MAXIMIZING THE SHANNON CAPACITY OF CONSTRAINED SYSTEMS WITH TWO CONSTRAINTS[*]

NAVIN KASHYAP[†]

Abstract. In this paper, we consider the problem of finding the set $\{A, B\} \subset \{0,1\}^m$ that maximizes, among all 2-subsets of $\{0,1\}^m$, the Shannon capacity, $H(A,B)$, of a constrained system of binary sequences that do not contain A or B as a contiguous subsequence. This problem is motivated by the problem of finding a pair of length-m binary sequences, called markers, that achieves the maximum rate, $R(2,m,n)$, of a $(2,m,n)$ periodic prefix-synchronized (PPS) code. A $(2,m,n)$ PPS code is a binary block code with two length-m markers A, B, and codewords of length n that inserts A and B alternately at regular intervals in the encoded bitstream, with the additional constraint that A and B may not appear anywhere in the encoded bitstream other than where inserted. We show that for any $m \geq 2$, $\lim_{n\to\infty} R(2,m,n) = \max\{H(A,B) : \{A,B\} \subset \{0,1\}^m\} = \log_2 \rho_{m-1}$, where ρ_{m-1} is the largest-magnitude zero of the polynomial $z^{m-1} - z^{m-2} - \ldots - 1$. Moreover, we completely characterize the sequences A and B that achieve $\max H(A,B)$, as well as those that achieve $R(2,m,n)$ for all sufficiently large n.

Key words. Shannon capacity, constrained codes, periodic prefix-synchronized codes, shifts of finite type

AMS subject classifications. 68P30, 94A45, 94A55, 37B10

1. Introduction.

We begin by defining the notion of Shannon capacity [1],[13] of a constrained system of binary sequences. Given a constraint set (or forbidden set) $\mathcal{F}$ of finite-length binary sequences, we define the corresponding *constrained system*, $\mathcal{S}(\mathcal{F})$, to be the set of finite-length binary sequences that do not contain any member of $\mathcal{F}$ as a contiguous subsequence. The *Shannon capacity* of the constrained system $\mathcal{S}(\mathcal{F})$ is defined as $H(\mathcal{F}) = \lim_{n\to\infty} n^{-1} \log_2 q_{\mathcal{F}}(n)$, where $q_{\mathcal{F}}(n)$ is the number of length-n sequences in $\mathcal{S}(\mathcal{F})$. In this paper, we shall be concerned with constraint sets $\mathcal{F} \subset \{0,1\}^m$ containing binary sequences of a fixed length m, and for the most part, sets $\mathcal{F}$ of cardinality 2. The main contribution of this paper is a complete solution to the problem of finding the set $\{A, B\}$ that maximizes $H(A,B)$[1] among all 2-subsets of $\{0,1\}^m$.

The motivation for this problem comes from two sources. The first source is the area of symbolic dynamics [10], where the problem may be reformulated in terms of characterizing those shifts of finite type that have the maximum entropy among all shifts that forbid 2-subsets of $\{0,1\}^m$. Indeed, a related problem was considered by Lind [11], who provided computable bounds on the change in the entropy of a shift of finite type when an extra sequence is added to the original set of forbidden sequences.

The second source for the problem is a closely related question that arises in the context of periodic prefix-synchronized codes which were introduced recently [8] as a family of sync-timing codes. Sync-timing codes are needed in most communication systems where data synchronization is needed (*cf.* [12]), *i.e.*, when a sequence of data symbols must be encoded into bits and transmitted across a channel that can make arbitrary insertion, deletion and substitution errors. These codes not only enable the decoder to resynchronize rapidly upon cessation of such errors to correctly reproduce data symbols, but also allow the decoder to produce estimates of the time indices of

[*] This work was supported by NSF Grants NCR-9415754 and CCR-9815006.

[†] Dept. of Electrical and Computer Engineering, University of California, San Diego, CA 92093. Email: nkashyap@ece.ucsd.edu.

[1] For ease of notation, we use $H(A)$, $H(A,B)$ etc. instead of $H(\{A\})$, $H(\{A,B\})$ etc.

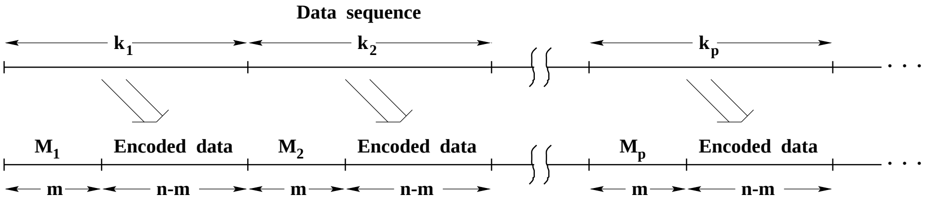

FIG. 1. *A fragment of the encoded bit stream.*

the decoded data symbols, in order to determine their positions in the original source sequence.

Periodic prefix-synchronized (PPS) codes are binary block codes that insert synchronizing markers at regular intervals (see Fig. 1). They are characterized by positive integers p, m, n, distinct binary length-m sequences called markers $M_1, \ldots, M_p$, and codebooks $C_1, \ldots, C_p$. Each C_i contains all binary codewords of length $n > m$, with the following properties: (i) each codeword begins with the marker M_i, and (ii) if $M_i = a_1^{(i)} \ldots a_m^{(i)}$, $i = 1, 2, \ldots, p$, and $a_1^{(i)} \ldots a_m^{(i)} b_1 \ldots b_{n-m}$ is a codeword from C_i, then none of the markers $M_1, \ldots, M_p$ can be found as a contiguous subsequence of $a_2^{(i)} \ldots a_m^{(i)} b_1 \ldots b_{n-m} a_1^{(i+1)} \ldots a_{m-1}^{(i+1)}$ (the superscript $(p+1)$ is to be interpreted as the superscript (1)). The idea here is that if a codeword from C_i were to be followed by one from C_{i+1}, then no marker can appear at any place except at the beginning of each codeword. A specific PPS code with *period p* and markers of length m, whose codebooks contain sequences of length n, will be referred to as a (p, m, n) PPS code.

The sequence of data symbols (which we assume to be binary) to be encoded is first divided into blocks of length $K = \sum_{i=1}^{p} k_i$, where $k_i = \lfloor \log_2 |C_i| \rfloor$, with $|C_i|$ denoting the cardinality of C_i. Each such block is then encoded by the PPS code as follows: the first k_1 data symbols are encoded using the codebook C_1, the next k_2 data symbols are encoded using C_2, and so on until the last k_p data symbols are encoded using C_p. Since the encoding procedure has a block structure with input blocklength K and output blocklength $N = pn$, the rate of the code is $R = K/N$. Note that $R \leq 1$ since $K \leq N$, and it is desirable to have codes with rates as close to 1 as possible, so as to minimize the redundancy introduced.

Due to the constraints on the codewords, it is clear that markers can only appear at specific places in the sequence of encoded bits, and hence any marker can be used by the decoder to recover synchronization. The idea behind inserting multiple markers in a periodic manner in the encoded sequence is that, as explained in [8], this periodicity allows the decoder to estimate the time index of each decoded data symbol, relative to the beginning of the "period" to which it belongs, without compromising on the delay in recovering synchronization.

For $p = 1$, the above description is simply that of a prefix-synchronized code, first studied by Gilbert [3] and further analyzed by Guibas and Odlyzko [4]. The PPS code with markers $M_1 = 000$, $M_2 = 111$, and codebooks $C_1 = \{0001100, 0001010\}$, $C_2 = \{1110011, 1110101\}$, is an example of a $(2, 3, 7)$ PPS code.

Let us define $R(p, m, n)$ to be the maximum rate achievable by a (p, m, n) PPS code, if such a code exists, and to be zero otherwise. To find $R(p, m, n)$, it is necessary to find markers $M_1, \ldots, M_p$ that maximize the sizes of the codebooks $C_1, \ldots, C_p$ subject to the constraints defining the code. Due to the similarity of the constraints involved, it is reasonable to expect that, asymptotically in n, $R(p, m, n)$ is closely related to $H_{p,m} = \max_{\mathcal{F}} H(\mathcal{F})$, with the maximum being taken over all p-subsets

$\mathcal{F} \subset \{0,1\}^m$.

The work of Gilbert [3], along with that of Guibas and Odlyzko [4],[5], showed that such a relationship does indeed hold for $p=1$, *i.e.*, for prefix-synchronized codes. Specifically, using generating functions, Gilbert showed that if $m \geq 1$ is fixed, then for all sufficiently large n, $R(1,m,n)$ is achieved by choosing the single marker, M_1, in the code to be either the length-m all-zeros sequence, 0^m, or the length-m all-ones sequence, 1^m. It also follows from his results that $\lim_{n\to\infty} R(1,m,n) = \log_2 \rho_m$, where ρ_m is the largest-magnitude zero of the polynomial $z^m - z^{m-1} - \ldots - 1$. Guibas and Odlyzko subsequently used generating functions to show (among other things) that for any $A \in \{0,1\}^m$, $H(A) \leq H(1^m) = \log_2 \rho_m$ with equality iff $A = 0^m$ or 1^m. Thus, we see that for $m \geq 1$, $\lim_{n\to\infty} R(1,m,n) = H_{1,m} = \log_2 \rho_m$.

In this paper, we derive a corresponding relationship for the case when $p=2$. The major part of this derivation lies in a proof of the fact that for all $m \geq 2$, $H_{2,m} = \log_2 \rho_{m-1}$, where ρ_{m-1} is the largest-magnitude zero of the polynomial $z^{m-1} - z^{m-2} - \ldots - 1$. This fact is then used to show that $\lim_{n\to\infty} R(2,m,n) = \log_2 \rho_{m-1}$ as well. Moreover, we identify (Theorem 1) all the pairs of length-m sequences that achieve $H_{2,m}$, as well as those (Theorem 2) that achieve $R(2,m,n)$ for all sufficiently large n. We also provide a partial result (Theorem 18) for $p=3$, for which we show that $H_{3,m} = \log_2 \rho_{m-1}$ as well, and is achieved by the set $\{10^{m-1}, 0^{m-1}1, 0^m\}$. However, this set of sequences cannot be used as markers in a $(3,m,n)$ PPS code, as any codeword that begins with 0^m must have an occurrence of 0^m or $0^{m-1}1$ starting at the second bit, which violates the constraints defining the code. Hence, it may not be true that $\lim_{n\to\infty} R(3,m,n) = H_{3,m}$.

For higher values of p, not even $H_{p,m}$ is known, although we conjecture that if $p = 2^k$ for any $k \geq 0$, then for $m \geq k+1$, $H_{p,m} = \log_2 \rho_{m-k}$, where ρ_{m-k} is the largest-magnitude zero of the polynomial $z^{m-k} - \ldots - 1$. This conjecture is based on the following observation. Let $\mathcal{F}_0 = \{0^{m-k}\langle i \rangle_2 : i = 0,1,\ldots,2^{k-1}\}$, where $\langle i \rangle_2$ denotes the k-bit binary representation of i. Also, define the functions $\psi_1, \psi_2, \psi_3 : \{0,1\}^* \to \{0,1\}^*$, where $\{0,1\}^*$ denotes the set of all finite-length binary sequences, as follows: for $b_1 b_2 \ldots b_n \in \{0,1\}^*$,

$$\begin{aligned}
\psi_1(b_1 b_2 \ldots b_n) &= b_1 b_2 \ldots b_{n-1} \\
\psi_2(b_1 b_2 \ldots b_n) &= b_2 b_3 \ldots b_n \\
\psi_3(b_1 b_2 \ldots b_n) &= c_1 c_2 \ldots c_{n-1}
\end{aligned}$$

where $c_i = b_i \oplus b_{i+1}$, $\oplus$ being modulo-2 addition. Note that each ψ_i is a two-to-one function. Now, numerical evidence seems to suggest that when $p = 2^k$, for any p-subset $\mathcal{F} \subset \{0,1\}^m$, $q_{\mathcal{F}}(n) \leq q_{\mathcal{F}_0}(n)$ for all n, with equality (for all n) if and only if $\mathcal{F} = \psi_{i_1}^{-1} \circ \psi_{i_2}^{-1} \circ \ldots \circ \psi_{i_k}^{-1}(0^{m-k})$ or $\psi_{i_1}^{-1} \circ \psi_{i_2}^{-1} \circ \ldots \circ \psi_{i_k}^{-1}(1^{m-k})$, for any $i_1, i_2, \ldots, i_k \in \{1,2,3\}$. It is a simple matter to verify that when $\mathcal{F}$ is of the above form, we have $q_{\mathcal{F}}(n) = 2^k q_{0^{m-k}}(n-k)$, so that $H(\mathcal{F}) = H(0^{m-k}) = \log_2 \rho_{m-k}$, which leads us to the statement of the conjecture.

Before stating our main result, we define some notation that is used throughout this paper: $\langle 01 \rangle_m$ and $\langle 10 \rangle_m$ denote the two length-m sequences of alternating 0's and 1's. The main contribution of this paper is a proof of the following result:

THEOREM 1. *If A, B are distinct binary sequences of length $m \geq 5$, then*

$$H(A,B) \leq \log_2 \rho_{m-1}$$

with equality if and only if $\{A, B\}$ *or* $\{\overline{A}, \overline{B}\}$ *is one of the following:* $\{0^m, 1^m\}$, $\{\langle 01\rangle_m, \langle 10\rangle_m\}$, $\{0^m, 10^{m-1}\}$, $\{0^m, 0^{m-1}1\}$ *and* $\{10^{m-1}, 0^{m-1}1\}$. *(*$\overline{A}, \overline{B}$ *are the sequences obtained by complementing each bit of* A, B.*)*

In the terminology of symbolic dynamics, this theorem shows that the entropy of a shift of finite type that forbids some 2-subset $\mathcal{F} \subset \{0,1\}^m$ is at most $\log_2 \rho_{m-1}$. This maximum is achieved precisely when $\mathcal{F}$ is one of the sets listed in the statement of the theorem.

The approach we use to prove the above theorem is based on a generating function for the number, $q_{AB}(n)$, of length-n binary sequences that do not contain A or B as a contiguous subsequence. This generating function can be expressed in a simple form, based on the concept of correlation between two binary strings, which we now define.

The *correlation* between two binary sequences A and B (not necessarily distinct), denoted by $A \circ B$, is a binary sequence of the same length as A. The ith bit (from the left) of $A \circ B$ is determined as follows: place B under A in such a way that the first bit of B lies under the ith bit of A; if the segments that overlap are identical, then the ith bit of $A \circ B$ is 1, else it is 0. Note that if A and B have the same length, then the first bit of $A \circ B$ is a 1 if and only if $A = B$. For example, if $A = 110001$ and $B = 1000$, then $A \circ B = 010001$, $B \circ A = 0000$, $A \circ A = 100001$, and $B \circ B = 1000$. The correlation of a sequence A with itself is also called the *autocorrelation* of A.

If $A \circ B = (c_0 c_1 \ldots c_{n-1})$ is the correlation between two sequences A and B, then we define the corresponding *correlation polynomial*

$$\phi_{AB}(z) = \sum_{i=0}^{n-1} c_i z^{n-1-i} \tag{1}$$

With A and B as in the previous example, we have $\phi_{AB}(z) = z^4 + 1$, $\phi_{BA}(z) = 0$, $\phi_{AA}(z) = z^5 + 1$, and $\phi_{BB}(z) = z^3$. For the sake of notational simplicity, it shall henceforth be tacitly understood that correlation polynomials are functions of the complex variable z, and so the argument z will be dropped from their notation whenever deemed necessary.

Guibas and Odlyzko [5] showed that given two distinct sequences $A, B \in \{0,1\}^m$, the generating function for $q_{AB}(n)$, defined by $Q_{AB}(z) = \sum_{n=0}^{\infty} q_{AB}(n) z^{-n}$, can be expressed using correlation polynomials as

$$Q_{AB}(z) = \frac{z(\phi_{AA}\phi_{BB} - \phi_{AB}\phi_{BA})}{(z-2)(\phi_{AA}\phi_{BB} - \phi_{AB}\phi_{BA}) + \phi_{AA} + \phi_{BB} - \phi_{AB} - \phi_{BA}} \tag{2}$$

Thus, the generating function $Q_{AB}(z)$ is a rational function, and we show that it always has a positive real pole, ρ_{AB}, that is larger in magnitude than any other pole. It then follows from the theory of complex variables[2] (see *e.g.*, [14], Chap. 5) that $q_{AB}(n) = c(n)\,(\rho_{AB})^n\,(1 + o(1))$ for some $c(n)$ depending polynomially on n ($c(n)$ is a constant if ρ_{AB} is a simple pole). This shows that $H(A, B) = \log_2 \rho_{AB}$, and a careful analysis thereafter shows how ρ_{AB} varies with A and B, and what choice of A, B maximizes ρ_{AB}.

To connect the above theorem with $R(2, m, n)$, we use a rational generating function for the number, $f_{AB}(k)$, of length-k sequences that begin with A, end with B, but do not contain A or B anywhere else. Using Guibas and Odlyzko's methods, it is

[2] We need the largest pole here because we define $Q_{AB}(z)$ as a power series in z^{-1}.

shown in [9] that if A and B are distinct length-m binary sequences, then for $k > m$, $f_{AB}(k)$ is the coefficient of z^{-k} in the expansion of

$$(3)\qquad F_{AB}(z) = \frac{1}{z}\,\frac{(z-2)\phi_{AB}+1}{(z-2)(\phi_{AA}\phi_{BB}-\phi_{AB}\phi_{BA})+\phi_{AA}+\phi_{BB}-\phi_{AB}-\phi_{BA}}$$

as $F_{AB}(z) = \sum_{k=0}^{\infty} v_k z^{-k}$. Note that we can use $f_{AB}(k)$ to define the rate of a $(2,m,n)$ PPS code with markers $M_1 = A$ and $M_2 = B$ as follows:

$$(4)\qquad \widehat{R}(A,B,n) = \frac{\lfloor \log_2 f_{AB}(m+n)\rfloor + \lfloor \log_2 f_{BA}(m+n)\rfloor}{2n}$$

Hence, $R(2,m,n) = \max_{A,B}\widehat{R}(A,B,n)$, the maximum being taken over all pairs of distinct sequences $A, B \in \{0,1\}^m$. By showing that in most cases, the largest-magnitude pole of $F_{AB}(z)$ is the same as that for $Q_{AB}(z)$, we are able to prove the following result:

THEOREM 2. *If A, B are distinct binary sequences of length $m \geq 5$, then*

$$\lim_{n\to\infty} \widehat{R}(A,B,n) \leq \log_2 \rho_{m-1}$$

with equality if and only if $\{A,B\} = \{0^m, 1^m\}$ or $\{\langle 01\rangle_m, \langle 10\rangle_m\}$. Consequently, $\lim_{n\to\infty} R(2,m,n) = \log_2 \rho_{m-1}$.

This theorem shows that when $m \geq 5$, for all sufficiently large n, $R(2,m,n)$ is either $\widehat{R}(0^m,1^m,n)$ or $\widehat{R}(\langle 01\rangle_m, \langle 10\rangle_m, n)$. In fact, we show further that for nearly all (if not all) values of n, $\widehat{R}(0^m,1^m,n) = \widehat{R}(\langle 01\rangle_m, \langle 10\rangle_m, n)$.

The remainder of this paper is devoted to the proofs of the above results. In §2, we show that $Q_{AB}(z)$ has a real largest-magnitude pole, ρ_{AB}, by demonstrating that the poles of $Q_{AB}(z)$ are actually eigenvalues of a certain non-negative matrix, which allows us to utilize the powerful Perron-Frobenius theory. Theorem 1 is proved in §3 by studying the behavior of ρ_{AB} as A and B vary. In §4, we explore the relationship between $H(A,B)$ and $\widehat{R}(A,B,n)$, and prove Theorem 2.

2. Walks on Graphs. In this section, we show that $Q_{AB}(z)$ has a positive real pole that is largest in magnitude among all poles of $Q_{AB}(z)$. It is well known that $q_{AB}(n)$ is precisely the number of walks of length $n-m+1$ on a certain directed graph $\mathcal{G}_{AB}$ obtained by removing a pair of edges from the de Bruijn graph $\mathcal{G}^{(m-1)}$ of order $m-1$. $\mathcal{G}^{(m-1)}$ is a directed graph with vertex set $\{v_i : i = 0,1,\ldots,2^{m-1}-1\}$ (see Fig. 2). If we label each vertex v_i with the $(m-1)$-bit binary representation of i, then $\mathcal{G}^{(m-1)}$ has a directed edge from v_i to v_j if and only if there exists a binary m-sequence $(b_1, b_2, \ldots, b_m)$ whose first $m-1$ bits forms v_i's label, and whose last $m-1$ bits forms v_j's label. Moreover, this directed edge is labeled with the bit b_m. Thus, each walk of length $n-m+1$ on $\mathcal{G}^{(m-1)}$ has a unique binary n-sequence associated with it, namely the sequence formed by concatenating the label of the initial vertex with the labels of the $n-m+1$ edges constituting the walk. In fact, this establishes a one-to-one correspondence between walks of length $n-m+1$ on $\mathcal{G}^{(m-1)}$ and binary n-sequences. Since the edges of the graph are themselves walks of length 1, there is a one-to-one correspondence between the edge set of $\mathcal{G}^{(m-1)}$ and the set of binary m-sequences. Defining $\mathcal{G}_{AB}$ to be the graph obtained by removing the edges corresponding to the sequences A and B from $\mathcal{G}^{(m-1)}$, it is not hard to see that for $n \geq m$, $q_{AB}(n)$ is equal

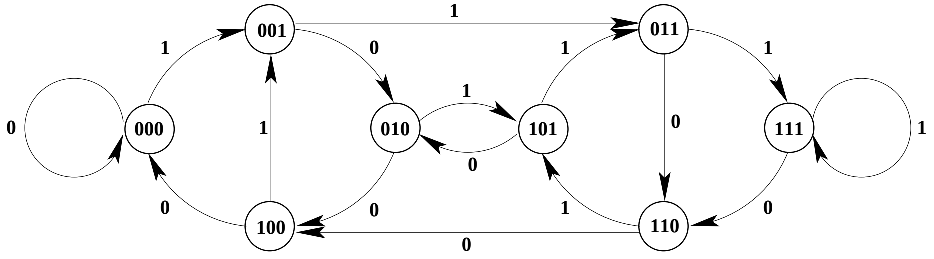

FIG. 2. *The de Bruijn graph $\mathcal{G}^3$.*

to the number of walks of length $n-m+1$ on $\mathcal{G}_{AB}$. Note that for $0 \leq n \leq m-1$, $q_{AB}(n) = 2^n$, as all the binary sequences of length n do not contain A or B.

Let $\mathcal{A}$ be the adjacency matrix of $\mathcal{G}_{AB}$. It is easy to show that the number of walks of length $n-m+1$ on $\mathcal{G}_{AB}$ is given by the sum of the entries of $\mathcal{A}^{n-m+1}$. Therefore, $q_{AB}(n) = \mathbf{1}^T \mathcal{A}^{n-m+1}\mathbf{1}$ for $n \geq m$, where $\mathbf{1}$ is a column vector of ones. Now if $\mathcal{A} = SJS^{-1}$, where J is the Jordan canonical form of $\mathcal{A}$, then $q_{AB}(n) = \mathbf{1}^T SJ^{n-m+1}S^{-1}\mathbf{1} = \mathbf{x}^T J^{n-m+1}\mathbf{y}$, for some column vectors $\mathbf{x}$ and $\mathbf{y}$. Utilizing the special structure of Jordan forms and working through the details, it can be shown that for $n \geq m$,

$$q_{AB}(n) = \sum_{i=1}^{r} p_i(n) \left(\lambda_i\right)^n \tag{5}$$

where $\lambda_1, \ldots, \lambda_r$ are the distinct non-zero eigenvalues of $\mathcal{A}$, and each p_i is a polynomial (possibly zero) of degree strictly less than the algebraic multiplicity of λ_i. Since $\mathcal{A}$ is a non-negative matrix, the Perron-Frobenius theorem cite[Theorem 8.3.1]horn shows that it has a real positive eigenvalue, say λ_1, such that $|\lambda_i| \leq \lambda_1$ for $i = 1, \ldots, r$. Note that $\lambda_1 \geq 1$ because if $0 < \lambda_1 < 1$, then (5) shows that we would have $0 < q_{AB}(n) < 1$ for some sufficiently large n. The case $\lambda_1 = 1$ is of little interest, as $q_{AB}(n)$ then grows polynomially with n, which shows that $H(A, B) = 0$. Therefore, we shall henceforth focus on the case when $\lambda_1 > 1$. We next show that in this case, λ_1 is the unique largest-magnitude eigenvalue of $\mathcal{A}$, *i.e.*, $|\lambda_i| < \lambda_1$ for $i \neq 1$, and is algebraically simple. This will then imply that $p_1(n)$ is a non-zero constant.

It is easy to see that since $\mathcal{G}_{AB}$ is obtained by removing two edges from $\mathcal{G}^{(m-1)}$, it is either irreducible (*i.e.*, it has a path connecting every ordered pair of vertices), or it has one vertex with either no incoming edges or no outgoing edges while the rest of the graph is irreducible. It is also a straightforward exercise to show that in the former case, $\mathcal{G}_{AB}$ is aperiodic as well (*i.e.*, the greatest common divisor of the lengths of all the cycles in the graph is 1), because it either contains a loop (an edge that connects a vertex to itself), or it contains two cycles of lengths 2 and 3 respectively. Therefore by Theorem 4.5.11 in [10], $\mathcal{A}$ has a unique largest-magnitude eigenvalue which is also algebraically simple. In the case when $\mathcal{G}_{AB}$ has an isolated vertex, if we let $\widehat{\mathcal{G}}_{AB}$ be the subgraph of $\mathcal{G}_{AB}$ obtained by removing that vertex and all the edges attached to it, then $\widehat{\mathcal{G}}_{AB}$ is also aperiodic as it always contains a loop. If $\hat{\mathcal{A}}$ is the adjacency matrix of $\widehat{\mathcal{G}}_{AB}$, then $\hat{\mathcal{A}}$ has a unique largest eigenvalue which is also algebraically simple. Now, the eigenvalues of $\mathcal{A}$ are precisely the eigenvalues of $\hat{\mathcal{A}}$, along with the eigenvalues of the adjacency matrix of the subgraph of $\mathcal{G}_{AB}$ formed by the isolated vertex and any edges connecting that vertex to itself [10, Section 4.4]. Since the latter adjacency matrix contains only one element, which is either 0 or 1, the only eigenvalue it contributes to $\mathcal{A}$ is 0 or 1. Hence, as $\lambda_1 > 1$, it must come from $\hat{\mathcal{A}}$ and so it is the unique largest-magnitude eigenvalue of $\mathcal{A}$ and is algebraically simple.

Since $p_1(n)$ has degree strictly less than the algebraic multiplicity of λ_1, we see that $p_1(n)$ is a constant c_1. Now, it is known that the number of length-n walks on any graph is at least as large as $c\,(\lambda_{\max})^n$, where $\lambda_{\max}$ is the largest eigenvalue of the adjacency matrix of the graph, and c is some positive constant. Therefore, we have $q_{AB}(n) \geq c(\lambda_1)^n$. Since $\lambda_1 > 1$ is the unique largest-magnitude eigenvalue of $\mathcal{A}$, (5) now shows that $p_1(n) = c_1 > 0$. (At this point, we would like to remark that even if $\lambda_1 = 1$, then it can be argued that $p_1(n)$ is non-zero, but not necessarily a constant, but we omit the argument as this fact is not essential for our results.)

Let us now define the function $\widehat{Q}_{AB}(z) = \sum_{n=m}^{\infty} q_{AB}(n)z^{-n}$. Since $q_{AB}(n)$, $n \geq m$, can be expressed in the form given in (5), using the formulae for summation of infinite series, it is easy to see that $\widehat{Q}_{AB}(z)$ is a rational function whose non-zero poles are eigenvalues of $\mathcal{A}$, and the multiplicity of the pole at λ_i is precisely one larger than the degree of p_i (if $p_i \equiv 0$, then we take $\deg(p_i)$ to be -1, which implies that there is no pole at λ_i). Since for $\lambda_1 > 1$, $\deg(p_1) = 0$, there is always a simple pole at λ_1. Note that $Q_{AB}(z) = \sum_{n=0}^{m-1} 2^n z^{-n} + \widehat{Q}_{AB}(z)$, but $\sum_{n=0}^{m-1} 2^n z^{-n}$ cannot contribute any non-zero poles to $Q_{AB}(z)$. Therefore, the non-zero poles of $Q_{AB}(z)$ are the same as those of $\widehat{Q}_{AB}(z)$. In particular, λ_1 is a pole of $Q_{AB}(z)$, and is in fact the unique largest-magnitude pole. (We further remark that in the case when $\lambda_1 = 1$, this argument would show that λ_1 is a largest-magnitude pole of $Q_{AB}(z)$, but it may not be a simple pole).

To summarize, we have shown that $Q_{AB}(z)$ always has a real largest-magnitude pole $\rho_{AB} \geq 1$. If $\rho_{AB} > 1$, then it is the unique largest-magnitude pole and is also simple, and hence as explained previously, $H(A,B) = \log_2 \rho_{AB}$. On the other hand, if $\rho_{AB} = 1$, then $H(A,B) = 0$. Thus, in order to maximize $H(A,B)$, we need to maximize ρ_{AB}. In the next section, we study how ρ_{AB} behaves as A and B vary.

3. Maximizing $H(A,B)$. Note that if we define $D_{AB}(z) = (z-2)(\phi_{AA}\phi_{BB} - \phi_{AB}\phi_{BA}) + \phi_{AA} + \phi_{BB} - \phi_{AB} - \phi_{BA}$, then by adding a zero at 2 to D_{AB}, we obtain a polynomial Δ_{AB} that is easier to handle. More precisely,

$$\Delta_{AB}(z) = (z-2)D_{AB}(z) = \gamma_{AA}\gamma_{BB} - \gamma_{AB}\gamma_{BA} \tag{6}$$

where γ_{**} is the polynomial defined by $\gamma_{**}(z) = (z-2)\phi_{**}(z) + 1$. Thus, the behavior of $D_{AB}(z)$ is intimately connected with the behavior of the polynomials γ_{**}, making it necessary to gain some understanding of the behavior of γ_{**}.

We shall also find it convenient to define the polynomials $p_k(z) = z^k - z^{k-1} - \ldots - 1$, for $k = 1, 2, \ldots$. It is known [15] that for $k \geq 2$, p_k has exactly one root, which is a simple root, in the region $1 < z < 2$, and all its other roots lie within the unit circle (for $k = 1$, the only root of p_k is 1). We shall denote the largest root of p_k by ρ_k. It is also known that ρ_k increases with k, and $2(1 - 2^{-k}) < \rho_k < 2$ for $k \geq 2$. We use these facts about p_k and ρ_k extensively in all that follows.

To any polynomial ϕ with coefficients 0 and 1, we can associate a *γ-polynomial* defined by $\gamma(z) = (z-2)\phi(z) + 1$. Since $\gamma(z) \geq 1$ for $z \geq 2$, all the real zeros of γ must be less than 2. Moreover, if ϕ is not identically zero, then γ has a real zero in $[1, 2)$ because $\gamma(1) = -\phi(1) + 1 \leq 0$. Thus, the largest positive zero of γ lies in $[1, 2)$. We now provide some more results concerning these γ-polynomials.

Lemma 3. *Let ϕ_1, ϕ_2 be polynomials in z with coefficients 0 and 1, with $\phi_2(2) = \phi_1(2) + 1$. Let γ_1, γ_2 be the corresponding γ-polynomials. Let k be the largest integer such that the coefficients of z^k in ϕ_1 and ϕ_2 are different. If $k = 0$, then $\gamma_1(z) > \gamma_2(z)$ for all $z < 2$, and if $k \geq 1$, then $\gamma_1(z) \geq \gamma_2(z)$ for $\rho_k \leq z < 2$, with equality iff $z = \rho_k$.*

Proof: Note that the sequence formed by the coefficients of each ϕ_i is the binary representation of the integer $\phi_i(2)$. Therefore, it is convenient to identify each ϕ_i with the binary sequence formed by its coefficients. Since ϕ_1 and ϕ_2 are identified with sequences that are binary representations of successive integers, it can be seen that the coefficient of z^k in ϕ_2 is 1, while that in ϕ_1 is 0. Moreover, for each $j < k$, the coefficient of z^j in ϕ_2 is 0, while that in ϕ_1 is 1. Therefore, $\phi_2(z) - \phi_1(z) = 1$ if $k = 0$, and if $k \geq 1$, $\phi_2(z) - \phi_1(z) = z^k - z^{k-1} - \ldots - 1$.

If $k = 0$, we have $\gamma_2(z) - \gamma_1(z) = (z-2)(\phi_2(z) - \phi_1(z)) = z - 2$, which is negative for $z < 2$. If $k \geq 1$, then $\gamma_2(z) - \gamma_1(z) = (z-2)(\phi_2(z) - \phi_1(z)) = (z-2)\,p_k$. Since ρ_k is the largest root of p_k and $p_k(2) = 1$, we must have $p_k(z) > 0$ for $\rho_k < z \leq 2$, from which the result follows. □

Lemma 4. *Let ϕ_1, ϕ_2 be non-zero polynomials with coefficients 0 and 1, and let γ_1, γ_2 be the corresponding γ-polynomials. Let r_1 and r_2 be the largest positive roots of γ_1 and γ_2 respectively. Suppose that $\phi_1(2) < \phi_2(2)$. Then, $r_1 \leq r_2$ with equality iff $\phi_1(z) = z^{m-1} + z^{m-2} + \ldots + 1$ and $\phi_2(z) = z^m$ for some $m \geq 1$. Moreover, for $r_2 < z < 2$, $\gamma_1(z) > \gamma_2(z)$.*

Proof: It suffices to consider the case when $\phi_2(2) - \phi_1(2) = 1$. The general case then follows by induction on $\phi_2(2) - \phi_1(2)$. Let k be the largest integer such that the coefficients of z^k in $\phi_1(z)$ and $\phi_2(z)$ differ. Note that $\gamma_1(2) = \gamma_2(2) = 1 > 0$. If $k = 0$, then by the previous lemma, since $r_1 < 2$, we have $\gamma_2(r_1) < \gamma_1(r_1) = 0$, which shows that γ_2 has a real root in $(r_1, 2)$. Therefore, $r_2 > r_1$.

If $k \geq 1$, then define $\hat{\phi}_1(z) = z^{k-1} + \ldots + 1$, $\hat{\phi}_2(z) = z^k$. The corresponding γ-polynomials are $\hat{\gamma}_1(z) = z^k - z^{k-1} - \ldots - 1$, and $\hat{\gamma}_2(z) = (z-1)(z^k - z^{k-1} - \ldots - 1)$. It is clear that if $\phi_i = \hat{\phi}_i$, $i = 1, 2$, then $r_1 = r_2 = \rho_k$.

Now, suppose that $\phi_1 \neq \hat{\phi}_1$, in which case since $\phi_2(2) = \phi_1(2) + 1$, we must have $\phi_2 \neq \hat{\phi}_2$ as well. Note that for $i = 1, 2$, we have $\gamma_i(z) = \hat{\gamma}_i(z) + (z-2)(\phi(z) - \hat{\phi}_i(z))$. Since all the coefficients of z^j, $j \leq k$, in ϕ_i are the same as those in $\hat{\phi}_i$, we see that $\phi_i(z) - \hat{\phi}_i(z)$ is itself a non-zero polynomial with coefficients 0 and 1. Therefore, at $z = \rho_k$, $\hat{\gamma}_i$ vanishes and $\phi_i - \hat{\phi}_i$ is positive, which implies that $\gamma_i(\rho_k) < 0$. Hence, we must have $r_1, r_2 > \rho_k$. Now, again by the previous lemma, we have $g_2(r_1) < g_1(r_1) = 0$, which shows that $r_2 > r_1$.

The fact that $\gamma_1(z) > \gamma_2(z)$ for $r_2 < z < 2$ also follows from the previous lemma, since we have shown that $r_2 \geq \rho_k$. □

Lemma 5. *Let $f(z) = (z-2)p(z)$, where $p(z)$ is a non-zero polynomial of degree m with non-negative coefficients. Then, $f'(z) > 0$ for $z > 2(1 - \frac{1}{m+1})$.*

Proof: Let $p(z) = \sum_{i=0}^{m} a_i z^i$, with $a_i \geq 0$ for $i = 0, 1, \ldots, m-1$, and $a_m > 0$. Then, $f(z) = \sum_{i=0}^{m} a_i (z-2) z^i$. Therefore, $f'(z) = \sum_{i=0}^{m} a_i ((i+1) z^i - 2 i z^{i-1})$. Noting that $(i+1)z^i - 2iz^{i-1} > 0$ for $z > 2(1 - \frac{1}{i+1})$, the lemma follows. □

Lemma 6. *Let ϕ be a polynomial of degree $m \geq 2$ with coefficients 0 and 1, and let γ be the corresponding γ-polynomial. Then, γ has exactly one real root r in the interval $(1, 2)$. Moreover, r is a simple root, $\gamma(z) < 0$ for $1 < z < r$, and $g(z) > 0$ for*

$r < z \leq 2$.

Remark: It is easily verified that the conclusions of the lemma are also valid for $\phi(z) = z + 1$. However, if $\phi(z) = z$, 1 or 0, then γ has no roots in (1,2).

Proof: Define $\hat{\phi}(z) = z^m$, and $\hat{\gamma}(z) = (z-2)z^m + 1 = (z-1)(z^m - z^{m-1} - \ldots - 1)$. If $\phi = \hat{\phi}$, then $\gamma = \hat{\gamma}$ has exactly one root, ρ_m, in (1,2), and it is simple. Since ρ_m is a simple root, $\hat{\gamma}(z)$ must undergo exactly one sign change in (1,2). Therefore, noting that $\hat{\gamma}(2) = 1 > 0$, we must have $\hat{\gamma}(z) > 0$ for $\rho_m < z \leq 2$, and $\hat{\gamma}(z) < 0$ for $1 < z < \rho_m$.

If $\phi \neq \hat{\phi}$, then $\phi(z) - \hat{\phi}(z) > 0$ for $z > 0$. Therefore, $\gamma(z) - \hat{\gamma}(z) = (z-2)(\phi(z) - \hat{\phi}(z)) < 0$ for $0 < z < 2$. Therefore, $\gamma(z) < \hat{\gamma}(z) < 0$ for $1 < z < \rho_m$, which shows that γ has no roots in $(1, \rho_m)$.

Now by the previous lemma, $\gamma'(z) > 0$ for $z > 2(1 - \frac{1}{m+1})$. Hence, if γ has a root in this range, it must be unique (since γ is strictly increasing), and it must have multiplicity 1 (since $\gamma'(z) \neq 0$). Now by Lemma 4, γ has a root r larger than the largest root, ρ_m, of $\hat{\gamma}(z)$. But as mentioned earlier, $\rho_m > 2(1 - 2^{-m}) > 2(1 - \frac{1}{m+1})$. The negative and positive regions for $\gamma(z)$ are determined using arguments identical to those used above for $\hat{\gamma}(z)$. □

We next prove a theorem that locates the largest positive zero of D_{AB}, which is the denominator of Q_{AB} in (2).

THEOREM 7. *Let A and B be distinct binary sequences of length $m \geq 5$. Then, $D_{AB}(z)$ has its largest positive real root ρ in (1,2). Moreover,* $\max\{r_{AB}, r_{BA}\} \leq \rho \leq \min\{r_{AA}, r_{BB}\}$, *where r_{**} denotes the largest real root of γ_{**}, and the following statements are all equivalent:*
(a) $\rho = \min\{r_{AA}, r_{BB}\}$
(b) $\rho = \max\{r_{AB}, r_{BA}\}$
(c) $\phi_{AA}(z)$ *or* $\phi_{BB}(z) = z^{m-1}$, *and* $\phi_{AB}(z)$ *or* $\phi_{BA}(z) = z^{m-2} + z^{m-3} + \ldots + 1$
(d) $\{A, B\}$ *or* $\{\overline{A}, \overline{B}\} = \{10^{m-1}, 0^m\}$, $\{10^{m-1}, 0^{m-1}1\}$ *or* $\{0^m, 0^{m-1}1\}$.
($\overline{A}, \overline{B}$ are the sequences obtained by complementing each bit of A, B.)

Remark: If ϕ_{AB} (or ϕ_{BA}) $\equiv 0$, then γ_{AB} (or γ_{BA}) $\equiv 1$, in which case we arbitrarily define r_{AB} (or r_{BA}) to be 0.

Proof: Observe first that since the correlations $A \circ A$ and $B \circ B$ begin with 1, while $A \circ B$ and $B \circ A$ begin with 0, we have $\phi_{AA}(2) - \phi_{AB}(2) \geq 1$ and $\phi_{BB}(2) - \phi_{BA}(2) \geq 1$. Hence, for any $z \geq 2$, we have $D_{AB}(z) \geq 2 > 0$, so that all the real roots of D_{AB} must be less than 2. Recall that $\Delta_{AB}(z) = (z-2)D(z) = \gamma_{AA}\gamma_{BB} - \gamma_{AB}\gamma_{BA}$.

Our first goal is to show $\rho \geq \max\{r_{AB}, r_{BA}\}$. Since $\deg(\phi_{AA}), \deg(\phi_{BB}) \geq 4$, Lemma 6 shows that r_{AA} and r_{BB} are the unique roots of g_{AA} and g_{BB} in (1,2). We first consider the case when $\max\{r_{AB}, r_{BA}\} \in (1,2)$. By Lemma 6, this is the case when $\max\{\deg(\phi_{AB}), \deg(\phi_{BA})\} \geq 2$. It can also be verified that this is the case when either $\phi_{AB}(z)$ or $\phi_{BA}(z)$ is $z+1$. Without loss of generality, suppose $r_{AB} \geq r_{BA}$. Note that by Lemma 4, $r_{AB} \leq r_{AA}$, r_{BB}, and hence by Lemma 6, we must have $g_{AA}(r_{AB}) \leq 0$. A similar argument shows that $g_{BB}(r_{AB}) \leq 0$. Therefore, $\Delta_{AB}(r_{AB}) \geq 0$, which implies that $D_{AB}(r_{AB}) \leq 0$, and since $D_{AB}(2) \geq 2 > 0$, we must have $\rho \geq r_{AB}$.

It remains to consider the case when the possible choices for $\phi_{AB}(z)$ and $\phi_{BA}(z)$

are 0, 1 and z. In all these cases, we have $0 < g_{AB}(1.5)g_{BA}(1.5) \leq 1$. Now,

$$g_{AA}(z) \leq (z-2)z^{m-1} + 1 = -0.5(1.5)^{m-1} + 1$$

for $z = 1.5$. Hence for $m \geq 5$, $g_{AA}(1.5) < -1$, and similarly, $g_{BB}(1.5) < -1$, so that $g_{AA}(1.5)g_{BB}(1.5) > 1$, Therefore, $\Delta_{AB}(1.5) > 0$, which shows that $D_{AB}(1.5) < 0$, and hence $\rho > 1.5$.

We now proceed to show that $\rho \leq \min\{r_{AA}, r_{BB}\}$. Without loss of generality, assume $r_{AA} \leq r_{BB}$. In the region $2 > z > r_{BB} = \max\{r_{AA}, r_{BB}, r_{AB}, r_{BA}\}$, all the γ's are positive. Moreover, Lemma 4 shows that for any z in this region, $\gamma_{AA}(z) < \gamma_{AB}(z)$ and $\gamma_{BB}(z) < \gamma_{BA}(z)$. Therefore, $\Delta_{AB}(z) < 0$ for all $z \in (r_{BB}, 2)$ which means that $\rho \notin (r_{BB}, 2)$.

If $r_{AA} < r_{BB}$ and $r_{AA} < z \leq r_{BB}$, then we must have $\gamma_{AA}(z)$, $\gamma_{AB}(z)$, $\gamma_{BA}(z) > 0$, and $\gamma_{BB}(z) \leq 0$. As a result, $\Delta_{AB}(z) < 0$ in this region, which means that $\rho \notin (r_{AA}, r_{BB}]$. Hence, $\rho \leq r_{AA}$.

It only remains to show that the statements (a), (b), (c) and (d) are all equivalent to one another. We first show that $(a) \Rightarrow (b)$. Let $\rho = \min\{r_{AA}, r_{BB}\}$ which means that either $\gamma_{AA}(\rho)$ or $\gamma_{BB}(\rho)$ is 0. Therefore, $0 = \Delta_{AB}(\rho) = -\gamma_{AB}(\rho)\gamma_{BA}(\rho)$. Thus, we must have $\rho = r_{AB}$ or r_{BA}. In either case, $\max\{r_{AB}, r_{BA}\} \geq \rho$. The reverse inequality is trivial since $\max\{r_{AB}, r_{BA}\} \leq \min\{r_{AA}, r_{BB}\}$.

$(b) \Rightarrow (a)$ is proved by a similar argument.

We next show that $(a) \Leftrightarrow (c)$. Note first that $\rho = \min\{r_{AA}, r_{BB}\}$ iff r_{AA} or r_{BB} is a root of Δ_{AB}. Since $\Delta_{AB}(r_{AA}) = -\gamma_{AB}(r_{AA})\gamma_{BA}(r_{AA})$, we see that r_{AA} is a root of Δ_{AB} iff r_{AB} or $r_{BA} = r_{AA}$. But by Lemma 4, $r_{AA} = r_{AB}$ or r_{BA} iff $\phi_{AA}(z) = z^{m-1}$ and $\phi_{AB}(z)$ or $\phi_{BA}(z) = z^{m-2} + z^{m-3} + \ldots + 1$. A similar argument shows that r_{BB} is a root of Δ_{AB} iff $\phi_{BB}(z) = z^{m-1}$ and $\phi_{AB}(z)$ or $\phi_{BA}(z) = z^{m-2} + z^{m-3} + \ldots + 1$.

Finally, we show that $(c) \Leftrightarrow (d)$. Now, $A \circ B = 01^{m-1}$ (*i.e.*, $\phi_{AB}(z) = z^{m-2} + z^{m-3} + \ldots + 1$) can happen if and only if the longest proper suffix of A is either 0^{m-1} or 1^{m-1}, and is the same as the longest proper prefix of B. Moreover, $A \circ A = 10^{m-1}$ (*i.e.*, $\phi_{AA}(z) = z^{m-1}$) implies that the first and last bits of A are different. Therefore, it easily follows that the correlations listed above can arise if and only if $\{A, B\}$ or $\{\overline{A}, \overline{B}\}$ is one of the sequence pairs listed in the statement of the proposition. □

Observe that in the above proof, the fact that the polynomials ϕ_{AA}, ϕ_{BB}, ϕ_{AB} and ϕ_{BA} are correlation polynomials for certain sequences is only used in showing that the statement (d) is equivalent to (c). The rest of the proof continues to work even if we only assume that ϕ_{AA}, ϕ_{BB}, ϕ_{AB} and ϕ_{BA} are polynomials with coefficients 0 and 1, with $\deg(\phi_{AA}) = \deg(\phi_{BB}) = m - 1 \geq 4$, and $\deg(\phi_{AB}), \deg(\phi_{BA}) < m - 1$. Therefore, the conclusions of the theorem, apart from statement (d), remain valid for *any* set of four polynomials ϕ_{AA}, ϕ_{BB}, ϕ_{AB} and ϕ_{BA} that satisfy the properties listed above. We will, in fact, utilize this observation later.

The following corollary is the first important consequence of the previous theorem.

COROLLARY 8. *For $m \geq 5$, the largest pole (in terms of absolute value) of Q_{AB} in (2) is precisely the largest positive real root of D_{AB}.*

Proof: We have already seen earlier that the largest-magnitude pole of Q_{AB} is real and positive. Note that all the poles of Q_{AB} must be roots of D_{AB}. Suppose that the largest positive real root ρ of D_{AB} is not a pole of Q_{AB}. Then, ρ must be a root of the numerator polynomial of Q_{AB}, *i.e.*, we must have $\phi_{AA}(\rho)\phi_{BB}(\rho) = \phi_{AB}(\rho)\phi_{BA}(\rho)$. But then, since $D_{AB}(\rho) = 0$, we also have $\phi_{AA}(\rho) + \phi_{BB}(\rho) = \phi_{AB}(\rho) + \phi_{BA}(\rho)$.

Now, if we have real numbers a, b, c and d such that $a+b=c+d$ and $ab=cd$, then the polynomials $(z-a)(z-b)$ and $(z-c)(z-d)$ must be identical. This implies that $\{a,b\}=\{c,d\}$. Thus, we have $\{\phi_{AA}(\rho),\phi_{BB}(\rho)\}=\{\phi_{AB}(\rho),\phi_{BA}(\rho)\}$, and hence, $\{\gamma_{AA}(\rho),\gamma_{BB}(\rho)\}=\{\gamma_{AB}(\rho),\gamma_{BA}(\rho)\}$.

By the previous theorem, $\rho\in(1,2)$ and $\max\{r_{AB},r_{BA}\}\le\rho\le\min\{r_{AA},r_{BB}\}$. Therefore, by Lemma 6, we must have $\gamma_{AA}(\rho)$, $\gamma_{BB}(\rho)\le 0$, and $\gamma_{AB}(\rho)$, $\gamma_{BA}(\rho)\ge 0$. Hence, $\{\gamma_{AA}(\rho),\gamma_{BB}(\rho)\}=\{\gamma_{AB}(\rho),\gamma_{BA}(\rho)\}$ iff all of them are 0, *i.e.*, $\rho=r_{AA}=r_{BB}=r_{AB}=r_{BA}$. But by Lemma 4, this is possible iff the correlations $A\circ A$ and $B\circ B$ are 10^{m-1}, and $A\circ B$ and $B\circ A$ are 01^{m-1}. However, it is easily seen that no pair of sequences A and B can have this set of correlations, leading to a contradiction that proves the result. □

From now on, we shall denote by ρ_{AB} the largest-magnitude pole of Q_{AB}, which (at least for $m\ge 5$) is also the largest positive root of D_{AB}. Since $\rho_{AB}>1$ for $m\ge 5$, ρ_{AB} must be a simple pole of $Q_{AB}(z)$, and hence is a simple root of $D_{AB}(z)$. Using the fact that $H(A,B)=\log_2\rho_{AB}$, we now identify a pair of sequences that *minimizes* $H(A,B)$ among all pairs of binary m-sequences $\{A,B\}$.

PROPOSITION 9. *For $m\ge 5$, $\min\{H(A,B):\, A,B\in\{0,1\}^m, A\ne B\}$ is achieved by $A=110^{m-2}$, $B=110^{m-4}10$.*

Proof: We shall show that if $\widehat{A},\widehat{B}$ is any pair of binary m-sequences, and $A=110^{m-2}$, $B=110^{m-4}10$, then $\rho_{\hat{A}\hat{B}}\ge\rho_{AB}$. Observe first that $\phi_{AA}(z)=\phi_{BB}(z)=z^{m-1}$, and $\phi_{AB}=\phi_{BA}\equiv 0$. Thus, $\Delta_{AB}={\gamma_{AA}}^2-1$. Since $\rho_{AB}\le r_{AA}$, and r_{AA} is the unique root of γ_{AA} in (1,2), we must have $\gamma_{AA}(\rho_{AB})\le 0$.

Now, for any pair of m-sequences $\widehat{A},\widehat{B}$, since the correlations $\widehat{A}\circ\widehat{A}$ and $\widehat{B}\circ\widehat{B}$ must always begin with 1, we have $\phi_{\hat{A}\hat{A}}(z)$, $\phi_{\hat{B}\hat{B}}(z)\ge z^{m-1}=\phi_{AA}(z)$ for all $z\ge 1$. Therefore, $\gamma_{\hat{A}\hat{A}}(z)$, $\gamma_{\hat{B}\hat{B}}(z)\le\gamma_{AA}(z)$ for all $z\in[1,2]$. In particular, we have $\gamma_{\hat{A}\hat{A}}(\rho_{AB})$, $\gamma_{\hat{B}\hat{B}}(\rho_{AB})\le\gamma_{AA}(\rho_{AB})\le 0$. Therefore, $\gamma_{\hat{A}\hat{A}}(\rho_{AB})\gamma_{\hat{B}\hat{B}}(\rho_{AB})\ge(\gamma_{AA}(\rho_{AB}))^2$.

On the other hand, since $\phi_{\hat{A}\hat{B}}$, $\phi_{\hat{B}\hat{A}}\ge 0$, we see that $\gamma_{\hat{A}\hat{B}}(z)$, $\gamma_{\hat{B}\hat{A}}(z)\le 1$ for all $z\in[1,2]$, and in particular at $z=\rho_{AB}$. Therefore,

$$\Delta_{\hat{A}\hat{B}}(\rho_{AB})\ge(\gamma_{AA}(\rho_{AB}))^2-1=\Delta_{AB}(\rho_{AB})=0$$

which shows that $D_{\hat{A}\hat{B}}(\rho_{AB})\le 0$. Since $D_{\hat{A}\hat{B}}(2)\ge 2>0$, we have $\rho_{\hat{A}\hat{B}}\ge\rho_{AB}$. □

The next lemma, which yields a lower bound to the minimum value of $H(A,B)$, is crucial to the proof of the important theorem that follows it.

LEMMA 10. *For $A=110^{m-2}$, $B=110^{m-4}10$, $H(A,B)>\log_2\rho_{m-3}$, where ρ_{m-3} is the largest zero of $z^{m-3}-z^{m-4}-\ldots-1$.*

Proof: With A,B as above, we have

$$\Delta_{AB}(z)=[(z-2)z^{m-1}+1]^2-1=(z-2)z^{m-1}[(z-2)z^{m-1}+2]$$

using $a^2-b^2=(a+b)(a-b)$. Therefore, the largest positive zero ρ_{AB} of D_{AB} is the largest positive zero of the polynomial $p(z)=(z-2)z^{m-1}+2$. Observe that $p(z)=(z-1)(z^{m-1}-z^{m-2}-\ldots-1)+1=(z-1)[z^2(z^{m-3}-z^{m-4}-\ldots-1)-z-1]+1$. Therefore,

$$p(\rho_{m-3})=-(\rho_{m-3}-1)(\rho_{m-3}+1)+1$$

$$\begin{aligned} &= 2 - (\rho_{m-3})^2 \leq 2 - (\rho_2)^2 \\ &= 2 - \left(\frac{1+\sqrt{5}}{2}\right)^2 < 0 \end{aligned}$$

the first inequality arising from the fact that $\rho_{m-3} \geq \rho_2$ for $m \geq 5$. Since $p(2) = 2 > 0$, we must have $\rho_{AB} > \rho_{m-3}$. □

At this point, we introduce a means of comparing two correlation sequences, which will make it easier to comprehend our next result which is a theorem of fundamental importance. Given two correlation sequences $A \circ B = (c_0 c_1 \ldots c_{m-1})$ and $\widehat{A} \circ \widehat{B} = (\hat{c}_0 \hat{c}_1 \ldots \hat{c}_{m-1})$ of the same length, we say that $\widehat{A} \circ \widehat{B}$ is *stronger* than $A \circ B$, and denote it by $\widehat{A} \circ \widehat{B} > A \circ B$, if $\hat{c}_i > c_i$ for the smallest i such that $\hat{c}_i \neq c_i$ (this is simply a lexicographic ordering of the correlation sequences). Equivalently, $\widehat{A} \circ \widehat{B} > A \circ B$ iff $\phi_{\hat{A}\hat{B}}(2) > \phi_{AB}(2)$. We also define $\widehat{A} \circ \widehat{B} \geq A \circ B$ in the obvious way. We now show that if all the correlations between sequences $\widehat{A}$ and $\widehat{B}$ are stronger than the corresponding correlations between sequences A and B, then $H(\widehat{A}, \widehat{B}) \geq H(A, B)$.

THEOREM 11. *Let $A, B, \widehat{A}, \widehat{B}$ be binary sequences of length $m \geq 5$ such that $\widehat{A} \circ \widehat{A} \geq A \circ A$, $\widehat{B} \circ \widehat{B} \geq B \circ B$, $\widehat{A} \circ \widehat{B} \geq A \circ B$ and $\widehat{B} \circ \widehat{A} \geq B \circ A$. Then $H(\widehat{A}, \widehat{B}) \geq H(A, B)$, with equality iff all the above correlation inequalities hold with equality.*

Instead of directly proving this theorem, we shall find it easier to prove a more general result, which yields the above theorem as a special case. The more general result is easier to state if we introduce the following definition.

Definition: A quadruple of polynomials $(\phi_1, \phi_2, \phi_3, \phi_4)$, each ϕ_i having coefficients 0 and 1, is called an *admissible m-quadruple* if $\deg(\phi_1) = \deg(\phi_2) = m$ and $\deg(\phi_3), \deg(\phi_4) < m$.

Observe that if A, B are binary m-sequences, then $(\phi_{AA}, \phi_{BB}, \phi_{AB}, \phi_{BA})$ is an admissible $(m-1)$-quadruple. As observed previously, Theorem 7 is essentially a result on the largest positive root ρ of the polynomial

$$D(z) = (z-2)(\phi_1\phi_2 - \phi_3\phi_4) + \phi_1 + \phi_2 - \phi_3 - \phi_4 \tag{7}$$

where $(\phi_1, \phi_2, \phi_3, \phi_4)$ is an admissible m-quadruple. The proof of that theorem shows that for $m \geq 4$, ρ lies in (1,2) and $\max\{r_3, r_4\} \leq \rho \leq \min\{r_1, r_2\}$, where r_i is the largest positive root of γ_i, the γ-polynomial associated with ϕ_i. Moreover, the equivalence of the corresponding statements (a), (b) and (c) is also established. Furthermore, it clearly follows from the proofs of Proposition 9 and Lemma 10 that for $m \geq 4$, $\rho > \rho_{m-2}$ (ρ_{m-2} being the largest zero of the polynomial $z^{m-2} - \ldots - 1$). These facts will be needed to prove our next result, which covers Theorem 11 as a special case.

THEOREM 12. *Let $(\phi_1, \phi_2, \phi_3, \phi_4)$ and $(\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \hat{\phi}_4)$ be admissible m-quadruples, $m \geq 4$, such that $\phi_i(2) \leq \hat{\phi}_i(2)$ for $i = 1, 2, 3, 4$. Let D and $\widehat{D}$ be the corresponding polynomials defined via (7), and let ρ and $\hat{\rho}$ be their respective largest positive roots. Then $\rho \leq \hat{\rho}$, with equality iff either $\phi_i(2) = \hat{\phi}_i(2)$ for $i = 1, 2, 3, 4$, or $\phi_i(z) = \hat{\phi}_i(z) =$*

z^m *for* $i = 1$ *or* 2 *and* $\phi_i(z) = \hat{\phi}_i(z) = z^{m-1} + z^{m-2} + \ldots + 1$ *for* $i = 3$ *or* 4.

Proof: It should be clear that the following four propositions, when patched together, yield the theorem:

1. If $\phi_1(2) < \hat{\phi}_1(2)$ and $\phi_i(2) = \hat{\phi}_i(2)$ for $i = 2, 3, 4$, then $\rho \leq \hat{\rho}$ with equality iff $\phi_2(z) = \hat{\phi}_2(z) = z^m$ and $\phi_i(z) = \hat{\phi}_i(z) = z^{m-1} + \ldots + 1$ for $i = 3$ or 4.
2. If $\phi_2(2) < \hat{\phi}_2(2)$ and $\phi_i(2) = \hat{\phi}_i(2)$ for $i = 1, 3, 4$, then $\rho \leq \hat{\rho}$ with equality iff $\phi_1(z) = \hat{\phi}_1(z) = z^m$ and $\phi_i(z) = \hat{\phi}_i(z) = z^{m-1} + \ldots + 1$ for $i = 3$ or 4.
3. If $\phi_3(2) < \hat{\phi}_3(2)$ and $\phi_i(2) = \hat{\phi}_i(2)$ for $i = 1, 2, 4$, then $\rho \leq \hat{\rho}$ with equality iff $\phi_i(z) = \hat{\phi}_i(z) = z^m$ for $i = 1$ or 2 and $\phi_4(z) = \hat{\phi}_4(z) = z^{m-1} + \ldots + 1$.
4. If $\phi_4(2) < \hat{\phi}_4(2)$ and $\phi_i(2) = \hat{\phi}_i(2)$ for $i = 1, 2, 3$, then $\rho \leq \hat{\rho}$ with equality iff $\phi_i(z) = \hat{\phi}_i(z) = z^m$ for $i = 1$ or 2 and $\phi_3(z) = \hat{\phi}_3(z) = z^{m-1} + \ldots + 1$.

We shall prove the first proposition alone, as the other propositions can be proved analogously. For the proof of the first proposition, we assume that $\hat{\phi}_1(2) = \phi_1(2) + 1$ and $\phi_i(2) = \hat{\phi}_i(2)$ for $i = 2, 3, 4$. The general case follows by induction on $\hat{\phi}_1(2) - \phi_1(2)$. As usual, note that $D(2), \widehat{D}(2) \geq 2 > 0$. Let $\gamma_i, \hat{\gamma}_i$ be the γ-polynomials corresponding to $\phi_i, \hat{\phi}_i$, $i = 1, 2, 3, 4$, and define the polynomials Δ and $\widehat{\Delta}$ analogous to (6).

We first prove the proposition based on the claim that $\hat{\gamma}_1(\rho) < \gamma_1(\rho)$, deferring the proof of this claim until later. Note that since ρ cannot exceed the largest positive root, r_2, of γ_2, Lemma 6 shows that $\gamma_2(\rho) \leq 0$ with equality iff $\rho = r_2$. Therefore, $\hat{\gamma}_1(\rho)\gamma_2(\rho) \geq \gamma_1(\rho)\gamma_2(\rho)$, and hence,

$$\begin{aligned} \widehat{\Delta}(\rho) &= \hat{\gamma}_1(\rho)\hat{\gamma}_2(\rho) - \hat{\gamma}_3(\rho)\hat{\gamma}_4(\rho) = \hat{\gamma}_1(\rho)\gamma_2(\rho) - \gamma_3(\rho)\gamma_4(\rho) \\ &\geq \gamma_1(\rho)\gamma_2(\rho) - \gamma_3(\rho)\gamma_4(\rho) = \Delta(\rho) = (\rho - 2)D(\rho) = 0 \end{aligned}$$

with equality holding iff $\rho = r_2$. Hence, $\widehat{D}(\rho) \leq 0$ which implies that $\hat{\rho} \geq \rho$, with equality iff $\rho = r_2$. Now, as shown in the proof of Theorem 7, $\rho = r_2$ iff $\phi_2(z) = z^m$ and $\phi_3(z)$ or $\phi_4(z) = z^{m-1} + z^{m-2} + \ldots + 1$.

It only remains to prove the claim that $\hat{\gamma}_1(\rho) < \gamma_1(\rho)$. As observed prior to the statement of the theorem, $\rho_{m-2} < \rho < 2$ for $m \geq 4$. Let k be the largest integer such that the coefficients of z^k in ϕ_1 and $\hat{\phi}_1$ are different. If $k = 0$, then Lemma 3 shows that $\hat{\gamma}_1(\rho) < \gamma_1(\rho)$ as $\rho < 2$. If $1 \leq k \leq m - 2$, then Lemma 3 again shows that $\hat{\gamma}_1(\rho) < \gamma_1(\rho)$, since $\rho > \rho_{m-2} \geq \rho_k$.

The case $k = m - 1$ arises only when $\phi_1(z) = z^m + z^{m-2} + z^{m-3} + \ldots + 1$ and $\hat{\phi}_1(z) = z^m + z^{m-1}$. Therefore, it will suffice to show that for any admissible m-quadruple $(\phi_1, \phi_2, \phi_3, \phi_4)$ with $\phi_1(z) = z^m + z^{m-2} + z^{m-3} + \ldots + 1$, the corresponding ρ exceeds ρ_{m-1}. Now, an argument similar to the proof of Proposition 9 can be used to show that the ρ corresponding to such an admissible m-quadruple is at least as large as the ρ corresponding to the quadruple $(\phi_1(z), z^m, 0, 0)$, with ϕ_1 as above. Therefore, it is sufficient to show that when ϕ_1 is as above, $\phi_2(z) = z^m$ and $\phi_3 = \phi_4 \equiv 0$, then the largest positive zero of $D(z)$ exceeds ρ_{m-1}. In this case, we have

$$\gamma_2(z) = (z - 2)z^m + 1 = (z - 1)\left[z(z^{m-1} - z^{m-2} - \ldots - 1) - 1\right]$$

which shows that $\gamma_2(\rho_{m-1}) = (\rho_{m-1} - 1)(-1)$. We also have

$$\begin{aligned} \gamma_1(z) &= (z - 2)z^m + (z - 2)(z^{m-2} + \ldots + 1) + 1 \\ &= \gamma_2(z) - 1 + (z^{m-1} - z^{m-2} - \ldots - 1) \end{aligned}$$

which means that $\gamma_1(\rho_{m-1}) = -\rho_{m-1}$. Therefore,

$$\Delta(\rho_{m-1}) = \rho_{m-1}(\rho_{m-1} - 1) - 1 = (\rho_{m-1})^2 - \rho_{m-1} - 1$$

which is strictly positive for $m \geq 4$, as $z^2 - z - 1 > 0$ for $z > \rho_2$. This means that $D(\rho_{m-1}) < 0$, and so $\rho > \rho_{m-1}$, thus concluding the proof of the claim and hence the theorem. $\square$

As mentioned previously, Theorem 11 is a special case of Theorem 12. Therefore, the proof of Theorem 11 will be complete if we show that under its hypotheses, we can have a situation where $\widehat{A} \circ \widehat{A}$ or $\widehat{B} \circ \widehat{B} = 10^{m-1}$ and $A \circ B$ or $B \circ A = 01^{m-1}$, only if all the correlation inequalities are satisfied with equality.

Consider the case when $\widehat{A} \circ \widehat{A} = 10^{m-1}$ and $A \circ B = 01^{m-1}$ (we can dispose of the other cases similarly). For $\widehat{A} \circ \widehat{A} \geq A \circ A$ and $\widehat{A} \circ \widehat{B} \geq A \circ B$ to be true, we must have $A \circ A = 10^{m-1}$ and $\widehat{A} \circ \widehat{B} = 01^{m-1}$ as well. But now, we must have (up to complementation of A, B or $\widehat{A}, \widehat{B}$) $A = 10^{m-1}$, $B = 0^{m-1}b$, $\widehat{A} = 10^{m-1}$ and $\widehat{B} = 0^{m-1}\hat{b}$, for some $b, \hat{b} \in \{0, 1\}$. It is easily verified that if $b \neq \hat{b}$, then either $B \circ B > \widehat{B} \circ \widehat{B}$ or $B \circ A > \widehat{B} \circ \widehat{A}$, both of which contradict the hypotheses of the theorem. This completes the proof of Theorem 11.

Having Theorem 11 in hand, we are in a position to begin our search for the binary m-sequences A and B that maximize $H(A, B)$. At this point, it should be noted that if there existed A and B such that $A \circ A = B \circ B = 1^m$ and $A \circ B = B \circ A = 01^{m-1}$, then Theorem 11 would imply that A and B are the sequences for which $H(A, B)$ is a maximum. However, it is a simple exercise to show that no pair of sequences can have these correlations.

In order to prove our next important result, which reduces the search space significantly, we need a couple of preliminary lemmas. Recall that $\langle 01 \rangle_m$ and $\langle 10 \rangle_m$ denote the two length-m sequences of alternating 0's and 1's.

LEMMA 13. *The only length-m autocorrelation sequence $A \circ A$ stronger than $\langle 10 \rangle_m$ is 1^m.*

Proof: If $(b_0 b_1 \ldots b_{m-1})$ is an autocorrelation sequence (so that $b_0 = 1$), then it is easily seen that whenever $b_j = 1$ for some $j \in [1, m-1]$, then $b_k = 1$ for any $k \in [1, m-1]$ that is a multiple of j. Moreover, the GCD rule for autocorrelation sequences ([6], Theorem 3.1) states that if $b_j = b_k = 1$ for some $j, k \in [1, m-1]$ such that $j + k \leq m + l$, where $l = \gcd(j, k)$, then $b_l = 1$ as well.

Note that any autocorrelation sequence $(b_0 b_1 \ldots b_{m-1})$ stronger than $\langle 10 \rangle_m$ must have either $b_1 = 1$ or $b_2 = 1$. In the first case, we must have $b_k = 1$ for all $k \leq m - 1$. Thus, the only autocorrelation sequence with $b_1 = 1$ is 1^m. On the other hand, if $b_2 = 1$, then $b_i = 1$ for all even i. In addition, if $b_k = 1$ for some odd k, then by the GCD rule applied to the pair of indices $(2, k)$, we must have $b_1 = 1$, and hence $b_k = 1$ for all $k \leq m - 1$. Thus, any autocorrelation sequence with $b_2 = 1$ must either be $\langle 10 \rangle_m$ or 1^m. $\square$

LEMMA 14. *The only correlation sequence $A \circ B$, between distinct binary m-sequences A and B, that is stronger than $\langle 01 \rangle_m$ is 01^{m-1}.*

Proof: Let $A \circ B = (c_0 c_1 \ldots c_{m-1})$, with $c_0 = 0$. If $A \circ B > \langle 01 \rangle_m$, then c_1 must be 1. Therefore, it follows that $(c_1 \ldots c_{m-1})$ must be the autocorrelation sequence for B', where B' is the sequence obtained from B by deleting its last bit. As a result, for $A \circ B > \langle 01 \rangle_m$ to be true, we must have $B' \circ B' > \langle 10 \rangle_{m-1}$ which, by the previous lemma, is possible only if $B' \circ B' = 1^{m-1}$. □

We now have the requisite tools to prove a result that considerably simplifies the problem of finding the pair of m-sequences (A, B) that maximizes $H(A, B)$. Recall that $H_{2,m} = \max\{H(A, B) : A, B \in \{0, 1\}^m, A \neq B\}$

Proposition 15. *For $m \geq 5$, if $A, B \notin \{0^m, 1^m\}$, then $H(A, B) \leq \log_2 \rho_{m-1}$ with equality iff $\{A, B\} = \{\langle 01\rangle_m, \langle 10\rangle_m\}$, $\{01^{m-1}, 1^{m-1}0\}$ or $\{10^{m-1}, 0^{m-1}1\}$. Consequently, $H_{2,m} = \max\{H(1^m, B) : B \in \{0, 1\}^m, B \neq 1^m\}$.*

Proof: Fix $\widehat{A} = \langle 01\rangle_m$, $\widehat{B} = \langle 10\rangle_m$, so that $\widehat{A} \circ \widehat{A} = \widehat{B} \circ \widehat{B} = \langle 10\rangle_m$ and $\widehat{A} \circ \widehat{B} = \widehat{B} \circ \widehat{A} = \langle 01\rangle_m$. It is easily verified using (2) that

$$Q_{\hat{A}\hat{B}}(z) = \frac{z^{m-1} + z^{m-2} + \ldots + 1}{z^{m-1} - z^{m-2} - \ldots - 1}$$

and hence $H(\widehat{A}, \widehat{B}) = \log_2 \rho_{m-1}$. Therefore, we have $H_{2,m} \geq \log_2 \rho_{m-1}$.

Let $A, B \notin \{0^m, 1^m\}$ be a pair of distinct binary m-sequences. Then, we either have $\widehat{A} \circ \widehat{A} \geq A \circ A$, $\widehat{B} \circ \widehat{B} \geq B \circ B$, $\widehat{A} \circ \widehat{B} \geq A \circ B$ and $\widehat{B} \circ \widehat{A} \geq B \circ A$, or at least one of these correlation inequalities is not satisfied. In the former case, Theorem 11 shows that $H(\widehat{A}, \widehat{B}) \geq H(A, B)$ with equality iff $A \circ A = B \circ B = \langle 10\rangle_m$ and $A \circ B = B \circ A = \langle 01\rangle_m$, which can happen iff $\{A, B\} = \{\langle 01\rangle_m, \langle 10\rangle_m\}$.

We next deal with the case when $A \circ B > \widehat{A} \circ \widehat{B}$ which, by Lemma 14, means that $A \circ B = 01^{m-1}$. Therefore, up to complementation of A and B, we must have $\{A, B\} = \{01^{m-1}, 1^{m-1}0\}$, $\{01^{m-1}, 1^m\}$ or $\{1^m, 1^{m-1}0\}$. Our assumption that $A, B \notin \{0^m, 1^m\}$ eliminates the last two sequence pairs, along with their complements. When $\{A, B\} = \{01^{m-1}, 1^{m-1}0\}$, it can be verified that $D_{AB}(z) = (z-1)z^{m-1}(z^{m-1} - z^{m-2} - \ldots - 1)$, so that $H(A, B) = \log_2 \rho_{m-1} = H(\widehat{A}, \widehat{B})$. This shows that when $A \circ B > \widehat{A} \circ \widehat{B}$, we must have $H(A, B) = H(\widehat{A}, \widehat{B}) = \log_2 \rho_{m-1}$. The case when $B \circ A > \widehat{B} \circ \widehat{A}$ is similar, and leads to the same conclusion.

We are left with the case when we have $A \circ A > \widehat{A} \circ \widehat{A}$ or $B \circ B > \widehat{B} \circ \widehat{B}$, so that by Lemma 13, either $A \circ A$ or $B \circ B$ is 1^m, *i.e.*, A or $B = 1^m$ or 0^m. But this is not possible, as we assumed that $A, B \notin \{0^m, 1^m\}$. Therefore, one of the previously considered cases must hold, and hence $H(A, B) \leq \log_2 \rho_{m-1}$ with equality iff $\{A, B\}$ is one of the pairs listed in the statement of the proposition.

Finally, it is easily verified that $Q_{0^m 1^m}(z) = Q_{\hat{A}\hat{B}}(z)$, and hence $H(0^m, 1^m) = H(\widehat{A}, \widehat{B}) = \log_2 \rho_{m-1}$. Therefore, if the inequality $H_{2,m} \geq \log_2 \rho_{m-1}$ is actually an equality, then one of the maximizing pairs $\{A, B\}$ contains 1^m. On the other hand, if this inequality is strict, then any of the sequence pairs that achieve the maximum must include either 1^m or 0^m. But since $H(A, B) = H(\overline{A}, \overline{B})$, where $\overline{A}$ and $\overline{B}$ are the sequences obtained by complementing each bit of A and B, at least one of the maximizing pairs includes 1^m, which concludes the proof of the proposition. □

We have thus reduced the problem of maximizing $H(A, B)$ to the problem of finding the sequence $B \neq 1^m$ that maximizes $H(1^m, B)$. We tackle this problem by considering the following two cases separately: (i) B begins or ends with a 0, and (ii) B begins and ends with a 1. In fact, as explained below, it is possible to reduce the search space even further in each of these cases.

Given a sequence $A = (a_1 a_2 \ldots a_{n-1} a_n)$, let A^R denote the sequence obtained by reversing A, *i.e.*, $A^R = (a_n a_{n-1} \ldots a_2 a_1)$. It is clear that if Z is a sequence counted by $q_n(A,B)$ for some A, B, then Z^R is a sequence counted by $q_n(A^R, B^R)$. Therefore, we must have $H(A,B) = H(A^R, B^R)$. This conclusion can also be reached from the observation that $A \circ B = B^R \circ A^R$. In particular, $H(1^m, B) = H(1^m, B^R)$. Thus, if B has a longer run of ones at the end than at the beginning, then the situation is reversed for B^R, but the resulting Shannon capacity is the same in both cases. As a result, for case (i), it suffices to consider only those sequences B that end with a 0, and for case (ii), it is enough to consider sequences that begin with a run of ones that is at least as long as the final run of ones. We deal with case (i) first.

LEMMA 16. *If B is a binary sequence of length $m \geq 5$ that begins or ends with a 0, then $H(1^m, B) \leq \log_2 \rho_{m-1}$ with equality iff $B = 1^{m-1}0$, 01^{m-1} or 0^m.*

Proof: It suffices to show that if B ends with a 0, then $H(1^m, B) \leq \log_2 \rho_{m-1}$ with equality iff $B = 1^{m_1}0$ or 0^m. Suppose that $B = 1^{m_1}\underline{b}0$, where $\underline{b}$ is a binary sequence of length $m - m_1 - 1$ that begins with a 0, and $0 \leq m_1 \leq m-1$. Setting $A = 1^m$, we see that $A \circ A = 1^m$, $A \circ B = 0^{m-m_1}1^{m_1}$ and $B \circ A = 0^m$. Moreover, $B \circ B$ must end in a run of m_1 zeros, because that is when, in the procedure used to determine $B \circ B$, one of the initial m_1 1's in B overlaps with the final 0. Thus, we have $\phi_{AA}(z) = z^{m-1} + z^{m-2} + \ldots + 1$, $\phi_{AB}(z) = z^{m_1-1} + z^{m_1-2} + \ldots + 1$, $\phi_{BA} \equiv 0$, and $\phi_{BB}(z) = z^{m-1} + z^{k_1} + z^{k_2} + \ldots + z^{k_r}$ for some $k_1, k_2, \ldots, k_r \geq m_1$.

Using the fact that $\gamma_{BA} \equiv 1$, we see that

$$\begin{aligned}
\Delta_{AB}(z) &= \gamma_{AA}\gamma_{BB} - \gamma_{AB} = \gamma_{AA}\left[(z-2)\phi_{BB} + 1\right] - \gamma_{AB} \\
&= (z-2)\gamma_{AA}\,\phi_{BB} + \gamma_{AA} - \gamma_{AB} \\
&= (z-2)\left[\gamma_{AA}\,\phi_{BB} + (\phi_{AA} - \phi_{AB})\right]
\end{aligned}$$

Therefore, $D_{AB} = \gamma_{AA}\,\phi_{BB} + (\phi_{AA} - \phi_{AB})$. Now, $\phi_{AA}(z) - \phi_{AB}(z) = z^{m-1} + z^{m-2} + \ldots + z^{m_1} = \phi_{BB}(z) + \phi(z)$ for some polynomial ϕ with coefficients 0 and 1, since the coefficient of z^k, $0 \leq k \leq m_1 - 1$, in $\phi_{BB}(z)$ is zero. Hence, we can write

$$D_{AB} = (\gamma_{AA} + 1)\phi_{BB} + \phi$$

Now, $\gamma_{AA}(z) = z^m - z^{m-1} - \ldots - 1 = z(z^{m-1} - z^{m-2} - \ldots - 1) - 1$. Therefore, $\gamma_{AA}(z) \geq -1$ for all $z \geq \rho_{m-1}$, with equality iff $z = \rho_{m-1}$. Hence for any $z \geq \rho_{m-1}$, we have

$$\begin{aligned}
D_{AB}(z) &\geq \phi(z) && \text{with equality iff } z = \rho_{m-1} \\
&\geq 0 && \text{with equality iff } \phi \equiv 0
\end{aligned}$$

This shows that if $z > \rho_{m-1}$, then $D_{AB}(z) \neq 0$, since the first of the above inequalities is strict. Therefore, $\rho_{AB} \leq \rho_{m-1}$, which shows that $H(A,B) = \log_2 \rho_{AB} \leq \log_2 \rho_{m-1}$.

The above inequalities also show that $D_{AB}(\rho_{m-1}) = 0$ (*i.e.* $\rho_{AB} = \rho_{m-1}$) iff $\phi \equiv 0$. It is easily verified that with $B = 1^{m-1}0$ or 0^m, we obtain $\phi \equiv 0$. Conversely, if $\phi \equiv 0$, then we must have $\phi_{BB}(z) = z^{m-1} + z^{m-2} + \ldots + z^{m_1}$, or equivalently, $B \circ B = 1^{m-m_1}0^{m_1}$. Now, we can either have $m_1 = m-1$ or $0 \leq m_1 < m-1$. In the former case, B must be $1^{m-1}0$. In the latter case, since $B \circ B$ begins with two 1's, as shown in the proof of Lemma 13, we must have $B \circ B = 1^m$, which means that B must be 0^m. Therefore, if $\phi \equiv 0$, then B can only be $1^{m-1}0$ or 0^m. This shows that if B ends in a zero, then $H(1^m, B) = \log_2 \rho_{m-1}$ iff $B = 1^{m-1}0$ or 0^m, which concludes the proof of the lemma. □

The only case remaining is when B begins and ends with a 1. We now show that no such B distinct from 1^m can maximize $H(1^m, B)$.

LEMMA 17. *If $B \neq 1^m$ is a binary sequence of length $m \geq 5$ that begins and ends with a 1, then $H(1^m, B) < \log_2 \rho_{m-1}$.*

Proof: As observed prior to the statement of the previous lemma, it is sufficient to consider the case when B is of the form $1^{m_1}\underline{b}1^{m_2}$ with $m_1 \geq m_2 \geq 1$, where $\underline{b}$ is a binary sequence of length $m - m_1 - m_2 > 0$ that begins and ends with a 0. With $A = 1^m$, we have $A \circ A = 1^m$, $A \circ B = 0^{m-m_1}1^{m_1}$ and $B \circ A = 0^{m-m_2}1^{m_2}$. Thus, $\phi_{AA}(z) = z^{m-1} + \ldots + 1$, $\phi_{AB}(z) = z^{m_1-1} + \ldots + 1$ and $\phi_{BA}(z) = z^{m_2-1} + \ldots + 1$. Now, note that since $B \neq 1^m$, $B \circ B$ must begin with 10. Moreover, $B \circ B$ must end with $0^{m_1-m_2}1^{m_2}$, as that is when, in the procedure for determining $B \circ B$, some part of the prefix $1^{m_1}0$ of B overlaps with some part of the suffix 01^{m_2}. Thus, $\phi_{BB}(z) = \phi(z) + \phi_{BA}(z)$, where $\phi(z) = \sum_{k=0}^{m-1} c_k z^k$ is some polynomial with $c_{m-1} = 1$, $c_{m-2} = 0$, $c_k = 0$ for $0 \leq k \leq m_1 - 1$, and the remaining c_k's being either 0 or 1.

With the correlation polynomials being as above, we see that $\gamma_{AA}(z) = z^m - z^{m-1} - \ldots - 1$, $\gamma_{AB}(z) = z^{m_1} - z^{m_1-1} - \ldots - 1$, $\gamma_{BA}(z) = z^{m_2} - z^{m_2-1} - \ldots - 1$, and $\gamma_{BB}(z) = (z-2)\phi + \gamma_{BA}(z)$. Also, note that

$$\begin{aligned}\Delta_{AB}(z) &= \gamma_{AA}[(z-2)\phi + \gamma_{BA}] - \gamma_{AB}\gamma_{BA} \\ &= (z-2)\gamma_{AA}\,\phi + (\gamma_{AA} - \gamma_{AB})\gamma_{BA} \\ &= (z-2)[\gamma_{AA}\,\phi + (\phi_{AA} - \phi_{AB})\gamma_{BA}]\end{aligned}$$

which shows that $D_{AB} = \gamma_{AA}\,\phi + (\phi_{AA} - \phi_{AB})\gamma_{BA}$.

Or goal is to show that $D_{AB}(z) \neq 0$ for all $z \geq \rho_{m-1}$. We claim that D_{AB} is, in fact, an increasing function of z in this region, and so it will suffice to show that $D_{AB}(\rho_{m-1}) > 0$. To justify this claim, we first note that $\gamma_{AA}\,\phi = (z-2)\phi_{AA}\,\phi + \phi$. Since ϕ is a polynomial with coefficients 0 and 1, and $\phi_{AA}\,\phi$ is a polynomial of degree $2m-2$ with non-negative coefficients, it follows from Lemma 5 that $\gamma_{AA}\,\phi$ is an increasing function of z for $z > 2(1 - \frac{1}{2m-1})$. Now, $\rho_{m-1} > 2(1 - 2^{-(m-1)}) > 2(1 - \frac{1}{2m-1})$ for $m \geq 5$. Therefore, $\gamma_{AA}\,\phi$ is an increasing function of z for $z \geq \rho_{m-1}$. A similar argument shows that $(\phi_{AA} - \phi_{AB})\gamma_{BA}$ is also increasing in this region, which proves that D_{AB} is an increasing function in the region $z \geq \rho_{m-1}$.

The remainder of the proof just involves finding a positive lower bound for $D_{AB}(\rho_{m-1})$. From now on, for notational simplicity, we shall drop the subscript from ρ_{m-1}. Note first that $(\phi_{AA} - \phi_{AB})(\rho) = \rho^{m-1} + \rho^{m-2} + \ldots + \rho^{m_1} = (\rho^m - \rho^{m_1})/(\rho - 1)$. Next, we have

$$\begin{aligned}\gamma_{BA}(\rho) &= \rho^{m_2} - \rho^{m_2-1} - \ldots - 1 \\ &= \rho^{m_2-m+1}(\rho^{m-1} - \rho^{m-2} - \ldots - 1) + \rho^{-1} + \rho^{-2} + \ldots + \rho^{-(m-m_2-1)} \\ &= \frac{\rho^{-1} - \rho^{-(m-m_2)}}{1 - \rho^{-1}} = \frac{1 - \rho^{-(m-m_2-1)}}{\rho - 1}\end{aligned}$$

Hence, we have

$$\begin{aligned}(\phi_{AA} - \phi_{AB})\,\gamma_{BA}(\rho) &= \frac{1}{(\rho-1)^2}\,(\rho^m - \rho^{m_1} - \rho^{m_2+1} + \rho^{m_1+m_2-(m-1)}) \\ &= \frac{1}{(\rho-1)^2}\,[\rho^m - \rho^{m_1} - \rho^{m_1+m_2-(m-1)}(\rho^{m-m_1} - 1)]\end{aligned}$$

$$\geq \frac{1}{(\rho-1)^2}\left[\rho^m - \rho^{m_1} - (\rho^{m-m_1} - 1)\right]$$

the last inequality being a consequence of the fact that $m_1 + m_2 \leq m - 1$, which implies that $\rho^{m_1+m_2-(m-1)} \leq 1$. Noting that $\gamma_{AA}(\rho) = \rho(\rho^{m-1} - \ldots - 1) - 1 = -1$, we obtain

$$D_{AB}(\rho) \geq -\phi(\rho) + \frac{1}{(\rho-1)^2}\left[\rho^m - \rho^{m_1} - (\rho^{m-m_1} - 1)\right] \tag{8}$$

We shall first consider the case when $3 \leq m_1 \leq m-3$. In this case, we see that

$$\begin{aligned}
\frac{1}{(\rho-1)^2}\left[\rho^m - \rho^{m_1} - (\rho^{m-m_1} - 1)\right] &\geq \frac{1}{(\rho-1)^2}\left[\rho^m - \rho^{m-3} - (\rho^{m-3} - 1)\right] \\
&= \frac{1}{\rho-1}\left[\rho^{m-1} + \rho^{m-2} + \rho^{m-3} - (\rho^{m-4} + \ldots + 1)\right] \\
&> \frac{1}{\rho-1}\left(\rho^{m-1} + \rho^{m-2}\right) \qquad (9)
\end{aligned}$$

the last inequality arising from the fact that $z^{m-3} - z^{m-4} - \ldots - 1 > 0$ for $z > \rho_{m-3}$. Moreover,

$$\phi(\rho) \;=\; \rho^{m-1} + \sum_{k=0}^{m-3} c_k \rho^k \;\leq\; \rho^{m-1} + \sum_{k=0}^{m-3} \rho^k \;<\; \rho^{m-1} + \rho^{m-2} \tag{10}$$

since $z^{m-2} - z^{m-3} - \ldots - 1 > 0$ for $z > \rho_{m-2}$. Putting (9) and (10) into (8), we get

$$\begin{aligned}
D_{AB}(\rho) &> -(\rho^{m-1} + \rho^{m-2}) + \frac{1}{\rho-1}(\rho^{m-1} + \rho^{m-2}) \\
&= \frac{2-\rho}{\rho-1}(\rho^{m-1} + \rho^{m-2}) > 0
\end{aligned}$$

This shows that when $3 \leq m_1 \leq m-3$, $D_{AB}(z) > 0$ for $z = \rho$, and hence for all $z \geq \rho$ as well, which implies that $\rho_{AB} < \rho \; (= \rho_{m-1})$.

It only remains to deal with the cases when $m_1 = 1$, 2 and $m - 2$. Suppose that $m_1 = m - 2$. We then have

$$\begin{aligned}
\frac{1}{(\rho-1)^2}\left[\rho^m - \rho^{m_1} - (\rho^{m-m_1} - 1)\right] &= \frac{1}{(\rho-1)^2}\left[\rho^m - \rho^{m-2} - (\rho^2 - 1)\right] \\
&= \frac{1}{\rho-1}\left[\rho^{m-1} + \rho^{m-2} - (\rho+1)\right] \\
&\geq \rho^{m-1} + \rho^{m-2} - \rho - 1 \qquad (11)
\end{aligned}$$

Now, with $m_1 = m - 2$, the only possibility for B is $1^{m-2}01$, so that $\phi_{BB}(z) = z^{m-1} + 1$, and hence $\phi(z) = z^{m-1}$. Putting this and (11) into (8), we see that $D_{AB}(\rho) \geq \rho^{m-2} - \rho - 1 > \rho^2 - \rho - 1 > 0$.

Next, suppose $m_1 = 2$, which implies that B begins with 110, which in turn means that $B \circ B$ begins with 100. As a result, we have

$$\phi(\rho) \leq \rho^{m-1} + \rho^{m-4} + \rho^{m-5} + \ldots + 1 < \rho^{m-1} + \rho^{m-3} \tag{12}$$

since $z^{m-3} - z^{m-4} - \ldots - 1 > 0$ for $z > \rho_{m-3}$. Note that $\rho^m - \rho^{m_1} - (\rho^{m-m_1} - 1)$ is the same for $m_1 = 2$ and $m_1 = m - 2$. Therefore, putting (12) and (11) into (8), we get $D_{AB}(\rho) > \rho^{m-2} - \rho^{m-3} - \rho - 1 > 0$, since $z^{m-2} - z^{m-3} - \ldots - 1 > 0$ for $z > \rho_{m-2}$.

Finally, consider $m_1 = 1$, in which case we also have $m_2 = 1$. Therefore, $\phi_{AB} = \phi_{BA} \equiv 1$, and hence $\gamma_{AB}(z) = \gamma_{BA}(z) = z - 1$. Therefore,

$$(\phi_{AA} - \phi_{AB})\gamma_{BA}(z) = (z^{m-1} + z^{m-2} + \ldots + z)(z-1) = z^m - z$$

Also, as in (10), we have $\phi(\rho) < \rho^{m-1} + \rho^{m-2}$. Therefore, since $D_{AB} = \gamma_{AA}\,\phi + (\phi_{AA} - \phi_{AB})\gamma_{BA}$, and $\gamma_{AA}(\rho) = -1$, we see that

$$\begin{aligned} D_{AB}(\rho) &> -(\rho^{m-1} + \rho^{m-2}) + \rho^m - \rho \\ &= \rho(\rho^{m-1} - \rho^{m-2} - \rho^{m-3} - 1) \end{aligned}$$

which is strictly positive for $m \geq 5$, because $z^{m-1} - z^{m-2} - \ldots - 1 = 0$ at $z = \rho = \rho_{m-1}$.

We have thus shown that when B begins and ends with a 1, we have $D_{AB}(z) > 0$ for all $z \geq \rho_{m-1}$, and hence $\rho_{AB} < \rho_{m-1}$, which proves the lemma. □

Putting together the last three results, we obtain Theorem 1.

When $2 \leq m \leq 4$, it can be verified by computing $H(A, B)$ for all possible A, B of length m, that Theorem 1 remains valid when $m = 2$ or 4. When $m = 3$, all but the "only if" part of the theorem remains true. It turns out that for $m = 3$, the maximum Shannon capacity of $\log_2 \rho_2$ is also achieved by two other pairs, namely $\{000, 010\}$ and its complementary pair $\{111, 101\}$.

Interestingly, the answer to the problem of maximizing $H(A, B)$ also provides us with a means of determining the maximum Shannon capacity $H(A, B, C)$ of a constrained system that forbids three distinct binary m-sequences A, B and C. Formally, let $q_{ABC}(n)$ denote the number of binary n-sequences that do not contain A, B or C as a contiguous subsequence, and define $H(A, B, C) = \lim_{n\to\infty}(\log_2 q_{ABC}(n))/n$. We now show that $\max\limits_{A,B,C} H(A, B, C) = \max\limits_{A,B} H(A, B) = \log_2 \rho_{m-1}$.

THEOREM 18. *For $m \geq 2$,*

$$\max\{H(A, B, C) : A, B, C \in \{0,1\}^m, A \neq B, B \neq C, C \neq A\} = \log_2 \rho_{m-1}$$

Proof: It is clear that for any three distinct sequences A, B and C, we have $q_{ABC}(n) \leq q_{AB}(n)$ for all n. Therefore, $H(A, B, C) \leq H(A, B)$, from which it follows that $\max_{A,B,C} H(A, B, C) \leq \max_{A,B} H(A, B)$, where the maximum on the left is taken over all triples of distinct binary m-sequences, and that on the right is taken over all pairs of distinct binary m-sequences.

Now, from Theorem 1 (and, for $2 \leq m \leq 4$, the remarks following the proof of Lemma 17), we know that one of the sequence pairs that achieves $\max_{A,B} H(A, B) = \log_2 \rho_{m-1}$ is $\{10^{m-1}, 0^{m-1}1\}$. Let $\widehat{A} = 10^{m-1}$, $\widehat{B} = 0^{m-1}1$ and $\widehat{C} = 0^m$. For any $\mathcal{F} \subset \{0,1\}^m$, we define $\mathcal{B}_n(\mathcal{F})$ to be the set of all binary n-sequences that do not contain any element of $\mathcal{F}$ as a contiguous subsequence. It is easy to verify that $\mathcal{B}_n(\widehat{A}, \widehat{B}) = \mathcal{B}_n(0^{m-1}) \cup \{0^n\}$. Since the only sequence in $\mathcal{B}_n(\widehat{A}, \widehat{B})$ that contains 0^m is the all-zeros sequence, it is clear that $\mathcal{B}_n(\widehat{A}, \widehat{B}, \widehat{C}) = \mathcal{B}_n(0^{m-1})$.

Thus, we see that $q_{\hat{A}\hat{B}\hat{C}}(n) = q_{\hat{A}\hat{B}}(n) - 1$ which shows that $H(\widehat{A}, \widehat{B}, \widehat{C}) = H(\widehat{A}, \widehat{B})$. Since $H(\widehat{A}, \widehat{B}) = \log_2 \rho_{m-1}$, we obtain the following chain of inequalities

$$\log_2 \rho_{m-1} = H(\widehat{A}, \widehat{B}, \widehat{C}) \leq \max_{A,B,C} H(A, B, C) \leq \max_{A,B} H(A, B) = \log_2 \rho_{m-1}$$

which proves the theorem. □

4. Connection between $H(A,B)$ and $\widehat{R}(A,B,n)$. We now explore the relationship between the Shannon capacity $H(A,B)$ and the PPS code rate $\widehat{R}(A,B,n)$ defined in (4). We shall show that for nearly all choices of $A, B \in \{0,1\}^m$, $H(A,B) = \lim_{n\to\infty} \widehat{R}(A,B,n)$, and as a result, $\max_{A,B} H(A,B) = \lim_{n\to\infty} R(2,m,n)$, where $R(2,m,n)$ is the maximum possible rate of a $(2,m,n)$ PPS code.

We know from (3) that $F_{AB}(z) = \frac{\gamma_{AB}(z)}{z\,D_{AB}(z)}$ and $F_{BA}(z) = \frac{\gamma_{BA}(z)}{z\,D_{AB}(z)}$ are generating functions for $f_{AB}(k)$ and $f_{BA}(k)$, respectively. Now as noted previously, for $m \geq 5$, the largest positive root, ρ_{AB}, of $D_{AB}(z)$ is simple. Hence, if we establish that ρ_{AB} is also the largest-magnitude pole of $F_{AB}(z)$ and $F_{BA}(z)$, then it would follow that $f_{AB}(k) = c_{AB}\,(\rho_{AB})^k\,(1+o(1))$ and $f_{BA}(k) = c_{BA}\,(\rho_{AB})^k\,(1+o(1))$ for some constants c_{AB} and c_{BA}. This would clearly imply that $\lim_{n\to\infty} \widehat{R}(A,B,n) = \log_2 \rho_{AB} = H(A,B)$. We shall show that ρ_{AB} is almost always the largest-magnitude pole of both $F_{AB}(z)$ and $F_{BA}(z)$, and we shall characterize the exceptional cases.

The first step in this process is to show that ρ_{AB}, which we know is the largest positive root of $D_{AB}(z)$, is in fact the largest-magnitude root of $D_{AB}(z)$. Recall that we have previously shown using Perron-Frobenius theory that whenever $\rho_{AB} > 1$, ρ_{AB} is the unique largest-magnitude pole of $Q_{AB}(z)$ (which is defined by (2)), *i.e.*, if ρ is any other pole of $Q_{AB}(z)$, then $|\rho| < \rho_{AB}$.

Lemma 19. *For $m \geq 5$, if $\rho \neq \rho_{AB}$ is a root of $D_{AB}(z)$, then $|\rho| < \rho_{AB}$.*

Proof: We shall first show that $\rho_{AB} > 1.7$ which, apart from implying that ρ_{AB} is the unique largest pole of $Q_{AB}(z)$, will be important later in the proof. When $m \geq 5$, Proposition 9 shows that $H(A,B)$ is minimized by choosing $A = 110^{m-2}$ and $B = 110^{m-4}10$, and the proof of Lemma 10 shows that for this choice of A and B, $H(A,B) = \log_2 \zeta$, where ζ is the largest real zero of the polynomial $(z-2)z^{m-1}+2$. Therefore, for any $A, B \in \{0,1\}^m$, $\rho_{AB} \geq \zeta$. For $m \geq 6$, Lemma 10 shows that $\zeta > \rho_3 \approx 1.84$. For $m = 5$, it can be verified that $\zeta \approx 1.816$.

Now, suppose that ρ is a root of $D_{AB}(z)$ such that $|\rho| \geq \rho_{AB}$ and $\rho \neq \rho_{AB}$. We shall first show that ρ must be real (and hence negative), and then reach a contradiction by showing that ρ cannot be less than $-\rho_{AB}$. Since ρ cannot be a pole of $Q_{AB}(z)$, it must be a root of the numerator polynomial of $Q_{AB}(z)$, *i.e.*, $\phi_{AA}(\rho)\phi_{BB}(\rho) = \phi_{AB}(\rho)\phi_{BA}(\rho)$. An argument similar to that in the proof of Corollary 8 now shows that $\{\phi_{AA}(\rho), \phi_{BB}(\rho)\} = \{\phi_{AB}(\rho), \phi_{BA}(\rho)\}$. Thus, ρ must be a root of one of the polynomials $\phi_{AA} - \phi_{AB}$ and $\phi_{AA} - \phi_{BA}$, both of which are polynomials of degree $m-1$ whose coefficients take values in the set $\{0,1,-1\}$.

A result of Bloch and Pólya [2] states that if $p(z)$ is any polynomial whose coefficients take values in $\{0,1,-1\}$, then for any $q \in (1,2)$, the number N of roots of $p(z)$ in the region $|z| > q$ can be bounded as follows:

$$N \leq \frac{1}{2}\left(\log \frac{4q^2}{(3q+1)(q-1)}\right) \Big/ \log\left(1 + \frac{q-1}{2}\right)$$

Evaluating this expression with $q = 1.7$, we see that $p(z)$ has at most one root in the region $|z| > 1.7$.

Thus, since $\rho_{AB} > 1.7$, the polynomials $\phi_{AA} - \phi_{AB}$ and $\phi_{AA} - \phi_{BA}$ can have at most one root in the region $|z| \geq \rho_{AB}$. Since ρ is a root of one of these polynomials, it is the unique root in $|z| > \rho_{AB}$, and hence must be real. Since ρ_{AB} is the largest positive root of D_{AB}, ρ must be negative. Recall from Proposition 9 and Lemma 10 that $\rho_{AB} > \rho_{m-3}$. Thus, we shall reach a contradiction if we can show that no negative root of $\phi_{AA} - \phi_{AB}$ or $\phi_{AA} - \phi_{BA}$ can be less than $-\rho_{m-3}$. We now provide the sketch of an argument that shows this.

Let $(\phi_{AA} - \phi_{AB})(z) = z^{m-1} + \sum_{k=0}^{m-2} c_k z^k$, with the c_k's taking values in $\{0, 1, -1\}$ (the argument for $\phi_{AA} - \phi_{BA}$ is identical). Suppose first that $m-1$ is even, and further that c_{m-2} is 0 or -1. In this case, for any $z < -\rho_{m-3}$, we have

$$\begin{aligned}(\phi_{AA} - \phi_{AB})(z) &\geq |z|^{m-1} - \sum_{k=0}^{m-3} |z|^k \\ &= |z|^2 \left(|z|^{m-3} - \sum_{k=0}^{m-4} |z|^k \right) - |z| - 1 + |z|^{m-2} \\ &> |z|^{m-2} - |z| - 1\end{aligned}$$

with the first inequality holding for any $z < 0$, and the last inequality holding for $|z| > \rho_{m-3}$. But, $|z|^{m-2} - |z| - 1 > 0$ for $m \geq 5$ and $|z| > \rho_2$, which shows that $\phi_{AA} - \phi_{AB}$ has no zeros less than $-\rho_{m-3}$.

Next, suppose that $m-1$ is even and $c_{m-2} = 1$. Then, the correlation $A \circ A$ must begin with 11, and hence must be 1^m. Therefore, $\phi_{AA}(z) = \sum_{k=0}^{m-1} z^k$, which means that $c_k \in \{0, 1\}$ for $k = 0, 1, \ldots, m-3$. We then have for any $z < 0$,

$$\begin{aligned}(\phi_{AA} - \phi_{AB})(z) &\geq |z|^{m-1} - \sum_{\substack{1 \leq k \leq m-2 \\ k \text{ odd}}} |z|^k \\ &= |z|^2 \left(|z|^{m-3} - \sum_{k=0}^{m-4} |z|^k \right) - |z| + \sum_{\substack{2 \leq k \leq m-3 \\ k \text{ even}}} |z|^k \\ &> -|z| + \sum_{\substack{2 \leq k \leq m-3 \\ k \text{ even}}} |z|^k\end{aligned}$$

with the last inequality holding for $|z| > \rho_{m-3}$. Since $|z|^{m-3} + |z|^{m-5} + \ldots + |z|^2 - |z|$ is clearly positive for $m \geq 5$ and $|z| > 1$, we see that $\phi_{AA} - \phi_{AB}$ has no zeros less than $-\rho_{m-3}$ whenever $m-1$ is even.

A similar argument as above shows that when $m-1$ is odd, then $(\phi_{AA} - \phi_{AB})(z) < 0$ for all $z < -\rho_{m-3}$, which completes the proof of the lemma. □

We have thus shown that for $m \geq 5$, ρ_{AB} is the unique largest-magnitude root of D_{AB}. We are now in a position to determine exactly when ρ_{AB} is the largest-magnitude pole of $F_{AB}(z)$ and $F_{BA}(z)$. Note that ρ_{AB} cannot be a pole of *both* $F_{AB}(z)$ and $F_{BA}(z)$ if and only if ρ_{AB} is a root of γ_{AB} as well as γ_{BA}. But by Theorem 7, this can happen if and only if $\{A, B\}$ or $\{\overline{A}, \overline{B}\} = \{10^{m-1}, 0^m\}$, $\{10^{m-1}, 0^{m-1}1\}$ or $\{0^m, 0^{m-1}1\}$. This leads us to the following proposition.

Proposition 20. *For all $m \geq 5$, the following are true:*
(a) If $\{A, B\}$ or $\{\overline{A}, \overline{B}\} = \{0^m, 0^{m-1}1\}$ or $\{0^m, 10^{m-1}\}$, then $\lim_{n\to\infty} \widehat{R}(A, B, n) = 0$.
(b) If $\{A, B\}$ or $\{\overline{A}, \overline{B}\} = \{10^{m-1}, 0^{m-1}1\}$, then $\lim_{n\to\infty} \widehat{R}(A, B, n) = \frac{1}{2} H(A, B)$.
(c) For all other pairs of distinct binary m-sequences A, B, $\lim_{n\to\infty} \widehat{R}(A, B, n) = H(A, B)$.

Proof: The discussion preceding the statement of the proposition shows that if $\{A,B\}$ or $\{\overline{A},\overline{B}\}$ is not one of the pairs listed in (a) and (b), then ρ_{AB} is the unique largest pole, in terms of absolute value, of both $F_{AB}(z)$ and $F_{BA}(z)$. Therefore, as noted prior to the statement of Lemma 19, it follows that $\lim_{n\to\infty}\widehat{R}(A,B,n) = \log_2\rho_{AB} = H(A,B)$, which proves (c).

To prove (a), note that if $A = 0^m$ and $B = 0^{m-1}1$, then we can have no binary sequence of length $k \geq m+2$ that begins with A and ends with B, but does not contain A or B elsewhere. In other words, $f_{AB}(k) = 0$ for all $k \geq m+2$, and so by definition, $\widehat{R}(A,B,n) = 0$ for all $n \geq m$. The other cases can be similarly dismissed.

Finally, if $A = 10^{m-1}$ and $B = 0^{m-1}1$, then it is clear that the only sequence that can be counted by $f_{AB}(k)$, $k \geq m+2$, is $10^{k-2}1$. Hence, $f_{AB}(k) = 1$ for all $k \geq m+2$. However, $f_{BA}(k) = c_{BA}\left(\rho_{AB}\right)^k(1+o(1))$ for some positive constant c_{BA} because, as can easily be verified, ρ_{AB} is the unique largest-magnitude pole of $F_{BA}(z)$ in this case. As a result, we have $\lim_{n\to\infty}\widehat{R}(A,B,n) = \frac{1}{2}\log_2\rho_{AB}$, which completes the proof of the proposition. □

Theorem 2 is an immediate consequence of the above proposition and Theorem 1. Theorem 2 shows that when $m \geq 5$, for all sufficiently large n, $R(2,m,n)$ is either $\widehat{R}(0^m,1^m,n)$ or $\widehat{R}(\langle 01\rangle_m,\langle 10\rangle_m,n)$. In fact, as we show next, $|f_{\langle 01\rangle_m\langle 10\rangle_m}(k) - f_{0^m1^m}(k)| \leq 1$ for all k, and hence due to the floor function used in defining $\widehat{R}(A,B,n)$, for nearly all (if not all) values of n, $\widehat{R}(0^m,1^m,n) = \widehat{R}(\langle 01\rangle_m,\langle 10\rangle_m,n)$. Thus, for nearly all (if not all) sufficiently large integers n,

$$R(2,m,n) = \widehat{R}(0^m,1^m,n) = \widehat{R}(\langle 01\rangle_m,\langle 10\rangle_m,n)$$

Note that $F_{\langle 01\rangle_m\langle 10\rangle_m}(z) - F_{0^m1^m}(z)$ is a generating function for $f_{\langle 01\rangle_m\langle 10\rangle_m}(k) - f_{0^m1^m}(k)$. Using (3) to get explicit expressions for $F_{\langle 01\rangle_m\langle 10\rangle_m}(z)$ and $F_{0^m1^m}(z)$, we find after some algebraic manipulations that

$$F_{\langle 01\rangle_m\langle 10\rangle_m}(z) - F_{0^m1^m}(z) = \begin{cases} \frac{1}{z(z^m-1)} & \text{if } m \text{ is even} \\ \frac{z^{m-1}-1}{z^{2m}-1} & \text{if } m \text{ is odd} \end{cases}$$

It is easily verified that the coefficients in the power series expansions (in the variable z^{-1}) of both $\frac{1}{z(z^m-1)}$ and $\frac{z^{m-1}-1}{z^{2m}-1}$ belong to the set $\{-1,0,1\}$. Thus, for each k, $f_{\langle 01\rangle_m\langle 10\rangle_m}(k) - f_{0^m1^m}(k)$ is either -1, 0 or 1.

For the sake of completeness, we would like to mention that when $m = 4$, it can be shown that Theorem 2 remains true in its entirety. When $m = 3$, the theorem remains valid if its statement is modified as follows: $\lim_{n\to\infty}\widehat{R}(A,B,n) \leq \log_2\rho_2$ with equality if and only if $\{A,B\} = \{000,111\}$, $\{010,101\}$, $\{000,010\}$ or $\{111,101\}$. However, it can be shown that if $\{A,B\}$ is one of the last two sequence pairs, then $f_{0^31^3}(k) - f_{AB}(k) = c\left(\rho_2\right)^k(1+o(1))$, where c is approximately 0.0034. Thus, for all sufficiently large n, we have $\widehat{R}(A,B,n) \leq \widehat{R}(0^m,1^m,n)$. Finally, when $m = 2$, $\lim_{n\to\infty}\widehat{R}(A,B,n) = 0$ for all sequence pairs A,B.

REFERENCES

[1] J.J. Ashley and P.H. Siegel, *A note on the Shannon capacity of run-length-limited codes*, IEEE Trans. Inform. Theory, 33 (1987), pp. 601–605.

[2] A. Bloch and G. Pólya, *On the roots of certain algebraic equations*, Proc. London Math. Soc. (2), 33 (1930), pp. 102–114. Reproduced in G. Pólya, *Collected Papers: Location of Zeros*, vol. II, pp. 336–346, MIT Press, Cambridge, MA, 1974.

[3] E.N. Gilbert, *Synchronization of binary messages*, IRE Trans. Inform. Theory, 6 (1960), pp. 470–477.

[4] L.J. Guibas and A.M. Odlyzko, *Maximal prefix-synchronized codes*, SIAM J. Appl. Math., 35 (1978), pp. 401–418.

[5] L.J. Guibas and A.M. Odlyzko, *String overlaps, pattern matching, and nontransitive games*, J. Combin. Theory Ser. A, 30 (1981), pp. 183–208.

[6] L.J. Guibas and A.M. Odlyzko, *Periods in strings*, J. Combin. Theory Ser. A, 30 (1981), pp. 19–42.

[7] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, UK, 1985.

[8] N. Kashyap and D.L. Neuhoff, *Data synchronization with timing*, IEEE Trans. Inform. Theory, 47 (2001), pp. 1444–1460.

[9] N. Kashyap and D.L. Neuhoff, *Periodic prefix-synchronized codes: a generating function approach*, submitted to IEEE Trans. Inform. Theory. Available via anonmyous ftp at `ftp.eecs.umich.edu/people/neuhoff/period2_codes.submit.{ps,pdf}`.

[10] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge Univ. Press, Cambridge, UK, 1995.

[11] D.A. Lind, *Perturbation of shifts of finite type*, SIAM J. Discrete Math., 2 (1989), pp. 350–365.

[12] R.A. Scholtz, *Frame synchronization techniques*, IEEE Trans. Commun., 28 (1980), pp. 1204–1212.

[13] C.E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. J., 27 (1948), pp. 379–423.

[14] H.S. Wilf, *generatingfunctionology*, 2nd ed., Academic Press, San Diego, CA, 1994.

[15] D.A. Wolfram, *Solving generalized Fibonacci recurrences*, Fibonacci Quart., 36.2 (1998), pp. 129–145.