Enumeration of Parallelograms in Permutation Matrices for Improved Bounds on the Density of Costas Arrays

Christopher N. Swanson
Ashland University
Ashland, OH, U.S.A.
cswanson@ashland.edu

Bill Correll, Jr.
MDA Information Systems LLC
Ypsilanti, MI, U.S.A.
william.correll@mdaus.com

Randy W. Ho
Garmin International
Chandler, AZ, U.S.A.
googlypower@gmail.com

Submitted: Oct 2, 2015; Accepted: Feb 2, 2016; Published: Mar 4, 2016
Mathematics Subject Classifications: 05A05, 05B30

Abstract

A Costas array of order n is an $n \times n$ permutation matrix such that all vectors between pairs of ones are distinct. Thus, a permutation matrix fails to be a Costas array if and only if it contains ones that form a (possibly degenerate) parallelogram. In this paper, we enumerate parallelograms in an $n \times n$ permutation matrix. We use our new formulas to improve Davies's $O(n^{-1})$ result for the density of Costas arrays.

Keywords: Costas array; Permutation; Enumeration

1 Introduction

In 1965, John P. Costas introduced a special class of permutations of n elements with applications to improving the target detection performance of radar and sonar systems [6], [7]. Edgar Gilbert also wrote about this special class of permutations in the same year, referring to them as permutations with distinct differences and devising a construction method for certain values of n [14]. Today, these permutations are referred to as Costas arrays and mathematicians and engineers alike have been studying these structures in an attempt to answer some fundamental questions about them, such as how to construct these arrays, whether these arrays exist for all n, and how many distinct Costas arrays of order n exist. Readers interested in more details regarding the applications of Costas arrays in target detection should read Section III of [1].

A *Costas array* of order n is an $n \times n$ matrix of ones and zeroes such that there is exactly one 1 per row and column (i.e., it is a permutation matrix) and such that all vectors between pairs of ones are distinct, i.e., if the permutation matrix has 1's at (i_k, j_k) for $k = 1, 2, 3, 4$ and if $i_1 - i_2 = i_3 - i_4$ and $j_1 - j_2 = j_3 - j_4$, then $(i_1, j_1) = (i_3, j_3)$ and $(i_2, j_2) = (i_4, j_4)$. Thus, a permutation matrix fails to be a Costas array if and only if it contains (at least 3) 1's that form a (possibly degenerate) parallelogram.

One easy way to check if a given permutation matrix is a Costas array is to consider its corresponding permutation and form a triangular difference table of values where the first row of the table contains the differences between the elements of the permutation that are adjacent, the second row contains the differences between the elements that are 2 places apart, etc. Then the permutation matrix is a Costas array if and only if the entries in each row of the triangular difference table are distinct. For example, consider the permutation 1243. Its triangular difference table and corresponding matrix are

$$\begin{array}{ccc} -1 & -2 & 1 \\ -3 & -1 & \\ -2 & & \end{array} \qquad A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Since no row of this table has repeated entries, the permutation matrix A for 1243 is a Costas array. Alternatively, consider the permutation 3214. Its triangular difference table and corresponding matrix are

$$\begin{array}{ccc} 1 & 1 & -3 \\ 2 & -2 & \\ -1 & & \end{array} \qquad B = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Since the first row of this table contains the difference 1 twice, the permutation matrix B for 3214 is not a Costas array.

In this paper we count how many parallelograms of 1's are possible in an $n \times n$ permutation matrix. At first glance, this might seem like a tangential enumerative problem. However, by understanding the ways a permutation matrix fails to be a Costas array, we are led towards new insights about density among permutation matrices and enumeration. The proofs of our formulas make use of a geometric fact about quadrilaterals and a binomial coefficient identity known as the hockey-stick identity.

2 Challenges and Enumerative Progress

In 1984, S. Golomb, using previously available empirical constructions, stated and proved two algebraic construction methods for Costas arrays based on finite fields, known as the

Table 1: Number of Costas arrays $C(n)$ as a function of order n

n	$C(n)$	n	$C(n)$	n	$C(n)$	n	$C(n)$	n	$C(n)$
1	1	7	200	13	12828	19	10240	25	88
2	2	8	444	14	17252	20	6464	26	56
3	4	9	760	15	19612	21	3536	27	204
4	12	10	2160	16	21104	22	2052	28	712
5	40	11	4368	17	18276	23	872	29	164
6	116	12	7852	18	15096	24	200	30	?

Golomb and Welch methods [15], [16] (Gilbert's construction in [14] is the logarithmic Welch construction.) These remain the only general constructions available and work for infinitely many orders due to the infinitude of primes, but not all orders. It was then that Costas arrays acquired their present name and became an object of study (see [8] for more details and a summary of results).

The challenges of enumerating Costas arrays and improving estimates of their asymptotic density are two of the four core challenges of Costas arrays[10]. It is not known if Costas arrays of all orders exist. Besides arrays from the constructions that have been devised, sporadic arrays not arising from known constructions have also been found in exhaustive computer searches. Massively-distributed backtracking searches [13][11][12] in recent years have helped to identify all Costas arrays for $n \leqslant 29$. The search for $n = 29$ required the equivalent of 366.55 years for a single CPU [12]. Define $C(n)$ to be the number of Costas arrays of order n. All known values of $C(n)$ are given in Table 1. Correll [3] used Möbius inversion to establish a theoretical formula for $C(n)$, but the resulting summands are impractical to evaluate. James Beard has tabulated all known Costas arrays out to $n = 500$ [2]. Discovery of arrays up to and including order 31 has been achieved but it is not known if any Costas arrays of order 32 or 33 exist. Aggressive computing with an efficient new recursive approach [17] to $C(n)$ may change this, however. Verification of the Costas property for a candidate permutation of order n involves $O(n^3)$ comparisons; therefore, the decision problem "Is $C(n) > 0$?" is in NP [10] [1].

3 Violating Configurations

Davies, within [8], notes that a permutation matrix fails to be a Costas array if it contains 3 equally-spaced collinear ones. Davies referred to these configurations as L_3-*configurations.* An L_3-configuration is not the only configuration of ones whose existence prevents a permutation matrix from being a Costas array. Four ones that form a non-degenerate parallelogram have this property. We will refer to such sets of ones as P_4-*configurations.* Two equally-spaced pairs of ones forming four distinct ones, all of which are collinear, form a degenerate parallelogram and also have this property. We will refer to such sets of ones as L_4-*configurations.* One may think of an L_3-configuration as a degenerate L_4-configuration in which two ones overlap.

For the 4 permutation matrices below, matrix A is a Costas array, matrix B is not a Costas array as it contains an L_3-configuration, matrix C is not a Costas array as it contains a P_4-configuration and matrix D is not a Costas array as it contains an L_4-configuration (as well as two L_3-configurations).

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

For $n \geqslant 3$, Davies showed that the number of L_3-configurations when considering permutations of n elements is

$$l_3(n) = \begin{cases} \frac{1}{8}n^2(n-2)^2, & n \text{ even} \\ \frac{1}{8}(n-1)^4, & n \text{ odd.} \end{cases}$$

He derived this by considering the placement of the endpoints of an L_3-configuration– see [8] for details. We created an entry in the Online Encyclopedia of Integer Sequences (OEIS, http://oeis.org/) for $l_3(n)$ which was merged into A212892. The two distinct descriptions of the same sequence suggest the following:

Problem: Let $n \geqslant 3$ be a positive integer. Let $A(n)$ be the number of ordered 4-tuples (w, x, y, z) of positive integers such that $w, x, y, z \in \{0, 1, \ldots, n\}$ with the differences $w - x, x - y$, and $y - z$ all odd. Show that $A(n - 2) = l_3(n)$.

Proof: Consider a 4-tuple (w, x, y, z) with the desired property. By parity arguments, the 4-tuple must be of the form (E,O,E,O) or (O,E,O,E) where O and E denote odd and even integers drawn with replacement from $\{0, 1, \ldots, n\}$. It follows that $A(n) = 2e(n)^2 o(n)^2$ where $e(n)$ is the number of even integers in the set and $o(n) = n + 1 - e(n)$ is the number of odd integers in the set. It is readily seen that $e(n)$ equals $(n + 2)/2$ if n is even and $(n + 1)/2$ if n is odd and that $o(n)$ equals $n/2$ if n is even and $(n + 1)/2$ if n is odd, so the result follows.

So why should we care about violating configurations? Davies studied the the random variable X counting L_3-configurations in a permutation matrix to prove the strongest-known result about the density $\frac{C(n)}{n!}$ of Costas arrays among permutation matrices. He showed that $\frac{C(n)}{n!} = O(\frac{1}{n})$ as $n \to \infty$ using the second moment method, thus settling Unsolved Problem 5 in [16]. Davies reasoned that

$$\frac{C(n)}{n!} \leqslant P(X = 0) \leqslant P(|X - \mathbb{E}(X)| \geqslant \mathbb{E}(X)) \leqslant \frac{\mathbb{E}(X^2)}{\mathbb{E}^2(X)} - 1$$

where the right-hand inequality follows from Chebyshev's Inequality. From the formula for $l_3(n)$ it follows that $\mathbb{E}(X) = \frac{l_3(n)}{n(n-1)(n-2)}$. Davies completed his proof by devising an upper bound on the second moment $\mathbb{E}(X^2)$ of X as a function of n by reasoning about how two L_3-configurations intersect. Table 1 supports the stronger conjecture that $\frac{C(n)}{n!}$ has exponential decay [4], [9].

As Davies's result only used L_3-configurations, it may be possible to improve on the upper bound of the asymptotic behavior of the density of Costas arrays by applying a similar technique that incorporates P_4-configurations and/or L_4-configurations. In order to do this, it is first necessary to determine how many L_4- and P_4-configurations exist in permutation matrices of order n, the theorems we present in this paper.

4 Configuration Enumeration

Let $l_4(n)$ and $p_4(n)$ denote the numbers of L_4- and P_4-configurations when considering permutations of n elements, respectively. Our main results are

Theorem 1. *For $n > 3$,*

$$l_4(n) = 2\sum_{a=1}^{n-1}\sum_{b=1}^{n-1}(n-a)(n-b)\left\lfloor \frac{\gcd(a,b)-1}{2} \right\rfloor$$

Theorem 2. *For $n > 3$,*

$$l_4(n) + p_4(n) = \begin{cases} 8[4\binom{\frac{n}{2}}{3} + \binom{\frac{n}{2}}{2}]^2 = \frac{1}{72}n^2(n-2)^2(2n-5)^2, & n \text{ even} \\ 8[4\binom{\frac{n+1}{2}}{3} - \binom{\frac{n-1}{2}}{2}]^2 = \frac{1}{72}(n-1)^2(n-3)^2(2n-1)^2, & n \text{ odd.} \end{cases}$$

Theorem 2 refines the result of Appendix A of [17], which shows that $l_3(n) + l_4(n) + p_4(n) = O(n^6)$. We made use of Theorem 2 in [5] where it was stated with proof in conjunction with the Lovász Local Lemma to study the existence of Costas arrays and partial Costas arrays. Our proofs for Theorems 1 and 2 make use of the geometric result that a quadrilateral is a parallelogram if and only if its diagonals bisect each other. We apply this result bijectively to enumerate violating configurations by counting pairs of diagonals of a violating configuration with the same midpoint such that each of the 4 endpoints has distinct x- and y-coordinates (with values between 1 and n, inclusive) on an $n\times n$ grid. If the 4 endpoints are collinear, then they correspond to an L_4-configuration while if they are not collinear, they correspond to a P_4-configuration.

Focusing on counting the L_4-configurations, let (x_1, y_1) and (x_2, y_2) be two points on the $n \times n$ grid with $x_1 < x_2$ and $y_1 < y_2$ and let M be the midpoint of the segment connecting these two points. Then the slope of the segment connecting these two points is positive and the number of additional grid points on this line segment is $\gcd(x_2 - x_1, y_2 - y_1) - 1$. If $\gcd(x_2 - x_1, y_2 - y_1)$ is even, then M is one of these grid points. These additional grid points (excluding M) can be paired together to form line segments with midpoint M as well. Thus, the number of shorter line segments having midpoint M is $\left\lfloor \frac{\gcd(x_2-x_1,y_2-y_1)-1}{2} \right\rfloor$. Summing this expression over all points (x_1, y_1) and (x_2, y_2) with $x_1 < x_2$ and $y_1 < y_2$ counts the number of L_4-configurations on lines with positive slopes. Due to symmetry, there is the same number of L_4-configurations on lines with negative slopes. Thus,

$$l_4(n) = 2\sum_{x_1=1}^{n-1}\sum_{x_2=x_1+1}^{n}\sum_{y_1=1}^{n-1}\sum_{y_2=y_1+1}^{n}\left\lfloor\frac{\gcd(x_2-x_1,y_2-y_1)-1}{2}\right\rfloor$$

Let $a = x_2 - x_1$ and $b = y_2 - y_1$. Then

$$\begin{aligned} l_4(n) &= 2\sum_{x_1=1}^{n-1}\sum_{a=1}^{n-x_1}\sum_{y_1=1}^{n-1}\sum_{b=1}^{n-y_1}\left\lfloor\frac{\gcd(a,b)-1}{2}\right\rfloor \\ &= 2\sum_{a=1}^{n-1}\sum_{x_1=1}^{n-a}\sum_{b=1}^{n-1}\sum_{y_1=1}^{n-b}\left\lfloor\frac{\gcd(a,b)-1}{2}\right\rfloor \\ &= 2\sum_{a=1}^{n-1}\sum_{b=1}^{n-1}(n-a)(n-b)\left\lfloor\frac{\gcd(a,b)-1}{2}\right\rfloor \end{aligned}$$

Turning our attention to proving Theorem 2, we need to count pairs of line segments with the same midpoint and with each endpoint having distinct x- and y-coordinates on the $n\times n$ grid. Note the set of all midpoints of line segments having endpoints on this grid is $\{(\frac{x}{2},\frac{y}{2})|3\leqslant x\leqslant 2n-1, 3\leqslant y\leqslant 2n-1\}$. Given any one of these midpoints $M=(\frac{x}{2},\frac{y}{2})$, any line segment with endpoints (x_1,y_1) and (x_2,y_2) with $x_1\neq x_2$ and midpoint M on the grid must have x_1 or x_2 less than $\frac{x}{2}$ and the other coordinate greater than $\frac{x}{2}$. Without loss of generality, assume $x_1<\frac{x}{2}<x_2$. Note the number of positive integers less than $\frac{x}{2}$ is $\lfloor\frac{x-1}{2}\rfloor$ and the number of positive integers greater than $\frac{x}{2}$ and less than or equal to n is $n-\lfloor\frac{x}{2}\rfloor$. Thus, there are

$$f_n(x) = \min\left\{\left\lfloor\frac{x-1}{2}\right\rfloor, n-\left\lfloor\frac{x}{2}\right\rfloor\right\}$$

possible values for x_1. Similarly, given x_1, the number of possible values for y_1 is $2f_n(y)$ as y_1 can be less than or greater than $\frac{y}{2}$. Given any one of these line segments, the number of ways to select a second line segment with the same midpoint and endpoints with x- and y-coordinates different than (x_1,y_1) and (x_2,y_2) is $(f_n(x)-1)(2f_n(y)-2)$ as we lose one possible value for the x-coordinate (since x_1 must be less than $\frac{x}{2}$) and two possible values for the y-coordinate. Thus, the number of pairs of line segments with midpoint $M=(\frac{x}{2},\frac{y}{2})$ and with each endpoint having distinct x- and y-coordinates on the grid equals

$$\frac{1}{2}(f_n(x)\cdot 2f_n(y))(f_n(x)-1)(2f_n(y)-2)$$

where we have multiplied by $\frac{1}{2}$ for otherwise we would count each pair of segments twice - once for each order the segments are selected. We can rewrite the product with binomial coefficients as

$$\frac{1}{2} \cdot 2\binom{f_n(x)}{2} \cdot 8 \cdot \binom{f_n(y)}{2}.$$

Thus, summing over all possible midpoints M (see above), we have

$$\begin{aligned} l_4(n) + p_4(n) &= \sum_{x=3}^{2n-1} \sum_{y=3}^{2n-1} 8\binom{f_n(x)}{2}\binom{f_n(y)}{2} \\ &= 8\sum_{x=5}^{2n-3} \binom{f_n(x)}{2} \sum_{y=5}^{2n-3} \binom{f_n(y)}{2} \end{aligned}$$

The indices of the summations in the last expression change since there cannot be two line segments with distinct x- and y-coordinates for their endpoints if the x- or y-coordinate of the midpoint is in the set $\{\frac{3}{2}, 2, \frac{2n-1}{2}, n-1\}$.

If n is even, then

$$\begin{aligned} \sum_{y=5}^{2n-3} \binom{f_n(y)}{2} &= \binom{2}{2} + \binom{2}{2} + \binom{3}{2} + \binom{3}{2} + \cdots + \binom{\frac{n}{2}-1}{2} + \binom{\frac{n}{2}-1}{2} \\ &\quad + \binom{\frac{n}{2}}{2} + \binom{\frac{n}{2}-1}{2} + \binom{\frac{n}{2}-1}{2} + \cdots + \binom{3}{2} + \binom{3}{2} + \binom{2}{2} + \binom{2}{2} \\ &= 4 \cdot \left[\binom{2}{2} + \binom{3}{2} + \cdots + \binom{\frac{n}{2}-1}{2}\right] + \binom{\frac{n}{2}}{2} \\ &= 4\binom{\frac{n}{2}}{3} + \binom{\frac{n}{2}}{2} \end{aligned}$$

The final simplification follows from a binomial coefficient identity known as the hockey-stick identity, namely $\sum_{i=r}^{n} \binom{i}{r} = \binom{n+1}{r+1}$.

Repeating this for the summation involving x, we see that if n is even, then $l_4(n) + p_4(n) = 8 \cdot \left(4\binom{\frac{n}{2}}{3} + \binom{\frac{n}{2}}{2}\right)^2 = \frac{1}{72}n^2(n-2)^2(2n-5)^2$.

If n is odd, then

$$\begin{aligned} \sum_{y=5}^{2n-3} \binom{f_n(y)}{2} &= \binom{2}{2} + \binom{2}{2} + \binom{3}{2} + \binom{3}{2} + \cdots + \binom{\frac{n-3}{2}}{2} + \binom{\frac{n-3}{2}}{2} \\ &\quad + \binom{\frac{n-1}{2}}{2} + \binom{\frac{n-1}{2}}{2} + \binom{\frac{n-1}{2}}{2} \\ &\quad + \binom{\frac{n-3}{2}}{2} + \binom{\frac{n-3}{2}}{2} + \cdots + \binom{3}{2} + \binom{3}{2} + \binom{2}{2} + \binom{2}{2} \\ &= 4 \cdot \left[\binom{2}{2} + \binom{3}{2} + \cdots + \binom{\frac{n-1}{2}}{2}\right] - \binom{\frac{n-1}{2}}{2} \\ &= 4\binom{\frac{n+1}{2}}{3} - \binom{\frac{n-1}{2}}{2} \end{aligned}$$

Thus, if n is odd then

$$l_4(n) + p_4(n) = 8 \cdot \left(4\binom{\frac{n+1}{2}}{3} - \binom{\frac{n-1}{2}}{2}\right)^2 = \frac{1}{72}(n-1)^2(n-3)^2(2n-1)^2,$$

hence completing the proof of Theorem 2.

We created OEIS entries A234471 and A243643 for $l_4(n)$ and $p_4(n)$, respectively, for $1 \leqslant n \leqslant 1000$.

Direct computation reveals that

$$\frac{C(n)}{n!} < \left(1 - \frac{1}{n(n-1)(n-2)}\right)^{l_3(n)} \left(1 - \frac{1}{n(n-1)(n-2)(n-3)}\right)^{l_4(n)+p_4(n)}$$

for $4 \leqslant n \leqslant 29, n \neq 5$. The right-hand expression can be interpreted as the probability that none of the violating configurations of order n occurs if we make the assumption that the occurrences of violating configurations are mutually independent events. We have no proof that this inequality holds for all sufficiently large n. If it is valid for $n > 5$, the inequality can be used to quantify exponential decay. For $n > 5$, define $a(n) = n(n-1)(n-2)$ and $b(n) = a(n) \cdot (n-3)$. Rewrite the right-hand expression as

$$\left[\left(1 - \frac{1}{a(n)}\right)^{a(n)}\right]^{\frac{l_3(n)}{a(n)}} \left[\left(1 - \frac{1}{b(n)}\right)^{b(n)}\right]^{\frac{l_4(n)+p_4(n)}{b(n)}}.$$

Using the formulas for $l_3(n)$ and $l_4(n) + p_4(n)$ and taking the limit of this expression as $n \to \infty$ gives

Conjecture 3.

$$\frac{C(n)}{n!} = O\left(e^{-\frac{n^2}{18} - \frac{n}{8}}\right).$$

5 Configuration Enumeration for Identity Permutations

While a Costas array is a permutation matrix that contains no violating configurations, we may also ask what is the permutation matrix that contains the maximal number of violating configurations. Intuition and numerical evidence suggest that the identity matrix and the anti-diagonal matrix of 1's contain the maximal number of violating configurations, but how many such violating configurations do these matrices contain? Clearly they contain no P_4-configurations. Let $Idl_3(n)$ and $Idl_4(n)$ denote the number of L_3-configurations and L_4-configurations I_n contains. The results are

Theorem 4. *For $n \geqslant 3$,*

$$Idl_3(n) = \begin{cases} (\frac{n}{2})(\frac{n-2}{2}), & n \text{ even} \\ (\frac{n-1}{2})^2, & n \text{ odd.} \end{cases}$$

Theorem 5. *For $n \geqslant 4$,*

$$Idl_4(n) = \begin{cases} 4\binom{\frac{n}{2}}{3} + \binom{\frac{n}{2}}{2} = \frac{1}{24}n(n-2)(2n-5), & n \text{ even} \\ 4\binom{\frac{n+1}{2}}{3} - \binom{\frac{n-1}{2}}{2} = \frac{1}{24}(n-1)(n-3)(2n-1), & n \text{ odd.} \end{cases}$$

Focusing on Theorem 4, note that (a, a) and (b, b) are both endpoints of the same L_3-configuration if and only if $\frac{a+b}{2}$ is an integer. Thus, a and b must have the same parity. If n is even, the number of pairs of even (or pairs of odd) integers less than or equal to n is $\binom{\frac{n}{2}}{2}$. Thus, if n is even, $Idl_3(n) = \binom{\frac{n}{2}}{2} \cdot 2 = (\frac{n}{2})(\frac{n-2}{2})$. If n is odd the number of pairs of even integers less than or equal to n is $\binom{\frac{n-1}{2}}{2}$ while the number of pairs of odd integers less than or equal to n is $\binom{\frac{n+1}{2}}{2}$. Thus, if n is odd, $Idl_3(n) = \binom{\frac{n-1}{2}}{2} + \binom{\frac{n+1}{2}}{2} = (\frac{n-1}{2})^2$, thus completing the proof of Theorem 4.

Turning our attention to proving Theorem 5, we essentially repeat the technique that we used to prove Theorem 2 noting that our midpoints are of the form $M = (\frac{x}{2}, \frac{x}{2})$ and the y-coordinates are equal to the x-coordinates for any L_4-configuration on I_n. Thus, we are only free to choose 2 x-coordinates less than $\frac{x}{2}$ for any midpoint $M = (\frac{x}{2}, \frac{x}{2})$, which can be done in $\binom{f_n(x)}{2}$ ways. Summing over the potential midpoints M, the number of L_4-configurations I_n contains is

$$\begin{aligned} Idl_4(n) &= \sum_{x=5}^{2n-3} \binom{f_n(x)}{2} \\ &= \begin{cases} 4\binom{\frac{n}{2}}{3} + \binom{\frac{n}{2}}{2} = \frac{1}{24}n(n-2)(2n-5), & n \text{ even} \\ 4\binom{\frac{n+1}{2}}{3} - \binom{\frac{n-1}{2}}{2} = \frac{1}{24}(n-1)(n-3)(2n-1), & n \text{ odd} \end{cases} \end{aligned}$$

as derived in Section 4.

Sequences $Idl_3(n)$ and $Idl_4(n)$ correspond to OEIS entry A002620 and the $(n-4)^{th}$ term of OEIS entry A002623, respectively. We also noticed that the total number of violating configurations that I_n contains, $Idl_4(n) + Idl_3(n)$, is the $(n-3)^{rd}$ term of OEIS entry A002623 and thus also equals $Idl_4(n+1)$. While this can be verified algebraically, we provide a combinatorial proof instead.

Theorem 6. *For $n \geqslant 4$,*

$$Idl_4(n) + Idl_3(n) = Idl_4(n+1)$$

The number of L_4-configurations contained by I_{n+1} is the number of L_4-configurations contained by this matrix that do not include the point $(n+1, n+1)$ plus the number that do contain the point $(n+1, n+1)$. Those L_4-configurations that do not contain $(n+1, n+1)$ are precisely the L_4-configurations of I_n and hence there are $Idl_4(n)$ such configurations. For simplicity of notation, we represent the location of a 1 in the identity matrix simply by its corresponding integer, so $n+1$ will represent the point $(n+1, n+1)$. Given an L_4-configuration with points $a < b < c < n+1$, we have $n+1-c = b-a$. Then the points $c-b < n+1-b < n+1-a$ form an L_3-configuration in I_n. Conversely, given an L_3-configuration with points $x < y < z$ in the I_n, we have $z - y = y - x$. Then the points $n+1-z < n+1-y < n+1-y+x < n+1$ form an L_4-configuration in I_{n+1}. As these two mappings are inverses of each other, the number of L_4-configurations that contain the point $(n+1, n+1)$ is $Idl_3(n)$.

6 Density Revisited

Let the random variables X, Y and Z be the number of L_3-, L_4- and P_4-configurations in a permutation matrix and let $W = X + Y + Z$. Then the permutation matrix is a Costas array if and only if $W = 0$. Can we use the second moment method on W to improve Davies's result? To find out, we begin by calculating $\mathbb{E}(W)$. By the additivity of the expectation operator, it suffices to consider $\mathbb{E}(Y)$ and $\mathbb{E}(Z)$. Let $\mathcal{P}_n$ denote the set of $n \times n$ permutation matrices. By $L_4 \subset A$ we denote the fact that the particular L_4-configuration appears in a fixed permutation matrix A. For $n \geqslant 4$, we have

$$\begin{aligned}
\mathbb{E}(Y) &= \frac{1}{n!}\sum_{A\in\mathcal{P}_n} Y(A) \\
&= \frac{1}{n!}\sum_{A\in\mathcal{P}_n}\sum_{L_4\subset A} 1 \\
&= \frac{1}{n!}\sum_{L_4}\sum_{\{A\in\mathcal{P}_n \mid L_4\subset A\}} 1 \\
&= \frac{1}{n!}\sum_{L_4}(n-4)! \\
&= \frac{l_4(n)}{n(n-1)(n-2)(n-3)}
\end{aligned}$$

Similar reasoning shows that

$$\mathbb{E}(Z) = \frac{p_4(n)}{n(n-1)(n-2)(n-3)}$$

Combining these results with the expression for $\mathbb{E}(X)$ gives

$$\mathbb{E}(W) = \begin{cases} \frac{n(n-2)(4n^2-11n-2)}{72(n-1)(n-3)} & , n \text{ even} \\ \frac{(n-1)(4n-3)(n+1)}{72n} & , n \text{ odd} \end{cases}$$

which yields

$$\mathbb{E}(W) \sim \frac{n^2}{18}.$$

At this point we suspect that an improvement over Davies's result in [8] could be possible from applying the second moment method to W since $\mathbb{E}^2(W) = O(n^4)$ instead of $O(n^2)$. We will have to show that $\mathbb{E}(W^2) = O(n^{\alpha})$ for some $\alpha < 3$ to make an improvement to his power-law decay result. Bounding $\mathbb{E}(W^2)$ is made more complicated by the fact that W is a sum and thus we would be obliged to consider cross terms upon squaring W. The formula for $l_4(n)$ in Theorem 1 makes it clear that $l_4(n) < n^5$ and thus $\mathbb{E}(Y) = O(n)$. As $\mathbb{E}(X) = O(n)$ and $\mathbb{E}(W) = O(n^2)$, it follows that $\mathbb{E}(Z) = O(n^2)$ and thus $\mathbb{E}^2(Z) = O(n^4)$. Hence Z is the dominant term in W and $\mathbb{E}(Z) \sim \frac{n^2}{18}$. We consider applying the second moment method to Z:

$$\frac{C(n)}{n!} = P(W = 0) \leqslant P(Z = 0) \leqslant P(|Z - \mathbb{E}(Z)| \geqslant \mathbb{E}(Z)) \leqslant \frac{\mathbb{E}(Z^2)}{\mathbb{E}^2(Z)} - 1$$

We approach bounding the second moment of Z in the same way that Davies did within [8] when bounding the second moment of X:

$$\mathbb{E}(Z^2) = \frac{1}{n!} \sum_{A \in \mathcal{P}_n} (Z(A))^2 = \frac{1}{n!} \sum_{A \in \mathcal{P}_n} \left(\sum_{P_4 \subset A} 1 \right)^2 = \frac{1}{n!} \sum_{A \in \mathcal{P}_n} \sum_{(P_4^1 \subset A, P_4^2 \subset A)} 1$$

The inner sum in the last expression is over ordered pairs (P_4^1, P_4^2) of P_4-configurations appearing in the permutation matrix A. Interchanging the order of summation allows us to write

$$\mathbb{E}(Z^2) \leqslant \frac{1}{n!} \{m_4(n-4)! + m_3(n-5)! + m_2(n-6)! + m_1(n-7)! + m_0(n-8)!\}$$

where $m_i = m_i(n)$ is the number of ordered pairs of P_4-configurations with exactly i points in common. Thus, the utility of the second moment method applied to Z for improving density results depends on any formulas and upper bounds for the m_i. In finding upper bounds for $m_i(n)$, for $i < 4$, the ones of the second P_4-configuration that are not also ones of the first P_4-configuration must be in distinct rows and columns from the ones of the first configuration. Why is this? If we picked one of the remaining $4 - i$ points as being in the same row or column as a 1 for the first configuration, then since there is only one 1 in each row and column in the permutation matrix, it would have to be a one of the first configuration. Thus there would be at least $i + 1$ points in common.

Without extensive effort we can derive upper bounds for the $m_i(n)$. Clearly $m_4(n) = p_4(n)$. To bound $m_3(n)$, note given a P_4-configuration, there are 4 ways to select exactly 3 points from it to share with another P_4-configuration. Given these 3 points, at most 2 additional P_4-configurations are possible. Hence

$$m_3(n) \leqslant 4 \cdot 2 \cdot p_4(n)$$

We claim that

$$m_2(n) \leqslant 9p_4(n)(n-4)^2$$

To see this, first note that there are $\binom{4}{2} = 6$ ways to pick two common points after fixing the first P_4-configuration. There are at most $(n-4)^2$ valid ways to pick a third one in the permutation matrix. Given the third one, the fourth one in the second P_4-configuration is not uniquely determined, but rather there are 3 possibilities for the fourth one. Since we do not care about the order in which the last 2 points are selected, we divide by 2, avoiding double counting in the case that these last two points are both in the permutation matrix, and reducing the error in allowing the second configuration in the case that the the fourth point falls outside the permutation matrix.

Similar reasoning gives us the inequality

$$m_1(n) \leqslant \frac{4}{6}(3)p_4(n)(n-4)^2(n-5)^2.$$

Finally, it is clear that $m_0 \leqslant p_4(n)p_4(n-4)$. Combining these inequalities gives us

$$\frac{C(n)}{n!} < \frac{1 + \frac{8}{n-4} + \frac{9(n-4)}{(n-5)} + \frac{2(n-4)(n-5)}{n-6} + \frac{p_4(n-4)}{(n-4)(n-5)(n-6)(n-7)}}{\frac{p_4(n)}{n(n-1)(n-2)(n-3)}} - 1$$

Letting $n \to \infty$ gives

$$\frac{C(n)}{n!} < \frac{1 + \frac{8}{n-4} + 9 + 2n + \frac{(n-4)^2}{18}}{\frac{n^2}{18}} - 1$$

which reduces to

$$\frac{C(n)}{n!} < \frac{28}{n} + \frac{196}{n^2} + \frac{144}{n^2(n-4)}$$

As mentioned in Section 3, Davies (Theorem 4.2, [8]) showed $\frac{C(n)}{n!} = O(\frac{1}{n})$ as $n \to \infty$ by studying L_3-configurations in permutation matrices, thereby proving the strongest known result about the asymptotic rate of decay of the density $\frac{C(n)}{n!}$ of Costas arrays. Our result above, proven using P_4-configurations, confirms this rate of decay. Looking more closely at Davies's proof, he shows that

$$\frac{C(n)}{n!} \leqslant \frac{c_1(n)}{\mu} + \frac{9n^2 - 45n + 60}{(n-3)(n-4)(n-5)}$$

where $\mu = \mathbb{E}(X) \sim \frac{n}{8}$ is the mean number of L_3-configurations in a permutation matrix and $c_1(n) = 1 + \frac{4}{n-3} + \frac{9n^2}{4(n-3)(n-4)}$. Note that $c_1(n) \sim 13/4$ and thus

$$\frac{c_1(n)}{\mu} + \frac{9n^2 - 45n + 60}{(n-3)(n-4)(n-5)} \sim \frac{13/4}{n/8} + \frac{9n^2 - 45n + 60}{(n-3)(n-4)(n-5)} \sim \frac{26}{n} + \frac{9}{n} = \frac{35}{n}$$

while our upper bound above is asymptotically equal to $\frac{28}{n}$. Thus, our asymptotic upper bound above is a slight improvement over Davies's.

The presence of the term $\frac{28}{n}$ in our bound suggests that a tighter upper bound on m_1 or a different probabilistic inequality should be used. If the bound on the function $m_1(n)$ could be shown to be asymptotically at most 2/9 of what it currently is then it would be canceled in the limit by the linear term of the quadratic in the numerator: $(\frac{2}{9})2n - \frac{8}{18}n = 0$. The derivation of our upper bound on $m_1(n)$ is based on assumptions that do allow room for improvement. For instance, the selection of the second and third vertices on the second P_4-configuration can be such that the fourth point falls outside the grid. The function $m_1(n) = 0$ for $n < 7$ because two P_4-configurations would overlap in at least two rows or columns. We wrote a C++ program to compute the values of $m_1(n)$ for $7 \leqslant n \leqslant 18$. We found that $m_1(7) = 636, m_1(8) = 25888, m_1(9) = 401432, m_1(10) = 2788608, m_1(11) = 14302888, m_1(12) = 57333552, m_1(13) = 190675736, m_1(14) = 556245344$, $m_1(15) = 1457381068, m_1(16) = 3498250624, m_1(17) = 7799395304$, and that $m_1(18) = 16399389728$. The ratio

$$\frac{m_1(n)}{2p_4(n)(n-4)^2(n-5)^2}$$

is monotone increasing for $7 \leqslant n \leqslant 18$ and equals $0.2262\ldots > \frac{2}{9}$ for $n = 18$. This numerical evidence suggests that it may not be possible to establish quadratic decay for density from the second moment method applied to the random variable Z. However, it is not clear that the ratio must be monotone increasing for larger n.

References

[1] L. Barker, K. Drakakis, and S. Rickard. On the complexity of the verification of the Costas property. *Proc. IEEE*, 97(3):586–593, March 2009.

[2] J. K. Beard. CD-ROM images of Costas arrays from CISS-2006/2008/2010 papers. http://jameskbeard.com/jameskbeard/Files.html.

[3] B. Correll Jr. A closed form expression for the number of Costas arrays of arbitrary order. In *Proceedings of IEEE Asilomar Conference on Signals, Systems, and Computers*, pages 2218 – 2222, Monterey, CA, November 2007.

[4] B. Correll Jr. The density of Costas arrays and three-free permutations. In *Proceedings of the IEEE Statistical Signal Processing Workshop*, pages 492 – 495, Ann Arbor, MI, 2012.

[5] B. Correll Jr., C. N. Swanson, and R. W. Ho. Costas arrays and the Lovász local lemma. In *Proceedings of 2015 IEEE International Radar Conference*, Washington, DC, May 2015.

[6] J. P. Costas. Medium constraints on sonar design and performance. Technical Report Class 1 Rep. R65EMH33, GE Co., 1965.

[7] J. P. Costas. A study of detection waveforms having nearly ideal range-Doppler ambiguity properties. *Proc. IEEE*, 72(8):996–1009, August 1984.

[8] K. Drakakis. A review of Costas arrays. *J. Appl. Math.*, 2006:1–32, 2006. Article ID 26385.

[9] K. Drakakis. On the degrees of freedom of Costas permutations and other constraints. In *Adv. Math. Commun.*, 5(3):435-448, 2011.

[10] K. Drakakis. Open problems in Costas arrays. Preprint, 2011. arXiv:1102.5727.

[11] K. Drakakis, F. Iorio, and S. Rickard. The enumeration of Costas arrays of order 28 and its consequences. In *Adv. Math. Commun.*, 5(1): 69-86, 2011.

[12] K. Drakakis, F. Iorio, S. Rickard, and J. Walsh. Results of the enumeration of Costas arrays of order 29. *Adv. Math. Commun.*, 5(3):547–553, 2011.

[13] K. Drakakis et al. Results of the enumeration of Costas arrays of order 27. *IEEE Trans. Inf. Theory*, 54(10):4684–4687, October 2008.

[14] E. N. Gilbert. Latin squares which contain no repeated digrams. *SIAM Review*, 7(2):189–198, April 1965.

[15] S. Golomb. Algebraic constructions for Costas arrays. *J. Comb. Theory Series A*, 37(1):13–21, 1984.

[16] S. Golomb and H. Taylor. Constructions and properties of Costas arrays. *Proc. IEEE*, 72(9):1143–1163, September 1984.

[17] M. Soltanalian, P. Stoica, and J. Li. A recursive method for enumeration of Costas arrays. In *Proceedings of 22nd European Signal Processing Conference*, Lisbon, Portugal, 2014.