Constellation Shaping for OFDM-ISAC Systems: From Theoretical Bounds to Practical Implementation

Benedikt Geiger, *Graduate Student Member, IEEE*, Fan Liu, *Senior Member, IEEE*, Shihang Lu, *Graduate Student Member, IEEE*, Andrej Rode, *Graduate Student Member, IEEE*, Daniel Gil Gaviria, *Graduate Student Member, IEEE*, Charlotte Muth, *Graduate Student Member, IEEE*, and Laurent Schmalen, *Fellow, IEEE*

***Abstract*—Integrated sensing and communications (ISAC) promises new use cases for mobile communication systems by reusing the communication signal for radar-like sensing. However, sensing and communications (S&C) impose conflicting requirements on the modulation format, resulting in a trade-off between their corresponding performance. This paper investigates constellation shaping as a means to simultaneously improve S&C performance in orthogonal frequency division multiplexing (OFDM)-based integrated sensing and communications (ISAC) systems. We begin by deriving how the transmit symbols affect detection performance and derive theoretical lower and upper bounds on the maximum achievable information rate under a given sensing constraint. Using an autoencoder-based optimization, we investigate geometric, probabilistic, and joint constellation shaping, where joint shaping combines both approaches, employing both optimal maximum a-posteriori decoding and practical bit-metric decoding. Our results show that constellation shaping enables a flexible trade-off between S&C, can approach the derived upper bound, and significantly outperforms conventional modulation formats. Motivated by its practical implementation feasibility, we review probabilistic amplitude shaping (PAS) and propose a generalization tailored to ISAC. For this generalization, we propose a low-complexity log-likelihood ratio computation with negligible rate loss. We demonstrate that combining conventional and generalized PAS enables a flexible and low-complexity trade-off between S&C, closely approaching the performance of joint constellation shaping.**

***Index Terms*—Constellation shaping, integrated sensing and communication (ISAC), 6G, OFDM, end-to-end learning.**

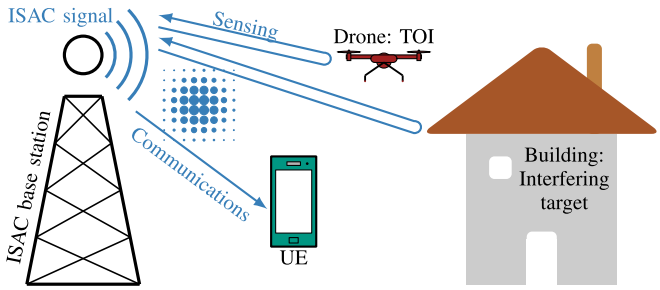


Fig. 1. Considered ISAC scenario. A base station employs a unified transmit signal with an optimized constellation to communicate with a UE and to sense the environment in a radar-like manner. The objective of the ISAC base station is to transmit data to the UE and to detect a target of interest (TOI), such as a drone, in the presence of an interfering object like a building. Constellation shaping enables a flexible trade-off between communication throughput and detection probability of potential targets, and it improves both simultaneously compared with legacy modulation formats.

Received 4 September 2025; revised 17 December 2025 and 2 February 2026; accepted 19 February 2026. Date of publication 2 March 2026; date of current version 19 March 2026. This work has received funding from the German Federal Ministry of Research, Technology and Space (BMFTR) within the projects Open6GHub, Open6GHub+, KOMSENS-6G (grant agreements 16KISK010, 16KIS2405, and 16KISK123) and the National Natural Science Foundation of China (NSFC) under Grant 62522107. Benedikt Geiger acknowledges the Networking Grant from the Karlsruhe House of Young Scientists (KHYS). An earlier version of this paper was presented in part at the IEEE International Symposium on Joint Communications and Sensing (JC&S), Oulu, Finland, January 2025 [DOI: 10.1109/JCS64661.2025.10880638]. The associate editor coordinating the review of this article and approving it for publication was X. Yuan. *(Corresponding author: Benedikt Geiger.)*

Benedikt Geiger, Andrej Rode, Daniel Gil Gaviria, Charlotte Muth, and Laurent Schmalen are with the Communications Engineering Laboratory (CEL), Karlsruhe Institute of Technology (KIT), 76187 Karlsruhe, Germany (e-mail: benedikt.geiger@kit.edu).

Fan Liu is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: fan.liu@seu.edu.cn).

Shihang Lu is with the School of Automation and Intelligent Manufacturing, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: lush2021@mail.sustech.edu.cn).

Digital Object Identifier 10.1109/TCOMM.2026.3669491

I. Introduction

AUTONOMOUS drones, intelligent transportation, and smart city infrastructures are novel applications that require precise environmental awareness seamlessly integrated with communications. Consequently, ISAC has emerged as a key research direction for 6G, reusing communication signals for radar-like sensing of passive objects. Integrated sensing and communications (ISAC) is expected not only to reduce hardware redundancy compared to two separate systems but also to improve energy efficiency and overall system reliability of existing cellular networks [2], [3].

Current 5G systems use OFDM due to its numerous and significant advantages, e.g., high spectral efficiency, low-complexity equalization, easy rate adaptation, and easy multi-user downlink handling [4]. Beyond its role in communications, OFDM offers also key advantages for radar systems. In particular, it has the lowest average ranging sidelobe for quadrature amplitude modulation (QAM) constellations

[5] among all orthogonal signaling bases, and enables low-complexity range and velocity estimation of targets through fast Fourier transform (FFT)-based processing [6]. This makes OFDM also a compelling choice for future 6G ISAC systems [2]. Fig. 1 shows a sketch of the ISAC scenario considered in this paper, which follows a standard setup from the literature [2], [7], [8]. The key idea is to use the same transmit signal of, e.g., the base station, for both communications to the user equipment (UE) and sensing. In particular, the ISAC transmit signal is used to convey information to a communication user, and to detect a potential target in a radar-like manner.

However, a fundamental difference between OFDM-based radar and communication systems lies in the modulation format. Orthogonal frequency division multiplexing (OFDM)-radar systems commonly use quadrature phase shift keying (QPSK) [6], while communication systems employ higher-order modulation formats, such as 64-QAM, to increase spectral efficiency [8]. This is a result of the underlying trade-off between S&C, known as the deterministic-random trade-off (DRT) [9], [10], which becomes particularly evident when considering the optimal channel input distributions for each task. To elaborate, a Gaussian distributed input maximizes the mutual information (MI) for an additive white Gaussian noise (AWGN) channel under an average power constraint, whereas constant modulus constellations achieve optimal sensing performance [9]. Consequently, both S&C performance are heavily influenced by the constellation. Therefore, constellation shaping was recently investigated for ISAC to trade off MI against sidelobe levels in the ambiguity function, aiming to balance S&C performance [7], [11]. While constellation design plays a critical role in communication performance, its influence on sensing-oriented metrics, e.g., detection probability, remains insufficiently explored. In practical OFDM-ISAC systems, it is still unclear how different constellations impact detection performance, or what trade-offs arise when balancing sensing accuracy with communication efficiency.

Constellation shaping, long studied in communications theory, can offer up to 1.53 dB signal-to-noise ratio (SNR) gain over uniform signaling [12] in an AWGN channel with an average power constraint. Yet, in the context of ISAC, the fundamental limits remain unknown. In particular, the maximum achievable MI constrained by sensing performance requirement has yet to be characterized.

Generally, there are three widely used approaches to constellation shaping: *Geometric shaping* optimizes the location of constellation points, assuming equal probability for each point [13]. *Probabilistic shaping* uses a conventional QAM modulation format but optimizes the probability distribution of these points [14]. *Joint shaping* simultaneously optimizes both the location and probability of the constellation points [15].

These different shaping methods have been successfully optimized and compared for communications using an autoencoder (AE) framework [15], [16]. Here, the communication system is modeled by differentiable blocks, and the constellation points and their probabilities are treated as trainable parameters. In contrast, within the context of ISAC, only probabilistic constellation shaping has been studied in detail [7], [11], [17], and a systematic comparison between different shaping approaches is still missing. In particular, it remains an open question which shaping method is best suited for ISAC systems, and how closely it can approach the maximum MI under a given sensing constraint?

While previous work has addressed low-complexity constellation optimization [11], the optimization itself can be done offline, with the resulting constellations stored as look-up tables (LUTs). Therefore, the key challenge is to design constellations that enable low-complexity implementation in practical systems.

Various schemes like Gallager's scheme, trellis shaping, or shell mapping have been proposed to bring probabilistic constellation shaping to practice, see e.g., [18] and references therein. However, these often suffer from severe error propagation unless mitigated by computationally intensive algorithms. A key breakthrough was PAS, which efficiently integrates forward error correction (FEC) into shaping, mitigating error propagation [18], [19]. PAS has become the state-of-the-art in commercial fiber-optic communication systems [20] and is now under investigation for wireless applications [21], [22]. However, conventional PAS has been designed for communication systems, without considering sensing requirements. As a result, its structural constraints may limit the sensing performance, raising critical questions for ISAC: How do the structural constraints of PAS affect the trade-off between S&C? Can PAS be generalized to improve S&C performance while preserving its low-complexity implementation advantage? How does the achievable information rate (AIR) in practical ISAC systems with bit-metric decoding (BMD) compare to the MI?[1]

In this paper, we extend our previous work [1] investigating constellation shaping as a tool to improve S&C performance in a monostatic ISAC system with a unified waveform for both S&C. Specifically, we optimize and systematically compare constellation shaping methods to maximize the AIR under constraints on the sensing detection probability and the false alarm rate. The overarching goal of this work is, first, to establish fundamental performance limits and, second, to bridge the gap between theory and practice by demonstrating how these limits can be approached using low-complexity, practical systems. The key contributions of this paper are:

- **Theoretical characterization of the DRT:** We derive the constellation-dependent detection probability and show that it depends only on the kurtosis of the constellation. Based on this result, we derive lower and upper bounds on the maximum MI given a detection probability constraint. This establishes a fundamental limit on the ISAC performance and provides insight into the maximum achievable gain through constellation shaping.

[1]We note that the authors in [17] mention that constellation shaping can be implemented using PAS, however they do not incorporate the structural constraints associated with PAS into their optimization problem; the analysis is based on a loose upper bound on the MI in the relevant operating regime; and performance is evaluated solely in terms of MI overestimating the AIR in practical systems with BMD.

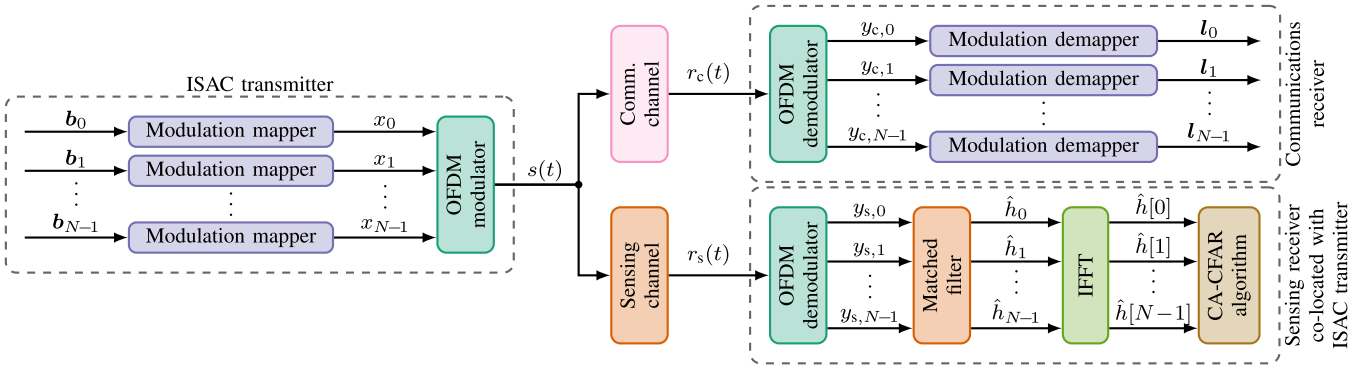


Fig. 2. Block diagram of the considered monostatic OFDM-ISAC system. The unified ISAC signal is employed for achieving both S&C tasks.

- **Autoencoder-based constellation optimization and systematic analysis:** We use a binary AE to optimize constellations shaped geometrically, probabilistically, and jointly for both ideal receivers and practical systems with BMD, while satisfying detection and false alarm constraints. We show that constellations optimized for ideal receivers closely approach the theoretical upper bound and quantify the performance loss when transitioning to practical BMD. Furthermore, we demonstrate that joint constellation shaping outperforms geometric and probabilistic constellation shaping, enables flexible control over the S&C trade-off, and significantly outperforms conventional formats, e.g., 64-QAM.
- **Low-complexity implementation of constellation shaping for practical ISAC systems:** We review PAS and demonstrate its limitations in the ISAC context. Therefore, we generalize PAS to improve sensing performance and propose low-complexity log-likelihood ratio (LLR) computation with negligible AIR loss. Simulation results show that a combination of conventional and generalized PAS achieves S&C performance close to that joint shaping across a wide range of kurtosis constraints while remaining compatible with practical implementation requirements.

Notations: Bold lowercase letters (e.g., $\boldsymbol{x}$) denote vectors, bold uppercase letters represent matrices (e.g., $\boldsymbol{J}$), and scalars are written in normal font (e.g., N). We do not explicitly distinguish between time and frequency domains; instead, the index implicitly indicates the respective domain: t denotes continuous time, k denotes discrete time/delay, and n refers to the sub-carrier index in the frequency domain. Sans-serif symbols (e.g., x) indicate random variables, and calligraphic letters (e.g., $\mathcal{X}$) represent sets. The operator $(\cdot)^\star$ denotes complex conjugation. The Dirac delta function is written as $\delta(\cdot)$. The expectation with respect to the random variable x is denoted by $\mathbb{E}_{\mathsf{x}}\{\cdot\}$. The functions $\mathbb{h}(\cdot)$ and $\mathbb{H}(\cdot)$ denote the differential and discrete entropy, respectively. The real and imaginary parts of a complex number are denoted by $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$, respectively. The operator $\mathrm{vec}(\cdot)$ denotes row-wise vectorization, and $\otimes$ denotes the outer product. The softmax function is written as $\mathrm{Softmax}(\cdot)$.

II. System Model

In this work, we consider a monostatic OFDM-ISAC system as illustrated in Fig. 2. To reduce computational complexity and simplify the analysis, all N sub-carriers use the same constellation, and the sensing target delays are modeled as integer multiples of the sampling period and are shorter than the duration of the cyclic prefix (CP). Furthermore, we assume that the targets are static. We focus on the transmission of single OFDM symbols, and neglect Doppler frequency, i.e., velocity, estimation.

A. ISAC Transmitter

For each sub-carrier $n \in \{0, 1, \ldots, N-1\}$, a constellation mapper encodes M bits $\boldsymbol{b}_n \in \{0,1\}^M$, onto a modulation symbol x_n taken from the modulation alphabet $\mathcal{X}$, i.e., $x_n \in \mathcal{X} \subset \mathbb{C}$, where $|\mathcal{X}| = \widetilde{M} = 2^M$. Since the constellation symbols appear randomly selected and independent of each other, the transmit symbols x_n can be considered as i.i.d. random variables (RVs) $\mathsf{x}_n \sim P_{\mathsf{x}}(x_n)$.

The OFDM modulator transforms the N frequency domain symbols x_n, into the time domain using the orthonormal inverse fast Fourier transform (IFFT). Next, a CP is added before the signal $\mathsf{s}(t)$ is transmitted over an equivalent baseband channel.

B. Communications Channel and Receiver

For communications, the influence of a potential multi-path channel is assumed to be mitigated through equalization. The communication receiver removes the CP from the received communication signal $\mathsf{r}_{\mathrm{c}}(t)$ and converts it into the frequency domain using an orthonormal FFT. This simplifies the OFDM system into a set of N parallel AWGN channels, one for each sub-carrier. Therefore, the received communications symbol of the nth sub-carrier is given by

$$\mathsf{y}_{\mathrm{c},n} = \mathsf{x}_n + \mathsf{w}_{\mathrm{c},n}, \tag{1}$$

where $\mathsf{w}_{\mathrm{c},n} \sim \mathcal{CN}(0, \sigma^2_{\mathrm{c},n})$ is AWGN with variance $\sigma^2_{\mathrm{c},n}$.

At the receiver, we consider an optimal maximum a-posteriori decoder using symbol-metric decoding (SMD), as well as BMD, which is typically used in practical systems with binary error correcting codes [18]. In BMD, a demapper computes a vector of LLRs $\boldsymbol{l}_n$, where the individual LLRs are

evaluated separately for each bit position $m = \{1, \dots, M\}$ [24]

$$l_{n,m}(y_{\mathrm{c},n}) = \log \frac{\sum_{x_n \in \mathcal{X}_m^{(0)}} f_{\mathsf{y}_{\mathrm{c},n}|\mathsf{x}}(y_{\mathrm{c},n}|x_n) P_{\mathsf{x}}(x_n)}{\sum_{x_n \in \mathcal{X}_m^{(1)}} f_{\mathsf{y}_{\mathrm{c},n}|\mathsf{x}}(y_{\mathrm{c},n}|x_n) P_{\mathsf{x}}(x_n)}, \quad (2)$$

where $\mathcal{X}_m^{(b)}$ represents the set of all constellation symbols labeled with bit $b \in \{0,1\}$ at bit position m, and $f_{\mathsf{y}_{\mathrm{c},n}|\mathsf{x}}$ denotes the communications channel transition probability density function (PDF).

C. Achievable Information Rates

The AIR depends on the decoding metric, and we refer the reader to [18], [19], and [25] and the references therein for a comprehensive overview.

Under SMD, where decoding is performed directly on the constellation symbols, the AIR is given by the MI [26]

$$\mathbb{I}_{\mathrm{MI}}(\mathsf{x}_n; \mathsf{y}_{\mathrm{c},n}) = \mathbb{E}_{\mathsf{x}_n, \mathsf{y}_{\mathrm{c},n}} \left\{ \log_2 \left(\frac{f_{\mathsf{y}_{\mathrm{c},n}|\mathsf{x}}(\mathsf{y}_{\mathrm{c},n}|\mathsf{x}_n)}{f_{\mathsf{y}_{\mathrm{c},n}}(\mathsf{y}_{\mathrm{c},n})} \right) \right\}, \quad (3)$$

which defines the highest achievable rate with an arbitrarily complex receiver, where $f_{\mathsf{y}_{\mathrm{c},n}}(\mathsf{y}_{\mathrm{c},n})$ denotes the PDF of the received communication symbols $\mathsf{y}_{\mathrm{c},n}$ on sub-carrier n.

In contrast, practical communication systems employ BMD, where the AIR is lower bounded by the generalized mutual information (GMI) [18], [25]

$$\mathbb{I}_{\mathrm{GMI}}(\mathsf{b}_n; \mathsf{y}_{\mathrm{c},n}) = \left[\mathbb{H}(\mathsf{b}_n) - \sum_{m=1}^{M} \mathbb{H}(\mathsf{b}_{n,m}|\mathsf{y}_{\mathrm{c},n}) \right]^+, \quad (4)$$

where $[\,\cdot\,]^+$ is $\max(\cdot, 0)$. The GMI is upper bounded by the MI $\mathbb{I}(\mathsf{x}_n; \mathsf{y}_{\mathrm{c},n})$ [24], [25].

Consequently, we consider the MI to explore the fundamental DRT between S&C in idealized systems using SMD, and use the GMI to assess the S&C trade-off in practical communication systems employing BMD.

D. Sensing Channel and Receiver

We consider a sensing scenario with J targets, where one target is designated as the TOI, e.g., a drone, and the remaining $J-1$ targets act as strong interferers, e.g., static buildings or residual self-interference from the finite isolation between the transmitter and receiver paths, which can be modeled as a static target at zero delay with an artificial radar cross section (RCS). The RCS of the TOI follows a Swerling-1 model with PDF

$$f_{\sigma_{\mathrm{RCS,TOI}}}(\sigma_{\mathrm{RCS,TOI}}) = \begin{cases} \frac{1}{\bar{\sigma}_{\mathrm{RCS,TOI}}} \exp\left(-\frac{\sigma_{\mathrm{RCS,TOI}}}{\bar{\sigma}_{\mathrm{RCS,TOI}}}\right), & \sigma_{\mathrm{RCS,TOI}} \geqslant 0, \\ 0, & \sigma_{\mathrm{RCS,TOI}} < 0, \end{cases} \quad (5)$$

i.e., the RCS fluctuates independently from OFDM symbol to OFDM symbol according to an exponential distribution with mean $\bar{\sigma}_{\mathrm{RCS,TOI}}$. The interfering targets follow a non-fluctuating Swerling-0 model and therefore have constant RCS values $\sigma_{\mathrm{RCS},j}$, which are larger than the mean RCS of the TOI, i.e., $\sigma_{\mathrm{RCS},j} \gg \bar{\sigma}_{\mathrm{RCS}}$ for all $j \neq$ TOI. This detection setting is widely regarded as particularly challenging, since the sidelobes of the strong interfering targets can obscure the weak TOI. For each OFDM symbol, we draw one realization of the RCS of the TOI and treat it as fixed. Consequently, within a single symbol, the sensing channel can be modeled as the sum of J static point targets with arbitrary, but fixed amplitudes

$$\mathsf{r}_{\mathrm{s}}(t) = h(t) * \mathsf{s}(t) + \mathsf{w}_{\mathrm{s}}(t) = \sum_{j=1}^{J} a_j \mathsf{s}(t - \tau_j) + \mathsf{w}_{\mathrm{s}}(t), \quad (6)$$

where $\mathsf{w}_{\mathrm{s}}(t) \sim \mathcal{CN}(0, \sigma_{\mathrm{s}}^2)$ denotes AWGN with variance σ_{s}^2. The amplitudes a_j and delays τ_j follow from the radar equation

$$a_j = \sqrt{\frac{\sigma_{\mathrm{RCS},j} c_0^2 P_{\mathrm{Tx}} G_{\mathrm{Tx}} G_{\mathrm{Rx}}}{(4\pi)^3 R_j^4 f_{\mathrm{c}}^2}} \mathrm{e}^{\mathrm{j}\varphi_j}, \qquad \tau_j = \frac{2R_j}{c_0}, \quad (7)$$

where P_{Tx}, G_{Tx}, G_{Rx}, f_{c}, and c_0 denote transmit power, antenna gains, carrier frequency, and speed of light. The phases φ_j are independent across targets and, for each target, drawn i.i.d. from a uniform distribution over $[0, 2\pi)$. The sensing receiver samples the baseband signal $\mathsf{r}_{\mathrm{s}}(t)$, removes the CP, and transforms it into the frequency domain using the orthonormal FFT. The frequency domain received symbols are

$$\mathsf{y}_{\mathrm{s},n} = \mathsf{x}_n h_n + \mathsf{w}_{\mathrm{s},n} = \mathsf{x}_n \sum_{j=1}^{J} a_j \mathrm{e}^{-\mathrm{j}2\pi \frac{n}{N} \tau_j} + \mathsf{w}_{\mathrm{s},n}, \quad (8)$$

where h_n denotes the channel frequency response and $\mathsf{w}_{\mathrm{s},n}$ is the FFT of the sampled AWGN $\mathsf{w}_{\mathrm{s}}(t)$, and follows $\mathcal{CN}\left(0, \sigma_{\mathrm{s}}^2\right)$.

The sensing receiver applies a sensing matched filter (MF)

$$\hat{\mathsf{h}}_n = \mathsf{y}_{\mathrm{s},n} \mathsf{x}_n^\star = (h_n \mathsf{x}_n + \mathsf{w}_{\mathrm{s},n}) \mathsf{x}_n^\star = h_n |\mathsf{x}_n|^2 + \mathsf{w}_{\mathrm{s},n} \mathsf{x}_n^\star, \quad (9)$$

which yields an unbiased estimate of the sensing channel

$$\mathbb{E}_{\mathsf{x}, \mathsf{w}_{\mathrm{s},n}}\{\hat{\mathsf{h}}_n\} = \mathbb{E}_{\mathsf{x}}\{h_n |\mathsf{x}_n|^2\} + \mathbb{E}_{\mathsf{x}, \mathsf{w}_{\mathrm{s},n}}\{\mathsf{w}_{\mathrm{s},n} \mathsf{x}_n^\star\} = h_n \quad (10)$$

for unit power constellations $\mathbb{E}_{\mathsf{x}}\{|\mathsf{x}_n|^2\} = 1$.

The delay domain channel estimate $\hat{\mathsf{h}}[k]$ is obtained by applying the orthonormal IFFT to the frequency domain channel estimate $\hat{\mathsf{h}}_n$

$$\hat{\mathsf{h}}[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \hat{\mathsf{h}}_n \mathrm{e}^{\mathrm{j}2\pi \frac{n}{N} k}, \qquad k = \{0, \dots N-1\}. \quad (11)$$

Finally, the cell-averaging (CA)-constant false alarm rate (CFAR) algorithm, which maximizes the detection probability P_{D} given a maximum false alarm rate P_{FA}, determines whether a target is present at a delay k [27].

E. The Constellation-Dependent Detection Probability

We derive the constellation-dependent detection probability of the TOI by analyzing each signal processing block of the sensing receiver. For the CA-CFAR algorithm with a window length $N_{\mathrm{win}} \to \infty$, the detection probability is

$$P_{\mathrm{D}} = P_{\mathrm{FA}}^{\frac{1}{1+\bar{\gamma}_{\mathrm{TOI}}}}, \quad (12)$$

assuming Gaussian distributed noise and interference [27]. The detection probability P_{D} depends only on the false alarm rate P_{FA} and the average signal-to-interference-and-noise ratio

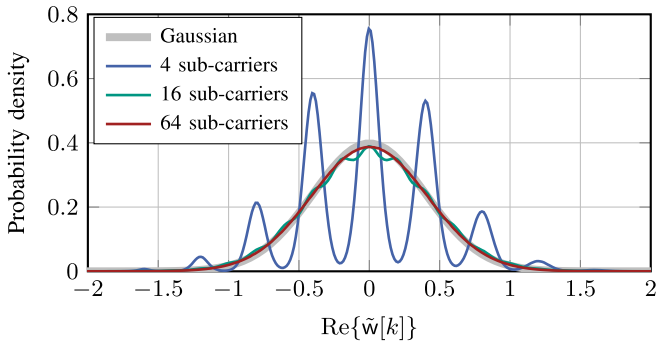


Fig. 3. PDF of the noise at the input of the target detector, i.e., at the output of the IFFT assuming a 16-QAM and a sensing SNR of 20 dB in a single target scenario for various numbers of sub-carriers.

(SINR) $\bar{\gamma}_{\mathrm{TOI}}$ at the input of the detector. To prevent clutter-induced bursts of false alarms and to enable robust and predictable detection performance, the false alarm rate P_{FA} is usually fixed during system design [27]. Consequently, increasing the detection probability P_{D} requires increasing the average SINR $\bar{\gamma}_{\mathrm{TOI}}$ at the input of the CA-CFAR, i.e., at the output of the IFFT. Note that the average is taken with respect to the RCS fluctuations of the TOI, as modeled by the Swerling-1 model. Following standard practice [27], we first derive the SINR for a fixed sensing-channel realization (6) and subsequently average over the sensing-channel fluctuations.

When random constellation symbols are transmitted, the noise at the output of the MF, i.e., at the input of the IFFT, may no longer be Gaussian as can be observed from (9). To show that the noise is still approximately Gaussian at the output of the IFFT, we decompose the channel transfer function estimate $\hat{\mathsf{h}}_n$ into a deterministic part h_n and a random part $\widetilde{\mathsf{w}}_n = h_n(|\mathsf{x}_n|^2 - 1) + \mathsf{w}_{\mathrm{s},n}\mathsf{x}_n^\star$, which accounts for AWGN and the randomness of the modulation, i.e., $\hat{\mathsf{h}}_n = h_n + \widetilde{\mathsf{w}}_n$.

For the random part $\widetilde{\mathsf{w}}_n$, the IFFT acts as a summation of N independent and scaled RVs $\widetilde{\mathsf{w}}_n$. According to the central limit theorem, the random part of the channel estimate $\widetilde{\mathsf{w}}[k] = \hat{\mathsf{h}}[k] - h[k]$, which is the IFFT of $\widetilde{\mathsf{w}}_n$ approximates a Gaussian if the number of sub-carriers N is sufficiently large. Fig. 3 shows the PDF of $\mathrm{Re}\{\widetilde{\mathsf{w}}[k]\}$ for an increasing number of sub-carriers, demonstrating that the Gaussian assumption is already reasonable for 64 sub-carriers. This justifies modeling the detection probability using (12).

Since the frequency-domain channel estimation errors are zero-mean and uncorrelated on different sub-carriers, and using the linearity of the FFT, the variance

$$\sigma^2_{\hat{\mathsf{h}}[k]} = \left(\frac{1}{\sqrt{N}}\right)^2 \sum_{n=0}^{N-1} \sigma^2_{\hat{\mathsf{h}}_n} \tag{13}$$

of the delay domain channel estimate equals the mean variance of the frequency domain channel estimates

$$\begin{aligned}\sigma^2_{\hat{\mathsf{h}}_n} &= \mathbb{E}_{\mathsf{x},\mathsf{w}_{\mathrm{s},n}}\{|\hat{\mathsf{h}}_n|^2\} - |\mathbb{E}_{\mathsf{x},\mathsf{w}_{\mathrm{s},n}}\{\hat{\mathsf{h}}_n\}|^2 \\ &= \mathbb{E}_{\mathsf{x},\mathsf{w}_{\mathrm{s},n}}\Big\{\Big(h_n|\mathsf{x}_n|^2+\mathsf{w}_{\mathrm{s},n}\mathsf{x}_n^\star\Big)\Big(h_n|\mathsf{x}_n|^2+\mathsf{w}_{\mathrm{s},n}\mathsf{x}_n^\star\Big)^\star\Big\} - |h_n|^2 \\ &= \mathbb{E}_{\mathsf{x},\mathsf{w}_{\mathrm{s},n}}\Big\{|h_n|^2\,|\mathsf{x}_n|^4 + h_n\,|\mathsf{x}_n|^2\,\mathsf{w}_{\mathrm{s},n}^\star\mathsf{x}_n \\ &\qquad + \mathsf{w}_{\mathrm{s},n}\mathsf{x}_n^\star h_n^\star\,|\mathsf{x}_n|^2 + |\mathsf{w}_{\mathrm{s},n}\mathsf{x}_n|^2\Big\} - |h_n|^2 \\ &= |h_n|^2(\mathbb{E}_{\mathsf{x}}\{|\mathsf{x}_n|^4\} - 1) + \sigma_{\mathrm{s}}^2 \\ &= |h_n|^2\,(\kappa_n - 1) + \sigma_{\mathrm{s}}^2, \end{aligned} \tag{14}$$

where κ_n is the kurtosis of the constellation

$$\kappa_n = \frac{\mathbb{E}_{\mathsf{x}}\Big\{|\mathsf{x}_n - \mathbb{E}_{\mathsf{x}}\{\mathsf{x}_n\}|^4\Big\}}{\Big(\mathbb{E}_{\mathsf{x}}\Big\{|\mathsf{x}_n - \mathbb{E}_{\mathsf{x}}\{\mathsf{x}_n\}|^2\Big\}\Big)^2} \overset{\substack{\mathbb{E}_{\mathsf{x}}\{\mathsf{x}_n\}=0 \\ \mathbb{E}_{\mathsf{x}}\{|\mathsf{x}_n|^2\}=1}}{=} \mathbb{E}_{\mathsf{x}}\{|\mathsf{x}_n|^4\}, \tag{15}$$

which is equivalent to the 4th-order moment for unit power zero mean constellations. Moreover, since all subcarriers employ the same constellation, it follows that $\kappa_n = \kappa$. Inserting (14) into (13) yields the noise-and-interference power

$$\sigma^2_{\hat{\mathsf{h}}[k]} = \sum_{j=1}^{J} |a_j|^2(\kappa - 1) + \sigma_{\mathrm{s}}^2, \tag{16}$$

where we exploit that all targets are independent with uniformly distributed phases

$$\frac{1}{N}\sum_{n=1}^{N-1} |h_n|^2 = \sum_{j=1}^{J} |a_j|^2. \tag{17}$$

For the deterministic part of the channel estimate, the IFFT leads to an integration gain, increasing the power of the targets and consequently the SINR by a factor of N, as

$$\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\sum_{j=1}^{J} a_j \mathrm{e}^{-\mathrm{j}2\pi\frac{n\tau_j}{N}}\mathrm{e}^{\mathrm{j}2\pi\frac{nk}{N}} = \begin{cases}\sqrt{N}a_j, & k=\tau_j, \\ 0, & k\neq\tau_j.\end{cases} \tag{18}$$

The fraction of the power in the deterministic (18) and random part (14) yields the SINR

$$\gamma_{\mathrm{TOI}} = \frac{N\cdot|a_{\mathrm{TOI}}|^2}{\sum_{j=1}^{J}|a_j|^2(\kappa-1)+\sigma_{\mathrm{s}}^2} \tag{19}$$

for a fixed sensing channel realization at the input of the CFAR algorithm.

We observe that the SINR depends only on the kurtosis κ of the constellation for a given sensing channel realization. Using the identity $\mathrm{Var}\{|\mathsf{x}|^2\} = \mathbb{E}\{|\mathsf{x}|^4\} - |\mathbb{E}\{|\mathsf{x}|^2\}|^2$ and the unit-power constraint $\mathbb{E}\{|\mathsf{x}|^2\} = 1$, the kurtosis can be written as

$$\kappa = \mathrm{Var}\{|\mathsf{x}|^2\} + 1, \tag{20}$$

showing that the kurtosis directly measures the variance of the power of the constellation. Table I lists the kurtosis of common modulation formats. Unit-modulus constellations such as phase shift keying (PSK) achieve the minimum kurtosis $\kappa = 1$, preventing interference since the interference term in (16) vanishes and only the AWGN noise power σ_{s}^2 remains.

TABLE I
KURTOSIS VALUES OF MODULATION FORMATS

Constellation	$\widetilde{M}$-PSK	16-QAM	64-QAM	256-QAM	$\mathcal{CN}(0,1)$
Kurtosis κ	1	1.320	1.381	1.395	2

In contrast, conventional modulation formats such as $\widetilde{M}$-QAM and a circular complex Gaussian PDF, which maximizes the MI for an AWGN channel under an average transmit power constraint, introduce interference. For conventional $\widetilde{M}$-QAM, the kurtosis approaches $\kappa \approx 1.4$ as $M \to \infty$ (see Appendix A), while a circular complex Gaussian PDF has $\kappa = 2$.

We observe that the interference introduced by non-constant-modulus constellations, i.e., $\kappa > 1$, scales with the sum of the powers of all targets. Since the TOI is assumed to be much weaker than the interfering targets, the interference power is dominated by the non-fluctuating interferers, such that the resulting noise-and-interference term is effectively independent of the RCS fluctuations of the TOI. In contrast, the received target power itself fluctuates according to the Swerling-1 model. Inserting (7) into (19) and averaging with respect to these fluctuations yields the average SINR of the TOI

$$\bar{\gamma}_{\mathrm{TOI}} = \frac{N \cdot |\bar{a}_{\mathrm{TOI}}|^2}{\sum_{j=1}^{J} |a_j|^2(\kappa - 1) + \sigma_{\mathrm{s}}^2}, \tag{21}$$

where $\bar{a}_{\mathrm{TOI}}$ denotes the average amplitude corresponding to the mean RCS $\bar{\sigma}_{\mathrm{RCS,TOI}}$. Substituting (21) into (12) gives the constellation-dependent detection probability. Consequently, the RCS fluctuations of the TOI average out, and the detection probability depends only on the mean amplitude of the TOI, the amplitude of the interfering targets, and the kurtosis κ of the constellation.

F. Optimization Problem

In this paper, we aim to find constellations that maximize the communication performance, i.e., the (G)MI $\mathbb{I}_{\mathrm{(G)MI}}$ of the overall ISAC system, subject to a minimum detection probability constraint α_{D}, i.e., $P_{\mathrm{D}} \geqslant \alpha_{\mathrm{D}}$. Since all sub-carriers employ the same constellation, the total (G)MI is maximized if the (G)MI per sub-carrier is maximized. To ease optimization, we assume that all sub-carriers observe the same noise variance σ_{c}^2, i.e., $\sigma_{\mathrm{c},n}^2 = \sigma_{\mathrm{c}}^2$, $\forall n$ and we omit the sub-carrier index n in the following.[2]

Furthermore, we reformulate the detection-probability constraint $P_{\mathrm{D}} \geqslant \alpha_{\mathrm{D}}$ as a kurtosis constraint $\kappa \leqslant \tilde{\kappa}$. This follows from the fact that the detection probability increases monotonically with the average SINR (12), and the average SINR decreases monotonically with increasing kurtosis (21). Consequently, the detection probability P_{D} is a monotonically decreasing function of the kurtosis κ of the constellation, which directly justifies the reformulation of the detection-probability constraint as $\kappa \leqslant \tilde{\kappa}$. Although the average SINR (21), and thus the detection probability, depend on the unknown target amplitudes $|a_j|^2$, these amplitudes are not required during the constellation optimization. We exploit the monotonic relationship between kurtosis and detection probability, which holds for any given realization of the target amplitudes. In other words, decreasing the kurtosis always improves the sensing performance for any possible set of target amplitudes, so that no prior knowledge of the unknown target amplitudes is required during the optimization.[3] This leads to the following optimization problem

$$\max_{\mathcal{X}, P_{\mathsf{x}}(x)} \ \mathbb{I}_{\mathrm{(G)MI}} \tag{22}$$

$$\text{s.t.} \ \sum\nolimits_x P_{\mathsf{x}}(x) = 1, \quad P_{\mathsf{x}}(x) \geqslant 0, \quad \forall x \in \mathcal{X}, \tag{C0}$$

$$\mathbb{E}_{\mathsf{x}}\{|\mathsf{x}|^2\} = 1, \quad \mathbb{E}_{\mathsf{x}}\{\mathsf{x}\} = 0, \tag{C1}$$

$$\kappa \leqslant \tilde{\kappa}. \tag{C2}$$

Here, the constraints (C0) enforce that $P_{\mathsf{x}}(x)$ satisfies the properties of a probability mass function, while the constraints (C1) ensure that the constellation has unit power $E_{\mathrm{s}} = \mathbb{E}_{\mathsf{x}}\{|\mathsf{x}|^2\} = 1$ and zero mean.

III. A Lower and Upper Bound for the Maximum Mutual Information Under a Kurtosis Constraint

In this section, we derive a lower and upper bound on the maximum MI $\mathbb{I}(\mathsf{x}; \mathsf{y}_{\mathrm{c}})$ under the constraints from (22). These bounds serve as performance benchmarks for our optimized constellations. To facilitate our study, we model the communication channel input as a continuous complex-valued RV $\mathsf{x} \in \mathbb{C}$. A lower and upper bound on the MI

$$\mathbb{I}(\mathsf{x}; \mathsf{y}_{\mathrm{c}}) = \mathbb{h}(\mathsf{y}_{\mathrm{c}}) - \mathbb{h}(\mathsf{y}_{\mathrm{c}}|\mathsf{x}) = \mathbb{h}(\mathsf{y}_{\mathrm{c}}) - \mathbb{h}(\mathsf{w}_{\mathrm{c}}) \tag{23}$$

can be obtained by bounding the entropy of the received signal $\mathbb{h}(\mathsf{y}_{\mathrm{c}})$ as the entropy of the complex-valued AWGN $\mathbb{h}(\mathsf{w}_{\mathrm{c}}) = \log_2\left(\pi \mathrm{e} \sigma_{\mathrm{c}}^2\right)$ is constant [26].

For the lower bound, we use the entropy power inequality (EPI) [26, Chap. 17.8] expressed in bits

$$\mathbb{h}(\mathsf{y}_{\mathrm{c}}) \geqslant \log_2\left(2^{\mathbb{h}(\mathsf{x})} + 2^{\mathbb{h}(\mathsf{w}_{\mathrm{c}})}\right), \tag{24}$$

which is tight when x is Gaussian distributed.

Since we are interested in a lower and upper bound on the maximum MI, we need to maximize the entropy of the transmit and received signal, respectively

$$\underbrace{\mathbb{h}(\mathsf{y}_{\mathrm{max,c}}) - \mathbb{h}(\mathsf{w}_{\mathrm{c}})}_{\text{Upper bound}} \geqslant \max_{f_{\mathsf{X}}(x):\,(\mathrm{C0}),(\mathrm{C1}),(\mathrm{C2})} \mathbb{I}(\mathsf{x}; \mathsf{y}_{\mathrm{c}})$$

$$\geqslant \underbrace{\log_2\left(2^{\mathbb{h}(\mathsf{x}_{\max})} + 2^{\mathbb{h}(\mathsf{w}_{\mathrm{c}})}\right) - \mathbb{h}(\mathsf{w}_{\mathrm{c}})}_{\text{Lower Bound}}. \tag{25}$$

Here $\mathsf{x}_{\max}$ and $\mathsf{y}_{\mathrm{max,c}}$ denote the RVs associated with the entropy-maximizing PDFs.

[2] Frequency-selective fading results in a sub-carrier-dependent SNR, which substantially complicates the optimization and obscures the fundamental S&C trade-off that we aim to characterize. If the transmitter has no channel state information (CSI), the shaped constellation would need to generalize over a broad range of SNR values. Since the GMI-optimal constellation depends on the operating SNR, this would introduce not only the S&C trade-off studied in this work, but also an additional communications-only trade-off between low and high communication SNR operation, thereby preventing a clean characterization of S&C trade-off. Conversely, if transmitter-side CSI is available, adaptive modulation and power allocation, e.g., waterfilling, should be applied, which in turn affects the sensing performance and requires a joint optimization of constellation and power distribution. However, power allocation has only been explored recently [23] and constitutes a research topic of its own. Including power allocation in this work would significantly complicate the optimization and obscure the shaping-induced trade-off, and is therefore beyond the focus of this paper. Consequently, we assume a constant noise power across the sub-carriers. This corresponds to an AWGN communications channel and provides a clean first step for analyzing and characterizing the shaping-induced S&C trade-off, which is the focus of this work. Note that including the sub-carrier dependent SNR into the optimization is an important direction for future work.

[3] The target amplitudes $|a_j|^2$ are required only to evaluate the resulting detection probability P_{D}, but not to improve it by reducing the kurtosis κ.

TABLE II
MOMENT CONSTRAINTS FOR ENTROPY MAXIMIZATION

Bound	Domain z	C_0	C_1	C_2
Lower	Input x	1	E_s	$\tilde{\kappa}$
Upper	Output y_c	1	$E_\mathrm{s}+\sigma_\mathrm{c}^2$	$\tilde{\kappa}+4E_\mathrm{s}\sigma_\mathrm{c}^2+2\sigma_\mathrm{c}^4$

In both cases, the entropy maximization problem has the same structure[4]

$$\max_{f_\mathsf{z}(z)} \ \mathbb{h}(\mathsf{z}) \tag{26}$$
$$\text{s.t.} \ \ \mathbb{E}_\mathsf{z}\{|\mathsf{z}|^0\} = C_0, \quad f_\mathsf{z}(z) \geqslant 0, \forall z \in \mathbb{C}, \tag{C_0}$$
$$\mathbb{E}_\mathsf{z}\{|\mathsf{z}|^2\} = C_1, \tag{C_1}$$
$$\mathbb{E}_\mathsf{z}\{|\mathsf{z}|^4\} = C_2, \tag{C_2}$$

where z denotes either the transmit signal x (lower bound) or the received communication signal y_c (upper bound). The values of the corresponding moment constraints are summarized in Table II. For the upper bound, the transmit-side constraints on x are mapped into receive-side constraints on y_c and follow directly from substituting $\mathsf{y}_\mathrm{c} = \mathsf{x} + \mathsf{w}_\mathrm{c}$ into the respective moment definitions.

Remark: The optimization problem for the upper bound does not include any constraint ensuring that the resulting output PDF is realizable by the system model. Specifically, there may not exist a transmit PDF f_x such that passing it through the AWGN channel yields the entropy-maximizing received PDF $f_{\mathsf{y}_\mathrm{max,c}}$. Consequently, the corresponding MI represents a theoretical upper bound rather than an AIR. Nevertheless, the upper bound gives insights into the gap between the achieved MI of the optimized constellations and the theoretically maximum MI under the given constraints.

According to the maximum entropy principle [26, Ch. 12.1], the PDF that maximizes the entropy under moment constraints takes the exponential form. For constraints on the zeroth, second and fourth moments, the optimal PDF has the form

$$f_{\mathsf{z}_\mathrm{max}}(z) = \mathrm{e}^{\gamma_0+\gamma_2|z|^2+\gamma_4|z|^4}, \tag{27}$$

where the real-valued parameters γ_0, γ_2, and γ_4 are chosen to satisfy the constraint equations

$$C_q = \int_\mathbb{C} |z|^{2q}\mathrm{e}^{\gamma_0+\gamma_2|z|^2+\gamma_4|z|^4}\,\mathrm{d}z, \qquad q = 0, 1, 2. \tag{28}$$

We note that the optimal PDF $f_{\mathsf{z}_\mathrm{max}}(z)$ is circularly symmetric with respect to the origin and, as a consequence, has zero mean, thus fulfilling constraint (C1). The resulting system of equations can be solved numerically. However, this system involves integrals over the complex plane and comprises three coupled equations in the three unknowns γ_0, γ_2, and γ_4, which leads to a comparatively high computational complexity. To reduce computational complexity, we reformulate the system of equations (28) in Appendix B such that γ_2 can be obtained by numerically solving a single nonlinear equation

$$C_0 = \left(\frac{C_1(\gamma_2 C_1+1)}{C_2} - \gamma_2\right)\sqrt{\frac{\pi C_2}{2(\gamma_2 C_1+1)}} \cdot \exp\left(\frac{\gamma_2^2 C_2}{2(\gamma_2 C_1+1)}\right)\operatorname{erfc}\left(-\gamma_2\sqrt{\frac{C_2}{2(\gamma_2 C_1+1)}}\right). \tag{29}$$

Then, γ_0 and γ_4 follow immediately in closed form

$$\gamma_0 = \ln\left(\frac{1}{\pi}\left[\frac{C_1(\gamma_2 C_1+1)}{C_2} - \gamma_2\right]\right), \tag{30}$$
$$\gamma_4 = -|\gamma_4| = \frac{-1}{2C_2}(\gamma_2 C_1 + 1). \tag{31}$$

Consequently, determining γ_2 is the key step and can be achieved by solving a single nonlinear equation (29) numerically using standard one-dimensional root-finding methods.

Note that for $\kappa = 2$, the optimal distribution reduces to a circular complex Gaussian, which corresponds to $\gamma_2 = -1/C_1$ and $\gamma_4 = 0$. Starting from this point, we can continuously and smoothly transition from the Gaussian case ($\kappa = 2$) to a unit-modulus distribution ($\kappa = 1$) by gradually shifting probability mass toward the unit circle (see Fig. 14), achieved by increasing γ_2 and decreasing γ_4. This continuous transition ensures that every target value $\kappa \in (1, 2]$ is attainable along this path. Moreover, if there exists a solution for the entropy-maximizing distribution, it is unique [26]. Since (29) is a reformulation of the corresponding moment-matching condition (28), the solution for γ_2 is likewise unique for the considered kurtosis range $\kappa \in [1, 2]$, apart from potential numerical instabilities caused by vanishing or exploding terms. In such cases, the system of equations (28) can be solved instead. This does not pose a practical issue, since the bounds are computed offline and are used solely as a reference for evaluating the optimized constellations.

Once the parameters for the transmit and receive PDFs are known, the respective entropy $\mathbb{h}(\mathsf{z}_\mathrm{max})$ is

$$\mathbb{h}(\mathsf{z}_\mathrm{max}) = \frac{-\gamma_0 - C_1\cdot\gamma_2 - C_2\cdot\gamma_4}{\ln(2)}, \tag{32}$$

which follows immediately by inserting (27) into the definition of the entropy.

To obtain the lower and upper bounds, the following steps must be carried out separately for each bound. First, compute the constraints C_0, C_1, and C_2 in Table II. Second, solve (29) numerically to obtain γ_2. Third, compute γ_0 (30) and γ_4 (31). Fourth, determine the entropy (32). Finally, insert this entropy into (25) to obtain the respective bound. We provide source code to reproduce these bounds in [28].

Remark: Both bounds are tight for $\kappa = 2$, where they coincide with the Shannon capacity of an AWGN channel $\log_2(1+E_\mathrm{s}/\sigma_\mathrm{s}^2)$. For $\kappa < 2$, the bounds are generally loose: the lower bound is tight only when x is Gaussian, and the upper bound may not be attainable because the entropy-maximizing receive PDF $f_{\mathsf{y}_\mathrm{max,c}}$ is not realizable as discussed in the previous remark. Nevertheless, the bounds become asymptotically tight as the communication SNR increases, since the noise power σ_c^2 decreases and the constraints C_1 and C_2 in Tab. II converge. Moreover, our numerical results in Sec. VI and Fig. 13 show that the MI of the optimized constellations approaches the upper bound closely.

[4]Note that we omit the zero-mean constraint $\mathbb{E}_\mathsf{z}\{z\} = 0$ here and show later that the entropy-maximizing PDF satisfies it without explicit enforcement.

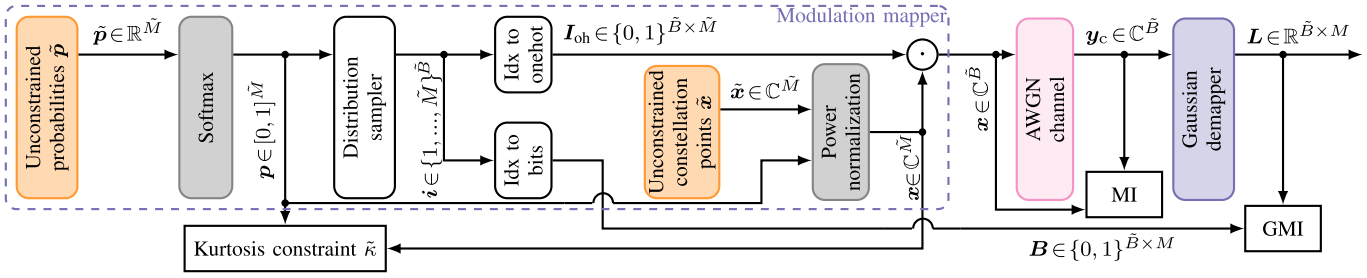


Fig. 4. Proposed AE framework to shape constellations for ISAC. The trainable parameters are marked in orange and their normalization in gray.

IV. Constellation Shaping Using Autoencoders

In this section, we propose to incorporate the sensing constraint (C2) into the bitwise AE framework [15], enabling the joint end-to-end optimization of both geometric and probabilistic shaping for S&C. Autoencoders have become a standard and state-of-the-art tool for constellation shaping in communication systems, see, e.g., [29], [30]. Although other machine-learning-based optimization techniques could in principle be used, we are not aware of any approach that systematically outperforms AE-based end-to-end optimization for constellation shaping. A further key advantage of the AE approach is that it naturally supports probabilistic, geometric, and joint shaping within a unified architecture, whereas many existing methods focus on only one of these aspects.

A classical AE consists of an encoder that learns an internal representation of the input and a decoder that attempts to reconstruct the input, typically both implemented as neural networks. When applied to communication systems, in particular to constellation shaping, the encoder takes the role of the transmitter by learning a representation of the binary input in the form of an optimized constellation, while the decoder takes the role of the receiver by recovering the transmitted bits. During training, encoder and decoder are optimized jointly in an end-to-end manner such that the AIR is maximized. For more details on AE-based constellation optimization, we refer the reader to [15], [16], and [31].

The block diagram of the AE is shown in Fig. 4. In our setup, the receiver is a demapper with a Gaussian noise assumption [24], and only the constellation of the transmitter is trainable. Depending on the shaping method (geometric, probabilistic, or joint), we optimize the constellation points, their probabilities, or both. For geometric shaping, only the unconstrained constellation points $\tilde{\boldsymbol{x}}$ are trainable. The unconstrained probabilities $\tilde{\boldsymbol{p}}$ are initialized uniformly and remain fixed, i.e., they are not trainable. Thus, the AE optimizes the position of the constellation points while the probability of occurrence remains unchanged, i.e., uniform. For probabilistic shaping, the constellation points are initialized, e.g., as conventional $\tilde{M}$-QAM, and remain fixed during training, while only the unconstrained probabilities $\tilde{\boldsymbol{p}}$ are trainable. This corresponds to optimizing the probability of occurrence of the constellation points while keeping their position fixed. For joint, i.e., both geometric and probabilistic, constellation shaping, the unconstrained constellation points $\tilde{\boldsymbol{x}}$ and the unconstrained probabilities $\tilde{\boldsymbol{p}}$ are trainable, which yields the most degrees of freedom. The trainable parameters, shown as orange blocks in Fig. 4, are implemented as linear layers without bias.

For a batch size $\tilde{B}$, the distribution sampler generates random indices $\boldsymbol{i} \in \{1,\dots,\tilde{M}\}^{\tilde{B}}$ according to the input distribution $P_{\mathsf{x}}(x) \triangleq \boldsymbol{p}$. To ease optimization, we use the Gumbel-softmax trick and optimize the unconstrained probabilities $\tilde{\boldsymbol{p}} \in \mathbb{R}^{\tilde{M}}$ which are passed through a softmax layer to ensure that $\boldsymbol{p}$ satisfies the probability distribution properties (C0) [15].

Then, these indices are mapped to bits $\boldsymbol{B}$ and one-hot vectors $\boldsymbol{I}_{\text{oh}}$. To select the transmit symbols, the one-hot vectors are multiplied with the normalized constellation points $\boldsymbol{x}$. Similar to the probabilities, we optimize the unconstrained constellation points $\tilde{\boldsymbol{x}}$, which are passed through a power normalization layer to obtain the unit power constellation points $\mathcal{X} \triangleq \boldsymbol{x}$.

Next, the selected constellation symbols are transmitted over an AWGN channel (1). For BMD, the LLRs (2) are computed for each bit using a classical Gaussian demapper [24]. The resulting GMI (4) is computed from the bits $\boldsymbol{B}$ and LLRs $\boldsymbol{L}$ [25]. For SMD, the MI is directly computed from the transmit symbols $\boldsymbol{x}$ and receive symbols $\boldsymbol{y}_{\text{c}}$.

We showed in Sec. II that the detection probability constraint can be reformulated as a kurtosis constraint (C2). Therefore, we propose a sensing loss term that depends on the kurtosis κ of the constellation to satisfy the detection probability constraint

$$L_{\text{sens}} = \begin{cases} 0, & \kappa \leqslant \tilde{\kappa}, \\ d(\kappa - \tilde{\kappa}), & \kappa > \tilde{\kappa}. \end{cases} \tag{33}$$

The penalty factor $d \in \mathbb{R}^{+}$ controls the strength of the penalty if the kurtosis threshold $\tilde{\kappa}$ is violated. To improve S&C performance simultaneously, the overall non-negative loss function combines both S&C performance

$$L_{S\&C} = \underbrace{\frac{M - \mathbb{I}_{\text{(G)MI}}}{M}}_{\text{Communications}} + L_{\text{sens}}. \tag{34}$$

Note that this loss function does not strictly enforce the sensing constraint. However, the communications loss term is normalized between $[0,1]$ and by selecting the penalty factor d sufficiently large, the sensing loss term dominates if the kurtosis constraint is violated, effectively enforcing the sensing constraint.

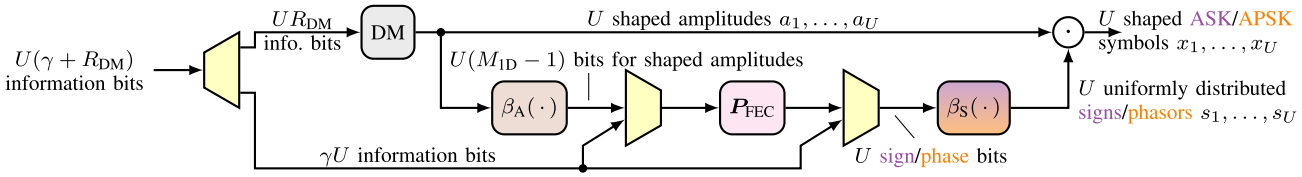


Fig. 5. Encoding structure of conventional PAS, based on [18], [19], and the proposed generalized PAS. In both schemes, a distribution matcher (DM) generates shaped amplitudes, while sign or phase bits are derived from a combination of uniformly distributed parity bits $\boldsymbol{P}_{\text{FEC}}$ and information bits. In conventional PAS, a single sign bit modulates the sign of real-valued amplitude shift keying (ASK) symbols, requiring two parallel encoders for the real and imaginary parts. In generalized PAS, multiple phase bits are mapped to discrete phases, directly generating complex-valued, circularly symmetric constellations. The generalization preserves the advantages of the original scheme while enhancing sensing performance.

V. Low-Complexity Implementation of Constellation Shaping for ISAC

In the previous sections, we focused on the theoretical limits and the optimization procedure for constellation shaping without addressing its practical implementation. A promising candidate to implement probabilistic constellation shaping is PAS, which integrates shaping and FEC in a structured, hardware-efficient manner, effectively mitigating error propagation. We begin this section by revisiting PAS, referring the reader to [18] and [19] for a detailed introduction. Our focus lies on how the PAS structure constrains constellation design, and how these constraints can be incorporated into our AE framework to optimize PAS-compatible constellations. Then, we show how PAS enables low-complexity LLR computation at the receiver.

A. Conventional Probabilistic Amplitude Shaping

1) Constellation Design and Encoding Scheme: Conventional PAS constructs a complex-valued constellation point by independently shaping and combining two real-valued $2^{M_{\text{1D}}}$-ASK components, where $M_{\text{1D}} = M/2$. Fig. 5 shows the encoding procedure for a single real-valued dimension corresponding to one of the two components of a complex-valued modulation mapper in Fig. 2.

PAS generates blocks of $U \in \mathbb{N}$ ASK symbols with target amplitude distribution $P_{\mathsf{a}}(a_u)$. To this end, the information bits are split into two parts. The first fraction, consisting of $U \cdot R_{\text{DM}}$ bits, is passed through a distribution matcher (DM), e.g., a constant composition DM [18]. The rate R_{DM} asymptotically approaches $\mathbb{H}(\mathsf{a})$ and we assume that $U \cdot R_{\text{DM}}$ is integer [18]. The DM maps the bits onto U shaped amplitudes $a_u \in \mathcal{A} = \Delta \cdot \{1, 3, \ldots, 2^{M_{\text{1D}}-1} - 1\}$, for $u \in \{0, 1, \ldots, U-1\}$, following the target amplitude distribution $P_{\mathsf{a}}(a_u)$, where Δ is a normalization factor. Each amplitude a_u is labeled using $(M_{\text{1D}}-1)$ bits via binary reflected Gray coding $\beta_{\text{A}} : \mathcal{A} \to \{0,1\}^{M_{\text{1D}}-1}$.

The remaining γU information bits are first appended to the $U(M_{\text{1D}}-1)$ bits representing the U shaped amplitudes, where $\gamma > 0$ controls how many bits bypass the DM. The combined bitstream is then encoded using a systematic FEC code with rate $(M_{\text{1D}} - 1 + \gamma)/M_{\text{1D}}$. Since the encoder is systematic, the shaped amplitude distribution is preserved at the output. The resulting parity bits, which are approximately uniformly distributed [18], are next combined with the γU information bits to form the U sign bits. These determine the signs $s_u \in \mathcal{S}_{\text{CPAS}} = \{-1, +1\}$ of the ASK symbols via the mapping $\beta_{\text{S}} : \{0,1\} \to \mathcal{S}_{\text{CPAS}}$.

Consequently, each one-dimensional ASK transmit symbols is constructed via amplitude-sign factorization $x_{\text{1D},u} = a_u \cdot s_u$ and is represented by one sign bit and $(M_{\text{1D}}-1)$ amplitude bits.

Since shaping is applied only to amplitudes and the sign bits remain uniform, this structure imposes a symmetry constraint on the one-dimensional ASK distribution, i.e.,

$$P_{\mathsf{x}_{\text{1D}}}(x_{\text{1D}}) = P_{\mathsf{x}_{\text{1D}}}(-x_{\text{1D}}), \quad \forall x_{\text{1D}} \in \mathcal{X}_{\text{1D}}, \tag{35}$$

where $\mathcal{X}_{\text{1D}} = \{s \cdot a : s \in \mathcal{S}_{\text{CPAS}}, a \in \mathcal{A}\}$. The same one-dimensional distribution is typically used for both dimensions of the complex constellation [18], [19].

2) Optimizing Conventional PAS-Compatible Constellations Using an AE Framework: In general probabilistic shaping, the AE optimizes $\widetilde{M}$ unconstrained probabilities $\tilde{\boldsymbol{p}}$ directly (see Sec. IV and Fig. 4). However, as discussed in the previous paragraph, conventional PAS imposes a specific structure on the probability of occurrence of the constellation points: (i) the probability of a complex-valued symbol factorizes into one-dimensional probabilities for the in-phase and quadrature components, and we enforce identical marginal distributions to reduce demapping complexity, (ii) the amplitude distribution is symmetric to enforce that the sign bits are uniformly distributed. Thus, to optimize PAS-compatible constellations using the AE framework, this structure must be enforced inside the AE. Instead of learning all $\widetilde{M}$ unconstrained probabilities $\tilde{\boldsymbol{p}}$, only the unconstrained parameters $\tilde{\boldsymbol{p}}_{\mathsf{a}}$, which define the one-dimensional target amplitude distribution $\text{Softmax}(\tilde{\boldsymbol{p}}_{\mathsf{a}}) = [P_{\mathsf{a}}(a_1), \ldots, P_{\mathsf{a}}(a_{2^{M_{\text{1D}}-1}})]$, are learned. Based on this optimized target amplitude distribution, the probability of occurrence of all constellation points is then generated. To impose symmetry, we construct an unconstrained, symmetric one-dimensional probability vector

$$\tilde{\boldsymbol{p}}_{\mathsf{x}_{\text{1D}}} = [\tilde{\boldsymbol{p}}_{\mathsf{a}}, \boldsymbol{J}_{\text{N}} \tilde{\boldsymbol{p}}_{\mathsf{a}}] \in \mathbb{R}^{\left(2^{M_{\text{1D}}}\right)}, \tag{36}$$

by concatenating $\tilde{\boldsymbol{p}}_{\mathsf{a}}$ with its flipped version, where $\boldsymbol{J}_{\text{N}}$ denotes the anti-identity matrix that reverses the order of the vector entries. The final two-dimensional probability mass vector $\boldsymbol{p}_{\text{CPAS}}$ is obtained by applying softmax to each dimension and taking the outer product, followed by row-wise vectorization

$$\boldsymbol{p}_{\text{CPAS}} = \text{vec}\left(\text{Softmax}(\tilde{\boldsymbol{p}}_{\mathsf{x}_{\text{1D}}}) \otimes \text{Softmax}(\tilde{\boldsymbol{p}}_{\mathsf{x}_{\text{1D}}})\right). \tag{37}$$

The probability mass vector $\boldsymbol{p}_{\text{CPAS}}$ satisfies the constraints of a conventional PAS constellation and replaces the probabilities

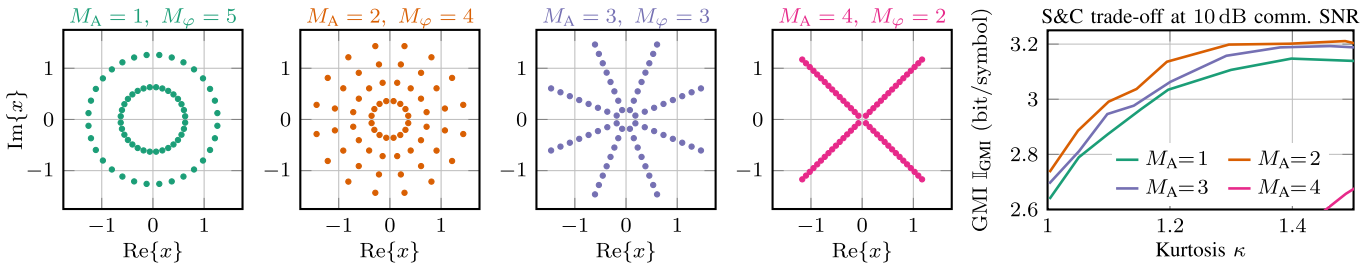


Fig. 6. Generalized PAS constellations for different combinations of M_A and M_φ with $M=6$ total bits, and their corresponding S&C trade-off.

$\boldsymbol{p}$ in the AE shown in Fig. 4. This also significantly reduces the number of trainable parameters: instead of $\widetilde{M}$ unconstrained probabilities, the $2^{M_\mathrm{ID}-1}$ unconstrained parameters of the one-dimensional target amplitude distribution $\tilde{\boldsymbol{p}}_\mathsf{a}$ define the probability of occurrence of all constellation points.

3) Low-Complexity LLR Calculation: With PAS, the receiver still uses classical BMD, which relies on (2). However, while (2) can be easily evaluated offline, i.e., during the optimization of the constellations, it is computationally too complex for a real-time implementation in 6G systems. Therefore, we briefly discuss how the structural properties of PAS can be exploited to reduce the computational complexity. Since the constellation points and their probabilities are fixed after optimization, the LLRs can be precomputed and stored in LUTs. However, in the general case, this requires a separate two-dimensional LUT per bit position, which covers both real and imaginary components, resulting in prohibitive memory requirements. Here, the structure of PAS offers a key advantage: due to the independent modulation of the real and imaginary parts and the separability of the channel [25], the LLRs can be computed separately per dimension and stored in one-dimensional LUTs, yielding low-complexity, memory-efficient LLR computation suitable for practical 6G receivers.

B. Generalized Probabilistic Amplitude Shaping

In conventional PAS, the real and imaginary components are shaped independently, and the joint distribution is the Cartesian product of two one-dimensional marginals. However, constructing PSK-like constellations, which are preferred under strong sensing constraints, from such a product is not possible. As a result, the Cartesian structure of conventional PAS inherently limits circular symmetry, resulting in suboptimal S&C performance under strong kurtosis constraints.

To overcome this limitation, we propose a generalization of the conventional PAS scheme. In particular, we extend the original PAS concept from ASK to amplitude and phase-shift keying (APSK)-like constellations, resulting in circularly symmetric constellations that are better suited for sensing. Furthermore, we show that the corresponding LLRs can still be approximated with low complexity and minimal memory, thereby preserving the implementation benefits of the original PAS architecture.

1) Constellation Design and Encoding Scheme: Conventional PAS operates on ASK constellations, where a single bit denotes the sign. We generalize this concept by introducing multiple "sign" bits, interpreted as phase bits, which no longer yield bipolar constellations, as in ASK, but enable multiple complex-valued phase factors, resulting in APSK constellations. These phase bits are mapped onto a set of uniformly spaced phases

$$s_u \in \mathcal{S}_\mathrm{GPAS} = \left\{ \exp\left(\mathrm{i} \frac{\pi(2m_\varphi+1)}{\tilde{M}_\varphi} \right) \middle| \, m_\varphi = 0, \ldots, \tilde{M}_\varphi - 1 \right\} \tag{38}$$

using binary reflected Gray mapping $\beta_\mathrm{S}: \{0,1\}^{M_\varphi} \to \mathcal{S}_\mathrm{GPAS}$, where $\tilde{M}_\varphi = 2^{M_\varphi}$ denotes the number of phase levels. Thus, each s_u is a complex phasor with unit magnitude and quantized angle. The remaining $M_\mathrm{A} = M - M_\varphi$ bits are used to label the amplitude using binary reflected Gray coding. This generalization transforms the conventional grid-like constellation into a circularly symmetric constellation, resembling APSK. Fig. 6 shows how different combinations of amplitude bits M_A and phase bits M_φ, with a total of $M=6$ bits, influence the generalized PAS constellations and their corresponding S&C trade-off for a communication SNR of 10 dB. The configuration with $M_\mathrm{A}=2$ amplitude bits and $M_\varphi=4$ phase bits achieves the highest GMI across all considered kurtosis values and therefore provides the best S&C trade-off among the evaluated combinations of amplitude bits M_A and phase bits M_φ. For this reason, it is selected for comparison in the following analysis. Unlike conventional PAS, generalized PAS directly produces complex-valued symbols.

Despite this change, the amplitude-phase factorization $x_u = a_u \cdot s_u$ still holds, where the amplitude $a_u \in \{1, 2, \ldots, 2^{M_\mathrm{A}}\}$ can be shaped independently, while the phasor s_u is uniformly distributed on a unit circle. This allows for assigning the approximately uniformly distributed parity-check bits to the phase. As a result, the encoding structure from conventional PAS can be reused, requiring only minor modifications such as allocating more bits to the phase and adapting the inverse sign mapping accordingly.

2) Optimizing Generalized PAS-Compatible Constellations Using an AE Framework: Similarly to conventional PAS, constellations for generalized PAS can be efficiently optimized using our AE framework by including the structural constraints of generalized PAS into the AE. The probability of occurrence of the constellation points is the outer product of the trainable target amplitude distribution and a fixed uniform distribution over the phase

$$\boldsymbol{p}_\mathrm{GPAS} = \frac{1}{\tilde{M}_\varphi} \cdot \mathrm{vec}\left(\mathrm{Softmax}\left(\tilde{\boldsymbol{p}}_\mathsf{a}\right) \otimes \mathbf{1}_{\tilde{M}_\varphi} \right), \tag{39}$$

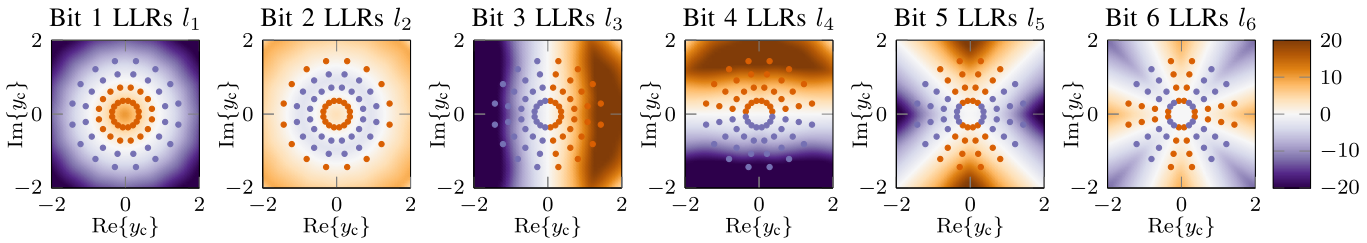


Fig. 7. Two-dimensional LLR values of received symbols for generalized PAS with a total of $M = 6$ bits, composed of $M_\mathrm{A} = 2$ amplitude bits (Bits 1 and 2) and $M_\varphi = 4$ phase bits (Bits 3 to 6), assuming equal probability of all constellation points at an SNR of 10 dB. The constellation points are colored orange if the corresponding bit is 0 and purple if it is 1.

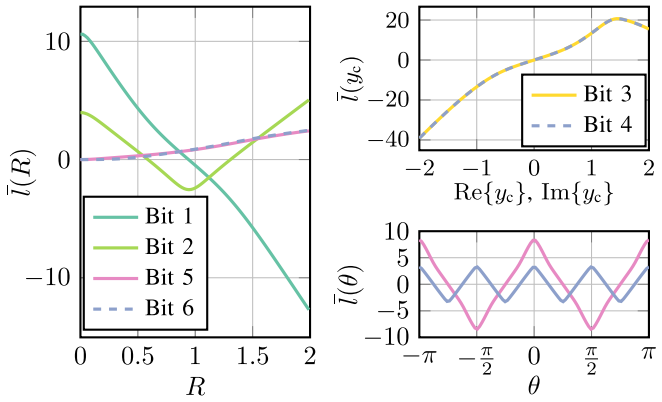


Fig. 8. One-dimensional LLRs for the proposed low-complexity LLR calculation, obtained by marginalizing the two-dimensional LLRs shown in Fig. 7. The LLRs are factorized into three domains: (left) amplitude R for bits 1, 2, 5, and 6; (top right) Cartesian components for bits 3 and 4; and (bottom right) phase θ for bits 5 and 6.

where $\mathbf{1}_{\tilde{M}_\varphi}$ is an all-one vector of length $\tilde{M}_\varphi$. The probability mass vector $\boldsymbol{p}_\mathrm{GPAS}$ satisfies the constraints of a conventional PAS constellation and replaces the probabilities $\boldsymbol{p}$ in the AE shown in Fig. 4.

3) Low-Complexity LLR Calculation: One key advantage of conventional PAS is the low-complexity LLR computation enabled by one-dimensional LUTs. In generalized PAS, while the constellation points and their probabilities factor into amplitude and angular components, this factorization does not hold for the channel transition probabilities. As a result, the LLR computation is not inherently separable. To address this, we analyze the structure of the exact 2D LLRs, and identify the feature of the received value that primarily determines the LLR values, such as amplitude, phase, or the real or imaginary part. Then, we approximate each LLR using only the dominant feature, reducing the computation to a one-dimensional LUT. Figure 7 shows the exact LLRs for a generalized PAS constellation with $M_\mathrm{A} = 2$ amplitude bits (bit 1 & bit 2) and $M_\varphi = 4$ phase bits (bit 3 - bit 6).

For convenience, we recall the LLR of bit position m, see (2)

$$l_{n,m}(y_{\mathrm{c},n}) = \log \frac{\sum_{x_n \in \mathcal{X}_m^{(0)}} f_{\mathsf{y}_\mathrm{c}|\mathsf{x}}(y_{\mathrm{c},n}|x_n) P_\mathsf{x}(x_n)}{\sum_{x_n \in \mathcal{X}_m^{(1)}} f_{\mathsf{y}_\mathrm{c}|\mathsf{x}}(y_{\mathrm{c},n}|x_n) P_\mathsf{x}(x_n)}, \tag{40}$$

which indicates whether bit m is more likely to be 0 or 1 and quantifies the reliability of this decision. We observe that the LLRs of the first and second bit depend mainly on the amplitude (and not the phase) of the received symbol. Therefore, we transform these LLRs $l_1(\mathrm{Re}\{y_\mathrm{c}\}, \mathrm{Im}\{y_\mathrm{c}\})$ and $l_2(\mathrm{Re}\{y_\mathrm{c}\}, \mathrm{Im}\{y_\mathrm{c}\})$ from Cartesian into polar coordinates, where $R = |y_\mathrm{c}|$ and $\theta = \sphericalangle\{y_\mathrm{c}\}$. Then, we average over the angular dimension θ

$$\bar{l}_m(R) \approx \frac{1}{2\pi} \int_0^{2\pi} l_m(R, \theta)\, \mathrm{d}\theta, \quad m \in \{1, 2\}, \tag{41}$$

resulting in LLRs that depend solely on the received magnitude R. Bits 3 and 4 correspond to the sign of the real and imaginary components, respectively. Thus, we can observe in Fig. 7 that their LLRs are primarily influenced by $\mathrm{Re}\{y_\mathrm{c}\}$ and $\mathrm{Im}\{y_\mathrm{c}\}$, respectively. Therefore, we average over the orthogonal component

$$\bar{l}_3(\mathrm{Re}\,\{y_\mathrm{c}\}) = \frac{1}{2g} \int_{-g}^{g} l_3(\mathrm{Re}\{y_\mathrm{c}\} + \mathrm{j} \cdot y)\, \mathrm{d}y\,, \tag{42}$$

$$\bar{l}_4(\mathrm{Im}\,\{y_\mathrm{c}\}) = \frac{1}{2g} \int_{-g}^{g} l_4(x + \mathrm{j} \cdot \mathrm{Im}\{y_\mathrm{c}\})\, \mathrm{d}x\,, \tag{43}$$

where g is set to 2 covering most received values y_c. Bits 5 and 6 encode further phase information, and we can observe in polar coordinates an oscillatory behavior over the angle θ, with their LLRs also increasing with magnitude R. We propose to approximate the 2D LLR functions in polar coordinates $l_5(R, \theta)$ and $l_6(R, \theta)$ each as a product of two one-dimensional functions

$$l_m(R, \theta) \approx \bar{l}_{m,R}(R) \cdot \bar{l}_{m,\theta}(\theta), \quad m \in \{5, 6\}, \tag{44}$$

where we first compute the phase-dependent term by averaging over the radius with $g = 2$

$$\bar{l}_{m,\theta}(\theta) \approx \frac{1}{g} \int_0^{g} l_m(R, \theta)\, \mathrm{d}R, \quad m \in \{5, 6\}\,. \tag{45}$$

Subsequently, we estimate the radial dependence by averaging

$$\bar{l}_{m,R}(R) \approx \frac{1}{2\pi} \int_0^{2\pi} \frac{l_m(R, \theta)}{\bar{l}_{m,\theta}(\theta)}\, \mathrm{d}\theta, \quad m \in \{5, 6\}\,. \tag{46}$$

Note that it is crucial to estimate the phase-dependent component first, as the oscillations lead to a zero average when first averaging over the phase.

Fig. 8 shows the marginalized LLR functions for all six bits, which can now be approximated using one-dimensional LUTs, enabling low-complexity implementation. In the top-right plot, we observe that bits 3 and 4 depend on their respective Cartesian component in an equivalent manner. This follows from the rotational symmetry of the shaped constellation (see

Fig. 7) and the fact that they represent coarse phase information, corresponding to the sign of the real and imaginary parts, respectively. Consequently, a single LUT suffices for both bits. Furthermore, the radial components of bits 5 and 6 are similar (see left plot) because both bits encode further phase information. Consequently, a single LUT suffices for the radial part of both bits. Therefore, only six one-dimensional look-up tables are needed in total: one each for bits 1 and 2, one for bits 3 and 4, one for the radial component shared by bits 5 and 6, and one for their respective phase components. We note that the one-dimensional LLRs can be approximated using piecewise linear functions, further reducing complexity. However, the specific implementation details are beyond the scope of this paper. Nevertheless, assuming V entries per dimension per LUT, this reduces the total storage requirement from $\mathcal{O}(6V^2)$ to $\mathcal{O}(7V)$, offering a significant memory saving. We remark that the proposed approximation generalizes to larger constellations: Amplitude bits remain largely insensitive to phase, the first two phase bits correspond to real and imaginary half-planes, and higher-order phase bits exhibit oscillatory angular patterns.

VI. Simulation Results

In this section, we first validate our derivation of the constellation-dependent detection probability through simulations. Furthermore, we evaluate the S&C performance and resulting trade-offs of the five constellation shaping methods under both SMD and BMD, and compare the results to lower and upper bounds on the maximum MI. Throughout this section, we optimize constellations with $M = 6$ bit/symbol for each kurtosis constraint $\tilde{\kappa} \in [1, 2]$ independently, under a communications SNR of $\mathrm{SNR_c} = 10$ dB and an increasing penalty factor d. The constellations are initialized as QAM and optimized by minimizing the loss function (34) using the Adam optimizer. During training, we increase the batch size from 500 to 10,000, while we decrease the learning rate, with the initial value depending on the shaping method. As discussed in Sec. II, $\tilde{\kappa} = 1$ should maximize sensing performance. On the contrary, a larger kurtosis constraint $\tilde{\kappa}$ is expected to improve communications performance, which is maximized for a circular symmetric Gaussian, that has a kurtosis of $\kappa = 2$. A kurtosis constraint $\tilde{\kappa}$ between these two extremes should yield a trade-off between S&C performance.

A. Optimized Constellations

Fig. 9 shows the optimized constellations for BMD using geometric, probabilistic (generic, conventional & generalized PAS), and joint constellation shaping for three kurtosis constraints $\tilde{\kappa} \in \{1.0, 1.2, 1.6\}$. We note that the strongest sensing constraint $\tilde{\kappa} = 1.0$ results in unit modulus constellations to reduce the sensing loss term (33). In this case, both geometrically and jointly shaped constellations resemble a PSK with overlapping constellation points. For generic probabilistic constellation shaping, a PSK is only approximated because the constellation points that have the same power do not have equal distances. For a looser sensing constraint $\tilde{\kappa} = 1.6$, the constellations approximate a Gaussian PDF to maximize communications performance. In between ($\tilde{\kappa} = 1.2$), a balance between Gaussian and unit modulus PDF is learned. Interestingly, the jointly shaped constellations exhibit more geometric shaping characteristics for a small kurtosis constraint $\tilde{\kappa}$ and more probabilistic behavior as the kurtosis constraint $\tilde{\kappa}$ increases.

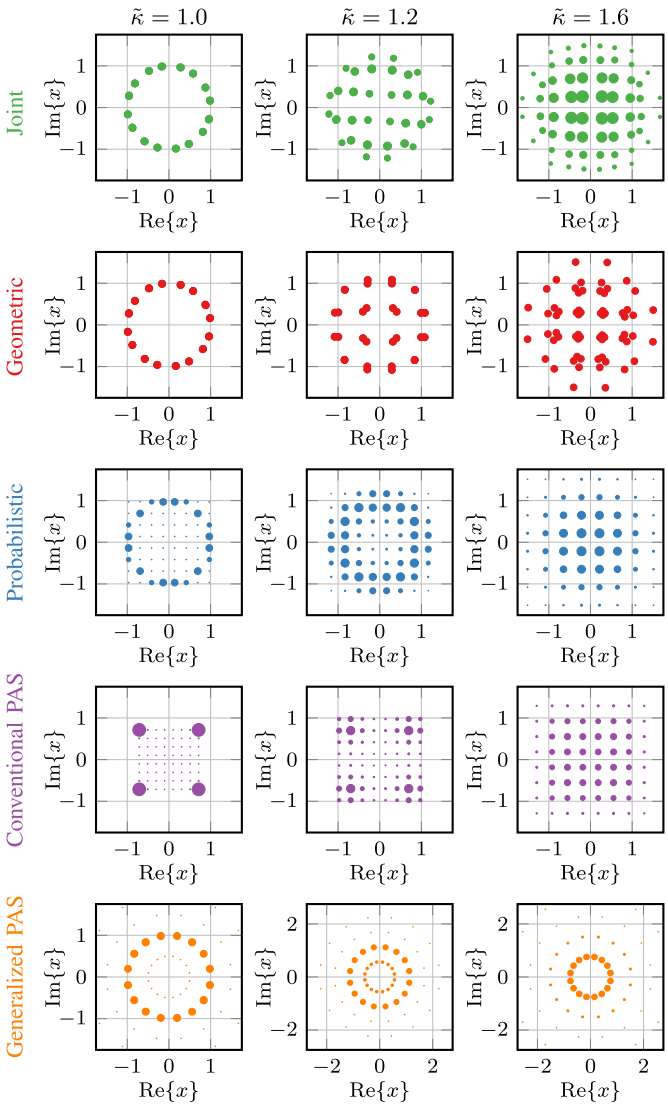


Fig. 9. Optimized constellations for geometric, probabilistic, and joint constellation shaping as well as conventional and generalized PAS under various kurtosis constraints $\tilde{\kappa}$. The size of each constellation point is proportional to its probability and each constellation point has an associated bit label, which is omitted for clarity.

On the one hand, for conventional PAS, the obtained constellation resembles a Gaussian for a weak sensing constraint (large $\tilde{\kappa}$), and is similar to (unconstrained) probabilistic constellation shaping, indicating that the structural constraints imposed by PAS do not significantly constrain optimization. However, for a strong kurtosis constraint, probabilistic constellation shaping yields a (blurry) ring, whereas PAS resembles a QPSK, clearly diverging from the expected circle-like PDF. This mismatch arises due to the independent shaping of the real and imaginary components, where the joint PDF is

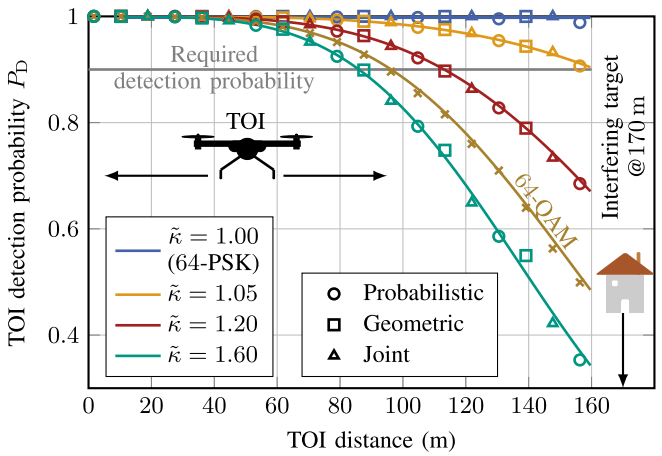


Fig. 10. Derived detection probability and simulated detection rates of a TOI, e.g., a drone, with a RCS of $\sigma_{\mathrm{RCS}} = 0.1$ m^2 following a Swerling-1 model in the presence of an interfering target, e.g., a building, at 170 m with an RCS of $\sigma_{\mathrm{RCS}} = 500$ m^2 following a Swerling-0 model [27].

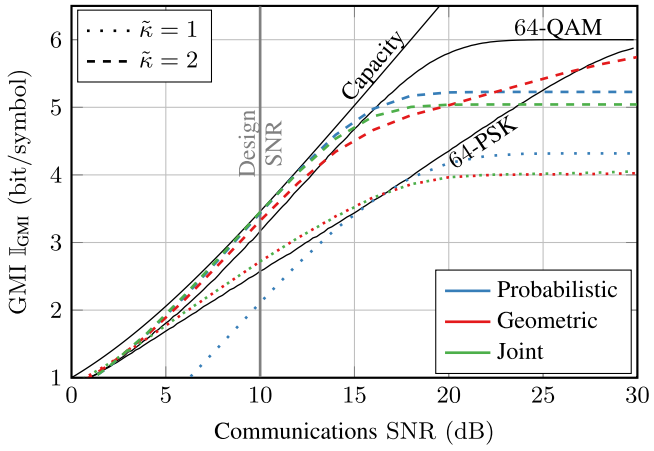


Fig. 11. Simulated communications performance (AIR) of the optimized constellations in comparison to legacy constellation formats.

the Cartesian product of two one-dimensional marginals and Cartesian structure of conventional PAS constrains optimization. On the other hand, generalized PAS with $M_{\mathrm{A}} = 2$ and $M_{\varphi} = 4$ is able to approximate the expected PSK-like constellation for a strong sensing constraint.

B. Sensing Performance

In Fig. 10, we consider a scenario with two targets: a distant interfering target and a nearby TOI, which should be detected and whose distance is varied. The RCS of the TOI follows a Swerling-1 model, whereas the interfering target follows a non-fluctuating Swerling-0 model. Multi-target scenarios are of particular interest because the average SINR (21) depends on the kurtosis and the power reflected by all targets. To verify the impact of the constellation on the detection performance, we evaluate the resulting detection rate for each of the optimized transmit constellations. The simulation setup follows [6], with the simulation parameters being the FR2 case from [32] with $N = 12672$ sub-carriers. For simplicity, we use an FFT size of 12672, and only one OFDM symbol. For CA-CFAR, we assume a false alarm rate of $P_{\mathrm{FA}} = 10^{-3}$ and a sliding window length of $N_{\mathrm{win}} = 100$. To obtain the detection probability, we perform a Monte-Carlo simulation with 100 independent Swerling-1 channel realizations. This inherently averages over the RCS fluctuations of the TOI, such that the mean target power converges to that determined by the mean RCS, see (21). Hence, the impact of the Swerling-1 fluctuations vanishes, and the detection probability is governed by the noise-and-interference power, which depends on the kurtosis κ of the constellation, as described in Sec. II-E.

We found that the simulated detection rates (markers) align well with the analytical (12) constellation-dependent detection probabilities (curves). This verifies our derivation in Sec. II and shows the impact of the kurtosis κ of the constellation on the detection probability. The minor discrepancies stem from the fact that the kurtosis κ of the optimized constellations slightly differs from the kurtosis constraint $\tilde{\kappa}$. As expected from (12) and (21), the detection probability decreases with increasing kurtosis κ and depends only on the kurtosis of the constellation, irrespective of the specific shaping method. Moreover, our simulations demonstrate that constellation shaping enables a dynamic adjustment of the detection range. For example, for a required target detection probability of $P_{\mathrm{D}} = 0.9$, the detection range can be varied from 90 m to beyond 160 m by modifying the kurtosis κ of the constellation.

C. Communications Performance

Fig. 11 shows the GMI [24, Eq. 14] as a function of the communications SNR for various sensing constraints $\tilde{\kappa}$. For $\tilde{\kappa} = 2$, all shaping methods reduce the gap to capacity and outperform conventional 64-QAM across an SNR range of 10 dB. We note that probabilistic and joint shaping reduce the entropy of the transmit symbols at the design SNR of $\mathrm{SNR}_{\mathrm{c}} = 10$ dB, resulting in a saturation of the GMI below 6 bit/symbol for large SNR values. For $\tilde{\kappa} = 1$, the GMI of the shaped constellations is similar to that of the PSK, although probabilistic constellation shaping performs slightly worse due to unequal distances between the constellation points and the absence of Gray coding. Both geometric and joint constellations slightly outperform the 64-PSK because their optimized constellations collapse into several tightly spaced point groups (see Fig. 9), effectively resembling a 16-PSK, which achieves a higher GMI at low SNR. However, within the SNR range of Fig. 11, these clusters are not distinguishable, practically reducing the number of reliably resolvable points and causing the GMI to follow that of a lower-order constellation. In principle, however, these points remain distinguishable as $\mathrm{SNR} \to \infty$. For the geometric constellation with a large kurtosis constraint, the points do not collapse but form locally dense groups (see Fig. 9). These constellation points are not distinguishable at low and moderate SNR values, but once the SNR exceeds approximately 20 dB, even these tightly spaced points become reliably detectable. Consequently, geometric shaping continues to increase its GMI toward the maximal achievable entropy of $\log_2(M)$ and eventually surpasses probabilistic and joint shaping. The GMI of the latter converges to the reduced

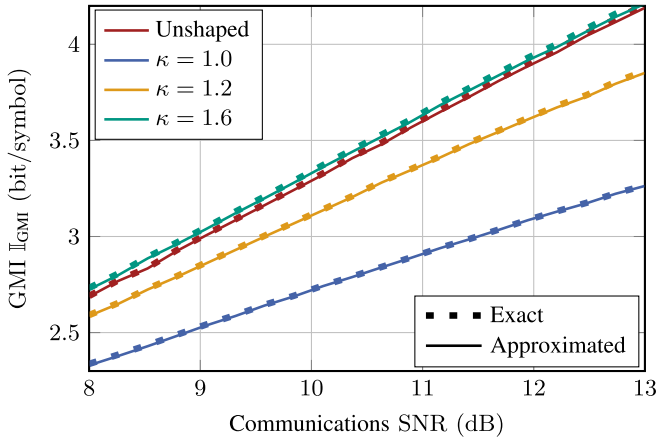


Fig. 12. Communication performance comparison between exact and low-complexity approximate LLR computation for unshaped and shaped generalized PAS constellations.

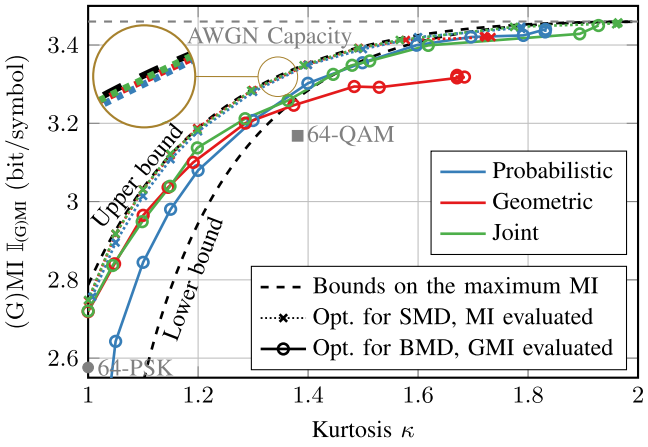


Fig. 13. Comparison of the derived bounds on the maximum MI and the S&C trade-off in terms of GMI and MI for different shaping methods and legacy modulation formats. Markers represent individually optimized constellations. For large kurtosis constraints, the kurtosis of the optimized constellations saturates, indicating that the maximum AIR is achieved with a kurtosis lower than the constraint. The kurtosis saturates below 2 for all shaping methods.

input entropy due to the non-uniform symbol probabilities in probabilistic shaping, which is optimized for a communications $\mathrm{SNR_c}$ of 10 dB. This explains the eventual dominance of geometric shaping for large SNR. Note that, in practice, communication systems do not operate over such a wide SNR range with a single constellation format. Instead, the constellation format is selected depending on the operating SNR.

We remark that the gap between the performance of PSK-like constellations with $\kappa = 1$ and the AWGN capacity reflects the S&C trade-off. Unit modulus, potentially low-order, constellations achieve optimal sensing performance with only a small gap to AWGN capacity at low SNR. At higher SNR, higher-order modulations are required to increase spectral efficiency for communications, resulting in an increasing S&C trade-off. Yet, the SNR in wireless systems is typically limited by power constraints and multipath fading. As a result, many practical systems operate in moderate SNR regimes, where 64-QAM is commonly used, making our parametrization particularly relevant for real-world deployments.

Next, we assess our proposed low-complexity LLR calculation for generalized PAS, and compare it against the exact LLR calculation (2). Fig. 12 shows the GMI as a function of the communications SNR for both unshaped (equiprobable transmit symbols) and shaped generalized PAS constellations, respectively. It can be observed that the curves corresponding to the exact and approximated LLR calculations match closely with a maximum deviation of less than 0.016 bit/symbol. This confirms that the proposed approximation enables low-complexity LLR computation with negligible performance loss.

D. Sensing & Communications Trade-off

Finally, we demonstrate how constellation shaping can govern the DRT in ISAC systems by analyzing the MI and GMI as a function of the kurtosis κ for geometric, probabilistic, and joint shaping. These results are compared in Fig. 13 to conventional modulation formats as well as the lower and upper bounds on the maximum MI derived in Sec. III to assess both the derived bounds and the effectiveness of our optimized constellations.

First, we observe that the AIR increases monotonically with increasing kurtosis across all shaping methods and for both bounds. It exhibits a particularly steep rise for lower kurtosis values and approaches the AWGN capacity for a kurtosis of $\kappa = 2$. Notably, the gap between the upper and lower bounds on the maximum MI narrows with increasing kurtosis, dropping below 0.1 bit/symbol for $\kappa \geqslant 1.35$ and vanishing at $\kappa = 2$. This is because for $\kappa = 2$, the involved PDFs become Gaussian and the EPI is tight. With decreasing kurtosis constraint, the deviation of the entropy-maximizing transmit PDF from a Gaussian increases, making the EPI and, therefore, the lower bound increasingly loose.

Next, we observe that the constellations optimized for SMD yield a MI that closely approaches the upper bound (25). In Fig. 14, we compare the empirical PDF of the received signal resulting from the optimized constellations to the maximum-entropy PDF used to derive the upper bound (27). For $\kappa = 2$, both PDFs converge to the circularly symmetric Gaussian PDF, which achieves the capacity for an AWGN channel. At lower kurtosis values, the two PDFs diverge slightly, as the upper bound does not constrain the output to be realizable under the system model, resulting in a gap between the achieved MI and the derived bound (25). Nevertheless, the gap in Fig. 13 is less than 0.04 bit/symbol at $\kappa = 1$ and below 0.01 bit/symbol for $\kappa \geqslant 1.05$. This demonstrates the effectiveness of the AE-based constellation optimization and shows that the upper bound can be closely approached by constellation shaping.

Now, we focus on the GMI as an AIR in practical communication systems using BMD. As expected, the GMI is consistently lower than the MI, with the performance gap depending on both the kurtosis constraint and the shaping method.

We observe that probabilistic constellation shaping effectively approaches the capacity and outperforms geometric constellation shaping if the sensing constraint is loose (large $\tilde{\kappa}$). On the contrary, geometric constellation shaping outperforms probabilistic constellation shaping for strict sensing constraints $\tilde{\kappa} < 1.3$. While this is also true for the MI, the

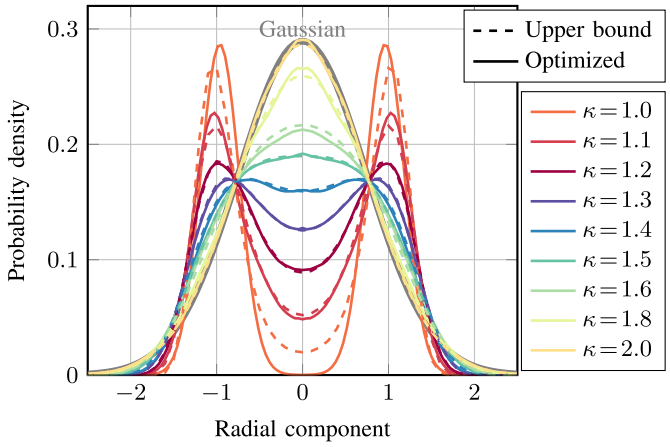


Fig. 14. Comparison of the empirical received signal PDF resulting from the optimized constellations and the theoretical maximum-entropy PDF used to derive the upper bound for various kurtosis constraints, along with a Gaussian as reference. Note that the original PDFs are two-dimensional. The displayed PDFs are obtained by averaging one-dimensional marginals computed along 64 uniformly spaced radial directions in the range $[0, \pi]$.

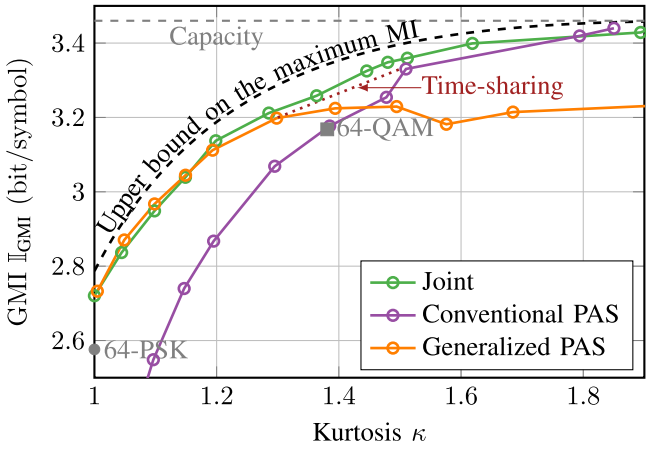


Fig. 15. Trade-off between communication performance (GMI) and sensing performance (kurtosis) for conventional and generalized PAS, compared to joint constellation shaping and conventional modulation formats.

performance differences are far more pronounced in terms of GMI. This highlights the importance of directly optimizing the GMI, rather than the MI as done in prior work [7], [11], [17].

Unlike MI, which measures symbol-wise reliability, GMI reflects bit-level reliability and is therefore sensitive to the binary labeling of the constellation points. However, for geometrically shaped constellations a Gray labeling may not exist. Gaussian-like constellations at high kurtosis values lack the regular grid structure needed for Gray coding, resulting in a performance gap at high kurtosis values. In probabilistic shaping, the bits are correlated, but BMD computes the LLRs independently, see (2), (4), leading to an AIR loss, especially under strong shaping, i.e., at low kurtosis. Moreover, the resulting constellations are non-uniformly spaced and lack Gray coding between diagonally placed symbols (see Fig. 9), which further degrades the GMI for low kurtosis values. To validate that the observed gap is indeed a BMD artefact rather than a result of the optimization process, we computed the MI of the constellations optimized for BMD and found it closely approaches the MI of the constellations optimized for SMD, confirming that the degradation stems from BMD.

Remark: This makes geometric and probabilistic constellation shaping well-suited for applications where sensing or communications performance is prioritized, respectively. Joint shaping consistently exhibits strong performance across all regimes, effectively leveraging the advantages of both geometric and probabilistic shaping. It performs similarly to geometric and probabilistic constellation shaping at low and high kurtosis values, respectively. Compared to conventional 64-QAM, joint constellation shaping achieves a 0.16 bit/symbol higher GMI at the same kurtosis κ. This means that joint shaping increases the communication throughput while maintaining the same detection probability for sensing. Conversely, joint shaping can reduce the kurtosis κ by 0.19 while achieving the same GMI, which directly translates into a higher detection probability of the TOI at equal communication performance. In addition, while 64-PSK and 64-QAM offer only two discrete operating points, joint shaping enables a continuous adjustment of the AIR-kurtosis operating point. This allows for a flexible and fine-grained trade-off between communication throughput and sensing detection probability. Furthermore, remember that the self-interference from the limited isolation between the transmitter and receiver paths manifests as a target with a large virtual RCS, and that the noise-and-interference (16) floor scales with the product of the kurtosis and the summed power of all targets. Consequently, if the self-interference is stronger than the received reflections of the targets, it dominates the noise-and-interference floor, which scales with the kurtosis of the employed constellations. Hence, reducing the kurtosis through constellation shaping can reduce the noise-and-interference floor caused by the self-interference and, consequently, increase the detection probability of a TOI with small RCS.

To quantify the performance gains of low-complexity constellation shaping, Fig. 15 compares conventional and generalized PAS in terms of their S&C trade-off and benchmarks them against joint constellation shaping and standard modulation formats. Conventional PAS performs best under weak kurtosis constraints, i.e., high kurtosis, closely approaching the AWGN capacity. However, as discussed previously, it suffers from a significant penalty when circular constellations with low kurtosis are required. In this regime, generalized PAS matches the performance of joint constellation shaping. In the intermediate regime, a small gap remains between the PAS variants and joint shaping, which could, in principle, be closed by time-sharing between both shaping configurations.

Remark: A combination of conventional and generalized PAS outperforms legacy modulation formats and achieves nearly the same S&C performance as joint shaping, which fully exploits all available degrees of freedom. However, in contrast to joint shaping, both conventional and generalized PAS maintain low complexity at the transmitter and receiver. This makes them well-suited for 6G ISAC systems, where maximizing performance while enabling a flexible trade-off

between S&C under tight implementation constraints will be crucial.

VII. Conclusion

We optimized and systematically compared geometric, probabilistic, and joint constellation shaping for OFDM-ISAC using a bitwise AE framework, maximizing the MI and the GMI under a sensing target detection probability constraint. We showed analytically and confirmed through simulations that the detection probability depends solely on the kurtosis of the constellation. Next, we derived lower and upper bounds on the maximum MI and showed that the optimized constellations achieve MI values close to the upper bound.

In practical systems using BMD, geometric shaping achieves a higher GMI under strict sensing constraints, while probabilistic shaping performs better under relaxed sensing constraints. Notably, our proposed joint shaping approach combines the strengths of both geometric and probabilistic constellation shaping, significantly outperforming legacy constellation formats and offering a flexible S&C trade-off.

Furthermore, we revisited PAS as a promising candidate for a practical implementation of constellation shaping and showed that its independent shaping of in-phase and quadrature components restricts the circular symmetry needed for constellations with low kurtosis. To overcome this, we generalized PAS to yield circularly symmetric constellations and proposed a low-complexity LLR approximation with negligible performance loss, preserving the original low-complexity advantage. Our results show that combining conventional and generalized PAS achieves near-joint shaping S&C performance with low complexity.

Thus, constellation shaping enables dynamically adjustable, high-performance S&C at low complexity, making it a compelling solution for flexible and efficient 6G ISAC systems.

Future research includes extending the proposed framework to frequency-selective communication channels, where the communications SNR varies across sub-carriers and power allocation becomes relevant, as discussed in Sec. II. Additional directions include incorporating the Doppler effect, i.e., velocity estimation, for sensing and studying constellation shaping under practical hardware impairments.

Appendix A

To show that the kurtosis κ of a conventional $\widetilde{M}$-QAM approaches $\kappa \approx 1.4$ as $\widetilde{M} \to \infty$ (see Sec. II-E), we use that, in this limit, the constellation converges to an independently and uniformly distributed random variable on $\mathcal{U}[-\sqrt{3/2}, \sqrt{3/2}]$ in both dimensions, i.e.,

$$\mathsf{x} = \mathsf{x}_{\mathrm{Re}} + \text{æ}\mathsf{x}_{\mathrm{Im}}, \quad \mathsf{x}_{\mathrm{Re}}, \mathsf{x}_{\mathrm{Im}} \sim \mathcal{U}\left[-\sqrt{3/2}, \sqrt{3/2}\right], \quad \mathsf{x}_{\mathrm{Re}} \perp \mathsf{x}_{\mathrm{Im}}. \tag{47}$$

The kurtosis then directly follows as

$$\mathbb{E}\{|\mathsf{x}|^4\} = \mathbb{E}\left\{\left(\mathsf{x}_{\mathrm{Re}}^2 + \mathsf{x}_{\mathrm{Im}}^2\right)^2\right\} = 2 \cdot \mathbb{E}\left\{\mathsf{x}_{\mathrm{Re}}^4\right\} + 2 \cdot \left(\mathbb{E}\left\{\mathsf{x}_{\mathrm{Re}}^2\right\}\right)^2 \tag{48}$$

$$= 2 \cdot \frac{9}{20} + 2 \cdot \left(\frac{1}{2}\right)^2 = \frac{7}{5}, \tag{49}$$

where we used $\mathbb{E}\{\mathsf{s}^4\} = a^4/5$ and $\mathbb{E}\{\mathsf{s}^2\} = a^2/3$ for $\mathsf{s} \sim \mathcal{U}[-a, a]$ with $a = \sqrt{3/2}$.

Appendix B

In this appendix, we show that solving the nonlinear system (28) is equivalent to solving (29) and inserting the result into (30) and (31). We begin by rewriting (28) for $q = 0, 1, 2$ as

$$C_q = \int_0^{2\pi} \int_0^{\infty} r^{2q+1} e^{\gamma_0 + \gamma_2 r^2 + \gamma_4 r^4} \, \mathrm{d}r \, \mathrm{d}\theta \tag{50}$$

$$= 2\pi \mathrm{e}^{\gamma_0} \int_0^{\infty} r^{2q+1} e^{\gamma_2 r^2 + \gamma_4 r^4} \, \mathrm{d}r \tag{51}$$

$$= \pi \mathrm{e}^{\gamma_0} \int_0^{\infty} u^q e^{\gamma_2 u + \gamma_4 u^2} \, \mathrm{d}u, \tag{52}$$

$$= \pi \mathrm{e}^{\gamma_0} \int_0^{\infty} u^q \mathrm{e}^{\frac{\gamma_2^2}{4\gamma_4}} \mathrm{e}^{-|\gamma_4|\left(u - \frac{\gamma_2}{2\gamma_4}\right)^2} \, \mathrm{d}u \tag{53}$$

$$= \pi \mathrm{e}^{\gamma_0} \mathrm{e}^{\frac{\gamma_2^2}{4\gamma_4}} \int_{-\gamma_2/(2|\gamma_4|)}^{\infty} \left(v + \frac{\gamma_2}{2|\gamma_4|}\right)^q \mathrm{e}^{-|\gamma_4| v^2} \, \mathrm{d}v \tag{54}$$

where we transform a 2D Cartesian complex-plane integral to polar coordinates with $Z = r\mathrm{e}^{\mathrm{j}\theta}$ in (50), and substitute $u = r^2$ in (52). In (53), we complete the square in the exponent and set $\gamma_4 = -|\gamma_4|$, since the integral diverges for $\gamma_4 > 0$. Finally, we substitute $v = u - \frac{\gamma_2}{2\gamma_4}$ in (54).

Using the integrals [33, entries: 3.321.2, 3.321.4, 3.321.5], we obtain closed-form expressions

$$C_0 = \pi \mathrm{e}^{\gamma_0} \frac{1}{2} \sqrt{\frac{\pi}{|\gamma_4|}} \operatorname{erfc}\left(\frac{-\gamma_2}{2\sqrt{|\gamma_4|}}\right) \mathrm{e}^{\frac{\gamma_2^2}{4|\gamma_4|}},$$

$$C_1 = \pi \mathrm{e}^{\gamma_0} \left[\frac{1}{2|\gamma_4|} + \frac{\gamma_2 \sqrt{\pi}}{4|\gamma_4|^{3/2}} \operatorname{erfc}\left(\frac{-\gamma_2}{2\sqrt{|\gamma_4|}}\right) \mathrm{e}^{\frac{\gamma_2^2}{4|\gamma_4|}}\right], \tag{55}$$

$$C_2 = \pi \mathrm{e}^{\gamma_0} \left[\frac{\gamma_2}{4|\gamma_4|^2} + \left(\frac{1}{4|\gamma_4|} + \frac{\gamma_2^2}{8|\gamma_4|^2}\right)\right. \tag{56}$$

$$\left. \cdot \sqrt{\frac{\pi}{|\gamma_4|}} \operatorname{erfc}\left(\frac{-\gamma_2}{2\sqrt{|\gamma_4|}}\right) \mathrm{e}^{\frac{\gamma_2^2}{4|\gamma_4|}}\right]. \tag{57}$$

Using $C_0 = 1$ (Tab. II), we substitute the common term in (55) into (56) and (57) yielding

$$C_1 = \frac{1}{2|\gamma_4|} \left(\pi \mathrm{e}^{\gamma_0} + \gamma_2\right), \tag{58}$$

$$C_2 = \frac{\gamma_2}{4|\gamma_4|^2} \left(\pi \mathrm{e}^{\gamma_0} + \gamma_2\right) + \frac{1}{2|\gamma_4|}. \tag{59}$$

Inserting (58) into (59) and then back into (58) yields (31) and (30). Finally, inserting both (30) and (31) in (55) yields (29).

References

[1] B. Geiger, F. Liu, S. Lu, A. Rode, and L. Schmalen, "Joint optimization of geometric and probabilistic constellation shaping for OFDM-ISAC systems," in *Proc. IEEE 5th Int. Symp. Joint Commun. Sens.*, Jan. 2025, pp. 1–6.

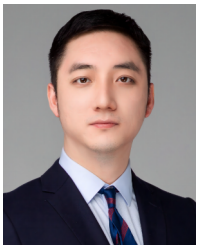
[2] T. Wild, A. Grudnitsky, S. Mandelli, M. Henninger, J. Guan, and F. Schaich, "6G integrated sensing and communication: From vision to realization," in *Proc. 20th Eur. Radar Conf. (EuRAD)*, Berlin, Germany, Sep. 2023, pp. 355–358.

[3] F. Liu et al., "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[4] T. Hwang, C. Yang, G. Wu, S. Li, and G. Ye Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[5] F. Liu et al., "CP-OFDM achieves the lowest average ranging sidelobe under QAM/PSK constellations," *IEEE Trans. Inf. Theory*, vol. 71, no. 9, pp. 6950–6967, Sep. 2025.
[6] M. Braun, "OFDM radar algorithms in mobile communication networks," Ph.D. dissertation, Karlsruhe Inst. Technol., Karlsruhe, Germany, 2014.
[7] Z. Du et al., "Probabilistic constellation shaping for OFDM-based ISAC signaling," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 509–514.
[8] M. F. Keskin et al., "Fundamental trade-offs in monostatic ISAC: A holistic investigation toward 6G," *IEEE Trans. Wireless Commun.*, vol. 24, no. 9, pp. 7856–7873, Sep. 2025.
[9] Y. Xiong, F. Liu, K. Wan, W. Yuan, Y. Cui, and G. Caire, "From torch to projector: Fundamental tradeoff of integrated sensing and communications," *IEEE BITS Inf. Theory Mag.*, vol. 4, no. 1, pp. 73–90, Mar. 2024.
[10] S. Lu et al., "Random ISAC signals deserve dedicated precoding," *IEEE Trans. Signal Process.*, vol. 72, pp. 3453–3469, 2024.
[11] Z. Du, F. Liu, Y. Xiong, T. X. Han, Y. C. Eldar, and S. Jin, "Reshaping the ISAC tradeoff under OFDM signaling: A probabilistic constellation shaping approach," *IEEE Trans. Signal Process.*, vol. 72, pp. 4782–4797, 2024.
[12] G. D. Forney and L.-F. Wei, "Multidimensional constellations. I. introduction, figures of merit, and generalized cross constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 877–892, Jun. 1989.
[13] K. Gümüs, A. Alvarado, B. Chen, C. Häger, and E. Agrell, "End-to-end learning of geometrical shaping maximizing generalized mutual information," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, Mar. 2020, pp. 1–3.
[14] J. Cho and P. J. Winzer, "Probabilistic constellation shaping for optical fiber communications," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1590–1607, Mar. 2019.
[15] M. Stark, F. A. Aoudia, and J. Hoydis, "Joint learning of geometric and probabilistic constellation shaping," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2019, pp. 1–6.
[16] V. Aref and M. Chagnon, "End-to-end learning of joint geometric and probabilistic constellation shaping," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, San Diego, CA, USA, Mar. 2022, pp. 1–3.
[17] Y. Liu, Y. Guo, Y. Gu, M. Wang, J. Liu, and B. Xia, "Probabilistic shaping-based ISAC systems with finite constellations: Analysis and optimization," *IEEE Trans. Wireless Commun.*, vol. 24, no. 12, pp. 9868–9881, Dec. 2025.
[18] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.
[19] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *J. Lightw. Technol.*, vol. 34, no. 7, pp. 1599–1609, Apr. 1, 2016.
[20] H. Sun et al., "800G DSP ASIC design using probabilistic shaping and digital sub-carrier multiplexing," *J. Lightw. Technol.*, vol. 38, no. 17, pp. 4744–4756, Sep. 2020.
[21] Y. C. Gültekin, W. J. van Houtum, A. G. C. Koppelaar, and F. M. J. Willems, "Enumerative sphere shaping for wireless communications with short packets," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1098–1112, Feb. 2020.
[22] S. Hu, H. Wang, and S. Semenov, "Supporting probabilistic constellation shaping in 5G-NR evolution," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 3586–3599, Apr. 2024.
[23] Y. Zhang, F. Liu, T. Liu, and S. Jin, "Optimal power allocation for OFDM-based ranging using random communication signals," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 6460–6473, 2025.
[24] M. Ivanov, C. Häger, F. Brännström, A. Graell i Amat, A. Alvarado, and E. Agrell, "On the information loss of the max-log approximation in BICM systems," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3011–3025, Jun. 2016.
[25] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," *J. Lightw. Technol.*, vol. 34, no. 21, pp. 5063–5073, Nov. 1, 2016.
[26] T. Cover and J. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2006.
[27] M. A. Richards, *Fundamentals of Radar Signal Processing*. New York, NY, USA: McGraw-Hill, 2014.
[28] B. Geiger. (2026). *Constellation Shaping for OFDM-ISAC Systems: From Theoretical Bounds To Practical Implementation*. [Online]. Available: https://github.com/kit-cel/isac_constellation_shaping
[29] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.
[30] S. Cammerer, F. A. Aoudia, S. Dörner, M. Stark, J. Hoydis, and S. Ten Brink, "Trainable communication systems: Concepts and prototype," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5489–5503, Sep. 2020.
[31] A. Rode, B. Geiger, S. Chimmalgi, and L. Schmalen, "End-to-end optimization of constellation shaping for Wiener phase noise channels with a differentiable blind phase search," *J. Lightw. Technol.*, vol. 41, no. 12, pp. 3849–3859, Jun. 2023.
[32] S. Mandelli, M. Henninger, M. Bauhofer, and T. Wild, "Survey on integrated sensing and communication performance modeling and use cases feasibility," in *Proc. 2nd Int. Conf. 6G Netw. (6GNet)*, Paris, France, Oct. 2023, pp. 1–8.
[33] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. New York, NY, USA: Academic, 2007.

Benedikt Geiger (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2020 and 2023, respectively, where he is currently pursuing the Ph.D. degree. From 2021 to 2022, he was an ERASMUS+ Research Intern with the Optical Networks Group, University College London, London, U.K. He wrote his master's thesis with Nokia, Stuttgart, Germany, in 2022. His research interests include joint communication and sensing for 6G, ultra high-speed fiber optic communication systems, and the application of machine learning in communications. He is a Student Member of the VDE and his awards and honors include the first place of the Rohde & Schwarz Engineering Competition in 2021. During his studies, he was a Scholar of the Konrad Adenauer Foundation and is an alumnus of KIT's Leadership Talent Academy.



Fan Liu (Senior Member, IEEE) received the B.Eng. and Ph.D. degrees from Beijing Institute of Technology (BIT), Beijing, China, in 2013 and 2018, respectively. He is currently a Full Professor with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, China. Prior to that, he was an Assistant Professor with the Southern University of Science and Technology, Shenzhen, China, from 2020 to 2024. Previously, he has held academic positions in the University College London (UCL), London, U.K., as a Visiting Researcher from 2016 to 2018; and a Marie Curie Research Fellow from 2018 to 2020. His research interests lie in the general area of signal processing and wireless communications and in particular in the area of integrated sensing and communications (ISAC). He is also the Founding Academic Chair of the IEEE ComSoc ISAC Emerging Technology Initiative (ISAC-ETI); the Vice Chair and a Founding Member of the IEEE SPS ISAC Technical Working Group (ISAC-TWG); an elected member of the IEEE SPS Sensor Array and Multichannel Technical Committee (SAM-TC); an Associate Editor of IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON MOBILE COMPUTING, and IEEE OPEN JOURNAL OF SIGNAL PROCESSING; and a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS, and *IEEE Vehicular Technology Magazine*. He was the TPC Co-Chair of the Inaugural Edition of the IEEE ISAC Conference 2026 and the Second and Fourth IEEE Joint Communication and Sensing (JC&S) Symposium; and the Symposium Co-Chair for the IEEE ICC 2026, IEEE/CIC ICCC 2025, IEEE WCNC 2024, and IEEE GLOBECOM 2023. He is a member of the IMT-2030 (6G) ISAC Task Group. He was listed among the 2025 Clarivate Highly Cited Researcher. He was a recipient of numerous best paper awards, including the 2025 IEEE Communications Society & Information Theory Society Joint Paper Award, 2024 IEEE Signal Processing Society Best Paper Award, 2024 IEEE Signal Processing Society Donald G. Fink Overview Paper Award, 2024 IEEE Communications Society Asia–Pacific Outstanding Paper Award, 2023 IEEE Communications Society Stephan O. Rice Prize, and 2021 IEEE SPS Young Author Best Paper Award.

Shihang Lu (Graduate Student Member, IEEE) received the B.E. degree from Anhui University (AHU) in 2018, the M.E. degree from the University of Science and Technology of China (USTC), Hefei, China, in 2021, and the Ph.D. degree from the Southern University of Science and Technology (SUSTech), Shenzhen, China, in 2025, with the Outstanding Doctoral Dissertation Award. His research interests include integrated sensing and communications (ISAC), low-altitude wireless networks, and convex optimization methods. He was a recipient of the Exemplary Reviewer of IEEE COMMUNICATIONS LETTERS and IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING. He was a recipient of the 2020 IEEE Sixth ICCC Best Presentation Award and the 2023 IEEE/CIC ICCC Best Paper Award. He served as the Session Chair for the IEEE ICC 2023 on SPC-5 ISAC. He also serving as a Guest Editor for IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY and the Workshop Technical Program Co-Chair for the IEEE ICC 2026.

Andrej Rode (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2016 and 2019, respectively, where he is currently pursuing the Ph.D. degree with the Communications Engineering Laboratory (CEL). His research focuses on optimizing communication systems using novel machine learning techniques with KIT. From 2019 to 2021, he held a position as a Young Graduate Trainee with the European Space Agency (ESA) with the European Space Research and Technology Centre (ESTEC), working in the section responsible for telemetry, tracking and command, and payload data transmission (TT&C and PDT) systems and techniques. His research interests encompass the application of machine learning to optimize communication systems, the optimization of geometric and probabilistic constellation shaping, the development of novel physical layer formats for multi-carrier and multi-user wireless communications, and joint sensing and communication for fiber-optical and wireless systems. Additionally, he serves as the Chair for the IEEE Student Branch Karlsruhe and is involved with the GNU Radio Project.

Daniel Gil Gaviria (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2018 and 2020, respectively, where he is currently pursuing the Ph.D. degree. He is also a Research and Teaching Assistant with the Communications Engineering Laboratory (CEL), KIT. His research interests include modern modulation schemes and signal processing for integrated sensing and communication (ISAC), specially in the context of automotive radar-communication and future cellular networks.

Charlotte Muth (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2019 and 2021, respectively, where she is currently pursuing the Ph.D. degree with the Communications Engineering Laboratory. Her research interests include the integration of sensing into communication systems, signal processing, and estimation using trainable algorithms and technologies for wireless communication. She has been part of the Executive Committee of the IEEE Student Branch at KIT since 2022.

Laurent Schmalen (Fellow, IEEE) received the Dipl.-Ing. degree in electrical engineering and information technology and the Dr.-Ing. degree from RWTH Aachen University, Germany, in 2005 and 2011, respectively. From 2011 to 2019, he was with Alcatel-Lucent Bell Laboratories and Nokia Bell Laboratories. From 2014 to 2019, he was also a Guest Lecturer with the University of Stuttgart, Stuttgart, Germany. Since 2019, he has been a Full Professor with Karlsruhe Institute of Technology, where he co-heads the Communications Engineering Laboratory. His research interests include channel coding, modulation formats, and optical communications. He was a recipient and a co-recipient of several awards, including the E-Plus Award for his Ph.D. thesis, the 2013 Best Student Paper Award from the IEEE Signal Processing Systems Workshop, the 2016 and 2018 Journal of Lightwave Technology Best Paper Awards, and a Bell Laboratories President Award. He served as an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS between 2021 and 2026.