Construction of Polar Codes with Sublinear Complexity

Marco Mondelli, S. Hamed Hassani, and Rüdiger Urbanke

***Abstract*—Consider the problem of constructing a polar code of block length N for the transmission over a given channel W. Typically this requires to compute the reliability of all the N synthetic channels and then to include those that are sufficiently reliable. However, we know from [1], [2] that there is a partial order among the synthetic channels. Hence, it is natural to ask whether we can exploit it to reduce the computational burden of the construction problem.**

We show that, if we take advantage of the partial order [1], [2], we can construct a polar code by computing the reliability of roughly a fraction $1/\log^{3/2} N$ of the synthetic channels. In particular, we prove that $N/\log^{3/2} N$ is a lower bound on the number of synthetic channels to be considered and such a bound is tight up to a multiplicative factor $\log\log N$. This set of roughly $N/\log^{3/2} N$ synthetic channels is universal, in the sense that it allows one to construct polar codes for any W, and it can be identified by solving a maximum matching problem on a bipartite graph.

Our proof technique consists of reducing the construction problem to the problem of computing the maximum cardinality of an antichain for a suitable partially ordered set. As such, this method is general and it can be used to further improve the complexity of the construction problem in case a new partial order on the synthetic channels of polar codes is discovered.

***Keywords*—Polar codes; partial order; construction problem; antichain; chain.**

I. Introduction

Polar codes, introduced by Arıkan [3], achieve the capacity of any binary memoryless symmetric (BMS) channel with encoding and decoding complexity $\Theta(N\log_2 N)$, where N is the block length of the code. A unified characterization of the performance of polar codes in several regimes is presented in [4]. Let us mention the following basic facts: the error probability scales with the block length roughly as $2^{-\sqrt{N}}$ [5]; the gap to capacity scales with the block length as $N^{-1/\mu}$, and bounds on the scaling exponent μ are provided in [4], [6], [7]; and polar codes are not affected by error floors [4]. A successive cancellation list (SCL) decoder with space complexity $O(LN)$ and time complexity $O(LN\log_2 N)$ is proposed in [8], where L is the size of the list. Empirically, the use of several concurrent decoding paths yields an error probability comparable to that under optimal MAP decoding with practical values of the list size. Furthermore, by adding only a few extra bits of cyclic redundancy check (CRC) precoding, the resulting performance is comparable with state-of-the-art LDPC codes. Because of their attractive features, polar codes are being considered for use in future wireless communication systems (e.g., 5G cellular systems).

The idea of channel polarization is to take independent copies of the transmission channel and to transform them into a set of reliable channels and a set of unreliable channels, in such a way that the overall capacity is preserved. Then, the information bits are transmitted in the positions corresponding to the reliable channels and the remaining positions are *frozen* (i.e., their value is shared between the encoder and the decoder). Therefore, in order to construct a polar code, we need to identify the positions corresponding to the reliable synthetic channels. Several techniques have proposed to estimate the reliability of the synthetic channels: Monte Carlo simulations [3], density evolution [9], [10], Gaussian approximation of density evolution [11], efficient degrading and upgrading methods [12], [13].

In general, the ranking of the synthetic channels depends on the specific transmission channel. Hence, one solution to the problem of code construction is to evaluate the reliability of *all* synthetic channels. However, it was observed that there is a partial order between the synthetic channels, which holds for any transmission channel. A first partial order was described in [10] and it was combined with a different partial order in the two independent works [1], [2]. In [1], it is also empirically shown that, by exploiting this combined partial order, the complexity of the code construction can be significantly reduced.

In this paper, we give a tight characterization of the complexity reduction guaranteed by the exploitation of this partial order. In particular, we derive *universal* bounds on the number of synthetic channels whose reliability has to be computed in order to construct the polar code. The bounds are *universal* in the sense that they hold for any transmission channel. Our main result consists in proving an upper and a lower bound that differ by a factor of $\log\log N$, where $N = 2^n$ is the block length of the code. The lower bound is equal to a known integer sequence, i.e., the maximal number of subsets of $\{1, 2, \dots, n\}$ that share the same sum (sequence A025591 in [14]). Such a sequence scales as $N/\log^{3/2} N$, which means that we need to compute the reliability of roughly a fraction $1/\log^{3/2} N$ of the synthetic channels. In other words, in order to construct a polar code, it suffices to know the reliability of a sublinear number of synthetic channels.

The remainder of this paper is organized as follows. In Section II, we set up the notation, describe the partial order derived in [1], [2], and formalize the construction problem.

M. Mondelli and R. Urbanke are with the School of Computer and Communication Sciences, EPFL, CH-1015 Lausanne, Switzerland (e-mail: {marco.mondelli, ruediger.urbanke}@epfl.ch).

S. H. Hassani is with the Computer Science Department, ETH Zürich, Switzerland (e-mail: hamed@inf.ethz.ch).

In Section III, we state the main result about the complexity of the construction problem and we present its immediate implications. In Section IV, we give the proof and we describe how to actually find the channels whose reliability has to be computed. In Section V, we provide some concluding remarks. The proofs of some intermediate results are deferred to the Appendix.

II. Preliminaries

A. Reliability Measures and Degradation

Let $W : \mathcal{X} \to \mathcal{Y}$ be a BMS channel with input alphabet $\mathcal{X} = \{0, 1\}$, output alphabet $\mathcal{Y}$, and transition probabilities $p_{Y|X}(y \mid x)$ for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The random variables representing the input and the output of the channel are denoted by X and Y, respectively. Since the channel is symmetric, we impose a uniform prior on the input, i.e., $p_X(0) = p_X(1) = 1/2$.

There are several measures of the reliability of a channel, as specified by the following definition.

Definition 1 (Reliability Measures)**:** Let $W : \mathcal{X} \to \mathcal{Y}$ be a BMS channel with transition probabilities $p_{Y|X}(y \mid x)$ for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The reliability of W is measured by one of the following quantities:

- The *mutual information* $I(W)$, defined as
$$I(W) = I(X; Y); \tag{1}$$
- The *Bhattacharyya parameter* $Z(W)$, defined as
$$Z(W) = \sum_{y \in \mathcal{Y}} \sqrt{p_{Y|X}(y \mid 0) p_{Y|X}(y \mid 1)}; \tag{2}$$
- The *MAP error probability* $P_{\mathrm{e}}(W)$, defined as
$$P_{\mathrm{e}}(W) = \mathbb{P}(X \neq \hat{x}(Y)), \tag{3}$$
where $\hat{x}(y) = \operatorname{argmax}_x p_{X|Y}(x \mid y)$ is the MAP decision of X given Y.

Note that a channel is *reliable* when it has a *large* mutual information, a *small* Bhattacharyya parameter, and a *small* MAP error probability.

Let us now define the concept of stochastic degradation.

Definition 2 (Stochastic Degradation)**:** Let $W_1 : \mathcal{X} \to \mathcal{Y}_1$ and $W_2 : \mathcal{X} \to Y_2$ be two BMS channels with respective transition probabilities $p_{Y_1|X}(y_1 \mid x)$ and $p_{Y_2|X}(y_2 \mid x)$, for $x \in \mathcal{X}$, $y_1 \in \mathcal{Y}_1$, and $y_2 \in \mathcal{Y}_2$. We say that W_1 is *stochastically degraded* with respect to W_2 and we write $W_1 \preceq W_2$ if there exists a memoryless channel with transition probabilities $p_{Y_1|Y_2}(y_1 \mid y_2)$ such that for all $x \in \mathcal{X}$ and $y_1 \in \mathcal{Y}_1$,

$$p_{Y_1|X}(y_1 \mid x) = \sum_{y_2 \in \mathcal{Y}_2} p_{Y_1|Y_2}(y_1 \mid y_2) p_{Y_2|X}(y_2 \mid x). \tag{4}$$

If a channel is stochastically degraded, all the reliability measures defined in Definition 1 become worse. This means that the mutual information decreases, the Bhattacharyya parameter increases, and the error probability increases. Such a fact is formalized by the following proposition (see Theorem 4.76 of [15] or Lemma 3 of [12]).

Proposition 1 (Stochastic Degradation and Reliability Measures)**:** Let $W_1 : \mathcal{X} \to \mathcal{Y}_1$ and $W_2 : \mathcal{X} \to Y_2$ be two BMS channels and assume that $W_1 \preceq W_2$. Then,

$$I(W_1) \leq I(W_2), \tag{5}$$
$$Z(W_1) \geq Z(W_2), \tag{6}$$
$$P_{\mathrm{e}}(W_1) \geq P_{\mathrm{e}}(W_2). \tag{7}$$

B. Synthetic Channels

The basis of channel polarization consists in mapping two identical copies of the channel $W : \mathcal{X} \to \mathcal{Y}$ into the pair of channels $W^0 : \mathcal{X} \to \mathcal{Y}^2$ and $W^1 : \mathcal{X} \to \mathcal{X} \times \mathcal{Y}^2$, defined as

$$W^0(y_1, y_2 \mid x_1) = \sum_{x_2 \in \mathcal{X}} \frac{1}{2} W(y_1 \mid x_1 \oplus x_2) W(y_2 \mid x_2), \tag{8}$$

$$W^1(y_1, y_2, x_1 \mid x_2) = \frac{1}{2} W(y_1 \mid x_1 \oplus x_2) W(y_2 \mid x_2). \tag{9}$$

Then, W^0 is a worse channel in the sense that it is degraded with respect to W, hence less reliable than W; and W^1 is a better channel in the sense that it is upgraded with respect to W, hence more reliable than W.

By iterating this operation n times, we map $N = 2^n$ identical copies of the transmission channel W into the synthetic channels $\{W_N^{(i)}\}_{i \in \{0,\dots,N-1\}}$. More specifically, given $i \in \{0, \dots, N-1\}$, let $(i_1, i_2, \dots, i_n)$ be its binary expansion over n bits, where i_1 is the most significant bit and i_n is the least significant bit, i.e.,

$$i = \sum_{k=1}^{n} i_k 2^{n-1-k}. \tag{10}$$

Then, we define the synthetic channels $\{W_N^{(i)}\}_{i \in \{0,\dots,N-1\}}$ as

$$W_N^{(i)} = (((W^{i_1})^{i_2})^{\cdots})^{i_n}. \tag{11}$$

Example 1 (Synthetic Channel)**:** Take $n = 4$ and $i = 10$. Then, the synthetic channel $W_{16}^{(10)} = (((W^1)^0)^1)^0$ is obtained by applying first (9), then (8), then (9), and finally (8).

C. Partial Order

In order to describe the partial order, it is helpful to define two operators, i.e., the addition and the left-swap operator, that map the index of a synthetic channel into the index of another synthetic channel.

Definition 3 (Addition Operator)**:** Let $i \in \{0, \dots, N-1\}$ and denote by $(i_1, i_2, \cdots, i_n)$ its binary expansion over n bits, defined in (10). Given $k \in \{1, \dots, n\}$, the *addition operator* at position k maps i into $A^{(k)}(i) \in \{0, \dots, N-1\}$. The binary expansion over n bits of $A^{(k)}(i)$ is defined as

$$(A^{(k)}(i))_\ell = \begin{cases} 1, & \ell = k, \\ i_\ell, & \ell \neq k. \end{cases} \tag{12}$$

In words, the addition operator $A^{(k)}$ takes the input i and sets to 1 the k-th of its binary expansion. Note that, if $i_k = 1$, the addition operation $A^{(k)}$ simply copies the input into the output.

Example 2 (Addition Operator)**:** Take $n = 4$ and $i = 10$. Note that i has binary expansion $(1, 0, 1, 0)$. Then, $A^{(2)}(10) = 14$ and its binary expansion is $(1, 1, 1, 0)$. Furthermore, $A^{(3)}(10) = 10$ and its binary expansion is $(1, 0, 1, 0)$.

Definition 4 (Left-swap Operator)**:** Let $i \in \{0, \dots, N-1\}$ and denote by $(i_1, i_2, \cdots, i_n)$ its binary expansion over n bits, defined in (10). Given $k \in \{2, \dots, n\}$, the *left-swap operator* at position k maps i into $L^{(k)}(i) \in \{0, \dots, N-1\}$. If $i_k \neq 1$ *or* $i_{k-1} \neq 0$ then $L^{(k)}(i) = i$. Otherwise, the binary expansion over n bits of $L^{(k)}(i)$ is defined as

$$(L^{(k)}(i))_\ell = \begin{cases} 1, & \ell = k-1, \\ 0, & \ell = k, \\ i_\ell, & \ell \notin \{k-1, k\}. \end{cases} \tag{13}$$

In words, the left-swap operator $L^{(k)}$ takes the input and, if possible, it swaps the 1 in the k-th position with the bit on its left. This means that, if $i_k = 1$ *and* $i_{k-1} = 0$, the left-swap operator $L^{(k)}$ swaps position $k-1$ with position k. Otherwise, it simply copies the input into the output.

Example 3 (Left-swap Operator)**:** Take $n = 4$ and $i = 10$. Note that i has binary expansion $(1, 0, 1, 0)$. Then, $L^{(2)}(10) = 10$ and its binary expansion is $(1, 0, 1, 0)$. Furthermore, $L^{(3)}(10) = 12$ and its binary expansion is $(1, 1, 0, 0)$.

We are now ready to describe the partial order introduced in [1], [2].

Proposition 2 (Partial Order)**:** Let W be a BMS channel and consider the $N = 2^n$ synthetic channels $\{W_N^{(i)}\}_{i \in \{0,\dots,N-1\}}$ obtained from W by applying (11). Then, for any $i \in \{0, \dots, N-1\}$,

$$W_N^{(i)} \preceq W_N^{(A^{(k)}(i))}, \quad \forall\, k \in \{1, \dots, n\}, \tag{14}$$

$$W_N^{(i)} \preceq W_N^{(L^{(k)}(i))}, \quad \forall\, k \in \{2, \dots, n\}. \tag{15}$$

For the proof of (14), see Section V of [10] and, for the proof of (15), see Theorem 1 of [1]. Furthermore, note that (14) and (15) hold for any BMS channel W. For this reason, we say that the partial order of Proposition 2 is *universal*.

Example 4 (Partial Order)**:** Take $n = 4$ and $i = 10$. By applying Proposition 2 and recalling Examples 2 and 3, we immediately conclude that $W_{16}^{(10)} \preceq W_{16}^{(12)}$ and $W_{16}^{(10)} \preceq W_{16}^{(14)}$.

D. Construction Problem

Given a BMS channel W and a block length N, the problem of the construction of polar codes consists in selecting the set of the most reliable synthetic channels defined as in (11). According to Definition 1, there are several notions of reliability. Since all these reliability measures become worse under stochastic degradation by Proposition 1, it does not really matter which one we choose. To fix the ideas, let us consider the Bhattacharyya parameter and define the construction problem to be the selection of the set of synthetic channels with the lowest Bhattacharyya parameters. However, keep in mind that the arguments and the conclusions of this paper remain valid when we choose the mutual information or the MAP error probability as reliability measures. Indeed, in this paper we exploit the partial order of Proposition 2, which is an ordering of the synthetic channels in the sense of the stochastic degradation.

Definition 5 (Construction Problem)**:** Let W be a BMS channel and consider the $N = 2^n$ synthetic channels $\{W_N^{(i)}\}_{i \in \{0,\dots,N-1\}}$ obtained from W by applying (11). In order to construct a polar code of block length N, we need to solve either the *fixed rate (FR) problem* or the *fixed performance (FP) problem* that are defined as follows.

- *Fixed rate (FR) problem.* Given a block length N and a rate $R \in (0, 1)$, output the set of $\lfloor NR \rfloor$ synthetic channels with the smallest Bhattacharyya parameters.
- *Fixed performance (FP) problem.* Given a block length N and a threshold $\gamma \in (0, 1)$, output all the synthetic channels whose Bhattacharyya parameter is smaller than γ.

In the sequel we will limit our discussion to the FP construction problem. Note that if we can solve the FP construction problem, we can also solve the FR construction problem by simply performing a bisection on the values of the threshold.

III. Statement of the Main Result

Theorem 1 (Complexity of FP Construction Problem)**:** Let W be a BMS channel and $N = 2^n$ be the block length. Let $M(n)$ be the maximal number of subsets of $\{1, \dots, n\}$ that share the same sum. Consider the partial order of Proposition 2 and use it to solve the FP construction problem with threshold γ of Definition 5. Then, the complexity of such a task can be bounded as follows.

- *Upper bound:* it suffices to compute the Bhattacharyya parameter of *at most*

$$M(n) \cdot \log\left(\frac{2^{n+1}}{M(n)}\right)$$

  synthetic channels, for any $\gamma \in (0, 1)$.
- *Lower bound:* it is necessary to compute the Bhattacharyya parameter of *at least* $M(n)$ synthetic channels, for some $\gamma \in (0, 1)$.

Let us point out that the results above are *universal* in the sense that they hold for any BMS channel W. Note also that the upper bound holds for any choice of the threshold $\gamma \in (0, 1)$. On the contrary, the lower bound holds for some $\gamma \in (0, 1)$. Indeed, for some specific values of the threshold, the task might be easier. For example, if γ is very small (e.g. $\gamma < Z(W)^N$), then none of the synthetic channels have a Bhattacharyya parameter smaller than γ. Similarly, if γ is very large (e.g., $\gamma > 1 - (1 - Z(W))^N$), then all the synthetic channels have a Bhattacharyya parameter smaller

than γ. Hence, it is interesting to provide a lower bound for the "hard" instances of γ.

The sequence $M(n)$ is the integer sequence A025591 in [14]. The following lemma, stated below and proved in Appendix B, provides an asymptotic formula for it.

Theorem 2 (Asymptotic Formula for $M(n)$)**:** Let $M(n)$ be the maximal number of subsets of $\{1,\dots,n\}$ that share the same sum. Then,

$$M(n) = \sqrt{\frac{6}{\pi}}\frac{2^n}{n^{3/2}}(1+o(1)). \tag{16}$$

By applying Theorem 1 and 2, we immediately conclude that, in order to solve the FP construction problem, we need to compute the Bhattacharyya parameter of roughly $N/\log^{3/2} N$ synthetic channels. Furthermore, the upper and the lower bound differ by a multiplicative factor of $\log(2N/M(n))$, which scales as $\log\log N$. In words, this means that we need to compute the Bhattacharyya parameter of a sublinear number of channels. This is possible only because we exploit the partial order of Proposition 2.

Indeed, assume that we do not use any partial order between the synthetic channels. Then, the only way to solve the FP construction problem is to compute the Bhattacharyya parameter of all the N synthetic channels. On the contrary, suppose that there was a total order among the synthetic channels. Then, we could rank them from best to worst and, by using a binary search algorithm, we need to compute the Bhattacharyya parameter of at most $n+1 = \log N + 1$ synthetic channels. The main result of this paper is that by using the partial order of Proposition 2 we need to compute the Bhattacharyya parameter of roughly $N/\log^{3/2} N$ synthetic channels. Furthermore, as detailed at the end of Section IV, these $N/\log^{3/2} N$ synthetic channels can also be identified efficiently by solving a maximum matching problem on a bipartite graph.

Let us highlight that the bounds of Theorem 1 hold when we exploit *only* the partial order of Proposition 2. This partial order relies on the addition and left-swap operators of Definitions 3-4, and these represent the only known operators that imply stochastic degradation. If one finds another operator that implies stochastic degradation, by exploiting the induced partial order, in principle it is possible to further reduce the number of Bhattacharyya parameters to be computed.

IV. Proof and Discussion

In order to prove Theorem 1, we need some definitions about partially ordered sets (or posets, for short). For a general introduction to the subject of posets, we refer the interested reader to [16, Chapter 1] and [17, Chapter 3].

Let us associate the synthetic channel $W_N^{(i)}$ with the binary expansion $(i_1, i_2, \dots, i_n)$ of the index i defined in (10). Then, the partial order of Proposition 2 induces a partial order over the Hamming cube $\{0,1\}^n$. We will denote such a partial order by $\prec_1$.

The *Hasse diagram* of the poset $\{0,1\}^n$ equipped with the order $\prec_1$ is represented in Figure 1 for $n=4$ and $n=5$. Recall that an element x is connected via an edge to an element y *if and only if* they are ordered, i.e., $x \prec_1 y$ (respectively, $y \prec_1 x$) *and* there is no other element z such that $x \prec_1 z \prec_1 y$ (respectively, $y \prec_1 z \prec_1 x$). In words, the Hasse diagram connects only "nearest neighbors".

Let us now define the concepts of chain and antichain that play a central role in our analysis.

Definition 6 (Chain and Antichain)**:** Let P be a poset. We say that a subset of P is a *chain* if it is totally ordered. We say that a subset of P is an *antichain* if no two elements in it are comparable.

Example 5 (Chain and Antichain)**:** Consider the partial order over $\{0,1\}^4$ whose Hasse diagram is represented in Figure 1a. Define

$$\mathcal{C} = \{(0,0,1,0),(0,0,1,1),(0,1,0,1),(1,0,0,1)\},$$
$$\mathcal{A} = \{(1,0,0,0),(0,1,1,1)\}.$$

Then, $\mathcal{C}$ is a chain and $\mathcal{A}$ is an antichain. Indeed, the elements $(1,0,0,0)$ and $(0,1,1,1)$ are not comparable and we have that

$$(0,0,1,0) \prec_1 (0,0,1,1) \prec_1 (0,1,0,1) \prec_1 (1,0,0,1).$$

Analogously, consider the partial order over the set of synthetic channels $\{W_N^{(i)}\}_{i\in\{0,\dots,N-1\}}$ given by Proposition 2. Define

$$\mathcal{C}' = \{W_{16}^{(2)}, W_{16}^{(3)}, W_{16}^{(5)}, W_{16}^{(9)}\},$$
$$\mathcal{A}' = \{W_{16}^{(8)}, W_{16}^{(7)}\}.$$

Then, $\mathcal{C}'$ is a chain and $\mathcal{A}'$ is an antichain. Indeed, the synthetic channels $W_{16}^{(8)}$ and $W_{16}^{(7)}$ are not comparable and we have that

$$W_{16}^{(2)} \prec W_{16}^{(3)} \prec W_{16}^{(5)} \prec W_{16}^{(9)}.$$

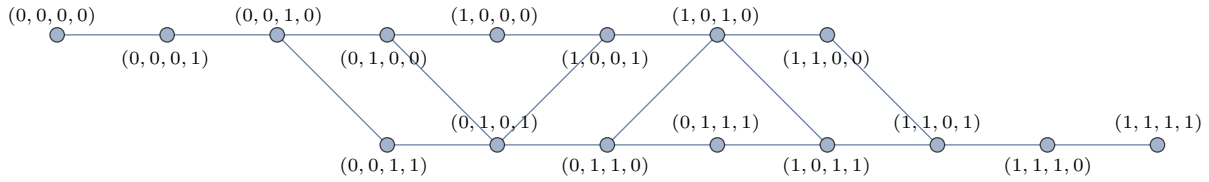
The maximum cardinality of an antichain is equal to the minimum number of chains that form a partition of the poset by Dilworth's theorem [18, Theorem 1.2], [19, Theorem 12.5].

Theorem 3 (Dilworth)**:** The minimum number of chains into which the elements of a poset P can be partitioned is equal to the maximum number of elements in an antichain of P.

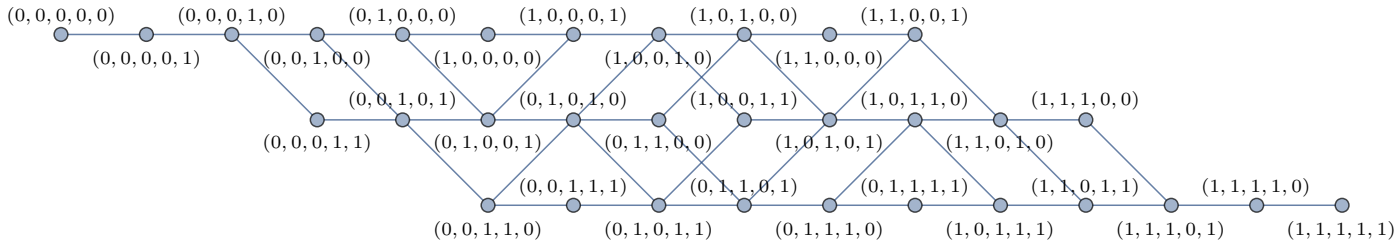
Example 6 (Partition into Chains)**:** Consider the partial order over $\{0,1\}^4$ whose Hasse diagram is represented in Figure 1a. As the set is not totally ordered, we cannot find a chain that contains all its elements. However, we can find a partition of it into two chains. For example, we can pick the following two chains:

$$\mathcal{C}_1 = \{(0,0,0,0),(0,0,0,1),(0,0,1,0),(0,1,0,0),$$
$$(1,0,0,0),(1,0,0,1),(1,0,1,0),(1,1,0,0)\}$$
$$\mathcal{C}_2 = \{(0,0,1,1),(0,1,0,1),(0,1,1,0),(0,1,1,1),$$
$$(1,0,1,1),(1,1,0,1),(1,1,1,0),(1,1,1,1)\}.$$

Note that this decomposition is not unique, and there are other ways of partitioning the set $\{0,1\}^n$ into two chains. As predicted by Dilworth's theorem, the maximum cardinality of an antichain is 2. Indeed, $\mathcal{A} = \{(1,0,0,0),(0,1,1,1)\}$ is an antichain and it is easy to verify that there is no antichain of cardinality 3.



(a) $n = 4$



(b) $n = 5$

Figure 1: Hasse diagram of the partial order over the Hamming cube $\{0,1\}^n$ induced by Proposition 2.

The following lemma, whose proof is deferred to Appendix A, characterizes the maximum cardinality of an antichain of the poset $\{0,1\}^n$ with the order $\prec_1$.

Lemma 1 (Maximum Cardinality of an Antichain)**:** Let $M(n)$ be the maximal number of subsets of $\{1,\dots,n\}$ that share the same sum. Consider the set $\{0,1\}^n$ with the partial order $\prec_1$ and let $\mathcal{A}$ be an antichain. Then,

$$\max_{\mathcal{A}} |\mathcal{A}| = M(n), \tag{17}$$

where the maximum is computed over the set of all antichains.

We are now ready to prove the main result of this paper.

Proof of Theorem 1: Consider the set of synthetic channels $\{W_N^{(i)}\}_{i\in\{0,\dots,N-1\}}$ with the partial order given by Proposition 2. Let $\mathcal{C}' \subseteq \{W_N^{(i)}\}$ be a chain. By Definition 6, $\mathcal{C}'$ is totally ordered. Hence, in order to establish which elements of $\mathcal{C}'$ have a Bhattacharyya parameter smaller than γ, we can use a binary search algorithm, which requires the computation of at most $\lfloor \log|\mathcal{C}'| + 1 \rfloor$ Bhattacharyya parameters.

Let $(\mathcal{C}'_1,\dots,\mathcal{C}'_K)$ be a partition of $\{W_N^{(i)}\}_{i\in\{0,\dots,N-1\}}$ into a minimum number of chains. Clearly, the FP construction problem is equivalent to the problem of establishing which elements of $\mathcal{C}'_i$ have a Bhattacharyya parameter smaller than γ for all $i \in \{1,\dots,K\}$. In order to solve this last problem, the number of Bhattacharyya parameters to be computed is bounded as follows:

$$\begin{aligned}\sum_{i=1}^{K} \lfloor \log|\mathcal{C}'_i| + 1 \rfloor &\le \sum_{i=1}^{K} (\log|\mathcal{C}'_i| + 1) \\ &= K \cdot \left(1 + \sum_{i=1}^{K} \frac{1}{K} \log|\mathcal{C}'_i|\right) \\ &\overset{(a)}{\le} K \cdot \left(1 + \log \sum_{i=1}^{K} \frac{1}{K} |\mathcal{C}'_i|\right) \\ &= K \cdot \left(1 + \log \frac{2^n}{K}\right) = K \cdot \log \frac{2^{n+1}}{K},\end{aligned} \tag{18}$$

where the inequality (a) is an application of Jensen's inequality.

Let $\mathcal{A}'$ be an antichain with the maximum cardinality. By Definition 6, every pair of synthetic channels in $\mathcal{A}'$ is not comparable. Hence, in order to solve the FP construction problem, we necessarily need to compute the Bhattacharyya parameter of all the elements of $\mathcal{A}'$. This requires the computation of

$$|\mathcal{A}'| = K \tag{19}$$

Bhattacharyya parameters, where the equality comes from Dilworth's theorem.

The set $\{W_N^{(i)}\}_{i\in\{0,\dots,N-1\}}$ with the partial order given by Proposition 2 is order-isomorphic to the set $\{0,1\}^n$ with the partial order $\prec_1$. Hence, by applying Lemma 1, we have that

$$K = M(n). \tag{20}$$

By combining (18), (19), and (20), the thesis immediately follows. ∎

The result that we have just proved tightly bounds the *number* of synthetic channels whose Bhattacharyya parameter has to be computed in order to solve the FP construction problem. Let us now describe *how to find* these synthetic channels.

Following the reasoning of the proof above, in order to solve the FP construction problem, we need to find a partition into chains of the set of synthetic channels. Then, for each chain, we establish which of the synthetic channels is reliable via a binary search algorithm. It remains to discuss how to find the partition into chains.

Consider the bipartite graph $G = (U, V, E)$, where $U = V = \{0,1\}^n$ and where (u, v) is an edge in G if and only if $u \prec_1 v$. The graph G is represented in Figure 2 for $n = 4$. Recall that a matching is a set of edges without common vertices. Given a matching M containing m edges, we can associate to it the partition of $\{0,1\}^n$ defined as follows: for each edge (x, y) in M, include x and y in the same subset.

Suppose that x and y belong to the same subset. Then, there are two possibilities: either $(x, y) \in M$, which implies that $x \prec_1 y$; or there exists a set of intermediate vertices $z_1, \dots, z_k$ such that $(x, z_1), (z_1, z_2), \dots, (z_{k-1}, z_k), (z_k, y) \in M$, which implies that $x \prec_1 z_1 \prec_1 \cdots \prec_1 z_k \prec_1 y$. In both cases, x and y are comparable. Hence, P is a partition of $\{0,1\}^n$ into chains. Note also that the partition P contains $|U| - |M| = 2^n - m$ chains.

Similarly, given a partition P of $\{0,1\}^n$ into p chains, we can associate to it the set of edges M defined as follows. For $i \in \{1, \dots, p\}$, let $P_i = \{x_1^{(i)}, \dots, x_k^{(i)}\}$ be a chain of P. Then, we include the edges $(x_1^{(i)}, x_2^{(i)}), \dots, (x_{k-1}^{(i)}, x_k^{(i)})$ in M. Clearly, M is a matching and it contains $|U| - |P| = 2^n - p$ edges.

In conclusion, we have described a way to associate the matchings of the graph G to the partitions of $\{0,1\}^n$ into chains and vice versa. Therefore, in order to find the partition of $\{0,1\}^n$ containing the smallest number of chains, it suffices to find a maximum matching for the bipartite graph G. The last one is a classical problem in graph theory and it can be solved, e.g., via the Ford-Fulkerson algorithm in $O(|U| \cdot |E|) \leq O(N^3)$ [20] or via the Hopcroft-Karp algorithm in $O(\sqrt{|U|} \cdot |E|) \leq O(N^{5/2})$ [21].

V. Conclusions

In this work, we consider the problem of constructing a polar code of block length $N = 2^n$ and we show that, by taking advantage of the partial order described in [1], [2], we need to compute the reliability of roughly a fraction $1/\log^{3/2} N$ of the synthetic channels. This result is universal in the sense that it holds for the transmission over any BMS channel.

The idea of the proof consists in relating the construction problem to the problem of computing the maximum cardinality of an antichain for a suitably defined poset. In particular, we prove that a lower bound to the number of synthetic channels whose reliability has to be computed is equal to the maximum cardinality of an antichain. Furthermore, this bound is tight up to a multiplicative factor scaling as $\log \log N$. Eventually, we show that the maximum cardinality of an antichain for the poset taken into account is equal to the maximal number of subsets of $\{1, \dots, n\}$ that share the same sum. Such a sequence is the integer sequence A025591 in [14] and it scales as $N/\log^{3/2} N$.

In order to establish which are the indices of these $N/\log^{3/2} N$ synthetic channels, we need to solve a maximum matching problem on a bipartite graph, which can be done in $O(N^{5/2})$. Note that this operation has to be performed only once, since by computing the reliability of those synthetic channels we can solve the construction problem for any BMS channel.

Let us point out that the main idea of the proof technique is completely general in the sense that it does not depend on the particular order described in [1], [2]. Indeed, suppose that a new partial order on the synthetic channels of polar codes is found. Then, in order to improve the bounds on the the number of synthetic channels to be considered, it suffices to compute the maximum cardinality of an antichain for the poset induced by the new order.

Acknowledgement

The work of M. Mondelli and R. Urbanke was supported by grant No. 200021_166106 of the Swiss National Science Foundation. M. Mondelli was also supported by the Dan David Foundation.

Appendix

A. Maximum Cardinality of Antichain: Proof of Lemma 1

Let $\mathcal{P}([n])$ denote the set of subsets of $\{1, \dots, n\}$. Consider the following order relation: let $x = \{x_1, x_2, \dots, x_k\}$ and $y = \{y_1, y_2, \dots, y_j\}$ be elements of $\mathcal{P}([n])$ with $x_1 < x_2 < \cdots < x_k$ and $y_1 < y_2 < \cdots < y_j$; then we define $x \preceq_2 y$ if and only if $k \leq j$ and $x_i \leq y_i$ for all $i \in \{1, \dots, k\}$. In words, $x \preceq_2 y$ if and only if x has at most as many elements as y and, by ordering them in an increasing fashion, the i-th element of x is not larger than the i-th element of y.

The following lemma proves that the set $\mathcal{P}([n])$ with the order $\prec_2$ and the set $\{0,1\}^n$ with the order $\prec_1$ are essentially the same.

Lemma 2 (Order-Isomorphism)**:** The set $\mathcal{P}([n])$ with the order $\prec_2$ is order-isomorphic to the set $\{0,1\}^n$ with the order $\prec_1$.

Proof: By definition of order-isomorphism, in order to prove the claim, we need to find a bijective function f from $\{0,1\}^n$ to $\mathcal{P}([n])$ such that $x \prec_1 y$ if and only if $f(x) \prec_2 f(y)$ for every $x, y \in \{0,1\}^n$.

Consider the function $f : \{0,1\}^n \to \mathcal{P}([n])$ defined as follows. Given $x = (x_n, x_{n-1}, \dots, x_1) \in \{0,1\}^n$, we have that $i \in f(x)$ if and only if $x_i = 1$. In words, we associate to a sequence of n bits the set of indices corresponding to the 1s. Note that we index the sequence from right to left, i.e., the left-most bit of the sequence has index n and the right-most bit of the sequence has index 1.

Assume that $x \prec_1 y$. This means that y is obtained from x by applying addition and left-swap operators. Then, x has at most as many 1s as y, which implies that $f(x)$ has at most as many elements as $f(y)$. Furthermore, the 1s of y are

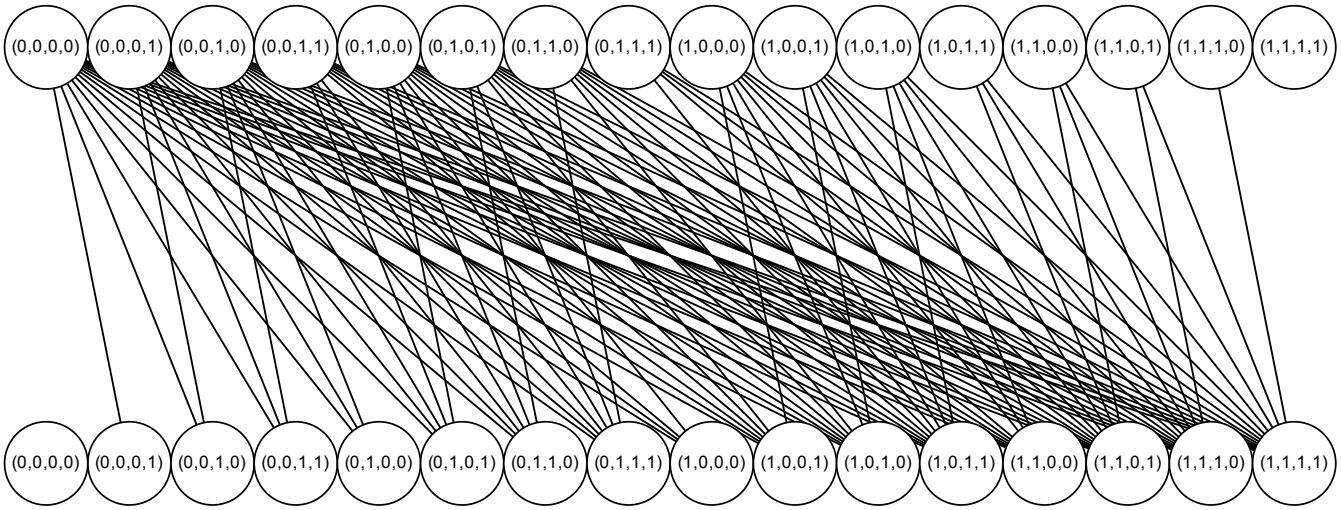


Figure 2: Bipartite graph $G = (U, V, E)$, where $U = V = \{0,1\}^4$ and where (u, v) is an edge in G if and only if $u \prec_1 v$.

placed more to the left than the 1s of x, which implies that i-th element of x is not larger than the i-th element of y for all i. Hence, $f(x) \prec_2 f(y)$. Analogously, we can prove that $f(x) \prec_2 f(y)$ implies that $x \prec_1 y$, which yields the desired claim. ■

Since $(\mathcal{P}([n]), \prec_2)$ is order-isomorphic to $(\{0,1\}^n, \prec_1)$, it suffices to compute the maximum cardinality of an antichain of the former poset. To do so, let us define the concept of rank function.

Definition 7 (Rank Function)**:** Given a poset P with the order $\prec$, a *rank function* is a function $\rho : P \to \mathbb{N}$ that fulfills the following properties:

1) if x is a minimal[1] element of P, then $\rho(x) = 0$;
2) if y covers[2] x , then $\rho(y) = \rho(x) + 1$.

If a poset is equipped with a rank function ρ, we say that the element x has rank $\rho(x)$. The set $\mathcal{P}([n])$ with the order $\prec_2$ is equipped with a rank function, as proved in the following lemma.

Lemma 3 (Rank Function for $\mathcal{P}([n]), \prec_2)$)**:** Given $x = \{x_1, x_2, \ldots, x_k\} \in \mathcal{P}([n])$, define

$$\rho(x) = \sum_{i=1}^{k} x_i. \tag{21}$$

Then, ρ is the rank function for the set $\mathcal{P}([n])$ with the order $\prec_2$.

Proof: Clearly, $x = \emptyset$ is the unique minimal element of the poset $\mathcal{P}([n])$ with the order $\prec_2$. Furthermore, $\rho(x) = 0$, which proves the first property of Definition 7.

Assume now that y covers x. Then, either $y = x \cup \{1\}$ or x and y differ only in 1 element, say the i-th element, and $y_i = x_i + 1$. In both these cases, $\rho(y) = \rho(x) + 1$, which proves the second property of Definition 7. ■

Let P be a poset equipped with a rank function. If every maximal[3] element has the same rank, call it $r_{\max}$, then we say that P is a *graded* poset and we can decompose it as

$$P = P_0 \cup P_1 \cup \cdots \cup P_{r_{\max}}, \tag{22}$$

where P_i contains all the elements of P with rank i. Every chain with the maximum cardinality passes through exactly one element of each of the subsets P_i, starting from P_0, then P_1, and so on.

Note that if two elements of a poset have the same rank, then they are not comparable. Therefore, for all $i \in \{1, \ldots, r_{\max}\}$, the subset P_i is an antichain. The following definition relates these antichains to the antichain with the maximum cardinality.

Definition 8 (Sperner Property)**:** Let P be a graded poset and let P_i be the antichain that contains all the elements of P with rank i. We say that P has the *Sperner property* if the maximum cardinality of an antichain is equal to $\max_i |P_i|$.

Lemma 4 (Sperner Property for $(\mathcal{P}([n]), \prec_2)$)**:** The poset $\mathcal{P}([n])$ with the order $\prec_2$ has the Sperner property.

The proof of the result above is algebraic and it follows from Theorem 4.1 of [22] (see also [22, Section 4.1.2] for a more detailed discussion). Eventually, we are ready to prove Lemma 1.

Proof of Lemma 1: Consider the poset $\mathcal{P}([n])$ with the order $\prec_2$. By Lemma 3, its rank function is given by (21). Furthermore, the maximal element is unique, hence the poset

[1]We say that x is a *minimal* element of P if there is no $y \in P$ such that $y \prec x$.

[2]We say that y *covers* x if $x \prec y$ and there is no other element z such that $x \prec z \prec y$.

[3]We say that x is a *maximal* element of P if there is no $y \in P$ such that $x \prec y$.

is graded. This maximal element is $\{1, 2, \ldots, n\}$ and it has rank

$$r_{\max} = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}. \tag{23}$$

Furthermore, by Lemma 4, $(\mathcal{P}([n]), \prec_2)$ has the Sperner property. Hence, the cardinality of the largest antichain is given by

$$\max_{i \in \{0,\ldots,r_{\max}\}} |P_i| = M(n),$$

where $M(n)$ is defined as the maximal number of subsets $\{1, \ldots, n\}$ that share the same sum. Since $(\mathcal{P}([n]), \prec_2)$ is order-isomorphic to $(\{0,1\}^n, \prec_1)$ by Lemma 2, the thesis immediately follows. ∎

As a final remark, let us point out other two interesting properties of $(\mathcal{P}([n]), \prec_2)$ that follow from the discussion in [22, Section 4.1.2]:

- The poset $\mathcal{P}([n])$ with the order $\prec_2$ is *rank symmetric*, i.e., $|P_i| = |P_{r_{\max}-i}|$ for all $i \in \{0, \ldots, r_{\max}\}$, where $r_{\max}$ is given by (23).
- The poset $\mathcal{P}([n])$ with the order $\prec_2$ is *rank unimodal*, i.e., there is a j such that $|P_0| \leq |P_1| \leq \cdots \leq |P_j| \geq |P_{j+1}| \geq \cdots \geq |P_{r_{\max}}|$.

As a result, we have that

$$\max_{i \in \{0,\ldots,r_{\max}\}} |P_i| = |P_{\lfloor r_{\max}/2 \rfloor}| = |P_{\lceil r_{\max}/2 \rceil}|.$$

In words, the subset(s) P_i with the maximum cardinality correspond to the middle rank(s). This means that $M(n)$ is equal to the number of subsets of $\{1, \ldots, n\}$ that have sum equal to $\lfloor n(n+1)/4 \rfloor$ or to $\lceil n(n+1)/4 \rceil$.

B. Asymptotic Formula for $M(n)$: Proof of Theorem 2

Proof of Theorem 2: Recall that $M(n)$ is defined as the maximal number of subsets of $\{1, \ldots, n\}$ that share the same sum. Clearly, for any $K \in \mathbb{N}$, the number of subsets of $\{1, \ldots, n+1\}$ with sum K is no smaller than the number of subsets of $\{1, \ldots, n\}$ with sum K. Hence, $M(n)$ is an increasing sequence and its limit is equal to the limit of any of its subsequences. The rest of the proof consists in showing that a suitably defined subsequence of $M(n)$ has the asymptotic behavior given by (16).

From the discussion at the end of Appendix A, we have that the integer K that maximizes the number of subsets of $\{1, \ldots, n\}$ with sum K is $\lfloor n(n+1)/4 \rfloor$ or $\lceil n(n+1)/4 \rceil$. Assume now that $n \equiv 0$ or $n \equiv 3$ modulo 4. Then, we have that

$$\frac{n(n+1)}{4} \in \mathbb{N}.$$

Furthermore, we claim that $M(n)$ is equal to the number of choices of $+$ and $-$ signs such that

$$\pm 1 \pm 2 \ldots \pm n = 0. \tag{24}$$

To see this, let A be a subset of $\{1, \ldots, n\}$ with sum $n(n+1)/4$. Then, the set $\{1, \ldots, n\} \setminus A$ has also sum $n(n+1)/4$. By associating the positive sign to the elements of A and the negative sign to the elements of $\{1, \ldots, n\} \setminus A$, we have that the overall sum is 0. As a result, we have found a bijection between the set of subsets of $\{1, \ldots, n\}$ with sum $n(n+1)/4$ and the set of choices of $+$ and $-$ signs such that (24) holds.

Define $S(n)$ as the number of choices of $+$ and $-$ signs such that (24) holds. From the discussion in the previous paragraph, we conclude that $S(n) = M(n)$ for $n \equiv 0$ or $n \equiv 3$ modulo 4. Thus, the asymptotic behavior of $S(n)$ for $n \equiv 0$ or $n \equiv 3$ modulo 4 is the same as the asymptotic behavior of $M(n)$.

In [23, Theorem 2.1], it is proved that $S(n)$ is equal to the coefficient of $x^{n(n+1)/4}$ in the expansion of $\prod_{i=1}^{n}(1+x^i)$ and it is it conjectured that

$$S(n) = \sqrt{\frac{6}{\pi}} \frac{2^n}{n^{3/2}}(1 + o(1)).$$

This conjecture is proved in [24], which implies our desired result. ∎

REFERENCES

[1] C. Schürch, "A partial order for the synthesized channels of a polar code," in *Proc. of the IEEE Int. Symposium on Inform. Theory (ISIT)*, Barcelona, Spain, July 2016, pp. 220–224.

[2] M. Bardet, V. Dragoi, A. Otmani, and J.-P. Tillich, "Algebraic properties of polar codes from a new polynomial formalism," in *Proc. of the IEEE Int. Symposium on Inform. Theory (ISIT)*, Barcelona, Spain, July 2016, pp. 230–234.

[3] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[4] M. Mondelli, S. H. Hassani, and R. Urbanke, "Unified scaling of polar codes: Error exponent, scaling exponent, moderate deviations, and error floors," *IEEE Trans. Inform. Theory*, vol. 62, no. 12, pp. 6698–6712, Dec. 2016.

[5] E. Arıkan and I. E. Telatar, "On the rate of channel polarization," in *Proc. of the IEEE Int. Symposium on Inform. Theory (ISIT)*, Seoul, South Korea, July 2009, pp. 1493–1495.

[6] S. H. Hassani, K. Alishahi, and R. Urbanke, "Finite-length scaling for polar codes," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 5875–5898, Oct. 2014.

[7] D. Goldin and D. Burshtein, "Improved bounds on the finite length scaling of polar codes," *IEEE Trans. Inform. Theory*, vol. 60, no. 11, pp. 6966–6978, Nov. 2014.

[8] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inform. Theory*, vol. 61, no. 5, pp. 2213–2226, May 2015.

[9] R. Mori and T. Tanaka, "Performance and construction of polar codes on symmetric binary-input memoryless channels," in *Proc. of the IEEE Int. Symposium on Inform. Theory (ISIT)*, Seoul, South Korea, July 2009, pp. 1496–1500.

[10] ——, "Performance of polar codes with the construction using density evolution," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 519–521, July 2009.

[11] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3221–3227, Nov. 2012.

[12] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, pp. 6562–6582, Oct. 2013.

[13] R. Pedarsani, H. Hassani, I. Tal, and E. Telatar, "On the construction of polar codes," in *Proc. of the IEEE Int. Symposium on Inform. Theory (ISIT)*, St. Petersberg, Russia, Aug. 2011, pp. 11–15.

[14] N. J. A. Sloane, "The On-Line Encyclopedia of Integer Sequences," published electronically at `https://oeis.org`.

[15] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.

[16] K. Engel, *Sperner Theory*. Cambridge University Press, 1997.

[17] R. P. Stanley, *Enumerative Combinatorics*, 2nd ed., ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2011, vol. 1.

[18] R. P. Dilworth, "A decomposition theorem for partially ordered sets," *Annals Math.*, vol. 51, no. 1, pp. 161–166, 1950.

[19] J. A. Bondy and U. S. R. Murty, *Graph Theory*. Springer, 2008.

[20] L. R. Ford and D. R. Fulkerson, "Maximal flow through a network," *Canadian Journal of Mathematics*, vol. 8, no. 3, pp. 399–404, 1956.

[21] J. E. Hopcroft and R. M. Karp, "An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs," *SIAM Journal on Computing*, vol. 2, no. 4, pp. 225–231, 1973.

[22] R. P. Stanley, "Some applications of algebra to combinatorics," *Discrete Applied Math.*, vol. 34, pp. 241–277, 1991.

[23] D. Andrica and I. Tomescu, "On an integer sequence related to a product of trigonometric functions, and its combinatorial relevance," *Journal of Integer Sequences*, vol. 5, 2002, Article 02.2.4.

[24] B. D. Sullivan, "On a conjecture of Andrica and Tomescu," *Journal of Integer Sequences*, vol. 16, 2013, Article 13.3.1.