# Twin Tree Hierarchy: A Regularized Approach to Construction of Signature Matrices for Overloaded CDMA

Amiya Singh, Poonam Singh, Arash Amini, Farokh Marvasti

**Abstract**

We consider two crucial problems to feature a new set of Signature Matrices with Orthogonal Subsets (SMOS) for overloaded Code Division Multiple Access (CDMA). The first one concerns whether the existence of non-ternary version of SMOS is realizable, and if so, what difference lies in the pattern of interference and error performance, as compared to ternary. The second one is about identifying the superior one. To address the first, we propose the $2k$-ary SMOS, where the binary alphabets in each of the $k$ constituent (orthogonal) subsets are unique. Despite the similarity in twin tree hierarchy of interference, its non-uniformity brings contrast to the analysis, and outcome including the existence of optimal signatures. In response to the second, based on the behavior of interference, the whole set is split into $2k$ number of subsets thereby segregating the simulations too. At higher loading, where for larger and smaller subsets the superiority is detected for the $2k$-ary and ternary respectively, counter-intuitive deviations are noticed once the loading gets reduced. For the maximization in user capacity to be $50\%$, superiority is featured by the $2k$-ary, but beyond, it becomes a conditional entity. To validate, we present few logical anomalies and trade-off involving suitable operational metrics followed by rigorous simulations.

**Keywords**

*Overloaded CDMA, Uniquely Decodable Codes, Orthogonality, Multiple Access Interference (MAI), Twin Tree Hierarchy, Decoder, Bit Error Rate (BER).*

---

A. Singh and P. Singh are with Department of Electronics and Communication Engg., National Institute of Technology Rourkela, India. (amiyasingh87@gmail.com, psingh@nitrkl.ac.in)

A. Amini and F. Marvasti are with Advanced Communication Research Institute (ACRI), Department of Electrical Engg., Sharif University of Technology, Tehran, Iran. (aamini@sharif.edu, marvasti@sharif.edu)

# I. INTRODUCTION

The system using Code Division Multiple Access (CDMA) assigns each user a distinct code or signature to access a common communication channel and is generally represented as

$$\mathbf{y} = \mathbf{r} + \mathbf{n} \tag{1}$$

where $\mathbf{y}$ and $\mathbf{r} = \mathbf{CAx}$ are the noisy and noiseless received vector for $\mathbf{C} \triangleq [\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \cdots, \mathbf{c}_M]$ being a $N \times M$ matrix representing $M$ signatures of length $N$, $\mathbf{x} = [x_1, x_2, x_3, \cdots, x_M]$ represents the input data vector corresponding to all $M$ users or signatures, $\mathbf{A} = \mathbf{I}_{M\times M}$ = Identity matrix with diagonal elements representing the amplitudes assuming the system to be perfectly power controlled, and $\mathbf{n}$ denotes zero mean AWGN vector. While for conventional (underloaded) CDMA the maximum achievable value of $M$ is $N$ (i.e.; $M \leq N$), this capacity limit can be maximized beyond $N$ (i.e.; $M > N$) using the feature of overloaded CDMA, which subsequently leads to the significant rise in the loading (overloading) factor $\beta = ({}^{M}\!/_{N})$.

Over a decade, the popularity of CDMA besides the military applications can be well-speculated from its growing commercialization of the wireless communication systems like Third Generation (3G) cellular architecture using Wideband CDMA (WCDMA). Although for the Fourth Generation (4G) architecture, the scope of participation of CDMA has been fully obliterated, its further advancement in terms of Sparse Coded Multiple Access (SCMA) [1] using Low Density Signature (LDS) [2] has regained its priority in terms of being prospectively provisioned for the Fifth Generation (5G) [3] cellular architecture.

## A. Research Overview

The concept of overloading in CDMA as a solution to the strict capacity limit of its conventional counterpart has drawn frequent interest of the researchers in the past and continuously evolved though several phases. The extensive research in the field can be broadly classified into problems like: (a) designing of the signature matrices (Uniquely Decodable (UD) [4]–[11] and non-UD [12]–[14]), (b) development of reusing techniques [15]–[17] for the existing non-overloaded matrices, (c) designing of Multi User Decoders (MUD) [18]–[22], (d) adding system level improvisations for improved error performance [23], [24], and (e) study of capacity and performance under different channel and input conditions [25]–[27], for synchronous or asynchronous [28] transmission. Where a matrix $\mathbf{C}$ is considered UD over $\mathbf{x}$, if for $\mathbf{x}_1 \neq \mathbf{x}_2$

the inequality $\mathbf{C}\mathbf{x}_1 \neq \mathbf{C}\mathbf{x}_2$ is true for $\mathbf{x}_1$ and $\mathbf{x}_2$ denoting two different input vectors, its major applications in literature are shown to be in the field of multiuser encoding for Physical Layer Network Coding (PLNC) [29], overloaded CDMA [4]–[11], [25], [26], and T-user channels [30]–[34].

### *B. Background Study*

Over a decade, the group based approach for the construction of signature matrices for overloaded CDMA has drawn the attention, significantly. Be it UD or non-UD, the usual practice adopted in this method for the formation of the overloaded matrix $\mathbf{C}_{N\times M} = \left[\mathbf{H}_{N\times N}|\mathbf{O}_{N\times(M-N)}\right]$ $(M > N)$ is to add another group (set) of random or quasi-orthogonal signature sequences $\mathbf{O}_{N\times(M-N)}$ to the existing orthogonal matrix (usually Hadamard denoted as $\mathbf{H}_{N\times N}$) and thus achieve suitable gain in the value of $\beta$. For CDMA system using $\mathbf{C}_{N\times M}$ as the signature matrix, the metrics [28] primarily affecting the recovery performance of an arbitrary user (for noisefree transmission) are the level of total peak cross-correlation (TPC) and peak auto-correlation (PA) e.g.; $\text{TPC}(\mathbf{c}_a) \triangleq \sum_{i=1}^{M-1}\sum_{n=1}^{N} c_{an}c_{in}$ and $\text{PA}(\mathbf{c}_a) \triangleq \sum_{n=1}^{N} c_{an}c_{an}$, where $\mathbf{C}_{N\times M} = [\mathbf{c}_1\mathbf{c}_2\cdots\mathbf{c}_M]$ and $\mathbf{c}_a = [c_{a1}c_{a2}\cdots c_{aN}]$. With the enhancement in loading condition, the level of TPC representing the total multiple access interference (MAI) on the signatures in each group also increases and can be classified into two types: intergroup MAI (on each signature in one group due to all signatures in the other), intragroup MAI (on each signature of one group due to other signatures of the same group). To avoid the adverse effect of excess MAI on the recovery performance, the technique of iterative interference cancellation is implemented in [12], [23] followed by the use of matched filter (MF) detection in each iteration. Still the impact of MAI can hardly have its complete elimination and therefore, achieving the errorfree performance, even in the absence of noise, becomes impossible. Additionally, the massive complexity of such detectors also drives its implementation aspect to further doubtfulness.

In 2009 [9], the UD version of the group based construction is introduced where the signature matrices of larger dimensions from its smaller counterparts (core matrix) are formed using the tensor product with an invertible matrix (mostly Hadamard). For decoding, a two stage simplified decoder (unlike the complex iterative structures for the non-UD case) is also introduced where the detection is based on the logic of a simplified maximum likelihood decoding (MLD). The result shows that the error performance of the system using core matrix of moderate or larger dimension is found to be better than that of smaller dimension. But, selection of the core matrix with relatively large dimension, on the other hand, leads to the rise in decoder's complexity. This inconvenience is due to the rise in the number of secondary users (Hadamard refers to primary users), usually detected through the method of joint MLD [35]. Very recently, in [5], further improvisations are added towards the modification of the simplified MLD (SMLD), but it is related to the improvement in error performance only. Afterwards, hardly, any further attempt has been made to linearize the decoding structure so that the variance in its complexity in comparison to the conventional decoding can be maintained at a moderate level.

### *C. Motivation*

The concept of UD based construction in the context of multi-user communication is no more new to literature, and exist in the form of several recursive and non-group based [11], [30]–[34] architecture. However, most of them are hardly accompanied by an efficient decoder (both in complexity and error performance) for noisy transmission. Where the recovery performance of a decoder in the multi-user environment like CDMA is directly influenced by the level of total MAI on the signatures, most of the proposed systems [5], [6], [9]–[12], [15], [21], [23], [24], [26], [30]–[34], [36]–[39] in literature deal with a type of MAI pattern, which is completely random in nature. Furthermore, due to the unpredictability and non-uniformity in its distribution among the signatures, it's accurate estimation (for each signature) for the purpose of cancellation is considered nearly impossible due to the demand of complex methods involving MLD or iterative decoders. Consequently, the overall complexity of the system shows a gigantic rise. Under such constraints, concentrating more on the tactical designing of the signature matrices to generate a regularized, predictable, and balanced pattern involving the PA and TPC (i.e.; the sum of inter and intra-group TPC or MAI) can yield better recovery efficiency. Not only it can bring noticeable improvement in the form of complete elimination of the available MAI and offer better error performance, but also in retaining the simplicity of the decoder due to the approval to use linear decoding blocks like MF. While looking for such unique and efficient constructions, the *tree structured* correlation pattern of the optimal construction in [18] immediately draws our attention. However, once again the challenge related to massive complexity blocks the scope of its implementation aspect. But, interestingly enough, the above motivation has been recently addressed in [4] to propose the ternary signature matrices with orthogonal subsets (SMOS).

Here, the proposed decoder exploits the defined hierarchy of correlation to attain simplicity as well as, offering better performance over the binary random and Welch bound equality (WBE) sequences using a low complex multi-user decoder (MUD) [19].

In [4], the introductory discussion on SMOS is limited to the recursive construction of the ternary matrices, the illustration of the *twin tree hierarchy* of correlation pattern, designing of the simplified decoder followed by the validation of its errorless decoding, and above all, the performance analysis for noisy transmission. Furthermore, the ternary SMOS being a subset of the matrices of [30]–[32], [38] implies that their efficient decoding for noisy channel is also possible, but at the cost of the sacrifice in terms of the loss in asymptotic equality (see section II (A), [4]). As a distinct observation, the unprecedented behavior in error performance of the individual orthogonal subsets is recorded, for which the progressive variation in cardinality and pattern of MAI are shown to be collectively responsible. However, the important queries those have remained unanswered and hence, become our source of encouragement for further research in this paper are:

- if there exist any non-ternary constructions portraying the similar hierarchy in correlation,
- if yes, then how its recursive approach of construction is going to vary from the ternary,
- how differently the error performance of its individual subsets do respond to noisy channel ,
- whether there prevails any scope to achieve optimality, either fully or partially, and
- ultimately, which of them (ternary or non-ternary)has the superiority in error performance.

### *D. Contribution and Structure of the Paper*

In this paper, we address the above queries chronologically. First, we propose the non-ternary version of SMOS, where the elements used in the construction of each subset are binary, and the set of binary alphabets vary from one subset to the other. In total, $2k$ number of alphabets are involved in the construction of the SMOS with index-$k$ (i.e.; $\mathbf{C}^k$). So, we call it $2k$-ary SMOS. Unlike the case of ternary SMOS, $50\%$ signatures of its largest subset (binary) are recognized to be optimal. While both types of constructions, in overall, replicates a twin tree structured correlation among the signatures, the difference in the value and pattern of their correlation coefficients is found to be noteworthy. Hence expecting the proportional deviation in error performance of their individual subsets over ternary is easily predictable and also verified through simulation. The fact that makes this contribution more productive is its approach of analysis pf the error performance, which strongly emphasizes on the overall study of error performance to be partitioned into that of the smaller subsets, rather than just accepting the average bit error rate (BER) of the whole SMOS for evaluation. While for ternary, this classification results in $k$ separate subsets, that for the $2k$-ary, to validate the prevailing non-uniformity in MAI split it further into two smaller counterparts, and thus, produces $2k$ number of subsets in total. With this extensive segregation, the objective to compare the error performance of $2k$-ary and ternary demands a large number of simulations to be run. Moreover, due to multiple abnormalities and non-uniformity captured in the simulation results, the extraction of a concrete decision about their superiority (ternary versus $2k$-ary) also gets complex. To offer a simplified explanation, we introduce few logical anomalies and trade-off that collectively bring sufficient insight towards the apprehension of their overall performance. By and large, the rigorous analysis of the MAI hierarchy, comparative elaboration of the error performance under different loading conditions, the interesting outcomes and above all, their logical validations imbibe further gravity into the whole contribution.

Rest of the paper is organized as follows. Section II describes the brief introduction to the construction of several structures of ternary SMOS followed by their generalization. In Section III, we elaborate the discussion on the construction and features of the $2k$-ary SMOS. Section IV, being an important part, emphasizes on the error performance of $2k$-ary and its significant deviation from that of ternary followed by appropriate explanation. In Section V, the simulation results are logically analyzed. Finally, the conclusion is presented in Section VI.

## II. Review of Ternary SMOS

Let us start from the introductory literature on SMOS in [4]. Based on its discussion, the matrix $\mathbf{C}_{N_k \times M_k}$ is said to be SMOS over the input $\{0, 1, -1\}$, if the following conditions are satisfied.

- $\mathbf{C}^k$ is uniquely decodable over $\{0, 1, -1\}^{M_k}$.
- $\mathbf{C}^k$ comprises of $k$ orthogonal subsets, such that $\mathbf{C}^k = \left[\mathbf{C}_1^k | \mathbf{C}_2^k | \cdots | \mathbf{C}_k^k\right]$, where the number of signatures in subset-$\mathbf{C}_i^k$ and the value of the effective spreading gain ($N_{ef}$) for a signature in it (defined as the cardinality of a signature) is $\frac{N_k}{2^{i-1}}$.
- $\mathbf{Det}(\rho) = 0$, for $\rho = (\mathbf{C}^k)^T \mathbf{C}^k$

• The level of PA of an arbitrary signature in subset-$\mathbf{C}_i^k$ must be greater than the TPC from $(k-i)$ successive subsets: $\mathbf{C}_{i+1}^k, \mathbf{C}_{i+2}^k, \ldots, \mathbf{C}_k^k$. In fact, this is the sufficient condition for the MF to decode each subset with no error. Mathematically, it is described by the relation

$$\rho_{ii}(u,u) > \sum_{j=i+1}^{k} \sum_{v=1}^{\frac{N_k}{2^{j-1}}} \rho_{ij}(u,v). \tag{2}$$

In the next section and onwards, our objective not only just lies in proposing a new set of SMOS, but also in presenting a comparative overview of the features of the new construction in contrast to its existing counterpart (i.e.; Ternary SMOS [4]). Apart from refining the concepts of SMOS through a generalized perspective, our approach also provides an opportunity to reveal about, which of them carries the ultimate superiority in error performance for noisy transmission.

## III. Construction of 2k-ary SMOS and Optimality

### A. Basis Matrix and its Correlation Structure

In [4], the role of the basis (fundamental) matrix $\mathbf{B} = \frac{1}{\sqrt{2}} \left[ \begin{matrix} + & + \\ + & - \end{matrix} \middle| \begin{matrix} + \\ 0 \end{matrix} \right] = [\mathbf{c}_{11}\mathbf{c}_{12}|\mathbf{c}_{21}]$ in recursively driving the construction and decoding has been explained in detail. So, prior to proposing the 2k-ary SMOS, diverting equal emphasis on analyzing the structure of the its basis matrix $\mathbf{B}'$ is of high importance, where

$$\mathbf{B}' = \frac{1}{\sqrt{2}} \left[ \begin{matrix} + & + \\ + & - \end{matrix} \middle| \begin{matrix} 1/2 \\ 1/2 \end{matrix} \right]. \tag{3}$$

First, let us draw a comparison of the correlation matrices ($\rho$ and $\rho\prime$) of the respective basis sets i.e.; $\mathbf{B}$ and $\mathbf{B}'$, where both

$$\rho = \mathbf{B}^T\mathbf{B} = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 1/2 & 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \end{bmatrix}, and \tag{4}$$

$$\rho\prime = (\mathbf{B}')^T\mathbf{B}' = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/4 \end{bmatrix} = \begin{bmatrix} \rho\prime_{11} & \rho\prime_{12} & \rho\prime_{13} \\ \rho\prime_{21} & \rho\prime_{22} & \rho\prime_{23} \\ \rho\prime_{31} & \rho\prime_{32} & \rho\prime_{33} \end{bmatrix}, \tag{5}$$

carrying an organized correlation pattern, can be translated to a meaningful two level tree hierarchy, as shown in Fig. 1. Below, we arrange its important outcomes.

• The node (signature) $\mathbf{c}_{21}$ at the bottom level can be interpreted as the root to two signatures (i.e.; $\mathbf{c}_{11}$: left child, $\mathbf{c}_{12}$: right child) of the top level.

• For $\mathbf{B}$, there exists equal level of correlation between $\mathbf{c}_{11}$, $\mathbf{c}_{21}$, and $\mathbf{c}_{12}$, $\mathbf{c}_{21}$ i.e.; $\rho_{13} = \rho_{23} = 1/2$ (uniform). In contrast, the correlation is not equal for $\mathbf{B}'$ i.e.; $\rho\prime_{13} = 1/2$, $\rho\prime_{23} = 0$ (non-uniform)

• Unlike $\mathbf{B}$, the right child ($\mathbf{c}_{12}$) available in the topmost level is always *optimal* due to its complete orthogonality i.e.; $\rho\prime_{23} = 0$

• Besides cross-correlation, the difference in the level of auto-correlation is also noteworthy e.g; $\rho_{11} = \rho_{22} = \rho\prime_{11} = \rho\prime_{22} = 1$ (i.e.; $\rho\prime_{11} = \rho_{11}, \rho\prime_{22} = \rho_{22}$ ), where as $\rho_{33} = 1/2$, $\rho\prime_{33} = 1/4$ (i.e.; $\rho\prime_{33} \neq \rho_{33}$). It clearly indicates the mismatch existing in the level of auto-correlation for the signature that is not a part of the largest subset (binary Hadamard matrix).

• Similar to $\mathbf{B}$, for each signature in $\mathbf{B}'$, the criterion for errorless decoding (2) is also verified.

### *B. Recursive Method of Construction*

Based on the above developments in the structure of $\mathbf{B}'$, we propose the following recurrent approach for the construction of the matrices with larger dimension i.e.; $\mathbf{C}_{N_k \times M_k}$ for $N_k = 2^k$ where $k \in \mathbf{Z}^+$ e.g.;

• Initialize $\mathbf{C}^0 = [1]$ and find $\mathbf{C}^1 = \left(\frac{1}{\sqrt{2}}\mathbf{H}_{2\times 2} \otimes \mathbf{C}^0\right) = \frac{1}{\sqrt{2}}\mathbf{H}_2$

• For $k > 1$, $\mathbf{C}^k = \left(\frac{1}{\sqrt{2}}\mathbf{H}_{2\times 2} \otimes \mathbf{A}\right)$, where

$$\mathbf{A} = \left[\, \mathbf{C}^{k-1} \quad | \quad \left[\, 1/2^{k-1} \quad 1/2^{k-1} \quad \cdots \quad 1/2^{k-1} \,\right]^T_{1\times 2^{k-1}} \,\right]$$

Table II illustrates the construction of $2k$-ary SMOS for $k = 1, 2, 3$. As evident, for $\mathbf{C}^3 = \left[\mathbf{C}^3_1 | \mathbf{C}^3_2 | \mathbf{C}^3_3\right]$, the elements of each subset are binary and these binary elements or alphabets or symbols differ from one subset to other e.g.; $\mathbf{C}^3_1 \in \{1, -1\}^8$, $\mathbf{C}^3_2 \in \{1/2, -1/2\}^4$,$\mathbf{C}^3_3 \in \{1/2^2, -1/2^2\}^2$. Also, we may define it as $\mathbf{C}^3 \in \{1, 1/2, 1/2^2, -1/2^2, -1/2, -1\}^{14}$. In general, the matrix $\mathbf{C}^k \in \{1, 1/2, 1/2^2, 1/2^3, \cdots, -1/2^3, -1/2^2, -1/2, -1\}^{M_k}$ for its construction, requires $2k$ number of elements, where the nature of each subset, individually, is binary e.g.; $\mathbf{C}^k_1 \in \{1, -1\}^{N_k}$, $\mathbf{C}^k_2 \in \{1/2, -1/2\}^{N_k/2}$, $\mathbf{C}^k_3 \in \{1/2^2, -1/2^2\}^{N_k/2^2}$, $\cdots$, $\mathbf{C}^k_k \in \{1/2^{k-1}, -1/2^{k-1}\}^2$. The steps to prove the UD nature of the above construction are similar to that of the proof in section III (A) in [4].

### C. Twin Tree Hierarchy of MAI

From the discussion of the basis matrices (in section III A), as an important inference, the *two level* tree hierarchy of $\mathbf{B}'$ in comparison to that of $\mathbf{B}$ is found to be non-uniform. To have a broader vision of the statistics involved in the correlation pattern, let us look upon the *multi level* hierarchy for $2k$-ary SMOS $\mathbf{C}^k$, where the nodes of the tree at a particular level (depth) $l = 1, 2, \ldots, k$ (i.e., $l = k - i + 1$ for $1 \leq i \leq k$) collectively form the subset-$\mathbf{C}^k_{k-l+1}$. The following observations collaboratively provide a factual summarization.

• There exist two identical (twin) trees, each of which has its origin or root from the smallest orthogonal subset ( i.e.; $\mathbf{C}^k_k$ at the lowest level of the tree, $l = 1$). The nodes at the highest level of the tree ($l = k$) represent the largest subset: $\mathbf{C}^k_1$.

• Each node (parent) at level-$l$ can be interpreted to generate two child nodes for its next higher level i.e.; level-$(l+1)$, of which, the *left child* (connected by *solid lines* in Fig 2) is correlated to the parent node, where as no such correlation exists for the *right child* (connected by *dotted lines*).

• Each left child at level-$l$ is correlated to its left child in the subsequent upper levels (i.e.; level-$(l+1)$ to $k$) and lower levels (i.e.; level-$(l-1)$ to 1) e.g.;

$$\mathbf{c}^k_{(k)(2j-1)} = 2\mathbf{c}^k_{(k-1)(j)}, \ldots, \mathbf{c}^k_{(l+1)(2j-1)} = 2\mathbf{c}^k_{(l)(j)}, \mathbf{c}^k_{(l)(2j-1)} = 2\mathbf{c}^k_{(l-1)(j)}, \ldots,$$
$$\mathbf{c}^k_{(l-(l-2))(2j-1)} = 2\mathbf{c}^k_{(l-(l-1))(j)} \text{ for } 1 \leq j \leq 2^l$$

and *orthogonal* (zero correlation) to all its right child lying in the subsequent upper levels.

• On the other hand, each right child at level-$l$ is correlated to its left child in the subsequent upper levels only. To all other nodes of the whole tree, it is fully orthogonal. Therefore, the left child at each level confronts a relatively high level of MAI over the right and subsequently for $l = k$, the right child corresponding to each root node in $l = (k-1)$ is under *zero MAI* and hence, justifies its *optimality*.

With an aim to offer better visualization of the differences existing in the correlation pattern, between the ternary and $2k$-ary, we present the correlation matrices for $\mathbf{C}^3 = \left[\mathbf{C}^3_1 | \mathbf{C}^3_2 | \cdots | \mathbf{C}^3_3\right]$ and the corresponding twin tree structured hierarchy in Fig. 2 and 3 respectively. From the perspective of the basis matrix, it is important to note that simply substituting $\mathbf{c}_{21} = [1/2 \;\; -1/2]^T$ in (3) also retains the power of $\mathbf{B}'$ to generate the desired SMOS, with the only variation added in the form of the switching of the behavior between the left and right child. In other words, the pattern of MAI related to the left and right child corresponding to the present construction (Table II) will be conveyed by the right and left child of the new one.

### D. Decoding

In [4], the proof to the errorfree decoding of the ternary SMOS is explained in Lemma 1, where the criterion in (2) is found to be regulating the errorfree validation at each stage of detection. Due to the resemblance in construction and pattern of MAI, an analogous approach can also be adopted to justify the errorless decoding of $2k$-ary. Fig. 5 describes the block diagram of the decoder where the decoding of each subset is achieved using the simple logic of MF, thus, keeping a high degree of linearity in decoder's overall design over the MLD [35] and SMLD [5], [9].

## IV. Error Performance Analysis: A Comparative Approach

### A. Expression for Average BER

Now, based on the discussion from the previous section, let us first rewrite the expression in (1) so as to support our analysis involving the left and right child existing in each subset of $2k$-ary e.g.;

$$\mathbf{y} = \sum_{u=1}^{k} \mathbf{C}_{uL}^{k} \mathbf{x}_{uL} + \sum_{v=1}^{k} \mathbf{C}_{vR}^{k} \mathbf{x}_{vR} + \mathbf{n}. \tag{6}$$

To have the expression for the average BER of the left and right child of subset-$\mathbf{C}_i^k$, it is logical to modify the expression of BER for the individual subsets of ternary SMOS i.e.;

$$P_e^{ij} = Q\left(\sqrt{\frac{\left(\frac{N_{ef_i}}{N}\right)^2 \mathrm{E}\left(x_{ij}^2\right)}{\sum\limits_{u=i+1}^{k} \rho_{iu(j)}^2 + 4\sum\limits_{v=1}^{i-1} \rho_{iu(j)}^2 P_e^v + \mathrm{E}\left(n_{i(j)}^2\right)}}\right) \text{ (Equation (17), [4]).}$$

Finally, we present the corresponding expressions in Table III, which clearly validates the difference existing in the average error performance of different smaller subsets, thereby embracing the existing outcomes related to the non-uniformity in MAI. Afterwards, regardless of the type, our approach to analyze the behavior of each subset (say $\mathbf{C}_i^k$) is therefore divided in terms of its constituents subsets (i.e.; $\mathbf{C}_{iL}^k$ and $\mathbf{C}_{iR}^k$).

### B. *Analysis Involving Fundamental Metrics: A Closer Overview*

Before we switch to the error performance analysis in Section V, it is highly imperative to consider variation in the behavior of the basic metrics (i.e.; TPC and PA) controlling the quality of recovery of the decoder. According to Fig. 5, for the errorless MF detection of different subsets (i.e.; $\mathbf{C}_1^k, \mathbf{C}_2^k, \mathbf{C}_3^k, \ldots, \mathbf{C}_k^k$) from the respective received vectors (i.e.; $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, …, $\mathbf{y}_k$.), the inequality in (2) must be satisfied. So, in Table IV, we describe the statistics involving these metrics for the decoding of $\mathbf{C}^6 = \left[\mathbf{C}_1^6|\mathbf{C}_2^6|\mathbf{C}_3^6|\mathbf{C}_4^6|\mathbf{C}_5^6|\mathbf{C}_6^6\right]$ ($N_k$=64 and $M_k$=126). We position them separately for both ternary and $2k$-ary SMOS, so that the difference in their values can be carefully scrutinized. Even if the smaller subsets ($\mathbf{C}_{iL}^6, \mathbf{C}_{iR}^6$) comprising the left and right child of subset-$\mathbf{C}_i^6$ (e.g.; $\mathbf{C}_i^6 = \left[\mathbf{C}_{iL}^6|\mathbf{C}_{iR}^6\right]$) are detected in the same stage of the decoder (stage-$i$), they are presented in separate (consecutive) rows, only to vindicate their difference in the pattern of MAI, as already discussed in section III (C). Below, we list the crucial remarks extracted from Table IV.

*a) TPC and PA: (Left Child versus Right Child):* For Ternary, both the left and right child in a subset (i.e.; $\mathbf{C}_{iL}^6$ and $\mathbf{C}_{iR}^6$) confronts identical level of TPC and PA. On the other hand, for $2k$-ary, the level of TPC on the right child is always lower than that of the left. But, no such mismatch is noted for the PA.

*b) TPC and PA: (Ternary versus $2k$-ary):* Except $\mathbf{C}_{1L}^6$ and $\mathbf{C}_{1R}^6$, the level of PA and TPC for all other subsets of $2k$-ary lies lower than that of the ternary i.e.; $(\text{PA})_{2k-ary} = \left(\sqrt{\text{PA}}\right)_{ternary}$ and $(\text{TPC})_{2k-ary} << (\text{TPC})_{ternary}$. For $\mathbf{C}_{1L}^6$ and $\mathbf{C}_{1R}^6$, the corresponding values are fully equal i.e.; $(\text{PA})_{2k-ary} = (\text{PA})_{ternary}$ and $(\text{TPC})_{2k-ary} = (\text{TPC})_{ternary}$.

*c) (PA−TPC): (Ternary versus $2k$-ary):* For $2k$-ary, first, the difference ((PA − TPC)) for the left and right child is significant for the larger subsets (i.e.; the higher level of the tree) and it shows a continuous fall as we proceed towards the smaller subsets (i.e.; lower level of the tree). Secondly, the gap between the difference is also more for the larger subset and gradually decreases towards the smaller ones. For the smallest subset (the bottommost level of the tree), the gap becomes zero, thus showing the difference to attain equality for both left and right child.

After the above description on the behavior of the metrics, an immediate retrospection may yield multiple estimations about their error performance. But, hardly any of them leads to a concrete projection of the reality. So, for a clear-cut summarization, we present the following inferences to provide an appropriate explanation towards the match between the theory (so far) and observations (simulation results in Section V).

### *C. Trade-off: PA versus TPC*

Consider the explanation in Section IV (Bb). In the context of $2k$-ary SMOS, it actually points to the trade-off associated with the level of PA and TPC for an arbitrary signature (left or right child) in a subset (except $\mathbf{C}_{1L}^{k}$). In fact, the structure of construction of $2k$-ary strictly allows the TPC on its signature in any subset to remain always lower than that of ternary. So, for noisy transmission, the level of BER of the former becoming lower than that of the later is expected. But, at the same time, significant fall in the level of PA for the same signature of $2k$-ary is also discovered, which on the other hand indicates the rise in BER level. Therefore, leading to a firm conclusion about their superiority becomes practically infeasible since, for a particular value of the signal to noise ratio (SNR), realization of both the outcomes are considered as mutually exclusive. Purposefully, we leave it to be addressed in the simulation section.

### *D. Anomaly 1: Impact of Effective Peak Correlation ($\rho_{ef}$)*

According to Table IV, regardless of being the left or right child ($\mathbf{C}_{iL}^{6}$ or $\mathbf{C}_{iR}^{6}$), the level of PA and TPC on a signature in a subset of $2k$-ary SMOS (except $\mathbf{C}_{1L}^{6}$) always remains lower than that of ternary. For $\mathbf{C}_{1L}^{6}$, the deviation regarding the equality of the level of the corresponding metrics is realized and hence, anticipating their error performance to be identical is quite apparent. However, keeping in mind about the variation of the signature alphabets and their positioning across each subset, the probability of our expectation turning into a counter-intuitive outcome can not be fully overlooked. The following explanation offers a logical validation to this inference.

Assume two different signatures belonging to two different code sets (not necessarily SMOS) e.g.; $\mathbf{c}_{a1} \in \mathbf{C}_a \triangleq [\mathbf{c}_{a1}, \mathbf{c}_{a2}, \cdots, \mathbf{c}_{aN}]$ and $\mathbf{c}_{b1} \in \mathbf{C}_b \triangleq [\mathbf{c}_{b1}, \mathbf{c}_{b2}, \cdots, \mathbf{c}_{bN}]$, where $\mathbf{C}_a$, $\mathbf{C}_b \in \{1, -1, 0\}^N$. Also, assume the net level of MAI on $\mathbf{c}_{a1}$ and $\mathbf{c}_{b1}$ (due to the non-zero cross-correlation from remaining $(N-1)$ signatures) to be equal. While for $\mathbf{C}_a$ and $\mathbf{C}_b$ to be binary i.e.; $\mathbf{C}_a$, $\mathbf{C}_b \in \{1, -1\}^N$ , the expression $\rho_{ij} = \sum_{n=1}^{N} c_{in} c_{jn}$ should be considered to measure the peak cross-correlation level between any two signatures ($\mathbf{c}_i = [c_{i1}, c_{i2}, \cdots, c_{iN}]$ and $\mathbf{c}_j =$

$[c_{j1}, c_{j2}, \cdots, c_{jN}]$) of length $N$ (for binary code set, $N_{ef_i} = N$ for $1 \leq i \leq N$), its importance looses its precision, if the code domain translates to ternary. This is because, for a ternary code space, the net peak cross-correlation on a signature is due to the non-zero elements (only) of other active signature sequences. Therefore, a more accurate tool of analysis is to define a new metric i.e.; *effective peak cross-correlation*, which is nothing but the normalized version of the peak cross-correlation. For instance, the effective peak-cross correlation due to $c_j$ on $c_i$ can be expressed as

$$\rho_{ef_{ij}} = \left(\tfrac{1}{N_{ef_i}}\right) \sum_{n=1}^{N} c_{in}c_{jn} \text{ ( for } N_{ef_i} < N_{ef_j})$$
$$=\left(\tfrac{1}{N_{ef_j}}\right) \sum_{n=1}^{N} c_{in}c_{jn} \text{ ( for } N_{ef_i} > N_{ef_i}).$$

Equivalently, the level of TPC on $\mathbf{c}_{a1}$ and $\mathbf{c}_{b1}$ can be expressed as $\rho_{ef_{a1}} = \sum_{i=1}^{N-1} \rho_{ef_{a1i}}$ and $\rho_{ef_{b1}} = \sum_{i=1}^{N-1} \rho_{ef_{b1i}}$ respectively. Without loss of generality, if the concept of $\rho_{ef}$ is imposed on $\mathbf{C}_{1L}^{6}$, the existing equality based on the concept of peak cross-correlation i.e.;

$$\left(\sum_{j=2}^{k}\sum_{v=1}^{\frac{N_k}{2^{j-1}}} \rho_{1j}(u,v)\right)_{ternary} = \left(\sum_{j=2}^{k}\sum_{v=1}^{\frac{N_k}{2^{j-1}}} \rho_{1j}(u,v)\right)_{2k-ary} \tag{7}$$

gets modified to the following inequality.

$$\left(\sum_{j=2}^{k}\sum_{v=1}^{\frac{N_k}{2^{j-1}}} \frac{\rho_{1j}(u,v)}{N_{ef_i}}\right)_{ternary} > \left(\sum_{j=2}^{k}\sum_{v=1}^{\frac{N_k}{2^{j-1}}} \frac{\rho_{1j}(u,v)}{N_{ef_i}}\right)_{2k-ary}. \tag{8}$$

Surprisingly, now the expression in (8) finally breaks its ambiguity (7) related to superiority between the ternary and $2k$-ary for $\mathbf{C}_{1L}^{6}$, as the above inequality favors for the better error performance of $2k$-ary. It can also be clarified from the simulation results (in Fig. 8 (a)). Please note that, for the left child in all other subsets (say $\mathbf{C}_i^k$ for $2 \leq i \leq k$), the relationship

$$\left(\sum_{j=i}^{k}\sum_{v=1}^{\frac{N_k}{2^{j-1}}} \rho_{ij}(u,v)\right)_{ternary} > \left(\sum_{j=2}^{k}\sum_{v=1}^{\frac{N_k}{2^{j-1}}} \rho_{1j}(u,v)\right)_{2k-ary} \tag{9}$$

always holds true, thus always satisfying the equivalent relation of (8), in default.

### *E. Anomaly 2: Impact of Free Diversity ($N_{fd_i} = \frac{N_{ef_i}}{2}$)*

First, let us revisit the structure of the ternary SMOS in Table I. With reference to the expression of $\mathbf{y}$ for $\mathbf{C}^k = \left[\mathbf{C}_1^k|\mathbf{C}_2^k|\ldots|\mathbf{C}_k^k\right]$ in (1), for subset-$\mathbf{C}_i^k$ with effective spreading gain $N_{ef_i}$, there exists $\frac{N_{ef_i}}{2}$ number of elements (chips) in $\mathbf{y}_i$ (see Fig. 5) that carries the spread data of $\mathbf{C}_i^k$ with no MAI, *provided* the previous $(i-1)$ subsets are decoded with no error. A close retrospection interprets the existence of such chips as the outcome of the tactical presence of the zero elements in the whole ternary structure. We call it *MAI free diversity* or simply *free diversity* for $\mathbf{C}_i^k$. Even if, the detection of the previous subsets go erroneous (for transmission to be noisy) leading to the inaccuracy in interference cancellation, the level of MAI in these chips still continues to be significantly lower than that of its remaining counterparts. Ultimately, having 50% of the transmitted diversities (chips) with zero or reduced MAI level, in deed, becomes an open scope for improving the BER of ternary SMOS. On a note, no such advantage is spotted for the case of $2k$-ary due to the absence of the zero element in the available chips.

With the alleviation of loading condition by subsequent removal of the smaller subsets (lower level of tree hierarchy) in $\mathbf{C}^k$ (ternary or $2k$-ary), the level of MAI on an arbitrary signature of each active subset (based on the concept of $\rho_{ef}$ in Anomaly-1) also decreases. While following *Anomaly-1*, for the left child, the superiority of $2k$-ary over ternary should prevail for noisy transmission, deflection in this regard may be perceived for relatively low loading condition. This translation is evidently due to the impact of free diversity that empowers the ternary to dominate. We call this transition of superiority (from $2k$-ary to ternary) as the *superiority crossover*. An elaborate picture of this transformation can be tracked in Table V presenting a comparative overview of the outcomes of Fig. 11 and 12.

## V. Simulation Results

In this section, we concentrate on the BER versus $(\mathrm{E}_b/N_0)$ performance of the system, assuming the channel to be AWGN. The system is supposed to be BPSK modulated, perfectly power controlled, and operated for synchronous transmission. Each figure is meant to illustrate the comparison between ternary and $2k$-ary with a specific purpose. Finally, we assimilate their outcomes in Table V so as to reach a conclusion about the superiority.

Fig. 6 (a) and (b) presents the comparison of the average BER of the left child. For simulation, we choose $\mathbf{C}^6 = \mathbf{C}_{64\times 126} = \left[\mathbf{C}_1^6|\mathbf{C}_2^6|\mathbf{C}_3^6|\mathbf{C}_4^6|\mathbf{C}_5^6|\mathbf{C}_6^6\right]$ as the encoding matrix. The decoder shown in Fig. 5 is selected for detection. In section V (b) in [4], the role of $N_{ef}$ and net MAI (or TPC) being responsible for the non-uniformity in response of the error performance of the individual subsets has already been explained, following which a dramatic lowering in the level of BER is observed for the smaller subsets (with lower $N_{ef}$ or detected in the later stages of decoder) over the larger ones (higher $N_{ef}$ or detected in the earlier stages), at higher values of $E_b/N_o$ (see Fig 6 (a)). As the reason of this irregularity, the gradual reduction in the level of TPC with the progress of the decoding stages is reported to be accountable.

On the other hand, similar behavior is not observed for the $2k$-ary (Fig 6 (b)), even if the effect of reduced level of TPC on the subsequent smaller subsets (similar to ternary) can still be found to exist. For this contrast, the effect of significant reduction in the level of PA for each subset (except $\mathbf{C}_1^6$) (e.g.; $(\rho_{ii})_{2k-ary} = \left(\sqrt{\rho_{ii}}\right)_{ternary}$, see Section IV (Bb)) is to be held responsible. Furthermore, for higher values of $E_b/N_o$, the convergence of the curves of different subsets is also captured, which is intrinsically due to the UD nature of construction that allows them to approach to the errorfree performance.

Fig. 7 (a) and (b), being the right counterpart of Fig. 6 (a) and (b), present the BER performance of the right child for ternary and $2k$-ary SMOS respectively. Where for ternary, the behavior of the right child is fully identical to that of the left, observing a dramatic variation for $2k$-ary is well precedented, and the explanation of this difference is straight i.e.; before MF detection, the level of TPC on the right child is notably lesser than that of the left (Section IV (Ba)). The explanation to the noticeable degradation in BER performance of the smaller subsets is same as that of the left child in Fig. 6 (a) i.e.; significant lowering of the level of PA.

Fig. 8 being the distributed version of Fig 6, projects a one-to-one comparison of the performance of the left child. Our intention is to study the variation in their (ternary versus $2k$-ary) performance more vividly, followed by appropriate rationalization. Fig 8 (a) to (f) chronologically correspond to $\mathbf{C}_{1L}^6$,$\mathbf{C}_{2L}^6$, $\mathbf{C}_{3L}^6$,$\mathbf{C}_{4L}^6$, $\mathbf{C}_{5L}^6$, and $\mathbf{C}_{6L}^6$ respectively. Where for $\mathbf{C}_{1L}^6$, the superiority of the $2k$-ary is due to the impact of low level of $\rho_{ef}$ (8) (*Anomaly-1*, Section IV (C)), its influence on the subsequent subsets gradually gets suppressed due to the substantial fall in the level of PA (*Trade-off*, Section IV (A)). Consequently, the edge in the level of BER over the moderate range of $E_b/N_o$ starts to fall and finally attends the equality with ternary for $\mathbf{C}_{3L}^6$. Eventually, from $\mathbf{C}_{4L}^6$ onwards, the level of BER of the ternary dominates and the difference continuously grows, from $\mathbf{C}_{4L}^6$ to $\mathbf{C}_{6L}^6$.

Fig. 9 being the distributed version of Fig 7 illustrates the one-to-one comparison of the BER for the right child. Fig 9 (a) to (f), in sequence, correspond to $\mathbf{C}_{1R}^{6}$,$\mathbf{C}_{2R}^{6}$, $\mathbf{C}_{3R}^{6}$,$\mathbf{C}_{4R}^{6}$, $\mathbf{C}_{5R}^{6}$, and $\mathbf{C}_{6R}^{6}$ respectively. The explanation to the non-uniformity in their superiority is same as that of the left and hence, needs no further detailing. However, as a crucial observation, the optimality of $\mathbf{C}_{1R}^{6}$ must be highlighted.

In Fig. 11, we extend the comparison of the left child (in Fig. 6 and 8), a step further. While Fig. 6 and 8 illustrates the comparison for the left child at maximum loading condition (i.e.; $\mathbf{C}_{64\times126}^{6}$ or $\beta = 1.97$), the study of their performance for different (reduced) loading condition will further supplement our analysis. So, we start with $\mathbf{C}_{64\times126}^{6}$ (see Fig. 8) and subsequently, smaller subsets ($\mathbf{C}_{6}^{6}$, $\mathbf{C}_{5}^{6}$, $\mathbf{C}_{4}^{6}$, $\mathbf{C}_{3}^{6}$, and $\mathbf{C}_{2}^{6}$) are removed one-by-one to realize five different loading conditions: $\mathbf{C}_{64\times124}^{6}$, $\mathbf{C}_{64\times120}^{6}$, $\mathbf{C}_{64\times112}^{6}$, $\mathbf{C}_{64\times96}^{6}$. While *each column* of the grid (Fig. 11 and 12) corresponds to the available subsets for a particular loading condition, *each row* corresponds to a specific subset under different loading. The size of the subset decreases from top to bottom e.g.; Fig. 11 (a1) and (a6) represents the BER of $\mathbf{C}_{1L}^{6}$ and $\mathbf{C}_{6L}^{6}$ respectively. Likewise, (b1) to (b5), (c1) to (c4), (d1) to (d3), and (e1) to (e2) illustrate for $\mathbf{C}_{1L}^{6}$ to $\mathbf{C}_{5L}^{6}$, $\mathbf{C}_{1L}^{6}$ to $\mathbf{C}_{4L}^{6}$, $\mathbf{C}_{1L}^{6}$ to $\mathbf{C}_{3L}^{6}$, $\mathbf{C}_{1L}^{6}$ to $\mathbf{C}_{2L}^{6}$ respectively. For better perception of this grid structure, its replication in Fig. 10 briefing the description about the change in the behavior of different operational metrics can be referred. From the observation, for a fixed loading condition (column wise), for the larger subsets the performance of $2k$-ary outsmarts that of ternary and it has already been explained in the context of Fig. 8, for $\mathbf{C}_{64\times126}^{6}$ (the first column in Fig. 11). Therefore, identical reasoning also suffices to understand the behavior of the other columns. However, with the reduction in loading (or level of TPC), the deviation is noticed in the order of their superiority. To explain this, let us shift our perspective of analysis to row-wise, since the level of MAI shows a continuous fall along a row (see Fig. 10). For example, take the case of the first row (Fig. 11 (a1), (b1), (c1), (d1), (e1)), where the superiority of the $2k$-ary can be found to have gradual degradation (from left to right). In particular, while for Fig. 11 (a1) and (b1) the superiority is traced for the $2k$-ary SMOS, a transition to the ternary in terms of the superiority cross-over is captured in Fig. 11 (c1), and for Fig. 11 (d1) onwards, the ternary finally dominates. Without any variation, similar tendency can be recorded for other rows too.

Fig. 12 is the extrapolation to Fig. 7, in the same way, Fig. 11 is meant to Fig. 6. Therefore, it is logical to expect a similar apprehension towards the demonstration of their behavior across the grid structure. As a noteworthy point, the description of various metrics in Fig. 10 (controlling the BER performance of the left child in Fig. 11) is also applicable to this case.

After the relative analysis of the BER for the left and right child under different loading conditions in Fig 11 and 12, now, it is essential to collect their outcome in a single frame so as to attain a firm conclusion of the superiority. So, we introduce Table V. Following its observation, at a particular loading condition, concluding any of them as superior, in overall, is quite contradictory due to mixed nature of their outcomes. Nonetheless, as the only difference, for $\mathbf{C}^2_{64\times 96}$ (last column in Table V) (i.e.; $\beta = 1.5$) it is no more illogical to consider $2k$-ary as superior.

In Fig. 13 (a) to (f), the comparison lies in between the left and right child for the individual subsets of $2k$-ary. For an arbitrary signature in CDMA, to gain better (optimal) error performance, the most coveted criterion to be satisfied is the higher value of PA ($\approx$ 1) in conjunction with the lower value of TPC ($\approx$ 0). Since the importance of PA serves exactly opposite to that of the TPC, it is clearly acceptable to treat their difference as a suitable metric for the analysis of error performance. For this difference to be higher for a signature, the better performance is expected and vice versa. For a system using $2k$-ary SMOS, it can be defined as $d_i = \frac{\rho_{ii}(u,v)}{N_{ef_i}} - \sum\limits_{j=i+1}^{k} \sum\limits_{v=1}^{\frac{N_k}{2^{j-1}}} \frac{\rho_{ij}(u,v)}{N_{ef_i}}$, where $d_i$ indicates the difference corresponding to an arbitrary signature in subset $\mathbf{C}^k_i$. Now, based on the change in the magnitude of $d_i$ (see Table IV), two crucial observations are spotted. First, for each subset, the value of $d_i$ for the right child ($d_{iR}$) is always higher than that of the left ($d_{iL}$) and its value gradually decreases as we move from top ($\mathbf{C}_1$) to bottom ($\mathbf{C}_6$) of the tree hierarchy. As a result, the level of BER for the right child always remains lower than that of the left. Second, the difference in its magnitude ( i.e.; $(d_i)_L - (d_i)_R$) also manifests a continuous reduction, as we move from $\mathbf{C}_1$ to $\mathbf{C}_6$ (i.e.; from Fig. 13 (a) to (f)). Finally, for $\mathbf{C}_6$, the value of $((d_i)_L - (d_i)_R)$ almost gets to zero that results in the overlapping of the curves in Fig. 13 (f).

## VI. Conclusion

In this paper, our attempt to extend the domain of SMOS beyond ternary culminated in a new set of SMOS i.e.; $2k$-ary SMOS. Along with the similarity regarding recursive construction and hierarchy of correlation with ternary, also, the difference is observed in the form of non-uniformity in the twin tree hierarchy and the existence of optimal users for the largest subset.

In contrast to ternary SMOS, where the level of MAI within a subset was found to be uniform, the deviation in this regard is was discovered for $2k$-ary in terms of two separate levels of MAI (based on the left or right child generated). Subsequently, the error performance analysis of its individual subsets was split into two sections, and the simulation results also confirmed this difference. Using the concept of suitable anomalies and trade-off, we reasonably validated the difference between the performance of the ternary and $2k$-ary. Despite the prominent edge in terms of optimality of the specific users, proclaiming the $2k$-ary to be superior over ternary appeared to be illogical, since the outcome of our attempt for recognizing either of them as superior was later found not to be mutually exclusive. This inference became more vivid from our observations when for a fixed loading condition, the superiority of any of them hardly retained the uniformity for all the constituent subsets. Therefore, evidently enough, superiority in this context became a conditional entity. However, with a difference, for the maximization in loading capacity to be 50% ($\beta = 1.5$ i.e.; the first two subsets being active), the superiority was found to be exclusively possessed by the $2k$-ary. In overall, the demonstration, so far, concentrating on the multiple attributes and aspects related to each and every individual signature (of the smallest possible subset), rather than just considering their average statistics, was truly profound for the better understanding of the generalization of SMOS.

## REFERENCES


[1] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE, 2013, pp. 332–336.

[2] R. Hoshyar, R. Razavi, and M. Al-Imari, "Lds-ofdm an efficient multiple access technique," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*. IEEE, 2010, pp. 1–5.

[3] S. Zhang, X. Xu, L. Lu, Y. Wu, G. He, and Y. Chen, "Sparse code multiple access: An energy efficient uplink approach for 5g wireless systems," in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 4782–4787.

[4] A. Singh, A. Amini, P. Singh, and F. Marvasti, "A new set of uniquely decodable codes for overloaded synchronous cdma," *IET Communications (accepted)*, 2016.

[5] M. Li, "Fast code design for overloaded code-division multiplexing systems," 2015.

[6] S. Dashmiz, M. R. Takapoui, S. Moazeni, M. Moharrami, M. Abolhasani, and F. Marvasti, "Generalisation of code division multiple access systems and derivation of new bounds for the sum capacity," *IET Communications*, vol. 8, no. 2, pp. 153–162, 2014.

[7] M. H. Khoozani, F. Marvasti, M. Azghani, and M. Ghassemian, "Finding sub-optimum signature matrices for overloaded code division multiple access systems," *IET Communications*, vol. 7, no. 4, pp. 295–306, 2013.

[8] O. Mashayekhi and F. Marvasti, “Uniquely decodable codes with fast decoder for overloaded synchronous cdma systems,” *Communications, IEEE Transactions on*, vol. 60, no. 11, pp. 3145–3149, 2012.

[9] P. Pad, F. Marvasti, K. Alishahi, and S. Akbari, “A class of errorless codes for overloaded synchronous wireless and optical cdma systems,” *Information Theory, IEEE Transactions on*, vol. 55, no. 6, pp. 2705–2715, 2009.

[10] M. Kulhandjian and D. A. Pados, “Uniquely decodable code-division via augmented sylvester-hadamard matrices,” in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. IEEE, 2012, pp. 359–363.

[11] G. H. Khachatrian and S. S. Martirossian, “A new approach to the design of codes for synchronous-cdma systems,” *Information Theory, IEEE Transactions on*, vol. 41, no. 5, pp. 1503–1506, 1995.

[12] H. Sari, F. Vanhaverbeke, and M. Moeneclaey, “Extending the capacity of multiple access channels,” *Communications Magazine, IEEE*, vol. 38, no. 1, pp. 74–82, 2000.

[13] G. N. Karystinos and D. A. Pados, “The maximum squared correlation, sum capacity, and total asymptotic efficiency of minimum total-squared-correlation binary signature sets,” *Information Theory, IEEE Transactions on*, vol. 51, no. 1, pp. 348–355, 2005.

[14] S. P. Ponnaluri and T. Guess, “Signature sequence and training design for overloaded cdma systems,” *Wireless Communications, IEEE Transactions on*, vol. 6, no. 4, pp. 1337–1345, 2007.

[15] H. H. Nguyen and E. Shwedyk, “A new construction of signature waveforms for synchronous cdma systems,” *Broadcasting, IEEE Transactions on*, vol. 51, no. 4, pp. 520–529, 2005.

[16] Y. H. Tahir, C. K. Ng, N. K. Noordin, B. M. Ali, and S. Khatun, “Superposition coding with unequal error protection for the uplink of overloaded ds-cdma system,” *Wireless Personal Communications*, vol. 65, no. 3, pp. 567–582, 2012.

[17] F. H. Ali and I. Shakya, “Collaborative spreading for the downlink of overloaded cdma,” *Wireless Communications and Mobile Computing*, vol. 10, no. 3, pp. 383–393, 2010.

[18] R. E. Learned, *Low complexity optimal joint detection for oversaturated multiple access communications*. Mass. Inst Technol, Cambridge, 1997.

[19] F. Marvasti, P. Pad, and M. F. Naeiny, “Iterative synchronous and asynchronous multi-user detection with optimum soft limiter,” May 17 2008, uS Patent App. 12/122,668.

[20] P. Pad, A. Mousavi, A. Goli, and F. Marvasti, “Simplified map-mud for active user cdma,” *Communications Letters, IEEE*, vol. 15, no. 6, pp. 599–601, 2011.

[21] P. Azmi and T. S. Zand, “An iterative multiuser detector for overloaded ldpc coded cdma systems,” *Wireless Personal Communications*, vol. 66, no. 1, pp. 41–56, 2012.

[22] A. Kapur and M. K. Varanasi, “Multiuser detection for overloaded cdma systems,” *Information Theory, IEEE Transactions on*, vol. 49, no. 7, pp. 1728–1742, 2003.

[23] F. Vanhaverbeke and M. Moeneclaey, “An improved ocdma/ocdma scheme based on displaced orthogonal user sets,” *Communications Letters, IEEE*, vol. 8, no. 5, pp. 265–267, 2004.

[24] F. Vanhaverbeke and M. Moeneclaey, “Sequences for oversaturated cdma channels,” in *Spread Spectrum Techniques and Applications, 2008 IEEE 10th International Symposium on*. IEEE, 2008, pp. 735–739.

[25] P. Kabir, M. Shafinia, and F. Marvasti, “Capacity bounds of finite dimensional cdma systems with power allocation, fading, and near-far effects,” *Communications Letters, IEEE*, vol. 17, no. 1, pp. 15–18, 2013.

[26] K. Alishahi, S. Dashmiz, P. Pad, and F. Marvasti, “Design of signature sequences for overloaded cdma and bounds on the sum capacity with arbitrary symbol alphabets,” *Information Theory, IEEE Transactions on*, vol. 58, no. 3, pp. 1441–1469, 2012.

[27] T. Guess, “Cdma with power control and sequence design: the capacity region with and without multidimensional signaling,” *Information Theory, IEEE Transactions on*, vol. 50, no. 11, pp. 2604–2619, 2004.

[28] D. Chakraborty, M. K. Tarafder, and A. Chandra, “A new walsh-like near orthogonal (wno) sequence for asynchronous cdma system,” *Wireless Personal Communications*, pp. 1–19.

[29] Y.-T. Li, Q.-Y. Yu, K. He, and W. Xiang, “Apply uniquely-decodable codes to multiuser physical-layer network coding based on amplify-and-forward criterion,” in *2015 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2015, pp. 1–6.

[30] S.-C. Chang and E. Weldon, “Coding for t-user multiple-access channels,” *Information Theory, IEEE Transactions on*, vol. 25, no. 6, pp. 684–691, 1979.

[31] T. Ferguson, “Generalized t-user codes for multiple-access channels (corresp.),” *Information Theory, IEEE Transactions on*, vol. 28, no. 5, pp. 775–778, 1982.

[32] S.-C. Chang, “Further results on coding for t-user multiple-access channels,” *IEEE transactions on information theory*, vol. 30, no. 2, pp. 411–415, 1984.

[33] W. H. Mow, “Recursive constructions of detecting matrices for multiuser coding: A unifying approach,” *Information Theory, IEEE Transactions on*, vol. 55, no. 1, pp. 93–98, 2009.

[34] Y.-W. Wu and S.-C. Chang, “Coding scheme for synchronous-cdma systems,” *Information Theory, IEEE Transactions on*, vol. 43, no. 3, pp. 1064–1067, 1997.

[35] S. Verdu, *Multiuser detection*. Cambridge university press, 1998.

[36] G. Khachatrian and S. Martirossian, “Codes for t-user noiseless adder channel.” *PROB. CONTROL INF. THEORY.*, vol. 16, no. 3, pp. 187–192, 1987.

[37] P. Kumar and S. Chakrabarti, “An analytical model of iterative interference cancellation receiver for orthogonal/orthogonal overloaded ds-cdma system,” *International Journal of Wireless Information Networks*, vol. 17, no. 1-2, pp. 64–72, 2010.

[38] G. Khachatrian and S. Martirossian, “Code construction for the t-user noiseless adder channel,” *Information Theory, IEEE Transactions on*, vol. 44, no. 5, pp. 1953–1957, 1998.

[39] S. Sasipriya and C. Ravichandran, “Performance analysis of overloaded cdma system under imperfect synchronization using parallel/successive interference cancellation,” *Telecommunication Systems*, vol. 56, no. 4, pp. 509–518, 2014.

TABLE I: Construction of SMOS (Type I) Matrices for $k = 1, 2, 3$ e.g., $\mathbf{C}^1 = [\mathbf{C}_1^1]$, $\mathbf{C}^2 = [\mathbf{C}_1^2 | \mathbf{C}_2^2]$, $\mathbf{C}^3 = [\mathbf{C}_1^3 | \mathbf{C}_2^3 | \mathbf{C}_3^3]$

$$\mathbf{C}^1 = \frac{1}{\sqrt{2}} \begin{bmatrix} + & + \\ + & - \end{bmatrix}$$

$$\mathbf{C}_1^2 = \frac{1}{\sqrt{4}} \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix}, \mathbf{C}_2^2 = \frac{1}{\sqrt{4}} \begin{bmatrix} + & + \\ 0 & 0 \\ + & - \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{C}_1^3 = \frac{1}{\sqrt{8}} \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & - & + & - & + & - & + & - \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ + & + & + & + & - & - & - & - \\ + & - & + & - & - & + & - & + \\ + & + & - & - & - & - & + & + \\ + & - & - & + & - & + & + & - \end{bmatrix}, \mathbf{C}_2^3 = \frac{1}{\sqrt{8}} \begin{bmatrix} + & + & + & + \\ 0 & 0 & 0 & 0 \\ + & - & + & - \\ 0 & 0 & 0 & 0 \\ + & + & - & - \\ 0 & 0 & 0 & 0 \\ + & - & - & + \\ 0 & 0 & 0 & 0 \end{bmatrix}, \mathbf{C}_3^3 = \frac{1}{\sqrt{8}} \begin{bmatrix} + & + \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ + & - \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

TABLE II: Constuction of $2k$-ary SMOS for $k = 1, 2, 3$ e.g., $\mathbf{C}^1 = [\mathbf{C}_1^1]$, $\mathbf{C}^2 = [\mathbf{C}_1^2 | \mathbf{C}_2^2]$, $\mathbf{C}^3 = [\mathbf{C}_1^3 | \mathbf{C}_2^3 | \mathbf{C}_3^3]$, where '+' and '-' actually denotes the '+1' and '-1' respectively.

$$\mathbf{C}^1 = \frac{1}{\sqrt{2}} \begin{bmatrix} + & + \\ + & - \end{bmatrix}$$

$$\mathbf{C}_1^2 = \frac{1}{\sqrt{4}} \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix}, \mathbf{C}_2^2 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

$$\mathbf{C}_1^3 = \frac{1}{\sqrt{8}} \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & - & + & - & + & - & + & - \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ + & + & + & + & - & - & - & - \\ + & - & + & - & - & + & - & + \\ + & + & - & - & - & - & + & + \\ + & - & - & + & - & + & + & - \end{bmatrix}, \mathbf{C}_3^2 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}, \mathbf{C}_3^2 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1/2^2 & 1/2^2 \\ 1/2^2 & 1/2^2 \\ 1/2^2 & 1/2^2 \\ 1/2^2 & 1/2^2 \\ 1/2^2 & -1/2^2 \\ 1/2^2 & -1/2^2 \\ 1/2^2 & -1/2^2 \\ 1/2^2 & -1/2^2 \end{bmatrix}$$

TABLE III: Expression for the Average BER of different subsets in SMOS ($2k$-ary)

| Subsets | $\mathbf{C}_{1R}^k$ | $\mathbf{C}_{iR}^k$ $(2 \leq i \leq k)$ | $\mathbf{C}_{iL}^k$ $(1 \leq i \leq k)$ |
|---|---|---|---|
| Average BER | $Q\left(\sqrt{\frac{\mathrm{E}\left(x_{1\mathrm{R_j}}^2\right)}{\mathrm{E}\left(x_{\mathrm{i(j)}}^2\right)}}\right)$ | $Q\left(\sqrt{\frac{\left(\frac{1}{4^{i-1}}\right)^2 \mathrm{E}\left(x_{1\mathrm{R_j}}^2\right)}{4\sum_{v=1}^{i-1}\left(\rho'_{iRv(j)}\right)^2 + \mathrm{E}\left(n_{\mathrm{i(j)}}^2\right)}}\right)$ | $Q\left(\sqrt{\frac{\left(\frac{1}{4^{i-1}}\right)^2 \mathrm{E}\left(x_{\mathrm{iL_j}}^2\right)}{\sum_{u=i+1}^{k}\left(\rho_{iLu(j)}^2\right)^2 + 4\sum_{v=1}^{i-1}\rho_{iLv(j)}^2 P_e^v + \mathrm{E}\left(n_{\mathrm{i(j)}}^2\right)}}\right)$ |

TABLE IV: Comparison of the inequality $\rho_{ii}(u,u) > \sum_{j=i+1}^{k} \sum_{v=1}^{\frac{N_k}{2^{j-1}}} \rho_{ij}(u,v)$ for each subset in $\mathbf{C}^6_{64\times126} = \left[\mathbf{C}^6_{1L}|\mathbf{C}^6_{1R}|\mathbf{C}^6_{2L}|\mathbf{C}^6_{2R}|\mathbf{C}^6_{3L}|\mathbf{C}^6_{3R}|\mathbf{C}^6_{4L}|\mathbf{C}^6_{4R}|\mathbf{C}^6_{5L}|\mathbf{C}^6_{5R}|\mathbf{C}^6_{6L}|\mathbf{C}^6_{6R}\right]$ (ternary and $2k$-ary), prior to applying the MF decoding in the respective stages (see Fig. 5)

| Subset (↓) | Ternary SMOS | $2k$-ary SMOS |
|---|---|---|
| $\mathbf{C}^6_{1L}$ | $1 > \left(\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $1 > \left(\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}\right)$ |
| $\mathbf{C}^6_{1R}$ | $1 > \left(\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $1 > 0$ |
| $\mathbf{C}^6_{2L}$ | $\frac{1}{2} > \left(\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $\frac{1}{2^2} > \left(\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}+\frac{1}{2^6}\right)$ |
| $\mathbf{C}^6_{2R}$ | $\frac{1}{2} > \left(\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $\frac{1}{2^2} > 0$ |
| $\mathbf{C}^6_{3L}$ | $\frac{1}{2^2} > \left(\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $\frac{1}{2^4} > \left(\frac{1}{2^5}+\frac{1}{2^6}+\frac{1}{2^7}\right)$ |
| $\mathbf{C}^6_{3R}$ | $\frac{1}{2^2} > \left(\frac{1}{2^3}+\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $\frac{1}{2^4} > 0$ |
| $\mathbf{C}^6_{4L}$ | $\frac{1}{2^3} > \left(\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $\frac{1}{2^6} > \left(\frac{1}{2^7}+\frac{1}{2^8}\right)$ |
| $\mathbf{C}^6_{4R}$ | $\frac{1}{2^3} > \left(\frac{1}{2^4}+\frac{1}{2^5}\right)$ | $\frac{1}{2^6} > 0$ |
| $\mathbf{C}^6_{5L}$ | $\frac{1}{2^4} > \left(\frac{1}{2^5}\right)$ | $\frac{1}{2^8} > \left(\frac{1}{2^9}\right)$ |
| $\mathbf{C}^6_{5R}$ | $\frac{1}{2^4} > \left(\frac{1}{2^5}\right)$ | $\frac{1}{2^8} > 0$ |
| $\mathbf{C}^6_{6L}$ | $\frac{1}{2^5} > 0$ | $\frac{1}{2^{10}} > 0$ |
| $\mathbf{C}^6_{6R}$ | $\frac{1}{2^5} > 0$ | $\frac{1}{2^{10}} > 0$ |

TABLE V: Comparison of Superioty in average BER: $2k$-ary versus Ternary, where '$\times$','$\approx$', and '$-$' denote the "superiority crossover", "approximate equality", and "Not Applicable" respectively.

| Loading (→) | $\mathbf{C}^6_{64\times126}$ | | $\mathbf{C}^6_{64\times124}$ | | $\mathbf{C}^6_{64\times120}$ | | $\mathbf{C}^6_{64\times112}$ | | $\mathbf{C}^6_{64\times96}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| i (↓) | $\mathbf{C}^6_{iL}$ | $\mathbf{C}^6_{iR}$ | $\mathbf{C}^6_{iL}$ | $\mathbf{C}^6_{iR}$ | $\mathbf{C}^6_{iL}$ | $\mathbf{C}^6_{iR}$ | $\mathbf{C}^6_{iL}$ | $\mathbf{C}^6_{iR}$ | $\mathbf{C}^6_{iL}$ | $\mathbf{C}^6_{iL}$ |
| 1 | $2k$ | $2k$ | $2k$ | $2k$ | $\times$ | $2k$ | $\times$ | $2k$ | $\approx$ | $2k$ |
| 2 | $2k$ | $2k$ | $2k$ | $2k$ | Ter. | $2k$ | Ter. | $2k$ | $\approx$ | $\approx$ |
| 3 | $\times$ | $2k$ | Ter. | $2k$ | Ter. | Ter. | Ter. | Ter. | - | - |
| 4 | Ter. | $2k$ | Ter. | Ter. | Ter. | Ter. | - | - | - | - |
| 5 | Ter. | Ter. | Ter. | Ter. | - | - | - | - | - | - |
| 6 | Ter. | Ter. | - | - | - | - | - | - | - | - |