Structured Low-Density Parity-Check Codes: Algebraic Constructions

Shu Lin
Department of Electrical and Computer Engineering
University of California, Davis
Davis, California 95616
Email:shulin@ece.ucdavis.edu

I. Introduction

- LDPC codes were discovered by Gallager in 1962 and rediscovered in late 1990's. These codes form another class of Shannon limit approaching codes, besides turbo codes.
- Well designed LDPC codes perform amazingly well and close to the Shannon limit with iterative decoding using the sum-product-algorithm (SPA). Long LDPC codes have been constructed and they perform only a few tenths (or hundredths) of a dB from the Shannon limit.
- LDPC codes have some advantages over the turbo codes and have a great potential for error control in digital communication and storage systems.

Definitions and Basic Concepts

- A binary LDPC code **C** is given by the null space of a sparse matrix H, called the parity-check matrix. If H has constant column weight γ and constant row weight ρ, it is said to be (γ, ρ)-regular and the code **C** generated by it is called a (γ, ρ)-regular LDPC code. Otherwise, **H** is said to be irregular and **C** is called an irregular LDPC code.
- Suppose that **H** satisfies the constraint that no two rows (or two columns) have more than one 1-component in common. This constraint is called the *row-column* (*RC)-constraint*. The RC-constraint on **H** ensures that the Tanner graph of the parity-check matrix is free of cycles of length 4 and hence has a girth of at least 6.

Definitions and Basic Concepts

- The null space of a sparse parity-check matrix **H** that satisfies the RC-constraint gives an LDPC code whose Tanner graph has a girth of at least 6. The girth of the Tanner graph of an LDPC code is simply called the girth of the code.

- Let $\gamma_{\min}$ be the minimum column weight of **H**. If **H** satisfies RC-constraint, the LDPC code generated by **H** has a minimum distance of at least $\gamma_{\min} + 1$.

Classifications of Construction of LDPC Codes

- Constructions of LDPC codes can be classified into two general categories: random and algebraic constructions.

- Random construction is to construct codes using computer search based on a set of design rules (or guidelines) and required structures of their Tanner graphs, such as the degree distributions of the variable and check nodes. Random LDPC codes in general do not have sufficient structures to allow simple encoding. However, they do perform well in the waterfall region.

Classifications of Construction of LDPC Codes

- Algebraic construction is to construct structured LDPC with algebraic and combinatorial methods. Structured LDPC codes in general have encoding (or decoding) advantage over the random codes in terms of hardware implementation.

- Well designed structured codes can perform just as well as random codes in terms of bit-error performance, frame-error performance and error floor, collectively.

Cyclic and Quasi-Cyclic LDPC Codes

- If a sparse matrix **H** consists of a single circulant or a column of circulants, then the null space of **H** gives a cyclic LDPC code. Then the code is uniquely specified by a generator polynomial and its encoding can be implemented with a simple feedback shift-register.

- If a sparse matrix **H** consists of an array of circulants of the same size, then the null space of **H** gives a QC-LDPC code whose encoding can also be encoded with simple shift-registers.

- Cyclic and QC-LDPC codes have encoding advantage over all the other types of LDPC codes.

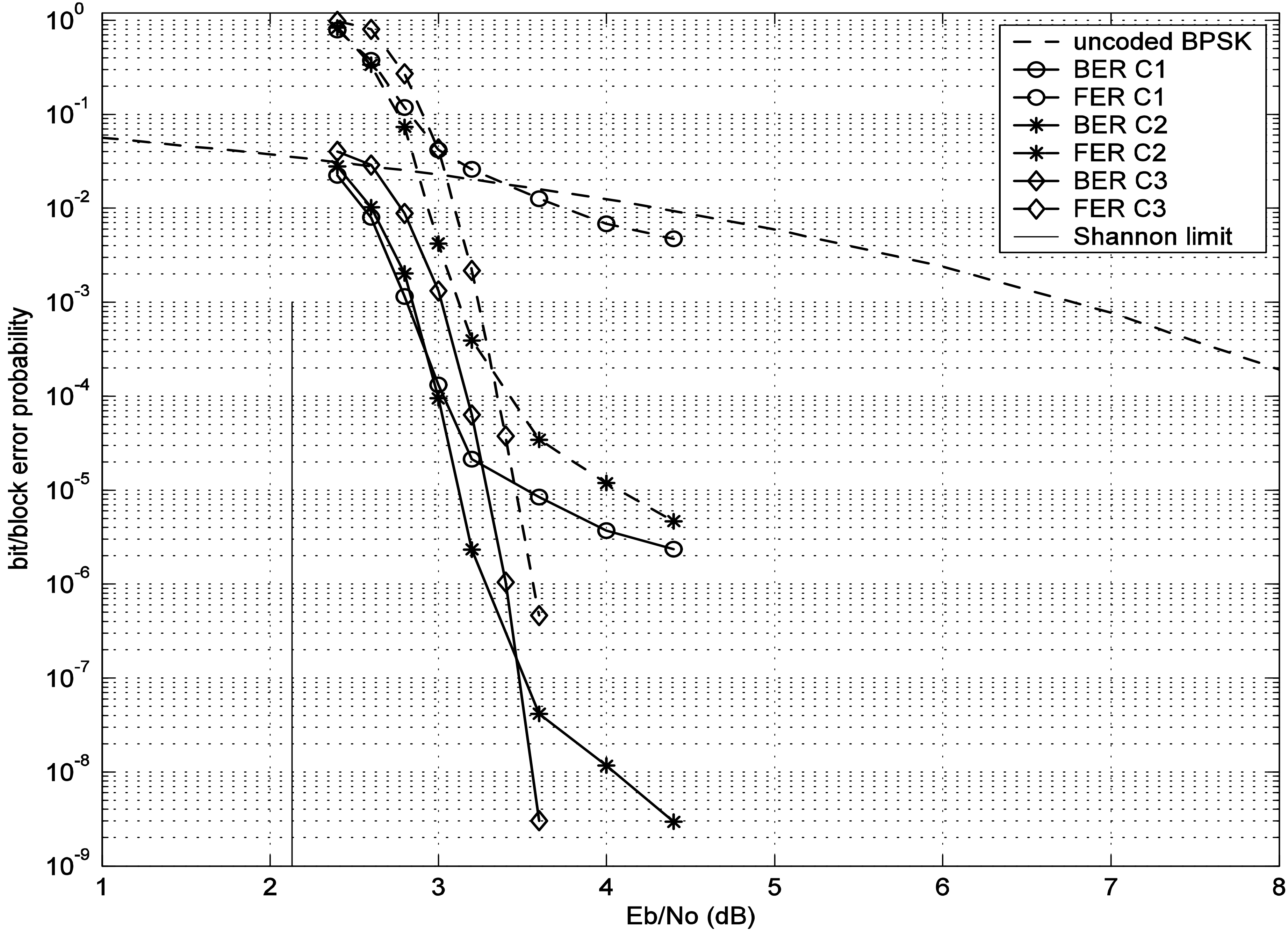
Performance of LDPC Codes with Iterative Decoding

- The error performance of an LDPC codes with iterative decoding using the SPA depends on a number of code structures.
- The most important structures are: girth, cycle distributions, cycle structure of the code graph and the minimum distance of the code.
- How does the error performance of an LDPC code depend on these structural properties, collectively, is basically unknown.
- In many applications in communication and digital storage systems, a major concern is the error-floor. It is desired to design (or construct) LDPC codes either with no error-floor or a very low error-floor.

Performance of LDPC Codes with Iterative Decoding (Cont'd)

- Based on our many experimental results, the error floor very much depends on the column weight of the parity-check matrix. The error floor can be pushed down by increasing the column weight. However, as the error floor being pushed down by increasing the column weight, the waterfall performance of the code is pushed away from the Shannon limit. Figure 1 displays this phenomenon. For given code length and rate, a proper choice of the column weight is needed to achieve a low error-floor while maintain a close to Shannon limit waterfall performance.

Figure 1. error performance of the 3 irregular LDPC(4032,3264) codes



II. Major Algebraic and Combinatorial Methods --- Construction based on finite geometries

- Y. Kou, S. Lin, and M. Fossorier," Low-density parity-check codes based on finite geometries: a rediscovery and new results," *IEEE Trans. on Inform. Theory*, vol.47, no.7, pp. 2711-2736, Nov. 2001.
- H. Tang, J. Xu, Y. Kou, S. Lin, and K. Abdel-Ghaffar,"On algebraic construction of Gallager and circulant low-density parity-check codes," *IEEE Trans. on Inform. Theory*, vol. 50, no. 6, pp. 1269-1279, June 2004
- H. Tang, J. Xu, S. Lin, K. Abdel-Ghaffar,"Codes on finite geometries," accepted for publication in *IEEE Trans. on Inform. Theory*, 2004.
- J. Xu, L. Chen, I. Djurdjevic, S. Lin, and K. Abdel-Ghaffar,"Construction of regular and irregular LDPC codes: geometry decomposition and masking," submitted to *IEEE Trans. on Inform. Theory*, 2004.

Construction based on combinatorial designs

- B. Ammar, B. Honary, Y. Kou, J. Xu, and S. Lin"Construction of low-density parity-check codes based on balanced incomplete block designs," *IEEE Trans. Inform. Theory*, vol. 50, no.6, pp. 1157-1268, June 2004.

- B. Vasic and O. Milenkovic,"Combinatorial construction of low-density parity-check codes for iterative decoding," *IEEE Trans. on Inform. Theory*, vol. 50, no. 6, pp. 1156-1176, June 2004.

Construction based on Reed-Solomon (RS) codes

- I. Djurdjevic, J. Xu, K. Abdel-Ghaffar, and S. Lin,"Construction of low-density parity-check codes based on Reed-Solomon codes with two information symbols," *IEEE Communications Letters*, vol. 7, no. 7, pp. 317-319, July 2003.

- L. Chen, I. Djurdjevic, J. Xu, S. Lin, and K. Abdel-Ghaffar,"Construction of quasi-cyclic LDPC codes based on the minimum weight codewords of Reed-Solomon Codes," *Proc. 2004 IEEE Int. Symp. Inform. Theory*, p. 239, Chicago, IL, June 27-July 2, 2004.

Construction Based on Circulant Decomposition

- L. Chen, J. Xu, I. Djurdjevic, and S. Lin,"Near Shannon limit quasi-cyclic low-density parity-check codes," *IEEE Trans. Communications*, vol. 52, no. 7, July 2004.

- S. Lin, L. Chen, I. Djurdjevic, J. Xu,"Near Shannon limit quasi-cyclic low-density parity-check codes," *Proc. IEEE GlobeCom"2003*, pp. 2030-2035, San Francisco, CA, Dec. 2003.

Construction based on Superposition

- S. Lin, J. Xu, I. Djurdjevic, and H. Tang,"Hubrid construction of LDPC codes," *Proc. 40th Annual Allerton Conf. on Commu. Control, and Computing*, pp. 1149-1158, Monticello, IL, October 2-4, 2002.

- J. Xu, L. Chen, I. Djurdjevic, L. -Q. Zeng,"Construction of low-density parity-check codes by superposition," accepted for publication in *IEEE Trnas. Commun.*, 2004.

Construction Based on Graphs

- J. Rosenthal and P. O. Vontobel," Construction of LDPC codes using Ramanujan graphs and ideas from Margulis," *Proc. the 38th Allerton Conf. on Commun., Control and Computing*, pp. 248-257, Monticello, IL, Oct. 4-6, 2001.

III. Construction of Structured LDPC Codes Based on RS Codes With Two Information Symbols

- Consider the Galois field GF(q) where q is a power of prime p, i.e., $q = p^m$. Let α be a primitive element in GF(q). Then
  $0 = \alpha^{-\infty}, \; 1 = \alpha^0, \; \alpha, \; \ldots, \; \alpha^{(q-2)}$ give all the q elements of GF(q).
- For $-\infty \le i < q-1$, represent α^i by a unit q-tuple over GF(2),

$$\mathbf{z}(\alpha^i) = (z_{-\infty}, \; z_0, \; \ldots, \; z_{q-2}) \, ,$$

  whose components correspond to the q elements of GF(q), where $z_i = 1$ and all the other components are equal to 0. This unit q-tuple is called the location vector of α^i. It is clear that the 1-components of the location vectors of two elements in GF(q) are at two different locations.

III. Construction of Structured LDPC Codes Based on RS Codes With Two Information Symbols (Cont'd)

- Form a $q \times q$ square matrix **A** over GF(2) with the location vectors of the q elements of GF(q) as the rows. Then **A** is a $q \times q$ permutation matrix with column and row weights equal to 1.

- Consider a (q, 2, q-1) RS code $\mathbf{C}_b$ over GF(q) obtained by adding an overall parity-check symbol to each codeword of the (q-1, 2, q-2) cyclic RS code. $\mathbf{C}_b$ has q^2 codewords, one codeword with weight 0, q(q-1) codewords with minimum weight q-1, and q-1 codewords with weight q. Two codewords differ in at least q-1 positions, in other words, they have at most one position with the same code symbol.

III. Construction of Structured LDPC Codes Based on RS Codes With Two Information Symbols (Cont'd)

- Let $\mathbf{C}_1$ be the $(q, 1, q)$ linear subcode of $\mathbf{C}_b$. Partition $\mathbf{C}_b$ into q cosets with respect to $\mathbf{C}_1$. Denote these cosets with $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_q$, each consisting of q codewords. Two codewords in the same coset $\mathbf{C}_i$ differ at every position.
- For $1 \leq i \leq q$, form a $q \times q$ matrix $\mathbf{G}_i$ over GF(q) with the codewords in $\mathbf{C}_i$ as rows.

$$\mathbf{G}_i = \begin{bmatrix} v_{1,\infty} & v_{1,0} & \cdots & v_{1,q-2} \\ v_{2,\infty} & v_{2,0} & \cdots & v_{2,q-2} \\ \vdots & \vdots & \cdots & \vdots \\ v_{q,\infty} & v_{q,0} & \cdots & v_{q,q-2} \end{bmatrix} \qquad (1)$$

Any two rows of $\mathbf{G}_i$ differ at every position. The q components of each column are all different and they form all the q elements of GF(q).

III. Construction of Structured LDPC Codes Based on RS Codes With Two Information Symbols (Cont'd)

- Replacing each entry of $\mathbf{G}_i$ by its location vector, we obtain a $q \times q^2$ matrix $\mathbf{B}_i$ over GF(2) which consists of q submatrices,

$$\mathbf{B}_i = \left[\mathbf{A}_{i,\infty} \ \mathbf{A}_{i,0} \ \cdots \ \mathbf{A}_{i,q\text{-}2} \right],$$

  where each submatrix $\mathbf{A}_{\mathrm{i,j}}$ is a $q \times q$ permutation matrix.

- Form a $q \times q$ array of $q \times q$ permutation matrices as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_{\mathrm{q}} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{1,\infty} & \mathbf{A}_{1,0} & \cdots & \mathbf{A}_{1,q-2} \\ \mathbf{A}_{2,\infty} & \mathbf{A}_{2,0} & \cdots & \mathbf{A}_{2,q-2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{A}_{q,\infty} & \mathbf{A}_{q,0} & \cdots & \mathbf{A}_{q,q-2} \end{bmatrix} \qquad (2)$$

III. Construction of Structured LDPC Codes Based on RS Codes With Two Information Symbols (Cont'd)

- **H** is a $q^2 \times q^2$ matrix over GF(2) with both column and row weights q. Since the rows of **H** correspond to the codewords in the RS code $\mathbf{C}_b$, no two rows have more than one 1-component in common, which also implies that no two columns of **H** have more than one 1-component in common. Therefore, **H** satisfies the RC-constraint and hence its Tanner graph has a girth of at least 6.
- For $1 \le \gamma,\ \rho \le q$, let $\mathbf{H}(\gamma,\rho)$ be a $\gamma \times \rho$ subarray of **H**. $\mathbf{H}(\gamma,\rho)$ is a $\gamma q \times \rho q$ regular matrix with column and row weights γ and ρ and satisfies the RC-constraint. The null space of $\mathbf{H}(\gamma,\rho)$ gives a regular LDPC code $\mathbf{C}_{rs}$ of length $n = \rho q$ with rate at least $(\rho$-$\gamma)/\gamma$ and girth at least 6.
- The minimum distance of $\mathbf{C}_{rs}$ is at least γ+1 for odd γ and γ+2 for even γ.

III. Construction of Structured LDPC Codes Based on RS Codes With Two Information Symbols (Cont'd)

- The above construction gives a family of structured regular LDPC codes with various lengths, rates and minimum distances. The girths of the codes in this family are at least 6. We call the codes in this family, RS-based LDPC codes.

Example I

- Suppose we choose the extended (32,2,31) RS code over GF(2^5) for code construction. Based on this code, we can construct a 32x32 array of 32x32 permutation matrices

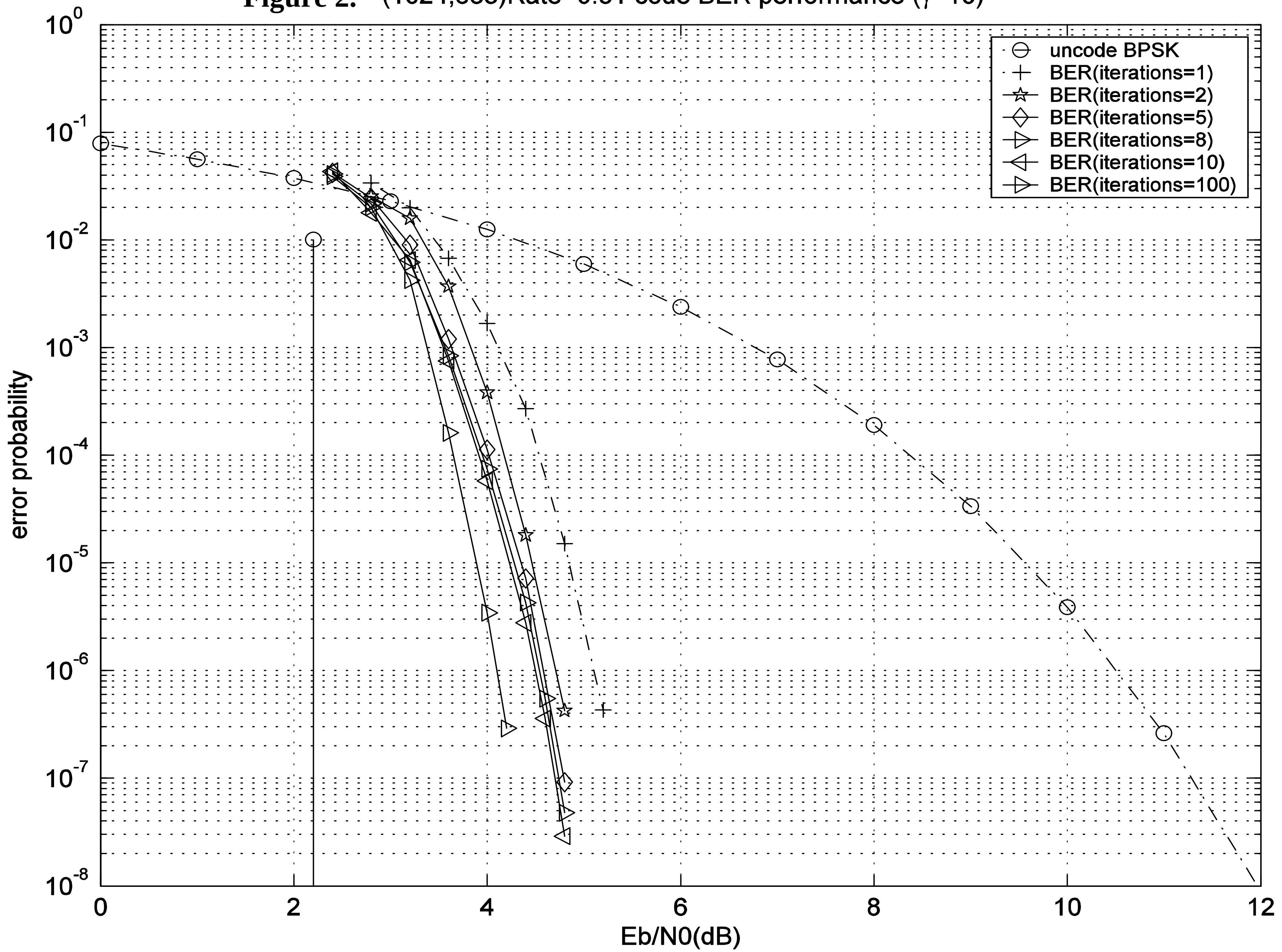
$$\mathbf{H} = \begin{bmatrix} \mathbf{A}_{1,\infty} & \mathbf{A}_{1,0} & \cdots & \mathbf{A}_{1,30} \\ \mathbf{A}_{2,\infty} & \mathbf{A}_{2,0} & \cdots & \mathbf{A}_{2,30} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{A}_{32,\infty} & \mathbf{A}_{32,0} & \cdots & \mathbf{A}_{32,30} \end{bmatrix}$$

- Set γ=10 and ρ=32. Let **H**(10,32) be the 10x32 subarray that consists of the first 10 rows of permutation matrices of **H**. **H**(10,32) is a 320x1024 matrix over GF(2) with column and row weights 10 and 32, respectively. The null space of **H**(10,32) gives a (1024,833) LDPC code with rate 0.8134 and minimum distance at least 12.

Example I

- Assume BPSK transmission over an AWGN channel with iterative decoding using the SPA. Set the maximum number decoding iterations to 100. The performance of the code is shown in Figure 1. At the BER of 10^{-6}, it achieves more than 6 dB coding gain over the uncoded BPSK and performs only 1.9 dB from the Shannon limit. The decoding converges very fast. At the BER of 10^{-6}, the performance gap between 5 and 100 iterations is within 0.4 dB.

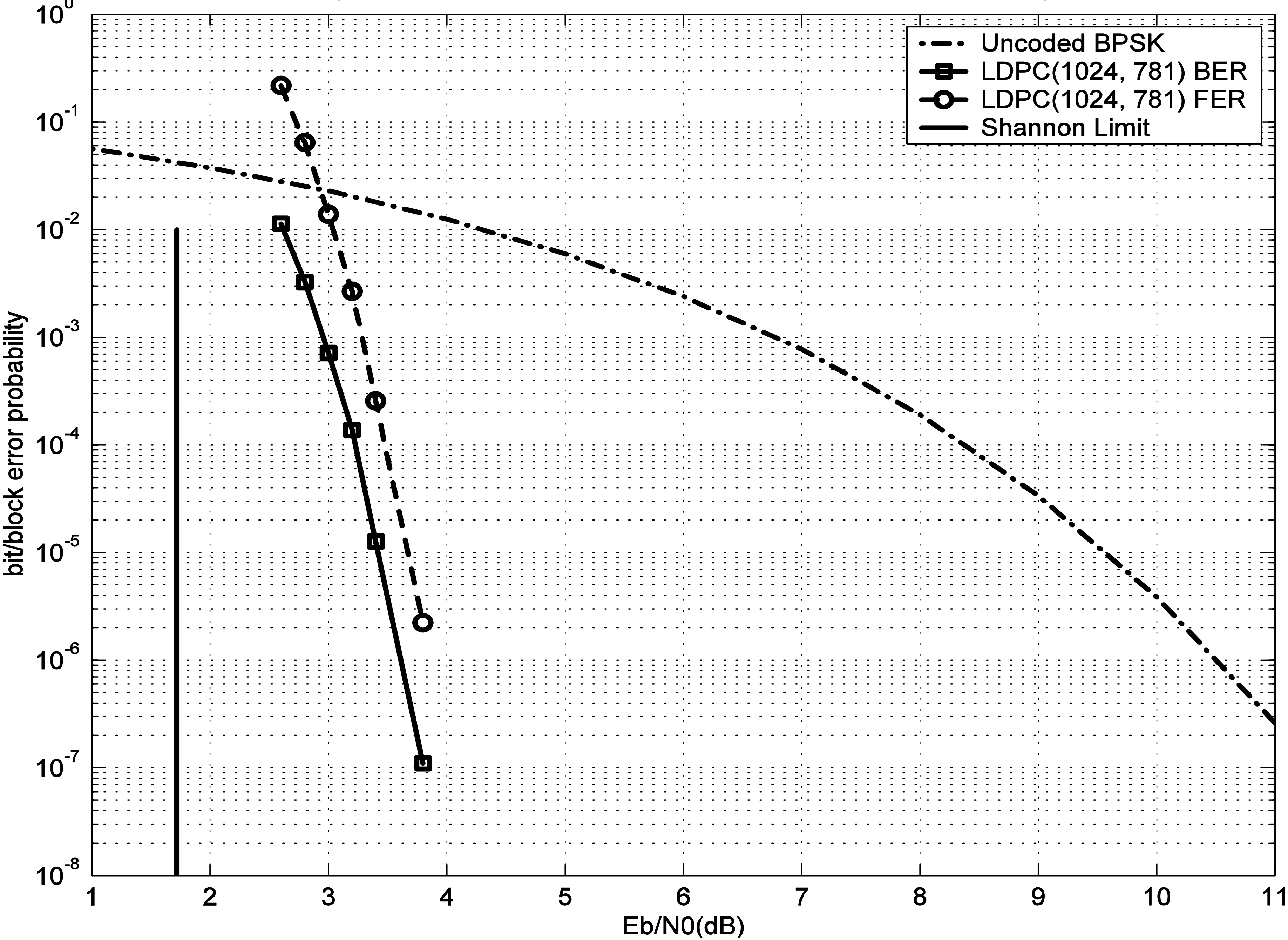
Figure 2. (1024,833)Rate=0.81 code BER performance (γ=10)



Example II

- Again we use the extended (32,2,31) RS code over GF(2^5) for code construction. Set $\gamma = \rho$=32. Then **H**(32,32) is the entire array of **H** given in Example I. **H**(32,32) is a 1024x1024 matrix over GF(2) with both column and row weights 32. The null space of **H**(32,32) gives a (1024,781) LDPC code with rate 0.7626 with minimum distance exactly 34 (a large minimum distance for an LDPC code of length 1024). The performance of this code is shown in Figure 2. At the BER of 10^{-6}, it achieves almost 7 dB coding gain over the uncoded BPSK and performs only 1.9 dB from the Shannon limit. For such a short LDPC code, the code performs very well. Since it has large minimum distance, it is not expected to have an error floor.

Figure 3. LDPC(1024, 781) Code Performance (SPA Decoding)



- Table 1 gives a list LDPC codes of length 1024 constructed based on the (32,2,31) extended RS code over GF(2^5).

Table 1 LDPC codes of length 1024 constructed based on the (32,31) extended RS code

Codes	Rates	γ	Minimum Distantce
(1024,845)	0.8252	8	$\geq$10
(1024,833)	0.8134	10	$\geq$ 12
(1024,821)	0.8017	12	$\geq$ 14
(1024,809)	0.7900	14	$\geq$ 16
(1024,797)	0.7783	16	$\geq$ 18
(1024,793)	0.7744	20	$\geq$ 22
(1024,783)	0.7646	30	$\geq$ 32
(1024,781)	0.7626	32	34

IV. Construction of QC-LDPC Codes Based on the Minimum Weight Codewords of RS Codes with Two Information Symbols

- For constructing QC-LDPC codes, we need to redefine the location vectors of elements of a finite field. Again we consider the elements $0=\alpha^{\infty}$, $1=\alpha^{0}$, α, . . . , α^{q-2}. For $0 \le i < q-1$, the location vector of an nonzero element α^{i} of GF(q) is a (q-1)-tuple,

$$\mathbf{z}(\alpha^{i}) = (z_0,\ z_1, \ldots ,\ z_{q-2}),$$

where z_i=1 and all the other components are equal to zero. The location vector for the 0 element of GF(q) is represented by the all zero (q-1)-tuple, (0 0 . . . 0).

IV. Construction of QC-LDPC Codes Based on the Minimum Weight Codewords of RS Codes with Two Information Symbols (Cont'd)

- Consider the $(q,2,q\text{-}1)$ extended cyclic RS code $\mathbf{C}_b$. Each minimum weight (m-w) codeword has one and only one 0-component. For $i = -\infty$, 0, 1, . . . , q-2, let $\mathbf{v}_i = (v_{i,\infty}, v_{i,0}, v_{i,1}, \ldots, v_{i,q\text{-}2})$ be a m-w codeword with the ith component $v_{i,i}$=0. Let $\mathrm{U}_i = \{\mathbf{v}_i, \alpha\mathbf{v}_i, \ldots, \alpha^{q\text{-}2}\mathbf{v}_i\}$ be the set of q-1 m-w codewords of $\mathbf{C}_b$ with the ith components equal to 0.

- The m-w codewords of $\mathbf{C}_b$ can be partitioned into q sets, U_∞, U_0, U_1, . . . , $\mathrm{U}_{q\text{-}2}$, each consisting of q-1 m-w codewords. These sets are called uniform classes of m-w codewords of $\mathbf{C}_b$. Two m-w codewords in the same uniform class U_i differ in all the q-1 nonzero positions. Two m-w codewords from two different classes differ in at least q-1 positions.

IV. Construction of QC-LDPC Codes Based on the Minimum Weight Codewords of RS Codes with Two Information Symbols (Cont'd)

- For the ith uniform class U_i of m-w codewords, form a $(q-1)\times q$ matrix $\mathbf{G}_i$ over GF(q) with the q-1 m-w codewords in U_i as rows. For $j \neq i$, the q-1 entries of the jth column of $\mathbf{G}_i$ are nonzero and they form the q-1 nonzero elements of GF(q), and the q-1 entries of ith column of $\mathbf{G}_i$ are all zero.

- Replacing the entries of $\mathbf{G}_i$ by their location vectors, we obtain a $(q-1)\times q(q-1)$ matrix $\mathbf{B}_i$ which consists of a row of q $(q-1)\times(q-1)$ submatrices,

$$\mathbf{B}_i = \left[\mathbf{A}_{i,\infty} \ \mathbf{A}_{i,0} \ \cdots \ \mathbf{A}_{i,q-2}\right],$$

where $\mathbf{A}_{i,i}$ is a $(q-1)\times(q-1)$ zero matrix and all the other submatrices $\mathbf{A}_{i,j}$s are circulant permutation matrices.

IV. Construction of QC-LDPC Codes Based on the Minimum Weight Codewords of RS Codes with Two Information Symbols (Cont'd)

- Form the following $q \times q$ array of $(q-1)\times(q-1)$ circulant permutation and zero matrices:

$$\mathbf{H}_{qc,1} = \begin{bmatrix} \mathbf{B}_{\infty} \\ \mathbf{B}_{0} \\ \vdots \\ \mathbf{B}_{q-2} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{\infty,\infty} & \mathbf{A}_{\infty,0} & \cdots & \mathbf{A}_{\infty,q-2} \\ \mathbf{A}_{0,\infty} & \mathbf{A}_{0,0} & \cdots & \mathbf{A}_{0,q-2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{A}_{q-2,\infty} & \mathbf{A}_{q-2,0} & \cdots & \mathbf{A}_{q-2,q-2} \end{bmatrix}$$

  where the submatrices on the main diagonal are zero matrices and all the other submatrices are circulant permutation matrices. $\mathbf{H}_{qc,1}$ is a $q(q-1)\times q(q-1)$ matrix over GF(2) with both column and row weights q-1. It satisfies the RC-constraint and hence its Tanner graph has a girth at least 6.

IV. Construction of QC-LDPC Codes Based on the Minimum Weight Codewords of RS Codes with Two Information Symbols (Cont'd)

- For $1 \leq \gamma, \rho \leq q$, let $\mathbf{H}_{qc,1}(\gamma, \rho)$ be a $\gamma \times \rho$ subarray of $\mathbf{H}_{qc,1}$. If $\mathbf{H}_{qc,1}(\gamma, \rho)$ does not contain zero matrices, then the column and row weights of $\mathbf{H}_{qc,1}(\gamma, \rho)$ are γ and ρ, respectively. Then null space of $\mathbf{H}_{qc,1}(\gamma, \rho)$ gives a regular QC-LDPC code of length $n = \rho(q-1)$ with rate at least $(\rho - \gamma)/\rho$ and minimum distance at least $\gamma+1$ for odd γ and $\gamma+2$ for even γ.

- If $\mathbf{H}_{qc,1}(\gamma, \rho)$ contains zero matrices, then it has two column weights, γ-1 and γ, and two row weights ρ-1 and ρ. Then the null space of $\mathbf{H}_{qc,1}(\gamma, \rho)$ gives a near regular QC-LDPC code.

IV. Construction of QC-LDPC Codes Based on the Minimum Weight Codewords of RS Codes with Two Information Symbols (Cont'd)

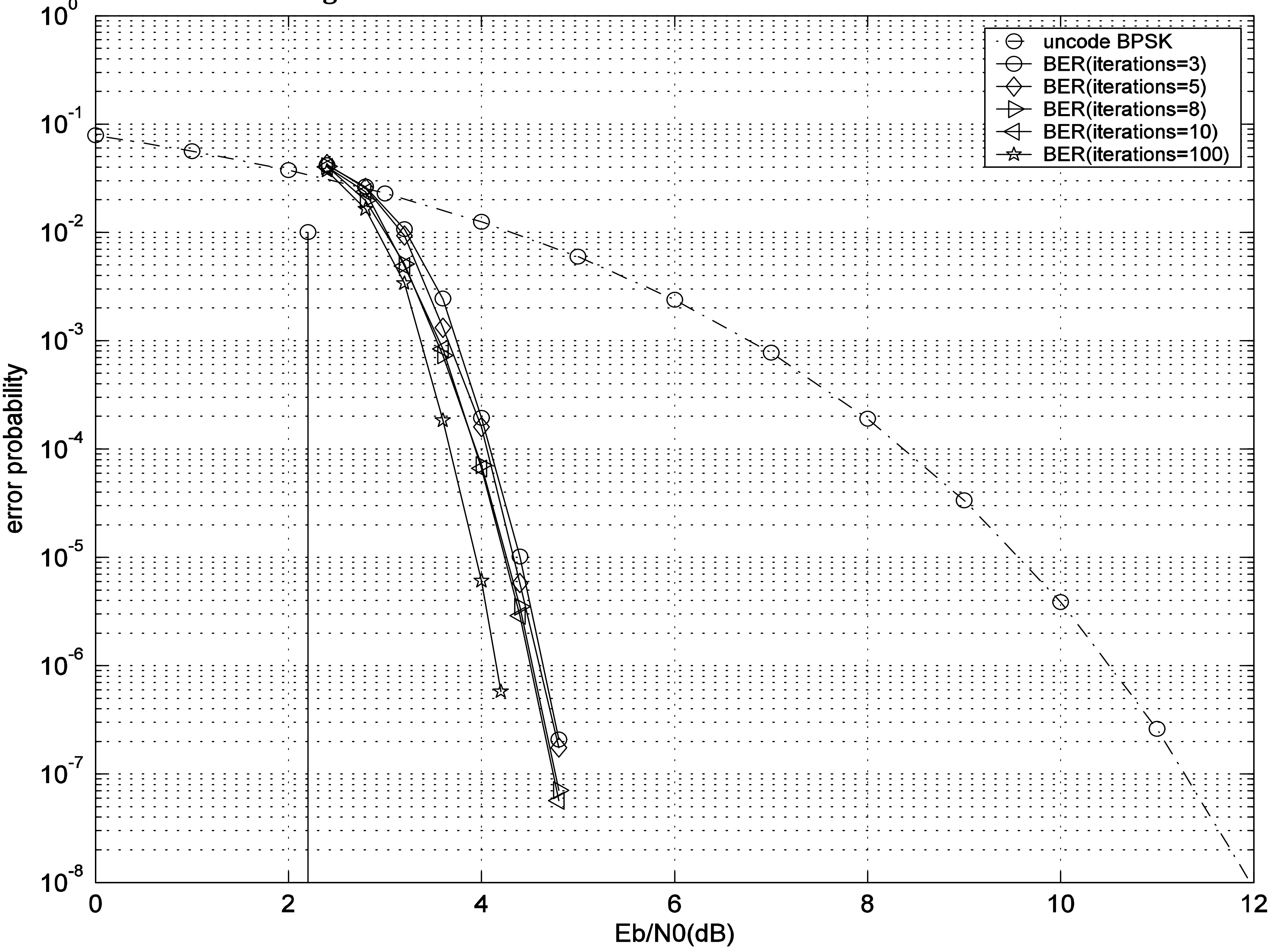
- The above construction gives a family of RS-based QC-LDPC codes with various lengths, rates and minimum distances, whose Tanner graph have girth at least 6.
- QC-LDPC codes can be encoded using simple shift-register with complexity linearly proportional to the number of parity-check bits.

  Z. -W. Lee, L. Chen, S. Lin, W. Fong and P. -S. Yeh, "Efficient encoding of quasi-cyclic LDPC codes," submitted to *IEEE Trans. Commun*.,

Example III

- In this example, the m-w codewords of the (32,2,31) extended cyclic RS code over GF(2^5) is used for code construction. Based on the m-w codewords of this RS code, we form a 32x32 array of 31x31 circulant permutation and zero matrices $\mathbf{H}_{qc,1}$.
- Set γ=10 and ρ=32, Let $\mathbf{H}_{qc,1}(10,32)$ be the subarray that consists of the first 10 rows of $\mathbf{H}_{qc,1}$. It is a 310x992 matrix over GF(2) with row weight 31 and two column weights 9 and 10. The null space of $\mathbf{H}_{qc,1}(10,32)$ gives a (992,802) QC-LDPC code with rate 0.8084 and minimum distance at least 10. The error performance of this code is shown in Figure 3. At the BER of 10^{-6}, it achieves 6 dB coding gain over the uncoded BPSK and performs within 2.0 dB from the Shannon limit.

Figure 4. (992,802)Rate=0.81 code BER performance (γ=9,10)

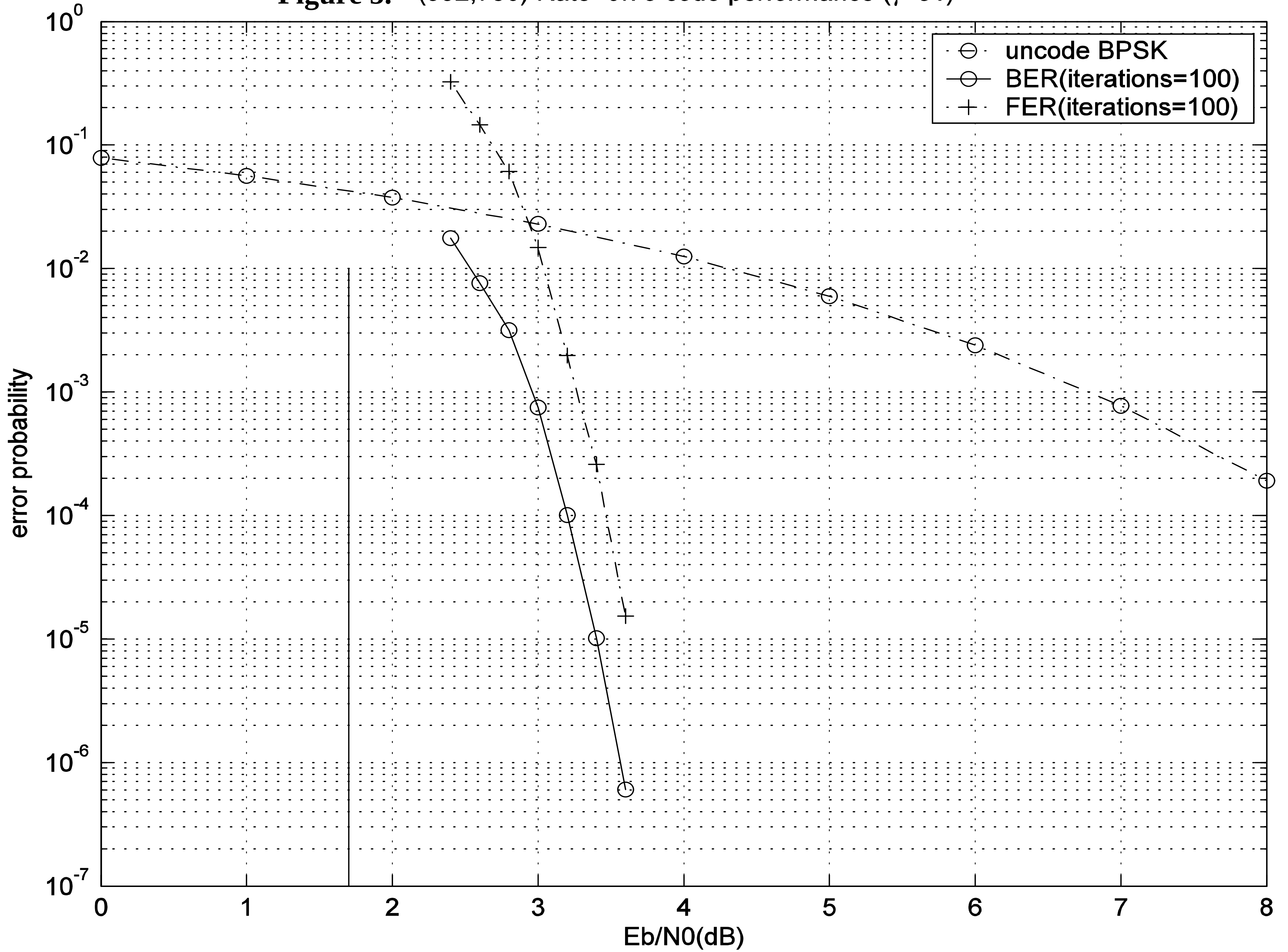


Example IV

- For code construction, we use the (32,2,31) extended cyclic RS code over GF(2^5). Set $\gamma = \rho = 32$. Then $\mathbf{H}_{qc,1}(32,32)$ is the full array $\mathbf{H}_{qc,1}$ constructed based on all the m-w codewords of the (32,2,31) RS code. The column and row weights of $\mathbf{H}_{qc,1}(32,32)$ are both 31.

- The null space of $\mathbf{H}_{qc,1}(32,32)$ gives a (992,750) QC-LDPC code with rate 0.756 and minimum distance at least 32. Its performance is shown in Figure 4. At the BER of 10^{-6}, it achieves almost 7 dB coding gain over the uncoded BPSK and performs only 1.9 dB from the Shannon limit.

Figure 5. (992,750) Rate=0.76 code performance (γ=31)



V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form

- RS codes were originally defined in polynomial form in frequency domain. Using the polynomial form, arrays of circulant permutation matrices that satisfy the RC-constraint can also be constructed from the codewords of an RS code over a prime field GF(p) with two information symbols, where p is a prime.
- Since GF(p) is a prime field, the set of integers, $\{0, 1, \ldots, p\text{-}1\}$, gives the set of elements of GF(p). The addition and multiplication of GF(p) are modulo-p addition and multiplication.

V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form (Cont'd)

- Let $\mathcal{P} = \left\{ \boldsymbol{a}(X) = a_1 X + a_0 : a_1, a_0 \in \mathrm{GF}(p) \right\}$ be the set of p^2 polynomials of degree 1 or less with coefficients from GF(p). For each polynomial in $\mathcal{P}$, define the following p-tuple over GF(p):

$$\mathbf{v} = \left(\boldsymbol{a}(0), \boldsymbol{a}(1), \ldots, \boldsymbol{a}(p-1) \right) \quad ,$$

where $\boldsymbol{a}(j) = a_1 \cdot j + a_0$ with $j \in \mathrm{GF}(p)$. Then the set of p^2 p-tuples over GF(p),

$$\mathbf{C}_b = \left\{ \mathbf{v} = \left(\boldsymbol{a}(0), \boldsymbol{a}(1), \ldots, \boldsymbol{a}(p-1) \right) : \boldsymbol{a}(X) \in \mathcal{P} \right\} \qquad (3)$$

gives a (p, 2, p-1) RS code over GF(p) with two information symbols. The RS code $\mathbf{C}_b$ given by (3) is not cyclic.

V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form (Cont'd)

- Consider the subset $\mathcal{P}_0 = \{\boldsymbol{a}(X) = a_0 : a_0 \in \mathrm{GF}(p)\}$ of zero-degree polynomials of $\mathcal{P}$. Then the set of p-tuples,

$$\begin{aligned} \mathbf{C}_0 &= \left\{\left(\boldsymbol{a}(0), \boldsymbol{a}(1), \ldots, \boldsymbol{a}(p-1)\right) : \boldsymbol{a}(X) \in \mathcal{P}_0\right\} \\ &= \left\{(a_0, a_0, \ldots, a_0) : a_0 \in \mathrm{GF}(p)\right\}, \end{aligned} \qquad (4)$$

  constructed from the zero-degree polynomials in $\mathcal{P}_0$ forms a subcode of $\mathbf{C}_b$ and is a $(p, 1, p)$ RS code over GF(p).

V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form (Cont'd)

- Partition $\mathbf{C}_b$ with respect to $\mathbf{C}_0$ into p subsets, $\mathbf{C}_0$, $\mathbf{C}_1$, … , $\mathbf{C}_{p\text{-}1}$, where

$$\mathbf{C}_i = \left\{ \left(\boldsymbol{a}(0), \boldsymbol{a}(1), \ldots, \boldsymbol{a}(p-1) \right) : \boldsymbol{a}(X) = iX + a_0, a_0 \in \mathrm{GF}(p) \right\}, \quad (5)$$

for $0 \leq i < p$, $\mathbf{C}_i$ contains p codewords in $\mathbf{C}_b$ of the following form:

$$\left(i \cdot 0 + a_0, i \cdot 1 + a_0, \ldots, i \cdot (p-1) + a_0 \right). \quad (6)$$

- $\mathbf{C}_i$ is called a *cloud* of codewords of $\mathbf{C}_b$. The codeword

$$(i \cdot 0,\ i \cdot 1, \ldots,\ i \cdot (p-1))$$

in $\mathbf{C}_i$ is called the *center* of $\mathbf{C}_i$ and the other p-1 codewords in $\mathbf{C}_i$ are called *satellites*.

V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form (Cont'd)

- For each element $j \in \mathrm{GF}(p)$, we define its *location vector* as a p-tuple, $\mathbf{z}_j = (z_0, z_1, \ldots, z_{p-1})$, with z_j=1 and all the other components equal to zero. For $0 \le i < p$, form a $p \times p$ matrix $\mathbf{G}_i$ over GF(p) with the codewords in the ith cloud $\mathbf{C}_i$ as rows. For $0 \le k < p$, the kth column of $\mathbf{G}_i$ consists of the following components: $i \cdot k + 0, i \cdot k + 1, \ldots, i \cdot k + (p-1)$, which form all the p elements of GF(p). From (4) and (5), we readily see that any two rows in $\mathbf{G}_i$ differ in all p positions.

V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form (Cont'd)

- Replacing each entry in $\mathbf{G}_i$ by its location vector, we obtain a row of p $p \times p$ submatrices, $\mathbf{B}_i = \left[\mathbf{A}_{i,0} \mathbf{A}_{i,1} \cdots \mathbf{A}_{i,p-1} \right]$, where the kth submatrix has the location vectors of $i \cdot k + 0, i \cdot k + 1, \ldots, i \cdot k + (p-1)$ as the rows,

$$\mathbf{A}_{i,k} = \begin{bmatrix} \mathbf{z}(i \cdot k + 0) \\ \mathbf{z}(i \cdot k + 1) \\ \vdots \\ \mathbf{z}(i \cdot k + (p-1)) \end{bmatrix}. \qquad (7)$$

- Under modulo-p addition and multiplication, the location vector $\mathbf{z}(i \cdot k + (j+1))$ of the field element $i \cdot k + (j+1)$ is the cyclic-shift of the location vector $\mathbf{z}(i \cdot k + j)$ of the field element $i \cdot k + j$ for $0 \le j < p$. Therefore, $\mathbf{A}_{i,k}$ is a $p \times p$ circulant permutation matrix for $0 \le k < p$ and $\mathbf{B}_i$ is a row of p circulant permutation matrices.

V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form (Cont'd)

- Form the following $p \times p$ array of $p \times p$ circulant permutation matrices:

$$\mathbf{H}_{qc,2} = \begin{bmatrix} \mathbf{A}_{0,0} & \mathbf{A}_{0,1} & \cdots & \mathbf{A}_{0,p-1} \\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} & \cdots & \mathbf{A}_{1,p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{p-1,0} & \mathbf{A}_{p-1,1} & \cdots & \mathbf{A}_{p-1,p-1} \end{bmatrix} \qquad (8)$$

$\mathbf{H}_{qc,2}$ is a $p^2 \times p^2$ matrix with constant column weight p and constant row weight p. Since the rows of $\mathbf{H}_{qc,2}$ correspond to codewords of $\mathbf{C}_b$ and two codewords in $\mathbf{C}_b$ can have at most one location with the same code symbol, no two rows (or two columns) in $\mathbf{H}_{qc,2}$ have more than one 1-component in common. Consequently, $\mathbf{H}_{qc,2}$ satisfies the RC-constraint.

V. Construction of QC-LDPC Codes Based on RS Code in Polynomial Form (Cont'd)

- For $1 \leq \gamma, \rho \leq p$, let $\mathbf{H}_{qc,2}(\gamma, \rho)$ be a $\gamma \times \rho$ subarray of $\mathbf{H}_{qc,2}$. Then $\mathbf{H}_{qc,2}(\gamma, \rho)$ is a $\gamma p \times \rho p$ matrix over GF(2) with column and row weights γ and ρ.
- The null space of $\mathbf{H}_{qc,2}(\gamma, \rho)$ gives a QC-LDPC code with girth at least 6.

VI. Construction by Masking

- Given a $\gamma \times \rho$ array of permutation (or circulant permutation) matrices, say $\mathbf{H}(\gamma, \rho)$, $\mathbf{H}_{qc,1}$ or $\mathbf{H}_{qc,2}$, a set of permutation matrices can be masked (i.e., replaced by zero matrices) to generate a new structured LDPC code with good performance.

- Masking operation can modeled mathematically as a special matrix product.

VI. Construction by Masking (Cont'd)

- To illustrate the masking operation, we use the $\gamma \times \rho$ array $\mathbf{H}_{qc,1}(\gamma, \rho)$ = [$\mathbf{A}_{i,j}$] of circulant permutation matrix as the base matrix for masking. Let $\mathbf{W}(\gamma, \rho) =$ [$w_{i,j}$] be a $\gamma \times \rho$ matrix over GF(2). Define the following matrix product:

$$\mathbf{M}_{qc,1}(\gamma, \rho) = \mathbf{W}(\gamma, \rho) \otimes \mathbf{H}_{qc,1}(\gamma, \rho) = \left[w_{i,j} \mathbf{A}_{i,j} \right],$$

where $w_{i,j}\mathbf{A}_{i,j} = \mathbf{A}_{i,j}$ for $w_{i,j}$=1 and $w_{i,j}\mathbf{A}_{i,j} = \mathbf{O}$ (a $(q-1) \times (q-1)$ zero matrix) for $w_{i,j}$=0. We call $\mathbf{W}(\gamma, \rho)$ the masking matrix, $\mathbf{H}_{qc,1}(\gamma, \rho)$ the base matrix, and $\mathbf{M}(\gamma, \rho)$ the masked matrix. In masking, a set of circulant permutation matrices in the base matrix $\mathbf{H}_{qc,1}(\gamma, \rho)$ is masked by the 0-entries of the masking matrix $\mathbf{W}(\gamma, \rho)$. If $\mathbf{H}_{qc,1}(\gamma, \rho)$ contains zero submatrices, we avoid to mask these zero submatrices.

VI. Construction by Masking (Cont'd)

- The masked matrix $\mathbf{M}_{qc,1}(\gamma,\rho)$ is an array of circulant permutation and zero matrices. The distribution of circulant matrices in $\mathbf{M}_{qc,1}(\gamma,\rho)$ is identical to the distribution of the 1-entry in the base matrix $\mathbf{W}(\gamma,\rho)$.
- Masking operation preserves the RC-constraint on the rows and columns of the base matrix and hence $\mathbf{M}_{qc,1}(\gamma,\rho)$ also satisfies the RC-constraint. Furthermore, masking reduces the density of 1-entries of the base matrix and hence the masked matrix $\mathbf{M}_{qc,1}(\gamma,\rho)$ is a sparser matrix. Consequently, the Tanner graph of $\mathbf{M}_{qc,1}(\gamma,\rho)$ has either larger girth or smaller number of short cycles than that of the base matrix.

VI. Construction by Masking (Cont'd)

- If the girth of the Tanner graph of the masking matrix $\mathbf{W}(\gamma, \rho)$ is $g \geq 6$, then the girth of the Tanner graph of the masked matrix $\mathbf{M}_{qc,1}(\gamma, \rho)$ is at least g. Since the size of a masking matrix is in general small, it is quite easy to construct masking matrices with relatively large girth, say 6, 8, 10, and 12, either by computer search or algebraic methods.
- The null space of the masked matrix $\mathbf{M}_{qc,1}(\gamma, \rho)$ gives a QC-LDPC code $\mathbf{C}_{qc,1}$ with girth at least 6.
- If the masking matrix $\mathbf{W}(\gamma, \rho)$ is a regular matrix, $\mathbf{C}_{qc,1}(\gamma, \rho)$ is a regular QC-LDPC code. If the masking matrix $\mathbf{W}(\gamma, \rho)$ has varying column and varying row weights, then $\mathbf{C}_{qc,1}$ is an irregular QC-LDPC code.

VI. Construction by Masking (Cont'd)

- Masking is particularly effective for constructing structured irregular codes which have encoding advantage over random irregular codes.
- One approach to construct irregular LDPC codes is based on variable- and check-node degree distributions of the code graphs derived from density evolution of the messages passed between the two types of nodes in a belief propagation decoder.

T. J. Richardson, M. A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, no.2, pp. 619-637, Feb. 2001.

VI. Construction by Masking (Cont'd)

- Let $\mathbf{v}(X) = \sum_{i=1}^{d_v} v_i X^{i-1}$ and $\mathbf{c}(X) = \sum_{i=1}^{d_c} c_i X^{i-1}$, be the variable- and check-node degree distributions of a code graph designed for a given rate R, where v_i and c_i are the fractions of variable- and check-nodes that have degree i, and d_v anc d_c are the maximum variable- and check-node degrees, respectively.
- Construct a masking matrix $\mathbf{W}(\gamma, \rho)$ with column and row weight distributions identical (or close) to the degree distributions $\mathbf{v}(X)$ and $\mathbf{c}(X)$, respectively, by computer search. Masking the base matrix $\mathbf{H}_{qc,1}(\gamma, \rho)$ with $\mathbf{W}(\gamma, \rho)$, then the masked matrix $\mathbf{M}_{qc,1}(\gamma, \rho)$ has column and row weight distributions identical (or close) to $\mathbf{v}(X)$ and $\mathbf{c}(X)$, respectively.

VI. Construction by Masking (Cont'd)

- The masked matrix $\mathbf{M}_{qc,1}(\gamma,\rho)$ is an array of circulant permutation and zero matrices. The null space of $\mathbf{M}_{qc,1}(\gamma,\rho)$ gives an irregular QC-LDPC code that can be encoded with simple shift-registers.
- Proper masking gives very good structured regular and irregular LDPC codes that perform just as well as random LDPC codes.
- J. Xu, L. Chen, I. Djurdjevic, S. Lin and K. Abdel-Ghaffar, "Construction of regular and irregular LDPC codes: geometry decomposition masking," submitted to *IEEE Trans. Inform. Theory*, 2004.

VI. Construction by Masking (Cont'd)

- The following degree distributions

$$\mathbf{v}(X) = 0.4410X + 0.3603X^2 + 0.00171X^5 + 0.03543X^6 +$$
$$0.09331X^7 + 0.0204X^8 + 0.0048X^9 + 0.000353X^{27} + 0.04292X^{29}$$
$$\mathbf{c}(X) = 0.00842X^7 + 0.99023X^8 + 0.00135X^9$$

  are derived based on density evolution for a code of rate 1/2.

- The next three examples give three long irregular QC-LDPC codes of rate 1/2 constructed based on the above degree distributions.

Figure 6. LDPC(16002, 8001) Code Performance

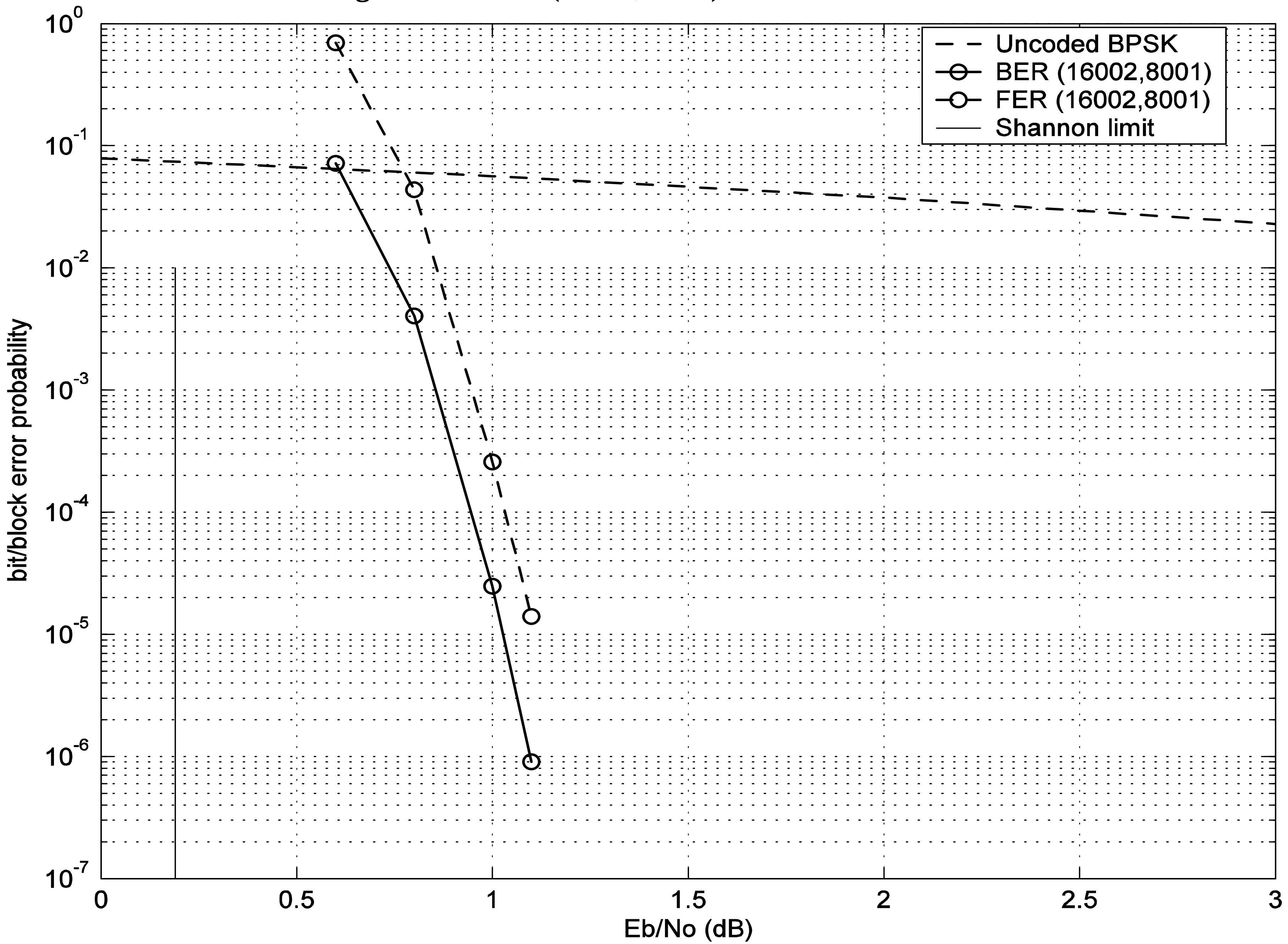


Figure 7. LDPC(32130, 16065) Code Performance

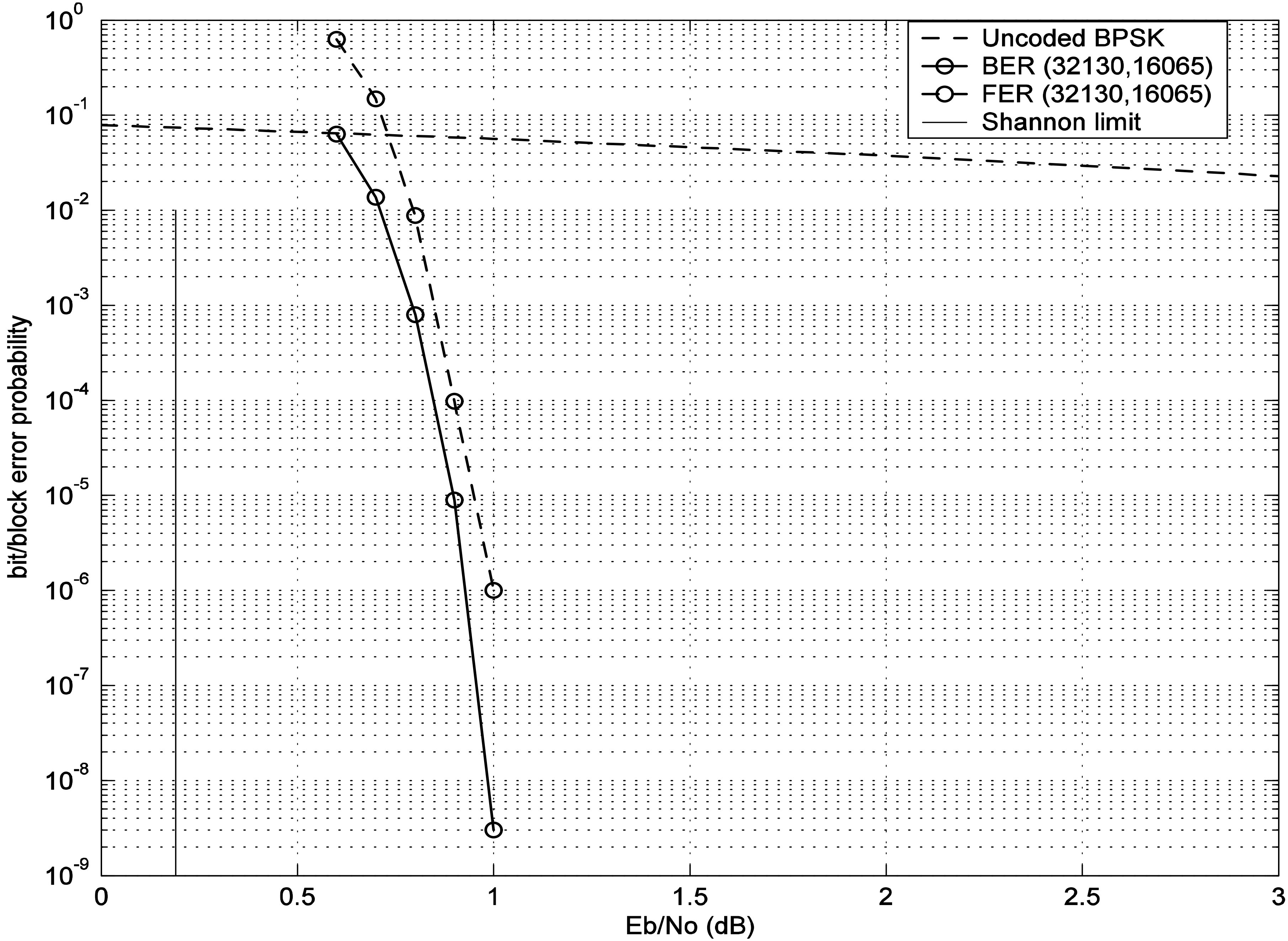
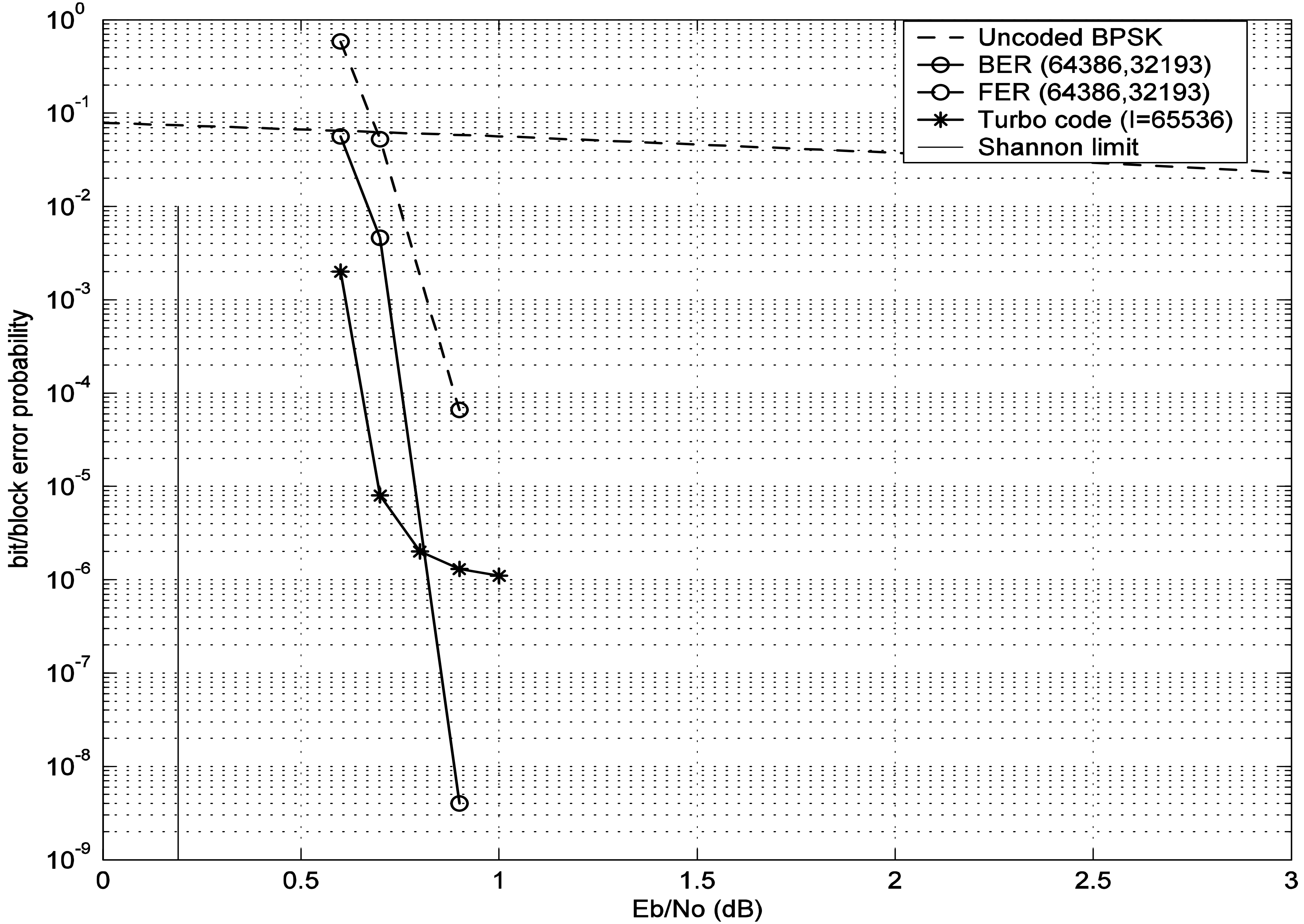


Figure 8. LDPC(64386, 32193) Code Performance



VII. Construction Based on Finite Geometries

- Construction based on the hyperplane, lines and points of either Euclidean and projective geometries.

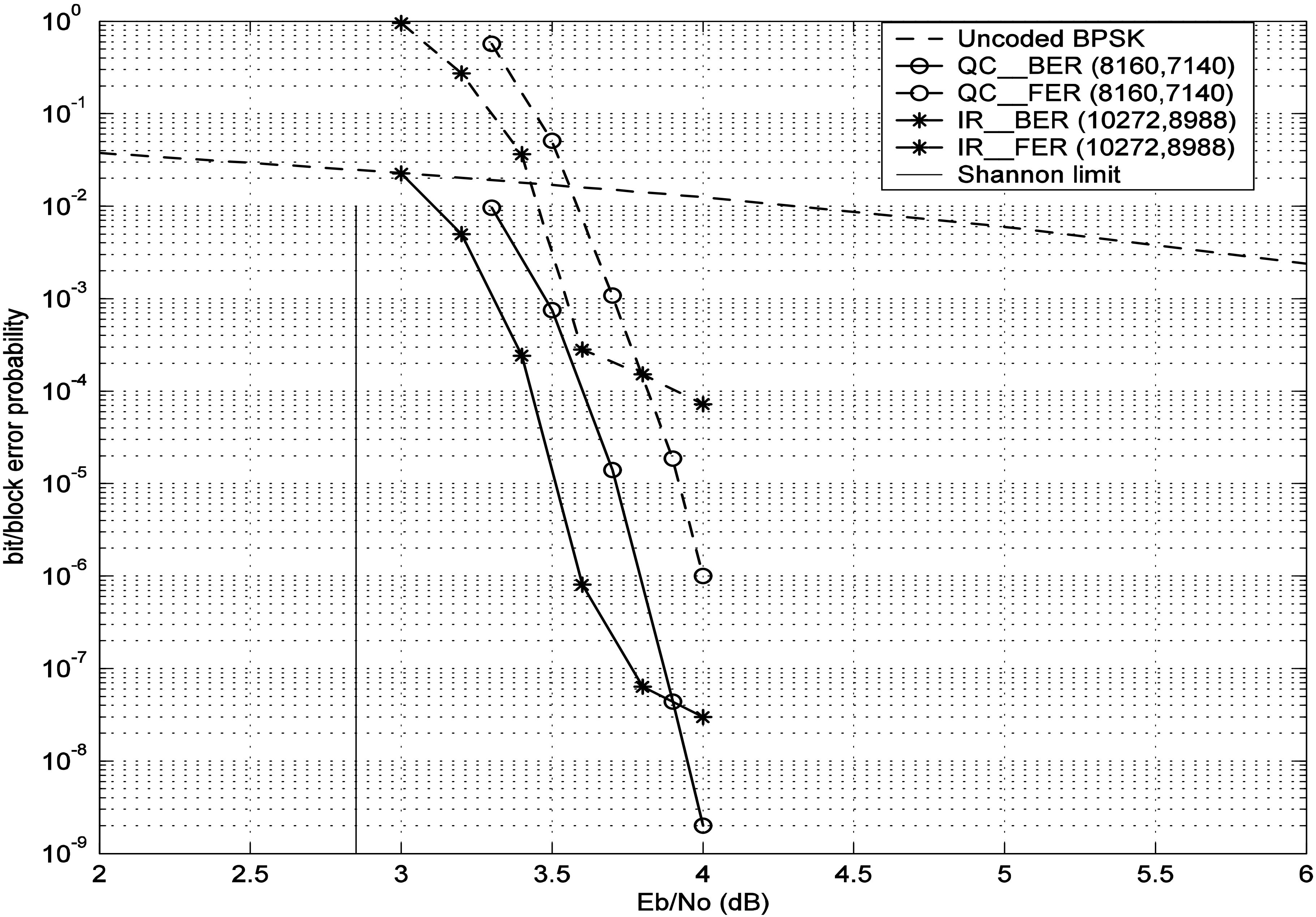
- LDPC codes constructed are either cyclic or quasi-cyclic with large minimum distance and girth at least 6.

- No or very low error-floor.

Example VIII (NASA/GSFC Code)

- Construction geometry: 3-dimensional Euclidean geometry EG($3,2^3$)
- Parity-check matrix **H**: a 2x16 array of 511x511 circulant matrices, each having weight 2. The column and row weights of H are 4 and 32, respectively.
- Code: a (8176,7156) QC-LDPC code with rate 7/8 and girth 6
- Shannon gap at the 10^{-6}: 1 dB
- Error-floor: no down to the BER of 10^{-12} (verified by FPGA)
- Decoding convergence: very fast, only 5 iteration are needed
- Encoding: Two 511-stage shift-register-adder-accumulator (SRAA) units for serial encoding

Figure 9. LDPC(8160, 7140) Code Performance



VIII. Reed-Solomon Codes V.S. LDPC Codes

- RS codes by far form the best class of codes.
- Decoding methods: algebraic decoding, reliability-based algebraic decoding, list decoding and turbo decoding through decomposition and self-concatenation.
- If an effective soft-decision scheme (or algorithm) for decoding RS codes can be devised, then RS codes will outperform all the other codes, including LDPC codes. There is no such decoding algorithm.
- The next few graphs show the performance of two popular RS codes and some short LDPC codes proposed for 10GBASE-T. The purpose of these graphs is not for comparison, because the code lengths, rates, and types of decoding are different.

Figure 10 Code Performance

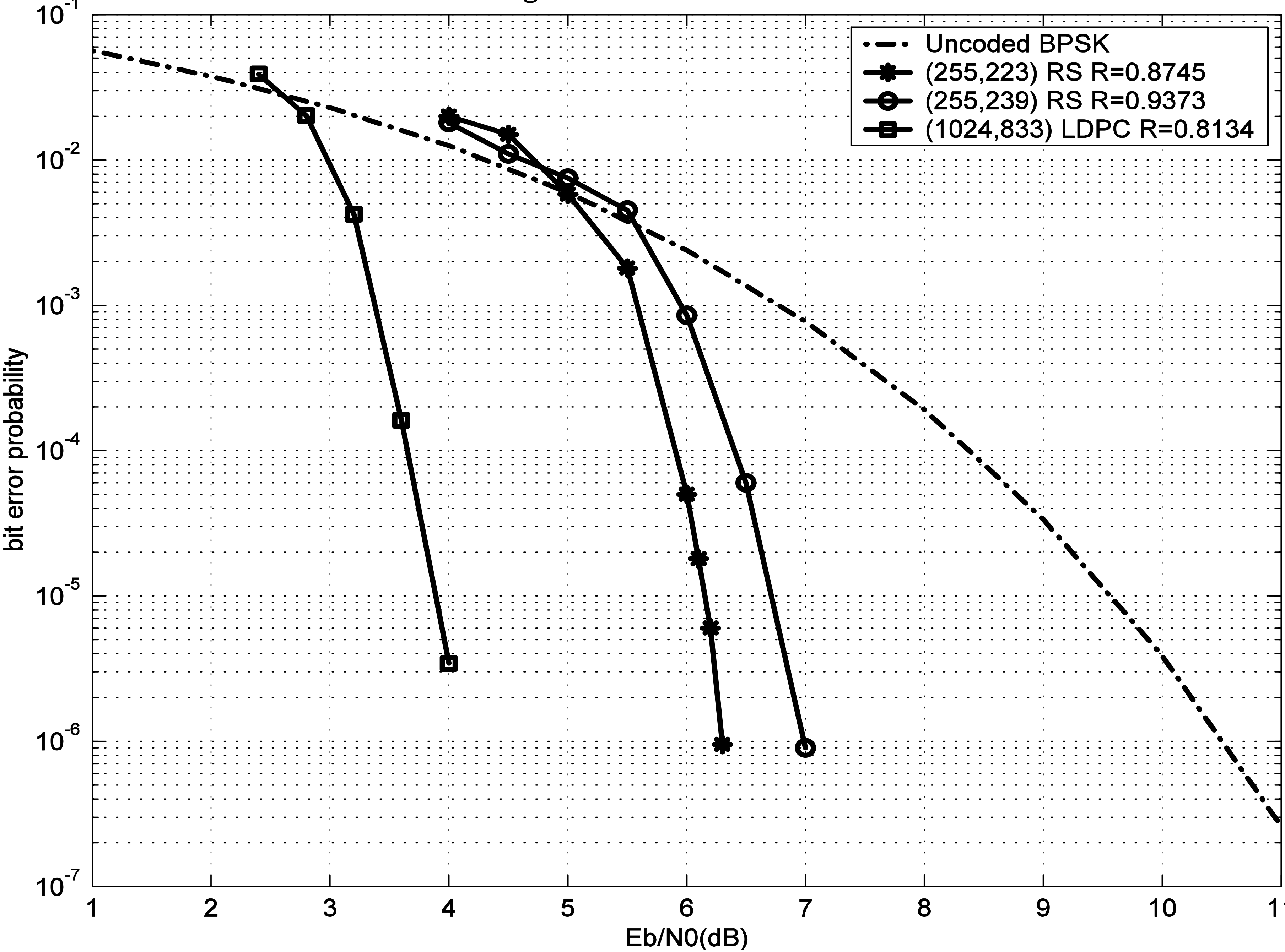


Figure 11 Code Performance

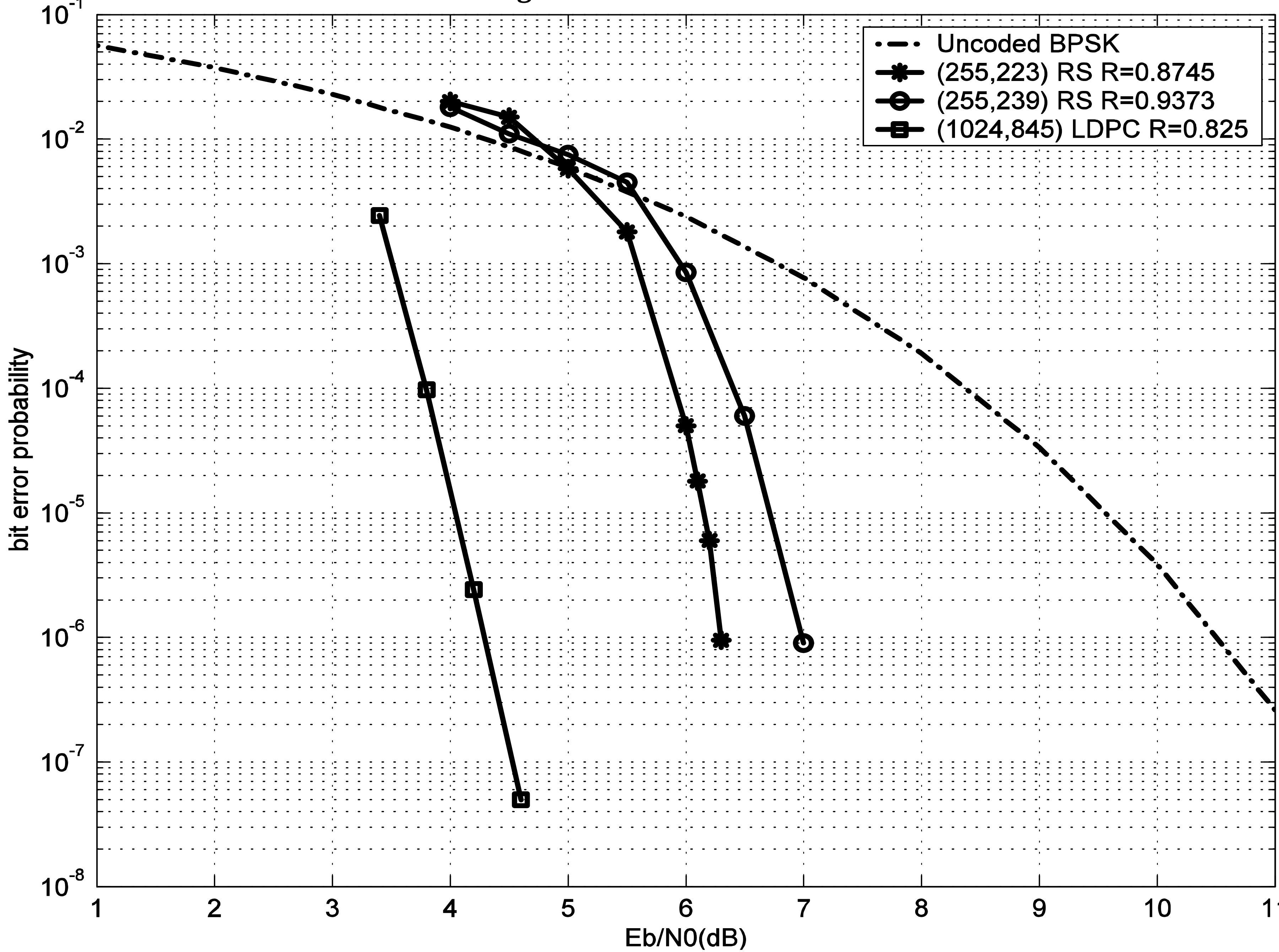
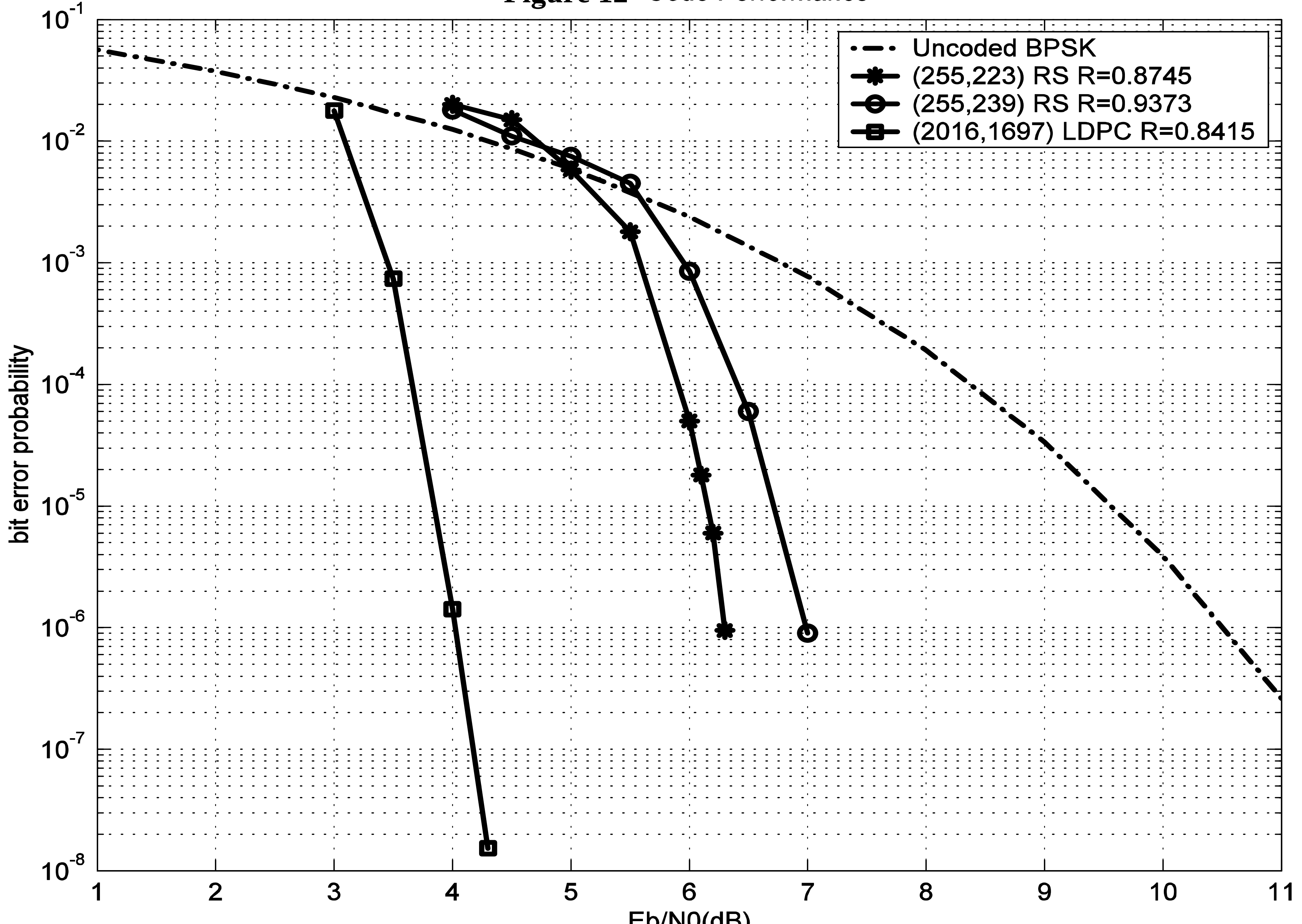


Figure 12 Code Performance



IX. Turbo Decoding of RS Codes

- RS codes can be turbo encoded and decoded through decomposition and self-concatenation.

  C. Y. Liu and S. Lin, "Turbo encoding and decoding of RS codes through binary decomposition and self-concatenation," to appear in *IEEE Trans. Communications*, vol. 52, no. 9, September 2004.

- This turbo decoding of RS codes can achieve large coding over algebraic, reliability-based and current list decoding algorithms or schemes.

IX. Turbo Decoding of RS Codes

- Figure 13 shows the performances of the (127,113,15) RS code over $GF(2^7)$ with algebraic, GMD, Chase-GMD and turbo decoding. We see that at BER of 10^{-6}, Turbo decoding achieves almost 2 dB coding gain over the algebraic decoding.

- The (255,239,17) RS code over $GF(2^8)$ can be practically decoded based on two component codes, each having a trellis with 256 states.

Figure 13. Bit error performance of the (127, 113, 15) RS code