sensors

MDPI

Article

Bhattacharyya Parameter of Monomial Codes for the Binary Erasure Channel: From Pointwise to Average Reliability

Vlad-Florin Drăgoi [1,2,*] and Gabriela Cristescu [1]

[1] Faculty of Exact Sciences, Aurel Vlaicu University of Arad, 2 Elena Dragoi Street, 310130 Arad, Romania; gabriela.cristescu@uav.ro
[2] LITIS, University of Rouen Normandie, Avenue de l'Université, 76801 Saint-Etienne-du-Rouvray, France
* Correspondence: vlad.dragoi@uav.ro

Abstract: Monomial codes were recently equipped with partial order relations, a fact that allowed researchers to discover structural properties and efficient algorithm for constructing polar codes. Here, we refine the existing order relations in the particular case of the binary erasure channel. The new order relation takes us closer to the ultimate order relation induced by the pointwise evaluation of the Bhattacharyya parameter of the synthetic channels, which is still a partial order relation. To overcome this issue, we appeal to a related technique from network theory. Reliability network theory was recently used in the context of polar coding and more generally in connection with decreasing monomial codes. In this article, we investigate how the concept of average reliability is applied for polar codes designed for the binary erasure channel. Instead of minimizing the error probability of the synthetic channels, for a particular value of the erasure parameter p, our codes minimize the average error probability of the synthetic channels. By means of basic network theory results, we determine a closed formula for the average reliability of a particular synthetic channel, that recently gain the attention of researchers.

Keywords: Bhattacharyya parameter; average reliability; monomial code; polar code; order relation; binary erasure channel

Citation: Drăgoi, V.-F.; Cristescu, G. Bhattacharyya Parameter of Monomial Codes for the Binary Erasure Channel: From Pointwise to Average Reliability. *Sensors* **2021**, *21*, 2976. https://doi.org/10.3390/s21092976

Academic Editor: Dimitrie Popescu

Received: 25 January 2021
Accepted: 21 April 2021
Published: 23 April 2021

1. Introduction

One of the most striking developments in coding theory in the last two decades is probably the theory around polar codes. In his seminal article [1], Arikan demonstrated, for the first time, that one could achieve the capacity of binary discrete memoryless channels (BDMC), using both efficient encoding and efficient decoding algorithms. The so-called polar codes are now present in the fifth generation (5G) technology [2]. Indeed, polar code was elected as the standard coding technique for the control channel in support of the enhanced mobile broadband service, one of the major parts in 5G wireless network technology. Getting back to the three principal directions on which coding theory evolved, polar coding seemed to be unrelated to classical algebraic coding. Typically, the construction of polar codes does not come from any particular structure in the code, but rather from the process of channel polarization. However, polar codes are closely related to Reed–Muller codes, as pointed out even by Arikan [1]. Hence, polar and Reed-Muller code share a common algebraic description [3,4]. More precisely, they are sub-classes of a larger family of algebraic codes called decreasing monomial codes (DMC). The structure underlying DMCs and its algebraic formalism was applied in conjunction with other fields, e.g., in the context of quantum error correcting codes [5–7], post-quantum cryptography [8–10] and network reliability [11–13].

Several challenges regarding polar coding, among which the efficient construction of polar codes given a specific BDMC, were proposed. Arikan's initial technique [1] was improved by several authors [14–22]. Let W denote a BDMC, m a fixed integer and $\boldsymbol{u}$ a binary vector of length m. The main idea of the construction of polar codes is to estimate

the reliability of the synthetic channels $\{W^{\boldsymbol{u}} \mid \boldsymbol{u} \in \{0,1\}^m\}$. For that, one might use the Bhattacharyya parameter $\mathcal{B}(W^{\boldsymbol{u}})(p)$, where p denotes the error probability of the channel W. The message parts of a polar code of length 2^m and dimension k are allocated to the k sub-channels $W^{\boldsymbol{u}}$ having the smallest $\mathcal{B}(W^{\boldsymbol{u}})$. Hence, one might classify the set of $W^{\boldsymbol{u}}$ into "good" (reliable) or "bad" (non-reliable). For a fixed value of p, the values $\mathcal{B}(W^{\boldsymbol{u}})(p)$ are put in order. In other words, when the parameter p is fixed, any distinct pair of channels $W^{\boldsymbol{u}}, W^{\boldsymbol{v}}$ satisfies either $\mathcal{B}(W^{\boldsymbol{u}})(p) \leq \mathcal{B}(W^{\boldsymbol{v}})(p)$ or $\mathcal{B}(W^{\boldsymbol{v}})(p) \leq \mathcal{B}(W^{\boldsymbol{u}})(p)$. In this case, we say that a channel $W^{\boldsymbol{u}}$ is point-wise more reliable than a channel $W^{\boldsymbol{v}}$. However, when considering the whole interval $p \in [0,1]$, ranking the synthetic channels becomes complicated. In this case, we say that $W^{\boldsymbol{u}}$ is globally more reliable than $W^{\boldsymbol{v}}$, and write $\boldsymbol{u} \leq \boldsymbol{v}$, if and only if $\forall p \in [0,1]$, $\mathcal{B}(W^{\boldsymbol{u}})(p) \leq \mathcal{B}(W^{\boldsymbol{v}})(p)$.

Estimating how reliable a synthetic channel is can be done in several ways, Monte Carlo simulations being among the most common ones. Arikan, in his seminal paper [1], proposed such a method for determining the error probability of each synthetic channel which yields a relatively efficient generating algorithm. It could be possible to employ recent development in Monte Carlo such as those in [23,24] in order to improve Arikan's idea. However, the most efficient techniques for constructing polar codes are exploiting order relations between the synthetic channels. One of the most efficient techniques that orders the set of synthetic channels (with respect to the concept of globally more reliable), provides sub-linear complexity construction [14]. It exploits the existence of a partial order (denoted by $\preceq$) on the set of synthetic channels [4]. This partial order is compatible with the notion of being globally more reliable, i.e., $\boldsymbol{u} \preceq \boldsymbol{v} \Rightarrow \boldsymbol{u} \leq \boldsymbol{v}$. Even though $\preceq$ provided a contribution to understanding polar codes, i.e., their structure and construction, simulations show that $\preceq$ is far from ordering $\mathcal{B}(W^{\boldsymbol{u}})$ optimally. Hence, in a recent article, $\preceq$ was refined [25]. Compared to Monte Carlo methods, the techniques based on order relations valid for polar codes require a small number of computations, only for a fraction of the synthetic channels, as they exploit the rules induced by the order relations.

In the analysis of the performance of several families of codes, among which are polar, Reed–Muller, cyclic and BCH codes, the communication channel that received a lot of attention is the binary erasure channel (BEC). When polar codes are designed for BEC(p) (in this particular case, p denotes the erasure probability), all the synthetic channels $\{W^{\boldsymbol{u}} \mid \boldsymbol{u} \in \{0,1\}^m\}$ are also BEC. In this case, the erasure probability of $W^{\boldsymbol{u}}$ is equal to the Bhattacharyya parameter of $W^{\boldsymbol{u}}$. Here, we analyze this particular channel. Our choice is motivated by several results and methods. First of all, the simplicity of this channel makes the theoretical proofs significantly simpler and easier. Moreover, many of the properties that hold for the BEC turn out to be valid for more general channel models. For example, the proof of Reed–Muller codes achieving the capacity of a communication channel started with the BEC [26,27]. Codes that admit a doubly-transitive automorphism group or having large orbits under the action of their permutation group achieve the capacity of the BEC [27,28]. In [29], the authors analyze threshold points for $W^{\boldsymbol{u}}$ in the case of BEC, a fact that allows them to propose sets of asymptotically "good" channels. Recently, in [30] the authors analyzed the Bhattacharyya parameter of polar codes for the BEC using network reliability theory. They have proposed simple approximations of $\mathcal{B}(W^{\boldsymbol{u}})$. These were used to determine sub-intervals of $[0,1]$, where polar codes coincide with Reed–Muller codes. They have also managed to determine new sets of asymptotically "good" channels.

1.1. Polar Codes are Strongly Decreasing Monomial Codes

Polar codes over the BEC satisfy an order relation that is finer than $\preceq$ .. Hence, we define another order relation $\preceq_d$ on the set of monomials on m variables $\mathcal{M}_m = \{\mathbf{1}, x_0, \ldots, x_{m-1}, x_0x_1, \ldots, x_0x_1 \ldots x_{m-1}\}$, coming closer to the $\leq$ relation, i.e., we have

$$\forall f, g \in \mathcal{M}_m \quad f \preceq g \Rightarrow f \preceq_d g \Rightarrow f \leq g.$$

The relation $\preceq_d$ allows one to compare monomials with equal degrees that were not comparable with respect to $\preceq$, e.g., $x_1x_2 \preceq_d x_0x_3$. The idea of $\preceq_d$ came from the link between the set of monomials of degree d in $\mathcal{M}_m$ and the set of partitions/Young diagrams inside the $d \times (m-d)$ grid (see Proposition 3.7.8 in [3]). From that, looking at order relations on partitions came as a natural idea, and the most common one is the dominance order [31]. The order $\preceq_d$ is exactly defined as the dominance order on partitions inside a fixed grid.

The main result in this section can be stated as follows

Theorem 1. *Polar codes over the binary erasure channel are strongly decreasing monomial codes.*

In the proof of this theorem, we will need to demonstrate two useful properties of this new order

- given two monomials in $f, g \in \mathcal{M}_m$ such that $f \preceq_d g$, then, for any multiples fh, gh with $\gcd(h, f) = 1$ and $\gcd(g, h) = 1$ we have $fh \preceq_d gh$, where $\gcd(f, g)$ denotes the greatest common divisor of f, g.
- two particular monomials are the key ingredients in the proof, $x_1x_2 \preceq_d x_0x_3$. We show that for all $p \in [0,1], \mathcal{B}(W^{x_1x_2})(p) \leq \mathcal{B}(W^{x_0x_3})(p)$, and in general that any pair of monomials of degree 2 f, g, satisfying that $f \preceq_d g$ has the property for all $p \in [0,1], \mathcal{B}\left(W^f\right)(p) \leq \mathcal{B}(W^g)(p)$.

Even though $\preceq_d$ gets us closer to the ultimate order relation $\leq$, we know that $\preceq_d$ is a partial order relation. $\preceq_d$ seems to perform as well as the order relation from [25], being much simpler to describe and analyze than the order in [25]. Furthermore, in [25] the authors determine new order relations based on some hypotheses which are not algebraically easy to express, and which are to be tested each time we change the parameters of the code. Figure 1 illustrates how our results fit into state-of-the-art order relations in conjunction with polar coding.

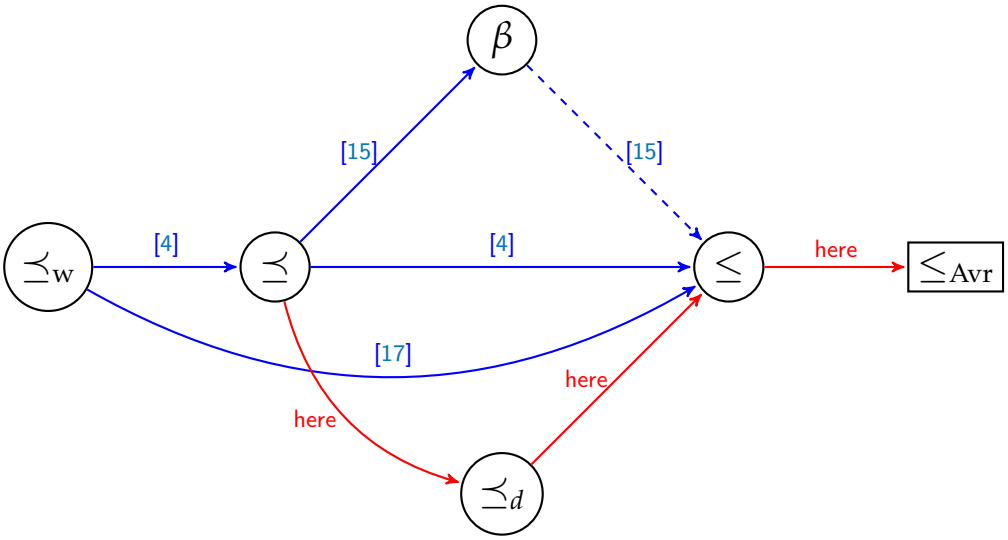


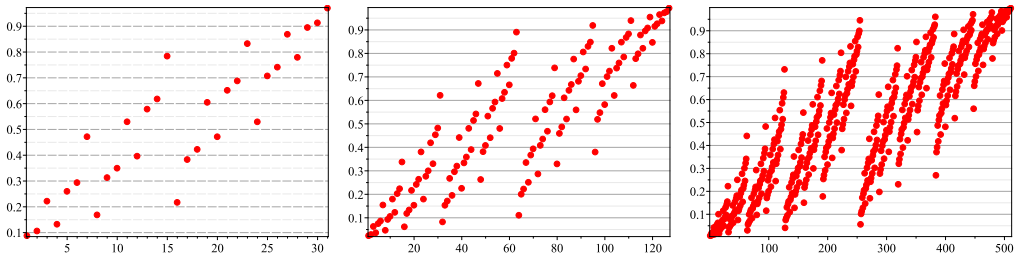
Figure 1. Order ($\preceq_{\mathrm{w}}, \preceq, \preceq_d, \leq$) and preorder relations ($\leq_{\mathrm{Avr}}$) for monomials codes over the BEC. The connections in red are the results coming from this article. The dotted edge from β to $\leq$ represents an order relation that is valid only for a sub-interval of $[0,1]$.

Now, as $\preceq_d$ is thinner than $\preceq$, one could use it in order to determine new sets of comparable monomials and thus generate polar codes in a more efficient manner. Indeed, as more monomials are comparable with respect to $\preceq_d$, they induce new chains in the poset $\{\mathcal{M}_m, \preceq_d\}$. This enables us to reduce the number of non comparable elements and also the number of strongly decreasing monomial codes compared to decreasing monomial codes.

These two ideas are being illustrated as possible methods for reducing the complexity of the construction algorithm, hence directing towards possible practical applications.

1.2. Average Reliability of Synthetic Channels

Hence, we are still left with elements that are not comparable and for which we need to compute $\mathcal{B}(W^{\boldsymbol{u}})$. In order to overcome this issue, we propose an alternative solution. Suppose that the erasure probability of the channel p changes with respect to the uniform distribution over the closed interval $[0,1]$. Instead of constructing, for each p, the corresponding polar code, we propose to construct the best polar code in average. More exactly, we consider the average reliability of the synthetic channels $W^{\boldsymbol{u}}$, $\mathrm{Avr}(W^{\boldsymbol{u}}) = \int_0^1 \mathcal{B}(W^{\boldsymbol{u}})dp$, and choose those $\boldsymbol{u}$ that minimize this quantity. As the average reliability induces a total order relation (see Figure 2), there is only one polar code for a given dimension and length. It is the linear code that minimizes the average error probability for all $p \in [0,1]$. Hence, it might be less efficient than polar codes designed for a particular value of p, but it has the best performance on average.



(**a**) $m = 5$ (**b**) $m = 7$ (**c**) $m = 9$

Figure 2. Average Bhattacharyya parameter. On the x-axis are the integer values of the binary vectors $\boldsymbol{u} \in \{0,1\}^m$, and on the y-axis are the values $\mathrm{Avr}(\mathcal{B}(W^{\boldsymbol{u}}))$.

The preorder $\boldsymbol{u} \preceq_{\mathrm{Avr}} \boldsymbol{v} \Leftrightarrow \mathrm{Avr}(W^{\boldsymbol{u}}) \leq \mathrm{Avr}(W^{\boldsymbol{v}})$ induces a complementarity property with respect to the integral operator over $[0,1]$, as defined in [32,33] in the case of two-terminal networks. We retrieve a similar property, i.e., $\mathrm{Avr}(W^{\boldsymbol{u}}) = 1 - \mathrm{Avr}(W^{\overline{\boldsymbol{u}}})$, where $\overline{\boldsymbol{u}}$ is the bit-wise complement of $\boldsymbol{u}$, in the context of monomial codes. Our simulations have shown that, considering the relation $\preceq_{\mathrm{Avr}}$ in the set of the synthetic channels, in each sub-interval $(i/10, (i+1)/10)$, for $0 \leq i \leq 9$, we have a rough proportion of $2^m/10$ binary vectors $\boldsymbol{u}$. So, roughly speaking, a uniform distribution could be used to approximate the number of $\boldsymbol{u}$ inside each sub-interval, with respect to Avr. However, our result is not constructive, in the sense that it does not fully characterize exactly the $\boldsymbol{u}$ that belong to a specific interval. An answer to this question might provide an extremely efficient method for constructing polar codes and give much more insight into the synthetic channels $W^{\boldsymbol{u}}$.

Threshold Points for Sharp Transitions

Determining the threshold point of $\mathcal{B}(W^{\boldsymbol{u}})$ is in general a difficult task [29,34]. In [29], the authors analyze a particular synthetic channel $W^{(1^i 0^{m-1})}$, for which asymptotic threshold points were determined. The conditions on i and m were further improved in [25]. Based on some basic notions and facts from network theory, we determine an exact formula for the average reliability of $W^{(1^i 0^{m-1})}$. The main result is

Theorem 2. *Let* $\boldsymbol{u} = (1^i 0^{m-i})$. *Then*

$$\mathrm{Avr}(W^{\boldsymbol{u}}) = 1 - \frac{1}{\binom{2^i + 2^{i-m}}{2^i}} \tag{1}$$

This allows us to determine the exact threshold point of this particular channel. Moreover, we demonstrate that for any $i \leq m - \log_2(m) - \log_2(\log_2(m))$, the channel $W^{(1^i 0^{m-i})}$ has an average Bhattacharyya parameter that tends to zero when m goes to infinity, i.e., $W^{(1^i 0^{m-i})}$ is asymptotically "good" on average. Another consequence of our formula is that for any monomial $g \preceq_d x_{m-i+1} \dots x_m$ with $i \leq \log_2(\log_2(m))$ is such that $\mathrm{Avr}(W^g)$ tends to zero when m goes to infinity.

Another significant implication of our result is that any synthetic channel in the $\mathcal{RM}(i, m)$ is asymptotically "good" on average, for any $i \leq \log_2(\log_2(m))$.

1.3. Outline of the Article

The main concepts, notations and properties that are used all over the paper are introduced in Section 2. We introduce the background on coding, referring to the monomial codes in Section 2.1, the polar codes in Section 2.2 and to various manners of comparing them in Section 2.3. The two-terminal networks and their reliability are discussed in Section 2.4. The similarity between the behavior of the Bhattacharyya parameters and the reliability polynomial in assessing the monomial codes is emphasized in Section 2.5. The relationship between the order relations and the corresponding order structures over the set of polar codes over the binary erasure channel is studied in Section 3. The main result in this section proves that the polar codes over the binary erasure channel are strongly decreasing monomial codes. In the same section, two ideas pointing to possible practical applications of the order relation $\preceq_d$ are developed. The concept of average reliability of a synthetic channel is introduced in Section 4. The properties of this operator are studied in Section 4.1 and the relation to the β-expansion is presented in Section 4.2. Finally, the threshold points of the binary erasure polarization sub-channels are determined in Section 4.3. Simulations and numerical examples are included to illustrate all the new results. We conclude our article in Section 5.

2. Background and Preliminary Results

Let us begin by listing some of the usual notations from coding theory that are going to be used in this article. $\mathbb{F}_2$ will denote the finite field with two elements $\{0, 1\}$. Let k, n be two strictly positive integers and $k \leq n$. A code $\mathscr{C}$ of length n and dimension k is a vector sub-space of $\mathbb{F}_2^n$ of dimension k. In this article, we focus out attention on a particular family or linear codes, namely monomial codes. W will be used to denote a communication channel with binary input $x \in \mathbb{F}_2$ and output from an alphabet $y \in \mathcal{Y}$. In particular, we will focus on BEC(p), where the output is $\mathcal{Y} = \{0, 1, ?\}$, ? denoting an erasure and p being the erasure probability. For a more detailed reading of the subject, we recommend [35,36].

2.1. Monomial Codes

Monomial codes are a special class of structured codes. Informally, any code that admits a basis, in which each vector is the evaluation of a monomial, is called a monomial code. In general, monomial codes have a predefined length, i.e., 2^m. Many of the notations, definitions, properties, and results presented in this section are taken from [3].

In this article, binary vectors of length m will be denoted using bold small letters, e.g., $\boldsymbol{u} = (u_0, \dots, u_{m-1}) \in \mathbb{F}_2^m$, with the convention that bits are ordered from left to right, u_0 being the least significant bit. We also define the bit-wise complement of $\boldsymbol{u} \in \{0, 1\}^m$ by $\overline{\boldsymbol{u}} = \mathbf{1}_{\boldsymbol{m}} \oplus \boldsymbol{u}$ (as in [4]), where $\mathbf{1}_{\boldsymbol{m}}$ is the all-ones vector. The set $\{\boldsymbol{u} \in \mathbb{F}_2^m\}$ will be ordered in a natural manner, using the mapping

$$(u_0, \dots, u_{m-1}) \to u = \sum_{i=0}^{m-1} u_i 2^i,$$

and the natural order on the integers. Notice that we compute the value u regardless of the fact that $u_i \in \mathbb{F}_2$. Notice that the relation between $\boldsymbol{u}$ and $\overline{\boldsymbol{u}}$ induces $u + \overline{u} = 2^m - 1$.

We consider multivariate polynomials and monomials defined over the polynomial ring $\mathbb{R}_m = \mathbb{F}_2[x_0, x_1, \ldots, x_{m-1}]/(x_0^2 - x_0, \ldots, x_{m-1}^2 - x_{m-1})$. The usual operators will be employed, i.e., for $f, g \in \mathbb{R}_m$, we denote by $\deg f$ the degree of f, $\gcd(f, g)$ the greatest common divisor of f and g. f/g denotes the quotient of f and g.

Notation 1. *Let m be a strictly positive integer. We denote*

- *monomials: $\boldsymbol{x}^{\boldsymbol{u}} = x_0^{u_0} \cdots x_{m-1}^{u_{m-1}}$, where $\boldsymbol{u} \in \mathbb{F}_2^m$.*
- *support of a monomial: $\mathrm{ind}(g) = \{l_1 \ldots, l_s\}$, where $g = x_{l_1} \ldots x_{l_s}$ and $0 \leq l_1 < l_2 \cdots < l_s \leq m-1$.*
- *a subset of the support of a monomial: $g_{[0,s]} = \gcd(g, \prod_{i=0}^{s} x_i)$.*
- *the set of monomials: $\mathcal{M}_m \stackrel{def}{=} \{\boldsymbol{x}^{\boldsymbol{u}} \mid \boldsymbol{u} = (u_0, \ldots, u_{m-1}) \in \mathbb{F}_2^m\}$.*

Proposition 1 ([37]). *Let $g \in \mathbb{R}_m$ and order the elements in $\mathbb{F}_2^m$ with respect to the decreasing index order. Define the evaluation function*

$$\begin{array}{ccc} \mathbb{R}_m & \to & \mathbb{F}_2^{2^m} \\ g & \mapsto & \mathsf{ev}(g) = \big(g(\boldsymbol{u})\big)_{\boldsymbol{u} \in \mathbb{F}_2^m} \end{array}$$

Then, ev is a bijection defining an isomorphism between the vector spaces $(\mathbb{R}_m, +, \cdot)$ and $(\mathbb{F}_2^n, +, \cdot)$.

Now, we are ready to define the concept of monomial codes.

Definition 1 (Monomial code). *Let $I \subseteq \mathcal{M}_m$ be a finite set of monomials in m variables. The linear code defined by I is the vector subspace $\mathscr{C}(I) \subseteq \mathbb{F}_2^{2^m}$ generated by $\{\mathsf{ev}(f) \mid f \in I\}$ that is called monomial code.*

Proposition 2 ([3]). *For all $I \subseteq \mathcal{M}_m$, the dimension of the monomial code $\mathscr{C}(I)$ is equal to $|I|$.*

Remark 1. *The r^{th} order Reed–Muller code $\mathcal{RM}(r, m) \stackrel{def}{=} \{\mathsf{ev}(g) \mid g \in \mathbb{R}_m, \deg g \leq r\}$ is a monomial code with dimension $k = \sum_{i=0}^{r} \binom{m}{i}$.*

2.2. Polar Codes

In order to define polar codes, we have to introduce the concept of synthetic channels. Consider the channel transformation $W \to (W_2^{(0)}, W_2^{(1)})$ defined in the following manner.

Definition 2 (Synthetic channels). *Let W be a BDMC with output alphabet $\mathcal{Y}$and $x_1, x_2 \in \mathbb{F}_2$ be the inputs and $y_1, y_2 \in \mathcal{Y}$ be the outputs of two copies of W. Define two new channels*

$$\begin{aligned} W^{(1)}(y_1, y_2|x_2) &\stackrel{def}{=} \frac{1}{2} \sum_{x_1 \in \mathbb{F}_2} W(y_1|x_1) W(y_2|x_1 \oplus x_2) \\ W^{(0)}(y_1, y_2, x_2|x_1) &\stackrel{def}{=} \frac{1}{2} W(y_1|x_1) W(y_2|x_1 \oplus x_2). \end{aligned}$$

For any $\boldsymbol{u} = (u_0, \ldots, u_{m-1}) \in \{0,1\}^m$, we define $W^{\boldsymbol{u}} = ((W^{u_{m-1}})^{\cdots})^{u_0}$ as in [4]. Moreover, we extend the notation to monomials, by $W_m^f = W^{\boldsymbol{u}}$ where $f = \boldsymbol{x}^{\boldsymbol{u}} \in \mathcal{M}_m$. We are using the index m in W_m^f to precisely identify the number of variables on which f is expressed. For example, if $\boldsymbol{u} = (1,0,0,1,1)$, we have $W_5^f = W_5^{x_0 x_3 x_4} = \big(W_1^{x_4}\big)_4^{x_0 x_3} = \Big(\big(W_1^{x_4}\big)_1^{x_3}\Big)_3^{x_0}$.

Definition 3. *Let W be a BDMC with output alphabet $\mathcal{Y}$. Then, the Bhattacharyya parameter of the channel W is*

$$\mathcal{B}(W) = \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}. \tag{2}$$

Remark 2. *Let W be a* BEC(p)*, then, we have that $\mathcal{B}(W^{x_0}) = \mathcal{B}\left(W^{(1)}\right) = 2p - p^2$ and $\mathcal{B}(W^{\mathbf{1}}) = \mathcal{B}\left(W^{(0)}\right) = p^2$.*

Definition 4. *The polar code of length $n = 2^m$ and dimension k devised for the channel W is the linear code obtained by selecting the set of k synthetic channels with the smallest $\mathcal{B}(W^{\boldsymbol{u}})$ values among all $\boldsymbol{u} \in \{0,1\}^m$.*

Moreover, we define the relation

$$f \leq g \Leftrightarrow \boldsymbol{u} \leq \boldsymbol{v} \Leftrightarrow \mathcal{B}(W^{\boldsymbol{u}})(p) \leq \mathcal{B}(W^{\boldsymbol{v}})(p), \forall p \in [0,1]. \tag{3}$$

The relation (3) is called universal, i.e., two monomials f, g satisfying $f \leq g$ are always comparable for any p and any m. This property can be used when constructing polar codes by storing a table with all such monomials. However, there might be several monomials bigger than f which are not comparable pairwise (see, for example, [4,14]). Indeed, one can easily verify that $\leq$ is a well-defined order relation (reflexive, anti-symmetric, and transitive), and thus, induces a poset on the set of monomials. In some particular cases, the order $\leq$ becomes total (all elements are ordered in a chain), e.g., when $W =$ BEC and $m \leq 4$. However, in general, $\leq$ is a partial order, even in the case of $W =$ BEC (starting from $m = 5$), as pointed out in [3,11,30].

Proposition 3. *Let $m \geq 5$ be an integer and $W =$ BEC. Then, $\{\mathcal{M}_m, \leq\}$ is a poset.*

For simplification, when we refer to ordering the Bhattacharyya parameters, we will just write $\mathcal{B}(W^{\boldsymbol{u}}) \leq \mathcal{B}(W^{\boldsymbol{v}})$.

2.3. Weakly Decreasing and Decreasing Monomial Codes

Definition 5. *Let f and g be two monomials in $\mathcal{M}_m$.*

- *The $\preceq_{\mathrm{w}}$ order between f and g is defined as*

$$f \preceq_{\mathrm{w}} g \quad \textit{iff} \quad f|g.$$

- *The $\preceq$ order between f and g is defined as*
  - *when $\deg(f) = \deg(g) = s$ and $f = x_{i_1} \dots x_{i_s}$, $g = x_{j_1} \dots x_{j_s}$ we have*

$$f \preceq g \quad \textit{iff} \quad \forall\, 1 \leq \ell \leq s \quad i_\ell \leq j_\ell.$$

  - *when $\deg(f) < \deg(g)$ we have*

$$f \preceq g \quad \textit{iff} \quad \exists g^* \in \mathcal{M}_m \textit{ s.t. } f \preceq g^* \preceq_{\mathrm{w}} g.$$

The two order relations $\preceq_{\mathrm{w}}$ and $\preceq$ are well defined. $\preceq_{\mathrm{w}}$ was already used in the case of polar codes, but in a completely different context by Mori and Tanaka in [17]. In their case, the purpose was to tighten the bounds of the error block probability of a polar code designed for the BEC family.

Notice that $\preceq_{\mathrm{w}}$ is weaker than $\preceq$, meaning that $\forall f, g \in \mathcal{M}_m f \preceq_{\mathrm{w}} g \Rightarrow f \preceq g$. The inverse is not always true: taking, for example, $f = x_0x_2$ and $g = x_1x_2$ it follows by definition that $f \preceq g$ but $f \npreceq_{\mathrm{w}} g$. We also remark that $\mathbf{1}$ is the smallest element both for $\preceq$ and for $\preceq_{\mathrm{w}}$, and we have

$$\mathbf{1} \preceq x_0 \preceq x_1 \preceq \cdots \preceq x_{m-1}.$$

Definition 6. *Let f and g be two monomials in $\mathcal{M}_m$ such that $f \preceq g$ and $I \subset \mathcal{M}_m$.*

- *We define the closed interval $[f,g]_{\preceq} = \{h \in \mathcal{M}_m \mid f \preceq h \preceq g\}$.*
- *$I \in \mathcal{M}_m$ is called a decreasing set if and only if ($f \in I$ and $g \preceq f$) implies $g \in I$.*
- *Let $I \in \mathcal{M}_m$ be a decreasing set. Then, $\mathscr{C}(I)$ is called a decreasing monomial code.*

Polar codes were recently related to network theory. In [30], the authors make a connection between the Bhattacharyya parameter of a synthetic channel and the reliability polynomial of a two-terminal network. Following the same path, we introduce in the next subsection all the required preliminaries in reliability and network theory.

2.4. Two-Terminal Networks

Definition 7. *Let n be a strictly positive integer. We say that $\boldsymbol{N}$ is a two-terminal network (2TN) of size n if $\boldsymbol{N}$ is a network made of n identical devices, that has two distinct terminals: an input S, and an output T.*

To any network, $\boldsymbol{N}$ made of n devices we associate two parameters: width (w) and length (l), where w is the cardinal of a "minimal cut" separating S from T, and l is the cardinal of a "minimal path" from S to T, that satisfy

$$n \geq wl \tag{4}$$

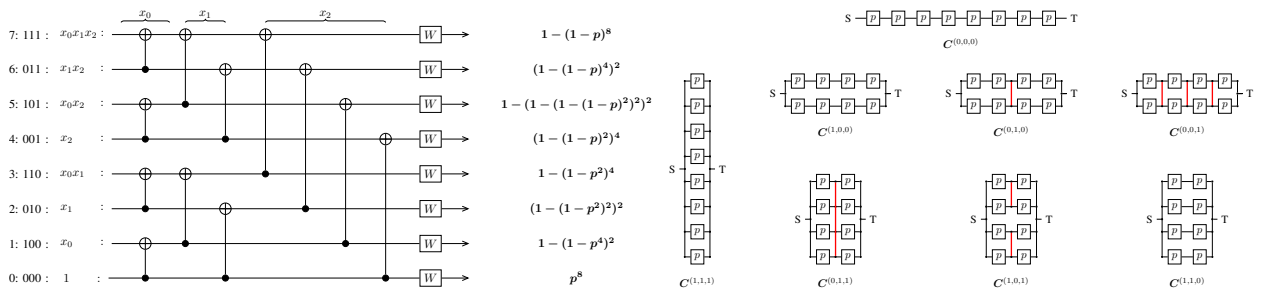
(see Theorem 3 in [38]). The number of devices n is known in the literature as the size of the network. When $n = wl$, we say that $\boldsymbol{N}$ is a minimal 2TN [38].

The composition of $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ can be defined as in [38]. The resulting network is obtained by replacing each device in $\boldsymbol{N}_1$ by a copy of $\boldsymbol{N}_2$. We will denote a composition by $\boldsymbol{C}$, the simplest possible being two devices in series $\boldsymbol{C}^{(0)}$, and two devices in parallel $\boldsymbol{C}^{(1)}$. The composition of $\boldsymbol{C}^{(0)}$ with $\boldsymbol{C}^{(1)}$ is $\boldsymbol{C}^{\boldsymbol{u}} = \boldsymbol{C}^{(0)} \bullet \boldsymbol{C}^{(1)}$, where $\boldsymbol{u} = (0,1)$. The set of all 2^m-size compositions will be denoted by $\mathcal{C}_{2^m}$, and the set of all compositions of width 2^i and length 2^{m-i} by $\mathcal{C}_{2^i,2^{m-i}}$ (see Figure 3b).

Proposition 4 ([39])**.** *Let $m > 0$ and $\boldsymbol{C}^{\boldsymbol{u}} \in \mathcal{C}_{2^m}$. Then, $\boldsymbol{C}^{\boldsymbol{u}}$ is a minimal 2TN of size 2^m, length $l = 2^{m-|\boldsymbol{u}|}$ and width $w = 2^{|\boldsymbol{u}|}$. We also have $\mathcal{C}_{2^m} = \bigcup_{i=0}^{m} \mathcal{C}_{2^i,2^{m-i}}$.*

Theorem 3 ([11])**.** *There is a natural bijection between $\mathcal{C}_{2^m}$ and the set of all $W^{\boldsymbol{u}}$, for any fixed positive integer m.*

In Figure 3, we illustrate the bijection between the two aforementioned sets. More significant is the equality between the reliability polynomial of a composition $\boldsymbol{C}^{\boldsymbol{u}}$ and the Bhattacharyya parameter of $W^{\boldsymbol{u}}$, a fact that is visible from Figure 3b and proven in the next paragraph.



(**a**) $W^{\boldsymbol{u}}$ and $\mathcal{B}(W^{\boldsymbol{u}})$ (**b**) $\mathcal{C}_3$

Figure 3. Combined circuit as defined by Arikan [1], the Bhattacharyya parameter of the corresponding synthetic channels and the compositions in $\mathcal{C}_3$.

Reliability Polynomial

The reliability of $\boldsymbol{N}$ is defined as the probability that S and T are connected (also known as s,t-connectivity) [40]. One of the most common hypotheses considered in network theory is that devices are uniformly and identically supposed to close with a probability $p \in [0,1]$. Hence, the reliability of $\boldsymbol{N}$, denoted by $\mathrm{Rel}(\boldsymbol{N};p)$, can be expressed as a polynomial

$$\mathrm{Rel}(\boldsymbol{N};p) = \sum_{i=0}^{n} N_i(\boldsymbol{N})\ p^i(1-p)^{n-i}. \tag{5}$$

The coefficients $N_i(\boldsymbol{N})$ represent the number of paths from S to T of length i. Several properties regarding the coefficients $N_i(\boldsymbol{N})$, as well as complementarity relations between a 2TN $\boldsymbol{N}$ and its dual $\boldsymbol{N}^{\perp}$, are detailed in [32,41] in the case of hammock networks.

2.5. Bhattacharyya Parameters and Reliability Polynomials

Theorem 4 ([11]). *Let $m > 0$, $\boldsymbol{u} \in \{0,1\}^m$, and $W = \mathrm{BEC}(p)$*

$$\mathcal{B}(W^{\boldsymbol{u}})(p) = \mathrm{Rel}(\boldsymbol{C}^{\boldsymbol{u}};p) \tag{6}$$

where $\mathrm{Rel}(\boldsymbol{C}^{(0)};p) = p^2$ and $\mathrm{Rel}(\boldsymbol{C}^{(1)};p) = 1-(1-p)^2$.

Proposition 5 ([25]). *Let $m > 0$ and $\boldsymbol{u} \in \{0,1\}^m$. Then*

$$\mathcal{B}\left(W^{\overline{\boldsymbol{u}}}\right)(p) = 1 - \mathcal{B}(W^{\boldsymbol{u}})(1-p). \tag{7}$$

This condition expresses the duality of the two corresponding networks, namely $\boldsymbol{C}^{\overline{\boldsymbol{u}}}$, the dual of $\boldsymbol{C}^{\boldsymbol{u}}$ (see [11,32]). Notice that by (7), one has to analyze only $\boldsymbol{u}$ with $|\boldsymbol{u}| \leq m/2$.

3. Polar Codes are Strongly Decreasing Monomial Code Over the BEC

3.1. Definitions and Results

Definition 8. *The $\preceq_d$ order between f and g is defined as*

- *when $\deg(f) = \deg(g) = s$ and $f = x_{i_1} \dots x_{i_s}$, $g = x_{j_1} \dots x_{j_s}$ we have*

$$f \preceq_d g \quad \textit{iff} \quad \forall \ell \in \{1,\dots,s\} \textit{we have} \sum_{k=0}^{\ell} i_{s-k} \leq \sum_{k=0}^{\ell} j_{s-k}.$$

- *when $\deg(f) < \deg(g)$ we have*

$$f \preceq_d g \quad \textit{iff} \quad \exists g^* \in \mathcal{M}_m \textit{ s.t. } f \preceq_d g^* \preceq_{\mathrm{w}} g.$$

Definition 9. *Let f and g be two monomials in $\mathcal{M}_m$, such that $f \preceq g$ and $I \subset \mathcal{M}_m$.*

- *We define the closed interval $[f,g]_{\preceq_d} = \{h \in \mathcal{M}_m \mid f \preceq_d h \preceq_d g\}$.*
- *$I \in \mathcal{M}_m$ is called a strongly decreasing set if, and only if, ($f \in I$ and $g \preceq_d f$) implies $g \in I$.*
- *Let $I \in \mathcal{M}_m$ be a strongly decreasing set. Then, $\mathscr{C}(I)$ is called strongly decreasing monomial code.*

Lemma 1. *The order $\preceq_d$ is a well-defined order relation and $\{\mathcal{M}_m, \preceq_d\}$ forms a Poset.*

The proof of this lemma comes directly from the definition of $\preceq_d$.

Remark 3. *Notice that $x_{i_1} \dots x_{i_s} \preceq x_{j_1} \dots x_{j_s}$ implies that $x_{i_1} \dots x_{i_s} \preceq_d x_{j_1} \dots x_{j_s}$. The converse is no longer true, take for example, the monomials x_0x_3 and x_1x_2.*

Proposition 6. *Let f and g be two monomials with the same degree and x_h be such that $x_h \not| f$ and $x_h \not| g$. Then, we have*

$$f \preceq_d g \text{ iff } x_h f \preceq_d x_h g. \tag{8}$$

The proof of Proposition 6 can be found in Appendix A. In particular, notice that if f, g are co-prime, i.e., $\gcd(f,g) \neq 1$, then

$$f \preceq_d g \text{ iff } f/\gcd(f,g) \preceq_d g/\gcd(f,g). \tag{9}$$

Remark that when the condition on variable x_h is not satisfied, the result does not hold, e.g., $x_2x_3 \preceq_d x_1x_4$, but $x_1x_2x_3$ and x_1x_4 are not comparable with respect to $\preceq_d$. Before we get to our main theorem of this section, the following lemma is required.

Lemma 2. *Let $f, g \in \mathcal{M}_m$, $\deg f = \deg g = 2$, such that $f \preceq_d g$. Then, $\mathcal{B}\left(W_m^f\right) \leq \mathcal{B}\left(W_m^g\right)$.*

Corollary 1. *Let $f = x_{i_1}x_{i_2}$ and $g = x_{j_1}xj_2$ s.t. $f \preceq_d g$. Then, for any monomial $h = x_{l_1} \dots x_{l_t}$ satisfying $i_1 < l_1 < \dots < l_t < i_2$, we have $fh \preceq_d gh$ and $\mathcal{B}\left(W_m^{fh}\right) \leq \mathcal{B}\left(W_m^{gh}\right)$.*

Theorem 5. *Polar codes over the binary erasure channel are strongly decreasing monomial codes.*

The proof of Theorem 5 is given in Appendix A.

Theorem 6. *Reed–Muller codes are strongly decreasing monomial codes, i.e.,*

$$\mathcal{RM}(i,m) = \mathscr{C}([\mathbf{1}, x_{m-i} \dots x_{m-1}]_{\preceq_d}). \tag{10}$$

Proof. The proof follows from $[\mathbf{1}, x_{m-i} \dots x_{m-1}]_{\preceq_d} = [\mathbf{1}, x_{m-i} \dots x_{m-1}]_{\preceq}$ and $\mathcal{RM}(i,m) = \mathscr{C}([\mathbf{1}, x_{m-i} \dots x_{m-1}]_{\preceq})$ (see Proposition 3.3.12 in [3]). □

3.2. Perspectives of Application of $\preceq_d$ in the Construction of Polar Codes

State-of-the-art algorithms for constructing polar codes [14] are using the structure induced by the existing partial order relations on the set of monomials. As explained in [14], the complexity of the algorithm for construction of polar codes is dominated by the cardinality of the largest set of non comparable monomials with respect to $\preceq$. Hence, a finer order relation than $\preceq$ could potentially decrease the complexity of such an algorithm. As $\preceq_d$ is thinner than $\preceq$, we will seek, through examples, how many non-comparable monomials with respect to $\preceq$ are comparable with respect to $\preceq_d$. Typically, our procedure can be used for a more efficient enumeration of sets of non-comparable elements in the new poset $\{\mathcal{M}_m, \preceq_d\}$. The longest antichain in the poset gives a direct intuition on how efficient the construction algorithm can be. Indeed, when estimating the reliability of the synthetic channels in order to construct a polar code, one need to estimate the reliability of the non comparable elements. Hence, having a finner poset, where the maximum length antichain becomes smaller, induces a more efficient construction algorithm. In order to make things clear, we will explain, in view of two distinct perspectives (reducing the number of non comparable elements and reducing the number of codes of fixed dimension), how $\preceq_d$ induces more efficient construction rules for the polar code.

3.2.1. Reducing the Number of Non-Comparable Monomials

For $\{\mathcal{M}_m, \preceq\}$, the middle of the poset is also a maximum length antichain. Let us explain the concept of middle of $\{\mathcal{M}_m, \preceq\}$. A monomial g is situated in the middle of the poset if any chain from g to $\mathbf{1}$ (the infimum of the poset) has length equal to any chain from g to $x_0 \dots x_{m-1}$ (the supremum of the poset). Clearly, two distinct monomial f, g that are in the middle of the poset are non-comparable. Moreover, notice that not every poset admits a middle, with respect to our definition. For example, $\{\mathcal{M}_m, \preceq\}$ has a middle, since

it is a graded poset (see [11,14,31] for more details). However, $\{\mathcal{M}_m, \preceq_d\}$ is not graded and it does not admit a middle. When $m = 4$ (see Figure 4), the middle of $\{\mathcal{M}_m, \preceq\}$ is the set $\{x_1x_2, x_0x_3\}$. As $\preceq_d$ is thinner than $\preceq$, it could be possible to reduce the number of non-comparable elements from the middle of $\{\mathcal{M}_m, \preceq\}$. Indeed, this fact can be validated through simulations, as we point out in Table 1.

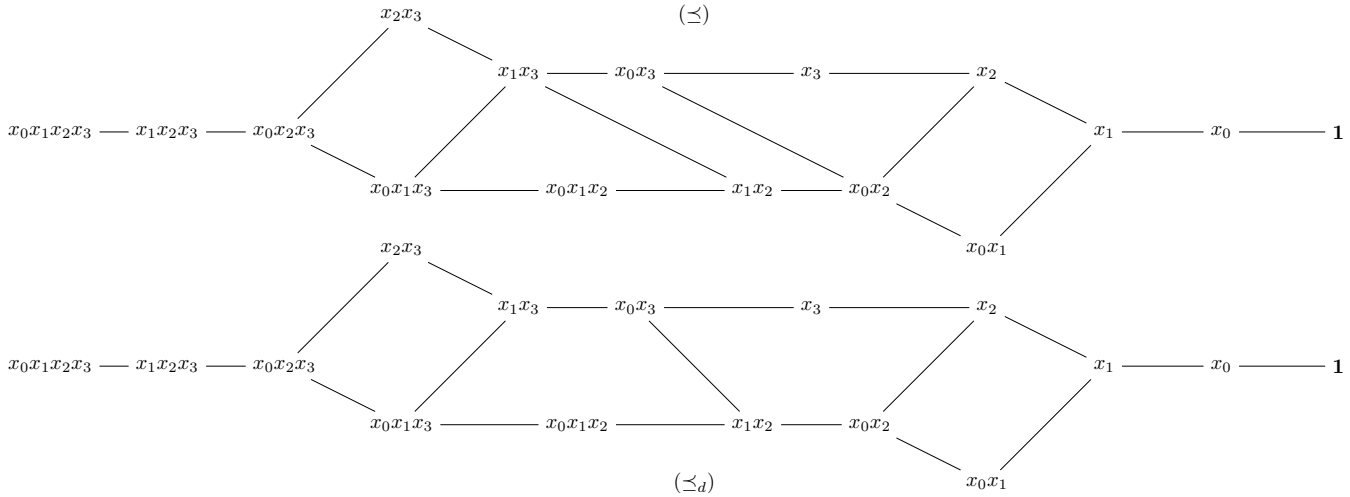


Figure 4. The two-order relations $\preceq$ and $\preceq_d$ for $m = 4$.

Table 1. Non comparable elements in the middle of $\{\mathcal{M}_m, \preceq\}$.

	$m = 6$
$\preceq$	$(4,3,2,1), (5,3,2), (5,4,1), (6,3,1), (6,4)$
$\preceq_d$	$(4,3,2,1)$ $(5,3,2) \preceq_d (5,4,1) \preceq_d (6,3,1)$ $(6,4)$
	$m = 7$
$\preceq$	$(5,4,3,2), (6,4,3,1), (6,5,2,1), (6,5,3), (7,4,2,1), (7,4,3), (7,5,2), (7,6,1)$
$\preceq_d$	$(5,4,3,2) \preceq_d (6,4,3,1) \preceq_d (6,5,2,1) \preceq_d (7,4,2,1)$ $(6,5,3) \preceq_d (7,4,3) \preceq_d (7,5,2) \preceq_d (7,6,1)$
	$m = 8$
$\preceq$	$(8,4,3,2,1), (7,5,3,2,1), (6,5,4,2,1), (8,7,3), (8,6,4), (7,6,5)$ $(6,5,4,3), (7,5,4,2), (7,6,3,2), (7,6,4,1), (8,5,3,2), (8,5,4,1), (8,6,3,1), (8,7,2,1)$
$\preceq_d$	$(6,5,4,2,1) \preceq_d (7,5,3,2,1) \preceq_d (8,4,3,2,1)$ $(7,6,5) \preceq_d (8,6,4) \preceq_d (8,7,3)$ $(6,5,4,3) \preceq_d (7,5,4,2) \preceq_d (7,6,3,2) \preceq_d$ $(7,6,4,1)$ $(8,5,3,2)$ $\preceq_d (8,5,4,1) \preceq_d (8,6,3,1) \preceq_d (8,7,2,1)$

Simulations

We have implemented an algorithm for generating all elements in the middle of $\{\mathcal{M}_m, \preceq\}$. For each value of m in $\{6..8\}$, the elements in the set are displayed in Table 1. We choose to display the Shift(ind(g)) instead of g, where for $g = x_{i_1} \dots x_{i_l}$, Shift(ind($g$)) $= (i_1 + 1, \dots, i_l + 1)$. The main reason for this convention is that the elements in the middle of $\{\mathcal{M}_m, \preceq\}$ are the answers of a well-known problem in computer science, i.e., perfect subset sum problem. Indeed, if we carefully check the elements for each value of m, we discover that $\sum_{j \in \text{Shift(ind}(g))} j = \lfloor \binom{m+1}{2}/2 \rfloor$ for any g in the middle of the poset $\{\mathcal{M}_m, \preceq\}$.

As one can notice from Table 1, the number of non comparable elements from the middle of $\{\mathcal{M}_m, \preceq\}$ decreases rapidly when the order relation $\preceq_d$ is used. For example, when $m = 8$, we have decreased this number from 14 non comparable monomials, with respect to $\preceq$, to 3 distinct non comparable chains, with respect to $\preceq_d$.

3.2.2. Reducing the Number of Codes

Another pertaining aspect when we deal with decreasing and strongly decreasing monomial codes is the estimation of codes for a fixed length 2^m and dimension $k \in \{1..2^m\}$. It seems quite natural, in view of the relation between $\preceq$ and $\preceq_d$, to state that there are fewer strongly decreasing monomial codes of fixed length and dimension than decreasing monomial codes. Formally, we have

Lemma 3. *Let m be a strictly positive integer and $0 \leq k \leq 2^m$. Then the number of decreasing monomial code $\mathscr{C}(I)$ with $I \in \mathcal{M}_m$ and $|I| = k$ greater than or equal to the number of strongly decreasing monomial codes $\mathscr{C}(J)$ with $J \in \mathcal{M}_m$ and $|J| = k$.*

The proof of this lemma is obvious and comes directly from the fact that any strongly decreasing monomial set I is necessarily a decreasing monomial set. However, the inverse is not always true.

Simulations

We have written an algorithm that computes the number of decreasing monomial codes for a fixed length and dimension. Our algorithm works recursively, by adding new monomials from $\{Mon, \preceq\}$ to the previous sets of monomials of cardinality $k-1$, in such a manner that the cardinality of the new sets does not exceed k. The algorithm outputs all possible decreasing monomials sets of cardinality k and then it checks which of them are also strongly decreasing monomial sets.

For small values of k, i.e., $k = O(1)$ when $n \to \infty$, there are almost no differences between strongly decreasing and decreasing sets. However, for bigger values of k, we observe a significant reduction of the number of strongly decreasing monomial codes compared to decreasing monomial codes. In Table 2, we illustrate on a small example $m = 5$ and $k \in \{9, 10, 11, 12\}$ the difference between the two sets. We choose to represent each monomial set I by $\text{ind}(I) = \{\text{ind}(g), g \in I\}$.

Table 2. Decreasing and strongly decreasing monomial sets for $m = 5$.

k	ind(I) for $\preceq$	ind(J) for $\preceq_d$
9	$\{0,1,2,3,4,01,02,03,04\}$	
	$\{0,1,2,3,4,01,02,03,012\}$	
	$\{0,1,2,3,4,01,02,03,12\}$	$\{0,1,2,3,4,01,02,03,12\}$
	$\{0,1,2,3,01,02,03,12.012\}$	$\{0,1,2,3,01,02,03,12,012\}$
	$\{0,1,2,3,01,02,03,12,13\}$	$\{0,1,2,3,01,02,03,12,13\}$
10	$\{0,1,2,3,4,01,02,03,04,12\}$	
	$\{0,1,2,3,4,01,02,03,12,012\}$	$\{0,1,2,3,4,01,02,03,12,012\}$
	$\{0,1,2,3,4,01,02,03,12,13\}$	$\{0,1,2,3,4,01,02,03,12,13\}$
	$\{0,1,2,3,01,02,03,12,13,012\}$	$\{0,1,2,3,01,02,03,12,13,012\}$
	$\{0,1,2,3,01,02,03,12,13,23\}$	

Table 2. *Cont.*

k	ind(I) for $\preceq$	ind(J) for $\preceq_d$
11	$\{0,1,2,3,4,01,02,03,04,12,012\}$	
	$\{0,1,2,3,4,01,02,03,04,12,13\}$	$\{0,1,2,3,4,01,02,03,04,12,13\}$
	$\{0,1,2,3,4,01,02,03,12,13,012\}$	$\{0,1,2,3,4,01,02,03,12,13,012\}$
	$\{0,1,2,3,4,01,02,03,12,13,23\}$	
	$\{0,1,2,3,01,02,03,12,13,23,012\}$	
	$\{0,1,2,3,01,02,03,12,13,012,013\}$	$\{0,1,2,3,01,02,03,12,13,012,013\}$
12	$\{0,1,2,3,4,01,02,03,04,12,13,012\}$	$\{0,1,2,3,4,01,02,03,04,12,13,012\}$
	$\{0,1,2,3,4,01,02,03,04,12,13,23\}$	$\{0,1,2,3,4,01,02,03,04,12,13,23\}$
	$\{0,1,2,3,4,01,02,03,04,12,13,14\}$	
	$\{0,1,2,3,4,01,02,03,12,13,012,013\}$	$\{0,1,2,3,4,01,02,03,12,13,012,013\}$
	$\{0,1,2,3,4,01,02,03,12,13,23,012\}$	
	$\{0,1,2,3,01,02,03,12,13,23,012,013\}$	

4. Average Reliability of the Synthetic Channels

The geometric approach of the properties of a function by means of its subgraph and/or epigraph generated useful mathematical tools from the very beginning of the theory of functions. Measure, intersection, support and shape properties lead to applications in various domains: optimization, shape description and recognition, etc. Here, we propose a geometric approach in the field of polar coding. Recently, the concept of average reliability was introduced and analyzed in the context of all terminal reliability [42]. In view of Theorem 4, the Bhattacharyya parameter of a synthetic channel can be mapped into the reliability polynomial of a minimal two-terminal network. As a consequence, almost all constructive and efficient methods from network reliability can be applied to polar codes over BEC, by means of the Bhattacharyya parameter.

As the set of the synthetic channels cannot be totally ordered [4,25], we propose a different method to define the optimality of a synthetic channel. For that, we will check how reliable a channel is on average, i.e., we define

Definition 10. *Let m be a strictly positive integer and $\boldsymbol{u} \in \{0,1\}^m$. The average reliability of $W^{\boldsymbol{u}}$ is*

$$\mathrm{Avr}(W^{\boldsymbol{u}}) = \int_0^1 \mathcal{B}(W^{\boldsymbol{u}})(p)dp.$$

Moreover, we define the relation $\leq_{\mathrm{Avr}}$

$$\boldsymbol{u} \leq_{\mathrm{Avr}} \boldsymbol{v} \Leftrightarrow \mathrm{Avr}(W^{\boldsymbol{u}}) \leq \mathrm{Avr}(W^{\boldsymbol{v}})$$

This notion of optimality has a meaning in the following context. Imagine that the communication channel is a BEC with variable erasure probability, coming from different physical reasons. This means that either we choose a different polar code in function of the variations of p and in this case we obtain the best performance for each instance, or we choose a polar code and hope that on average it performs in an optimal way. The former strategy comes with the cost of computing for each value of p the corresponding polar code; as for the latter, the cost is minimal, since we construct only one polar code.

4.1. Properties

Lemma 4. *The relation $\leq_{\mathrm{Avr}}$ is reflexive and transitive. In other words, $\leq_{\mathrm{Avr}}$ is a preorder relation.*

Our simulations have shown that up to $m = 13$, $\leq_{\text{Avr}}$ is also antisymmetric. However, this property might not be true in general. Indeed, one can easily find two distinct polynomials with integer coefficients defined over $[0,1]$ with values in $[0,1]$, such that their integrals are equal.

Conjecture 1. *The relation $\leq_{\text{Avr}}$ is a preorder relation on the set of all synthetic channels.*

All the same, we can overcome this by applying the following procedure.

Remark 4. *Let $\boldsymbol{u} \equiv_{\text{Avr}} \boldsymbol{v}$ if, and only if, $\text{Avr}(\mathcal{B}(W^{\boldsymbol{u}})) = \text{Avr}(\mathcal{B}(W^{\boldsymbol{v}}))$. Let us extend the relation $\leq_{\text{Avr}}$ to the factor set $\mathcal{M}_m / \equiv_{\text{Avr}}$ naturally, using the relation between class representatives. Then, $\leq_{\text{Avr}}$ is a total order relation over $\mathcal{M}_m / \equiv_{\text{Avr}}$. Indeed, one can easily check that $\leq_{\text{Avr}}$ is antisymmetric over $\mathcal{M}_m / \equiv_{\text{Avr}}$.*

Lemma 5. *Let m be a strictly positive integer, $n = 2^m$, and $\boldsymbol{u} \in \{0,1\}^m$. Then,*

$$\text{Avr}(W^{\boldsymbol{u}}) = \frac{1}{n+1} \sum_{i=2^{m-|\boldsymbol{u}|}}^{n} \frac{N_i(\boldsymbol{C}^{\boldsymbol{u}})}{\binom{n}{i}}. \tag{11}$$

$$\text{Avr}(W^{\boldsymbol{u}}) + \text{Avr}\left(W^{\overline{\boldsymbol{u}}}\right) = 1. \tag{12}$$

Proposition 7. *Let m be a strictly positive integer and $\boldsymbol{u}, \boldsymbol{v}$ be two binary vectors of length m, such that $\boldsymbol{u} \leq \boldsymbol{v}$. Then,*

$$\boldsymbol{u} \leq \boldsymbol{v} \Rightarrow \text{Avr}(W^{\boldsymbol{u}}) \leq \text{Avr}(W^{\boldsymbol{v}}). \tag{13}$$

In Table 3, we compute the $\text{Avr}(W^{\boldsymbol{u}})$ of all the binary vectors $\boldsymbol{u} \in \{0,1\}^m$ for $m \in \{2,3,4\}$. Notice that in this case, Proposition 7 applies, since we know that up to $m = 4$, the synthetic channels can be totally ordered over the BEC [3,25]. Starting from $m = 5$, this property is no longer true. When $\boldsymbol{u}$ and $\boldsymbol{v}$ are no longer comparable, i.e., there is $p_0 \in (0,1)$ such that $\mathcal{B}(W^{\boldsymbol{u}})(p_0) = \mathcal{B}(W^{\boldsymbol{v}})(p_0)$, we can still decide whether on average $\boldsymbol{u}$ is optimal compared with $\boldsymbol{v}$. The set of non-comparable pairs $(\boldsymbol{u}, \boldsymbol{v})$ for $m = 5$ is $\{(3,16),(12,17),(7,20),(7,24),(11,24),(14,19),(15,28)\}$. Notice that half of the pairs are coming from duality, i.e., if $(\boldsymbol{u}, \boldsymbol{v})$ are not comparable, then $(\overline{\boldsymbol{u}}, \overline{\boldsymbol{v}})$ are also non-comparable. However, these are ordered with respect to average reliability. The average reliability for the first 4 non-comparable pairs are $(0.221, 0.216)$, $(0.396, 0.383)$, $(0.4712, 0.4710)$, $(0.4712, 0.5288)$. Hence, for $m = 5$ the ordering with respect to the average reliability is 0,1,2,4,8,16,3,5,6,9, 10,17,12,18,20,7,24, and the rest can be completed by symmetry.

Table 3. Average reliability of the synthetic channels.

$m = 2$															
0	1	2	3												
0.20	0.47	0.53	0.80												
$m = 3$															
0	1	2	4	3	5	6	7								
0.11	0.29	0.34	0.41	0.59	0.66	0.71	0.89								
$m = 4$															
0	1	2	4	8	3	5	6	9	10	12	7	11	13	14	15
0.06	0.16	0.20	0.24	0.30	0.38	0.44	0.48	0.52	0.56	0.62	0.70	0.76	0.80	0.84	0.94

Example 1. *The ordering induced by the average reliability.*

- $m = 5$

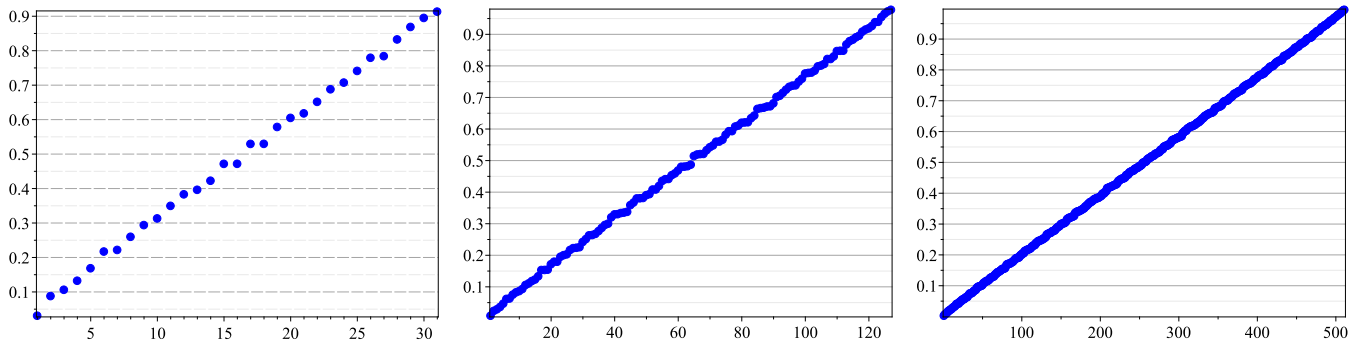
$$0, \underbrace{1,2,4,8,16}_{\mathcal{RM}(1,5)}, \overbrace{3,5,6,9,10,17,12,18,20}^{\mathcal{RM}(2,5)}, \underbrace{7}_{\mathcal{RM}(3,5)}, \overbrace{24}^{\mathcal{RM}(2,5)}$$

- $m = 6$

$$0, \underbrace{1,2,4,8,16}_{\mathcal{RM}(1,6)}, \overbrace{3,5}^{\mathcal{RM}(2,6)}, \underbrace{32}_{\mathcal{RM}(1,6)}, \overbrace{6,9,10,17,12,18,33,20}^{\mathcal{RM}(2,6)}, \underbrace{7}_{\mathcal{RM}(3,6)}, \overbrace{34,24}^{\mathcal{RM}(2,6)}, \underbrace{11}_{\mathcal{RM}(3,6)}, \overbrace{36}^{\mathcal{RM}(2,6)}, \underbrace{13,19,14}_{\mathcal{RM}(3,6)}, \overbrace{40}^{\mathcal{RM}(2,6)},$$

$$\underbrace{21}_{\mathcal{RM}(3,6)}, \overbrace{48}^{\mathcal{RM}(2,6)}, \underbrace{22,35,25,37,26,38,28,41}_{\mathcal{RM}(3,6)}$$

Our simulations have shown that, considering the relation $\leq_{\text{Avr}}$ in the set of the synthetic channels, in each sub-interval $(i/10, (i+1)/10)$, for $0 \leq i \leq 9$, we have a rough proportion of $2^m/10$ binary vectors $\boldsymbol{u}$. So, roughly speaking, a uniform distribution could be used to approximate the number of $\boldsymbol{u}$ inside each sub-interval (illustrated in Figure 5), with respect to Avr (see Table 4 for $5 \leq m \leq 11$.)



(**a**) $m = 5$ (**b**) $m = 7$ (**c**) $m = 9$

Figure 5. Sorted $\text{Avr}(\mathcal{B}(W^{\boldsymbol{u}}))$ for all $\boldsymbol{u} \in \{0,1\}^m$.

Table 4. Number of $\boldsymbol{u} \in \{0,1\}^m$ that satisfy $\text{Avr}(\mathcal{B}(W^{\boldsymbol{u}})) \in (i/10, (i+1)/10]$, for $0 \leq i < 5$, $\epsilon_m = 2^{m-4}/10$.

m	(0,0.1]	(0.1,0.2]	(0.2,0.3]	(0.3,0.4]	(0.4,0.5]	$[\lfloor 2^m/10 - \epsilon_m \rfloor, \lceil 2^m/10 + \epsilon_m \rceil]$
5	2	3	4	4	3	$[3,4]$
6	5	7	6	8	6	$[6,7]$
7	11	13	14	13	13	$[12,14]$
8	23	25	27	27	26	$[24,28]$
9	49	51	50	55	51	$[48,55]$
10	99	104	98	107	104	$[97,109]$
11	199	209	204	204	208	$[194,218]$

4.2. Relation to β-Expansion

β-expansion [15] is a well-known method for an efficient construction of polar codes. Hence, it is with no surprise that our results on average reliability determine possibly more refined choices of the variable β. Let us begin by defining the method.

$$\beta(\boldsymbol{u}) = \sum_{i=0}^{m-1} u_i \beta^i \tag{14}$$

In [15], the authors proved that for any $\beta \in (1, \infty)$, the order induced by β on the sequence of synthetic channels respects the order relation $\preceq$. In particular, this means that if $\boldsymbol{u} \preceq \boldsymbol{v}$ then $\beta(\boldsymbol{u}) \leq \beta(\boldsymbol{v})$ and this for any value of $\beta > 1$. Some values of β are of high interest, in particular $\beta = 2^{1/4}$, when W is designed for additive white Gaussian noise (AWGN). In the case of AWGN, the authors in [15] proposed a procedure in which an interval for β is determined, an interval that converges to a value close to $2^{1/4}$. Notice that in [15], the order induced by β is not valid for any signal to noise ratio value, but it tries to cover as much as possible the interval $[0, 1]$. A natural question that one could raise is whether there is a β-expansion for the average reliability, i.e., is there a real value β such that β and Avr are identical over the set of binary vectors of length m. There is a significant difference between the two relations. In our case, not only that W is a BEC, but also the preorder induced by the average reliability is total over $\mathcal{M}_m / \equiv_{\text{Avr}}$ and holds for the entire interval $[0, 1]$.

Remark 5. *By computer simulations, one can easily prove that for $m \leq 5$,, there is $\beta \in (1, \infty)$, such that the order induced by beta and the preorder induced by the average reliability coincide. It can be done by simply tacking $\beta = 1.22$.*

Conjecture 2. *For $m > 6$, we did not find a value of β for which the two aforementioned relations are equal. Moreover, for $\beta \sim 1.22$, the number of elements with similar mutual relations with respect to the two relations is minimized (see Table 5).*

Table 5. Number of pairs $(\boldsymbol{u}, \boldsymbol{v})$ satisfying $\text{Avr}(\mathcal{B}(W^{\boldsymbol{u}})) \leq \text{Avr}(\mathcal{B}(W^{\boldsymbol{v}}))$ for which $\nexists \beta \in (1, \infty)$ s.t. $\beta(\boldsymbol{u}) \leq \beta(\boldsymbol{v})$.

m	β	Number of Incompatible Pair of Elements	2^m
4	$(1, 1.32]$	–	16
5	$(1.18, 1.22]$	–	32
6	1.22	2	64
7	1.22	10	128
8	1.22	36	256
9	1.22	99	512

4.3. Threshold Points of the Binary Erasure Polarization Sub-Channels

The fact that when m goes to infinity the Bhattacharyya polynomial has a sharp transition from zero to one when m goes to infinity has already been proven ([34]). More exactly, for any $\boldsymbol{u}$, there exists a point $p_0(\boldsymbol{u}) \in (0, 1)$ for which in its vicinity $\mathcal{B}(W^{\boldsymbol{u}})$ passes from very small values (close to zero) to very high values (close to one). Formally speaking, we have

Lemma 6 ([34])**.**

$$\lim_{m \to \infty} \mathcal{B}(W^{\boldsymbol{u}}) = \begin{cases} 0 & p \in [0, p_0(\boldsymbol{u})) \\ 1 & p \in (p_0(\boldsymbol{u}), 1] \end{cases} \tag{15}$$

However, finding the point $p_0(\boldsymbol{u})$ where this transition holds is not trivial (see [15,29]). Here, we will use the average reliability to determine this point for some specific channels.

Lemma 7.

$$\lim_{m\to\infty} \mathrm{Avr}(W^{\boldsymbol{u}}) = 1 - p_0(\boldsymbol{u}). \tag{16}$$

A particular interesting channel analyzed in [25,29] is the synthetic channel $W^{(1^i0^{m-i})}$. More exactly, the authors analyze the sharp transition of $W^{(1^i0^{m-i})}$ from 0 to 1 when m tends to infinity, in function of the limit $i/m - i$. Here, we will give an exact formula for the average reliability of $W^{('1^i0^{m-i})}$. This result combined with Lemma 7 will allow us to obtain a finer approximation of $p_0(\boldsymbol{u})$. To achieve our goal, we will look at the corresponding 2TN, namely at $\boldsymbol{C}^{(1^i0^{m-i})}$. For simplification, we use $l = 2^{m-i}, w = 2^i$ and $n = 2^m$. Notice that

$$\mathrm{Rel}\left(\boldsymbol{C}^{(1^i0^{m-i})}; p\right) = 1 - (1-p^l)^w. \tag{17}$$

Theorem 7.

$$\mathrm{Rel}\left(\boldsymbol{C}^{(1^i0^{m-i})}; p\right) = \sum_{i=l}^{n} \sum_{j=1}^{\lfloor \frac{i}{l} \rfloor} (-1)^{j+1} \binom{w}{j} \binom{n-jl}{n-i} p^i (1-p)^{n-i}. \tag{18}$$

Proof. In order to prove our result, we need to demonstrate that $\forall\ l \leq i \leq n$

$$N_i(\boldsymbol{C}^{(1^i0^{m-i})}) = \sum_{j=1}^{\lfloor \frac{i}{l} \rfloor} (-1)^{j+1} \binom{w}{j} \binom{n-jl}{n-i} \tag{19}$$

The proof is based on an inclusion-exclusion argument. Denote by $\mathcal{P}_i$ the set of paths of length i from S to T for the $\boldsymbol{C}^{(1^i0^{m-i})}$. This leads to $|\mathcal{P}_i| = N_i(\boldsymbol{C}^{(1^i0^{m-i})})$.

Any path of length i with $l \leq i$ is composed of at least one path of length l, hence we have w choices for fixing a path of length l and $\binom{n-l}{i}$ choices for the remaining positions. However, in the $\binom{n-l}{i}$ choices, we might count other l length paths. Hence, we need to subtract the over-counting, which is all the combinations of two length l paths, i.e., $\binom{w}{2}$, times the number of choices for the remaining positions, i.e., $\binom{n-2l}{i-2l}$. Now, we need to add all the paths that are composed of at least 3 l paths which equals $\binom{w}{3}\binom{n-3l}{i-3l}$., and so on till we reached the last level, i.e., $\binom{w}{\lfloor \frac{i}{l} \rfloor}\binom{n-l\lfloor \frac{i}{l} \rfloor}{i-l\lfloor \frac{i}{l} \rfloor}$. □

Theorem 8.

$$\mathrm{Avr}\left(W^{(1^i0^{m-i})}\right) = 1 - \frac{1}{\binom{2^i+2^{i-m}}{2^i}} \tag{20}$$

Proof.

$$\begin{aligned}
\mathrm{Avr}(W^{\boldsymbol{u}}) &= \frac{1}{n+1}\sum_{i=l}^{n}\frac{N_i(\boldsymbol{C}^{\boldsymbol{u}})}{\binom{n}{i}} = \frac{1}{n+1}\sum_{i=l}^{n}\sum_{j=1}^{\lfloor \frac{i}{l} \rfloor}(-1)^{j+1}\frac{\binom{w}{j}\binom{n-jl}{n-i}}{\binom{n}{i}} \\
&= \frac{1}{n+1}\sum_{j=1}^{w}\sum_{i=jl}^{n}(-1)^{j+1}\frac{\binom{w}{j}\binom{n-jl}{n-i}}{\binom{n}{i}} = \frac{1}{n+1}\sum_{j=1}^{w}\sum_{i=jl}^{n}(-1)^{j+1}\frac{\binom{w}{j}\binom{i}{jl}}{\binom{n}{jl}} \\
&= \frac{1}{n+1}\sum_{j=1}^{w}(-1)^{j+1}\frac{\binom{w}{j}}{\binom{n}{jl}}\sum_{i=jl}^{n}\binom{i}{jl} = \frac{1}{n+1}\sum_{j=1}^{w}(-1)^{j+1}\frac{\binom{w}{j}\binom{n+1}{jl+1}}{\binom{n}{jl}} \\
&= \sum_{j=1}^{w}(-1)^{j+1}\binom{w}{j}\frac{1}{jl+1} = 1-\sum_{j=0}^{w}(-1)^{j}\binom{w}{j}\frac{1}{jl+1} = 1-\frac{1}{\binom{\frac{n+1}{l}}{w}}
\end{aligned}$$

□

Basically, we have

Corollary 2.

$$\mathrm{Avr}\left(W^{(0^i1^{m-i})}\right) = \frac{1}{\binom{2^i+2^{i-m}}{2^i}} \tag{21}$$

Based on Theorem 8, we can establish new classes of asymptotically "good" channels. For that, we will need the following result.

Lemma 8.

$$\lim_{n\to\infty}\begin{pmatrix} \frac{n}{\log_2(n)(\log_2(\log_2(n)))} + \frac{1}{\log_2(n)(\log_2(\log_2(n)))} \\ \frac{n}{\log_2(n)(\log_2(\log_2(n)))} \end{pmatrix} = 1. \tag{22}$$

$$\lim_{n\to\infty}\begin{pmatrix} \frac{n}{\log_2(\log_2(n))} + \frac{1}{\log_2(\log_2(n))} \\ \frac{n}{\log_2(\log_2(n))} \end{pmatrix} = \infty. \tag{23}$$

$$\lim_{n\to\infty}\begin{pmatrix} \frac{n}{\log_2(n)} + \frac{1}{\log_2(n)} \\ \frac{n}{\log_2(n)} \end{pmatrix} = 2. \tag{24}$$

Theorem 8, Lemma 8 and Lemma 7 imply the following result.

Corollary 3. *Let m be a strictly positive integer and $\boldsymbol{u} = \left(1^i0^{m-i}\right)$ Then,*

- *for any $i \leq m - \log_2(m) - \log_2(\log_2(m))$ we have $p_0(\boldsymbol{u}) \to 1$ and $p_0(\overline{\boldsymbol{u}}) \to 0$.*
- *for any $i \geq m - \log_2(\log_2(m))$ we have $p_0(\boldsymbol{u}) \to 0$ and $p_0(\overline{\boldsymbol{u}}) \to 1$.*

Another direct consequence of our results is that for any $i \leq m - \log_2(m) - \log_2(\log_2(m))$, the monomial $f = x_0 \dots x_{i-1}$ is highly reliable on average. Hence, all the monomials $g \preceq_d f$ are also highly reliable in average, as their average reliability tends to zero when m goes to infinity. Moreover, f becomes unreliable on average for $i \geq m - \log_2(\log_2(m))$. The values $m - \log_2(m) - \log_2(\log_2(m)) < i < m - \log_2(\log_2(m))$ are to be considered in more detail.

Corollary 4. *Let m be a strictly positive integer and $i \leq \log_2(\log_2(m))$. Then, for any $f \in \mathcal{M}_m$ with $f \preceq_d x_{m-i+1} \dots x_m$, we have that $\mathrm{Avr}(W^f) \to 0$ when $m \to \infty$. In other words, any synthetic channel in the $\mathcal{RM}(i,m)$ is asymptotically "good" on average.*

5. Conclusions and Perspectives

A complete characterization of the Bhattacharyya parameter of synthetic channels of a monomial code is an open problem that has attracted a lot of attention in the last decade. Even for the particular case of binary erasure channel, the question remains unanswered. However, the implications of such a result are of high importance in coding theory, especially in polar coding. In this article, we make a step forward by proposing an order relation $\preceq_d$ that decreases the gap between state-of-the-art and the ultimate partial order relation for the Bhattacharyya parameter of synthetic channels. The advantage of this approach is that our algebraic description is rather easy to implement and analyze, compared to other order relations such as [25]. Simulations show that $\preceq_d$ is a valid order relation on binary symmetric channel, and a deeper inspection of [25], and our work could potentially determine an algebraic description that fits the latest results.

The order relation proposed here could be employed as a new construction rule for polar codes. As $\preceq_d$ is thinner than $\preceq$, it enables two reductions:

- the number of non-comparable monomials in the middle of $\{\mathcal{M}_m, \preceq\}$ is significantly reduced by means of $\preceq_d$,
- the number of strongly decreasing monomial codes is less than the number of decreasing monomial codes for fixed length and dimension.

Hence, these two properties open the perspectives of a more efficient construction algorithm for strongly decreasing monomial codes, and hence for polar codes.

As the relations on the Bhattacharyya parameter are all partial orders, we have proposed an alternative solution for ordering the synthetic channels. For that, we have used the concept of average reliability, borrowed from network theory. Instead of the local evaluation of the Bhattacharyya parameter, we propose a global one, by evaluating the integral, i.e., by measuring its global average behavior. Hence, we rank the synthetic channels using a preorder relation $\leq_{\text{Avr}}$, given by the value of the integral. Our result is not constructive, in the sense that it does not fully characterize the channels that belong to a specific interval. An answer to this question might provide an extremely efficient method for constructing polar codes and give much more insight into the synthetic channels $W^{\boldsymbol{u}}$.

Author Contributions: Conceptualization, V.-F.D. and G.C.; methodology, V.-F.D.; software, V.-F.D.; validation, V.-F.D. and G.C.; formal analysis, V.-F.D. and G.C.; investigation, V.-F.D.; resources, V.-F.D. and G.C.; writing—original draft preparation, V.-F.D. and G.C. All authors have read and agreed to the published version of the manuscript.

Funding: V-F. Dragoi is supported by a grant of the Romanian Ministry of Education and Research, CNCS- UEFISCDI, project number PN-III-P1-1.1-PD-2019-0285, within PNCDI III.

Data Availability Statement: The data presented in this study are available within the article.

Acknowledgments: The authors are thankful to anonymous reviewers for their valuable suggestions and comments to improve the quality of the article.

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

Appendix A. Proofs of Results from Section 3

Appendix A.1. Proof of Proposition 6

Proposition (6). *Let f and g be two monomial having the same degree and x_h be such that $x_h \not| f$ and $x_h \not| g$. Then we have*

$$f \preceq_d g \text{ iff } x_h f \preceq_d x_h g. \tag{A1}$$

Proof. Let $f = x_{i_1} \dots x_{i_s}$ and $g = x_{j_1} \dots x_{j_s}$ with $f \preceq_d g$. Also, let $f^* = f x_h$ and $g^* = g x_h$. There are several cases to be examined here:

- If $x_{j_s} \preceq x_h$ or $x_h \preceq x_{i_1}$ then the relation $f^* \preceq_d g^*$ can easily be verified by using the definition of $\preceq_d$.

- If there is an integer $r \in \{1, \dots, s\}$ s.t. $x_{i_r} \preceq x_h \preceq x_{i_{r+1}}$ and $x_{j_r} \preceq x_h \preceq x_{j_{r+1}}$ then the relation $f^* \preceq_d g^*$ can easily be verified as in the previous step.
- If there are two distinct integers $r, t \in \{1, \dots, s\}$ s.t. $x_{i_r} \preceq x_h \preceq x_{i_{r+1}}$ and $x_{j_t} \preceq x_h \preceq x_{j_{t+1}}$ then two cases have to be considered.
  - If $t < r$ then we have $x_{i_h} \preceq x_{j_{t+1}} \preceq \cdots \preceq x_{j_r}$ and $x_{i_{t+1}} \preceq \cdots \preceq x_{i_r} \preceq x_h$, which implies the following ordering

$$x_{i+t+1} \preceq \cdots \preceq x_{i_r} \preceq x_h \preceq x_{j_{t+1}} \preceq \cdots \preceq x_{j_r}. \tag{A2}$$

    Combining Equation (A2) with the definition of $\preceq_d$ we obtain the desired result, i.e., $f^* \preceq_d g^*$.
  - $t > r$ then we have $x_{i_h} \preceq x_{i_{r+1}} \preceq \cdots \preceq x_{i_t}$ and $x_{j_{r+1}} \preceq \cdots \preceq x_{j_t} \preceq x_h$ which implies the following ordering

$$x_{j+r+1} \preceq \cdots \preceq x_{j_t} \preceq x_h \preceq x_{i_{r+1}} \preceq \cdots \preceq x_{i_t}. \tag{A3}$$

    Now, since $x_h \preceq x_{i_t}$ it might be possible to have $j_s + \dots j_{t+1} + h < i_s + \cdots + i_{t+1} + i_t$, which implies a violation of the partial sum conditions in the definition of $\preceq_d$. If the next partial sum changes the sign, i.e., $j_s + \dots j_{t+1} + h + j_t \geq i_s + \cdots + i_{t+1} + i_t + i_{t-1}$, by setting $\Delta_{t+1} = (j_s - i_s) + \cdots + (j_{t+1} - i_{t+1})$ we have the following inequalities

$$\Delta_{t+1} + h - i_t < 0 \tag{A4}$$
$$\Delta_{t+1} + h - i_t + j_t - i_{t-1} \geq 0. \tag{A5}$$

    This implies

$$\Delta_{t+1} + h < i_t \leq \Delta_{t+1} + h + j_t - i_{t-1}. \tag{A6}$$

    However, since $j_t < i_{t-1}$ the Equation (A6) becomes impossible, which completes the proof.

□

Appendix A.2. Proof of Lemma 2

Lemma (2). *Let $f, g \in \mathcal{M}_m$, $\deg f = \deg g = 2$, such that $f \preceq_d g$. Then $\mathcal{B}\left(W_m^f\right) \leq \mathcal{B}\left(W_m^g\right)$.*

Proof. Let $f = x_{i_1} x_{i_2}$ and $g = x_{j_1} x_{j_2}$. If $\gcd(f, g) \neq 1$ then $f \preceq g$, which implies $\mathcal{B}\left(W_m^f\right) \leq \mathcal{B}\left(W_m^g\right)$.

Now suppose $\gcd(f, g) = 1$, and let $j_1 - i_1 \leq j_2 - i_2$. Denote $\epsilon = j_1 - i_1$. By definition of $\preceq_d$ and $\preceq$ we have

$$f = x_{i_1} x_{i_2} \preceq_d x_{i_1 - 1} x_{i_2 + 1} \preceq_d \cdots \preceq_d x_{j_1} x_{i_2 + \epsilon} \preceq x_{j_1} x_{j_2} = g. \tag{A7}$$

If we prove $x_{i_1} x_{i_2} \preceq_d x_{i_1 - 1} x_{i_2 + 1} \Rightarrow \mathcal{B}\left(W_m^{x_{i_1} x_{i_2}}\right) \leq \mathcal{B}\left(W_m^{x_{i_1 - 1} x_{i_2 + 1}}\right)$ the proof is finished. By definition, one can easily notice that $\mathcal{B}\left(W_4^{x_1 x_2}\right) \leq \mathcal{B}\left(W_4^{x_0 x_3}\right) \Leftrightarrow \mathcal{B}\left(W_m^{x_{i_1} x_{i_2}}\right) \leq \mathcal{B}\left(W_m^{x_{i_1 - 1} x_{i_2 + 1}}\right)$. Hence we are left to prove that $\mathcal{B}\left(W_4^{x_1 x_2}\right) \leq \mathcal{B}\left(W_4^{x_0 x_3}\right)$. We have that $\mathcal{B}\left(W_4^{x_0 x_3}\right)(p) = 1 - (1 - (1 - (1 - p)^2)^4)^2$ and $\mathcal{B}\left(W_4^{x_1 x_2}\right)(p) = (1 - (1 - p^2)^4)^2$.

By writing the two polynomials in the Bernstein basis, and using Theorem 4 we have $\mathcal{B}\left(W_4^{x_1 x_2}\right) = \mathrm{Rel}(\boldsymbol{C}^{(0,1,1,0)})$ and $\mathcal{B}\left(W_4^{x_0 x_3}\right) = \mathrm{Rel}(\boldsymbol{C}^{(1,0,0,1)})$ with

$$N_i(\boldsymbol{C}^{(0,1,1,0)}) = 0, 0, 0, 0, 16, 192, 1008, 3040, 5828, 7456, 6552, 4048, 1788, 560, 120, 16, 1$$
$$N_i(\boldsymbol{C}^{(1,0,0,1)}) = 0, 0, 0, 0, 32, 320, 1456, 3984, 7042, 8400, 7000, 4176, 1804, 560, 120, 16, 1$$

As $N_i(\boldsymbol{C}^{(0,1,1,0)}) \leq N_i(\boldsymbol{C}^{(1,0,0,1)})$ for all $i \in \{0,\ldots,16\}$ we conclude the proof. □

Appendix A.3. Proof of Theorem 5

Theorem (5)**.** *Polar codes over the Binary Erasure Channel are strongly decreasing monomial codes.*

Proof. The proof is based on two induction steps. First the parameter m is fixed and we prove that the result holds for any $1 \leq s \leq m$. Secondly we use induction on m.

Firstly, fix m and use an induction argument on the degree of monomial, namely on s. We also suppose that $\gcd(f,g) = 1$. For $s = 1$ we have that $\preceq = \preceq_d$ so the result is obvious. For $s = 2$ use Lemma 2.

Now suppose that for any $f \preceq_d g$ with $\deg f = \deg g = s-1$ we have that $\mathcal{B}\left(W_m^f\right) \leq \mathcal{B}\left(W_m^g\right)$. Let $f = x_{i_1}\ldots x_{i_s}$ and $g = x_{j_1}\ldots x_{j_s}$ such that $g \preceq_d f$ with the usual convention $i_1 < \cdots < i_s$ and $j_1 < \cdots < j_s$. Then we have two cases. Either if $f/x_{i_s} \preceq_d g/x_{j_s}$ or if $x_{i_1} \leq x_{j_1}$ then we have $\mathcal{B}\left(W_m^f\right) \leq \mathcal{B}\left(W_m^g\right)$. Indeed, in the first case we have that

$$\begin{aligned}\mathcal{B}\left(W_m^f\right) &= \mathcal{B}\left(\left(W_{m-j_{s-1}-1}^{x_{i_s}}\right)_{j_{s-1}+1}^{f/x_{i_s}}\right) \leq \mathcal{B}\left(\left(W_{m-j_{s-1}-1}^{x_{j_s}}\right)_{j_{s-1}+1}^{f/x_{i_s}}\right)\\ &\leq \mathcal{B}\left(\left(W_{m-j_{s-1}-1}^{x_{j_s}}\right)_{j_{s-1}+1}^{g/x_{j_s}}\right) = \mathcal{B}\left(W_m^g\right).\end{aligned}$$

In the second case when $x_{i_1} \leq x_{j_1}$ the proof works in the same way. If we are not in the previous case it means that $j_1 < i_1$ and $i_s < j_s$. We know that there is $l \in \{1,\ldots,s-1\}$ for which $i_l > j_l$. First we treat the two extreme cases $l = 1$ or $l = s-1$. If $l = 1$ this implies that $j_k \geq i_k$ for all $k > 1$. Let $\delta = \min\{j_2 - i_2, i_1 - j_1\}$. Then

$$f = x_{i_1}x_{i_2}\ldots x_{i_s} \preceq_d x_{i_1-\delta}x_{i_2+\delta}x_{i_3}\ldots x_{i_s} \preceq_d x_{j_1}\ldots x_{j_s} = g.$$

In this case, either $x_{i_1-\delta} = x_{j_1}$ or $x_{i_2+\delta} = x_{j_2}$. Hence, by Lemma 2 and using the order relation $\preceq$ we obtain

$$\begin{aligned}\mathcal{B}\left(W_m^f\right) &= \mathcal{B}\left(\left(W_{m-j_2-1}^{f/(x_{i_1}x_{i_2})}\right)_{j_2+1}^{x_{i_2}x_{i_1}}\right) \leq \mathcal{B}\left(\left(W_{m-j_2-1}^{f/(x_{i_1}x_{i_2})}\right)_{j_2+1}^{x_{i_2+\delta}x_{i_1-\delta}}\right)\\ &\leq \mathcal{B}\left(\left(W_{m-j_2-1}^{x_{i_s}\ldots x_{i_3}}\right)_{j_2+1}^{x_{j_2}x_{j_1}}\right) \leq \mathcal{B}\left(\left(W_{m-j_2-1}^{x_{j_s}\ldots x_{j_3}}\right)_{j_2+1}^{x_{j_2}x_{j_1}}\right) = \mathcal{B}\left(W_m^g\right).\end{aligned}$$

If $l = s-1$, by putting $\delta = i_{s-1} - j_{s-1}$ and taking into account that $\delta \leq j_s - i_s$ we obtain

$$\begin{aligned}\mathcal{B}\left(W_m^f\right) &= \mathcal{B}\left(\left(W_{m-j_{s-2}-1}^{x_{i_s}x_{i_{s-1}})}\right)_{j_{s-2}+1}^{f/(x_{i_s}x_{i_{s-1}})}\right) \leq \mathcal{B}\left(\left(W_{m-j_{s-2}-1}^{x_{i_s+\delta}x_{i_{s-1}-\delta})}\right)_{j_{s-2}+1}^{f/(x_{i_s}x_{i_{s-1}})}\right)\\ &= \mathcal{B}\left(\left(W_{m-j_{s-2}-1}^{x_{i_s+\delta}x_{j_{s-1}})}\right)_{j_{s-2}+1}^{x_{i_{s-2}}\ldots x_{i_1}}\right) \leq \mathcal{B}\left(\left(W_{m-j_{s-2}-1}^{x_{j_s}x_{j_{s-1}})}\right)_{j_{s-2}+1}^{x_{j_{s-2}}\ldots x_{j_1}}\right) = \mathcal{B}\left(W_m^g\right).\end{aligned}$$

When $1 < l < s-1$ suppose that $i_s - j_l < j_{l+1} - i_s$ and denote by $\delta_{l,s} = i_s - j_l$. Using the definition of $\preceq_d$ we have

$$h = x_{j_1}\ldots x_{j_l+\delta_{l,s}}x_{j_{l+1}-\delta_{l,s}}\ldots x_{j_s} \preceq_d \cdots \preceq_d x_{j_1}\ldots x_{j_l+1}x_{j_{l+1}-1}\ldots x_{j_s} \preceq_d x_{j_1}\ldots x_{j_s} = g.$$

Notice that $h = x_{j_1} \dots x_{j_{l-1}} x_{i_s} x_{j_{l+1}-i_s+j_l} \dots x_{j_s}$. Next, we prove that $f \preceq_d h$. Since $\gcd(f,h) = x_{i_s}$, we can use Lemma 6 and demonstrate $f/x_{i_s} \preceq_d h/x_{i_s}$, i.e., $x_{i_1} \dots x_{i_{s-1}} \preceq_d x_{j_1} \dots x_{j_{l-1}} x_{j_{l+1}-i_s+j_l} x_{j_{l+2}} \dots x_{j_s}$. As,

$$x_{i_{l+1}} \preceq \cdots \preceq x_{i_s} \preceq x_{j_{l+1}} \preceq \cdots \preceq x_{j_s}, \tag{A8}$$

we obtain $x_{i_{l+1}} \dots x_{i_{s-1}} \preceq_d x_{j_{l+2}} \dots x_{j_s}$, simply by verifying

$$\forall\, t \in \{0, \dots, s-l-2\} \quad \sum_{k=0}^{t} i_{s-1-k} \leq \sum_{k=0}^{t} j_{s-k}. \tag{A9}$$

The next partial sums inequalities,

$$(i_k + \cdots + i_{l-1}) + i_l + i_{l+1} + \cdots + i_{s-1} < (j_k + \cdots + j_{l-1}) + j_{l+1} - i_s + j_l + j_{l+2} + \cdots + j_s \tag{A10}$$

are verified from the relation $f \preceq_d g$. So we check the partial sums step by step:

1. $i_{s-1} < i_s < j_{l+1} < j_s$ and thus $x_{i_{s-1}} \preceq_d x_{j_s}$
2. $i_{s-1} + i_{s-1} < j_{s-1} + j_s$ and thus $x_{i_s-2} x_{i_{s-1}} \preceq_d x_{j_{s-1}} x_{j_s}$
3. ...
4. $i_{l+1} + \cdots + i_{s-1} < j_{l+2} + \cdots + j_s$ and thus $x_{i_{l+1}} \dots x_{i_{s-1}} \preceq_d x_{j_{l+2}} \dots x_{j_s}$
5. $i_l + i_{l+1} + \cdots + i_{s-1} < j_{l+1} - i_s + j_l + j_{l+2} + \cdots + j_s$ by definition of $f \preceq_d g$ and thus $x_{i_l} \dots x_{i_{s-1}} \preceq_d x_{j_{l+1}-i_s+j_l} \dots x_{j_s}$
6. ...

Hence we have that $f/x_{i_s} \preceq_d h/x_{i_s}$ which implies, using the induction hypothesis, that $\mathcal{B}\left(W_m^{f/x_{i_s}}\right) \leq \mathcal{B}\left(W_m^{h/x_{i_s}}\right)$, from which we deduce $\mathcal{B}\left(W_m^f\right) \leq \mathcal{B}\left(W_m^h\right)$. Also,

$$\begin{aligned}
\mathcal{B}\left(W_m^h\right) &= \mathcal{B}\left(\left(\left(W^{x_{j_s} \dots x_{j_{l+2}}}\right)^{x_{j_{l+1-\delta_{l,s}}} x_{j_{l+\delta_{l,s}}}}\right)^{x_{j_{l-1}} \dots x_{j_1}}\right) \\
&= \mathcal{B}\left(\left((W^*)^{x_{j_{l+1-\delta_{l,s}}} x_{j_{l+\delta_{l,s}}}}\right)^{x_{j_{l-1}} \dots x_{j_1}}\right) \leq \mathcal{B}\left(\left((W^*)^{x_{j_{l+1}} x_{j_l}}\right)^{x_{j_{l-1}} \dots x_{j_1}}\right) \\
&= \mathcal{B}\left(\left(\left(W^{x_{j_s} \dots x_{j_{l+2}}}\right)^{x_{j_{l+1}} x_{j_l}}\right)^{x_{j_{l-1}} \dots x_{j_1}}\right) = \mathcal{B}\left(W_m^g\right).
\end{aligned}$$

Secondly, we use induction on the number of variables m. For the first values of m, i.e., $m \leq 4$ it is straightforward to check the result.

Let $f = x_{i_1} \dots x_{i_s}$ and $g = x_{j_1} \dots x_{j_s}$ such that $g \preceq_d f$ with the usual convention $i_1 < \cdots < i_s$ and $j_1 < \cdots < j_s$. The following cases are possible

- If $i_s = j_s = m$ then we have $W_{m+1}^f = \left(W_1^{f_m}\right)_m^{f_{[0,m-1]}}$ and $W_{m+1}^g = \left(W_1^{g_m}\right)_m^{g_{[0,m-1]}}$. Since $f_{[0,m-1]} \preceq_d g_{[0,m-1]}$ we have by the induction hypothesis

$$\mathcal{B}\left(\left(W_1^{x_m}\right)_m^{f_{[0,m-1]}}\right) \leq \mathcal{B}\left(\left(W_1^{x_m}\right)_m^{g_{[0,m-1]}}\right). \tag{A11}$$

- Else, by the definition of the order we necessary have $j_s > i_s$. We also have that

$$h = x_{j_1} \dots x_{j_{s-1}+1} x_{j_s-1} \preceq_d x_{j_1} \dots x_{j_{s-1}} x_{j_s} = g.$$

Which implies that $\mathcal{B}\left(W_{m+1}^h\right) \leq \mathcal{B}\left(W_{m+1}^g\right)$. In the same time notice that $f \preceq_d h$ and $\mathrm{ind}(f), \mathrm{ind}(h) \in \{0, \dots, m-1\}$. Therefore we obtain

$$\mathcal{B}\left(W_{m+1}^f\right) \leq \mathcal{B}\left(W_{m+1}^h\right) \leq \mathcal{B}\left(W_{m+1}^g\right). \tag{A12}$$

□

References

1. Arıkan, E. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Trans. Inform. Theory* **2009**, *55*, 3051–3073. [CrossRef]
2. Bioglio, V.; Condo, C.; Land, I. Design of Polar Codes in 5G New Radio. *IEEE Commun. Surv. Tutor.* **2020**, *23*, 29–40. [CrossRef]
3. Dragoi, V. Algebraic Approach for the Study of Algorithmic Problems Coming from Cryptography and the Theory of Error Correcting Codes. Ph.D. Thesis, Université de Rouen, Normandie, France, 2017.
4. Bardet, M.; Dragoi, V.; Otmani, A.; Tillich, J. Algebraic properties of polar codes from a new polynomial formalism. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016; pp. 230–234. [CrossRef]
5. Rengaswamy, N.; Calderbank, R.; Newman, M.; Pfister, H.D. Classical Coding Problem from Transversal *T* Gates, 2020, In Proceedings of the 2020 IEEE International Symposium on Information Theory (ISIT), Los Angeles, CA, USA, 21–26 June 2020; pp. 1891–1896. [CrossRef]
6. Rengaswamy, N. Classical Coding Approaches to Quantum Applications. *arXiv* **2020**, arXiv:2004.06834.
7. Krishna, A.; Tillich, J.P. Magic state distillation with punctured polar codes. *arXiv* **2019**, arXiv:1811.03112.
8. Bardet, M.; Chaulet, J.; Dragoi, V.; Otmani, A.; Tillich, J.P. Cryptanalysis of the McEliece Public Key Cryptosystem Based on Polar Codes. In *Post-Quantum Cryptography, PQCrypto 2016*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2016; Volume 9606, pp. 118–143. [CrossRef]
9. Drăgoi, V.; Beiu, V.; Bucerzan, D. Vulnerabilities of the McEliece Variants Based on Polar Codes. In *Innovative Security Solutions for Information Technology and Communications, SecITC 2018*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11359, pp. 376–390. [CrossRef]
10. Bucerzan, D.; Dragoi, V.; Kalachi, H.T. Evolution of the McEliece Public Key Encryption Scheme. In *Innovative Security Solutions for Information Technology and Communications, SecITC 2017*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, Volume 10543, pp. 129–149. [CrossRef]
11. Drăgoi, V.F.; Beiu, V. Fast Reliability Ranking of Matchstick Minimal Networks. *arXiv* 2019, arXiv:1911.01153.
12. Dragoi, V.; Cowell, S.; Beiu, V. Ordering series and parallel compositions. In Proceedings of the 2018 IEEE 18th International Conference on Nanotechnology (IEEE-NANO), Cork, Ireland, 23–26 July 2018; pp. 1–4. [CrossRef]
13. Beiu, V.; Cowell, S.R.; Drăgoi, V.F. On Posets for Reliability: How Fine Can They Be? In *Soft Computing Applications SOFA 2018, Advances in Intelligent Systems and Computing*; Springer: Cham, Switzerland, 2021; Volume 1221, pp. 115–129. [CrossRef]
14. Mondelli, M.; Hassani, S.H.; Urbanke, R. Construction of polar codes with sublinear complexity. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 25–30 June 2017; pp. 1853–1857. [CrossRef]
15. He, G.; Belfiore, J.; Land, I.; Yang, G.; Liu, X.; Chen, Y.; Li, R.; Wang, J.; Ge, Y.; Zhang, R.; et al. Beta-Expansion: A Theoretical Framework for Fast and Recursive Construction of Polar Codes. In Proceedings of the GLOBECOM 2017-2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017; pp. 1–6. [CrossRef]
16. Tal, I.; Vardy, A. How to Construct Polar Codes. *IEEE Trans. Inform. Theory* **2013**, *59*, 6562–6582. [CrossRef]
17. Mori, R.; Tanaka, T. Performance and construction of polar codes on symmetric binary-input memoryless channels. In Proceedings of the 2009 IEEE International Symposium on information theory, Seoul, Korea, 28 June–3 July 2009; pp. 1496–1500. [CrossRef]
18. Mahdavifar, H.; El-Khamy, M.; Lee, J.; Kang, I. On the construction and decoding of concatenated polar codes. In Proceedings of the 2013 IEEE International Symposium on Information Theory, Istanbul, Turkey, 7–12 July 2013; pp. 952–956. [CrossRef]
19. Korada, S.B.; Sasoglu, E.; Urbanke, R.L. Polar Codes: Characterization of Exponent, Bounds, and Constructions. *IEEE Trans. Inform. Theory* **2010**, *56*, 6253–6264. [CrossRef]
20. Afşer, H.; Deliç, H. On the Channel-Specific Construction of Polar Codes. *IEEE Commun. Lett.* **2015**, *19*, 1480–1483. [CrossRef]
21. Trifonov, P.; Trofimiuk, G. A randomized construction of polar subcodes. In Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 25–30 June 2017; pp. 1863–1867. [CrossRef]
22. Huang, L.; Zhang, H.; Li, R.; Ge, Y.; Wang, J. AI Coding: Learning to Construct Error Correction Codes. *IEEE Trans. Commun.* **2020**, *68*, 26–39. [CrossRef]
23. Romano, G. ; Ciuonzo, D. Minimum-Variance Importance-Sampling Bernoulli Estimator for Fast Simulation of Linear Block Codes over Binary Symmetric Channels. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 486–496. [CrossRef]
24. Minja, A. ; Šenk, V. Quasi-Analytical Simulation Method for Estimating the Error Probability of Star Domain Decoders *IEEE Trans. Commun.* **2019**,*67*, 3101–3113. [CrossRef]
25. Wu, W.; Siegel, P.H. Generalized Partial Orders for Polar Code Bit-Channels. *IEEE Trans. Inf. Theory* **2019**, *65*, 7114–7130. [CrossRef]
26. Saptharishi, R.; Shpilka, A.; Volk, B.L. Efficiently Decoding Reed–Muller Codes From Random Errors. *IEEE Trans. Inf. Theory* **2017**, *63*, 1954–1960. [CrossRef]
27. Kudekar, S.; Kumar, S.; Mondelli, M.; Pfister, H.D.; Sasoglu, E.; Urbanke, R. Reed-Muller Codes Achieve Capacity on Erasure Channels. *IEEE Trans. Inf. Theory* **2017**, *63*, 4298–4316. [CrossRef]
28. Kumar, S.; Calderbank, R.; Pfister, H.D. Beyond double transitivity: Capacity-achieving cyclic codes on erasure channels. In *Proceedings of the* 2016 IEEE Information Theory Workshop (ITW), Cambridge, UK, 11–14 September 2016; pp. 241–245. [CrossRef]
29. Ordentlich, E.; Roth, R.M. On the Pointwise Threshold Behavior of the Binary Erasure Polarization Subchannels. *IEEE Trans. Inf. Theory* **2019**, *65*, 6044–6055. [CrossRef]

30. Drăgoi, V.F.; Beiu, V. Studying the Binary Erasure Polarization Subchannels Using Network Reliability. *IEEE Commun. Lett.* **2020**, *24*, 62–66. [CrossRef]
31. Stanley, R.P. *Enumerative Combinatorics*; Cambridge University Press: Cambridge, NY, USA, 2012.
32. Cristescu, G.; Drăgoi, V.F. Cubic Spline Approximation of the Reliability Polynomials of Two Dual Hammock Networks. *Transylv. J. Math. Mech.* **2019**, *11*, 77–90.
33. Cristescu, G.; Drăgoi, V.F. Efficient approximation of two-terminal networks reliability polynomials using cubic splines. *IEEE Trans. Reliab.* **2021**, *1–11*. [CrossRef]
34. Mondelli, M. From Polar to Reed-Muller Codes: Unified Scaling, Non-standard Channels, and a Proven Conjecture. Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2016.
35. Richardson, T.; Urbanke, R. *Modern Coding Theory*; Cambridge University Press: New York, NY, USA, 2008.
36. Roth, R.M. *Introduction to Coding Theory*; Cambridge University Press: New York, NY, USA, 2006.
37. Carlet, C. Boolean functions for cryptography and error correcting codes. In *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*; Cambridge University Press: Cambridge, NY, USA, 2010; Chapter 8; pp. 257–397.
38. Moore, E.F.; Shannon, C.E. Reliable circuits using less reliable relays—Part I. *J. Frankl. Inst.* **1956**, *262*, 191–208. [CrossRef]
39. Drăgoi, V.; Cowell, S.R.; Beiu, V.; Hoară, S.; Gaşpar, P. How Reliable are Compositions of Series and Parallel Networks Compared with Hammocks? *Int. J. Comput. Commun. Control* **2018**, *13*, 772–791. [CrossRef]
40. Colbourn, C.J. *The Combinatorics of Network Reliability*; Oxford University Press: New York, NY, USA, 1987.
41. Dăuş, L.; Jianu, M. The shape of the reliability polynomial of a hammock network. In *Intelligent Methods in Computing, Communication and Control. ICCC 2020, Advances in Intelligent Systems and Computing*; Springer: Berlin/Heidelberg, Germany, 2021; Volume 1243, pp. 93–105. [CrossRef]
42. Brown, J.; Cox, D.; Ehrenborg, R. The average reliability of a graph. *Discret. Appl. Math.* **2014**, *177*, 19–33. [CrossRef]