symmetry

MDPI

Article

Symmetry and Attention Dynamics in Ducci-Generated Jacobsthal Circulant Matrices

Bahar Kuloğlu [1,†], Taras Goy [2,†] and Engin Özkan [3,*,†]

[1] Department of Engineering Basic Sciences, Sivas University of Science and Technology, 58000 Sivas, Türkiye; bkuloglu@sivas.edu.tr
[2] Faculty of Mathematics and Computer Science, Vasyl Stefanyk Carpathian National University, 76018 Ivano-Frankivsk, Ukraine; taras.goy@pnu.edu.ua
[3] Department of Mathematics, Marmara University, 34722 Istanbul, Türkiye
* Correspondence: engin.ozkan@marmara.edu.tr
† These authors contributed equally to this work.

Abstract

A Ducci sequence generated by the vector $A = (a_1, a_2, \ldots, a_n) \in \mathbb{Z}^n$ is defined by $(A, D_A, D_A^2, D_A^3, \ldots)$, where the Ducci map $D : \mathbb{Z}^n \to \mathbb{Z}^n$ is given by $D_A = (|a_2 - a_1|, |a_3 - a_2|, \ldots, |a_n - a_{n-1}|, |a_1 - a_n|)$. In this paper, we examine the impact of iterative Ducci transformations on Jacobsthal numbers and construct circulant and skew-circulant matrices generated by the resulting sequences. Their properties are investigated through matrix norms (Euclidean (Frobenius), spectral, and ℓ_p), determinants, and eigenvalues. To extend the classical analysis, we incorporate the Convolutional Block Attention Module (CBAM) from deep learning and interpret the structured matrices as simulated image inputs. By analyzing channel-attention vectors and their variances, we assess how successive Ducci transformations influence attention distribution. The first-order transformation produces greater variance in attention weights, indicating enhanced feature discrimination, whereas higher-order transformations promote a more balanced distribution. The results highlight how Ducci transformations influence attention variance in structured matrices.

Keywords: Ducci map; Ducci sequence; Jacobsthal sequence; circulant matrix; skew circulant matrix; Convolutional Block Attention Module; channel attention

MSC: 15B05; 11B83; 15A18; 68T07

1. Introduction and Motivation

Ducci sequences are generated by taking an initial sequence of integers and repeatedly forming a new sequence consisting of the absolute differences of consecutive terms arranged cyclically. More precisely, given an initial vector $A = (a_1, a_2, \ldots, a_n) \in \mathbb{Z}^n$, the associated Ducci sequence is defined as $(A, D_A, D_A^2, D_A^3, \ldots)$, where the Ducci map $D : \mathbb{Z}^n \to \mathbb{Z}^n$ is defined by

$$D_A := (|a_2 - a_1|, |a_3 - a_2|, \ldots, |a_n - a_{n-1}|, |a_n - a_1|),$$

and $D_A^k = D_{D_A^{k-1}}$ for a positive integer k.

For example, starting with a four-tuple $A = (1, 5, 9, 13)$, we obtain

Check for updates

Academic Editor: Qing-Wen Wang

Received: 15 February 2026
Revised: 9 March 2026
Accepted: 16 March 2026
Published: 18 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

$$D_A = (|1-5|, |5-9|, |9-13|, |13-1|) = (4,4,4,12),$$
$$D_A^2 = D(4,4,4,12) = (|4-4|, |4-4|, |4-12|, |12-4|) = (0,0,8,8),$$
$$D_A^3 = D(0,0,8,8) = (|0-0|, |0-8|, |8-8|, |8-0|) = (0,8,0,8).$$

Such sequences were first introduced by Enrico Ducci in 1937. Over the years, these sequences have been the subject of extensive research, with most studies focusing on their limiting behavior when starting from an initial vector $A = (a_1, a_2, \ldots, a_n)$; see [1–6].

A fundamental result states that a Ducci sequence converges to the zero vector if and only if n is a power of two [2,3]; this phenomenon is known as vanishing. If n is not a power of 2, the sequence instead enters a periodic cycle [7] and stabilizes at a positive constant $k(x_1, x_2, \ldots, x_n)$, where $x_i \in \{0; 1\}$ [3,5]. Because binary vectors are mapped to binary vectors under the Ducci transformation, many structural properties can be studied within the framework of $\mathbb{Z}_2^n$; see [2,5].

Furthermore, Ducci matrix sequences have been explored as a means of representing real numbers, and recent research has increasingly focused on studying their properties [8,9]. As highlighted in [8], despite a history spanning nearly 75 years, the study of the Ducci map remains an active and evolving area of mathematical inquiry. For a more detailed historical perspective on Ducci sequences, we refer the reader to Brockman and Zerr [3]. Conway and Guy [10] investigated number sequences and combinatorial properties related to Ducci sequences, while Allouche and Shallit [11] examined their chaotic behavior and recurrence properties. Furthermore, Chamberland and Thomas [12] introduced the Ducci map acting on vector spaces $\mathbb{Z}_2^n$ and $\mathbb{R}_n$, posing several open questions regarding its transient and cyclic behavior. In [13], Clausing explores a generalization of the Ducci sequence using matrix theory.

Ducci sequences have been studied under modular arithmetic and within other number systems to explore their broader mathematical properties. They have been applied in both theoretical and applied mathematics. The interplay between their periodic structure and chaotic behavior makes them particularly relevant for modeling cryptographic functions, as well as for analyzing both chaotic and periodic number sequences. Ducci sequences also play a role in studying the evolution of numerical structures in discrete mathematics. They continue to be an active area of research, particularly in the fields of chaos theory, modular analysis, and cryptography.

Circulant matrices form another well-studied class of structured objects. Given an n-tuple of complex or real numbers $S = (s_0, s_1, \ldots, s_{n-1})$, the matrix

$$\mathrm{Circ}(S) = \begin{bmatrix} s_0 & s_1 & \cdots & s_{n-2} & s_{n-1} \\ s_{n-1} & s_0 & \cdots & s_{n-3} & s_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_1 & s_2 & \cdots & s_n & s_0 \end{bmatrix}$$

is called a circulant matrix. Circulant matrices are closely connected to a wide range of problems in physics, image processing, probability theory, numerical analysis, and number theory [14]. Their versatility has led to extensive study across various applications. For instance, Kuloğlu et al. [15] investigated r-circulant matrices associated with k-Fermat and k-Mersenne numbers. Andrade et al. [16] studied circulant-like matrices involving Horadam, Fibonacci, Jacobsthal, and Pell numbers. Shannon et al. [17] analyzed Rhaly terraced sequences, their generalizations, properties, and applications. Finally, Solak and Bahşi [18] explored properties of circulant matrices with Ducci sequences. By expressing

the sequences obtained through the Ducci transformation in the form of circulant matrices, we derive the following matrices:

$$\mathrm{Circ}(A) = \begin{bmatrix} 1 & 5 & 9 & 13 \\ 13 & 1 & 5 & 9 \\ 9 & 13 & 1 & 5 \\ 5 & 9 & 13 & 1 \end{bmatrix}, \quad \mathrm{Circ}(D_A) = \begin{bmatrix} 4 & 4 & 4 & 12 \\ 12 & 4 & 4 & 4 \\ 4 & 12 & 4 & 4 \\ 4 & 4 & 12 & 4 \end{bmatrix},$$

$$\mathrm{Circ}(D_A^2) = \begin{bmatrix} 0 & 0 & 8 & 8 \\ 8 & 0 & 0 & 8 \\ 8 & 8 & 0 & 0 \\ 0 & 8 & 8 & 0 \end{bmatrix}.$$

Another important line of research involves incorporating number sequences into matrix structures. In these studies, circulant matrices have been the primary focus, with many of their fundamental properties being thoroughly investigated. For example, Kızılateş and Tuglu [19] examined norms of geometric and symmetric geometric circulant matrices associated with Tribonacci numbers. Kuloğlu et al. [20] studied the properties of r-circulant matrices involving generalized Fermat numbers. Kumari et al. [21] investigated norms and eigenvalues of r-circulant matrices with k-Mersenne and k-Mersenne–Lucas numbers. Finally, Özkan and Akkuş [22] introduced a new approach to k-Oresme and k-Oresme–Lucas sequences. In [23], Akbıyık et al. introduced a symmetric matrix family constructed using harmonic Pell numbers, along with its associated Hadamard exponential matrix. They examined various linear algebraic properties and inequalities using matrix norms. Shi and Kızılateş [24] proposed a new generalization of the Frank matrix and studied several of its properties. Durna and Yıldırım [25] defined a generalized terraced matrix based on the generalized Cesàro matrix and explored its mathematical characteristics. In [26,27], Goy and Shattuck evaluated the determinants and permanents of several families of Hessenberg and Toeplitz–Hessenberg matrices whose nonzero entries correspond to various subsequences of the Jacobsthal sequence.

Recently, attention-based techniques have transformed numerous areas of artificial intelligence (AI) by enabling models to dynamically focus on the most informative elements of the data. A prominent mechanism in this domain is the Convolutional Block Attention Module (CBAM), which has demonstrated notable success in enhancing convolutional neural networks through the sequential application of channel and spatial attention [28]. Recent developments have shown that attention mechanisms extend beyond visual and sequential inputs into structured data domains. For example, Graph Attention Networks (GATs) leverage attention over graph-structured inputs to weight feature contributions adaptively [29]. Moreover, comprehensive surveys of attention models applied to structural data confirm that attention mechanisms can effectively highlight relevant structural patterns across non-visual domains [30]. The growing influence of attention mechanisms in deep learning originally popularized by Vaswani et al. [31] through the transformer model has led to increasingly effective methods for extracting and prioritizing meaningful features. Building on this momentum, our work applies CBAM to structurally defined circulant matrices generated from Jacobsthal and Ducci sequences. These matrices exhibit inherent algebraic properties that make them well-suited for attention-based modeling.

This paper extends the conceptual framework for constructing circulant matrices using Ducci sequence and Jacobsthal numbers by incorporating attention-based pooling techniques. Specifically, we investigate the implications of focused attention on structured numerical patterns and their interpretability. Our objectives are as follows:

1. To examine how attention mechanisms can be adapted to matrices possessing algebraic structure;
2. To evaluate CBAM-based weighting and pooling strategies on circulant matrices derived from Ducci and Jacobsthal sequences;
3. To compare their structural and feature-preserving characteristics;
4. To identify which structured matrices are most compatible with attention-based transformations.

Traditionally, CBAM has been employed in image processing tasks, where attention is computed over pixel-based feature maps. In this study, we adopt a novel perspective by simulating image-like structures using skew circulant matrices formed from Jacobsthal and Ducci sequences. Within this framework, matrix elements are treated as structured analogs of pixel intensities, allowing us to investigate how CBAM distributes attention across channels in a mathematically interpretable setting. This approach provides a bridge between structured matrix theory and attention-based deep learning, facilitating a deeper understanding of feature selection dynamics in controlled environments.

CBAM consists of two sequential sub-modules, the Channel Attention Module (CAM) and the Spatial Attention Module (SAM), which together enable the network to adaptively focus on informative features [28]. CAM emphasizes inter-channel relationships by applying both average pooling and max pooling across spatial dimensions to capture complementary global statistics: average pooling reflects the overall presence of features, while max pooling highlights the most salient ones. These pooled representations are then passed through a shared multi-layer perceptron (MLP) with fully connected layers to model non-linear dependencies. The output is processed by a sigmoid activation function to produce a channel attention map, which enhances relevant channels and suppresses less informative ones.

Although CBAM is predominantly used in visual domains, its principles can be extended to structured numerical data such as matrices. Skew circulant matrices, in particular, offer rich algebraic characteristics due to their cyclic and sign-alternating structure. By embedding sequences like Ducci arrays and Jacobsthal numbers into these matrices, we construct highly organized numerical structures that stand to benefit from attention-driven weighting and pooling.

In this study, the use of CBAM (Convolutional Block Attention Module) is motivated by its ability to perform channel-wise feature weighting, which provides a fine-grained assessment of structural relevance in matrix representations. Unlike simpler pooling methods or standard spectral analysis, CBAM captures not only global norms or eigenvalue distributions, but also the relative importance of individual channels, highlighting which rows/columns (or combinations thereof) carry the most discriminative information in number-theoretic sequences. This is particularly relevant for structured matrices derived from sequences such as Jacobsthal, Fibonacci, or Lucas numbers, where certain elements inherently dominate the structural patterns.

Regarding symmetry, circulant and skew-circulant matrices possess inherent algebraic symmetries rotational or sign-shifted repetitions which are preserved under Ducci transformations. CBAM interacts with these symmetries by assigning attention weights in a manner sensitive to repeated or structured patterns. For instance, symmetric or quasi-symmetric blocks often receive correlated attention scores, reflecting their structural equivalence. Thus, the attention dynamics reveal how algebraic symmetries influence feature importance and channel-wise discriminative power, providing a quantitative bridge between matrix-theoretic properties and learned attention distributions.

Jacobsthal numbers serve as a fundamental component of our analysis. Recall that the Jacobsthal numbers are defined recursively by $J_n = J_{n-1} + 2J_{n-2}$ for $n \geq 2$, with initial conditions $J_0 = 0$ and $J_1 = 1$, and admit the closed form

$$J_n = \frac{2^n - (-1)^n}{3}, \qquad n \geq 0. \tag{1}$$

Jacobsthal numbers possess numerous combinatorial and algebraic properties and appear in number theory, graph theory, and discrete mathematics. See sequence A001045 from the On-Line Encyclopedia of Integer Sequences [32] for further information on this sequence.

We will use the following basic properties of Jacobsthal numbers, derived from (1):

$$J_{n+1} + J_n = 2^n, \tag{2}$$

$$J_{n+1}^2 + J_n^2 = 4^n - 2J_{n+1}J_n, \tag{3}$$

$$J_{n+1}^2 - J_n^2 = 2^{n+1}J_{n-1}, \tag{4}$$

$$\sum_{k=1}^{n} J_k = \frac{J_{n+2} - 1}{2}, \tag{5}$$

$$\sum_{k=1}^{n} J_k^2 = \frac{1}{9}\left(J_{2n+2} - (-1)^n J_{n+2} + n\right). \tag{6}$$

Consider the vector $J = (J_1, J_2, \ldots, J_n)$. Then, consecutive images under the Ducci map for the Jacobsthal sequence are defined as follows:

$$\begin{aligned} D_J &= (J_2 - J_1, J_3 - J_2, \ldots, J_n - J_{n-1}, J_n - J_1) = (2J_0, 2J_1, \ldots, 2J_{n-2}, J_n - 1), \\ D_J^2 &= \left(2(J_1 - J_0), 2(J_2 - J_1), \ldots, 2(J_{n-2} - J_{n-3}), J_n - 1 - 2J_{n-2}, J_n - 1 - 2J_0\right) \\ &= (4J_{-1}, 4J_0, \ldots, 4J_{n-4}, J_{n-1} - 1, J_n - 1), \end{aligned}$$

where $J_{-1} = 1/2$ as defined in (1). Thus, the above starting vector and its consecutive images under the Ducci map give rise to the following $n \times n$ matrices:

$$\mathrm{Circ}(J) = \begin{bmatrix} J_1 & J_2 & \cdots & J_{n-1} & J_n \\ J_n & J_1 & \cdots & J_{n-2} & J_{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ J_2 & J_3 & \cdots & J_n & J_1 \end{bmatrix}, \tag{7}$$

$$\mathrm{Circ}(D_J) = \begin{bmatrix} 2J_0 & 2J_1 & \cdots & 2J_{n-2} & J_n - 1 \\ J_n - 1 & 2J_0 & \cdots & 2J_{n-3} & 2J_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2J_1 & 2J_2 & \cdots & J_n - 1 & 2J_0 \end{bmatrix}, \tag{8}$$

$$\mathrm{Circ}(D_J^2) = \begin{bmatrix} 4J_{-1} & 4J_0 & \cdots & 4J_{n-4} & J_{n-1} - 1 & J_n - 1 \\ J_n - 1 & 4J_{-1} & \cdots & 4J_{n-5} & 4J_{n-4} & J_{n-1} - 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 4J_0 & 4J_1 & \cdots & J_{n-1} - 1 & J_n - 1 & 4J_{-1} \end{bmatrix}. \tag{9}$$

2. Norms of Circulant Matrices with Ducci Sequences and Jacobsthal Numbers

In this section, we examine the matrices $\mathrm{Circ}(J)$, $\mathrm{Circ}(D_J)$, and $\mathrm{Circ}(D_J^2)$, and explore relationships among their norms, determinants, and eigenvalues.

The following definitions will be used throughout this section.

Let $X = [x_{ij}]$ be any $m \times n$ matrix with complex (or real) entries. The l_p norm of X is defined by

$$\|X\|_p = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |x_{ij}|^p \right)^{1/p}, \qquad p \geq 1.$$

In particular, for $p = 2$, this norm coincides with the Frobenius (Euclidean) norm, denoted by $\|X\|_F$.

Theorem 1. *The following identities hold:*

$$\|\mathrm{Circ}(J)\|_F^2 = \frac{n}{27}\left(4^{n+1} - (-1)^n 2^{n+2} + 3n\right), \tag{10}$$

$$\|\mathrm{Circ}(D_J)\|_F^2 = \frac{n}{27}\left(2 \cdot 4^{n+1} - (9 + (-1)^n 5)\, 2^{n+1} + 18(-1)^n + 6 + 12n\right), \tag{11}$$

$$\left\|\mathrm{Circ}(D_J^2)\right\|_F^2 = \frac{n}{27}\left(4^{n+1} - ((-1)^n 7 + 27)\, 2^n - 24 + 48n\right). \tag{12}$$

Proof. From the definition of the Frobenius norm and Equations (7) and (8), together with identities (1) and (6), we obtain

$$\begin{aligned}
\|\mathrm{Circ}(J)\|_F^2 &= n \sum_{k=1}^{n} J_k^2 = \frac{n}{9}\left(J_{2n+2} - (-1)^n J_{n+2} + n\right) = \frac{n}{27}\left(4^{n+1} - (-1)^n 2^{n+2} + 3n\right), \\
\|\mathrm{Circ}(D_J)\|_F^2 &= 4n \sum_{k=0}^{n-2} J_k^2 + n(J_n - 1)^2 = \frac{n}{9}\left(4J_{2n-2} + 9J_n^2 - 2\left((-1)^n 2 + 9\right) J_n + 1 + 4n\right) \\
&= \frac{2n}{27}\left(2 \cdot 4^n - \left(9 + (-1)^n 5\right) 2^n + 9(-1)^n + 3 + 6n\right).
\end{aligned}$$

For the Frobenius norm $\left\|\mathrm{Circ}(D_J^2)\right\|_F$, using Equations (2), (3) and (6), we have

$$\begin{aligned}
\left\|\mathrm{Circ}(D_J^2)\right\|_F^2 &= 16n \sum_{k=0}^{n-3} J_{k-1}^2 + n(J_{n-1} - 1)^2 + n(J_n - 1)^2 \\
&= 4n + 16n \sum_{k=1}^{n-4} J_k^2 + n(J_{n-1} - 1)^2 + n(J_n - 1)^2 \\
&= \frac{16n}{9}\left(J_{2n-6} - (-1)^n J_{n-2} + n - 4\right) + n(J_{n-1}^2 + J_n^2) - 2n(J_{n-1} + J_n) + 6n \\
&= \frac{16n}{9}\left(J_{2n-6} - (-1)^n J_{n-2} + n\right) + n\left(4^{n-1} - 2J_{n-1}J_n\right) - n2^n + 6n,
\end{aligned}$$

and applying (1) yields (12). □

From (10)–(12), we deduce the following relationships among the Frobenius norms $\|\,\mathrm{Circ}(J)\|_F$, $\|\,\mathrm{Circ}(D_J)\|_F$, and $\|\,\mathrm{Circ}(D_J^2)\|_F$.

Corollary 1. *We have*

$$\begin{aligned}
\|\mathrm{Circ}(D_J)\|_F^2 - \left\|\mathrm{Circ}(D_J^2)\right\|_F^2 &= \frac{n}{9}\left(2^n\left(3 - (-1)^n\right) + 6(-1)^n - 12n + 10\right), \\
\|\mathrm{Circ}(J)\|_F^2 - \|\mathrm{Circ}(D_J)\|_F^2 &= \frac{n}{9}\left(2^{n+1}\left(3 + (-1)^n\right) - 6(-1)^n - 3n - 2\right), \\
\|\mathrm{Circ}(J)\|_F^2 - \left\|\mathrm{Circ}(D_J^2)\right\|_F^2 &= \frac{n}{9}\left(2^n\left(9 + (-1)^n\right) - 15n + 8\right).
\end{aligned}$$

The following theorem concerns the ℓ_p norms of the $n \times n$ matrices $\mathrm{Circ}(J)$, $\mathrm{Circ}(D_J)$, and $\mathrm{Circ}(D_J^2)$.

Theorem 2. *For $n \geq 3$, we have*

$$\|\mathrm{Circ}(J)\|_p^p = \frac{n}{3^p}\sum_{k=1}^{n}\sum_{j=1}^{p}(-1)^{(k-1)(p-j)}\binom{p}{j}2^{kj},$$

$$\left\|\mathrm{Circ}(D_J)\right\|_p^p = \frac{n}{3^p}\sum_{j=0}^{p}(-1)^{p-j}\binom{p}{j}$$

$$\times\left(2^p\sum_{k=1}^{n-2}(-1)^{k(p-j)}2^{kj} + (-1)^{(n-1)j}3^{p-j}\sum_{s=0}^{j}(-1)^{(n-1)s}\binom{j}{s}2^{ns}\right),$$

$$\left\|\mathrm{Circ}(D_J^2)\right\|_p^p = n2^p + \frac{n}{3^p}\sum_{j=0}^{p}(-1)^{p-j}\binom{p}{j}$$

$$\times\left(4^p\sum_{k=1}^{n-4}(-1)^{k(p-j)}2^{kj} + (-1)^{(n-1)j}3^{p-j}\sum_{s=0}^{j}(-1)^{(n-1)s}\binom{j}{s}2^{(n-1)s}\left(2^s+(-1)^{j-s}\right)\right).$$

Proof. By definition of l_p norm, we have $\|\mathrm{Circ}(J)\|_p^p = n\sum_{k=1}^{n}J_k^p$. Using (1) and the binomial formula yields

$$\|\mathrm{Circ}(J)\|_p^p = n\sum_{k=1}^{n}\left(\frac{2^k-(-1)^k}{3}\right)^p = \frac{n}{3^p}\sum_{k=1}^{n}\sum_{j=1}^{p}(-1)^{(k-1)(p-j)}\binom{p}{j}2^{kj}.$$

Similarly,

$$\begin{aligned}\left\|\mathrm{Circ}(D_J)\right\|_p^p &= n\left(2^p\sum_{k=1}^{n-2}J_k^p + (J_n-1)^p\right)\\ &= \frac{n}{3^p}\sum_{j=0}^{p}(-1)^{p-j}\binom{p}{j}\left(\sum_{k=1}^{n-2}(-1)^{k(p-j)}2^{kj+p} + 3^{p-j}\sum_{s=0}^{j}(-1)^{(n-1)(j-s)}\binom{j}{s}2^{ns}\right)\end{aligned}$$

and

$$\begin{aligned}\left\|\mathrm{Circ}(D_J^2)\right\|_p^p &= n\left(2^p + 4^p\sum_{k=1}^{n-4}J_k^p + (J_{n-1}-1)^p + (J_n-1)^p\right) = n2^p + \frac{n}{3^p}\sum_{j=0}^{p}(-1)^{p-j}\binom{p}{j}\\ &\times\left(4^p\sum_{k=1}^{n-4}(-1)^{k(p-j)}2^{kj} + 3^{p-j}\sum_{s=0}^{j}(-1)^{(n-1)(j-s)}\binom{j}{s}2^{(n-1)s}\left(2^s+(-1)^{j-s}\right)\right).\end{aligned}$$

□

As immediate consequences of Theorem 2, we obtain the following relations among the corresponding ℓ_p norms.

Corollary 2. *For $n \geq 3$, we have*

$$\left\|\mathrm{Circ}(D_J)\right\|_p^p - 2^p\|\mathrm{Circ}(J)\|_p^p = n\left((J_n-1)^p - 2^p(J_n^p + J_{n-1}^p)\right),$$

$$\begin{aligned}&\left\|\mathrm{Circ}(D_J^2)\right\|_p^p - 4^p\|\mathrm{Circ}(J)\|_p^p\\ &\qquad = n\left(2^p + (J_n-1)^p + (J_{n-1}-1)^p - 4^p(J_n^p + J_{n-1}^p + J_{n-2}^p + J_{n-3}^p)\right),\end{aligned}$$

$$\begin{aligned}&\left\|\mathrm{Circ}(D_J^2)\right\|_p^p - 2^p\|\mathrm{Circ}(D_J)\|_p^p\\ &\qquad = n\left(2^p + (1-2^p)(J_n-1)^p - 4^p(J_{n-2}^p + J_{n-3}^p) + (J_{n-1}-1)^p\right).\end{aligned}$$

Lemma 1 ([14]). *The eigenvalues of the circulant matrix* $\mathrm{Circ}(s_0, s_1, \ldots, s_{n-1})$ *are given by*

$$\lambda_j = \sum_{k=0}^{n-1} s_k w_n^{-jk}, \quad j = 0, 1, \ldots, n-1,$$

where $w_n = e^{2\pi i/n}$ *with* $i^2 = -1$.

Theorem 3. *Let* α_k, β_k, *and* γ_k *denote the eigenvalues of the* $n \times n$ *matrices* $\mathrm{Circ}(J)$, $\mathrm{Circ}(D_J)$, *and* $\mathrm{Circ}(D_J^2)$, *respectively. Then, for* $m = 0, 1, \ldots, n-1$, *the following relations hold:*

$$\beta_m = 2\alpha_m w_n^{-m} - \left(1 + (-1)^n\right) w_n^m - 2J_n, \tag{13}$$

$$\gamma_m = 4\alpha_m w_n^{-2m} - \left(1 + (-1)^n\right) w_n^{2m} - \left(1 - (-1)^n\right) w_n^m - 4J_n w_n^{-m} - 4J_{n-1} + 2, \tag{14}$$

where $w_n = e^{2\pi i/n}$ *and* $i^2 = -1$.

Proof. From (7), (8), and Lemma 1, the eigenvalues of the matrices $\mathrm{Circ}(J)$ and $\mathrm{Circ}(D_J)$ are given by

$$\alpha_m = \sum_{k=0}^{n-1} J_{k+1} w_n^{-mk} = w_n^m \sum_{k=1}^{n} J_k w_n^{-mk}, \tag{15}$$

$$\beta_m = \sum_{k=0}^{n-2} 2J_k w_n^{-mk} + (J_n - 1) w_n^{-m(n-1)} = 2\sum_{k=1}^{n-2} J_k w_n^{-mk} + (J_n - 1) w_n^m, \tag{16}$$

from which it follows that $\beta_m = 2\alpha_m w_n^{-m} + (J_n - 2J_{n-1} - 1)\, w_n^m - 2J_n$, and hence, by applying (1), we obtain (13).

On the other hand, in view of (9), we have

$$\begin{aligned} \gamma_m &= \sum_{k=0}^{n-3} 4J_{k-1} w_n^{-mk} + (J_{n-1} - 1) w_n^{-m(n-2)} + (J_n - 1) w_n^{-m(n-1)} \\ &= 2 + 4w_n^{-m} \sum_{k=1}^{n-4} J_k w_n^{-mk} + (J_{n-1} - 1) w_n^{2m} + (J_n - 1) w_n^m. \end{aligned} \tag{17}$$

Comparing (17) with (15), we obtain

$$\gamma_m = 4\alpha_m w_n^{-2m} + (J_{n-1} - 4J_{n-3} - 1) w_n^{2m} + (J_n - 4J_{n-2} - 1) w_n^m - 4J_n w_n^{-m} - 4J_{n-1} + 2,$$

and hence, again using (1), we have (14). □

Formulas (13) and (14) demonstrate an important computational advantage: the eigenvalues of the matrix can be obtained directly from the generating sequence using discrete Fourier transform evaluations, avoiding explicit characteristic polynomial calculations or matrix diagonalization.

Remark 1. *The factors* $(1 + (-1)^n)$ *appearing in* (13) *and* (14) *reflect a parity effect inherent in the spectral structure of circulant matrices. When n is even, the set of nth roots of unity contains the real root* $w_n^{n/2} = -1$, *which contributes an additional term in the discrete Fourier representation of the Ducci image. In this case,* $(1 + (-1)^n) = 2$. *When n is odd, such a root does not exist, and the corresponding contribution vanishes, yielding* $(1 + (-1)^n) = 0$. *Thus, the parity dependence is a natural consequence of Fourier analysis on cyclic groups and does not indicate a different dynamical behavior of the Ducci transformation itself.*

Proposition 1. *Let* $\omega_n = e^{\frac{2\pi i}{n}}$ *and let* $C = \mathrm{SCirc}(J_0, J_1, \ldots, J_{n-1})$ *be the skew circulant matrix generated by the Jacobsthal numbers. Then the eigenvalues* α_m *of* C *admit the explicit form*

$$\alpha_m = \frac{1}{3}\left(\frac{1-2^n}{1-2\omega_n^{-m}} - \frac{1-(-1)^n}{1+\omega_n^{-m}}\right), \qquad m = 0, 1, \ldots, n-1.$$

Proof. The eigenvalues of a skew circulant matrix are given by

$$\alpha_m = \sum_{k=0}^{n-1} J_k\, \omega_n^{-mk}, \qquad m = 0, 1, \ldots, n-1,$$

where $\omega_n = e^{\frac{2\pi i}{n}}$. Using the Binet formula (1), we obtain

$$\alpha_m = \frac{1}{3}\left(\sum_{k=0}^{n-1} (2\omega_n^{-m})^k - \sum_{k=0}^{n-1} \left((-1)\omega_n^{-m}\right)^k\right).$$

Both summations are finite geometric series. Hence,

$$\sum_{k=0}^{n-1} (2\omega_n^{-m})^k = \frac{1-(2\omega_n^{-m})^n}{1-2\omega_n^{-m}},$$

and

$$\sum_{k=0}^{n-1} \left((-1)\omega_n^{-m}\right)^k = \frac{1-\left((-1)\omega_n^{-m}\right)^n}{1+\omega_n^{-m}}.$$

Since $(\omega_n)^n = 1$, we have $(2\omega_n^{-m})^n = 2^n$ and $\left((-1)\omega_n^{-m}\right)^n = (-1)^n$. Substituting these identities into the above expressions yields the desired closed form. □

3. Attention-Based Pooling Methods Applied to Circulant Matrices with Ducci Sequences and Jacobsthal Numbers

In this section, we explore how attention mechanisms inspired by CBAM can be adapted and applied to skew circulant matrices constructed from special integer sequences, such as Ducci sequences and Jacobsthal numbers. This integration aims to enhance our understanding of how attention-based pooling interacts with algebraic matrix structures, potentially yielding more expressive numerical representations and inspiring novel applications in structured data analysis.

Given an n-tuple of complex or real numbers $S = (s_0, s_1, \ldots, s_{n-1})$, the matrix

$$\mathrm{SCirc}(S) = \begin{bmatrix} s_0 & s_1 & \cdots & s_{n-2} & s_{n-1} \\ -s_{n-1} & s_0 & \cdots & s_{n-3} & s_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -s_2 & -s_3 & \cdots & s_0 & s_1 \\ -s_1 & -s_2 & \cdots & -s_{n-1} & s_0 \end{bmatrix}$$

is called a skew circulant matrix [33].

Let us consider the vector $J = (J_1, J_2, \ldots, J_n)$. Applying the Ducci map to this initial vector and to its successive images yields the following $n \times n$ skew circulant matrices:

$$\mathrm{SCirc}(J) = \begin{bmatrix} J_1 & J_2 & \cdots & J_{n-1} & J_n \\ -J_n & J_1 & \cdots & J_{n-2} & J_{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -J_2 & -J_3 & \cdots & -J_n & J_1 \end{bmatrix},$$

$$\mathrm{SCirc}(D_J) = \begin{bmatrix} 2J_0 & 2J_1 & \cdots & 2J_{n-2} & J_n - 1 \\ -(J_n - 1) & 2J_0 & \cdots & 2J_{n-3} & 2J_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -2J_1 & -2J_2 & \cdots & -(J_n - 1) & 2J_0 \end{bmatrix},$$

$$\mathrm{SCirc}(D_J^2) = \begin{bmatrix} 4J_{-1} & 4J_0 & \cdots & 4J_{n-4} & J_{n-1} - 1 & J_n - 1 \\ -(J_n - 1) & 4J_{-1} & \cdots & 4J_{n-5} & 4J_{n-4} & J_{n-1} - 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ -4J_0 & -4J_1 & \cdots & -(J_{n-1} - 1) & -(J_n - 1) & 4J_{-1} \end{bmatrix}.$$

To further investigate the impact of attention-based mechanisms on algebraically structured matrices, we focus on applying CBAM to skew circulant matrices constructed from Jacobsthal numbers. In particular, we consider fourth-order skew circulant matrices whose entries are derived from the Jacobsthal sequence. These matrices provide a compact yet informative framework for observing the effects of channel and spatial attention operations in a controlled numerical setting.

The following computations simulate the core stages of the CBAM pipeline, namely, global pooling, multilayer perceptron (MLP) transformation, sigmoid-based attention refinement, and final matrix reweighting applied directly to structured numerical matrices rather than raw image features. This numerical adaptation allows us to examine how attention mechanisms selectively emphasize or suppress certain regions of the matrix and how this behavior depends on the underlying number-theoretic structure.

To illustrate the structure and behavior of skew circulant matrices, we begin with the numerical case $n = 4$. We first apply CBAM operations to the matrix $\mathrm{SCirc}(J)$:

$$\mathrm{SCirc}(J) = \begin{bmatrix} J_1 & J_2 & J_3 & J_4 \\ -J_4 & J_1 & J_2 & J_3 \\ -J_3 & -J_4 & J_1 & J_2 \\ -J_2 & -J_3 & -J_4 & J_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 & 5 \\ -5 & 1 & 1 & 3 \\ -3 & -5 & 1 & 1 \\ -1 & -3 & -5 & 1 \end{bmatrix}, \tag{18}$$

followed by the same operations to $\mathrm{SCirc}(D_J)$, and finally to $\mathrm{SCirc}(D_J^2)$, all based on the same underlying numerical data.

With the skew circulant matrix constructed from Jacobsthal numbers defined, we proceed to apply the CBAM framework in order to analyze the interaction between attention-based mechanisms and the matrix's algebraic structure.

4. Channel Attention Module: Average and Max Pooling Followed by MLP and Sigmoid Activation

The attention-variance results are generally robust to standard initialization, scaling, or normalization procedures applied to the structured matrices. Before being input to CBAM, matrices are typically normalized to maintain comparable magnitude across channels, ensuring that the attention mechanism responds to structural differences rather than absolute values. Minor variations in initialization or scaling do not qualitatively alter the observed patterns: high-variance channels remain distinct, and low-variance channels remain relatively suppressed. Therefore, the discriminative effects highlighted by CBAM

primarily reflect the inherent structure of the Jacobsthal-based or skew-Ducci matrices rather than artifacts of preprocessing.

In the first stage of the CBAM framework, we apply channel attention to the constructed skew circulant matrix populated with Jacobsthal numbers. This mechanism aims to emphasize the most informative channels by computing attention weights from pooled statistics using a shared multilayer perceptron (MLP).

Let $\mathrm{SCirc}(J) \in \mathbb{R}^{4\times 4}$ denote the skew circulant matrix with Jacobsthal elements under the Ducci map. We compute two feature descriptors along the spatial dimension global average pooling and global max pooling defined by

$$F_{\text{avg}} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{SCirc}(J)_{i,j}, \quad F_{\max} = \max_{1\le i\le n} \mathrm{SCirc}(J)_{i,j}.$$

These operations produce two channel-wise vectors, $F_{\text{avg}}, F_{\max} \in \mathbb{R}^{4\times 1}$, summarizing spatial information. Both descriptors are passed through a shared MLP consisting of two fully connected layers. Let $W_1 \in \mathbb{R}^{4\times r}$ and $W_2 \in \mathbb{R}^{r\times 4}$ denote the weight matrices of the MLP, where the reduction ratio is set to $r = 2$. The MLP transformation is

$$\mathrm{MLP}(F) = W_2 \cdot \mathrm{ReLU}(W_1 \cdot F),$$

where $\mathrm{ReLU}(x) = \max(0, x)$ is the rectified linear unit activation function, commonly used in neural networks to introduce non-linearity.

The outputs corresponding to average and max pooling are summed and passed through a sigmoid activation, $\mathrm{SCirc}(J)_c \in \mathbb{R}^{1\times 4}$:

$$\mathrm{SCirc}(J)_c = \sigma\big(\mathrm{MLP}(F_{\text{avg}}) + \mathrm{MLP}(F_{\max})\big),$$

where $\sigma(x) = \frac{1}{1+e^{-x}}$ is the element-wise sigmoid function.

Finally, the original matrix $\mathrm{SCirc}(J)$ is reweighted by broadcasting the attention map across rows, $\mathrm{SCirc}(J)' = \mathrm{SCirc}(J) \odot \mathrm{SCirc}(J)_c$, where $\odot$ denotes element-wise multiplication with broadcasting along the channel axis. This completes the channel attention refinement step, preparing the matrix for subsequent spatial attention operations.

Let us now consider each row as a channel and compute the corresponding averages and maxima:

$$F_{\text{avg}} = \begin{bmatrix} \frac{1-5-3-1}{4} \\ \frac{1+1-5-3}{4} \\ \frac{3+1+1-5}{4} \\ \frac{5+3+1+1}{4} \end{bmatrix} = \begin{bmatrix} -2 \\ -1.5 \\ 0 \\ 2.5 \end{bmatrix}, \qquad F_{\max} = \begin{bmatrix} \max(1,-5,-3,-1) \\ \max(1,1,-5,-3) \\ \max(3,1,1,-5) \\ \max(5,3,1,1) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 5 \end{bmatrix}.$$

In the CBAM framework, channel attention begins with applying average pooling and max pooling to the input matrix along the spatial dimension. This step produces a single representative value for each channel (i.e., each column of the matrix).

These values (e.g., a 4-dimensional vector for 4-channel data) are passed through MLP consisting of two fully connected layers. The first layer reduces the channel dimensionality (performing information compression into an embedded space), while the second layer restores it to the original size. This means that in CBAM, we initially work with 4-channel input data. In the CBAM literature, dimensionality reduction within the MLP is typically controlled by a reduction ratio denoted by r (commonly set to 2 or 4). Accordingly, the MLP employs weight matrices $W_1 \in \mathbb{R}^{4\times 2}$ and $W_2 \in \mathbb{R}^{2\times 4}$, which reduce and then restore the dimensionality of the input features, respectively.

The channel attention mechanism begins by applying global average pooling and global max pooling along the spatial dimension of the matrix. For the given 4×1 Jacobsthal-based circulant matrix, these operations yield two 4×1 vectors, F_{avg} and $F_{\max}$, which summarize channel-wise information.

Let us denote the following:

$$W_1 = \begin{bmatrix} 0.5 & 0.1 & 0.4 & 0.6 \\ 0.2 & 0.3 & 0.1 & 0.5 \end{bmatrix}^T, \quad W_2 = \begin{bmatrix} 0.3 & 0.2 & 0.4 & 0.1 \\ 0.6 & 0.1 & 0.3 & 0.5 \end{bmatrix}.$$

Then we obtain

$$K_{\text{avg}} = W_1^T \cdot F_{\text{avg}} = \begin{bmatrix} 0.5 & 0.1 & 0.4 & 0.6 \\ 0.2 & 0.3 & 0.1 & 0.5 \end{bmatrix} \begin{bmatrix} -2 & -1.5 & 0 & 2.5 \end{bmatrix}^T = \begin{bmatrix} 0.35 \\ 0.4 \end{bmatrix},$$

$$K_{\max} = W_1^T \cdot F_{\max} = \begin{bmatrix} 0.5 & 0.1 & 0.4 & 0.6 \\ 0.2 & 0.3 & 0.1 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 & 5 \end{bmatrix}^T = \begin{bmatrix} 4.8 \\ 3.3 \end{bmatrix}.$$

Hence,

$$K = K_{\text{avg}} + K_{\max} = \begin{bmatrix} 5.15 \\ 3.7 \end{bmatrix}.$$

For sigmoid calculation, the first value is $x_1 = 5.15$: $\sigma(5.15) \approx 0.9942$, and the second value is $x_2 = 3.7$: $\sigma(3.7) \approx 0.9761$. The resulting attention weight is

$$\text{SCirc}(J)_c = \sigma(K) = \begin{bmatrix} 0.9942 \\ 0.9761 \end{bmatrix}.$$

These weights reflect the relative importance of each channel, with higher values indicating stronger attention. This vector contains only 2 values because the reduction ratio was $r = 2$, so the MLP output was 2-dimensional. However, since the original matrix had 4 channels (i.e., 4 columns), the final attention vector must be 4-dimensional. To restore the original dimensionality, we apply the transposed weight matrix W_2^T to K:

$$K_{\text{final}} = W_2^T \cdot K = \begin{bmatrix} 0.3 & 0.2 & 0.4 & 0.1 \\ 0.6 & 0.1 & 0.3 & 0.5 \end{bmatrix}^T \begin{bmatrix} 5.15 \\ 3.7 \end{bmatrix} = \begin{bmatrix} 3.765 & 1.4 & 3.17 & 2.365 \end{bmatrix}^T.$$

If we apply the sigmoid function to the resulting matrix, we obtain the following values: $\sigma(3.765) \approx 0.977357$, $\sigma(1.4) \approx 0.802184$, $\sigma(3.17) \approx 0.95969$, and $\sigma(2.365) \approx 0.914119$. Thus, we obtain the channel attention vector as follows:

$$\text{SCirc}(J)_{c^*} = \begin{bmatrix} 0.977357 & 0.802184 & 0.95969 & 0.914119 \end{bmatrix}.$$

To obtain the final result matrix, we multiply the channel attention vector $\text{SCirc}(J)_{c^*}$ with the skew circulant matrix $\text{SCirc}(J)$ constructed using Jacobsthal numbers. In this operation, each column (i.e., channel) of $\text{SCirc}(J)$ is scaled by the corresponding coefficient from $\text{SCirc}(J)_{c^*}$ as follows:

$$\text{SCirc}(J) = \begin{bmatrix} J_1 & J_2 & J_3 & J_4 \\ -J_4 & J_1 & J_2 & J_3 \\ -J_3 & -J_4 & J_1 & J_2 \\ -J_2 & -J_3 & -J_4 & J_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 & 5 \\ -5 & 1 & 1 & 3 \\ -3 & -5 & 1 & 1 \\ -1 & -3 & -5 & 1 \end{bmatrix}.$$

The channel attention coefficients applied to each column are: 1st column × 0.977357, 2nd column × 0.802184, 3rd column × 0.95969, and 4th column × 0.914119. Then, the resulting matrix is

$$\mathrm{SCirc}(J)_{\mathrm{new}} = \mathrm{SCirc}(J) \odot \mathrm{SCirc}(J)_{c^*} = \begin{bmatrix} 0.977357 & 0.802184 & 2.87907 & 4.570595 \\ -4.886785 & 0.802184 & 0.95969 & 2.742357 \\ -2.932071 & -4.01092 & 0.95969 & 0.914119 \\ -0.977357 & -2.406552 & -4.79845 & 0.914119 \end{bmatrix},$$

where $\odot$ denotes element-wise multiplication with appropriate broadcasting along the channel axis. This new matrix, $\mathrm{SCirc}(J)_{\mathrm{new}}$, is the reweighted output obtained after applying the channel attention module of CBAM.

By applying the channel attention mechanism, the original Jacobsthal-based skew circulant matrix is adaptively rescaled column-wise according to learned importance weights. These attention scores, derived from sigmoid-activated MLP outputs, enable selective enhancement or suppression of individual channels. For example, the first channel received the highest attention weight (0.977357), indicating its relatively greater contribution to the overall pattern recognition process. In contrast, the second channel was more significantly down-weighted (approximately 0.802), suggesting a lower relevance. The third and fourth channels had intermediate attention values (0.959690 and 0.914119, respectively), indicating moderate importance.

This dynamic reweighting helps highlight informative features while reducing the influence of less critical ones. In summary, we analyzed the channel attention mechanism applied to a skew circulant matrix constructed from Jacobsthal numbers. After computing the attention scores through the CBAM framework, the matrix underwent channel-wise reweighting based on learned importance values. These dynamic adjustments demonstrate CBAM's effectiveness in emphasizing dominant patterns and filtering out redundant or noisy components.

The motivation for applying CBAM to Jacobsthal-based skew circulant matrices lies in their inherent periodic and recursive nature, which often encodes latent symmetry and directionality. Unlike traditional matrix decompositions (such as SVD), CBAM provides a learnable mechanism to selectively enhance or suppress information across channels, offering deeper insights into the discriminative potential of such structured numerical sequences. This attention-based transformation not only reflects the relative importance of each channel but also prepares the matrix for downstream AI-related applications, such as feature extraction, classification, or signal enhancement.

We will now compare all the calculations presented above with the matrices obtained by applying both the first- and second-order transformations of the skew circulant matrices under the Ducci map to derive their new, final states. Additionally, we will compare these with matrices derived from the standard Jacobsthal sequence and its first- and second-order Ducci transformations. The first-order Ducci transformation is given by

$$\mathrm{SCirc}(D_J) = \begin{bmatrix} 2J_0 & 2J_1 & 2J_2 & J_4 - 1 \\ -(J_4 - 1) & 2J_0 & 2J_1 & 2J_2 \\ -2J_2 & -(J_4 - 1) & 2J_0 & 2J_1 \\ -2J_1 & -2J_2 & -(J_4 - 1) & 2J_0 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 2 & 4 \\ -4 & 0 & 2 & 2 \\ -2 & -4 & 0 & 2 \\ -2 & -2 & -4 & 0 \end{bmatrix}. \quad (19)$$

Let us consider each row as a channel and first compute the averages and maxima:

$$F_{\text{avg}} = \begin{bmatrix} \frac{0-4-2-2}{4} \\ \frac{2+0-4-2}{4} \\ \frac{2+2+0-4}{4} \\ \frac{4+2+2+0}{4} \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \qquad F_{\max} = \begin{bmatrix} \max(0,-4,-2,-2) \\ \max(2,0,-4,-2) \\ \max(2,2,0,-4) \\ \max(4,2,2,0) \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 4 \end{bmatrix}.$$

To ensure a fair comparison between the Jacobsthal-based and Ducci-transformed matrices, the same shared MLP weights (W_1 and W_2) were used in the channel attention module. This setup isolates the structural influence of each matrix type on the resulting attention weights, facilitating a consistent analytical interpretation.

The channel attention mechanism begins by applying global average pooling and global max pooling operations along the spatial dimension of the matrix. For the given fourth-order Jacobsthal-based circulant matrix under the Ducci transformation, this results in two 4×1 vectors, F_{avg} and $F_{\max}$.

Let us

$$W_1 = \begin{bmatrix} 0.5 & 0.1 & 0.4 & 0.6 \\ 0.2 & 0.3 & 0.1 & 0.5 \end{bmatrix}^T \qquad \text{and} \qquad W_2 = \begin{bmatrix} 0.3 & 0.2 & 0.4 & 0.1 \\ 0.6 & 0.1 & 0.3 & 0.5 \end{bmatrix}.$$

Here, to ensure a fair and consistent comparison of the attention mechanisms across the original Jacobsthal matrix and its Ducci-transformed counterparts, we employed the same shared MLP weight matrices (W_1 and W_2) throughout all CBAM computations. This approach isolates the effects of the matrix structural variations on the resulting attention scores, eliminating confounding factors related to differing learned parameters. Then

$$K_{\text{avg}} = W_1^T \cdot F_{\text{avg}} = \begin{bmatrix} 0.1 \\ 0.3 \end{bmatrix}, \qquad K_{\max} = W_1^T \cdot F_{\max} = \begin{bmatrix} 3.4 \\ 2.8 \end{bmatrix},$$

$$K = K_{\text{avg}} + K_{\max} = \begin{bmatrix} 0.1 \\ 0.3 \end{bmatrix} + \begin{bmatrix} 3.4 \\ 2.8 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 3.1 \end{bmatrix}.$$

To compute the sigmoid activations, $\sigma(3.5) \approx 0.970688$ and $\sigma(3.1) \approx 0.956893$. Hence, the attention vector becomes

$$\text{SCirc}(D_J)_c = \sigma(K) = \begin{bmatrix} 0.970688 \\ 0.956893 \end{bmatrix}.$$

These weights reflect the relative importance of each channel, with higher values indicating stronger attention. This vector contains only two values because the reduction ratio was $r = 2$, and thus the output of the MLP was 2-dimensional. However, the original matrix was 4 channels (i.e., 4 columns), so the final attention vector had to be 4-dimensional as well.

We now compute the final transformation using the learned MLP weights:

$$K_{\text{final}} = W_2^T \cdot K = \begin{bmatrix} 2.91 & 1.01 & 2.33 & 1.9 \end{bmatrix}^T$$

Applying the sigmoid function to each element of the resulting vector, $\sigma(2.91) \approx 0.948339$, $\sigma(1.01) \approx 0.73302$, $\sigma(2.33) \approx 0.911331$, and $\sigma(1.9) \approx 0.869892$. Hence, we obtain the channel attention vector:

$$\text{SCirc}(D_J)_{c^*} = \begin{bmatrix} 0.948339 & 0.73302 & 0.911331 & 0.869892 \end{bmatrix}$$

To obtain the resulting matrix, we multiply each column of the original skew circulant matrix $\mathrm{SCirc}(D_J)$ by the corresponding attention coefficient (19) in an element-wise manner.

Applying the attention vector to each column,

$$\mathrm{SCirc}(D_J)_{\mathrm{new}} = \mathrm{SCirc}(D_J) \odot \mathrm{SCirc}(D_J)_{c^*} = \begin{bmatrix} 0 & 1.46604 & 1.822662 & 3.479568 \\ -3.793356 & 0 & 1.822662 & 1.739784 \\ -1.896678 & -2.93208 & 0 & 1.739784 \\ -1.896678 & -1.46604 & -3.645324 & 0 \end{bmatrix}.$$

This new matrix, $\mathrm{SCirc}(D_J)_{\mathrm{new}}$, is the reweighted output obtained after the channel attention stage of CBAM. A refined interpretation of the channel attention outputs reveals critical insights into how CBAM discriminates between informative and less significant channels within different matrix structures. The Jacobsthal-based skew circulant matrix produced a relatively uniform attention vector, with all channel weights exceeding 0.80. This indicates that CBAM identified meaningful contributions across all channels, suggesting a homogeneously distributed informational structure where no particular channel was drastically suppressed. In contrast, the Ducci transformed matrix exhibited a more uneven attention profile, most notably with the second channel receiving a significantly lower weight of 0.733. The increased disparity among the attention scores suggests that CBAM detected sharper contrasts between channels, emphasizing those with stronger feature relevance while attenuating others considered redundant or noisy. This divergence implies that the Jacobsthal matrix supports a more balanced and generalizable feature representation, whereas the Ducci structure leads to a more localized and selective concentration of information. Such behavior underscores CBAM's capacity to perform fine-grained channel-wise filtering based on the latent structure of the input, making the Jacobsthal formulation more favorable for applications requiring uniform feature extraction. We will now compare all the calculations given above with the matrices obtained by applying second-order transformations of the skew circulant matrices under the Ducci transformations to obtain the new matrices, which are their final states, and with the matrices obtained by the normal Jacobsthal and second-order Ducci transformations. The second-order Ducci-transformed skew circulant fourth-order matrix has the following form:

$$\mathrm{SCirc}(D_J^2) = \begin{bmatrix} 4J_{-1} & 4J_0 & J_3 - 1 & J_4 - 1 \\ -(J_4 - 1) & 4J_{-1} & 4J_0 & J_3 - 1 \\ -(J_3 - 1) & -(J_4 - 1) & 4J_{-1} & 4J_0 \\ -4J_0 & -(J_3 - 1) & -(J_4 - 1) & 4J_{-1} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 2 & 4 \\ -4 & 2 & 0 & 2 \\ -2 & -4 & 2 & 0 \\ 0 & -2 & -4 & 2 \end{bmatrix}.$$

Let us think of each line as a channel and first take the averages and maxima:

$$F_{\mathrm{avg}} = \begin{bmatrix} -1 & -1 & 0 & 2 \end{bmatrix}^T, \qquad F_{\mathrm{max}} = \begin{bmatrix} 2 & 2 & 2 & 4 \end{bmatrix}^T.$$

To ensure a fair comparison between the Jacobsthal-based and second-order Ducci-transformed matrices, the same shared MLP weights (W_1 and W_2) were used in the channel attention module. This allows for isolating the structural influence of each matrix type on the resulting attention weights, thereby facilitating a consistent analytical interpretation.

The channel attention mechanism begins by applying global average pooling and global max pooling operations along the spatial dimension of the matrix. For the given fourth-order Jacobsthal-based circulant matrix under the second-order Ducci transformation, this results in two 4×1 vectors, F_{avg} and F_{max}.

Let us obtain the following:

$$W_1 = \begin{bmatrix} 0.5 & 0.1 & 0.4 & 0.6 \\ 0.2 & 0.3 & 0.1 & 0.5 \end{bmatrix}^T \quad \text{and} \quad W_2 = \begin{bmatrix} 0.3 & 0.2 & 0.4 & 0.1 \\ 0.6 & 0.1 & 0.3 & 0.5 \end{bmatrix}.$$

Then

$$K_{\text{avg}} = W_1^T \cdot F_{\text{avg}} = \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix}, \qquad K_{\max} = W_1^T \cdot F_{\max} = \begin{bmatrix} 4.4 \\ 3.2 \end{bmatrix},$$

and the total score is

$$K = K_{\text{avg}} + K_{\max} = \begin{bmatrix} 5.0 \\ 3.7 \end{bmatrix}.$$

For the first input to the sigmoid function, $x_1 = 5.0$, we compute $\sigma(5.0) \approx 0.993307$. For the second value, $x_2 = 3.7$, we have $\sigma(3.7) \approx 0.975873$. The resulting attention weight is

$$\text{SCirc}(D^2 J)_c = \sigma(K) = \begin{bmatrix} 0.993307 \\ 0.975873 \end{bmatrix}.$$

These weights reflect the relative importance of each channel, with higher values indicating stronger attention. The intermediate attention vector contains only 2 values because the reduction ratio was $r = 2$, resulting in a 2-dimensional output from the MLP. However, since the original matrix had 4 channels (i.e., 4 columns), the attention vector ultimately needed to be projected back to 4 dimensions. Therefore,

$$K_{\text{final}} = W_2^T K = \begin{bmatrix} 3.72 & 1.37 & 3.11 & 2.35 \end{bmatrix}^T.$$

If we apply sigmoid to the obtained matrix, the following values are obtained: $\sigma(3.72) \approx 0.976339$, $\sigma(1.37) \approx 0.79738$, $\sigma(3.11) \approx 0.957303$, and $\sigma(2.35) \approx 0.912934$. Thus, we get the channel attention vector as follows:

$$\text{SCirc}(D_J^2)_{c^*} = \begin{bmatrix} 0.976339 & 0.79738 & 0.957303 & 0.912934 \end{bmatrix}.$$

To obtain the result matrix, let us multiply the channel attention vector we obtained last with the skew circulant matrix we obtained under the Jacobsthal numbers $\text{SCirc}(D_J^2)$. In this multiplication, we will multiply each column (channel) by the relevant $\text{SCirc}(D_J^2)_{c^*}$ coefficient.

Applying the attention vector to each column,

$$\text{SCirc}(D_J^2)_{\text{new}} = \text{SCirc}(D_J^2) \odot \text{SCirc}(D_J^2)_{c^*} = \begin{bmatrix} 1.952678 & 0 & 1.914606 & 3.651736 \\ -3.905356 & 1.59476 & 0 & 1.825868 \\ -1.952678 & -3.18952 & 1.914606 & 0 \\ 0 & -1.59476 & -3.829212 & 1.825868 \end{bmatrix}.$$

This new matrix $\text{SCirc}(D_J^2)_{\text{new}}$ is the reweighted output matrix obtained after the channel attention part of CBAM.

A comparative analysis of the attention vectors derived from the original Jacobsthal matrix and its successive Ducci transformations reveals a non-linear, though not arbitrary, evolution. The original Jacobsthal matrix exhibits a well-balanced distribution of attention across all channels, reflecting a non-skewed and uniform information structure. Interestingly, the first-order Ducci transformation introduces pronounced intra-channel variation; this sharper contrast suggests that the transformed matrix accentuates certain features,

potentially making it more effective for AI tasks that benefit from feature emphasis, such as classification or pattern recognition.

Following the second-order Ducci transformation, the attention distribution returns to a more balanced state, closely resembling the original Jacobsthal configuration. This indicates a reorganization in which the initial transformation causes a structural perturbation, leading to a regularized pattern that can be effectively identified by the CBAM framework. Notably, the final transformation demonstrates the system's ability to suppress noise and recover the latent structure embedded in the data.

In summary, while the original and intermediate matrices support the broad representation of information, the first-order Ducci matrix provides greater focus on distinctive features, enabling enhanced feature separation.

The attention variance is computed from the channel-attention vector produced by the CBAM module. Let $a = (a_1, a_2, \ldots, a_m)$ denote the attention weights across m channels. The variance is calculated using the population variance formula

$$\mathrm{Var}(a) = \frac{1}{m}\sum_{i=1}^{m}(a_i - \bar{a})^2,$$

where $\bar{a}$ denotes the mean of the attention weights and m is the number of channels. In our experiments, $m = 4$, corresponding to the four channel attention values shown in the heatmap.

Figure 1 illustrates these observations in concrete detail.

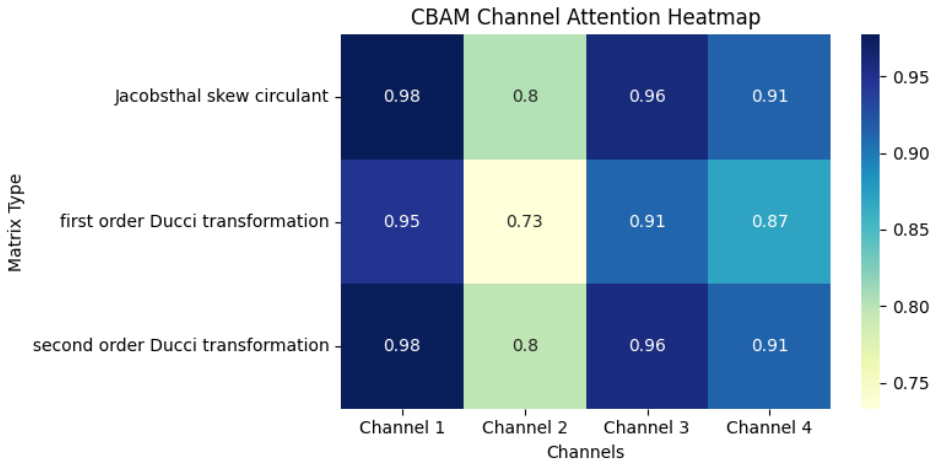


Figure 1. CBAM Channel Attention Heatmap.

Figure 1 shows the CBAM attention weights for three types of matrices: the original Jacobsthal-based skew circulant matrix, its first-order Ducci transformation, and the second-order transformation. Each cell's intensity indicates how much attention CBAM assigns to that channel.

The first Ducci transformation reduces attention in certain channels (e.g., Channel 2), while the second transformation restores a more balanced distribution, similar to the original Jacobsthal matrix. This demonstrates how iterative transformations influence channel-wise discriminative power, balancing between feature sparsity and structural regularity.

We quantify this effect using the variance of the CBAM attention vector:

- **Low variance:** Attention values are similar across channels, meaning all channels contribute roughly equally.
- **High variance:** Attention values differ widely, showing that CBAM focuses selectively on channels carrying more informative features while suppressing less relevant ones.

In attention-based AI models, highlighting important features while suppressing noise is crucial for learning and generalization. Variance measures how sharply CBAM discriminates between important and less important channels. Higher variance indicates stronger feature sparsity and sharper discrimination, which can improve model performance by concentrating capacity on key patterns [34]. This idea aligns with structured feature pruning strategies, such as Optimal Brain Damage, where removing low-variance weights improves both efficiency and generalization [35]. In our study, attention variance quantifies the impact of different circulant matrix configurations on channel selectivity.

Figure 2 summarizes the variance of the CBAM channel attention vectors computed for the original Jacobsthal matrix and its first- and second-order Ducci transformations. The first-degree Ducci matrix exhibits the highest variance, indicating a greater differentiation in attention scores across channels and thus a more selective focus on distinct features. In contrast, the original Jacobsthal and second-degree Ducci matrices show lower variances, reflecting a more uniform distribution of attention. This numerical evidence supports the interpretation that the first Ducci transformation enhances feature sparsity and channel discrimination, while the second transformation restores balance and structural regularity.

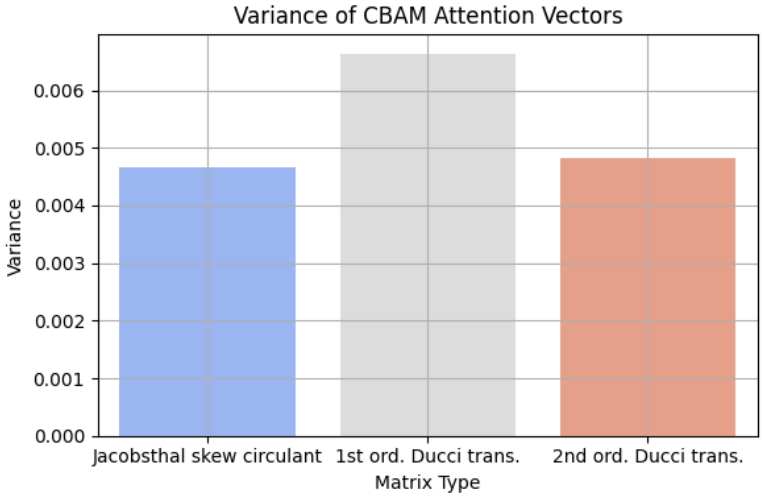


Figure 2. Variance of CBAM Attention Vectors.

5. Conclusions

Structured matrices, such as circulant and skew-Ducci matrices based on sequences like Jacobsthal numbers, can be interpreted as image-like inputs for CBAM because their entries form a 2D grid where both rows and columns encode meaningful numerical relationships. The choice of Jacobsthal and skew-Ducci sequences is motivated by their inherent recursive and periodic properties, which create structured patterns that CBAM can exploit. Each matrix element acts like a pixel, representing structural information from the sequence, while the skew-Ducci transformations introduce controlled variation that highlights specific channels. This mapping preserves relative differences and correlations among entries, allowing CBAM to learn attention patterns that reflect the underlying mathematical structure. Consequently, the model emphasizes important structural features while suppressing less relevant ones, maintaining interpretability and fidelity to the original sequence properties. This study explores the interaction between Ducci transformations and Jacobsthal numbers, extending their application to circulant matrices. By applying the Ducci transformation iteratively, we generated new sequences and examined their structural properties within matrix representations. Our findings include explicit computations of various matrix norms, such as the Frobenius, spectral, and l_p norms. The novelty of this

work lies in combining Ducci sequences with Jacobsthal numbers within the framework of circulant matrices, offering new perspectives in matrix theory and numerical analysis.

Moreover, we investigated the use of the CBAM to generate warped circulant matrices and Ducci-transformed forms derived from Jacobsthal numbers. By incorporating channel attention mechanisms, we effectively enhanced the most informative channels in the matrices, thereby improving feature discrimination. Our results showed that the first Ducci shift leads to more significant changes in focus, resulting in a complex and diverse suppression pattern. Interestingly, the second-order Ducci transformation exhibited a partial recovery of attention intensity, suggesting a nuanced interaction between iterative transformations and attention dynamics. The proposed framework provides a descriptive analysis of how structural transformations influence attention variance.

This interdisciplinary approach bridges classical matrix theory with contemporary AI techniques, opening new avenues for research in numerical analysis, algebraic signal processing, and machine learning.

For future work, extending this framework to include spatial attention and higher-order Ducci transformations may enhance our understanding of hierarchical feature extraction. Furthermore, applying this method to various numerical sequences and evaluating its impact on practical domains such as cryptography and matrix-based image processing could provide valuable contributions to real-world applications.

Author Contributions: Conceptualization, B.K., T.G. and E.Ö.; methodology, B.K., T.G. and E.Ö.; software, B.K.; validation, B.K. and T.G.; formal analysis, T.G. and E.Ö.; investigation, B.K., T.G. and E.Ö.; resources, B.K. and E.Ö.; data curation, B.K.; writing—original draft preparation, B.K. and T.G.; writing—review and editing, B.K., T.G. and E.Ö.; visualization, B.K.; supervision, E.Ö.; project administration, E.Ö.; funding acquisition, B.K. and T.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data used to support the findings of this study are included within the article. The numerical values used to compute the attention weights and variance are provided in the figures and tables of the manuscript. Additional details can be made available from the authors upon reasonable request.

Acknowledgments: We would like to sincerely thank the Editor and the Reviewer(s) for their careful reading of our manuscript and for their constructive and insightful comments. Their valuable suggestions have helped us significantly improve the clarity, rigor, and presentation of the paper. We have carefully addressed all comments and revised the manuscript accordingly. We greatly appreciate the time and effort invested in reviewing our work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Breuer, F. Ducci sequences in higher dimensions. *Integers* **2007**, *7*, 24.
2. Breuer, F.; Lötter, E.; van der Merwe, B. Ducci-sequences and cyclotomic polynomials. *Finite Fields Appl.* **2007**, *13*, 293–304. [CrossRef]
3. Brockman, G.; Zerr, R.J. Asymptotic behavior of certain Ducci sequences. *Fibonacci Quart.* **2007**, *45*, 155–163. [CrossRef]
4. Brown, R.; Merzel, J.L. Limiting behavior in Ducci sequences. *Period. Math. Hungar.* **2003**, *47*, 45–50. [CrossRef]
5. Calkin, N.J.; Stevens, J.G.; Thomas, D.M. A characterization for the length of cycles of the *n*-number Ducci game. *Fibonacci Quart.* **2005**, *43*, 53–59. [CrossRef]
6. Wong, F. Ducci processes. *Fibonacci Quart.* **1982**, *20*, 97–105. [CrossRef]
7. Ehrlich, A. Periods in Ducci's *n*-number game of differences. *Fibonacci Quart.* **1990**, *28*, 302–305. [CrossRef]
8. Hogenson, K.; Negaard, S.; Zerr, R.J. Matrix sequences associated with the Ducci map and the mediant construction of the rationals. *Linear Algebra Appl.* **2012**, *437*, 285–293. [CrossRef]
9. Odegard, I.; Zerr, R.J. The quadratic irrationals and Ducci matrix sequences. *Linear Algebra Appl.* **2015**, *484*, 344–355. [CrossRef]

10. Conway, J.H.; Guy, R.K. *The Book of Numbers*; Springer: New York, NY, USA, 1996.
11. Allouche, J.-P.; Shallit, J. *Automatic Sequences: Theory, Applications, Generalizations*; Cambridge University Press: Cambridge, UK, 2023.
12. Chamberland, M.; Thomas, D.M. The *n*-number Ducci game. *J. Differ. Equ. Appl.* **2004**, *10*, 339–342. [CrossRef]
13. Clausing, A. Ducci matrices. *Amer. Math. Monthly* **2018**, *125*, 901–921. [CrossRef]
14. Davis, P.J. *Circulant Matrices*; Wiley: New York, NY, USA, 1979.
15. Kuloğlu, B.; Eser, E.; Özkan, E. The *r*-circulant matrices associated with *k*-Fermat and *k*-Mersenne numbers. *WSEAS Trans. Math.* **2023**, *22*, 531–543. [CrossRef]
16. Andrade, E.; Carrasco-Olivera, D.; Manzaneda, C. On circulant like matrices properties involving Horadam, Fibonacci, Jacobsthal and Pell numbers. *Linear Algebra Appl.* **2021**, *617*, 100–120. [CrossRef]
17. Shannon, A.G.; Kuloğlu, B.; Özkan, E. Rhaly terraced sequences their generalizations, properties and applications. *Comp. Appl. Math.* **2025**, *44*, 226. [CrossRef]
18. Solak, S.; Bahşi, M. Some properties of circulant matrices with Ducci sequences. *Linear Algebra Appl.* **2018**, *542*, 557–568. [CrossRef]
19. Kızılateş, C.; Tuglu, N. On the norms of geometric and symmetric geometric circulant matrices with the Tribonacci number. *Gazi Univ. J. Sci.* **2018**, *31*, 555–567.
20. Kuloğlu, B.; Eser, E.; Özkan, E. On the properties of *r*-circulant matrices involving generalized Fermat numbers. *Sakarya Univ. J. Sci.* **2023**, *27*, 956–964. [CrossRef]
21. Kumari, M.; Prasad, K.; Özkan, E.; Jagmohan, T. On the norms and eigenvalues of *r*-circulant matrices with *k*-Mersenne and *k*-Mersenne–Lucas numbers. *Math. Notes* **2023**, *114*, 522–535. [CrossRef]
22. Özkan, E.; Akkuş, H. A new approach to *k*-Oresme and *k*-Oresme–Lucas sequences. *Symmetry* **2024**, *16*, 1407. [CrossRef]
23. Akbiyik, S.Y.; Akbiyik, M.; Yilmaz, F. One type of symmetric matrix with harmonic Pell entries, its inversion, permanents and some norms. *Mathematics* **2021**, *9*, 539. [CrossRef]
24. Shi, B.; Kızılateş, C. A new generalization of the Frank matrix and its some properties. *Comp. Appl. Math.* **2024**, *43*, 19. [CrossRef]
25. Durna, N.; Yildirim, M. Generalized terraced matrices. *Miskolc Math. Notes* **2016**, *17*, 201–208. [CrossRef]
26. Goy, T. On determinants and permanents of some Toeplitz-Hessenberg matrices whose entries are Jacobsthal numbers. *Eurasian Math. J.* **2018**, *9*, 61–67. [CrossRef]
27. Goy, T.; Shattuck, M. Determinant formulas of some Hessenberg matrices with Jacobsthal entries. *Appl. Appl. Math.* **2021**,*16*, 191–213.
28. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In *Computer Vision—ECCV 2018, Proceedings of the 15th European Conference, Munich, Germany, 8–14 September 2018*; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11211. [CrossRef]
29. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph Attention Networks. In Proceedings of the 6th International Conference on Learning Representations (ICLR–2018), Vancouver, BC, Canada, 30 April–3 May 2018; pp. 1–12. [CrossRef]
30. Lee, J.B.; Rossi, R.A.; Kim, S.; Ahmed, N.K.; Koh, E. Attention models in graphs: A survey. *arXiv* **2018**, arXiv:1807.07984. [CrossRef]
31. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Conference "Advances in Neural Information Processing Systems 30", Long Beach, CA, USA, 4–9 December 2017.
32. OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. Available online: https://oeis.org (accessed on 14 January 2026).
33. Karner, H.; Schneid, J.; Ueberhuber, C.W. Spectral decomposition of real circulant matrices. *Linear Algebra Appl.* **2003**, *367*, 301–311. [CrossRef]
34. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective kernel networks. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019.
35. Le Cun, Y.; Denker, J.S.; Solla, S.A. Optimal Brain Damage. In Proceedings of the Conference "Advances in Neural Information Processing Systems 2" (NIPS 1989), Denver, CO, USA, 27–30 November 1989.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.