

Efficient Information Set Decoding Based on Genetic Algorithms

Ahmed Azouaoui, Idriss Chana, Mostafa Belkasmi

SIME Lab., ENSIAS, Mohammed V-Souissi University, Rabat, Morocco
Email: aazouaoui@gmail.com

Received May 17, 2012; revised June 8, 2012; accepted June 20, 2012

ABSTRACT

In this paper, we describe a hard-decision decoding technique based on Genetic Algorithms (HDGA), which is applicable to the general case of error correcting codes where the only known structure is given by the generating matrix G . Then we present a new soft-decision decoding based on HDGA and the Chase algorithm (SDGA). The performance of some binary and non-binary Linear Block Codes are given for HDGA and SDGA over Gaussian and Rayleigh channels. The performances show that the HDGA decoder has the same performances as the Berlekamp-Massey Algorithm (BMA) in various transmission channels. On the other hand, the performances of SDGA are equivalent to soft-decision decoding using Chase algorithm and BMA (Chase-BMA). The complexity of decoders proposed is also discussed and compared to those of other decoders.

Keywords: Genetic Algorithms (GA); Error Correcting Codes; RS Codes; Information Set Decoding; Chase Algorithm

1. Introduction

In digital communication one of the important issues is how to transmit the message from the source to the destination as faithfully as possible. One of the most used techniques and also the most convenient is the adoption of error-correcting codes. Indeed the codes are used to improve the reliability of data transmitted over communication channels susceptible to noise. The Coding techniques are based on the following principle: Add the redundancy to the message to obtain a vector called "codeword". Decoding techniques are based on the algorithms which try to find the most likely transmitted code word related to the received (see **Figure 1**).

Decoding algorithms are classified into two categories: hard-decision and soft decision algorithms. Hard decision algorithms work on a binary form of the received information. In contrast, soft decision algorithms work directly on the received symbols [1].

Soft-decision decoding is an NP-hard problem and was approached in different ways. Recently artificial intelligence (AI) techniques were introduced to solve this problem. These techniques show very good results. Among related works, one work A* algorithm to decode linear block codes [2], another one uses genetic algorithms for decoding linear block codes [3] and a third one uses Compact Genetic Algorithms to decode BCH codes [4]. Maini *et al.* were the first, to our knowledge, to introduce Genetic algorithms [5] in the decoding of linear block

codes. Hebbes *et al.* [6] worked on the integration of genetic algorithms in a classical turbo codes decoder, and Durand *et al.* [7] worked on the optimization of turbo decoding by optimizing the interleaver with a genetic algorithm. Furthermore the deployment of Artificial Neural Networks (ANN), to train the system for higher fault tolerance in OFDM is used by Praveenkumar [8]. There are also other works [9-11] based on AI trying to solve problems related to coding theory.

We have investigated the use of genetic algorithms in different ways. In [12], GA is used to search of good double-circulant codes. In [13], a new soft decoding of block codes based on the genetic algorithms with very good performances was presented.

The purpose of this paper is to complete the study of the decoder proposed in [14]. Indeed, at first we study the complexity of the decoder by comparing it with the complexity of the BMA and Chase decoders. Secondly, we apply it to the non-binary codes family such as Reed Solomon (RS) codes. And finally, we evaluate its performances on a Rayleigh fading channel.

The rest of this paper is organized as follows: Section 2 introduces the decoder based on genetic algorithms for linear block codes. The Section 3 presents the Chase algorithms decoding and our soft decoding algorithm based on GA. In Section 4 we study the complexity of the hard and the soft decoders. Simulation result are given in Section 5, Section 6 concludes this paper.

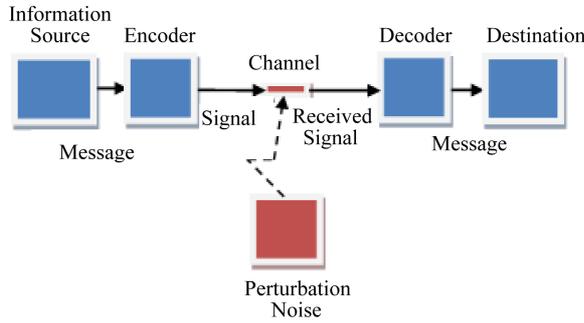


Figure 1. A simplified communication system model.

And before concluding this paper we present the experimentation results.

2. Hard-Decision Decoder Based on GA

2.1. Information Set (IS)

In a (n, k) linear block code, an information set is defined as set of k symbols of a codeword that can be independently specified. The remaining $(n - k)$ elements are redundant parity-check symbols. Thus, if a received word has been corrupted by noise in some bits of the transmitted word, and if it is possible to find an error-free (IS), so that all the wrong bits are paritycheck bits, then the received vector can be successfully decoded.

In this work, an (IS) is represented by an n -dimensional vector in which k bits are equal to 1 and the others are set to 0. The non-zero bits of the (IS) correspond to linearly independent columns of the matrix G such an information set is represented by the symbol p .

2.2. HDGA Algorithm

The purpose of the presented GA is to convert the decoding problem into search problem for a feasible solution around the received vector $r = (r_0, r_1, \dots, r_{n-1})$. In the presented case, the population will always consists of individuals which are intrinsically associated with codewords. Besides, the evolutionary operations over the individuals will always result in new individuals which are associated with codewords. The evolution, therefore, will take place within the constrained search space. After a finite number of generations, the best individual (codeword) in the population will become a highly evolved solution to the problem.

Step 1: Initial Population

The initial population is obtained by random generation of N pairs of vectors r and p_i to form N individuals (r, p_i) .

Step 2: For i from 1 to N_g

Step 2.1: Crossover and Replication

The crossover operator provides $2N$ offspring as follows:

Choose form the intermediate population tow individuals at random, for example, $c_i = (r, p_i)$ and $c_k = (r, p_k)$.

Pair the IS chromosome p_i with p_k to generate a crossover q_i as described bellow.

Set all n bits of q_i to zero.

For each $l = 1, 2, \dots, n$, do:

If $p_i[l] = p_k[l] = 0$ do $q_i[l] \leftarrow 1$;

For each location of q_i that is equal to 1, get the corresponding columns of the generator matrix G to form a vectorial set. If this set is linearly dependent (LD), do:

$q_i[l] \leftarrow 0$;

End of for.

Since the code is (n, k) , if only m positions were set to 1 in the IS chromosome, then the remaining $k - m$ positions equal to 0 must be used to find the remaining $k - m$ information bits. Thus, set at random $k - m$ in the $n - m$ remaining zero bits, in such a way that q_i represents a valid information set of the code.

These offspring and parents are put together to form an intermediate population.

Step 2.2: Fitness Evaluation

The objective function is the Hamming distance between a codeword c_i and the received vector r , given by $D_H(c_i, r)$. The fitness function can be simply defined as

$$a(c_i) = \frac{1}{[1 + D_H(c_i, r)]} \quad (1)$$

Step 2.3: Saving the Best Individuals

The M fittest individuals are copied and saved.

Step 2.4: Mutation

This operator alters an individual by bit inversion of chromosome X . However, such an inversion takes place only in one bit and only when an improvement in the individual's fitness is achieved. If it is not possible to improve the individual's fitness, then no alteration is performed.

Step 2.5: Fitness Evaluation

Step 2.6: Saving the Best Individuals

The M fittest individuals are copied and saved.

Step 2.7: Fitness Evaluation

Step 2.8: Selection

The natural selection is used and the probability of selecting the i^{th} individual of c_i the population, will be defined as

$$p(c_i) = \frac{a(c_i)}{\sum_{j=0}^{N-1} a(c_j)} \quad (2)$$

Step 2.9: Elitist Strategy

Restore the M best individuals.

Step 3: Solution

The best member from the last generation is returned

as the final decision.

3. Soft-Decision Decoding Based on Genetic Algorithm

3.1. Chase Algorithm

The Chase algorithm [15] is a suboptimal decoding procedure that uses a set of most likely error patterns. These error patterns are selected on the basis of the reliability of the received symbols. Each error pattern is added to the hard-decision received word and decoded using a hard-decision decoder. Each decoded code word is scored by computing its metric with respect to the received sequence. The code word with the best metric is selected as the most likely.

3.2. SDGA Algorithm

Let C be a binary linear (n, k, d) block code where d is its minimum distance. C is capable to correct any combination of t or less random bit errors. Let $\bar{r} = (r_1, r_2, \dots, r_n)$ be the received word from the output of the channel, $r_i = (-1)^{v_i} + w_i$, where w_i is a zero-mean Gaussian random variable with variance $N_0/2$, $i = 1, 2, \dots, n$.

The sign bits of the received values represent the hard-decision

$$\bar{z}_0 = (z_{0,1}, z_{0,1}, \dots, z_{0,n}), z_{0,j} = \text{sign}(r_j), 0 \leq j < n,$$

where

$$\text{sign}(x) = \begin{cases} 0, & \text{if } x \geq 0; \\ 1, & \text{otherwise} \end{cases}$$

The reliabilities of the received channel values, for binary transmission over an AWGN channel, are the amplitudes $|r_i|$. The received symbol reliabilities are ordered with a sorting algorithm (e.g., quick sort). The output of the algorithm is a list of indexes I_j , $j = 1, 2, \dots, n$, such that

$$|r_{I_1}| \leq |r_{I_2}| \leq \dots \leq |r_{I_n}|$$

In the first round of decoding, the hard-decision received word \bar{z}_0 is fed into a HDGA. Let \bar{v}_0 denote the decoded code word, stored as the initial code word guess. Then the metric of \bar{v}_0 with respect to the received word

$$v_0 = \sum_{j=1}^n (-1)^{v_{0j}} \quad (4)$$

is computed and its value is stored as the maximum.

Test the error patterns with at most $\lfloor d/2 \rfloor$ errors located within the bit positions having the $\lfloor d/2 \rfloor$ lowest reliabilities.

- For $i = 1, 2, \dots, 2^{t-1}$, an error pattern \bar{e}_i is added to the hard-decision received word: $\bar{z}_i = \bar{e}_i \oplus \bar{z}_0$. The

word “data” is plural, not singular.

- The error patterns are generated among the t least reliable positions (LRPs), that is, Positions $\{I_1, I_2, \dots, I_t\}$, for which reliabilities (amplitudes) are the smallest.
- Each test vector \bar{z}_i is input to a HDGA, producing a codeword \bar{v}_i , $i = 1, 2, \dots, 2^{t-1}$.
- The metric is computed according to Equation (3) and if maximum, code word \bar{v}_i stored as the most likely.

The **Figure 2** below illustrates the steps of our SDGA algorithm.

4. Complexity Analysis

Let n be the code length, k be the code dimension, t be the error correction capability of a linear bloc code C , N_i be the population size which must be equal to the total number of individuals in the population, and let N_g be the number of generation. The **Table 1** shows the complexity of the four algorithms.

For BMA, the complexity is polynomial in n . Similarly, the complexity of HDGA is polynomial in n , k , N_i , N_g and $\log(N_i)$. Although, the complexity of the

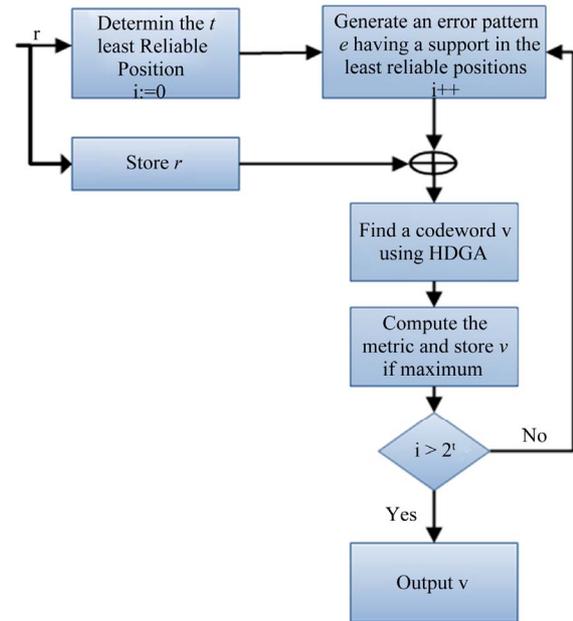


Figure 2. Flow diagram of SDGA.

Table 1. Complexity of BMA, HDGA, Chase-BMA, SDGA.

Algorithm	Complexity
BMA	$O(n^2)$
HDGA	$O(N_i N_g [kn^2 + kn + \log(N_i)])$
Chase-BMA	$O(2^t n^2)$
SDGA	$O(2^t (N_i N_g [kn^2 + kn + \log(N_i)]))$

Chase-BMA and SDGA is exponentially in t . For the non-binary codes, such as RS codes, all the expression in **Table 1** must be multiplied by the term $\log(n)$.

We also note that BMA is less complex than HDGA. While the complexity of BMA is less than the decoder based on genetic algorithms, it can decode all linear codes while this is not the case for BMA and Chase-BMA.

5. Simulations Results

Computer simulations are carried out in order to evaluate the performance of the HDGA and SDGA decoding algorithm. Several RS and BCH codes are considered.

The simulations were made with the default parameters outlined in the **Table 2** except where indicated. For transmission we used an AWGN/Rayleigh channel with a BPSK modulation.

5.1. HDGA versus BMA Performances

5.1.1. AWGN Channel

The **Figures 3-5** below compare the performances of HDGA and BMA for the BCH(63,30,13), BCH(63,45,7) and the RS(15,7,9) codes.

Table 2. Default parameters.

Parameter	Value
Ng (generations)	20
Ni (population)	20
Selection type	Roulette
M	2
Channel	AWGN
Modulation	BPSK
Minimum number of transmitted blocks	1000
Minimum number of residual bit errors	200

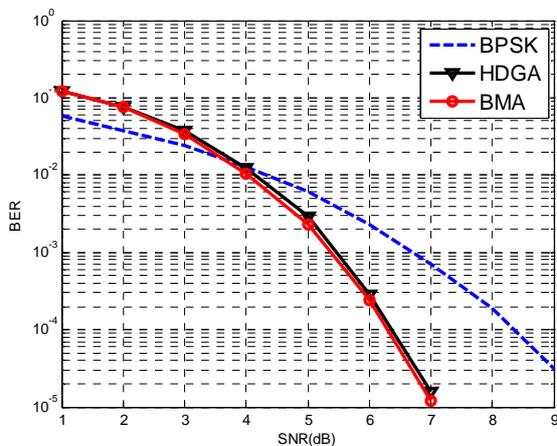


Figure 3. Performances of HDGA and BM algorithm for BCH(63,30,13).

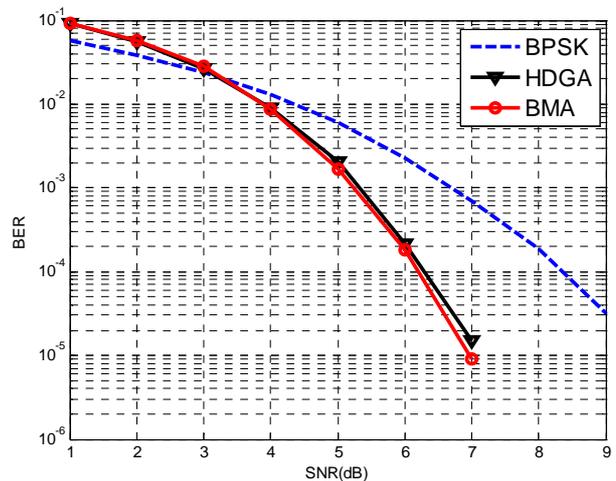


Figure 4. Performances of HDGA and BM algorithm for BCH(63,45,7).

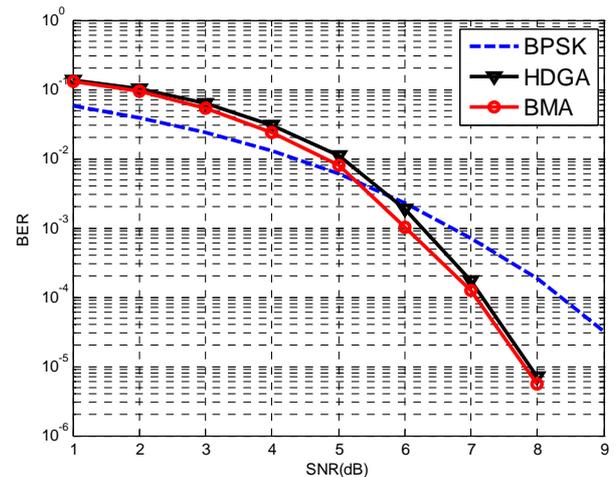


Figure 5. Performances of HDGA and BM algorithm for RS(15,7,9).

We notice in general the equality in terms of performance between the two decoders.

The performances of HDGA for RM(32,12,8) and Golay(23,12,7) codes, are shown in **Figure 6**. From the simulation results, we observe that the performances of the decoder HDGA realize about 2 dB coding gain from the non-coded BPSK curve. In addition the codes simulated in **Figure 6** can't be decoded by BMA.

5.1.2. Rayleigh Channel

HDGA has been described for the AWGN channel. The decoding is particularly desirable over channels with more severe impairments. Therefore a flat Rayleigh fading channel without side information is used for the performance investigation of HDGA.

The performances in terms of BER of HDGA and BMA for BCH(63,30,13) code are shown in **Figure 7**. From the later, we observe that the performances of both

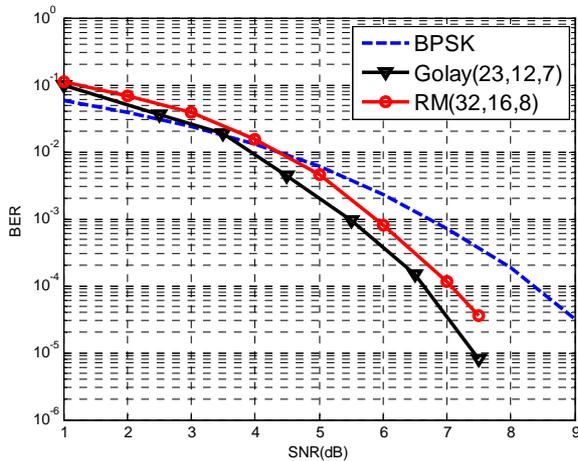


Figure 6. Performances of HDGA for Golay(23,12,7) and RM(32,16,8).

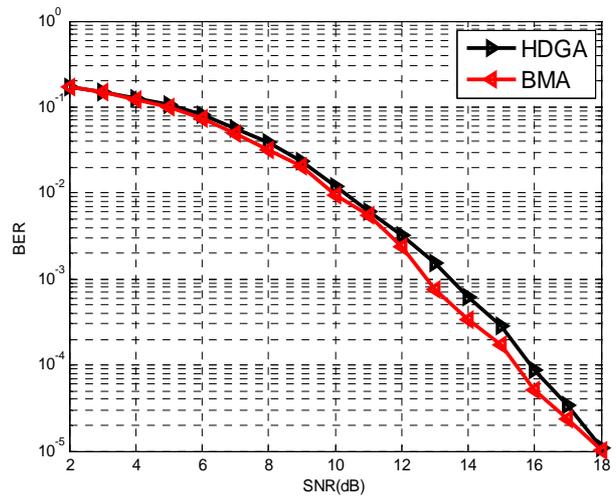


Figure 8. Performances of HDGA and BMA for RS(15,7,9).

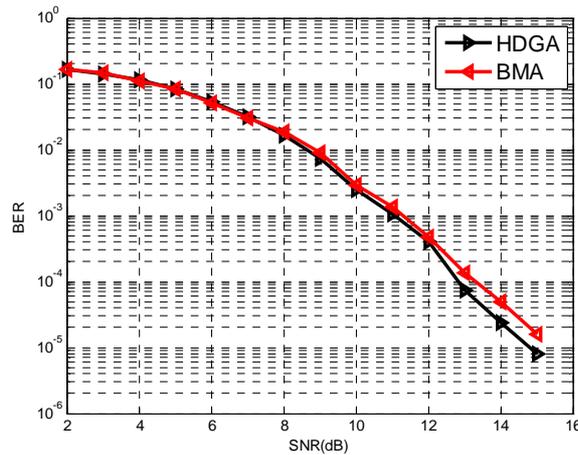


Figure 7. Performances of HDGA and BMA for BCH(63,30,13).

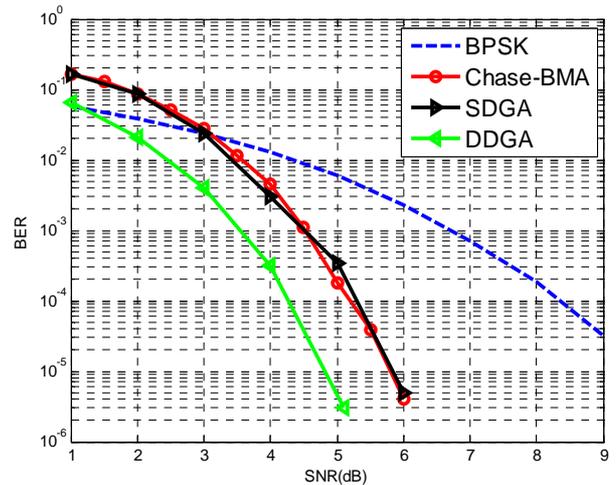


Figure 9. Performances of SDGA and Chase-BMA for BCH(63,45,7).

algorithms are similar.

Again, the Figure 8 shows that the two decoders present the same performances also for non-binary codes namely RS(15,7,9).

5.2. SDGA versus Chase-BMA Performance

In order to compare the SDGA with Chase-BMA and DDGA [13] algorithms, we carried out some simulations using transmission over AWGN channel and then over Rayleigh channel.

5.2.1. AWGN Channel

The Figures 9 and 10 compare the performance of the three decoders namely SDGA, Chase-BM and DDGA applied on the BCH(63,45,7) and RS(15,7,9) codes. We show that the performances of DDGA are better than both SDGA and Chase-BM algorithms which present relatively the same performances. This is can be explain by the fact that DDGA explore about 30,000 test vectors,

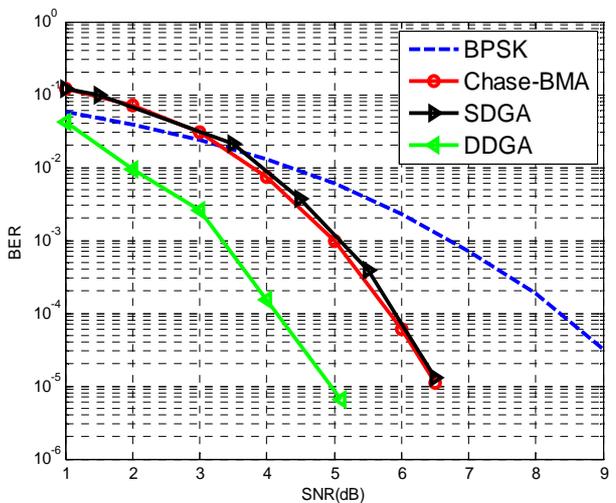


Figure 10. Performances of SDGA and Chase-BMA for RS(15,7,9).

whoever, the other decoders operate respectively 8 and 16 test vectors for BCH(63,45,7) and RS(15,7,9) codes.

The performances of SDGA for RM(32,12,8) and Golay(23,12,7) codes, are shown in **Figure 11**. From the very figure, we observe that the performance of our algorithm gives an important coding gain. Furthermore, in our algorithm, the algebraic decoding is not required as is the case of the Chase algorithm.

5.2.2. Rayleigh Channel

To evaluate more our soft decoding algorithm, we process some simulations in the Rayleigh channel, and we compare them with the performances of Chase-BMA using respectively BCH and RS(15,7,5) codes. The curves plotted in **Figures 12** and **13** show that the performance of our soft decoding algorithm SDGA are comparables with those of Chase-BMA one. So we have in the case of Rayleigh fading channel the same behavior as in the case

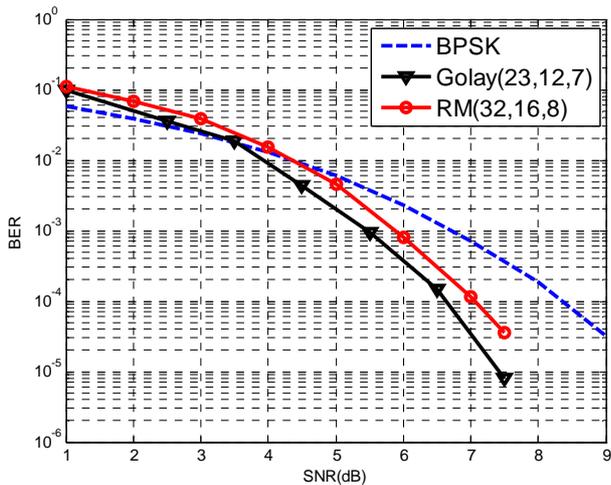


Figure 11. Performances of HDGA for Golay(23,12,7) and RM(32,16,8).

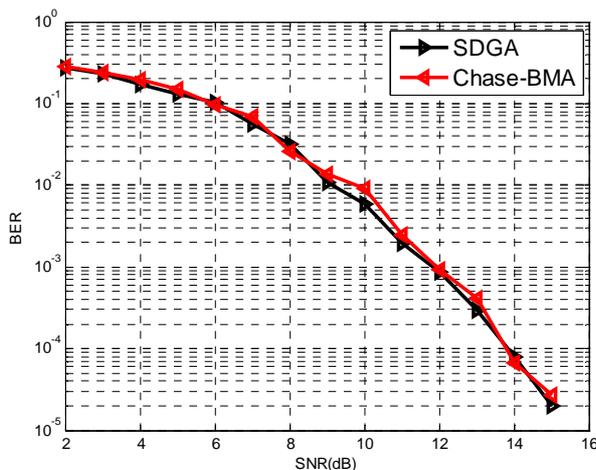


Figure 12. Performances of HDGA and BMA for BCH(63,30,13).

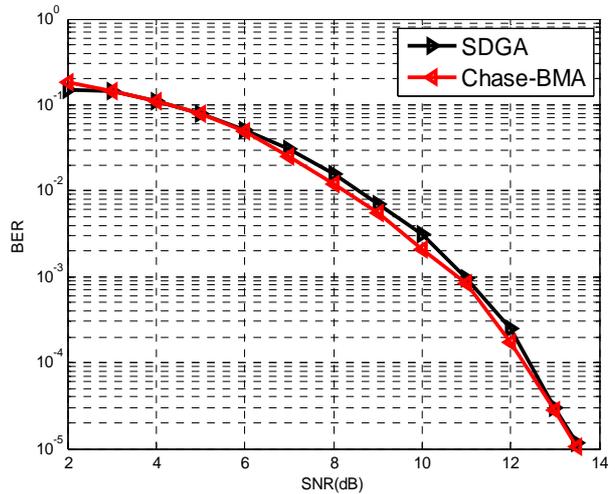


Figure 13. Performances of HDGA and BM algorithm for RS(15,7,9).

of AWGN channel.

6. Conclusions

In this paper, we have presented a hard and soft decoders based on GA. These decoders can be applied on binary and non-binary linear block codes; and constitute in fact evolutionary versions of the well known “information set decoding”. Their computational complexity is also studied and compared with BMA and Chase-BMA.

The simulations are carried out for some binary and non-binary Linear Block Codes over Gaussian and Rayleigh channels. They show that our SDGA decoder presents the same performances as Chase-BMA. In addition our SDGA decoder does not require algebraic decoding as is the case of the Chase-BM algorithm. The obtained results will open new horizons for the artificial intelligence algorithms in the coding theory field.

REFERENCES

- [1] G. C. Clark and J. B. Cain, “Error-Correction Coding for Digital Communications,” Plenum, New York, 1981.
- [2] Y. S. Han, C. R. P. Hartmann and C.-C. Chen, “Efficient Maximum-Likelihood Soft-Decision Decoding of Linear Block Codes Using Algorithm A*,” Technical Report, Syracuse University, Syracuse, 1991.
- [3] A. C. M. Cardoso and D. S. Arantes, “Genetic Decoding of Linear Block Codes,” *Congress on Evolutionary Computation*, Washington DC, 1999.
- [4] I. Shakeel, “GA-Based Soft-Decision Decoding of Block Codes,” *IEEE 17th International Conference on Telecommunications*, Doha, 4-7 April 2010, pp. 13-17.
- [5] H. S. Maini, K. G. Mehrotra, C. Mohan and S. Ranka, “Genetic Algorithms for Soft Decision Decoding of Linear Block Codes,” *Journal of Evolutionary Computation*, Vol. 2, No. 2, 1994, pp. 145-164.

- [doi:10.1162/evco.1994.2.2.145](https://doi.org/10.1162/evco.1994.2.2.145)
- [6] L. Hebbes, R. Malyan and A. Lenaghan, "Genetic Algorithms for Turbo Codes," *IEEE EUROCON*, Belgrade, 21-24 November 2005.
- [7] N. Durand, J.-M. Alliot and B. Bartolomé, "Turbo Codes Optimization Using Genetic Algorithms," *IEEE Congress on Evolutionary Computation*, Vol. 2, 1999, pp. 119-125.
- [8] P. Praveenkumar, R. Amirtharajan, K. Thenmozhi and J. B. B. Rayappan, "Regulated OFDM-Role of ECC and ANN: A Review," *Journal of Applied Sciences*, Vol. 12, 2012, pp. 301-314. [doi:10.3923/jas.2012.301.314](https://doi.org/10.3923/jas.2012.301.314)
- [9] R. Sujan, *et al.*, "Adaptive 'soft' Sliding Block Decoding of Convolutional Code Using the Artificial Neural Network," *Transactions on Emerging Telecommunications Technologies*, 2012.
- [10] J. W. H. Kao, S. M. Berber and A. Bigdeli, "A General Rate K/N Convolutional Decoder Based on Neural Networks with Stopping Criterion," *Advances in Artificial Intelligence*, Vol. 2009, 2009, Article ID 356120. [doi:10.1155/2009/356120](https://doi.org/10.1155/2009/356120)
- [11] J. Orth and S. Houghten, "Optimizing the Salmon Algorithm for the Construction of DNA Error-Correcting Codes," *IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, Paris, 11-15 April 2011, pp. 1-7.
- [12] A. Azouaoui, M. Askali and M. Belkasmi, "A Genetic Algorithm to Search of Good Double-Circulant Codes," *International Conference on Multimedia Computing and Systems (ICMCS)*, Ouarzazate, 7-9 April 2011, pp. 1-5.
- [13] A. Azouaoui, M. Belkasmi and A. Farchane, "Efficient Dual Domain Decoding of Linear Block Codes Using Genetic Algorithms," *Journal of Electrical and Computer Engineering*, Vol. 2012, 2012, Article ID 503834. [doi:10.1155/2012/503834](https://doi.org/10.1155/2012/503834)
- [14] A. Azouaoui and Dr. M. Belkasmi, "A Soft Decoding of Linear Block Codes by Genetic Algorithms," *International Conference on Intelligent Computational Systems (ICICS'2012)*, Dubai, 7-8 January 2012.
- [15] D. Chase, "A Class of Algorithms for Decoding Block Codes with Channel Measurement," *IEEE Transactions on Information Theory*, Vol. 18, 1972, pp. 170-181. [doi:10.1109/TIT.1972.1054746](https://doi.org/10.1109/TIT.1972.1054746)