Pseudo-Ternary Run-Length Limited Spectrum Shaped Codes

Oleg Kurmaev, Yanxing Zeng, Sergei Fedorenko, Vladimir Ivanov, and Jianqiang Shen
Huawei Russia Research Centre,
Moscow, Russia
Email: Oleg.Kurmaev,zengyanxing,Sergei.Fedorenko,vladimir.ivanov,shenjianqiang@huawei.com

***Abstract*—A method to convert binary to pseudo-ternary sequences is proposed. The method keeps run-length constraints imposed on finite-length binary sequence and also provides a unit value of a channel transition. The resulting finite-length pseudo-ternary sequence has exponential sums falling within specified regions. This allows spectrum shaping. The average power spectral density of infinite sequences obtained by concatenating spectrally-constrained sequences is examined. Thus we obtain the pseudo-ternary spectrum shaped code, which generally yields higher capacity than the binary one.**

I. Introduction

Binary run-length limited (RLL) sequences are widely used for resolving problems of synchronization recovery and suppressing dc in data transmission and storage systems [1]. Spectrum control is a related but more common and important problem. An enumerative method for computing the number of binary run-length limited spectrum shaped sequences was suggested in [2]. Using a 3-level signal system instead of the binary one provides many advantages beginning with lower bandwidth [3] and also allows increasing the recording density [4], [5]. The evolution of these codes seems to be useful in the real engineering world. To this end, we choose a pseudo-ternary [3] code in which each signal element only represents one bit of information.

In the binary case, the Non Return to Zero, Inverted (NRZI) is a method of mapping a binary sequence $\boldsymbol{x} \in \{0,1\}^n$ to a physical signal $\boldsymbol{z} \in \{-1,1\}^n$ for transmission over some transmission media

$$z_j = \prod_{s=1}^{j} (-1)^{x_s}. \tag{1}$$

Generally, the discrete Fourier transform (DFT)

$$z_m^* = \sum_{j=0}^{n-1} z_{j+1} e^{-2\pi i \frac{mj}{n}}, \quad m = 0, \ldots, n-1$$

is used for the spectral analysis of $\boldsymbol{z}$ (see, e.g., [6]). In this paper we expand the method described in [2] to computing the number of the pseudo-ternary run-length constrained sequences whose values of the DFT components fall within some given regions in the complex plane. Then encoding and decoding procedures become possible with known algorithms [7], [8]. Below, for convenience we use the term *ternary* instead of *pseudo-ternary*.

The rest of this paper is organized as follows. First, in Section II, we recall spectrum shaped binary sequences and compare the binary and ternary sequences in terms of spectral properties. Next, in Section III, we introduce a construction that allows converting from binary to ternary sequences while retaining the run-length constraints and providing the unit value of channel transitions. Further, in Section IV, we consider a charge distribution for these ternary sequences, next, in Section V, we turn to the more common case and propose an enumerative scheme for computing the number of spectrum shaped ternary sequences. Subsequently, in Section VI, we consider spectra of infinite-length sequences composed of an endless concatenation of the finite-length spectrum shaped sequences chosen uniformly at random. We also state a relationship between the complex-plane constraints and the PSD of the overall sequence of channel signals.

II. Binary Spectrum Shaped RLL Sequences

Usually, spectral properties of a sequence to be transmitted are described in terms of power spectral density (PSD). In [2] it was shown that a set of points in the complex plane for each of the DFT components is a more convenient form for use of the enumerative technique [9], [8]. This allows us to consider a purely combinatorial problem that also requires considering both real and imaginary parts of the DFT components. In any case, it is not so hard to obtain the PSD from the complex form of the DFT.

In binary dk-constrained sequences each pair of consecutive ones is separated by at least d and at most k zeros [1]. A dkr-sequence ia a dk-sequence, ending in a run of at most r trailing zeros. A $dklr$-sequence ia a dkr-sequence, beginning with a run of at most l leading zeros.

Let $\{0,1\}^n$ be the set of all binary sequences of length n and let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ denote a generic element of this set. Let $\mathscr{S}(n) = \{\boldsymbol{x} \in \{0,1\}^n \mid$ satisfy the d, k, r constraints and $x_1 = 1\}$ and let $\hat{\mathscr{S}}(n) = \{\boldsymbol{x} \in \{0,1\}^n \mid$ satisfy the d, k, l, r constraints$\}$.

A simplest example of spectrum shaping is so called dc-control. The 0th or dc component of the DFT is defined as $z_0^* = \sum_{j=1}^{n} z_j$. Binary sequences having predefined dc are called constant-charge sequences. A method for enumerating these sequences was suggested in [8].

Now consider a way that led us from enumeration of the constant-charge to spectrum shaped run-length constrained sequences[1]:

A. The number of constant-charge dkr-sequences

Let $C_n^{\sigma_0}$ be the number of sequences from $\mathscr{S}(n)$ with charge $\sigma_0 = z_0^*$. Using Cover's method [9], the number of these sequences can be computed [8] using recurrence relation in implicit mutual form

$$C_n^{\sigma_0} = \begin{cases} 0, & n < d+1, \\ \sum_{j=d+1}^{\min(n,k+1)} C_{n-j}^{-\sigma_0 - j}, & \text{otherwise} \end{cases}$$

with initial conditions

$$C_n^{-n} = \begin{cases} 1, & n \le r+1, \\ 0, & \text{otherwise.} \end{cases}$$

B. Regions instead of exact values of the DFT components

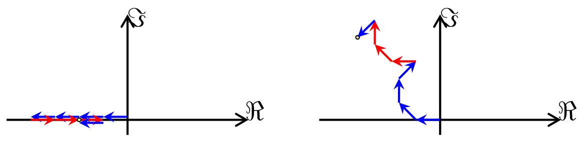
If we expand this method to $z_1^* \dots z_{n-1}^*$ components of the DFT, then we obtain a trivial case that is not of interest. Indeed, recall that the DFT is a one-to-one transformation. If we proceed from the single spectral component to the vector $\boldsymbol{z}^* = (z_0^*, z_1^*, \dots, z_{n-1}^*)$, then for an exact value of the corresponding vector $\boldsymbol{\sigma} = (\sigma_0, \sigma_1, \dots, \sigma_{n-1}) \in \mathbb{C}^n$ we see that the number $C_n^{\boldsymbol{\sigma}}$ of sequences from $\mathscr{S}(n)$ with the spectral components $\boldsymbol{\sigma}$ admits only two values:

$$C_n^{\boldsymbol{\sigma}} = \begin{cases} 1, & \text{if } \boldsymbol{\sigma} = \boldsymbol{z}^*, \\ 0, & \text{otherwise.} \end{cases}$$

As a solution we can consider those number of sequences $C_{n,m}^{\sigma_m}$ for which spectral components z_m^* fall within some region; let this region be a closed disk $\mathscr{D}_{\sigma_m,\rho_m}$ of radius $\rho_m \in \mathbb{R}$ centered at σ_m, i.e., the disk $\mathscr{D}_{\sigma_m,\rho_m}$ is a set $\mathscr{D}_{\sigma_m,\rho_m} = \{s \in \mathbb{C} : \ |z_m^* - s| \le \rho_m\}$.

C. A trigonometric recurrence relation and the number of the DFT component bounded dkr-sequences

In the case of 0th (dc) component of the DFT we only sum the charge with proper signs. For the other DFT components, we must account the angle of each vector (see Fig. 1). Now consider a method for this.



(a) A diagram of $z_0^* = \sigma$. (b) A diagram of z_1^*.

Fig. 1. Vector diagrams of 0th and 1st DFT components of (-1,-1,-1,-1, 1, 1, 1,-1).

Let $\omega_n = e^{\frac{2\pi i}{n}}$ denote a primitive nth root of unity.

Let $0 \le m \le n-1$, $\tilde{n} \le n$, $z \in \{-1, 1\}$, and let $\theta \in \mathbb{C}$ be some constant. Then an order j recurrence relation

$$\sigma_{\tilde{n},m} = z \sum_{t=0}^{j-1} \omega_n^{-mt} + \theta \sigma_{\tilde{n}-j,m} \tag{2}$$

is called a *trigonometric recurrence relation*.

The constant θ is said to be the linear phase if $\theta = z_m \omega_n^{-mj}$.

In [2] it was necessary to consider two kinds of length: The length of the entire sequence, for which the DFT is computed, is denoted by n. The lengths of the nested subsequences are denoted by $\tilde{n}$. In such a case we must add another subscript $\tilde{n}$ to σ_m and write $\sigma_{\tilde{n},m}$.

Consider bipolar run-length constrained $\tilde{n}$-length subsequences of $\boldsymbol{z}$ having an mth trigonometric recurrence relation disk $\mathscr{D}_{\sigma_{\tilde{n},m},\rho_m}$ of radius ρ_m centered at $\sigma_{\tilde{n},m}$.

Let $C_{n,\tilde{n},m}^{\sigma_{\tilde{n},m},\rho_m}(d,k,r)$ be the number of the $\tilde{n}$-length subsequences that begin with one. (Below for convenience, we use $C_{n,\tilde{n},m}^{\sigma_{\tilde{n},m},\rho_m}$ instead of $C_{n,\tilde{n},m}^{\sigma_{\tilde{n},m},\rho_m}(d,k,r)$.)

The number $C_{n,\tilde{n},m}^{\sigma_{\tilde{n},m},\rho_m}$ can be obtained as:

$$C_{n,\tilde{n},m}^{\sigma_{\tilde{n},m},\rho_m} = \begin{cases} a_{\tilde{n},m}, & \tilde{n} < d+1, \\ \sum\limits_{j=d+1}^{\min(\tilde{n},k+1)} C_{n,\tilde{n}-j,m}^{\sigma_{\tilde{n}-j,m},\rho_m} + b_{\tilde{n},m}, & \text{otherwise} \end{cases}$$

where

$$\sigma_{\tilde{n}-j,m} = -\omega_n^{mj} \left(\sigma_{\tilde{n},m} + \frac{\sin\left(\frac{mj}{n}\pi\right)}{\sin\left(\frac{m}{n}\pi\right)} \omega_n^{-m(j-1)/2} \right).$$

The initial condition

$$a_{\tilde{n},m} = \begin{cases} 1, & \tilde{n} \le r+1 \text{ and } |\sigma_{\tilde{n},m} - \tilde{\sigma}_{\tilde{n},m}| \le \rho_m, \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

and additional condition

$$b_{\tilde{n},m} = \begin{cases} -1, & r+1 < \tilde{n} \le k+1 \\ & \text{and } |\sigma_{\tilde{n},m} - \tilde{\sigma}_{\tilde{n},m}| \le \rho_m, \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where

$$\tilde{\sigma}_{\tilde{n},m} = -\frac{\sin\left(\frac{m\tilde{n}}{n}\pi\right)}{\sin\left(\frac{m}{n}\pi\right)} \omega_n^{-m(\tilde{n}-1)/2}. \tag{5}$$

Here $0 \le k \le \tilde{n}$, $0 \le d \le k$, $0 \le r \le k$.

D. Entire set of the spectral components

In this case we can consider a set $\boldsymbol{D}_{\boldsymbol{\sigma}_{\tilde{n}},\boldsymbol{\rho}}$ of the disks $\mathscr{D}_{\sigma_{\tilde{n},0},\rho_0}$, $\mathscr{D}_{\sigma_{\tilde{n},1},\rho_1}, \dots,$ $\mathscr{D}_{\sigma_{\tilde{n},n-1},\rho_{n-1}}$, where $\boldsymbol{\sigma}_{\tilde{n}} = (\sigma_{\tilde{n},0}, \sigma_{\tilde{n},1}, \dots, \sigma_{\tilde{n},n-1}) \in \mathbb{C}^n$ is a vector of centres of these disks, $\boldsymbol{\rho} = (\rho_0, \rho_1, \dots, \rho_{n-1}) \in \mathbb{R}^n$ is a vector of their radii, respectively [2].

Then consider those $\tilde{n}$-length subsequences from $\mathscr{S}(n)$ whose vectors of the exponential sums belong to $\boldsymbol{D}_{\boldsymbol{\sigma}_{\tilde{n}},\boldsymbol{\rho}}$.

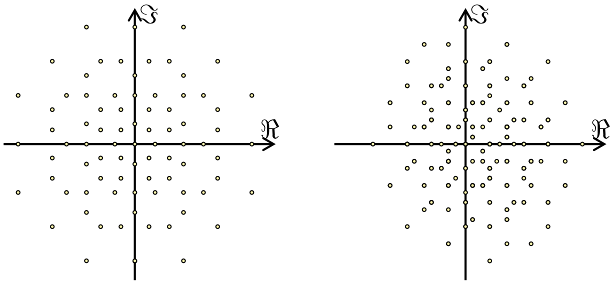
To compute the number of these subsequences an indicator δ was introduced [2]. This indicator is the logical conjunction ($\bigwedge$) of n statements each of which predicates that the mth exponential sum of trailing run (5) falls within the disk $\mathscr{D}_{\sigma_{\tilde{n},m},\rho_m}$, i.e.,

$$\delta = \bigwedge_{m=0}^{n-1} \Big[|\sigma_{\tilde{n},m} - \tilde{\sigma}_{\tilde{n},m}| \le \rho_m \Big].$$

Next the indicator is involved in conditional in (3) and (4).

[1]In this paper we focus on dkr-sequences; $dklr$-sequences deliver a more trivial case. In our examples, however, we refer to $dklr$-sequences.

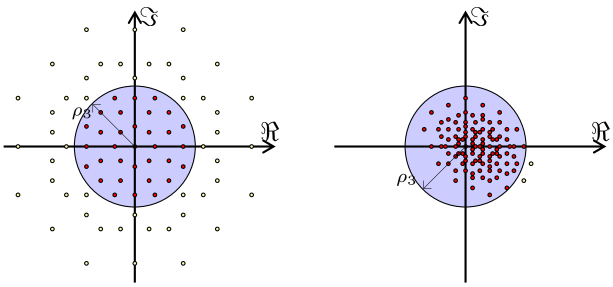
In [10] advantages of alphabet $\{-1,0,1\}$ over $\{-1,1\}$ was shown in terms of the PSD. We can evolve this in terms of distribution of the spectral components. To this end we must compare density of spectral points in the complex plane for a given spectral component. Consider an example. In Fig. 2 distributions of spectral points are depicted for binary as well as for ternary cases. For ternary sequences these points are



(a) binary case (b) ternary case

Fig. 2. Distributions of spectral points of z_1^* for all sequences of length 8.

more concentrated around the origin. This means that already for the 1st spectral component we see some energy gain which becomes essential towards $\lfloor n/2 \rfloor$ component. For our example we skip the 2nd spectral component and consider the 3rd (see Fig. 3) as more descriptive one[2].



(a) binary case (b) ternary case

Fig. 3. Distributions of spectral points of z_3^* for all sequences of length 8.

III. Conversion from Binary to Ternary

The main idea of this conversion is preventing transitions from -1 to 1 and *vice versa*. Indeed, this is what provides such a significant energy gain (see Fig. 2 and 3) together with the alphabet $\{-1,0,1\}$.

A construction for converting binary sequences to ternary ones is shown in Fig. 4 for odd length sequences and in Fig. 5 for even length sequences.

Conceptually the source binary sequence is split in two sequences such that $\hat{x}_j = x_{2j-1}$ and $\check{x}_j = x_{2j}$, where $j = 1,\ldots,\lceil n/2 \rceil$. Next, we provide intermediate results $\hat{\boldsymbol{z}}$ and $\check{\boldsymbol{z}}$ using (1). Finally, we obtain a ternary sequence $\boldsymbol{z}$ by the following rule $z_j = (\hat{z}_{\lceil j/2 \rceil} + \check{z}_{\lfloor j/2 \rfloor})/2$, where $j = 1,\ldots,n$.

The construction also requires an element $\check{z}_0 = 1$. As a result the plots in Figs. 2(b) and 3(b) are not symmetric that is particularly appreciable at small length n.

[2]Observe, for binary sequences of length 8 it coincides with the distribution for z_1^* (see Fig. 2).

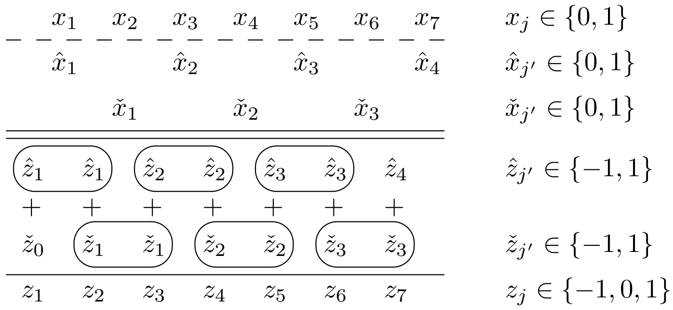


Fig. 4. Splitting odd length sequences.

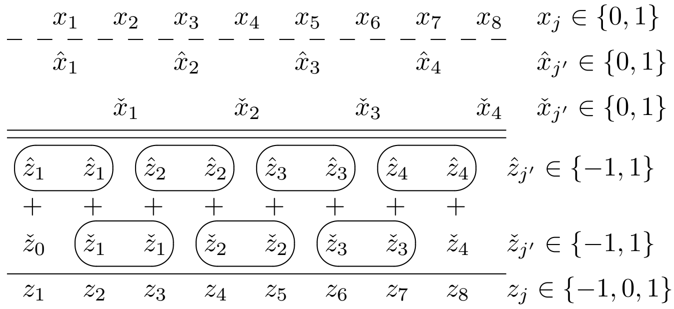


Fig. 5. Splitting even length sequences.

Proposition 1. *If the source binary sequence $\boldsymbol{x}$ satisfies some run-length constraints, then the resulting ternary sequence $\boldsymbol{z}$ also satisfies these constraints.*

The proof[3] directly follows from the construction (see Fig. 4 and 5).

To apply (2) we must consider $\hat{\boldsymbol{z}}$ and $\check{\boldsymbol{z}}$ separately. We state the following theorem.

Theorem 1. *The DFT of the ternary sequences can be expressed in terms of $\hat{\boldsymbol{z}}$ and $\check{\boldsymbol{z}}$ as follows:*

$$z_m^* = \varphi_m \sum_{j=0}^{\lfloor n/2 \rfloor - 1} \tilde{z}_{m,j+1} \omega_n^{-2mj} + \tilde{z}_{m,0}/2$$

where

$$\varphi_m = \omega_n^{-m} \cos\frac{\pi m}{n},$$
$$\tilde{z}_{m,j} = (\hat{z}_j + \check{z}_j)\cos\frac{\pi m}{n} + i(\hat{z}_j - \check{z}_j)\sin\frac{\pi m}{n}, \quad j > 0,$$
$$\tilde{z}_{m,0} = \check{z}_0 + \begin{cases} -\check{z}_{n/2}, & n \text{ even}, \\ \hat{z}_{(n+1)/2}\omega_n^m, & n \text{ odd}. \end{cases}$$

This theorem can be proved by direct calculations.

For the ternary sequences it remains to repeat the steps described in Section II-A … II-D. In the result we should obtain the number of the spectrum shaped run-length constrained ternary sequences.

IV. Constant-Charge Ternary Sequences

We presume the crux of enumeration lies in determining the number of constant-charge sequences. This means that if we find an enumerative scheme for constant-charge ternary

[3]Here and in what follows, we omit proofs due to space limitations.

TABLE I
AN EXAMPLE OF CHARGE DISTRIBUTIONS. CONSTRAINTS: $d = 2$, $k = 4$, $l = 1$, AND $r = 3$.

	$\widehat{C}_n^{\sigma_0}$ (binary)																	$\widehat{C}_n^{\sigma_0}$ (ternary)														
σ / n	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
0									1																1							
1								1		1														0	1	1						
2							1		1		0												0	0	1	1	0					
3						1		1		0		0										0	0	0	1	1	0	0				
4					1		2		0		0		0								0	0	0	1	1	1	0	0	0			
5				0		2		2		0		0		0						0	0	0	1	0	1	2	0	0	0	0		
6			0		1		2		2		0		0		0				0	0	0	1	0	2	0	0	2	0	0	0	0	
7		0		0		2		3		2		0		0		0		0	0	0	1	1	2	0	1	0	0	2	0	0	0	0
8	0		0		0		4		4		1		0		0		0	0	0	0	1	3	1	1	0	0	0	1	2	0	0	0

sequences, like we did in the binary case (see Section II-A), then it is not so hard to expand this method to enumeration of the spectrum shaped ternary sequences.

Let $C_n^{\hat{\sigma}_0,\check{\sigma}_0}$ be the number of sequences from $\mathscr{S}(n)$; these sequences have charge $\sigma_0 = z_0^* = (\hat{\sigma}_0 + \check{\sigma}_0)/2$.

Proposition 2. *The number $C_n^{\hat{\sigma}_0,\check{\sigma}_0}$ can be obtained as:*

$$C_n^{\hat{\sigma}_0,\check{\sigma}_0} = \begin{cases} 0, & n < d+1, \\ \sum_{j=d+1}^{\min(n,k+1)} C_{n-j}^{\hat{\vartheta}-j,\check{\vartheta}-j}, & \text{otherwise}, \end{cases}$$

where

$$\hat{\vartheta} = \begin{cases} -\hat{\sigma}_0, & j \text{ even}, \\ \check{\sigma}_0, & j \text{ odd}, \end{cases}$$

$$\check{\vartheta} = \begin{cases} \check{\sigma}_0, & j \text{ even}, \\ -\hat{\sigma}_0, & j \text{ odd} \end{cases}$$

with initial conditions

$$C_n^{-n,n} = \begin{cases} 1, & n \le r+1, \\ 0, & \text{otherwise}. \end{cases}$$

Here $0 \le k \le \tilde{n}$, $0 \le d \le k$, $0 \le r \le \tilde{n}$.

The value $2\sigma_0$ is a composition of $\hat{\sigma}_0$ and $\check{\sigma}_0$ (including negative numbers). Therefore,

$$C_n^{\sigma_0} = \sum_{\substack{-n \le \{\hat{\sigma}_0,\check{\sigma}_0\} \le n \\ \hat{\sigma}_0+\check{\sigma}_0 = 2\sigma_0}} C_n^{\hat{\sigma}_0,\check{\sigma}_0}. \tag{6}$$

By using these relations, we can obtain the charge distribution of our sequences. An example of such a distribution is shown in Table I. We have borrowed the left triangle from [8] to compare it with our ternary case.

V. THE NUMBER OF SPECTRUM SHAPED TERNARY SEQUENCES

Let $C_{n,\tilde{n},m}^{\hat{\sigma}_{\tilde{n},m},\check{\sigma}_{\tilde{n},m},\hat{\rho}_m,\check{\rho}_m}$ be the number of $\tilde{n}$-length subsequences from $\mathscr{S}(n)$; the corresponding subsequences $\hat{\boldsymbol{z}}$ and $\check{\boldsymbol{z}}$ have mth exponential sums $\hat{z^*}_m$ and $\check{z^*}_m$ falling within the disks $\hat{\mathscr{D}}_{\hat{\sigma}_{\tilde{n},m},\hat{\rho}_m}$ and $\check{\mathscr{D}}_{\check{\sigma}_{\tilde{n},m},\check{\rho}_m}$ centered at $\hat{\sigma}_{\tilde{n},m} \in \mathbb{C}$ and $\check{\sigma}_{\tilde{n},m} \in \mathbb{C}$ respectively.

Proposition 3. *The number $C_{n,\tilde{n},m}^{\hat{\sigma}_{\tilde{n},m},\check{\sigma}_{\tilde{n},m},\hat{\rho}_m,\check{\rho}_m}$ can be obtained as:*

$$C_{n,\tilde{n},m}^{\hat{\sigma}_{\tilde{n},m},\check{\sigma}_{\tilde{n},m},\hat{\rho}_m,\check{\rho}_m} = \begin{cases} a_{\tilde{n},m}, & \tilde{n} < d+1, \\ \sum\limits_{j=d+1}^{\min(\tilde{n},k+1)} C_{n,\tilde{n}-j,m}^{\hat{\vartheta},\check{\vartheta},\hat{\rho}_m,\check{\rho}_m} + b_{\tilde{n},m}, & \text{otherwise}, \end{cases}$$

where

$$\hat{\vartheta} = \begin{cases} -\omega_n^{mj}\left(\hat{\sigma}_{\tilde{n},m} + \frac{\sin\left(\frac{mj}{n}\pi\right)}{\sin\left(\frac{m}{n}\pi\right)}\omega_n^{-m(j-1)/2}\right), & j \text{ even}, \\ \omega_n^{mj}\left(\check{\sigma}_{\tilde{n},m} - \frac{\sin\left(\frac{mj}{n}\pi\right)}{\sin\left(\frac{m}{n}\pi\right)}\omega_n^{-m(j-1)/2}\right), & j \text{ odd}, \end{cases}$$

$$\check{\vartheta} = \begin{cases} \omega_n^{mj}\left(\check{\sigma}_{\tilde{n},m} - \frac{\sin\left(\frac{mj}{n}\pi\right)}{\sin\left(\frac{m}{n}\pi\right)}\omega_n^{-m(j-1)/2}\right), & j \text{ even}, \\ -\omega_n^{mj}\left(\hat{\sigma}_{\tilde{n},m} + \frac{\sin\left(\frac{mj}{n}\pi\right)}{\sin\left(\frac{m}{n}\pi\right)}\omega_n^{-m(j-1)/2}\right), & j \text{ odd}, \end{cases}$$

initial condition

$$a_{\tilde{n},m} = \begin{cases} 1, & \tilde{n} \le r+1, \\ & \text{and } |\hat{\sigma}_{\tilde{n},m} - \tilde{\sigma}_{\tilde{n},m}| \le \hat{\rho}_m, \\ & \text{and } |\check{\sigma}_{\tilde{n},m} + \tilde{\sigma}_{\tilde{n},m}| \le \check{\rho}_m, \\ 0, & \text{otherwise}, \end{cases}$$

and additional condition

$$b_{\tilde{n},m} = \begin{cases} -1, & r+1 < \tilde{n} \le k+1, \\ & \text{and } |\hat{\sigma}_{\tilde{n},m} - \tilde{\sigma}_{\tilde{n},m}| \le \hat{\rho}_m, \\ & \text{and } |\check{\sigma}_{\tilde{n},m} + \tilde{\sigma}_{\tilde{n},m}| \le \check{\rho}_m, \\ 0, & \text{otherwise}, \end{cases}$$

where $\tilde{\sigma}_{\tilde{n},m}$ is defined by (5).

Here $0 \le k \le \tilde{n}$, $0 \le d \le k$, $0 \le r \le k$.

Let $C_{n,\tilde{n},m}^{\sigma_{\tilde{n},m},\rho_m}$ be the number of $\tilde{n}$-length subsequences from $\mathscr{S}(n)$; these sequences have mth exponential sum z_m^* falling within a disk $\mathscr{D}_{\sigma_{\tilde{n},m},\rho_m}$ of radius $\rho_m = (\hat{\rho}_m + \check{\rho}_m)/2$ centered at $\sigma_{\tilde{n},m} = (\hat{\sigma}_{\tilde{n},m} + \check{\sigma}_{\tilde{n},m})/2$, then

$$C_{n,\tilde{n},m}^{\sigma_{\tilde{n},m},\rho_m} = \sum_{\substack{\sigma_{\tilde{n},m} = (\hat{\sigma}_{\tilde{n},m}+\check{\sigma}_{\tilde{n},m})/2 \\ \rho_m = (\hat{\rho}_m+\check{\rho}_m)/2}} C_{n,\tilde{n},m}^{\hat{\sigma}_{\tilde{n},m},\check{\sigma}_{\tilde{n},m},\hat{\rho}_m,\check{\rho}_m}. \tag{7}$$

To estimate the number of compositions in (6) and (7) we can consider a bipartite graph in which the vertices on one side of the bipartition represent the DFT points of $\hat{\boldsymbol{z}}$ while the vertices on the other side represent the DFT points of $\check{\boldsymbol{z}}$. Then the problem reduces to the problem on matching [11].

As a simple and practically sufficient solution we suggest usage of coinciding disks of equal radii. This method becomes more efficient as n increases.

It remains to consider the entire range of the spectral components. We follow Section II-D and get

$$\delta = \bigwedge_{m=0}^{n-1} \left(|\hat{\sigma}_{\tilde{n},m} - \tilde{\sigma}_{\tilde{n},m}| \le \hat{\rho}_m \text{ and } |\check{\sigma}_{\tilde{n},m} - \tilde{\sigma}_{\tilde{n},m}| \le \check{\rho}_m \right).$$

VI. Spectrum of Sequences of Infinite Length

For practical usage it remains to consider how satisfaction of a complex-plane constraint, which is imposed on DFT components of finite-length sequences, influences the spectral characteristics of the infinite concatenation of the finite-length spectrum shaped sequences chosen uniformly at random from $\hat{\mathscr{S}}(n)$. For the sake of simplicity, we consider the binary case that can be trivially expanded to the ternary sequences $\boldsymbol{z}$.

The expression

$$Z^{(n)}(f) = \sum_{j=0}^{n-1} z_{j+1} e^{-2\pi i f j}, \quad |f| \leq 1/2$$

is called a finite-length Discrete-Time Fourier Transform (DTFT) of the sequence $\boldsymbol{z}$.

The average power spectral density $\Phi(f)$ of the signal is defined as

$$\Phi(f) = \lim_{M\to\infty} \mathbb{E}\left\{ \frac{1}{M} \left| Z^{(M)}(f) \right|^2 \right\} \tag{8}$$

where $\mathbb{E}\{\cdot\}$ denotes the expectation over the ensemble of realizations of the signal.

We can readily reconstruct the DTFT from DFT. Also it is not so hard to find a relationship between spectra of sequences from $\hat{\mathscr{S}}(n)$ and the DTFT of the overall signal. To this end, we can write the resulting DTFT as

$$Z^{(Mn)}(f) = \sum_{t=0}^{M-1} \sum_{j=0}^{n-1} z_{j+1}(N_t) e^{-2\pi i f (nt+j)} \tag{9}$$

where by $N_t \in \mathbb{N}_0$, $N_t < |\hat{\mathscr{S}}(n)|$ we denote the lexicographic index of a sequence $\boldsymbol{x}$ randomly chosen from $\hat{\mathscr{S}}(n)$.

The inverse DFT is defined as

$$z_j = \frac{1}{n} \sum_{m=0}^{n-1} z_m^* e^{2\pi i \frac{mj}{n}}, \quad j = 1, \ldots, n.$$

Substituting it for $z_j(N_t)$ in (9) we readily obtain

$$Z^{(Mn)}(f) = \frac{e^{-\pi i f (n-1)}}{n} \sum_{t=0}^{M-1} e^{-2\pi i f n t} \times \sum_{m=0}^{n-1} z_m^*(N_t) \frac{\sin\left((m - fn)\pi\right)}{\sin\left(\left(\frac{m}{n} - f\right)\pi\right)} e^{\pi i m \left(1 + \frac{1}{n}\right)}. \tag{10}$$

We could estimate the PSD now, but this does not solve our problem. Indeed, the relationship between the DFT constraints and the PSD is yet to be stated. More precisely, we must define a range in the frequency domain $0 \leq f \leq 1/2$; the range depends on the spectral constraints $\boldsymbol{\sigma}$, $\boldsymbol{\rho}$, and it should be defined such that the PSD of the signal falls within this range.

First, assume $\boldsymbol{\sigma} = 0$ and $\boldsymbol{D}_{\boldsymbol{\sigma},\boldsymbol{\rho}}$ becomes a set $\boldsymbol{A}_{\boldsymbol{\rho}}$ of circles $\mathscr{A}_{\rho_0}, \mathscr{A}_{\rho_1}, \ldots, \mathscr{A}_{\rho_{n-1}}$ centered at the origin of the complex plane. Secondly, assume the run-length constraints are relaxed, i.e., $d = 0$, $k = l = r = n$. Next, we can choose a set $\hat{\mathscr{S}}_{\boldsymbol{\rho}}$ such that the DFT components of all sequences $\boldsymbol{z}$ lie on the circles $\mathscr{A}_{\rho_j}$, $0 \leq j < n$.

Lemma 1. *The average power spectral density of any NRZI encoded infinite concatenation of the sequences $\boldsymbol{x} \in \hat{\mathscr{S}}_{\boldsymbol{\rho}}$ is*

$$\Phi_n^{\boldsymbol{\rho}}(f) = \frac{1}{n} \sum_{m=0}^{n-1} \left(\frac{\rho_m}{n} \frac{\sin\left((m - fn)\pi\right)}{\sin\left(\left(\frac{m}{n} - f\right)\pi\right)} \right)^2, \quad \rho_m = |z_m^*| \tag{11}$$

if and only if the sequences $\boldsymbol{x}$ are chosen uniformly at random from $\hat{\mathscr{S}}_{\boldsymbol{\rho}}$.

Now let $\hat{\mathscr{S}}_{\boldsymbol{\rho}} = \bigcup_{u=1}^{L} \hat{\mathscr{S}}_{\boldsymbol{\rho}_u}$, $L \in \mathbb{N}$. For each $\hat{\mathscr{S}}_{\boldsymbol{\rho}_u}$ Lemma 1 holds true.

Lemma 2. *The average power spectral density of any NRZI encoded stochastic concatenation of the sequences $\boldsymbol{x} \in \hat{\mathscr{S}}_{\boldsymbol{\rho}}$ is as follows:*

$$\Phi_n(f) = \frac{1}{|\hat{\mathscr{S}}_{\boldsymbol{\rho}}|} \sum_{\hat{\mathscr{S}}_{\boldsymbol{\rho}_u} \subset \hat{\mathscr{S}}_{\boldsymbol{\rho}}} |\hat{\mathscr{S}}_{\boldsymbol{\rho}_u}| \Phi_n^{\boldsymbol{\rho}_u}(f).$$

Lemma 3. *If $f_m = m/n$, then*

$$\Phi_n(f_m) \leq \max_{1 \leq u \leq L} \frac{\rho_{u,m}^2}{n}.$$

Theorem 2. *The PSD of any NRZI encoded stochastic concatenation of the sequences $\boldsymbol{x}$ at frequency $f = m/n$ falls within the interval $\left[0, \frac{\rho_m^2}{n}\right]$ if and only if the value of mth DFT component of the NRZI encoded sequence $\boldsymbol{x} \in \hat{\mathscr{S}}(n)$ falls within the disk $\mathscr{D}_{\sigma_m,\rho_m}$ where $\sigma_m = 0$.*

VII. Conclusion

The proposed pseudo-ternary sequences are mostly suitable for communication systems and have a significant advantage over the binary ones in terms of capacity.

References

[1] K. A. S. Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2260–2299, Oct 1998.

[2] O. F. Kurmaev, "Enumeration of spectrum shaped binary run-length constrained sequences," in *Proc. 3rd Intl. Castle Meeting on Coding Theory and Applic. (3ICMCTA)*, Cardona (Barcelona). - Spain., Sep. 11–15 2011, pp. 175–180.

[3] A. Croisier, "Introduction to pseudoternary transmission codes," *IBM J. Res. Dev.*, vol. 14, no. 4, pp. 354–367, Jul. 1970.

[4] C. French, G. Dixon, and J. Wolf, "Results involving (d,k) constrained m-ary codes," *IEEE Trans. Magn.*, vol. 23, no. 5, pp. 3678–3680, Sep. 1987.

[5] S.-K. Lee, H.-Y. Chen, M.-C. Lin, T.-H. Chen, H.-F. Chou, and Y.-H. Ku, "On soft iterative decoding for ternary recording systems with rll constraints," in *IEEE Intl. Symp. on Inform. Theory Applic. (ISITA 2010)*, Taichung, Taiwan, Oct. 17–20 2010, pp. 484–489.

[6] E. Gorog, "Redundant alphabets with desirable frequency spectrum properties," *IBM J. Res. Dev.*, vol. 12, no. 3, pp. 234–241, Jul. 1968.

[7] V. Braun and K. A. S. Immink, "An enumerative coding technique for DC-free runlength-limited sequences," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2024–2031, Dec. 2000.

[8] O. F. Kurmaev, "Constant-weight and constant-charge binary run-length limited codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4497–4515, July 2011.

[9] T. M. Cover, "Enumerative source coding," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 1, pp. 73–77, Jan. 1973.

[10] I. J. Fair, "Construction and characteristics of codewords and pulse shapes that satisfy spectral constraints," *IEEE Trans. Commun.*, vol. 56, no. 10, pp. 1585–1590, Oct. 2008.

[11] A. Gibbson, *Algorithmic Graph Theory*. Cambridge University Press, 1985.