A Nonlinear Sum of Squares Search for CAZAC Sequences

Mark Magsino  
Department of Mathematics  
U.S. Naval Academy  
magsino@usna.edu

Yixin Xu  
Department of Mathematics  
Ohio State University  
xu.3520@buckeyemail.osu.edu

Abstract—We report on a search for CAZAC sequences by using a nonlinear sum of squares optimization procedure. Up to equivalence, we found all length 7 CAZAC sequences. We obtained evidence suggesting there are finitely many length 10 CAZAC sequences with a total of 3040 sequences. Last, we compute longer CAZAC sequences and compare their aperiodic autocorrelation properties to the well-known Zadoff-Chu and Björck sequences. The code and results of this search are publicly available through GitHub.

I. INTRODUCTION

Given \( x \in \mathbb{C}^n \), we define the periodic ambiguity function by

\[
A_p(x)[k, \ell] := \frac{1}{n} \sum_{j=0}^{n-1} x_{j+k}\overline{x_j} e^{-2\pi i j \ell/n},
\]

where \( 0 \leq k, \ell \leq n - 1 \), and indices are taken modulo \( n \). The periodic autocorrelation is the case with no frequency shift, i.e. when \( \ell = 0 \). In this case it is convenient to suppress the second input and write

\[
A_p(x)[k] := A_p(x)[k, 0] = \frac{1}{n} \sum_{j=0}^{n-1} x_{j+k}\overline{x_j}.
\]

A CAZAC sequence of length \( n \) is a vector \( x \in \mathbb{C}^n \) with the properties

(i) \( |x_j| = 1 \) for \( 0 \leq j \leq n - 1 \),

(ii) \( A_p(x)[k] = 0 \) for \( 1 \leq k \leq n - 1 \).

The first property is known as the constant amplitude property and the second is known as the zero autocorrelation property which gives rise to the acronym CAZAC.

These sequences have several interpretations and applications which motivate their study. In communication theory they are used to reduce the cross-correlation of signals, perform uplink synchronization [9], and OFDM for 5G communication [7][15]. It is also studied as an idealized waveform with regards to the narrowband ambiguity function in radar [14]. In general, the constant amplitude allows one to encode information purely in terms of phase and the zero autocorrelation ensures no interference with shifted copies of the signal. CAZAC sequences are also known as perfect polyphase sequences when they are comprised of roots of unity and are also well studied under that name [11][12][13].

Although CAZAC sequences have perfect periodic autocorrelation, many applications require good aperiodic autocorrelation. The aperiodic ambiguity function of a sequence \( x \in \mathbb{C}^n \) is given by

\[
A_a(x)[k, \ell] := \sum_{j=1}^{n-1} x^{(a)}_{j+k}\overline{x^{(a)}_j} e^{-2\pi i j \ell/n},
\]

where \( 0 \leq k, \ell \leq n - 1 \) and \( x^{(a)}_j \) is defined by

\[
x^{(a)}_j = \begin{cases} x_j, & \text{if } 0 \leq j \leq n - 1, \\ 0, & \text{otherwise}. \end{cases}
\]

The key difference is instead of taking indices modulo \( n \), we set the value to zero when the index falls outside of the range. The aperiodic autocorrelation corresponds to \( \ell = 0 \) and we write

\[
A_a(x)[k] := A_a(x)[k, 0] = \sum_{j=1}^{n-1} x^{(a)}_{j+k}\overline{x^{(a)}_j}.
\]
In particular, we use aperiodic autocorrelation to study two properties of interest. We define the peak sidelobe level (PSL) of \( x \in \mathbb{C}^n \) by
\[
\text{PSL}(x) = \frac{1}{|A_n(x)[0]|} \max_{k \neq 0} |A_n(x)[k]|, \tag{6}
\]
and the integrated sidelobe level (ISL) by
\[
\text{ISL}(x) = \frac{1}{|A_n(x)[0]|^2} \sum_{k=1}^{n-1} |A_n(x)[k]|^2. \tag{7}
\]

CAZAC sequences are well studied but many of their properties are still unknown. We say two CAZAC sequences, \( x \) and \( y \), are equivalent if there exists a complex scalar \( c \) with \( |c| = 1 \) so that \( y = cx \). The representative of each equivalence class is the sequence whose first entry is 1. With this in mind, it is natural to ask: For each \( n \), how many CAZAC sequences of length \( n \) are there? Haagerup proved if \( n \) is prime, then there are at most \( \binom{2n-2}{n-1} \) CAZAC sequences [8]. Björck and Saffari constructed an infinite family of sequences when \( n \) is composite and divisible by a perfect square [6][10]. If \( n \) is composite and not divisible by any perfect square, then it is unknown how many there are in general. A brute force calculation shows there are finitely many for \( n = 6 \) [5].

Proposition 1 describes additional transformations under which CAZAC sequences are closed [3]. They are a finite set of transformations so it does not fundamentally change the question of whether the set of CAZAC sequences of a given length is finite. We use these to filter out known length 7 CAZAC sequences.

**Proposition 1** (Theorem 2.1 in [3]). Let \( x \in \mathbb{C}^n \) be a CAZAC sequence and let \( \omega = e^{2\pi i/n} \). Then, the following transformations of \( x \) are also CAZAC sequences:

(i) \( (T_kx)_j := x_{j+k}, \ 0 \leq k \leq n-1 \),
(ii) \( (M_{\ell}x)_j := \omega^{\ell j}x_j, \ 0 \leq \ell \leq n-1 \),
(iii) \( (D_m x)_j := x_{mj} \), gcd \((m, n) = 1\),
(iv) \( (\overline{x})_j := \overline{x_j} \).

**II. CAZAC Sequences of Length 7**

The CAZAC sequences of length 7 can be split into quadratic phase sequences and non-quadratic phase sequences. Suppose \( x \in \mathbb{C}^n \) is defined by
\[
x_j = e^{\pi ip(j)/n},
\]
where \( p(j) \) is a quadratic polynomial. In this case, we say that \( x \) is a quadratic phase sequence. The polynomials associated with the known quadratic phase CAZAC sequences are

- **Zadoff-Chu:** \( p(j) = j(j-1), \) (\( n \) odd)
- **P4:** \( p(j) = j(j-3) \)
- **Wiener:** \( p(j) = 2kj^2, \) \( \gcd(k, n) = 1, \) odd
- \( p(j) = kj^2, \) \( \gcd(k, 2n) = 1, \) even

When \( n \) is prime, there are at least \( n(n-1) \) CAZAC sequences comprised of roots of unity, including the quadratic phase sequences [2]. When \( n = 7 \), this gives at least 42 roots of unity sequences. Moreover, the transformations described in Proposition 1 will keep the sequence a root of unity sequence. On the other hand, in [4] Björck constructed CAZAC sequences comprised of non roots of unity for each prime \( p > 5 \). The construction is as follows.

Given an odd prime \( p \), let \( \left( \frac{j}{p} \right) \) denote the Legendre symbol defined by
\[
\left( \frac{j}{p} \right) = \begin{cases} 
0, & \text{if } j \equiv 0 \mod p, \\
1, & \text{if } j \equiv x^2 \mod p, \text{ for some } x \neq 0, \\
-1, & \text{if } p \equiv j \neq x^2 \mod p, \text{ for any } x \neq 0.
\end{cases}
\]

We define the Björck sequence of length \( p \) by
\[
\theta(j) = \begin{cases} 
\arccos \left( \frac{1-p}{1+p} \right), & \text{if } \left( \frac{j}{p} \right) = -1, \\
0, & \text{otherwise.}
\end{cases}
\]

Since \( p \equiv 3 \mod 4 \), then \( \theta(j) \) is given by
\[
\theta(j) = \begin{cases} 
\arccos \left( \frac{1-p}{1+p} \right), & \text{if } \left( \frac{j}{p} \right) = -1, \\
0, & \text{otherwise.}
\end{cases}
\]

Since \( 7 \equiv 3 \mod 4 \), the Björck sequence of length 7 is
\[
x = (1, 1, e^{i\theta_7}, 1, e^{i\theta_7}, e^{i\theta_7}), \tag{11}
\]
where \( \theta_7 = \arccos(3/4) \).
In [8] Haagerup studies a system of polynomial equations which give rise to so-called cyclic $p$-roots. CAZAC sequences always give rise to cyclic $p$-roots but the reverse is not true. Björck and Fröberg used computer algebra to compute a Gröbner basis for cyclic 7-roots and filtered them further to show there are only 532 length 7 CAZAC sequences [5], which is much lower than the bound implied by Haagerup’s work.

The transformations in Proposition 1 on the length 7 Björck sequence at most 252 sequences, leaving at most 294 accounted for after adding the roots of unity sequences. In principle, the Gröbner basis discovered by Björck and Fröberg can be used to find the remaining sequences exactly, but this is not recorded in prior work. Instead, we find the remaining sequences numerically to record and study them computationally more easily.

III. CAZACs AS A REAL ALGEBRAIC VARIETY

The conditions defining a CAZAC sequence can be viewed as a system of $2n-1$ equations with $n$ variables. The dimension of the set of solutions helps determine if the set is finite. If the dimension is zero, then it is a discrete set and may be finite. Conversely, if the dimension is positive, then there are infinitely many of them. This puts the problem in the realm of algebraic geometry.

The techniques of algebraic geometry require systems of polynomials. However, the conjugations in the sums prevent those conditions from being interpreted as polynomials. We get around this by expressing entries in real and imaginary parts and expressing the equations to polynomial equations of real variables. Given a vector $x \in \mathbb{C}^n$ and expressing its entries as $x_j = a_j + ib_j$, CAZAC sequences arise as solutions to the system

$$a_j^2 + b_j^2 = 1, \quad 0 \leq j \leq n-1,$$

$$\sum_{j=0}^{n-1} a_{j+k}a_j + b_{j+k}b_j = 0, \quad 1 \leq k \leq n-1,$$

$$\sum_{j=0}^{n-1} a_jb_{j+k} - b_ja_{j+k} = 0, \quad 1 \leq k \leq n-1. \quad (14)$$

Constraining each amplitude gives $n$ conditions, and each of the $n-1$ autocorrelation constraints has been split into a constraint on the real and imaginary parts, respectively. Thus, there are $3n-2$ equations in this system. Each entry of the sequence was split into a real and imaginary part so the system has $2n$ variables. Hence, the system is a real algebraic variety with $3n-2$ equations and $2n$ variables.

IV. NONLINEAR LEAST SQUARES SEARCH

The system of equations (12) - (14) can be converted into an unconstrained nonlinear sum of squares optimization problem. Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$, and for each $1 \leq k \leq n-1$ and $0 \leq \ell \leq n-1$, define the functions

$$f_\ell(a, b) = a_\ell^2 + b_\ell^2 - 1,$$

$$g_k(a, b) = \sum_{j=0}^{n-1} a_{j+k}a_j + b_{j+k}b_j, \quad (16)$$

$$h_k(a, b) = \sum_{j=0}^{n-1} a_jb_{j+k} - b_ja_{j+k}. \quad (17)$$

Note that $(a, b)$ are the real and imaginary parts, respectively, of a CAZAC sequence precisely when all $3n-2$ functions are all zero at $(a, b)$. We can obtain CAZAC sequences as solutions to the unconstrained optimization problem

$$\arg \min_{(a, b)} \sum_{\ell=0}^{n-1} f_\ell^2 + \sum_{k=1}^{n-1} g_k^2 + h_k^2, \quad (18)$$

for which the objective function is zero.

This objective is nonlinear and non-convex so there may be local minima where the objective function is positive and optimization fails to find a CAZAC sequence. Nonetheless, in our experiments we found very few such local minima and generated as many sequences as desired for each test.

V. METHOD AND RESULTS

The experimental design was implemented in a Jupyter Notebook. All trials were run on a 2020 MacBook Pro with the 10-core M1 Pro CPU and 16 GB of physical memory. The optimization problem (18) was implemented using the least squares solver in SciPy’s Optimization package. By default the
solver uses the Trust Region Reflective (TRR) algorithm and a finite difference operation to estimate the Jacobian at each point. The code is available through GitHub at: https://github.com/magsinos/usna/IEEE-SoS-CAZAC.git.

A. Length 7 CAZACS

We perform the optimization on 10,000 random initial starting points in $[-1,1]^4$. For increased precision, we used stronger tolerances of $10^{-12}$ instead of the default $10^{-8}$ for changes in the gradient, input variables, and cost function. If the resulting cost function was lower than $10^{-10}$, we considered the point found as a CAZAC sequence stored it as a new row in an array.

After generating the list of sequences, we rounded each one to 8 decimal places and pass the rounded list through NumPy’s algorithm for determining unique rows in arrays and record the number of sequences found. After that, we used the transformations listed in Proposition 1 using the Björck and the Wiener sequences as base sequences to filter out all previously explicitly known CAZAC sequences.

This process enumerated all 532 CAZAC sequences, handling equivalence by dividing to make the first entry 1. It took 15 min to run all 100,000 attempts on the aforementioned MacBook Pro. The maximum final value of the objective function across all trials was on the order of $10^{-17}$ which implies the objective function has no spurious local minima in the length 7 case. We picked out the new CAZAC with the best PSL to illustrate how its aperiodic autocorrelation compares to the length 7 Zadoff-Chu and Björck sequence in Fig. 1.

B. Length 10 CAZACs

Since 10 is composite and not divisible by a perfect square, it is not known if there are finitely many of them. Our idea for exploring whether the set of solutions is finite is based on the following intuition.

If sum of squares minimization successfully finds a zero of the objective function, then we have projected the random initial point into the set of solutions to the CAZAC equations. If the set of solutions has positive dimension, it is a continuous set and most of the initialization points should project to different points on the set of solutions. Conversely, if the set of solutions is zero dimensional and finite, then the points will project to that finite set of points. Thus, if the number of unique sequences found is far smaller than the number of successful trials, the set is likely finite. This intuition is illustrated by the example in Fig. 2.

![Fig. 1. Aperiodic autocorrelation of the best performing length 7 CAZAC sequence plotted against the aperiodic autocorrelations of the length 7 Zadoff-Chu and Björck sequences. The new CAZAC sequence achieves a lower PSL than the Zadoff-Chu sequence.](image)

![Fig. 2. Experiments taking 1000 random initial points in $\mathbb{R}^2$ and using nonlinear sum of squares to project onto two sets. For both trials, all 1000 points successfully found a zero to the cost function. Left: The points were projected onto the one dimensional curve $x^3 - x - y = 0$. Right: The points were projected to the 8 points at the intersection of $x^2 + y^2 - 1 = 0$ and $8x^4 - 10x^2 - x + 2 = 0$.](image)
as in the length 7 case. We ran three sets of these trials and in all three the same 3040 CAZAC sequences were found. This heavily implies that there are 3040 CAZAC sequences of length 10, although it is still possible that points corresponding to a highly unstable minimum were missed. However, the average runtime for each set of these trials was about 13 hours, so efficiency can likely be improved. The maximum cost function value across all trials was to the order of $10^{-24}$, which implies once again that the objective function has no spurious local minima.

C. Aperiodic Autocorrelations of Longer CAZACs

We used the nonlinear sum of squares optimization method to compute 1000 CAZAC sequences of prime lengths 11, 13, 17, 23, 29, 37, 43, and 47. We then computed the PSL and ISL of those CAZAC sequences and compared them to the PSL and ISL of the Björck and Zadoff-Chu sequences using a boxplot. This is depicted in Fig. 3 and gives a sense of how generic CAZACs compare to Björck and Zadoff-Chu sequences. For these larger cases a few spurious local minima were found. This means more minimizations than the number of desired sequences were needed which increased the computation time required.

To further illustrate the properties of CAZAC sequences generated by these methods, we juxtapose the (modulus) of the ambiguity function of a length 43 numerical CAZAC sequence and the typically used used Zadoff-Chu sequence. This is depicted in Fig. 4.

VI. FUTURE DIRECTIONS

This work provides a framework for computing CAZAC sequences by using nonlinear sum of squares optimization. In addition to the potential problem of spurious local minima, scalability is an issue as well. It is possible that other optimization methods or equation solvers would lead to improved efficiency. The code has been made available through GitHub to help ensure replicability of the numerical experiments and to provide a base of code for future directions of work.

It is also likely that the system of equations defining CAZAC sequences has some symmetries or redundancies. Exploring these could make it possible to reduce the number of equations and find new sequences more efficiently. Alternatively, the transformations of Proposition 1 could be used to add constraints to the search space and reduce the number of target sequences.

Another alternative is to convert the aperiodic autocorrelations into a system of polynomials of real variables. We could then use the resulting function to create objective functions that act as a proxy for PSL or ISL and run an optimization search. This could also give us insight on what the optimal PSL and ISL values are for phase-coded waveforms.

Since this work involves an algebraic variety, it is useful to consider techniques from numerical algebraic geometry. For example, homotopy continuation [1] is a framework for analyzing isolated solutions of algebraic varieties. This could be useful in determining whether there are finitely many CAZAC sequences in currently unknown cases.

ACKNOWLEDGEMENTS

The views expressed in the paper are those of the first author and do not reflect the official policy or
position of the Department of the Navy, Department of Defense, or the U.S. Government.

The first author would also like to thank Bill Correll, Ram Narayanan, and Greg Coxson for encouraging me to submit this work for the 2023 IEEE Radar Conference.

REFERENCES

[1] JC Alexander and James A Yorke. The homotopy continuation method: numerically implementable topological procedures. Transactions of the American Mathematical Society, 242:271–284, 1978.

[2] John J Benedetto, Katherine Cordwell, and Mark Magsino. CAZAC sequences and Haagerup’s characterization of cyclic N-roots. In New Trends in Applied Harmonic Analysis, Volume 2, pages 1–43. Springer, 2019.

[3] John J Benedetto and Jeffrey J Donatelli. Ambiguity function and frame-theoretic properties of periodic zero-autocorrelation waveforms. IEEE Journal of Selected Topics in Signal Processing, 1(1):6–20, 2007.

[4] Göran Björck. Functions of modulus 1 on Zn whose Fourier transforms have constant modulus, and “cyclic n-roots”. In Recent Advances in Fourier Analysis and its Applications, pages 131–140. Springer, 1990.

[5] Göran Björck and Ralf Fröberg. A faster way to count the solutions of inhomogeneous systems of algebraic equations, with applications to cyclic n-roots. Journal of Symbolic Computation, 12(3):329–336, 1991.

[6] Göran Björck and Bahman Saffari. New classes of finite unimodular sequences with unimodular Fourier transforms. Circulant Hadamard matrices with complex entries. Comptes rendus de l'Académie des sciences. Série 1, Mathématique, 320(3):319–324, 1995.

[7] Zhenhua Feng, Ming Tang, Songnian Fu, Lei Deng, Qiong Wu, Rui Lin, Ruoxiu Wang, Ping Shum, and Deming Liu. Performance-enhanced direct detection optical OFDM transmission with CAZAC equalization. IEEE Photonics Technology Letters, 27(14):1507–1510, 2015.

[8] Uffe Haagerup. Cyclic p-roots of prime lengths p and related complex Hadamard matrices. arXiv preprint arXiv:0803.2629, 2008.

[9] Mohammad M Mansour. Optimized architecture for computing Zadoff-Chu sequences with application to LTE. In GLOBECOM 2009-2009 IEEE Global Telecommunications Conference, pages 1–6. IEEE, 2009.

[10] Andrzej Milewski. Periodic sequences with optimal properties for channel estimation and fast start-up equalization. IBM Journal of Research and Development, 27(5):426–431, 1983.

[11] Wai Ho Mow. A unified construction of perfect polyphase sequences. In Proceedings of 1995 IEEE International Symposium on Information Theory, page 459. IEEE, 1995.

[12] Wai Ho Mow and S-YR Li. Aperiodic autocorrelation properties of perfect polyphase sequences. In Proceedings Singapore ICCS/ISITA92, pages 1232–1234. IEEE, 1992.

[13] Ki-Hyeon Park, Hong-Yeop Song, Due San Kim, and Solomon W Golomb. Optimal families of perfect polyphase sequences from the array structure of fermat-quotient sequences. IEEE Transactions on Information Theory, 62(2):1076–1086, 2015.

[14] DE Vakman. The ambiguity function in the statistical theory of radar. In Sophisticated Signals and the Uncertainty Principle in Radar, pages 98–125. Springer, 1968.

[15] Krzysztof Wesolowski, Adrian Langowski, and Krzysztof Bakowski. A novel pilot scheme for 5g downlink transmission. In 2015 International Symposium on Wireless Communication Systems (ISWCS), pages 161–165. IEEE, 2015.