# De Bruijn Sequences for DS/CDMA Transmission: Efficient Generation, Statistical Analysis and Performance Evaluation

Susanna Spinsante(*), Madhyiar Sarayloo(*), Ennio Gambi(*), Chirag Warty(**), **Claudio Sacchi**(***)

*(*) Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, Ancona (Italy)*
*(**) Intelligent Communication Lab, Mumbai (India)*
*(***) Dept. of Information Engineering and Computer Science (DISI), University of Trento, Trento (Italy)*

# Outline

- *Introduction and aims of the paper;*
- *Efficient generation of large sets of De Bruijn sequences;*
- *De Bruijn sequences properties;*
- *Statistical analysis of DS/CDMA system performance;*
- *Numerical results;*
- *Conclusion.*

# Introduction

- **Random (or quasi-random) spreading sequences for DS/CDMA**
  - Direct Sequence Code Division Multiple Access (DS/CDMA) still represents a core technology for the physical layer of commercially remunerative applications and standards (radiolocalization, automotive radar, 3G UMTS);
  - A very critical issue of Spread Spectrum and CDMA: keeping the probability of intercept the lowest possible;
  - Secure information hiding must be guaranteed at the physical layer level: random spreading sequences should be applied;
  - Due to complexity of generating truly random sequences, deterministic sequences (i.e. pseudorandom) are used in real applications;
  - Required features: pseudo-noise auto-correlation patterns, quasi-orthogonal cross-correlation.

# Introduction

- **Gold and De Bruijn sequence sets**
  - Typical choice: Gold codes, generated as logical combination of linear shift register (LSR) sequences (*preferred pairs*) of span *n* (= number of LSR cells);

  - Gold codes features:
    - favorable statistical properties;
    - small cardinality = *N* + 2, where *N* (sequence length) = $2^n - 1$

  - In the literature, the alternative use of **De Bruijn** binary sequences for DS/CDMA has been recently proposed [**SPI11**]. Their features are:
    - generation by nonlinear shift register;
    - maximal length ($N = 2^n$);
    - very large cardinality $2^{2^{(n-1)}-n}$
    - interesting correlation-related features [**AND10, SPI11, SPI13, WAR13, SAR14**]

# Aims of the paper and advancement with respect to related work

- Propose an efficient sequence generation algorithm based on De Bruijn graphs theory and Eulerian cycles;

- Formal statistical analysis of De Bruijn sequences in DS/CDMA with explicit computation of 2nd and 4th order statistics (variance and normalized kurtosis) of multi-user interference (MUI), in asynchronous BPSK-modulated DS/ CDMA transmission;

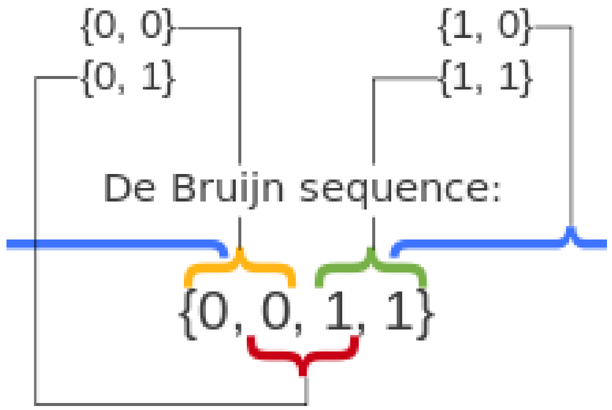
- Closed form computation of average bit-error-probability (BEP):
  - by Gaussian approximation [**PUR76**]
  - by non-Gaussian evaluation [**TES99**], based on the Generalized Gaussian modeling of the global detection noise affecting the CDMA receiver (Gaussian noise + MUI)

- MUI statistics and BEP performance comparison to Gold codes with and without code selection driven by a formal criterion.

At the end of this analysis we may have more insights about the use of De Bruijn sequences in real DS/CDMA systems

# Efficient generation of large sets of De Bruijn sequences

***DEF***: in a binary De Bruijn sequence viewed cyclically over a period, each binary n-tuple appears exactly once, included the all-zero n-tuple, due to the non-linear nature of the generating register

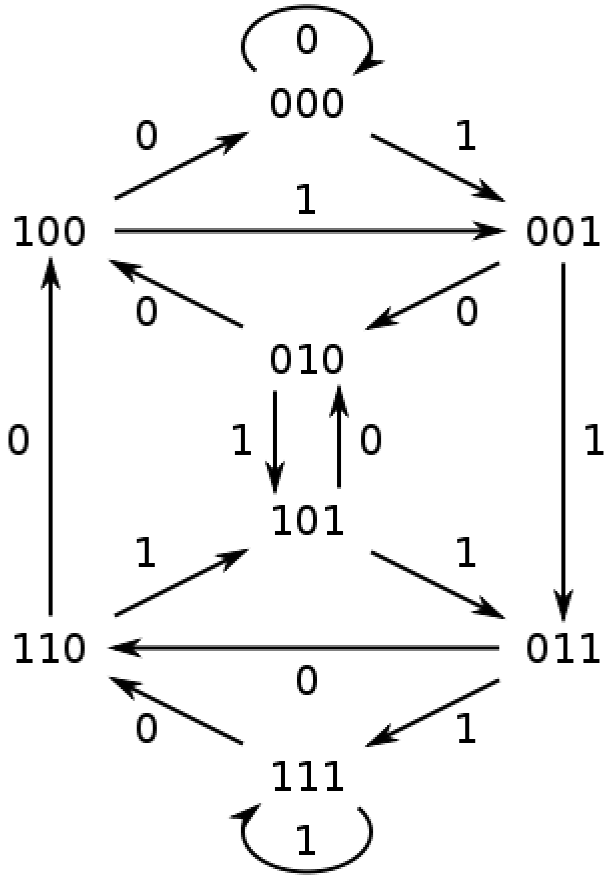



- Generation by Non Linear Feedback Shift Registers (NLFSRs)
- NLFSR state at time t: $\boldsymbol{s}(t) = (s_1(t), s_2(t), \dots, s_3(t)),\ s_i(t) \epsilon A = \{0,1\}, for\ i = 1,2,\dots,n$ where
- At each clock transition:
  - each memory cell content shifted one position to the right
  - leftmost cell $s_n(t)$ updated by the output of a nonlinear feedback function *g(.)*
  - *g(.)* defines a mapping of $A^n \rightarrow A$
- At time *(t+1),* the state of the register is given by: $s_i(t+1) = \begin{cases} s_{i+1}(t), for\ i = 1,2,\dots,n-1 \\ g(\boldsymbol{s}(t)), for\ i = n \end{cases}$

# Efficient generation of large sets of De Bruijn sequences

- Various generation methods proposed in the literature:
  - use of a lower-order and n-bit pattern initial stage [**TUR11**]
  - prefer-one, prefer-opposite [**ALH10**], and prefer-same [**FRE82**] approaches for bit insertion
  - n-stage FSRs [**CHA90, ZHA09**]

- To **increase time efficiency**, De Bruijn graphs theory and Eulerian cycles are here exploited to <u>ignore a large number of sequences</u> that do not verify the definition given above:
  - each acceptable sequence starts from an arbitrary vertex and walks through the graph, by crossing each degree no more than once;
  - a rotated sequence to either the right or the left is not a different De Bruijn sequence;
  - the bitwise *not* of each De Bruijn sequence originates a distinct De Bruijn sequence of the same family: generation of half the set allows to obtain the whole family;
  - decimal representation of the array used to store the sequences.




*A De Bruijn graph: every four-digit sequence occurs exactly once if one traverses every edge exactly once and returns to one's starting point (an Eulerian cycle)*

# Efficient generation of large sets of De Bruijn sequences

- Generating algorithm pseudocode:

1.Parameters initialization: $n$, $L\_Seq$ (length of sequence), $N\_Seq$ (number of distinct sequences), $T\_Seq$ (matrix to store the generated De Bruijn sequences)
2.Set $Poss\_Seq$ to $\{1, \{0\}^n, 1\}$ or $\{0, \{1\}^n, 0\}$
3.Calculate $Dir\_Vectors$ according to $Poss\_Seq$
4.Calculate $Next\_Num$ according to $Dir\_Vectors$
5.**Loop** (# generated sequences $< N\_Seq$)
5.1Calculate next possible vertex w.r.t. $Dir\_Vectors$
5.2Update $Poss\_Seq$ according to the next possible vertex
5.3Update $Next\_Num$ according to $Dir\_Vectors$
5.4**If** (achieved sequence meets De Bruijn definition) **then**
Calculate bitwise NOT of the generated sequence
Rotate the generated sequence and its complementary such that it starts with $0^n$
Store the decimal value of both the generated sequence and its complementary one, in $T\_Seq$
6.**End Loop**

- Generation time for different span:

| Span | 3 | 4 | 5 | 6 |
|---|---|---|---|---|
| Time (sec.) | 0.263 | 0.416 | 70.764 | $\approx$ 4 days |
| # generated sequences | 2 | 16 | 2048 | 4000000 |
| # sequences | 2 | 16 | 2048 | 67108864 |

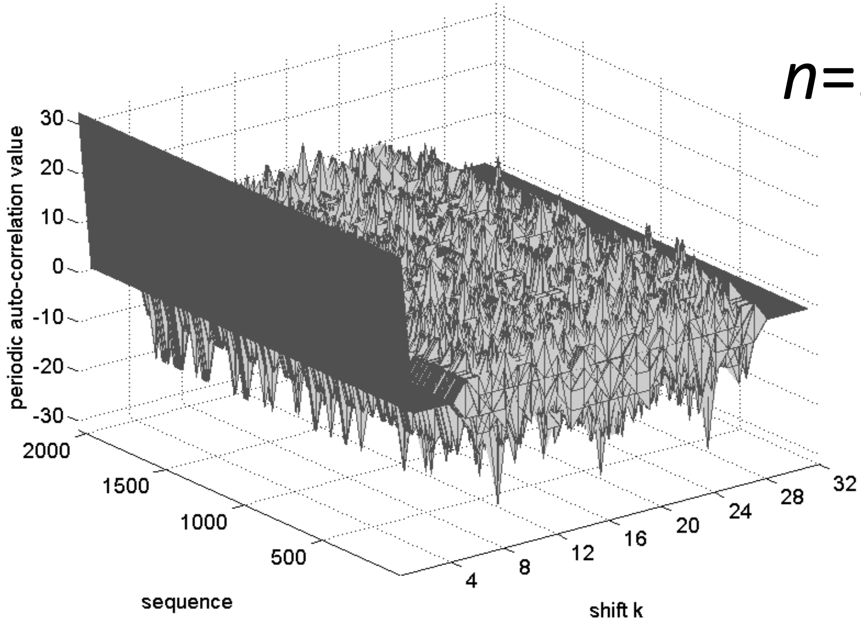
- Sequence sets: length and cardinality comparison:

| | $m$-sequences | | Gold | | De Bruijn | |
|---|---|---|---|---|---|---|
| $n$ | length | # seq. | length | # seq. | length | # seq. |
| 5 | 31 | 6 | 31 | 33 | 32 | 2048 |
| 6 | 63 | 6 | 63 | 65 | 64 | $2^{26}$ |
| 7 | 127 | 18 | 127 | 129 | 128 | $2^{57}$ |
| 8 | 255 | 16 | 255 | 257 | 256 | $2^{120}$ |
| 9 | 511 | 48 | 511 | 513 | 512 | $2^{247}$ |
| 10 | 1023 | 60 | 1023 | 1025 | 1024 | $2^{502}$ |

# De Bruijn sequences properties

- Periodic auto-correlation ${C_{aa}}^{P}[k]$ of span n De Bruijn sequence **a** for a shift $k$:
  - ${C_{aa}}^{P}[k] = 2^n$, for $k = 0$
  - ${C_{aa}}^{P}[k] = 0$, for $1 \leq |k| \leq n-1$ (Zero Correlation Zone)
  - ${C_{aa}}^{P}[k] \neq 0, |k| = n$
  - ${C_{aa}}^{P}[k] \equiv 0 (\mathrm{mod}\ 4), \forall \mathrm{k}, \mathrm{n} \geq 2$
- Bound on periodic auto-correlation sidelobes values:
  - $0 \leq \max {C^{P}}_{aa}[k] \leq 2^n - 4\left[\frac{2^n}{2n}\right]^{+}$, $1 \leq k \leq N-1, N = 2^n$

*n*=5




Bound on $max{C_{aa}}^{P}[k]$ sidelobe value for 5≤n≤10

| span $n$ | length $N$ | null samples around peak | bound on $max\ C^{P}_{\mathbf{aa}}[k]$ | ratio $\frac{max\ C^{P}_{\mathbf{aa}}[k]}{N}$ |
|---|---|---|---|---|
| 5 | 32 | $1 \leq \|k\| \leq 4$ | $[0, 16]$ | 0.5 |
| 6 | 64 | $1 \leq \|k\| \leq 5$ | $[0, 40]$ | 0.625 |
| 7 | 128 | $1 \leq \|k\| \leq 6$ | $[0, 88]$ | 0.687 |
| 8 | 256 | $1 \leq \|k\| \leq 7$ | $[0, 192]$ | 0.75 |
| 9 | 512 | $1 \leq \|k\| \leq 8$ | $[0, 396]$ | 0.77 |
| 10 | 1024 | $1 \leq \|k\| \leq 9$ | $[0, 816]$ | 0.79 |

# De Bruijn sequences properties

- Cross-correlation function $C_{a_1 a_2}[k]$ for a shift $k$:

  - $C_{a_1 a_2}[k] = C_{a_1 a_2}[N-k], 0 \leq k \leq N-1$
  - $\sum_{k=0}^{N-1} C_{a_1 a_2}[k] = 0$
  - $C_{a_1 a_2}[k] \equiv 0 (\mathrm{mod}\ 4), \mathrm{n} \geq 2, \forall k$

- Bound on cross-correlation sidelobes values:

$$-2^n \leq C_{a_1 a_2}[k] \leq 2^n - 4, 0 \leq k \leq N-1$$

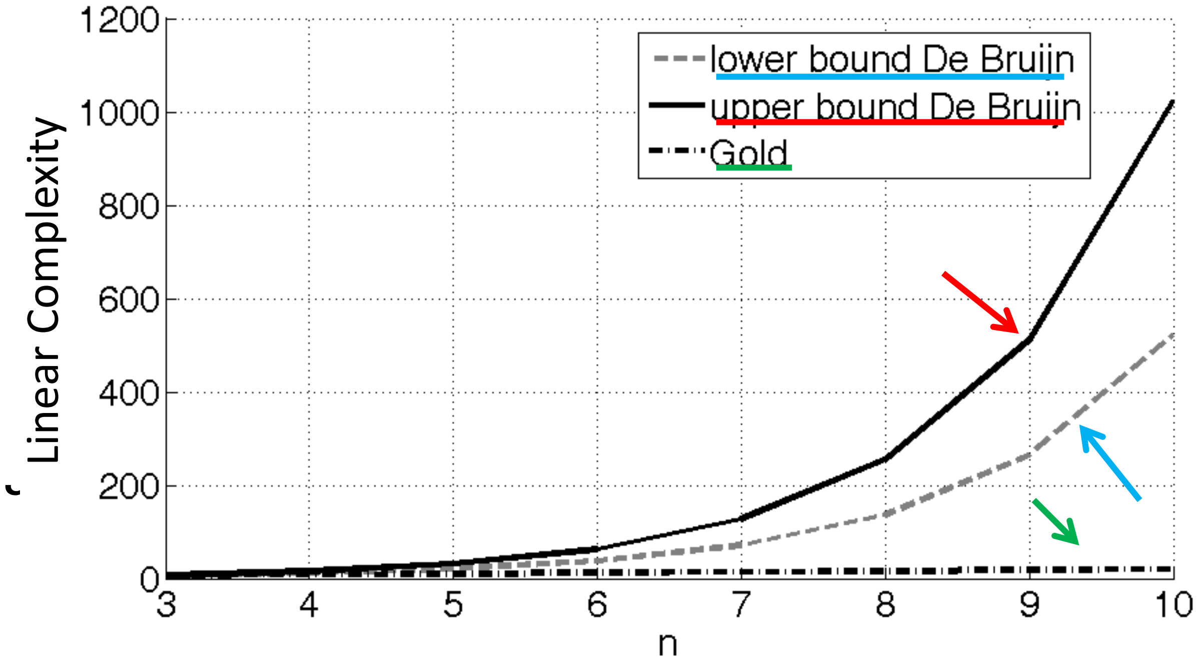
| | Max abs. value | Mean | $\sigma$ |
|---|---|---|---|
| $m$-sequence | 11 | 0.032258 | 5.65391 |
| Gold31 | 9 | -0.0447 | 5.4064 |
| De Bruijn | 32 | 0 | 6.0703 |

*Maximum absolute value, mean, and standard deviation of the cross-correlation, for De Bruijn, Gold, and m-sequences of span n = 5*

# De Bruijn sequences properties

- Randomness analysis: **Golomb's postulates [GOL82]**
  - 1st and 2nd postulates (balance and run properties): always verified;
  - 3rd postulate (ideal 2-level auto-correlation): not verified BUT a Zero Correlation Zone is exhibited;
- **Linear complexity** (**C**) = estimated length of the shortest LFSR which would be able to generate the sequence itself;
- According to Berlekamp [**BER68**], **C** provides a numerical description of the amount of information needed to infer the structure of the spreading codes generation algorithm
- De Bruijn sequences: $2^{n-1} + n \le \boldsymbol{C} \le 2^n - 1$
- Gold codes: $\boldsymbol{C} = n$

Linear complexity profiles and bounds of sequences, 3≤ n ≤ 10



# Statistical analysis of DS/CDMA system performance

- **Received multi-user DS/CDMA signal (1):**
  - From the channel (supposed to be AWGN):

$$r(t)=\sqrt{2P}\sum_{k=1}^{K} b_k\left(t-\tau_k\right)a_k\left(t-\tau_k\right)\cos\left(2\pi f_0 t+\theta-\phi_k\right)+z(t)$$

  - After coherent demodulation and de-spreading (ref. user 1), sampled at *t=T*:

$$R=\sqrt{\frac{P}{2}}Tb_{1,0}+\sqrt{\frac{P}{2}}T\left(\frac{1}{N}\sum_{k=2}^{K} I_{k,1}\right)+\xi$$

$$I_{k,1}=\left\{\overline{\chi}_{k,1}\left(\alpha_k\right)+\left[\overline{\chi}_{k,1}\left(\alpha_k+1\right)-\overline{\chi}_{k,1}\left(\alpha_k\right)\right]\upsilon_k\right\}\cos\left(\phi_k\right) \qquad \overline{\chi}_{k,1}\left(\alpha_k\right)=\begin{cases} C_{a_k,a_1}\left(\alpha_k\right) & if\ b_{k,-1}=b_{k,0} \\ \hat{C}_{a_k,a_1}\left(\alpha_k\right) & if\ b_{k,-1}\neq b_{k,0} \end{cases}$$

***Multi-User Interference (MUI) term***

$$\alpha_k T_c \le \tau_k < \left(\alpha_k+1\right)T_c \qquad \upsilon_k \triangleq \left(\tau_k-\alpha_k T_c\right)/T_c$$

**Even and odd PN cross correlations**

# Statistical analysis of DS/CDMA system performance

- **Received multi-user DS/CDMA signal (2):**
  - More in details [**PUR76,TES99**]:

$$C_{a_k,a_1}(\alpha_k)=\Psi_{a_k,a_1}(\alpha_k)+\Psi_{a_k,a_1}(\alpha_k-N)$$

$$\hat{C}_{a_k,a_1}(\alpha_k)=\Psi_{a_k,a_1}(\alpha_k)-\Psi_{a_k,a_1}(\alpha_k-N)$$

$$\Psi_{a_k,a_1}(l)=\begin{cases}\sum_{j=0}^{N-1-l} a_k(l)a_1(j+l) & 0\le l\le N-1\\ \sum_{j=0}^{N-1+l} a_k(j-l)a_1(l) & 1-N\le l<0\\ 0 & |l|\ge N\end{cases}$$

  - Considering a BPSK modulation and deterministic (known) spreading sequences, the BEP computation is as follows:

$$P_{be}=\Pr\left\{error \big/ b_{k,0}=-1\right\}=\Pr\left\{R>0 \big/ b_{k,0}=-1\right\}=\Pr\left\{\left[\sqrt{\frac{P}{2}}T\left(\frac{1}{N}\sum_{k=2}^{K}I_{k,1}\right)+\xi\right]>\sqrt{\frac{P}{2}}T\right\}$$

Practically: $P_{be}=\int_{\sqrt{P/2}T}^{+\infty} f_{Z_G}(z)dz$ where: $Z_G=\sqrt{\frac{P}{2}}T\left(\frac{1}{N}\sum_{k=2}^{K}I_{k,1}\right)+\xi$

# Statistical analysis of DS/CDMA system performance

- **How can we NUMERICALLY compute DS/CDMA BEP?**
  - In other words: can we express in closed form the probability density function of the random variable $Z_G$?
  - The answer is **NO**, therefore, we should resort to some approximation:
    - ***Gaussian Approximation (GA)***: it simply considers a Gaussian distribution for $Z_G$. It is reasonable when the number of users is large [**PUR76**];
    - ***Generalized Gaussian Approximation (GG)***: as the pdf of MUI for real-valued binary sequence has an impulsive pseudo-Laplace distribution (*leptokurtic*), we can suppose that the pdf of $Z_G$ fits well with the Generalized Gaussian pdf model [**TES99**], expressed in terms of its normalized kurtosis:

$$f_{Z_G}(z) = \frac{c\gamma}{\Gamma(1/c)} \exp\left(-|\gamma z|^c\right)$$

$$\gamma = \sqrt{\frac{\Gamma(3/c)}{\Gamma(1/c)\,\mathrm{var}(Z_G)}}$$

$$\kappa(Z_G) \triangleq \frac{E\left(Z_G^4\right)}{\left\{E\left(Z_G^2\right)\right\}^2} = \frac{\frac{3}{4}\left(\frac{E_b}{\eta}\right)^{-2} + \frac{E\left(I^4\right)}{N^4} + 3\left(\frac{E_b}{\eta}\right)^{-1}\frac{E\left(I^2\right)}{N^2}}{\left[\frac{E\left(I^2\right)}{N^2} + \frac{\eta}{2E_b}\right]^2}$$

$$c = F\left(\kappa_{Z_G}\right) \approx \sqrt{\frac{5}{\kappa_{Z_G} - 1.865} - 0.12}$$

$$2 < \kappa_{Z_G} < 10$$

(an alternative, more precise expression of $F$, valid for a wider range of values of the normalized kurtosis is in eq.30 of the paper)

$\Gamma$ = Euler's Gamma function

$\kappa$ = 3 and $c$=2 for Gaussian-distributed r.v.

# Statistical analysis of DS/CDMA system performance

- **Approximated analytical expressions for DS/CDMA BEP:**
  - Using GA approximation, BEP is given as follows:

$$P_{be} \approx Q\left(\sqrt{SINR}\right) \quad SINR \triangleq \left(\frac{E\left(I^2\right)}{N^2} + \frac{\eta}{2E_b}\right)^{-1}$$

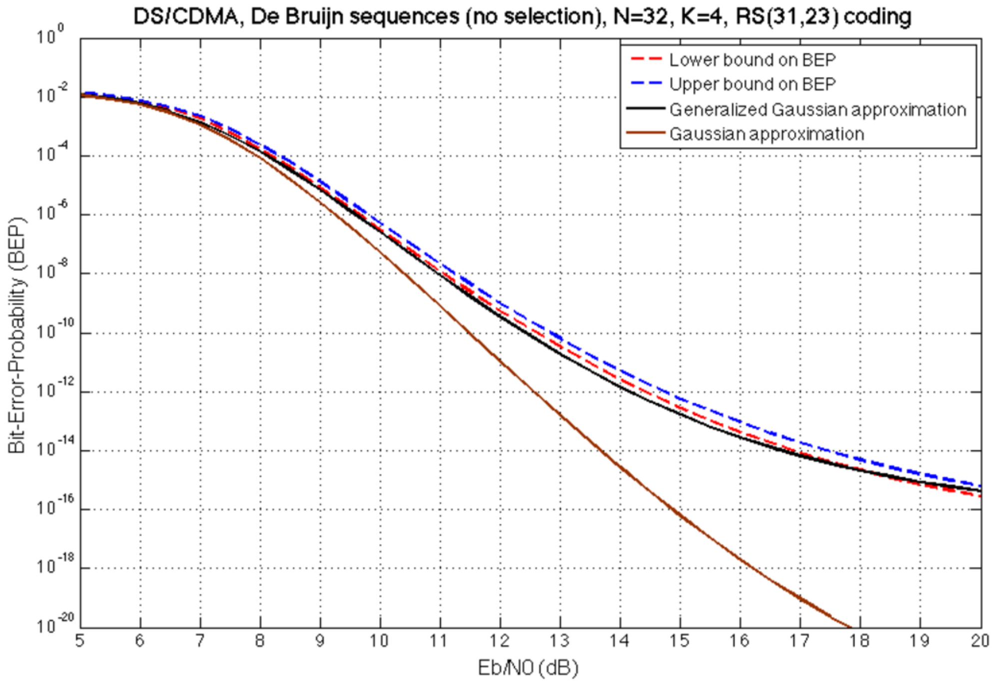
  - Using GG approximation, we obtain after some mathematical manipulations:

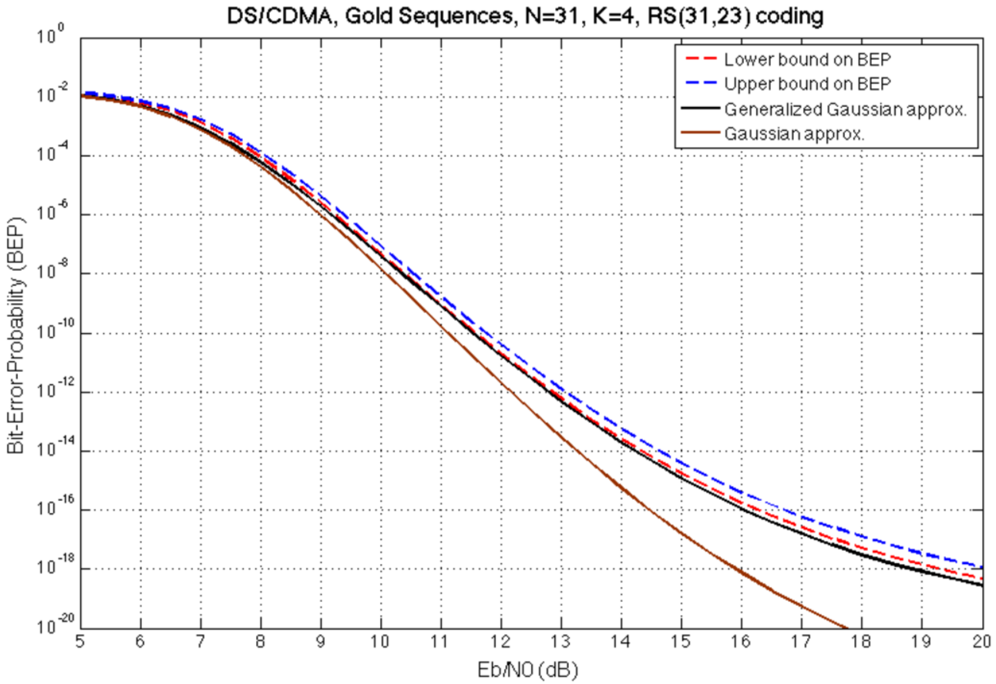
$$P_{be} \approx \frac{1}{2} - \frac{1}{2}\Gamma_{inc}\left(\left[\frac{\Gamma(3/c)}{\Gamma(1/c)} SINR\right]^{c/2}, \frac{1}{c}\right) \qquad \Gamma_{inc}(x,s) = \frac{1}{\Gamma(s)}\int_0^x t^{(s-1)} e^{-t} dt$$

«Incomplete» Gamma function

# Numerical results

- **Random sequence selection**
  - This really means: no selection criterion applied, random indices of the De Bruijn matrix have been used to select the sequences;
  - GA and GG approximations have been compared with tight upper and lower bounds on DS/CDMA BEP computed as in [**LEH89**];
  - BPSK modulation with Reed-Solomon coding (RS(31,23)) have been considered in deriving numerical results (an analytical lower bound on BER is available for RS coding);
  - *N*=32 and *K*=4 users have been considered in BEP computations.




***MUI Statistics***

| | $E(I^2)/N^2$ | $\kappa(I)$ |
|---|---|---|
| **De Bruijn** | 0.0349 | 3.56 |
| **Gold** | 0.0306 | 3.43 |

Gold sequences performs better than De Bruijn ones thanks to their superior “Gaussianity” (the GA curve is closer to GG and upper and lower bounds)

# Numerical results

- **Making things smarter: sequence selection criterion** (*K* sequences selected, for each value of *N*)
  i. all the sequences in each set are assessed for their minimum aperiodic auto-correlation sidelobe $C_{a_k,a_1}(\alpha_k)$
  ii. looking at the lowest minimum aperiodic auto-correlation sidelobe values found in i), the subset featuring the lowest sidelobe value joint a number of sequences *K* is selected;
  iii. K sequences are extracted from the subset obtained in ii), by looking at sequence pairs featuring the most favorable aperiodic cross-correlation;
  iv. iv) if it is not possible to find a close subset of K sequences as per iii), they are selected randomly over the subset obtained in ii);

  Number of groups of non-duplicated K sequences out of *M* (cardinality of the set):

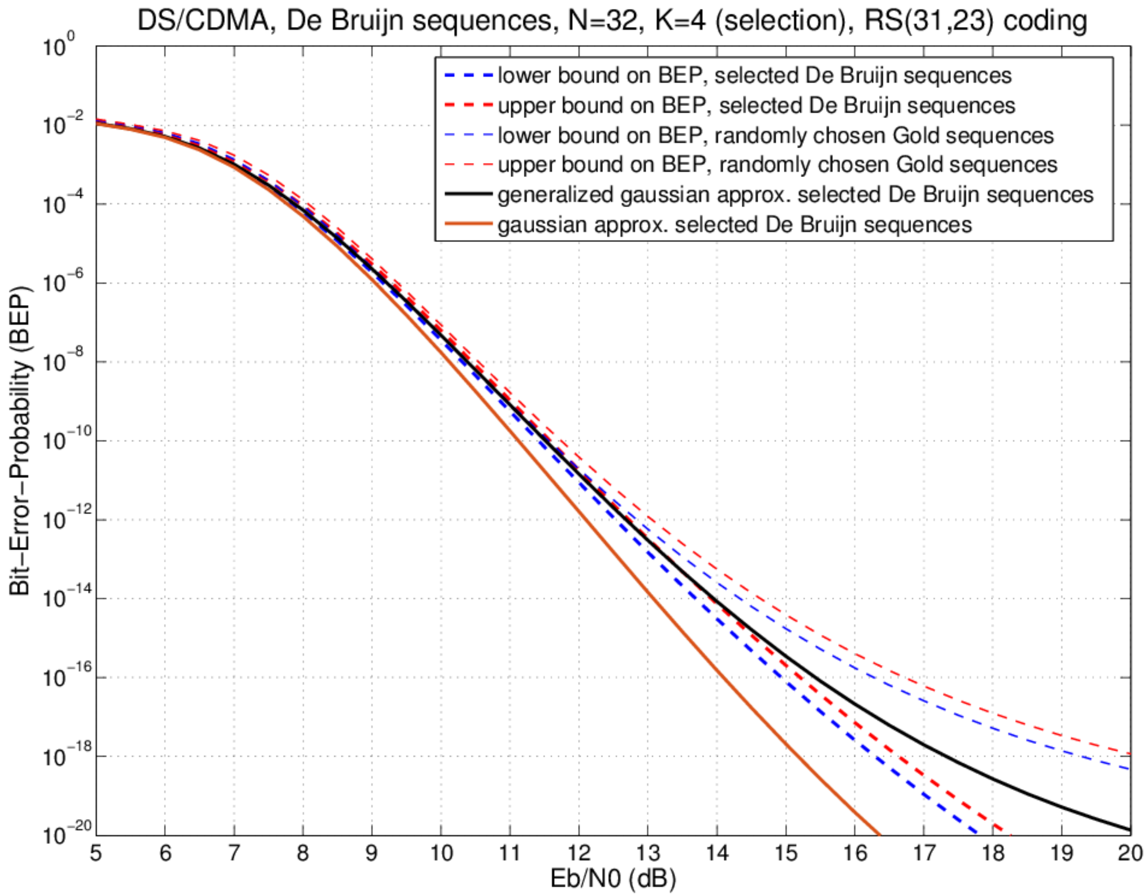
$$G_{M,K} = \binom{M}{K} = \frac{M!}{K!(M-K)!}$$

Check the **whole family** for sequences featuring the **MINIMUM** auto-correlation sidelobe

select the subset **S** for which:
(lowest auto-correlation sidelobe)
**AND**
(# seq. >= K)

select K sequences from **S** with best aperiodic cross-correlation
**OR**
randomly select K sequences in S

- **Sequence selection**
  - Minimum aperiodic cross-correlation sidelobe criterion;
  - Things are changing: selected De Bruijn sequences decreases both variance and normalized kurtosis of MUI;
  - As result, BEP is noticeable decreased with respect to Gold sequences (the selection criterion is not effective for small Gold sets).



# Conclusion

- Performance of binary De Bruijn sequences assessed, as spreading codes in multiple users DS/CDMA systems, through a formal statistical analysis of link performance, in comparison with traditionally used Gold codes;
- The formal statistical analysis shows that De Bruijn codes exhibit performance comparable to Gold codes and even worse if no selection criterion is applied;
- On the other hand, the selection criterion based on the minimization the pairwise aperiodic cross-correlation among the sequences associated to different users may lead to remarkably improved performance of De Bruijn sequences;
- The much greater cardinality, and better randomness-related properties of De Bruijn sequences, could anyway improve the robustness of the communication system against interception or security attacks.

# References

**[AND10]** S. Andrenacci, E. Gambi, C. Sacchi, and S. Spinsante, “Application of de Bruijn sequences in automotive radar systems: Preliminary evaluations,” *in Proc. IEEE RADARCON’10*.

**[SPI11]** S. Spinsante, S. Andrenacci, and E. Gambi, “Binary de bruijn sequences for ds/cdma systems: analysis and results,” *EURASIP Jour. on Wir. Comm. and Networking*, vol. 4, 2011.

**[SPI13]** S. Spinsante and E. Gambi, “De Bruijn binary sequences and spread spectrum applications: a marriage possible?”, *IEEE Aerosp. Electron. Syst. Mag*., vol.28, no.11, pp. 28 – 39, Nov. 2013.

**[WAR13]** C. Warty, E. Gambi, and S. Spinsante, “Secured scrambling codes for vehicular control and navigation,*” in Connected Vehicles and Expo (ICCVE), 2013 International Conference on*, Dec 2013, pp. 920–925.

**[SAR14]** M. Sarayloo, E. Gambi, and S. Spinsante, “De bruijn sequences as zero correlation zone codes for satellite navigation systems,” *in Telecommunications (ICT), 2014 21st International Conference on*, May 2014, pp. 216–220.

**[PUR76]** M. Pursley, “Performance evaluation of phase-coded spread spectrum multiple access communication – part 1: System analysis,” *IEEE Trans. Commun*., vol.COMM-25, Aug. 1977, pp. 816-825.

**[TES99]** A. Teschioni, C. Sacchi, and C. Regazzoni, “Non Gaussian characterization of DS/CDMA noise in few-user systems with complex signature sequences,” *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 234–237, Jan 1999.

**[LEH89]** J. Lehnert, “An efficient technique for evaluating direct sequence spread-spectrum multiple-access communications,” *IEEE Transactions on Communications*, vol. 37, no. 8, pp. 851–858, Aug 1989.

**[TUR11]** M. Turan, “Evolutionary construction of de Bruijn sequences,” *in Proc. 4th ACM Workshop on Security* AISec’11.

**[ALH10]** A. Alhakim, “A simple combinatorial algorithm for de Bruijn sequences,” *American Mathematical Monthly*, vol. 117.

**[FRE82]** H. Fredricksen, “A survey of full-length nonlinear shift register cycle algorithms,” *SIAM Review*, vol. 24, pp. 195–221, Apr. 1982.

**[CHA90]** A. Chan and R. Games, “On the quadratic spans of de Bruijn sequences,*” IEEE Trans. Inf. Theory*, vol. 36, pp. 822–829, Jul. 1990.

**[GOL82]** S. Golomb, Ed., Shift Register Sequences. Laguna Hills, CA: Aegean Park Press, 1982.

**[BER68]** E. Berlekamp, Ed., Algebraic coding theory. New York: McGraw-Hill, 1968.

**[ZHA09]** W. Zhang, S. Liu, and H. Huang, “An efficient implementation algorithm for generating de Bruijn sequences,” *Computer Standards & Interfaces*, vol. 31, pp. 1190–1191, Nov. 2009.