $\boldsymbol{S}_n$ by solving the following problem

$$\min_{\|\boldsymbol{W}_n\|_F^2 = P_{\rm T}} \mathrm{Tr}\left[\left(\boldsymbol{R}_{\mathbf{H}}^{-1} + \frac{1}{\sigma_{\rm s}^2 N_{\rm s}} \boldsymbol{W}_n \boldsymbol{S}_n \boldsymbol{S}_n^{\rm H} \boldsymbol{W}_n^{\rm H}\right)^{-1}\right]. \quad (43)$$

While problem (43) is non-convex, it is provable that it admits an optimal closed-form solution [28]. Consequently, one may minimize the E-LMMSE $\xi_{\mathbf{H}_{\rm s}}$ by minimizing every instance of $\xi_{\mathbf{H}_{\rm s}|\mathbf{X}}$.

2) Data-Independent Precoding: The optimality of the data-dependent precoder is achieved at the price of high complexity, as one has to solve for $\boldsymbol{W}_n$ for every instance $\boldsymbol{S}_n$. To ease the computational burden, an alternative option would be to conceive a data-independent precoder, where a single $\boldsymbol{W}$ is leveraged for every instance of $\mathbf{S}$, leading to the following stochastic optimization problem

$$\min_{\|\boldsymbol{W}_n\|_F^2 = P_{\rm T}} \mathbb{E}\left\{\mathrm{Tr}\left[\left(\boldsymbol{R}_{\mathbf{H}}^{-1} + \frac{1}{\sigma_{\rm s}^2 N_{\rm s}} \boldsymbol{W}\mathbf{S}\mathbf{S}^{\rm H}\boldsymbol{W}^{\rm H}\right)^{-1}\right]\right\}, \quad (44)$$

which can be solved via stochastic gradient descent (SGD) algorithm in an offline manner, where massive training data samples may be locally generated based on the adopted communication codebook.

To validate the performance of the proposed sensing precoding design for random signals, we show their corresponding estimation errors in Fig. 8b with $M = 64, T = 32$ and $N_{\rm s} = 32$, where a Gaussian codebook is employed again for generating $\mathbf{S}$. As expected, both precoding designs significantly outperform the classical water-filling approach (41). In particular, the computationally expensive data-dependent design achieves better average estimation performance (0.4-1.4 dB gain) compared to its data-independent counterpart, while the latter attains a favorable performance-complexity tradeoff. This provides strong evidence that the data randomness is non-negligible in ISAC signaling. To achieve the S&C performance boundary, DRT-aware ISAC precoding techniques are yet to be implemented in practical ISAC systems.

B. Frequency-domain ST: Valuating Sensing Resources

Let us now turn our attention to the ST. In previous sections, we have illustrated the ST using spatial-domain examples. In these examples, roughly speaking, the ST is manifested as the preference for direct-illuminating beams, which holds for both the communication user and the sensing target. By contrast, when we consider non-spatial scenarios, the ST might take a less intuitive form, but it would also be more interesting and non-trivial.

To see this, let us consider the example of ranging waveform design, characterized by the simple observation model

$$y(t) = s(t - \tau) + n(t),$$

where $y(t)$, $s(t)$ and $n(t)$ represent the received signal, the transmitted signal, and the noise with constant power spectral density (PSD) N_0, respectively. The term $\tau = d/c$ denotes the propagation delay, with c being the propagation speed, and d being the distance to be estimated. For this model, the CRB reads

$$\mathbb{E}\{(d - \hat{d})^2\} \geqslant c^2 (8\pi^2 \beta^2 \mathrm{SNR})^{-1}, \quad (45)$$

where $\beta = \int_{-\infty}^{\infty} f^2 |S(f)|^2 \mathrm{d}f / \int_{-\infty}^{\infty} |S(f)|^2 \mathrm{d}f$ is referred to as the "root-mean-square (RMS) bandwidth" [29], while $\mathrm{SNR} = \frac{1}{N_0} \int_{-\infty}^{+\infty} s^2(t) \mathrm{d}t$.

What does (45) imply? Upon assuming that the signal is constrained to reside in the frequency interval of $[0, f_{\rm high}]$, we find immediately that the CRB-optimal signal is in fact a sinusoidal signal with frequency $f_{\rm high}$, which maximizes the RMS bandwidth. The intuition behind this result is that, as long as the integer ambiguity can be resolved, ranging methods based on carrier phase sensing would yield the optimal performance, as has been recognized in the practice of global navigation satellite system (GNSS)-based positioning.

Of course, if not supplemented by further information, the integer ambiguity of a single-tone signal can never be resolved. However, CRB is known to be unable to capture the ambiguity phenomenon. To this end, we may use the Ziv-Zakai bound (ZZB) given by [30]

$$\mathbb{E}\left\{(d - \hat{d})^2\right\} \geqslant \int_0^{\epsilon_{\max}} x Q\left(\sqrt{2^{-1} \mathrm{SNR}(1 - \widetilde{R}(x))}\right) \mathrm{d}x, \quad (46)$$

where $\epsilon_{\max}$ is the maximum possible ranging error, $Q(\cdot)$ denotes the Marcum Q-function, $\widetilde{R}(x)$ denotes the normalized autocorrelation function (ACF) defined as $\widetilde{R}(x) =$

This article has been accepted for publication in IEEE BITS the Information Theory Magazine. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MBITS.2024.3376638

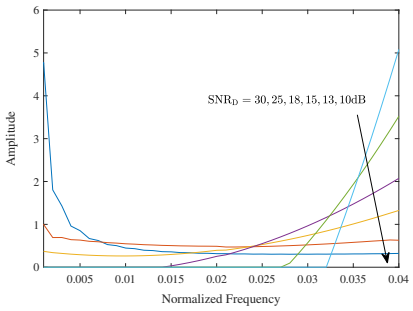
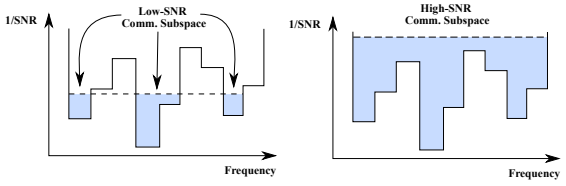
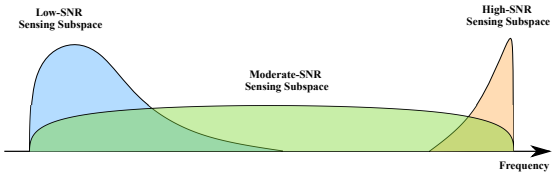


Figure 9. The PSDs of ZZB-optimal waveforms at different SNRs.



(a) Communication Subspace



(b) Sensing Subspace

Figure 10. Frequency-domain communication and sensing subspaces (for the ranging task) in low- and high-SNR regimes.

$R(x/c)/R(0)$, with $R(\tau) = \int_{-\infty}^{\infty} s(t-\tau)s(t)\mathrm{d}t$. Although it does not admit a closed-form expression, we may observe from (46) that the ZZB can reflect the ambiguity phenomenon: It is a decreasing function with respect to $(1 - \widetilde{R}(x))$. Since the normalized ACF $\widetilde{R}(x)$ achieves its maximum at $\widetilde{R}(0) = 1$, $(1 - \widetilde{R}(x))$ can be viewed as a measure of the sidelobe level. Intuitively, given a fixed noise level, a higher sidelobe level would make it less distinguishable from the mainlobe, and hence cause larger errors.

With the aid of ZZB, we are now able to understand the behaviour of waveforms that achieve (near-) optimal sensing performance. Numerically computed PSDs of ZZB-optimal waveforms (using the method in [30]) are plotted in Fig. 9. Observe that as the total SNR increases, the ZZB-optimal waveform would refocus its power from the low-frequency band to the high-frequency band. The reason is that when the total SNR is sufficiently high, one may effectively resolve the ambiguity caused by relatively high sidelobes, and hence the power is focused on the high-frequency band. By contrast, when the total SNR is lower, one needs to lower the sidelobe level to fight against the ambiguity issue, which would inevitably widen the mainlobe, leading to low-frequency waveforms.[7]

Note that this is a remarkable observation suggesting that sensing tasks have a unique preference on the subspaces where the ISAC signal resides. This is in stark contrast to communication tasks, for which the optimal frequency-domain power allocation scheme is the water-filling strategy. In particular, the water-filling strategy assigns more power to frequency bands having higher SNR, and might even abandon some low-SNR bands when the total power constraint is stringent, or equivalently, when the total SNR is relatively low. In the language of ST, we may say that in the frequency domain, the communication subspace corresponds to those frequency bands with high SNR. By contrast, the sensing subspace is not solely dependent on the SNR; rather, as the total SNR increases, the sensing subspace would move from the low-frequency band to the high-frequency band, as portrayed in Fig. 10.

We may obtain further insights from the perspective of resource allocation. One of the resources of communication tasks is the DoF, which is manifested as the bandwidth in the frequency domain. For communication tasks, the value of the frequency bands only depends on their *quality* (in terms of SNR). However, for sensing tasks, the value of frequency bands depends not only on their quality, but on their *location* as well. In light of this, we may say that the DoF is not, in its nature, a sensing resource. We are thus motivated to ask the following question: *What do we really mean when we say "sensing resources"?*

VI. Concluding Remarks and Open Challenges

Unfortunately, at the moment of writing, we do not have a well-stated answer to this question. After all, in contrast to communication tasks always aiming to deliver information, sensing tasks have vastly diverse purposes, and hence may rely on different resources. Apart from this question, many important problems remain open in the context of S&C tradeoffs. To name a few:

1) What can we learn from a more general formulation of detection-oriented ISAC scenarios (e.g., those in which the channel state information is unavailable)? How does the DRT manifest itself under generic sensing performance metrics?
2) How do we characterize the S&C tradeoff for generic channels with memory? What are the performances of online and offline estimators in such scenarios?
3) How do we design ISAC systems that are capable of achieving the entire capacity-distortion boundary (not just the corner points)[8]?
4) Can we unify S&C performance metrics?

These challenging questions remind us that there is still a long way to go before the merit of ISAC can be fully understood and utilized. Nevertheless, the ST-DRT decomposition (i.e., the "projector metaphor") is likely to be a useful meta-intuition in future investigations of ISAC systems: the fundamental tradeoff in ISAC is manifested as the preference discrepancies between S&C tasks, concerning both the resources (ST) and the signal patterns (DRT).

[7]The curvature of the ACF mainlobe at $\tau = 0$ is proportional to the RMS bandwidth [29]. Therefore, wider mainlobes correspond to lower-frequency waveforms.

[8]A special case where we do know the entire boundary is the detection-communication scenario when the error exponent is chosen as the sensing performance metric, for which only the ST persists.

REFERENCES

[1] ITU-R WP5D, "Draft New Recommendation ITU-R M. [IMT.FRAMEWORK FOR 2030 AND BEYOND]," 2023.

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

[3] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proc. IEEE*, vol. 99, no. 7, pp. 1236–1259, Jul. 2011.

[4] Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Network*, vol. 35, no. 5, pp. 158–167, Sep. 2021.

[5] F. Liu, Y.-F. Liu, A. Li, C. Masouros, and Y. C. Eldar, "Cramér-Rao bound optimization for joint radar-communication beamforming," *IEEE Trans. Signal Process.*, vol. 70, pp. 240–253, 2022.

[6] H. Zhang, H. Zhang, B. Di, M. D. Renzo, Z. Han, H. V. Poor, and L. Song, "Holographic integrated sensing and communication," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2114–2130, 2022.

[7] D. W. Bliss, "Cooperative radar and communications signaling: The estimation and information theory odd couple," in *Proc. 2014 IEEE Radar Conf.*, Covington, KY, USA, May 2014, pp. 0050–0055.

[8] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[9] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An information-theoretic approach to joint sensing and communication," *IEEE Trans. Inf. Theory*, pp. 1–1, *early access*, 2022.

[10] 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 Trans. Inf. Theory*, vol. 69, no. 9, pp. 5723–5751, 2023.

[11] H. Joudeh and F. M. J. Willems, "Joint communication and binary state detection," *IEEE J. Sel. Areas Inf. Theory*, vol. 3, no. 1, pp. 113–124, 2022.

[12] H. Hua, T. X. Han, and J. Xu, "MIMO integrated sensing and communication: CRB-rate tradeoff," *IEEE Trans. Wireless Commun.*, pp. 1–1, *early access*, 2023.

[13] M. Bică and V. Koivunen, "Radar waveform optimization for target parameter estimation in cooperative radar-communications systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2314–2326, 2019.

[14] T. M. Cover, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[15] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Int. Conv. Rec.*, vol. 7, no. 4, pp. 142–163, 1993.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ, USA: Prentice Hall, 1998.

[17] S. Li and G. Caire, "On the capacity and state estimation error of "beam-pointing" channels: The binary case," *IEEE Trans. Inf. Theory*, vol. 69, no. 9, pp. 5752–5770, Sep. 2023.

[18] A. El Gamal and Y.-H. Kim, *Network Information Theory*, 1st ed. Cambridge university press, 2011.

[19] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460–473, 1972.

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

[21] T. Jiao, Y. Geng, Z. Wei, K. Wan, Z. Yang, and G. Caire, "Information-theoretic limits of bistatic integrated sensing and communication," *arXiv preprint*, 2023. [Online]. Available: https://arxiv.org/abs/2306.06648

[22] V. Anantharam, V. Jog, and C. Nair, "Unifying the Brascamp-Lieb inequality and the entropy power inequality," *IEEE Trans. Inf. Theory*, vol. 68, no. 12, pp. 7665–7684, Dec. 2022.

[23] R. Miller and C. Chang, "A modified Cramér-Rao bound and its applications (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 398–400, May 1978.

[24] Y. Polyanskiy and Y. Wu, *Information Theory: From Coding to Learning*, 1st ed. Cambridge University Press, 2023.

[25] J. Li, M. Zhang, D. Wang, S. Wu, and Y. Zhan, "Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication," *Optics express*, vol. 26, no. 8, pp. 10 494–10 508, 2018.

[26] S. Lu, F. Liu, F. Dong, Y. Xiong, J. Xu, Y.-F. Liu, and S. Jin, "Random ISAC signals deserve dedicated precoding," *arXiv preprint*, 2023. [Online]. Available: https://arxiv.org/abs/2311.01822

[27] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 330–343, 2007.

[28] B. Tang, J. Tang, and Y. Peng, "MIMO radar waveform design in colored noise based on information theory," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4684–4697, 2010.

[29] A. Liu *et al.*, "A survey on fundamental limits of integrated sensing and communication," *IEEE Commun. Surv. Tuts.*, vol. 24, no. 2, pp. 994–1034, 2022.

[30] Y. Xiong and F. Liu, "SNR-adaptive ranging waveform design based on Ziv-Zakai bound optimization," *IEEE Signal Process. Lett.*, *early access*, 2023.