The Combinatorics of Stopping Sets in Product Codes

Fanny Jardel and Joseph J. Boutros

Telecom ParisTech
&
Texas A&M University at Qatar

July 5, 2016

Motivation

- Edge coloring of binary product codes (Boutros, Zémor, et al. 2008) but no stopping sets analysis.
- Social networks are using Reed-Solomon codes for storage.
- We are not building codes for locality. We are only interested by rate and performance. However, iterative row-column decoding of product codes is simple enough.
- We consider product codes with MDS components.
  F. Jardel and J.J. Boutros, Edge Coloring and Stopping Sets Analysis in Product Codes with MDS components, submitted to the *IEEE Trans. Inf. Theory*, Dec. 2015. ArXiv 1603.01468.

Some Literature

- Elias, Error-free coding, 1954.
- Reddy and J.P. Robinson, Random Error and Burst Correction by Iterated Codes, 1972.
- Sendrier, Codes correcteurs d'erreurs à haut pouvoir de correction, 1991.
- Pyndiah, Near-optimum decoding of product codes: Block turbo codes, 1998.
- Sella and Be'ery, Convergence analysis of turbo decoding of product codes, 2001.
- Tolhuizen, More results on the weight enumerator of product codes, 2002.
- Schwartz, Siegel, and Vardy, On the asymptotic performance of iterative decoders for product codes, 2005.
- Rosnes, Stopping set analysis of iterative row-column decoding of product codes, 2008.

Definition: Product Code

See MacWilliams & Sloane 1977 and Kschischang 2003.

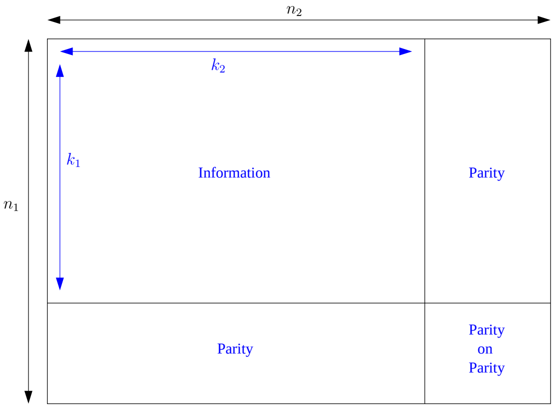
- Column code C_1: linear block code over $\mathbb{F}_q$ with parameters $[n_1, k_1, d_1]_q$.
- Row code C_2: linear block code with parameters $[n_2, k_2, d_2]_q$.
- Let G_1 and G_2 be two generator matrices of size $k_1 \times n_1$ and $k_2 \times n_2$ for C_1 and C_2 respectively.
- A product code C_P is constructed as a subspace of $\mathbb{F}_q^N$ with generator matrix $G_P = G_1 \otimes G_2$, where $N = n_1 n_2$ and $\otimes$ denotes the Kronecker product.
- C_P has dimension $K = k_1 k_2$ and minimum Hamming distance $d_P = d_1 d_2$.

Definition: Product Code

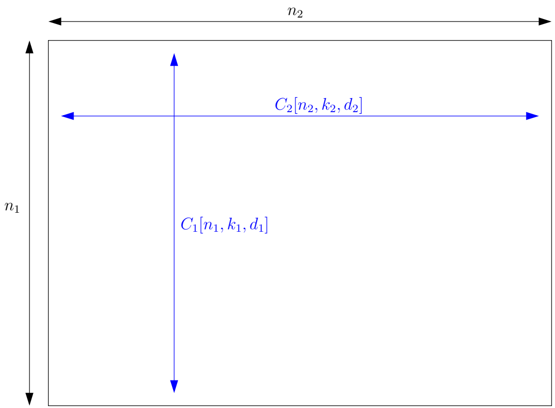
See MacWilliams & Sloane 1977 and Kschischang 2003.

- Column code C_1: linear block code over $\mathbb{F}_q$ with parameters $[n_1, k_1, d_1]_q$.
- Row code C_2: linear block code with parameters $[n_2, k_2, d_2]_q$.
- Let G_1 and G_2 be two generator matrices of size $k_1 \times n_1$ and $k_2 \times n_2$ for C_1 and C_2 respectively.
- A product code C_P is constructed as a subspace of $\mathbb{F}_q^N$ with generator matrix $G_P = G_1 \otimes G_2$, where $N = n_1 n_2$ and $\otimes$ denotes the Kronecker product.
- C_P has dimension $K = k_1 k_2$ and minimum Hamming distance $d_P = d_1 d_2$.

Definition: Product Code



Definition: Product Code



Decoders for Product Codes

- Type I: ML decoder. This is a non-iterative decoder. It is based on a Gaussian reduction of the parity-check matrix of the product code.
- Type II: Iterative algebraic decoder. At odd decoding iterations, component codes C_1 on each column are decoded via an algebraic decoder (bounded-distance) that fills up to $d-1$ erasures. Similarly, at even decoding iterations, component codes C_2 on each row are decoded via an algebraic decoder.
- Type III: Iterative ML-per-component decoder. This decoder was considered by E. Rosnes (2008) for binary product codes. At odd decoding iterations, column codes C_1 are decoded via an optimal decoder (ML for C_1). At even decoding iterations, row codes C_2 are decoded via a similar optimal decoder (ML for C_2).
- Type IV: Iterative belief-propagation decoder based on the Tanner graph of $C_{\mathcal{P}}$ as studied by Schwartz and Vardy (2006) for general linear block codes and for low-density parity-check codes by Di, Proietti, Telatar, Richardson, Urbanke (2002).

Decoders for Product Codes

- Type I: ML decoder. This is a non-iterative decoder. It is based on a Gaussian reduction of the parity-check matrix of the product code.
- Type II: Iterative algebraic decoder. At odd decoding iterations, component codes C_1 on each column are decoded via an algebraic decoder (bounded-distance) that fills up to $d-1$ erasures. Similarly, at even decoding iterations, component codes C_2 on each row are decoded via an algebraic decoder.
- Type III: Iterative ML-per-component decoder. This decoder was considered by E. Rosnes (2008) for binary product codes. At odd decoding iterations, column codes C_1 are decoded via an optimal decoder (ML for C_1). At even decoding iterations, row codes C_2 are decoded via a similar optimal decoder (ML for C_2).
- Type IV: Iterative belief-propagation decoder based on the Tanner graph of $C_{\mathcal{P}}$ as studied by Schwartz and Vardy (2006) for general linear block codes and for low-density parity-check codes by Di, Proietti, Telatar, Richardson, Urbanke (2002).

Decoders for Product Codes

- Type I: ML decoder. This is a non-iterative decoder. It is based on a Gaussian reduction of the parity-check matrix of the product code.
- Type II: Iterative algebraic decoder. At odd decoding iterations, component codes C_1 on each column are decoded via an algebraic decoder (bounded-distance) that fills up to $d-1$ erasures. Similarly, at even decoding iterations, component codes C_2 on each row are decoded via an algebraic decoder.
- Type III: Iterative ML-per-component decoder. This decoder was considered by E. Rosnes (2008) for binary product codes. At odd decoding iterations, column codes C_1 are decoded via an optimal decoder (ML for C_1). At even decoding iterations, row codes C_2 are decoded via a similar optimal decoder (ML for C_2).
- Type IV: Iterative belief-propagation decoder based on the Tanner graph of $C_{\mathcal{P}}$ as studied by Schwartz and Vardy (2006) for general linear block codes and for low-density parity-check codes by Di, Proietti, Telatar, Richardson, Urbanke (2002).

Decoders for Product Codes

- Type I: ML decoder. This is a non-iterative decoder. It is based on a Gaussian reduction of the parity-check matrix of the product code.
- Type II: Iterative algebraic decoder. At odd decoding iterations, component codes C_1 on each column are decoded via an algebraic decoder (bounded-distance) that fills up to $d-1$ erasures. Similarly, at even decoding iterations, component codes C_2 on each row are decoded via an algebraic decoder.
- Type III: Iterative ML-per-component decoder. This decoder was considered by E. Rosnes (2008) for binary product codes. At odd decoding iterations, column codes C_1 are decoded via an optimal decoder (ML for C_1). At even decoding iterations, row codes C_2 are decoded via a similar optimal decoder (ML for C_2).
- Type IV: Iterative belief-propagation decoder based on the Tanner graph of $C_{\mathcal{P}}$ as studied by Schwartz and Vardy (2006) for general linear block codes and for low-density parity-check codes by Di, Proietti, Telatar, Richardson, Urbanke (2002).

The Rectangular Support

Useful later to characterize a stopping set in a row-column (bi-dimensional) product code.

- Let $\mathcal{S} \subseteq \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ be a set of symbol positions in the product code.
- The set of row positions associated to $\mathcal{S}$ is $\mathcal{R}_1(\mathcal{S}) = \{i_1, \ldots, i_{\ell_1}\}$ where $|\mathcal{R}_1(\mathcal{S})| = \ell_1$ and for all $i \in \mathcal{R}_1(\mathcal{S})$ there exists $(i, \ell) \in \mathcal{S}$.
- The set of column positions associated to $\mathcal{S}$ is $\mathcal{R}_2(\mathcal{S}) = \{j_1, \ldots, j_{\ell_2}\}$ where $|\mathcal{R}_2(\mathcal{S})| = \ell_2$ and for all $j \in \mathcal{R}_2(\mathcal{S})$ there exists $(\ell, j) \in \mathcal{S}$.
- The rectangular support of $\mathcal{S}$ is

$$\mathcal{R}(\mathcal{S}) = \mathcal{R}_1(\mathcal{S}) \times \mathcal{R}_2(\mathcal{S}), \tag{1}$$

i.e. the smallest $\ell_1 \times \ell_2$ rectangle including all columns and all rows of $\mathcal{S}$.

Type-III Stopping Set

Definition

Consider a product code $C_P = C_1 \otimes C_2$. Let $\mathcal{S} \subseteq \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ with $|\mathcal{R}_1(\mathcal{S})| = \ell_1$ and $|\mathcal{R}_2(\mathcal{S})| = \ell_2$. Consider the ℓ_1 rows of $\mathcal{S}$ given by $\mathcal{S}_r^{(i)} = \{j : (i,j) \in \mathcal{S}\}$ and the ℓ_2 columns of $\mathcal{S}$ given by $\mathcal{S}_c^{(j)} = \{i : (i,j) \in \mathcal{S}\}$. The set $\mathcal{S}$ is a stopping set of type III for C_P if there exist linear subcodes $C_c^{(j)} \subseteq C_1$ and $C_r^{(i)} \subseteq C_2$ such that $\mathcal{X}(C_c^{(j)}) = \mathcal{S}_c^{(j)}$ and $\mathcal{X}(C_r^{(i)}) = \mathcal{S}_r^{(i)}$ for all $i \in \mathcal{R}_1(\mathcal{S})$ and for all $j \in \mathcal{R}_2(\mathcal{S})$.

Type-II Stopping Set

Definition

Consider a product code $C_P = C_1 \otimes C_2$. Let $\mathcal{S} \subseteq \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ with $|\mathcal{R}_1(\mathcal{S})| = \ell_1$ and $|\mathcal{R}_2(\mathcal{S})| = \ell_2$. Consider the ℓ_1 rows of $\mathcal{S}$ given by $\mathcal{S}_r^{(i)} = \{j : (i,j) \in \mathcal{S}\}$ and the ℓ_2 columns of $\mathcal{S}$ given by $\mathcal{S}_c^{(j)} = \{i : (i,j) \in \mathcal{S}\}$. The set $\mathcal{S}$ is a stopping set of type II for C_P if $|\mathcal{S}_r^{(i)}| \geq d_2$ and $|\mathcal{S}_c^{(j)}| \geq d_1$, for all $i \in \mathcal{R}_1(\mathcal{S})$ and for all $j \in \mathcal{R}_2(\mathcal{S})$.

Basic Results for MDS codes (1)

Proposition

Let $C[n,k,d]_q$ be a linear code with $q \geq 2$. Assume that C is not MDS and the n symbols of a codeword are transmitted on an erasure channel. Then, there exists an erasure pattern of weight greater than $d-1$ that is ML-correctable.

Proof.

Let H be an $(n-k) \times n$ parity-check matrix of C with rank $n-k > d-1$. For any integer w in the range $[d, n-k]$, there exists a set of w linearly independent columns in H. Choose an erasure pattern of weight w with erasures located at the positions of the w independent columns. Then, the ML decoder is capable of solving all these erasures by simple Gaussian reduction of H. □

Corollary

Let $C[n,k,d]_q$ be an MDS code. All erasure patterns of weight greater than $d-1$ are not ML-correctable.

Basic Results for MDS codes (2)

We conclude from the previous corollary that an algebraic decoder for an MDS code attains the word-error performance of its ML decoder.

What about symbol-error performance?

Proposition

Let $C[n,k,d]_q$ be a non-binary MDS code ($q > n > 2$). For any w satisfying $d \leq w \leq n$ and any support $\mathcal{X} = \{i_1, i_2, \ldots, i_w\}$, where $1 \leq i_j \leq n$, there exists a codeword in C of weight w having $\mathcal{X}$ as its own support.

Basic Results for MDS codes (2)

Proof.

Any set of $n-k$ columns of H is full-rank (MDS code). $w = n-k+\ell = r+\ell$, where $\ell = 1 \dots k$. Name r independent columns of H as $h_1 \dots h_r$. The remaining ℓ columns are denoted $\zeta_1 \dots \zeta_\ell$. For any $j = 1 \dots \ell$, we have $\zeta_j = \sum_{i=1}^{r} a_{i,j} h_i$, where $a_{i,j} \in \mathbb{F}_q \setminus \{0\}$ otherwise it contradicts $d = n-k+1$. Now, select $\alpha_1 \dots \alpha_\ell$ from $\mathbb{F}_q \setminus \{0\}$ such that: α_1 is arbitrary, α_2 is chosen outside the set $\{-\alpha_1 a_{i,1}/a_{i,2}\}_{i=1}^{r}$, then α_3 is chosen outside the set $\{(-\alpha_1 a_{i,1} - \alpha_2 a_{i,2})/a_{i,3}\}_{i=1}^{r}$, and so on, up to α_ℓ which is chosen outside the set $\{-\sum_{u=1}^{\ell-1} \alpha_u a_{i,u}/a_{i,\ell}\}_{i=1}^{r}$. The equality

$$\sum_{j=1}^{\ell} \alpha_j \zeta_j = \sum_{i=1}^{r} \sum_{j=1}^{\ell} \alpha_j a_{i,j} h_i$$

produces a codeword of Hamming weight w. Hence, there exists a codeword of weight w with non-zero symbols in all positions given by $\mathcal{X}$. □

Type-II and Type-III Stopping Sets are Identical for MDS-Based Product Codes

Stopping sets are identical when dealing with algebraic and ML-per-component decoders, i.e. type-II and type-III stopping sets are identical thanks to Corollary 3 and Proposition 2.

In the sequel, component codes C_1 (column) and C_2 (row) of a product code are assumed to be MDS.

Characterization of Stopping Sets

Stopping sets can be characterized as follows.

- Obvious or not obvious sets, also known as rank-1 sets. A stopping set $\mathcal{S}$ is obvious if $\mathcal{S} = \mathcal{R}(\mathcal{S})$.
- Primitive or non-primitive stopping sets. A stopping set is primitive if it cannot be partitioned into two or more smaller stopping sets. Notice that all stopping sets, whether they are primitive or not, are involved in the code performance.
- Codeword or non-codeword. A stopping set $\mathcal{S}$ is said to be a codeword stopping set if there exists a codeword c in C_P such that $\mathcal{X}(c) = \mathcal{S}$.
- ML-correctable or non-ML-correctable. A stopping set $\mathcal{S}$ cannot be corrected via ML decoding if it includes the support of a non-zero codeword.

Illustration: type-II stopping set with $w = 9$

Example

Consider a $[n_1, n_1 - 2, 3]_q \otimes [n_2, n_2 - 2, 3]_q$ product code. A stopping set $\mathcal{S}$ of size $w = 9$ is shown as a weight-9 matrix of size $n_1 \times n_2$, where 1 corresponds to an erased position:

$$\mathcal{S} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{R}(\mathcal{S}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

We took $n_1 = n_2 = 7$ for illustration.The rectangular support is shown in a compact representation as a matrix of size $\ell_1 \times \ell_2 = 3 \times 3$, This stopping set is obvious. It is not ML-correctable because it is a product-code codeword.

Illustration: type-II stopping set with $w = 12$

Example

Same $[n_1, n_1 - 2, 3]_q \otimes [n_2, n_2 - 2, 3]_q$ product code as in the previous example.

$$\mathcal{S}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{S}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$\mathcal{S}_1$ and $\mathcal{S}_2$ are not obvious. $\mathcal{S}_1$ is ML-correctable. $\mathcal{S}_2$ is not ML-correctable.

$$\mathcal{R}(\mathcal{S}_1) = \mathcal{R}(\mathcal{S}_2) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}. \qquad (2)$$

For $w = 12$, it is possible to build obvious stopping sets 3×4 and 4×3.

The Rectangle Size and Number of Zeros

Lemma

Given a weight $w \leq (d_1+1)(d_2+1)$ *and assuming* $\tau_w > 0$*, then* $\exists \mathcal{S}^0$ *such that* $\forall \mathcal{S}$ *with* $|\mathcal{S}| = w$*, we have* $\|\mathcal{R}(\mathcal{S})\| \leq \|\mathcal{R}(\mathcal{S}^0)\| = (\ell_1^0, \ell_2^0)$*, where*

$$\ell_1^0 \leq d_1 + 1 + \left\lfloor \frac{d_1+1}{d_2} \right\rfloor, \qquad \ell_2^0 \leq d_2 + 1 + \left\lfloor \frac{d_2+1}{d_1} \right\rfloor.$$

Lemma

Let $\mathcal{R}(\mathcal{S})$ *be the* $\ell_1 \times \ell_2$ *rectangular support of a stopping set* $\mathcal{S}$ *of size* w*. Let* $\beta = \ell_1 \ell_2 - w$ *be the number of zero positions, or equivalently* β *is the size of the set* $\mathcal{R}(\mathcal{S}) \setminus \mathcal{S}$*. Then*

$$\beta \leq \min((\ell_1 - d_1)\ell_2, \ell_1(\ell_2 - d_2)).$$

Stopping Sets and Bipartite Graphs

A stopping set of weight w and having a $\ell_1 \times \ell_2$ rectangular support shall be represented by a bipartite graph with ℓ_1 left vertices, ℓ_2 right vertices, and a total of $\beta = \ell_1\ell_2 - w$ edges.

Notice that these bipartite graphs have no length-2 cycles because parallel edges are forbidden.

For finite ℓ_1 and ℓ_2, given the left degree distribution and the right degree distribution, there exists no exact formula for counting bipartite graphs. The best recent results are asymptotic in the graph size for sparse and dense matrices (E.R. Canfield and B.D. McKay 2009, C. Greenhill and B.D. McKay 2012) and cannot be applied in our enumeration.

The following two lemmas solve two cases encountered in our Theorem on stopping sets enumeration for $w = d(d+2)$ and $w = (d+1)(d+1)$ both inside a $(d+2) \times (d+2)$ rectangular support. The definition of special partitions is required before introducing the two lemmas.

Special Partitions

Definition

Let $\ell \geq 2$ be an integer. A *special partition* of length j of ℓ is a partition defined by a tuple $(\ell_1, \ell_2, \ldots, \ell_j)$ such that its integer components satisfy:

- $\ell_1 \leq \ell_2 \leq \ldots \leq \ell_j$.
- $\sum_{i=1}^{j} \ell_i = \ell$.
- $\ell_i \geq 2$, $\forall j$.
- $1 \leq j \leq \ell/2$.

A special partition shall be denoted by $((\ell_1, \ldots, \ell_j))$.

Definition

The *group number* of a special partition, denoted by $\kappa = \kappa(\ell_1, \ell_2, \ldots, \ell_j)$, is the number of different integers ℓ_j, for $j = 1 \ldots \ell/2$. In other words, following set theory, the set including the j integers ℓ_i's is $\{\ell_{i_1}, \ell_{i_2}, \ldots, \ell_{i_\kappa}\}$. The group number divides the partition of ℓ into κ groups where the m^{th} group includes ℓ_{i_m} repeated g_m times, and $\sum_{m=1}^{\kappa} g_m = j$.

Degree-2 Bipartite Graphs

Lemma

Consider bipartite graphs defined as follows: ℓ left vertices, ℓ right vertices, all vertices have degree 2, and no length-2 cycles are allowed. For $\ell \geq 2$, the total number x_ℓ of such bipartite graphs is given by the expression

$$x_\ell = \sum_{((\ell_1,\ldots,\ell_j))} \frac{1}{\prod_{m=1}^{\kappa(\ell_1,\ldots,\ell_j)} g_m!} \prod_{k=1}^{j} \frac{\prod_{u=0}^{\ell_k-1} (\ell - \sum_{i=1}^{k-1} \ell_i - u)^2}{2\ell_k} \qquad (3)$$

where $\sum_{((\ell_1,\ldots,\ell_j))}$ is a summation over all special partitions of the integer ℓ, $\kappa(\ell_1,\ldots,\ell_j)$ is the group number of the special partition $((\ell_1,\ldots,\ell_j))$, and g_m is the size of the m^{th} group.

Degree-2 Bipartite Graphs

Proof. Firstly, let us find the number of Hamiltonian bipartite graphs having ℓ_k left vertices, ℓ_k right vertices, all vertices of degree 2, and no length-2 cycles allowed. This number is:

$$\frac{(\ell_k!)^2}{2\ell_k}. \tag{4}$$

Secondly, given the half-size ℓ of the bipartite graph stated in this lemma, all special partitions of ℓ are considered. The number of choices for selecting the vertices of the j Hamiltonian graphs is

$$\prod_{k=1}^{j} \binom{\ell - \sum_{i=1}^{k-1} \ell_i}{\ell_k}^2. \tag{5}$$

The above number should be multiplied by the number of Hamiltonian graphs for each selection of vertices to get

$$\prod_{k=1}^{j} \binom{\ell - \sum_{i=1}^{k-1} \ell_i}{\ell_k}^2 \frac{(\ell_k!)^2}{2\ell_k}.$$

Degree-2 Bipartite Graphs

But for a given special partition, each group of size g_m is creating $g_m!$ identical bipartite graphs. Hence, the final result for a fixed partition becomes

$$\frac{1}{\prod_{m=1}^{\kappa(\ell_1,\ldots,\ell_j)} g_m!}\prod_{k=1}^{j}\binom{\ell-\sum_{i=1}^{k-1}\ell_i}{\ell_k}^2\frac{(\ell_k!)^2}{2\ell_k}.$$

Then, x_ℓ is obtained by summing over all special partitions of the integer ℓ to yield

$$x_\ell=\sum_{((\ell_1,\ldots,\ell_j))}\frac{1}{\prod_{m=1}^{\kappa(\ell_1,\ldots,\ell_j)} g_m!}\prod_{k=1}^{j}\binom{\ell-\sum_{i=1}^{k-1}\ell_i}{\ell_k}^2\frac{(\ell_k!)^2}{2\ell_k}.$$

The simplification of the factors $(\ell_k!)^2$ yields the expression stated by this lemma.□

Degree-2 Bipartite Graphs (One vertex of Degree 1)

Lemma

Consider bipartite graphs defined as follows: ℓ left vertices, ℓ right vertices, all left vertices have degree 2 except one vertex of degree 1, all right vertices have degree 2 except one vertex of degree 1, and finally no length-2 cycles are allowed. For $\ell \geq 3$, the total number y_ℓ of such bipartite graphs is

$$y_\ell = \ell^2 \cdot \left((2\ell - 1) \cdot x_{\ell-1} + (\ell - 1)^2 \cdot x_{\ell-2}\right), \tag{6}$$

where x_ℓ is determined via the previous lemma and $x_1 = 0$.

Degree-2 Bipartite Graphs (One vertex of Degree 1)

Proof. Let the first $\ell - 1$ left vertices and the first $\ell - 1$ right vertices be of degree 2. There exists two ways to complete this bipartite graph such that the two remaining vertices have degree 1.

1) Each of the $x_{\ell-1}$ sub-graphs has $2(\ell-1)$ edges. Break one edge into two edges and connect them to the remaining left and right vertices, the number of such graphs is $2(\ell-1)x_{\ell-1}$. Another set of $x_{\ell-1}$ bipartite graphs is built by directly connecting the last two vertices together without breaking any edge in the upper sub-graph. Now, we find $2(\ell-1)x_{\ell-1} + x_{\ell-1} = (2\ell-1)x_{\ell-1}$ bipartite graphs.

2) Fix a vertex among the $\ell - 1$ upper left vertices and fix one among the $\ell - 1$ upper right vertices ($(\ell-1)^2$ choices). Consider a length-2 cycle including these two vertices. One edge of this cycle can be broken into two edges and then attached to the degree-1 vertices at the bottom. The remaining $\ell - 2$ left and right vertices may involve $x_{\ell-2}$ sub-graphs. Consequently, the number of graphs in this second case is $(\ell-1)^2 x_{\ell-2}$.

The total number of bipartite graphs enumerated in the above cases is

$$(2\ell-1)x_{\ell-1} + (\ell-1)^2 x_{\ell-2}.$$

Finally, the degree-1 left vertex has ℓ choices and so has the degree-1 right vertex. The number of graphs should be multiplied by ℓ^2. □

Tables of Special Partitions and Bipartite Graphs

The first table below shows the number of special partitions for $\ell = 2 \ldots 32$. The number of standard partitions (the partition function) can be found by a recursion resulting from the pentagonal number theorem. To our knowledge, there exists no such recursion for special partitions. The number of bipartite graphs under the assumptions of the previous lemmas is found in the second table for a graph half-size up to 8.

1, 1, 2, 2, 4, 4, 7, 8, 12, 14, 21, 24, 34, 41, 55, 66, 88, 105, 137, 165, 210, 253, 320, 383, 478, 574, 708, 847, 1039, 1238, 1507

ℓ	2	3	4	5	6	7	8
x_ℓ	1	6	90	2040	67950	3110940	187530840
y_ℓ	0	45	816	22650	888840	46882710	3199593600

Finally, we are ready to state and prove the first theorem on stopping sets enumeration.

Stopping Sets Enumeration (1)

Theorem: Let C_P be a product code $[n_1, k_1, d_1]_q \otimes [n_2, k_2, d_2]_q$ built from row and column MDS component codes, where the alphabet size q is greater than $\max(n_1, n_2)$. Let τ_w be the number of stopping sets of size w. We write $\tau_w = \tau^a + \tau^b$, where τ^a counts obvious stopping sets and τ^b counts non-obvious stopping sets. Under (type-II) iterative algebraic decoding and for $d_1 = d_2 = d \geq 2$, stopping sets are characterized as follows:

- For $w < d^2$, $\tau^a = \tau^b = 0$.
- For $w = d^2$, $\tau^a = \binom{n_1}{d}\binom{n_2}{d}$ and $\tau^b = 0$.
- For $w \in]d^2, d(d+1)[$, $\tau^a = \tau^b = 0$.
- For $w = d(d+1)$,

$$\tau^a = \binom{n_1}{d}\binom{n_2}{d+1} + \binom{n_1}{d+1}\binom{n_2}{d}, \quad \tau^b = (d+1)!\binom{n_1}{d+1}\binom{n_2}{d+1}.$$

- For $w \in\ ]d(d+1), d(d+2)[$.
  Let us write $w = d^2 + d + \lambda$, where $\lambda \in [1, d-1]$.

$$\tau^a = 0, \quad \tau^b = (d+1-\lambda)!\binom{d+1}{\lambda}^2\binom{n_1}{d+1}\binom{n_2}{d+1}.$$

Stopping Sets Enumeration (2)

- For $w = d(d+2)$,

$$\tau^a = \binom{n_1}{d}\binom{n_2}{d+2} + \binom{n_1}{d+2}\binom{n_2}{d},$$

$$\begin{aligned}\tau^b = {} & (d+1)^2 \binom{n_1}{d+1}\binom{n_2}{d+1} \\ & + \sum_{2r_0+r_1=d} \binom{d+1}{r_0}\binom{d+1-r_0}{r_1}\frac{(d+2)!}{2^{r_2}} \\ & \left[\binom{n_1}{d+1}\binom{n_2}{d+2} + \binom{n_1}{d+2}\binom{n_2}{d+1}\right] \\ & + x_{d+2}\binom{n_1}{d+2}\binom{n_2}{d+2},\end{aligned}$$

where $\sum_{2r_0+r_1=d}$ is a summation over r_0 and r_1, both being non-negative and satisfying $2r_0 + r_1 = d$, $r_2 = d + 1 - r_0 - r_1$, and x_{d+2} is the number of degree-2 bipartite graphs as given by the previous lemma.

Stopping Sets Enumeration (3)

- For $w = (d+1)(d+1)$

$$\tau^a = \binom{n_1}{d+1}\binom{n_2}{d+1},$$

$$\tau^b = \sum_{2r_0+r_1=d+1} \binom{d+1}{r_0}\binom{d+1-r_0}{r_1}\frac{(d+2)!}{2^{r_0}}$$

$$\left[\binom{n_1}{d+1}\binom{n_2}{d+2} + \binom{n_1}{d+2}\binom{n_2}{d+1}\right]$$

$$+\, y_{d+2}\binom{n_1}{d+2}\binom{n_2}{d+2},$$

where y_{d+2} is the number of degree-2 bipartite graphs, except for one left vertex and one right vertex having degree 1. The number y_{d+2} is given by the previous lemma.

The detailed proof of this Theorem is found in (Jardel and Boutros 2015, Th.2).

Illustration: enumeration of stopping sets for $w = 15$

How to compute τ_w for $w = 15$ and $d = 3$?
Stopping set of size $w = d(d+2)$.

- $d \times (d+2)$ rectangular support corresponds to obvious sets τ_a.
  $\tau^a = \binom{n_1}{d}\binom{n_2}{d+2} + \binom{n_1}{d+2}\binom{n_2}{d}$.
- Non-obvious sets counted by τ_b correspond to three different sizes of rectangular support: $(d+1) \times (d+1)$, $(d+1) \times (d+2)$, and $(d+2) \times (d+2)$.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

$$(r_0, r_1) = (1, 1) \qquad (r_0, r_1) = (0, 3)$$

The number of 4×5 matrices is given by:

$$\sum_{2r_0 + r_1 = d} \binom{d+1}{r_0} \binom{d+1-r_0}{r_1} \frac{(d+2)!}{2^{r_2}} = 600.$$

Performance with Independent Erasures (1)

Proposition

Let $C_P = C_1 \otimes C_2$ be a product code with non-binary MDS components. All obvious stopping sets are supports of product code codewords.

Proof.

Consider an $\ell_1 \times \ell_2$ obvious stopping set. Its rectangular support is $\mathcal{R}(\mathcal{S}) = \mathcal{R}_1(\mathcal{S}) \times \mathcal{R}_2(\mathcal{S})$. We have $\ell_1 \geq d_1$ and $\ell_2 \geq d_2$. There exists a column codeword $x = (x_1, x_2, \ldots, x_{n_1}) \in C_1$ of weight ℓ_1 with support $\mathcal{R}_1(\mathcal{S}) \times \{j_1\}$, where $j_1 \in \mathcal{R}_2(\mathcal{S})$. Similarly, there exists a row codeword $y = (y_1, y_2, \ldots, y_{n_2}) \in C_2$ of weight ℓ_2 with support $\{i_1\} \times \mathcal{R}_2(\mathcal{S})$, where $i_1 \in \mathcal{R}_1(\mathcal{S})$. The Kronecker product of x and y satisfies $\mathcal{X}(x \otimes y) = \mathcal{S}$. □

Performance with Independent Erasures (2)

Corollary

Consider a product code $C_P = C_1 \otimes C_2$ with non-binary MDS component codes. Assume the symbols of C_P are transmitted over a $SEC(q, \epsilon)$ channel. Let $P_{ew}^{\mathcal{G}}$ be the word error probability of an iterative (type-II) decoder and P_{ew}^{ML} be the word error probability of ML decoding. Then, for $\epsilon \ll 1$, the error probabilities satisfy $P_{ew}^{\mathcal{G}} \sim P_{ew}^{ML}$.

Performance with Independent Erasures (2)

Proof.

On the $SEC(q, \epsilon)$, the word error probabilities are given by (Schwartz and Vardy 2006),

$$P_{ew}^{ML} = \sum_{i=d_1 d_2}^{N} \Psi_i(ML)\epsilon^i (1-\epsilon)^{N-i}, \tag{7}$$

where $\Psi_i(ML)$ is the number of weight-i erasure patterns covering a product code codeword, and

$$P_{ew}^{\mathcal{G}} = \sum_{i=d_1 d_2}^{N} \Psi_i(\mathcal{G})\epsilon^i (1-\epsilon)^{N-i}, \tag{8}$$

where $\Psi_i(\mathcal{G})$ is the number of weight-i erasure patterns covering a stopping set. Asymptotic length analysis is not considered in this paper, i.e. $N = n_1 n_2$ is fixed. We write $P_{ew}^{ML} = \Psi_{d_1 d_2}(ML)\epsilon^{d_1 d_2} + o(\epsilon^{d_1 d_2})$ and $P_{ew}^{\mathcal{G}} = \Psi_{d_1 d_2}(\mathcal{G})\epsilon^{d_1 d_2} + o(\epsilon^{d_1 d_2})$. From the previous proposition, we get the equality $\Psi_{d_1 d_2}(\mathcal{G}) = \Psi_{d_1 d_2}(ML)$ and so $\lim_{\epsilon \to 0} P_{ew}^{\mathcal{G}} / P_{ew}^{ML} = 1$. □

Performance with Independent Erasures (3)

For $P_{ew}^{\mathcal{G}}$, thanks to the enumeration Theorem, a union bound can be easily established. Indeed, we have

$$\begin{aligned} P_{ew}^{\mathcal{G}} &= Prob(\exists \mathcal{S}\ covered) \\ &\leq \sum_{w} Prob(\exists \mathcal{S} : |\mathcal{S}| = w, \mathcal{S}\ covered), \end{aligned}$$

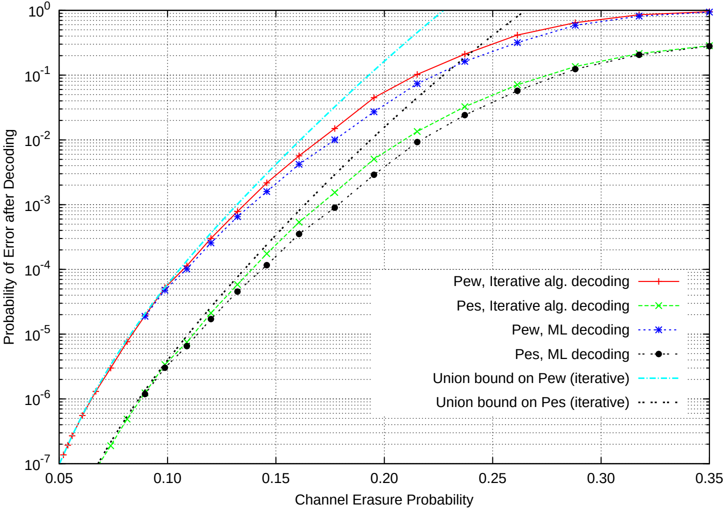
leading to

$$P_{ew}^{\mathcal{G}} \leq P^{U}(\epsilon) = \sum_{w=d_1 d_2}^{N} \tau_w \epsilon^w. \tag{9}$$

From the theorem, the union bound $P^U(\epsilon)$ for the $[12, 10, 3]_q^{\otimes 2}$ product code is

$$\begin{aligned} P^U(\epsilon) =& 48400\epsilon^9 + 6098400\epsilon^{12} + 23522400\epsilon^{13} + 17641800\epsilon^{14} \\ &+ 1754335440\epsilon^{15} + 9126691200\epsilon^{16} + o(\epsilon^{16}). \end{aligned}$$

Performance with Independent Erasures (4)



Product code $[12,10]_q^{\otimes 2}$, no edge coloring. Word and symbol error rate performance for iterative decoding versus its union bound and ML decoding.

Finite-Length Regime: $[14,12]_q \otimes [16,14]_q$ *product code*

Finite-regime BEC bounds from Polyanskiy-Poor-Verdù (2010) are directly applicable to our product codes over the $SEC(q,\epsilon)$. The BEC channel dispersion is $V=\epsilon(1-\epsilon)$ and its maximal achievable rate is

$$R=(1-\epsilon)-\sqrt{\frac{V}{n}}Q^{-1}(P_{ew})+O(\frac{1}{n}), \tag{10}$$

where n is the code length, $Q(x)$ is the Gaussian tail function, ϵ is the channel erasure probability, and P_{ew} is the target word error probability. The next table shows how good is the proposed product code based on MDS components. The value of ϵ in the third column is given for $P_{ew}=10^{-2}$ at all rows.

	Coding Rate R for $\epsilon=0.15$	Erasure Prob. ϵ for $R=0.75$
Polyanskiy-Poor-Verdú	$0.794:\ P_{ew}=1.0\cdot10^{-2}$	0.189
$[14,12]_q \otimes [16,14]_q$	$0.750:\ P_{ew}=1.0\cdot10^{-2}$	0.150
Regular-$(3,12)$ LDPC	$0.750:\ P_{ew}=2.9\cdot10^{-2}$	0.135