C4-Sequences: Rate Adaptive Coded Modulation for Few Bits Message

E. Boutillon

C4-Sequences: Rate Adaptive Coded Modulation for Few Bits Message

Emmanuel Boutillon,
Université Bretagne Sud and IMT-Atlantique, Lab-STICC UMR 6285, CNRS
Email: emmanuel.boutillon@univ-ubs.fr

***Abstract*—This paper introduces a novel type of sequences called C4-sequences. C4-sequences share similar optimal autocorrelation properties with Zadoff-Chu sequences. However, C4-sequences offer the additional advantage of having also optimal (in the sense of minimal Euclidean distance between sequences) for several truncation lengths, providing flexibility in adapting to different channel conditions without compromising performance. Moreover, unlike Zadoff-Chu sequences, the points of a constellation associated with a C4-sequence are not limited to the unit circle. This opens up possibilities for achieving shaping gain, leading to enhanced spectral efficiency. By combining a truncated C4-sequence modulation as an inner code with a fixed-rate non-binary outer code, flexible and performant rate-adaptive communication systems can be achieved.**

***Index Terms*—Low SNR, Rate-adaptive, CCSK, truncated, sequence, very short message.**

I. Introduction

The Cyclic Code Shift Keying (CCSK) modulation is a well-known spreading technique [2]. It is used to increase the spectral efficiency of a length-q spreading sequence by utilizing its q circularly rotated versions to encode $m = \log_2(q)$ bits per sequence transmission. Additionally, [2] suggested truncating the CCSK sequence to its first l elements to further enhance the spectral efficiency. Recently, Marchand et al. introduced a coded-modulation scheme that combines a fixed-rate non-binary outer code with a variable-rate inner code based on variable-length truncated CCSK modulation. They employ a binary CCSK sequence in [4] and a q-ary CCSK sequence in [5] to construct efficient rate-adaptive communication schemes.

This paper presents a systematic mathematical construction method for generating optimal q-ary CCSK sequences of length q, which are referred to as C4-sequences. The term "C4" stands for the first "C" letter of the four words: "Constellation", "Cross", "Circular", and "Correlation". It also refers to the four truncation lengths that yield an inner code with an optimal distance property, namely $q/4$, $q/2$, $3q/4$, and q. Additionally, the paper demonstrates that when the sequence is not truncated (i.e., a sequence of length q), C4-sequences exhibit the same autocorrelation property as the well-known Zadoff-Chu sequences [1], [3].

This work has been funded by the French ANR under grants ANR-19-CE25-0013-01 (https://qcsp.univ-ubs.fr/) and ANR-21-CE25-0006 (https://ai4code.projects.labsticc.fr/). The author would like to thank Cédric Marchand and Alexandru Olteanu for their indirect, but fundamental, contribution: their machine learning development to optimize T-CCSK sequences gives the mathematical clues to construct the C4-sequences.

Notations: The complex vector $\mathbf{x} = (x(0), x(1), \ldots, x(n), \ldots)$ is a q-periodic infinite vector of complex numbers. The vector $\mathbf{x}_a$ represents the vector $\mathbf{x}$ left-shifted by a positions, i.e., for all n, $x_a(n) = x(n+a)$. The vector $\mathbf{x}_a^{a+l-1}$ denotes the truncated vector obtained by taking the first l values of $\mathbf{x}_a$, i.e., $\mathbf{x}_a^{a+l-1} = (x(n+a))_{n=0,1,\ldots,l-1}$. The notation $\langle \mathbf{x}, \mathbf{y} \rangle$ represents the complex scalar product over a period between the vectors $\mathbf{x}$ and $\mathbf{y}$. It is defined as $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=0}^{q-1} x(n)y(n)'$, where $y(n)'$ denotes the complex conjugate of $y(n)$.

II. Definition and properties of the C4-sequences

This section provides the definition of the C4-sequences along with some properties. In this paper, C4-sequences of length $q = 2^m$ are considered, where m is an integer with typical values ranging from 3 to 10. From a given length-q C4-sequence, it is possible to generate q distinct sequences $\mathbf{x}_a$, $a = 0, 1, \ldots, q-1$, and thus, to encode $\log_2(q) = m$ bits of information. The variable p is defined as $p = q/4$, i.e., $p = 2^{m-2}$, as this value holds a particular significance. It is important to note that there are multiple equivalent ways to define a C4-sequence, such as through its auto-correlation function, its Discrete Fourier Transform (DFT) function, or directly in the time domain by its distance property. In this paper, a C4-sequence is defined based on its circular auto-correlation property.

Let $\mathbf{x}$ be a sequence of length q, its circular auto-correlation $\mathbf{R_x}$ is the length q vector defined as

$$R_{\mathbf{x}}(\tau) = \langle \mathbf{x}, \mathbf{x}_\tau \rangle, \tag{1}$$

for $\tau = 0, 1, \ldots, q-1$.

Definition 2.1: A complex sequence of constellation points $\mathbf{x} = (x(n))_{n=0,1,\ldots,q-1}$, $q \geqslant 4$, is said to be a C4-sequence if and only if its circular auto-correlation $\mathbf{R_x}$ sequence verifies

$$\begin{cases} R_{\mathbf{x}}(\tau = lq/4) = qj^{-cl},\ l = 0, 1, 2, 3 \\ R_{\mathbf{x}}(\tau) = 0 \text{ otherwise,} \end{cases} \tag{2}$$

with c a sign value, i.e., $c \in \{-1, 1\}$ and j the imaginary number verifying $j^2 = -1$. By convention, when $c = 1$, the C4-sequence is referred to as clockwise C4-sequence since the non-null values of $R_{\mathbf{x}}(\tau)$ take sequentially the values 1, $-j$, -1 and j (i.e., a clockwise rotation direction). Symmetrically, when $c = -1$, the non-null value of a C4-sequence will take

the value 1, j, -1, $-j$. This type of sequences is thus referred to as a counter-clockwise C4-sequences.

Theorem 2.2: Let $\mathbf{x}$ be a C4-sequence of length q, and let $\mathbf{X}$ be the Discrete Fourier Transform (DFT) of $\mathbf{x}$ denoted as $\mathbf{X} = \mathcal{F}(\mathbf{x})$. Then, the squared ℓ_2-norm $\|X(k)\|^2 = X(k)X(k)'$ of the k-th element of $\mathbf{X}$ satisfies the following property

$$\begin{cases} \|X(k)\|^2 = 4q \text{ if } k + c \mod 4 = 0, \\ \|X(k)\|^2 = 0 \text{ otherwise.} \end{cases} \tag{3}$$

Proof: Consider a C4-sequence $\mathbf{x}$ of length q. By computing the circular auto-correlation $\mathbf{R_x}$ defined in (2) in the frequency domain, the following expression is obtained

$$\mathbf{R_x} = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{x}) \odot \mathcal{F}(\mathbf{x})'), \tag{4}$$

where $\odot$ represents the term-by-term component multiplication of the two vectors. Considering the DFT of both terms in (4) yields $\mathcal{F}(\mathbf{R_x}) = (\mathcal{F} \circ \mathcal{F}^{-1})(\mathcal{F}(\mathbf{x}) \odot \mathcal{F}(\mathbf{x})')$, and thus

$$\mathcal{F}(\mathbf{R_x}) = \mathcal{F}(\mathbf{x}) \odot \mathcal{F}(\mathbf{x})' = \mathbf{X} \odot \mathbf{X}'. \tag{5}$$

Using the formal expression of the k^{th} terms $\mathcal{F}(R_\mathbf{x})(k)$ of $\mathcal{F}(\mathbf{R_x})$ and by permuting the left and the right terms, (5) gives

$$X(k)X(k)' = \|X(k)\|^2 = \sum_{\tau=0}^{q-1} R_\mathbf{x}(\tau) e^{\frac{-2j\pi\tau k}{q}}. \tag{6}$$

According to (2), $R_\mathbf{x}(\tau)$ contains only 4 non-null terms for $\tau = 0$, $q/4$, $q/2$ and $3q/4$, thus, (6) gives

$$\|X(k)\|^2 = \sum_{l=0}^{3} q j^{-cl} e^{\frac{-2j\pi l \frac{q}{4} k}{q}} \tag{7}$$

$$= q \sum_{l=0}^{3} j^{-(k+c)l} \tag{8}$$

According to (8), $\|X(k)\|^2$ is equal to the product of q with the sum of the first 4 terms of a geometric series with a common ratio $\rho = j^{-(k+c)}$. This sum equals 4 if the common ratio ρ is equal to 1, which occurs when $k + c \mod 4 = 0$, and the sum equals 0 otherwise □

Theorem 2.2 provides a simple constructive rule for generating a C4-sequence of length q, which is summarized in Algorithm 1.

Algorithm 1 Generation of a C4-sequence of length q by the function $\mathbf{x} = G(\mathbf{s})$

Input A seed vector $\mathbf{s}$ of size $p = q/4$ composed of $q/4$ reals on the interval $[0, q[$, a value of c in the set $\{-1, 1\}$.
Output A clockwise ($c = 1$) or anti-clokwise ($c = -1$) C4-sequence $\mathbf{x}$ of length q

for $k \leftarrow 0$ **to** $q/4 - 1$ **do**
  $E_s(k) \leftarrow 4\sqrt{p} \times \exp(2j\pi \frac{s(k)}{q})$
end for
$\mathbf{X} \leftarrow \text{kron}(\mathbf{E}_s, [0, \frac{1-c}{2}, 0, \frac{1+c}{2}])$
$\mathbf{x} \leftarrow \mathcal{F}^{-1}(\mathbf{X})$
Return $\mathbf{x}$

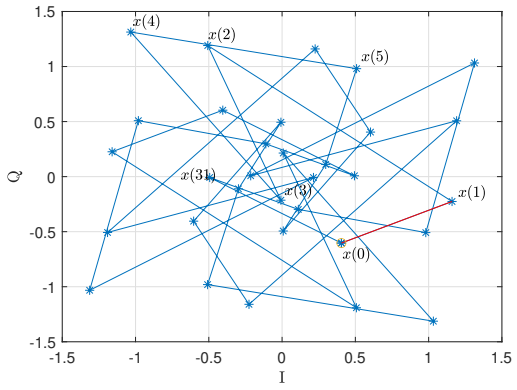

Fig. 1. Example of randomly generated C4-sequence

In Algorithm 1, the operator $\text{kron}(\mathbf{a}, \mathbf{b})$ represents the Kronecker product between vectors $\mathbf{a}$ and $\mathbf{b}$. The value of c determines whether the Kronecker product is performed with the vector $[0, 0, 0, 1]$ ($c = 1$, clockwise C4-sequence) or with $[0, 1, 0, 0]$ ($c = -1$, counter-clockwise C4-sequence). The scaling factor $4\sqrt{p}$ is applied to ensure that $\|X(k)\|^2 = 4q$ when $k + c \equiv 0 \mod 4$.

Figure 1 illustrates the construction of a C4-sequence of length $q = 32$ using $\mathbf{x} = G(\mathbf{s})$ with $\mathbf{s} = (7, 5, 3, 5, 6, 15, 31, 24)$ a length $p = 8$ vector taken randomly. The plot labels the first five elements $x(0), x(1), \ldots, x(5)$ and the last element $x(31)$ of the C4-sequence. Each consecutive pair of points in $\mathbf{x}$ is connected by a line, and the last point $x(31)$ is also connected to the first point $x(0)$.

Before providing a theorem that explicitly states the optimality of the C4-sequences, a few lemmas are established.

Lemma 2.3: The square of the ℓ_2-norm of a C4-sequence $\mathbf{x}$ of length q is given by $\|\mathbf{x}\|^2 = q$. This means that the average energy of the components of $\mathbf{x}$ is equal to 1.

Proof: This is a direct application of Parseval's theorem, which states that $\sum_{n=0}^{q-1} \|x(n)\|^2 = \frac{1}{q} \sum_{k=0}^{q-1} \|X(k)\|^2$. According to Theorem 2.2, $\|X\|^2$ contains exactly $\frac{q}{4}$ non-zero values, each equal to $4q$. Thus, $\|X\|^2 = \frac{q}{4} \cdot 4q = q^2$, and consequently, $\|x\|^2 = q$. Therefore, the square of the ℓ_2-norm of a C4-sequence $\mathbf{x}$ of length q is equal to q □

Lemma 2.4: Let $\mathbf{x}$ be a C4-sequence of length q, Then, for all n, $x(n + q/4) = j^{-c} x(n)$.

Note: The proof of this lemma is omitted here due to space limitations.

This lemma implies that $\mathbf{x}$ possesses a 4-fold rotational symmetry, i.e., $\mathbf{x}$ remains unchanged under a rotation of $\pi/2$, as can be observed in Fig. 1.

Definition 2.5: Let us consider a vector (or sequence) $\mathbf{x}$ of length q. The notation $\mathbf{x}_a^{a+l-1}$ indicates the length l subsequence of $\mathbf{x}$ that spans from index a to index $a + l - 1$.

From the definition of a length l truncated sequence, the following theorem is derived:

Theorem 2.6: Let $\mathbf{x}$ be a C4-sequence of length q and let $p = q/4$. For any $l \in \{p, 2p, 3p, 4p\}$ and any pair (a, b) of integers between 0 and $q - 1$, $a \neq b \Rightarrow \left\|\mathbf{x}_a^{a+l-1} - \mathbf{x}_b^{b+l-1}\right\|^2 \geqslant 2l$.

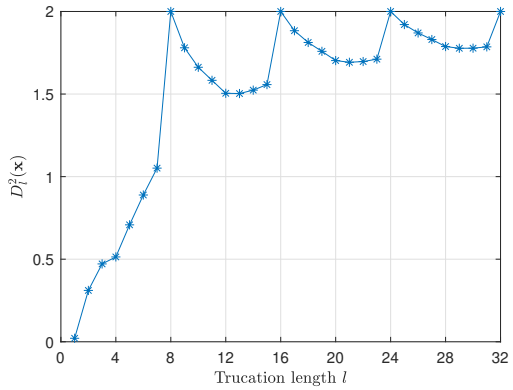

Fig. 2. Normalize square minimum distance $D_l^2(\mathbf{x})$ as a function of l for the C4-sequence given in Fig. 1.

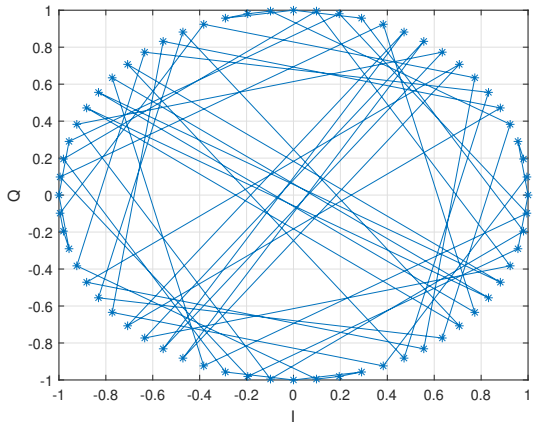

Fig. 3. C4-unitary sequence of size $q = 64$ generated with sequence $\mathbf{s}_u$ defined with $\mathbf{d}_u = (1, 11, 17, 27)$ and $\rho = (0, 1, 2, 3)$

Note: Due to space constraints, the proof of this theorem is not provided. Let the normalize square minimum distance $D_l^2(\mathbf{x})$ between two sequences of the set $\{\mathbf{x}_a^{a+l-1}\}_{a=0,1,\ldots,q-1}$ [5] be defined as

$$D_l^2(\mathbf{x}) = \frac{1}{l}\min_{a,b,a\neq b}\{\left\|\mathbf{x}_a^{a+l-1} - \mathbf{x}_b^{b+l-1}\right\|^2\}. \quad (9)$$

According to Theorem 2.6, $D_l^2(\mathbf{x}) = 2$ for $l \in \{p, 2p, 3p, 4p = q\}$. Fig. 2 shows the function $D_l^2(\mathbf{x})$ associated with the sequence shown in Fig. 1.

To summarize, C4-sequences are simple to build and are also optimal for the truncation lengths p, $2p$, $3p$ and $4p$. In the next sections, the construction of unitary C4-sequence is described, i.e., C4 sequences with points on the unit circle.

III. Unitary C4-sequences

In this section, an explicit construction method is provided for building C4-sequences on the unit circle. The principle is to impose constraints on the seed sequence s such that the resulting C4-sequence $\mathbf{x} = G(s)$ (refer to Algorithm 1) becomes a unitary C4-sequence. At first, the case of unitary C4-sequences of size $q = 2^{2t}$ is discussed before addressing the case of size $q = 2^{2t+1}$.

A. Unitary C4-sequence of size $q = 2^{2t}, t > 1$

Let a unitary seed vector $\mathbf{s}_u$ be defined as a real vector of size p constructed from a real vector $\mathbf{d}_u$ of size 2^{t-1} taking its values in the interval $[0, q[$ and a permutation γ over the set $(0, 1, \ldots, 2^{t-1} - 1)$ as

$$s_u(2^{t-1}a + r) = d_u(r) + a\gamma(r)2^{t+1}, \quad (10)$$

with $0 \leq r < 2^{t-1}$ and $0 \leq a < 2^{t-1}$ so that $k = 2^{t-1}a + r$ spans all the integer values between 0 and $2^{2t-2} - 1 = p - 1$.

Note that there is no constraint on the choice of the vector $\mathbf{d}_u$. For example, for $q = 64$ (i.e. $t = 3$), taking $\mathbf{d}_u = (1, 11, 17, 27)$ and $\gamma = (0, 1, 2, 3)$ gives the unitary seed vector $\mathbf{s}_u = (1, 7, 21, 39, 1, 23, 53, 23, 1, 39, 21, 7, 1, 55, 53, 55)$, with the components given modulo $q = 64$. The plot of $\mathbf{x}_u = G(\mathbf{s}_u)$ is given in Fig. 3. As it can be noticed, all the points of $\mathbf{x}_u$ are located on the unit circle. This result is predicted by the following theorem.

Theorem 3.1: If $\mathbf{s}_u$ is a unitary sequence, then $\mathbf{x}_u = G(\mathbf{s}_u)$ is a unitary C4-sequence. Moreover, the n^{th} term of $\mathbf{x}$ is given as

$$x(n) = \exp\left(\frac{2\pi j(n + d(\gamma^{-1}(-n)) + 4\gamma^{-1}(-n))}{q}\right), \quad (11)$$

with $\gamma^{-1}(-n)$ the unique solution of the equation $\gamma(\gamma^{-1}(-n)) = -n \mod 2^{t-1}$

The proof of this theorem is not provided here due to space limitations.

Finally, it is worth highlighting that (11) presents an alternative approach for constructing unitary C4 sequences directly in the time domain.

B. Link with Epicycloid sequence

In paper [5], the authors utilize specific sequences known as "Epicycloid" or "four-cups sequences", which are particular instances of unitary C4 sequences. For instance, the recursion defined as $\psi(1) = 0, x(1) = 1$, $\psi(i) = \psi(i-1)5 + 1$, $x(i) = \exp\left(\frac{2\pi j\psi(i)}{64}\right)$ for $i = 1, 2, \ldots, 64$ represents the length-64 Epicycloid four-cups sequence. This sequence corresponds to the unitary length-64 C4 sequence with parameters $\mathbf{d}_u = (62, 0, 62, 40)$ and $\rho = (3, 0, 1, 2)$.

C. Unitary C4-sequence of size $q = 32$

In this section, an explicit method is provided for constructing unitary C4 sequences of length $q = 32$. The construction method for length $q = 128$ is not given due to space limitation. However, the construction method is not yet generalized for size $q = 512$ and higher in the frequency domain. It is nevertheless possible to build them directly in the time domain using a similar approach than the one given in (11).

If $q = 32$, then $t = 2$ to satisfy $q = 2^{2t+1} = 32$. Let us define $\mathbf{d}_u = (d_u(0), d_u(1))$ a vector of size $2^{t-1} = 2$, ρ a permutation of the set of the first $2^{t-1} = 2$ naturals ($\rho = (0, 1)$

or $(1,0)$) and $(\epsilon_0,\epsilon_1) \in \{-1,1\}^2$ two extra parameters. Let us define the length-8 unitary seed vector $\mathbf{s}_u$ as

$$\begin{cases} s_u(i\&6+\rho(0)) = d_u(0) + 8i_1\epsilon_0 \\ s_u(i\&6+\rho(1)) = d_u(1) + 16i_2 + 8i_1 + 8i_1\epsilon_1 \end{cases}, \quad (12)$$

where $i = 0,1,\ldots,7$ and $(i\&6)$ represents the binary masking of $i = (i_2 i_1 i_0)_2$ with the binary vector $6 = (110)_2$ (i.e. $(i\&6) = i - (i \mod 2)$). The C4-sequence $\mathbf{x}_u = G(\mathbf{s}_u)$ is a unitary C4-sequence. For example, by considering $\mathbf{d}_u = (12,27)$, $\rho = (0,1)$, $\epsilon_0 = -1$, and $\epsilon_1 = 1$, the obtained unitary seed vector is $\mathbf{s}_u = (12,27,4,27,12,11,4,11)$. It can be verified that $\mathbf{x} = G(\mathbf{s}_u)$ is a unitary C4-sequence.

D. Comparison with Zadoff-Chu sequences

The interest of Unitary C4-sequences becomes apparent when compared with the well-known Zadoff-Chu sequences [1], [3], which are widely used in various communication standards. Consider a Zadoff-Chu sequence of length p, denoted as $\mathbf{z} = (z(l))_{l=0,1,\ldots,p-1}$. Similar to Definition 2.5, $\mathbf{z}_a$ represents the sequence $\mathbf{z}$ with a circular shift of a positions. The Zadoff-Chu sequence satisfies the property that for all $a \neq b$, $\langle \mathbf{z}_a, \mathbf{z}_b \rangle = 0$. Since the p elements of $\mathbf{z}_a$ have a magnitude of 1, $\|\mathbf{z}_a\|^2 = p$. Consequently, for $a \neq b$, $\|\mathbf{z}_a - \mathbf{z}_b\|^2 = \|\mathbf{z}_a\|^2 + \|\mathbf{z}_b\|^2 - 2\mathcal{R}(< \mathbf{z}_a, \mathbf{z}_b >) = 2p$. In summary, a Zadoff-Chu sequence of length $p = \frac{q}{4}$ generates a set of p vectors, each at a distance equal to $2p$ from one another.

Now, consider a C4-sequence $\mathbf{x}$ of length $q = 4p$. When this C4-sequence is truncated to a length of p, a set of $q = 4p$ vectors $\mathbf{x}_a^{a+p-1}$ is obtained, where $a = 0,1,\ldots,q-1$, each of length p. According to Theorem 2.6, the minimum distance between these vectors is greater than or equal to $2p$. Comparing it to the Zadoff-Chu sequences of length p, the minimum distance of the generated set is the same (i.e., $2p$), but the cardinality is four times higher. This demonstrates the potential advantage of C4-sequences in terms of cardinality when compared to Zadoff-Chu sequences.

IV. Geometrical shaping of C4-sequences

Using operational research techniques, it is possible to optimize a C4-sequence according to a given objective, such as maximizing the Mutual Information (MI) for a specific Signal-to-Noise Ratio (SNR). For example, Figure 4.a shows the optimized constellations for MI at an SNR of 5 dB using a greedy optimization algorithm. For this constellation, the achieved mutual information is 2.0536 bit/s/Hz, very close to the channel capacity of 2.057 bit/s/Hz. Note that 12 points are regularly spread over an external circle. Similarly, Fig. 4.b presents the optimized constellations for an SNR of 0 dB. For this SNR, the channel capacity is 1 bit/s/Hz, while the achieved mutual information is 0.9998 bit/s/Hz. The shape of the constellation is different compared to the 0 dB case, since the outer circle contains only 8 points.

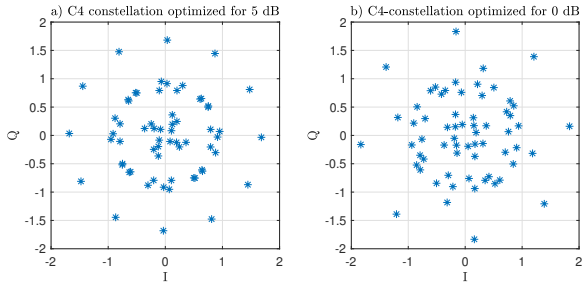

Fig. 4. Optimized C4-Constellations: a) Optimized for 5 dB of SNR, MI of 2.0536 bit/s/Hz (99.81 % of the channel capacity), b) Optimized for 0 dB, MI of 0.9998 bit/s/Hz (99.98 % of the channel capacity).

	$l=1$	$l=2$	$l=3$	$l=6$
$D_l^2(\mathbf{c})$	0.0952	0.2929	0.6667	0.6667
$D_l^2(\mathbf{x})$	0.0023	0.2577	0.4701	1.0420
MI($\mathbf{c}$) @ 0 dB	0.9918	1.9617	2.9158	4.3286
MI($\mathbf{x}$) @ 0 dB	0.9974	1.9753	2.9082	4.9290
MI($\mathbf{c}$) @ 5 dB	1.9926	3.7243	5.1551	5.8571
MI($\mathbf{x}$) @ 5 dB	2.0292	3.8548	5.1433	5.9843
MI($\mathbf{c}$) @ 10 dB	3.2686	5.3549	5.9805	≈ 6
MI($\mathbf{x}$) @ 10 dB	3.3108	5.5044	5.9745	≈ 6

TABLE I
COMPARISON OF CLASSICAL MODULATIONS WITH A TRUNCATED C4-SEQUENCE FOR TRANSMITTING A LENGTH-6 MESSAGE. THE MUTUAL INFORMATION (MI) IS EXPRESSED IN BIT/S/HZ.

.

V. C4-sequences for transmission of short packets

A C4-sequence can be used alone to encode a 3 to 10 bits message. However, it is also possible to concatenate a non-binary error correcting code over GF(q) as an outer code and a Truncated-C4 sequence as an inner code [4], [5]. This scheme allows constructing a flexible rate-adaptive coding scheme for short packets.

A. Performance of C4-sequence alone

A C4-sequence alone can be effectively used to send very short-length messages (between 3 and 10 bits, typically) in the Additive White Gaussian Noise (AWGN) channel. Table I characterizes a GF(64) C4-sequence $\mathbf{x}$ built to jointly optimize the normalize square minimum distance for lengths $l = 1,2,3$, and 6[1]. Table I presents the normalize square minimum distance $D_l^2(\mathbf{x})$ for those truncation lengths, as well as the resulting mutual information (MI, in bit/s/Hz) in the AWGN channel. These characteristics are compared with the use of classical modulations, denoted by the generic term $\mathbf{c}$, for transmitting 6 bits of information. For instance, $\mathbf{c}$ refers to 64-QAM for $l = 1$, a couple of 8-PSK symbols for $l = 2$, a triplet of QPSK symbols for $l = 3$, and 6 BPSK symbols for $l = 6$.

According to Table I, for truncation lengths $l = 1,2$, and 3, the minimum distances of classical constellations are better. However, the Mutual Information (MI) is better for the C4-sequence for $l = 1$ and $l = 2$. The advantage of

[1]achieved using the optimization process with the cost function $\Psi(\mathbf{x}) = 6D_1^2(\mathbf{x}) + 3D_2^2(\mathbf{x}) + 2D_2^2(\mathbf{x}) + D_6^2(\mathbf{x})$

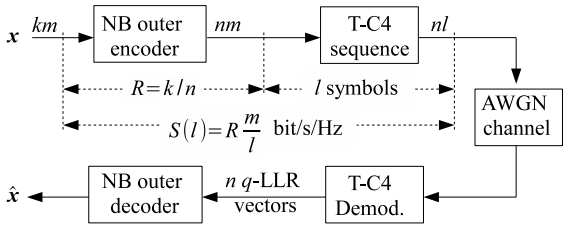

Fig. 5. Concatenation of an NB outer code and a Truncated-C4 sequence.

the C4-sequence becomes predominant for $l = 6$, as both the normalize square minimum distance and the MI become significantly higher compared to the classical constellation.

B. Principle of concatenated scheme

As proposed in [4], [5], it is possible to concatenate an outer Non-Binary Error Correcting Code (NB-ECC) over GF(q) with an inner code composed of a Truncated-C4 sequence, as shown in Fig. 5. In this scheme, the outer code takes k GF(q) symbols (i.e., m-tuple binary vector, with $m = \log_2(q)$) to generate n GF(q) symbols (coding rate $R = k/n$). Each of the GF(q) symbols is then used to modulate a Truncated-C4 sequence of length l, where l is a parameter that allows flexibility to precisely match the spectral efficiency to the channel condition. The total spectral efficiency of this coding scheme is thus

$$S(l) = \frac{Rm}{l} \text{ bit/s/Hz.} \tag{13}$$

The decoding of the T-CCSK sequence consists of computing the Log Likelihood Ratio (LLR) of all the q possible codewords of the Truncated-C4 sequence based on the length-l noisy received sequence. The LLRs are then used by the outer code to retrieve the transmitted message.

C. Simulation results

In this section, an outer code constituted by a single parity check of degree 4 over GF(64) is considered. It allows coding a 3 GF(64) symbol message (thus 18 bits of information) into a codeword of 4 GF(64) symbols. For a given truncation length l, the spectral efficiency $S(l)$ of this coding scheme is thus given by $S(l) = \frac{9}{2l}$, as mentioned in (13).

Fig. 6 shows the performance of the C4-sequence used in Table I for truncation lengths varying from $l = 1$ up to $l = 60$, resulting in spectral efficiencies ranging from 4.5 bit/s/Hz ($l = 1$) to 0.075 bit/s/Hz ($l = 60$) as indicated by (13). This concatenated coding scheme can be effectively used to finely adapt the spectral efficiency to the channel condition. Moreover, it is well-suited for use in a Hybrid Automatic Request communication scheme. In the event of decoding failure, the receiver can request the emitter to send the following symbols of the truncated sequences to increase the effective truncation length.

More complex outer error codes can be utilized. For example, [5] gives performance for rate 1/3 GF(64) Non-Binary Low Density Parity Check (NB-LDPC) outer code of size $(k, n) = (20, 60)$. It also uses a unitary Truncated C4-sequence of length 256 with a rate 2/3 GF(256) NB-LDPC outer code of size $(k, n) = (16, 24)$.

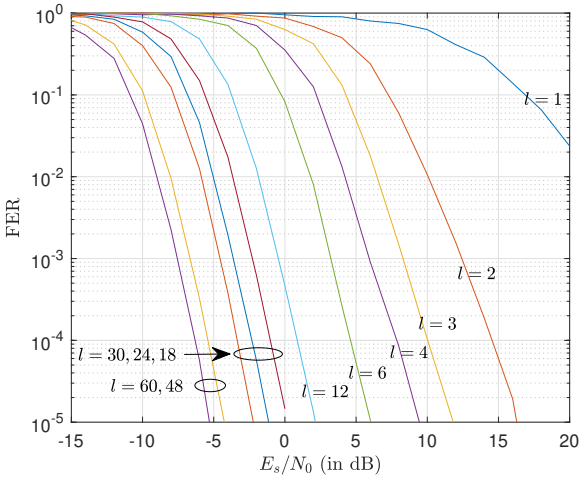

Fig. 6. Performance of an outer code composed of a degree 4 parity check over GF(64) and a Truncated-C4 sequence as the inner code.

VI. Conclusion

In this paper, the notion of C4-sequences has been defined. C4-sequences share similar optimal autocorrelation properties with Zadoff-Chu sequences. However, C4-sequences offer the additional advantage of having optimal properties (in the sense of minimal Euclidean distance between sequences) for several truncation lengths. Moreover, unlike Zadoff-Chu sequences, they are not limited to having their points on the unitary circle.

C4 sequences can have several applications in a communication system. Firstly, they can be used on their own as an alternative to Zadoff-Chu sequences. Secondly, the constellation associated with a C4-sequence can be shaped to maximize the mutual information through the AWGN channel, thereby providing geometric shaping gain. Finally, the concatenation of an outer non-binary code with a truncated C4-sequence as an inner code presents a very efficient and flexible communication scheme. While having a fixed outer code, the choice of truncation length provides a versatile tool for closely adapting the overall coding rate to the channel condition. Finally, it's worth mentioning that this flexibility can be effectively exploited in a hybrid-automatic request (H-ARQ) communication system.

References

[1] D. Chu. Polyphase codes with good periodic correlation properties (corresp.). *IEEE Transactions on Information Theory*, 18(4):531–532, 1972.

[2] G.M. Dillard, M. Reuter, J. Zeiddler, and B. Zeidler. Cyclic code shift keying: a low probability of intercept communication technique. *IEEE Transactions on Aerospace and Electronic Systems*, 39(3):786–798, 2003.

[3] R. Frank, S. Zadoff, and R. Heimiller. Phase shift pulse codes with good periodic correlation properties (corresp.). *IRE Transactions on Information Theory*, 8(6):381–382, 1962.

[4] Cédric Marchand and Emmanuel Boutillon. Rate-adaptive inner code for non-binary decoders. In *2021 11th International Symposium on Topics in Coding (ISTC)*, pages 1–5, 2021.

[5] Cédric Marchand and Emmanuel Boutillon. Rate-adaptive cyclic complex spreading sequence for non-binary decoders. In *International Symposium on Topics in Coding (ISTC'2023), Brest*, 2023.