

Article

# Unitary Space-Time Modulation Based on Coherent Codes

Daniel Calabuig 

Instituto de Telecomunicaciones y Aplicaciones Multimedia (iTEAM), Universitat Politècnica de València, 46022 Valencia, Spain; dacaso@iteam.upv.es

**Abstract:** This paper analyzes the relationship between pilot symbol-assisted modulation (PSAM) and unitary space-time modulation (USTM). In particular, we present a map that transforms any PSAM into a USTM and vice versa. USTMs are known to be capacity-achieving. However, most of the proposed USTM construction methods in the literature are computationally expensive, and the resulting constellations do not have a known structure that could simplify their decoding. Using the relationship between PSAM and USTM, and inspired by a graphical representation of these constellations used in this paper, we propose new USTM construction methods, which ensure that the USTM has a good performance compared to the corresponding PSAM, and a feasible construction and decoding, even for high data rates.

**Keywords:** constellation design; noncoherent multiple input–multiple-output (MIMO) communication; pilot symbol-assisted modulation

## 1. Introduction

Most of the current communication systems are based on coherent reception where channel state information (CSI) should be estimated for equalization, demodulation, etc. The CSI can be obtained via training by inserting pilots in the data signals. These schemes, however, increase the signaling overhead, which can be counterproductive in systems with many antennas or fast movement of either side of the communication link [1].

This drawback increased the interest in schemes that do not require full CSI at the transmitter and the receiver. An information-theoretic analysis of multi-antenna Rayleigh block-fading channels was performed in [2]. Motivated by this work, a communication method using unitary space-time modulation (USTM) and noncoherent reception was proposed in [3]. In this method, assuming that the receiver has  $N$  antennas, the transmitter has  $M \leq N$  antennas, and the channel block length is  $T \geq M + N$  channel uses, the transmitted signals, viewed as matrices with  $T \times M$  elements, form a unitary matrix, i.e., one with orthonormal columns. This signal structure was shown to be capacity-achieving for high signal-to-noise ratio (SNR) or  $T \gg M$  [3]. USTM constellations can be constructed minimizing the pairwise error probability of the constellation elements, although this is, in general, a difficult task. Other construction methods were inspired by the geometrical interpretation, given in [4], of the capacity expression, i.e., sphere packing in the Grassmann manifold  $\mathcal{G}(T, M)$ : the set of all  $M$ -dimensional subspaces of  $\mathbb{C}^T$ . In particular, the columns of the USTM matrices are viewed as a basis that spans an  $M$ -dimensional subspace. At the receiver, the channel modifies the basis but keeps the subspace unchanged. Therefore, information should be encoded into subspaces and not into the particular basis. This idea is used in [5] to define a superposition coding of several USTM constellations in Grassmann manifolds.

These results motivated the USTM constellation design as sphere packing in the Grassmann manifold by using either numerical optimization tools [3,6–9] or algebraic constructions [10–12]. The main advantage of the first approach is that it does not restrict the constellation to have a specific structure. However, the direct optimization is computationally expensive, and the lack of structure implies that all the USTM matrices have to be



**Citation:** Calabuig, D. Unitary Space-Time Modulation Based on Coherent Codes. *Sensors* **2022**, *22*, 9049. <https://doi.org/10.3390/s22239049>

Academic Editor: Francisco J. González-Castaño

Received: 25 October 2022

Accepted: 16 November 2022

Published: 22 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

stored and that the optimum receiver has to test all signals. The second approach equips the constellations with a certain structure, although with either poor performance [10], intractable decoding [11], or intractable constellation construction for a high number of degrees of freedom [12]. A different approach to constructing USTM constellations is to map coherent codes, i.e., those designed for coherent reception, into the Grassmann manifold. An example in which the exponential map is used can be found in [13]. This map has the drawback of not being one-to-one, requiring an appropriate scaling of the coherent codes which is, in general, not straightforward. Another example can be found in [14]. In this example, the coherent codes are mapped in the faces of a hypercube, which are projected into a hypersphere. Although the final USTM has a good performance, this technique can only be used with one transmit antenna. A similar approach was proposed in [15]. In this case, the authors use a different hypercube, but the technique can only be used with, again, one transmit antenna.

Due to the inability of producing sphere packings in the Grassmann manifold with efficient decoding and good performance, various works proposed the combination of training pilots and coherent codes, known as pilot symbol-assisted modulation (PSAM), as a meaningful alternative [4,16–19]. However, in this case, the resulting matrices are no longer unitary, and hence, the constellations are not capacity-achieving, although, in some cases, they can be shown to be similar to other USTMs [20]. This fact suggests that USTMs and PSAMs might not be as different as originally thought. In fact, the USTM constellation in [21] was shown to be more easily interpreted and analyzed as a PSAM [20].

In this paper, we use a map that transforms any PSAM into a USTM and vice versa, and hence generalize the observation done in [20] for one particular case. To obtain this map, we propose a new definition of the space-time matrices that compose a PSAM. Moreover, this definition enables a graphical representation of both USTMs and PSAMs that will draw insight into their relationship. Inspired by this representation and the PSAM matrix definition, we propose a new USTM constellation construction method that outperforms and inherits the structure of a coherent code. This construction method ensures that the USTM has a good performance compared to the coherent code, and a feasible construction and decoding, even for high data rates.

The rest of the paper is organized as follows. In Section 2, we present the system model and introduce both the graphical representation and the new definition of PSAM matrices. In Section 3, we present the map to transform PSAMs into USTMs and vice versa. In Section 4, we present a first USTM construction method and use the graphical representation to show the limitations of this method. In order to overcome these limitations, we propose a second USTM construction method in Section 5. In Section 6, we compare the performance of particular constellations constructed with the previous methods. Finally, main conclusions are drawn in Section 7.

## 2. System Model and Constellation Structure

Consider a wireless communication system with  $M$  transmit and  $N$  receive antennas, and a block-fading channel with a coherence interval of  $T$  channel uses. Let  $\mathbf{X} \in \mathbb{C}^{T \times M}$  and  $\mathbf{Y} \in \mathbb{C}^{T \times N}$  be the transmitted and received signals during  $T$  constant-fading channel uses, respectively, where the transmitted signal satisfies the power constraint

$$\mathbb{E}[\|\mathbf{X}\|_{\text{F}}^2] = M, \quad (1)$$

and  $\|\mathbf{X}\|_{\text{F}} = \sqrt{\text{tr}(\mathbf{X}^* \mathbf{X})}$  is the Frobenius norm. The relationship between  $\mathbf{X}$  and  $\mathbf{Y}$  is given by

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \sqrt{\frac{M}{\rho T}} \mathbf{Z}, \quad (2)$$

where  $\mathbf{H} \in \mathbb{C}^{M \times N}$  is the channel matrix whose entries are i.i.d. drawn from the standard complex Gaussian distribution  $\mathcal{CN}(0, 1)$ ,  $\mathbf{Z} \in \mathbb{C}^{T \times N}$  is the additive noise whose entries are also i.i.d. drawn from  $\mathcal{CN}(0, 1)$ , and  $\rho$  is the signal-to-noise ratio (SNR).

The transmitted signals are selected from a constellation of matrices with  $L = 2^{RT}$  elements, where  $R$  is the transmission rate. The constellation  $\mathcal{C}^U = \{\mathbf{X}_i^U\}_{i=1}^L$  is a USTM if the elements are  $T \times M$  unitary matrices, i.e.,  $\mathbf{X}_i^{U*} \mathbf{X}_i^U = \mathbf{I}_M$ , where  $\mathbf{I}_M$  is the  $M \times M$  identity matrix, and they span different subspaces, i.e.,  $\mathcal{S}(\mathbf{X}_i^U) \neq \mathcal{S}(\mathbf{X}_j^U)$  for all  $i \neq j$ , where  $\mathcal{S}(\mathbf{X}) = \{\mathbf{X}\mathbf{x}, \mathbf{x} \in \mathbb{C}^M\}$  is the subspace spanned by the columns of  $\mathbf{X}$ . In the case of a PSAM, a certain quantity of channel uses are reserved to carry the pilot signals. If optimization over the training and data powers is allowed, the optimal number of training channel uses is  $M$  [1]. In this case,  $T - M$  channel uses are used to transmit matrices from a coherent code  $\mathcal{A} = \{\mathbf{A}_i\}_{i=1}^L$ , where  $\mathbf{A}_i \in \mathbb{C}^{T-M \times M}$ . Therefore, the PSAM can be constructed concatenating the training and data signals in  $\mathcal{A}$ . However, it is also possible to allow these signals to spread throughout all the coherence intervals by means of a unitary matrix. In particular, let  $[\mathbf{B}_1 \ \mathbf{B}_2]$  be a  $T \times T$  unitary matrix, where  $\mathbf{B}_1$  and  $\mathbf{B}_2$  are the matrices with the first  $M$  columns and the last  $T - M$  columns, respectively. The constellation  $\mathcal{C}^P = \{\mathbf{X}_i^P\}_{i=1}^L$ , where

$$\mathbf{X}_i^P = \mathbf{B}_1 \mathbf{P} + \mathbf{B}_2 \mathbf{A}_i, \quad i = 1, \dots, L, \quad (3)$$

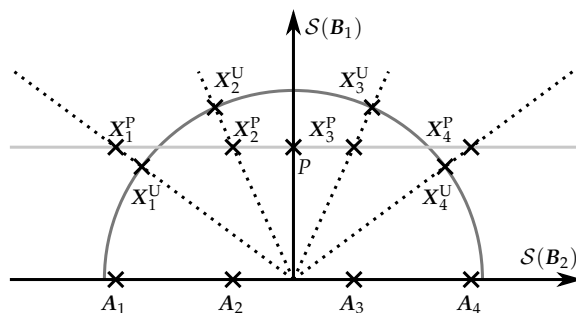
and  $\mathbf{P} \in \mathbb{C}^{M \times M}$  is a pilot matrix, is a PSAM. In order to maximize the achievable rates,  $\mathbf{P}$  should be a scaled unitary matrix [1]. Note that, in order to satisfy the power constraint in (1), assuming that all the matrices in the constellation are equiprobable, the coherent code and the pilot matrix must satisfy  $\frac{1}{L} \sum_{i=1}^L \text{tr}(\mathbf{P}^* \mathbf{P} + \mathbf{A}_i^* \mathbf{A}_i) = M$ , and hence,  $\|\mathbf{P}\|_F^2 < M$ . For any  $\mathbf{A} \in \mathcal{A}$  and  $\mathbf{X} = \mathbf{B}_1 \mathbf{P} + \mathbf{B}_2 \mathbf{A}$ , the received pilot and data signals can be obtained from  $\mathbf{Y}$  as follows:

$$\mathbf{B}_1^* \mathbf{Y} = \mathbf{P} \mathbf{H} + \sqrt{\frac{M}{\rho T}} \mathbf{B}_1^* \mathbf{Z}, \quad (4)$$

$$\mathbf{B}_2^* \mathbf{Y} = \mathbf{A} \mathbf{H} + \sqrt{\frac{M}{\rho T}} \mathbf{B}_2^* \mathbf{Z}. \quad (5)$$

The matrix  $\mathbf{B}_1^* \mathbf{Y}$  can then be used for channel estimation, and the matrix  $\mathbf{B}_2^* \mathbf{Y}$  for data reception.

If  $[\mathbf{B}_1 \ \mathbf{B}_2] = \mathbf{I}_T$ , the PSAM is constructed by concatenating  $\mathbf{P}$  and the matrices in  $\mathcal{A}$ . Although these are the most extensively used values for  $\mathbf{B}_1$  and  $\mathbf{B}_2$ , the (more general) definition of PSAM matrices in (3) provides a new interpretation of a PSAM, which will be used in Section 5 to design new constellations that combine coherent codes and USTMs. In addition to this, the orthogonal matrices  $\mathbf{B}_1$  and  $\mathbf{B}_2$  enable a two-dimensional graphical representation of the different constellations, as shown in Figure 1. In the figure, the  $(T - M)$ -dimensional subspace spanned by  $\mathbf{B}_2$ ,  $\mathcal{S}(\mathbf{B}_2)$ , is represented in the horizontal axis, and the  $M$ -dimensional subspace spanned by  $\mathbf{B}_1$ ,  $\mathcal{S}(\mathbf{B}_1)$ , in the vertical axis. The PSAM matrices can be viewed as shifts of the coherent code signals through the subspace spanned by  $\mathbf{B}_1$ . As a result, these matrices have, in general, different energy, although this will ultimately depend on the coherent code. However, the USTM matrices have always the same energy, independently of how the USTM was constructed. In Figure 1, we highlight an interesting case in which the elements of a PSAM and a USTM span the same subspaces. In the following section, we will show that this case is not exceptional, i.e., it is always possible to find this type of PSAM-USTM pairs.



**Figure 1.** Graphical representation of a coherent code  $\mathcal{A} = \{A_i\}_{i=1}^4$ , a PSAM  $\mathcal{C}^P = \{X_i^P\}_{i=1}^4$ , and a USTM  $\mathcal{C}^U = \{X_i^U\}_{i=1}^4$ . The signals of both the PSAM and the USTM have the same mean power and span the same subspaces, which are drawn with dotted lines.

### 3. PSAM-USTM Equivalence

In this section, we will show that there is certain equivalence between PSAM and USTM constellations in terms of the spanned subspaces. We begin with the following result, which shows that, under very general conditions, the PSAM matrices span different subspaces.

**Theorem 1.** Let  $\mathbf{P}$  be invertible, and let the constellation  $\mathcal{C}^P = \{X_i^P\}_{i=1}^L$  be constructed from  $\mathcal{A}$  as in (3). Then,  $A_i = A_j$  if and only if  $\mathcal{S}(X_i^P) = \mathcal{S}(X_j^P)$ .

**Proof.** Assume that  $A_i = A_j$ . From (3), we have that  $X_i^P = X_j^P$ , and hence,  $\mathcal{S}(X_i^P) = \mathcal{S}(X_j^P)$ . Assume now that  $\mathcal{S}(X_i^P) = \mathcal{S}(X_j^P)$ . In this case, the columns of  $X_i^P$  can be expressed as linear combinations of the columns of  $X_j^P$ , and vice versa. In particular, there is an  $M \times M$  matrix  $\mathbf{R}$  such that  $X_i^P \mathbf{R} = X_j^P$ , which, pre-multiplying by  $\mathbf{B}_1^*$  and using (3), yields  $\mathbf{P}\mathbf{R} = \mathbf{P}$ , and, since  $\mathbf{P}$  is invertible, we have that  $\mathbf{R} = \mathbf{I}_M$ . Now,  $\mathbf{B}_2^* X_i^P \mathbf{R} = \mathbf{B}_2^* X_j^P$  yields  $A_i = A_j$ .  $\square$

Theorem 1 shows that the subspaces spanned by PSAM matrices constructed from different coherent signals,  $A_i$  and  $A_j$ , are different, and hence, that a PSAM could be decoded detecting the transmitted subspace, as a USTM receiver. The following corollary follows from Theorem 1.

**Corollary 1.** Let  $\mathcal{C}^P$  be a PSAM constructed from  $\mathcal{A}$  as in (3), where  $\mathbf{P}$  is invertible. Then, there is a USTM,  $\mathcal{C}^U$ , which elements span the same subspaces than the elements of  $\mathcal{C}^P$ .

In other words, any PSAM has an equivalent USTM in terms of the spanned subspaces. The elements of the USTM can be obtained from the elements of the PSAM by using any kind of orthonormalization. In particular, it is possible to find matrices  $\{\mathbf{R}_i\}_{i=1}^L$  whose inverses orthonormalize the elements of the PSAM. The equivalent USTM is, in this case, composed of the matrices

$$X_i^U = X_i^P \mathbf{R}_i^{-1}, \quad i = 1, \dots, L. \tag{6}$$

The orthonormalization is, hence, a linear map. As an example, the matrices  $X_i^U$  and  $\mathbf{R}_i$  can be obtained by means of a QR decomposition of  $X_i^P$ . It is important to note that the performance of the equivalent USTM does not depend on the particular orthonormalization. This fact is due to the following result.

**Proposition 1.** Let the inverse of  $\mathbf{R}$  orthonormalize  $\mathbf{X} \in \mathbb{C}^{T \times M}$ ,  $T \geq M$ . Then, the inverse of  $\bar{\mathbf{R}}$  orthonormalizes  $\mathbf{X}$  if and only if  $\bar{\mathbf{R}} = \mathbf{U}\mathbf{R}$ , for some squared unitary matrix  $\mathbf{U}$ .

**Proof.** Since  $\mathbf{R}$  is invertible,  $\mathbf{U} = \bar{\mathbf{R}}\mathbf{R}^{-1}$  is well defined. The inverse of  $\bar{\mathbf{R}}$  orthonormalizes  $\mathbf{X}$  if and only if

$$\begin{aligned} \mathbf{I}_M &= \bar{\mathbf{R}}^{-1*} \mathbf{X}^* \mathbf{X} \bar{\mathbf{R}}^{-1} = \mathbf{U}^{-1*} \mathbf{R}^{-1*} \mathbf{X}^* \mathbf{X} \mathbf{R}^{-1} \mathbf{U}^{-1} \\ &= \mathbf{U}^{-1*} \mathbf{U}^{-1}. \end{aligned} \quad (7)$$

Therefore,  $\mathbf{U}$  is unitary.  $\square$

Proposition 1 implies that any orthonormalization of  $\mathbf{X}$  can be expressed using a given matrix  $\mathbf{R}$  which inverse orthonormalizes  $\mathbf{X}$  and a unitary matrix  $\mathbf{U}$ . In this case and using (2), the received signal, if  $\mathbf{X}$  is orthonormalized, is  $\mathbf{X}\mathbf{R}^{-1}\mathbf{U}^*\mathbf{H} + \sqrt{M/(\rho T)}\mathbf{Z}$ . Since the entries of  $\mathbf{H}$  are i.i.d. drawn from  $\mathcal{CN}(0, 1)$ , the entries of  $\mathbf{U}^*\mathbf{H}$  follow exactly the same distribution. Using this fact, it can be shown that the knowledge of the particular orthonormalization, i.e.,  $\mathbf{U}$ , does not provide additional information on  $\mathbf{X}$ . Therefore, the achievable rate is independent of  $\mathbf{U}$ .

The equivalence pointed out in Corollary 1 also exists in the other direction, as stated in the following result.

**Theorem 2.** Let  $\mathcal{C}^U = \{\mathbf{X}_i^U\}_{i=1}^L$  be a USTM and  $\mathbf{P}$  an  $M \times M$  invertible matrix. Then, there is a  $T \times M$  unitary matrix  $\mathbf{B}_1$ , with orthogonal complement  $\mathbf{B}_2$ , such that  $\mathbf{B}_1^* \mathbf{X}_i^U$ ,  $i = 1, \dots, L$ , are invertible and  $\{\alpha \mathbf{X}_i^U \bar{\mathbf{R}}_i\}_{i=1}^L$  is a PSAM which elements span the same subspaces than the elements of  $\mathcal{C}^U$ , where

$$\bar{\mathbf{R}}_i = (\mathbf{B}_1^* \mathbf{X}_i^U)^{-1} \mathbf{P}, \quad (8)$$

$$\alpha = \sqrt{\frac{LM}{\sum_{i=1}^L \|\bar{\mathbf{R}}_i\|_F^2}}. \quad (9)$$

In addition, if all the matrix signals in  $\mathcal{C}^U$  are equiprobable, then the PSAM satisfies the power constraint in (1).

**Proof.** This proof is divided into four steps, in which we will prove that i) matrix  $\mathbf{B}_1$ , such that  $\mathbf{B}_1^* \mathbf{X}_i^U$ ,  $i = 1, \dots, L$ , are invertible, exists; ii)  $\{\alpha \mathbf{X}_i^U \bar{\mathbf{R}}_i\}_{i=1}^L$  is a PSAM; iii) the elements of  $\mathcal{C}^U$  and  $\{\alpha \mathbf{X}_i^U \bar{\mathbf{R}}_i\}_{i=1}^L$  span the same subspaces; and iv) the constellation  $\{\alpha \mathbf{X}_i^U \bar{\mathbf{R}}_i\}_{i=1}^L$  satisfies the power constraint in (1). The proofs of these steps will complete the proof of the theorem.

(i) We start showing that  $\mathbf{B}_1$  exists such that  $\mathbf{B}_1^* \mathbf{X}_i^U$ ,  $i = 1, \dots, L$ , are invertible. To do this, we will use the following lemma.

**Lemma 1** ([22] (Exercise 14, p. 57)). If  $\{\mathcal{X}_i\}_{i=1}^L$  are subspaces of equal dimension of a finite-dimensional vector space  $\mathcal{V}$  over an infinite field  $\mathcal{F}$ , then there is a subspace  $\mathcal{W}$  of  $\mathcal{V}$  such that the direct sum of  $\mathcal{W}$  and each individual subspace in  $\{\mathcal{X}_i\}_{i=1}^L$  is  $\mathcal{V}$ , that is,  $\mathcal{V} = \mathcal{W} \oplus \mathcal{X}_i$ ,  $i = 1, \dots, L$ . In other words,  $\mathcal{W}$  is a common complement of the subspaces  $\{\mathcal{X}_i\}_{i=1}^L$ .

We will use this lemma with the subspaces  $\mathcal{W} = \mathcal{S}(\mathbf{B}_2)$  and  $\mathcal{X}_i = \mathcal{S}(\mathbf{X}_i^U)$ ,  $i = 1, \dots, L$ . In particular, Lemma 1 implies that we can find  $\mathbf{B}_2$  such that  $\mathcal{C}^T = \mathcal{S}([\mathbf{X}_i^U \mathbf{B}_2])$ ,  $i = 1, \dots, L$ . We will use this result with a contradiction argument. Let us assume that we can find  $i$  such that  $\mathbf{B}_1^* \mathbf{X}_i^U$  is not invertible. This implies that  $\mathcal{S}(\mathbf{X}_i^U)$  has, at least, one orthogonal direction to  $\mathcal{S}(\mathbf{B}_1)$ , or, in other words, that there is a vector  $\mathbf{x} \in \mathbb{C}^M$  such that  $\mathbf{b}^* \mathbf{B}_1^* \mathbf{X}_i^U \mathbf{x} = 0$  for all  $\mathbf{b} \in \mathbb{C}^M$ . Since  $\mathbf{B}_2$  is the orthogonal complement of  $\mathbf{B}_1$ , then the direction orthogonal to  $\mathcal{S}(\mathbf{B}_1)$  is in the subspace spanned by this basis, i.e.,  $\mathbf{X}_i^U \mathbf{x} \in \mathcal{S}(\mathbf{B}_2)$ . In this case,  $\mathbf{X}_i^U$  and  $\mathbf{B}_2$  cannot span  $\mathbb{C}^T$ , which contradicts Lemma 1.

(ii) In the second step, we will show that  $\{\alpha \mathbf{X}_i^U \bar{\mathbf{R}}_i\}_{i=1}^L$  is a PSAM by showing that these matrices can be written as in (3). In particular, we will find the pilot matrix and the coherent code  $\mathcal{A} = \{\mathbf{A}_i\}_{i=1}^L$  to be used in (3). From (3), the pilot matrix used to construct

the PSAM can be obtained by left-multiplying the PSAM matrices by  $B_1^*$ . In this case, this leads to

$$\alpha B_1^* X_i^U \bar{R}_i = \alpha P, \quad (10)$$

for  $i = 1, \dots, L$ . Now, using (3) again, the coherent code  $\mathcal{A}$  can be obtained by left-multiplying the PSAM matrices by  $B_2^*$ , that is,

$$A_i = \alpha B_2^* X_i^U \bar{R}_i = \alpha B_2^* X_i^U (B_1^* X_i^U)^{-1} P. \quad (11)$$

The PSAM can then be constructed as in (3) using the pilot matrix  $\alpha P$  and the coherent code  $\mathcal{A}$  with elements in (11).

(iii) Since the elements of  $\{\alpha X_i^U \bar{R}_i\}_{i=1}^L$  are generated by scaling and right-multiplying by  $M \times M$  invertible matrices the elements of the USTM, both span the same subspaces.

(iv) If the transmitted signal,  $X$ , takes values from the constellation  $\{\alpha X_i^U \bar{R}_i\}_{i=1}^L$  with equal probability, the average consumed energy in one channel block is

$$E[\|X\|_{\mathbb{F}}^2] = \frac{\alpha^2}{L} \sum_{i=1}^L \|X_i^U \bar{R}_i\|_{\mathbb{F}}^2 = M, \quad (12)$$

where the second equality follows from (9) and the fact that  $X_i^{U*} X_i^U = I_M$ ,  $i = 1, \dots, L$ . Therefore, the power constraint in (1) is satisfied.  $\square$

Corollary 1 and Theorem 2 describe a PSAM-USTM equivalence in terms of the spanned subspaces. In particular, any PSAM has an equivalent USTM which elements span the same subspaces, and vice versa.

In the following section, we will use the results of this section to design a new USTM construction method. The resulting USTM constellations exhibit a certain structure inherited from a coherent code.

#### 4. USTM Construction from PSAM Orthonormalization

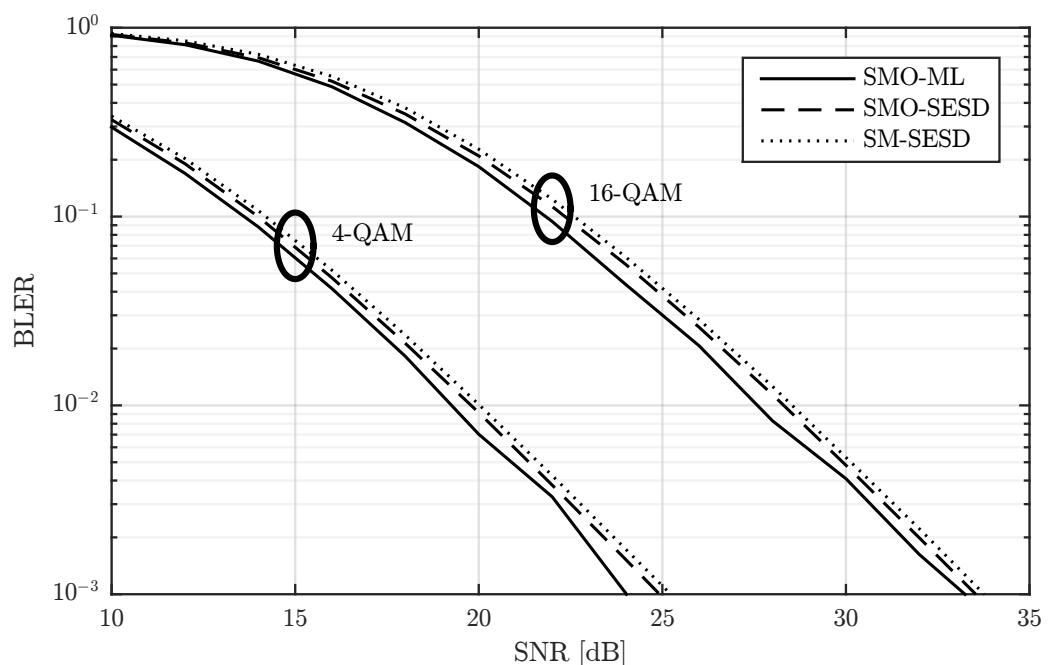
In this section, we present a new USTM construction method based on the results in Section 3. In particular, a USTM,  $\mathcal{C}^U = \{X_i^U\}_{i=1}^L$ , can be constructed by orthonormalizing the elements of a PSAM,  $\mathcal{C}^P = \{X_i^P\}_{i=1}^L$ , as in (6). In this case, the received signal when the  $i$ -th matrix of the constellation is transmitted results in

$$Y = X_i^U H + \sqrt{\frac{M}{\rho T}} Z = X_i^P R_i^{-1} H + \sqrt{\frac{M}{\rho T}} Z. \quad (13)$$

The constellation construction is computationally simple, since it only requires an orthonormalization (by means of, e.g., a QR decomposition) in addition to the generation of the PSAM matrix. However, in general, the orthonormalization does not leave an apparent structure in  $\mathcal{C}^U$  that could be used to efficiently decode these signals. This fact implies that the optimum receiver should test all signals. In order to circumvent this issue, we note that the inverse of the  $R_i$  matrix in (13) left-multiplies the channel matrix, but does not affect the noise. Therefore, it would be possible to interpret the  $R_i^{-1} H$  term as an effective channel, and use a PSAM reception technique to detect  $X_i^P$ . It is important to note that this effective channel is not independent of the transmitted signal, which implies that using a PSAM reception technique is suboptimal. Despite this, we will show a case in which the combination of these transmission and reception techniques is better than directly transmitting the PSAM matrices. In particular, we tested this constellation design method with a PSAM constructed from a scaled identity pilot matrix and a coherent code whose matrix entries are independent and selected from a rectangular quadrature amplitude modulation (QAM), i.e., spatial multiplexing (SM) of QAM symbols. The available power was split between the pilot matrix and the coherent code following the indications in [1].

In Figure 2, we show the block error rate (BLER) for  $T = 4$ ,  $M = N = 2$ , modulations of 4-QAM and 16-QAM, and three different transmitter-receiver configurations. The modulations of 4-QAM and 16-QAM lead to transmission rates of  $R = 2$  and  $R = 4$  bits per channel use (bpcu), respectively, whereas the transmitter-receiver configurations are detailed next:

- SM-SESD: In this case, the transmitter sent PSAM matrices constructed from a scaled identity pilot matrix and SM of QAM symbols. The receiver used a least squares (LS) channel estimator [23] and a Schnorr-Euchner sphere decoder (SESD) [24].
- SMO-SESD: This configuration is similar to the previous one. The only difference is that the signal matrices were orthonormalized before transmission. The receiver had no knowledge of this orthonormalization and assumed that the transmitted matrices were from the original PSAM constellation.
- SMO-ML: In this configuration, the transmitter orthonormalized the PSAM matrices too. The receiver tested all constellation matrices and performed a maximum likelihood (ML) detection [7].



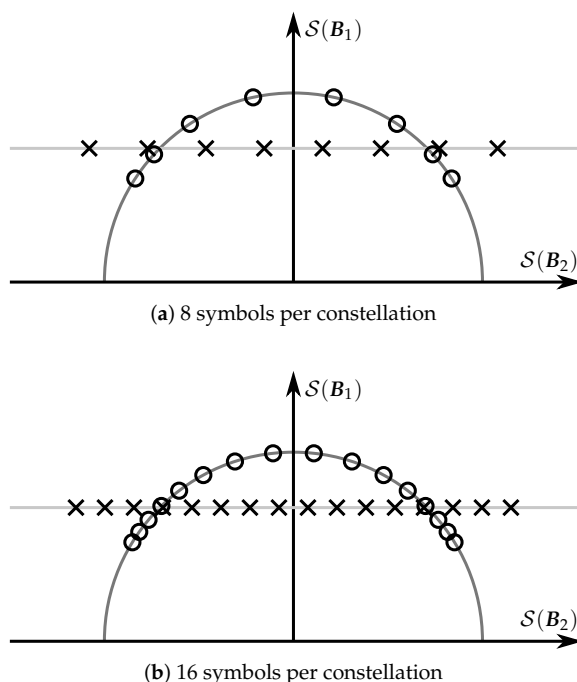
**Figure 2.** Performance comparison of a PSAM and its element orthonormalization with two different receivers for  $T = 4$ ,  $M = N = 2$ , and two transmission rates.

The results in Figure 2 show that the SMO-ML configuration exhibits around 1 dB gain with respect to the other two configurations, albeit by means of using a computationally expensive receiver. In particular, since the SMO-ML receiver has to test all constellation matrices, the computational cost is proportional to the number of matrices in the constellation, i.e.,  $L = 2^{RT}$ , which follows an exponential law with respect to  $R$  and  $T$ . This fact makes the SMO-ML receiver extremely costly for relatively small values of  $R$  and/or  $T$ . As an illustrative example, for the cases drawn in Figure 2, the number of matrices are 256 and 65,536 for the  $R = 2$  and  $R = 4$  cases, respectively, and this value increases to 1,048,576 for  $R = 5$ . The gain of the SMO-SESD configuration with respect to the SM-SESD is around 0.25 dB. In this case, the receivers are exactly the same, and the computational cost of the orthonormalization performed by the transmitter in the SMO-SESD configuration is polynomial with respect to  $T$  and  $M$ . Therefore, this gain is completely free for the receiver and relatively inexpensive for the transmitter.

The previous result shows that the orthonormalization of PSAM matrices is a simple idea that can provide some performance gain. In the rest of this section, we will examine

how the orthonormalization modifies the structure of the PSAM and how this modification affects the performance. To do this, we graphically represented in Figure 3 the structures of two PSAMs and the equivalent USTMs using the  $\mathcal{S}(\mathbf{B}_1)$  and  $\mathcal{S}(\mathbf{B}_2)$  subspaces.

It can be easily inferred from the figure that the PSAM elements whose energy is lower than the constellation average, i.e., those inside the semicircle in Figure 3, are separated after the orthonormalization, and that the rest are bunched together. The separation between constellation elements is related to the protection against noise. In other words, Figure 3 suggests that the orthonormalization is beneficial for the PSAM elements with less energy than the average, but disadvantageous for the rest.

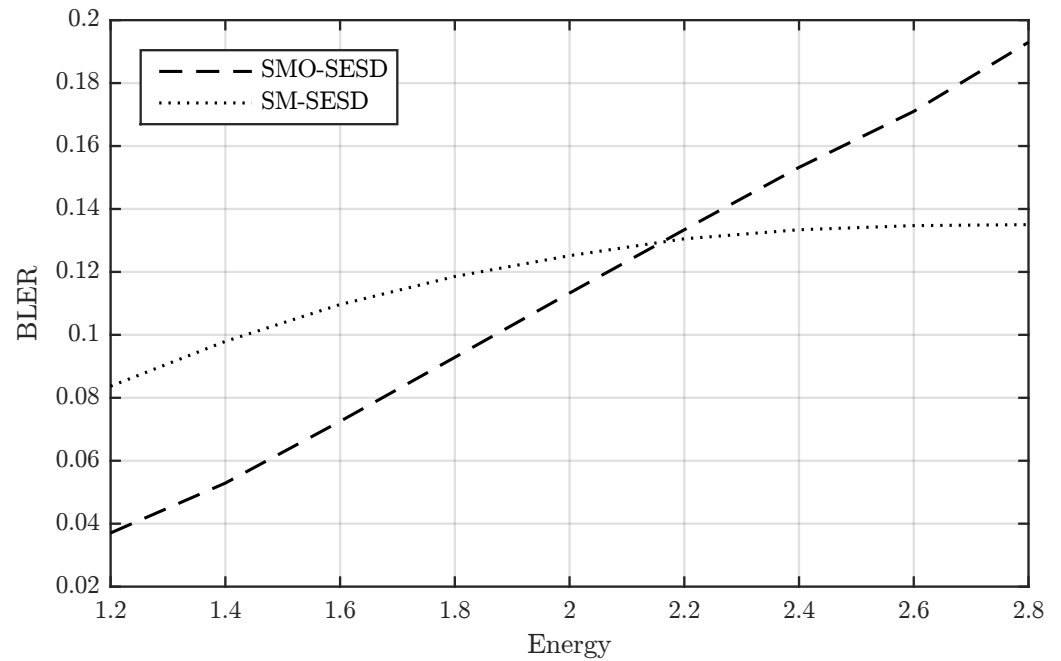


**Figure 3.** Two-dimensional graphical representation of the elements of a PSAM (with ‘x’ symbols) and the equivalent USTM (with ‘o’ symbols).

We verified this effect in the constellations of Figure 2. In particular, we studied how the PSAM matrix energy affects the BLER of the SM-SESD and SMO-SESD transmitter-receiver configurations with  $R = 4$ . This result is depicted in Figure 4, where the energy of the  $i$ -th PSAM matrix was computed as  $\|\mathbf{X}_i^P\|_F^2$ . The figure confirms that the orthonormalization is beneficial below certain energy of the PSAM matrix.

Looking at Figure 3, it is possible to identify a better USTM in which the bunched elements are stretched out towards the  $\mathcal{S}(\mathbf{B}_2)$  subspace. Unfortunately, this type of USTMs do not correspond to a simple orthonormalization of the elements of a PSAM. In the following section, we will propose another USTM constellation construction method to circumvent the limits of the PSAM orthonormalization.





**Figure 4.** BLER vs. PSAM matrix energy of a PSAM and its element orthonormalization for  $T = 4$ ,  $M = N = 2$ ,  $R = 4$ , and an SNR of 22 dB.

## 5. USTM Construction Combining a Coherent Code and a Small USTM

### 5.1. General Description

In this section, we present another USTM construction method that is inspired by the definition of PSAM matrices in (3), and based on the combination and orthonormalization of a USTM with  $L^U$  elements and a coherent code of  $L^P$  elements to generate a constellation of  $L = L^U L^P$  elements. The main idea is to compose  $L^U$  PSAMs using the USTM elements to carry the pilot matrix in place of the  $\mathbf{B}_1$  matrix in (3). In particular, letting  $\mathcal{C}^U = \{\mathbf{X}_i^U\}_{i=1}^{L^U}$ ,  $\mathcal{C}^{U\perp} = \{\mathbf{X}_i^{U\perp}\}_{i=1}^{L^U}$ ,  $\mathbf{P}$ , and  $\mathcal{A} = \{\mathbf{A}_i\}_{i=1}^{L^P}$  be the USTM, a set of orthogonal complements of the USTM elements, a pilot matrix, and the coherent code, respectively, we propose to compose the PSAMs  $\mathcal{C}_i^P = \{\mathbf{X}_{i,j}^P\}_{j=1}^{L^P}$ ,  $i = 1, \dots, L^U$ , where

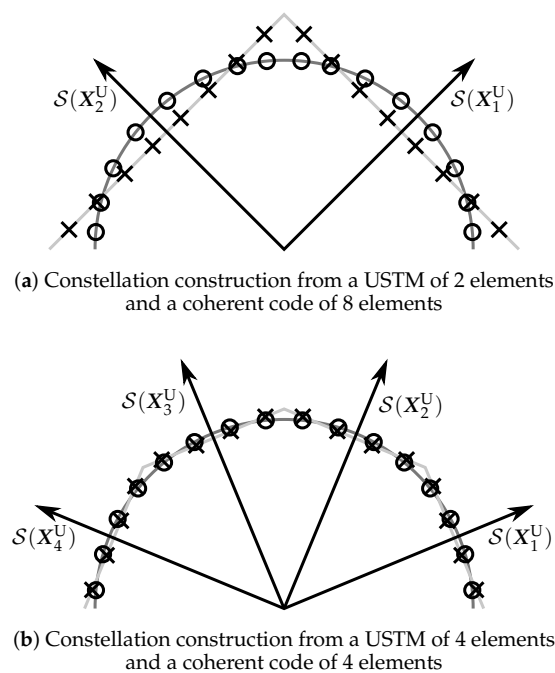
$$\mathbf{X}_{i,j}^P = \mathbf{X}_i^U \mathbf{P} + \mathbf{X}_i^{U\perp} \mathbf{A}_j. \quad (14)$$

Subsequently, we use these PSAMs to construct the constellation  $\mathcal{C}^O = \{\{\mathbf{X}_{i,j}^O\}_{i=1}^{L^U}\}_{j=1}^{L^P}$ , where  $\mathbf{X}_{i,j}^O$  is the orthonormalization of  $\mathbf{X}_{i,j}^P$ , i.e., we can find a matrix  $\mathbf{R}_{i,j}$  such that  $\mathbf{X}_{i,j}^O = \mathbf{X}_{i,j}^P \mathbf{R}_{i,j}^{-1}$  is unitary. Following the general PSAM description of Section 2, we assume that the PSAMs  $\{\mathcal{C}_i^P\}_{i=1}^{L^U}$  satisfy the power constraint in (1), i.e., the pilot matrix power  $P_P = \|\mathbf{P}\|_F^2$  and the coherent code power  $P_A = \frac{1}{L^P} \sum_{j=1}^{L^P} \|\mathbf{A}_j\|_F^2$  satisfy  $P_P + P_A = M$ .

A two-dimensional graphical representation of this constellation construction method is depicted in Figure 5. The figure illustrates one of the benefits of this method: the constellation elements are more homogeneously distributed in the semicircle than in Figure 3. In addition to a composition of  $L^U$  orthonormalized PSAMs, the constellation  $\mathcal{C}^O$  admits another interpretation. In particular, it can be seen as a generalization of a PSAM orthonormalization in which data are encoded also in the pilots, or, more specifically, in the location of the pilot matrix within a channel block. This location is given by the USTM matrices in  $\mathcal{C}^U$ . It is important to note that, assuming  $L^U$  is small enough, both constructing  $\mathcal{C}^U$  using numerical optimization tools and storing  $\mathcal{C}^U$  and  $\mathcal{C}^{U\perp}$  is feasible.

The receiver can estimate the transmitted signal using a PSAM and a USTM receiver in two phases. In the first phase, we propose to select one candidate from each PSAM  $\mathcal{C}_i^P$ ,  $i = 1, \dots, L^U$ , using the PSAM receiver. More specifically, for the  $i$ -th PSAM,  $\mathcal{C}_i^P$ , the PSAM receiver should assume that the  $i$ -th matrix from  $\mathcal{C}^U$  was used to carry the pilot matrix and estimate the transmitted coherent matrix, say,  $A_{j_i}$ . In the second phase, the USTM receiver assumes that the constellation used by the transmitter was the set of candidates, i.e.,  $\{X_{i,j_i}^O\}_{i=1}^{L^U}$ , and selects the received matrix from this set. The computational cost of this reception technique is equivalent to the cost of  $L^U$  PSAM receivers for constellations of  $L^P$  elements, plus the cost of a USTM receiver for constellations of  $L^U$  elements.

Apart from the power constraint, this construction method does not impose any other limitation on how the available power is distributed between  $P$  and  $\mathcal{A}$ . However, this distribution affects the performance of  $\mathcal{C}^O$ , and, hence, it should be carefully designed. In the following two sections, we will propose two methods to select  $P_P$  and  $P_A$ . The first method, in Section 5.2, is a general approach that can be used for any system, although it is very conservative and has a limited performance, especially for certain coherent codes. The second method, in Section 5.3, is valid for certain  $T$ ,  $M$ , and  $L^U$  values, and the generated constellations present some performance gains as compared with other constellation constructions.



**Figure 5.** Two-dimensional graphical representation of the elements of the PSAMs constructed combining a USTM and a coherent code (with ‘x’ symbols), and the orthonormalization of these elements (with ‘o’ symbols).

### 5.2. General Method to Select $P_P$ and $P_A$

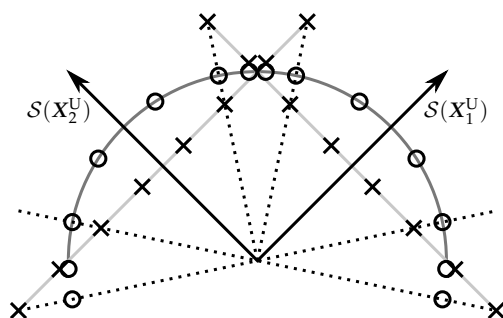
Although the authors of [1] obtained the optimal power distribution between the pilot matrix and the coherent code for one PSAM, the combination of several PSAMs to compose a unique constellation limits the performance of this power distribution. In particular, the elements of the resulting constellation  $\mathcal{C}^O$  are intermingled, and even some of them could span the same subspace, as depicted in Figure 6. In order to avoid this, the distance between the subspaces spanned by the constellation elements has to be controlled. For instance, let  $d_S(X_1, X_2)$  be the distance between the subspaces spanned by the matrices  $X_1$  and  $X_2$ , if  $P_P$  and  $P_A$  are selected in such a way that

$$\max_{i,j} d_S(X_i^U, X_{i,j}^O) < \frac{1}{2} \min_{k,\ell \neq k} d_S(X_k^U, X_\ell^U), \tag{15}$$

then it can be shown that the constellation elements span different subspaces. In particular, using (15) and the triangle inequality, i.e.,  $d_S(\mathbf{X}_1, \mathbf{X}_3) \leq d_S(\mathbf{X}_1, \mathbf{X}_2) + d_S(\mathbf{X}_2, \mathbf{X}_3)$ , we have that

$$d_S(\mathbf{X}_i^U, \mathbf{X}_k^U) \leq d_S(\mathbf{X}_i^U, \mathbf{X}_{i,j}^O) + d_S(\mathbf{X}_{i,j}^O, \mathbf{X}_{k,\ell}^O) + d_S(\mathbf{X}_k^U, \mathbf{X}_{k,\ell}^O) < d_S(\mathbf{X}_i^U, \mathbf{X}_k^U) + d_S(\mathbf{X}_{i,j}^O, \mathbf{X}_{k,\ell}^O), \tag{16}$$

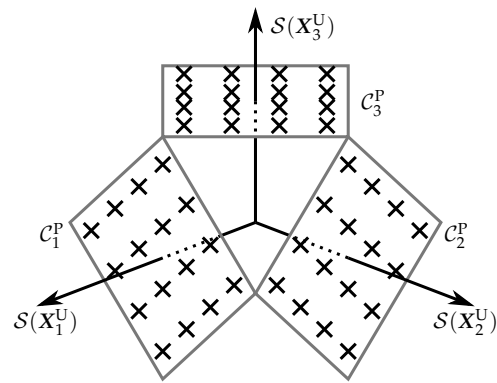
for all  $i = 1, \dots, L^U, k = 1, \dots, i - 1, i + 1, \dots, L^U, j = 1, \dots, L^P$ , and  $\ell = 1, \dots, L^P$ . The inequality in (16) implies that  $d_S(\mathbf{X}_{i,j}^O, \mathbf{X}_{k,\ell}^O) > 0$  and, hence, that  $\mathbf{X}_{i,j}^O$  and  $\mathbf{X}_{k,\ell}^O$  span different subspaces. It is important to note that the previous result is independent of the particular metric used to measure the subspace distance.



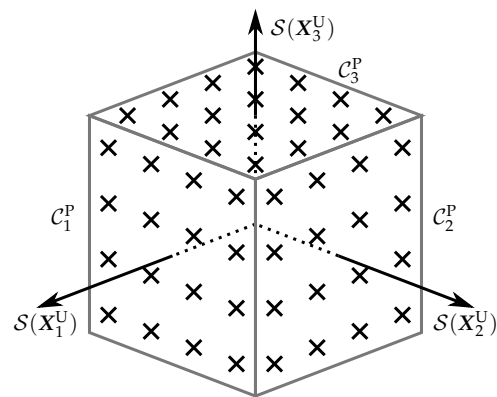
**Figure 6.** Two-dimensional graphical representation of the potential consequence of a wrong power distribution between the pilot matrix and the coherent code. Four pairs of constellation elements span the same subspaces, which are drawn with dotted lines. (element: ‘o’ symbols, coherent code: ‘x’ symbols).

The main drawback of this method is that it does not take into account the shape of the coherent code,  $\mathcal{A}$ , or the orientation of the orthogonal complements in  $\mathcal{C}^{U\perp}$ . In fact, this method ensures that the constellation elements span different subspaces even in cases in which a bad combination of coherent code and orthogonal complements is used. This implies that this method is, in general, conservative and not very efficient. Figure 7 illustrates this drawback with two examples. In particular, the example in Figure 7a shows three PSAMs constructed from a square-shaped coherent code, which power has to be reduced to avoid that the elements near the square corners intersect with those of the other PSAMs. This power level generates a lot of empty space, although, in this case, the orthonormalization of the elements in Figure 7a satisfy (15). In Figure 7b, the orthogonal complements are rotated 45°, which allows more power to be allocated to the coherent code. In this case, the large empty spaces are eliminated and the constellation elements are more homogeneously distributed, although their orthonormalization does not necessarily satisfy (15).

In the following section, we will describe a method to produce constellations similar to that depicted in Figure 7b. In particular, prior to the orthonormalization, the constellation elements are located in the surface of a hypercube defined in a special vector set: a module over a ring.



(a) Bad orientation of the orthogonal complements



(b) Good orientation of the orthogonal complements

**Figure 7.** Three-dimensional graphical representation of the effect of the orthogonal complements orientation and the coherent code shape. A USTM of three elements is used to construct three PSAMs, which elements are surrounded by quadrilaterals. (coherent code: ‘x’ symbols).

### 5.3. USTM Construction from a Hypercube

In this section, we present a USTM construction method in which the constellation elements are first placed in the surface of a hypercube, and then they are orthonormalized. Hypercubes are generally defined in vector spaces over the field of the real numbers  $\mathbb{R}$ , i.e.,  $\mathbb{R}^n$ . However, for this construction method, we need to define the hypercube in a different vector set in which the field  $\mathbb{R}$  is substituted by a ring. This type of vector sets are known as modules over a ring. In this section, (i) we present the module in which the hypercube is defined, (ii) we introduce a hypercube definition that can be used in this module, and (iii) we describe the method itself.

Assuming that  $T$  is a multiple of  $M$ , i.e.,  $\frac{T}{M}$  is an integer, we can interpret  $\mathbb{C}^{T \times M}$  as the Cartesian product of  $\frac{T}{M}$  sets  $\mathcal{R} = \mathbb{C}^{M \times M}$ , i.e.,  $\mathcal{M} = \mathbb{C}^{T \times M} = \mathcal{R}^{\frac{T}{M}}$ . It is important to note that  $\mathcal{R}$  is not commutative and, hence, it is not a field, which implies that  $\mathcal{M}$  is not a vector space over  $\mathcal{R}$ . However, it can be easily verified that  $\mathcal{R}$  is a ring equipped with the operations of matrix sum and product. Consequently,  $\mathcal{M}$  is a module over  $\mathcal{R}$ . More specifically,  $\mathcal{M}$  is a right  $\mathcal{R}$ -module equipped with the matrix product between elements of the two sets. Note that, using this interpretation,  $\mathbf{X} \in \mathcal{M}$  is a vector with  $\frac{T}{M}$  components, and that each component is a matrix in  $\mathcal{R}$ . In addition to this, this module has a basis, which can be built from any  $T \times T$  unitary matrix. In particular, let  $\mathcal{B} = \{\mathbf{B}_i\}_{i=1}^{\frac{T}{M}}$  be a collection of  $T \times M$  matrices such that  $[\mathbf{B}_1 \cdots \mathbf{B}_{\frac{T}{M}}]$  is unitary, i.e.,

$$[\mathbf{B}_1 \cdots \mathbf{B}_{\frac{T}{M}}][\mathbf{B}_1 \cdots \mathbf{B}_{\frac{T}{M}}]^* = \sum_{i=1}^{\frac{T}{M}} \mathbf{B}_i \mathbf{B}_i^* = \mathbf{I}_T, \quad (17)$$

then  $\mathcal{B}$  is a basis of  $\mathcal{M}$ . In other words, for any  $\mathbf{X} \in \mathcal{M}$ , we can find coordinates  $\mathbf{C}_i \in \mathcal{R}$ ,  $i = 1, \dots, \frac{T}{M}$ , such that  $\mathbf{X} = \sum_{i=1}^{\frac{T}{M}} \mathbf{B}_i \mathbf{C}_i$ . In particular, it can be shown from (17) that the matrices  $\mathbf{C}_i = \mathbf{B}_i^* \mathbf{X}$  satisfy the previous equality. Since the elements of  $\mathcal{B}$  also satisfy  $\mathbf{B}_i^* \mathbf{B}_j = \mathbf{0}_M$ , for all  $i \neq j$ , and  $\mathbf{B}_i^* \mathbf{B}_i = \mathbf{I}_M$ , for all  $i$ , where  $\mathbf{0}_M$  is the  $M \times M$  null matrix, we say that  $\mathcal{B}$  is an orthonormal basis.

We define the hypercube in  $\mathcal{M}$  defined by the orthonormal basis  $\mathcal{B}$  as the set

$$\mathcal{H}_{\mathcal{B}}(\xi) = \left\{ \mathbf{X} = \sum_{i=1}^{\frac{T}{M}} \mathbf{B}_i \mathbf{C}_i \mid \|\mathbf{C}_i\|_{\mathbb{F}} \leq \xi, \forall i \right\}, \quad (18)$$

where  $\xi > 0$  is a constant that defines the size of the hypercube. It can be readily verified that, in the case  $\mathcal{R} = \mathbb{R}$ ,  $\mathcal{H}_{\mathcal{B}}(\xi)$  is an hypercube centered at the origin, oriented following the elements in the basis  $\mathcal{B}$ , and whose sides expand from  $-\xi$  to  $\xi$  in the coordinate system defined by  $\mathcal{B}$ . The surface of the hypercube is the set of points in which the constraint in (18) is satisfied as an equality for at least one matrix  $\mathbf{C}_i$ , i.e., the surface of  $\mathcal{H}_{\mathcal{B}}(\xi)$  is

$$\mathcal{H}_{\mathcal{B}}^{\mathcal{S}}(\xi) = \left\{ \mathbf{X} \in \mathcal{H}_{\mathcal{B}}(\xi) \mid \exists i, \|\mathbf{B}_i^* \mathbf{X}\|_{\mathbb{F}} = \xi \right\}. \quad (19)$$

We say that all the matrices  $\mathbf{X}$  that satisfy  $\|\mathbf{B}_i^* \mathbf{X}\|_{\mathbb{F}} = \xi$  for the same coordinate  $i$  are in the same face of the hypercube.

After the introduction of the hypercube definition, we are able to describe the construction of the USTM. To do that, we start composing  $\frac{T}{M}$  PSAMs with elements in different faces of a hypercube. These PSAMs are composed using an orthonormal basis  $\mathcal{B}$  of the module  $\mathcal{M}$ , a coherent code  $\mathcal{A}^{\text{H}} = \{\mathbf{A}_i^{\text{H}}\}_{i=1}^{L^{\text{P}}}$ , where  $\mathbf{A}_i^{\text{H}} \in \mathbb{C}^{M \times M}$ , and a pilot matrix  $\mathbf{P}$ . Note that, in this case, the coherent code elements are drawn from  $\mathbb{C}^{M \times M}$  and not from  $\mathbb{C}^{T-M \times M}$  as before. The reason for this is that the coherent code elements are going to be used as coordinates of matrices in  $\mathcal{M}$ . Using the previous pieces, we can compose the PSAMs  $\mathcal{C}_i^{\text{HP}} = \{\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HP}}\}_{\boldsymbol{\iota} \in \mathcal{I}}$ ,  $i = 1, \dots, \frac{T}{M}$ , where  $\boldsymbol{\iota} = [\iota_1 \cdots \iota_{\frac{T}{M}-1}]^*$  is a vector of indices,  $\mathcal{I} = \{1, \dots, L^{\text{P}}\}^{\frac{T}{M}-1}$ ,

$$\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HP}} = \mu \sum_{j=1}^{i-1} \mathbf{B}_j \mathbf{A}_j^{\text{H}} + \mathbf{B}_i \mathbf{P} + \mu \sum_{j=i+1}^{\frac{T}{M}} \mathbf{B}_j \mathbf{A}_{j-1}^{\text{H}}, \quad (20)$$

and  $\mu > 0$  is a factor that scales the coherent code. As mentioned before, it is clear from (20) that  $\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HP}}$  has coordinates in the module  $\mathcal{M}$  that are scaled versions of the elements of the coherent code  $\mathcal{A}^{\text{H}}$ , except for the  $i$ -th coordinate, which is the pilot matrix. It can be shown that, with a proper  $\mu$  value, the PSAM elements are in different faces of the hypercube  $\mathcal{H}_{\mathcal{B}}(\|\mathbf{P}\|_{\mathbb{F}})$ . In particular,  $\mu$  has to satisfy

$$\mu \|\mathbf{A}_i^{\text{H}}\|_{\mathbb{F}} < \|\mathbf{P}\|_{\mathbb{F}}, \quad (21)$$

for all  $i = 1, \dots, L^{\text{P}}$ . Note that the strict inequality is required to ensure that the PSAM elements are not placed in two hypercube faces at the same time. We now use the PSAMs  $\mathcal{C}_i^{\text{HP}}$ ,  $i = 1, \dots, \frac{T}{M}$ , to construct the constellation  $\mathcal{C}^{\text{HO}} = \{\{\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HO}}\}_{\boldsymbol{\iota} \in \mathcal{I}}\}_{i=1}^{\frac{T}{M}}$ , where  $\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HO}}$  is the orthonormalization of  $\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HP}}$ . The number of elements in this constellation is  $\frac{T}{M} (L^{\text{P}})^{\frac{T}{M}-1}$ .

The following result ensures that the elements in  $\mathcal{C}^{\text{HO}}$  span different subspaces.

**Theorem 3.** Let the pilot matrix used to compose  $\mathcal{C}^{\text{HO}}$  be a scaled unitary matrix, i.e.,  $\mathbf{P} = \eta \mathbf{U}$ , where  $\eta > 0$  and  $\mathbf{U}$  is unitary; and let  $\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HO}} \in \mathcal{C}^{\text{HO}}$  and  $\mathbf{X}_{j,\boldsymbol{\kappa}}^{\text{HO}} \in \mathcal{C}^{\text{HO}}$ . Then,  $\mathcal{S}(\mathbf{X}_{i,\boldsymbol{\iota}}^{\text{HO}}) = \mathcal{S}(\mathbf{X}_{j,\boldsymbol{\kappa}}^{\text{HO}})$  if and only if  $i = j$  and  $\boldsymbol{\iota} = \boldsymbol{\kappa}$ .

**Proof.** We start by noting that  $\mathbf{X}_{i,\iota}^{\text{HO}}$  and  $\mathbf{X}_{i,\iota}^{\text{HP}}$  span the same subspaces. Therefore, it is equivalent to show that  $\mathcal{S}(\mathbf{X}_{i,\iota}^{\text{HP}}) = \mathcal{S}(\mathbf{X}_{j,\kappa}^{\text{HP}})$  if and only if  $i = j$  and  $\iota = \kappa$ . Using Theorem 1, it is straightforward to prove this theorem for the case  $i = j$ . Therefore, we continue assuming  $i \neq j$ . We will proceed by assuming that  $\mathcal{S}(\mathbf{X}_{i,\iota}^{\text{HP}}) = \mathcal{S}(\mathbf{X}_{j,\kappa}^{\text{HP}})$ , and conclude that, in this case, the inequality in (21) cannot be satisfied, which concludes the proof. In particular, if  $\mathcal{S}(\mathbf{X}_{i,\iota}^{\text{HP}}) = \mathcal{S}(\mathbf{X}_{j,\kappa}^{\text{HP}})$ , there is some  $\mathbf{R} \in \mathbb{C}^{M \times M}$  such that  $\mathbf{X}_{i,\iota}^{\text{HP}} = \mathbf{X}_{j,\kappa}^{\text{HP}} \mathbf{R}$ . Premultiplying both sides of this equality by  $\mathbf{B}_i^*$  and  $\mathbf{B}_j^*$ , we obtain  $\mathbf{P} = \mu \mathbf{D} \mathbf{R}$  and  $\mu \mathbf{E} = \mathbf{P} \mathbf{R}$ , respectively, for some  $\mathbf{D}, \mathbf{E} \in \mathcal{A}^{\text{H}}$ . Therefore, since  $\mathbf{P}$  is invertible, then all  $\mathbf{D}, \mathbf{E}$  and  $\mathbf{R}$  are invertible. Moreover, we can combine both equalities to obtain  $\mu \mathbf{E} = \frac{1}{\mu} \mathbf{P} \mathbf{D}^{-1} \mathbf{P} = \frac{\eta^2}{\mu} \mathbf{U} \mathbf{D}^{-1} \mathbf{U}$ . Both  $\mathbf{D}$  and  $\mathbf{E}$  must satisfy the inequality in (21), and thus, we should have that

$$\mu^2 \text{tr}(\mathbf{D}^* \mathbf{D}) < \text{tr}(\mathbf{P}^* \mathbf{P}) = \eta^2 M, \quad (22)$$

$$\begin{aligned} \mu^2 \text{tr}(\mathbf{E}^* \mathbf{E}) &= \frac{1}{\mu^2} \text{tr}(\mathbf{P}^* \mathbf{D}^{*-1} \mathbf{P}^* \mathbf{P} \mathbf{D}^{-1} \mathbf{P}) \\ &= \frac{\eta^4}{\mu^2} \text{tr}(\mathbf{D}^{*-1} \mathbf{D}^{-1}) < \eta^2 M. \end{aligned} \quad (23)$$

Let  $\lambda_1, \dots, \lambda_M$  be the eigenvalues of  $\mathbf{D}^* \mathbf{D}$ . The two previous inequalities imply that

$$\mu^2 \sum_{m=1}^M \lambda_m < \eta^2 M, \quad (24)$$

$$\frac{\eta^4}{\mu^2} \sum_{m=1}^M \frac{1}{\lambda_m} < \eta^2 M. \quad (25)$$

For a given sum of the eigenvalues, the left-hand side of (25) is minimum if all the eigenvalues are equal. Therefore,

$$\frac{\eta^4}{\mu^2} \sum_{m=1}^M \frac{1}{\lambda_m} > \frac{\eta^4}{\mu^2} \sum_{n=1}^M \frac{M}{\sum_{m=1}^M \lambda_m} = \frac{\eta^4}{\mu^2} \frac{M^2}{\sum_{m=1}^M \lambda_m}. \quad (26)$$

Using (24) in (26) we obtain

$$\frac{\eta^4}{\mu^2} \sum_{m=1}^M \frac{1}{\lambda_m} > \eta^2 M, \quad (27)$$

which contradicts (25).  $\square$

As a consequence of Theorem 3, if  $\mathbf{P}$  is a scaled unitary matrix and  $\mu$  satisfies (21) for all the elements in the coherent code  $\mathcal{A}^{\text{H}}$ , then the elements in  $\mathcal{C}^{\text{HO}}$  span different subspaces. This fact highlights the importance of a good selection of the scaling factor  $\mu$ . In addition to this, it is important to note that  $\mu$  significantly impacts the performance of the constellation  $\mathcal{C}^{\text{HO}}$ . In order to visualize this impact, we refer to Figure 7b, which is a graphical representation of the PSAMs  $\mathcal{C}_i^{\text{HP}}, i = 1, \dots, \frac{T}{M}$ , for the case  $\frac{T}{M} = 3$ . Note that the 'x' symbols in one of the faces of the cube in Figure 7b represent one constellation element, and that the coherent code  $\mathcal{A}^{\text{H}}$  used for this graphical representation is composed of four elements. The scaling factor  $\mu$  controls how the elements of each PSAM are distributed throughout the cube faces. In particular, a low value causes the PSAM elements to be concentrated around the axes, whereas a large value makes the elements approach the face borders, and thus, the elements in a neighbouring cube face. In order to obtain a good performance, we have to set the value of  $\mu$  taking into account the protection against noise and the reception technique.

Regarding the protection against noise, we should equalize the elements' separation in each PSAM, and the separation between elements in different PSAMs. To this end, we propose to select  $\mu$  in such a way that

$$\min_{i \neq j} d_F(\mu A_i^H, \mu A_j^H) = \sqrt{2} \min_{i, Q} d_F(\mu A_i^H, Q), \tag{28}$$

where  $d_F(X, Y)$  is the Euclidean distance between  $X$  and  $Y$ , i.e.,  $d_F(X, Y) = \|X - Y\|_F$ , and  $Q$  is restricted to  $\|Q\|_F = \|P\|_F$ . The left-hand side of (28) is the minimum distance between two coherent code elements, which is, at the same time, the minimum distance between two elements of the same PSAM. The minimization in the right hand side of (28) provides the minimum distance between the coherent code elements and a matrix with Frobenius norm equal to  $\|P\|_F$ . Figure 8 provides a graphical representation of the previous distances for a  $\frac{T}{M} = 2$  case.

Regarding the reception technique, if a PSAM receiver is used to select candidates from the PSAMs, we should consider the power split between the pilot matrix and the coherent code obtained in [1]. In this case, we propose to set  $\mu$  to the minimum value between the one that satisfies (28), and the one that satisfies the power split in [1].

In the following section, we will use the hypercube method described in this section to obtain USTM constellations from coherent codes composed of matrices with rectangular QAM symbols in their entries.

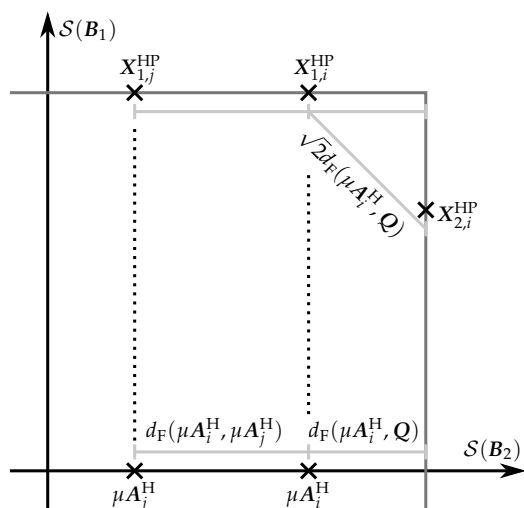


Figure 8. Graphical representation of the distances used in (28) to set the value of  $\mu$  for  $\frac{T}{M} = 2$ .

#### 5.4. Application to Rectangular QAM-Based Coherent Codes

In this section, we focus on a special case of the hypercube method to construct USTM constellations. In particular, the coherent codes of this case are based on rectangular QAM, i.e., the real and imaginary parts of each matrix entry in the coherent codes are independently and equiprobably drawn from an equispaced set of values centered at the origin. More specifically, the set of values is

$$Q = \left\{ q \left( i - \frac{L^Q + 1}{2} \right) \right\}_{i=1}^{L^Q}, \tag{29}$$

where  $L^Q$  is the number of elements in  $Q$ , and  $q > 0$  is a factor that scales the QAM constellation. In the rest of this section, we assume that a PSAM receiver is used to select candidates from the PSAMs, and we obtain the value of  $\mu$  following the method exposed in the previous section.

We start finding the value of  $\mu$  that satisfies (28), which we express as  $\mu_1$ . To this aim, we note that, for coherent codes based on rectangular QAM, the left-hand side of (28) is  $\mu_1 q$ . In order to find the right-hand side of (28), we first solve

$$\begin{aligned} \min_{\mathbf{Q}} d_{\text{F}}^2(\mathbf{X}, \mathbf{Q}), \\ \text{s.t., } \|\mathbf{Q}\|_{\text{F}}^2 = \zeta^2. \end{aligned} \quad (30)$$

The Lagrangian of (30) is  $L(\mathbf{Q}) = \text{tr}((\mathbf{X} - \mathbf{Q})^*(\mathbf{X} - \mathbf{Q})) + \sigma(\text{tr}(\mathbf{Q}^*\mathbf{Q}) - \zeta^2)$ , where  $\sigma$  is a Lagrange multiplier. Using the differential of the Lagrangian, it can be shown that the matrix  $\mathbf{Q}$  that solves (30) must satisfy  $-(\mathbf{X} - \mathbf{Q})^* + \sigma\mathbf{Q}^* = 0$ , which yields  $\mathbf{Q} = \frac{1}{1+\sigma}\mathbf{X}$ . This implies that the solution of (30) is proportional to the given matrix  $\mathbf{X}$ . Using these results, the equality in (28) can be expressed as

$$\begin{aligned} \mu_1 q &= \sqrt{2} \min_i d_{\text{F}} \left( \mu_1 \mathbf{A}_i^{\text{H}}, \frac{\|\mathbf{P}\|_{\text{F}}}{\|\mathbf{A}_i^{\text{H}}\|_{\text{F}}} \mathbf{A}_i^{\text{H}} \right) \\ &= \sqrt{2} \min_i \left| \mu_1 - \frac{\|\mathbf{P}\|_{\text{F}}}{\|\mathbf{A}_i^{\text{H}}\|_{\text{F}}} \right| \|\mathbf{A}_i^{\text{H}}\|_{\text{F}}. \end{aligned} \quad (31)$$

Since  $\mu_1$  must satisfy (21), we have that

$$\mu_1 < \frac{\|\mathbf{P}\|_{\text{F}}}{\|\mathbf{A}_i^{\text{H}}\|_{\text{F}}}, \quad (32)$$

and hence, the equality in (31) results in

$$\begin{aligned} \mu_1 q &= \sqrt{2} \min_i \left( \frac{\|\mathbf{P}\|_{\text{F}}}{\|\mathbf{A}_i^{\text{H}}\|_{\text{F}}} - \mu_1 \right) \|\mathbf{A}_i^{\text{H}}\|_{\text{F}} \\ &= \sqrt{2} \min_i (\|\mathbf{P}\|_{\text{F}} - \mu_1 \|\mathbf{A}_i^{\text{H}}\|_{\text{F}}). \end{aligned} \quad (33)$$

The coherent code matrix that minimizes the last term in (33) is the one with the maximum norm. For the case of rectangular QAM, we have that  $\max_i \|\mathbf{A}_i^{\text{H}}\|_{\text{F}} = \sqrt{2}M^{\frac{L^{\text{Q}}-1}{2}}q$ . Using this result in (33) yields

$$\mu_1 q = \sqrt{2}(\|\mathbf{P}\|_{\text{F}} - M(L^{\text{Q}} - 1)\mu_1 q). \quad (34)$$

Consequently,

$$\mu_1 = \frac{\|\mathbf{P}\|_{\text{F}}}{q} \frac{\sqrt{2}}{1 + M(L^{\text{Q}} - 1)}. \quad (35)$$

We continue computing the value of  $\mu$  that satisfies the power split in [1], which we express as  $\mu_2$ . This power split is given in [1] in terms of the portion of power dedicated to the coherent code,

$$\alpha = \frac{P_{\text{A}}}{P_{\text{A}} + P_{\text{P}}} = \frac{P_{\text{A}}}{P_{\text{A}} + \|\mathbf{P}\|_{\text{F}}^2}. \quad (36)$$

Since all the QAM symbols are equiprobable, the average power used in the coherent code is

$$\begin{aligned} P_{\text{A}} &= \frac{2(T-M)M}{L^{\text{Q}}} \sum_{i=1}^{L^{\text{Q}}} \mu_2^2 q^2 \left( i - \frac{L^{\text{Q}} + 1}{2} \right)^2 \\ &= 2(T-M)M\mu_2^2 q^2 \frac{(L^{\text{Q}})^2 - 1}{12}. \end{aligned} \quad (37)$$



Using (36) and (37), we obtain

$$\mu_2 = \frac{\|\mathbf{P}\|_F}{q} \sqrt{\frac{\alpha}{1-\alpha} \frac{1}{(T-M)M} \frac{6}{(L^Q)^2 - 1}}, \quad (38)$$

where, as indicated in [1],  $\alpha = \frac{1}{2}$  if  $T = 2M$  and  $\alpha = \gamma - \sqrt{\gamma(\gamma - 1)}$  if  $T > 2M$ , where  $\gamma = \frac{(T-M)(M+\rho T)}{\rho T(T-2M)}$ .

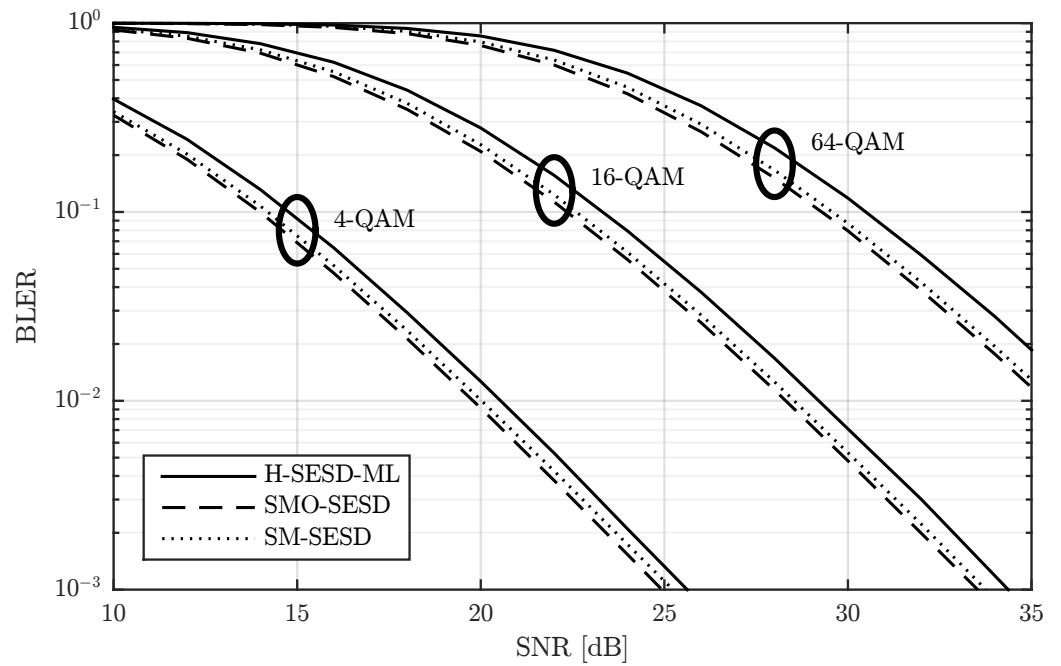
Finally, as proposed in the previous section, we scale the coherent code by  $\mu = \min(\mu_1, \mu_2)$ . In the following section, we compare the performance of these constellations with the performance of a PSAM and its orthonormalization.

## 6. Constellation Performance Analysis

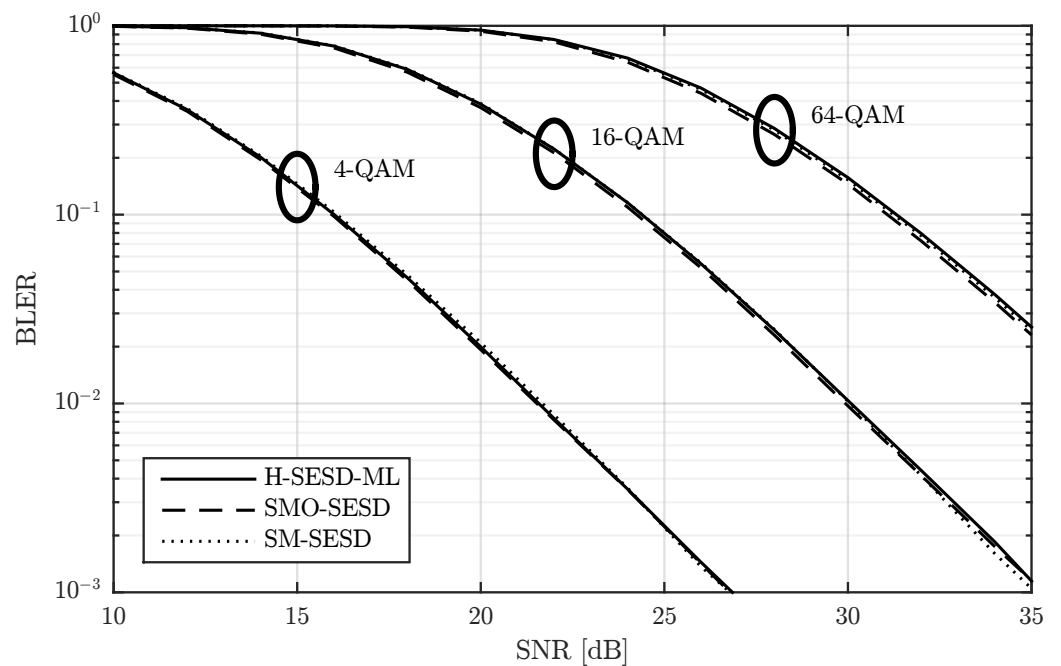
In this section, we analyze and compare the performance of three transmitter-receiver configurations. Two of those configurations were briefly analyzed in Section 4. In particular, the three configurations are as follows:

- SM-SESD: The transmitter sends PSAM matrices constructed from a scaled identity pilot matrix and SM of QAM symbols. The receiver uses an LS channel estimator and a SESD.
- SMO-SESD: The same configuration as the previous one, but the matrices are orthonormalized before transmission.
- H-SESD-ML: In this configuration, the transmitter uses a constellation generated using the hypercube method with an identity pilot matrix and rectangular QAM-based coherent codes, as described in Sections 5.3 and 5.4. The receiver uses the reception technique described in Section 5.1, i.e., it estimates the transmitted signal using a PSAM and a USTM receiver in two phases. In the first phase, the receiver selects one candidate from each PSAM. To this end, it assumes that the transmitted signal belongs to a given PSAM and uses a LS channel estimator and a SESD to estimate the transmitted signal (one for each PSAM). In the second phase, the USTM receiver assumes that the constellation used by the transmitter was the set of candidates and selects the received matrix from this set using an ML detection.

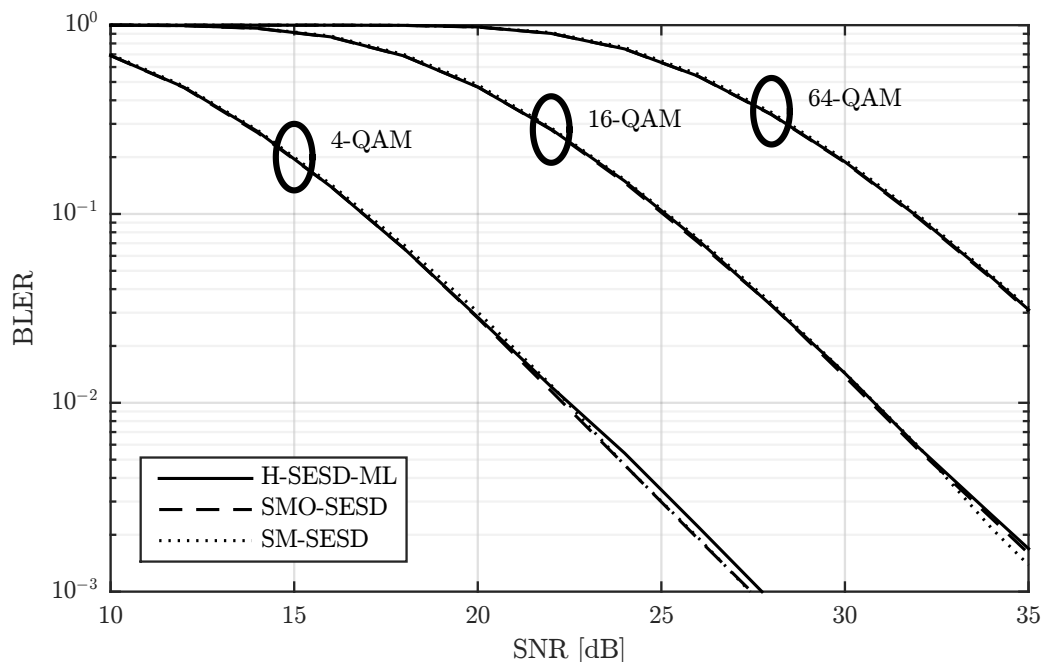
We first compare the BLER of these configurations for  $M = N = 2$ ,  $T$  equal to 4, 8 and 16, and for three QAM sizes: 4-QAM, 16-QAM, and 64-QAM. The BLER is depicted in Figures 9–11 for  $T = 4$ ,  $T = 8$ , and  $T = 16$ , respectively. The figures show that the performance of the three configurations is very similar. SMO-SESD is slightly better than SM-SESD and the performance of H-SESD-ML with respect to the other two configurations depends on the particular case. However, it is important to note that the constellation used in the H-SESD-ML configuration has more elements than those used in the other configurations. In particular, the constellations of the SM-SESD and SMO-SESD have  $(L^Q)^{2M(T-M)}$  elements, and the constellation used in the H-SESD-ML configuration has  $\frac{T}{M}(L^Q)^{2M(T-M)}$  elements, where  $L^Q$  equals 2, 4 and 8 for 4-QAM, 16-QAM and 64-QAM, respectively. As a consequence, the bits transmitted per channel use are different for each configuration, as shown in Table 1. Therefore, the BLER obtained with the H-SESD-ML configuration cannot be directly compared.



**Figure 9.** Performance comparison of three PSAMs constructed from rectangular QAM, their orthonormalization, and a constellation constructed using the hypercube method with the same QAM. All results are obtained for a channel with  $T = 4$  and  $M = N = 2$ .



**Figure 10.** Performance comparison of three PSAMs constructed from rectangular QAM, their orthonormalization, and a constellation constructed using the hypercube method with the same QAM. All results are obtained for a channel with  $T = 8$  and  $M = N = 2$ .

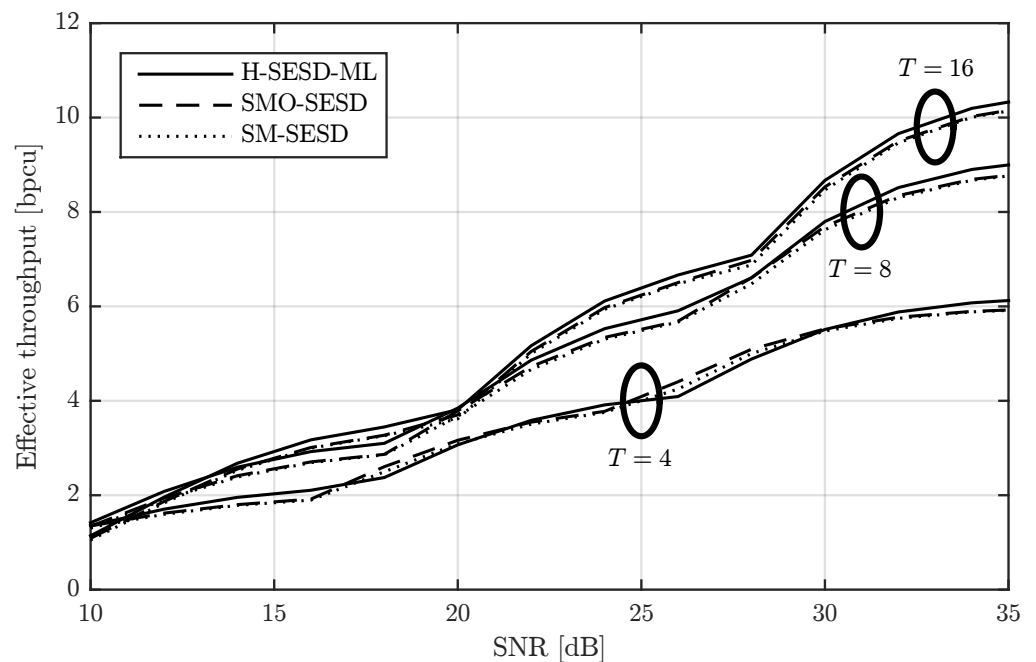


**Figure 11.** Performance comparison of three PSAMs constructed from rectangular QAM, their orthonormalization, and a constellation constructed using the hypercube method with the same QAM. All results are obtained for a channel with  $T = 16$  and  $M = N = 2$ .

**Table 1.** Bits per channel use transmitted for each configuration.

$T$	QAM	H-SESD-ML	SMO-SESD	SM-SESD
4	4	2.25	2	2
4	16	4.25	4	4
4	64	6.25	6	6
8	4	3.25	3	3
8	16	6.25	6	6
8	64	9.25	9	9
16	4	3.6875	3.5	3.5
16	16	7.1875	7	7
16	64	10.6875	10.5	10.5

In order to perform a fairer comparison, we computed the effective throughput of the three configurations as the nominal throughput in Table 1 times one minus the BLER in Figures 9–11. The maximum of the effective throughput from modulations 4-QAM, 16-QAM and 64-QAM is shown in Figure 12. Mathematically, letting  $R_i$  and  $\text{BLER}_i(\rho)$  be the nominal throughput and the BLER for an SNR  $\rho$  of the  $i$ -th modulation, respectively, the effective throughput  $R_{\text{eff}} = \max_i R_i(1 - \text{BLER}_i(\rho))$  is shown in Figure 12. The figure shows that H-SESD-ML is the best configuration for  $T = 8$  and  $T = 16$ , and it is also the best configuration for some SNR ranges for  $T = 4$ . With respect to SM-SESD, SMO-SESD increases the effective throughput around 0.04 bpcu on average for the three values of the channel block length  $T$ , whereas the increment provided by H-SESD-ML is around 0.06 bpcu for  $T = 4$ , 0.2 bpcu for  $T = 8$  and 0.17 bpcu for  $T = 16$ .



**Figure 12.** Effective throughput comparison of a PSAM constructed from rectangular QAM, its orthonormalization, and a constellation constructed using the hypercube method with the same QAM. This throughput is the maximum from that obtained with 4-QAM, 16-QAM and 64-QAM, and for  $M = N = 2$ .

## 7. Conclusions

We studied the relationship between PSAMs and USTMs. To do this, we introduced a new definition of PSAM space-time matrices, in which the pilot matrix can be spread out across the coherence interval using a tall unitary matrix. This unitary matrix and its orthogonal complement enables a graphical representation of these constellations. Using this definition, we found a map that transforms any PSAM into a USTM and vice versa. This map indicates a PSAM-USTM equivalence in terms of the spanned subspaces.

We used this equivalence and the graphical representation to define new USTM construction methods based on different PSAMs and their equivalent USTMs. In particular, firstly, we proposed the use of the equivalent USTM of a PSAM. Secondly, inspired by the graphical representation, we proposed to combine the equivalent USTMs of several PSAMs of which the pilots are carried by different tall unitary matrices. The resulting constellations can be interpreted as a PSAM where information is also encoded in the pilot position inside the channel coherence interval. Using the correct power distribution between pilots and signals, our results show that the obtained USTMs have higher effective throughput than the PSAMs used to construct them.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data will be provided upon request.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Hassibi, B.; Hochwald, B.M. How much training is needed in multiple-antenna wireless links? *IEEE Trans. Inf. Theory* **2003**, *49*, 951–963. [\[CrossRef\]](#)
2. Marzetta, T.L.; Hochwald, B.M. Capacity of a Mobile Multiple-Antenna Communication Link in Rayleigh flat fading. *IEEE Trans. Inf. Theory* **1999**, *45*, 139–157. [\[CrossRef\]](#)
3. Hochwald, B.M.; Marzetta, T.L. Unitary Space-Time Modulation Multiple-Antenna Communications in Rayleigh flat fading. *IEEE Trans. Inf. Theory* **2000**, *46*, 543–564. [\[CrossRef\]](#)
4. Zheng, L.; Tse, D.N.C. Communication on the Grassmann Manifold: A geometric Approach to the noncoherent multiple-antenna channel. *IEEE Trans. Inf. Theory* **2002**, *48*, 359–383. [\[CrossRef\]](#)
5. Schwarz, S. Non-Coherent Broadcasting based on Grassmannian Superposition Transmission. In Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT), Melbourne, Australia, 12–20 July 2021; pp. 2882–2887.
6. McCloud, M.L.; Brehler, M.; Varanasi, M.K. Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel. *IEEE Trans. Inf. Theory* **2002**, *48*, 1186–1194. [\[CrossRef\]](#)
7. Gohary, R.H.; Davidson, T.N. Non-coherent MIMO communication: Grassmannian constellations and efficient detection. *IEEE Trans. Inf. Theory* **2009**, *55*, 1176–1205. [\[CrossRef\]](#)
8. Cuevas, D.; Beltrán, C.; Santamaria, I.; Tuček, V.; Peters, G. A fast algorithm for designing Grassmannian constellations. In Proceedings of the 25th International ITG Workshop on Smart Antennas (WSA), French Riviera, France, 10–12 November 2021.
9. Ngo, K.H.; Yang, S.; Guillaud, M.; Decurninge, A. Joint constellation design for noncoherent MIMO multiple-access channels. *IEEE Trans. Inf. Theory* **2022**, *68*, 7281–7305. [\[CrossRef\]](#)
10. Hochwald, B.M.; Marzetta, T.L.; Richardson, T.J.; Sweldens, W.; Urbanke, R. Systematic design of unitary space-time constellations. *IEEE Trans. Inf. Theory* **2000**, *46*, 1962–1973. [\[CrossRef\]](#)
11. Marzetta, T.L.; Hassibi, B.; Hochwald, B.M. Structured unitary space-time constellations. *IEEE Trans. Inf. Theory* **2002**, *48*, 942–950. [\[CrossRef\]](#)
12. Jing, Y.; Hassibi, B. Unitary Space-Time Modulation via Cayley Transform. *IEEE Trans. Signal Process.* **2003**, *51*, 2891–2904. [\[CrossRef\]](#)
13. Kammoun, I.; Cipriano, A.M.; Belfiore, J.C. Non-coherent codes over the Grassmannian. *IEEE Trans. Wirel. Commun.* **2007**, *6*, 3657–3667. [\[CrossRef\]](#)
14. Ngo, K.H.; Decurninge, A.; Guillaud, M.; Yang, S. Cube-Split: A Structured Grassmannian Constellation for Non-Coherent SIMO Communications. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 1948–1964. [\[CrossRef\]](#)
15. Cuevas, D.; Álvarez-Vizoso, J.; Beltrán, C.; Santamaria, I.; Tuček, V.; Peters, G. A measure preserving mapping for structured Grassmannian constellations in SIMO channels. In Proceedings of the 2022 IEEE Global Communications Conference (GLOBECOM), Rio de Janeiro, Brazil, 4–8 December 2022.
16. El Gamal, H.; Damen, M.O. Universal space-time coding. *IEEE Trans. Inf. Theory* **2003**, *49*, 1097–1119. [\[CrossRef\]](#)
17. Krishnamoorthy, A.; Anastasopoulos, A. Code and Receiver Design for the Noncoherent Fast-Fading Channel. *IEEE J. Sel. Areas Commun.* **2005**, *23*, 1769–1778. [\[CrossRef\]](#)
18. Chen, Y.M.; Ueng, Y.L. Turbo Coded Noncoherent Space-Time Modulation Using Information-Bearing Pilots and Spatial Multiplexing. *IEEE Trans. Commun.* **2011**, *59*, 1543–1554. [\[CrossRef\]](#)
19. Zhu, Y.J.; Zhang, J.K.; Yu, H.Y. Linear Receivers for Full-Diversity Training Space-Time Block Codes. *IEEE Trans. Veh. Technol.* **2012**, *61*, 2884–2889. [\[CrossRef\]](#)
20. Dayal, P.; Brehler, M.; Varanasi, M.K. Leveraging coherent space-time codes for noncoherent communication via training. *IEEE Trans. Inf. Theory* **2004**, *50*, 2058–2080. [\[CrossRef\]](#)
21. Zhao, W.; Leus, G.; Giannakis, G.B. Algebraic Design of Unitary Space-Time Constellations. In Proceedings of the IEEE International Conference on Communications, 2003, ICC '03, Anchorage, AK, USA, 11–15 May 2003; pp. 3180–3184.
22. Roman, S. *Advanced Linear Algebra*, 3rd ed.; Springer: New York, NY, USA, 2008.
23. Biguesh, M.; Gershman, A.B. Training-Based MIMO Channel Estimation: A Study of Estimator Tradeoffs and Optimal Training Signals. *IEEE Trans. Signal Process.* **2006**, *54*, 884–893. [\[CrossRef\]](#)
24. Garcia-Molla, V.M.; Vidal, A.M.; Gonzalez, A.; Roger, S. Improved maximum likelihood detection through sphere decoding combined with box optimization. *Signal Process.* **2014**, *98*, 284–294. [\[CrossRef\]](#)