Coset graphs and LDPC codes

Josef Lauri[1] and Cen J Tjhai[2]

[1] University of Malta || [2]University of Plymouth

July 21, 2010

Software used

Software used

- GAP: Groups, Algorithms, Programming
  System for doing Computational Discrete Algebra

Software used

- GAP: Groups, Algorithms, Programming
  System for doing Computational Discrete Algebra
- GRAPE: GRaph Algorithms using PErmutation groups
  A GAP package

Software used

- GAP: Groups, Algorithms, Programming
  System for doing Computational Discrete Algebra
- GRAPE: GRaph Algorithms using PErmutation groups
  A GAP package
- guava: A GAP package for error-correcting codes

Software used

- GAP: Groups, Algorithms, Programming
  System for doing Computational Discrete Algebra
- GRAPE: GRaph Algorithms using PErmutation groups
  A GAP package
- guava: A GAP package for error-correcting codes
- Software developed by Cen and the group he works with in Plymouth for simulating the codes' behaviour on the BER

Linear binary error-correcting codes

A linear binary error-correcting code is a subspace C of $\mathbb{B}^n$. The dimension of C is denoted by k.

Linear binary error-correcting codes

A linear binary error-correcting code is a subspace C of $\mathbb{B}^n$. The dimension of C is denoted by k.

The code C can be defined in terms of a generator matrix whose k rows are simply a basis of C.

Linear binary error-correcting codes

A linear binary error-correcting code is a subspace C of $\mathbb{B}^n$. The dimension of C is denoted by k.

The code C can be defined in terms of a generator matrix whose k rows are simply a basis of C.

But in general more information on the error-correcting capabilities of C can be obtained by considering it as the kernel of a linear transformation.

Linear binary error-correcting codes

A linear binary error-correcting code is a subspace C of $\mathbb{B}^n$. The dimension of C is denoted by k.

The code C can be defined in terms of a generator matrix whose k rows are simply a basis of C.

But in general more information on the error-correcting capabilities of C can be obtained by considering it as the kernel of a linear transformation.

So we let H be a $(n-k) \times n$ matrix and define C to be all those $1 \times n$ vectors c such that

$$cH^T = 0.$$

The matrix H is called the *parity check matrix* of the code.

Linear binary error-correcting codes

A linear binary error-correcting code is a subspace C of $\mathbb{B}^n$. The dimension of C is denoted by k.

The code C can be defined in terms of a generator matrix whose k rows are simply a basis of C.

But in general more information on the error-correcting capabilities of C can be obtained by considering it as the kernel of a linear transformation.

So we let H be a $(n-k) \times n$ matrix and define C to be all those $1 \times n$ vectors c such that

$$cH^T = 0.$$

The matrix H is called the *parity check matrix* of the code.

Usually error-corrections makes use of the *syndrome*: Let c' be a received codeword. The syndrome is defined to be

$$s = c'H^T.$$

If $s = 0$ then c' is taken to be correct. Otherwise a look-up table is used to determine from s the codeword which is the nearest to c'.

Low Density Parity Check codes

Gallager, in 1960, was the first to find that good codes can be constructed if the check matrix is sparse. Here we consider *regular* LDPC codes for which every column has a constant number of 1's (usually 3) and every row has a constant number of 1's (usually 4, 5 or 6).

Low Density Parity Check codes

Gallager, in 1960, was the first to find that good codes can be constructed if the check matrix is sparse. Here we consider *regular* LDPC codes for which every column has a constant number of 1's (usually 3) and every row has a constant number of 1's (usually 4, 5 or 6).

Instead of syndrome decoding iterative methods over graphs are used.

Tanner graphs

Represent the check matrix H by a bipartite graph G with bipartition $V_b \cup V_c$. Each vertex in V_b corresponds to a column (bit) of H and each vertex in V_c corresponds to a row (check equation) of H. Two vertices are adjacent iff there is a one in the intersection of the corresponding row and column.

Tanner graphs

Represent the check matrix H by a bipartite graph G with bipartition $V_b \cup V_c$. Each vertex in V_b corresponds to a column (bit) of H and each vertex in V_c corresponds to a row (check equation) of H. Two vertices are adjacent iff there is a one in the intersection of the corresponding row and column.

Example of a non-regular check matrix and its Tanner graph:

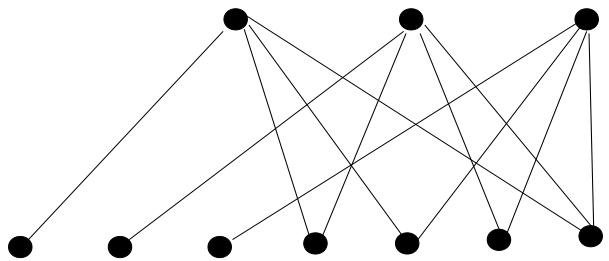
$$H = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Tanner graphs

Represent the check matrix H by a bipartite graph G with bipartition $V_b \cup V_c$. Each vertex in V_b corresponds to a column (bit) of H and each vertex in V_c corresponds to a row (check equation) of H. Two vertices are adjacent iff there is a one in the intersection of the corresponding row and column.

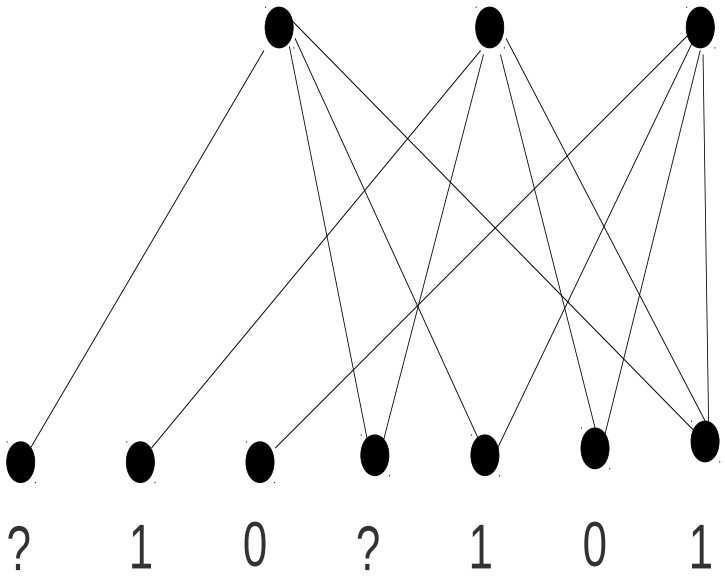
Example of a non-regular check matrix and its Tanner graph:

$$H = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

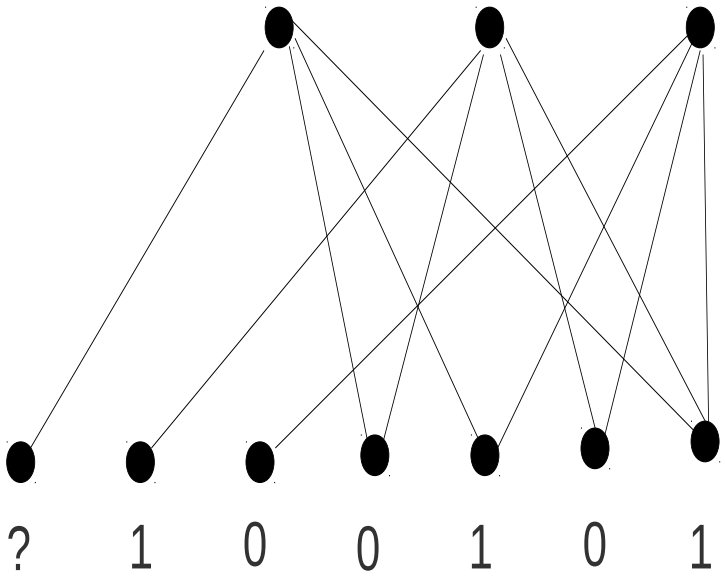


Iterative correction of erasures

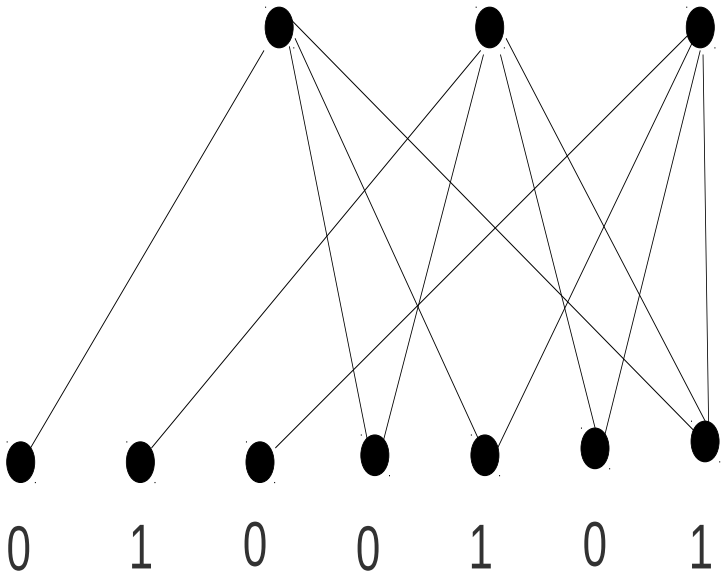
Iterative correction of erasures



Iterative correction of erasures(2)

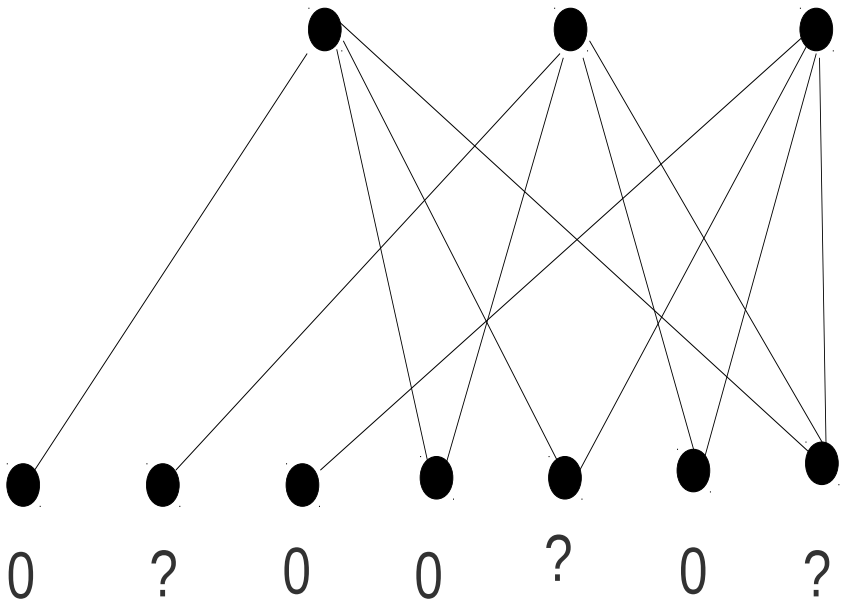


Iterative correction of erasures(3)



Stopping sets!

Stopping sets!



Coset graphs

Let $(\Gamma, \mathcal{H}, \mathcal{K})$ be a group with two subgroups such that $\Gamma = \langle \mathcal{H}, \mathcal{K} \rangle$. Define the graph $\mathrm{Cos}(\Gamma, \mathcal{H}, \mathcal{K})$ as follows: its vertices are the right cosets of $\mathcal{H}$ and $\mathcal{K}$, and two cosets $\mathcal{H}x, \mathcal{K}y$ are adjacent if and only if their intersection is non-empty.

Coset graphs

Let $(\Gamma, \mathcal{H}, \mathcal{K})$ be a group with two subgroups such that $\Gamma = \langle \mathcal{H}, \mathcal{K} \rangle$. Define the graph $\mathrm{Cos}(\Gamma, \mathcal{H}, \mathcal{K})$ as follows: its vertices are the right cosets of $\mathcal{H}$ and $\mathcal{K}$, and two cosets $\mathcal{H}x, \mathcal{K}y$ are adjacent if and only if their intersection is non-empty.

Theorem

The graph $Cos(\Gamma, \mathcal{H}, \mathcal{K})$ is a connected edge-transitive bipartite graph with vertex degrees $|\mathcal{H}|/|\mathcal{H} \cap \mathcal{K}|$ and $|\mathcal{K}|/|\mathcal{H} \cap \mathcal{K}|$ and with the two sets of cosets of $\mathcal{H}$ and $\mathcal{K}$ being the bipartition of $Cos(\Gamma, \mathcal{H}, \mathcal{K})$.

Conversely, if G is a graph on which the group Γ acts edge-transitively but not vertex-transitively, then G is isomorphic to $Cos(\Gamma, \mathcal{H}, \mathcal{K})$ where $\mathcal{H}$ and $\mathcal{K}$ are the stabilisers of two adjacent vertices.

A non-vertex-transitive edge-regular graph

A non-vertex-transitive edge-regular graph

A graph of order 465 whose automorphism group acts regularly on its edges and whose girth is 8. The group Γ was

$$\langle a, b, c | a^5 = b^3 = c^{31} = 1, ba = abc, ca = ac^2, cb = bc^{25} \rangle$$

with $\mathcal{H} = \langle a \rangle$ and $\mathcal{K} = \langle b \rangle$.

A non-vertex-transitive edge-regular graph

A graph of order 465 whose automorphism group acts regularly on its edges and whose girth is 8. The group Γ was

$$\langle a, b, c | a^5 = b^3 = c^{31} = 1, ba = abc, ca = ac^2, cb = bc^{25} \rangle$$

with $\mathcal{H} = \langle a \rangle$ and $\mathcal{K} = \langle b \rangle$.
At about the same time, Tanner, Sridhara and Fuja investigated quasi-cyclic LDPC codes. It turns out that the Tanner graphs of these codes are the coset graphs of the group $\Gamma(p, q, r)$ where p, q, r are primes with $r = 1 \bmod pq$ and such that $\Gamma(p, q, r)$ is

$$\langle a, b, c | a^p = b^q = c^r = 1, ba = abc, ca = ac^s, cb = bc^t \rangle$$

where s^p and t^q are equal to 1 mod r. Therefore the above graph is the coset graph of $\Gamma(3, 5, 31)$.

Girth

Girth

Tanner et al's QC LDPC codes have girth at most 12.

Girth

Tanner et al's QC LDPC codes have girth at most 12.

In general, there is no upper bound for the girth of coset graphs.

Girth

Tanner et al's QC LDPC codes have girth at most 12.

In general, there is no upper bound for the girth of coset graphs.

For example, if

$$\Gamma = \langle (1\ 2\ 3\ 4)(6\ 7\ 8\ 9), (4\ 5\ 6) \rangle$$

and $\mathcal{H} = \langle (1\ 2\ 3\ 4)(6\ 7\ 8\ 9) \rangle$, $\mathcal{K} = \langle (4\ 5\ 6) \rangle$, then the girth of $\mathrm{Cos}(\Gamma, \mathcal{H}, \mathcal{K})$ is 16.

Girth

Tanner et al's QC LDPC codes have girth at most 12.

In general, there is no upper bound for the girth of coset graphs.

For example, if

$$\Gamma = \langle (1\ 2\ 3\ 4)(6\ 7\ 8\ 9), (4\ 5\ 6) \rangle$$

and $\mathcal{H} = \langle (1\ 2\ 3\ 4)(6\ 7\ 8\ 9) \rangle$, $\mathcal{K} = \langle (4\ 5\ 6) \rangle$, then the girth of $\mathrm{Cos}(\Gamma, \mathcal{H}, \mathcal{K})$ is 16.

The girth g of $\mathrm{Cos}(\Gamma, \mathcal{H}, \mathcal{K})$ can be defined as the length of a shortest reduced word $k_1 h_2 k_3 \ldots k_{g-1} h_g$ in Γ which is equal to 1, where, by a *reduced word* we mean an alternating product of elements of $\mathcal{H} - \mathcal{H} \cap \mathcal{H}$ and $\mathcal{K} - \mathcal{H} \cap \mathcal{K}$ and by length of a word we mean the number of such elements.

Some small LDPC codes constructed from coset graphs

Table: Parameters of small LDPC codes

Name	Generator a	Generator b	Girth	s_{min}
Small-Cage $[9, 4, 4]$	(1 2 3)	(1 4)(2 5)(3 6)	8	4
Gen-Quad $[16, 9, 4]$	(1 2 3 4)	(1 5)(2 6) (3 7)(4 8)	8	4
Gray $[27, 8, 8]$	(1 2 3)	(1 4 7)(2 5 8) (3 6 9)	8	8

Simulations of small LDPC codes over the Binary Erasure Channel (BEC)

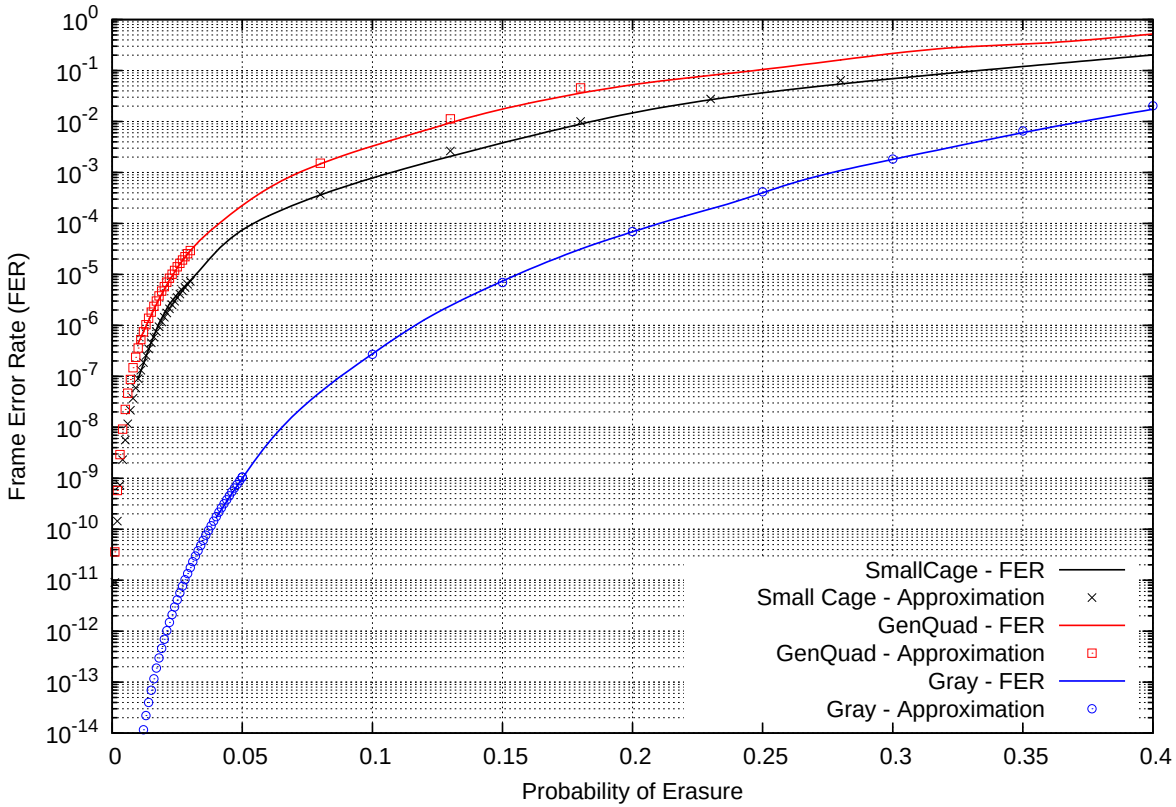


Figure: Frame Error Probability of Small LDPC Codes over the BEC

Some $(3, 4)$ LDPC codes constructed from coset graphs

Table: Parameters of $(3, 4)$ LDPC codes in Figure 2

Name	Generator a	Generator b	Girth	s_{min}	S_i
$[40, 15, 8]$	(1 2 3 4 5) (8 9 10) (11 12 13)	(1 2 3 4 5) (8 9 10 11)	8	8	$S_8 = 45$ $S_8 = 54$ $S_{12} = 1200$ $S_{13} = 1440$ $S_{14} = 4260$
$[720, 270, 8]$	(1 2 3)(4 5 6)	(3 7 8 9) (4 10)	8	8	$S_8 = 270$ $S_{12} = 2160$ $S_{14} = 2160$ $S_{15} = 2160$ $S_{16} = 50355$
$[360, 117, 24]$	(1 2 3)(4 5 6)	(5 6 7 8)(3 9) (2 10)(1 11)	12	24	$S_{24} = 1755$
$[720, 216, 24]$	(1 2 3)(4 5 6)	(3 7 8 9)	12	24	$S_{24} = 450$
$[2520, 701, 24]$	(1 2 3)(4 5 6)	(1 7) (2 8 9 10)	12	24	$S_{24} = 105$
$[2160, 591, \geq 32]$	(1 2 3)(4 5 6)	(5 6 7 8)(3 9) (2 10)(1 11) (4 12)	12	≤ 48	

Simulations of $(3, 4)$ codes over the BEC

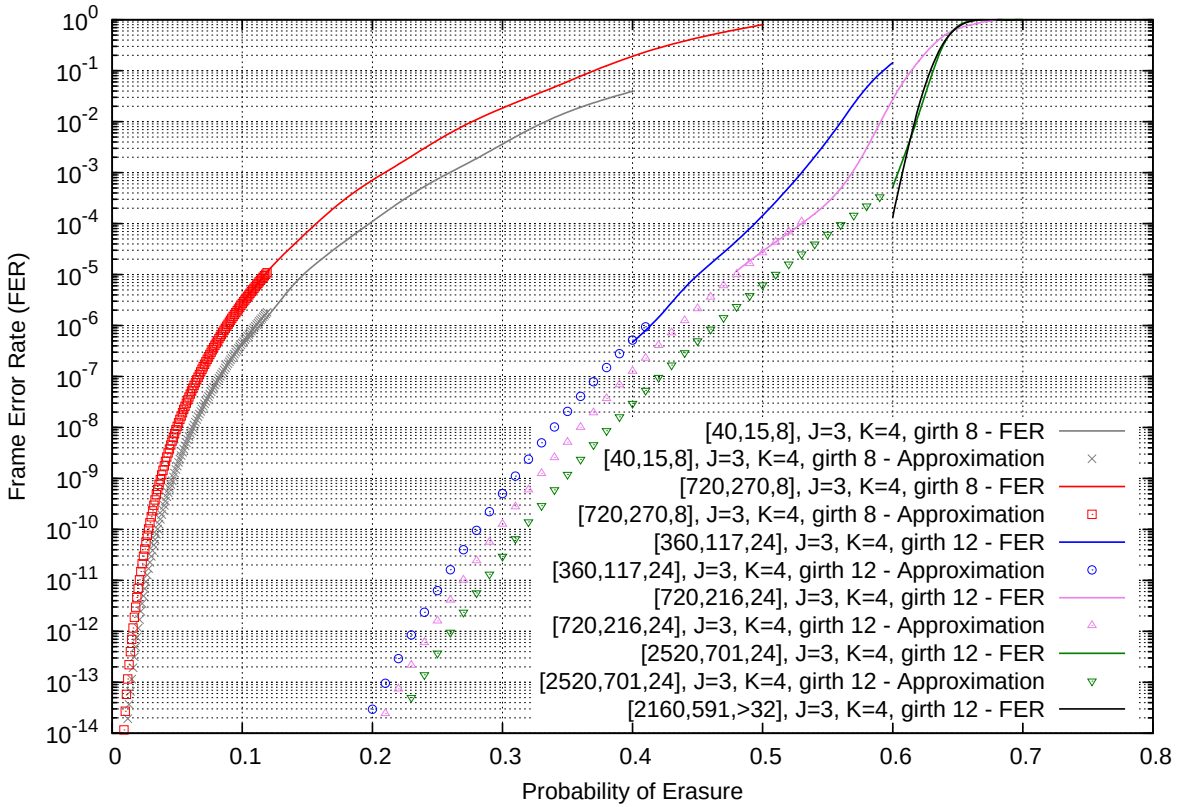


Figure: Frame Error Probability of $(3, 4)$ coset graph LDPC Codes over the BEC

Some $(3, 5)$ LDPC codes constructed from coset graphs

Table: Parameters of $(3, 5)$ LDPC codes in Figure 3

Name	Generator a	Generator b	Girth	s_{min}	S_i
$[155, 64, 20]$	$\langle a, b, c : a^3 = b^5 = c^{31}$ $ba = abc,\ ca = ac^{25},$ $cb = bc^2\rangle$		8	18	$S_{18} = 465$ $S_{19} = 2015$ $S_{20} = 9548$ $S_{21} = 23715$ $S_{22} = 106175$
$[755, 334, 14]$	$\langle a, b, c : a^3 = b^5 = c^{151}$ $ba = abc,\ ca = ac^{32},$ $cb = bc^8\rangle$		10	14	$S_{14} = 755$ $S_{18} = 755$ $S_{20} = 3020$ $S_{22} = 9815$ $S_{24} = 30200$
$[840, 342, 28]$	$(1\ 2\ 3\ 4\ 5)$	$(2\ 6\ 7)(1\ 4\ 3)$	8	28	$S_{28} = 120$
$[905, 364, 24]$	$\langle a, b, c : a^3 = b^5 = c^{181}$ $ba = abc,\ ca = ac^{48},$ $cb = bc^{42}\rangle$		12	24	$S_{24} = 905$
$[6720, 2695, 24]$	$(1\ 2\ 3)(4\ 5\ 6)$	$(3\ 4\ 5\ 7\ 8)$	10	24	$S_{24} = 1120$

Simulations of $(3, 5)$ codes over the Binary Erasure Channel (BEC)

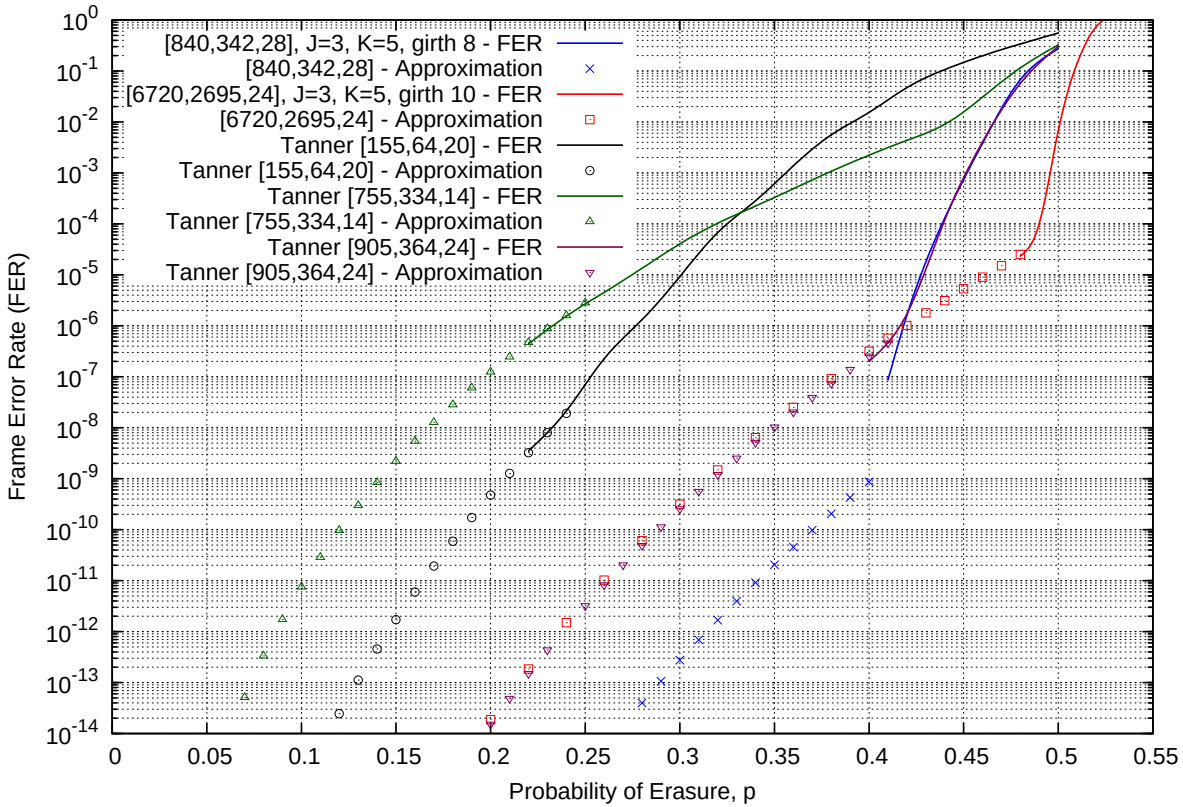


Figure: Frame Error Probability of $(3, 5)$ coset graph LDPC Codes over the BEC

Is this a useful generalisation of Tanner's QC codes?

Is this a useful generalisation of Tanner's QC codes?

- Loses the convenient quasi-cyclic structure of the check matrix,

Is this a useful generalisation of Tanner's QC codes?

- Loses the convenient quasi-cyclic structure of the check matrix, but:
- The QC structure of a (J, K) LDPC code places an upper bound of $(J+1)!$ on the minimum Hamming distance. So our examples $[840, 342, 28]$ and $[2160, 591, \geq 32]$ would have had minimum Hamming distance at most 24.

Is this a useful generalisation of Tanner's QC codes?

- Loses the convenient quasi-cyclic structure of the check matrix, but:
- The QC structure of a (J, K) LDPC code places an upper bound of $(J+1)!$ on the minimum Hamming distance. So our examples $[840, 342, 28]$ and $[2160, 591, \geq 32]$ would have had minimum Hamming distance at most 24.
- Conceptually quite simple;

Is this a useful generalisation of Tanner's QC codes?

- Loses the convenient quasi-cyclic structure of the check matrix, but:
- The QC structure of a (J, K) LDPC code places an upper bound of $(J+1)!$ on the minimum Hamming distance. So our examples $[840, 342, 28]$ and $[2160, 591, \geq 32]$ would have had minimum Hamming distance at most 24.
- Conceptually quite simple;
- Provides a compact algebraic description of the code which might make coding and decoding more efficient;

Is this a useful generalisation of Tanner's QC codes?

- Loses the convenient quasi-cyclic structure of the check matrix, but:
- The QC structure of a (J, K) LDPC code places an upper bound of $(J+1)!$ on the minimum Hamming distance. So our examples $[840, 342, 28]$ and $[2160, 591, \geq 32]$ would have had minimum Hamming distance at most 24.
- Conceptually quite simple;
- Provides a compact algebraic description of the code which might make coding and decoding more efficient;
- Finding codes with large girth and not so large length places the problem within an important area of graph theory: cages (here bi-regular edge-transitive graphs with given girth and minimal size);

Is this a useful generalisation of Tanner's QC codes?

- Loses the convenient quasi-cyclic structure of the check matrix, but:
- The QC structure of a (J, K) LDPC code places an upper bound of $(J+1)!$ on the minimum Hamming distance. So our examples $[840, 342, 28]$ and $[2160, 591, \geq 32]$ would have had minimum Hamming distance at most 24.
- Conceptually quite simple;
- Provides a compact algebraic description of the code which might make coding and decoding more efficient;
- Finding codes with large girth and not so large length places the problem within an important area of graph theory: cages (here bi-regular edge-transitive graphs with given girth and minimal size);
- Also, the use of coset graphs places the search for good codes squarely into an important problem in group theory: amalgams for finite groups (as developed by Djoković and Miller and Goldschmidt). Maybe this powerful algebraic machinery can help to obtain not only codes with large girth but, more importantly, without small stopping sets

THANK YOU!