Information-Theoretic Analysis of a Family of Improper Discrete Constellations

Ignacio Santamaria [1,*] , **Pedro M. Crespo** [2] , **Christian Lameiro** [3] **and Peter J. Schreier** [3]

[1] Departamento de Ingeniería de Comunicaciones, Universidad de Cantabria, 39005 Santander, Spain
[2] TECNUN, Universidad de Navarra, 20018 San Sebastián, Spain; pcrespo@.tecnun.es
[3] Signal & System Theory Group, Paderborn University, 33098 Paderborn, Germany;
     christian.lameiro@sst.upb.de (C.L.); peter.schreier@sst.upb.de (P.J.S.)
*   Correspondence: i.santamaria@unican.es; Tel.: +34-942-201552

Abstract: Non-circular or improper Gaussian signaling has proven beneficial in several interference-limited wireless networks. However, all implementable coding schemes are based on finite discrete constellations rather than Gaussian signals. In this paper, we propose a new family of improper constellations generated by widely linear processing of a square M-QAM (quadrature amplitude modulation) signal. This family of discrete constellations is parameterized by κ, the circularity coefficient and a phase ϕ. For uncoded communication systems, this phase should be optimized as $\phi^*(\kappa)$ to maximize the minimum Euclidean distance between points of the improper constellation, therefore minimizing the bit error rate (BER). For the more relevant case of coded communications, where the coded symbols are constrained to be in this family of improper constellations using $\phi^*(\kappa)$, it is shown theoretically and further corroborated by simulations that, except for a shaping loss of 1.53 dB encountered at a high signal-to-noise ratio (snr), there is no rate loss with respect to the improper Gaussian capacity. In this sense, the proposed family of constellations can be viewed as the improper counterpart of the standard proper M-QAM constellations widely used in coded communication systems.

1. Introduction

In the theoretical analysis of wireless communication systems, the transmitted symbols are typically assumed to be proper complex Gaussian random variables, i.e., the transmitted symbols are complex Gaussian and uncorrelated with their complex conjugate. This is motivated by the fact that such an input distribution achieves the capacity of the point-to-point, broadcast and multiple-access channels [1]. In more complex interference-limited scenarios, such as the interference channel, transmitting proper Gaussian signals is in general suboptimal, so assuming this input distribution in such cases is not theoretically justified. As a matter of fact, recent works have shown that, when treating interference as noise, a higher transmitted rate may be achieved in interference-limited networks by transmitting improper Gaussian signals, which are correlated with their complex conjugate [2–8].

Improper signals are common in communications since many important digital modulation schemes yield improper signals. Examples are binary phase shift keying (BPSK) or Gaussian minimum shift keying (GMSK) [9]. Additionally, power imbalance between the in-phase and in-quadrature components also yields an improper signal. A substantial body of work has focused on the design of widely linear receivers to exploit the impropriety of a received improper signal (we refer the interested reader to [10–13] and the references therein). While these works devise techniques to deal with signals whose impropriety was not intentionally introduced, recent works consider the deliberate transmission of improper signals as a novel way to handle interference in a multiuser channel.

Cadambe et al. showed the benefits of improper signaling for the first time in their pioneering work [2]. They considered a single-antenna three-user interference channel with constant channel extensions and showed that the number of degrees of freedom (DoF), which is the number of interference-free signals that can be transmitted, was larger when users transmit improper signals as opposed to proper ones. Similar DoF results were derived for the four-user interference channel in [14]. After that, a number of new works has emerged, and the payoffs of improper signaling have been revealed for different multiuser scenarios such as the interference channel [3–5,15], Z-interference channel [6,16–18], broadcast channels with linear precoding [19,20], underlay and overlay cognitive radio networks [7,8,21] and relay channels [22,23]. More recently, in [24], the authors showed the advantages of improper signaling in a multiuser scenario where a point-to-point link coexists with a multiple access channel. Even though these works present new and intriguing results that pave the way for more efficient interference management schemes, the vast majority of them assume Gaussian codebooks (the only exception is [15], where the symbol error rate is minimized for arbitrary digital modulation schemes) and, hence, can never be realized in practice since practical codes use non-Gaussian finite signaling constellations.

It is well known that the penalty in signal-to-noise ratio (snr) paid by constraining the complex symbols to a uniform M-ary quadrature amplitude modulation (QAM) constellation is given by the shaping loss, which for large M is 1.53 dB in the proper case [25]. When dealing with proper signals, this shaping loss is due to enclosing the uniformly-transmitted codewords inside a hypercube rather than a hypersphere, which would be the optimal region that minimizes the transmitted power.

When using non-Gaussian improper signals, different questions arise: What is the optimal way to generate an improper constellation with a given circularity coefficient (i.e., degree of impropriety)? In coded communication systems, what is the reduction in rate compared to the improper Gaussian capacity when the coded channel symbols are constrained to be in a given discrete improper constellation? Is this rate reduction due only to a shaping loss? Should the channel encoder be modified for discrete improper constellations? In this paper, we address these questions and consider the design of improper discrete constellations for the additive white Gaussian noise channel (AWGN). We focus on coded communication systems, and therefore, our aim is the design of discrete improper constellations for the coded bits to approach the improper Gaussian capacity (i.e., maximize the throughput). Consequently, mutual information is the figure of merit to be optimized. The design of practical capacity-achieving encoders-decoders will be considered in a future work. Therefore, the main contribution of this paper is to propose a class of improper constellations and to show that, in terms of mutual information, this family of constellations behaves very similarly to the standard family of M-QAM proper constellations.

Starting with a proper M-QAM constellation, we analyze the design of a widely linear transformation that generates a family of improper M-QAM constellations with a prescribed circularity coefficient κ. The only free parameter of these improper M-ary constellations is a phase $\phi \in [0, \pi/2]$. For uncoded communication systems, this phase should be optimized to maximize the minimum Euclidean distance between points of the improper constellation, so that the bit error rate (BER) is minimized. Note that for a given M (i.e., for a given spectral efficiency), maximizing d_{min} is equivalent, at high snr, to maximizing the mutual information (MI) between the transmit and the received signals [26]. On the other hand, for coded improper communication systems with a circularity coefficient κ, one should find the ϕ that maximizes, for each snr and M, the envelope of the resulting family of mutual information values of these M-ary improper constellations. In this way, the rate gap compared to the improper Shannon capacity is minimized.

One of the main contributions of this paper is to show that except for a shaping loss of 1.53 dB encountered at high snr, there is no rate loss with respect to the improper Gaussian Shannon capacity when using this family of constellations. We conclude that the proposed family of improper constellations behaves much the same as the standard family of proper M-QAM constellations. Finally,

note that as in the case of proper QAM constellations, the shaping loss can be reduced by designing codes that result in a Gaussian-like rather than uniform distribution of the coded symbols.

The rest of the paper is organized as follows. Section 2 provides preliminaries about improper random variables and introduces the considered model. The design of improper constellations is formulated in Section 3, and its information-theoretic analysis is done in Section 4. Section 5 provides some numerical examples. Finally, concluding remarks are presented in Section 6.

2. Background and System Model

2.1. Notation

Scalar random variables are denoted by uppercase letters X. Vectors and matrices are denoted by lowercase boldface letters, $\mathbf{x}$, and uppercase boldface letters, $\mathbf{X}$, respectively. The determinant and trace of a matrix $\mathbf{A}$ will be denoted as $\det(\mathbf{A})$ and $\operatorname{tr}(\mathbf{A})$. $\mathbf{I}$ is the identity matrix. Probability density functions are denoted by $p_Y(y)$. The superscript $(\cdot)^*$ denotes complex conjugate, and $(\cdot)^T$ denotes transpose. $E[\cdot]$ denotes mathematical expectation, and $h(Y)$ denotes the differential entropy of a continuous random variable. $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$ indicates that $\mathbf{x}$ is a complex circular Gaussian random vector of zero mean and covariance $\mathbf{R}$, while $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}, \tilde{\mathbf{R}})$ indicates that $\mathbf{x}$ is a complex non-circular Gaussian random vector of zero mean, covariance $\mathbf{R}$ and complementary covariance $\tilde{\mathbf{R}}$.

2.2. Improper Random Variables

This section introduces some preliminaries about improper random variables that will be used throughout the paper. The interested reader is referred to [9] for a comprehensive treatment of the topic.

Definition 1 ([9]). *A complex-valued random variable X is said to be circular (or circularly symmetric) if $e^{j\theta} X$ has the same probability distribution as X for all real θ.*

Definition 2 ([27]). *A complex-valued random variable X is said to be proper if X is uncorrelated with its complex conjugate, i.e., $E[(X - E[X])^2] = 0$.*

A circular X is proper, but the converse is not necessarily true. However, for complex Gaussian random variables, propriety implies circularity, so both terms can be used interchangeably. The next definitions characterize an improper random variable.

Definition 3 ([9]). *The complementary variance of a zero-mean complex random variable X is defined as $\tilde{\sigma}_X^2 = E[X^2]$.*

Furthermore, σ_X^2 and $\tilde{\sigma}_X^2$ are a valid pair of variance and complementary variance if and only if $\sigma_X^2 \geq 0$ and $|\tilde{\sigma}_X^2| \leq \sigma_X^2$.

Definition 4 ([9]). *The circularity coefficient of a complex random variable X, which measures the degree of impropriety, is defined as:*

$$\kappa = \frac{|\tilde{\sigma}_X^2|}{\sigma_X^2}. \tag{1}$$

The circularity coefficient satisfies $0 \leq \kappa \leq 1$. If $\kappa = 0$, then X is proper, otherwise improper. If $\kappa = 1$, we call X maximally improper.

There are two models that are typically used to express an improper complex random variable $X = X_I + jX_Q$: the augmented complex model and the real composite model. The former defines the

augmented random variable $\mathbf{x} = \begin{bmatrix} X & X^* \end{bmatrix}^T$. This way, its covariance matrix (called the augmented covariance matrix) contains all the second-order information of X, i.e.,

$$\mathbf{R}_X = \begin{bmatrix} \sigma_X^2 & \tilde{\sigma}_X^2 \\ (\tilde{\sigma}_X^2)^* & \sigma_X^2 \end{bmatrix}. \tag{2}$$

Alternatively, the real composite model defines the real random vector $\tilde{\mathbf{x}} = \begin{bmatrix} X_I & X_Q \end{bmatrix}^T$, and its covariance matrix also contains the complete second-order information of X. We now present some well-known results about the differential entropy of a complex Gaussian random variable.

Theorem 1 ([9])**.** *The differential entropy of an improper complex Gaussian random variable X is:*

$$h(X) = \frac{1}{2} \log \left[(\pi e)^2 \det(\mathbf{R}_X) \right] = \log \left(\pi e \sigma_X^2 \right) + \frac{1}{2} \log \left(1 - \kappa^2 \right). \tag{3}$$

Taking $\kappa = 0$ in the above expression yields the entropy of a proper Gaussian random variable [27]. Indeed, it is well known that the entropy of a complex random variable is maximized if it follows a proper complex Gaussian distribution [27].

2.3. A Motivating Example

In this subsection, we present a motivating example for the study conducted in this paper. Let us consider an underlay cognitive radio (UCR) system transmitting Gaussian codewords in which one single-antenna secondary user (SU) wishes to access the channel in the presence of a single-antenna primary user (PU) as shown in Figure 1, where the top and bottom links represent the PU and the SU, respectively. Without loss of generality, we consider the standard or canonical form of this scenario for which the direct channels are normalized to one, and the noise at both receivers is Gaussian and proper with zero mean and unit variance. This scenario is known in the information theory literature as the Z (or one-sided) interference channel [28].
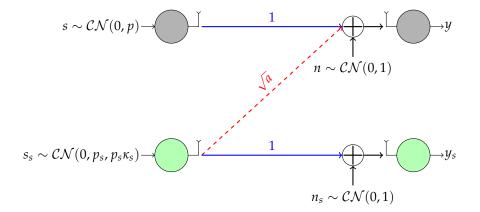

Figure 1. Underlay cognitive radio scenario with a secondary user that may transmit improper Gaussian signals.

In the UCR paradigm, the unlicensed SU is allowed to coexist with the PU as long as they ensure a minimum transmission rate $\bar{R}_{PU} \geq \xi \log(1 + p)$ for the PU, where $\xi \in [0, 1]$ is the loading factor. That is, the transmit power and signaling scheme of the SU must be selected to control the interference level at the primary receiver such that the PU's rate requirement is guaranteed. Since the PU is typically unaware of the SU, it is also assumed that the PU always transmits proper Gaussian signals, whereas the SU may transmit either proper or improper Gaussian signals depending on which signaling scheme

performs better. This problem has been analyzed in [7], where it has been proven that the SU improves its rate by transmitting Gaussian improper signals if and only if (cf. Theorem 1 in [7]):

$$a \geq \beta = 1 - \frac{p}{(1+p)^{2\xi} - 1},$$

(4)

where p is the power transmitted by the PU. Furthermore, Theorem 1 in [7] also gives us the optimal circularity coefficient κ_s that must be used by the SU depending on its power budget.

For example, let us suppose that the cross-link channel is $a = 4$; the power budgets for the PU and SU are 100 and 20, respectively; and the loading factor is $\xi = 0.4$. With these parameters $\beta = -1.56$ and, since $a \geq \beta$, improper signaling with $\kappa_s = 0.96$ (close to maximally improper) improves the rate of the SU. Specifically, the rate achieved by the SU using improper Gaussian codewords is $R_{SU} = 3.09$ b/s/Hz, whereas using proper Gaussian codewords, the rate would be only $R_{SU} = 2.44$ b/s/Hz.

As we have already argued, practical constraints dictate the use of discrete constellations instead of Gaussian signals. Nevertheless, the previous analysis with Gaussian signals is still valuable as an upper bound in the sense that any finite constellation with the same circularity coefficient will have lower differential entropy than a Gaussian and, consequently, will be less harmful for the PU. Therefore, in our example, for any improper discrete constellation, we might transmit with $\kappa_s = 0.96$, and the rate requirement of the PU would still be satisfied.

When the PU transmits Gaussian proper signals and the interference at the SU is treated as noise, the SU link is an AWGN channel over which we have to send improper digital signals with a given κ_s. In a more realistic scenario, the PU would also transmit a discrete non-Gaussian constellation. In this case, the interference power tolerated by the PU could potentially vary with respect to Gaussian signaling, and the SU would receive non-Gaussian interference, as well. While all Gaussian signals with a given circularity coefficient are identical in terms of performance, this is not the case with our family of discrete constellations. As we will see shortly, with discrete improper constellations, we have one degree of freedom to optimize the SU performance. In this work, we study a family of improper constellations and design its parameters to maximize the mutual information over the SU link modeled as an AWGN channel. Obviously, if we maximize the mutual information, the SU will cause more harm to the PU, but still, this harm is upper bounded by the harm caused by a Gaussian signal with the same κ_s, so the PU will sustain its rate.

2.4. Capacity Results for Gaussian Improper Signals

We consider a complex AWGN channel:

$$Y = g\sqrt{\text{snr}}\, V + N,$$

(5)

where $N \sim \mathcal{CN}(0,1)$ is circular, V is the transmitted signal and the channel is a deterministic and known constant, which we take as $g = 1$ without loss of generality.

The capacity (in bits per channel use) of this channel is [1]:

$$C = \sup_{p_V(v)} \text{I}(Y;V) = \sup_{p_V(v)} h(Y) - h(Y|V) = \sup_{p_V(v)} h(Y) - \log(\pi e),$$

(6)

where the supremum is taken over all possible input distributions of V with $E[|V|^2] = 1$, and the logarithms have base two. It is well known that the supremum is achieved when $V \sim \mathcal{CN}(0,1)$ [27].

In this paper, we consider the case where the input signal is unit-power improper with circularity coefficient κ; then, the channel output is also improper Gaussian with variance $E[|Y|^2] = 1 + \text{snr}$ and circularity coefficient $\kappa_y = \kappa\text{snr}/(1 + \text{snr})$. By Theorem 1, the differential entropy of Y is:

$$h(Y) = \log(\pi e(1 + \text{snr})) + \frac{1}{2}\log(1 - \kappa_y^2),$$

(7)

and the capacity is:

$$C(\text{snr}, \kappa) = \log(1 + \text{snr}) + \frac{1}{2} \log \left(1 - \frac{\kappa^2}{(1 + \text{snr}^{-1})^2} \right) \quad \text{bits per complex dimension.} \tag{8}$$

Observe that for $0 < \kappa \le 1$, the second term in (8) is always negative, and therefore, there is a capacity loss with respect to proper Gaussian signaling. In Appendix A, we present a heuristic derivation of this formula based on a packing argument similar to [29].

Remark 1. *As* snr $\to \infty$, *the term:*

$$\frac{1}{2} \log \left(1 - \frac{\kappa^2}{(1 + \text{snr}^{-1})^2} \right)$$

in (8) accounts for the loss in capacity due to enclosing the codewords in a multidimensional ellipsoid rather than in a hypersphere (as proper Gaussian signaling does) [30].

3. A Family of Improper Constellations Generated by a Widely Linear Transformation

In this section, we propose a family of improper discrete constellations that are generated by a widely linear transformation of a standard M-QAM constellation, where $M = 4s^2$ with s a positive integer. To this end, let $X = X_I + jX_Q$ be a symbol in a proper M-QAM constellation of unit power ($E[|X|^2] = 1$). Then, we propose to generate the corresponding improper symbol as:

$$V = V_I + jV_Q = h_1 X + h_2 X^*, \tag{9}$$

where $h_1 = h_{1I} + jh_{1Q}$ and $h_2 = h_{2I} + jh_{2Q}$ are complex scalars such that $|h_1|^2 + |h_2|^2 = 1$ (to maintain unit transmit power). Alternatively, the widely linear precoding operation (9) can be expressed in terms of real variables as:

$$\begin{bmatrix} V_I \\ V_Q \end{bmatrix} = \begin{bmatrix} h_{1I} + h_{2I} & h_{2Q} - h_{1Q} \\ h_{1Q} + h_{2Q} & h_{1I} - h_{2I} \end{bmatrix} \begin{bmatrix} X_I \\ X_Q \end{bmatrix}. \tag{10}$$

Under the constraint $|h_1|^2 + |h_2|^2 = 1$, the circularity coefficient of the generated symbol is:

$$\kappa = \frac{|E[V^2]|}{E[|V|^2]} = 2|h_1||h_2|. \tag{11}$$

The question that now arises is: given κ, what is the simplest parametrization of this family of improper constellations in terms of the complex variables h_1 and h_2? All values of h_1 and h_2 producing a unit variance improper V with a given κ must satisfy:

$$|h_1|^2 |h_2|^2 = \frac{\kappa^2}{4}, \tag{12}$$
$$|h_1|^2 + |h_2|^2 = 1.$$

The nonlinear system of equations in (12) has a unique solution for the squared modulus $|h_1|^2$ and $|h_2|^2$, which can be written in terms of the variable $\alpha = \sqrt{1 - \kappa^2}$ as:

$$|h_1|^2 = \frac{1}{2}(1 + \alpha), \tag{13}$$
$$|h_2|^2 = \frac{1}{2}(1 - \alpha).$$

Since for the AWGN channel considered in (5), the noise is Gaussian and proper (circular), it is clear that the mutual information does not change when the constellation is rotated by an arbitrary angle. This means that we can always rotate the constellation V such that h_1 is a real and positive parameter, whereas h_2 is a complex value. With this restriction, h_1 and h_2 can be expressed as:

$$h_1 = \sqrt{\frac{1}{2}(1+\alpha)}, \tag{14}$$

$$h_2 = \sqrt{\frac{1}{2}(1-\alpha)}e^{j\phi}. \tag{15}$$

Furthermore, the original proper M-QAM constellation X has a rotational symmetry of $\pm n\pi/2$ for any integer n, which means that ϕ in (15) can be restricted to the interval $\phi \in [0, \pi/2]$. Finally, the proposed family of improper constellations with circularity coefficient κ can be parameterized as:

$$V = \sqrt{\frac{1}{2}(1+\alpha)}X + \sqrt{\frac{1}{2}(1-\alpha)}e^{j\phi}X^*, \tag{16}$$

where $\alpha = \sqrt{(1-\kappa^2)}$, $0 \leq \kappa \leq 1$ and $\phi \in [0, \pi/2]$. In what follows, to stress the dependency of the improper constellation on κ, ϕ and M, we will denote this family of constellations as $\mathcal{A}(\phi, \kappa, M)$.

In terms of real variables, the transformation in (16) can be written as:

$$\begin{bmatrix} V_I \\ V_Q \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{0.5(1+\alpha)} + \sqrt{0.5(1-\alpha)}\cos(\phi) & \sqrt{0.5(1-\alpha)}\sin(\phi) \\ \sqrt{0.5(1-\alpha)}\sin(\phi) & \sqrt{0.5(1+\alpha)} - \sqrt{0.5(1-\alpha)}\cos(\phi) \end{bmatrix}}_{\mathbf{H}(\phi,\kappa)} \begin{bmatrix} X_I \\ X_Q \end{bmatrix}, \tag{17}$$

where $\mathbf{H}(\phi, \kappa)$ denotes the 2×2 transformation matrix.

Observe that by setting $\phi = 0$ in (16), one obtains a rather naive rectangular QAM improper constellation, where its real and imaginary components are independent and have a power imbalance determined by the required circularity coefficient κ. For any other $\phi \neq 0$ the constellation points are enclosed in a non-rectangular parallelogram. Figure 2 shows four improper constellations generated according to (16) for different values of κ and ϕ when X is proper QPSK ($M = 4$). For $\kappa = 1$ (maximally improper signal), in the right bottom panel, the points of the constellation lie on a line. Similarly, Figure 3 shows the constellation $\mathcal{A}(\pi/2, 0.7, 256)$, where the shape of the constellation, a parallelogram, can be more clearly observed.

We now provide an alternative geometric interpretation of the proposed improper constellations $\mathcal{A}(\phi, \kappa, M)$. Consider the 2D integer lattice $Z^2 \subset \mathbb{R}^2$, consisting of all two-dimensional vectors $\mathbf{z} = (a, b)^T$ with a and b integers. Then, the M-QAM constellation $\frac{1}{2}[\pm 1, \pm 3, \ldots, \pm(\sqrt{M}-1)]^2$ can be obtained as the intersection of the translated lattice $Z^2 + (\frac{1}{2}, \frac{1}{2})^T$ and the square region $\mathcal{R} = [-\frac{\sqrt{M}}{2}, \frac{\sqrt{M}}{2}]^2 \subset \mathbb{R}^2$.

If we apply the linear transformation $\mathbf{H}(\phi, \kappa)$ to the lattice $Z^2 + (\frac{1}{2}, \frac{1}{2})^T$, we obtain the new translated lattice $\Lambda(\phi, \kappa) + \mathbf{H}(\phi, \kappa)(\frac{1}{2}, \frac{1}{2})^T$, where:

$$\Lambda(\phi, \kappa) = \{\boldsymbol{\lambda} \in \mathbb{R}^2 : \boldsymbol{\lambda} = \mathbf{H}(\phi, \kappa)\mathbf{z}, \ \forall \mathbf{z} \in Z^2\}. \tag{18}$$

Reasoning as before, the unnormalized family of improper constellations $\mathcal{A}(\phi, \kappa, M)$ can be obtained as the intersection of the translated lattice $\Lambda(\phi, \kappa) + \mathbf{H}(\phi, \kappa)(\frac{1}{2}, \frac{1}{2})^T$ and the parallelogram region $\mathcal{P} \subset \mathbb{R}^2$ that results from the $\mathbf{H}(\phi, \kappa)$-transformation of a centered square with side length $\sqrt{M}$, that is,

$$\mathcal{A}(\phi, \kappa, M) = \left\{ \Lambda(\phi, \kappa) + \mathbf{H}(\phi, \kappa)\left(\frac{1}{2}, \frac{1}{2}\right)^T \right\} \cap \mathcal{P}. \tag{19}$$
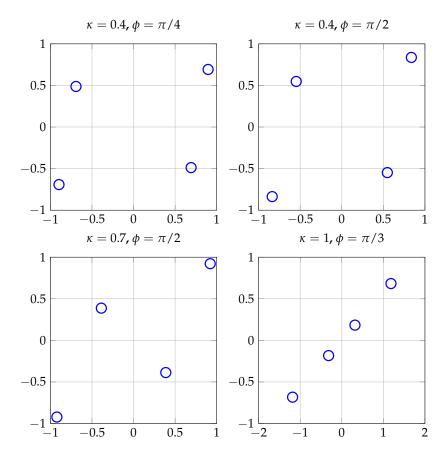
Figure 2. Improper constellations generated according to (16) for different values of κ and ϕ when X is proper Quadrature Phase Shift Keying (QPSK).
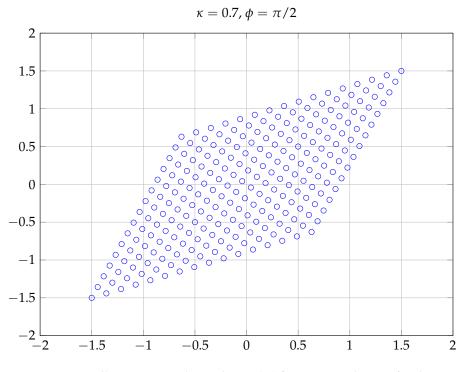

Figure 3. Improper constellation generated according to (16) for $\kappa = 0.7$ and $\phi = \pi/2$ when X is proper 256-QAM (Quadrature Amplitude Modulation).

Note that by doing this, only M points of the lattice $\Lambda(\phi,\kappa) + \mathbf{H}(\phi,\kappa)(\frac{1}{2},\frac{1}{2})^T$ are enclosed, and their average energy is $(M-1)/6$. Since translation of a lattice does not affect its minimum distance between signal points, d_{min}, we conclude that the minimum distance of the unnormalized family of constellations $\mathcal{A}(\phi,\kappa,M)$ (19) is given by the minimum distance of the lattice $\Lambda(\phi,\kappa)$. For a given κ, the optimal ϕ that maximizes $d_{min}(\Lambda(\phi,\kappa))$ can be obtained as:

$$\phi^*(\kappa) = \arg\max_{\phi \in [0,\pi/2]} d_{min}(\Lambda(\phi,\kappa)). \tag{20}$$

This problem will be solved in Section 4. Taking into account that $\phi^*(\kappa)$ does not depend on the scaling of the lattice $\Lambda(\phi,\kappa)$, we arrive at the following proposition:

Proposition 1. *The angle ϕ^* that maximizes the minimum distance of the family constellations $\mu \times \mathcal{A}(\phi,\kappa,M)$, where μ is an arbitrary normalization factor, only depends on the circularity coefficient κ and is given by (20), where $\Lambda(\phi,\kappa)$ in (18) is the lattice that results from the $\mathbf{H}(\phi,\kappa)$-transformation of the integer lattice $\mathbb{Z}^2$.*

4. Information-Theoretic Analysis

4.1. Introduction

To assess the performance of the proposed family of improper constellations $\mathcal{A}(\phi,\kappa,M)$ in coded communication systems, one has to check whether the maximum reliable communication rates achievable by channel encoders designed with coded symbols in these constellations are close to the rates dictated by the improper Gaussian capacity (8). To that end, let us define the mutual information envelope, denoted by $I^*(\text{snr},\kappa)$, as follows. Given κ and M, we obtain for each snr the optimal phase $\hat{\phi}(\kappa,\text{snr},M)$ that maximizes the mutual information $I(Y,V)$. By repeating this optimization process for each M, a family of mutual information curves is obtained. $I^*(\text{snr},\kappa)$ is now obtained as the envelope of these curves, that is,

$$I^*(\text{snr},\kappa) = \max_M \{I(Y;V) : V \in \mathcal{A}(\hat{\phi}(\kappa,\text{snr},M),\kappa,M)\}.$$

The closer $I^*(\text{snr},\kappa)$ is to the improper Gaussian capacity $C(\kappa,\text{snr})$, the better the proposed family of constellations will be. A way of quantifying performance is to obtain the signal-to-noise ratio gap, $\Delta(\text{snr},\kappa)$:

$$C(\kappa,\text{snr}) = I^*(\text{snr} + \Delta(\text{snr},\kappa),\kappa).$$

Although the analytical derivation of $\Delta(\text{snr},\kappa)$ looks intractable, useful insights into the behavior of $\Delta(\text{snr},\kappa)$ versus snr can be gained by analyzing in the high- and low-snr regimes the family of mutual information values, $\{I(Y;V) : V \in \mathcal{A}(\hat{\phi}(\kappa,\text{snr},M),\kappa,M),$ for all M and snr values$\}$. This is studied in the following sections.

Section 4.2 shows that at low snr, the behavior of $C(\kappa,\text{snr})$ and $I^*(\text{snr},\kappa)$ is the same up to the second order in snr. In Section 4.3, we show that at high snr, the optimal phase $\hat{\phi}(\kappa,\text{snr},M)$ that maximizes the mutual information $I(Y;V)$ with $V \in \mathcal{A}(\phi,\kappa,M)$ does not depend on M and is given by $\phi^*(\kappa)$, i.e., the phase that maximizes the minimum distance of the lattice $\Lambda(\phi,\kappa)$ (refer to Proposition 1). Observe that maximizing the mutual information in this context is equivalent to minimizing the BER for uncoded systems.

In Section 4.4, we compute the shaping loss of the proposed family of constellations $\mathcal{A}(\phi,\kappa,M)$ with respect to improper Gaussian signaling. It is shown that this loss is 1.53 dB, for all ϕ and κ. Although this result can easily be checked for rectangular QAM improper constellations (when $\phi = 0$), it has to be proven for any other value of $\phi \in (0,\pi/2]$, which generates non-rectangular constellations.

Section 5 shows by computer simulation that the gap, $\Delta(\text{snr},\kappa)$, increases monotonically with the snr up to a saturation value of 1.53 dB. Since at low snr, the value of $\Delta(\text{snr},\kappa)$ is negligible, we conclude

that the only loss in rate is due to shaping. In this sense, the proposed family of constellations can be viewed as the improper counterpart of the standard proper *M*-QAM constellations.

4.2. Asymptotic Results at Low SNR

The following theorem characterizes the mutual information at low snr up to the second order.

Theorem 2. *Let V be a discrete improper random variable with circularity coefficient $0 \leq \kappa \leq 1$ taking values in $\mathcal{A}(\phi, \kappa, M)$ and constructed as in (16) with X being a square proper M-QAM constellation. Then, the mutual information between Y and V in (5) admits the following second-order expansion around* snr = 0:

$$I(\sqrt{\text{snr}}\, V + N; V) = \text{snr} - \frac{1}{2}(1 + \kappa^2)\text{snr}^2 + O(\text{snr}^3). \tag{21}$$

Proof. From Expression (17), the random variable $V = V_I + jV_Q$ can be expressed in terms of real variables as:

$$\underbrace{\begin{bmatrix} V_I \\ V_Q \end{bmatrix}}_{\mathbf{v}} = \underbrace{\begin{bmatrix} \sqrt{0.5(1+\alpha)} + \sqrt{0.5(1-\alpha)}\cos(\phi) & \sqrt{0.5(1-\alpha)}\sin(\phi) \\ \sqrt{0.5(1-\alpha)}\sin(\phi) & \sqrt{0.5(1+\alpha)} - \sqrt{0.5(1-\alpha)}\cos(\phi) \end{bmatrix}}_{\mathbf{H}(\phi,\kappa)} \underbrace{\begin{bmatrix} X_I \\ X_Q \end{bmatrix}}_{\mathbf{x}}, \tag{22}$$

where $\alpha = \sqrt{1 - \kappa^2}$ and $X_I + jX_Q$ is a proper uniformly-distributed *M*-QAM random variable. Then, the complex-valued AWGN channel reduces to the real-valued vector channel:

$$\mathbf{y} = \mathbf{H}(\phi, \kappa)\mathbf{x} + \mathbf{n}, \tag{23}$$

where $\mathbf{y} = (Y_I, Y_Q)^T$ and $\mathbf{n} = (N_I, N_Q)^T$, with N_I and N_Q being independent real Gaussian random variables with zero mean and variance $\frac{1}{2}$.

We exploit the known relationship between mmse and mutual information (MI) [31], namely,

$$\frac{d}{d\,\text{snr}}I(\sqrt{\text{snr}}) = \text{mmse}(\sqrt{\text{snr}}), \tag{24}$$

and use a result in [31] showing that the mmse of a real-valued vector channel can be expanded around snr = 0 up to the first order as (notice that the factor 1/2 in the expression is because, according to our formulation, $E[\mathbf{x}\mathbf{x}^T] = \frac{1}{2}\mathbf{I}$):

$$\text{mmse}(\text{snr}, \phi, \kappa) = \frac{1}{2}\text{tr}(\mathbf{H}(\phi, \kappa)\mathbf{H}(\phi, \kappa)^T) - \frac{1}{2}\text{snr} \cdot \text{tr}((\mathbf{H}(\phi, \kappa)\mathbf{H}(\phi, \kappa)^T)^2) + O(\text{snr}^2). \tag{25}$$

Since the matrix $\mathbf{H}(\phi, \kappa)$ is real and symmetric, it is always diagonalizable as:

$$\mathbf{H}(\phi, \kappa) = \mathbf{Q}\Lambda\mathbf{Q}^T,$$

where **Q** is an orthogonal matrix ($\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$) and:

$$\Lambda = \text{diag}\{\lambda_1(\phi, \kappa), \lambda_2(\phi, \kappa)\},$$

with $\lambda_j(\phi, \kappa), j = 1, 2$, being its real eigenvalues. It is easy to check that:

$$\lambda_1(\phi, \kappa) = \sqrt{0.5(1+\alpha)} - \sqrt{0.5(1-\alpha)},$$

$$\lambda_2(\phi, \kappa) = \sqrt{0.5(1+\alpha)} + \sqrt{0.5(1-\alpha)}.$$

Note that these eigenvalues do not depend on ϕ. By the properties of the trace operator $\text{tr}(\mathbf{H}(\phi,\kappa)\mathbf{H}(\phi,\kappa)^T) = \lambda_1^2(\phi,\kappa) + \lambda_2^2(\phi,\kappa) = 2$ and $\text{tr}((\mathbf{H}(\phi,\kappa)\mathbf{H}(\phi,\kappa)^T)^2) = \lambda_1^4(\phi,\kappa) + \lambda_2^4(\phi,\kappa) = 2(1+\kappa^2)$.

Therefore, Expression (25) reduces to:

$$\text{mmse}(\text{snr},\kappa) = 1 - (1+\kappa^2)\text{snr} + O(\text{snr}^2), \tag{26}$$

and using the relationship (24), we obtain:

$$\text{I}(\sqrt{\text{snr}}\,V + N; V) = \text{snr} - \frac{1}{2}(1+\kappa^2)\text{snr}^2 + O(\text{snr}^3), \tag{27}$$

as we wanted to prove.   □

Remark 2. *Theorem 2 is an extension to improper signaling (and up to the second-order term) of the well-known result that states that at low* snr, *the proper Gaussian capacity behaves linearly with* snr.

Remark 3. *Equation (21) shows that at low* snr, *the behavior of the proposed family of improper constellations* $\mathcal{A}(\phi,\kappa,M)$ *only depends on the circularity coefficient* κ, *but neither on the phase* ϕ, *nor on the cardinality* M *of* V.

A more elaborate proof derived directly on the complex field gives us the MI third-order expansion around snr $= 0$. This is stated in the next theorem, which is proven in Appendix B.

Theorem 3. *Let* V *be a discrete improper random variable with circularity coefficient* $0 \leq \kappa \leq 1$ *taking values in* $\mathcal{A}(\phi,\kappa,M)$ *and constructed as in (16) with* X *being a square proper M-QAM constellation. Then, the mutual information between* Y *and* V *in (5) admits the following third-order expansion around* snr $= 0$:

$$\text{I}(\sqrt{\text{snr}}\,V + N; V) = \text{snr} - \frac{1}{2}(1+\kappa^2)\text{snr}^2 + \frac{1}{3}(1-\kappa^2)\text{snr}^3 + O(\text{snr}^4). \tag{28}$$

Proof. See Appendix B.   □

Remark 4. *The asymptotic expansion of the capacity (8) of the AWGN channel when transmitting improper Gaussian codewords is:*

$$C(\text{snr},\kappa) = \text{snr} - \frac{1}{2}(1+\kappa^2)\text{snr}^2 + \frac{1}{3}(1+3\kappa^2)\text{snr}^3 + \mathcal{O}(\text{snr}^4). \tag{29}$$

Comparing Equations (28) and (29), we see that at low snr, *the behavior of the mutual information for the discrete improper constellation* V *and the improper Gaussian capacity is the same up to the second-order term.*

4.3. Asymptotic Results at High SNR

The optimum $\hat{\phi}$ that maximizes $\text{I}(\sqrt{\text{snr}}\,V + N; V)$ depends in general on M, κ and the snr. The analytical derivation of $\hat{\phi}(\text{snr},\kappa,M)$ for all snr appears to be intractable in general. However, at high snr, the optimal phase depends only on κ and is given by the following theorem.

Theorem 4. *At high* snr, *the optimal* $\hat{\phi}$ *that maximizes* $\text{I}(\sqrt{\text{snr}}\,V + N; V)$ *over* $0 \leq \phi \leq \frac{\pi}{2}$ *is equal to* $\phi^*(\kappa)$ *in Expression (20) and is given by:*

$$\lim_{\text{snr}\to\infty} \hat{\phi}(\text{snr},\kappa,M) = \phi^*(\kappa) = \begin{cases} \pi/2, & 0 < \kappa \leq 0.5, \\ \phi^*, & 0.5 < \kappa \leq 1, \end{cases} \tag{30}$$

where ϕ^* *is the solution of the equation* $2\sin(\phi) - \cos(\phi) = 1/\kappa$.

Proof. In order to proceed, we use the fact [26] that for any discrete constellation, the mutual information, the minimum mean-squared error (mmse) and the symbol error probability all have an asymptotic behavior proportional to $Q\left(\sqrt{\text{snr}}d_{min}/2\right)$, where $Q(\cdot)$ denotes the Gaussian Q-function and d_{min} is the minimum Euclidean distance of the constellation.

In Proposition 1, we showed that the value of $\phi^*(\kappa)$ that maximizes the minimum distance of the family of constellations $\mathcal{A}(\phi, \kappa, M)$ is given by:

$$\phi^*(\kappa) = \arg \max_{\phi \in [0, \pi/2]} d_{min}(\Lambda(\phi, \kappa)) \tag{31}$$

where $\Lambda(\phi, \kappa)$ is the lattice that results from the $\mathbf{H}(\phi, \kappa)$-transformation of the integer lattice Z^2 (refer to (18)). From (18) and exploiting the fact that $0 \leq \phi \leq \pi/2$, it can be computed that:

$$d_{min}^2 = \min\{(1 - \kappa \cos(\phi)), 2(1 - \kappa \sin(\phi))\},$$

which can be obtained from the real representation given by (17) when the signal points of the original proper constellation belong to a regular square lattice. It is easy to check that $0 \leq \kappa \leq 0.5 \Rightarrow (1 - \kappa \cos(\phi)) \leq 2(1 - \kappa \sin(\phi))$. This, in turn, means that the optimal solution of (31) when $0 \leq \kappa \leq 0.5$ is $\phi^*(\kappa) = \pi/2$. On the other hand, when $\kappa > 0.5$, the optimal phase satisfies the nonlinear equation:

$$(1 - \kappa \cos(\phi)) = 2(1 - \kappa \sin(\phi)) \Leftrightarrow 2\sin(\phi) - \cos(\phi) = \frac{1}{\kappa}, \tag{32}$$

which proves the theorem. □

Remark 5. *For $\kappa = 1$ (maximally improper, rectilinear, constellations), the optimal phase in Equation (30) is $\phi^* \approx 0.9273$. This value results in a rectilinear constellation with equidistant points, which would be equivalent to a rotated M-ary pulse amplitude modulation (PAM). As an example, starting from a standard QPSK constellation and using $\kappa = 1$ and $\phi^* \approx 0.9273$ in (16), we obtain a unit-norm rectilinear constellation with equidistant signal points:*

$$V = \begin{bmatrix} 1.2 + 0.6i & 0.4 + 0.2i & -0.4 - 0.2i & -1.2 - 0.6i \end{bmatrix}.$$

4.4. Shaping Loss

Next, we compute the shaping loss defined as the ratio of the powers required to achieve the same transmission rate when using symbols that are either uniformly distributed in $\mathcal{A}(\phi, \kappa, M)$ (by assuming M to be an unbounded large number) or Gaussian distributed. The shaping loss in the improper case can also be interpreted as the power penalty due to using codewords uniformly distributed inside a multidimensional parallelogram rather than inside a multidimensional ellipsoid.

Theorem 5. *The shaping loss γ_s of the proposed family of improper constellations with respect to improper Gaussian signaling is:*

$$\gamma_s = \frac{\pi e}{6} \quad (1.53\, dB)$$

which is independent of κ and ϕ and coincides with the shaping loss of proper M-QAM constellations [25].

Proof. The shaping gain can be obtained as follows. Let us consider two memoryless complex sources. The first source generates discrete values $V = V_I + jV_Q$ uniformly distributed over the constellation

$\mathcal{A}(\phi, \kappa, M)$ where M is assumed to be a very large number. Under this assumption, its uniform distribution can be approximated by the probability density function:

$$f_V(v_I, v_Q) = \begin{cases} \frac{1}{A(\mathcal{P})} & \text{if } (v_I, v_Q) \in \mathcal{P}; \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathcal{P}$ denotes the parallelogram determined by the boundaries of the constellation $\mathcal{A}(\phi, \kappa, M)$ and $A(\mathcal{P})$ denotes its area. Based on the Jacobian of the transformation $\mathbf{H}(\phi, \kappa)$, this area can be computed as:

$$A(\mathcal{P}) = \det(\mathbf{H}(\phi, \kappa)) A(\mathcal{R})$$

where $A(\mathcal{R}) = 1$ denotes the area of the square $\mathcal{R} = [-0.5, 0.5]^2$. Note that the second-order moment of $\mathcal{R}$ is given by $P_\mathcal{R} = 1/6$ and so will be the second-order moment, P_V, of the parallelogram $\mathcal{P}$, since the transformation $\mathbf{H}(\phi, \kappa)$ keeps its input-output power invariant. Therefore, $A(\mathcal{P})$ can also be written as:

$$A(\mathcal{P}) = \det(\mathbf{H}(\phi, \kappa)) A(\mathcal{R}) = \sqrt{1 - \kappa^2} \, 6 \, P_V.$$

The second source, $G = G_I + jG_Q$, generates continuous values in $\mathbb{C}$, with the improper complex Gaussian distribution:

$$f_G(g_I, g_Q) = \mathcal{N}(0, \mathbf{K}_G),$$

where $\mathbf{K}_G = \text{diag}\left(\frac{P_G}{2}(1 + \kappa), \frac{P_G}{2}(1 - \kappa)\right)$ and P_G denotes the second-order moment of this Gaussian source.

Note that in order to generate the same number of bits per source symbol (i.e., the same information rate), the entropy of both sources should be the same. Since:

$$h_V(V_I, V_Q) = \int\int_{A(\mathcal{P})} \log A(\mathcal{P}) \frac{1}{A(\mathcal{P})} dv_I dv_Q = \log A(\mathcal{P}) = \log\left(\sqrt{1 - \kappa^2} \, 6 \, P_V\right)$$

and:

$$h_G(G_I, G_Q) = \log\left((2\pi e)\sqrt{\det(\mathbf{K}_G)}\right) = \log\left((2\pi e)\frac{P_G}{2}\sqrt{1 - \kappa^2}\right),$$

we find that γ_s is:

$$\gamma_s = \frac{P_V}{P_G} = \frac{\pi e}{6},$$

as we wanted to prove. □

5. Numerical Results

5.1. Mutual Information as a Function of ϕ

In this section, we present some numerical results that validate our findings. First, we compute numerically the MI of improper constellations generated according to (16) and study its behavior with respect to ϕ. Figure 4 plots the MI of an improper QPSK constellation with $\kappa = 0.4$ for increasing values of the snr. The MI always peaks at $\phi = \pi/2$, which is the optimal value according to Theorem 4. Note the different scale of the y-axis for each subplot. Figure 5 shows the results for an improper QPSK signal with a higher circularity coefficient ($\kappa = 0.95$). In this case, as the snr increases, the optimal phase decreases until reaching, for sufficiently high snr, the optimal ϕ^* in Theorem 4. For $\kappa = 0.95$, the optimal phase shown in Figure 5 is $\phi^* = 0.9538$.
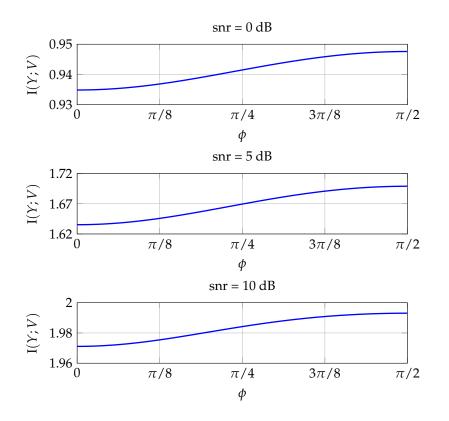
Figure 4. Mutual information of an improper QPSK constellation with $\kappa = 0.4$ as a function of ϕ, for three different signal-to-noise ratio (snr).
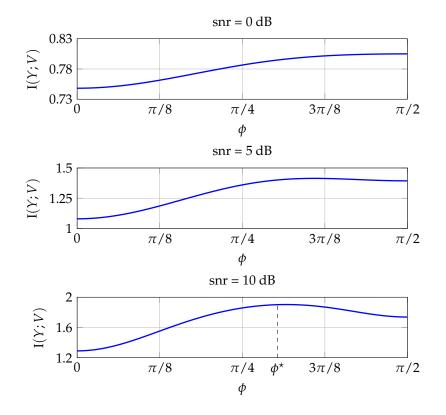

Figure 5. Mutual information of an improper QPSK constellation with $\kappa = 0.95$ as a function of ϕ, for three different snr. The optimal phase is $\phi^* = 0.9538$.

Finally, Figure 6 shows the value of ϕ that maximizes $I(Y; V)$ for different values of κ for an improper QPSK constellation. In the interval $0 \leq \kappa \leq 0.5$, the optimum value is $\pi/2$ regardless of the snr. When $\kappa > 0.5$, $\pi/2$ is still the optimum value for the low-snr regime, whereas at high snr the optimum value is $\phi_{opt} = \phi^*$ given by Equation (30), which depends only on κ. The curves in Figure 6 for $\kappa > 0.5$ are monotonically non-increasing. However, for high values of both κ and snr, the estimated curves present some small fluctuations, which are due to numerical errors in the simulation. These errors can be explained because the mutual information is computed by numerical integration over a finite region for a discrete grid of equally-spaced phases.
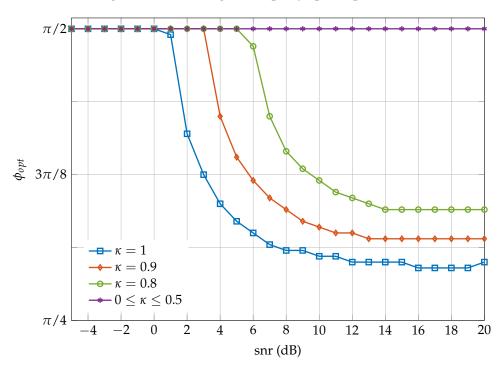

Figure 6. ϕ_{opt} versus snr in dBs for different values of κ.

5.2. Capacity Curves

Figure 7 plots the mutual information achieved by improper M-QAM signals with $\kappa = 0.95$ when the optimum $\phi^* = 0.9538$ at high snr is used (black solid lines), in comparison to the case when we use fixed $\phi = \pi/2$ (dashed blue lines) or $\phi = 0$ (red dashed lines). We consider three increasing values of the constellation size: $M = 16$, 64 and 144. The constellation with $\phi = 0$ is a rectangular constellation with uncorrelated real and imaginary parts, but with a power imbalance to achieve the desired circularity coefficient. We observe that this naive family $\mathcal{A}(\phi = 0, \kappa, M)$ of constellations attains lower MI than using the optimum $\phi^* = 0.9538$ given by Theorem 4. This rectangular constellation is also outperformed by the constellation generated with $\phi = \pi/2$.

The figure also depicts the Gaussian capacity curve with improper signaling and the envelope of the capacity curves with M-QAM constellations for M sufficiently large. The distance between both curves is the shaping gain, which as proven in Theorem 5 is 1.53 dB. It should be mentioned that, since the shaping loss does not depend on ϕ, the mutual information envelope $I^*(\text{snr}, \kappa)$ obtained from the set of constellations with $\phi = 0$, $\phi = \pi/2$ or with the optimum ϕ^* are all the same as M increases. However, for a given snr and κ, the minimum M that makes $I(Y; V) \approx I^*(\text{snr}, \kappa)$, will be smaller when $\phi = \phi^*(\kappa)$ than when $\phi = 0$. That is, $\phi = 0$ needs a larger constellation expansion ratio [30] to achieve the same rate.
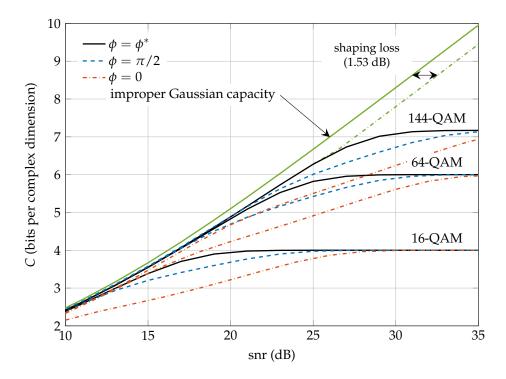
Figure 7. Mutual information vs. snr when $\kappa = 0.95$ for $\phi = \pi/2$ (blue dashed lines), $\phi = 0$ (red dashed lines) and $\phi^* = 0.9538$ (black solid lines) for improper M-QAM signals. The capacity of the improper Additive White Gaussian Noise (AWGN) channel using Gaussian codewords and the asymptotic envelope of the curves are also depicted.

The situation is different for maximally improper signals ($\kappa = 1$) (i.e., rectilinear constellations). In this case, using $\phi = \pi/2$ makes some signal points of the original proper constellation collapse into the same symbol, thus losing rate at high snr. This effect becomes evident in Figure 8, which plots the mutual information achieved by the proposed improper constellation with $\kappa = 1$, and $M = 4$ and $M = 16$ signal points, for $\phi = \pi/2$ and for the optimal $\phi^* = 0.9273$ at high snr. From a practical point of view, we advocate the use of the optimal ϕ^* given in (30), which maximizes the minimum distance between constellation points.

Finally, Figure 9 depicts the capacity curves for proper and improper (with $\kappa = 0.95$) M-QAM discrete constellations. For the improper M-QAM, the optimal $\phi^* = 0.9538$ at high snr is used. The proper and improper Gaussian capacity is also depicted for comparison.

From these results, the following conclusions are drawn:

1. For low-rate transmissions ($M = 4$ or $M = 16$), there is no rate loss with respect to the improper Gaussian capacity when using coded symbols uniformly drawn from the proposed family of improper constellations. Therefore, like in the proper case, the proposed improper QPSK and 16-QAM constellations are optimal in the wideband regime [32].

2. For high-rate transmissions ($M > 16$), the proposed family of improper constellations incurs a shaping loss of up to 1.53 dB. As in the case of proper M-QAM signaling, this snr loss can be reduced by designing codes with some shaping mechanism to make the coded symbols in $\mathcal{A}(\phi, \kappa, M)$ more Gaussian-like distributed.

3. The optimal ϕ^*, derived at high snr, can be safely used without any observable degradation in $\Delta(\mathrm{snr}, \kappa)$, independently of the value of M and snr.

Therefore, we arrive at the main conclusion of the paper, namely that the proposed family of improper constellations is a good candidate to be used in communication systems with improper signaling. In fact, it can be viewed as the improper counterpart of the standard proper M-QAM constellations.
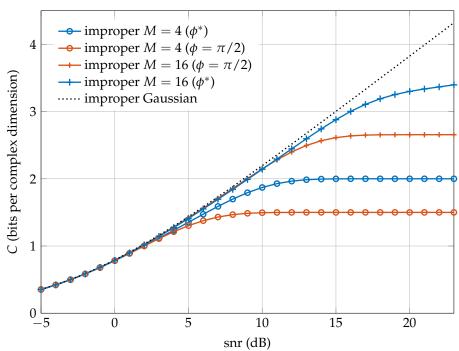
Figure 8. Mutual information vs. snr when $\kappa = 1$ (maximally improper) for $\phi = \pi/2$ and ϕ^* for improper constellations with $M = 4$ and $M = 16$ signal points. The capacity of the improper AWGN channel using Gaussian codewords is also depicted.
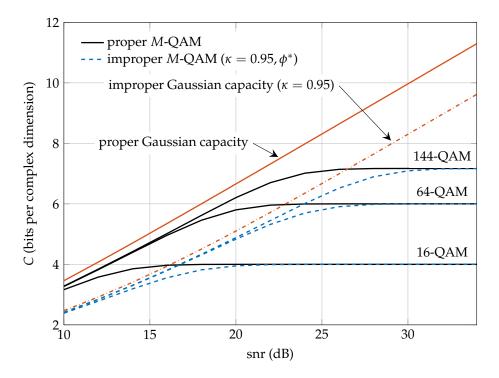

Figure 9. Capacity comparison between proper and improper (with $\kappa = 0.95$) M-QAM discrete constellations. For the improper M-QAM, the optimal $\phi^* = 0.9538$ at high snr is used. The proper and improper Gaussian capacity is also depicted.

6. Conclusions

In this paper, we have proposed a family of improper constellations constructed by widely linear processing of proper M-QAM signals. We have shown that, when the coded channel symbols are uniformly drawn from these constellations, the achievable transmission rates are close to the ones dictated by the improper Gaussian capacity. In particular, except for a shaping loss of 1.53 dB encountered at high snr, there is no rate loss of the proposed improper discrete constellations with respect to the improper Shannon capacity. In practice, we can select the optimal phase at high snr, $\phi^*(\kappa)$, obtained by minimizing the BER of the uncoded system without significant loss in performance at lower snr. Observe that this fact facilitates the practical implementation of these constellations since $\phi^*(\kappa)$ does not depend on the snr and M.

These results are obtained by analyzing at the high- and low-snr regimes the family of mutual information, $\{I(Y;V) : V \in \mathcal{A}(\phi^*(\kappa, \mathrm{snr}, M), \kappa, M), \text{ for all } M \text{ and snr values}\}$, and by obtaining the shaping loss computed as the ratio of the normalized second moments of a multidimensional parallelogram and a multidimensional ellipsoid.

We conclude that the proposed family of improper constellations is a good candidate to be used in communication systems when improper signaling is required over the AWGN channel. In fact, it behaves like the standard family of proper M-QAM constellations and can be viewed as its improper counterpart. In perspective, these results extend the analysis of improper signaling in interference-limited scenarios to non-Gaussian discrete constellations with equiprobable symbols.

Author Contributions: Ignacio Santamaria and Christian Lameiro conceived of the research question. Pedro M. Crespo, Ignacio Santamaria and Christian Lameiro proved the main results. Pedro M. Crespo, Ignacio Santamaria, Christian Lameiro and Peter J. Schreier wrote the paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AWGN	Additive white Gaussian noise
BPSK	Binary phase shift keying
GMSK	Gaussian minimum shift keying
MI	Mutual information
MMSE	Minimum mean-squared error
PAPR	Peak-to-average-power ratio
PU	Primary user
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase shift keying
SNR	Signal-to-noise ratio
SU	Secondary user
UCR	Underlay cognitive radio

Appendix A. Derivation of the Capacity for Improper Gaussian Signals by an Elliptic Packing Argument

We want to reliably transmit 2^{NR} uniformly-distributed messages across an AWGN channel. To that end, we associate with each message $m \in \{1, 2, \ldots, 2^{NR}\}$ a complex point $\mathbf{x}(m)$ belonging to some N-dimensional regular lattice with minimum Euclidean distance d_{min} and enclosed in some region $\mathcal{R} \subset \mathbb{C}^N$. This region is such that the induced empirical distribution on each of the N complex components satisfies the improper circularity constraint:

$$\frac{|E[X_{k,I}^2] - E[X_{k,Q}^2]|}{E[X_{k,I}^2] + E[X_{k,Q}^2]} = \kappa, \ \forall k = 1, 2, \ldots, N, \tag{A1}$$

where $X_{k,I}$ and $X_{k,Q}$ denote the real and complex components of symbol k. Without any loss of generality, we also assume $E[X_{k,I} X_{k,Q}] = 0$

The total average energy of such constellation for large N can be approximated as:

$$P_{\mathrm{T}} = \frac{1}{2^{nR}} \sum_{k=1}^{2^{nR}} \|\mathbf{x}(k)\|^2 \approx \int_{\mathcal{R}} \|\mathbf{r}\|^2 \frac{d\mathbf{r}}{\mathrm{Vol}(\mathcal{R})},$$

The region $\mathcal{R}$ that minimizes P_{T} (i.e., the second moment of $\mathcal{R}$), under the constraint (A1), is the N-dimensional ellipsoid:

$$\mathcal{E}_X = \left\{ \mathbf{r} = (r_{1,I}, r_{1,Q}, \ldots, r_{N,I}, r_{N,Q})^T \in \mathbb{R}^{2N} : \frac{1}{\sigma_I^2} \|\mathbf{r}_I\|^2 + \frac{1}{\sigma_Q^2} \|\mathbf{r}_Q\|^2 = 2N(1+\epsilon) \right\}. \tag{A2}$$

where $E[X_{k,I}^2] = \sigma_I^2$, $E[X_{k,Q}^2] = \sigma_Q^2$ for all $k = 1, 2, \ldots, N$ and $\sigma_Q^2 = \sigma_I^2 \frac{1-\kappa}{1+\kappa}$.

By the law of large numbers, the N-dimensional received complex vector $\mathbf{y} = \mathbf{x} + \mathbf{n}$ will lie, with high probability, within the ellipsoid:

$$\mathcal{E}_Y = \left\{ \mathbf{r} = (r_{1,I}, r_{1,Q}, \ldots, r_{N,I}, r_{N,Q})^T \in \mathbb{R}^{2N} : \frac{1}{\sigma_I^2 + N_0/2} \|\mathbf{r}_I\|^2 + \frac{1}{\sigma_Q^2 + N_0/2} \|\mathbf{r}_Q\|^2 = 2N(1+\epsilon) \right\}, \tag{A3}$$

and hence, without loss of generality, we only need to focus on what happens inside this $\mathcal{E}_Y$ ellipsoid. On the other hand,

$$\lim_{N \to \infty} \frac{1}{N} \sum |N_k|^2 = N_0,$$

again by the law of large numbers. That is, the noise realizations of the channel are with high probability, inside a sphere S_N of radius $\sqrt{N N_0}$.

Therefore, when the codeword $\mathbf{x}(m)$ is sent, the received vector at the output of the channel will be contained with high probability in the sphere $\mathcal{S}_N$ centered at $\mathbf{x}(m)$. If we assign as the decision region for message m the region $\mathcal{D}_m = \mathbf{x}(m) + \mathcal{S}_N$, an error will be produced only if the received vector falls outside $\mathcal{D}_m$, which has low probability. By choosing the other decision regions in the same way, i.e., $\mathcal{D}_l = \mathbf{x}(l) + \mathcal{S}_N$, the overall probability of error will be negligible, as long as there is no intersection among the spheres. Therefore, the maximum number of possible messages, i.e., 2^{nC} is given by the ratio:

$$M = \frac{\mathrm{Vol}(\mathcal{E}_Y)}{\mathrm{Vol}(S_N)}. \tag{A4}$$

The volumes of the $\mathcal{E}_Y$ ellipsoid and the S_N sphere are [33]:

$$\mathrm{Vol}(\mathcal{E}_Y) = \frac{(2N\pi)^N}{N!} (\sigma_Q^2 \sigma_I^2)^{\frac{N}{2}} \quad \text{and} \quad \mathrm{Vol}(\mathcal{S}_N) = \frac{(2N\pi)^N}{N!} \left(\frac{N_0}{2}\right)^N, \tag{A5}$$

respectively.

Substituting (A5) into (A4) yields:

$$M = \frac{\text{Vol}(\mathcal{E}_Y)}{\text{Vol}(\mathcal{S}_N)} = \frac{(P(1+\kappa)+N_0)^{\frac{N}{2}}(P(1-\kappa)+N_0)^{\frac{N}{2}}}{N_0^N} = 2^{\frac{N}{2}\log((1+\text{snr})^2-\kappa^2\text{snr}^2)},$$

where $P = \sigma_I^2 + \sigma_Q^2$ is the average power per complex dimension. Therefore,

$$C = \frac{\log M}{N} = \frac{1}{2}\log\{(1+\text{snr})^2 - \kappa^2\text{snr}^2\} = \log(1+\text{snr}) + \frac{1}{2}\log\left(1 - \frac{\kappa^2}{(1+\text{snr}^{-1})^2}\right),$$

which is the capacity of the AWGN channel with improper Gaussian codewords in Equation (8).

Appendix B. Proof of Theorem 3

For the proof, we follow the technique used in [31,34]. Let us denote the standard complex Gaussian density as $\psi(y) = \frac{1}{\pi}e^{-|y|^2}$, and let us define the functions:

$$g_0(y;\text{snr}) = E\left[\psi(y - \sqrt{\text{snr}}V)\right], \tag{A6}$$

$$g_1(y;\text{snr}) = E\left[V\psi(y - \sqrt{\text{snr}}V)\right], \tag{A7}$$

where the expectation is taken over V. The conditional mean estimate can be expressed as [31]:

$$E[V|Y = y] = \frac{g_1(y;\text{snr})}{g_0(y;\text{snr})}, \tag{A8}$$

and (taking into account that $E[|V|^2] = 1$), the mmse can be calculated as [34]:

$$\text{mmse}(\text{snr}) = 1 - \int \frac{|g_1(y;\text{snr})|^2}{g_0(y;\text{snr})}dy, \tag{A9}$$

where the integral is calculated over the complex field.

When $\text{snr} \to 0$, the Gaussian complex density can be expanded as:

$$\psi(y - \sqrt{\text{snr}}V) = \psi(y)\left[1 + (y^*V + yV^*)\sqrt{\text{snr}} + \frac{1}{2}\left((y^*V)^2 + (yV^*)^2 + 2|V|^2(|y^2|-1)\right)\text{snr} + \mathcal{O}(\text{snr}^{3/2})\right]. \tag{A10}$$

Using this Taylor series expansion, the functions $g_0(y;\text{snr})$ and $g_1(y;\text{snr})$ defined in (A6) and (A7), respectively, can be computed in terms of the moments of V as:

$$g_0(y;\text{snr}) = \psi(y)\left[1 + \frac{1}{2}\left((y^*)^2E[V^2] + y^2E[V^2]^* + 2E[|V|^2](|y^2|-1)\right)\text{snr}\right], \tag{A11}$$

$$g_1(y;\text{snr}) = \psi(y)\left[\left(y^*E[V^2] + yE[|V|^2]\right)\sqrt{\text{snr}} + \frac{1}{2}(y^*)^2E[V^3]\text{snr}\right], \tag{A12}$$

where in (A12), we have used $E[|V|^2V^*] = 0$. We also have the expansion:

$$\frac{1}{g_0(y;\text{snr})} = \frac{1}{\psi(y)}\left[1 - \frac{1}{2}\left((y^*)^2E[V^2] + y^2E[V^2]^* + 2E[|V|^2](|y^2|-1)\right)\text{snr}\right]. \tag{A13}$$

We are assuming $|E[V^2]|^2 = 1$, and in addition, it is easy to check that $E[V^3] = 0$. Therefore, substituting these values in (A12) and (A13), we get:

$$g_1(y;\text{snr}) = \psi(y)\left(y^*E[V^2] + y\right)\sqrt{\text{snr}} \tag{A14}$$

and:

$$\frac{1}{g_0(y;\mathrm{snr})} = \frac{1}{\psi(y)}\left[1 - \frac{1}{2}\left((y^*)^2 E[V^2] + y^2 E[V^2]^* + 2(|y^2| - 1)\right)\mathrm{snr}\right], \tag{A15}$$

which only depend on $E[V^2]$.

Substituting (A14) and (A15) in (A9) and evaluating the resulting integral over y, we obtain the result:

$$\mathrm{mmse}(\mathrm{snr},\kappa) = 1 - (1+\kappa^2)\mathrm{snr} + (1-\kappa^2)\mathrm{snr}^2 + \mathcal{O}(\mathrm{snr}^3). \tag{A16}$$

Finally, by integrating Expression (A16), we obtain the expansion around $\mathrm{snr} = 0$ of the mutual information as:

$$\mathrm{I}(\mathrm{snr},\kappa) = \mathrm{snr} - \frac{(1+\kappa^2)}{2}\mathrm{snr}^2 + \frac{(1-\kappa^2)}{3}\mathrm{snr}^3 + \mathcal{O}(\mathrm{snr}^4). \tag{A17}$$

which completes the proof.

References

1. Cover, T.M.; Thomas, J.A. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*; John Wiley & Sons: New York, NY, USA, 2006.
2. Cadambe, V.; Jafar, S.; Wang, C. Interference alignment with asymmetric complex signaling—Settling the Høst-Madsen-Nosratinia conjecture. *IEEE Trans. Inf. Theory* **2010**, *56*, 4552–4565.
3. Ho, Z.; Jorswieck, E. Improper Gaussian signaling on the two-user SISO interference channel. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 3194–3203.
4. Zeng, Y.; Yetis, C.; Gunawan, E.; Guan, Y.; Zhang, R. Transmit optimization with improper Gaussian signaling for interference channels. *IEEE Trans. Signal Process.* **2013**, *61*, 2899–2913.
5. Lagen, S.; Agustin, A.; Vidal, J. Coexisting linear and widely linear transceivers in the MIMO interference channel. *IEEE Trans. Signal Process.* **2016**, *64*, 652–664.
6. Lameiro, C.; Santamaria, I.; Schreier, P.J. Rate region boundary of the SISO Z-interference channel with improper signaling. *IEEE Trans. Commun.* **2017**, *65*, 1022–1034.
7. Lameiro, C.; Santamaria, I.; Schreier, P.J. Benefits of improper signaling for underlay cognitive radio. *IEEE Wirel. Commun. Lett.* **2015**, *4*, 22–25.
8. Amin, O.; Abediseid, W.; Alouini, M.S. Underlay cognitive radio systems with improper Gaussian signaling: Outage performance analysis. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 4875–4887.
9. Schreier, P.J.; Scharf, L.L. *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*; Cambridge University Press: Cambridge, UK, 2010.
10. Gerstacker, W.; Schober, R.; Lampe, A. Receivers with widely linear processing for frequency-selective channels. *IEEE Trans. Commun.* **2003**, *51*, 1512–1523.
11. Buzzi, S.; Lops, M.; Sardellitti, S. Widely linear reception strategies for layered space-time wireless communications. *IEEE Trans. Signal Process.* **2006**, *54*, 2252–2262.
12. Chevalier, P.; Pipon, F. New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK, and GMSK signals corrupted by noncircular interferences-application to SAIC. *IEEE Trans. Signal Process.* **2006**, *54*, 870–883.
13. Schober, R.; Gerstacker, W.; Lampe, L.J. Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA. *IEEE Trans. Signal Process.* **2004**, *52*, 746–756.
14. Lameiro, C.; Santamaria, I. Degrees-of-freedom for the 4-user SISO interference channel with improper signaling. In Proceedings of the 2013 IEEE International Conference on Communications (ICC), Budapest, Hungary, 9–13 June 2013.
15. Nguyen, H.; Zhang, R.; Sun, S. Improper signaling for symbol error rate minimization in *K*-user interference channel. *IEEE Trans. Commun.* **2015**, *63*, 857–869.
16. Kurniawan, E.; Sun, S. Improper Gaussian signaling scheme for the Z-interference channel. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 3912–3923.
17. Lagen, S.; Agustin, A.; Vidal, J. Improper Gaussian signaling for the Z-interference channel. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 1145–1149.

18. Lagen, S.; Agustin, A.; Vidal, J. On the superiority of improper Gaussian signaling in wireless interference MIMO scenarios. *IEEE Trans. Commun.* **2016**, *64*, 3350–3368.

19. Hellings, C.; Utschick, W. Performance gains due to improper signals in MIMO broadcast channels with widely linear receivers. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vancouver, BC, Canada, 26–31 May 2013; pp. 4379–4383.

20. Zeng, Y.; Zhang, R.; Gunawan, E.; Guan, Y.L. Optimized transmission with improper Gaussian signaling in the K-user MISO interference channel. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 6303–6313.

21. Amin, O.; Abediseid, W.; Alouini, M.S. Overlay spectrum sharing using improper Gaussian signaling. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 50–62.

22. Gaafar, M.; Khafagy, M.G.; Amin, O.; Alouini, M.S. Improper Gaussian signaling in full-duplex relay channels with residual self-interference. In Proceedings of the IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–7.

23. Gaafar, M.; Amin, O.; Ikhlef, A.; Chaaban, A.; Alouini, M.S. On alternate relaying with improper Gaussian signaling. *IEEE Commun. Lett.* **2016**, *20*, 1683–1686.

24. Kariminezhad, A.; Chaaban, A.; Sezgin, A. Improper signaling and symbol extensions: How far can we go with Gaussian P2P codebooks in the interfering MAC with TIN? *arXiv* **2017**, arXiv:1607.01995.

25. Forney, G.; Ungerboek, G. Modulation and coding for linear Gaussian channels. *IEEE Trans. Inf. Theory* **1998**, *44*, 2384–2415.

26. Alvarado, A.; Brännström, F.; Agrell, E.; Koch, T. High-SNR asymptotics of mutual information for discrete constellations with applications to BICM. *IEEE Trans. Inf. Theory* **2014**, *60*, 1061–1076.

27. Neeser, F.D.; Massey, J.L. Proper complex random processes with applications to information theory. *IEEE Trans. Inf. Theory* **1993**, *39*, 1293–1302.

28. Costa, M. On the Gaussian interference channel. *IEEE Trans. Inf. Theory* **1985**, *31*, 607–615.

29. Tse, D.; Viswanath, P. *Fundamentals of Wireless Communications*; Cambridge University Press: Cambridge, UK, 2005.

30. Fisher, R.F.H. *Precoding and Signal Shaping for Digital Transmissions*; John Wiley & Sons: New York, NY, USA, 2002.

31. Guo, D.; Shamai, S.; Verdu, S. Mutual information and minimum mean-square error in Gaussian channels. *IEEE Trans. Inf. Theory* **2005**, *51*, 1261–1282.

32. Verdu, S. Spectral efficiency in the wideband regime. *IEEE Trans. Inf. Theory* **2002**, *48*, 1319–1343.

33. Sykora, S. Volume Integrals over n-Dimensional Ellipsoids. *Stan's Library* **2005**, *1*, 101–106.

34. Guo, D.; Wu, Y.; Shitz, S.S.; Verdu, S. Estimation in Gaussian noise: Properties of the minimum mean-square error. *IEEE Trans. Inf. Theory* **2011**, *57*, 2371–2385.