IECO

Volume 7 No 2 Year 2024

■ Volume 7

■ No 2

■ Year 2024

University of Sistan and Baluchestan

Iranian Association of Electrical and Electronics Engineers

International Journal Of

Industrial Electronics Control and Optimization

In This Issue:

Research Articles:

IECO

International Journal Of
Industrial Electronics Control and Optimization

About Journal

The University of Sistan and Baluchestan entered into strategic partnership with Iranian Association of Electrical and Electronic Engineers (IAEEE) to publish the **International Journal of Industrial Electronics Control and Optimization (IECO)**. The IECO is a refereed international journal which presents to the international scientific community important results of work in these fields, whether in the form of modeling simulation, analysis, fundamental research, development, application, design or real-time implementation. The scope of IECO is broad, encompassing all aspects of Industrial Electronics, Control and Optimization.

Note: International Journal of Industrial Electronics, Control and Optimization (IECO) has qualified to **ACADEMIC RESEARCH JOURNAL (ELMI-PAJOHESHI)** status certified by the ministry of Science, Research and Technology of Iran (No. 231566/3/18 dated 1396/10/09), and is published by the University of Sistan and Baluchestan through a formal partnership (No. 952/2/1500 dated 1395/11/04) with Iranian Association of Electrical and electronic Engineers (IAEEE) in order to develop scientific and research cooperation.

Aims and Scope

International Journal of Industrial Electronics, Control and Optimization (IECO) is a Peer reviewed journal of advanced and state-of-the-art in the science and engineering of Industrial Electronics, Control and Optimization. Its Scope encompasses the applications of Industrial Electronics, power systems, control, optimization and computational intelligence for the enhancement of industrial and manufacturing system and processes. The scope of the journal include the following:

I. Industrial Electronics

- Low and high-power converters
- Renewable energy
- Drive control techniques
- Techniques for advanced power semiconductor devices
- Power quality and utility applications
- Communications
- Flexible AC Transmission Systems (FACTS)
- Control in power electronics
- Electromagnetic and thermal performance of electronic power converters
- Motion control, robotics, sensors and actuators
- Fault detection and diagnosis
- Power systems
- Factory automation, communication, and computer networks

II. Control

- Adaptive control
- Control of process systems
- Control theory
- Data processing
- Design of control systems
- Hybrid systems
- Identification and observation
- Intelligent systems
- Model-predictive control
- Optimal control

- Robust control
- Fractional order systems

III. Optimization

- Ant Colony
- Chaos Theory
- Evolutionary Computing
- Fuzzy Computing
- Hybrid Methods
- Immunological Computing
- Neuro Computing
- Particle Swarm
- Probabilistic Computing
- Rough Sets
- Wavelet

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Binary Wyner-Ziv and CEO Problem Coding Under the Hamming Distortion Criterion

Mahdi Nangir

Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran.
Corresponding author's email: nangir@tabrizu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 19-December-2023
Received in revised form: 12-April-2024
Accepted: 12-April-2024
Published online: 21-June-2024

Keywords:
Wyner-Ziv Problem,
CEO Problem,
LDGM,
LDPC,
Hamming Distortion.

ABSTRACT

In this paper, we present a practical encoding and decoding scheme for the binary Wyner-Ziv problem based on graph-based codes. Our proposed scheme uses low-density generator-matrix (LDGM) codes in lossy source coding part and low-density parity-check (LDPC) codes in syndrome generation and decoding part. Actually, we apply Bias-Propagation algorithm for lossy source coding or binary quantization and Sum-Product algorithm for syndrome-based channel decoding. Using appropriate degree distributions for LDGM codes and optimized degree distributions for LDPC codes, we will be able to achieve close rate-distortion performance to the theoretical Wyner-Ziv bound. Also, we extend our proposed scheme for presenting a practical coding scheme for the binary Chief Executive Officer (CEO) problem. In our scheme, encodig is based on binary-quantization and Slepian-Wolf coding using source-splitting technique. It is shown that, source-splitting technique is an efficient strategy for achieving non-corner points in Slepian-Wolf rate region. We show that, this technique along with iterative message-passing algorithms can be efficient for having close rate-distortion performance to the Berger-Tung inner bound of binary CEO problem for non-corner points too.

I. Introduction

Graph-based codes, i.e., codes with a graphical representation, are mostly utilized in channel coding, source coding, and also in joint source-channel coding schemes, because of their powerful performance in achieving theoretical bounds and low complexity. Turbo codes and LDPC codes with iterative decoding algorithms are able to achieve the capacity of the most known communication channel models [1], [2]. In addition, LDGM codes, which are the source-code dual of LDPC codes, have a good performance in achieving rate-distortion (RD) bound of the binary-symmetric source (BSS) [3] and the binary-erasure source (BES) [4].

These codes can be joined elaborately for solving network information theory problems such as Wyner-Ziv problem or CEO problem that are considered in this paper. Actually, quantization using LDGM codes and noiseless binning using LDPC codes play important role in these problems, if these part are efficiently designed then the total performance of structure will be efficient too.

Source coding with side information available at the decoder scenario, which is known as asymmetric Slepian-Wolf problem in lossless case [5], and the Wyner-Ziv problem in lossy case have own theoretical bounds [6], which determine the lowest achievable rates. These theoretical bounds depend on the structure of problem and its parameters. Using LDPC and LDGM codes elaborately, results in a structure which has good rate-distortion performance close to the Wyner-Ziv bound. Briefly, we show that when a good source code at the rate of R_s and a good channel code at the

rate of R_c are utilized in the compound coding structure [7], [8], the Wyner-Ziv theoretical bound can be saturated. These codes should be selected such that the following inequalities are satisfied for arbitrarily small positive values of ε_s and ε_c ,

$$1-h(D) \le R_s < 1-h(D)+\varepsilon_s, \quad (1)$$

$$1-h(D*p)-\varepsilon_c < R_c \le 1-h(D*p), \quad (2)$$

where D is acceptable distortion and p relates to the side information model that is available at the decoder of the Wyner-Ziv problem model, ε_s and ε_c relate to the source and channel coding bounds, respectively.

For the multi-terminal source coding problems, we can consider different distortion criteria to evaluate their performance. The Hamming distance measure is employed in binary and discrete spaces, the mean square error criterion is employed and applicable in continuous spaces, and the logarithmic loss can be used in both discrete and continuous spaces. Under the Hamming distortion, there is no closed form expression for the rate-distortion bound of the binary CEO problem, while under logarithmic loss, we have exact rate-distortion bound of the binary CEO problem, but under this criterion, the encoding and decoding scheme are more complex than the cases that we propose in this paper.

Different source and channel codes were utilized for solving Wyner-Ziv problem, such as Convolutional codes and Turbo codes were used in [8]. Nested lattice codes were employed for coding of the Gaussian source when an Additive White Gaussian Noise (AWGN) model is considered for the correlation channel between the side information and the information source, this model is called Quadratic Gaussian problem [9]. Polar codes were also used in [10] for the Quadratic Gaussian Wyner-Ziv problem. In addition, spatially-coupled codes using the belief-propagation guided decimation were utilized in [11]. Nested parity-check codes are introduced for coding of binary sources when the correlation channel is modeled by a BSC [12]. Moreover, universal Wyner-Ziv coding is considered in [13], [14]. We are especially interested in Low-Density Graph-based codes and BSC model for the correlation channel of Wyner-Ziv problem. Our proposed method has iterative graph-based and low-complex encoding and decoding algorithms, and realizes close rate-distortion performance to the Wyner-Ziv limit.

A natural extension of the Wyner-Ziv problem can be the multi-link scenario that is called multi-terminal lossy source coding problem. One special case of multi-terminal scenarios which has increasing application in Wireless Sensor Networks (WSN), is CEO problem. In CEO problem, several noisy observations of a single source are sensed by agents. These agents desire to encode their observations and sent them to a joint decoder with close performance to the information-theoretic bounds. This practical problem is a well-known problem in network information theory [15]. If source has binary alphabet, then we will face to the binary CEO problem which is the focus of this work.

We can consider research works on binary CEO problem in two category. The first one are works that focus on presenting tight bounds with the purpose of finding exact rate-distortion region for this problem, that is now an open problem in general case. The second one are works which try to design practical and low-complex coding schemes that achieve to existing bounds. Most of works in the second category are focused on Quadratic Gaussian CEO problem [16], [17] and [18]. Our work is in the second category for binary CEO problem.

We present a successive Wyner-Ziv coding scheme for the two-link binary CEO problem based on Source Splitting idea [19]. Using this idea, achieving an arbitrary point in rate-distortion plane of two-links case (L=2) is converted to achieving corner points in rate-distortion plane of three-links case (2L-1=3).

Our practical implemented scheme uses iterative low-complex graph-based codes that have close performance to the theoretical bounds of source coding and channel coding, we apply Bias-Propagation [20] and Sum-Product [21] for these purposes, respectively. Actually, these algorithms are the basic elements of our proposed scheme. Practical lossy source coding schemes are employed in a variety of real world applications such as multi-terminal video compression algorithms [22].

The organization of this paper is as follows, in section II, we present the problem definition and preliminaries, we review existing theoretical bounds for the binary Wyner-Ziv and the binary CEO problem, in this section. In section III, our encoding and decoding schemes are described in detail. We exhibit the results in section IV, these results show rate-distortion performance of our scheme. Finally, we conclude this paper in section V.

II. The Problem Definition and Preliminaries

In this section, the binary Wyner-Ziv problem and the binary CEO problem are reviewed. Also, the concept of binary quantization, the syndrome-based decoding and source splitting technique are presented.

A. The Binary Wyner-Ziv Problem

The lossy source coding scenario with side information that is available at the decoder is illustrated in Fig. 1, this structure is so called Wyner-Ziv problem. If *S* has binary alphabet, then we will face to a binary Wyner-Ziv problem. In general, in each solution for this problem, there exists an encoding function *f*(.), which maps *n*-vectors *S* emitting from the binary symmetric source to the compressed data Z_2 ($Z_2 = f(S)$). Moreover, a decoding function *g*(.,.) using compressed data Z_2 and side information *J* with the length of *n*, decodes *S* to $\hat{S} = g(Z_2, J)$. Note that, our intention in solving the Wyner-Ziv problem is presenting a practical low-complex coding scheme with rate-distortion performance near the Wyner-Ziv theoretical limit.

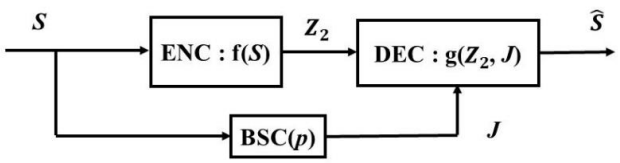


Fig. 1. The binary Wyner-Ziv problem.

According to [7], if the correlation between S and J is modeled by a BSC with the cross over probability of p and the distortion is evaluated by the Hamming distance measure, then we have the coding rate limit of (3), where D is the total distortion and D*p is the binary convolution of D and p.

$$R_{WZ} = l.c.e.\{h(D*p) - h(D), (p,0)\}, \qquad (3)$$

for $0 \le D \le p$, where $D*p=D(1-p)+p(1-D)\}$, $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$, and *l.c.e.* stands for the lower convex envelop of the term *h(D*p) - h(D)* and the point *(p,0)*.

The total distortion in this work stems from both the encoder and decoder. The encoder distortion is basically related to the quantization of the binary source *S* to a codeword, with a syndrome that is necessarily zero. At the decoder side, the distortion stems from inefficiency of the decoder which will be negligible if we apply an efficient channel error correcting code, with a rate that is smaller than the capacity of the correlation channel, in fact this inefficiency is measured by Bit Error Rate (BER) value which is equivalent with the distortion. Therefore, the total average distortion is as follows:

$$\begin{aligned} D_t &= \frac{1}{n} E\sum_{i=1}^{n} d(S_i, \hat{S}_i) \\ &= \frac{1}{n} E\sum_{i=1}^{n} d(S_i, X_i) * \frac{1}{n} E\sum_{i=1}^{n} d(X_i, \hat{S}_i) \le D, \end{aligned} \qquad (4)$$

where *d(.,.)* is the Hamming distortion measure, and *D* is acceptable distortion value.

Basically, the performance criterion in lossy source coding problems are the values of rate and distortion. They are compared with the theoretical limits depending on the configuration of problem. Because of this, we will evaluate performance of our proposed scheme from rate-distortion point of view and observe the gap between achieved points and theoretical bounds.

If the total rate of coding scheme in the nested structure equals to the difference $R_s - R_c$, then we will have,

$$R_{\text{total}} = R_s - R_c < h(D*p) - h(D) + \underbrace{\varepsilon_s + \varepsilon_c}_{\varepsilon}. \qquad (5)$$

Therefore, if R_s and R_c are chosen according to (1) and (2), then we can be close to *h(D*p) - h(D)* arbitrarily, because the term $\varepsilon = \varepsilon_s + \varepsilon_c$ can be made arbitrarily small by proper choices of source and channel codes in the compound structure. Note that, the Wyner-Ziv rate-distortion theoretical bound is formulated as (3).

B. The Binary CEO Problem

The configuration of CEO problem is depicted in two-link case in the Fig. 2. As it is seen, we have two noisy observation of a single binary source. These observations are encoded independently without any collaboration between encoders, and are sent to the joint decoder that is called CEO or fusion center. Because of lossy nature of CEO problem, if encoders collaborate with each other, then the rate-distortion region will be different from our case. Indeed, collaboration of encoders expands achievable rate region in lossy source coding scenarios.

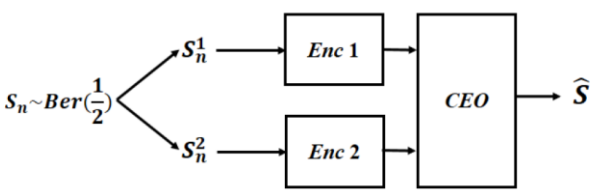


Fig. 2. Two-link binary CEO problem.

Inner and outer bounds for two-link case are provided in [23]. These bound play important role in our work, because we evaluate the performance of our scheme based on the gap value between achieved points and these bounds. In [23], binary CEO problem is reduced to a special case of binary multi-terminal source coding problem and mentioned bounds

are obtained. Inner bound is actually Berger-Tung inner bound, and the outer bound is obtained using converse proof techniques. For distortions D_1 and D_2, and the link rates R_1 and R_2, if the observation noises are $Ber(p_1)$ and $Ber(p_2)$ and $\rho = p_1 * p_2$, then the inner and outer bounds are according to (6) and (7), respectively.

$$\begin{cases} R_1^i \geq h(\rho * D_1 * D_2) - h(D_1), \\ R_2^i \geq h(\rho * D_1 * D_2) - h(D_2), \\ R_1^i + R_2^i \geq 1 + h(\rho * D_1 * D_2) - h(D_1) - h(D_2), \end{cases} \quad (6)$$

$$\begin{cases} R_1^o(D_1) \geq h[\rho * h^{-1}(1 - R_2^o(D_2))] - h(D_1), \\ R_2^o(D_2) \geq h[\rho * h^{-1}(1 - R_1^o(D_1))] - h(D_2), \\ R_1^o(D_1) + R_2^o(D_2) \geq 1 + h(\rho) - h(D_1) - h(D_2). \end{cases} \quad (7)$$

C. The Compound LDGM-LDPC Construction

The compound LDGM-LDPC construction is shown in Fig. 3. This construction consists of an LDGM code $C(G, H_1)$ nested with an LDPC code $C(G, H)$, both of these notations are related to the generator matrices and we have,

$$\begin{aligned} C(G, H_1) &= G_1 G, \\ C(G, H) &= \tilde{G} G, \end{aligned} \quad (8)$$

where G_1 and $\tilde{G}$ are the generator matrices of the codes with parity-check matrices of H_1 and $H = [H_1^T, H_2^T]^T$, respectively.

In our design, we consider the encoding procedure in two steps. It is shown in [24-25] that binary quantization using LDGM codes have a performance near rate-distortion limit for a binary symmetric source. We utilize this technique in the first step of the encoding by applying the Bias-Propagation algorithm with the compound code $C(G, H_1)$, that is an LDGM code having *n* check nodes and $m - k_1$ variable nodes.

In the second step of the encoding, we exploit the Syndrome-Based decoding method by utilizing H_2. We use the side information available at the decoder in this step. Our construction for solving the Wyner-Ziv problem is illustrated in Fig. 4.

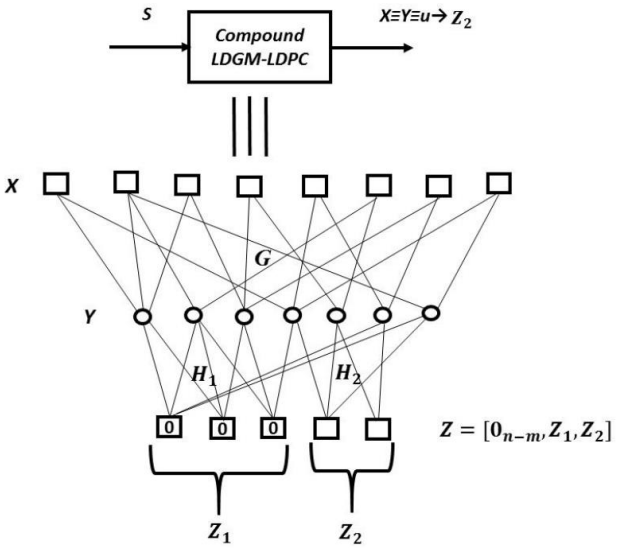


Fig. 3. The compound LDGM-LDPC code construction.

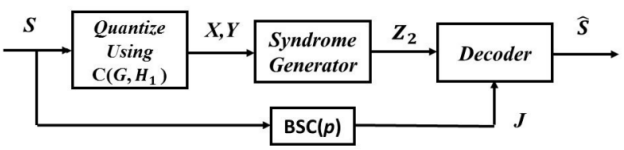


Fig. 4. Construction of the Wyner-Ziv solution.

D. The Binary Quantization

As mentioned, in our scheme much amount of distortion stems from the quantization. Since both of input and output in these quantizers are binary streams, they are called binary quantizers. Actually, the action of a binary quantizer is mapping an arbitrary $n-$vector to a codeword from a special code (for example LDGM code in our scheme) that has the minimum distance from that vector. These quantizers are the lossy source coding parts of scheme that have related rates and distortions. We can realize binary quantizations practically, using some iterative message passing algorithms, for example Bias-Propagation algorithm or Survey-Propagation algorithm. In these algorithms we will have specific compression rates and related distortions.

The Bias-Propagation Algorithm [25]:

It is shown that the Bias-Propagation algorithm has good performance in lossy compression of binary i.i.d. sources [25]. Moreover, lossy compression of binary sources was carried out using the Survey-Propagation algorithm in [26]. Suppose that an LDGM code with related bipartite graph is given, the check nodes of the graph are presented with *a,b,c,...* and the

variable nodes are presented with *i*,*j*,*k*,..., as illustrated in Fig. 4.

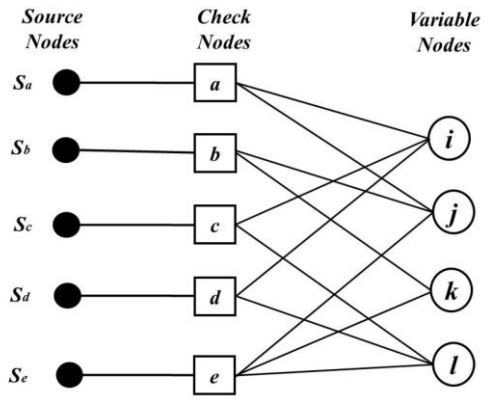


Fig. 5. Bipartite graph associated with an LDGM code for the binary Quantization.

In this algorithm, S_n is located in source nodes, which is directly connected to the check nodes. Source nodes have the bit values of the binary information source in the binary Quantization process, which are fixed and not changed during the process, while the bit values of the check nodes are changed and updated during the binary Quantization process. In each round of this algorithm, we have a constant number of iterations, the bias values in variable nodes are compared with the threshold of $0<t<1$. According to this comparison, some of the variable nodes with absolute bias values greater than t are fixed. If the bias value is positive, then the associated variable bit will be fixed to 0, and otherwise it will be fixed to 1. If there are no absolute bias values greater than t, then only a variable node with maximum absolute bias value is fixed. This procedure, continues until all the variable nodes are fixed. Finally, the compressed bit stream is obtained from the variable nodes. This algorithm is a lossy compression method. Now, we expound the update equations.

The message sent from the check node a to the variable node i in the l-th iteration of each round are as follows:

$$\phi_{a\to i}^{(l)} = \prod_{j\in\bar{N}(a)\setminus\{i\}} \theta_{j\to a}^{(l)}, \tag{9}$$

where $\bar{N}(a)$ is the set of nodes that are connected to the check node a including the source node S_a, and $\theta_{j\to a}^{(l)}$ is the message sent from the variable node j to the check node a in the l-th iteration. Moreover, the message sent from a source node S_a to the check node a is:

$$\theta_{S_a\to a}^{(l)} = (-1)^{S_a}\tanh(\gamma)\in[-1,1], \qquad \forall l \tag{10}$$

where γ is a real number depends on the LDGM code rate.

Similarly, the messages sent from the variable nodes to the check nodes in the l+1-th iteration of each round can be calculated from $\phi_{a\to i}^{(l)}$ values, according to (10),

$$\theta_{i\to a}^{(l+1)} = \frac{\prod_{b\in N(i)\setminus\{a\}}(1+\phi_{b\to i}^{(l)}) - \prod_{b\in N(i)\setminus\{a\}}(1-\phi_{b\to i}^{(l)})}{\prod_{b\in N(i)\setminus\{a\}}(1+\phi_{b\to i}^{(l)}) + \prod_{b\in N(i)\setminus\{a\}}(1-\phi_{b\to i}^{(l)})}, \tag{11}$$

where $N(i)$ is the set of check nodes that are connected to the variable node i. Initial bias values are,

$$\theta(i) = \frac{|N_0(i)|-|N_1(i)|}{|N_0(i)|+|N_1(i)|}, \tag{12}$$

where,

$$N_h(i) = \{a\in N(i)\,|\,(\forall j\in\bar{N}(a)\setminus\{i\}, v_j\neq *)\& \\ (\sum_{j\in\bar{N}(a)\setminus\{i\}} v_j = h, \bmod 2)\}, \qquad h=0,1$$

$$N_*(i) = \{a\in N(i)\,|\,(\exists j\in\bar{N}(a)\setminus\{i\}, v_j\neq *)\}. \tag{13}$$

where v_j is the j-th bit of variable nodes and its initial value is chosen randomly. Finally, at the end of each round, the bias values are calculated according to (14), then they are compared with the threshold value t.

$$\theta_i = \frac{\prod_{b\in N(i)}(1+\phi_{b\to i}^{(\hat{l})}) - \prod_{b\in N(i)}(1-\phi_{b\to i}^{(\hat{l})})}{\prod_{b\in N(i)}(1+\phi_{b\to i}^{(\hat{l})}) + \prod_{b\in N(i)}(1-\phi_{b\to i}^{(\hat{l})})}. \tag{14}$$

When there are some cycles of length 4 in the graph of codes, inevitably; we can equipped the Bias-Propagation algorithm with damping operation, it means that the equation (14) will be changed according to [26].

E. The Syndrome-Based Decoding Scheme

The parity and the syndrome approaches are efficient lossy source code design methods based on the channel coding methods [27]. In these methods, the encoder sends a subset of parity or syndrome bits to the decoder as lossy compressed data. Our proposed scheme is based on syndrome approach, that is more applicable in lossless source coding problems with side information at the decoder, a binary sequence $\hat{S}$ is decoded utilizing the received syndrome and side information J with the length of n [28]. This simple coding scheme is used for the asymmetric Slepian-Wolf structure

[29]. Inspired by this syndrome decoding idea, if we set some of syndrome bits equal to zero, and send the only determined non-zero part to the decoder, we will have a lossy coding scheme that can be employed for solving the Wyner-Ziv problem.

In this work, we use these original ideas with iterative message-passing algorithms on sparse graph-based codes for obtaining a practical solution of the binary Wyner-Ziv problem and the binary CEO problem.

The Sum-Product Algorithm [29]:

It is demonstrated that the Sum-Product algorithm from the family of iterative LDPC decoders has good performance in channel error correction and generally in syndrome-based decoding schemes. We briefly restate its iterative procedure and related parameters [29]. Similar to any iterative message passing algorithm, there are messages which are passed between variable nodes and check nodes according to Fig. 6.

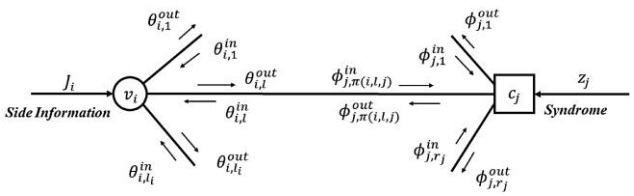


Fig. 6. Messages in Sum-Product algorithm over an LDPC code.

In each iteration, we have the following message-passing equations, the initial values of LLRs for each variable node in the corresponding bipartite graph of LDPC code are set according to:

$$\theta_{i,0} = \log \frac{\Pr[x_i = 0 \,|J_i]}{\Pr[x_i = 1 \,|J_i]} = (1 - 2J_i) \log \frac{1-p}{p} \qquad (15)$$

where J_i is the i-th bit of the side information J. Variable node update equation is as follows:

$$\theta_{i,l}^{out} = \theta_{i,0} + \sum_{j=1, j\neq l}^{l_i} \theta_{i,j}^{in}, \qquad l = 1,2,...,l_i \qquad (16)$$

where l_i is the degree of i-th variable node. The second subscript in messages shows the number of edge which they are passed over it. The initial values of $\theta_{i,j}^{in}$s are 0. Next, it is assigned $\theta_{i,l}^{out} = \theta_{j,\pi(i,l,j)}^{in}$, for all of edges according to the connections. Check update equation is as follows:

$$\tanh(\frac{\phi_{j,r}^{out}}{2}) = (1 - 2z_j) \prod_{i=1, i\neq r}^{r_j} \tanh(\frac{\phi_{j,i}^{in}}{2}), r = 1,2,...,r_j \quad (17)$$

where r_j is the degree of j-th check node, and z_j is the j-th bit of total syndrome $Z = [0_{n-m}\ Z_1\ Z_2]$. Now, it is assigned $\theta_{i,l}^{in} = \theta_{j,\pi(i,l,j)}^{out}$ and this process is run for a defined number of iterations. Finally, we will have the following decision making rule:

$$\hat{x}_i = \begin{cases} 0 & \theta_{i,0} + \sum_{j=1}^{l_i} \theta_{i,j}^{in} \geq 0 \\ 1 & \theta_{i,0} + \sum_{j=1}^{l_i} \theta_{i,j}^{in} < 0 \end{cases}, \quad \forall i. \qquad (18)$$

Generally, the complexity order of all message-passing algorithms is $O(nc)$; where, n is the number of nodes and c is the complexity of one complete passing. Note that optimal encoding and decoding schemes have an exponential complexity order with the block length of n, while message passing algorithms which are sub-optimal reduce complexity to the linear order.

F. Source Splitting Technique.

If a solution for the Wyner-Ziv problem is successively extended to a solution for the CEO problem in a straightforward manner, then only corner points in rate region can be achieved. In other words, successive Wyner-Ziv scheme for the CEO problem cannot be efficient by itself except for the corner points of rate region. For achieving non-corner points of rate region, we exploit from source splitting technique that is firstly introduced in [19]. Using this technique, we are able to achieve non-corner points with successive refinement. The details of this technique is discussed in the next section.

III. The Encoding and the Decoding Schemes

We present the processes that are done on the information sources, from the beginning to the end of proposed scheme, including both of encoder and decoder, in detail. We discuss about the Wyner-Ziv problem and the CEO problem separately, in the following.

A. The Wyner-Ziv Problem

In this subsection, we describe our proposed encoding and decoding schemes with more details. Briefly, we apply the Bias-Propagation algorithm for binary quantization of the

source sequence using an LDGM code. Thereafter, we obtain syndrome from quantized stream and send it to the decoder along with side information. At the decoder, the decoded stream is obtained using the Sum-Product algorithm. The formulation of these two iterative algorithms including their check update, variable update equations and related parameters are presented previously.

The Encoding Algorithm:

In our design, we consider the encoding procedure in two steps. It is shown in [3] that binary quantization using LDGM codes have a performance near rate-distortion limit for a binary symmetric source. We utilize this technique in the first step of the encoding by applying the Bias-Propagation algorithm with the compound code $C(G,H_1)$, that is an LDGM code having *n* check nodes and $m-k_1$ variable nodes. In the second step of the encoding, we exploit the Syndrome-Based decoding method by utilizing H_2. We use the side information available at the decoder in this step. Our construction for solving the Wyner-Ziv problem is illustrated in Fig. 7.

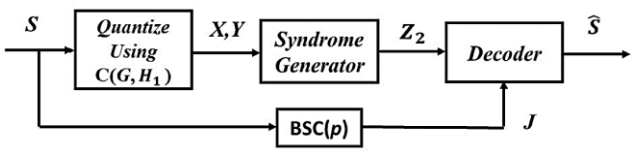


Fig. 7. Construction of the Wyner-Ziv solution.

A BSS is considered, i.e., $\Pr\{S_i=0\}=\Pr\{S_i=1\}=0.5$. As earlier mentioned, compression procedure of S_n for obtaining the syndrome consists of two steps.

Step 1: The source sequence S_n is quantized to the codeword $X_n \in C(G,H_1)$, using the compound code $C(G,H_1)$. From this step, $Y_m = u_{m-k_1} \times G_1$ is obtained, where u_{m-k_1} is information bits that is replaced in variable nodes after the quantization process, and G_1 is the generator matrix associated with H_1. Moreover, we have $X_n = Y_m \times G_{m\times n}$. This quantization is implemented by the Bias-Propagation algorithm.

Step 2: After the quantization, the syndrome of Y_m is obtained using H_2. Actually, the syndrome $Z_2 = Y_m H_2^T$, is sent to the decoder as compressed sequence. Note that the length of Z_2 is k_2, because the dimension of H_2 is $k_2 \times m$.

Note that the order of the binary Quantization and syndrome generation parts cannot be changed with each other because basically Quantization is a source encoding scheme, while the syndrome generation is used for the channel decoding.

The Decoding Algorithm:

The decoder receives side information $J = S \oplus Ber(p)$ and syndrome Z_2, then it finds nearest sequence to *J* in the coset with syndrome Z_2 using the sum-product algorithm that is an efficient sub-optimal LDPC decoder. This algorithm is performed by an LDPC code with the generator matrix $C(G,H)$. The formulations of this algorithm is presented, previously.

B. The Main Result on the Wyner-Ziv Solution

Our practical implemented coding scheme for binary Wyner-Ziv problem is based on an important theorem which was proven in [7]. In the following, we restate it.

Theorem. The compound LDGM-LDPC structure is able to achieve theoretical Wyner-Ziv bound if the following two conditions are satisfied:

1. An efficient lossy source coding algorithm is used for the Quantization, which achieves rate-distortion limit.

2. An efficient channel decoding algorithm is applied in the Syndrome-Decoder step, with low BER.

There are two important limitations in the proposed scheme which stem from the above conditions and should be considered. First one is that, for any quantization rate R_1, the resulting distortion d_1 should be greater than *D*, equivalently $R_1 = 1-h(D) > 1-h(d_1)$. Another limitation is related to the syndrome-decoding is that, the decoder rate R_2 should be smaller than the capacity of correlation channel model between quantized sequence X_n and side information J_n, which is really a BSC($d_1 * p$). In other words if $R_2 < 1-h(d_1 * p)$, then the distortion d_2 or BER in syndrome-decoding step can be arbitrary small.

These two limitations are actually theoretical. Furthermore, there are some practical limitations that should be considered in implementation and block length choosing. We express one of the practical limitations by the following lemma.

Lemma. In selecting block lengths, if $k_1 + k_2 < m < 2k_1 + k_2$ is satisfied, then the existence of at least one matrix for top layer *G* in the compound structure

is guaranteed, given $C(G,H_1)$, $C(G,H_2)$ and matrix H in the bottom layer.

Proof. The inequality $k_1+k_2<m$ is trivial due to the dimension of matrix H, we demonstrate that $m<2k_1+k_2$. Matrix G has mn unknown elements. Total number of known elements are $n(m-\{k\_1\})+n(m-\{k\_1\}-\{k\_2\})$, which are obtained from (8). Note that in these equations, G_1 and $\tilde{G}$ are known, because H is given. Therefore, if $mn>n(m-\{k\_1\})+n(m-\{k\_1\}-\{k\_2\})$, then there is at least one choice for G such that (8) is satisfied for given $C(G,H_1)$, $C(G,H_2)$ and matrix H. Therefore, we will have $m<2k_1+k_2$ and this completes the proof.

Actually, the Bias-Propagation and Sum-Product algorithms are able to satisfy conditions (1) and (2) in the above theorem, respectively, and because of this, our proposed scheme works efficiently close to the existing theoretical bounds. In general, iterative message-passing algorithms are sub-optimal and there is a gap between theoretical bounds and rate-distortion performance of these algorithms. This gap can be decreased using following methods:

1. Increasing length of blocks or number of iterations in algorithms.
2. Increasing threshold value for more accuracy.
3. Using Survey-Propagation algorithm instead of Bias-Propagation, knowing that the complexity increases.
4. Using exact ODDs for each rate.

C. The Binary CEO Problem

In this section our designed construction for the encoding and decoding side of binary CEO problem is described. For the sake of simplicity, only two-link case is considered in this paper. Our design can be extended to *L*-link case, easily.

The Encoding:

In the Fig. 8, the encoding process is depicted. As it is seen, first of all, agents quantize two noisy observations to X_n and Y_n, the rates of these quantizers are $R_{11}=\frac{k_1}{n}$ and $R_{21}=\frac{k_2}{n}$ respectively. These binary-quantization steps are done using LDGM codes $(G_1)_{k_1\times n}$ and $(G_2)_{k_2\times n}$, respectively.

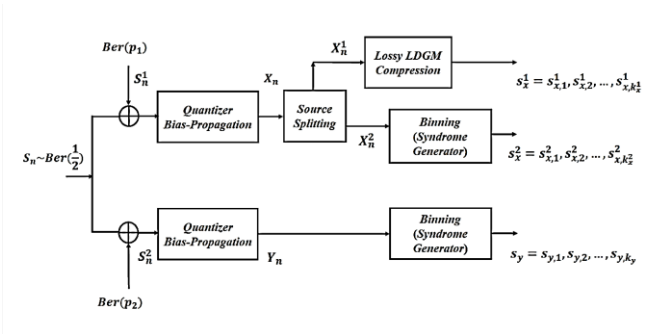


Fig. 8. The Encoding Scheme of the two-link Binary CEO Problem.

After that we split one of the quantized sources (for example X_n) to the two sources (X_n^1 and X_n^2), we use the following rule in source splitting,

$$\left.\begin{array}{l} X_i^1=\min(X_i,T) \\ X_i^2=\max(X_i,T)-T \end{array}\right\} \Rightarrow \left\{\begin{array}{l} X_i=X_i^1+X_i^2 \\ X_n \leftrightarrow (X_n^1,X_n^2) \end{array}\right. \qquad (19)$$

If we suppose that the alphabet of X_n has been mapped to $\{0,1,...,Q-1\}$ with $Q=2^{k_1}$, then $T\in\{1,2,...,Q-2\}$. So we will have,

$$X_n^1\in\{0,1,...,T\};X_n^2\in\{0,1,...,Q-1-T\} \qquad (20)$$

We suppose that $X_n\oplus X_n^l=Ber(\alpha_l);\ l=1,2$. The parameters α_1 and α_2 can be calculated from the source splitting rule.

Thereafter, we have syndrome generating or Slepian-Wolf encoding step. In this step, X_n^1 is compressed to the s_x^1 with the length of k_x^1 using an LDGM code with the rate R_{12} and resulted distortion d_{12}. Note that lossless compressions schemes like Entropy Coding (for example Huffman coding or Arithmetic coding) does not work properly in this step, because of their error propagation characteristic and their lossless essence. Parallel to this compression, syndromes s_x^2 and s_y is calculated from X_n^2 and Y_n respectively, using appropriate LDPC codes H_x and H_y.

$$(s_x^2)_{k_x^2}=X_n^2H_x^T,\quad (s_y)_{k_y}=Y_nH_y^T \qquad (21)$$

Here, we can exploit from parity-based methods as well as syndrome-based methods [30]. Therefore, these three compressed sequences (s_x^1, s_x^2 and s_y) are sent to the CEO decoder. Total rates in links are $R_1 = \frac{k_{x_1} + k_{x_2}}{n}$, $R_2 = \frac{k_y}{n}$. Using source splitting idea, two-link structure is converted to the equivalent three-link case. We have the following rate-distortion relations in binary-quantization steps,

$$\begin{aligned} R_{11} &= \tfrac{k_1}{n} = 1 - h(d_{11}) + \varepsilon_{11}, \quad d_{11} = \tfrac{1}{n}\left(\sum_{i=1}^{n} X_i \oplus S_i^1\right) \\ R_{21} &= \tfrac{k_2}{n} = 1 - h(d_{21}) + \varepsilon_{21}, \quad d_{21} = \tfrac{1}{n}\left(\sum_{i=1}^{n} Y_i \oplus S_i^2\right) \end{aligned} \quad (22)$$

The Decoding:

In the Fig. 9, the decoding process is depicted. As it is seen, the CEO decoder has a successive formation. In the decoder, first s_x^1 is decompressed to the $\hat{X}_1$, if Bias-Propagation or Survey-Propagation algorithm is applied in the lossy compressor using an LDGM code, then this decompressor will be a simple matrix multiplying. After that, $\hat{X}_1$ is used as side information in the Sum-Product algorithm to decode $\hat{Y}$ from the input s_y. Finally, $\hat{Y}$ and $\hat{X}_1$ are used as side information in the Sum-Product algorithm to decode $\hat{X}_2$ from the input s_x^2. By summing $\hat{X}_1$ and $\hat{X}_2$, $\hat{X}$ is obtained.

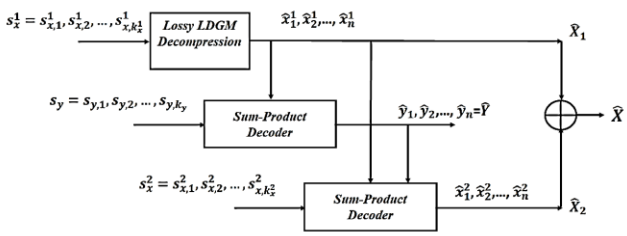


Fig. 9. The Decoding Scheme of the two-link Binary CEO Problem.

The total average hamming distortions are calculated as the following,

$$D_{t,l} = D_l * d_{l,1}, \qquad l = 1,2 \qquad (23)$$

where, $D_l \equiv Ber_l$ for $l=1,2$, $Ber(D_1) = X_n \oplus \hat{X}$ and $Ber(D_2) = Y_n \oplus \hat{Y}$, if appropriate channel decoding schemes are used, then we will have $D_1, D_2 \to 0$.

Similar to equations (22), we will have,

$$R_{12} = 1 - h(d_{12}) + \varepsilon_{12} \quad , \mathrm{BSC}(d_{12}) = \hat{X}_n^1 \oplus X_n^1. \qquad (24)$$

Parameters in our scheme should be designed in such a way that the following relations are satisfied. In other words, if in the first Sum-Product decoder we have,

$$\frac{n - k_v}{n} = 1 - h(A * \alpha_1) - \varepsilon_{22} \Rightarrow \frac{k_v}{n} = h(A * \alpha_1) + \varepsilon_{22} \qquad (25)$$

then the Bit Error Rate of decoding can be arbitrary small, where $A \triangleq d_{11} * \rho * d_{21}$, and $BSC(A) = X_n \oplus Y_n$. Similarly, in the second Sum-Product decoder we will have,

$$\frac{n - k_u^2}{n} = 1 - h(D_2) - \varepsilon_{21} \Rightarrow \frac{k_u^2}{n} = h(D_2) + \varepsilon_{22} \qquad (26)$$

because we have $\hat{Y}_n$ as side information in this decoder with distortion D_2, and $\hat{Y}_n \to \hat{X}_n \to \hat{X}_n^2$ forms a markov chain. Also, $\hat{X}_n^1$ is available at this decoder as side information.

IV. Simulation and Numerical Results and Performance Analysis

A. Results for the Binary Wyner-Ziv Problem

In this section, we present some simulation results which demonstrate our coding scheme performance for some special rates. We use the degree distributions which are optimized over BSC using density evolution technique. Our results are for two cases of low (p=0.3) and high (p=0.1) correlation between S_n and J_n.

In implementing the message-passing algorithms, we use time sharing between the points (p,0) and $(d_c + \varepsilon, R_c)$, where $(d_c + \varepsilon, R_c)$ is an achieved rate-distortion point. Obviously, this technique is efficient for rates in the linear part of Wyner-Ziv bound.

The block lengths which is employed in simulation, related rates and distortion values are presented in Tables 1 and 2, where d_1 and R_1 are distortion and rate of Binary-Quantization, d_2 and R_2 are distortion and rate of Syndrome-Decoding, D_t and R are the total distortion and rate, respectively. For more intuition, we represent D_{WZ} (Wyner-Ziv limit of distortion in rate R). We have $R_1 = \frac{m - k_1}{n}$, $R_2 = \frac{m - k_1 - k_2}{n}$, and $R = \frac{k_2}{n}$. Degree distributions which are utilized for producing LDGM and LDPC codes matrices are from [24] and [31]. We generate

matrices according to degree distribution ($\rho(x)$, $\lambda(x)$), randomly.

We use some approximations for producing matrices from ODDs, in such a way that for some rates with unavailable ODDs, we round them to the rates with available ODDs and use them instead. We run the encoding and decoding process for 100 randomly generated source sequence S_n. We run 25 iterations in each round of the Bias-Propagation algorithm, we set t =0.8 and $\gamma \approx 2R_{\text{LDGM}}$ in our simulations. Moreover, we run 80-100 iterations in the Sum-Product algorithm.

In Figures 10 and 11, rate-distortion performance of our proposed scheme for two different side information models are illustrated. As it is seen in these figures, rate-distortion performance of our scheme is very close to the theoretical Wyner-Ziv limit. This ability of compound codes has been proved in [7]. It is obvious that when time sharing is utilized, it is possible of being more close to the Wyner-Ziv limit for small rates in comparison with the regions where time sharing is not utilized (high rates).

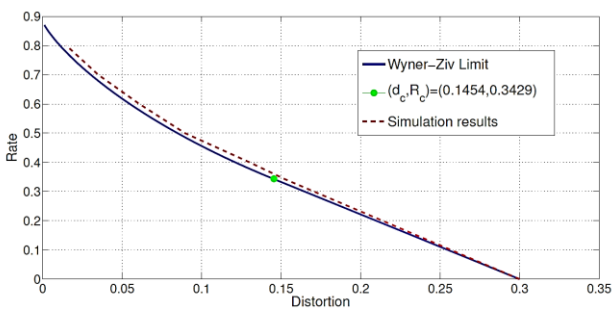


Fig. 10. Rate-distortion performance of the proposed scheme for p=0.3.

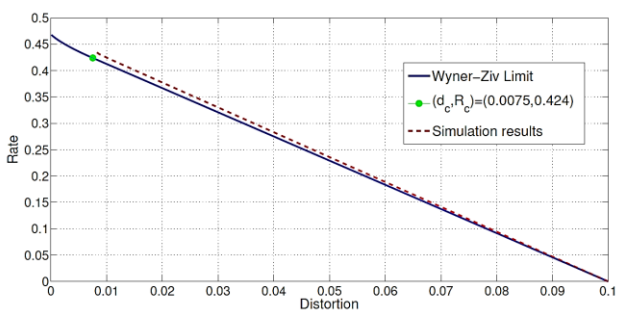


Fig. 11. Rate-distortion performance of the proposed scheme for p=0.1.

In these figures, the border points on Wyner-Ziv limit for p=0.3 and p=0.1 are $d_c = 0.1454$ and $d_c = 0.0075$, respectively. These points separate the Wyner-Ziv limit curve into two parts, rates in the high rate region are in the formation of $h(p*D_t)-h(D_t)$ for distortion D_t and for low rates the curve has linear formation.

B. Results for the Binary CEO Problem

Here, we present simulation and numerical results and related parameters that are used in implementation of iterative message passing algorithms. We apply Bias-Popagation algorithm in binary-quantization and lossy compression blocks using LDGM codes. Also, we apply Sum-Product decoder using LDPC codes in CEO decoders. Degree distributions for producing low-density matrices are from [25] and [31], these degree distributions are optimized over Binary Symmetric Channel using Density Evolution technique.

We set $(p_1, p_2) = (0.0625, 0.1)$, therefore we will have $\rho = p_1 * p_2 = 0.15$. Our results are for

$S_n \sim Ber(\frac{1}{2}), n = 10^4$, and $D_{t,1} = D_{t,2} = 0.005$. Theoretical bounds related to these parameters are shown in Fig. 12 according to [23].

Resulted values for parameters is presented in Tables 3 and 4 in details. Note that, in corner points, the binary CEO problem is reduced to a single binary Wyner-Ziv problem; in this case, one of the links should be lossless transmission link.

If we change the rule of two links with each other, i.e. $X \leftrightarrow Y$, then we achieve three new points that are symmetric with three achieved points that are described in tables. In other words we can achieve the points (R_1, R_2) =(0.85,0.71), (R_1, R_2) =(0.9,0.66) and (R_1, R_2) =(1,0.59) in a similar way. These six points obtained from simulation along with outer bound, Berger-Tung inner bound and Slepian-Wolf bound are shown in Fig. 12. Note that Slepian-Wolf bound is related to the lossless case, i.e. when we have $D_{t,1} = D_{t,2} = 0$.

The reason of the near the existing theoretical bounds performance returns to the intelligently selecting the block lengths in the parts of the binary Quantization and the syndrome decoding. In this paper, we divide the theoretical rate-distortion bound expression of the binary Wyner-Ziv and the binary CEO problem in such a way that a point-to-point source coding bound expression and a binary channel capacity expression appear. Then, in the design procedure, we choose block lengths such that those bound be achievable. Finally, by combining them, the general theoretical bounds of the binary Wyner-Ziv and the binary CEO problem will be achievable.

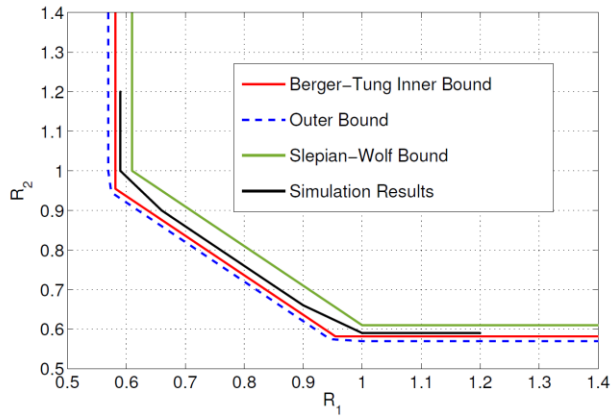


Fig. 12. Rate-distortion performance comparison of the proposed scheme.

Calculating Parameter α_1: We have the following relations in source splitting part of our proposed scheme,

$$R_y = H(Y|X_1) \tag{27}$$

$$R_x = \underbrace{H(X_1)}_{R_{X_1}} + \underbrace{H(X_2|X_1,Y)}_{R_{X_2}} = H(X,Y) - R_y \tag{28}$$

As it is mentioned $Ber(\alpha_1) = X_n \oplus X_n^1$. In the continue, we calculate the relation between α_1 and T using Gaussian approximation. For more intuition, we present some special cases as the following,

$$\begin{aligned} &T = 2 \Rightarrow \alpha_1 \approx 0.5 \\ &T = \tfrac{Q}{2} \Rightarrow \alpha_1 \approx 0.25 \\ &T = Q - 2 \Rightarrow \alpha_1 \approx 0 \end{aligned} \tag{29}$$

We can suppose $k_1 \approx n$. So, X_n^1 can possess all binary streams with length n. There are two possible states for X_n^1,

$$\begin{aligned} &X_n \le T \Rightarrow X_n^1 = X_n \Rightarrow X_n^1 \oplus X_n = 0 \\ &X_n > T \Rightarrow X_n^1 = T \Rightarrow X_n^1 \oplus X_n \neq 0 \end{aligned} \tag{30}$$

Therefore, α_1 is approximately average of term $d_H(X_n^1, X_n)$ over all of $X_n > T$, where $d_H(.,.)$ represents Hamming distance between two streams. So with a good approximation, the minimum possible value for α_1 is as follows,

$$\alpha_1 \approx \frac{1\binom{n}{1} + 2\binom{n}{2} + \ldots + k\binom{n}{k}}{n \times 2^n} \tag{31}$$

the integer k is selected such that we have,

$$\binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{k} \approx 2^n - (T+1) \tag{32}$$

the number of X_n s such that $X_n > T$. Because $i\binom{n}{i} = n\binom{n-1}{i-1}$, we will have,

$$\alpha_1 \approx \frac{\binom{n-1}{0} + \binom{n-1}{1} + \ldots + \binom{n-1}{k-1}}{2^n} \tag{33}$$

If $N \triangleq n-1$ and $K \triangleq k-1$, then,

$$2\alpha_1 \approx \sum_{i=0}^{K} \binom{N}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{N-i}. \tag{34}$$

Using Gaussian approximation $\binom{n}{i} p^i (1-p)^{n-i} \approx \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(i-np)^2}{2np(1-p)}}$, we can calculate K for a known α_1 and vice versa from (34). From (34) and Gaussian approximation, we will have,

$$\sum_{i=0}^{K} \binom{N}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{N-i} \approx \int_{-\infty}^{B} \frac{1}{\sqrt{0.5\pi N}} e^{-\frac{(t-\frac{N}{2})^2}{0.5N}} dt \tag{35}$$

where $B = \frac{20\sigma}{N} \times K + (-10\sigma + \mu)$, which $\sigma = \frac{\sqrt{N}}{2}$ and $\mu = \frac{N}{2}$. For standard Gaussian distribution, Q and G functions is defined as follows, and their table is available for required calculations,

$$Q(z) = \int_{z}^{\infty} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt = 1 - \int_{-\infty}^{z} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt = 1 - G(z), \tag{36}$$

where, $z \sim N(0,1)$ is standard normal distribution. Using (36), we can write (35) as follows,

$$G\left(z=\frac{B-\mu}{\sigma}\right)=1-Q\left(z=\frac{B-\mu}{\sigma}\right) \qquad (37)$$

The results of this work are based on these calculations. There are some hints in our proposed scheme that should be noted, they are listed in the following,

1. The degree distributions which are optimal in channel coding aren't necessarily optimal in source coding problems, but they have better performance, compared to regular degree distributions.

2. There is no necessity that the side information should be used only in variable nodes of the LDPC code; actually, the side information can be used in different ways; for example, it can be multiplied or added with received data and etc (e.g. other computational or combinatorial functions).

3. The output of Binary Quantization can be supposed to be uniform, this uniformity causes the calculating of α_1 has more accuracy. Also, the parameter α_2 don't affect to our design and calculations.

4. All of iterative message-passing algorithms are sub-optimal (with linear complexity), because optimal solution is really exhaustive search for finding nearest codeword (with exponential complexity), these iterative algorithm are done using a low-density graph-based code.

5. Source splitting idea for designing a practical coding scheme for the binary CEO problem results in a successive decoding scheme, this successive decoding scheme is in contrast with joint decoding scheme.

V. CONCLUSIONS

In this paper, we implemented a practical coding scheme for solving Wyner-Ziv problem, employing compound LDGM-LDPC construction, which is really a nested coding structure. When efficient source and channel coding schemes with low complexity are elaborately used in this construction, this structure will have the capability of coming close to the Wyner-Ziv limit with any arbitrary precision. We applied Bias-Propagation algorithms for binary quantization using LDGM codes and Sum-Product algorithm for syndrome-decoding using LDPC codes. Employing these algorithms and time sharing technique, the numerical and simulation results confirm close rate-distortion performance of proposed scheme to the Wyner-Ziv theoretical limit.

Furthermore, we proposed an encoding scheme based on binary quantization and source splitting idea in the Slepian-Wolf coding part for the binary CEO problem, in this paper. Our CEO decoder has successive structure and decoded sequence in each link is used as side information in the next links. All of binary quantization, lossy compressors and link decoders are graph based and utilize iterative message-passing algorithms. Because of good performance of these major components, the total rate-distortion performance of our proposed scheme is close to the theoretical bounds of CEO problem.

REFERENCES

[1] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," IEEE Transactions on Communications, vol. 44, no. 10, pp. 1261–1271, Oct 1996.
[2] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity approaching irregular low-density parity-check codes," IEEE Transactions on Information Theory, vol. 47, no. 2, pp. 619–637, Feb 2001.
[3] Z. Sun, M. Shao, J. Chen, K. M. Wong, and X. Wu, "Achieving the rate-distortion bound with low-density generator matrix codes," IEEE Transactions on Communications, vol. 58, no. 6, pp. 1643–1653, June 2010.
[4] E. Martinian and J. Yedidia, "Iterative quantization using codes on graphs," in Allerton Conference on Control, Computing, and Communication, October 2003.
[5] D. Slepian, and J. Wolf, "Noiseless coding of correlated information sources", IEEE Transactions on information Theory, vol. 19, no. 4, pp. 471-480, 1973.
[6] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Transactions on Information Theory, vol. 22, no. 1, pp. 1–10, Jan 1976.
[7] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information," IEEE Transactions on Information Theory, vol. 55, no. 3, pp. 1061–1079, March 2009.
[8] X. Chen, Z. Lin, C. Chen, and X. Deng, "Low-Delay Digital and Hybrid Digital/Analog Schemes for Wyner-Ziv Problem over Gaussian Broadcast Channel", IEEE Transactions on Vehicular Technology, 2023.
[9] R. Zamir, Lattice coding for signals and networks. Cambridge: Cambridge University Press, 2014.
[10] S. Eghbalian-Arani and H. Behroozi, "Polar codes for a quadratic gaussian wyner-ziv problem," in Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on, Aug 2013, pp. 1–5.
[11] S. Kumar, A. Vem, K. Narayanan, and H. D. Pfister, "Spatially-coupled codes for side-information problems," in IEEE International Symposium on Information Theory, June 2014, pp. 516–520.
[12] N. Zhang, and M. Tao, "An adaptive distributed source coding design for distributed learning", In 13th International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1-5, 2021, October.
[13] S. Watanabe and S. Kuzuoka, "Universal wyner-ziv coding for distortion constrained general side information,"

IEEE Transactions on Information Theory, vol. 60, no. 12, pp. 7568–7583, Dec 2014.
[14] E. Dupraz, A. Roumy, and M. Kieffer, “Universal wyner-ziv coding for gaussian sources,” in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, May 2013, pp. 5132–5135.
[15] A. El Gamal and Y.-H. Kim, Network information theory. Cambridge university press, 2011.
[16] M. Pathegama, and A. Barg, "Smoothing of binary codes, uniform distributions, and applications", Entropy, vol. 25, no. 11, 2023.
[17] J. Chen and T. Berger, “Successive wyner-ziv coding scheme and its application to the quadratic gaussian ceo problem,” IEEE Transactions on Information Theory, vol. 54, no. 4, pp. 1586–1603, April 2008.
[18] H. Behroozi and M. R. Soleymani, “Distortion sum-rate performance of successive coding strategy in quadratic gaussian ceo problem,” IEEE Transactions on Wireless Communications, vol. 6, no. 12, pp. 4361–4365, December 2007.
[19] V. Kostina, and B. Hassibi, "The CEO problem with inter-block memory", IEEE Transactions on Information Theory, vol. 67, no. 12, pp. 7752-7768, 2021.
[20] T. Filler and J. Fridrich, “Binary quantization using belief propagation with decimation over factor graphs of ldgm codes,” in Allerton Conference on Control, Computing, and Communication, September 2007.
[21] A. D. Liveris, Z. Xiong, and C. N. Georghiades, “Compression of binary sources with side information at the decoder using ldpc codes,” IEEE Communications Letters, vol. 6, no. 10, pp. 440–442, Oct 2002.
[22] H. Esmaeeli and M. Rezaei. “Methods and Criteria for Evaluating Controllability of Video Bit Rate in HEVC-SCC,” International Journal of Industrial Electronics Control and Optimization vol. 4, no. 3, pp. 313-320, 2021.
[23] X. He, X. Zhou, P. Komulainen, M. Juntti, and T. Matsumoto, “A lower bound analysis of hamming distortion for a binary ceo problem with joint source-channel coding,” IEEE Transactions on Communications, vol. 64, no. 1, pp. 343–353, Jan 2016.
[24] T. Filler and J. Fridrich, “Binary quantization using belief propagation with decimation over factor graphs of ldgm codes,” in Allerton Conference on Control, Computing, and Communication, September 2007.
[25] T. Filler, “Minimizing embedding impact in steganography using low density codes,” Master’s thesis, SUNY Binghamton, Jul. 2007.
[26] M. J. Wainwright, E. Maneva, and E. Martinian, “Lossy source compression using low-density generator matrix codes: Analysis and algorithms,” IEEE Transactions on Information Theory, vol. 56, no. 3, pp. 1351–1368, March 2010.
[27] P. L. Dragotti and M. Gastpar, Distributed source coding: Theory, Algorithms and Applications. Elsevier, 2009.
[28] S. S. Pradhan and K. Ramchandran, “Distributed source coding using syndromes (discus): design and construction,” IEEE Transactions on Information Theory, vol. 49, no. 3, pp. 626–643, Mar 2003.
[29] S. Yuxuan, S. Shuo, and W. Yongpeng, "Rate-Distortion Analysis for Semantic-Aware Multi-Terminal Source Coding Problem", arXiv preprint arXiv:2303.06391, 2023.
[30] M. Sartipi and F. Fekri, “Lossy distributed source coding using ldpc codes,” IEEE Communications Letters, vol. 13, no. 2, pp. 136–138, February 2009.
[31] D. H. Schonberg, “Practical distributed source coding and its application to the compression of encrypted data,” Ph.D. dissertation, Univ. California, Berkeley, 2007.

Mahdi Nangir received the B.Sc. degree with first rank in Electrical Engineering from University of Tabriz and M.Sc. degree in Communication System Engineering from Sharif University of Technology, Tehran, Iran, in 2010 and 2012, respectively. He received the Ph.D. degree from K. N. Toosi University of Technology, Tehran, Iran, in 2018. In 2017, he joined McMaster University, Hamilton, Ontario, Canada as a research visiting student. He is now an associate professor in Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran. His research interests include wireless communication, distributed source coding and multi-user information theory, data compression and optimization algorithms. Email: nangir@tabrizu.ac.ir.

IECO

This page intentionally left blank.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Robust QCA Full Adders Using a Novel Fault Tolerant Five-Input Majority Gate

Farzaneh Jahanshahi Javaran[1] | Somayyeh Jafarali Jassbi[1] | Hossein Khademolhosseini[1,2] | Razieh Farazkish[3]

Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.[1]
Department of Computer Engineering, Beyza Branch, Islamic Azad University, Beyza, Iran.[2]
Department of Computer Engineering- South Tehran Branch, Islamic Azad University, Tehran, Iran.[3]
Corresponding author's email: h.khademolhosseini@srbiau.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 2024-Jan-23
Received in revised form: 2024-Mar-03
Accepted: 2024-Mar-05
Published online: 21-June-2024

Keywords:
Quantum Cellular Automata,
Nano Electronics,
Fault-Tolerant,
Majority Gate.

ABSTRACT

A novel technique for creating logic gates and digital circuitry at the nanoscale is quantum cellular automata (*QCA*). The sensitivity of the circuit is enhanced and quantum circuits are more susceptible to unfavorable external conditions when component size are reduced. In this article, we offer a five-input majority gate with fault-tolerant feature in QCA technology, taking into account the significance of constructing circuits that can withstand flaws. We also assess all potential defects in the process of arranging cells in specific locations on the surface. These errors consist of extra cells, rotation, deletion, and displacement. The gate under study is subjected to the aforementioned four failure categories in the first stage. The QCADesigner simulator engine is then used to assess the accuracy of the circuit performance in the second step. *41* quantum cells have been used to make the gate of this five-input majority gate with fault-tolerant feature in QCA technology. Several techniques are explored to discover such a majority gate, such as adding cells (i.e., introducing redundancy into the circuit) and particular cell layout techniques. The goal is to come up with a design that can ideally withstand possible faults with the least amount of overhead on the circuit for fault-tolerant through a certain cell layout. The findings demonstrate the implemented majority gate's notable advantage over comparable scenarios.

I. Introduction

According to the trend of moving towards nano technology, quantum cellular automata (*QCA*) is a new technology for implementing logic gates and digital circuits in the nano scale [1]. This technology has many features such as small area, high processing speed and low power consumption, and also the elimination of electric current and the elimination of the capacitor element is one of the reasons for its popularity. In QCA, the information is expressed by binary numbers, and with this difference, a cell has taken the place of the current switch. One state of the cell indicates a binary *1* and the other a binary *0*, while no current enters or leaves the cell. The field caused by the charge shape of one cell changes the charge shape of another cells [2]. This mutual connection is associated with a timing scheme that regulates the effect between states and this is sufficient for calculation with the lowest amount of power loss. In recent years, QCA have received much attention because designs and simulations show that QCA will have many capabilities compared to complementary metal-oxide semiconductor *(CMOS)*, but it is still not clear that this technology can replace *CMOS*, even though many of the problems and obstacles of building QCA are being solved with the passage of time [3]. So far, a large number of implementations for three- and five-input majority gates have been presented, but a limited number of fault-tolerant five-

input majority gates have been introduced and research in this particular field is appropriate.

II. Basic Concepts of QCA and Related Works

The basic and important principle in quantum technology is the cell and in Figure 1 shows a four-point QCA-based cell. These four quantum dots are placed in the shape of a square. Exactly two mobile electrons are placed in the cell, which have the ability to move between different quantum points in a cell using electron tunneling, and stable states are established when the holes are occupied diagonally. According to Coulomb repulsion law, the location of these two electrons in the holes will be in opposite corners, which creates two structures. These two structures show the two poles *+1 and -1*, which in the calculations, we assign logical values *1 and 0* to them, respectively [4].

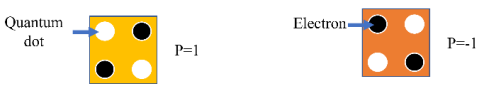


Fig. 1. QCA cell.

QCA has a clock-based mechanism that includes a four-clock signal with equal frequencies. We consider one of the clock signals as a reference phase (*0*) and the rest with a delay of one ($\frac{\pi}{2}$), two (π) and three ($3\frac{\pi}{2}$) quarters of the period. Clock is an electronic factor that controls the movement of electrons inside the cell [5, 6]. In fact, the way it is controlled is that if information reaches a part of the circuit that needs to be combined with several other inputs to produce an output, if the other inputs arrive at that part of the circuit later, it prevents the dissemination of information in that part until the arrival of the other inputs [7]. As shown in Figure 2, each clock signal is divided into four parts: *switch, hold, release, and relax.*

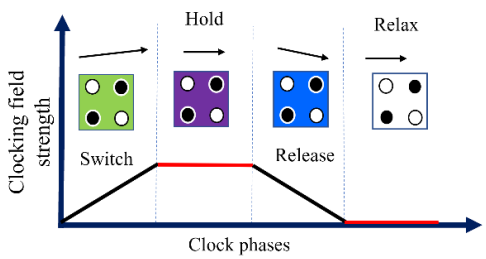


Fig. 2. QCA Clock phases.

Each cell also affects the neighboring cells. If two cells are placed next to each other, they always end up in a state where the repulsive force is minimized. An array of adjacent cells can be used like a wire to transmit information. Also, the cells can be placed in a 45- or 90-degree position, where the 45-degree wire is formed by rotating the cells 45 degrees [8]. In this type of wire, the binary value changes sequentially within the cells. The two wire configurations in QCA are depicted in Figure 3.

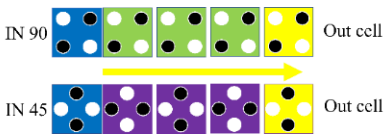


Fig. 3. QCA wires.

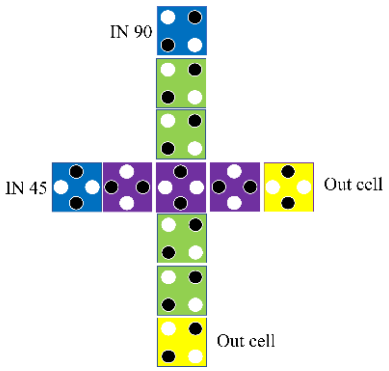


Fig. 4. QCA wiring.

Another benefit of this technology is the capacity to cross wires [9]. Additionally, cables may be arranged in several layers without negatively affecting one another, or two distinct kinds of cells can be bridged by utilizing both kinds of cells. Figure 4 shows wiring with two distinct kinds of cells.

Inverter gate and majority gate are two basic gates in QCA technology. In majority gate, considering all the input states, the output always shows the majority. Most circuits are made based on these two gates [10, 11]. Figure 5 shows two 3- and 5-input majority gates and an inverting gate.

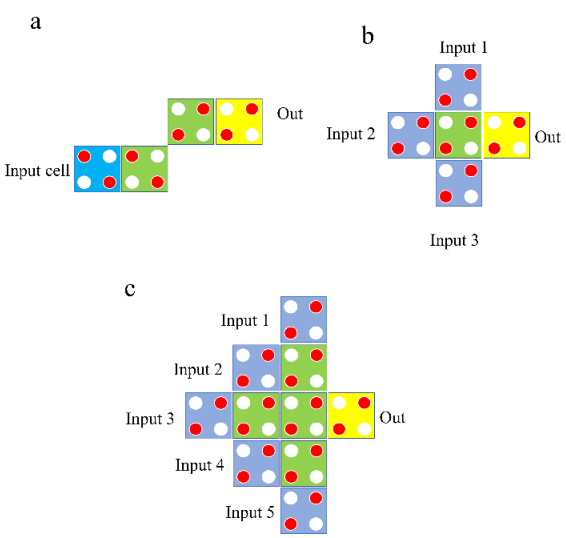


Fig. 5. QCA basic gates: (a) A QCA inverter, (b) a three-input majority gate and (c) a five-input majority gate.

In addition to the cases mentioned, there are types of defects in QCA technology. In the QCA fabrication stage, the occurrence of faults is possible in both the synthesis phase (where individual cells are constructed) and the placement phase (where cells are positioned at specific locations).

Defects in the synthesis phase may cause the cell or cells to lose some of their holes or electrons or gain some extra holes or electrons. Such defects disrupt the correct operation of the cell and are usually easily detectable, but defects in the process of placement are more likely than in the process of synthesis [12]. The defects created in the placement phase are called misplacement. The possible faults that occur in this technology include displacement faults, cell displacement, cell omission, cell rotated, and extra-cell. Displacement faults are faults that quantum cells move in their original direction. The fault of omission is the fault that a cell has been removed from the circuit. The fault of rotation means that the cell rotates in proportion to the adjacent cells in such a way that the direction of the cell changes, and the fault of the extra-cell is the shape that during the process of placing the cells, one or a number of additional cells may be added to the circuit [13]. Figure 6 shows the types of faults of QCA technology.

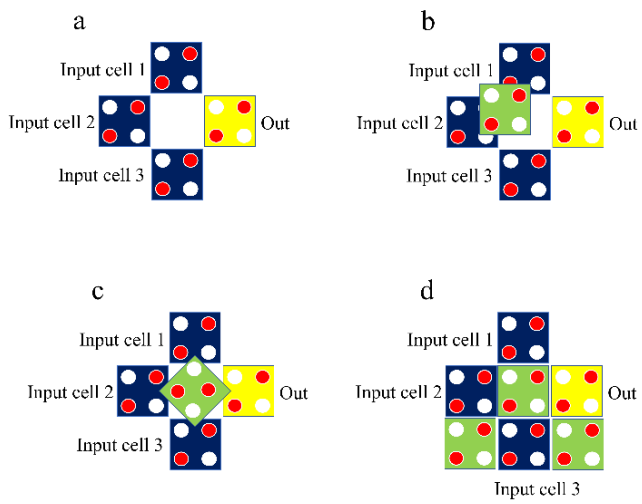


Fig. 6. Type of faults: (a) Cell omission, (b) cell displacement, (c) cell rotation, and (d) extra-cell deposition [14].

In the following, we have reviewed some of the latest designs of fault-tolerant five-input majority gate.

Farazkish [15] created a fault-tolerant five-input majority gate with *50* cells and an area of *0.0352* μm^2. The gate exhibits *13.6%* fault-tolerant to single-cell omission (Figure 7(a)). One of the disadvantages of this gate was that the output was in the middle, necessitating a multilayer structure to construct the circuit. Du, et al. [16] suggested a fault-tolerant five-input majority gate with *22* cells and an occupied area of *0.0163* μm^2, making it *12.5%* fault resistant to the single-cell omission defect (Figure 7 (b)). A fault-tolerant four-input majority gate with *27* cells and an occupied area of *0.0318* μm^2 was given by Goswami, et al. [17]. It is *28%* fault resistant to the single-cell omission defect (Figure 7 (c)). A fault-tolerant five-input majority gate with 27 cells and an occupied area of *0.0135* μm^2 was suggested by Sun, et al. [18]; it has a fault tolerance of up to *47.62%* (Figure 7 (d)).

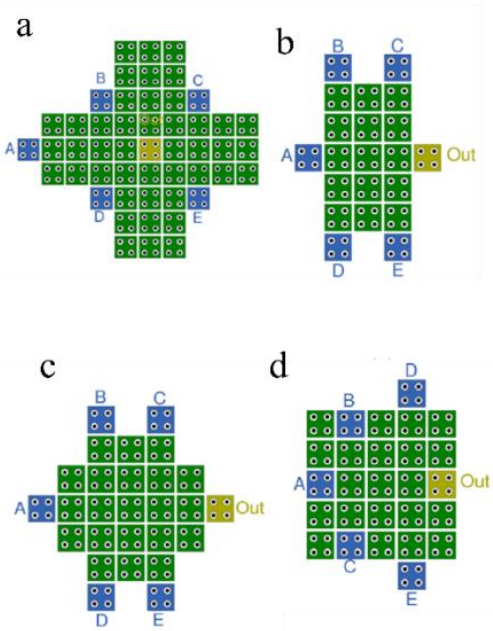


Fig. 7. Cellular structures of fault-tolerant, five-input majority gates: (a) [15], (b) [16], (c) [17], and (d) [18].

III. The Proposed Five-Input Majority Gate

The proposed fault-tolerant five-input majority gate is shown in Figure 8. This gate is designed with *41* quantum cells in the occupied space of *38804 square nanometers*. From the structural point of view, the proper arrangement of the input cells and the length of the used wires have caused their coordinated effect and as a result, the majority of the five input gates have been created. The simulation results from the QCADesigner simulation engines show the correct performance of the circuit for all *32* input combinations. The test vectors are sequentially applied to the inputs and the resulting waveform is examined by each engine separately. Matching the waveforms shows the accuracy of the designed circuit. Figure 9 shows the result of bistable approximation simulation of the circuit.

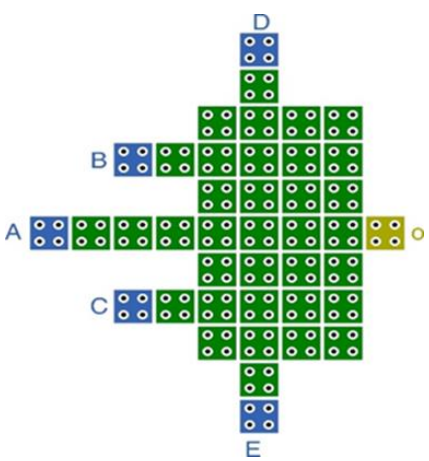


Fig. 8. Proposed defect tolerant five-input majority gate designed with minimal hardware complexity in QCA.

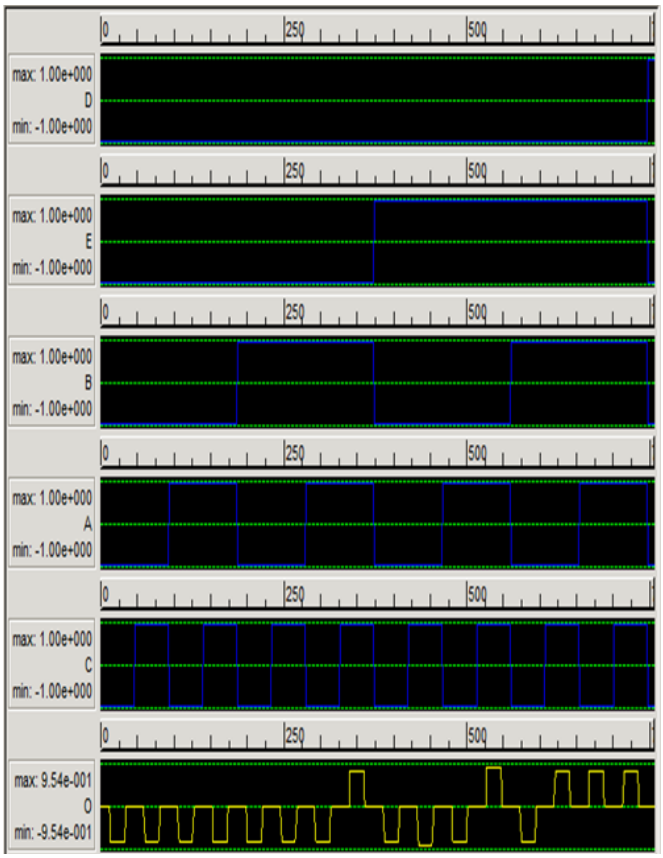


Fig. 9. The simulation results of proposed fault-tolerant five-input majority gate.

IV. Fault Tolerance Analysis

As shown in Figure 10, for applying defects to the majority gate, a number of cells in the circuit are considered as test cells. These cells have been selected for examination in more sensitive areas. The specified numbers indicate the area of application of the faults.

The results of the analysis of the displacement faults of the proposed gate are shown in Table 1. As it is evident, the displacement of *10* selected cells along with input and output cells of the circuit have been investigated as test cells in the four directions of north, south, east and west. The maximum amount of displacement faults allowed in the process of cell arrangement for each of the cells while maintaining the correct function of the output cell (*O*) has been studied. Considering the high importance of the input and output cells, we examine the performance of two cells, for example cell A can move up to *5 and 1 nm* in the west and east (horizontal movement) and on the other hand, 6 nm movement in the vertical direction for this cell is imaginable. Cell D with 5 nm displacement in the horizontal direction and 4 nm in the north direction has significantly increased the tolerance of the circuit. It is worth noting that all the middle cells that are surrounded by other cells have tolerance of 1 nm displacement in the four directions.

TABLE 1
MAXIMUM TOLERANCE AGAINST DISPLACEMENT FAULTS (THE NUMBERS SHOWN ARE THE MAXIMUM ALLOWABLE DISPLACEMENT IN NANOMETERS WHILE MAINTAINING CORRECT OUTPUT PERFORMANCE)

North	South	East	West	Cell
6	6	1	5	**A**
3	3	1	2	**B**
3	3	1	2	**C**
4	1	5	5	**D**
1	4	5	5	**E**
3	3	3	1	**O**
5	1	1	5	**1**
1	1	1	1	**2**
1	1	6	1	**3**
1	1	1	1	**4**
1	1	1	1	**5**
4	4	1	1	**6**
1	1	1	1	**7**
1	1	1	1	**8**
1	7	6	1	**9**
1	1	1	1	**10**

As shown in Table 2, the tolerance of the proposed circuit against the omission faults have been evaluated respectively by measuring the output of the circuit. The waveform obtained from the simulation after removing each cell has been compared with the waveform obtained from the faultless circuit, and if the output is preserved, the result is reported. In Table 2, out of the *10* cells examined for omission faults, ignoring *4* cells did not affect the output, and therefore *40%* can be considered as the tolerance percentage of omission faults.

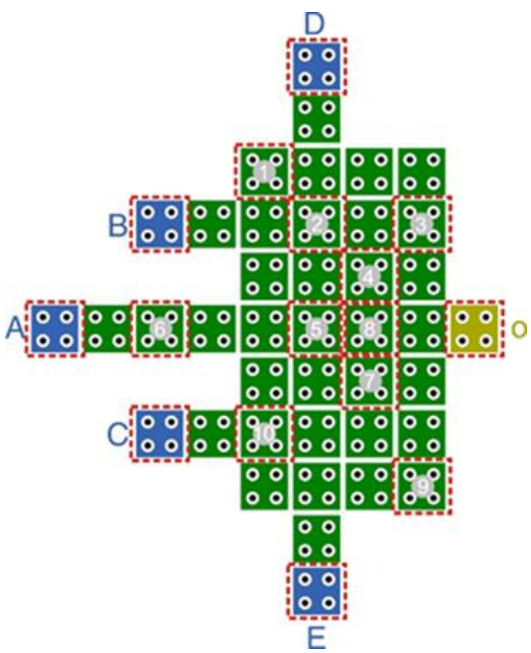


Fig. 10. The proposed majority gate along with the analyzed points for displacement, omission and rotated.

Also, as shown in Table 2, the tolerance of the proposed circuit against the omission faults have been evaluated respectively by measuring the output of the circuit. The waveform obtained from the simulation after removing each cell has been compared with the waveform obtained from the faultless circuit, and if the output is preserved, the result is reported. In Table 2, out of the *10* cells examined for omission faults, ignoring 4 cells did not affect the output, and therefore *40%* can be considered as the tolerance percentage of omission faults.

TABLE 2
TOLERABILITY OF OMISSION FAULTS AND ITS EFFECT ON THE OUTPUT

Output	Omission cell
MV5	**1**
Incorrect	**2**
Incorrect	**3**
Incorrect	**4**
MV5	**5**
Incorrect	**6**
Incorrect	**7**
MV5	**8**
MV5	**9**
Incorrect	**10**

For investigating the effects of cell rotation defects and their impact on the output, Table 3 has been considered. As it has been determined, in *40%* of the cases, the rotation faults have been covered.

TABLE 3
ANALYZING THE ACCEPTABLE TOLERANCE OF THE FAULTS BASED ON THE ACCEPTABLE TOLERANCE OF THE ROTATION FAULTS AND ITS EFFECT ON THE OUTPUT

Output	Rotated cell
MV5	**1**
Incorrect	**2**
Incorrect	**3**
Incorrect	**4**
MV5	**5**
Incorrect	**6**
Incorrect	**7**
MV5	**8**
MV5	**9**
Incorrect	**10**

In order to analyze the extra-cell faults on the majority gate circuit, Figure 11 is considered. The 9 ×11 rectangle shows all the places that can be used for adding cells to the proposed circuit. As shown in Table 4, among these places, *12* places have been selected and numbered according to the sensitivity of the test place. It seems that among all the locations examined, only adding a cell at one specific location has caused a change in the output value, demonstrating that the proposed circuit has a tolerance of *91.6%* against additional defects.

TABLE 4
TOLERABILITY OF REDUNDANCY FAULTS AND ITS EFFECT ON OUTPUT

Output	Additional cell
MV5	**1**
MV5	**2**
Incorrect	**3**
MV5	**4**
MV5	**5**
MV5	**6**
MV5	**7**
MV5	**8**
MV5	**9**
MV5	**10**
MV5	**11**
MV5	**12**

Table 5 shows a comprehensive hardware comparison of the five-input majority gates. In this table, the proposed gate has been compared with the seven previously presented circuits in terms of the number of consumed cells, the occupied area and accessing the input and output cells from the structural aspect. The proposed circuit has the lowest average tolerable percentage of faults compared to the lasted fault-tolerant gates. On the other hand, the new gate in this research along with two last beast designs have simultaneous single-layer access to input and output cells. This significantly facilitates the process of building QCA circuits because there will be no need for multi-layer structures or layered wiring and the inputs and outputs of the circuit can easily be connected to other circuits.

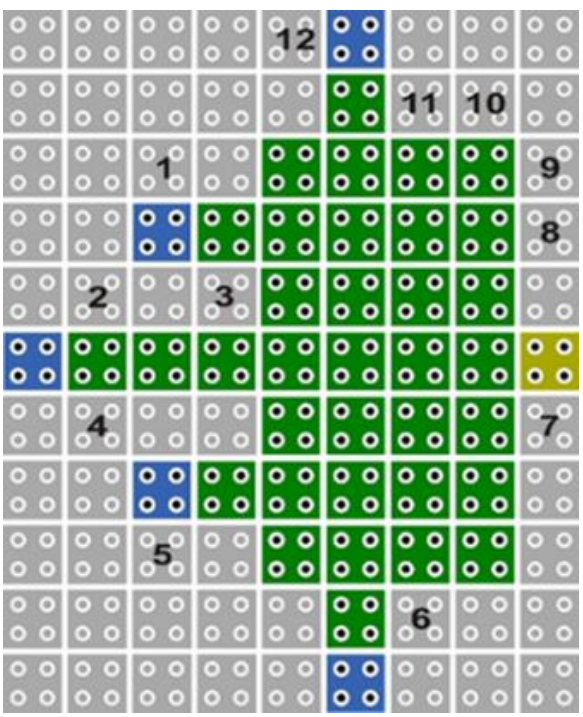


Fig. 11. The proposed five-input fault-tolerant majority gate circuit along with the analyzed points for redundancy fault.

TABLE 5
STRUCTURAL COMPARISON OF THE FIVE-INPUT MAJORITY GATES

Designs	Area (nm^2)	Cell count	Accessibility to the output cells	Accessibility to the input cells	Average Fault-tolerant
Proposed design	38804	41	Yes	Yes	60/75
[19]	34444	42	Yes	Yes	51/20
[15]	38804	51	No	Yes	48/30
[20]	7644	10	Yes	No	Less than 10
[21]	4524	10	No	Yes	Less than 10
[22]	16284	18	Yes	Yes	Less than 10
[23]	9604	13	Yes	Yes	Less than 10
[24]	22400	28	Yes	Yes	34

V. The Proposed Fault-Tolerant QCA Logic Circuits

In the five-input majority gate, by fixing two of the inputs to zero and one, it is possible to design three input AND gate and three input OR gate. In the case that two of the inputs are zero, the output of the function is one only when the other two inputs are one, otherwise the output of the function will be zero. So, the behavior of the system will be like an "*AND*" operator. In the same way, by fixing *+1* polarization to two of the inputs, a two input "*OR*" gate can be designed and they are implemented based on Eq. (1) and (2).

$$AND(A,B,C) = A.B.C = MV5(A,B,C,0,0) \quad (1)$$

$$OR(A,B,C) = A + B + C = MV5(A,B,C,1,1) \quad (2)$$

Figure 12 shows the symbols of a three input "*AND*" gate and a three input "*OR*" gate.

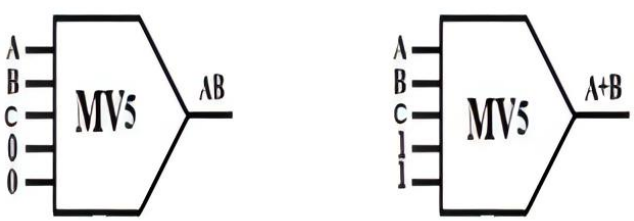


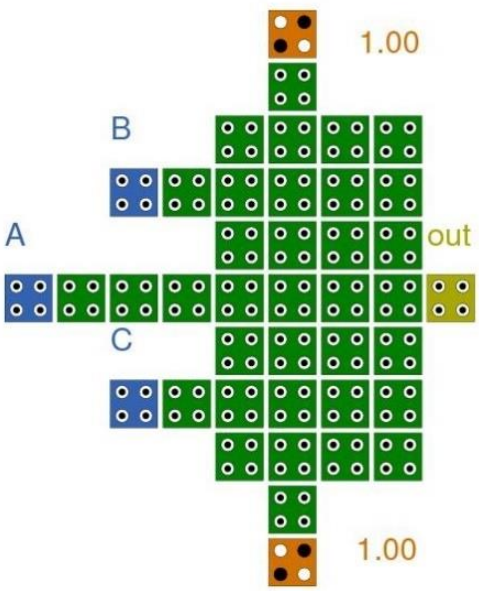
Fig. 12. Symbols of three input *AND* and *OR* gate.

Table 6 shows three input *"AND", "OR", "NAND" and "NOR"* gates.

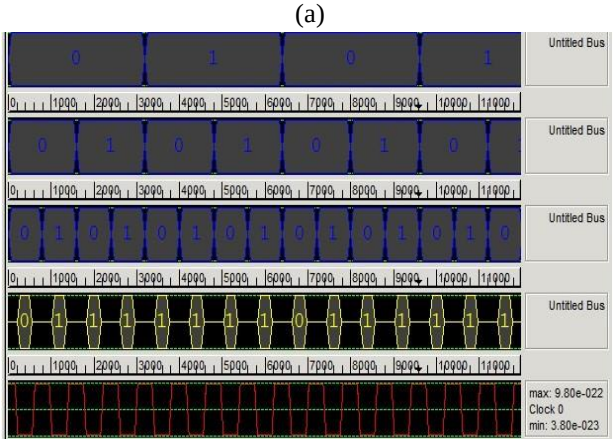
TABLE 6
IMPLEMENTATIONS OF 3-INPUT *"AND", "OR", "NAND" AND "NOR"* GATES

A	B	C	D	E	OUT
--	--	--	-	-	3-input AND
--	--	--	1	1	
--	--	--	1	1	3-input OR
--	--	--			
--	--	--	1	1	3-input NAND
--	--	--			
--	--	--	-	-	3-input NOR
--	--	--	1	1	

The three input "*AND*" gate using the proposed gate and the related simulation results are shown in Figure 13.



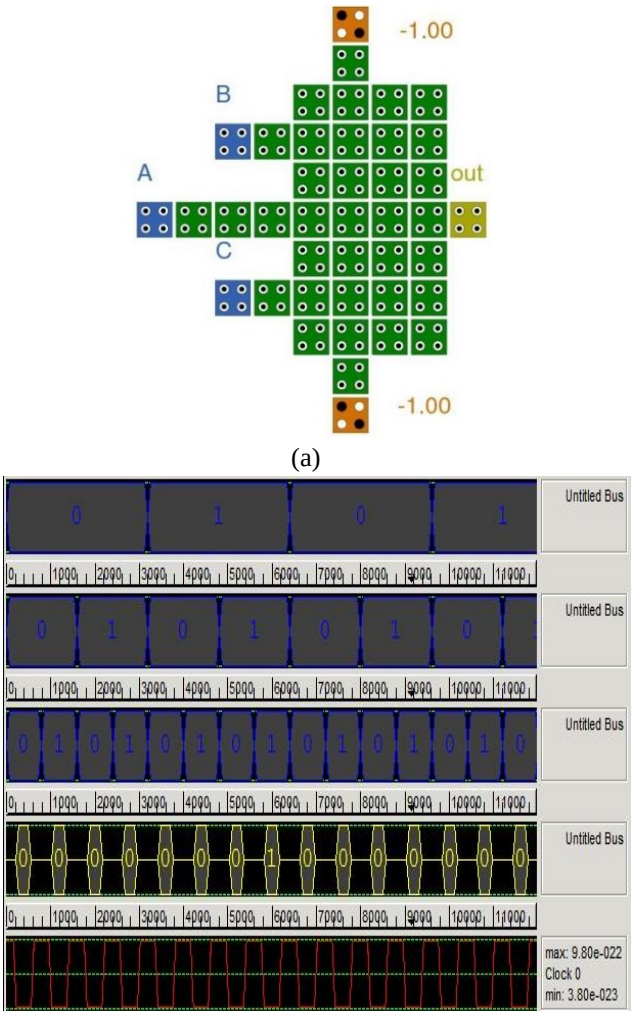
(a)



(b)

Fig. 13. (a) The QCA layout of proposed fault-tolerant 3-input AND gate, and (b) simulation results.

The three input "OR" gate using the proposed gate and the related simulation results are shown in Figure 14.



(a)

(b)

Fig. 14. (a) The QCA layout of proposed fault-tolerant 3-input OR gate, and (b) simulation results.

On the other hand; *"XOR"* gate is a logical gate that is applied to two operands, if exactly one of them is true, the output returns the true value or one, and otherwise the output value is false or zero. A two input *"XOR"* gate is implemented based on Eq. (3).

$$2-input\ XOR(A,B) = \bar{A}B + A\bar{B} = MV5(A,B,0,\overline{MV3(A,B,0)},MV3(A,B,0)) \quad (3)$$

The *"XOR"* gate is designed and implemented with the five-input majority gate proposed in this article and the three-input majority gate, so that two of the five input majority gate inputs named *A and B* and two of the other inputs are connected to the inverted output of three input majority gate and the remaining input of the five-input majority gate is connected to zero. Two of the inputs of the three-input gate are connected to *A and B*, and the third input is connected to *zero*. In Figure 15 the symbol of the two input *"XOR"* gate is shown.

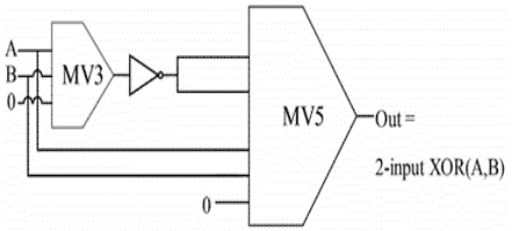
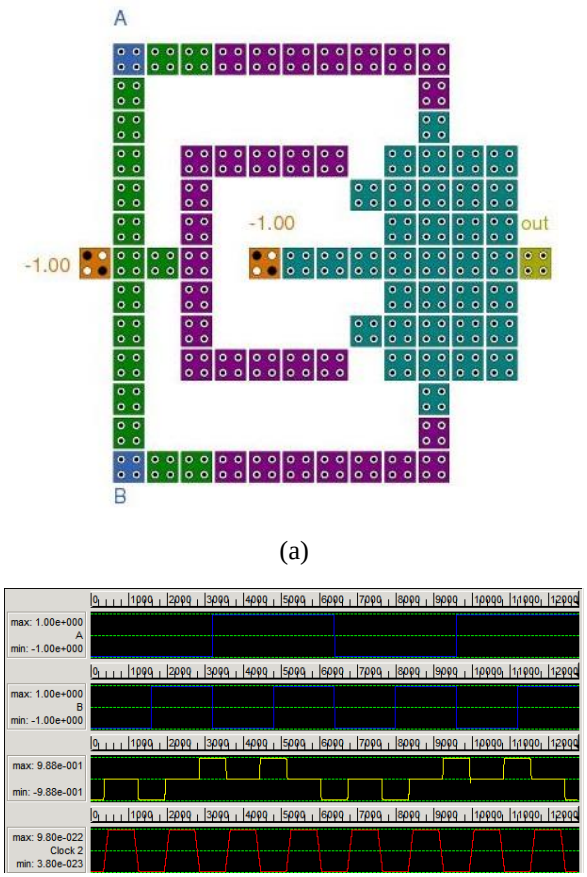


Fig. 15. Symbol of two input XOR gate.

The layout and simulation results of the 2-input "XOR" gate are shown in Figure 16.



(a)

(b)

Fig. 16. The proposed 2-input *"XOR"* gate; (a) QCA layout, and (b) simulation results.

Also, the *"XNOR"* gate is a digital logic gate and actually the complement of the *"XOR"* gate. This gate includes two or more inputs and one output, in the *"XNOR"* gate, the output becomes one if either both inputs are zero or both inputs are one. The two input *"XNOR"* gate is made based on Eq. (4).

$$2-input\ XNOR(A,B) = A.B + \bar{A}.\bar{B} = MV5(A,B,1,\overline{MV3(A,B,1)},MV3(A,B,1)) \quad (4)$$

The *"XNOR"* gate designed in this article is the same as the *"XOR"* gate with the difference that the inputs are connected

to one instead of zero. Figure 17 shows the symbol of the two input *"XNOR"* gate.

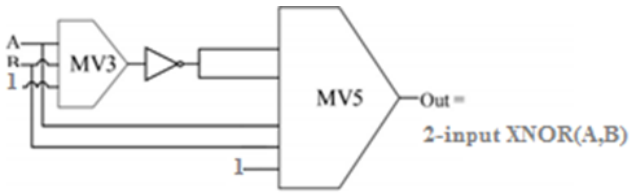
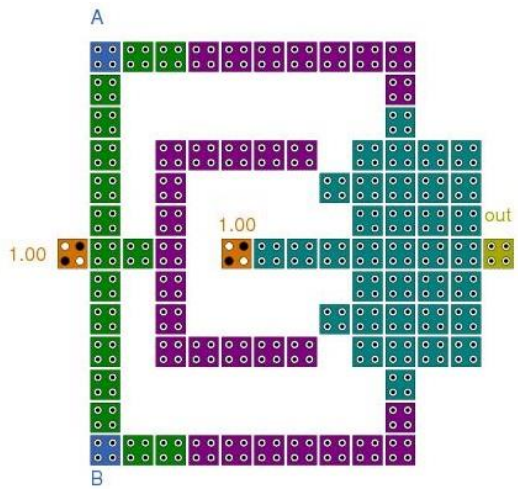
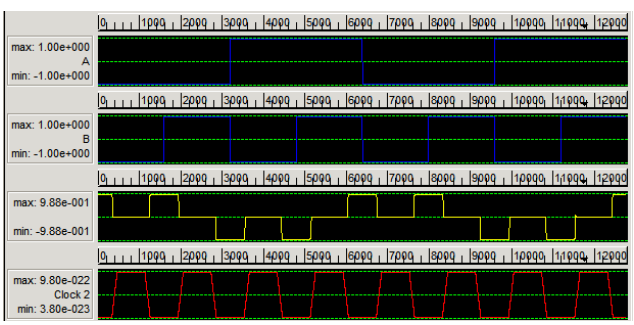


Fig. 17. Symbol of two input XNOR gate.

Figure 18 shows the layout of the two input *"XNOR"* gate and simulation results.



(a)



(b)

Fig. 18. The proposed 2-input *"XNOR"* gate; (a) QCA layout, and (b) simulation results.

Also, the full-adder is implemented based on relations (5) and (6) [25, 26]:

$$Cout = AB + AC + BC = MV5(A, B, C, 0, 0) \quad (5)$$

$$Sum = A \oplus B \oplus C = MV5(A, B, C, Cout, Cout) \quad (6)$$

The full-adder designed in this article includes a five-input majority gate and a three-input majority gate and an inverting gate. Figure 19 shows the symbol of the full-adder circuit.

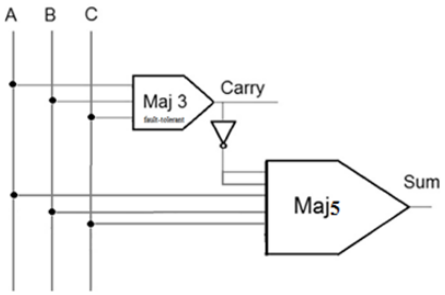
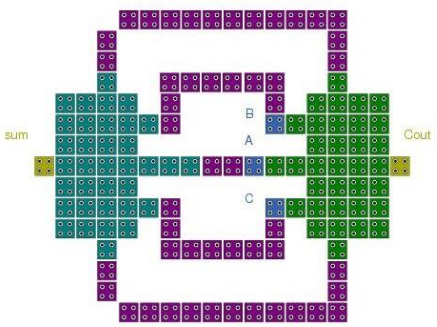
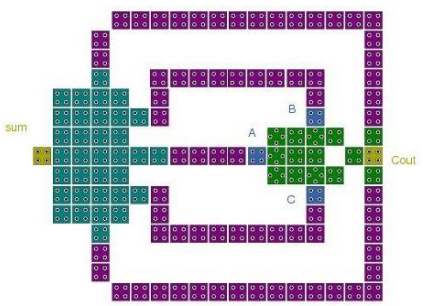


Fig. 19. The Symbol of full-adder.

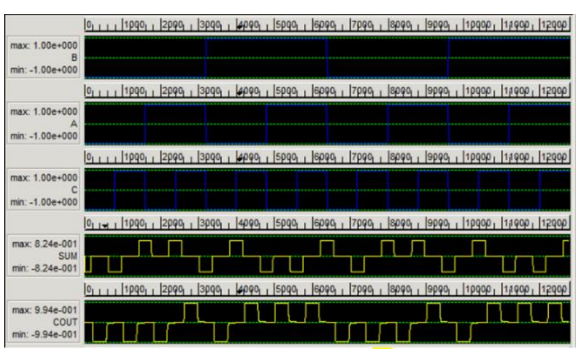
The full-adder circuit is designed and implemented in one layer and works in one clock phase, Figure 20 shows two layouts and simulation results for the full-adder circuit with QCADesigner software.



(a)



(b)



(c)

Fig. 20. The proposed full-adder (a) plan 1, (b) plan 2 and (c) simulation results.

VI. Conclusions

Because of its great density, low power, and rapid speed, QCA technology was seen as one of the finest alternatives to *CMOS* technology. Due to the novelty of QCA technology, some major concerns have yet to be properly explored. One of the most critical issues is fault-tolerant. As a result, the design of fault-tolerant QCA computational circuits is critical. The presentation of fault-tolerant designs in QCA technology, like other technologies, has been of great interest. In this article, a new fault-tolerant five input majority gate with minimal cell overhead has been designed and simulated in QCA technology. The new majority gate in this article is designed in one layer and with one clock phase. This gate is designed with *41* quantum cells in the occupied space of *38804* square nanometers. For this circuit, in addition to the structural analysis, the tolerance analysis of the circuit was investigated when four faults of displacement, deletion, rotation and redundancy were applied, and the correctness of the circuit performance was tested with QCADesigner simulator engines. It was demonstrated that the suggested gate was *60% and 75%* tolerant to single-cell omission and extra-cell deposition errors, respectively. It was also very stable under cell displacement and misalignment errors. Then three input *"AND" and "OR"* gates, two input *"XOR" and "XNOR" gates* and a *full-adder circuits* have been designed and simulated. The evaluation findings demonstrated that the circuits constructed using the suggested structure outperformed those designed using earlier fault-tolerant gates. Finally, we developed and built fault-tolerant single-layer full-adders employing the suggested fault-tolerant QCA majority gate. The circuits suggested in this article can be used as a foundation for future work to create more complicated and fault-tolerant designs such as multipliers, dividers, and so on.

REFERENCES

[1] F. Motalebi and S. Sayedsalehi, "Design of Low Power Full-Adder Circuit Using Quantum-dot Cellular Automata," International Journal of Industrial Electronics Control and Optimization, vol. 5, no. 1, pp. 99-108, 2022.

[2] S. F. Naz, S. Riyaz, and V. K. Sharma, "A Review of QCA Nanotechnology as an Alternate to CMOS," *Current Nanoscience*, vol. 18, no. 1, pp. 18-30, 2022.

[3] A. Shamsi, "A new Mismatch cancelation for Quadrature Delta Sigma Modulator," International Journal of Industrial Electronics Control and Optimization, vol. 3, no. 2, pp. 196-204, 2020.

[4] S. Seyedi and N. J. Navimipour, "A fault-tolerant image processor for executing the morphology operations based on a nanoscale technology," Multimedia Tools and Applications, vol. 82, no. 2, pp. 2489-2502, 2023.

[5] R. Marshal and G. Lakshminarayanan, "Fault resistant coplanar QCA full adder-subtractor using clock zone-based crossover," IETE Journal of Research, vol. 69, no. 1, pp. 584-591, 2023.

[6] K. Pandiammal and D. Meganathan, "Efficient design of QCA based hybrid multiplier using clock zone based crossover," Analog Integrated Circuits and Signal Processing, vol. 102, pp. 63-77, 2020.

[7] L. Wu, Z. Shen, and Y. Ji, "Using nano-scale QCA technology for designing fault-tolerant 2:1 multiplexer," Analog Integrated Circuits and Signal Processing, vol. 109, pp. 553-562, 2021.

[8] L. Zhou et al., "Narrowband photoblinking InP/ZnSe/ZnS quantum dots for super-resolution multifocal structured illumination microscopy enhanced by optical fluctuation," Nanophotonics, no. 0, 2023.

[9] S. Seyedi and B. Pourghebleh, "A new design for 4-bit RCA using Quantum Cellular Automata Technology," Optical and Quantum Electronics, vol. 55, no. 1, p. 11, 2023.

[10] J. Huang and S. Lale, "A novel nano-scale architecture of Vedic multiplier using majority logic in quantum-dot cellular automata technology," Electronics Letters, vol. 58, no. 17, pp. 660-662, 2022.

[11] C. Labrado and H. Thapliyal, "Design of adder and subtractor circuits in majority logic-based field-coupled QCA nanocomputing," Electronics letters, vol. 52, no. 6, pp. 464-466, 2016.

[12] Y. Yin, J. Liu, and C. She, "A new fault-tolerant single-bit comparator in QCA technology using a novel X-NOR gate," Optik, vol. 269, p. 169837, 2022.

[13] D. Pan, B.-N. Wu, Y.-L. Sun, and Y.-P. Xu, "A fault-tolerant and energy-efficient design of a network switch based on a quantum-based nano-communication technique," Sustainable Computing: Informatics and Systems, vol. 37, p. 100827, 2023.

[14] K. Das and D. De, "QCA defect and fault analysis of diverse nanostructure for implementing logic gate," International Journal of Recent Trends in Engineering and Technology, vol. 3, no. 1, p. N/A, 2010.

[15] R. Farazkish, "A new quantum-dot cellular automata fault-tolerant five-input majority gate," Journal of nanoparticle research, vol. 16, pp. 1-7, 2014.

[16] H. Du, H. Lv, Y. Zhang, F. Peng, and G. Xie, "Design and analysis of new fault-tolerant majority gate for quantum-dot cellular automata," Journal of Computational Electronics, vol. 15, pp. 1484-1497, 2016.

[17] M. Goswami, B. Sen, and B. K. Sikdar, "Design of low power 5-input majority voter in quantum-dot cellular automata with effective error resilience," in 2016 sixth international symposium on embedded computing and system design *(ISED)*, 2016: IEEE, pp. 101-105.

[18] M. Sun, H. Lv, Y. Zhang, and G. Xie, "The fundamental primitives with fault-tolerance in quantum-dot cellular automata," Journal of Electronic Testing, vol. 34, no. 2, pp. 109-122, 2018.

[19] R. Farazkish and K. Navi, "New efficient five-input majority gate for quantum-dot cellular automata," Journal of Nanoparticle Research, vol. 14, pp. 1-6, 2012.

[20] K. Navi, R. Farazkish, S. Sayedsalehi, and M. R. Azghadi, "A new quantum-dot cellular automata full-adder," Microelectronics Journal, vol. 41, no. 12, pp. 820-826, 2010.

[21] K. Navi, S. Sayedsalehi, R. Farazkish, and M. R. Azghadi, "Five-input majority gate, a new device for quantum-dot cellular automata," Journal of Computational and Theoretical Nanoscience, vol. 7, no. 8, pp. 1546-1553, 2010.

[22] R. Akeela and M. D. Wagh, "A five-input majority gate in quantum-dot cellular automata," in NSTI Nanotech, 2011, vol. 2, pp. 978-981.

[23] A. Roohi, H. Khademolhosseini, S. Sayedsalehi, and K. Navi, "A symmetric quantum-dot cellular automata design for 5-input majority gate," Journal of Computational Electronics, vol. 13, pp. 701-708, 2014.
[24] S.-S. Ahmadpour, M. Mosleh, and S. R. Heikalabad, "An efficient fault-tolerant arithmetic logic unit using a novel fault-tolerant 5-input majority gate in quantum-dot cellular automata," Computers & Electrical Engineering, vol. 82, p. 106548, 2020.
[25] H. Khademolhosseini and Y. Nemati, "A New Design for Two-input XOR Gate in Quantum-dot Cellular Automata," Journal of Advances in Computer Research, vol. 10, no. 1, pp. 89-96, 2019.
[26] H. Khademolhosseini, S. Angizi, and Y. Nemati, "A fault-tolerant design for 3-input majority gate in quantum-dot cellular automata," Journal of Nanoelectronics and Optoelectronics, vol. 13, no. 1, pp. 93-103, 2018.

Farzaneh Jahanshahi Javaran was born in Kerman, Iran She a PhD student in Computer Systems Architecture from Islamic Azad University, Science and Research Department. University of Science and Research is a prestigious university in the field of computer systems architecture in Iran. Currently, her researching the design of fault-tolerant majority gates in quantum cellular automata. Her current research interests include quantum dot cellular automata, fault tolerance.

Sommayeh Jafarali Jassbi was born in Tehran, Iran, in 1982. She received the MSc degree in computer architecture engineering in 2007, and the Ph.D. degree in computer architecture engineering in 2010 from the Islamic Azad University, Science and Research Branch. In 2010, she joined the Department of computer engineering, Islamic Azad University Science and Research Branch. She became an associate professor in 2011. Her interests are cloud computing, internet of things, wireless sensor network and computer architecture and cryptography. She was head of computer department in 2012. Now she is selected as a head of computer department again. She was also an active member of young researcher club from 2004. She has written, translate and published several professional books and paper in her fields.

Hossein Khademolhosseini received B.Sc. degree in computer engineering in 2008 from Shiraz University, Shiraz, Iran. He also received his M.Sc. and Ph.D. degree in computer architecture at Department of Computer Engineering, Science and Research Branch of Islamic Azad University, Tehran, Iran, in 2011 and 2016, respectively. He is currently an assistant professor with the Department of Computer Engineering, Islamic Azad University, Beyza Branch. His research interests are computer arithmetic, photonic NoC and electronics with emphasis on QCA and VLSI.

Razieh Farazkish received the B.S. degree in computer engineering from the IAU, Central Tehran Branch (2007) and the M.S. (2009) and Ph.D. (2012) degrees in computer engineering from the IAU, Science and Research Branch. In 2012, she joined the Department of Computer Engineering, IAU, South Tehran Branch, as a Professor. Her current research interests include quantum-dot cellular automata, fault tolerance, nanoelectronic circuits, nano computing, testing and design of digital systems.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

A Technical Review on the Proper Design of Gate Drivers in Industrial Power Electronics Applications

Saeid Ahmadi[1] | Kourosh Khalaj Monfared[1] | Mohammad Khalilzadeh[2] | Hossein Imaneini[1]

School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran[1]

Research Assistant at University of Tarbiat Modarres, Iran[2]

Corresponding author's email: kourosh.khalaj@ut.ac.ir

Article Info

Article type:
Review Articles

Article history:
Received: 7-Dec-2023
Received in revised form: 10-Feb-2023
Accepted: 20-April2024
Published online: 21-June-2024

Keywords:
Power Semiconductor Devices,
Gate Drive Circuit,
IGBT Gate Drive,
Desat Protection.

ABSTRACT

Power semiconductor devices are the most important components in power electronics applications. They are also the most fragile components of electronic circuits. A power semiconductor device's switching performance and protection depend on the gate drive circuit specifications. Therefore, choosing an appropriate gate-driver and designing its corresponding circuits is necessary. This paper is a technical review of the proper design of gate drivers for silicon power switches (like Si IGBT and Si MOSFET) in industrial power electronics applications. In this paper first conducts an overview of the main specifications of gate drivers for industrial power electronics applications. Then, concerning the protective role that a gate-driver can provide, crucial points of an effective design are discussed. Finally, a circuit is proposed to test the gate driver's short-circuit protection. The circuit is experimentally evaluated for three gate drives, and the results are discussed. A practical comparison of the protection performance of commercial gate drives ACPL-330J, ACPL331, and PC929 is also conducted.

I. INTRODUCTION

In the present era, power electronic components have become pervasive across many applications, spanning motor drives, switching power supplies, renewable energy systems, flexible AC transmission systems (FACTS), and more. The reliability of these components has become a focal point of concern. A key element within these systems is the power semiconductor device (PSD), encompassing many devices such as Thyristors, MOSFETs, IGBTs, GaN-based switches, and SiC switches [1]. Proper driving techniques for these devices are paramount, as the failure of even a single component could lead to a systemic operational breakdown [2]. Various protection strategies for PSDs have been proposed in the existing literature. Notably, due to the fast operation of GaN and SiC PSDs, protection methodologies for theses PSDs exhibit distinctions from those tailored for Si-based PSDs such as IGBTs and conventional MOSFETs. Comprehensive insights into these variances can be found in scholarly works [3],[4]. Given the prevalence of commercially available power electronic devices reliant on conventional Si-based PSDs, this study focuses primarily on them.

There are several gate drives for different PSDs. Gate drives are generally divided into two categories: isolated and non-isolated. In isolated gate drivers, the input command is isolated from the power. The common mode and noise transmission rejection capability are the most important advantages of isolated gate drivers. On the downside, they are more expensive than non-isolated ones [5]. The Bootstrap method is widely used among the non-isolated methods [6].

TABLE I. COMPARISON OF IGBT PROTECTION METHODS

Measurement Method	Implementation	Feature
V_{ce} Voltage	Requirement of a high voltage diode and simple sampling circuit	The best method with fewer additional circuits and more straightforward implementation (proper for gate drive ICs)
Gate voltage and current	Requirement of special sampling circuits of gate voltage and current and low voltage logic circuit	Due to the low gate current, special measuring circuits are needed, and the possibility of functional error is also high.
I_c Current	Requirement of high-cost power circuit current measurement	The cost of implementing the method is high, increasing the losses in the power section.
Changing the gate current	Requirement of special and expensive current measurement circuit	It requires special and complex measurement circuits and imposes costs and losses to the system.

Bootstrap configurations utilize the voltage across the low-side switch to generate a floating supply voltage for driving the high-side switch. This approach eliminates the need for a separate isolated power supply for the high-side driver, simplifying the overall circuit design and reducing component count and cost. However, this method is limited in high-power applications due to the problem of electromagnetic interference (EMI), and there is interest in using isolated gate drives.A suitable gate driver for the PSD should offer various functionalities. Among these, isolation and safeguarding the switch during short-circuit (SC) scenarios are crucial features. Additionally, essential attributes comprise under-voltage lockout (UVLO), the soft turn-off for the PSD during SC occurrences, high common-mode rejection, etc [7]. Numerous gate drive products in the market incorporate these functionalities, and selecting a specific gate drive type is contingent upon the application.

This paper presents a detailed review of the specifications of proper gate driver ICs for PSDs. In addition, the precise design of SC protection is studied, and a test circuit for evaluating the designed circuit is proposed. By using the proposed circuit, the correct protection functioning can be validated for every gate driver ICs. For instance, the protection circuits for three gate drivers are developed and assessed experimentally using the suggested circuit. The paper aims to enhance power electronic circuit designers' comprehension of the protective attributes of gate drivers for MOSFET and IGBT power switches. It also aims to aid them in creating and testing SC protection strategies to mitigate the potential system failure risks in safety-critical applications like motor drives.

II. DESIRABLE GATE DRIVE SPECIFICATIONS

This section presents some of the desired specifications of the IGBT/MOSFET PSD gate drivers. A gate drive may have some of the following specifications.

A. Electrical isolation

One essential requirement of the IEC 61800-5-1 standard is to ensure electrical isolation between the control and power sections, as outlined in references [8] and [9]. This practice enhances operational efficiency, safeguards electronic components against malfunctions, and enhances system reliability. The gate drive circuit addresses the electrical isolation challenge, separating the power and control sections. As a result, isolated gate drives have a higher priority in industrial power electronics applications [10].

B. Protections

IGBT protection (voltage and current) is important in power electronic converters. Managing voltage and current overshoot during hard switching processes and SC condition has historically relied on passive methods like fine-tuning gate resistor values and incorporating clamping circuits. However, recent advancements have led to more refined and efficient methodologies, predominantly through advanced gate drivers. This progression signifies a significant move towards enhancing the reliability of IGBT systems without incurring additional costs, capitalizing on the inherent capabilities of modern gate drivers. These advancements in gate driver technology have paved the way for various innovative techniques to enhance IGBT drive reliability [11]. Based on sampling schemes of IGBT current or voltage, cutting-edge protection mechanisms can be classified into distinct approaches: (1) the IGBT collector-emitter voltage-based (VCE) strategy [12], (2) gate sampling-based strategy [13], (3) the IGBT collector current-based strategy [14], and (4) the IGBT collector current rate of change (di/dt) strategy [15]. To facilitate a comprehensive comparison of these methodologies, Table I thoroughly analyzes the approaches above. Each of these approaches offers unique benefits and trade-offs in optimizing the performance and stability of power electronic systemss. This paper focuses on the VCE monitoring method because it is a standard method for commercial gate drives. Therefore, an important feature of a gate driver is the ability to protect the PSDs. Based on the type of protection, three important PSD protections are defined: SC protection (with VCE monitoring), protection of switches against dropping the voltage supply of gate below the operational value (under-voltage lockout capability (UVLO)), and soft turn-off capability in SC events [16].

1) Short circuit protection

Protection of IGBT/MOSFET during overcurrent conditions is a critical part of system reliability, both in terms of asset destruction and safety. PSDs such as IGBTs are not regarded as a fail-safe component, and their failure can result in a DC bus capacitor explosion, and complete power electronic system failure. SC protection in industrial gate drives is typically implemented based on desaturation (DESAT) detection (or VCE monitoring) [17], [18].

2) Undervoltage lockout (UVLO)

UVLO is a key feature that protects the system against bias supply failure. In isolated gate driver ICs, the voltage supply is provided at both the primary and secondary sides of the IC. The UVLO setting on the secondary side determines the minimum acceptable drive voltage for the PSD.

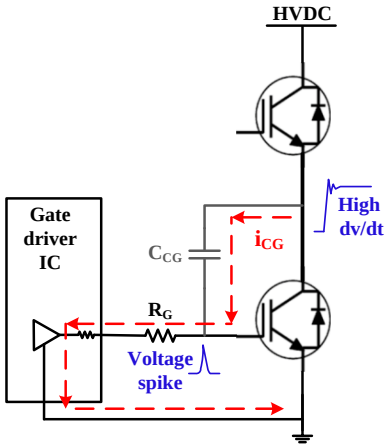


Fig. 1. Active miller effect.

By implementing UVLO in gate drivers, the gate voltage is monitored and prevented from dropping below a specified threshold. The UVLO rating is a determinant factor in high-power applications such as onboard chargers, EV traction inverters, and solar inverters [19]. Differing drive voltage requirements in these PSDs require different UVLO thresholds [20]. If the gate drives do not have UVLO capability, and if the gate supply voltage drops, the power switch will enter its linear region and cause a lot of losses, which can lead to irreparable damage.

3) Soft turn-off

Another important feature of industrial gate drivers is the soft turn-off of the switch in SC conditions. The soft-turn-off method protects the power switch in SC conditions. When the power switch is SC, the current increases to substantial values momentarily. Suppose the power switch is turned off at a high speed (as in normal conditions). In that case, it leads to a very large di/dt, and due to parasitic inductance in the path of the power switches, it leads to considerable voltage spikes that can damage the switch.

Therefore, turning off the switch slower than normal is necessary to avoid voltage spikes (VSPIKE). By implementing soft turn-off, can achieve smoother transitions during switching, reducing electromagnetic interference and enhancing the power electronic system's reliability during SC conditions [21]. The soft turn-off methods are very effective compared to the snubber circuits. This is because slow turn-off is employed only during overcurrent turn-off situations and does not affect the system's performance under normal operating conditions. This feature is achievable without the high cost and increasing the system's size. It increases the system's reliability and effectively achieves cost, performance, and reliability trade-offs [22].

4) Two-level turn-off

The two-level turn-off is another method to prevent the Collector-Emitter/Drain-Source voltage's sudden rise due to the switch turn-off, especially in SC protection conditions [23]. The two-level turn-off introduces a second turn-off voltage level at the gate-driver output between ON and OFF levels. This additional level ensures lower switch voltage overshoots at turn-off by reducing gate-emitter/source voltage at SC or over-current (OC) conditions.

C. Common-mode rejection

The common-mode (CM) noise is created in power electronic circuits due to switching action and creating dv/dt. This CM noise can be coupled to different parts of the circuit due to the parasitic capacitance in the circuit. Isolated gate drives that isolate the power and control sections have an important specification called CM transient immunity (CMTI). CMTI is defined as the quantification of CM rejection. CMTI criterion shows that the isolation is maintained up to the permissible CM level (dv/dt), and the noise of the power section is not coupled to the control. [24]. For example, if the DC bus is 1500V and the transition time of the gate-driver IC output signal is 100ns, the CMTI required by the IC is 15 V/ns. The gate drivers are specified for CMTI in their datasheet. For example, the ISO5851 and ISO5852S both have a minimum CMTI of 100 kV/μs. The higher CMTI for the IC is, the lower the false fault or the false output toggle due to the transient noise is ensured.

D. Gate Current Supply Capability

The IGBT/MOSFET switching time is controlled by charging and discharging its gate. The rise and fall times of the gate drive to the IGBT can be selectively controlled by selecting RG (on) and RG (off). When turned on, all current will flow from the IC through RG (on) and charge the IGBT gate capacitance, so increasing or decreasing RG (on) will increase or decrease the rise time in the application. With the addition of diode DRg, the fall time can be controlled independently because the turn-off current flows from the IGBT gate capacitance through RG (off) and DRg to the driver in the IC to GND. An increased gate peak current reduces switching losses by decreasing the turn-on and turn-off times. The gate resistors RG (on) and RG (off) can control the gate charge currents (Fig. 4). The theoretical gate peak current value can be calculated by:

$$i_{G,Peak} = \frac{V_{G(on)} - V_{G(off)}}{R_G + R_{G(\text{int})}} \quad , \tag{1}$$

where VG(on) and VG(off) are the gate's turn-on and turn-off supply voltage, respectively. Also, RG is an external gate resistor. The IGBT module's internal gate resistor RG(int) must be taken into account when calculating the peak gate current. In practice, stray inductance reduces the peak value below the possible theoretical value [25].

E. Other features

There are also some additional features that some gate drives may provide. They are introduced as follows.

1) Dead-time generation

In normal operation, two switches of a phase-leg of an inverter will be turned on and off one after the other. The simultaneous conduction of both switches results in a rise of current limited only by DC-link stray inductance and the switches' on-resistances.

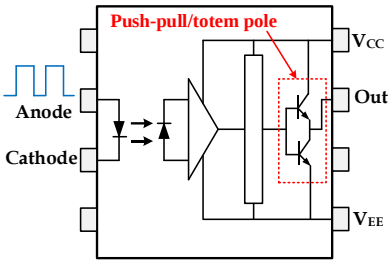


Fig. 2. Output stage with push-pull (or so-called totem-pole) structure.

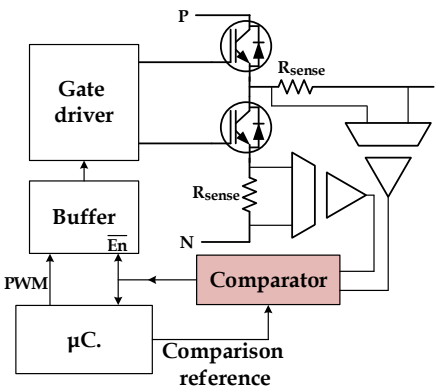


Fig. 3. Hardware-based SC protection circuit.

Although the switches would not be turned on at the same time intentionally, since they are not ideal switches, turn-on and turn-off times are not strictly identical. Adding a "dead time" into the control scheme is necessary to avoid shoot-through. With this additional time, one switch will always be turned off first, and the other will be turned on after the expired dead time. Hence, shoot-through can be avoided. The dead-time is typically generated in two ways; through the software or hardware. In the software-based method, the dead time is created to generate the PWM switching signal process based on the selected microcontroller. On the other hand, hardware-based methods use additional circuit elements to generate dead-time. In the case of hardware implementation, one solution is designing analog circuits [26]. Another solution is using special gate-driver ICs with built-in dead time generation. These ICs are more reliable and easy to implement than user-designed analog circuits. Some commercial examples are IR2104, IR21844, IR2110, etc.

2) Integrated Flyback controller

One of the interesting features of some gate-driver ICs is the integrated Flyback controller for isolated DC-DC converter.

Besides, these ICs contain a built-in MOSFET switch for the Flyback. An external transformer has to be connected to complete the Flyback circuit. This feature is beneficial in omitting an external isolated power supply for the gate driver's secondary side, resulting in a cost-effective design. One of the most famous commercial examples is the ACPL-302J gate drive.

3) Active Miller clamp

Parasitic turn-on due to the Miller capacitor is a common problem when operating an IGBT [27]. This effect is remarkable in 0 to +15 V type gate drivers (single supply driver). Due to the gate-collector coupling through the Miller capacitor, a high dv/dt transient created during IGBT turn-off can induce parasitic turn-on (gate voltage spike), which is potentially problematic. Fig.1 shows this phenomenon schematically.

When turning on the upper IGBT in a half-bridge, a voltage change dvCE/dt occurs across the lower IGBT. A current flows through the parasitic Miller capacitor CCG, the gate resistor RG, and the internal gate resistor. This current can be approximated by (2) and (3).

$$i_{CG} = C_{CG} \frac{dv_{CE}}{dt} \tag{2}$$

$$V_{CG} = i_{CG} R_G \tag{3}$$

The current creates a voltage drop across RG. An unwanted parasitic turn-on occurs if this voltage exceeds the IGBT gate threshold voltage. It should be noted that rising IGBT chip temperature would slightly reduce the gate threshold voltage. This parasitic turn-on can also be seen on the upper switch when the lower one is turned on.

III. STANDARD METHODS FOR PSD DRIVING

Various PSDs have been used in power electronic applications. One of them is the IGBT, which has found many applications in the industry. This section focuses on IGBT driving. There are various commercial methods to drive the gate of an IGBT [28]. Using a pulse-transformer is one of these methods, which provides isolation, as previously mentioned in section II. This is a simple and inexpensive method. However, pulse-transformers are more challenging to design than optically based isolation methods. When dealing with pulse transformers, the load becomes more critical than any other type of transformer. Leakage and primary inductance values have added significance because of their effect on the output wave shape [29]. If the load disturbs this wave shape, severe problems may occur. Also, this method cannot achieve a high and safe isolation level. It also does not provide SC protection. Meanwhile, a more advanced method is to exploit gate driver ICs to overcome the pulse-transformer drawbacks and provide most of the added benefits mentioned in the previous sections.

From the protective function point of view, commercial gate drives are divided into two main categories: 1) isolated gate drives without SC protection and 2) isolated gate drives with SC protection.

A. *Isolated gate-drives without SC protection*

Gate drives of this category provide galvanic isolation between the control and power sections, as shown in Fig.2. They usually have a push-pull (or so-called totem-pole) structure to supply the required sink and source current of the gate of the PSD. A gate driver sinks and sources current, meaning the driver can draw current from and supply current to the semiconductor device's gate terminal. The gate's sink and source current is proportional to PSD's switching speed. Some commercial examples are TLP250, TLP350, IX3180, HCPL-3120, HCPL-3150, etc. As shown in Fig.2, this type of IC has no protection against SC conditions. Therefore, the user should design an external SC protection circuit. One method is detecting the SC current, generating an error signal, and softly turning off the switch.

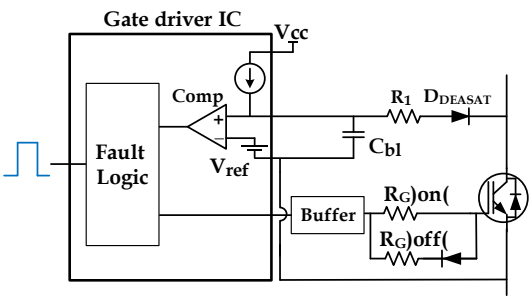


Fig. 4. DESAT circuit integrated with industrial gate drives ICs.

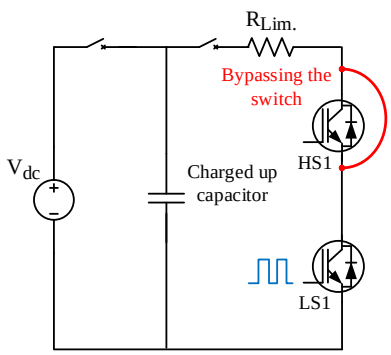


Fig. 5. Designed board for SC test condition.

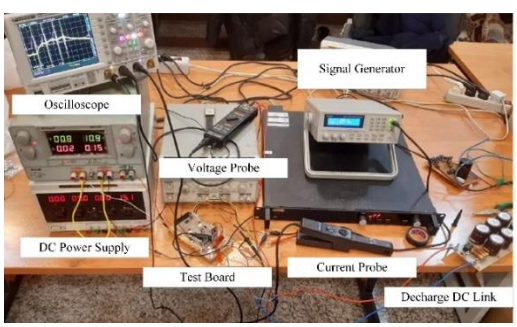


Fig. 6. Experimental setup.

The protection is provided by sending an error signal to the microcontroller and the shutdown command back via the microcontroller. Although the protection of the switch is not guaranteed due to the inherent delay in the method, it can protect the circuit.

A more effective protection strategy is to use hardware directly to detect SC and shut down the switch instead of using the software. This method has a small delay in protecting the switch if designed correctly. An example circuit is shown in Fig.3. As can be seen, the switch current is measured through the shunt resistors. After comparison with a threshold, if an SC is detected, the switch turn-off command is sent from the comparator, and then PWM is deactivated, resulting in the switch protection.

B. *Isolated gate-drives with SC protection*

In addition to isolation, some commercial gate drivers possess integrated circuits that are used to detect SC and protect switches. The most common method is monitoring the IGBT collector-emitter (VCE) voltage (or drain-source voltage, VDS, in MOSFET). This method is known as desaturation (DESAT) Protection. The VCE (or VDS) is proportional to the current flowing through the collector (IC) of the IGBT or drain (ID) of a MOSFET. The characteristic curve of VCE-IC or VDS-ID is given in the switch datasheets. These gate drivers monitor the VCE (or VDS) and compare it with an internal reference voltage. Once a SC is detected, which is an overvoltage of VCE (or VDS) concerning the reference voltage, the turn-off command is sent to the switch by the driver IC. Then, the switch is softly turned off, and the current is cut off through a mechanism described in section II. The design engineer's responsibility is to make reasonable adjustments to the SC detection threshold. An example of DESAT circuit integrated with industrial gate drives ICs is shown in Fig. 4. Some commercial example gate driver ICs of this category are HCPL316, ACPL-330J, ACPL331, ACPL336, FOD8316, etc. They have a 200 ~ 300 μA internal current source to charge the capacitor Cb1 to a voltage of VCE-VDESAT linearly. The capacitor voltage is continually compared with the internal Vref (usually 6.5V). The designer can adjust the circuit threshold for different switches by setting the VDESAT (using a Zener diode or several series diodes). According to Fig.4, resistor R1 and capacitor Cb1 filter out noises coming from the power switch side. The filter transfer function is expressed as (4).

$$V_{DESATF} = \frac{1}{1+R_1C_{b1}s}V_{DESAT} \quad (4)$$

where V_{DESATF} is the filtered value of voltage V_{DESAT}.

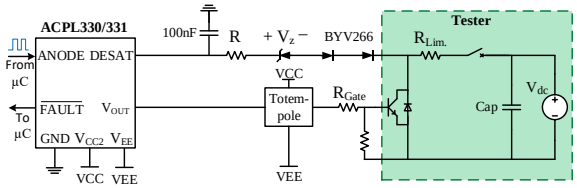


Fig. 7. Schematic of ACPL-330J and ACPL331 gate drive circuit

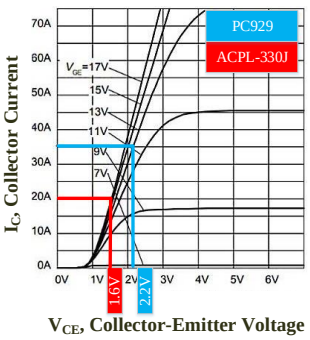


Fig. 8. V_{CE}-I_{CE} characteristic of IKW25T120 IGBT.

IV. DESIGN EXAMPLES AND THE PROPOSED TEST CIRCUIT

In this section, some commercial gate drives, including ACPL330J, ACPL331, and PC929, have been experimentally evaluated for further investigation. A general test circuit is proposed to determine if the designed SC protection can detect SC and shut down the switch correctly. The designed board for SC test condition, is shown in Fig. 5. In this board, the gate drive of high side PSD is the ACPL-330J, and the low side is ACPL-331J. On another similar board, the gate drive of PSD is PC929. The PSDs is an IGBT 1200V, 25A (IKW25T12). Two power supplies supply the gate driver (+15V and -9V). A signal generator is used to generate the switching commands for IGBT. To test each gate drive, SC has another IGBT, and with the gate command, the IGBT of the gate drive under test is placed in an SC condition (by discharging the charged capacitor). The amount of SC current as (5) can be adjusted through the series resistance (RLim) and the voltage of the charged capacitor (Vdc).

$$I_{Lim} \simeq \frac{V_{dc}}{R_{Lim}} \quad (5)$$

The SC tests are performed on the experimental system. The DC link uses a capacitive bank connected to the power supply via a relay. This capacitive bank discharges current into the IGBT, providing an intentional SC situation. The experimental test bench is shown in Fig. 6.

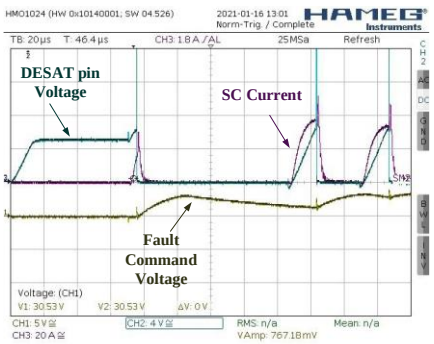


Fig. 9. ACPL-330J gate drive SC test.

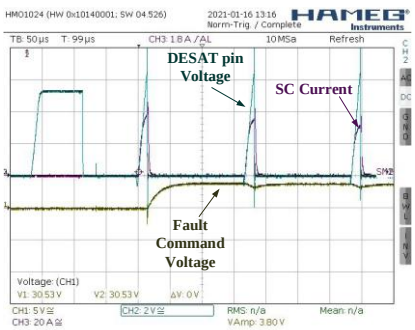


Fig. 10. ACPL-331J gate drive SC test.

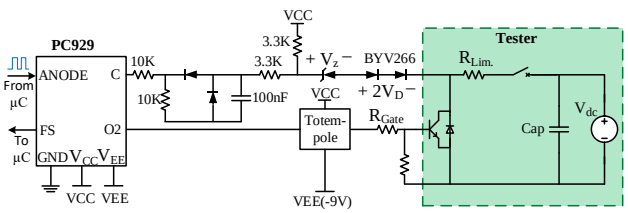


Fig. 11. Schematic of PC929 gate drive circuit

A. *ACPL-330J gate drive analysis*

ACPL-330J is used in the proposed test circuit, according to Fig. 7, to check its performance. DC link is charged, and the IGBT is commanded to start conduction, resulting in an SC condition. The gate drive must protect IGBT. In the designed circuit, the value of the Zener in the DESAT pin is set to 4.7V. According to Fig. 7 with a KVL in DESAT pin, (6) is obtained.

$$V_{DESAT,Fault} = R \times I_{CHG} + V_Z + 2 \times V_D + V_{CE}$$
$$6.5V = 100\Omega \times 250\mu A + 4.7V + 2 \times 0.1V + V_{CE} \quad (6)$$
$$V_{CE} = 1.6V$$

where VDESAT,Fault, and ICHG are internal fault reference voltages of IC and internal current sources. Also,

Vz and VD are Zener and diode voltages. According to the IKW25T12 IGBT (VCE-IC) curve (Fig.8), at about 20A to 30A, the IGBT must be turned off by the ACPL-330J. As shown in Fig.9, the current (purple waveform) is limited to around 30A when a short circuit occurs. In addition, the DESAT pin voltage (green waveform) is reached around 6.5V, and protection is activated.

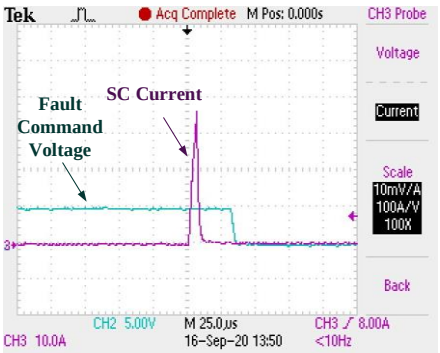


Fig. 12. PC929 gate drive SC test.

Table II. COMPARISON TABLES OF ACPL330, ACPL331, AND PC929 GATE DRIVES

Gate Driver	Desat Protection	Reset Type	Fault Command	Dependency of Protection on Voltage Supply
ACPL330	Yes	Automatic Reset After 25us	Isolated	No
ACPL331	Yes	Automatic Reset After First Pulse	Isolated	No
PC929	Yes	Automatic Reset After First Pulse	Isolated	Yes

The fault command (yellow waveform) also goes high and signals the SC conditions to the microcontroller. For ACPL-330J, the driver will automatically reset the FAULT pin after a fixed mute time of 25µs (typical). As a result, the SC conditions are repeated if the fault is not cleared during this mute time, as shown in Fig. 9.

In the first pulse, the applied SC fault is interrupted and then re-applied. In the last two pulses of Fig. 9, the SC fault (connected charged capacitor) remains on the system, and as it is clear, after 25µs, the current increases again.

B. ACPL-331J gate drive analysis

An intentional SC condition is created by replacing ACPL-331J gate drive IC in the proposed circuit, charging the DC link, and connecting the IGBT. The gate drive must protect IGBT. Zener setting in the DESAT pin is similar to one mentioned for the ACPL-330J. Similarly, as shown in Fig.10, the current (purple waveform) is limited to about 30A. Also, the DESAT pin voltage (green waveform) reaches about 6.5V, and protection is activated. The fault command (yellow waveform) also goes to a high level and indicates the SC conditions of the microcontroller. The next turn-on command can reset the fault mechanism after 5µs (minimum) mute time. As a result, the SC conditions are repeated if the fault is not cleared during the mute time between two turn-on commands, as shown in Fig. 10.

C. PC929 gate drive analysis

The structure of the test PC929 circuit is shown in Fig. 11. According to the PC929 datasheet, reaching the voltage of pin 9 to about V_{cc}-6V means an SC condition has happened, and the protection system operates. After the SC condition is detected, pin 8 (F_s) goes to low-level voltage (0V), and thus, the error is sent to the microcontroller. According to Fig.11, in the PC929 circuit, two diodes and a 6.5V zener are used. Also, the PC929 operates with a V_{cc}=25V, which, according to (4), the protection voltage is calculated as (7).

$$V_{cc} - 6V = V_Z + 2\times V_D + V_{CE} - V_{Gate-negative}$$
$$\rightarrow 25 - 6 = 65 + 2\times 0.6 + V_{CE} + 9 \rightarrow V_{CE} = 2.2V \quad (7)$$

According to the IKW25T12 IGBT (VCE-IC) curve (Fig.8), the current of VCE=2.2V is about 40A. Therefore, the IGBT must be turned off by the PC929 at about 40A. As shown in Fig.12, the current (purple waveform) is limited to about 40A. Also, the fault command (blue waveform) goes to a low level, indicating the short-circuit conditions of the microcontroller.

D. Dependence on supply voltage

Unlike ACPL-330J and ACPL-331J, PC929 gate drive SC protection level depends on the supply voltage (Vcc), which is a disadvantage.

In industrial power electronic circuits, the gate drive power supply may have a 10% tolerance, which can negatively affect the protection performance of PC929.

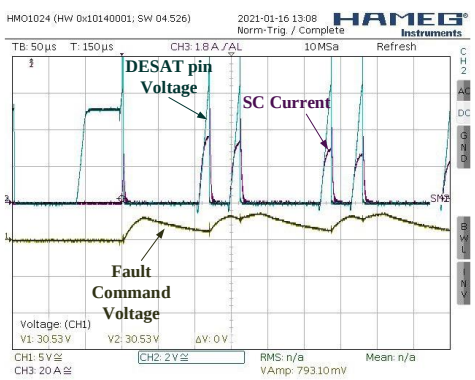


Fig. 13. SC test with 13.1V power supply for ACPL-330J.

TABLE III. COMPARISON OF THE SPECIFICATIONS OF THE MOST WIDELY USED INDUSTRIAL GATE DRIVES

Gate Drive	Peak Current	DESAT	UVLO	Reset Type
FOD8316	2.5A	Yes	Yes	Reset Pin
ACPL330	1.5A	Yes	Yes	Automatic Reset After 25us
ACPL331	1.5A	Yes	Yes	Automatic Reset After First Pulse
ACPL336	2.5A	Yes	Yes	Automatic Reset After First Pulse
HCPL316	2.5A	Yes	Yes	Reset Pin
PC929	0.4A	Yes	No	Automatic Reset After First Pulse
MC33153	1A	Yes	Yes	Automatic Reset After First Pulse
TD350	1.5A	Yes	Yes	Automatic Reset After First Pulse
2ED020I12	2A	No	Yes	-
Gate Drive	**Fault Command**	**Isolation**	**Dependency of Protection on Supply**	**Price**
FOD8316	Isolated	Yes	No	4$
ACPL330	Isolated	Yes	No	6$
ACPL331	Isolated	Yes	No	5.3$
ACPL336	Isolated	Yes	No	6$
HCPL316	Isolated	Yes	No	5.5$
PC929	Isolated	Yes	Yes	3.1$
MC33153	Non-Isolated	No	No	1.3$
TD350	Non-Isolated	No	No	2.2$

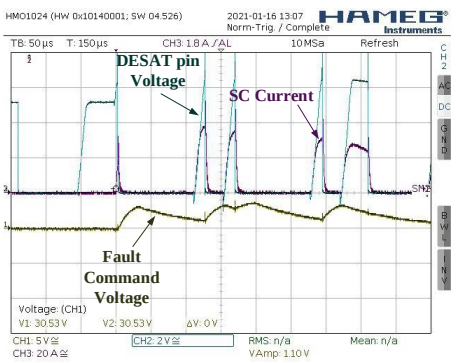


Fig. 14. SC test with 17.5V power supply for ACPL-330J.

To check the dependency of SC protection of ACPL-330J and ACPL-331J to supply voltage (Vcc), the voltage is set once to 13.1V and again to 17.5V. The SC protection operation is investigated. It is seen from Fig. 13 and Fig. 14, SC protection in ACPL-330J functions well independently from the supply. For a quick comparison, Table II summarizes the important characteristics of the compared gate drives (ACPL330. ACPL331, and PC929).

E. Comparing the important characteristics of widely used industrial gate drivers

This section provides a qualitative comparison based on widely used industrial gate drivers. Table III presents from the point of view of gate drive power, protection specifications (DESAT protection, UVLO protection, reset type, and fault command), isolation, and price. This comparison allows users to choose based on their desired application.

V. Conclusions

This paper presents a comprehensive overview of the industrial specifications of gate drives. This paper focuses on the gate drive of silicone power switches (Si IGBT and MOSFET). One of the important features of IGBT and MOSFET gate drives is short circuit protection. Gate drives with SC protection capability are prioritized more in industrial applications than in ordinary ones. As design examples, three commercial gate drive ICs are considered. A test circuit is proposed for experimentally evaluating SC protection's proper design and functioning. The dependency of the protection on the supply voltage of the ICs is also investigated. This dependency is a significant disadvantage for a gate drive. It may result in the malfunctioning of the SC protection and system failure due to a small supply voltage variation. As a result of the comparison between the three ICs, ACPL-330J and ACPL331 are better choices than PC929 for power electronic applications.

REFERENCES

[1] Disney, Don, and Z. John Shen. "Review of silicon power semiconductor technologies for power supply on chip and power supply in package applications." IEEE Transactions on Power Electronics 28.9 (2013): 4168-4181.

[2] Hernes, Magnar, et al. "Failure analysis and lifetime assessment of IGBT power modules at low temperature stress cycles." IET Power Electronics 14.7 (2021): 1271-1283.

[3] Mocevic, Slavko, et al. "Comparison and discussion on shortcircuit protections for silicon-carbide MOSFET modules: Desaturation versus Rogowski switch-current sensor." IEEE Transactions on Industry Applications 56.3 (2020): 2880-2893.

[4] Min, Sung-Soo, and Rae-Young Kim. "Improved gate-voltage-driven desaturation short-circuit protection circuit with robust switching noise immunity for WBG power semiconductors." IEEE Journal of Emerging and Selected Topics in Power Electronics (2022).

[5] Phukan, Ripun, Lixiang Wei, and Jiangang Hu. "A low profile gate drive power supply." 2019 IEEE Applied Power Electronics Conference and Exposition (APEC). IEEE, 2019.

[6] Zoubir, Abdefihak M., and D. Robert Iskandler. "Bootstrap methods and applications." IEEE Signal Processing Magazine 24.4 (2008): 10-19.

[7] Yin, Shan, et al. "Design considerations and comparison of high-speed gate drivers for Si IGBT and SiC MOSFET modules." 2016 IEEE Energy Conversion Congress and Exposition (ECCE). IEEE, 2016.

[8] IEC, IEC. "61800 adjustable speed electrical power drive systems–part 5-2: safety requirements–functional." (2007).

[9] Anurag, Anup, et al. "Gate drivers for medium-voltage SiC devices." IEEE Journal of Emerging and Selected Topics in Industrial Electronics 2.1 (2020): 1-12.

[10] Batra, Tushar, et al. "Isolation design considerations for power supply of medium voltage silicon carbide gate drivers." 2017 IEEE Energy Conversion Congress and Exposition (ECCE). IEEE, 2017.

[11] Radmehr, Mahdi, and Mohammad Mojibi. "Reliability assessment and thermal consideration of a step-down DC/DC converter." International Journal of Industrial Electronics

[12] Peng, Yingzhou, and Huai Wang. "A simplified on-state voltage measurement circuit for power semiconductor devices." IEEE Transactions on Power Electronics 36.10 (2021): 10993-10997.

[13] Horiguchi, Takeshi, et al. "Short-circuit protection method based on a gate charge characteristic." IEEJ Journal of Industry Applications 4.4 (2015): 360-369.

[14] Jiao, Chaoqun, et al. "Integrated Rogowski coil sensor for press-pack insulated gate bipolar transistor chips." Sensors 20.15 (2020): 4080.

[15] Yan, Wenyi, et al. "Variable Gate Resistance Drive Circuit Based on di/dt Feedback for IGBT." Journal of Physics: Conference Series. Vol. 2492. No. 1. IOP Publishing, 2023.

[16] Chen, Min, et al. "An improved IGBT short-circuit protection method with self-adaptive blanking circuit based on V CE measurement." IEEE Transactions on Power Electronics 33.7 (2017): 6126-6136.

[17] Wittig, Bjoern, Matthias Boettcher, and Friedrich W. Fuchs. "Analysis and design aspects of a desaturation detection circuit for low voltage power MOSFETs." Proceedings of 14th International Power Electronics and Motion Control Conference EPE-PEMC 2010. IEEE, 2010.

[18] Schlegel, David W., et al. "Characteristics, selection guidelines and performance of circuit protection devices for ASDs." 2013 IEEE Industry Applications Society Annual Meeting. IEEE, 2013.

[19] Karbalaei, Alireza, and Mohammad Mardaneh. "Improved symmetric switched-inductor/capacitor quasi Z-source inverter with ability uplifted-boost." International Journal of Industrial Electronics Control and Optimization 3.1 (2020): 47-58.

[20] Cho, Min-Hyeong, et al. "Development of undervoltage lockout (UVLO) circuit configurated Schmitt trigger." 2015 International SoC Design Conference (ISOCC). IEEE, 2015.

[21] Bolloju, Vijay, and Jun Yang. "Influence of short circuit conditions on IGBT short circuit current in motor drives." 2011 Twenty-Sixth Annual IEEE Applied Power Electronics Conference and Exposition (APEC). IEEE, 2011.

[22] Khalilzadeh, M., and A. Fereidunian. "A Markovian approach applied to reliability modeling of bidirectional DC-DC converters used in PHEVs and smart grids." Iranian Journal of Electrical & Electronic Engineering 12.4 (2016): 301.

[23] Yin, Shan, et al. "Comparative design of gate drivers with short-circuit protection scheme for SiC MOSFET and Si IGBT." Energies 12.23 (2019): 4546.

[24] Dalal, Dipen Narendra, et al. "Gate driver with high common mode rejection and self turn-on mitigation for a 10 kV SiC MOSFET enabled MV converter." 2017 19th European Conference on Power Electronics and Applications (EPE'17 ECCE Europe). IEEE, 2017.

[25] Pluschke, Norbert, and Niklas Hofstoetter. "How to turn off IGBTs with Unipolar Voltage (0V)?." PCIM Asia 2015; International Exhibition and Conference for Power Electronics, Intelligent Motion, Renewable Energy and Energy Management. VDE, 2015.

[26] Asad, Mohsin, Amit Kumar Singha, and Ravada Madhu Sudhan Rao. "Dead time optimization in a GaN-based buck converter." IEEE Transactions on Power Electronics 37.3 (2021): 2830-2844.

[27] Aeloiza, Eddy, Arun Kadavelugu, and Rostan Rodrigues. "Novel bipolar active miller clamp for parallel SiC MOSFET power modules." 2018 IEEE Energy Conversion Congress and Exposition (ECCE). IEEE, 2018.

[28] Zhang, Fan, et al. "Advanced active gate drive for switching performance improvement and overvoltage protection of high-power IGBTs." IEEE Transactions on Power Electronics 33.5 (2017): 3802-3815.

[29] Raki, A., Iman-Eini, H., Monfared, K. K., & Ahmadi, S. (2020, February). A Semi-Controlled Soft Charge Rectifier for Medium/High Power AC Drives: design and implementation. In 2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC) (pp. 1-6). IEEE.

Saeid Ahmadi received the B.Sc. and M.Sc. degrees in electrical engineering with honors from the University of Tehran, Tehran, Iran, in 2018 and 2020, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at University of Tehran, Tehran, Iran. His current research interests include advanced control for modular multilevel converters, and HVDC systems.

Kourosh Khalaj Monfared received the B.Sc. degree in electrical engineering from Shahid Beheshti University, Tehran, Iran, in 2015, M.Sc. and Ph.D. degrees in electrical engineering from the University of Tehran, Tehran, Iran, in 2017 and 2022, respectively. He is currently an Assistant Professor of electrical engineering at the School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran. His current research interests include advanced control in power electronics, multilevel converter applications, renewable energy systems, and pulsed power technology.

Mohammad Khalilzadeh was born in Urmia, Iran, in 1990. He received the B.Sc. and Ph.D. degrees in electrical engineering from the University of Tehran, in 2012 and 2021, respectively, and the M.Sc. degree from the K. N. Toosi University of Technology, in 2014. He is currently an Assistant Professor with Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran . His research interests include power electronics and motor drives.

Hossein Iman-Eini (Senior Member, IEEE) received the Ph.D. degree in electrical engineering jointly from the University of Tehran, Tehran, Iran, and Grenoble Alpes University, Grenoble, France, in 2009. He is a Professor of electrical engineering with the School of Electrical and Computer Engineering, University of Tehran. He currently follows a Heisenberg position as the Chair of Power Electronics with Kiel University, Kiel, Germany. His research interests include power electronics and applications of power electronics in power systems/grids.

IECO

This page intentionally left blank.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Addressing Dependent Data in Constrained Optimization Problems: A WOA-based Algorithm

Asieh Ghanbarpour[1] | Soheil Zaremotlagh[2] | Fahimeh Dabaghi-Zarandi[3]

Faculty of Electrical and Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran.[1]
Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran.[2]
Computer Engineering Department, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.[3]
Corresponding author's email: ghanbarpour@ece.usb.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 31-Dec-2023
Received in revised form: 23-March-2024
Accepted:03-April-2024
Published online: 21-June-2024

Keywords:
Compositional data,
Constraint,
Optimization,
Penalty function,
Swarm optimization.

ABSTRACT

Optimization methods are widely used in various fields to find the best possible solution to a given problem through the minimization or maximization of an objective function while adhering to specific constraints. In this paper, we present a new optimization algorithm, called WOADD, which was designed to handle the challenges of constrained optimization problems that involve dependent data. Unlike traditional algorithms that struggle with data dependencies and valid range constraints, WOADD uses a unique normalization process and a dynamic updating mechanism that accurately considers the interdependencies among features. By calculating a scaling parameter to navigate within feasible regions, WOADD adjusts its search strategy to ensure the preservation of data dependencies and adherence to constraints, ultimately leading to more efficient and precise optimization outcomes. Our extensive experimental analysis, which compared WOADD against other swarm optimization methods using a suite of benchmark functions, demonstrates its superior performance in terms of faster convergence rates, improved solution quality, and enhanced determinism in outcomes.

I. Introduction

Optimization methods are widely used in diverse fields, such as engineering, finance, and data science, with the aim of finding the best possible solution to a problem while meeting certain constraints [1-6]. These methods can be broadly classified into two categories: traditional optimization methods [7, 8] and swarm optimization methods [9, 10]. Traditional optimization methods, such as gradient descent [8] and Newton's method [11], rely on the derivatives of the objective function to approach the optimal solution iteratively. These methods are efficient for simpler problems, but may struggle in complex, high-dimensional or noisy scenarios. On the other hand, swarm optimization methods [1, 5, 9, 10, 12-16] are better suited for such complex problems as they explore the search space more extensively, but they can be computationally expensive and require careful parameter tuning for optimal performance.

According to [17], dependent data are "data where the observations are not independent of each other, and hence, the standard statistical methods that assume independence may not be applicable". Dependent variables are common in real-world scenarios across various fields including geochemistry (e.g. mineral compositions) [18, 19], Biology (e.g. microbiome research [20, 21]), tourism (e.g., distribution, share, allocation) [22, 23], society science (e.g. personal behavior) [24, 25], political science (e.g., voting proportions) [26], marketing (e.g. sale data) [27, 28], and Bioinformatics [29].

Assigning weights to the parameters of a function or model is a common problem defined on dependent data. This problem is known as feature weighting in machine learning, and it aims to identify important features and assign weights to the model parameters accordingly [30]. In linear regression, the weights assigned to the predictor variables determine their relative importance in predicting the outcome variable [31]. In neural networks, the weight assigned to each connection between neurons determines the strength of the connection and can significantly impact the network's performance [32, 33]. Optimizing the weights for parameters can be quite a challenge, particularly when dealing with high-dimensional data. One possible solution to address this issue is to use optimization

algorithms [1, 3, 9, 10, 12]. However, it's worth noting that optimization algorithms are primarily designed for unconstrained data, where variables can change independently. To handle constrained data, such as those with dependencies, optimization algorithms employ constraint-handling techniques such as penalty functions, repairing infeasible solutions, and multi-objective optimization.

In this research paper, we present a new method for handling dependent data in optimization systems. Our study shows that penalty functions often lead to convergence problems and suboptimal solutions. Instead, we propose a novel approach that uses the Whale Optimization Algorithm (WOA) to seamlessly integrate constraints into the optimization process. Our algorithm, called WOADD, normalizes the data and dynamically updates the search strategy to consider the interdependencies among the features. It calculates a scaling parameter to maneuver within feasible regions and ensure the preservation of data dependencies and adherence to constraints. Our method has been found to have superior performance in handling optimization problems characterized by dependent data, and it shows potential for more effective and efficient optimization solutions.

Throughout the remainder of this paper, we review related works in Section II, examine the problem statement in Section III, present the proposed method in Section IV, analyze experiments in Section V, and conclude our findings in Section VI.

II. Literature Review

Evolutionary Algorithms (EAs) often struggle to handle constraints directly during optimization in many problems. Various techniques have been used in EAs to solve constrained optimization problems, including penalty functions, repairing infeasible solutions, and multi-objective optimization. An overview of the literature on applying EAs to constraint optimization algorithms is presented in [34].

A penalty function is a method that converts a Constrained Optimization Problem (COP) into an unconstrained one by adding or subtracting a certain value to or from the objective function based on the number of constraint violations present in a given solution [35, 36]. However, determining the appropriate penalty factor for a problem can be challenging, because selecting an excessively high or low penalty can lead to issues [35]. Another group of techniques involves repair algorithms, which aim to transform infeasible solutions into feasible ones by applying specific procedures or mechanisms [37, 38]. Although repair algorithms are useful when converting infeasible solutions to feasible ones is computationally inexpensive, they may introduce biases and adversely impact the evolutionary process. Furthermore, designing a repair algorithm for each problem is a problem-dependent requirement.

A different way of dealing with constraints involves transforming a COP into a problem with m + 1 objectives, where m is the total number of constraints. This problem can then be solved using a multi-objective evolutionary algorithm [39] [40-43]. However, applying multi-objective optimization is not a simple process and requires careful consideration of various factors. One popular approach is to minimize both the original objective function and the degree of constraint violation simultaneously. The paper [39] used a modified non-dominated sorting genetic algorithm after transforming a COP into multi-objective optimization. On the other hand, [40] employed a Pareto-ranking-based Differential Evolution (DE) algorithm. Li et al. [41] addressed many-objective optimization problems through dynamic constraint handling, and [42] converted a COP into a dynamical bi-objective optimization problem. In [43], a multi-objective optimization problem is decomposed into sub-problems using the weighted sum approach, and dynamic adjustments to the weights enable each sub-problem to converge towards an equivalent objective. Recent advancements include techniques such as feasible-ratio control and dynamic constraint boundaries [44] to enhance EAs with dynamic constrained multi-objectives.

III. Problem Statements

A. Problem Definition

Let $\mathfrak{R}^D$ be the feasible region of the search space, contained within the D-dimensional rectangular space S. Denote $X = [X_1, \dots, X_f, \dots, X_D]$ as the feasible solution vector, where X is an element of $\mathfrak{R}^D$, where $\mathfrak{R}^D \subseteq S$. Each component of the feasible solution vector X, denoted as X_f, must satisfy the upper and lower boundary limits X_{fL} and X_{fU}, respectively, where $X_{fL} \le X_f \le X_{fU}$. The objective function value of solution X is defined as $f(X)$, while the functions $g_j(X)$ and $h_j(X)$ represent the *j*-th equality constraint and inequality constraint, respectively. A COP is defined as follows:

$$\min_{X=[X_1,\dots,X_f,\dots,X_D]\in\mathfrak{R}^D} f(X) \quad s.j.\begin{cases} g_j(X) \le 0, j = 1, \dots, p \\ h_j(X) = 0, j = 1, \dots, q \end{cases} \tag{1}$$

, where p and q represent the numbers of inequality and equality constraints of problem respectively.

In the problem of assigning weights to the parameters of a function, the objective is to assign weights to different parameters of a function such that the sum of the weights equals a constant κ, and each weight falls within the range of zero to one. This form of optimization problem can be defined as follows:

$$\min_{X=[X_1,\dots,X_f,\dots,X_D]\in\mathfrak{R}^D} \sum_{f=0}^{n} X_f \times V_f \quad s.j\begin{cases} X_f \ge 0, \ \ f = 0, \dots, n \\ \sum_{f=0}^{n} X_f = \kappa \end{cases} \tag{2}$$

where X_f shows the weighting parameter for $0 \le f \le n$, and V_f is the value of the element which should be weighted by X_f.

B. Whale Optimization Algorithm (WOA)

WOA is an algorithm that follows the social behavior of humpback whales. It works by initializing a set of random solutions (whales) and updating their positions based on whether they are in the exploration or exploitation phase. In

the beginning, each whale is assigned random values. The search process for finding the optimal solution involves two phases: exploration and exploitation. The exploitation phase models the bubble-net behavior of humpbacks when attacking prey. This behavior is modeled with two types of movement: shrinking encircling and spiral updating position. To encircle prey, the positions of search agents are updated towards the best search agent that is assumed to be the target or near the target. This update of the search agents is represented by the following equations, as in [45]:

$$\vec{D} = \left|\vec{C}.\overrightarrow{W^*}(t) - \vec{W}(t)\right| \quad (3)$$

$$\vec{W}(t+1) = \overrightarrow{W^*}(t) - \vec{A}.\vec{D} \quad (4)$$

where $\vec{W}(t)$ and $\vec{W}(t+1)$ shows the configuration of an search agent in the i-th iteration and (i+1)-iteration respectively, $\overrightarrow{W^*}(t)$ shows the best search agent regognized until iteration t. Vectors $\vec{A}$ and $\vec{C}$ are coefficient vectors which are calculated as follows:

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \quad (5)$$

$$\vec{C} = 2.\vec{r} \quad (6)$$

During iterations, the linearly decreasing vector $\vec{a}$ is reduced from 2 to zero, while the k-dimensional vector $\vec{r}$ has random values within the range of [0, 1]. The method for achieving shrinking encircling involves decreasing the $\vec{a}$ in Eq. (5). Since $\vec{r}$ is a random vector within the range of [0, 1], the values of $\vec{A}$ fall within the range of $[-a, a]$. This model demonstrates how the search agents move within hyper-cubes towards the current best solution.

The spiral updating position simulates the helix-shaped movement of humpback whales to exploit prey. This behavior is modeled as follows:

$$\vec{W}(t+1) = \overrightarrow{D'}.e^{bl}.\cos(2\pi l) + \overrightarrow{W^*}(t) \quad (7)$$

, where b is a constant which is set to one, and l is a random value in $[-1,1]$. Vector $\overrightarrow{D'}$ shows the distance between the search agent and the best current solution (the prey) and is defined as follows:

$$\overrightarrow{D'} = \left|\overrightarrow{W^*}(t) - \vec{W}(t)\right| \quad (8)$$

Humpback whales use two hunting strategies shrinking encircling and spiral updating positions together to efficiently catch their prey. To mimic this behavior, a new parameter (p) is introduced that randomly switches between the two strategies by taking on a random value between 0 and 1. During the exploitation phase, the configuration is updated using the following process:

$$\vec{W}(t+1) = \begin{cases} \overrightarrow{W^*}(t) - \vec{A}.\vec{D} & if\ p < 0.5 \\ \overrightarrow{D'}.e^{bl}.\cos(2\pi l) + \overrightarrow{W^*}(t) & if\ p \geq 0.5 \end{cases} \quad (9)$$

During the exploration phase, the search agents update their positions based on a randomly selected search agent $\overrightarrow{W^{rand}}(t)$ as follows:

$$\vec{D} = \left|C.\overrightarrow{W^{rand}}(t) - \vec{W}(t)\right| \quad (10)$$

$$\vec{W}(t+1) = \overrightarrow{W^{rand}}(t) - \vec{A}.\vec{D} \quad (11)$$

IV. Methodology

A. Normalization

Optimization algorithms often use randomly generated initial populations. However, when dealing with dependent data, the features of an object are interdependent. To maintain the dependency of the features while still generating a random initial population, a two-step process is proposed for individual initialization. Firstly, the population is randomly generated. Then, each individual is normalized to preserve the data's dependency.

Let $X = [X_1, \dots, X_f, \dots, X_D]$ represents a solution vector, where X is an element of $\Re^D$. Furthermore, let $\varkappa = \{X_p, \dots, X_q\}$, where $\varkappa \subseteq X$ shows a set of features of X that are dependant on each other according to the condition $M_c(\varkappa)$. This condition is an equality condition that takes the form of a sum function similar to those which are common in the weighting problems. To normalize Individual X based on $M_c(\varkappa)$, each $X_i \in \varkappa$ is first subtracted by the minimum value in the set $\varkappa$ as follows:

$$X_i' = X_i - X_{min} \quad (12)$$

By setting the value of $M_c(\varkappa)$ as α, Eq.(13) is used to transform the independent variables of $\varkappa$ into dependent variables [46].

$$(X_i)_N = \frac{\alpha}{\sum_{X_i \in \varkappa} X_i'} \times X_i' \quad (13)$$

If there exists equality conditions $M_{c_1}(\varkappa_1), M_{c_2}(\varkappa_2), \dots, M_{c_k}(\varkappa_k)$ for sets $\varkappa_1, \varkappa_2, .., \varkappa_k$, and these sets have no overlap, then each equality condition can be applied independently to the individuals of the population to obtain individuals that are within the valid domain of the search.

B. Updating Positions

In the WOA, the optimized solution is determined based on the leader individual, which is the best individual in all iterations. During iterations, the position of each individual X_i is updated based on three movements: two based on the position of the leader individual and the third based on the position of a randomly selected individual. While features of an individual are updated independently in the WOA, a filtering step is executed after updating positions to eliminate individuals that don't satisfy constraints for dependent data or other constraints. When the features of an individual are dependent, changes made to one feature may impact the other features. In such cases, because it is unlikely to generate a valid solution through random changes to the features, many solutions will be eliminated during the filtering stage. Therefore, it is crucial to consider this dependency when updating the individuals.

The Whale Optimization Algorithm for Dependent Data (WOADD) employs an appropriate policy for updating data that considers the dependence between features to avoid unnecessary processing. The changes are precisely controlled to preserve the equality constraints, and the features collaborate to satisfy these constraints.

The change in the value of a feature is entirely dependent on the type of whale movement in the corresponding iteration. The magnitude of change in feature $X_i \in X$ is calculated using the following formulas depending on whether it's a

circular movement, spiral movement, or exploratory movement:

$$\Delta X_{circular} = A_i \times |C_i \times X_i^* - X_i| \quad (14)$$

$$\Delta X_{spiral} = |X_i^* - X_i| \times e^{bl} \times \cos(2\pi l) \quad (15)$$

$$\Delta X_{expl} = A_i \times \left|C_i \times X_i^{rand} - X_i\right| \quad (16)$$

Where X_i shows the i-th feature of normalized individual X, b is a constant for defining the shape of the logarithmic spiral, and l is a random number in [-1,1]. The values of A_i and C_i show the i-th element of vectors $\vec{A}$ and $\vec{C}$ calculated by (5) and (6), respectively.

The change of feature $X_i \in \varkappa$ may violate the equlity condition $M_c(\varkappa)$. In this case, any change in one feature is completed with a candidate one. Let $\mathcal{L} \subseteq \varkappa$ shows the set of features that haven't been changed in the current iteration. The members of $\mathcal{L}$ are those that could be chosen as candidates. To ensure equal chances of being chosen, each member of $\mathcal{L}$ is assigned a random probability, and then sorted in descending order based on their random number. The member with the highest probability is checked first for candidacy. Let ΔX represent the amount of change of feature $X_i \in \varkappa$, and X_c represent the feature being considered as a candidate. There are three possible situations that may occur:

- If $\Delta X \geq 0$, an overflow has occurred in set $\varkappa$ based on the equality constraint. To resolve this issue, the candidate element should have the ability to reduce this additional load. If $(X_c - \Delta X)$ satisfies the defined constraints on X_c, then there will be a load transfer between X_i and X_c. If not, the next candidate feature is checked.
- If $\Delta X < 0$, an underflow has occurred in set $\varkappa$. Therefore, the candidate element should have the potential to carry additional load. If $(X_c + |\Delta X|)$ is validated according to the defined constraints, X_c is selected as the candidate for X_i. Otherwise the other potential candidates are checked.
- If a candidate cannot be found for X_i, it is returned to its previous value.

In each iteration, the features in an individual are updated at most once, and a flag is reserved for each feature to indicate its status. At the start of each iteration, all features of an individual are marked as unchanged. If a feature is modified or checked for an update and no candidate is found, it is marked as changed. Load transfers are continued until all the involved features are marked as changed.

In order to avoid exceeding the valid range of features, a parameter is introduced to limit the range of change for each feature. This parameter's value is determined based on the constraints of the problem and the specific requirements of the application. For simplicity, let all $X_i \in \varkappa$ satisfy $L \leq X_i \leq U$. The mathematical formulation of circular and exploratory movements are similar in that the position of an individual is updated based on a goal individual X^G which can be either the leader individual or a randomly selected individual. Let X_i^G denote the value of the i-th feature in X^G. The next value of X_i in the circular and exploratory movements is obtained as follows:

$$X_i(t+1) = X_i^G - \Delta X_i \quad (17)$$

, where the inequality conditions $L \leq X_i(t+1) \leq U$ should be met. Since $\Delta X_{expl} \geq 0$ and $\Delta X_{circle} \geq 0$ and X^G is an individual in the valid search space, the unequality $X_i(t+1) \leq U$ is always true. To have more control on the movement of objects, we consider a scaling parameter $\mu > 0$ on the change values, and solve the following unequality:

$$X_i^G - \mu \Delta X_i \leq U \quad (18)$$

This implies that:

$$\mu \leq \frac{X_i^G - L}{A\left|CX_i^G - X_i\right|} \quad (19)$$

On the other hands, when completing the change of X_i with candidate X_c, the following unequalities should be satisfied for X_c:

$$L \leq X_c + \mu \Delta X_i \leq U \quad (20)$$

Since $\Delta X_{expl} \geq 0$ and $\Delta X_{circle} \geq 0$ and X_c is in the valid range, the unequality $L \leq X_c + \mu \Delta X_i$ is always true and the unequality $L \leq X_c + \mu \Delta X_i$ should be examined. Solving this implies that:

$$\mu \leq \frac{U - X_c}{A\left|CX_i^G - X_i\right|} \quad (21)$$

Both the equations (19), (21) should be satisfied, therefore μ is determined as follows:

$$\mu \leq \frac{1}{A\left|CX_i^G - X_i\right|} \min\left(X_i^G - L\,, U - X_c\right) \quad (22)$$

As an example, if any variable falls within the range of [0,1], the maximum value of A and C is 2, and the maximum distance between two features of two whales is 1. Therefore, the range of μ would be as follows:

$$0 < \mu \leq \min\left(\frac{X_i^G}{4}, \frac{1 - X_i^G}{4}\right) \quad (23)$$

In the spiral movement, the sign of $\Delta X_{spriral}$ is relative to the sign of $\cos(2\pi l)$. If the value of feature X_i after spiral movement be equal to $\left(X_i^* + \mu \Delta X_{spiral}\right)$, the following unequalities should be met:

$$L \leq X_i^* + \mu(|X_i^* - X_i| \times e^{bl} \times \cos(2\pi l)) \leq U \quad (24)$$

By solving the inequalities, μ is calculated as follows:

$$\begin{cases} \mu \leq \dfrac{U - X_i^*}{|X_i^* - X_i| \times e^{bl} \times \cos(2\pi l)} & \text{if } \cos(2\pi l) \geq 0 \\ \mu \leq \dfrac{X_i^* - L}{|X_i^* - X_i| \times e^{bl} \times |\cos(2\pi l)|} & \text{if } \cos(2\pi l) < 0 \end{cases} \quad (25)$$

To ensure that the given set of inequalities are satisfied, μ is chosen accordingly.

$$0 < \mu \leq \frac{1}{|X_i^* - X_i| \times e^{bl} \times |\cos(2\pi l)|} \min(U - X_i^*, X_i^* - L) \quad (26)$$

It is noted that the candidate feature $X_c$in a spiral movement is changed to $\left(X_c - \mu \Delta X_{spiral}\right)$ which must fall within the range of L and U. However, this constraint also leads to the same range for μ as given in equation (26). The pseudo-code of the proposed algorithm is shown in Figure 1.

Algorithm 1: WOADD

Initialize the Whale population $W_i, i = 1, \dots, n$
Normalize the initial population
Initialize a, A, C, b , l
while $t < number\ of\ iteration$ **do**
  Calculate the fitness of each Whale
  Select the best Whale as the Leader Whale
  for $each\ Whale\ (X)$ **do**
    Randomly initialize r, p
    Update a , A , C, μ
    for each feature (X_i) in X **do**
      if ($p < 0.5$ and $|A| < 1$) **do**
        Calculate ΔX by (14)
      if ($p < 0.5$ and $|A| \geq 1$) **do**
        Select a random whale (X^{rand})
        Calculate ΔX by the (15)
      if ($p \geq 0.5$) **do**
        Calculate ΔX by the (16)
      X_c =Find a candidate for X_i
      if ($X_c! = null$)
        $X_i = X_i + \mu\Delta X$
        $X_c = X_c - \mu\Delta X$
  $t = t + 1$
return LeaderWhale

Fig. 1. Pseudo-code of the WOADD algorithm

V. Test Results

A. Benchmark functions

The performance of optimization algorithms is usually evaluated using a set of 24 benchmark functions [2, 3, 12, 14, 47, 48], as mentioned in CEC 2005 [49]. However, these functions aren't designed to account for dependent data. To test the proposed algorithm, we used a set of functions in Table 1 that include constraints imposing dependence among features. The focus is on minimizing the objective, and the last column of the table shows the optimal values for each function. As previously mentioned, most optimization algorithms lack explicit policies for handling problem constraints. Instead, they rely on penalty functions to manage them. The middle column of Table 1 shows the updated functions based on constraint penalties.

To evaluate optimization algorithms, two measures are commonly used: average fitness value and standard deviation. Suppose the initial population contains n objects. The average fitness value is obtained by summing the fitness value of each member of the population and dividing by the population size. The standard deviation is computed as the square root of the sum of squared distances of each fitness value from the average fitness value, divided by the population size minus one. The standard deviation represents the diversity among the fitness values of the population members. If the standard deviation is very small, it indicates that all search agents are near each other. In such cases, the algorithm may find it difficult to find the best solution, particularly if there are several local optima available.

To demonstrate the effectiveness of the proposed algorithm (WOADD), we conducted a comparative analysis with three other state-of-the-art optimization algorithms, namely Whale Optimization Algorithm (WOA) [45], Gray Wolf Optimization Algorithm (GWO) [12], and Aquila optimizer [10]. In all experiments, the number of search agents is set to 30, and the number of iterations is set to 300. We use state-of-the-art methods with default parameter values, but a penalty function is added to penalize invalid solutions, as shown in Table 1. The penalty function uses α, β, and γ values of 10, 10, and 20, respectively. The results of this comparison are presented in Table 2 for solutions with five and ten dimensions. For each test case, this table displays the algorithm's optimal $f(x)$ value, along with the average and standard deviation of $f(x)$ obtained by various search agents employed by the algorithm. As heuristic algorithms produce non-deterministic solutions, the presented results are the average of 20 runs.

The results show that the proposed algorithm has successfully obtained the minimum value for all functions f_2, f_3, f_4, f_5 and f_6 with particularly significant improvements observed for more complex functions f_5 and f_6. Notably, the proposed method achieved a result approximately 100% better than GWO and about 50% better than WOA for function f_5. However, the low standard deviation observed in GWO is not ideal for population-based approaches as it implies that all search agents are focused on a single target point.

The functions f_5 and f_6 have an additional parameter (c) to be determined in tests. Table 3 shows the performance of the optimization algorithms on these functions for c values of 0.1, 0.3, and 0.5. For instance, when $c = 0.1$, it means that m terms of the series carry 0.1 of the total weight, while the remaining terms carry 0.9 of the total weight. The results demonstrate that the proposed algorithm was able to achieve the best minimum value for the functions in all cases.

B. Analysis of Convergence curve

In population-based evolutionary algorithms, the search agents tend to extensively explore promising regions of the design space and gradually converge toward the best solution. During the initial stages of the optimization process, the search agents undergo abrupt changes, and then gradually converge towards the optimal solution. According to [5], this behavior guarantees that a population-based algorithm eventually converges to a point in the search space. The convergence curves of WOA, GWO, and WOADD for the benchmark functions are compared in Figure 2. Aquila is not included in the figure due to its large values in some cases. The results indicate that the WOADD algorithm is capable of quickly identifying promising areas within the search space during the initial stages of iteration and converges more rapidly towards the optimal solution.

Figure 3 illustrates the diversity of outcomes generated by different algorithms in 20 runs for functions f_5 and f_6. As shown in the figure, WOA and GWO display a high degree of diversity across different runs. In contrast, the results generated by WOADD are very similar to each other, and a consistent outcome is obtained in almost all runs. This makes WOADD more suitable for applications that require more deterministic results.

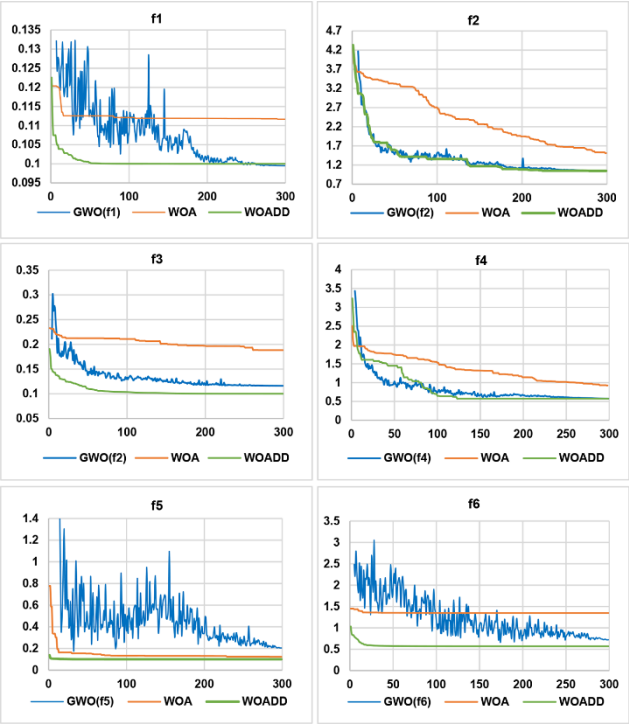


Fig. 2. Comparison of convergence curves of WOADD and literature algorithms obtained for the benchmark functions

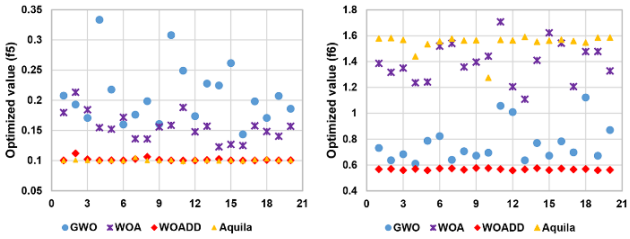


Fig. 3. The diversity of values produced by different algorithms in 20 runs for functions f_5 and f_6

VI. Conclusion

The paper presents an improved version of WOA that is specifically designed to handle dependent data. The proposed algorithm focuses on the valid range of data while searching the data space, leading to a more efficient search and faster attainment of the optimal solution. The experimental results demonstrate that this algorithm produces superior solutions, requiring fewer iterations to reach the optimal solution. Notably, for complex functions, the proposed algorithm achieves approximately 100% improvement over other tested algorithms. Additionally, the experiments show that the proposed algorithm's outcomes are more deterministic than other state-of-the-art ones, making it more suitable for applications that require precise outputs. Further development and refinement of such methods could unlock new insights and solutions from dependent data, an area that has received little attention thus far.

References

[1] Z. Lv, and R. Peng, "A novel meta-matching approach for ontology alignment using grasshopper optimization," *Knowledge-Based Systems*, vol. 201, pp. 106050, 2020.

[2] F. S. Gharehchopogh, and H. Gholizadeh, "A comprehensive survey: Whale Optimization Algorithm and its applications," *Swarm and Evolutionary Computation*, vol. 48, pp. 1-24, 2019.

[3] S. Mirjalili, S. M. Mirjalili, S. Saremi, and S. Mirjalili, "Whale optimization algorithm: theory, literature review, and application in designing photonic crystal filters," *Nature-Inspired Optimizers*, pp. 219-238, 2020.

[4] X. Xue, and J. Chen, "Optimizing sensor ontology alignment through compact co-firefly algorithm," *Sensors*, vol. 20, no. 7, pp. 2056, 2020.

[5] K. M. Ang, W. H. Lim, N. A. M. Isa, S. S. Tiang, and C. H. Wong, "A constrained multi-swarm particle swarm optimization without velocity for constrained optimization problems," *Expert Systems with Applications*, vol. 140, pp. 112882, Feb. 2020.

[6] F. A. Hashim, and A. G. Hussien, "Snake Optimizer: A novel meta-heuristic optimization algorithm," *Knowledge-Based Systems*, vol. 242, pp. 108320, 2022.

[7] L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz, and A. H. Gandomi, "The arithmetic optimization algorithm," *Computer methods in applied mechanics and engineering*, vol. 376, pp. 113609, 2021.

[8] I. Ahmadianfar, O. Bozorg-Haddad, and X. Chu, "Gradient-based optimizer: A new metaheuristic optimization algorithm," *Information Sciences*, vol. 540, pp. 131-159, 2020.

[9] M. Shehab, L. Abualigah, H. Al Hamad, H. Alabool, M. Alshinwan, and A. M. Khasawneh, "Moth–flame optimization algorithm: variants and applications," *Neural Computing and Applications*, vol. 32, pp. 9859-9884, 2020.

[10] L. Abualigah, D. Yousri, M. Abd Elaziz, A. A. Ewees, M. A. Al-Qaness, and A. H. Gandomi, "Aquila optimizer: a novel meta-heuristic optimization algorithm," *Computers & Industrial Engineering*, vol. 157, pp. 107250, 2021.

[11] S. Gholizadeh, M. Danesh, and C. Gheyratmand, "A new Newton metaheuristic algorithm for discrete performance-based design optimization of steel moment frames," *Computers & Structures*, vol. 234, pp. 106250, 2020.

[12] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, 2014/03/01/, 2014.

[13] G. Sun, Y. Shang, K. Yuan, and H. Gao, "An Improved Whale Optimization Algorithm Based on Nonlinear Parameters and Feedback Mechanism," *International Journal of Computational Intelligence Systems*, vol. 15, no. 1, pp. 1-17, 2022.

[14] S. Mirjalili, and A. Lewis, "S-shaped versus V-shaped transfer functions for binary particle swarm optimization," *Swarm and Evolutionary Computation*, vol. 9, pp. 1-14, 2013.

[15] W.-T. Pan, "A new fruit fly optimization algorithm: taking the financial distress model as an example," *Knowledge-Based Systems*, vol. 26, pp. 69-74, 2012.

[16] F. Zishan, E. Akbari, A. R. Sheikholeslami, and N. shafaghatian, "Optimization and Placement of DG Resources in the Network to Reduce Line Loading," *International Journal of Industrial Electronics Control and Optimization*, vol. 6, no. 2, pp. 89-100, 2023.

[17] N. Hashimzade, and M. A. Thornton, Handbook of research methods and applications in empirical microeconomics: Edward Elgar Publishing, 2021.

[18] D. Cicchella, M. Ambrosino, A. Gramazio, F. Coraggio, M. A. Musto, A. Caputi, D. Avagliano, and S. Albanese, "Using multivariate compositional data analysis (CoDA) and clustering to establish geochemical backgrounds in stream sediments of an onshore oil deposits area. The Agri River basin (Italy) case study," *Journal of Geochemical Exploration*, vol. 238, pp. 107012, 2022.

[19] G. Tepanosyan, L. Sahakyan, N. Maghakyan, and A. Saghatelyan, "Combination of compositional data analysis and machine learning approaches to identify sources and geochemical associations of potentially toxic elements in soil

and assess the associated human health risk in a mining city," *Environmental Pollution*, vol. 261, pp. 114210, 2020.
[20] H. Li, "Microbiome, metagenomics, and high-dimensional compositional data analysis," *Annual Review of Statistics and Its Application*, vol. 2, pp. 73-94, 2015.
[21] M. Greenacre, M. Martínez-Álvaro, and A. Blasco, "Compositional data analysis of microbiome and any-omics datasets: a validation of the additive logratio transformation," *Frontiers in microbiology*, vol. 12, pp. 727398, 2021.
[22] G. Coenders, and B. Ferrer-Rosell, "Compositional data analysis in tourism: review and future directions," *Tourism Analysis*, vol. 25, no. 1, pp. 153-168, 2020.
[23] L. Lalicic, E. Marine-Roig, B. Ferrer-Rosell, and E. Martin-Fuentes, "Destination image analytics for tourism design: An approach through Airbnb reviews," *Annals of Tourism Research*, vol. 86, pp. 103100, 2021.
[24] I. Janssen, A. E. Clarke, V. Carson, J. P. Chaput, L. M. Giangregorio, M. E. Kho, V. J. Poitras, R. Ross, T. J. Saunders, and A. Ross-White, "A systematic review of compositional data analysis studies examining associations between sleep, sedentary behaviour, and physical activity with health outcomes in adults," *Applied physiology, nutrition, and metabolism*, vol. 45, no. 10, pp. S248-S257, 2020.
[25] D. Dumuid, Ž. Pedišić, T. E. Stanford, J.-A. Martín-Fernández, K. Hron, C. A. Maher, L. K. Lewis, and T. Olds, "The compositional isotemporal substitution model: a method for estimating changes in a health outcome for reallocation of time between sleep, physical activity and sedentary behaviour," *Statistical methods in medical research*, vol. 28, no. 3, pp. 846-857, 2019.
[26] A. Alenazi, "A review of compositional data analysis and recent advances," *Communications in Statistics-Theory and Methods*, vol. 52, no. 16, pp. 5535-5567, 2023.
[27] D. Li, A. Srinivasan, Q. Chen, and L. Xue, "Robust covariance matrix estimation for high-dimensional compositional data with application to sales data analysis," *Journal of Business & Economic Statistics*, vol. 41, no. 4, pp. 1090-1100, 2023.
[28] T. Correa, M. Reyes, L. S. Taillie, C. Corvalán, and F. R. Dillman Carpentier, "Food advertising on television before and after a national unhealthy food marketing regulation in Chile, 2016–2017," *American journal of public health*, vol. 110, no. 7, pp. 1054-1059, 2020.
[29] T. P. Quinn, I. Erb, M. F. Richardson, and T. M. Crowley, "Understanding sequencing data as compositions: an outlook and review," *Bioinformatics*, vol. 34, no. 16, pp. 2870-2878, 2018.
[30] U. M. Khaire, and R. Dhanalakshmi, "Stability of feature selection algorithm: A review," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 4, pp. 1060-1073, 2022.
[31] D. Maulud, and A. M. Abdulazeez, "A review on linear regression comprehensive in machine learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 4, pp. 140-147, 2020.
[32] M. V. Narkhede, P. P. Bartakke, and M. S. Sutaone, "A review on weight initialization strategies for neural networks," *Artificial intelligence review*, vol. 55, no. 1, pp. 291-322, 2022.
[33] Y. Pan, Z. Su, A. Liu, W. Jingquan, N. Li, and Z. Xu, "A unified weight initialization paradigm for tensorial convolutional neural networks." pp. 17238-17257.
[34] C. Segura, C. A. C. Coello, G. Miranda, and C. León, "Using multi-objective evolutionary algorithms for single-objective constrained and unconstrained optimization," *Annals of Operations Research*, vol. 240, pp. 217-250, 2016.
[35] C. A. C. Coello, "Constraint-handling techniques used with evolutionary algorithms." In *Proceedings of the genetic and evolutionary computation conference companion*, pp. 1310-1333. 2022.
[36] G. Li, and Q. Zhang, "Multiple penalties and multiple local surrogates for expensive constrained optimization," *IEEE Transactions on Evolutionary Computation*, vol. 25, no. 4, pp. 769-778, 2021.
[37] F. Samanipour, and J. Jelovica, "Adaptive repair method for constraint handling in multi-objective genetic algorithm based on relationship between constraints and variables," *Applied Soft Computing*, vol. 90, pp. 106143, 2020.
[38] W. Liu, K. Liu, and T. Deng, "Modelling, analysis and improvement of an integrated chance-constrained model for level of repair analysis and spare parts supply control," *International Journal of Production Research*, vol. 58, no. 10, pp. 3090-3109, 2020.
[39] B. Ji, X. Yuan, and Y. Yuan, "Modified NSGA-II for solving continuous berth allocation problem: Using multiobjective constraint-handling strategy," *IEEE transactions on cybernetics*, vol. 47, no. 9, pp. 2885-2895, 2017.
[40] B.-C. Wang, H.-X. Li, Q. Zhang, and Y. Wang, "Decomposition-based multiobjective optimization for constrained evolutionary optimization," *IEEE Transactions on systems, man, and cybernetics: systems*, vol. 51, no. 1, pp. 574-587, 2018.
[41] X. Li, S. Zeng, C. Li, and J. Ma, "Many-objective optimization with dynamic constraint handling for constrained optimization problems," *Soft Computing*, vol. 21, pp. 7435-7445, 2017.
[42] R. Jiao, S. Zeng, J. S. Alkasassbeh, and C. Li, "Dynamic multi-objective evolutionary algorithms for single-objective optimization," *Applied Soft Computing*, vol. 61, pp. 793-805, 2017.
[43] T. Xu, J. He, and C. Shang, "Helper and equivalent objectives: Efficient approach for constrained optimization," *IEEE transactions on cybernetics*, vol. 52, no. 1, pp. 240-251, 2020.
[44] S. Zeng, R. Jiao, C. Li, and R. Wang, "Constrained optimisation by solving equivalent dynamic loosely-constrained multiobjective optimisation problem," *International Journal of Bio-Inspired Computation*, vol. 13, no. 2, pp. 86-101, 2019.
[45] S. Mirjalili, and A. Lewis, "The whale optimization algorithm," *Advances in engineering software*, Vol. 95, pp. 51-67, 2016.
[46] V. Pawlowsky-Glahn, J. J. Egozcue, and R. Tolosana-Delgado, Modeling and analysis of compositional data, John Wiley & Sons, 2015.
[47] G.-Y. Ning, and D.-Q. Cao, "Improved whale optimization algorithm for solving constrained optimization problems," *Discrete Dynamics in Nature and Society*, vol. 2021, pp.1-13 2021.
[48] S. Mostafa Bozorgi, and S. Yazdani, "IWOA: An improved whale optimization algorithm for optimization problems," *Journal of Computational Design and Engineering*, vol. 6, no. 3, pp. 243-259, July 2019.
[49] J.-J. Liang, P. N. Suganthan, and K. Deb, "Novel composition test functions for numerical global optimization," In *Proceedings 2005 IEEE Swarm Intelligence Symposium*, pp. 68-75, 2005.

Asieh Ghanbarpour received her M.S. degree in Computer Engineering from Sharif University in 2010. She obtained her Ph.D. in Computer Engineering from the Iran University of Science and Technology in 2018. Currently, she is working as an Assistant Professor in the Computer Department at the University of Sistan and Baluchestan. Her research interests lie in Information Retrieval, Data Mining, Databases, and Graph Processing topics.

Soheil Zaremotlagh received his M.S. degree from the Tehran University in 2003 and his Ph.D. degree from the Amirkabir University of Technology in 2016. He joined the Faculty of Engineering at the University of Sistan and Baluchestan in 2006. His research interests primarily focus on applying artificial intelligence methods to earth science data.

Fahimeh Dabaghi Zarandi is currently an assistant professor in computer engineering department of Vali-e-Asr University of Rafsanjan. She received the Ph.D. degree in software engineering from Iran University of Science and Technology, Tehran, Iran, in September 2018. She received the M.S. Degree in software engineering from the Sharif University of Technology, Tehran, Iran, in August 2010. She received the BSc Degree in software engineering from Ferdowsi university of Mashhad, Iran, in August 2008. Her research interests lie in Green Communication, Data Mining, Graph Processing, and Internet of Things.

Appendix:

TABLE 1 DESCRIPTION OF BENCHMARK CONSTRAINED FUNCTIONS

Function	Improved function with penalty	Optimum solution
$f_1(x) = \sum_{i=0}^{n} x_i^2$ $s.j: \forall x_i: 0 \le x_i \le 1, \sum_{i=0}^{n} x_i = 1$	$f_1(x) = \sum_{i=1}^{n} x_i - \alpha \sum_{i=1}^{n} \max(0, x_i - 1) - \beta \sum_{i=1}^{n} \max(0, -x_i) - \gamma \left(1 - \sum_{i=0}^{n} x_i\right)$	$f_1(x^*) = \frac{1}{n}$ $x^* = (\frac{1}{n}, \dots, \frac{1}{n})$
$f_2(x) = \sum_{i=0}^{n} (i+1) \times x_i$ $s.j \begin{cases} \forall x_i: 0 \le x_i \le 1 \\ \sum_{i=0}^{n} x_i = 1 \end{cases}$	$f_2(x) = \sum_{i=1}^{n} (i+1) \times x_i - \alpha \sum_{i=1}^{n} \max(0, x_i - 1)$ $- \beta \sum_{i=1}^{n} \max(0, -x_i) - \gamma \left(1 - \sum_{i=0}^{n} x_i\right)$	$f_2(x^*) = 1$ $x^* = (1, 0, \dots, 0)$
$f_3(x) = \sum_{i=0}^{n} \frac{1}{i+1} \times x_i$ $s.j \begin{cases} \forall x_i: 0 \le x_i \le 1 \\ \sum_{i=0}^{n} x_i = 1 \end{cases}$	$f_3(x) = \alpha \sum_{i=1}^{n} \frac{1}{i+1} \times x_i - \alpha \sum_{i=1}^{n} \max(0, x_i - 1)$ $- \beta \sum_{i=1}^{n} \max(0, -x_i) - \gamma \left(1 - \sum_{i=0}^{n} x_i\right)$	$f_3(x^*) = 1$ $x^* = (1, 0, \dots, 0)$
$f_4(x) = \sum_{i=0}^{n} \frac{(i+1)^2}{i+2} \times x_i$ $s.j \begin{cases} \forall x_i: 0 \le x_i \le 1 \\ \sum_{i=0}^{n} x_i = 1 \end{cases}$	$f_4(x) = \sum_{i=0}^{n} \frac{(i+1)^2}{i+2} \times x_i - \alpha \sum_{i=1}^{n} \max(0, x_i - 1)$ $- \beta \sum_{i=1}^{n} \max(0, -x_i) - \gamma \left(1 - \sum_{i=0}^{n} x_i\right)$	$f_4(x^*) = \frac{1}{2}$ $x^* = (1, 0, \dots, 0)$
$f_5(x) = \sum_{i=0}^{m} x_i^2 + \sum_{i=m+1}^{n} x_i^2$ $s.j \begin{cases} \forall x_i: 0 \le x_i \le 1 \\ \sum_{i=0}^{m} x_i = c \\ \sum_{i=m+1}^{n} x_i = 1 - c \end{cases}$	$f_5(x) = \alpha \left(\sum_{i=0}^{m} x_i + \sum_{i=m+1}^{n} x_i \right) - \beta \left(\left\lvert \frac{1}{2} - \sum_{i=0}^{m} x_i \right\rvert + \left\lvert \frac{1}{2} - \sum_{i=m+1}^{n} x_i \right\rvert \right) - \gamma \left(\sum_{i=1}^{n} \max(0, -x_i) + \sum_{i=1}^{n} \max(0, x_i - 1) \right)$	$f_5(x^*) = \frac{1}{m} + \frac{1}{n-m}$ $x^* = (\frac{1}{m}, \dots, \frac{1}{n-m}, \dots, \frac{1}{n-m})$
$f_6(x) = \sum_{i=0}^{m} (i+1) \times x_i + \sum_{i=m+1}^{n} \frac{1}{i+1} \times x_i$ $s.j \begin{cases} \forall x_i: 0 \le x_i \le 1, \sum_{i=0}^{m} x_i = c \\ \sum_{i=m+1}^{n} x_i = 1 - c \end{cases}$	$f_6(x) = \alpha \left(\sum_{i=0}^{m} (i+1) \times x_i + \sum_{i=m+1}^{n} \frac{1}{i+1} \times x_i \right)$ $- \beta \left(\left\lvert c - \sum_{i=0}^{m} x_i \right\rvert + \left\lvert 1 - c - \sum_{i=m+1}^{n} x_i \right\rvert \right)$ $- \gamma \left(\sum_{i=1}^{n} \max(0, -x_i) + \sum_{i=1}^{n} \max(0, x_i - 1) \right)$	$f_6(x^*) = \frac{1}{m} + \frac{1}{n-m}$ $x^* = (c, 0, \dots, 0, 1-c)$

TABLE 2 COMPARISON OF OPTIMIZATION RESULTS OBTAINED CONSTRAINED OPTIMIZATION FUNCTIONS

		n = 5			n = 10		
Function	Algorithm	opt	avg	std	opt	avg	std
f_1	GWO	**0.19801980**	0.19802	3.81E-09	**0.09951292**	0.09951	3.29E-06
	WOA	0.24697830	0.76064	0.473464	0.11257629	0.24801	0.136155
	Aquila	0.20901384	0.21568	0.049482	0.1011492	0.1102	0.03038
	WOADD	0.20000125	0.23054	0.025687	0.10013077	0.10685	0.014231
f_2	GWO	1.01336514	1.01336	0.000061	1.07093814	1.07101	4.58E-05
	WOA	1.08175004	4.45874	4.393632	1.49786942	2.51204	0.784425
	Aquila	2.54880186	2.91290	0.282923	5.49780352	5.40050	0.300903
	WOADD	**1.00046018**	1.34106	0.537621	**1.04863548**	2.14027	1.335185
f_3	GWO	0.21012547	0.20951	1.90E-04	0.11002846	0.11002	7.64E-05
	WOA	0.24386565	0.19516	0.025801	0.16962024	0.16114	0.008623
	Aquila	0.32208246	0.42317	0.08324	0.29050187	0.29944	0.050842
	WOADD	**0.20402433**	0.23263	0.051775	**0.10004526**	0.10433	0.016416
f_4	GWO	0.52344523	0.52378	5.12E-04	0.54897454	0.54874	6.76E-04
	WOA	0.52788077	1.46388	0.641348	0.92521517	1.76456	0.695900
	Aquila	1.68200241	2.16610	0.311962	4.42580623	4.63140	0.451971
	WOADD	**0.50056879**	0.81622	0.609542	**0.54862247**	3.64583	1.978977
f_5	GWO	0.25597429	0.25578	0.00025	0.20839877	0.21728	0.000266
	WOA	0.28083862	0.33046	0.08898	0.15356303	0.17039	0.017592
	Aquila	0.27418764	0.25381	0.082323	0.10035238	0.09692	0.00682
	WOADD	**0.20967541**	0.25214	0.03853	**0.10046314**	0.12067	0.019322
f_6	GWO	0.72157598	0.72146	5.43E-04	0.71640018	0.71629	4.53E-04
	WOA	0.88517801	0.86809	0.052462	1.25271829	1.23933	0.035963
	Aquila	0.94366280	1.02790	0.232031	1.56251892	1.5288	0.158072
	WOADD	**0.60162363**	0.73374	0.242773	**0.56118408**	1.00936	0.490113

TABLE 3 COMPARISON OF OPTIMIZATION RESULTS OBTAINED FOR f_5 AND f_6 BASED ON DIFFERENT VALUES OF C

		c = 0.1			c = 0.3			c = 0.5		
Fun	Algorithm	opt	avg	std	opt	avg	std	opt	avg	std
f_5	GWO	0.21	0.21	3.47E-04	0.21	0.21	3.3E-04	0.21	0.22	0.0002
	WOA	0.29	0.16	0.0204	0.17	0.17	0.0137	0.15	0.17	0.0176
	Aquila	0.39	0.38	0.2089	0.85	0.19	0.0175	0.10	0.09	0.0068
	WOADD	**0.16**	0.19	0.0334	**0.12**	0.14	0.0238	**0.10**	0.12	0.0193
f_6	GWO	0.34	0.34	1.59E-04	0.62	0.61	3.62E-04	0.72	0.72	4.5E-04
	WOA	0.46	0.46	0.013092	0.88	0.86	0.034454	1.25	1.24	0.0359
	Aquila	0.49	0.73	0.34441	0.92	0.94	0.41512	1.56	1.52	0.1580
	WOADD	**0.20**	0.28	0.098156	**0.38**	0.80	0.37041	**0.56**	1.01	0.4901

IECO

This page intentionally left blank.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Tuning Spring Constant of a Beam for a Promising Reduction in the Actuation Voltage of a V-Band RF MEMS Switch

Farid Khamouei Touli[1] , Javad Yavand Hasani[2]

School of Electrical Engineering, Iran University of Science and Technology,Tehran, Iran[1,2]
Corresponding author's email: yavand@iust.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 28-Dec-2023
Received in revised form: 6-March-2024
Accepted: 11-March-2024
Published online: 21-June-2024

Keywords:
Actuation Voltage,
Modified Energy Method,
RF-MEMS,
Switching Time,
Von Mises Stress.

ABSTRACT

Radiofrequency microelectromechanical system (RF-MEMS) switches are utilized across a broad spectrum of industries, including telecommunications, aerospace, defense, and smartphone technology.Herein, we proposed a new numerical and simulation analysis approach for spring constant (k) values as the characteristic mechanical parameters of RF MEMSs using the modified energy method (MEM). The proposed RF-MEMS switch was analyzed and simulated using the COMSOL package, and the findings confirmed that the alteration in the position and length (L) of the beams not only diminishes k significantly but also provides actuation-voltage VAC ultrasensitive structures and great concomitance between numerical and simulation k and VAC values. VAC value for the L-dependent numerical k (0.07 N m^{-1}) was calculated to be 1.61 V which was validated with simulation outputs at 0.08 N m^{-1} and 1.80 V for k and VAC, respectively. Additionally, the switching time (ts), Von Mises Stress (VMS), natural frequency (fn) and mass (m) characteristic mechanical parameters were found to be 25.60 µs, 4.50 MPa, 3118.60 Hz , and0.21 ng, respectively. RF analysis was conducted in HFSS, revealing promising simulation results for the studied RF-MEMS switches. The return loss demonstrates excellent performance, registering better than -1 dB at 46 GHz. Furthermore, the insertion loss is noteworthy, exceeding expectations with values better than -0.7 dB at 46 GHz. Importantly, the isolation is impressive, exceeding -25 dB across the frequency range from 40 GHz to 35 GHz, all achieved with a modest actuation voltage of 1.8 V. This study contributes valuable insights intothe design and application of low-actuation-voltage RF-MEMS switches.

NOMENCLATURE

$RFMEMS$	Radiofrequency micro Electromechanical system	f_n	Natural frequency
k	Spring constant	MEM	Modified energy method
V_{AC}	Actuation voltage	t_s	Switching time
m	Mass	VMS	Von Mises Stress

I. Introduction

In the pathway of the evolution of electrostatic switches, the paradigm of recent research has focused on improving structural challenges of field-effect transistors (FETs) and positive-intrinsic-negative (PIN) diodes which consist of excessive power consumption, high rate of return loss, low isolation,and limited bandwidth [1]. Microelectromechanical systems (MEMSs) as one of the most promising candidates which represented versatile properties such as superior isolation sensitivity and limited return loss have been interested in recent years. Several capacitors, switches, and phase shifters are benefited from MEMSs for wide irreplaceable applications including automatic test equipment, advanced communication technology, satellite communications, aerospace, defence, and smartphones [2, 3]. The MEMSs work based on electrostatic, magnetic, and piezoelectric actuators to provide the demand energy for the mechanical movement of these pieces [4]. Among the array of actuation methods, electrostatic actuation has emerged as the most efficient. popular method of switch actuation due to its high speed, easier construction, as well as tuneable size, high reliability and stability against environmental and temperature changes, satisfactory power consumption, high scalability, more straightforward structure and compatibility with CMOS technology [5]. The most important challenge in RF-MEMS switches is its limitation in accessingthe low actuation voltage (V_{AC}) synchronous with switch low input loss. Additionally, by decreasing V_{AC}, the air gap width is reduced in switch structure and leads to insertion loss increases. Some versatile methods such as increasing the section level, reducing the gap between the switch beam and the down electrode, and structure design with a low spring constant have been applied to reduce the V_{AC} in the RF-MEMS switches structures [6]. Amongst, increasing the section level and gap reduction between the beam and the down electrode to lead to theuntenability of the switch size and insertion loss, respectively [7]. Therefore, V_{AC} reduction is a promising and superior method than other approaches for designing a structure with low spring constant [8, 9]. In this regard, Sharma *et al*. decrease the spring constant by a new structure with non-uniform spiral curves and variable curve width which the obtained values for V_{AC} were between 15 to 3 V and for curve width was 100 µm to 280 µm. Additionally, MEM was applied to obtain displacement and eventually spring constant [10]. Ganji *et al.* report a new small area structure for RF-MEMS series switches by obtaining accurate spring constant and V_{AC} (4.05 V) [11]. In another study, Guha *et al.* applied capacitor analysis to the calculation of the electrostatic power effects and marginal fields and an authentic relationship was achieved; as well as a suitable algorithm was reported to improve the effectiveness of the hole which the proposed structure successfully decreases V_{AC} to 2.1 V [12].

In this study, we proposed a new numerical and simulation analysisapproach for spring constant (k) values as the characteristic mechanical parameter for highly flexible RF-MEMSs constructed from gold (Au) and silicon nitride (Si_3N_4) using the modified energy method (MEM). The as-proposed RF-MEMS switch was analyzed and simulated in the COMSOL package.

II. Methodology

A. *Understudy switch structure*

The as- designed RF-MEMS switch was developed based on the previous report by Guha *et al.* with a changing its square wave shape serpentine to non-uniform serpentine flexure with a two-meander section divided into seven beams Fig . 1(a) [13, 14]. In the suggested structure, there are four horizontal beams with length *a*, and three vertical beams L_i, Fig . 1(b).

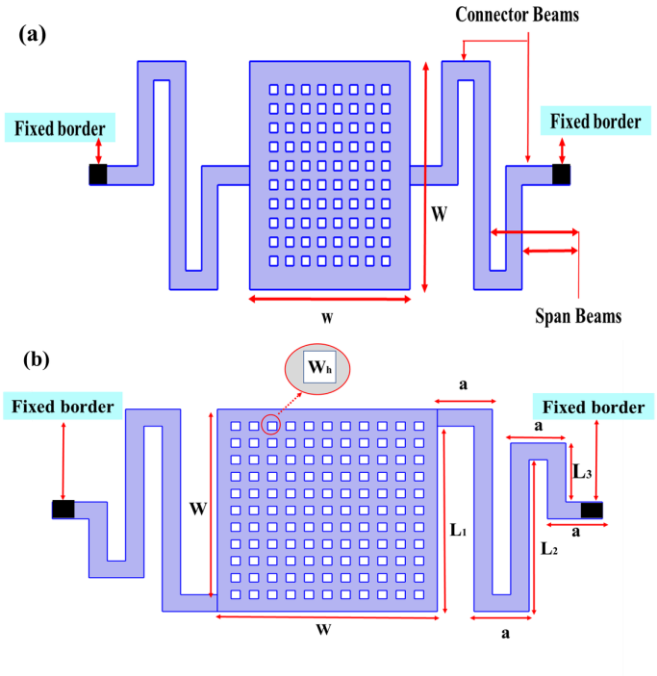


Fig. 1. The structure of the RF-MEMS switch by Guha *et al.* [14] (a); Proposed structure in this study (b)

The size of different sections of the suggested switch was summarized in Table 1. The RF-MEMS structure underwent optimization, and an analysis of its mechanical characteristics was conducted.The switch process is presented as follows: the bridge material is selected by considering Young's modulus and mass density so that Young's modulus has to be more than mass density by a selection of gold as a bridge. The Si_3N_4 and SiO_2 with a dielectric constant of 7.6 and 3.6, respectively were selected for the desirable comparison as dielectric materials by investigation of dielectric constant, dielectric resistance, leakage current, and surface roughness characters [15].

B. *Mathematical models for the suggested structure*

The modified energy method (MEM) was applied to analyse the spring constant of the proposed structure. First, the beam displacement was calculated based on the electrostatic force, and

then the spring constant was calculated by Eq. 1 [16].

$$k_z = \frac{F_z}{\delta_z} \quad (1)$$

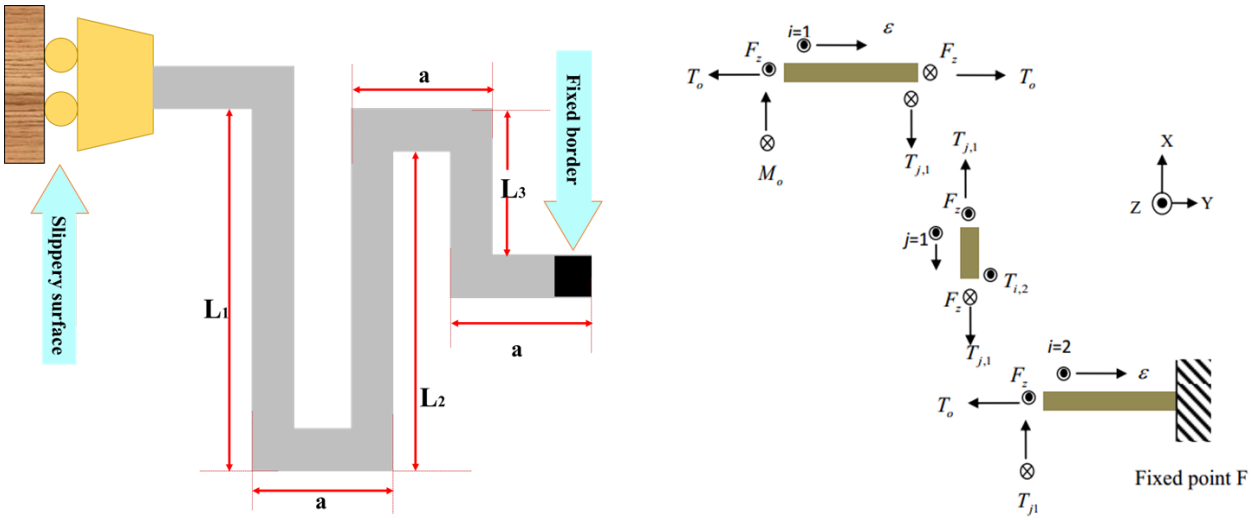
Where F_z is the exerted electrostatic force, and δ_z is the displacement of the beam. As in a static problem, it is supposed that the structure is perfectly balanced, hence, it is supposed that the sum of all internal and external forces (F), as well as the sum of all bending (M) and torsional (T) moments, are zero.Eq. 2 shows this supposition [16].

$$\sum T = 0,\ \sum M = 0,\ \sum F = 0 \quad (2)$$

In Fig .Fig . 2(a), in order to obtain static balance and to obtain bending and torsional anchors in different parts of the beam, the whole structure is assumed as a support as well as a sliding surface.

TABLE 1: SUGGESTED STRUCTURE CHARACTERISTICS

Parameter	Symbol	Size (μm)
Bridge width	w	120
Bridge length	W	120
Horizontal beams length	a	30
Vertical beam length	L1	110
Second vertical beam length	L2	90
Third vertical beam length	L3	30
Curve width	w_m	10
Dielectric thickness	td	0.15
Mobile bridge thickness	tb	0.60
Ari distance between two surfaces	g_0	2.50
Holes length size	w_h	5.0
Holes number	n_l .n_w	11*11
Dielectric constant	ε_r	7.60
Gold Young's modulus	E	79GPa



(a) (b)

Fig. 2. Characteristics of curve spiral sections for solving a static problem (a); Analysis of the suggested structures with different curves and their sections (b) [10].

Fig . 2(b) shows forces and moments decomposition in various parts of the beam. Furthermore, to make the structure remain static, M_0 and T_0 were suspended in the opposite direction of the constant head on the bridge side. The analysis begins withthe suspended part and moves to the T constant, finally M is calculated for each part of the beam considering the balance condition according to the related equations summarized in Table 2. Electrostatic force in RF-MEMS parallel capacitor switches is applied to change the switch condition from upstate to downstate. Therefore, it is possible to consider this force in line with the Z-axis. Using the Castiglione theorem, to simplify this problem, it is supposed that no other force in

no other direction is exerted on beams. In the MEM, the energy and spring constant were calculated from external force and internal moments according to Eq. 3. Hence, it is possible to define displacement as in Eq. 4 [16, 17].

$$U=\begin{pmatrix}\sum_{i=1}^{N}\int_{0}^{a}\left(\frac{M_{a,i}^2}{2EI_{x,a}}+\frac{T_{a,i}^2}{2GJ_a}\right)d\varepsilon+\\ \sum_{j=1}^{N}\int_{0}^{Lj}\left(\frac{M_{b,j}^2}{2EI_{x,b}}+\frac{T_{b,j}^2}{2GJ_b}\right)d\varepsilon\end{pmatrix} \quad (3)$$

$$\delta_z=\frac{\partial U}{\partial F_z} \quad (4)$$

where E is Young's modulus, G is thetorsional moment modulus, J is torsion constant, as well as I_{xa} and I_{xb} are the moment of inertia for vertical and horizontal curves whose values are calculated by Eq. 5 [16].

$$I_x=I_{x,a}=I_{x,b}=\int_{-w/2}^{w/2}\int_{-t/2}^{t/2}x^2dxdz=\frac{wt^3}{12} \quad (5)$$

The moment of inertia in line with z is as in Eq. 6 [16].

$$I_z=\int_{-t/2}^{t/2}\int_{-w/2}^{w/2}x^2dxdz=\frac{tw^3}{12} \quad (6)$$

Torsion modulus is defined in Eq. 7 [16].

$$G=\frac{E}{2(1+V)} \quad (7)$$

Where E is Young's modulus, and V is Poisson's ratio. The torsion constant was calculated according to the moment of inertia relation torsion constant for vertical and horizontal curves in the rectangular bridge, as Eq. 8 [16].

$$J=0.413(I_X+I_Z) \quad (8)$$

In the fixed-head state energy change to bending moment (θ_0) and energy change to torsional moment ($\emptyset_0$) were calculated as follows:

$$\theta_0=\frac{\partial U}{\partial M_0}=0 \qquad \emptyset_0=\frac{\partial U}{\partial T_0}=0 \quad (9)$$

In simple terms, the slope is zero in the fixed head for bending and torsional movements. To solve the problem, first, bending and torsional moments must be calculated in the fixed head. Therefore, $\theta_0=0$ and $\emptyset_0=0$ wereassumed for borders that [16].

TABLE 2: ANALYZING MOMENTS FOR DIFFERENT SECTIONS OF THE CURVE IN THE SUGGESTED STRUTURE

Vertical beam		Horizontal beam	
Torsional moment	**Bending moment**	**Torsional moment**	**Bending moment**
$T_{b1}=M_0-F_za$	$M_{b1}=-T_0-F_zx$	$T_{a1}=T_0$	$M_{a,1}=-M_0-F_zx$
$T_{b2}=M_0+2f_za$	$M_{b2}=-F_zx-T_0-F_zL_1$	$T_{a2}=T_0-F_zL_1$	$M_{a2}=-F_za-F_zx-M_0$
$T_{b3}=M_0-3aF_z$	$M_{b3}=-T_0-F_zx+F_z(L_1-L_2)$	$T_{a3}=T_0-F_z(L_1-L_2)$	$M_{a3}=M_0-F_z(2a)-F_zx$
0	0	$T_{a4}=T_0+F_z(L_1-L_2+L_3)$	$M_{a4}=M_0-3F_za-f_zx$

$$\theta_0=\frac{\partial U}{\partial M_0}=\begin{pmatrix}\sum_{i=1}^{2N}\int_{0}^{a}\left(\frac{M_{a,i}}{EI_{x,a}}\frac{\partial M_{a,i}}{\partial M_0}+\frac{T_{a,i}}{GJ_a}\frac{\partial T_{a,i}}{\partial M_0}\right)d\varepsilon+\\ \sum_{j=1}^{2N}\int_{0}^{Lj}\left(\frac{M_{b,j}}{EI_{x,b}}\frac{\partial M_{b,j}}{\partial M_0}+\frac{T_{b,j}}{GJ_b}\frac{\partial T_{b,j}}{\partial M_0}\right)d\varepsilon\end{pmatrix} \quad (10)$$

$$\emptyset_0=\frac{\partial U}{\partial T_0}=\begin{pmatrix}\sum_{i=1}^{2N}\int_{0}^{a}\left(\frac{M_{a,i}}{EI_{x,a}}\frac{\partial M_{a,i}}{\partial T_0}+\frac{T_{a,i}}{GJ_a}\frac{\partial T_{a,i}}{\partial T_0}\right)d\varepsilon+\\ \sum_{j=1}^{2N}\int_{0}^{Lj}\left(\frac{M_{b,j}}{EI_{x,b}}\frac{\partial M_{b,j}}{\partial T_0}+\frac{T_{bj}}{GJ_b}\frac{\partial T_{b,j}}{\partial T_0}\right)d\varepsilon\end{pmatrix} \quad (11)$$

Solving the above equations resulted inthe values of the M_0 and T_0 (based on the size of vertical and horizontal beams) as well as G, I, and F. Moreover, as F_z had to be derived to obtain the displacement, M_0 and T_0 are written as C_1F_z and C_2F_z.

$$M_0=f(a,L_1,L_2,L_3,G,I,F_z)=C_1F_z \quad (12)$$

$$T_0=f(a,L_1,L_2,L_3,G,I,F_z)=C_2F_z \quad (13)$$

Substituting the values for the parameters listed in Table 2 into Eq .12 and 13 yields the values of M_0 and T_0. Since the problem is parametric, MATLAB is utilized to symbolically solve the problem due to the complexity and length of the equations. Considering the size of different sections of the suggested structure in Table 1 the value of the spring constant was calculated. To investigate the spring constant of vertical and horizontal sections, the vertical and horizontal are scanned separately from 20 µm to 250 µm.

C. V_{AC} Analysis for the Suggested Switch

The most important challenge in RF-MEMS switches is its limitation to access to the low V_{AC} synchronous with switch low input loss. RF MEMS switch dynamic Eq. 14 is as follows [2].

$$m\frac{d^2x}{dt^2}+b\frac{dx}{dt}+kx=\frac{\varepsilon_0AV^2}{2\left(d-x+\frac{t_d}{\varepsilon_r}\right)} \quad (14)$$

Where *m* is switch mass, *b* is thedamping coefficient resulting from fluid viscosity, and *k* is theswitch spring constant. The right side of the equation represents the electrostatic power equation. Solving Eq. 14 yields the V_{AC} [2, 8].

There are three possible ways to decrease the actuation voltage:

- Increasing section level, which also increases switch size and insertion loss.
- Gap reduction between switch bar and low electrode, which is practically accompanied by problems like insertion loss increase.
- Structure design with low spring constant.

$$V_{AC} = \sqrt{\frac{8kg_0^3}{27A\varepsilon_0}} \quad (15)$$

As in Eq.15, V_{AC} depends on three factors includingspring constant, actuation surface size,and air distance. Change in any of these factors affects the V_{AC}. Applying the results of the last section, an analytical equation for actuation voltage can be derived that can be obtained by replacing the results of Eq. 1 in Eq. 15. Due to the equation length, the V_{AC} is analyzed in MATLAB software.

III. Suggested Simulation Structure, Results and Discussions

This section presents the results obtained from the mechanical structure simulation using COMSOL software, including the V_{AC}, switching time, natural frequencies, and the designed switch mass, as well as the Von Mises stress for the proposed structure.

A. Effect of vertical and horizontal sections of the beam on the spring constant and V_{AC}

Vertical and horizontal sections are important operational variable for optimizing the spring constant and V_{AC} which they were studied in the range of 20 µm to 250 µm. Results revealed that the increase in horizontal beams than that of vertical beams has superior effects on constant spring reduction which is comfortable with theMEM method (see Fig . 3(a)). By increasing the vertical beam from 20 µm to 100 µm spring constant was reduced from 0.17 N m^{-1} to 0.002 N m^{-1}. At first sight, it might seem useful in achieving our goal, but in the structure design, drag force in returning the switch to the initial state and consumption level values haveto be considered. In the range of 20 µm to 250 µm Vertical sections have fewer changes than horizontal beams. Scanning vertical beams changed the spring constant to approximately 0.02 N m^{-1}. The effect of different sections of the suggested structure on V_{AC} is shown in Fig . 3(b). It can be seen that by decreasing the horizontal beam, V_{AC} decreases with a steep slope. Changing the horizontal section from 20 µm to 100 µm the V_{AC} reduces from 2.5 V to 0.25 V. Vertical beams influence the V_{AC} less as compared to horizontal beams, and by increasing the beams from 20 µm to 250 µm, the actuation voltage changes from 1.6 V to 1.4 V. Table 3 summarizes the size change in horizontal and vertical beams as well as the effect of this change on spring constant and V_{AC}. Reducing air distance is another method to reduce voltage. Table (4) shows actuation voltage values for different air distances. The V_{AC} for a2.5 µm air distance is 1.6107 V. Air distance has to be selected such that the construction phase of the suggested structure is smoothly handled. Therefore, 2.5 µm air distance is selected to prevent any problems [2]. According to Table 4, when lower air distances are selected, lower V_{AC} is obtained. Another problem with lower air distance is the increase in input loss [2].

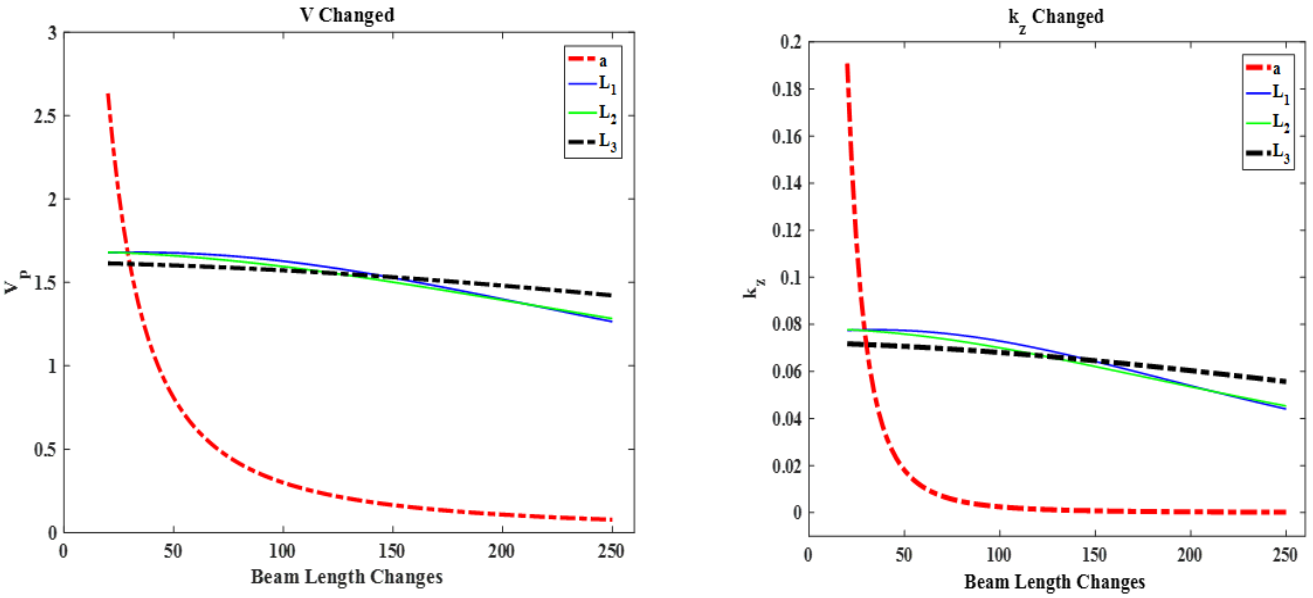


Fig. 3. The effect of different sections of the suggested structure on spring constant (a); The effect of changing different sections of the suggested structure on V_{AC} (b).

TABLE 3: SUMMARY OF THE EFFECT OF STRUCTURE CHANGE ON SPRING CONSTANT AND V_{AC}

Change range (20-250 µm)	Spring constant (N.m^{-1})	V_{AC}(V)
a	0.19-0.01	2.6-0.5
L_1	0.079-0.045	1.6-1.4
L_2	0.079-0.048	1.6-1.4
L_3	0.076-0.065	1.55-1.5

B. V_{AC} Calculation

Recently, researchers have used all three ways to reduce V_{AC}. Structure design with a low spring constant is the best method [18-20]. In order to decrease V_{AC} in the designed switch, a spring constant reduction method with non-uniform curves is used, and the replacement of a designed switch by voltage exertion in COMSOL software is demonstrated in Fig . 4(a). In this Fig ., switching is performed by replacing approximately 1/3 of the air distance [2].

The designed switch needs 1.8 V of actuation voltage for switching. Fig . 4(b) shows the simulation result in COMSOL software. For the simulation,thefirst mobile section and dielectric are illustrated. Then,a 2.5 µm air distance is applied between the mobile section and the fixed section. Finally, the mobile section is made up of gold and dielectric, and silicon nitride was utilized in the simulation. Actuation is electrostatically applied to the mobile section, and the ground is considered as the border between dialectic and air distance.

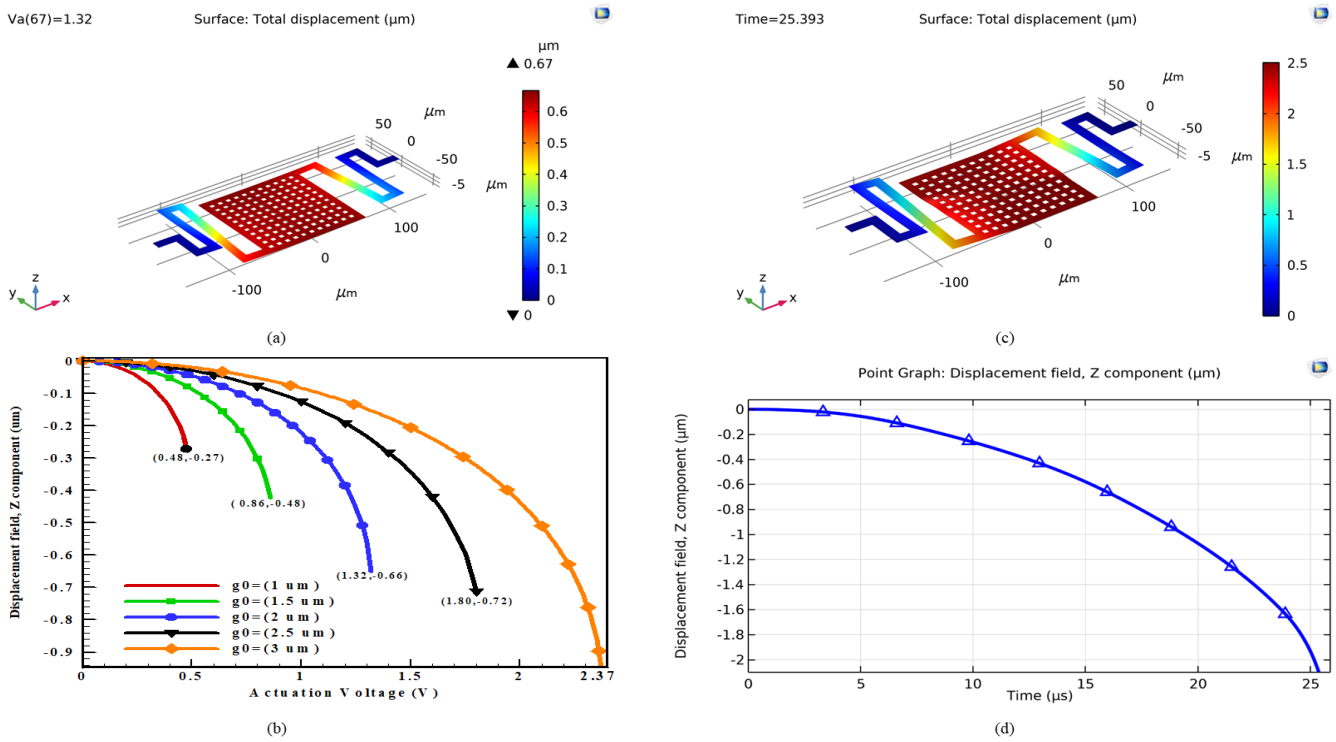


Fig. 4. Suggested switch replacement based on the electrostatic actuation (a); The V_{AC} for the designed switch (b); switching time of the designed switch (c); Replacement in switching time (d).

TABLE 4: COMPARISON BETWEEN MODEL AND SIMULATION V_{AC} VALUE VERSUS DIFFERENT AIR DISTANCES FOR THE SUGGESTED STRCTURE.

Air distance (µm)	Model V_{AC} (V)	Simulated V_{AC} (V)	Error (%)
1.0	0.41	0.48	14.0
1.5	0.75	0.86	12.0
2.0	1.15	1.32	12.0
2.5	1.61	1.80	10.0
3.0	2.12	2.37	10.0

C. Switching Time

Switching and RF-MEMS speed is less than semi-conductive and PIN diode and it is one of the limitations of RF-MEMS switches. Damping, mass, switch structure, and spring constant value impact the switching time. Using materials such as $ALSI_{0.04}$ with smaller mass may enhance switching speed [13] and other reported literature (see Table 6). Fig . 4(c)

shows the result of the switching time simulation for the designed structure in COMSOL software, which is 25.3 μs for this structure. Fig . 4(d) shows the movement in the whole air distance. The next section is dedicated to the calculation of mass and switch natural frequencies which greatly influence the switch function.

D. Natural Frequencies and the Designed Switch Mass

Switch natural frequencies and switch mass influence switch spring constant and finally V_{AC} such that the less the switch mass, the more the switching speed. Also, high natural frequencies enhance V_{AC} and reduce switching time. Thus, the switch is designed in a way that this negligence is considered [22]. Table 5 shows the first six natural frequencies of the designed switch and themass related to switching. The determination of the resonance frequency value is governed by the following relationship:

$$f_0 = \frac{1}{2*\pi}\sqrt{\frac{k}{m}} \tag{16}$$

In this equation, "k" represents the spring constant, and "m" signifies the effective mass of the cantilever beam. For the calculations presented in this article, the first mode, which aligns with the movement direction of the switch actuation, has been considered. The mass and spring constant values for the specified structure are determined to be 0.206 ng and 0.079 N $m^{(-1)}$ respectively. Consequently, the value of the first mode is confirmed to be 3118.8 Hz.

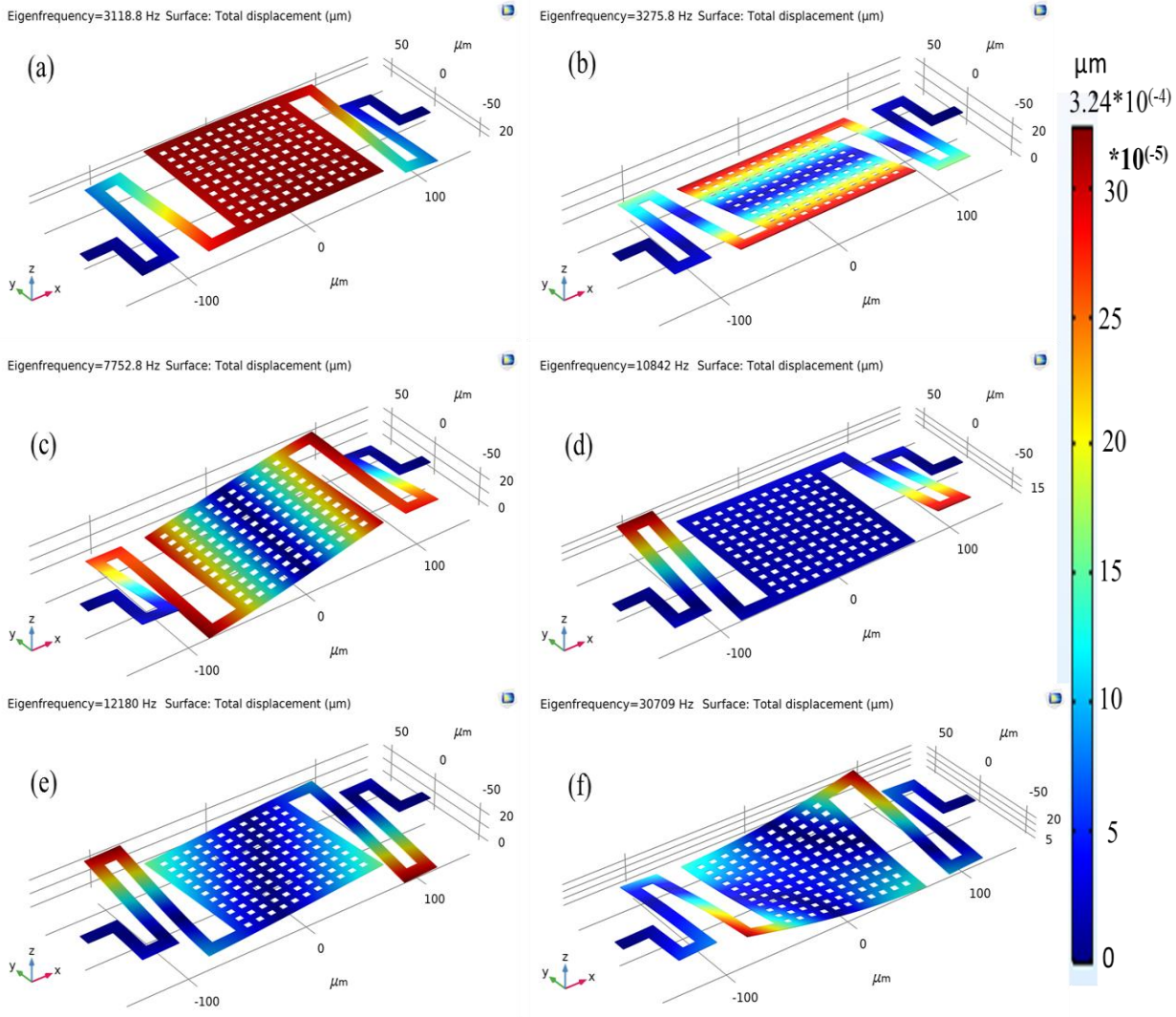


Fig. 5. Designed switch replacement in six natural frequencies, (a): 3118.63 Hz, (b): 3275.60 Hz, (c): 7752.09 Hz, (d): 10842.59 Hz, (e): 12180.50 Hz, (f): 30709.47 Hz

TABLE 5. SIX NATURAL FREQUENCIES OF THE DESIGNED SWITCH (MASS=0.206ng).

Frequency number	Natural frequency
1	3118.63
2	3275.60
3	7752.09
4	10842.59
5	12180.50
6	30709.48

E. Von Mises Stress for the Mechanical Structure

Von Mises Stress is one of the most important issues that must be studied in mechanical analysis, which is measured to guarantee the design safety. Comparing this value with material yield strength defines whether thestructure must be specially designed or not. In fact, if themaximum tolerable stress is more than material strength, the design fails, and if themaximum tolerable stress is less than material strength, the design is safe. For the suggested design, this value is 4.5 MPa, which is shown in Fig . 6, the maximum value obtained for silicon ground is within the acceptable range.

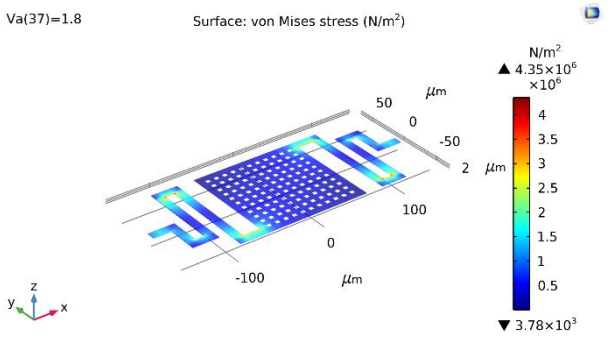


Fig. 6. Von Mises Stress for the suggested structure.

F. Spring Constant Value based on Simulation Results

A replacement proportionate to 1/3 of the air distance is necessary for the switch to function properly [1]. Hence, $g = \frac{2}{3} g_0$. On the one hand, according to Eq. 16, spring constant for the suggested structure is 0.075 N m^{-1}.

$$K = m * (2\pi f_1)^2 = (206 * 10^{-9}) * (2 * \pi * 3118.6)^2 = 0.079 \quad (17)$$

G. RF Performance

Traditional capacitive RF MEMS switches consist of a coplanar waveguide (CPW) connected to a transmission line, a floating membrane serving as an actuator, and an electrode for applying a bias voltage. This bias voltage enables the switch to transition between its on and off states. The up-state capacitance (C_u), down-state capacitance (C_d), and capacitive ratio (C_r) can be computed as follows:

$$C_r = \frac{C_d}{C_u} = \frac{\frac{\varepsilon_0 \varepsilon_r A}{t_d}}{\frac{\varepsilon_0 A}{g + \frac{t_d}{\varepsilon_r}} + C_f} = \frac{g_0 \varepsilon_r}{t_d} + 1 \quad (18)$$

A higher C_r indicates superior RF performance. To evaluate and enhance the RF performance, simulations were conducted using HFSS software. The RF performance is characterized by S parameters, which can be obtained as follows [21]:

$$S_{11} = -20\, log \left| \frac{-Z_h}{2Z_h + Z_0} \right| \quad (19)$$

$$S_{21} = -20\, log \left| \frac{2Z_h}{2Z_h + Z_0} \right| \quad (20)$$

The parameters Z_0 and Z_h represent the CPW impedance and cantilever beam impedance, respectively. The evaluation of RF performance involves defining return loss, insertion loss, and isolation. Isolation is defined as the ratio of transmitted power to submitted power in the down state, where a higher isolation indicates better RF performance. S_{12} signifies the isolation at the down state and insertion loss at the up state, while S_{11} represents the return loss at both the down and up states. In Fig. 7, the switch diagram is illustrated, and Fig. 8 depicts the simulated results for S_{11} and S_{12}, respectively. Simulation results indicate that the return loss is better than -1 dB at 46 GHz in the up state, and it remains more than -10 dB from 40 GHz to 50 GHz in the down state. The insertion loss is better than -0.7 dB at 46 GHz, and the isolation exceeds -25 dB from 40 GHz to 35 GHz with an actuation voltage of 1.8 V. Notably, the maximum isolation achieved is -29 dB at 46 GHz. These results highlight the favorable RF performance of the switch.

In this paper, the switch is based on silicon and a 0.15 µm-thich Si_3N_4 dielectric material is used. The CPW line has a size of 56/120/56 (G/S/G) µm and the size of the beam is 120 µm width and 300 µm length.

H. fabrication

The fabrication of the MEMS switch entails the utilization of a high-resistance silicon substrate, wherein successive layers of SiO_2, Si_3N_4, and Au are deposited. The manufacturing process encompasses photolithography, electron beam evaporation, sputtering, and electroplating techniques to fabricate the essential components. Notably, a sacrificial polyimide layer is strategically employed and subsequently eliminated to conclude the switch fabrication process[6].

The fabrication process of the MEMS switch is shown in Fig. 9. The switch is fabricated on 250 µm thickness high-resistance silicon, and a layer of SiO_2 is located above it. Furthermore, photolithograph negative glue was used to photoetching DC bias lines pattern, electron beam evaporation C_r DC bias lines, then stripped photolithograph negative glue, and 0.15 µm-thickness Si_3N_4 is deposited on top of the bias lines. Au layer of 0.3 µm thickness, as CPW lines, is sputtered on the substrate, and a layer of Si_3N_4 is located on the top of

the transmission lines as a dielectric layer. A 2 μm-thick polyimide is used as the sacrificial layer after the thermal curing process. The 0.6 μm-thick anchors and beams are formed by electroplating technology. The sacrificial layer is released using supercritical dry release[8, 22].

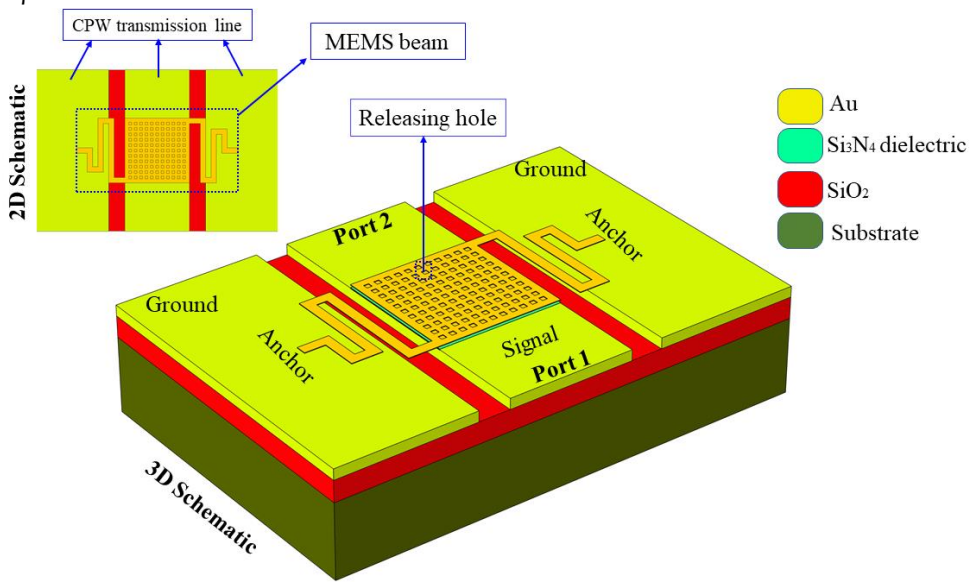


Fig.7. Schematic view of the structures of the RF MEMS switch.

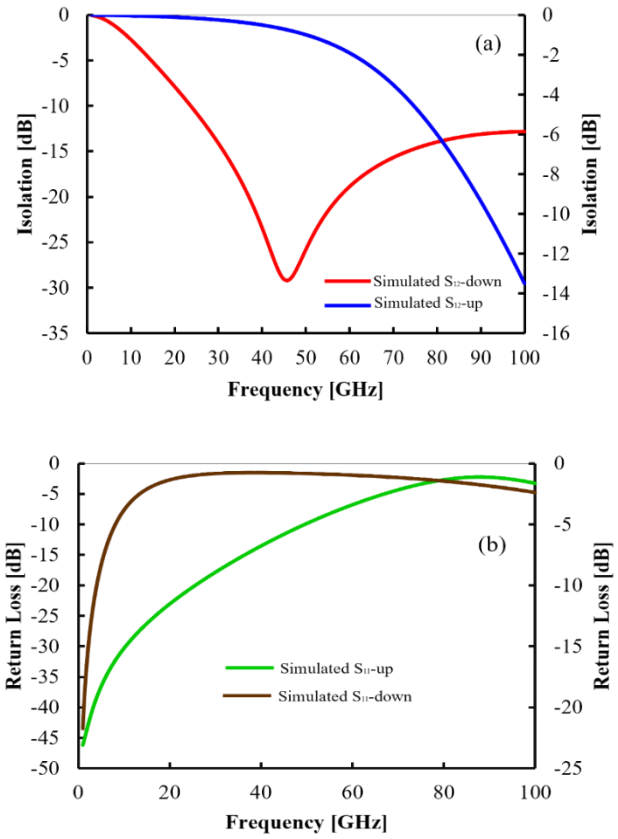


Fig.8. (a) S_{12} of the simulated results; (b) S_{11} of the simulated results

I. Conclusions

In this research, a novel architecture for radiofrequency microelectromechanical system (RF-MEMS) switches based on new numerical and simulation analyzing approaches. For the proposed RF-MEMS the spring constant (k) as the key characteristic mechanical parameter was analyzed for RF-MEMSs using the modified energy method (MEM). The proposed architecture provided ahighly flexible structure with low actuation voltage (V_{AC}) based on gold (Au) and silicon nitride (Si_3N_4) as the reticular beam and the dielectric layer, respectively. The COMSOL package was applied for analyze and simulate the structure from themechanical point of the view and the findings confirmed that the altering in position and length (L) values of beams not only diminishes the k outstandingly but also provides an ultra-sensitivity for V_{AC} and great concomitance between numerical and simulation k and V_{AC} values as well. The V_{AC} value for the L-dependent numerical k was obtained at1.61 V which closely was validated with simulation outputs that were at 0.08 N m^{-1} and 1.80 V, respectively. Additionally, the switching time (t_s), Von Mises Stress (VMS), natural frequency (f_n) and mass (m) have been introduced as the key characteristic mechanical parameters with values about 25.60 μs, 4.50 MPa, 3118.60 Hz , , and0.21 ng, respectively that could play significant cooperative factors. The results confirm that the return loss remains below -1 dB at 46 GHz during both the ON and OFF states and exceeds -10 dB from 40 GHz to 50 GHz. This study introduces a novel perspective into the design and application paradigm for trials targeting low actuation voltage (V_{AC}) for RF-MEMS switches. Compared with other capacitive RF-MEMS switch , as shown in Table 6.

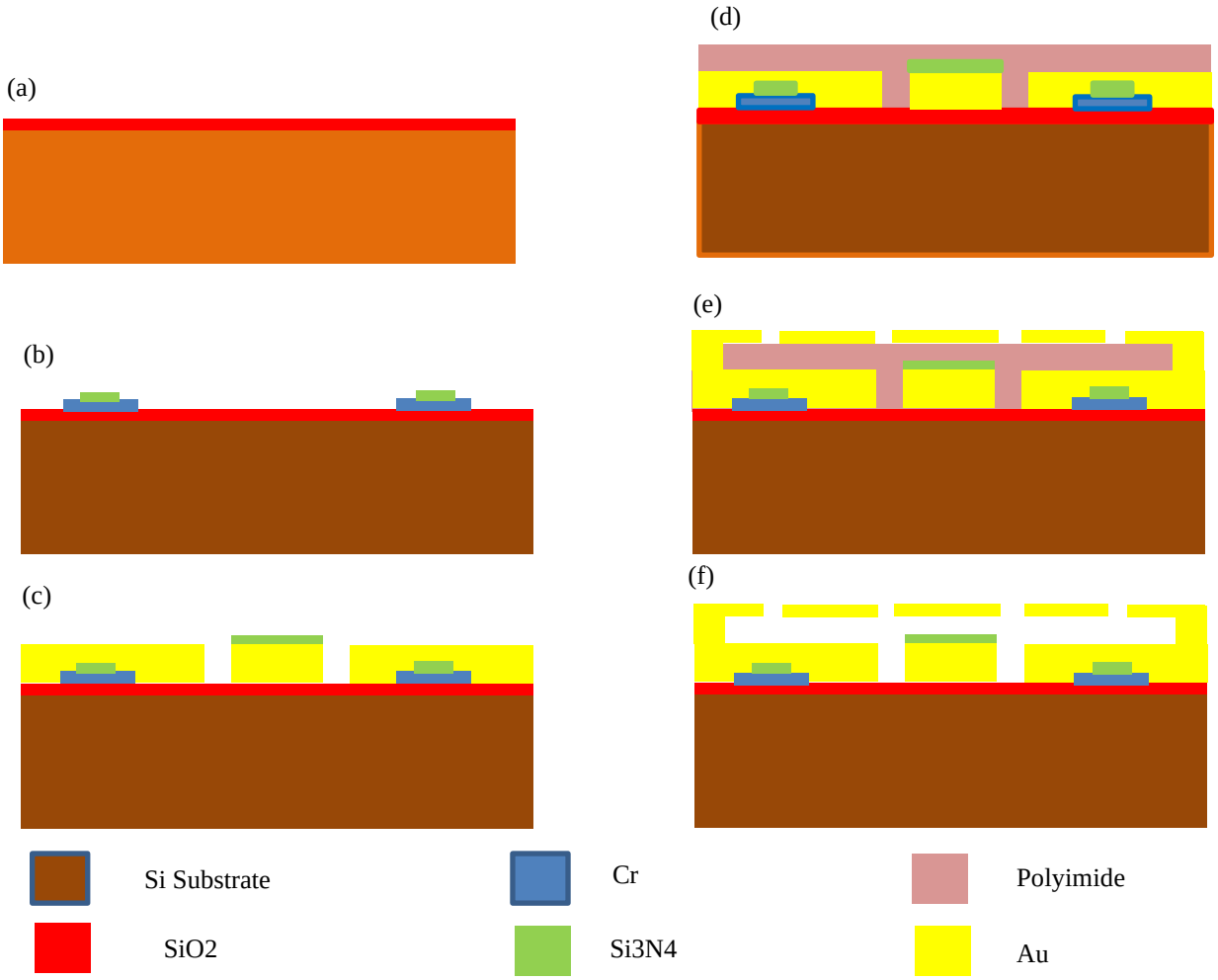


Fig. 9. The device fabrication process: (a) Silicon substrate post-oxidation; (b) Creation of high-resistance direct current (DC)-bias line and patterning of the Si3N4 dielectric layer; (c) Patterning of coplanar waveguide (CPW) lines; (d) Application of polyimide spinning to generate a sacrificial layer followed by chemical mechanical polishing (CMP) to diminish surface roughness of polyimide; (e) Patterning of gold layer to construct MEMS beams and anchors; (f) Removal of the sacrificial layer[8, 22].

TABLE 6: COMPARISON MECHANICAL PARAMETER OF PROPOSED SWITCH WITH PREVIOUS REPORTS

Ref.	Actuation mechanism	Beam material (μm)	Air gap (μm)	Dielectric material (μm)	Actuation voltage(V)	Switching time(μs)	Von Mises Stress
[22]	Electrostatic	1.60(Au)	2.00	(Si_3N_4)	4.40	22	18 MPa
[23]	Electrostatic	0.50	2.00	0.15 (Si_3N_4)	2.64	25.40	N/A
[24]	Electrostatic	0.50	2.00	0.10 (SiO_2)	1.80	11.20	N/A
[19]	Electrostatic	1.20 (Au)	3.70	1.00 (Si_3N_4)	10.60	N/A	41.70 MPa
[25]	Electrostatic	0.50 (Au)	3.00	(HfO_2)	2.40	96.00	N/A
[26]	Electrostatic	1.50 (Au)	2.00	1.50 (Si_3N_4)	3.30	N/A	N/A
[27]	Electrostatic	1.60 (Au)	N/A	0.10 (Si_3N_4)	2.54	N/A	N/A
[13]	Electrostatic	0.60 (Au)	2.50	0.15 (Si_3N_4)	2.10	58.72	N/A
Proposed switch	**Electrostatic**	**0.60 (Au)**	**2.5/2**	**0.15 (Si_3N_4)**	**1.80/1.15**	**25.60**	**4.50 MPa**

References

[1] B. Bakeer, A. Elsabbagh, and M. Hedaya, "Parametric optimization of non-prismatic micro-plates to reduce stiffening and curling initiated during fabrication," in *IOP Conference Series: Materials Science and Engineering*, 2021, vol. 1172, no. 1: IOP Publishing, p. 012021.

[2] G. M. Rebeiz, *RF MEMS: theory, design, and technology*. John Wiley & Sons, 2004.

[3] H. R. Ansari, S. Khosroabadi, and Y. Mafinejad, "Design and simulation of a RF MEMS shunt capacitive switch with low actuation voltage, low loss and high isolation," *Journal of Iranian Association of Electrical and Electronics Engineers*, vol. 18, no. 1, pp. 29-35, 2021.

[4] F. Khan and M. I. Younis, "RF MEMS electrostatically actuated tunable capacitors and their applications: a review," *Journal of Micromechanics and Microengineering*, vol. 32, no. 1, p. 013002, 2021.

[5] A. S. Khan and T. Shanmuganantham, "Arc-shaped cantilever beam RF MEMS switch for low actuation voltage," in *2017 IEEE International Conference on Circuits and Systems (ICCS)*, 2017: IEEE, pp. 302-305.

[6] K. Han, Y. Liu, X. Guo, Z. Jiang, N. Ye, and P. Wang, "Design, analysis and fabrication of the CPW resonator loaded by DGS and MEMS capacitors," *Journal of Micromechanics and Microengineering*, vol. 31, no. 6, p. 065004, 2021.

[7] N. Habbachi and K. Besbes, "RF MEMS filter based on dual liquid variations," *Journal of Micromechanics and Microengineering*, vol. 32, no. 6, p. 065002, 2022.

[8] K. Deng, F. Yang, Y. Wang, C. Lai, and K. Han, "Design and Fabrication of a Ka Band RF MEMS Switch with High Capacitance Ratio and Low Actuation Voltage," *Micromachines*, vol. 13, no. 1, p. 37, 2022.

[9] J. Li, Z. Chen, W. Liu, J. Yang, Y. Zhu, and F. Yang, "A novel piezoelectric RF-MEMS resonator with enhanced quality factor," *Journal of Micromechanics and Microengineering*, vol. 32, no. 3, p. 035002, 2022.

[10] A. K. Sharma and N. Gupta, "Investigation of actuation voltage for non-uniform serpentine flexure design of RF-MEMS switch," *Microsystem technologies*, vol. 20, no. 3, pp. 413-418, 2014.

[11] K. Khodadady and B. A. Ganji, "Design and modeling of a novel RF MEMS series switch with low actuation voltage," *Microsystem Technologies*, vol. 22, no. 12, pp. 2921-2929, 2016.

[12] N. Van Der Meijs and J. Fokkema, "VLSI circuit reconstruction from mask topology," *Integration*, vol. 2, no. 2, pp. 85-119, 1984.

[13] K. Guha, N. Laskar, H. Gogoi, K. Baishnab, and K. S. Rao, "A new analytical model for switching time of a perforated MEMS switch," *Microsystem Technologies*, vol. 26, no. 10, pp. 3143-3152, 2020.

[14] K. Guha, N. Laskar, H. Gogoi, A. Borah, K. Baishnab, and S. Baishya, "Novel analytical model for optimizing the pull-in voltage in a flexured MEMS switch incorporating beam perforation effect," *Solid-State Electronics*, vol. 137, pp. 85-94, 2017.

[15] Y. Mafinejad, A. Kouzani, K. Mafinezhad, and R. Hosseinnezhad, "Low insertion loss and high isolation capacitive RF MEMS switch with low pull-in voltage," *The International Journal of Advanced Manufacturing Technology*, vol. 93, no. 1, pp. 661-670, 2017.

[16] F. P. Beer, E. Johnston Jr, M. Russell, F. David, and E. R. Eisenberg, "Vector Mechanics for Engineers: Statics (SI Units)," *McGraw Hill Higher Education, 9th Revised edition (October 2010), Capítulo*, vol. 3, p. 75, 2019.

[17] F. Khamoii Toli and J. Yavand Hasani, "Design, simulation and optimization of an RF MEMS capacitive switch to reduce actuation voltage," *Journal of Iranian Association of Electrical and Electronics Engineers*, vol. 19, no. 2, pp. 1-11, 2022.

[18] V. K. Varadan, K. J. Vinoy, and K. A. Jose, *RF MEMS and their applications*. John Wiley & Sons, 2003.

[19] K. G. Sravani *et al.*, "Designing of RF-MEMS Capacitive Contact Shunt Switch and Its Simulation for S-band Application," in *Micro and Nanoelectronics Devices, Circuits and Systems*: Springer, 2022, pp. 439-448.

[20] A. Tkachenko, I. Lysenko, M. Denisenko, and O. Ezhova, "Design and Optimization of a Shunt RF MEMS Switch with a Hybrid Contact Type," in *International Youth Conference on Electronics, Telecommunications and Information Technologies*, 2022: Springer, pp. 281-299.

[21] A. Tkachenko, I. Lysenko, and A. Kovalev, "Investigation and Research of High-Performance RF MEMS Switches for Use in the 5G RF Front-End Modules," *Micromachines*, vol. 14, no. 2, p. 477, 2023.

[22] Z. Deng, C. Lai, J. Zhou, and Y. Wang, "Design and analysis of a novel low RF MEMS switch with low pull-in voltage and high capacitance ratio," *Microsystem Technologies*, pp. 1-13, 2023.

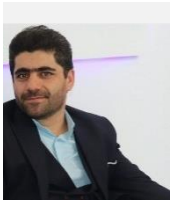
[23] M. Gaikwad, N. Deshmukh, and V. Sawant, "Electromagnetic Modelling and Parameters Extraction of Metal Contact and Capacitive Type RF MEMS Switch," in *Artificial Intelligence and Sustainable Computing*: Springer, 2022, pp. 143-153.

[24] P. Ashok Kumar, S. R. Karumuri, G. S. Kondavitee, and K. Guha, "Design and performance analysis of a low-pull-in-voltage RF MEMS shunt switch for millimeter-wave therapy, IoT, and 5G applications," *Journal of Computational Electronics*, pp. 1-8, 2022.

[25] L. N. Thalluri, M. K. Reddy, S. R. Hussain, G. C. Reddy, S. Kiran, and K. Guha, "Iterative Approach for Low Actuation Voltage RF MEMS Switch," in *Micro and Nanoelectronics Devices, Circuits and Systems*: Springer, 2022, pp. 117-128.

[26] K. S. Rao *et al.*, "Design and analysis of RF MEMS shunt switch for V-band applications," *Microsystem Technologies*, pp. 1-8, 2022.

[27] Y. Liu, K. Han, Z. Jiang, N. Ye, and P. Wang, "A High Capacitive Ratio and Low Actuation Voltage RF MEMS Switch for Multi Band: Design and Performance Analysis," in *2021 9th International Conference on Communications and Broadband Networking*, 2021, pp. 305-309.



Farid Khamouei Touli was born in Urmia, Iran, in 1990. He received his B.S. and his M.S. degree in Electronics Engineering from Urmia University of Technology, Urmia and Iran University of Science and Technology (IUST),Tehran, Iran, in 2015 and 2019 respectively. He is now an Ph.D. candidate in electrical engineering at Iran University of Science and Technology(IUST),Tehran, Iran. His research interests include Micro-fabrication, RF-MEMS Swithes, Charge Accumulation and Reliability Issue in RF MEMS. He focus on the modeling, design and fabrication of RF-MEMS passive components for 5G& 6G internet.

Dr. Javad Yavand Hasani is a professor in the School of Electrical Engineering at Iran University of Science and Technology (IUST), Tehran, Iran. In 2009 he received the PhD degree in electrical engineering from the University of Tehran and the PhD degree in high frequency and optics from the University of Joseph Fourier (UJF), Grenoble, France. From 2012–2014 he was the dean of the Electronic Research Center (ERC) in IUST. His current research interests include high frequency circuits and systems, as well as micro-fabrication theory and technology for RF and microwave devices, MEMS and NEMS..

IECO International Journal Of Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

Improvement of Battery State of Charge Estimation Using Recursive Least Squares-Based Adaptive Extended Kalman Filter

Ramazan Havangi[1]| Fatemeh Karimi[2]

Faculty of Electrical Engineering and Computer, University of Birjand, Birjand, Iran.[1,2]
Corresponding author's email: Havangi@Birjand.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 30-January-2024
Received in revised form: 26-March-2024
Accepted: 15-April-2024
Published online: 21-June-2024

Keywords:
Battery charge,
State of the Charge (SOC),
Kalman Filter,
Estimation.

ABSTRACT

This paper presents an advanced methodology for SOC estimation by integrating Recursive Least Squares (RLS) techniques with Adaptive Extended Kalman Filter (AEKF). The proposed methodology aims to mitigate the challenges associated with fluctuating battery parameters and varying noise characteristics over time, which can significantly impact the accuracy of SOC estimation. By dynamically adjusting to evolving system dynamics and noise statistics, the proposed approach exhibits enhanced robustness and accuracy compared to traditional techniques. The proposed methodology assumes that battery parameters, including internal resistance, capacitance, and noise information, undergo variations over time. To address this assumption, two distinct online identification algorithms for parameters and noise information are introduced. Specifically, the RLS algorithm is utilized to ascertain resistance and capacitance values. Process and measurement noise covariance is also estimated based on an iterative noise information identification algorithm. Subsequently, all updated values are incorporated into the EKF.
The results demonstrate that the RLS-AEKF approach achieves higher accuracy than the EKF. The results based on Fast Urban Driving Schedule (FUDS) and Urban Dynamometer Driving Schedule (UDDS) working current profiles validate the effectiveness of the proposed approach in enhancing SOC estimation accuracy under realistic operating conditions.

I. Introduction

The global community is confronted with substantial environmental challenges, notably greenhouse gas emissions and global warming, which stem predominantly from the widespread utilization of petrol and diesel in vehicular operations. These fossil fuels are major contributors to the annual CO_2 emissions [1, 2]. In response to this pressing issue, electric vehicles (EVs) have emerged as a promising solution to mitigate carbon emissions, garnering significant attention amid the prevailing energy crisis [3, 4]. Batteries, renowned for their comparatively high energy density, low noise emissions, and minimal maintenance demands, have garnered widespread utilization in electricity storage technologies on both small and medium scales [5, 6]. Among various battery types, lithium-ion batteries are extensively employed in diverse EV applications. Given the paramount importance of ensuring the safe and dependable operation of lithium-ion batteries, implementing a battery management system (BMS) is indispensable [6, 7].

Battery management technologies encompass a variety of estimations, including the assessment of parameters such as the state of charge (SOC), state of temperature (SOT), state of energy (SOE), state of power (SOP), state of health (SOH), and state of safety (SOS), among which SOC holds pivotal significance. SOC represents the proportion of a battery's remaining charge to its designated capacity. Precise anticipation of battery SOC serves as a preventive measure against both overcharging and over-discharging, thereby contributing to the overall extension of the battery's operational lifespan [1, 2].

Current prevalent methods for estimating SOC include the ampere-hour integration method, open circuit voltage

(OCV) method, data-driven method, and model-based method [1]. The ampere-hour integration method encounters challenges in accurately determining the initial SOC, and errors in current measurements exacerbate cumulative errors post-integration [8, 9]. The OCV method necessitates an extended resting period for obtaining the open circuit voltage, rendering it unsuitable for real-time SOC estimation [10, 11]. The data-driven method has garnered considerable attention for its adaptability and model-free advantages [11, 12]. However, this method relies heavily on the accuracy of training data, and it cannot guarantee precise SOC estimation under complex and unknown operational conditions [6, 7]. Artificial neural networks and fuzzy logic are the commonly employed data-driven methods [14, 15].

Unlike alternative approaches, model-based estimation methods offer enhanced robustness and accuracy, remaining unaffected by initial values and exhibiting lower computational overheads, rendering them well-suited for engineering applications [16]. The model-based approach requires the establishment of a battery model. Among model-based methods, the extended Kalman filter (EKF) is widely recognized as the most popular technique for estimating the SOC of batteries [17]. The EKF, introduced to enhance the performance of the linear Kalman filter through Taylor-series expansion, has been extensively utilized [17].

Different variants of the EKF are utilized for SOC estimation. For example, the dual-mode EKF is applied in [19, 20] to address aging phenomena. Additionally, adaptive EKF is employed in [21-23] to achieve rapid transient response.

Investigations into the SOC estimation show a close interdependence between the estimation accuracy and the values of the model parameters. It motivates the literature to propose a new method known as adaptive estimation. For instance, the problem of SOC estimation in the presence of the model uncertainty is discussed in [24], in which the authors have tried to solve the problem from both robust and adaptive estimation viewpoints. Similarly, in [25], the variation of the system parameters is considered in the adaptive EKF (AEKF) algorithm.

Although the literature commonly combines EKF and recursive least squares (RLS) methods, none have integrated noise information estimation within this framework. For example, [26] focuses on SOC estimation for EVs by employing a combination of RLS and nonlinear EKF. Their study emphasizes the importance of accurately updating battery parameters, particularly amidst fluctuating temperature and SOC conditions, to enhance SOC estimation precision. Similarly, [21] investigates SOC estimation for power batteries in unmanned aerial vehicles (UAVs), highlighting the significance of lithium-ion batteries in UAV development and the need for precise SOC estimation to optimize flight performance. They propose a joint estimation algorithm utilizing a second-order resistor-capacitor (RC) network equivalent circuit model and an EKF for SOC estimation. Additionally, [27] addresses SOC estimation for hybrid vehicles, proposing an adaptive forgetting factor regression least-squares–extended Kalman filter (AFFRLS–EKF) strategy to improve accuracy amid changing charge and discharge conditions. While these studies offer valuable insights into SOC estimation, our research in this paper advances these methodologies by incorporating noise information estimation into the SOC estimation framework, thereby enhancing the accuracy and robustness of our approach.

The literature review shows that a gap still exists in the adaptive estimation of SOC in the presence of the model uncertainty. To provide more detail, the literature does not use parallel filters for SOC estimation when dealing with model uncertainty. To do this, the RLS-AEKF method presented in this paper integrates the principles of RLS and EKF to estimate the SOC of batteries. Unlike conventional EKF approaches, which rely on fixed process and measurement noise covariance matrices, RLS-AEKF adaptively adjusts these covariance matrices based on real-time data and system dynamics. This adaptive mechanism enables the RLS-AEKF algorithm to effectively handle uncertainties and variations in battery behavior, resulting in improved estimation accuracy and robustness. RLS-AEKF dynamically adjusts the process and measurement noise covariance matrices based on the evolving system dynamics and measurement uncertainties. This adaptive mechanism enhances the algorithm’s resilience to changing operating conditions and improves estimation accuracy. Experimental studies and simulation results have demonstrated the robust performance of RLS-AEKF across various operating conditions. In summary, the superiority of the proposed method over other approaches is demonstrated in Table 1.

The paper is structured as follows. Section II provides the mathematical background of the system, including the battery equivalent model and relevant definitions. Section III discusses the fundamental concepts of the EKF design procedure. The adaptive SOC estimation is detailed in Section IV, where system identification and adaptive SOC estimation algorithms are presented. Section V examines the results to validate the proposed method. Finally, Section VI concludes the paper.

TABLE 1 ADVANTAGE OF THE PROPOSED METHOD

The proposed method	Other works
Introducing two distinct online identification algorithms for parameter and noise information: recursive least squares for determining resistance and capacitor values and an iterative method for estimating noise covariance	Utilizing existing methods without mentioning online identification algorithms for parameter and noise information.
Handling varying battery parameters and noise information over time to enhance SOC estimation accuracy	Specifically not addressing the challenge of handling time-varying battery parameters and noise information
Demonstrating improvement in SOC estimation accuracy by addressing uncertainties in model parameters and noise information through the proposed algorithms	It may not clearly show improvement in SOC estimation accuracy due to not addressing uncertainty in model parameters and noise information.

II. Mathematical Background

In model-based estimation, the model's accuracy leads to an accurate estimation, which is related to the dependence of the estimators on the model. Therefore, the first step in SOC estimation is to use a more accurate model. Different models have been addressed in the literature from the system identification point of view. For example, different models are explained in [1], where the main difference is to use a series of RC filters in the equivalent circuit.

Based on the literature, a common model of the battery is considered to be as shown in Fig. 1. The proposed method is in fact an equivalent model of a battery when driving a load. As can be seen, the battery is modeled as a DC voltage source connected to an RC filter and series resistance. Based on the equivalent circuit, the battery terminal voltage is achieved when the open circuit voltage passes through a filter. In a more accurate view, it can be seen that getting more power from the battery (at a constant voltage) leads to more loss in power when passing through the series filter. The mathematical model of the battery can be divided into two parts: an SOC model and an equivalent circuit model. This subsection presents these models separately, and then the final model is produced by integrating them.

A. Battery SOC definition

A battery's SOC is defined as the ratio of the available capacity to the maximum potential capacity that it can supply. The mathematical representation of SOC is as follows:

$$SOC(t) = SOC(t_0) - \frac{\eta}{C_n}\int_{t_0}^{t} I(t)\,dt \qquad (1)$$

where $SOC(t_0)$ and $SOC(t)$ represent SOC at initial time t and at time t_0, respectively, η represents the coulombic efficiency, which signifies the ratio of the battery discharge capacity to the charge capacity during the same cycle. Also, C_n is the battery's total capacitor and $I(t)$ shows its terminal current. Note that in Eq. (1), the positive sign is used for the battery charging, and correspondingly, the negative sign is used for the discharging. A discrete representation of (1) is as follows:

$$SOC(k) = SOC(k-1) - \frac{\eta I(k) T_s}{C_n}$$

where T_s denotes the sampling interval and $I(k)$ signifies the current load. $SOC(k)$ and $SOC(k-1)$ denote the battery's SOC at time step k and k-1, respectively.

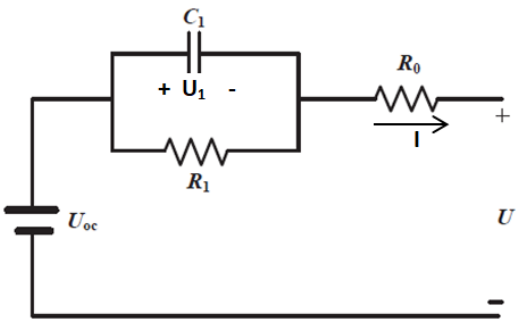


Fig. 1. Equivalent circuit of the battery

B. Battery Dynamical Model

To write the dynamical model of the battery, researchers have used the basic Kirshef laws to write the dynamical model. In more detail, note that in Fig.1, U_{oc} depicts the OCV. Also, U is the terminal voltage of the battery and the inner capacity and resistance is shown with C_1 and R_1, respectively. Therefore, the mathematical model of the battery can be written as:

$$U = U_{oc} - R_o I - U_1$$
$$\frac{dU_1}{dt} = \frac{-U_1}{R_1 C_1} + \frac{I}{C_1} \qquad (2)$$

in which U_1 is the voltage of the RC circuit. In addition, the total current of the battery is shown by I. The battery voltage can be written with the solution of the second differential equation as:

$$U_1(k) = U_1(k-1)e^{-\frac{T_s}{R_1 C_1}} + R_1 I(k)\left(1 - e^{-\frac{T_s}{R_1 C_1}}\right) \qquad (3)$$

in which $I(k)$ is the value of the total current in the sample of k. Similarly, the notations $U_1(k)$ and $U_1(k-1)$

shows the voltage of the RC circuit in the sampling time of k and $k-1$, respectively. If the current and voltage with the related parameters of the above equations are known, then the OCV can be written as:

$$U_{OC}(k)=U(k)+I(k)R_o+U_1(k-1)e^{-\frac{T_s}{R_1C_1}} +R_1I(k)(1-e^{-\frac{T_s}{R_1C_1}}) \quad (4)$$

Note that Eq. 3 and 4 can be described as:

$$x(k)=A(k)x(k)+B(k-1)u(k-1)+\omega(k)$$
$$y(k)=g\left(x(k),u(k)\right)+\upsilon(k) \quad (5)$$

in which $\omega(k)$ and $\upsilon(k)$ are the process noise and measurement noise, respectively. $A(k-1)$ and $B(k-1)$ are as follows:

$$A(k-1)=\begin{bmatrix}1 & 0\\ 0 & e^{-\frac{T_s}{R_1C_1}}\end{bmatrix} \quad B(k-1)=\begin{bmatrix}-\eta\frac{T_s}{C_n}\\ R_1(1-e^{-\frac{T_s}{R_1C_1}})\end{bmatrix}$$

Also, the state vector and the input vector are considered to be $x(k)=\left[SOC(k) \quad U_1(k)\right]^T$ and $u(k)=I(k)$, respectively. Furthermore, the measurement equation is the second equation of Eq. (5). This equation is indeed (4) when considering $y(k)=U_{OC}(k)$ and $g\left(x(k),u(k)\right)$ as the right-hand side of (4).

As can be seen from the final relation between the OCV and the terminal voltage of the battery, it can be concluded that there are parameters in the model whose time-varying behavior will affect the battery behavior. For example, assume the calculation of the SOC using (1). Under the time-varying behavior of the system, we cannot conclude an accurate value of the SOC due to the existence of uncertainty in the terminal current. Through the literature, some tests are defined to calculate the battery parameters as accurately as possible. For further explanation, the capacity-temperature test, the OCV-SOC temperature test, the internal resistance temperature test, and some other tests are defined in the literature to extract more accurate parameters and models for batteries. The next section describes the model of the battery proposed in this paper and introduces the parameters and the area of uncertainty in system modeling.

III. EKF for SOC estimation

The Kalman filter offers an optimal solution for filtering problems under linear and Gaussian conditions. However, SOC estimation encounters difficulties due to its nonlinear nature. Addressing nonlinear filtering problems effectively requires a thorough description of their conditional posterior probability. Yet, providing such a detailed description involves an infinite number of parameters, rendering it impractical for real-world applications. Consequently, various suboptimal approximation methods have been proposed to tackle this challenge. An example is the EKF used to linearize nonlinear systems for optimal estimation, which is achieved by expanding the nonlinear system using the Taylor series, disregarding high-order terms, and approximating the nonlinear system with the first-order term. Subsequently, the Kalman filter method is employed to estimate the state. Let us assume that the state equation and observation equation are as follows:

$$X(k)=AX(k-1)+BU(k)+W(k)$$
$$Y(k)=g(X(k),U(k))+V(k) \quad (6)$$

in which $X(k)$ is the system state, $U(k)$ is the system input, and $Y(k)$ is the system output. Furthermore, the process and measurement noises are shown with $W(k)$ and $V(k)$. Thus, the future prediction of the system state is written as:

$$X(k|k-1)=AX(k-1|k-1)+BU(k) \quad (7)$$

in which the predicted value is shown using $X(k|k-1)$. Moreover, the optimal value of the state is shown with $X(k-1|k-1)$. In the next step, the covariance matrix is predicted as:

$$P(k|k-1)=AP(k-1|k-1)A^T+Q(k) \quad (8)$$

in which $P(k|k-1)$ is the covariance matrix. Also, the parameter $Q(k)$ is the covariance matrix of the process noise $W(k)$. An optimal estimation for the states is as follows:

$$X(k|k)=X(k|k-1)+ K(k)\left(Y(k)-g\left(X(k|k-1),U(k)\right)\right) \quad (9)$$

in which the matrix of $K(k)$ is the filter gain which is calculated using the following equation:

$$K_k=P(k|k-1)H^T(k)(H(k)P(k|k-1)H^T(k) +R(k))^{-1}$$
$$H=\frac{\partial g}{\partial x} \quad (10)$$

where the matrix of $R(k)$ is the covariance matrix of measurement noise. The covariance is updated as follows:

$$P(k|k)=\left(I-K(k)H(k)\right)^{-1}P(k|k-1) \tag{11}$$

Also, the measurement equation is as follows:

$$U(k)=U_{OC}-U_1(k)-R_oI(k)+V(k) \tag{12}$$

In order to reduce the complexity of the above equation, the system linear model is rewritten as follows using the Taylor expansion equations:

$$\begin{bmatrix} SOC(k) \\ U_1(k) \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1-\frac{T_s}{\tau} \end{bmatrix}\begin{bmatrix} SOC(k-1) \\ U_1(k-1) \end{bmatrix} + \begin{bmatrix} -\eta\frac{T_s}{C_n} \\ \frac{T_s}{C_1} \end{bmatrix}I(k-1)+\begin{bmatrix} W_1(k) \\ W_2(k) \end{bmatrix} \tag{13}$$

in which C_n is the battery capacity in the n-th cycle and $\tau=R_1C_1$ is the time constant of the system. Based on the estimation theory, the accuracy of the estimation directly depends on the model accuracy and noise information. This can be understood from the estimation equations in different steps. Therefore, uncertainty in the dynamical model and noise information impairs the quality of SOC estimation. To solve this, as the main novelty of the paper, we present a new method to estimate the battery SOC in the presence of the mentioned uncertainties.

IV. Adaptive SOC estimation

This section is divided into two parts. The first part presents the AEKF. The second part introduces our novel approach known as RLS-AEKF.

A. SOC estimation using AEKF

A combination of the conventional EKF with an adaptive law is known as AEKF, which is given in [23-25]. The two steps of the model prediction and state correction construct the main framework of the KF theory. In the prediction step, the state values in the sampling time k are predicted. In this step, the prediction is accomplished using the system model. Therefore, obviously, the accuracy of the estimation depends on the model accuracy. In the correction step, the estimated values are updated based on the current measurement at the $k-th$ sampling time. Apart from the model accuracy, having accurate information about the process and measurement noise is a solution key to having high-quality estimation. In the AEKF of reference [25], the adaptive law is used to update the covariance matrix over time and parallel to the EKF algorithm. Therefore, the AEKF algorithm of reference [25] uses the following update law to estimate the covariance values over time.

$$\begin{aligned} e(k)&=Y(k)-g(X(k)) \\ L(k)&=\frac{1}{N}\sum_{i=k-N+1}^{k}e(k)e(k)^T \\ R_k&=L_k+H(k)P(k|k)H(k) \\ Q_k&=K(k)L_kK(k)^T \end{aligned} \tag{14}$$

where i =k −N+1 denotes the initial sample within the estimation window. The selection of the window size, denoted as N, is determined empirically to provide a degree of statistical smoothing. $e(k)$ represents the calculation of error innovation. Therefore, using Eq. (14), the noise information is updated over time. This method has some merits and drawbacks. As a merit, this method can estimate the time-varying behavior of the noise over time. On the other hand, its drawback is that it increases the computation burden. Specifically, a good estimation of noise information needs an adequately long data window (a larger value for N), which increases the computation burden. However, it seems to work acceptably for noises with slow time changes in information, as seen in the results of [25].

B. SOC estimation using RLS-AEKF

In the previous section, we discussed the common SOC background in the literature. The main gap in the SOC estimation using the aforementioned method is the use of a predefined model. The experimental results in the literature show a notable dependence between estimation and model accuracy. Consider the second row of Eq. (9), which is given as follows:

$$U_1(k)=(1-\frac{T_s}{\tau})U_1(k-1)+(\frac{T_s}{C_1})I(k-1)$$

Defining the model parameters as

$$\theta_1=1-\frac{T_s}{\tau}$$

$$\theta_2=\frac{T_s}{C_1}$$

results in Eq. (15) as follows:

$$U_1(k)=\theta_1U_1(k-1)+\theta_2I(k-1) \tag{15}$$

Eq. (6) can be written as

$$\phi(k)=[U_1(k-1)\quad I(k-1)]^T$$

$$\theta=\begin{bmatrix}\theta_1 & \theta_2\end{bmatrix}^T$$

Then, $U(k)$ can be written as:

$$U(k)=\phi^T(k)\theta(k)+\xi(k)$$

where $\xi(k)$ is set as the sampling error of the sensor at time k. Therefore, if it is possible to estimate $U_1(k),U_1(k-1)$

and $I(k-1)$, the parameters θ_1 and θ_2 can be estimated. In more detail, it leads to the estimation of R_1 and C_1. To do that, we use the RLS method using Eq. (16) as follows:

$$\begin{aligned}\hat{\theta}(k) &= \hat{\theta}(k-1) + K(k)\left(y(k) - \varphi^T(k)\hat{\theta}(k-1)\right)\\ K(k) &= P(k-1)\varphi(k)\left(\lambda + \varphi^T(k)P(k-1)\varphi(k)\right)^{-1}\\ P(k) &= \left(I - K(k)\varphi^T(k)\right)P(k-1)/\lambda\end{aligned} \quad (16)$$

in which $y(k) = U_1(k)$. The parameter $\lambda \in (0,1]$ is the forgetting factor. Furthermore, the matrix $P(k)$ is the covariance matrix of the algorithm. After identifying $\theta(k)$, the parameters of the Thevenin model can be obtained as follows:

$$C_1(k) = \frac{T_s}{\theta_2(k)}$$

$$R_1(k) = \frac{T_s}{C_1(k)(1-\theta_1(k))}$$

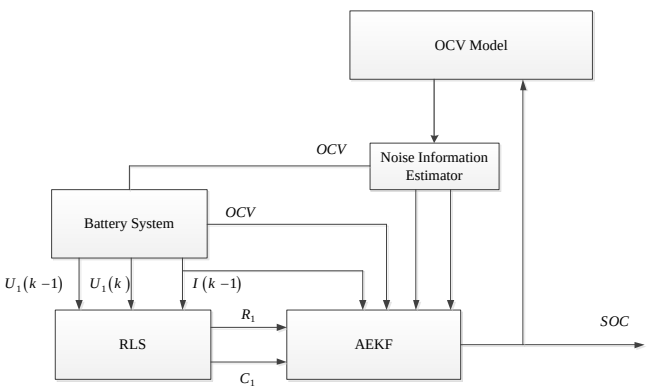


Fig. 2. The block diagram of the proposed SOC estimation

The combination of the proposed methods is presented in Fig. 2. As is seen in Fig. 2, the SOC is estimated with a composition of the RLS and adaptive EKF methods.

In more detail, the model's inner resistance and capacitor values are first estimated using the RLS identification method. Then, the values are delivered to the EKF to update the model. After that, the EKF estimates the SOC more accurately. As can be seen, the filter measurement does not depend on the other parts. Thus, the uncertainty of the measurement model also affects the estimation accuracy. To solve it, an accurate identification of the OCV curve is needed. Assuming information accuracy, we have calculated an interval for the estimation accuracy based on the OCV characteristics curve.

V. Results

The proposed method was implemented using MATLAB on a standard laptop system equipped with an Intel Core i7 processor and 16GB of RAM. The simulation result for SOC estimation is addressed for the proposed method compared with EKF and RLS-EKF. This makes comparison possible and shows how combining adaptive estimation and system identification can improve the results.

A. Model characteristics and time-varying behavior

In the mathematical model of the battery, we assume that there is 10% uncertainty in the SOC-OCV curve, which is the total effect of some uncertainty sources listed in the literature. Therefore, if assuming the characteristic curve as follows:

$$\begin{aligned}OCV &= 1.8666SOC^5 - 5.0687SOC^4 + 5.7086SOC^3\\ &-3.5125SOC^2 + 1.7145SOC + 3.4400\end{aligned} \quad (17)$$

then, the assumption of 10% uncertainty in the model parameters leads to Fig. (3) for the description of the characteristic curve.

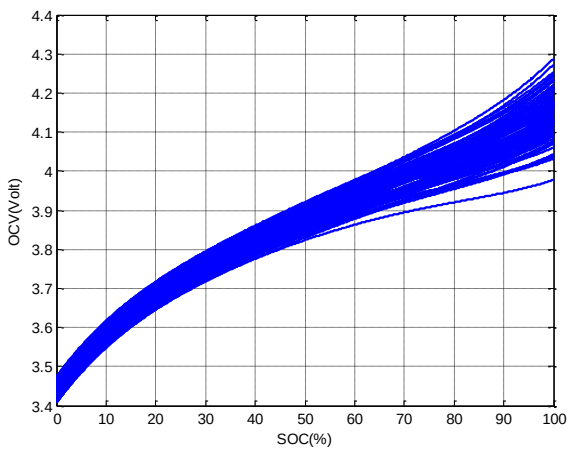


Fig. 3. Sensitivity of the battery characteristic curve with respect to 1% uncertainty in the 5-th order model parameters

As can be seen in Fig. 3, the effect of uncertainty appears to be more destructive in high voltages or high values of SOC. Similarly, the same is accomplished for the resistance and capacitor. To handle that, firstly, the nominal value of the mentioned parameters when there is no uncertainty in the values is given by:

$$\begin{aligned}R_1 &= 0.02346 - 0.10537SOC + 1.1371SOC^2\\ &-4.55188SOC^3 + 8.26827SOC^4 - 6.93032SOC^5\\ &+2.1787SOC^6\end{aligned} \quad (18)$$

$$\begin{aligned}C_1 &= 203.1404 + 3522.78847SOC\\ &-31392.66753SOC^2 + 122406.91269SOC^3\\ &-227590.94382SOC^4 + 198281.56406SOC^5\\ &-65171.90395SOC^6\end{aligned} \quad (19)$$

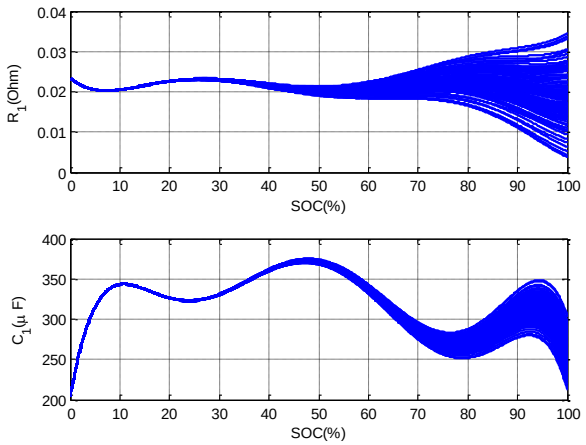


Fig. 4. The change in the resistance and capacitor values with respect to the battery SOC for 0.01% change in the polynomial curve for the calculation of the parameters

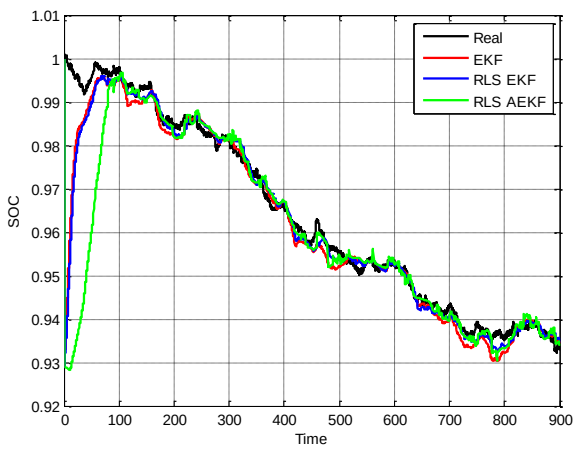


Fig. 5. The SOC estimation using the EKF, RLS-EKF, and RLS-AEKF

The above equations shall be used in simulations. However, we consider a notable uncertainty (10%) in the parameters. To have a more understandable view of the mentioned parameters and their behavior based on the operating point of the battery, we illustrate the resistance and capacitor bound in Fig. 4.

Based on the results displayed in Fig. 4, we can see that there is a notable sensitivity in the curve due to the identification error in the high values of the SOC. Therefore, it seems to be better to tune the filter robustness or adaptive behavior in the bigger values of the SOC. In other words, the identification error in small values of the SOC may not encounter the estimation with a notable error.

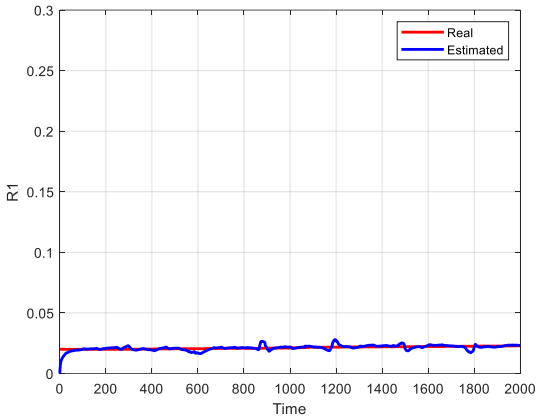


Fig. 6. The battery resistance estimation over the time using the RLS method

B. Estimation Results

This subsection presents the results for SOC estimation using RLS-AEKF, RLS-EKF [3], and EKF. It is worth noting that in the first three filters mentioned, the values of R and Q are set equal to $Q = diag\left[2\times10^{-8} \quad 1\times10^{-8}\right]$ and $R = 10^{-6}$, respectively. Finally, the simulation results for SOC estimation with the given system and estimator parameters are presented in Fig. 5. In these simulations, a 10% uncertainty is assumed in parameter identification.

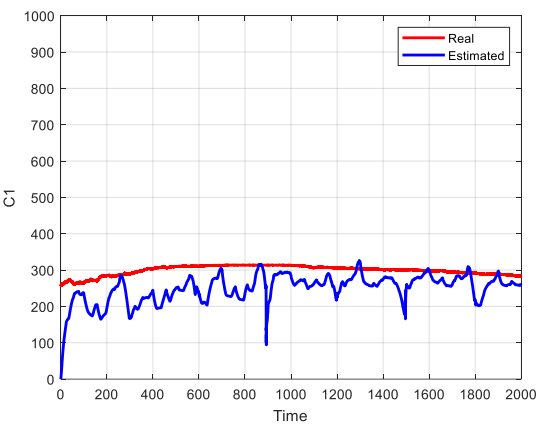


Fig. 7. The battery capacitor estimation over the time using the RLS method

Based on the results in Fig. 5, it can be concluded that the adaptive law in the estimation algorithm will remove the estimation bias. In other words, the uncertainty in the model parameters will conclude a notable bias in the estimation. The RLS estimation will update the model over time and give a more real model to the estimation algorithm (EKF). To complete the simulation results, the results of the RLS algorithm are shown below. As noted before, in the simulation, we assume that there is 10% uncertainty in the

model.

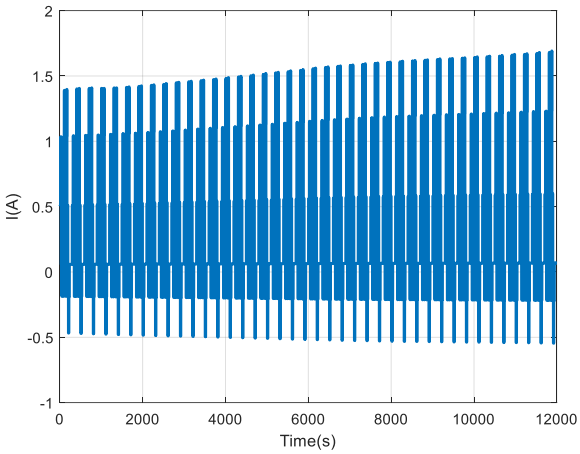


Fig. 8. The battery current during the SOC estimation

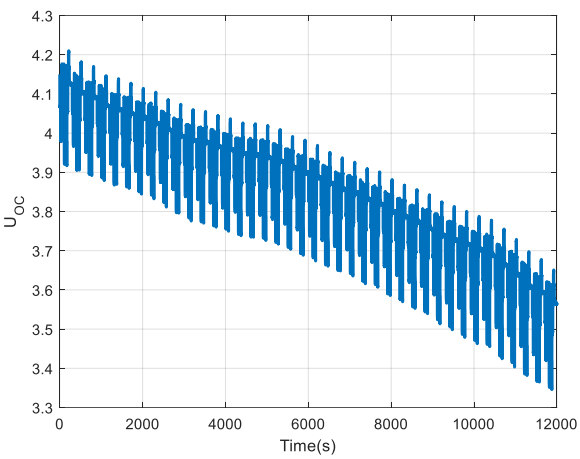


Fig. 9. The battery voltage during the SOC estimation

Figs. 6 and 7 show the parameter estimation of the algorithm. As can be seen, in the first times of the simulation, the unknown parameters are estimated, which is the main part of the paper. Furthermore, the battery terminal current and voltage are the inputs of the estimation algorithm, which is illustrated in Figs. 8 and 9, respectively. As seen in these figures, the online identification algorithm of RLS can successfully estimate the parameter values over time. Note that the ripple in the estimation value is caused by the ripples in the battery current, which is the algorithm's input. To reduce the ripple in the identification results, we used a first-order filter as $H(s)=1/(20s+1)$. The introduced filter deduces the fluctuation in the output of the RLS and causes the EKF to give far better results. Moreover, the terminal voltage and current behavior is illustrated in Figs. 8 and 9. Based on the figures, there is a ripple in the output, influencing the identification quality.

Also, for more details, the figure of the estimation error is given in Fig. 10, which proves the improvement of the SOC estimation using the online parameter estimation of RLS.

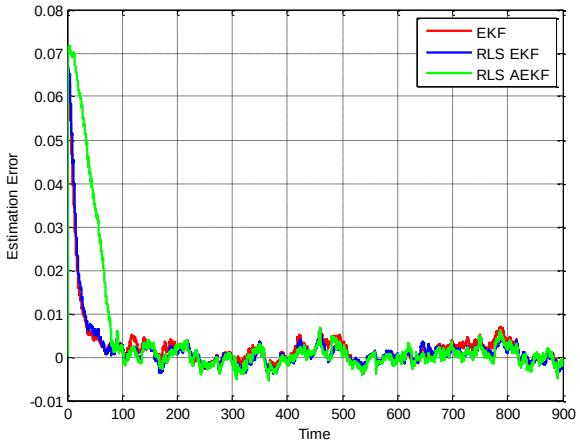


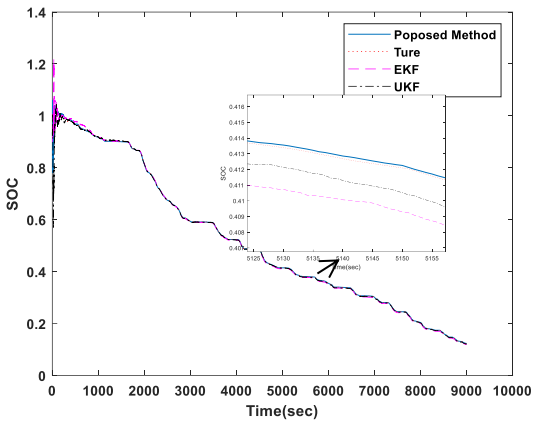
Fig. 10. The SOC estimation error using RLS-EKF and RLS-AEKF

To investigate the results more accurately, we use the root mean squared error (RMSE) criteria. Table 2 presents the RMSE associated with SOC estimation conducted by EKF, RLS-EKF [3], and RLS-AEKF. The data in Table 2 clearly demonstrate that the proposed approach surpasses alternative methods in terms of performance.

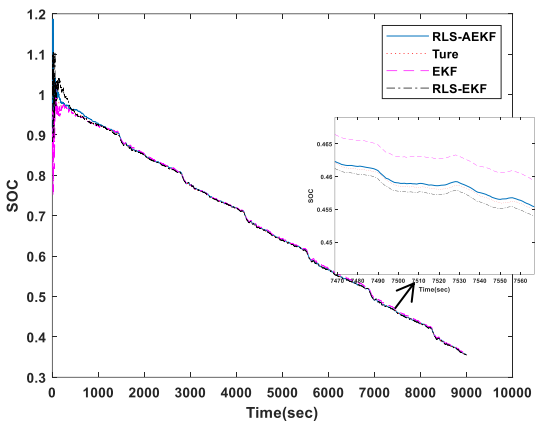
TABLE 2 RMSE OF SOC

	RMSE	**Prossocing Time**
EKF	0.033	0.8897
RLS-EKF[3]	0.0021	0.9928
RLS-AEKF	0.001	1.210

The accuracy and robustness of the proposed method are assessed using the working current profiles of the Fast Urban Driving Schedule (FUDS) and Urban Dynamometer Driving Schedule (UDDS), which are laboratory test methods used to characterize the performance of batteries, particularly for EVs. They involve applying a series of discharge and charge pulses at different magnitudes to the battery under controlled conditions. The experimental conditions are standardized with a fixed temperature of 0°C to streamline the procedures. Fig.11 illustrates the corresponding current profiles of FUDS and UDDS. Table 3 shows the estimated results of the methods under UDDS and FUDS conditions. Based on Table 3, the RMSE of the proposed method is lower than that of other methods under three working conditions.



(a)



(b)

Fig.11. The RMSE of SOC at 0◦C(a) UDDS (b) FUDS

TABLE 3 RMSE UNDER DIFFERENT CONDISTIONS

	RLS-AEKF	RLS- EKF	EKF
UDDS	0.0087	0.011	0.018
FUDS	0.0058	0.0089	0.0108

Ultimately, the scenario where measurement noise is erroneously configured is examined. Table 4 presents a comparative analysis between the RLS-AEKF and conventional techniques, such as EKF and RLS-EKF, under FUDS and UDDS scenarios. The results indicate that the RLS-AEKF exhibits superior accuracy compared to its counterparts, highlighting its robustness in handling unknown noise statistics. The RMSE is computed over 50 Monte Carlo iterations, justifying the adaptive adjustment of process and measurement noise covariance within the proposed framework.

TABLE 4 RMSE UNDER UNKHOWN NOISE STATISTICS

	RLS-AEKF	RLS- EKF	EKF
UDDS	0.0089	0.021	0.042
FUDS	0.006	0.017	0.033

C. Monte Carlo Analysis

This sub-section investigates the problem using a Mont Carlo simulation test to investigate the effect of the parameter uncertainties on the estimation quality. To this aim, we will run the simulation more times and then calculate the error bound of the SOC estimation error. Based on the results given in Fig. 13, we can conclude that the performance of the EKF is 2% (SOC estimation error), which is a notable error in the range of 0 to 100%. This shows the weak robustness of the filter against the parameter variation. The results for AEKF are given below. The results show that the AEKF can estimate the SOC with an accuracy of 0.4% in most cases.

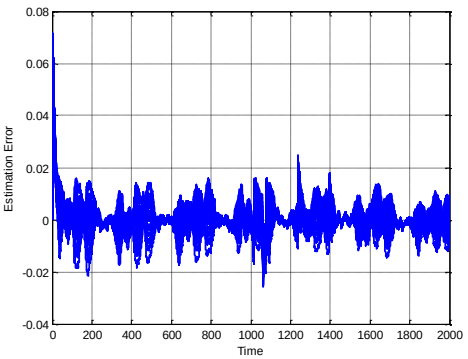


Fig. 12. Mont Carlo Simulation to investigate the effect of the parameter uncertainty on the AEKF accuracy

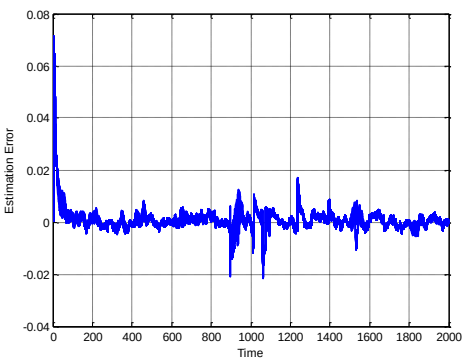


Fig. 13. Mont Carlo Simulation to investigate the effect of the parameter uncertainty on the RLS-AEKF accuracy

VI. Conclusion

This paper presented a new battery SOC estimation method in which the terminal voltage and current were measured. Due to the significance of SOC estimation in some applications, the ability of the estimation algorithm against model uncertainty is an important problem mentioned in the literature, which motivated us to find a solution to reduce the effect of uncertainty on SOC estimation. So, two algorithms were introduced to identify the deterministic uncertainty (model parameters) and stochastic parameters (noise information). To show the effect of the deterministic uncertainty, we first considered 0.02% uncertainty in the characteristics curve of the battery resistance and capacitor. In the first part of the simulation, we showed that just 0.02% uncertainty in the 5th-order polynomial causes a 10% error in

the observation of the parameters. Also, in the second part of the simulation, it was shown that the uncertainty of this parameter resulted in a notable error in SOC estimation. An online identification algorithm of RLS was introduced to identify the model's unknown parameters over time. In a parallel method, the noise information was estimated using an iterative method. Then, the overall information was given to the EKF, which was called RLS-AEKF. The results showed better accuracy of RLS-AEKF than the EKF. Finally, in the last part of the simulation, a Mont Carlo analysis was used to investigate the effect of uncertainty on the SOC estimation error bound. In case the adaptive method was used, the results of running the Mont Carlo simulation for 100 times showed the error bound of better that 0.4% in most times, i.e., about 5 times better than the non-adaptive method.

REFERENCES

[1] P.Shrivastava, P.Naidu , S.Sharma , B.K.Panigrahi , A.Garg , "Review on technological advancement of lithium-ion battery states estimation methods for electric vehicle applications", Journal of Energy Storage, vol.64, 2023.

[2] G.Mohebalizadeh, H.Alipour,L.Mohammadian, M.Sabahi, "An Improved Sliding Mode Controller for DC/DC Boost Converters Used in EV Battery Chargers with Robustness against the Input Voltage Variations", International Journal of Industrial Electronics, Control and Optimization,vol.4, no.2, pp.257-266 , 2021.

[3] C.Ge, Y.Zheng , Y.Yu, "State of charge estimation of lithium-ion battery based on improved forgetting factor recursive least squares-extended Kalman filter joint algorithm", Journal of Energy Storage, vol.55,2022.

[4] K.H.Wu , M.Seyedmahmoudian, A. tojcevski, "Lithium-ion battery state of charge estimation using improved coulomb counting method with adaptive error correction" Journal of Automobile Engineering,2023.

[5] P.Shrivastava, T.Soon , M.Y.I.B.Idris, S.Mekhilef , S. Bahari, R.S.Adnan, "Comprehensive co-estimation of lithium-ion battery state of charge, state of energy, state of power, maximum available capacity, and maximum available energy" Journal of Energy Storage,vol.6, Part B, 10 December 2022, 106049

[6] T. Bat-Orgil, B. Dugarjav, and T. Shimizu, "Battery Module Equalizer based on State of Charge Observation derived from Overall Voltage Variation," IEEJ J. Ind. Appl., vol. 9, no. 5, pp. 584–596, 2020.

[7] M. Lu, X. Zhang, J. Ji, X. Xu, and Y. Zhang, "Research progress on power battery cooling technology for electric vehicles," J. Energy Storage, vol. 27, pp. 101155, 2020.

[8] X. Xiong, S.-L. Wang, C. Fernandez, C.-M. Yu, C.-Y. Zou, and C. Jiang, "A novel practical state of charge estimation method: an adaptive improved ampere-hour method based on composite correction factor," Int. J. energy Res., vol. 44, no. 14, pp. 11385–11404, 2020.

[9] B.Zine,H.Bia,A.Benmouna,M.Becherif,and M.Iqbal, "Experimentally validated coulomb counting method for Battery State-of-Charge estimation under variable current profiles," Energies,vol.15,no.21,2022.

[10] Y.Xiong,Y.Zhu,H.Xing,S.Lin,J.Xiao,C.Zhang, "An improved state of charge estimation of lithium-ion battery based on a dual input model," Energy Sources, Part A,vol.45,no.1, 2023.

[11] X. Bian , L.Liu, J.Yan, Z.Zou, R. Zhao, "An open circuit voltage-based model for state-of-health estimation of lithium-ion batteries Model development and validation," Journal of Power Sources,vol.448,2020

[12] A.K.Birjandi, M.F. Alavi, M.Salem, M.E.H.Assad, N. Prabaharan, "Modeling carbon dioxide emission of countries in southeast of Asia by applying artificial neural network," International Journal of Low-Carbon Technologies,vol.17, pp.321–326, 2022.

[13] G.Javid, D.O.Abdeslam, M. Basset, "Adaptive online state of charge estimation of EVs Lithium-Ion batteries with deep recurrent neural networks," Energies, vol.14, 2021.

[14] F. Yang, W. Li, C. Li, and Q. Miao, "State-of-charge estimation of lithium-ion batteries based on gated recurrent neural network," *Energy*, vol. 175, pp. 66–75, 2019.

[15] M. Talha, F. Asghar, and S. H. Kim, "A neural network-based robust online SOC and SOH estimation for sealed lead--acid batteries in renewable systems," *Arab. J. Sci. Eng.*, vol. 44, no. 3, pp. 1869–1881, 2019.

[16] W. Kim, P.Y. Lee, J. Kim, K.S. Kim, "A robust state of charge estimation approach based on nonlinear battery cell model for lithium-ion batteries in electric vehicles", IEEE Trans. Veh. Technol., vol.70, no.6, pp. 5638–5647,.2021.

[17] Z.Huang, M. Best, J.Knowles; A.Fly, "Adaptive piecewise equivalent circuit model with SOC/SOH estimation based on extended Kalman filter," IEEE Transactions on Energy Conversion,vol.38,no.2,2023.

[18] T. Jarou, J. Abdouni, S. Benchikh, S. Elidrissi, and others, "The parameter update of Lithium-ion battery by the RSL algorithm for the SOC estimation under extended kalman filter (EKF-RLS)," Int. J. Eng. Appl. Phys., vol. 3, no. 2, pp. 706–719, 2023.

[19] Y. Fang, R. Xiong, and J. Wang, "Estimation of Lithium-ion battery state of charge for electric vehicles based on dual extended Kalman filter," Energy Procedia, vol. 152, pp. 574–579, 2018.

[20] R. Xiong, H. He, F. Sun, X. Liu, and Z. Liu, "Model-based state of charge and peak power capability joint estimation of lithium-ion battery in plug-in hybrid electric vehicles," J. Power Sources, vol. 229, pp. 159–169, 2013.

[21] S. Zhang, H. Tao, K. Bi, W. Yan, and H. Ni, "SOC Estimation of Lithium-ion Battery Based on RLS-EKF for Unmanned Aerial Vehicle," in Journal of Physics: Conference Series, pp. 12002.,2022.

[22] M. Li, Y. Zhang, Z. Hu, Y. Zhang, and J. Zhang, "A battery SOC estimation method based on AFFRLS-EKF," Sensors, vol. 21, no. 17, p. 5698, 2021.

[23] C. Ge, Y. Zheng, and Y. Yu, "State of charge estimation of lithium-ion battery based on improved forgetting factor recursive least squares-extended Kalman filter joint algorithm," J. Energy Storage, vol. 55, p. 105474, 2022.

[24] C. Zhang, X. Li, W. Chen, G. G. Yin, J. Jiang, and others, "Robust and adaptive estimation of state of charge for lithium-ion batteries," IEEE Trans. Ind. Electron., vol. 62, no. 8, pp. 4948–4957, 2015.

[25] D. Sun et al., "State of charge estimation for lithium-ion battery based on an Intelligent Adaptive Extended Kalman

Filter with improved noise estimator," Energy, vol. 214, p. 119025, 2021.
[26] T. Jarou, J. Abdouni, S. Benchikh, S. Elidrissi, and others, "The parameter update of Lithium-ion battery by the RSL algorithm for the SOC estimation under extended kalman filter (EKF-RLS)," Int. J. Eng. Appl. Phys., vol. 3, no. 2, pp. 706–719, 2023.
[27] M. Li, Y. Zhang, Z. Hu, Y. Zhang, and J. Zhang, "A battery SOC estimation method based on AFFRLS-EKF," Sensors, vol. 21, no. 17, pp. 5698, 2021.

Ramazan Havangi received his M.S. and Ph.D. degrees from the K.N. Toosi University of Technology, Tehran, Iran, in 2003 and 2012, respectively. He is currently an Associate Professor of control systems with the Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran. His main research interests are inertial navigation, integrated navigation, estimation and filtering, evolutionary filtering, simultaneous localization and mapping, fuzzy, neural network, and soft computing.

Fatemeh Karimi was born in Borujerd, Iran. She received her associate degree in telecommunications, her B.Sc. degree in electronics engineering, and her M.Sc. degree in control engineering all from Islamic Azad University in 2003, 2008, and 2013, respectively. She is currently a Ph.D. candidate of control engineering at the state-run University of Birjand. Her research interests include function fields, active power filters, multi-agents, and Kalman filters.

IECO

This page intentionally left blank.

International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

The Design and Simulation of an Efficient Quaternary Full-Adder Based on Carbon Nanotube Field Effect Transistor

Farzaneh Yousefzadeh Ahari[1] | Mousa Yousefi[2] | Khalil Monfaredi[3]

Department of Electrical Engineering, Faculty of Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran.[1,2,3]
Corresponding author's email: m.yousefi@azaruniv.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 16-December-2023
Received in revised form: 21-March-2024
Accepted: 30-March-2024
Published online: 21-June-2024

Keywords:
Multilevel Processing System, Carbon Nanotube Field Effect Transistor (CNTFET), Multiple-Valued Logic, Quaternary Full-adder, Low Power Consumption.

ABSTRACT

An essential reason for implementing multilevel processing systems is to reduce the number of semiconductor elements and hence the complexity of the system. Multilevel processing systems are realized much easier by carbon nanotube field effect transistors (CNTFET) than by MOSFET transistors due to the CNTFET transistors' adjustable threshold voltage capabilities. This paper presents an efficient quaternary full-adder based on CNTFET technology that consists of two half-adder blocks, a quaternary decoder, and a carry generator circuit. The proposed architecture combines the base-two and base-four circuit design techniques to take full advantage of both techniques, namely, simple implementation and low chip area occupation of the entire proposed quaternary full-adder. The proposed structure is evaluated using a Stanford 32nm CNTFET library in HSPICE software. The simulation results for the proposed full-adder structure that utilizes a 0.9-V supply voltage reveal that the power consumption, propagation delay, and energy index are equal to 2.67 μW, 40 ps, and 10.68 aJ, respectively.

NOMENCLATURE

$n_1.n_2$	Chiral vector
D_{cnt}	The diameter Carbon Nanotube (CNT)
E_π	π–π junction energy of carbon
V_{fb}	The flat band voltage
V_{dsi}	The drain to source voltage
δ	The Drain Induced-Barrier Lowering (DIBIL) coefficient in CNT

I. Introduction

While circuits and systems were primarily designed to operate by turning the vacuum tubes on and off, the idea of silicon transistor-based design was quickly considered by the electronics industry after its invention. Silicon (Si), which is an element with semiconductor properties that is found in abundance in nature, has become the main material in making transistors that are smaller and cheaper tools than previously used vacuum tubes, and it was possible to replace vacuum lamps in a short time [1].

Recently in an interesting attempt to follow the path of previous progress, some researchers have suggested replacing Si transistors with carbon nanotubes due to their better performance in realizing logic circuits [2-5]. This is due to the fact that the inevitable reduction in MOSFET transistor channel size and hence the appearance of the nanotechnology properties have posed major problems and challenges to implementing CMOS technology [6]. Issues, such as high leakage current, short channel length and very large channel length modulation, high power consumption, and high sensitivity to process variations have disrupted the dimensional process and degraded the proper performance of CMOS technology [7-9].

In binary logic, the power supply voltage of the circuit is divided into two levels of logic, i.e. 0 and 1. Making multi-level processing circuits reduces the consumption level and the connections between the blocks, but transistors are needed to make these circuits. It should be considered that the construction

of multi-level circuits with CMIS transistors is complicated.

In recent years, in order to provide high-speed processing systems, many researchers have focused on researching optical processing systems, including optical-based processing systems, including full-adders [10-13], multiplexers [14], encoders [15], and decoders [16]. Also, new technologies introduced in recent years, such as carbon nanotube transistors, have facilitated the implementation of this type of circuit and solved some of their problems, causing more and more designers to pay attention to this multilevel logic. One of the most important strengths of carbon nanotube transistors is that their threshold voltage can be adjusted by changing the diameter of the nanotubes, allowing Multi-Valued Logic (MVL) to be designed with less complexity and higher simplicity [2, 17-21]. Since the operation of CNTFET and MOSFET transistors are very similar, it is possible to design and implement simpler multilevel processing circuits by modeling the circuit design process using CMOS transistors. Quantum-dot Cellular Automata (QCA) is a new technology to eliminate some of the problems of existing technologies, such as CMOS. This study was concerned with designing an efficient full-adder circuit by using QCA technology [22].

With these explanations, MVL is a suitable alternative to binary processing systems [18, 24], which can solve and optimize the complexity of connections expressed by binary systems in VLSI circuits [9, 24]. Multilevel circuits can be used to solve binary problems [6].

The most common and efficient way to design MVL circuits is to use the known multi-threshold voltage CMOS technology experiences [11]. Due to the similarity of the MOSFET and CNTFET transistors in terms of their electrical properties, many CMOS structures that have already been designed can be implemented by CNTFET technology without any significant modifications [19]. In general, CNTFET transistors are faster and less energy-consuming than MOSFETs. In addition, in recent years, some multilevel circuits based on CNTFET transistors have been proposed that take advantage of the unique properties of carbon nanotube transistors [18, 13, 26-29]. A 2-trit ternary ALU using CNTFETs and RRAM as the basic design elements has been presented in [30]. The proposed ternary ALU modules are implemented, taking advantage of the variable multi-threshold design method of CNTFET and multilevel cell characteristics of Resistive Random Access Memory (RRAM) [30]. A design of quaternary multiplexer 4:1 with CNFETs has been proposed in which quaternary successor, predecessor, and second-level successor cells are introduced based on CNTFETs with single and three-supply voltage. All the above-mentioned designs have been applied to quaternary half-adder and full-adder circuits [31]. A one-bit full adder has been reported by Ebrahimy [32] designed by the proposed majority gates, which has only 0.75 clock cycle latency and suitable arrangements of inputs and outputs that give them the multi-bit extensibility feature.

A new memory element has reportedly been designed by magnetic tunnel junction spintronic device [33-34]. Magnetic tunnel junctions have special characteristics, such as high endurance, low leakage current, and ease of 3-D integration with CMOS and CNTFET technologies [35].

A fabrication process of CNTFET and magnetic tunnel junctions did not interfere with the fabrication of CMOS. Manufacturing of CNTFET has been reported in [36], which shows the compatibility of CNTFET with CMOS. As a result, it is possible to fabricate CNTFET, magnetic tunnel junction, and CMOS devices in a single chip [37-38].

In recent years, several modern analog and digital circuits have been fabricated based on CNTFET. Also, hybrid circuits in which CNTFET and Magnetic Tunnel Junction (MTJ) are used have been presented [39-40].

This paper proposes quaternary half-adder and full-adder structures realized by converting signals from quaternary to binary. Fewer transistors are used in the proposed architectures applying innovative circuit techniques, resulting in lower power consumption satisfactorily.

The paper is organized as follows. Section 2 explains the physical structure of CNTFET. Section 3 presents the conceptual circuitry and also the CNTFET transistor level realization of two different three-level adders based on CNTFET technology. A unique feature of adjusting the desired threshold voltage using suitable diameters and considering the mobility coefficient has been adopted [3]. In this section, the CNTFET-based quaternary full-adder and half-adder structures are explained in detail. The simulation results are provided in Section 5 and Section 6 concludes the paper.

II. A brief description of CNTFET

Carbon nanotubes are composed of graphite sheets tubed inside concentric cylinders with nanometer diameter and micrometer length. Carbon nanotubes are divided into single-walled (SWNT) and multi-walled (MWNT) based on their number of layers. Single-walled carbon nanotubes are the result of the complexity of a single layer of graphite, while multi-walled carbon nanotubes are the result of the complexity of several layers of nested graphite. The properties of nanotubes depend on their structure and act as metal or semiconductors depending on the chirality. The second characteristic of nanotubes that is affected by their electrical properties is the number of walls. The main difference between single-walled and multi-walled is the diameter of the nanotubes [22].

The arrangement of carbon atoms along the tube defines the chiral vector and is denoted by a correct pair (n1,n2). Based on the chiral vector, CNT can be conductive or semiconductor. If n1=n2 or n1-n2=3i, the nanotube is metal; otherwise, it is a semiconductor [41].

The diameter of CNT has a key role in the electrical characteristics of transistors. The equation of the diameter CNT is:

$$D_{CNT} = \frac{a\sqrt{n_1^2 + n_2^2 + n_1 n_2}}{\pi} \quad (1)$$

where a0=0.142 nm, π=3.14, and n1,n2 are chirality vectors [22].

In the structure of CNTFET components, one or more non-doped semiconductor SWCNTs are used as semiconductor channel material instead of Si bulk in the MOSFET structure [22].

The use of single-walled semiconductor carbon nanotubes as a substitute for Si channel field-effect transistors has major advantages, such as lower electron scattering, higher thermal conductivity, very high electrical conductivity, and high tensile strength.

Therefore, carbon nanotube circuits have much less latency than Si circuits. Also, a CNT transistor, like a MOSFET, has a

threshold voltage that is actually needed to turn on the transistor through the electrostatic gate. A very important advantage of CNTFET is that its threshold voltage can be adjusted by changing the diameter of the carbon nanotubes. In the CNTFET, the mobility of holes and electrons is the same.

This practical feature makes CNTFETs much more flexible for designing digital circuits than MOSFETs. It also makes the technology very suitable for designing multi-voltage threshold circuits.

The threshold voltage of a CNTFET can be calculated by Eq. 2:

$$V_{th} = \frac{2a_0E_\pi}{D_{CNT}} + V_{fb} - \delta \cdot V_{dsi} \quad (2)$$

in which a is the atomic distance of carbon to carbon, Eπ is π–π junction energy of carbon (Eπ =3.3eV), DCNT is the diameter of the carbon nanotube, Vfb is the flat band voltage, Vdsi is the drain to source voltage, and δ is the drain induced barrier lowering (DIBIL) coefficient in CNT [39-40].

This diameter can be calculated by Eq. 3 in which the CNTFET threshold voltage is an inverse function of the CNT diameter:

$$D_{CNT} = \frac{a\sqrt{n_1^2 + n_2^2 + n_1n_2}}{\pi} = 0.0783\sqrt{n_1^2 + n_2^2 + n_1n_2} \quad (3)$$

A carbon nanotube transistor looks like a MOSFET with the same pins. In other words, the principles of operation of carbon nanotube field effect transistors are similar to current Si transistors [22].

Fig. 1 shows the physical appearance of a carbon nanotube transistor. The nanotube remains not doped in the area below the gate.

The two ends of the nanotube that are connected to the source and drain connections are doped areas. The voltage applied to the gate of the transistor can control the electrical conductivity of the carbon nanotube in the area below the gate by changing the electron density in the channel.

The distance between the cross-sectional centers of two adjacent carbon nanotubes below the transistor gate is called a "pitch", which significantly affects the width of the gate and the drain and source of the transistor. Interestingly, the PCNTFET and NCNTFET have the same current rate.

The gate width of the transistor can be estimated based on Eq. 4, in which N is the number of nanotubes under the gate, DCNT is the diameter of the nanotube, and Wmin is the minimum possible width for the gate, which is determined based on lithographic constraints.

$$W_{gate} \approx Max(W_{min}(N-1)pitch + D_{CNT}) \quad (4)$$

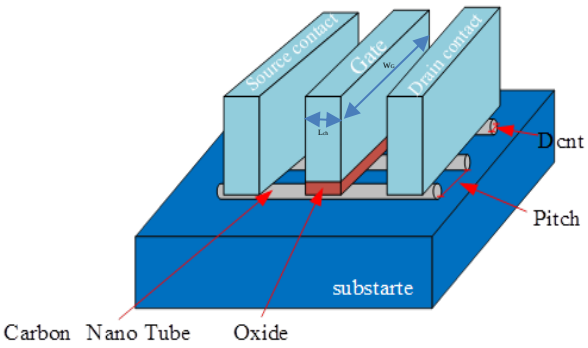


Fig. 1. The structure of CNTFET [23].

III. The Conceptual Scheme for quaternary Full-dder

The quaternary full-adder circuit, like the conventional full-adders, has three inputs, namely A, B, and Cin, and two outputs, namely, Sum and Cout. Quaternary digits A, B, Cin, and Sum have the same quaternary weight while Cout has a higher quaternary weight which must be used by the next addition block. In other words, the Carry input, Cin, of full-adders is supplied from the lower stage, while Cout, which indicates the carry-out, is transmitted to the upper stage, and Sum output is the sum of inputs A and B along with Cin. Unlike the conventional full-adders, which include two half-adders and a carry generator, the proposed quaternary full-adders have some complex design procedures and include half-adders, Q-Decoders, and carry generators. Two quaternary full-adders are proposed, and their half-adders are exhibited by Q- Decoders along with "Sum and Carry" and "Sum and Carry_bar" in Fig. 2 and Fig. 3, respectively [29]. The first quaternary full-adder shown in Fig. 2 that uses "Sum and Carry" in its half-adder is the initial not optimized version, while the second quaternary full-adder uses "Sum and Carry_bar" to reach an optimized design approach. The structure of a quaternary half-adder has two inputs and two outputs. Each input can be 0, 1, 2, or 3. In fact, input 0 means zero voltage, input 1 means VDD/3 voltage level, input 2 means 2VDD/3, and input 3 means VDD voltage level. In the following, we explain the architecture of the whole full-adder circuitry at the transistor level and block diagram.

A. The proposed full-adder architecture

The initial idea to realize a quaternary block is shown in Fig. 2.

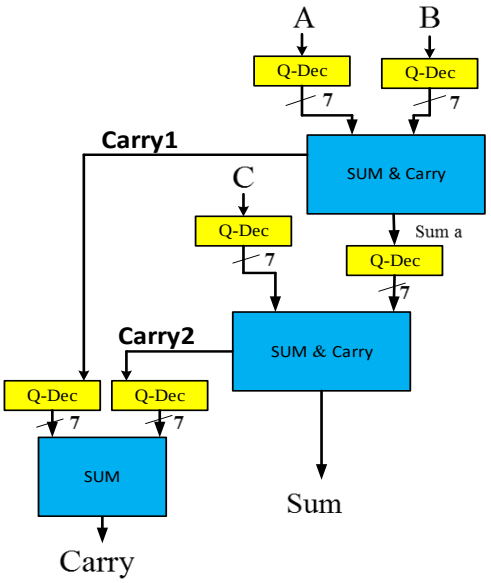


Fig. 2. The initial idea of the quaternary full-adder

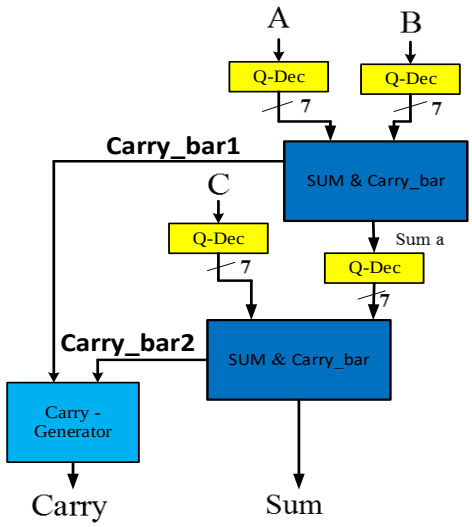


Fig. 3. The proposed optimized quaternary full-adder

To implement the full-adder, two inputs A and B can be added using one half-adder (depicted as SUM & Carry along with its Q-Decoders), and to add the third Cin input with the Sum output of the first half-adder, another half-adder can be used. Carry 1 and 2 related to both half-adders, as shown in Fig. 2, have the same digit weight and are applied to another half-adder to create the final carry digit of the full-adder. There is no need for a carry output for this half-adder because the sum of the carries will eventually be 2.

As previously noted, to implement a quaternary half-adder, a decoder gate, and a sum and carry generator circuit are used as shown in Fig. 2. The quaternary decoder takes one input and generates seven outputs, which will later be used to create a sum. The overall circuit design process is based on using the control signals provided by the decoder circuit output to control the keys that are actually needed to generate the sum and carry values. The basis of this design is adopted from the design procedure of the logic circuits using the CMOS transmission gate. In this regard, the sum and carry output look-up tables are considered to create the desired output values in the transmission paths.

The detailed explanations are provided in Subsections 2.3 and 2.4. The essential bottleneck with this type of design is that when 1 or 2 is needed at the output, the desired path must be designed from node VDD to the ground to supply the required voltage.

Fig. 3 depicts the proposed optimized quaternary full-adder, which has fewer devices.

Referring to Table 1, considering the input numbers in all states, it is clear that the half-adder's carry output does not interestingly have the values of 2 and 3 in any state. So, to reduce the design complexity, we create output 3 instead of generating 1 for caring output. This is due to the fact that the circuit needed to create output 3 is far simpler than the circuit needed to create outputs 1 or 2.

On the other hand, to generate the quaternary full-adder carry-out, the carry generator circuit uses two inputs, which have states 0 and 3. This approach for designing the circuit simplifies both the carry generator circuit of the half-adders and also the final carry generator circuit. With this proposal, two Q-decoders are eliminated, and the carry generator circuit is significantly simplified, so the number of transistors and the overall propagation delay of the full-adder will be reduced. In the following sections, the principle of the operation of the half-adder parts, namely the Q-decoder and sum and carry_bar, are explained, and their CNTFET level circuitry is provided.

B. Quaternary decoder structure

The four-level Quaternary Decoder (QDEC) decoder is shown in Fig. 4-a. It has one input and seven outputs with the corresponding input and output values shown in Table 1 [21]. After determining the logical input value, the appropriate transistor is selected based on Table 1, and the correct path is transmitted to the output. In Table 1, I is the input signal, I_x, and $\bar{I}_x$ are the decoder outputs where x ∈ (1, 2, 3, i). The required chiral vector and control signal connected to the gate of the transistors are given in Table 2.

The control signals which are created by a decoder, are applied to the transistors' gates.

TABLE 1 QDEC OUTPUT VALUE FOR DIFFERENT INPUTS

I	I_0	I_1	$\bar{I}_1$	I_i	$\bar{I}_2$	I_2	I_3	Number of transistors
0	3	0	3	3	3	0	3	18
1	0	3	0	3	3	0	0	
2	0	0	3	0	0	3	0	
3	0	0	3	0	3	0	0	

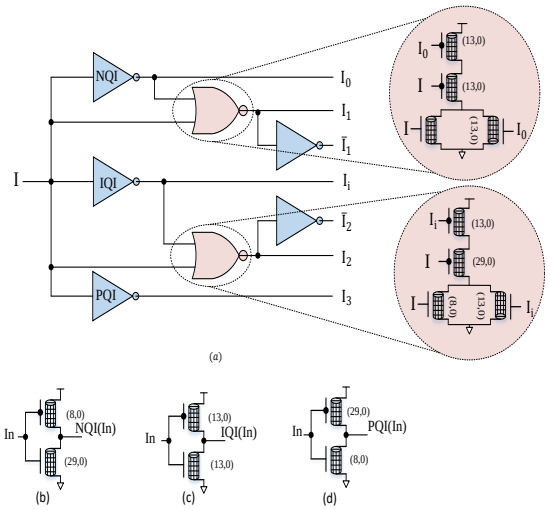


Fig. 4. a) QDEC - Quaternary Inverter; b) NQI c) IQI d) PQI [28]

Fig. 4 shows three types of four-level inverters called PQI (positive four-level standard), IQI (intermediate four-level standard), and NQI (negative four-level standard), which are described using Eq. 5.

$$NQI = \begin{cases} 3 & if \quad in = 0 \\ 0 & if \quad in \neq 0 \end{cases}$$

$$IQI = \begin{cases} 3 & if \quad in = 0 \; or \; 1 \\ 0 & if \quad in = 2 \; or \; 3 \end{cases} \qquad (5)$$

$$PQI = \begin{cases} 3 & if \quad in \neq 3 \\ 0 & if \quad in = 3 \end{cases}$$

TABLE 2 THE LOOK-UP TABLE OF FOUR-LEVEL INVERTERS

IN	Out		
	NQI	IQI	PQI
0	3	3	3
1	0	3	3
2	0	0	3
3	0	0	0

C. *Half-adder's sum and inverse carry output*

First of all, let's briefly explain the Sum formula. According to Eq. 5, the four states that the sum output is 3, are when A=3, B=0, A=1, B =2, A=2, B=1 and A=0, B=3.

Note that when A=i (i = 0, 1, 2, 3), Ai is activated. In other words, it is equal to VDD, and all other signals are zero. For example, when A=0, the output A0 is active and A1, A2, and A3 are zero. With these explanations, the Sum output is equal to 2 when (A=0, B=2), (A=1, B=1), (A=2, B=0), and (A=3, B=3), so in Eq. 5, the second sentence must be written as $A_0B_2 + A_1B_1 + A_2B_0 + A_3B_3$ to create output 2. With these conditions, the third sentence also creates the Sum output of 1, and the first sentence creates the Sum output of 3. Note that in any circumstance, only one of the AiBi components is true (or equal to one).

$$\begin{aligned} \text{Sum} = {} & 3(A_0B_3 + A_1B_2 + A_2B_1 + A_3B_0) \\ & + 2(A_0B_2 + A_1B_1 + A_2B_0 \\ & + A_3B_3) \\ & + 1(A_0B_1 + A_1B_0 + A_2B_3 \\ & + A_3B_2) \end{aligned} \qquad (6)$$

Since parallel transistors operate as OR and series transistors operate as AND, the Sum equation can be realized by transistors connected in series and parallels. As shown in Fig. 5, the half-adder sum output is controlled using QDEC outputs. Using Eq. 6, the Look Up Table of Sum output (Table 3) shows the values.

Table 3 shows the output SUM. For example, if the inputs (A and B) are equal to 2, that means that the output of the half-adder must be 6. If we show the number 6 in base three, the number will be 20, which means that the output SUM must be 2. The rest of the values in Table 3 can be obtained.

TABLE 3 LOOK-UP TABLE OF THE SUM OUTPUT OF HALF-ADDER

	B=0	B=1	B=2	B=3
A=0	0	1	2	3
A=1	1	2	3	0
A=2	2	3	0	1
A=3	3	0	1	2

Now, to fully describe the operation, the circuit of Fig. 5 can be used. Four-level logic is utilized with the voltages equal to 0, VDD/3, 2/3VDD, and VDD volt, which are the same as 0, 1, 2, or 3. Supposing that VDD equals 1 volt, the circuit equivalent to Fig. 6 can be used to create voltages of 0.3 and 0.6 volts. In the case of the output equal to 0.3 volts, keys K1 and K4 should be in the ON state, and keys K3 and K2 should be in the Off state. For the output to be equal to 0.6 volts, keys K2 and K3 should be the ON state, and keys K4 and K1 should be disconnected.

In order to design the sum circuit, the output should be logic 1 or 0.3 volts for (0,1), (1,0), (2,3), and (3,2) inputs. For example, in the case of half-adder inputs (2,3), the outputs (Sum, Carry) are 5 (Decimal); in this case, the value is Sum = '1' and the value Carry = '1'. In the circuit of Fig. 6, if K1 and K4 switches are turned on and K2 and K3 switches are turned off, the output of 0.3 volts or logic 1 is created. With this description, the output should be logic 2 or 0.6 volts for (0,2), (2,0), (1,1), and (3,3) inputs. In this situation, K3 and K1 switches are turned on, and K4 and K2 switches are turned off [30].

In the same way as the description of the sum 6, the inverse carry output (Cbar) can also be written as 6.

Note that the carry inverse is implemented to simplify the circuit design. In the reverse implementation of the carry (carry bar implementation), when the carry output is zero, we consider Cbar = 3 (or VDD), and when the carry output is one, we consider Cbar = 0.

According to Table 5, remember that in cases A=3, B=3 and A=3, B=1 and A=1, B=3 and A=2, B=2, the output of the carry output will be one; otherwise, it will be zero. With these explanations, the logical relation of the Cbar can be written as follows:

$$\text{Cbar} = 3\left(\left(\overline{A_1 B_1}\right) + (A_1 B_2) + (A_2 B_1) + \text{A} + \text{B}\right) \quad (7)$$

In this circuit, the NMOS transistor is used to transfer values 0 and 1, and the PMOS transistor is used for values 2 and 3. The value of their chiral vector is determined using the threshold voltage of each transistor. If we want to calculate the value of the chiral vector for these transistors using Eq. 2 and 3, the value is obtained as n = 7.

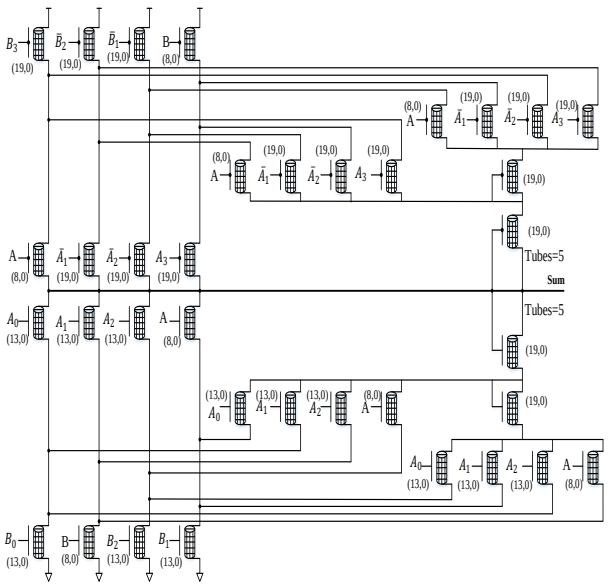


Fig. 5. The quaternary full-adder circuit's sum output realized at CNTFET transistor level [22]

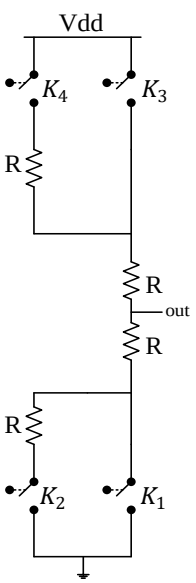


Fig. 6. The sum output equivalent circuit

The communication-aided adaptive protection methods [8, 9] and modified time-current curves for DOCRs [10] are some of the reported solutions to modify the MG protection system against changes in MG protection due to changes in system topologies. Since the adaptive scheme requires the hardware and software infrastructures to activate the suitable setting groups [11], a great deal of attention has also been paid to local measurement-based protection systems [12].

TABLE 4 THE LOOK-UP TABLE OF QUATERNARY SUM, FINAL CARRY, AND CBAR OF THE PROPOSED HALF-ADDER

Input values	Decimal Sum	Quaternary Sum	Carry_out	Cbar
0	0	0	0	3
(0,1) - (1,0)	1	1	0	3
(0,2) - (2,0)- (1-1)	2	2	0	3
(1,2) - (0,3) - (3,0)	3	3	0	3
(2,2) - (1,3) - (3,1)	4	0	1	0
(2,3) - (3,2)	5	1	1	0
(3,3)	6	2	1	0

The values in Table 4 are from 0 to 6 because, based on the quaternary logic's four-level values (i.e., 0, 1, 2, and 3), the maximum sum in the half-adder is 6. This is equal to 12 in the base four, in which digit 1 is carried, and digit 2 is the sum. Fig. 7 shows the transistor-level implementation of the Cbar.

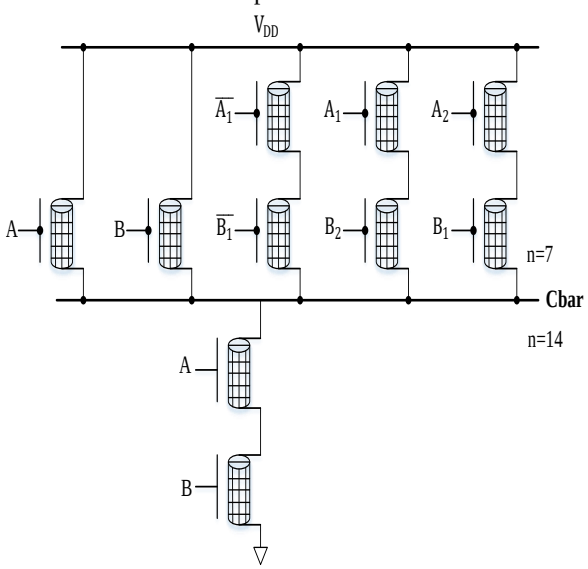


Fig. 7. Carry bar output implementation of the proposed half-adder

D. Final carry generator

Fig. 8 shows the transistor level of the full-adder final carry generator circuit. The carry output circuit consists of two inputs and one output. The inputs of this circuit are given using the reverse outputs of the carry output. In this circuit, the amplitude of the input signals will be zero (0) or VDD (3). Note that each input can be 0 or 3. In fact, the actual value of 3 is equivalent to

1 in the base of 4. Regarding the operation of this circuit, it can be said that this circuit calculates the sum of two inputs A and B.

According to Table 6, four states can occur for inputs A and B, and the carry creator circuit must be designed to operate according to Table 6. In order to realize the circuit, three different values 0, 1, and 2 must be generated at the output. If both inputs are 3, the output must be 0. For this purpose, two N-type series transistors are used. Note that zero-volt transmission is easily done with an N-type transistor. Two P-type series transistors are used to generate a value of 2 at the output when both inputs are zero. To transmit the value of 1 to the output when one of the inputs is 3 and the other is zero, two N-type series transistors and an inverter are used to invert the input with a value of zero. In fact, this inverter works like a binary inverter because inputs A and B have no more than two modes (0 and 3).

TABLE 5 THE LOOK-UP TABLE OF FINAL CARRY GENERATOR

inputs		output
B	A	Carry
0	0	2
0	3	1
3	0	1
3	3	0

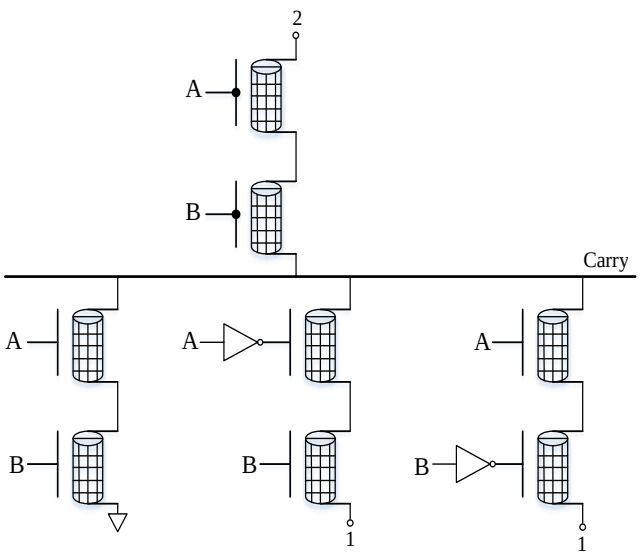


Fig. 8. The Carry generator transistor level circuit (n1=7, n2=0 for all of Transistor)

IV. Simulation results

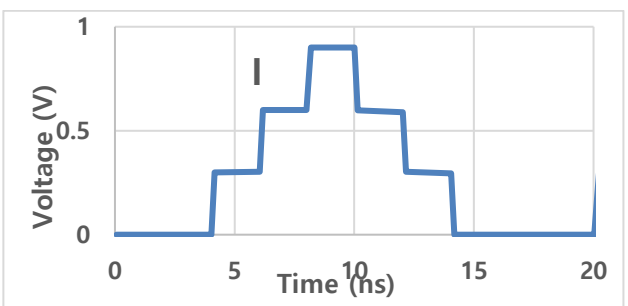
The proposed design was simulated using the Stanford 32 nm CNTFET library in HSPICE software [42-43]. The parameters of the CNTFET transistor and its values are shown in Table 6.

TABLE 6 THE PARAMETERS OF THE CNTFET TRANSISTOR

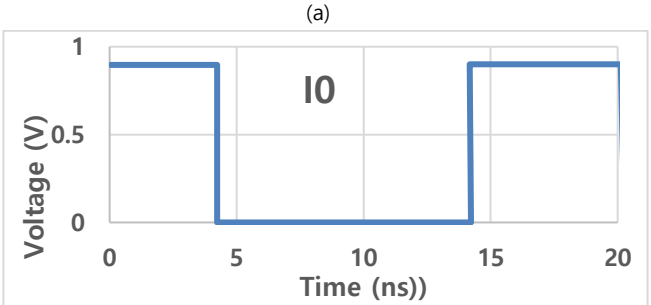
Parameters	Description	Value
$\mathbf{L_{ch}}$	Physical channel length	32 nm
$\mathbf{L_{ss}}$	The length of the doped CNT source extension region	32 nm
$\mathbf{L_{dd}}$	The length of the doped CNT drain extension region	32 nm
Lgeff	The mean-free-path in the intrinsic CNT channel	100 nm
Sub-pitch	The distance between the centers of two adjacent CNTs	4 nm
Kox	The dielectric constant of high-k top gate dielectric material (HfO_2)	16
Tox	The thickness of high-k top (planer) gate dielectric material (HfO_2	4 nm
$\mathbf{K_{sub}}$	The dielectric constant of the substrate	4
$\mathbf{C_{sub}}$	The Coupling capacitor between the channel area and the substrate	40 af/ µm
Efi	The fermi level of the doped source/drain tube	60 eV
Phi-M	The CNT work function	4.6 eV

The QDEC block is simulated by Hspice software. The timing diagram of the input and output signals of this block is shown in Fig. 9.

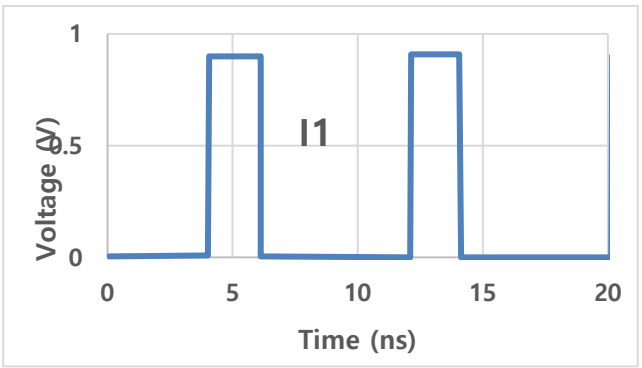
The proposed half-adder block consists of two circuits for creating Cbar and Sum, and the timing diagram of the input and output signals of this block is shown in Fig. 10. The simulation results show that this circuit consumes 1.3679 microwatts of power for proper operation, while the delay in the signal path is equal to 10 picoseconds. These results show that the PDP (Power Delay Product) index of this block is equal to 0.136 fJ.



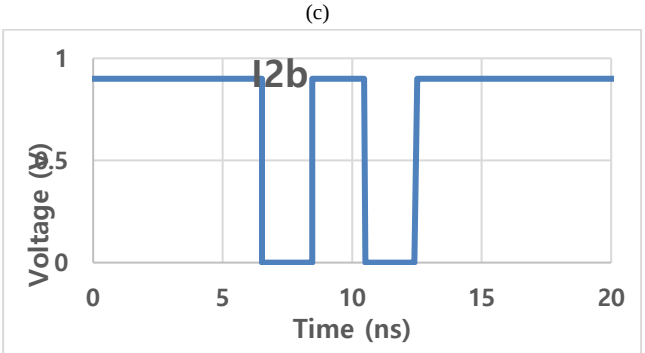
(a)



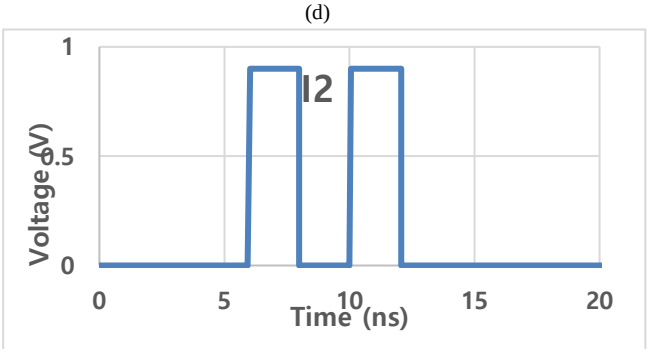
(b)



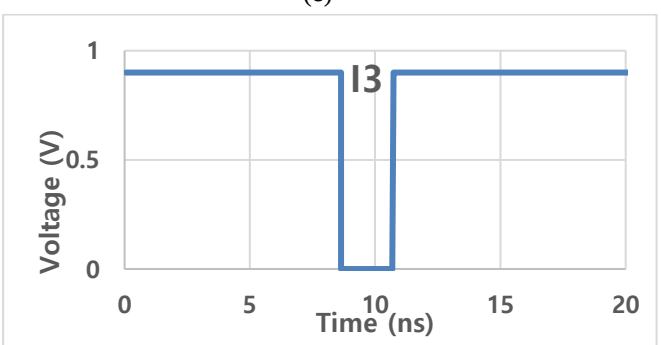
(c)



(d)

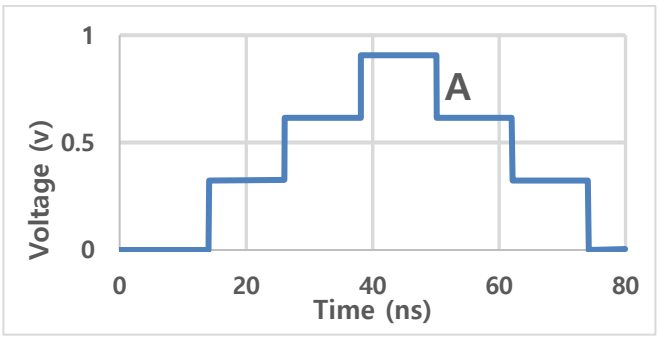


(e)

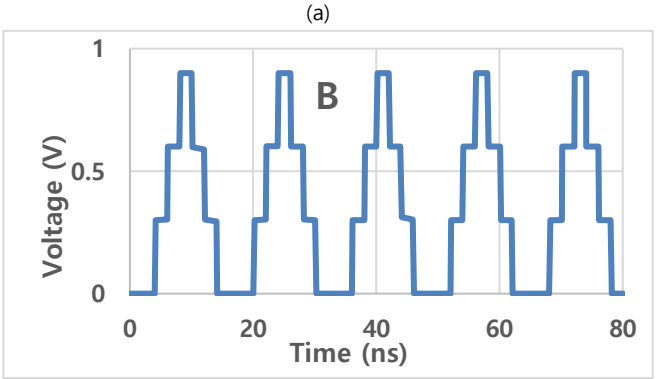


(f)

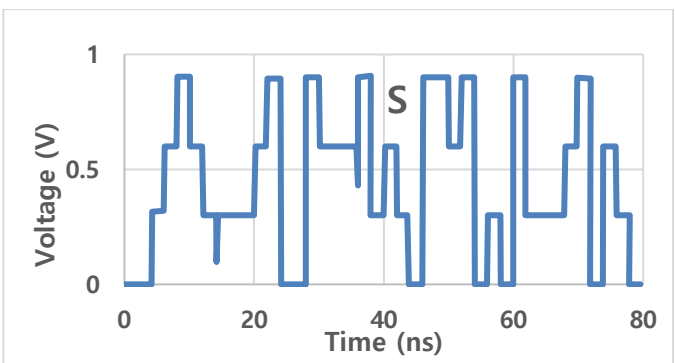
Fig. 9. The simulation result of QDEC- (a): I, (b): I0, (c): I1, (d): I2b, (e): I2, (f): I3



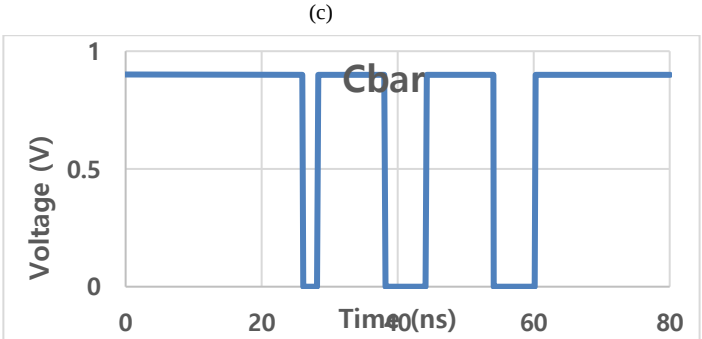
(a)



(b)



(c)



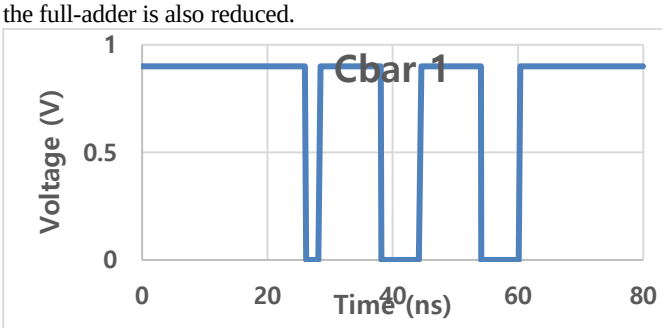
(d)

Fig. 10. The simulation results of the proposed half-adder-(a): A input, B Input, (c): Sum Output, (d): Car output.

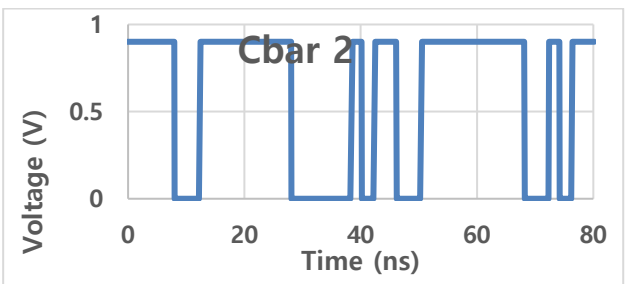
The time diagram of the input signals and the output signals of the carry generator block are shown in Fig. 11. The results show that the circuit operates correctly for all states. In Fig. 11, the Cbar1 and Cbar2 signals are the carry signals of the proposed half-adders.

The final performance of the proposed full-adder is correct in all possible states. The PDP index of the proposed full-adder is equal to 106 fJ.

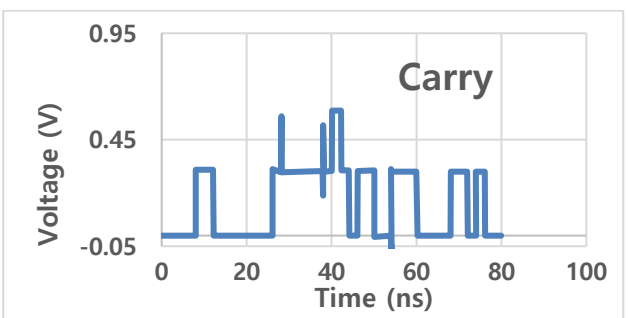
Table 7 shows the performance results of the proposed blocks in comparison with other works. It, indeed, compares the proposed four-level circuits and some three-level reports. There are several important points for a proper comparison of the design reports of the proposed circuits. In the processing circuits, the frequency and load capacitor of the circuits must be the same. Important parameters, including power consumption, propagation delay, and PDP coefficient, are reported in Table 7. Another important indicator of processing circuits is space consumption. For this type of comparison, valid outputs must be provided, and on the other hand, many conditions, including the load capacitor, must be considered. Finally, the results show that the proposed block is in better condition in terms of the energy index, while the space required to implement the final block of the full-adder is also reduced.



(a)

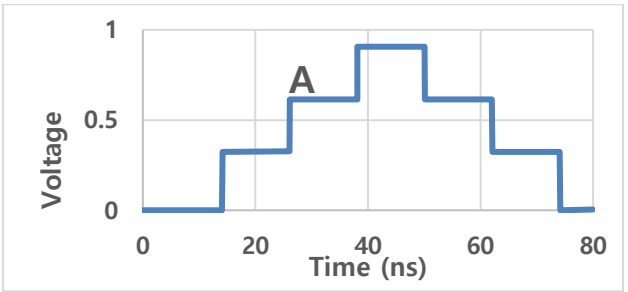


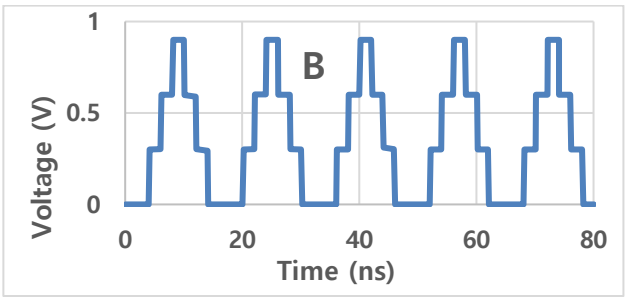
(b)

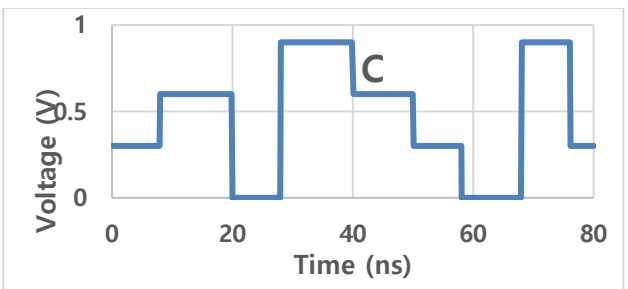


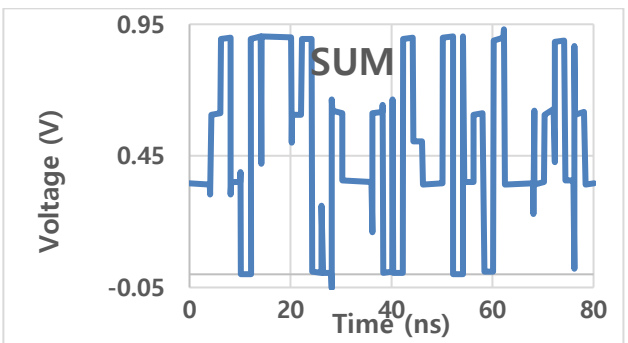
(c)

Fig. 11. The simulation results of the final carry generator-(a): Cbar1 output, (b): Cbar2 output, (c): Carry output.

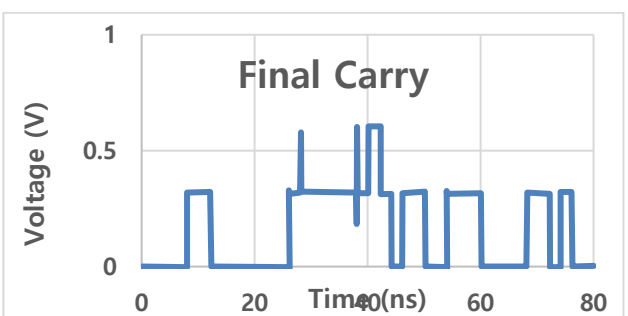








(a)



(b)

Fig. 12. The timing diagram of the proposed full-adder-(a): sum output (b): carry output.

V. Conclusion

Compared to binary circuits, MVL circuits have advantages and disadvantages. Binary processing circuits have a higher noise margin than MVL circuits. The reduction of connections between blocks and higher data density of multi-level circuits are the most important advantages of these circuits compared to processing circuits. In implementing quaternary circuits, it is complicated to produce levels 1 and 2 using transistors. In this article, the output of the Carry-out digit had two logical levels, and the approach adopted to implement the circuits was binary, which reduced the number of transistors and power consumption. The paper presented a quaternary full-adder structure consisting of two half-adders and a carry generator using a new structure. Using the techniques used to reduce the power and chip area in the design of this structure, the power and chip area of the proposed structure was finally optimized. To implement this structure, the four-level to binary conversion technique was used to implement the carry output circuit. The simulation results showed that the PDP of all the proposed full-adder was 10.68 aJ.

REFERENCES

[1] N. Yanamala, V. E. Kagan, and A. A. Shvedova, “Molecular modeling in structural nano-toxicology: Interactions of nano-particles with nano-machinery of cells," Advanced drug delivery reviews, Vol. 65: pp.2070-2077, 2013.

[2] S. Fakhari, N. H. Bastani, and M. H. Moaiyeri, “A low-power and area-efficient quaternary adder based on CNTFET switching logic,” Analog Integrated Circuits and Signal Processing, Vol. 98, pp.221-232, 2019.

[3] S. Lin, Y.-B. Kim, and F. Lombardi, “A novel CNTFET-based ternary logic gate design,” in 2009 52nd IEEE International Midwest Symposium on Circuits and Systems, 435-438, 2009.

[4] M. H. Moaiyeri, R. F. Mirzaee, A. Doostaregan, K. Navi, and O. Hashemipour, “A universal method for designing low-power carbon nanotube FET-based multiple-valued logic circuits, “IET Computers & Digital Techniques, Vol. 7, pp.167-181, 2013.

[5] S. Wind, J. Appenzeller, and P. Avouris, “Lateral scaling in carbon-nanotube field-effect transistors,” Physical Review Letters, vol. 91, pp.058301, 2003.

[6] E. Dubrova, “Multiple-valued logic in VLSI: challenges and opportunities," in Proceedings of NORCHIP, pp.340-350, 1999.

[7] M. H. Moaiyeri, R. F. Mirzaee, K. Navi, and O. Hashemipour, “Efficient CNTFET-based ternary full-adder cells for nanoelectronics," Nano-Micro Letters, Vol. 3: pp.43-50, 2011.

[8] A. D. Zarandi, M. R. Reshadinezhad, and A. Rubio, “A Systematic Method to Design Efficient Ternary High Performance CNTFET-Based Logic Cells,” IEEE Access, Vol. 8: pp.58585-58593, 2020.

[9] M. Shahangian, S. A. Hosseini, and R. F. Mirzaee, “A Universal Method for Designing Multi-Digit Ternary to Binary Converter Using CNTFET,” Journal of Circuits, Systems and Computers, p. 2050196, 2020.

[10] M. J. Maleki, A. Mir, and M. Soroosh, “Designing an ultra-fast all-optical full-adder based on nonlinear photonic crystal cavities,” Optical and Quantum Electronics, vol. 52, pp.1-11, 2020.

[11] F. Pakrai, M. Soroosh, M. and J. Ganji, “A novel proposal for an all optical combined half-adder/subtractor,” Optical and Quantum Electronics, Vol.54, p.518, 2022.

[12] M. J. Maleki, A. Mir, and M. Soroosh, “Design and analysis of a new compact all-optical full-adder based on photonic crystals," Optik, Vol. 227. p.166107, 2021.

[13] M. J. Maleki, A. Mir, and M. Soroosh, “Ultra-fast all-optical full-adder based on nonlinear photonic crystal resonant cavities,” Photonic Network Communications, Vol. 41, pp. 93-101, 2021.

[14] A. Rostami, A. Haddadpour, F. Nazari, and H. Alipour, "Proposal for an ultracompact tunable wavelength-division-multiplexing optical filter based on quasi-2D photonic crystals," Journal of Optics, Vol.12, p.015405, 2009.

[15] F. Mehdizadeh, M. Soroosh, and H. Alipour-Banaei, "Proposal for 4-to-2 optical encoder based on photonic crystals," IET Optoelectronics, Vol. 11, pp.29-35, 2017.

[16] T. Daghooghi, M. Soroosh and K. Ansari-Asl, "A novel proposal for all-optical decoder based on photonic crystals," Photonic Network Communications, Vol.35, pp.335-341, 2018.

[17] S. W. Dietrich, D. Goelman, C. M. Borror, and S. M. Crook, "An animated introduction to relational databases for many majors,"IEEE Transactions on Education, Vol. 58, pp.81-89, 2014.

[18] S. Lin, Y.-B. Kim, and F. Lombardi, "CNTFET-based design of ternary logic gates and arithmetic circuits, " IEEE transactions on nanotechnology, Vol. 10, pp.217-225, 2009.

[19] M. Mukaidono, "Regular ternary logic functions ternary logic functions suitable for treating ambiguity", IEEE transactions on computers, pp.179-183, 1986.

[20] S. A. Hosseini and S. Etezadi, "A novel very low-complexity multi-valued logic comparator in nanoelectronics," Circuits, Systems, and Signal Processing, Vol.39: pp.223-244, 2020.

[21] M. Ghelichkhan, S. A. Hosseini, and S. H. P. Komleh, "Multi-digit Binary-to-Quaternary and Quaternary-to-Binary Converters and Their Applications in Nanoelectronics," Circuits, Systems, and Signal Processing, Vol.39: pp.1920-1942, 2020.

[22] F. Motalebi, and S. Sayedsalehi, "Design of Low Power Full-Adder Circuit Using Quantum-dot Cellular Automata," International Journal of Industrial Electronics Control and Optimization, Vol.5, pp.99-108, 2022.

[23] D. Akinwande, J. Liang, S. Chong, Y. Nishi, and H. S. P. Wong, "Analytical ballistic theory of carbon nanotube transistors: experimental validation, device physics, parameter extraction, and performance projection," Journal of Applied Physics, Vol.104, pp.1–7, 2008.

[24] M. Yousefi, K. Monfaredi, Z. Moradi, "Design and Simulation of Pseudo Ternary Adder based on CNTFET, " AUT Journal of Electrical Engineering, Vol.54, pp.361-376, 2022.

[25] K. Navi, V. Foroutan, M. R. Azghadi, M. Maeen, M. Ebrahimpour, M. Kaveh, et al., "A novel low-power full-adder cell with new technique in designing logical gates based on static CMOS inverter, "Microelectronics Journal, Vol. 40, pp.1441-1448, 2009.

[26] P. C. Balla and A. Antoniou, "Low power dissipation MOS ternary logic family, "IEEE Journal of Solid-State Circuits, Vol. 19, pp. 739-749, 1984.

[27] J. T. Butler, Multiple-valued logic in VLSI design: IEEE Computer Society, 1991.

[28] M. H. Moaiyeri, A. Doostaregan, and K. Navi, "Design of energy-efficient and robust ternary circuits for nanotechnology, " IET Circuits, Devices & Systems, Vol. 5, pp.285-296, 2011.
[29] P. Keshavarzian and K. Navi, "Universal ternary logic circuit design through carbon nanotube technology," International Journal of Nanotechnology, Vol.6, pp.942-953, 2009.
[30] R. Rajaei, and A. Amirany, "Reliable, high-performance, and nonvolatile hybrid SRAM/MRAM-based structures for reconfigurable nanoscale logic devices," Journal of Nanoelectronics and Optoelectronics, vol.13, pp.1271-1283, 2018.
[31] E. Roosta, and S.A. Hosseini, "A novel multiplexer-based quaternary full adder in nanoelectronics," Circuits, Systems, and Signal Processing, Vol.38, pp.4056-4078, 2019.
[32] M. Ebrahimy, M. Gholami, H. Adarang, and R. Yosefi, "Novel robust quantum-dot cellular automata (QCA) full adder in the one-dimensional clock," International Journal of Nano Dimension, vol.13, no.4, pp.353-361, 2022.
[33] Z. Wang, W. Zhao, E. Deng,J.O. Klein and C. Chappert,"Perpendicular-anisotropy magnetic tunnel junction switched by spin-Hall-assisted spin-transfer torque," Journal of Physics D: Applied Physics, Vol.48, p.065001, 2015.
[34] E. Deng, Y. Zhang, W. Kang, B. Dieny, J.O. Klein, G. Prenat and W. Zhao, "Synchronous 8-bit non-volatile full-adder based on spin transfer torque magnetic tunnel junction, "IEEE Transactions on Circuits and Systems I: Regular Papers, Vol.62, pp.1757-1765, 2015.
[35] M.D. Bishop, G. Hills, T. Srimani, C. Lau, D. Murphy, S. Fuller, J. Humes, A. Ratkovich, M. Nelson and M.M. Shulaker, "Fabrication of carbon nanotube field-effect transistors in commercial silicon manufacturing facilities," Nature Electronics, Vol. 3, pp.492-501, 2020.
[36] M.M. Shulaker, G. Hills, R.S. Park, R.T. Howe, K. Saraswat, H.S.P Wong and S. Mitra, "Three-dimensional integration of nanotechnologies for computing and data storage on a single chip, "Nature, Vol.547(7661), pp.74-78, 2017.
[37] M. Krishna Gopi Krishna, A. Roohi, R. Zand and R.F. DeMara, "Heterogeneous energy-sparing reconfigurable logic: spin-based storage and CNFET-based multiplexing, " IET Circuits, Devices & Systems, Vol. 11, no.3, pp.274-279, 2017.
[38] F. Razi, M.H. Moaiyeri, R. Rajaei and S. Mohammadi, "A variation-aware ternary Spin-Hall assisted STT-RAM based on hybrid MTJ/GAA-CNTFET logic," IEEE Transactions on Nanotechnology, Vol.18, pp.598-605, 2019.
[39] A.A. Javadi, M. Morsali and M.H. Moaiyeri, "Magnetic nonvolatile flip-flops with Spin-Hall assistance for power gating in ternary systems," Journal of Computational Electronics, Vol. 19(3), pp.1175-1186, 2020.
[40] A. Amirany, M.H. Moaiyeri and K. Jafari, "Process-in-memory using a magnetic-tunnel-junction synapse and a neuron based on a carbon nanotube field-effect transistor," IEEE Magnetics Letters, Vol. 10, pp.1-5, 2019.
[41] P.L. McEuen, M.S. Fuhrer, and H. Park, "Single-walled carbon nanotube electronics,". IEEE transactions on nanotechnology, Vol.1, pp.78-85, 2002.
[42] J. Deng, and H.-S. P. Wong, "A compact SPICE model for carbon-nanotube field-effect transistors including nonidealities and its application—Part I: Model of the intrinsic channel region," IEEE Transactions on Electron Devices, Vol.54, No.12, 3186–3194, 2007.
[43] J. Deng, and H.-S. P. Wong, "A compact SPICE model for carbon-nanotube field-effect transistors including nonidealities and its application—Part II: Full device model and circuit per- formance benchmarking," IEEE Transactions on Electron Devi- ces, Vol.54, No.12, 3195–3205, 2007.
[44] A. Raychowdhury and K. Roy, "Carbon-nanotube-based voltage-mode multiple-valued logic design," IEEE Transactions on Nanotechnology, Vol. 4, pp.168-179, 2005.
[45] F. Sharifi, M. H. Moaiyeri, and K. Navi, "A novel quaternary full-adder cell based on nanotechnology," International Journal of Modern Education and Computer Science, Vol. 7, pp.19, 2015.

Farzaneh Yousefzadeh received the B.SC, M.SC degrees from Saraj University in 2017 and Azarbiaijan Shahid Madani University in 2019, respectively. She was Talented B.SC student and entrance without exam from B.SC to M.SC through scientific interview. His Current research interests include Design of Computing circuit based on CNTFET and CMOS transistors, design of analog and digital circuit. Farzaneh Yousefzadeh is M.SC Electronic Engineering , Azarbiaijan Shahid Madani university , Tabriz, Iran (e-mail: f.ahari167@gmail.com).

Mousa Yousefi received the B.Sc. degree from Urmia University, M.Sc., and PhD degrees from Tabriz University. He was with Electronic Research Center Group, during 2008 to 2011 and was also an academic staff with Islamic Azad University, Ilkhechi Branch from 2008 to 2012. He is currently with Electrical and Electronics Engineering Faculty, Azarbaijan Shahid Madani University, Tabriz, Iran. His current research interests include current mode integrated circuit design, low voltage, low power circuit, design RF circuit and systems and analog microelectronics and data converters. Mousa Yousefi is assistant professor with Department of Electrical and Electronic Engineering, Engineering Faculty, Azarbaijan Shahid Madani University, Tabriz, Iran (e-mail: m.yousefi@azaruniv.ac.ir).

Khalil Monfaredi received the B.Sc., M.Sc., and PhD degrees from Tabriz University in 2001 and Iran University of Science and Technology in 2003 and 2011, respectively. He was with Electronic Research Center Group, during 2001 to 2011 and was also an academic staff with Islamic Azad University, Miyandoab Branch from 2006 to 2012. He served as the Research and Educational Assistant of Miyandoab Sama College from 2009 to 2011 and vice chancellor during 2011 to 2012. He is currently with Electrical and Electronics Engineering Faculty, Azarbaijan Shahid Madani University, Tabriz, Iran. He is the associate dean of engineering faculty, Azarbaijan Shahid Madani University since 2017. He is the author

or coauthor of more than thirty national and international papers and also collaborated in several research projects. He is also the founder of electronic department in Islamic Azad University-Miyandoab Branch and was the chairman of 2010 electronic and computer scientific conference (ECSC2010) held in Islamic Azad University, Miyandoab Branch. His current research interests include current mode integrated circuit design, low voltage, low power circuit and systems and analog microelectronics and data converters. Khalil Monfaredi is associate professor with Department of Electrical and Electronic Engineering, Engineering Faculty, Azarbaijan Shahid Madani University, Tabriz, Iran (e-mail: khmonfaredi@azaruniv.ac.ir).