دانشگاه سمنان

Journal of Modeling & Simulation in Electrical & Electronics Engineering (MSEEE)

Semnan University

6

Volume 1, Number 4, February 2021

Online ISSN 2821-0786

Applying Decoupling Method to a Dual-Band Antenna Array for Element Isolation

Elham Atashpanjeh[1*] and Pejman Rezaei [1]

Abstract— **In this study, a feasible decoupling method is used for a binary-element closely spaced dual-band antenna array. The decoupling setup is used for array antenna element isolation. Decoupling and matching at two widely separated frequencies are accomplished by using a two-layer (level) network approach. Moreover, impedance matching is built into the decoupling network. The antenna array consists of two elements and works in dual-band (4.9-5.4 GHz and 2.2-3.1 GHz). This is considering that the elements of the antenna arrays are strongly coupled. A decoupling network is taken into account between the array elements and each element port. Thus the mutual coupling between the array elements is reduced well. It should be noted that other decupling methods cannot isolate and match the impedance simultaneously in the two bands. This technique has been applied to the desired array. The proposed structure is simulated with High-Frequency Structure Simulator (HFSS) software.**

Index Terms— **Antenna array, Decoupling Network, Dual-band, Isolation.**

I. INTRODUCTION

The use of antenna arrays has been on the rise due to technological developments [1]. The multi-input multi-output (MIMO) technology is widely used to improve the data throughput in a multipath environment, from 4G smartphones to Wi-Fi modules. However, in a mobile terminal, the distance between antennas is usually miniature in terms of wavelength; therefore, the strong electromagnetic coupling among the multiple antennas rigorously decreases the benefits of the MIMO system [2]. The coupling of the antenna arrays is an issue that has been known for a very long time. Several earlier papers addressed this issue like [3, 4].

To increase the number of components in large-scale optical integrated circuits, it is imperative to decrease the device size as much as possible [5]. The isolated antenna model cannot be considered when the antennas are close to each other, since the existence of the mutual coupling. In [6], a class of microstrip antennas known as reduced surface wave (RSW) antenna was proposed based on the principle that a ring of magnetic current in a substrate will not excite TM0 surface waves. However, the circular patch adopted in [6] has a larger radius than a conventional one, which may impose restrictions in array applications. Using periodic structures such as electromagnetic-frequencies (transmission zeros) have to be close to each other band gap (EBG) structures [7, 8], defected ground structures (DGS) [9], is another approach to suppress surface waves, due to their band stop features. Moreover, compact phased arrays may be produced from small tuned electrical antenna elements [10].

However, some difficulties may be encountered in practical applications. For example, EBG structures require enough units to maintain the periodic property, which will occupy much space between elements. For DGS, various patterns printed on the ground may lead to severe backward radiation. Except for those techniques to suppress surface waves, there are also a few investigations focused on decreasing space wave coupling.

The decoupling techniques can be divided into two classes: network-based and structural approaches. The most common structural models involve applying neutralization lines, electromagnetic band gaps, defected ground structures and defective wall structures, and parasitic scatters. Furthermore, the choice of antennas and their relative placement and orientation is part of the category of structural approaches. The network-based schemes are typically analytic and are based on the appliance of network parameters and matrix operations [11].

The dual-band decoupling network [12]–[18] enables the radiating elements and the feeding network to be developed in isolation with explicit formulas. Filter-like structures were reported in [12] and [13] to offer dual-frequency decoupling with single-band impedance matching. Under the assumption of reasonably matched antenna elements, dual-band circuit components (such as resonator and phase shifter) were then harnessed for enhanced port isolation [14]–[18]. The decoupling networks are mainly realized by replacing every single band component in a typical design with its dual-band counterpart. However, the decoupling capability highly relies on the inherent characteristics of these dual-band components. For instance, dual-mode resonators were adopted [14] as the bridge circuit with the substantial drawback of restrained frequency band ratio ($f_2/ f_1 \approx 2$). In [15], due to the use of dual-band phase shifters with limited phase control range, the choices of frequency ratio and the types of mutual coupling are further restricted [15].

In [19], a two-layer approach was described to offer wideband decoupling of a two-element array with post-matching circuitry circuitry. However, the two preselected

($f_2/ f_1 < 1.2$), and the proposed design theory only applies to

1: Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
Corresponding author: e.atashpanjeh@semnan.ac.ir

symmetrical arrays. Besides, the decoupling methods of [14-19] rely heavily on the assumption of slack coupling and pre-matched radiating elements. The decoupling network method in [20] is based on the Butler matrix and can be isolated beam-steering antenna array. In [21] has been investigated an Efficient SIW-feed network subdues mutual coupling. The analytical method in [22] has been employed that works based on mutual coupling reduction using plane spiral orbital angular momentum electromagnetic wave.

If the impedance matching is not achieved by applying the decoupling network to the antenna array, an external dual-band impedance matching network will be needed which increases design complexity. In sturdily coupled cases, even with pre-matched antenna elements, the input return loss of the resulting array attained will often be degraded [12].

In this study, a dual-band network-based decoupling structure is applied to an antenna array and has shown the efficiency of the mentioned network for isolation between the two antenna elements. This is accomplished by applying the method presented in [12] to an antenna presented by simulation with HFSS software. The isolation in two separate bands is done hand to hand with the insulation of the antenna by the decoupling network. In addition, it also achieves impedance matching. For this reason, the basic concepts of the decoupling network are introduced in Section 2. Section 3 is devoted to showing the structure of the antenna array and also the effectiveness of the introduced decoupling method. The simulation results show the efficiency of this approach.

II. Decoupling Method

As it is known, the decoupling methods are utilized for reducing the mutual coupling between the array elements. Therefore, the distances of the elements are reduced and cause the dimension reduction of the array. In this study, a decoupling network is used to isolate the symmetric array elements [12]. Also, this structure can realize the isolation of two frequency bands. Fig. 1 shows the block diagram of the mentioned structure. The components details of Fig. 1 are described in [12]. These components are realized by microstrip lines.

The decoupling conditions are described as follows [11]:

$$y_{11} = y_{22} = 0 \tag{1}$$

Thus:

$$\mathrm{Re}[y_{12}] = \mathrm{Re}[y_{21}] = 0 \tag{2}$$

And also,

$$\mathrm{Im}[y_{12}] = \mathrm{Im}[y_{21}] = 0 \tag{3}$$

whereby y_{11}, y_{12}, y_{21}, and y_{22} are the parameters of the admittance matrix, and indexes 1 and 2 refer to ports 1 and 2.

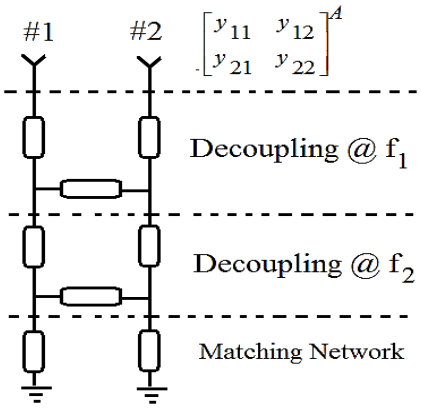


Fig. 1. Simple block diagram of the decoupling network

The mutual coupling is an inductive type; therefore, (2) is satisfied. On the other hand, the decoupling network should be designed to satisfy (3). Since the antenna array works in two frequency bands, these conditions should be established for each band, separately. Hence, the decoupling network consists of two levels. The first part is devoted to decoupling the array at the first frequency band and the second part for the second frequency band.

As indicated, the decoupling and matching are carried out at two arbitrarily chosen frequencies ($f_2 > f_1$). The two antenna elements (strongly coupled, asymmetric and unmatched) are connected to the decoupling network (two layers) followed by impedance matching circuitry. Each decoupling layer comprising of two transmission line sections and a bridge element. It is further assumed that the second bridge element exhibits zero mutual admittance at f_1.

III. Dual-Band Antenna Array with Decoupling Network Structure

In this section, the first symmetric dual-band antenna array is introduced. Then, a decoupling network structure is added to the array. Finally, the analyses of the simulation results are presented.These simulations are done using HFSS software. Consider a symmetric microstrip antenna array as shown in Fig. 2. This structure includes two elements. The monopole elements are placed close to each other on a 1.6 mm thick FR4 substrate. Also, the antenna array dimensions are shown in table 1.

TABLE I

The Dimension of proposed array antennas.

Variable	Value (mm)
L_G	80.0
W_G	55.0
L_A	30.0
D_A	9.8
L_D	6.3
W_A	3.0
W_D	3.5
L_E	5.0

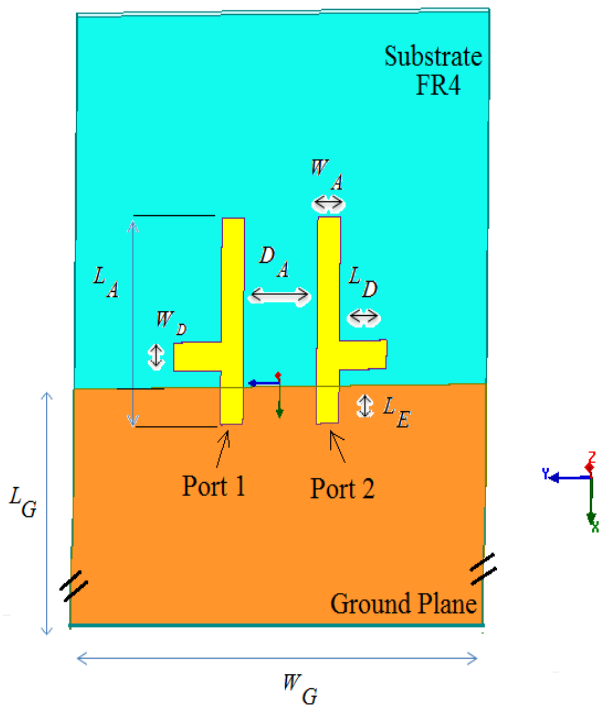


Fig. 2. The antenna array geometry and dimensions definition.

The antenna is simulated with HFSS software and the results are shown in Fig. 3 and 4. As it can be seen, the antenna array works in dual-band (4.9-5.4 GHz and 2.2-3.1 GHz) and the mutual coupling in these bands is inappropriate. The effect of using the decoupling network is considered in the rest of this section. Moreover, Fig. 4 represents the radiation sequence of the antenna array without considering the decoupling network.

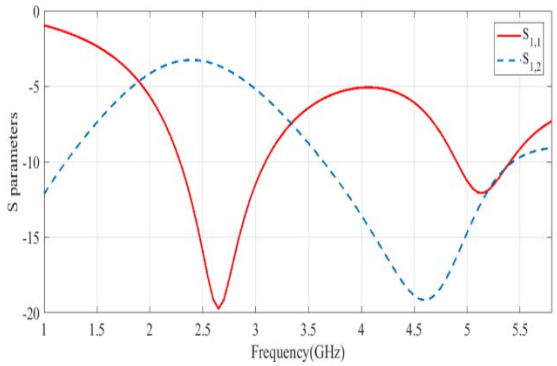


Fig. 3. The S-Parameters of the antenna array.

The reciprocal parameter related to the scattering matrix is undesirable even though the anticipated frequency band has the proper return loss amounts. It can be seen that According to Section 2, the decoupling network is added to the antenna arrays to improve the S_{12} and S_{21} parameters. The layout of the decoupling network is shown in Fig. 7.
In Table 2 circuit parameters of the decoupling network are presented.

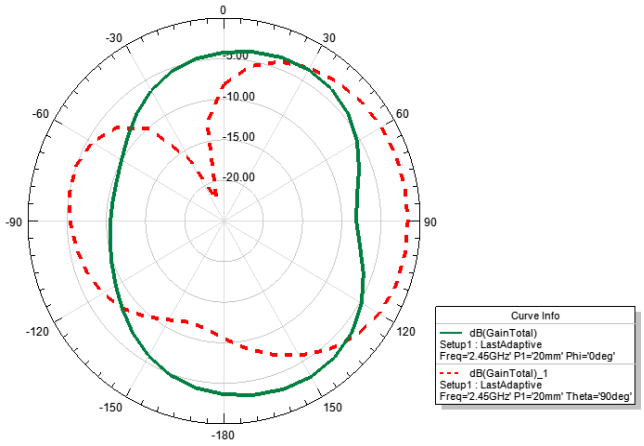


Fig. 4. The radiation pattern of the antenna array.

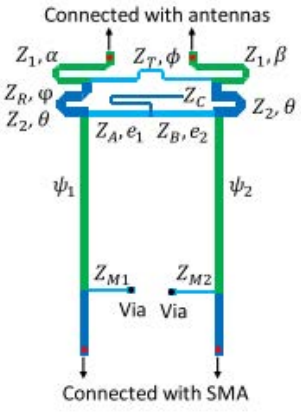


Fig. 5. The layout of decoupling and matching network.

TABLE II
The Dimension of circuit parameters of decoupling and matching network [12].

Unit: degree				Unit: Ohm			
α	135.5	θ	64.4	Z_1	62.3	Z_A	56.1
β	135.5	e_1, e_2	53.7	Z_T	79.8	Z_B	61.7
φ	116.4	ψ_1	148	Z_2, Z_R	48	Z_C	95
ϕ	1.32	ψ_2	150	Z_{M2}	89.9	Z_{M1}	91.2

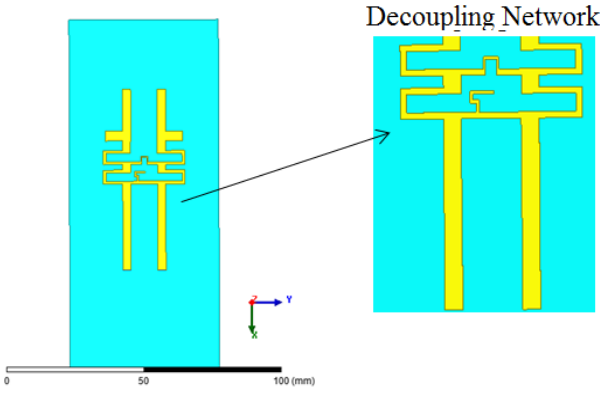


Fig. 6. Decoupled antenna array geometry.

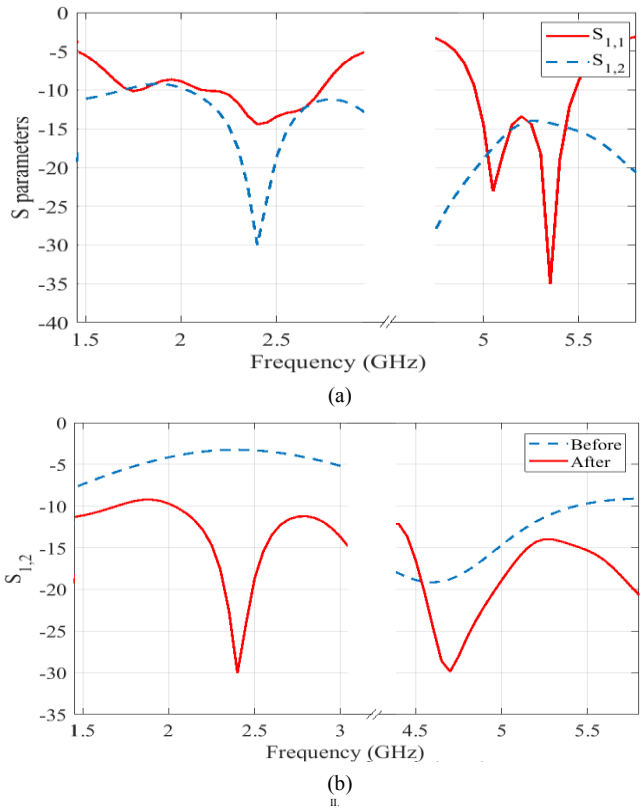


Fig. 7. (a) S-parameters of the decoupled antenna array, (b) the comparison of S_{12} before and after adding the decoupling network.

The new structure after adding the decoupling network is shown in Fig. 6. The results of the new structure simulation are depicted in Fig. 7. Fig. 7(a) represents the behaviors of the S_{11} and S_{12} parameters of the decoupled antenna array.

Also, Fig. 7(b) shows the comparison between the behaviors of the S_{12} parameter before and after applying the decoupling network. It can be seen that the S_{12} parameter is reduced which is expresses the reduction of the mutual coupling in both frequency bands. Also, Fig. 8 represents the radiation pattern of the antenna array considering the proposed decoupling network.

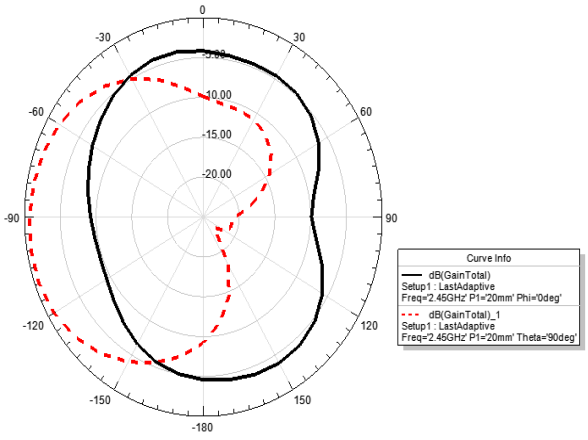


Fig. 8. The radiation pattern of the antenna array with decoupling network.

Fig. 8 shows the radiation pattern of the array antenna after isolation and adding the decoupling network. A comparison of Fig. 8 and Fig. 4 shows that the addition of a decoupling network did not adversely affect the expected radiation from the array. Table 3 contrasts the proposed method with similar techniques.

TABLE III
with other dual-band decoupling techniques.

Ref.	f2/f1	Enhancement in isolation	Improvement return loss	Decoupling Method
[15]	2.17 2.36	>15dB >20dB	Little	Dual-band T-junction
[14]	2.14	>20dB >15dB	Little	Dual-band resonator
[16]	2.14	>20dB >15dB	Little	Dual-band coupler
[19]	1.13	>20dB >20dB	Little	Dual-band two layer
Proposed	2 2.33	>20dB >15dB	>10dB >15dB	Dual-band two layer

IV. Conclusion

In this paper, one of the most important decoupling methods was implemented in the dual-band antenna array. The decoupling network was made up of two layers that isolated the antennas port and elements in two separated bands. Also, impedance matching was achieved simultaneously. This antenna array consists of two elements and works in dual-band (4.9-5.4 GHz and 2.2-3.1 GHz). The comparison of the mutual S parameters before and after using the decoupling method shows the efficiency of the decoupling network. It was seen that suitable isolation was obtained between the elements since the decoupling network was used.

V. Acknowledgment

The authors would like to thank the support from Semnan University. Also, the authors would like to thank the editor and reviewers for their constructive comments.

VI. References

[1] S. H. Chae, S. K. Oh, and S. O. Park, “Analysis of mutual coupling, correlations, and TARC in WiBro MIMO array antenna,” IEEE Antennas Wirel. Propag. Lett., vol. 6, pp. 122–125, 2007.

[2] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” IEEE Commun. Mag., vol. 52, no. 2, pp. 74–80, Feb. 2014.

[3] J. Andersen and H. Rasmussen, “Decoupling and descattering networks for antennas,” IEEE Trans. Antennas Propag., vol. 24, no. 6, pp. 841–846, Nov. 1976.

[4] P. Hannan, D. Lerner, and G. Knittel, “Impedance matching a phased-array antenna over wide scan angles by connecting circuits,” IEEE Trans. Antennas Propag., vol. 13, no. 1, pp. 28–34, Jan. 1965.

[5] S. Khani, M. Danaie, and P. Rezaei, “Size reduction of MIM surface plasmon based optical bandpass filters by the introduction of arrays of silver nano-rods,” Physica E: Low-dimensional Systems and Nanostructure, vol. 113, pp. 25-34, Sep. 2019.

[6] D. R. Jackson, J. T. Williams, A. K. Bhattacharyya, R. L. Smith, S. J. Buchheit, and S. A. Long, “Microstrip patch designs that do not excite surface waves,” IEEE Trans. Antennas Propag., vol. 41, no. 8, pp. 1026–1037, 1993

[7] M.M. Fakharian, P. Rezaei, "Mutual coupling reduction in microstrip array antenna using a novel compact electromagnetic band-gap structure," 2nd National Conf. New Technol. Electrical Computer Engin., Iran, Feb. 2019.

[8] F. Yang and Y. Rahmat-Samii, "Microstrip antennas integrated with electromagnetic band-gap (EBG) structures: A low mutual coupling design for array applications," IEEE Trans. Antennas Propag., vol. 51, no. 10, pp. 2936–2946, Oct. 2003.

[9] C.-Y. Chiu, C.-H. Cheng, R. D. Murch, and C. R. Rowell, "Reduction of Mutual Coupling Between Closely-Packed Antenna Elements," IEEE Trans. Antennas Propag., vol. 55, no. 6, pp. 1732–1738, Jun. 2007.

[10] M. Sharifi and P. Rezaei, "Conformal antenna array radiation pattern synthesis by tilt correction to improve Direction-of-Arrival estimation accuracy," Electromagnetics, vol. 40 no. 4, pp. 262-275, May 2020.

[11] R. Kormilainen, "Decoupling and matching network for a MIMO handset antenna," MSc Thesis in Radio Science and Engineering, Aalto University, Aug. 2018.

[12] Y.-F. Cheng and K.-K. M. Cheng, "A Novel Dual-Band Decoupling and Matching Technique for Asymmetric Antenna Arrays," IEEE Trans. Microw. Theory Tech., vol. 66, no. 5, pp. 2080–2089, May 2018.

[13] B. C. Pan and K. J. Cui, "Broadband decoupling network for dualband microstrip patch antennas," IEEE Trans. Antennas Propag., vol. 65, no. 10, pp. 5595–5598, Oct. 2017.

[14] L. Zhao and K.-L. Wu, "A dual-band coupled resonator decoupling network for two coupled antennas," IEEE Trans. Antennas Propag., vol. 63, no. 7, pp. 2843–2850, Jul. 2015.

[15] J. Sui and K.-L. Wu, "A general T-stub circuit for decoupling of two dual-band antennas," IEEE Trans. Microw. Theory Techn., vol. 65, no. 6, pp. 2111–2121, Jun. 2017.

[16] K.-C. Lin, C.-H. Wu, C.-H. Lai, and T.-G. Ma, "Novel dual-band decoupling network for two-element closely spaced array using synthesized microstrip lines," IEEE Trans. Antennas Propag., vol. 60, no. 11, pp. 5118–5128, Nov. 2012.

[17] P.-L. Chi, C.-J. Lee, and T. Itoh, "A compact dual-mand metamaterial based rat-race coupler for a MIMO system application," in IEEE MTT-S Int. Microw. Symp. Dig., Jun. 2008, pp. 667–670.

[18] P.-L. Chi and T. Itoh, "Miniaturized dual-band directional couplers using composite right/left-handed transmission structures and their applications in beam pattern diversity systems," IEEE Trans. Microw. Theory Techn., vol. 57, no. 5, pp. 1207–1215, May 2009.

[19] Y.-F. Cheng and K.-K. M. Cheng, "Novel wideband decoupling technique for MIMO antenna arrays with two independently controlled transmission zeros," IEEE MTT-S Int. Microw. Symp. Dig., pp. 853–856, June 2017,

[20] H. Nimehvari Varcheh, P. Rezaei, "Integration of the modified Butler matrix and decoupling network for beam-steering antenna array," Int. J. RF Microw Comput. Aided Engin., e23015, Early Access, 2021.

[21] A. Abdi Diman, F. Karami, et al., "Efficient SIW-feed network suppressing mutual coupling of slot antenna array," IEEE Trans. Antennas Propagat., vol. 32, no. 3, 23015, 2022.

[22] A. Habibi Daronkola, F. Tavakkol Hamedani, P. Rezaei, N. Monteseri, "Mutual coupling reduction using plane spiral orbital angular momentum electromagnetic wave," J. Electromag. Wav. Appl., Early Access, vol. 36, no. 3, pp. 346-355, 2022.

Convertible Perfect Absorber with Single Ring Resonator: Tunable Single Band/Dual-Band Visible

Pouria Zamzam[1] , Pejman Rezaei[2] ، Seyed Amin Khatami[3] and Zahra Mousavirazi[4]

Abstract: **This study proposes a single and dual-band tunable and convertible perfect absorber in the infrared band, consisting of a dielectric layer and a metallic bottom film. Primarily a tunable single-frequency absorber in the infrared region had been introduced. Subsequently, with the change in geometric structure, the proposed structure can be converted from infrared single-band to visible double-band frequency. The numerical simulation results indicate that the absorption spectrum of the single-band resonator is tuned from 337.4 THz to 210.2 THz, 227.3 THz, and 297.7 THz by changing effective parameters: ring width, ring height, and dielectric height. Next, by the parametric study of the proposed absorber dimensions, the absorption rate is obtained 99% more at the designed frequencies; lastly, the dual-band absorption with an average performance of 99.98% in the visible spectrum. The proposed plasmonic absorber in this research has a variety of applications, including sensing, imaging, wavelength-selective thermal emission, photodetectors, and so on.**

Index Terms— Metamaterial perfect absorber, Dual-band, Tunable, Infrared, Terahertz.

I. INTRODUCTION

Increasing bit rate and consequently the bandwidth, as the demand of today's telecommunications, motivates research towards higher frequencies, THz, infrared, and optics [1, 2]. Redesigning of well-known communications devices in microwaves, such as antennas [3-6], filters [7, 8], switches [9, 10], sensors [11-14], the antenna [15], etc, is a requirement of this process.

There is no doubt that any material that is not available in nature can be called metamaterial. Metamaterials are actually composite and artificial materials that have been engineered. They can also be described as materials that exhibit abnormal, unique, and peculiar properties. The first attempts at uncovering synthetic materials were made in 1898 by Jagadish Chandra Bose who researched materials with chiral properties. But one of the most important researches in this field dates back to 1968 when Veselago first presented a theoretical work on a material with a negative refractive index (NRI) with concurrent $\varepsilon, \mu < 0$.

Metamaterial perfect absorbers (MPAs) have recently been studied in microwaves, terahertz (THz), infrared and visible bands. Metamaterials, with the sub-wavelength, have attracted increasing attention because they have exhibited novel, unique and strange properties and are not found in nature at all such as negative permittivity, negative refractive index, cloaking behavior, invisibility, reverse doppler effects, and perfect absorption. Due to the significant benefits of full absorption, several studies have been reported on perfect absorbers and multi-layered MPAs after Landy. et al [16]. MPAs are usually realized by using lossy materials in periodic structures but in some periodic structures, materials are used without loss. The initial and basic mechanism of MPAs is that the incident electromagnetic wave fields are completely confined and gradually, inside these lossy materials, are consumed and become close to zero [17]. MPAs can be potentially used in many areas, such as plasmonic sensors, light-harvesting, thermal imaging, thermal emitter, and so on, while those with the narrowband are more desirable for single-pixel imaging and thermal measurement [18, 19].

There are usually two ways to achieve a multi-band perfect absorber. One is to Put several resonance structures next to each other and assemble with various geometric parameters in a coplanar such as the absorbers demonstrated by Huang et al [20] and Wang et al [21]. Another way is arranging the structures vertically and stacking, which is very effective in obtaining multi-band absorbers and it was applied in the absorbers presented by Mo et al. [22] and Grant et al. [23].

In the first model, the placement and assembling of the structure next to each other are called the unit cell of the absorber. One of its disadvantages is its extremely large size. The second model is faced with complexities and limitations in fabrication. It is known that classical (traditional) electromagnetic absorbers can be made based on a metal-dielectric-metal structure, which is made of a dielectric spacer sandwiched between two metals [24]. The first top layer is known as the resonator. This layer is primarily responsible for absorbing electromagnetic responses. In the paper [25], the electromagnetic force in the terahertz band generated by a cross-shaped absorber is introduced. Also in the paper [26, 27], ultra-wideband symmetric G-shape metamaterial-based microwave absorber, design and analysis of perfect metamaterial absorber in GHz and THz frequencies are introduced.

In [28], Quad-band polarization-insensitive metamaterial perfect absorber based on bilayer graphene metasurface is

1: Pouria Zamzam is the Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
2: Pejman Rezaei is the Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
3: Seyed Amin Khatami is Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
4: Zahra Mousavirazi is 1604, 5720 cavendish Cote Saint Luc.
Corresponding author: zahra.mousavirazi@emt.inrs.ca

introduced. In [29, 30], ultra-thin dual-band polarization-insensitive and wide-angled perfect metamaterial absorber based on a single circular sector resonator structure and tunable metamaterial dual-band terahertz absorber are introduced, respectively. Also, Graphene-based terahertz metamaterial absorbers for broadband application are introduced in [31]. Absorber and sensor applications of complimentary H-shaped fishnet metamaterial for sub-terahertz frequency region are proposed in [32].

The dielectric spacer, which usually uses lossy material, is analyzed as a cavity. One of the reasons that the middle dielectric layer is chosen as a lossy material is because it increases the absorption width. The third and lowest layer in classic absorber structures is a thin metal film that ensures that the light wave is not completely transmitted from the structure. Due to the widespread use and large investments of optical communication instruments in the NIR frequency range $(760nm < \lambda < 2\mu m)$, this frequency range is very useful in telecommunications engineering. This paper demonstrates a new design concept for achieving a tunable single-band absorber with convertible capability to infrared region dual-band perfect absorber which consists of a single circular ring on the top of the dielectric layer and a thin metal film on the bottom.

The numerical results show that the absorption of the single band is more than 99% and the absorption spectrum is tuned from 337.4 THz to 210.2 THz, 227.3 THz, and 297.7 THz. The changes in effective parameters were then successively shown such as ring width, ring height, and dielectric height, and their effects on absorption peak and frequency displacement. The process was followed by achieving perfect dual-band absorption with an average of 99.98% by creating four gaps in the single resonator ring ***and changing the outer radius (r)*** in the visible region. Finally, for better apprehension, the absorption mechanism of the proposed structure and the electromagnetic field distribution is plotted and analyzed. The proposed electromagnetic absorber in this paper has potential application in spectral imaging, photodetector, energy harvesting, and so on.

II. Absorber Theory and Design

One of the main principles in designing an absorber is based on the theory of electromagnetic wave transmission and the mechanism of losses. Gradually, metamaterials can be described by complex permittivity $\varepsilon(\omega) = \varepsilon'(\omega) + j\,\varepsilon''(\omega)$ and complex permeability $\mu(w) = \mu'(w) + j\mu''(w)$, and the imaginary part of $\varepsilon''(w)$ and $\mu''(w)$ shows the EM losses. In the study of metamaterials, researchers often try to reduce imaginary parts to reduce losses. However, in the case of metamaterial absorbers, losses play a significant role. The resonance between the top metal layer and the bottom metallic film can be inferred from the absorption mechanism of the metamaterial absorber. Surface currents on the top layer caused by incident EM waves can induce electric resonance. Hence, the top metal layer can also interact with the bottom metallic film to form the magnetic dipole resonance. Our proposed absorber consists of a single circular metallic ring and metallic bottom film separated by a dielectric spacer, as shown in Fig. 1(a-c). The geometric parameters of the unit cell are shown in Fig. 1.

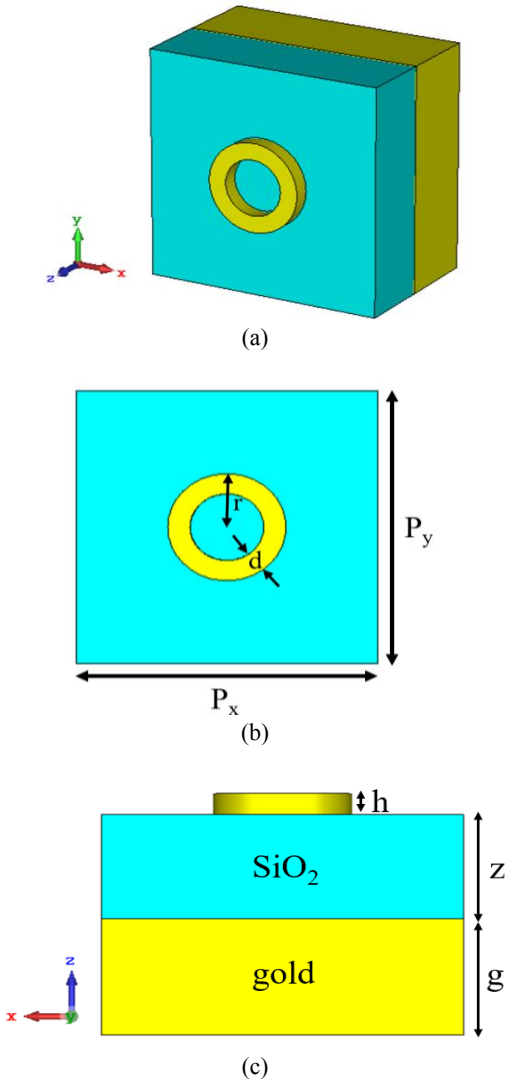


Fig. 1. The schematics of a single unit cell. (a) perspective view. (b) top view. (c) side view. The ring outer radius r=52 nm, d=20 nm, z=90 nm, g=100 nm, h=20 nm and $p_x = p_y = 286$ nm.

Based on the Finite Difference Time Domain (FDTD) method, the proposed absorber in this paper was calculated and optimized numerically. In all simulations, periodic boundary conditions were set in the x-and y-axis and a perfectly matched layer boundary condition was used along the z-axis. The dielectric layer is selected as SiO_2 substrate with a relative permittivity (ε =2.1+j0.0021) and two metallic layers on top and bottom made of gold by Drude model of $\varepsilon_{Au}=\varepsilon_\infty - \omega_p^2/\omega(\omega+j\mu)$ with plasma frequency $\omega_P=1.37\times10^{16}$ rad/s and collision frequency $\mu=4.07\times10^3$ rad/s and epsilon infinity is 9.5 [33, 34].

The absorptivity of the presented metamaterial absorber can be defined as A=1-T-R where T is transmissivity and R is reflectivity [35]. Because the thickness of the bottom Au film is much greater than the skin depth, the absorption can be obtained by A=1-R. Here it is assumed that the source is a plane wave that is normally incident onto the structure.

III. Result and Discussion

The unit-cell structure of the proposed design includes three functional layers: a ring resonator on the top, a dielectric spacer in the middle, and a bottom continuous metallic gold film. To study the efficiency of the proposed absorber, the model building and full-wave simulations were performed using the frequency domain solver in CST Microwave Studio.

Various parameters and the effect of their changes on peak absorption and frequency change had been investigated. The simulated absorption of the proposed absorber for different d values is shown in Fig. 2(a). Evidently, by decreasing the ring width by 2 nm, an increase in the absorption peak and a narrowing of the absorption band occurs, and then a perfect absorption with the absorptivity of more than 99.5% at 210 THz is obtained.

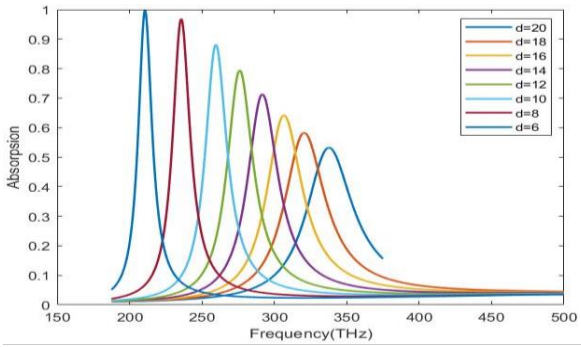
It is noteworthy mentioning that frequency change also occurs, but one of the advantages of checking multiple parameters separately is that it can be designed according to the desired frequency range and thus achieve a perfect narrow band absorber with a specific frequency.

Due to the constant value, d = 20 nm, the peak absorption is improved by a regular decrease of 10 nm from the dielectric thickness z respectively, while the other parameters are constant. As shown in Fig. 2(b), for z<40 nm; the peak absorption intensities are greater than 90% and the absorption is more than 98.8% when z=10 nm in 297 THz. Also, by adjusting the value of z, the frequency shifts and the absorption bandwidth can be further narrowed by reducing the thickness of the dielectric layer.

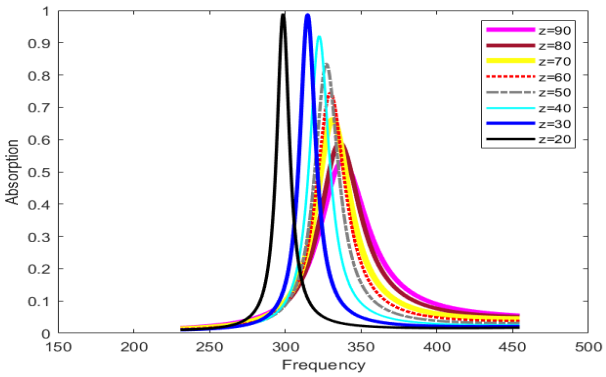
Also, the performance of the proposed absorber has been investigated by modifying ring height h from 20 nm to 6 nm, decreasing by 2 nm respectively when other parameters are constant (z=90, d=20), as shown in Fig. 2(c). The absorption peak intensities increase with decreased h. Then, with this decrease of ring height, the absorption band becomes narrower, which is desirable and the absorption is 99.5% at 227 THz.

Finally, the outer radius of the ring will be examined. Fig. 2(d), shows that by changing the outer radius (r) from 50 nm to 54 nm when the parameters of the structure are unchanged (d=6 nm, z=90 nm, and h=20 nm), the frequency changes from 218 THz to 206 THz. It should be noted that by changing this parameter, there is no change in the absorption peak and only allows it to have perfect absorption at the desired frequency by setting the above parameters.

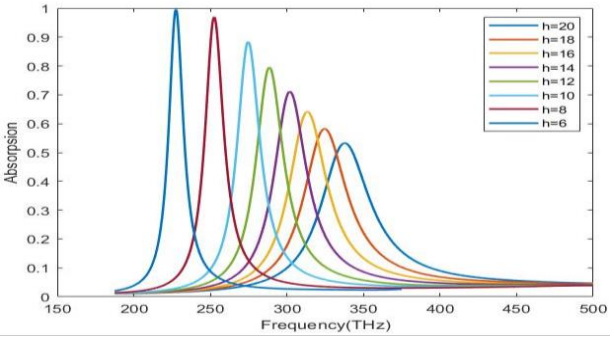
Also, this structure is insensitive to changes in incident wave polarization in TE mode, which is an advantage for the proposed structure. As shown in figure 3(a), when the polarization angle increases from 0 to 45 degrees, there is a change of about 0.0001 in the absorption rate at this frequency and no frequency shifts are occurring. According to the symmetry of the structure, polarization angle from 0 to 45 degrees has been studied and we are sure that the results will be repeated from 45 to 90 degrees. Therefore, the proposed structure is insensitive to polarization angle. As shown in Figure 3(b) another advantage of the proposed structure, in addition to being perfect single-banded, and insensitive to incident wave polarization, has a wide incident angle.



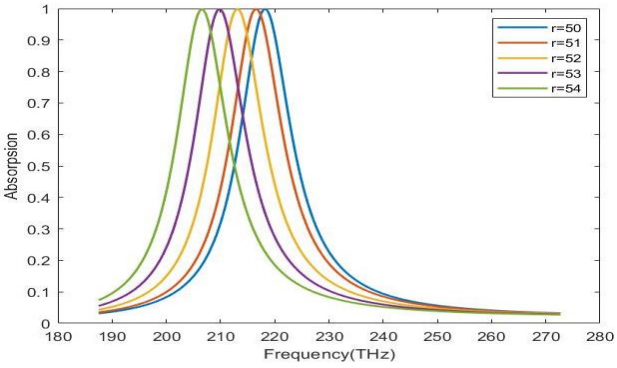
(a)



(b)

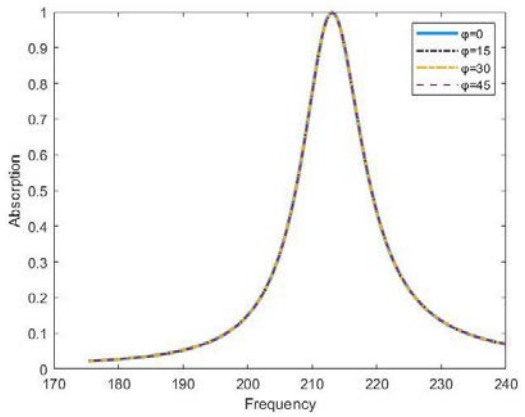


(c)

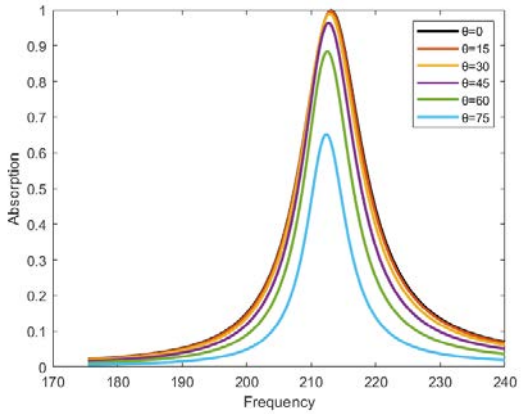


(d)

Fig. 2. The absorbance spectra of the proposed absorber for different geometric parameters: (a) Ring width "d". (b) dielectric thickness "z". (c) ring height "h". (d) outer radius "r".



(a)



(b)

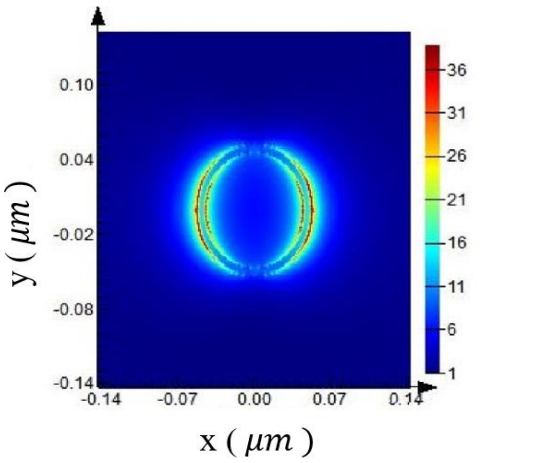
Fig. 3. (a) Dependence of the absorption spectra on the polarization angles of the normally incident wave. (b) Absorption spectra under different incidence angles with TE radiation.

To understand the physical mechanism underlying the absorption, we plotted the electric and magnetic field distribution in TM polarization for the resonant frequency. Note that accumulated surface charges can induce the magnetic field accountable for the magnetic resonance and resonance absorption.

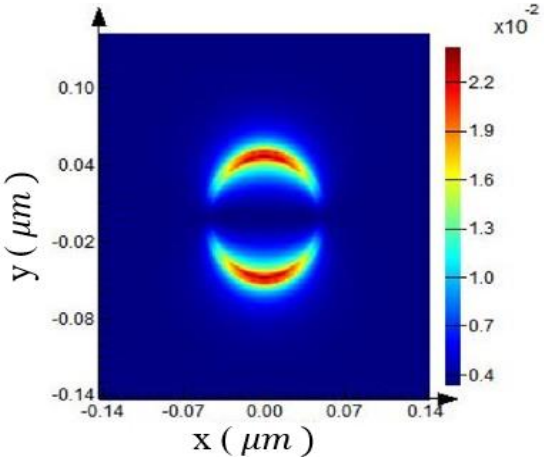
The electric and magnetic resonance are overlapped at the resonant frequency and then the perfect absorption of the normal incidence is realized. As shown in Fig. 4(a) the electric field is mainly concentrated at both edges of the ring in the right and left sides of the ring resonator along the x-axis for the resonant frequency.

It is easy to find that the classic absorption mechanism is composed of an electrical resonance (realized by the top resonator layer) and a magnetic resonance (realized by the top and bottom metal layer), and it matches well with the results published in [6]. Note that the ring only excites the dipolar resonance. The accumulation of large surface charges causes the strong electric field distribution, and the existence of the strong electric resonance in the metallic array is due to the strong coupling of the metallic resonance structure, the dielectric spacer, and the bottom metallic film. For different resonance frequencies, the strong coupling appears in different parameters. As shown in Fig. 4(b) the magnetic field focuses on the surface of the ring along the y-axis.

As shown in Fig. 5, four gaps were used in the ring, while the other parameters remained constant. Then, by increasing the gap spacing, the effect of these gaps was investigated.



(a)



(b)

Fig. 4. (a) Electric field distribution in the interface between the top patterned metallic layer and the air layer for the proposed absorber at resonant frequency (TE mode). (b) Magnetic field distribution at resonant frequency (TE mode).

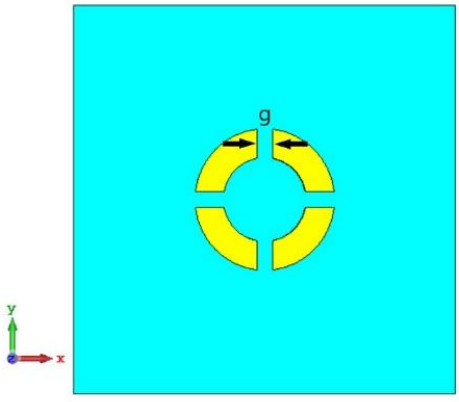


Fig.5. The schematics of a single unit cell with four gaps (top view).

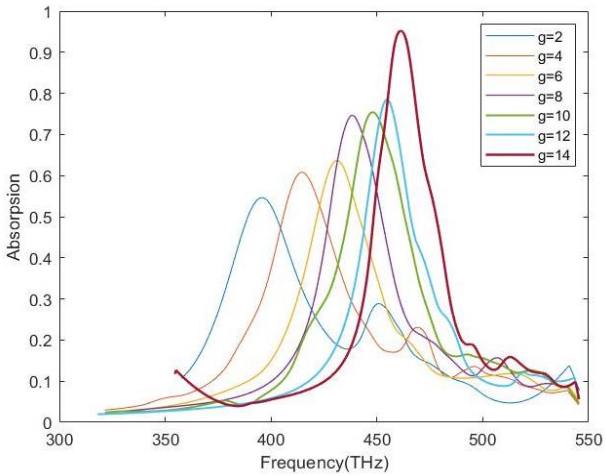


Fig. 6. The absorbance spectra of the proposed absorber with 4 gaps and with constant other parameters (d = 20 nm, h = 20 nm, z = 90 nm, r = 52 nm).

It can be seen in Fig. 6 by gradually increasing the gap spacing, the peak of absorption increases, and then 95% absorption is achieved at 461 THz.

Moreover, with a parametric study, we were able to obtain the structure of the dual-band using the same resonator ring. Using the same four gaps in the ring, a spacing of 2 nm (g = 2 nm), the value d = 6 nm, and with the change of the outer radius (r), finally, the average absorption is 99.7% at 446 THz and 470 THz, as shown in Fig. 7.

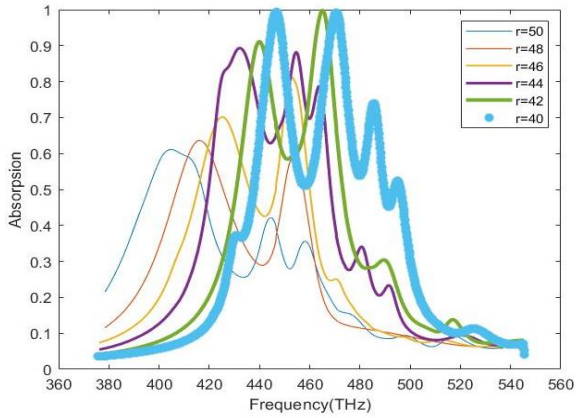
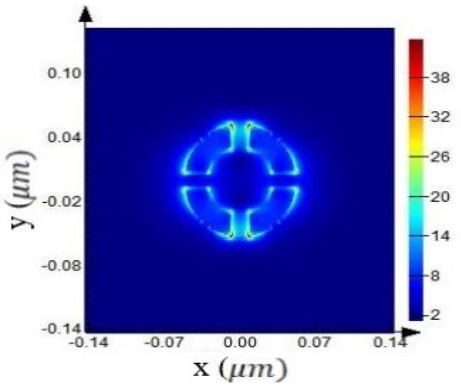
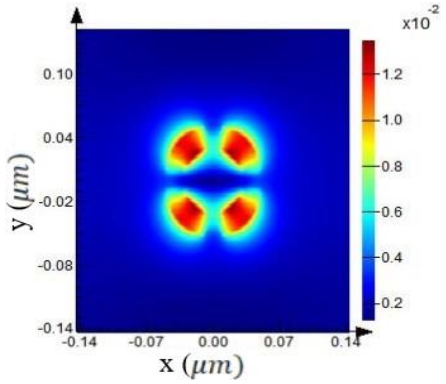


Fig. 7. The absorbance spectra of the proposed absorber with gaps and with changes of outer radius (r).



(a)



(b)

Fig. 8. (a) Electric field distribution of ring resonator with gaps at a resonant frequency. (b) Magnetic field distribution.

Also, as shown in Fig. 8(a) that the electric field concentrates near the cracks of gap rings along the y axis and unlike the electric field, there is a magnetic field on the four arms of this resonator, as you can see in Fig. 8(b). Finally, the compared performances of the proposed absorber and the previous works are collected in Table I.

TABLE I
They Compared Performances of the Proposed Absorber and the Previous Works

Ref.	Frequency (THz)	Cell dimension (μm)	Layers number	Absorptionrate
[31]	0.69	400 × 400	3	97.3%
[36]	4.24,5.89, 9.66, 10.62	3 × 3	3	97.74% Average
[37]	4.24, 6.86	9 × 9	3	97.18% Average
[38]	0.95	50 × 50	3	99.80%
[39]	1.6	50 × 50	2	95%
[40]	0.59	150 × 150	4	99%
[41]	1.26	150 × 150	3	99.93%
[42]	2.47-2.9	60 × 60	3	16% Bandwidth
Proposed	210.2 OR 227.3 OR 297.7 OR 337.4	0.286 × 0.286	3	99.98%

As shown in the table, the structure proposed in this paper has a perfect absorption peak, which is an advantage over other structures. Also as an important advantage, this structure in each of the frequencies 210.2, 227.3, 297.7, and 337.4 terahertz can provide perfect absorption depending on the different applications in these frequencies.

IV. Conclusions

In conclusion, we have illustrated a single-band absorber with convertible capability to dual-band absorber using single ring resonator based on metamaterial. The unit-cell structure of the design consists of a single circular ring on the top of the

dielectric spacer and the dielectric spacer on the continuous bottom metallic film. Results show with changed parameters in structure, absorption peaks with average absorptions higher than 99% can be gained. Finally, the circular ring was examined with 4 gaps and the effects of the spacing of these gaps. Then with a parametric study on the outer radius of the circular ring (r) with these gaps, a double-band absorber was obtained. Electric and magnetic field distributions are analyzed to understand the physical characteristics of the proposed metamaterial structure.

V. Acknowledgments

Also, the authors would like to thank the editor and reviewers for their constructive comments.

References

[1] Choi JH. (2015) High-speed devices and circuits with THz applications., CRC Press.

[2] Pereira MF, and Shulika O. (2011) Terahertz and Mid Infrared Radiation: Generation, Detection and Applications., Springer, Netherlands.

[3] Jafari Chashmi M., Rezaei P., and Kiani N. (2020) Y-shaped graphene-based antenna with switchable circular polarization., Optik, vol. 200, pp. 163321.

[4] Ullah S, Ruan C, Tanveer Ul Haq TUI and Zhang X. (2019) High performance THz patch antenna using photonic band gap and defected ground structure, J. Electromag. Wav. Appl., vol. 33, no. 15, pp. 1943-1954

[5] Jafari Chashmi M., Rezaei P., and Kiani N. (2020) Polarization controlling of multi resonant graphene-based microstrip antenna., Plasmonics, vol.15, no. 1, pp. 417-426.

[6] Jafari Chashmi M., Rezaei P., and Kiani N. (2019) Reconfigurable graphene-based V-shaped dipole antenna: From quasi-isotropic to directional radiation pattern., Optik, vol.184, pp. 421-427.

[7] Khani S., Danaie M., and Rezaei P. (2019) Size reduction of MIM surface plasmon based optical bandpass filters by the introduction of arrays of silver nano-rods., Phys. E: Low-dimens. Syst. Nanostruct., vol. 113, pp. 25-34.

[8] Khani S., Danaie M., and Rezaei P. (2019) Tunable single-mode bandpass filter based on metal-insulator-metal plasmonic coupled U-shaped cavities., IET Optoelec., vol. 13, no. 4, pp. 161-171.

[9] Khani S., Danaie M., and Rezaei P. (2020) All-optical plasmonic switches based on asymmetric directional couplers incorporating bragg gratings., Plasmonics, vol. 15, pp. 869-879.

[10] Khani S., Danaie M., and Rezaei P. (2020) Compact and low-power all-optical surface plasmon switches with isolated pump and data waveguides and a rectangular cavity containing nano-silver strips., Superlattices Microstruct., vol. 141, pp. 106481.

[11] Ghods MM., and Rezaei P. (2018) Graphene-based Fabry-Perot resonator for chemical sensing applications at mid-infrared frequencies. IEEE Photon. Technol. Lett., vol. 30, no. 22, pp. 1917-1920.

[12] Kiani S., Rezaei P., Fakhr M., An overview of interdigitated mic rowave resonance sensors for liquid samples permittivity detect o n, Ch. 7 in Interdigital Sensors, Springer, 2021. DOI: 10.1007/97 8-3-030-62684-6_7

[13] Sharma A.K., Kaur B. (2018) Analyzing the effect of graphene's chemical potential on the performance of a plasmonic sensor in infrared. Sol. State. Commun., Vol. 275, pp. 58-62.

[14] Nickpay M.R., Danaie M., Shahzadi A. (2021) Highly sensitive THz refractive index sensor based on folded split-ring metamaterial graphene resonators, Plasmonics. Doi:10.1007/s11468-021-01512-8.

[15] Nickpay M.R., Danaie M., Shahzadi A. (2019) Wideband Rectangular Double-Ring Nanoribbon Graphene-Based Antenna for Terahertz Communications. IETE J. Research, pp. 1-10.

[16] Landy NI., et al. (2008) Perfect Metamaterial Absorber. Phys. Rev. Lett., vol. 100, pp. 207402.

[17] Chen HT. (2012) Interference theory of metamaterial perfect absorbers., Opt. Express, vol. 20, pp. 7165-7172.

[18] Meng L., Zhao D., Ruan Z., Li Q., Yang Y., and Qiu M. (2014) Optimized grating as an ultra-narrow band absorber or plasmonic sensor., Opt. Lett., vol. 39, no. 5, pp. 1137-1140.

[19] He J., Ding P., Wang J., Fan C., and Liang E. (2015) Ultra-narrow band perfect absorbers based on plasmonic analog of electromagnetically induced absorption., Opt. Express, vol. 23, no. 5, pp. 6083-6091.

[20] Huang L., Chowdhury DR., Ramani S., Reiten MT., Luo SN., Taylor AJ., and Chen HT. (2012) Experimental demonstration of terahertz metamaterial absorbers with a broad and flat high absorption band., Opt. Lett., vol. 37, no. 2, pp. 154-156.

[21] Wang DG., Liu HM., Hu WX., Kong HL., Cheng LL., and Chen QZ., (2013) Broadband and ultra-thin terahertz metamaterial absorber based on multi-circular patches., Eur. Phys. J. B, vol. 86, no.7, pp. 1-9.

[22] Wen QY., et al. (2013) A polarization-independent and ultra-broadband terahertz metamaterial absorber studied based on circular-truncated cone structure. Acta. Phys. Sin-CH. ED., vol. 62, pp. 237801.

[23] Grant J., Ma Y., Saha S., Khalid A., and Cumming DRS. (2011) Polarization insensitive, broadband terahertz metamaterial absorber., Opt. Lett., vol. 36, no.17, pp. 3476-3478.

[24] Fu J., et al. (2015) The electromagnetic force in the terahertz band generated by a cross-shaped absorber., Sol. State. Commun., vol. 204, pp. 5-8.

[25] Ghods MM., and Rezaei P. (2017) Ultra-wideband microwave absorber based on uncharged graphene layers, Electromagnetic waves and applications., J. Electromag. Wav. Appl., vol. 32, no. 15, pp. 1950-1960 (2017).

[26] Naqavi SA., Baqir MA. (2018) ultra-wideband symmetric G-shape metamaterial based microwave absorber., J. Electromag. Wav. Appl., vol. 32, no. 16, pp. 1492976.

[27] Dincer F., Karaaslan M., and Sabah C. (2015) design and analysis of perfect metamaterial absorber in GHz and THz frequencies., J. Electromag. Wav. Appl., vol. 29. no.18, pp. 1043030.

[28] Zamzam P., Rezaei P., Khatami SA. (2021) Quad-band polarization-insensitive metamaterial perfect absorber based on bilayer graphene metasurface., Phys. E, vol. 128, pp. 114621.

[29] Lou C.Y., et al. (2015) tunable metamaterial dual-band terahertz absorber., Sol. State. Commun., Vol. 222, pp. 32-36.

[30] Ahmadi H., et al. (2021) Graphene-based terahertz metamaterial absorber for broadband application., Sol. State. Commun., vol. 323, pp. 114023.

[31] Aksimsek S. (2020) design of an ultra-thin, multiband, micro-slot based terahertz metamaterial absorber., J. Electromag. Wav. Appl., vol. 34, no. 16, pp.1809532.

[32] Faruk A., and Sabah C. (2018) Absorber and sensor applications of complimentary H-shaped fishnet metamaterial for sub-terahertz frequency region., Optik, vol. 177, pp. 64-70.
[33] Ordal MA., Long LL., Bell RJ., Bell SE., Bell RR., Alexander RW., and Ward CA. (1983) Optical properties of the metals al, co, cu, au, fe, pb, ni, pd, pt, ag, ti, and w in the infrared and far infrared., Appl. Opt., vol. 22, pp. 1099-1119.
[34] Norouzi Razani AH., Rezaei P. (2022) Broadband polarization insensitive and tunable terahertz metamaterial perfect absorber based on the graphene disk and square ribbon, Micro Nanostruct., vol. 163, 107153.
[35] Wang BX., Wang GZ., Wang LL., and Zhai X. (2015) Design of a five-band terahertz absorber based on three nested split-ring resonators., IEEE Photon. Technol. Lett., vol. 28, no. 3, pp. 307-310.
[36] Norouzi Razani A, Rezaei P. (2022) Multiband polarization insensitive and tunable terahertz metamaterial perfect absorber based on the heterogeneous structure of graphene., Opt. Quant. Electron., vol. 54, 407.
[37] Zamzam P., Rezaei P. (2021) A terahertz dual-band metamaterial perfect absorber based on metal-dielectric-metal multi-layer columns., Optic. Quant. Electron., vol. 53, pp. 109.
[38] Zheng W., Li W., and Chang SJ. (2015) A thermally tunable terahertz metamaterial absorber., Optoelec. Lett., vol. 11, pp. 18-21.
[39] Liu J., Fan L., Ku J., and Moe L. (2016) Absorber: a novel terahertz sensor in the application of substance identification., Opt. Quant. Electron., vol. 48, no. 2, pp. 80.
[40] Kong H., Li G., Jin Z., Ma G., Zhang Z., and Zhung C. (2012) Polarization-independent metamaterial absorber for terahertz frequency., J. Infrared, Millimeter, and Terahertz Wav., vol. 33, pp. 649-656.
[41] Ling X., Xhiao Z., and Zheng X. (2018) Tunable terahertz metamaterial absorber and the sensing application., J. Mater. Sci. Mater. in Electron., vol. 29, pp. 1497-1503.
[42] Nickpay M.R., Danaie M., Shahzadi A. (2021) A wideband and polarization-insensitive graphene-based metamaterial absorber., Superlatt. Microstru. vol. 150, pp. 106786.

Improvement of the Sum Rate and Energy Efficiency of IA-based Cognitive Radio Network by Successive Relaying and Power Allocation

Elahe Maddah[1] and Mohammad Lari[2]

***Abstract*-- In this paper, we propose an underlay cognitive radio network that consists of several secondary users and one successive relaying-aided primary user. Two half-duplex relays operate as full-duplex relays in the successive relaying technique. To improve spectral efficiency, the primary user utilizes the successive relaying technique. Inter relay interference and inter-user interference are challenges of the proposed network. For eliminating these interferences, the interference alignment method is utilized. Also, two power allocation algorithms are proposed to maximize the sum rate of secondary users and the energy efficiency of the network. In both power allocation algorithms, satisfying the quality of service of the primary user is considered. The closed-form solutions of these algorithms are obtained. We use the fractional programming approach to solve energy efficiency optimization in two steps.**

***Index Terms*-- Cognitive radio network, Interference suppression, Interference alignment, Power allocation, Successive relay, Inter relay interference.**

I. INTRODUCTION

These days, the number of devices connected to the wireless network has increased, among other reasons, because of some technologies such as the internet of things and device-to-device communications. Therefore, operators have to improve coverage, network capacity, reliability, and spectrum management while reducing operating costs. Thus, new approaches, for example, spectral sharing, full-duplex relays, and their combination, help improve coverage, energy efficiency, and spectral efficiency [1, 2].

Spectrum sharing in cognitive radio networks (CRN) is an approach to overcome spectral resource deficiency [3, 4]. CRN consists of primary users (PUs) and secondary users (SUs). SUs are allowed to use the spectrum of PUs to send their data while guaranteeing the performance of the PUs [5]. PUs share their spectrum in two ways: Overlay spectrum sharing and Underlay spectrum sharing. In the Overlay spectrum sharing, SUs can use the spectrum without PUs. In the Underlay spectrum sharing, SUs can simultaneously use the spectrum with PUs. Thus, interferences appear in all receivers. In this situation, the transmitted power of SUs should be controlled because of decreasing the interference in PU's receiver and guaranteeing their performance [6, 7]. Thus, improving the rate of secondary users by considering the quality of service (QoS) of PUs is one of the spectrum sharing challenges [8].

The CRN is an interference network so, cancellation of all interference is necessary. Interference alignment (IA) is one of the interference cancellation methods that has attracted enormous interest recently. The best definition of IA is aligning all interference in one subspace to increase free interference dimensions for the desired signal [9, 10]. IA provides a convenient transmission without any interference. The received SINR of PU is decreased compared to the situation without SUs and IA technique so, the QoS of PU is decreased [11]. Therefore using the relays in CRN is one of the powerful ways to improve CRN performance [12, 13]. Furthermore, power allocation is another effective approach to cope with this challenge [38].

In addition to the CRN, cooperative networks (relays) have also presented to overcome challenges such as the shortage of spectrum resources, higher QoS demands for users, and lower power consumption for transmission. A cooperative network in telecommunication systems increases network coverage and counteracts the effect of path loss. With a given power consumption, relay networks achieve higher capacity than networks without relays. [14, 15]. Therefore, relays help to save power. This feature improves the energy efficiency of the network. Accordingly, in a cognitive radio network, relays help to save more power to serve SUs [16-19]. Relays are divided into two groups, namely full-duplex and half-duplex relays. Two time-frequency resources are required in the end-to-end transmission between the source and destination with the help of half-duplex (HD) relays [20]. But in full-duplex (FD) relays transmission, one time-frequency resource is required. In other words, FD relays send and receive simultaneously in the same bandwidth. Accordingly, the required resources (bandwidth and time) are reduced to half also spectral efficiency is increased compared to HD relays. The main challenge of the FD relays is the interference from the transmitter side of the relay to the own receiver, which is known as self-interference (SI). Much research has been done to eliminate this interference, but it is not eliminated in practice [21, 22]. The receiver and transmitter of the relay are in the same device. As a result, self-interference from receiver to own receiver is strong. Hence, management and cancellation of self-interference are complex [22].

Authors in [23] proposed a successive relaying (SR) scheme that performs as the full-duplex relay. SR scheme consists of a source, a destination, and two relays in each

1: E. Maddah is with the Faculty of Electrical, and Computer Engineering Semnan University, Semnan, Iran.
2: M. Lari is with the Faculty of Electrical, and Computer Engineering Semnan University, Semnan, Iran.
Corresponding author: : m_lari@semnan.ac.ir

time slot, one of them is in transmit mode, and the other one is in receive mode. The receiver relay receives a new data frame from the source while the other forwards the previous data frame to the destination. Then, in the next time slot, relays change their mode. All the transmissions are simultaneously and have the same bandwidth; therefore, the receiver relay receives interference from the transmitter relay. This interference is called inter relay interference (IRI) [24].

In successive relay performance, the distance between the relays is more than the distance between transmitter and receiver of the FD relay, so IRI is weaker than SI. If the relays are far apart, the intensity of the IRI is very low and can be ignored. Furthermore, if relays are close to each other, IRI is strong; hence it can be decoded entirely and removed from the received signal in the receiver relay [25]. Besides, according to the complexity and cost of implementing SI cancellation in FD relays [24] and utilizing the present devices, SR can be effective, and thus some researchers are interested [26-29]. Therefore, the SR technique and power allocation in CRN are effective. In the following subsection, we express some research in this field.

Related Works:

The combination of SR and CRN (cooperative CRN) creates different system models that improve spectral efficiency and decrease the transmitted power. Additionally, an appropriate power allocation enhances the performance of this network [30-33]. In [30], spectrum sharing is done with the help of SR. Then, the transmission rate is maximized. In this paper, there are two Sus and one PU, where transmitters of the SUs operate as relays for PU in the SR technique. Authors in [31] proposed a network that the two relays of the SR method act as SU transceivers while serving as relays for PU. This paper uses the IRI between the two relays to transmit the SU data. Then, two optimization problems are formulated to minimize the BER and maximize the average achievable rate. Also, [32] proposed a hybrid satellite system with the help of the SR technique and then maximized the system's capacity. In [33], there are two secondary and one primary network. The secondary networks consist of a base station and many users. Base stations of secondary networks act as the relays for PU by SR. Then, by maximizing the rate of secondary networks, design the beamforming matrix.

Our Works:

In this paper, we propose an underlay spectrum sharing CRN that utilizes the SR technique in the PU to improve the QoS of PU while increasing the rate of SUs. Utilizing the SR technique in PU to save more power for SUs has not been proposed previously. Besides, we apply the power allocation technique to optimize the network's performance. In the SR technique, we need to manage the IRI. Some analyses of IRI partial or full cancellation [34, 35]. This paper applies IA to eliminate all interferences in CRN and IRI. One of our new works in this paper is adopting the IA method to manage IRI. Spectral efficiency and energy efficiency (EE) are significant parameters in green communication [36]. Therefore, the main goal of this paper is spectral and energy efficiency optimization of the network with considering QoS of PU. Our investigations for spectral efficiency and energy efficiency are as follows:

- In CRN, the quality of service of PU should be guaranteed. Therefore, we first obtain the minimum transmitted power to satisfy the QoS of PU. We apply an appropriate approximation in our equations for Simplification.
- Next, we propose a power allocation for maximizing the sum rate of SUs while guaranteeing the QoS of the PU. Then, we obtain the closed-form solution for the problem.
- Also, we propose a power allocation problem to maximize the energy efficiency of the network. Because of the complexity of the problem, we solve it in two steps. Therefore, the final solution is expressed as an algorithm. In this algorithm, the QoS of PU is considered.

The proposed optimization problems in this work have not been solved previously. Because of the appropriate approximation, the closed-form results are easily obtained.

This paper is organized as follows: In Section 2, the proposed system model, the SR scheme, the cognitive radio network, and the IA method are described. In Section 3, two power allocation algorithms optimize the network's sum rate of Sus and EE. Then in Section 4, the results are simulated, and section 5 is the paper's conclusion.

Notation: $\mathbf{I}_d$ represents the $d \times d$ identity matrix. $\mathbf{A}^{\dagger}$ and $|\mathbf{A}|$ are the Hermitian transpose and the determinant of matrix A, respectively. $\|a\|$ is the $\ell 2$-norm of vector a. |a| is the absolute value of complex number a. $\square^{M \times N}$ is the space of complex M × N matrices. $\mathcal{CN}(\mathrm{a}, \mathbf{A})$ is the complex Gaussian distribution with mean a and covariance matrix A.

II. System Model

As shown in Fig. 1, the proposed system model in this paper is a cognitive radio network including *K* users. One of them is a PU, and *K-1* of them are SUs. PU consists of one source, one destination, and two half-duplex amplify and forward (AF) relays that act as SR techniques. Each SU consists of one transmitter and one receiver. They use the PU spectrum for transmission. Therefore, we have an interference network (IRI and inter-users interference). We apply the IA technique to eliminate interferences of this network. Therefore, we have to consider that all nodes are multi-antennas.

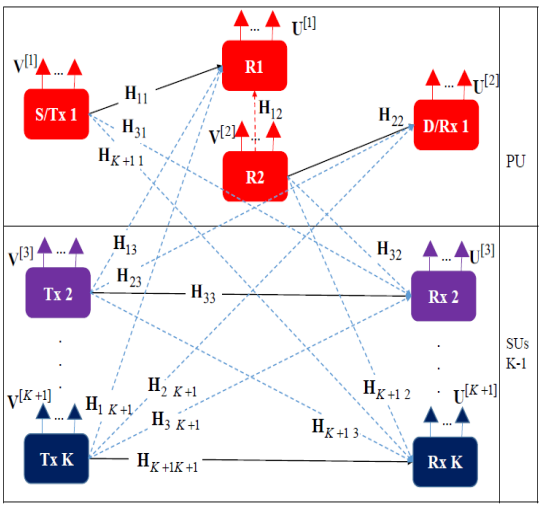


Fig. 1. IA-based CRN Model with 1 SR PU and K-1 SUs in Odd TS

In this section, we will describe the performance of the SR technique and the IA-based CRN model with SR PU. Then, the QoS requirement of the PU will be analyzed.

A. Successive Relaying Technique

As shown in Fig. 1, PU utilizes the SR technique. Relays in PU are shown as R1 and R2 in the figure. The SR performance for different time slots is as follows.

Here, the primary user sends L (L is even) messages from the source to the destination.

First Time Slot:

The source sends the first message to the R1 (relay in receive mode). R2 doesn't have any message to send in this time slot, so it's off.

Second-time Slot (Even Time Slots):

R2 is in receive mode during this time slot, and R1 is in transmit mode. Source and R1 send their massages simultaneously. Source sends a new message (second massage) to R2. Then, R1 amplifies the received message in the previous time slot (first massage) and then sends it to the destination. All transmissions are carried out at the same frequency. Thus, R2 receives interference from the R1. This Interference is known as inter relay interference (IRI).

Third Time Slot (Odd Time Slots):

R1 is in receive mode during this time slot, and R2 is in transmit mode. Source sends a new message (third massage) to the R1. R2 amplifies the received message in the previous time slot (second massage) and then transmits it to the destination. Thus, R1 receives interference from the R2 (red dotted line in Fig. 1).

This process continues until the L+1th time slot.

***L*+1th Time Slot (Last Time Slot):**

During this time slot, R2 sends the last message (the message received by R2 in the Lth time slot) to the destination, but the source doesn't have any message to send.

According to this description, L messages are sent in L + 1time slots. This scheme act as a full-duplex relay for a large number of L. Thus spectral efficiency is recovered compared to HD relays. In this paper, the first and the last time slots aren't considered, and the focus is on the even and odd time slot.

B. The Proposed Cognitive Radio Network

Cognitive radio is a network that PU shares its spectrum with the Sus. Hence, all the transmissions are at the same time and frequency. Thus, we have an interference network, especially when the PU utilizes the SR technique. Here, we use some antennas in all nodes to manage the interference. Accordingly, we propose a new cognitive radio network as below:

There are K users (one primary user and K-*1* secondary user), and all nodes are multi-antennas. All transmitters have N antennas, and receivers have M antennas. The relays have M antennas in receive mode, and they use N of M antennas when they're in transmit mode. We don't consider antenna selection in this paper.

Here, the channel coefficients are uncorrelated quasi-static flat fading [37]. The proposed network is expressed only in a one-time slot (odd time slot) due to the statistical similarity of the channel coefficients.

As shown in Fig. 1, $\mathbf{H}_{11} \in \mathbb{C}^{N\times M}$ and $\mathbf{H}_{22} \in \mathbb{C}^{N\times M}$ denote source-R1 and R1-destination channel coefficients. Further, $\mathbf{H}_{k+1\ k+1} \in \mathbb{C}^{N\times M}$ determines the channel coefficient from K+1th SU's transmitter to its receiver ($k \in \{2,...,K\}$). The interference channel coefficients of this network are shown in Table 1:

TABLE I

INTERFERENCE CHANNELS

Channel Coefficients	Channels
$\mathbf{H}_{21} \in \mathbb{C}^{N\times M}$	Channel between R1 and R2
$\mathbf{H}_{1\ k+1} \in \mathbb{C}^{N\times M}$	The channel between kth SU's transmitter and R1
$\mathbf{H}_{k+1\ 2} \in \mathbb{C}^{N\times M}$	Channel between R2 and kth SU's receiver
$\mathbf{H}_{2\ k+1} \in \mathbb{C}^{N\times M}$	The channel between kth SU's transmitter and PU's destination
$\mathbf{H}_{k+1\ j+1} \in \mathbb{C}^{N\times M}$	Channel between jth transmitter and kth Receiver of SUs

Where $k \neq j \in \{2,...,K\}$. Each of the coefficients entities is independent and identically distributed (i.i.d.), following $\mathcal{CN}$ (0, 1).

In this paper, the channel state information is entirely known in all nodes. Therefore, the linear IA technique eliminates all interferences (CRN and IRI).

In the n-1th time slot, the received signal at the Relay (R2), including d data streams, and the received signal at the PU's destination can be expressed as (1) and (2), respectively.

$$\begin{aligned}\mathbf{y}_R(n-1) = &\sqrt{p_t^{[S]}}\mathbf{U}^{\dagger[1]}(n-1)\mathbf{H}_{11}(n-1)\mathbf{V}^{[1]}(n-1)\mathbf{x}_S(n-1)\\ &+\sqrt{p_t^{[R]}}\mathbf{U}^{\dagger[1]}(n-1)\mathbf{H}_{21}(n-1)\mathbf{V}^{[2]}\mathbf{x}_R(n-1)+\\ &\sum_{k=2}^{K}\sqrt{p_t^{[k]}}\mathbf{U}^{\dagger[1]}(n-1)\mathbf{H}_{k+1\ 1}\mathbf{V}^{[k+1]}(n-1)\mathbf{x}_k(n-1)\\ &+\mathbf{U}^{\dagger[1]}(n-1)\mathbf{z}_R(n-1)\end{aligned} \quad (1)$$

$$\begin{aligned}\mathbf{y}_D(n) = &\sqrt{p_t^{[R]}}\mathbf{U}^{\dagger[2]}(n)\mathbf{H}_{22}(n)\mathbf{V}^{[2]}(n)\mathbf{x}_R(n)+\\ &\sum_{k=2}^{K}\sqrt{p_t^{[k]}}\mathbf{U}^{\dagger[2]}(n)\mathbf{H}_{k+1\ 2}(n)\mathbf{V}^{[k+1]}(n)\mathbf{x}_k(n)\\ &+\mathbf{U}^{[2]}(n)\mathbf{z}_D(n)\end{aligned} \quad (2)$$

Where $\mathbf{x}_R(n)$ is the amplified form of the received signal at the relay (R2) in the previous time slot ($\mathbf{x}_R(n) = \beta\mathbf{y}_R(n-1)$). In time slot n, the received signal at the kth SU's receiver can be denoted as:

$$\begin{aligned}\mathbf{y}_k(n) = &\sqrt{p_t^{[k]}}\mathbf{U}^{\dagger[k+1]}(n)\mathbf{H}_{k+1\ k+1}\mathbf{V}^{[k+1]}(n)\mathbf{x}_k(n)\\ &+\sqrt{p_t^{[R]}}\mathbf{U}^{\dagger[k+1]}(n)\mathbf{H}_{k+1\ 1}\mathbf{V}^{[2]}(n)\mathbf{x}_R\\ &+\sum_{\substack{j=2\\ j\neq k}}^{K}\sqrt{p_t^{[k]}}\mathbf{U}^{\dagger[k+1]}(n)\mathbf{H}_{k+1\ j+1}\mathbf{V}^{[j+1]}(n)\mathbf{x}_j\\ &+\mathbf{U}^{\dagger[k+1]}(n)\mathbf{z}_k(n); k=2,...,K\end{aligned} \quad (3)$$

Where, $\mathbf{U}^{[r]}, \mathbf{V}^{[r]}, r \in \{1,...,K+1\}$ are the unitary $M \times d$ interference suppression matrix in receivers and $N \times d$ precoding matrix in transmitters, respectively.

$\mathbf{x}_a, a \in \{k \in \{2,...,K\}, S, R\}$ includes d data streams of the SU's transmitter, the relay, and source of PU with the power of $\mathrm{E}\left[\|\mathbf{x}_a\|^2\right] = p_t^{[a]}$. $\mathbf{z}_b \in \mathbb{C}^{N\times 1}, b \in \{k \in \{2,...,K\}, d, R\}$ is an additive white Gaussian noise (AWGN) vector with distribution $\mathcal{CN}(0, \sigma^2\mathbf{I}_N)$ at the kth receiver, where σ^2 is the noise power at each antenna receiver's antenna.

When channel state information is available in all nodes, the interferences (inter-users interference and IRI) can be eliminated if the following conditions are met:

$$\mathrm{Rank}[\mathbf{U}^{\dagger[r]}\mathbf{H}_{r\,r}\mathbf{V}^{[r]}] = d; r = 1,...,K+1 \tag{4}$$

$$\mathbf{U}^{\dagger[r]}\mathbf{H}_{r\,t}\mathbf{V}^{[t]} = 0 \;; \forall r \neq t \;;\; t, r = 1,...,K+1 \tag{5}$$

Therefore, desired signals are received among the full rank channel $\mathbf{U}^{\dagger[r]}\mathbf{H}_{r\,r}\mathbf{V}^{[r]} = \bar{\mathbf{H}}_r \in \mathbb{C}^{d\times d}$. Also, we assume that the minimizing interference leakage (MinIL) Algorithm (39) is adopted to calculate the solutions of IA. Hence the received signals in (1), (2), and (3) can be rewritten as:

$$\mathbf{y}_R(n-1) = \sqrt{p_t^{[s]}}\bar{\mathbf{H}}_1\mathbf{x}_S + \bar{\mathbf{z}}_R \tag{6}$$

$$\mathbf{y}_D(n) = \sqrt{p_t^{[R]}}\bar{\mathbf{H}}_2\mathbf{x}_R + \bar{\mathbf{z}}_D \tag{7}$$

$$\mathbf{y}_k(n) = \sqrt{p_t^{[k]}}\bar{\mathbf{H}}_{k+1}\mathbf{x}_k + \bar{\mathbf{z}}_k(n); k = 2,...,K \tag{8}$$

Where $\bar{\mathbf{z}}_a = \mathbf{U}^{[r]}\mathbf{z}_a$ is AWGN with $CN(0, \sigma^2\mathbf{I}_d)$ distribution.

Since this paper mainly concentrates on power allocation and doesn't consider the degrees of freedom, the transmitter sends one data stream (d = 1). Hence, the number of users can be present in this network should follow:

$$K < M + N - 2 \tag{9}$$

When conditions (4) and (5) are met, the IA technique eliminates interferences. Due to d = 1, the signals are received among the full rank channel $h_r = \mathbf{u}^{\dagger[r]}\mathbf{H}_{r\,r}\mathbf{v}^{[r]} = \bar{\mathbf{H}}_r \in \mathbb{C}^{1\times 1}$. Consequently, the amplifying factor at the relay and the signal to noise ratio (SNR) of the PU and SUs can be obtained as (10), (11), and (12), respectively:

$$\beta = \frac{1}{\sqrt{|\,y_R(n-1)\,|^2}} = \frac{1}{\sqrt{p_t^{[S]}\,|\,h_1\,|^2 + \sigma^2}} \tag{10}$$

$$SNR^{[1]} = \frac{p_t^{[R]}p_t^{[S]}\,|\,h_2\,|^2\,|\,h_1\,|^2\,d_{RD}^{-\alpha}d_{SR}^{-\alpha}}{p_t^{[R]}\,|\,h_2\,|^2\,d_{RD}^{-\alpha}\sigma^2 + p_t^{[S]}\,|\,h_1\,|^2\,d_{SR}^{-\alpha}\sigma^2 + \sigma^4} \tag{11}$$

$$SNR^{[k]} = \frac{|\,h_{k+1}\,|^2\,d_k^{-\alpha}P_t^{[k]}}{\sigma^2}, k = 2,...,K \tag{12}$$

Where d_{RD}, d_{SR} are the distances between relay-destination and source-relay, respectively. d_k is the distance between the transmitter and receiver of the kth user and α is the channel attenuation factor. $h_r = \mathbf{u}^{\dagger[r]}\mathbf{H}_{r\,r}\mathbf{v}^{[r]}$ where $\mathbf{H}_{r\,r}$ is i.i.d with $\mathcal{CN}(0, 1)$ distribution. $\mathbf{u}$,$\mathbf{v}$ are the unitary vectors that they're independent of $\mathbf{H}_{r\,r}$. Therefore, h_r is i.i.d with $\mathcal{CN}(0, 1)$ distribution, too [38, Appendix E].

C. Primary User's Quality of Service

In an underlay spectrum sharing CR network, SUs cannot be present unless the interferences from SUs don't decrease the PU's performance. Consequently, the power of noise and interference in the primary user must be low. If IA is applied, interferences can all be eliminated. Therefore, IA can provide a convenient spectrum sharing, which the interference need not be considered any longer. But, the SINR of PU is reduced compared to MIMO PU without IA and SUs [11], and it doesn't guarantee the QoS of PU. A threshold rate for the PU () is defined to satisfy QoS to tackle this challenge. Accordingly, the minimum required transmitted power of the PU should follow:

$$R_{th}^{[1]} < \log_2\left(1 + SNR^{[1]}\right) \tag{13}$$

In the IA-based CR network, the SUs should try to satisfy the QoS requirement of the PU defined in (13) unless they will not be allowed to access the licensed spectrum.

III. Power Allocation Algorithms in the Cognitive Radio Network

In most previous works in this field, equal transmitted power is allocated to each user. However, this may not be an appropriate power allocation. In this section, an optimum power allocation among users is applied, under the condition that the sum transmitted power of the users should be lower than $p_{\max}$.

Here, the minimum transmitted power of the PU (power of the source and the relay) to guarantee the threshold rate is first presented. Then two power allocation algorithms are proposed.

A. Minimal Power of PU to Guarantee its QoS Requirement

In the proposed network, when the PA among users is considered, the threshold rate of the PU should be satisfied. In this section, the minimum required transmitted power of the relay and the source are obtained to ensure the PU's QoS. We assume the power of the source and relay are equal ($p_t^{[S]} = p_t^{[R]} = p_t^{[1]}$). Then, the minimum required power for the relay and source are obtained as (14) while solving (13).

$$p_t^{[S]} = p_t^{[R]} = p_t^{[1]} \geq p_{t\_\min} = \sigma^2 \frac{((2^{R_{th}^{[1]}}-1)(|\,h_2\,|^2\,d_{RD}^{-\alpha} + |\,h_1\,|^2\,d_{SR}^{-\alpha})) + \sqrt{(2^{R_{th}^{[1]}}-1)^2(|\,h_2\,|^2\,d_{RD}^{-\alpha} + |\,h_1\,|^2\,d_{SR}^{-\alpha})^2 + 4(2^{R_{th}^{[1]}}-1)\,|\,h_2\,|^2\,|\,h_1\,|^2\,d_{RD}^{-\alpha}d_{SR}^{-\alpha}}}{2\,|\,h_2\,|^2\,|\,h_1\,|^2\,d_{RD}^{-\alpha}d_{SR}^{-\alpha}} \tag{14}$$

In this paper, the power allocation algorithms mainly depend on two constraints. The first constraint is the maximum sum of transmitted powers. The second one is the minimal required transmitted power of the relay and the source. According to these constraints, two cases are introduced to assign the power to the PU and SUs. Thus $p_t^{[1]}$ follows (15).

$$\begin{cases} p_t^{[1]} = p_{\max}/2 & ;If\ \ 2p_{t\_\min} > p_{\max} \\ p_{t\_\min} \le p_t^{[1]} \le p_{\max}/2 & ;If\ \ 2p_{t\_\min} \le p_{\max} \end{cases} \tag{15}$$

According to (15), the cases are explained as follows:

$2p_{t\_\min} > p_{\max}$: It means the PU's power requirement cannot be satisfied. Therefore, the maximum power ($p_{\max}$) is allocated to the PU's relay and source to increase the rate. In this case, SUs cannot use the PU's spectrum, and they're off.

$2p_{t\_\min} \le p_{\max}$: It means the minimum required transmitted power of the source, relay, and then $R_{th}^{[1]}$ can be satisfied. Therefore, according to the power allocation, the transmitted power of the source and the transmitted power of the relay is in $\left[p_{t\_\min}, p_{\max}/2 \right]$ range. In this case, PU and probably SUs are active. The SUs should be off if there is no power to allocate them ($2p_{t\_\min} = p_{\max}$).

Algorithm 1 (SRPA)

1: $p_{t\_\min}$ is calculated according to (14).
2: if $2p_{t\_\min} < p_{\max}$, then
3: $p_t^{[1]} = p_{t\_\min}$.
4: $p_{\max} - 2p_{t\_\min}$ is allocated to SUs by (17) and (18)
5: else
6: Allocate $p_{\max}$ to the PU.
7: SUs are switched into sleep mode.
8: end if
9: Transmission for duration T with the power allocated.
10: The time slot ends.

B. Power Allocation Algorithm for Maximizing SU's Sum Rate

In the spectrum trading-based CRN, the income of PUs is proportional to the sum rate of SUs they provided. Besides, when multiple PUs sell spectrum to multiple SUs, SUs can adapt their behavior by observing the variations in price and quality of spectrum offered by these PUs [38]. In this section, the sum rate of the SUs is maximized while considering the PU's $R_{th}^{[1]}$ constraint. As a result, a power allocation algorithm is proposed as follows:

$$\max_{p_t^{[1]},\ldots,p_t^{[K]}} EE = \max \frac{\log_2(1+SNR^{[k]}) + \sum_{k=2}^{K} \log_2(1 + p_t^{[k]} \frac{d_k^{-\alpha} |h_{k+1}|^2}{\sigma^2})}{(K+1)(p_{ct}+p_{cr}) + 2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]}}$$
$$p_t^{[k]} \ge 0, p_t^{[1]} \ge p_{t\_\min}, k = 2,\ldots,K$$
$$2 \times p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} \le p_{\max} \tag{19}$$

If $2p_{t-\min} \le p_{\max}$ the required transmitted power of the source and the relay ($p_{t\_\min}$) is assigned to them. Next, the residual power is allocated to SUs to maximize their sum rate or spectral efficiency. The power allocation problem is denoted as:

$$\begin{cases} \max\limits_{p_t^{[2]},p_t^{[3]},\ldots,p_t^{[k]}} \sum\limits_{k=2}^{K} \log_2\left(1 + d_k^{-\alpha} |h_{k+1}|^2 \frac{p_t^{[k]}}{\sigma^2}\right) \\ s.t. \quad p_t^{[k]} \ge 0, \forall k = 2,\ldots,K \\ \sum\limits_{2}^{K} p_t^{[k]} = p_{\max} - 2p_{t\_\min} \end{cases} \tag{16}$$

which is similar to the PA problem in multiple parallel channels. Therefore, the water-filling PA method is exploited, and the closed-form solution for the optimal transmitted power of the SUs is obtained as (17).

$$p_t^{*[k]} = (\upsilon - \frac{\sigma^2}{d_k^{-\alpha} |h_{k+1}|^2})^+ \tag{17}$$

Where, $x^+ = \max(x,0)$ and υ should satisfy (18).

$$\sum_{k=2}^{K} (\upsilon - \frac{\sigma^2}{d_k^{-\alpha} |h_{k+1}|^2})^+ = p_{\max} - 2p_{t\_\min} \tag{18}$$

The closed-form solution of (15) that is expressed as (17) and (18) is easy to obtain. Hence, the computational complexity of the proposed algorithm (Algorithm 1) is reduced.

C. Power Allocation Algorithm for Maximizing EE of Network

In the future, since the number of devices is increased, energy management is critical to prevent economic and environmental problems. Power allocation is one of the effective ways to increase EE. The EE of the network can be defined as the transmitted information per unit frequency per Joule energy consumption (bits/Hz/Joule).

This section studies the EE of the CRN using the SR PU. Here, we want to show that the EE of the proposed network is increased, although relays add extra circuit power in each time slot. We propose a power allocation problem to maximize the EE. The problem is denoted as (19).

Where p_{ct}, p_{cr} is the circuit power of the transmitters and the receivers (the source and the relay in transmit mode is also included).

Problem (19) is complex because it's concave-convex fractional programming [40]. When $2p_{t-\min} \le p_{\max}$ (19) has an optimal solution. Part 1 and part 2 are first provided to obtain the closed-form solution of (19).

Part 1:

If the summation of the user's transmitted power is equal to $p_{\max}$ ($2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} = p_{\max}$), the fraction's denominator in (19) is constant. Therefore, we can optimize the fraction's numerator instead of the total fraction. The denominator $SNR^{[1]}$ in (19) includes the optimization parameter, and it's difficult to obtain the closed-form solution. To simplify the problem, an appropriate approximation is utilized in the form of $SNR^{[1]}$ as follows.

We can divide the numerator and denominator of $SNR^{[1]}$ into $p_t^{[1]}$. Since $p_t^{[R]} = p_t^{[S]} = p_t^{[1]}$ we can rewrite $SNR^{[1]}$ as (20).

$$SNR^{[1]} = \frac{p_t^{[1]} |h_2|^2 |h_1|^2 d_{SR}^{-\alpha} d_{RD}^{-\alpha}}{|h_2|^2 d_{RD}^{-\alpha}\sigma^2 + |h_1|^2 d_{SR}^{-\alpha}\sigma^2 + \frac{\sigma^4}{p_t^{[1]}}} \tag{20}$$

In $\sigma^4 / p_t^{[1]}$, $p_t^{[1]}$ is approximated by $p_{t-\min}$, so we can represent $SNR^{[1]} = p_t^{[1]}\gamma$ that γ is expressed as

$$\gamma = \frac{|h_2|^2 |h_1|^2 d_{SR}^{-\alpha} d_{RD}^{-\alpha}}{|h_2|^2 d_{RD}^{-\alpha}\sigma^2 + |h_1|^2 d_{SR}^{-\alpha}\sigma^2 + \frac{\sigma^4}{p_{t-\min}}} \tag{21}$$

Accordingly, the EE problem is rewritten as (22).

$$\begin{aligned} \max_{p_t^{[1]},\ldots,p_t^{[K]}} \quad & \log_2\left(1 + p_t^{[1]}\gamma\right) + \sum_{2}^{K} \log_2\left(1 + p_t^{[k]} \frac{d_k^{-\alpha} |h_{k+1}|^2}{\sigma^2}\right) \\ & p_t^{[1]} \ge p_{t\_\min}, p_t^{[k]} \ge 0 \quad k = 2,\ldots,K \\ & 2 \times p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} \le p_{\max} \end{aligned} \tag{22}$$

The closed-form solution of (22) is calculated as (23).

$$\begin{aligned} p_t^{*[1]} &= \left(\frac{1}{2\upsilon \ln 2} - \frac{1}{\hat{\gamma}}\right)^+ + p_{t\_\min} \\ p_t^{*[k]} &= \left(\frac{1}{\upsilon \ln 2} - \frac{\sigma^2}{d_k^{-\alpha} |h_{k+1}|}\right)^+, k = 2,\ldots,K \end{aligned} \tag{23}$$

where υ should be satisfied

$$2\left(\left(\frac{1}{2\upsilon \ln 2} - \frac{1}{\hat{\gamma}}\right)^+ + p_{t\_\min}\right) + \sum_{k=2}^{K}\left(\frac{1}{\upsilon \ln 2} - \frac{\sigma^2}{d_k^{-\alpha} |h_{k+1}|}\right)^+ \le p_{\max} \tag{24}$$

And $\hat{\gamma}$ is denoted as

$$\hat{\gamma} = \frac{\gamma}{1 + P_{t\_\min}} \tag{25}$$

Proof: **See Appendix A**

In part1, When $2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} = p_{\max}$ the solution of (19) is different from the solution of (16). This is because in (16), the transmitted power of relay and source is $p_{t\_\min}$ while it can be more than $p_{t\_\min}$ in (19).

When SNR is low, $2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} = p_{\max}$ it can be satisfied after optimization (16). Thus, part1 can be appropriate to obtain the optimum solution. However, when SNR becomes higher, $2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]}$ it can be smaller than $p_{\max}$ to maximize the network's EE. Thus, the water-filling strategy is not suitable. We will obtain the optimal solution of (19) by fractional programming as in part 2 and theorem 1.

Part 2:

In this part, we solve problem (19) by the fractional programming method [40]. Accordingly, we should optimize (26) instead of (19).

$$\begin{aligned} f(\lambda) = \max_{p_t^{[1]},\ldots,p_t^{[K]}} & \log_2(1 + SNR^{[1]}) + \sum_{k=2}^{K} \log_2(1 + P_t^{[k]} \frac{d_k^{-\alpha} |h_{k+1}|^2}{\sigma^2}) \\ & - \lambda\left(\left(K+1\right)\left(p_{ct} + p_{cr}\right) + 2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]}\right) \\ & p_t^{[k]} \ge 0, p_t^{[1]} \ge p_{t\_\min}, k = 2,\ldots,K \\ & 2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} \le p_{\max} \end{aligned} \tag{26}$$

In (26), we consider $SNR^{[1]} = p_t^{[1]}\gamma$ like part 1. The optimum solution of (26) is obtained as (28) by KKT conditions. But $2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} \le p_{\max}$ constraint is still present.

$$\begin{aligned} p_t^{*[1]} &= \max\left\{\frac{1}{2\lambda \ln 2} - \frac{1}{\gamma}, p_{t\_\min}\right\} \\ p_t^{*[k]} &= \max\left\{\frac{1}{\lambda \ln 2} - \frac{\sigma^2}{d_k^{-\alpha} |h_{k+1}|}, 0\right\}, k = 2,\ldots,K \end{aligned} \tag{27}$$

The optimum is obtained by substituting (27) in (26) and solving the equation. When $p_t^{*[k]}$ ($k = 2,\ldots,K$) is all expressed as $\frac{1}{\lambda \ln 2} - \frac{\sigma^2}{d_k^{-\alpha} |h_{k+1}|}$ and $p_t^{*[1]}$ is expressed as

$$\lambda^* = \frac{K\varphi\left(\frac{\ln 2}{eK}\left(\frac{\gamma}{2\ln 2}\prod_{k=2}^{K}\frac{d_k^{-\alpha}|h_{k+1}|^2}{\sigma^2 \ln 2}\right)^{\frac{1}{K}}\left(\left(K+1\right)\left(p_{ct}+p_{cr}\right) - \frac{2}{\gamma} - \sum_{k=2}^{K}\frac{\sigma^2}{d_k^{-\alpha}|h_{k+1}|^2}\right)\right)}{\left(\left(K+1\right)\left(p_{ct}+p_{cr}\right) - \frac{2}{\gamma} - \sum_{k=2}^{K}\frac{\sigma^2}{d_k^{-\alpha}|h_{k+1}|^2}\right)\ln 2} \tag{28}$$

$\frac{1}{2\lambda \ln 2} - \frac{1}{\gamma}$ the equation $f(\lambda) = 0$ is difficult to solve and λ^* can be calculated as (28). When the sum of all $p_t^{*[k]}$ is 0, and $p_t^{*[1]} = p_{t\_\min}$ in (27), $f(\lambda) = 0$ it is easier to solve.

Theorem 1:

In the end, we should apply the constraint $2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} \le p_{\max}$; therefore, according to part 1 and part 2, the closed-form solution of (19) can be discussed as:

1. If $2p_t^{*[1]} + \sum_{k=2}^{K} p_t^{*[k]} < p_{\max}$, the closed-form solution of (19) can be defined as (28) because the constraint is always satisfied.
2. If $2p_t^{*[1]} + \sum_{k=2}^{K} p_t^{*[k]} \ge p_{\max}$, $2p_{t\_\min} \le p_{\max}$, the Closed-form solution of (19) can be defined as (22).
3. If $2p_{t\_\min} \ge p_{\max}$, all of the constraints in (19) cannot be satisfied simultaneously thus, the problem doesn't have any solutions. In this case, the total power is allocated to the source and relay of the PU.

The maximizing EE power allocation problem is represented as Algorithm 2.

IV. SIMULATION

In this section, the results are illustrated. Each transmitter sends 1 data stream to its corresponding receiver. The Rayleigh block fading is adopted, and the perfect CSI is available at all nodes. The distance between the nodes is defined as $d_{RD} = d_{SR} = 0.5, d_{SD} = d_k = 1$ except in Fig.2) and the attenuation factor is 6.

Here, we will show the performance of the SR technique in the PU and the result of the proposed power allocations algorithms. We used the following conventions to label the curves in the plots: ***SRPU*** indicates the successive relay primary user. In ***Half-duplex*** mode, the PU uses just one half-duplex relay. ***Simple PU*** implies the primary user with only one transmitter and receiver (without relay). ***EEPA*** denotes the optimum power allocation for maximizing the EE of the network. ***SRPA*** and ***Eq PA*** indicate the optimum power allocation for maximizing the sum rate of SUs and the equal power allocation considering the PU's QoS, respectively.

At first, we want to show the effect of the SR technique in PU in the absence of the SUs when IRI still exists and all nodes have a single antenna. According to the previous sentence, (1) and (2), the SINR in the destination can be obtained as (29).

Where h_{12} and $d_{R_1R_2}$ are the IRI channel coefficient and the distance between the relays, respectively. We also consider μ as an IRI suppression coefficient (by any means) to show the changes by IRI values. Thus, we plot Fig. 2.

We assume that $p_t^{[s]} = p_t^{[r]}$, $d_{SR_1} = d_{R_1D} = 1/\sqrt{2}$, $d_{R_1R_2} = 1$ and $d_{SD} = 1$. In Fig. 2, the rate of the PU is plotted according to (29) for various μ . $\mu = 0$ indicates that IRI entirely exists, $\mu = 0.9$ means 0.9 of IRI is eliminated and $\mu = 1$ means IRI is wholly eliminated.

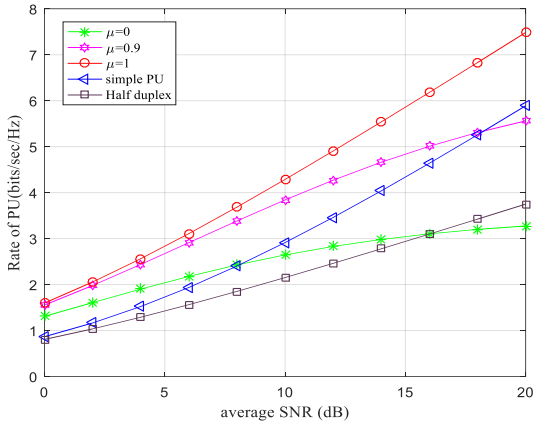


Fig. 2. Rate of PU in *Half-duplex* Mode, *Simple PU,* and *SRPU* for Various μ

Also, the rate of the PU in *simple PU* and *Half-duplex* mode is plotted. We can see that rate of the PU is upper than *Half-duplex* mode and *Simple PU* even when IRI is completely existed (in low average SNR). Moreover, the rate is more increased when IRI suppression is employed. Thus, it's better to utilize the IA technique to eliminate the IRI, especially when SUs are present.

Algorithm 2 (EEPA)

1: $p_{t\_\min}$ is calculated according to (14).
2: **if** $2p_{t\_\min} < p_{\max}$, **then**
3: Solve the EE power allocation problem in (19) through fractional programming according to **theorem 1**.
4: **else**
5: Allocate $p_{\max}$ to the PU.
6: SUs are switched into sleep mode.
7: **end if**
8: Transmission for duration T with the power allocated.
9: The time slot ends.

In Fig. 3, we compare the minimum required transmitted power of the PU to satisfy its Qos in *SRPU* and *Simple PU*. We assume and compare their PU's minimum required transmitted power and their rates in both cases. $2p_{t\_\min}/\sigma^2$ is plotted by considering (15) for the *SRPU* while $p_{t\_\min}/\sigma^2$ ($p_{t\_\min}$ is obtained in [38]) is plotted for the *Simple PU*. In both cases, the minimum required transmitted power varies dramatically over 200 time slots to guarantee the threshold rate of the PU ($R_{th}^{[1]}$). As shown in Fig.3b $R_{th}^{[1]}$ is achieved in both cases, but, as shown in Fig.3a in the *SRPU*, lower power is required ($2p_{t\_\min}$) compared to the *Simple PU*. In the *SRPU*, the minimum required power is lower than $p_{\max}$ most time slots. Thus, the SUs can use the PU's spectrum in most time slots. As shown in Fig. 3a, the largest value of $p_{t\_\min}$ is 1000 times more than its smallest value.

Consequently, if equal transmitted power is allocated to the users, the transmitted power of the PU may be much greater than its required power.

$$SINR^{[1]} = \frac{p_t^{[R]} p_t^{[S]} |h_{22}|^2 |h_{11}|^2 d_{R_1D}^{-\alpha} d_{SR_1}^{-\alpha}}{(\sigma^2 + p_t^{[R]} |h_{22}|^2 d_{R_1D}^{-\alpha}) + (\sigma^2 + (1-\mu) p_t^{[R]} |h_{12}|^2 d_{R_1R_2}^{-\alpha}) + p_t^{[S]} |h_{11}|^2 d_{SR_1}^{-\alpha} \sigma^2} \quad (29)$$

The opposite of this situation may also happen when the power of the PU is smaller than its required power. Therefore, both cases decrease the performance of the SUs and PU, respectively. Optimum power allocation can increase the performance of the SUs while guaranteeing the QoS of the PU.

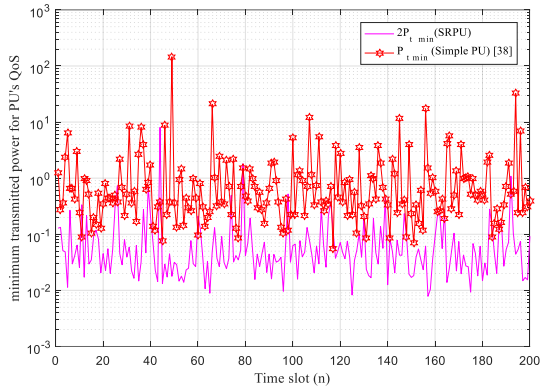


Fig. 3a. Minimal Transmitted Power of the PU to Guarantee the PU's QoS over 200 Time Slots

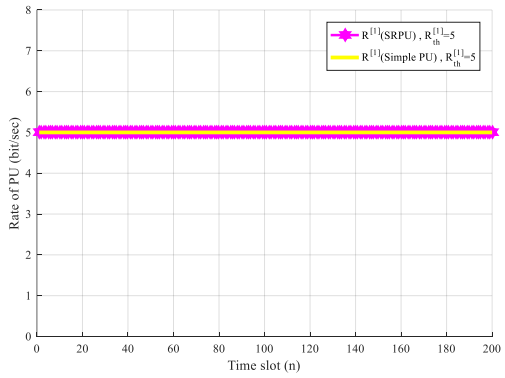


Fig. 3b. Rate of the PU over 200 Time Slots with Threshold Rate $R_{th}^{[1]} = 5$ bits/s/Hz

In Fig. 4 and Fig. 5, we assume $p_{\max} = K \times 0.1w$ and $K=3$.

In Fig. 4, the effect of the SR technique in the sum rate of the SUs is shown. In the *Simple PU*, the minimum required transmitted power of the PU is more than or equal to the network's total power in low SNR ($p_{t\_\min} \geq p_{\max}$), including 0-5 dB. Thus, $p_{\max}$ is only allocated to the PU, then the SUs must be off. As shown in Fig. 4, the sum rate of the SUs is equal to zero. When the SR technique is employed in the PU, in most of the time slots, $2p_{t\_\min}$ is lower than $p_{\max}$. Therefore, more power is allocated to the Sus and their sum-rate increases. In high SNR, due to the excellent channel conditions, the required powers of the PU in the *simple PU* and *SRPU* are very low and almost equal. Thus the sum rates of SUs in both cases are equal.

In Fig. 6 and Fig. 7, we assume that $p_{ct} = p_{cr} = 0.05W$.

Fig. 6 shows the EE of the network for *SRPU* and *Simple PU* in various $p_{\max}$. As shown in Fig. 6, by increasing $p_{\max}$ the EE of the network is decreased because of the Logarithmic property. In poor channel conditions, $p_{\max}$ is all allocated to the users, so the EE of the network is low. The EE of the network becomes higher by the SR PU in low SNR. This is because the relay cooperates with the source for transmission, and lower power is required. For example, in constant $p_{\max}$ and high SNR, SNR=30 dB, a little power $p_{\max}$ is allocated to the users. Accordingly, the total transmitted power of the users in both the *SRPU* and *Simple PU* is low and almost equal. Thus, the total circuit power consumption of the devices is determinative to compare the cases, especially in low $p_{\max}$. Due to using the relay in the PU, the circuit power consumption is increased. So in this situation, Simple PU (not using relay) is better, and the EE is higher. As shown in Fig. 6, in higher $p_{\max}$, the SR technique performs better in more regions of the SNR.

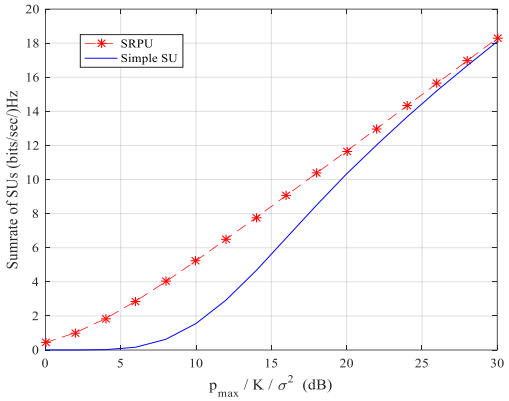


Fig. 4. The Sum Rate of SUs in the *SRPU* and *Simple PU*

Fig. 5 shows the performance of the proposed power allocation algorithm to maximize the sum rate of the SUs (Algorithm 1). The proposed power allocation improves the sum rate of SUs while satisfying the PU's QoS. The PU's QoS assurance is also considered in our simulation when equal power allocation is applied.

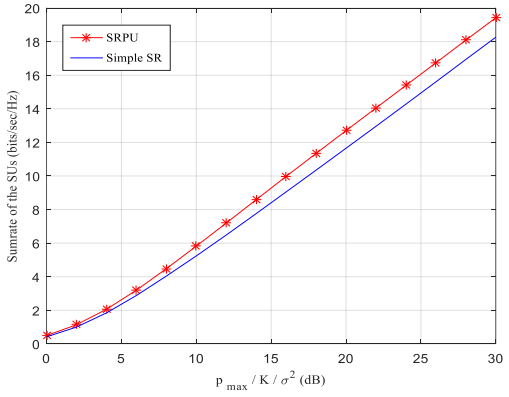


Fig. 5. The Sum Rate of SUs in *Opt SUPA* and *Eq PA* for the *SRPU*

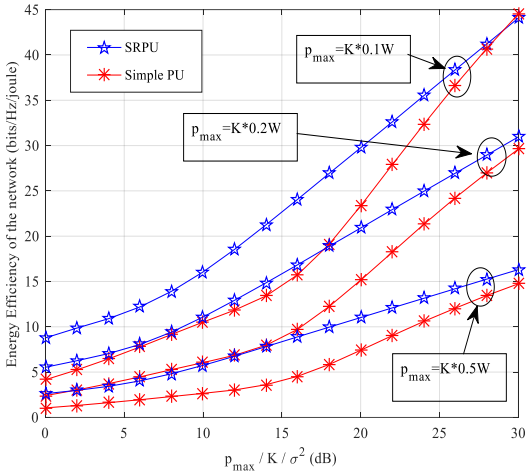


Fig. 6. The EE of the Network in *SRPU* and *Simple SR* for Various $p_{\max}$

Fig. 7 shows the EE power allocation algorithm (algorithm 2) for $p_{\max}=1\ W, p_{\max}=0.4W\ \ K=2$. As shown in the figure, the energy efficiency of the network is increased in high SNR. Also, in high, $p_{\max}$ the effect of power allocation is more than the lower one. In Fig. 7, we plot the EE power allocation problem (19) to indicate the result without approximation.

Therefore, we can see that the closed-form solution (obtained with the approximation) equals the numerical result. As shown in the figure, the approximation is appropriate in high SNR. But in lower SNR, including 10 dB, numerical and closed-form results are slightly different. If equal powers (PU's QoS is also considered) are assigned to the users, $p_{\max}$ is all consumed in the network, and EE is reduced.

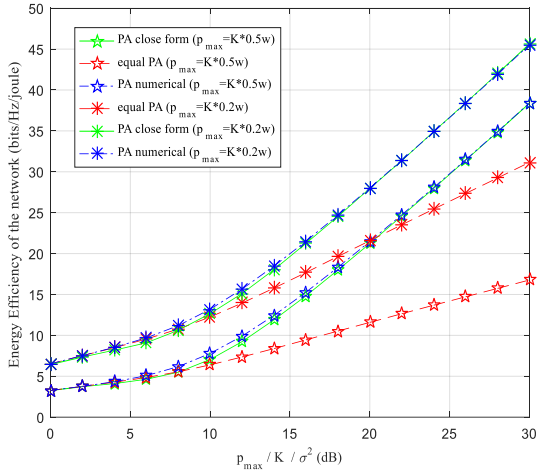


Fig. 7. The EE of the Network in *Eq PA*, *EEPA,* and Numerical Power Allocation for *SRPU*

In Fig. 8 and Fig. 9, we assume $R_{th}=7\ p_{\max}=1W$ the number of users is 3, and the PU has employed the SR technique. In these figures, we plot the energy efficiency of the network and the sum rate of the SUs by two proposed algorithms. Fig. 8 indicates that the sum rate of Sus in the *SUPA* algorithm is upper than the sum rate of Sus in the *EEPA* algorithm. This is because the transmitted power of users has decreased to increase energy efficiency in *EEPA*. Accordingly, Fig. 9 shows that the energy efficiency of the network in the *EEPA* algorithm is upper than the energy efficiency of the network in the *SUPA*. But, the comparison of both figures shows that by serving the EEPA algorithm, we will significantly improve energy efficiency while ignoring the small amount of SU's sum rate.

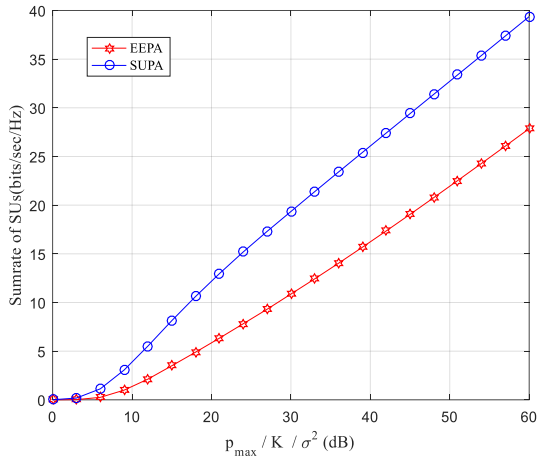


Fig. 8. Sum Rate of SUs by *EEPA* and *SUPA* Algorithms

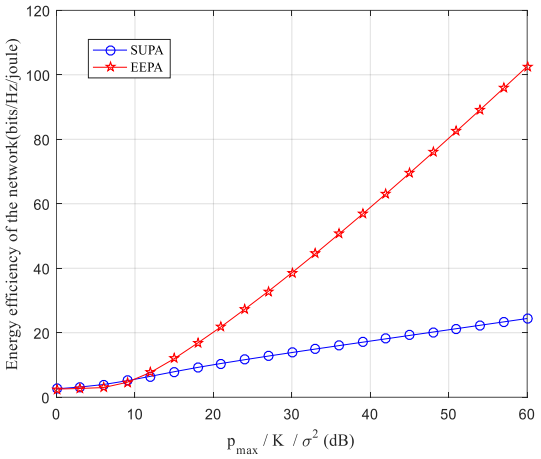


Fig. 9. Energy Efficiency of the Network by *EEPA* and *SUPA* Algorithms

V. Conclusions

This paper proposes a cooperative cognitive radio network that the primary user utilizes the SR technique. The IA method has been adopted to eliminate the interference (IRI in the SR and CRN interference). Then, a threshold rate has been defined for the primary user to satisfy its QoS. Accordingly, we have obtained the minimum required transmitted power for the source and the relay to satisfy the QoS of the PU. We have proposed two power allocation problems to increase the secondary users' sum rate and the network's energy efficiency. We have derived the closed-form solution for these problems. Next, we have expressed them as two algorithms. Finally, we have illustrated the obtained results. We have shown the improvement of the proposed network performance. Besides, we have shown that the SR technique in the PU has decreased the minimum transmitted power of PU. Consequently, the sum rate of secondary users and energy efficiency of the network has been increased. Also, we have proved that using the power allocations algorithms has improved the sum rate of secondary users and the energy efficiency of the network.

APPENDIX

Proof: when $2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} = p_{\max}$ we have

$$2P_c + 2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} + P_c = 2P_c + \sum_{k=2}^{K} P_c + p_{\max} = cte\,.$$

We also define $\hat{p}_t^{[1]} = p_t^{[1]} - p_{t\_\min}$ so we can rewrite (19) as (22) and then (30).

$$\begin{aligned}
&\max_{p_t^{[1]},\ldots,p_t^{[k]}} \log_2\left[\left(1+p_{t_\min}\right)\left(1+\hat{p}_t^{[1]}\hat{\gamma}\right)\right] + \sum_{2}^{K}\log_2\left(1+p_t^{[k]}\frac{|h_k|^2}{\sigma^2}\right) = \\
&\log_2\left(1+p_{t_\min}\right) + \log_2\left(1+\hat{p}_t^{[1]}\hat{\gamma}\right) + \sum_{2}^{K}\log_2\left(1+p_t^{[k]}\frac{|h_k|^2}{\sigma^2}\right) \\
&= \log_2\left(1+\hat{p}_t^{[1]}\hat{\gamma}\right) + \sum_{2}^{K}\log_2\left(1+p_t^{[k]}\frac{|h_k|^2}{\sigma^2}\right) \\
&st. \quad \hat{p}_t^{[1]} \ge 0, p_t^{[k]} \ge 0 \quad k = 2,\ldots,K \\
&\quad 2\hat{p}_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} \le p_{\max} - 2p_{t_\min}
\end{aligned} \tag{30}$$

Where $\hat{\gamma}$ is expressed as (25). The optimization problem in (29) is similar to the power allocation problem in multiple parallel channels. As a result, the water-filling method can be leveraged to obtain the closed-form optimal solution. Thus, when $2p_t^{[1]} + \sum_{k=2}^{K} p_t^{[k]} = p_{\max}$ the closed-form solution of (22) is denoted as (23) and υ should satisfy (24).

REFERENCES

[1] X. Hong, J. Wang, C.-X. Wang, and J. Shi, "*Cognitive radio in 5G: a perspective on energy-spectral efficiency trade-off*," *IEEE Communications Magazine*, vol. 52, pp. 46-53, 2014.

[2] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "*Full-duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection*," *IEEE Communications Magazine*, vol. 53, pp. 128-137, 2015.

[3] S. Haykin, "*Cognitive radio: brain-empowered wireless communications*," *IEEE Journal on selected areas in communications,* vol. 23, pp. 201-220, 2005.

[4] T. Luan, F. Gao, and X.-D. Zhang, "*Joint resource scheduling for relay-assisted broadband cognitive radio networks*," *IEEE Transactions on Wireless Communications*, vol. 11, pp. 3090-3100, 2012.

[5] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "*Breaking spectrum gridlock with cognitive radios: An information-theoretic perspective," proc. IEEE,* vol. 97, pp. 894-914, 2009.

[6] J. T. MacDonald and D. R. Ucci, "*Interference temperature limits of IEEE 802.11 protocol radio channels*," in *Electro/Information Technology, 2007 IEEE International Conference* on, pp. 64-69, 2007.

[7] B. Wang and K. R. Liu, "*Advances in cognitive radio networks: A survey,*" *IEEE Journal of selected topics in signal processing*, vol. 5, pp. 5-23, 2011.

[8] SJ. Kim, GB. Giannakis. "*Optimal resource allocation for MIMO ad hoc cognitive radio networks*." *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3117-3131, 2011.

[9] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," IEEE Transactions on Information Theory, vol. 54, pp. 3457-3470, 2008.

[10] N. Zhao, F. R. Yu, M. Jin, Q. Yan, and V. C. Leung, "Interference alignment and its applications: A survey, research issues, and challenges," IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1779-1803, 2016.

[11] N. Zhao, F. R. Yu, and H. Sun, "*Interference alignment with delayed channel state information and dynamic AR-model channel prediction in wireless networks*" in *Wireless Networks*, vol. 21, no. 4, pp. 1779--1803, 2015.

[12] J. Tang, S. Lambotharan, and S. Pomeroy, "Interference cancellation and alignment techniques for multiple-input and multiple-output cognitive relay networks," IET Signal Process, vol. 7, no. 3, pp. 188–200, May 2013.

[13] S. Arzykulov, G. Nauryzbayev, T. A. Tsiftsis, and M. Abdallah, "On the Performance of Wireless Powered Cognitive Relay Network with Interference Alignment," IEEE Transactions on Communications, vol. 66, no. 9, pp. 3825--3836, 2018.

[14] Z. Sheng, J. Fan, C. H. Liu, V. C. Leung, X. Liu, and K. K. Leung, "Energy-efficient relay selection for cooperative relaying in wireless multimedia networks," IEEE Transactions on Vehicular Technology, vol. 64, pp. 1156-1170, 2015.

[15] M. Hajiaghayi, M. Dong, and B. Liang, "*Jointly optimal channel and power assignment for dual-hop multi-channel multi-user relaying*," *IEEE Journal on Selected Areas in Communications*, vol. 30, no.9, pp. 1806-1814, 2012.

[16] F. Li, X. Tan, and L. Wang, "Power scheme and time-division bargaining for cooperative transmission in cognitive radio," Wireless Communications and Mobile Computing, vol. 15, no. 2, pp. 379-388, 2015.

[17] L. Lv, J. Chen, Q. Ni, Z. Ding, and H. Jiang, "*Cognitive Non-Orthogonal Multiple Access with Cooperative Relaying: A New Wireless Frontier for 5G Spectrum Sharing*," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 188-195, 2018.

[18] F. Li, X. Tan, and L. Wang, "Power scheme and time-division bargaining for cooperative transmission in cognitive radio," Wireless Communications and Mobile Computing, vol. 15, pp. 379-388, 2015.

[19] N. Zhang, N. Cheng, N. Lu, H. Zhou, J. W. Mark, & X. Shen, "*Risk-aware cooperative spectrum access for multi-channel cognitive radio networks*." *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 3, pp. 516-527, 2014.

[20] F. Gomez-Cuba, R. Asorey-Cacheda, and F. J. Gonzalez-Castano, "*A survey on cooperative diversity for wireless networks," IEEE Communications Surveys & Tutorials*, vol. 14, no.3, pp. 822-835, 2012.

[21] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "*Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," IEEE Trans. Wireless Communications,* vol. 13, no. 2, pp. 913-927, 2014.

[22] G. Liu, F. R. Yu, H. Ji, V. C. Leung, and X. Li, *"In-band full-duplex relaying: A survey, research issues, and challenges," Resource*, vol. 147, no.2, pp. 172, 2015.

[23] Y. Fan, C. Wang, J. Thompson, and H. Poor, "*Recovering Multiplexing Loss through Successive Relaying Using Repetition Coding," IEEE Transactions on Wireless Communications*, vol. 6, no.2, pp. 4484-4493, 2007.

[24] M. Lari, "*Power allocation and effective capacity of AF successive relays," Wireless Networks*, vol. 24, no.3, pp. 885-895, 2018.

[25] C. Wang, Y. Fan, I. Krikidis, J. S. Thompson, and H. V. Poor, *"Superposition-coded concurrent decode-and-forward relaying," in Information Theory*, 2008. ISIT 2008. IEEE International Symposium on, 2008, pp. 2390-2394.

[26] I. Orikumhi, C. Y. Leow, and Y. Li, "*Reliable Virtual Full-Duplex Relaying in the Presence of Interrelay Interference," IEEE Transactions on Vehicular Technology*, vol. 66, no.10, pp. 9098-9109, 2017.

[27] M. S. Gilan and A. Olfat, "*New beamforming and space-time coding for two-path successive decode and forward relaying," IET Communications*, vol. 12, pp. 1573-1588, 2018.

[28] Q. Y. Liau, C. Y. Leow, and Z. Ding, "*Amplify-and-Forward Virtual Full-Duplex Relaying-Based Cooperative NOMA*," *IEEE Wireless Communications Letters*, vol. 7, pp. 464-467, 2018.

[29] T. Charalambous, S. M. Kim, N. Nomikos, M. Bengtsson, and M. Johansson, "*Relay-pair selection in buffer-aided successive opportunistic relaying using a multi-antenna source," Ad Hoc Networks*, vol. 84, pp. 29-41, 2019.

[30] C. Zhai, W. Zhang, and P. Ching, "*Cooperative spectrum sharing based on two-path successive relaying," IEEE Transactions on Communications*, vol. 61, pp. 2260-2270, 2013.

[31] A. H. A. El-Malek and S. A. Zummo, *"A bandwidth-efficient cognitive radio with two-path amplify-and-forward relaying," IEEE Wireless Communications Letters*, vol. 4, pp. 66-69, 2015

[32] Z. Li, F. Xiao, S. Wang, T. Pei, and J. Li, *"Achievable rate maximization for cognitive hybrid satellite-terrestrial networks with af-relays," IEEE Journal on Selected Areas in Communications*, vol. 36, pp. 304-313, 2018.

[33] S. Masrour, A. H. Bastami, and P. Halimi, *"Spectrum sharing in cognitive radio networks using beamforming and two-path successive relaying," in Electrical Engineering (ICEE)*, 2017 Iranian Conference on, 2017, pp. 1810-1814.

[34] C. Luo, Y. Gong, and F. Zheng, *"Full interference cancellation for two-path relay cooperative networks," IEEE Transactions on Vehicular Technology*, vol. 60, pp. 343-347, 2011.

[35] Y. Ji, C. Han, A. Wang, and H. Shi, *"Partial inter-relay interference cancellation in two-path successive relay network," IEEE Communications Letters*, vol. 18, pp. 451-454, 2014.

[36] Chih-Lin, I., Rowell, C., Han, S., Xu, Z., Li, G., & Pan, Z. *"Toward green and soft: a 5G perspective." IEEE Communications Magazine,* vol. 52, no. 2, pp. 66–73, Feb. 2014.

[37] E. Biglieri, J. Proakis, and S. Shamai, *"Fading channels: Information-theoretic and communications aspects," IEEE Transactions on Information Theory*, vol. 44, pp. 2619-2692, 1998.

[38] N. Zhao, F. R. Yu, H. Sun, and M. Li, *"Adaptive power allocation schemes for spectrum sharing in interference-alignment-based cognitive radio networks," IEEE transactions on vehicular technology*, vol. 65, pp. 3700-3714, 2016.

[39] K. Gomadam, V. R. Cadambe, and S. A. Jafar, *"A distributed numerical approach to interference alignment and applications to wireless interference networks," IEEE Transactions on Information Theory*, vol. 57, p. 3309, 2011.

[40] A. Zappone and E. Jorswieck, "*Energy efficiency in wireless networks via fractional programming theory,*" *Foundations and Trends in Communications and Information Theory*, vol. 11, pp. 185-396, 2015.

N^+ Pocket Core-Shell Nanotube Tunnel Field-Effect Transistor

Amirsam Abad [1], Iman Chahardah Cherik[2] and Saeed Mohammadi[2]

Abstract— **One of the attractive candidates for improving the performance of tunnel transistors is cylindrical structures due to their impressive electrostatic control of the gate. But the on-state current of tunnel transistors is still very low compared to MOSFETs. An alternative is to use core-shell nanotubes rather than nanowires. In this article, we present a core-shell TFET nanotube based on a heterogeneous germanium/silicon structure. In our proposed structure, an N^+ pocket is employed to enhance the on-state current. A possible manufacturing method is also proposed that is fully compatible with CMOS technology. The main parameters of this transistor are 97.85 µA / µm on-state current, I_{on} / I_{off} ratio of 8.26×10^8, SS_{avg} mV/dec 21.15, and f_T of 878.95 GHz.**

***Index Terms*—Tunnel FET, germanium-source, core-shell nanotube, heterojunction, on-state current.**

I. Introduction

In recent years, MOSFET transistors have faced serious design and manufacturing challenges due to their short channel effects [1]. Lack of proper control over the source Fermi tail has restricted the subthreshold slope of these transistors. [2]. Due to the need of the semiconductor industry for a device with less power loss, different devices have been designed and proposed, in recent years [3]- [5]. One of the most important structures that do not pose a new challenge to manufacturing technology is Tunnel transistors due to their physical similarities to MOSFETs. The operating mechanism of these transistors, unlike MOSFETs, is to tunnel the carriers through the potential barrier. Like all semiconductor devices, they face challenges such as ambipolar conductivity and low on-state current [6], [7]. Different structures have been designed and proposed, to overcome these limitations [8] - [12]. Nanotube structures are one of the most interesting ideas to improve the performance of tunnel transistors [13].

In nanotubes, in addition to the shell gate, there is also a core gate which increases the electrostatic integrity of the gate and increases the intensity of the electric field at the tunneling junction. Musalgaonkar et al. presented a misaligned nanotube in which the shell gate overlaps with the source [14]. Apoorva et al. proposed dopingless nanotubes with high on-state current [15]. Hanna et al. Proposed germanium-based nanotubes with the on-state current of 18µA/µm at the V_{GS}=1.0V [16]. In this paper, we present a new core-shell nanotube that uses an n^+ pocket between source and channel. This n+ pocket increases the electric field and hence the band bending. This article consists of the following sections. In section II, we introduce the physics of structure, the manufacturing process, and the simulation method. In Section III, we assess the impact of different parameters on device performance. In section IV, we conclude the obtained results.

II. Device Structure, Fabrication Process, and Simulation Methodology

Fig.1 shows the cross-sectional (top) and cylindrical (bottom) shapes of the CSNT-TFET, respectively. As shown in the figure, our proposed device consists of core and shell gates that overlap the pocket and channel area. To increase the on-state current, germanium is used in the source area, which has better tunneling characteristics than silicon [17]. The work- function of the gate metal is 4.3eV. The contact radius of the core gate contact and the core gate are 5nm and 50nm, respectively. All other design parameters are listed in the Table I.

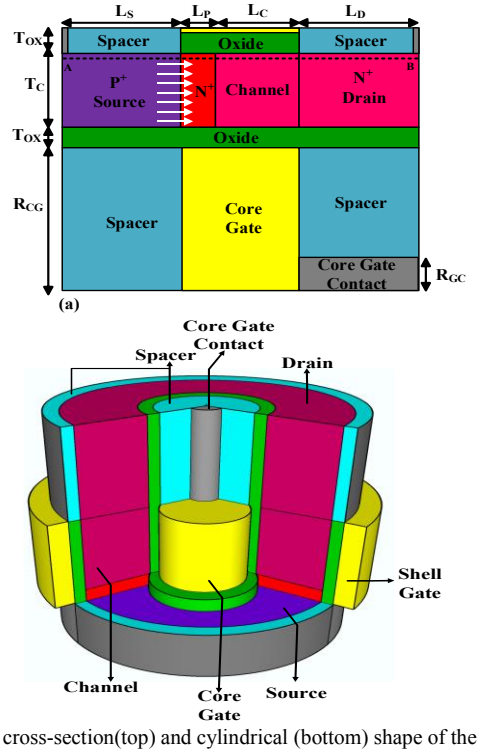


Fig. 1. A cross-section(top) and cylindrical (bottom) shape of the proposed TFET structure (CSNT-TFET).

1: Amirsam Abad is with the Electrical Engineering Department of Amirkabir University,Tehran 1591634311, Iran
Corresponding author: abadamirsam@aut.ac.ir

2: Iman Chahardah Cherik an Saeed Mohammadi are with the Department of Electrical and Computer Engineering,Semnan University, Semnan 3513119111, Iran

TABLE I
Default Values of Proposed TFED Parameters

Dimensions (nm)		Doping concentrations (cm^{-3})	
Epi-layer length, L_P	5	n^+ epi-layer, N_E	4×10^{19}
Source region length, L_S	30	Source region, N_S	2.5×10^{19}
Channel region length, L_C	50	Drain region, N_D	3×10^{18}
Drain region length, L_D	30	Channel region, N_C	1×10^{15}
Channel thickness, T_C,	10		

The operation mechanism of our device is described as follows. As the gate voltage increases, the electrons in the source region tunnel into the channel (Fig. 2 (a)) and then, with the increase of drain current move toward the drain region of the transistor (Fig. 2(b)).

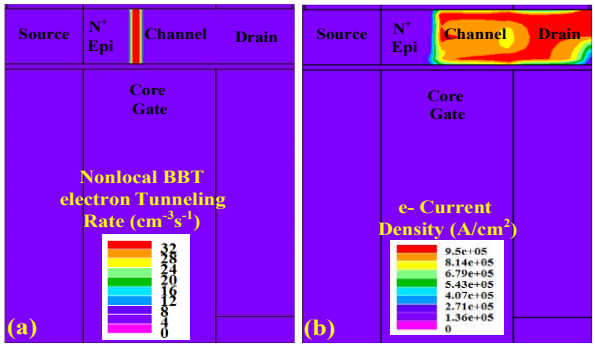


Fig. 2. (a) Electron BTBT tunneling rate, and (b) electron current density contour maps at V_{GS} = 0.5V and V_{DS} = 0.5V.

Fig. (3) indicates the proposed CSNT-TFET manufacturing process. This process begins with the epitaxial growth of germanium in the active silicon region, followed by the growth of n-epi, channel and drain region 3(a). Silicon and germanium layers are selectively etched, then sacrificial oxide and a dielectric layer are deposited 3(b). Gate oxide and gate metal are deposited 3(c). Gate metal is selectively etched, and a layer of SiO_2 and sacrificial oxide are deposited 3(d, e). The sacrificial layer is etched 3(f). A layer of dielectric is removed, the silicon region is etched, and a layer of gate oxide is deposited 3(g). Inner gate metal and a layer of SiO_2 are deposited 3(h), and finally, source, gate, and drain contacts are connected 3(i).

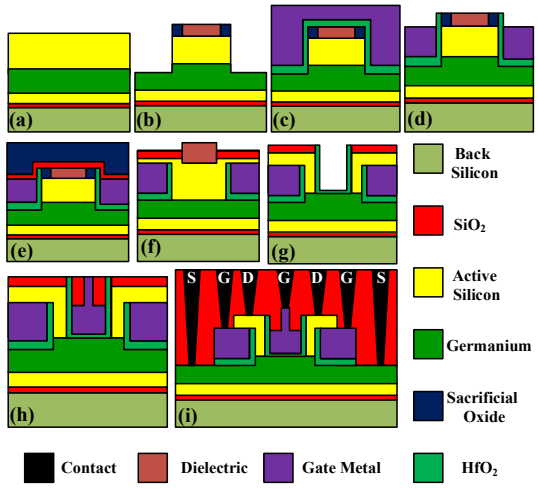


Fig. 3. Fabrication process flow for realizing proposed structure.

All simulations are performed with Silvaco ATLAS [18]. A dynamic non-local band-to-band tunneling model has been used for the proper calculation of the drain current. The generation-recombination of carriers is modeled with the SRH model. The effect of parameters such as doping and temperature on carrier mobility is modeled by CVT. The BGN model is activated due to high doping in the source region. The quantum confinement model is not activated since the length of the pocket region is more than 4 nm [19]. The gate leakage current model is also not activated.

III. Simulation Results and Discussion

Fig.4 indicates the effect of gate voltage on CSNT-TFET energy band diagrams. From the figure, it can be inferred that as the gate voltage increases, the bending of the band increases and we have a larger tunneling window for carriers to tunnel to the conduction band.

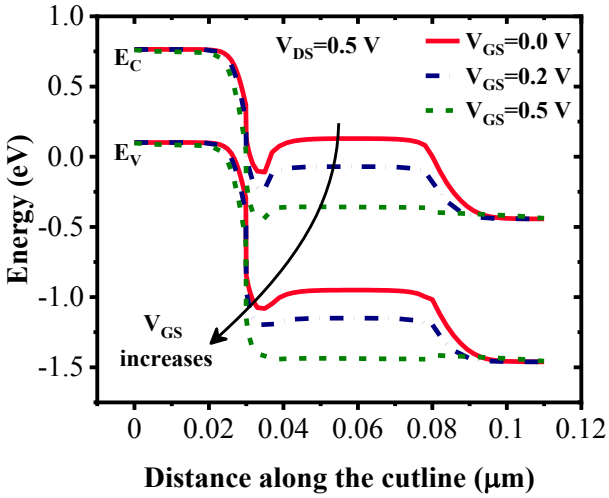


Fig. 4. Impact of gate voltage on the energy band diagram of CSNT-TFET.

In Fig. 5 (a), (b) impact of gate voltage and drain voltage on the drain current of the transistor is assessed. Fig. 5(a) shows that with the increase of gate voltage from V_{GS}=0.1V to V_{GS}=0.5V, drain current increases from 4.46×10^{-3} μA to 33.8 μA, which is mainly due to the reduction of band-to-band tunneling distance. Fig. 5(b) shows that increasing the drain voltage and, hence the increase of the density of states has a low impact on the drain current. It should also be noted that increasing the drain current does not affect the onset voltage of the transistor.

One of the important parameters that have a significant impact on the transfer characteristics of tunneling field-effect transistors is source doping. As shown in Fig.6, with the increase of N_S from 1×10^{19} to 2.5×10^{19}, drain current reaches from 4.02×10^{-1} μA to 33.8 μA which is attributed to the increases of available carriers for tunneling to the conduction band of the channel. Moreover, with the increment of source doping, onset voltage decreases from V_{GS}=0.08V to V_{GS}=0.02V.

One of the main drawbacks of tunneling field-effect transistors is traps in the semiconductor bandgap [20], which dramatically affect the off-state current in heavily doped p-n

junctions. Fig. 7(a) indicates the impact of the TAT model on the transfer characteristics of the CSNT-TFET. In the presence of TAT, the off-state current reaches from 5.9×10^{-12} µA to 4.09×10^{-8} µA. As shown in the inset of Fig (7(a), with the increase of off-state current, the I_{on}/I_{off} ratio decreases from 5.72×10^{12} to 8.26×10^{8}.

Another vital parameter that has a significant effect on the performance of tunneling field-effect transistors is temperature.

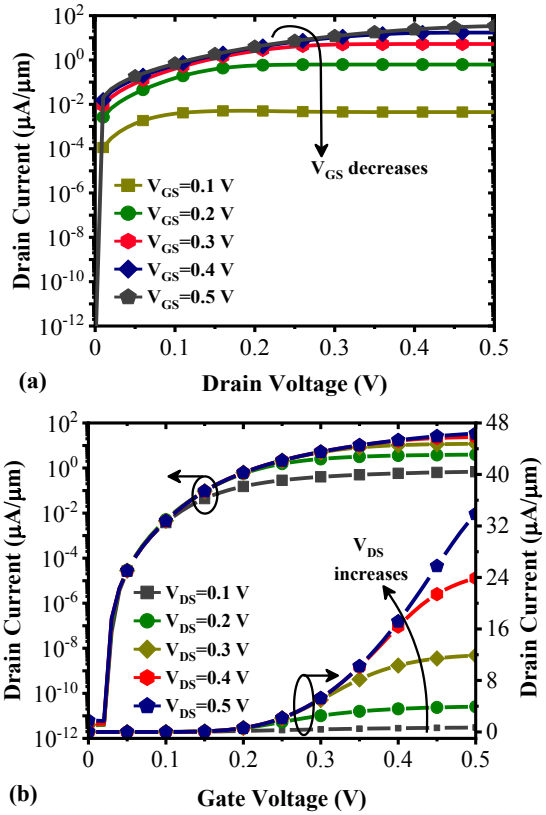


Fig.5. (a) output characteristics for different gate voltages, (b) transfer characteristics for different drain voltages of CSNT-TFET.

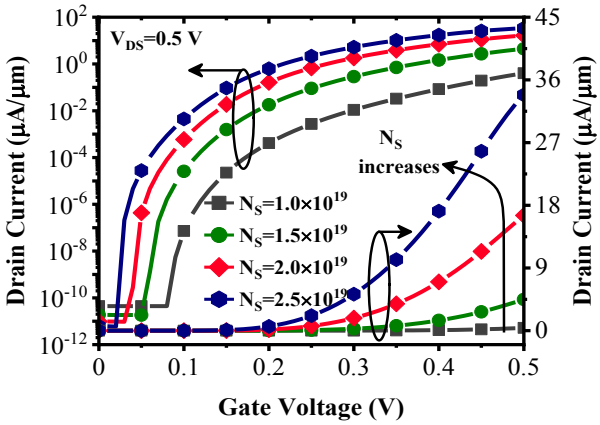


Fig. 6. Impact of N_S doping density on the transfer characteristics of CSNT-TFET.

Fig. 7(b) depicts that, with the increases of the temperature from 300ºK to 375ºK, off-state current with the three decades of current increment reaches 6.11×10-9 µA, which is mainly semiconductors. On the other hand, the on-state current has less sensitivity to the temperature since the tunneling equation has no direct relation to the temperature.

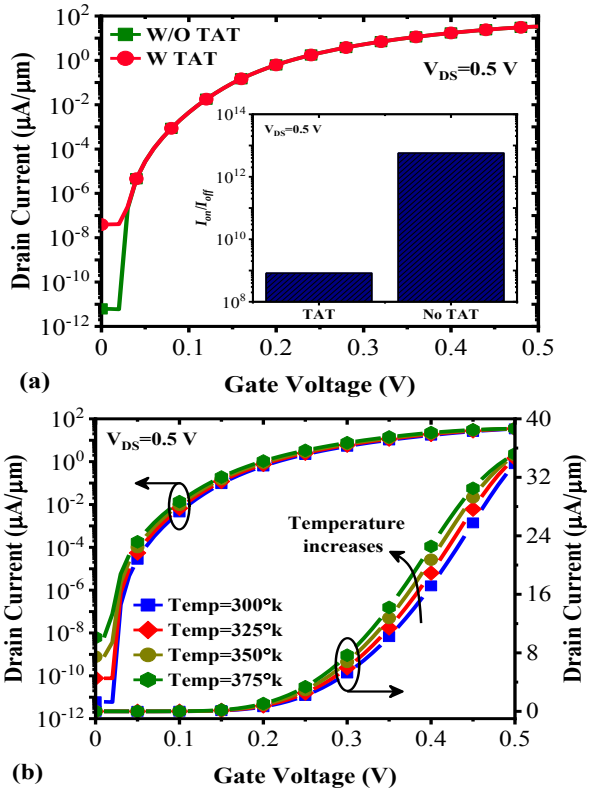


Fig. 7. Impact of TAT (a) and temperature (b) on the transfer characteristics of the CSNT-TFET.

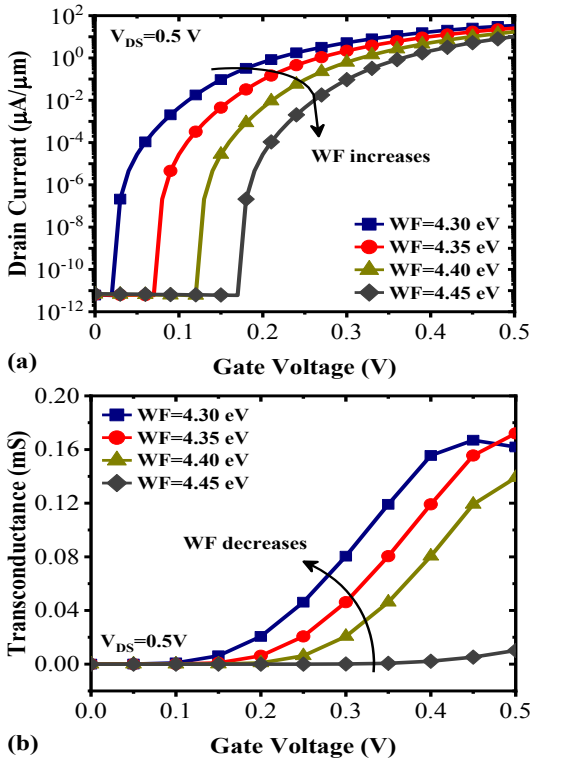


Fig. 8. Impact of the gate work function on (a) transfer characteristics, and (b) transconductance of CSNT-TFET.

Fig. 8(a) shows the effect of gate work- function on CSNT-TFET transmission characteristics. Selecting the appropriate gate work-function results in enough electric field flux at the tunneling junction, which decreases the onset voltage and increases the drain current. In this device, with increasing gate work- function from 4.3eV to 4.45eV, onset voltage increases from V_{GS} = 0.02V to V_{GS} = 0.17V.

Fig. 8(b) shows the impact of gate work-function on the transconductance of CSNT-TFET. Transconductance is given by $g_m = \partial I_D / \partial V_{GS}$ and, it is clear that lower gate work-function leads to a higher transconductance which is due to the impact of lower gate work-function on the decrement of the onset voltage (the gate voltage at which BTBT starts).

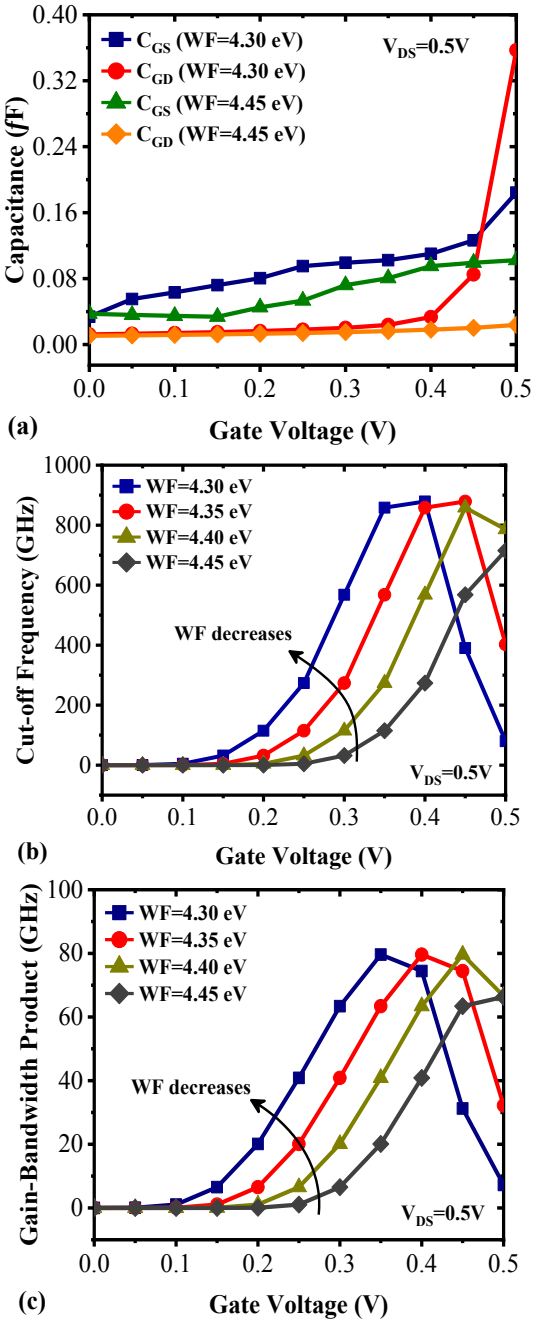


Fig. 9. Impact of the gate workfunction on (a) parasitic capacitance, (b) cut-off frequency, and (c) gain-bandwidth product of CSNT-TFET.

Fig. 9(a) examines the effect of gate work- function on CSNT-TFET parasitic capacitances. These parasitic capacitances have a great effect on the speed of the transistor, so it is better to design a device with fewer parasitic capacitances. Parasitic capacitance is composed of two main components: gate-to-source capacitance (C_{GS}) and gate-to-drain capacitance (C_{GD}).

From Fig. 9(a), it can be inferred that lower gate work-function leads to a higher C_{GS} and C_{GD}. This is because the lower gate work-function leads to a sharper band diagram at the tunneling junction, so we have more charge carriers on the source side of the transistor that results in higher C_{GS}. Then these charge carriers move toward the drain side of the transistor so, C_{GD} also increases.

The cut-off frequency and gain-bandwidth product are two crucial parameters of field-effect transistors. Cut-off frequency and gain-bandwidth product are defined by $f_T = g_m/2\pi(C_{GD}+C_{GS})$ and $GBW = g_m/2\pi(10C_{GD})$ respectively. From Fig. 9(b) and 9(c), it can be concluded that lower gate work function leads to a higher f_T and GBW. Thus we can understand that transconductance is the predominant component of parasitic capacitances. It should be mentioned that we have peak f_T and GBW of 878.95 GHz and 79.61 GHz for wf=4.35 eV.

In Table II and Table III, we compare the DC and AC/RF performance of our device with some similar structures. From tabs, 2 It can be inferred that our proposed device performs well at low voltage, and I_{on}/I_{off} has a reasonable value even in the presence of TAT. Table III also indicates that CSNT-TFET has a superior AC/RF performance concerning the other works.

TABLE II
Performance Comparison of Different TFET Architectures

Ref	SS_{min} (mV/dec)	SS_{avg} (mV/dec)	I_{on}/I_{off}	I_{on} (μA/μm)	V_{Bias} (V)
[13]	~18	-	$\rangle 10^{10}$	~0.01	1.0
[14]	~5	51	1.6×10^{8}	1.38	0.3
[15]	-	31.38	8.46×10^{11}	16.9	1.0
[16]	34	-	$\rangle 10^{6}$	18	1.0
[21]	10	32.01	3.92×10^{7}	52.19	0.5
[This work]	2	21.15	8.26×10^{8} (TAT included)	97.85	0.5

TABLE III
Comparison of Analog/RF Performance Of Different TFETS

Ref	V_{Bias} (V)	g_m (mS)	f_T (GHz)	GBW (GHz)
[21]	0.5	0.118	50.4	41.18
[22]	0.7	0.16	89.31	25.84
[23]	0.6	300	~75	~10
[24]	0.5	0.75	97.6	-
[25]	1.0	0.005	45	7.0
[26]	1.4	0.45	70	8.0
This Work	0.5	0.166	878.95	79.61

IV. Conclusion

In this article, we propose a CSNT-TFET that uses an n^+ pocket between the source and channel regions. Using this n^+ pocket along with the heterojunction improves the performance of the device and makes it a suitable candidate for low-power applications. We also showed that our proposed structure is

fully compatible with CMOS technology. All numerical simulations have been performed with the Silvaco Atlas, and the effect of nonidealities such as traps has been investigated for a fair comparison. Among the main parameters of the device, we can mention to I_{on}=97.85 μA/μm, SS_{avg}=21.15 mV/dec.

REFERENCES

[1] N. Z. Haron and S. Hamdioui, "Why is CMOS scaling coming to an END?," in *2008 3rd International Design and Test Workshop*, 2008, pp. 98-103: IEEE. DOI: https://doi.org/10.1109/IDT.2008.4802475

[2] J. Knoch and J. Appenzeller, "A novel concept for field-effect transistors-the tunneling carbon nanotube FET," in *63rd Device Research Conference Digest, 2005. DRC'05.*, 2005, vol. 1, pp. 153-156: IEEE, https://doi.org/10.1109/DRC.2005.1553099

[3] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proceedings of the IEEE,* vol. 98, no. 12, pp. 2095-2110, 2010.https://doi.org/10.1109/JPROC.2010.2070470

[4] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano Letters,* vol. 8, no. 2, pp. 405-410, 2008.https://doi.org/10.1021/nl071804g

[5] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, "Impact ionization MOS (I-MOS)-Part I: device and circuit simulations," *IEEE Transactions on electron devices,* vol. 52, no. 1, pp. 69-76, 2004.https://doi.org/10.1109/TED.2004.841344

[6] D. B. Abdi and M. J. Kumar, "Controlling ambipolar current in tunneling FETs using overlapping gate-on-drain," *IEEE Journal of the Electron Devices Society,* vol. 2, no. 6, pp. 187-190, 2014.https://doi.org/10.1109/JEDS.2014.2327626

[7] U. E. Avci, D. H. Morris, and I. A. Young, "Tunnel field-effect transistors: Prospects and challenges," *IEEE Journal of the Electron Devices Society,* vol. 3, no. 3, pp. 88-95, 2015.https://doi.org/10.1109/JEDS.2015.2390591

[8] N. Bagga, A. Kumar, and S. Dasgupta, "Demonstration of a novel two source region tunnel FET," *IEEE Transactions on Electron Devices,* vol. 64, no. 12, pp. 5256-5262, 2017.https://doi.org/10.1109/TED.2017.2759898

[9] I. C. Cherik and S. Mohammadi, "Enhanced on-state current and suppressed ambipolarity in germanium-source dual vertical-channel TFET," *Semiconductor Science and Technology,* 2020.https://doi.org/10.1088/1361-6641/abd63e

[10] I. C. Cherik and S. Mohammadi, "Germanium-source L-shaped TFET with dual-in-line tunneling junction," *Applied Physics A,* vol. 127, no. 7, pp. 1-8, 2021.https://doi.org/10.1007/s00339-021-04677-5

[11] S. Kim, W. Y. Choi, and B.-G. Park, "Vertical-structured electron-hole bilayer tunnel field-effect transistor for extremely low-power operation with high scalability," *IEEE Transactions on Electron Devices,* vol. 65, no. 5, pp. 2010-2015, 2018.https://doi.org/10.1109/TED.2018.2817569

[12] L. Lattanzio, L. De Michielis, and A. M. Ionescu, "Complementary germanium electron-hole bilayer tunnel FET for sub-0.5-V operation," *IEEE Electron Device Letters,* vol. 33, no. 2, pp. 167-169, 2011.https://doi.org/10.1109/LED.2011.2175898

[13] H. M. Fahad and M. M. Hussain, "High-performance silicon nanotube tunneling FET for ultralow-power logic applications," *IEEE transactions on electron devices,* vol. 60, no. 3, pp. 1034-1039, 2013.https://doi.org/10.1109/TED.2013.2243151

[14] G. Musalgaonkar, S. Sahay, R. S. Saxena, and M. J. Kumar, "A line tunneling field-effect transistor based on misaligned core-shell gate architecture in emerging nanotube FETs," *IEEE Transactions on Electron Devices,* vol. 66, no. 6, pp. 2809-2816, 2019.https://doi.org/10.1109/TED.2019.2910156

[15] N. Kumar, S. I. Amin, and S. Anand, "Design and Performance Optimization of Novel Core-Shell Dopingless GAA-Nanotube TFET With Si 0.5 Ge 0.5-Based Source," *IEEE Transactions on Electron Devices,* vol. 67, no. 3, pp. 789-795, 2020.https://doi.org/10.1109/TED.2020.2965244

[16] A. Hanna and M. M. Hussain, "Si/Ge hetero-structure nanotube tunnel field-effect transistor," *Journal of Applied Physics,* vol. 117, no. 1, p. 014310, 2015.https://doi.org/10.1063/1.4905423

[17] S. H. Kim, S. Agarwal, Z. A. Jacobson, P. Matheu, C. Hu, and T.-J. K. Liu, "Tunnel field-effect transistor with raised germanium source," *IEEE electron device letters,* vol. 31, no. 10, pp. 1107-1109, 2010.https://doi.org/10.1109/LED.2010.2061214

[18] Silvaco, ATLAS Device Simulation Software User's Manual, no. version 3.2. 2015

[19] D. Querlioz, J. Saint-Martin, K. Huet, A. Bournel, V. Aubry-Fortuna, C. Chassat, S. Galdin-Retailleau, and P. Dollfus, "On the ability of the particle Monte Carlo technique to include quantum effects in nano-MOSFET simulation," *IEEE transactions on electron devices,* vol. 54, no. 9, pp. 2232-2242, 2007.https://doi.org/10.1109/TED.2007.902713

[20] S. Sant, A. Schenk, K. Moselund, and H. Riel, "Impact of trap-assisted tunneling and channel quantization on InAs/Si hetero tunnel FETs," in *2016 74th Annual Device Research Conference (DRC)*, 2016, pp. 1-2: IEEE. DOI: https://doi.org/10.1109/DRC.2016.7548413.

[21] I. C. Cherik, S. Mohammadi and A. A. Orouji, "Switching Performance Enhancement in Nanotube Double-Gate Tunneling Field-Effect Transistor With Germanium Source Regions," in *IEEE Transactions on Electron Devices*, vol. 69, no. 1, pp. 364-369, Jan. 2022, DOI: 10.1109/TED.2021.3124984.

[22] I. C. Cherik and S. Mohammadi, "Vertical Cladding Layer Based Doping-Less Tunneling Field Effect Transistor, a Novel Low-Power High-Performance Device," IEEE Transactions on Electron Devices, 2021.https://doi.org/10.1109/TED.2021.3138669

[23] A. Bhattacharyya, M. Chanda, and D. De, "GaAs0. 5Sb0. 5/In0. 53Ga0. 47As heterojunction dopingless charge plasma-based tunnel FET for analog/digital performance improvement," *Superlattices and Microstructures,* vol. 142, p. 106522, 2020.https://doi.org/10.1016/j.spmi.2020.106522

[24] A. Acharya, A. B. Solanki, S. Dasgupta, and B. Anand, "Drain current saturation in line tunneling-based TFETs: An analog design perspective," *IEEE Transactions on Electron Devices,* vol. 65, no. 1, pp. 322-330, 2017.10.1109/TED.2017.2771249.

[25] M. R. Tripathy, A. K. Singh, K. Baral, P. K. Singh, and S. Jit, "III-V/Si staggered heterojunction based source-pocket engineered vertical TFETs for low power applications," *Superlattices and Microstructures,*p.106494,2020.https://doi.org/10.1016/j.spmi.2020.106494.

[26] C. Pandey, A. Singh, and S. Chaudhury, "Effect of asymmetric gate-drain overlap on ambipolar behavior of double-gate TFET and its impact on HF performances," *Applied Physics A,* vol. 126, no. 3, pp. 1-12, 2020.https://doi.org/10.1007/s00339-020-3402-2.

Ion-Sensitive Field-Effect Transistor-Based Biosensor for PSA Antigen Concentration Measurement Using Microfluidic System

Amir Azadi [1]*, Saeed Mohammadi [2] and Parviz Keshavarzi[3]

***Abstract*— Prostate cancer is one of the most common cancers in men. Prostate-Specific Antigen (PSA) is an important biomarker in the diagnosis of prostate cancer. In the present paper, an Ion-Sensitive Field-Effect Transistor (ISFET) is introduced, employing microfluidic technology to detect PSA antigens efficiently. PSA antigen is our analyte in this sensor. Due to the PSA antigen's acidity and sensor sensitivity to the hydrogen ions and pH index, absorbing the hydrogen ion by the OH receptor on the sensor surface modulates the drive current and the device's threshold voltage. The electroosmotic flow is induced inside the microchannel by applying a voltage to the electrodes on the walls of the microchannel. Consequently, turbulence in the fluid flow in the channel has occurred that effectively moves the intended analytes toward the sensor surface. The biosensor performance is investigated by the simulations carried on in COMSOL. Our simulation results indicate that the proposed structure facilitates rapid detection and measurement of PSA concentration.**

***Index Terms*—** ISFET, Prostate-Specific Antigen, Microfluidic, Electro-kinetics, Electroosmotic, Biosensor.

I. INTRODUCTION

Nowadays, biosensors are used in various applications such as the pharmaceutical industry, chemical industry, food industry, medical diagnosis, environmental monitoring, and production of health products to identify different biological molecules. The most common use of biosensors is in medical diagnoses and laboratory sciences. Significant improvements have been reported in silicon-based biosensors, including optical and field-effect transistor-based biosensors (ISFETs) [1]. Biological analysis of humans is an important technique for health monitoring. The values of the chemical parameters like pH, pNa, glucose, potassium, and calcium, along with the physical data of activity, body temperature, and the heartbeat, can effectively indicate someone's health condition. In addition, ISFET sensors are used to analyze human body sweat and plasma [2].

Ion-Sensitive Field-Effect Transistors (ISFETs) and Extended-Gate Field-Effect Transistors (EGFETs) are the first biosensors introduced. Employing biological materials as surface receptors, they exhibit more reliable detection of biological and chemical species [3]. One of the main advantages of the FET-based biosensors is their potential ability to be miniaturized while scaling other types of sensors (such as optical and surface acoustic wave sensors) has fundamental limitations. ISFET is a potentiometric sensor in which the field effect is created by the presence of electric charges at the electrolyte-insulator interface. ISFET-based biosensors are promising candidates for biological and chemical fluid analysis applications and laboratory diagnoses [4]. In addition to the ease of miniaturization, ISFETs can work at high temperatures [5]. Like MOSFET, ISFET is a three-terminal device. Still, in ISFET, the metal gate is replaced by an ion-sensitive membrane, electrolyte solution, and reference electrode, where the gate voltage is applied [6]. Although ISFETs exhibit high accuracy and fast response [7], they suffer from weak chemical stability [8].

ISFETs can measure the concentration of cancer biomarkers and detect cancer [9]. Cancer is a global health threat causing millions of deaths every year. The statistics reveal that prostate cancer is the world's second most common cancer in men [10]. Early detection of prostate cancer is of particular importance. One of the effective methods of diagnosing and treating this disease is its biomarker diagnosis. Currently, the best biomarker employed for this detection is a prostate-specific antigen (PSA) [11]. PSA is a protein made by prostate cells. Prostatic secretions are slightly acidic, with a pH of around 6.4. PSA is secreted by both healthy cells and prostate cancer cells. The normal level of PSA in human blood is about 4 ng/ml, but in patients with prostate cancer, this amount increases. Since about 30% of men with this type of cancer have PSA levels in the range of 4.1-9.9 ng/ml, accurately measuring PSA levels is critical [12].

In the present study, an Ion-Sensitive Field-Effect Transistor (ISFET) based biosensor is simulated to measure PSA concentration. Microfluidic technology is employed to improve the efficiency of the sensor. In section 2, the ISFET theory is

1- Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
2- Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
3- Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
Corresponding author: a_azadi@semnan.ac.ir

introduced, and in section 3, the microfluidic and electroosmotic flow theory is explained. The simulation approach is presented in section 4, and the simulation results are presented and discussed in sections 5 and 6. Finally, section 7 concludes the paper.

II. The ISFET theory

The ISFET operation is explained with the use of site-binding theory. To form the binding sites, the oxide layer surface is covered by hydroxyl groups (OH^-), which can be positively or negatively charged by absorption of H^+ or by losing H^+, respectively, depending on the concentration of hydrogen ions in the electrolyte. Accordingly, the insulating surface is charged with surface charge density depending on the hydrogen ion concentration (pH) of a solution [13]. The PSA concentration can also be measured by absorbing hydrogen ions on the sensor surface covered by OH^- [14]. An Ag electrode is employed as the reference electrode in sensor structure [15]. Fig.1 represents a three-dimensional scheme of the ISFET sensor in the COMSOL software framework.

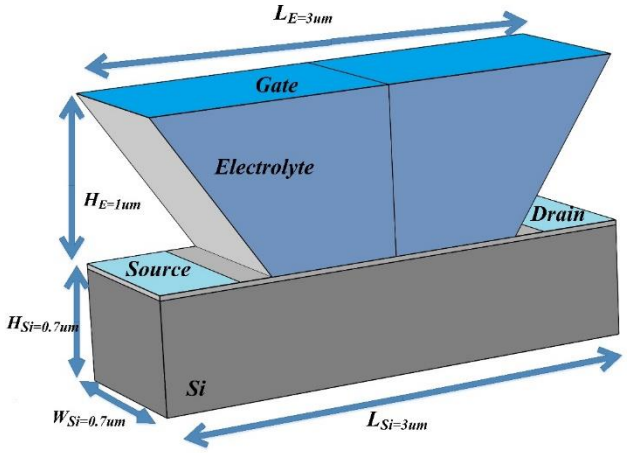


Fig1. Schematic 3D view of ISFET sensor.

When electrolyte enters between the reference electrode and the oxide layer, two potentials are induced at the reference electrode (E_{ref}) and solution-oxide interface (named interfacial potential $\psi_0 + \chi^{sol}$), where ψ_0 is the surface potential, which is a function of solution pH, and χ^{sol} is the surface dipole potential of solvent. The threshold voltage for ISFET is given by [16],

$$V_{TH(IS)} = E_{ref} - \psi_0 + \chi^{sol} - \phi_{si} - \frac{Q_{ox} + Q_{ss} + Q_B}{C_{ox}} + 2\phi_f \qquad (1)$$

where φ_{si} is the semiconductor work function, Q_{ox} is the oxide charge, Q_{ss} is the semiconductor surface charge at the insulator-semiconductor interface, Q_B is the depletion region charge, C_{ox} is the oxide capacitance, and φ_f is the Fermi potential. The potential surface changes with the absorption of hydrogen ions on the oxide layer surface, consequently affecting the device's performance. The value of surface potential is determined by [17],

$$\psi_0 = \chi^{sol} + \frac{\rho_e}{C_{stern}} \qquad (2)$$

Drain current of ISFET in the non-saturated region of operation may be expressed the same as that of MOSFET as,

$$I_{DS} = \frac{W}{L}\mu C_i \left[(V_G - V_{TH})V_{DS} - \frac{V_{DS}^2}{2} \right] \qquad (3)$$

where μ is the mobility of electrons in the inverted channel for an n-channel transistor, C_i is the capacitance per unit area of the gate, and W and L indicate the channel width and length of ISFET, respectively[18].

III. The Microfluidic and Electroosmotic Flow Theory

Research has been conducted to employ microfluidic technology in these sensors to increase the efficiency of ISFET sensors. Microfluidic technology is widely used in biosensors [19] and provides advantages such as less measure time, high sensitivity, transportability, and integration of laboratory methods in one device (lab-on-chip technology) [20]. Electrokinetics is an efficient technique to control liquids and samples in microfluidic systems. This technique is used to separate samples and create turbulence within the microchannel. By applying a driving potential to the microchannel's surface electrodes, the state of motion of liquid, which contains biomolecules, is changed. The velocity of the fluid in this electroosmotic flow is given by [21].

$$u_s = -\frac{\varepsilon_f \zeta}{U} E \qquad (4)$$

where ε_f is the dielectric constant, U is the viscosity constant, and E is the electric field. Electroosmotic flow is also used in micromixers, where fluid-induced turbulence leads to mixing operation in microchannel [22]. The fluid exhibits laminar flow without applied voltage with a small Reynolds number. In this situation, the absorption of target ions to the sensor surface is weakened; consequently, the sensor's response time and sensitivity are degraded. Inducing turbulence in the microchannel enhances the analyte absorption and improves the sensor efficiency. The Damkohler number is a dimensionless scale to study mass transfer rate [23].

$$Da = \frac{k_{on}\theta_0}{(D/h)} \qquad (5)$$

where Da is the Damkohler number, k_{on} is the reaction rate, θ_0 is the number of free receptors for antibody-antigen combination, D is the molecular diffusion coefficient, and h is the channel height. According to this equation, achieving a desirably high mass transfer rate requires a high reaction rate of the target ions and sufficient molecular diffusion.

IV. The ISFET sensor simulation

To predict the sensory operation of ISFET, we employ the COMSOL simulation framework. The sensitivity and conductivity of the sensor are evaluated in different modes. The substrate of the structure defined in the simulation environment

has 3µm length and 0.7µm height, and the electrolyte cavity has 1µm height and 1.6µm length. The gate length is 3µm. The physical models employed to simulate ISFETs are the same as those used for simulating MOSFETs. These models can describe and evaluate the electric charge concentration and their interactions, concentration of different species, and diffusion coefficient of ions. The microchannel that is simulated in COMSOL, has 25µm height and 100µm length, and two electrodes with 20µm length are located under the channel 10µm apart from each other. The sensor is located under the lower wall of the microchannel. This structure is indicated in Fig. 2.

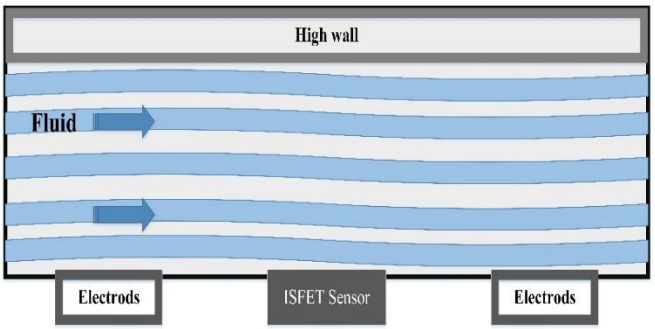


Fig 2. A schematic view of the microchannel and the locations of electrodes and ISFET sensor.

The incompressible Navier-Stokes and electrostatic physics are utilized to investigate the microchannel's fluid behavior. Navier-Stokes and continuity equations are used to analyze fluid motion in the presence of an electrical field, and electrostatic physics refers to the applied voltage to microchannel surface electrodes. The other boundaries of the channel are assumed insulated.

V. Results of ISFET sensor simulation

The current flows when applying a voltage to the reference electrode and increasing its value to a certain level, called threshold voltage. The threshold voltage changes according to the PSA value. Increasing PSA means decreasing pH value and the threshold voltage increases. While decreasing of PSA value leads to a decrease in the threshold voltage. Fig. 3. shows the transfer characteristics of ISFET for different values of PSA.

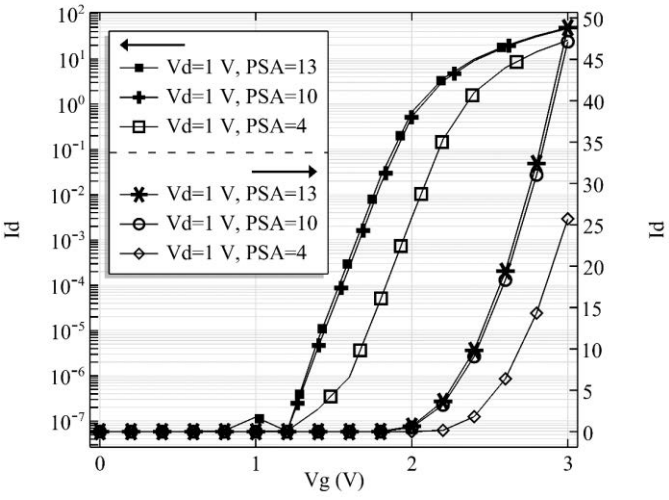


Fig 3. Transfer characteristics of ISFET for different values of PSA in logarithmic (left) and liner (right) scales.

Fig. 4 (a) shows the sensitivity curve for the ISFET PSA sensor operating in the constant current mode. A feedback circuit adjusts the gate voltage, so the drain current is maintained at a constant set point. The resulting gate voltage (the sensor's output) as a function of the PSA value (the sensor's input) is plotted in the figure. In Figure 4(b), the changes in the drain current for different PSA levels are depicted. The applied gate and drain voltages are kept constant to obtain this curve.

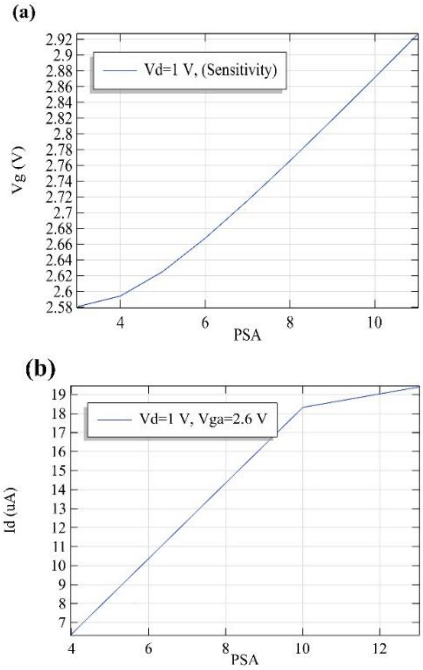


Fig. 4 (a) Sensitivity characteristics of ISFET sensor at a given drain current, and (b) variation of drain current for different PSA levels at a given bias point.

VI. Results of Microchannel Simulation

The simulation of the microchannel is carried out in two modes, with and without applying a voltage to the surface electrodes. As Fig. 5. indicates, the fluid flows without applying a voltage to the electrodes in a laminar state. It is shown in the figure that the fluid velocity is in layered distribution, and it flows faster in the microchannel center (red indicates higher speed and blue indicates lower speed). Applying an ac voltage with a frequency of 0.1Hz to the electrodes induces a turbulence state within the microchannel. The turbulence state changes over time, as is shown in Fig. 6. The higher the turbulence in the microchannel is, the more the analyte is pushed toward the sensor surface, and consequently, the adsorption time reduces.

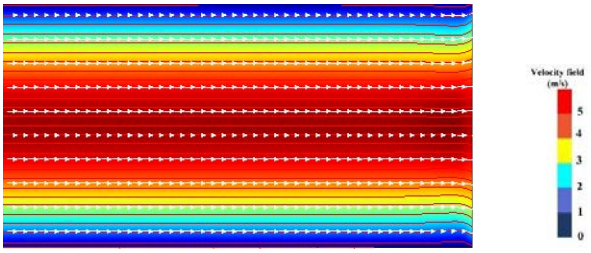


Fig 5. Fluid state in the microchannel without applying a voltage to the surface electrodes.

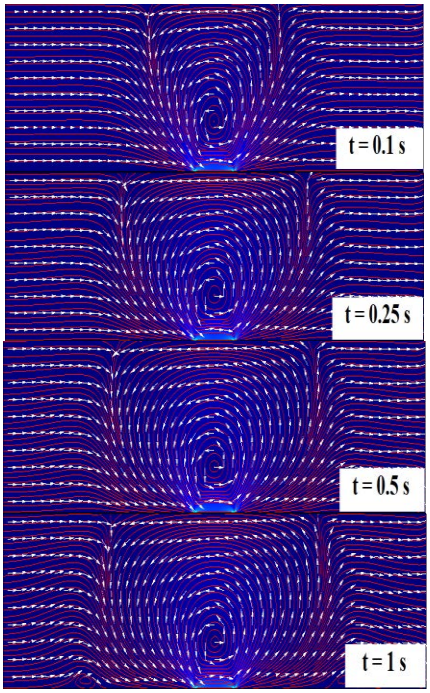


Fig 6. Changes in the turbulence state within the microchannel over time.

The required voltage to induce the most suitable turbulence depends on the microchannel dimensions, the electrode's distance from each other, fluid properties, and the initial velocity value.

VII. Conclusion

An ISFET-based biosensor for evaluation of PSA level is introduced and simulated in this work. Since the sensitivity and response time of this sensor depends on the proximity of analytes to the sensor surface, we have designed a microchannel on the sensor surface in which the electroosmotic flow is induced by applying external voltage. Adsorption of target ions on the ISFET surface modulated the channel's conductivity and changed the device's threshold voltage. Our simulation results have shown that the proposed structure facilitates rapid detection and measurement of PSA concentration.

REFERENCES

[1] Majji, Sankararao, Chandra Sekhar Dash, and Asisa Kumar Panigrahy. "Ion Sensitive Field Effect Transistor as a Bio-compatible Device: A Review." 2022 International Conference on Electronics and Renewable Systems (ICEARS). IEEE, 2022.

[2] Poghossian, Arshak, and Michael J. Schöning. "Recent progress in silicon-based biologically sensitive field-effect devices." Current Opinion in Electrochemistry 29 (2021): 100811.

[3] Ma, Xiaohao, et al. "Recent advances in ion‐sensitive field‐effect transistors for biosensing applications." Electrochemical Science Advances (2022): e2100163.

[4] Wadhera, Tanu, et al. "Recent advances and progress in development of the field-effect transistor biosensor: A review." *Journal of Electronic Materials* 48.12 (2019): 7635-7646.

[5] Mehta, Aditya, et al. "Machine Learning Techniques for Performance Enhancement of Si 3 N 4-gate ISFET pH Sensor." *2020 IEEE 17th India Council International Conference (INDICON)*. IEEE, 2020.

[6] Wang, Hui, and Naiyun Tang. "A modified TCAD simulation model for a-InGaZnO based ISFETs on GaAs substrate for pH sensing applications." *Materials Research Express* 8.9 (2021): 095901.

[7] Acar, Gizem, et al. "An ISFET Sensor-Integrated Micromixer for pH Measurements." *2020 21st International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)*. IEEE, 2020.

[8] Cao, Shengli, et al. "ISFET‐based sensors for (bio) chemical applications: A review." Electrochemical Science Advances: e2100207.

[9] Pindoo, Irfan Ahmad, and Sanjeet K. Sinha. "Increased sensitivity of biosensors using evolutionary algorithm for biomedical applications." *Radioelectronics and Communications Systems* 63.6 (2020): 308-318.

[10] Akbari jonous, Zahra, et al. "An electrochemical biosensor for prostate cancer biomarker detection using graphene oxide–gold nanostructures." *Engineering in Life Sciences* 19.3 (2019): 206-216.

[11] Crulhas, Bruno P., et al. "Electrochemical aptamer-based biosensor developed to monitor PSA and VEGF released by prostate cancer cells." *Analytical and bioanalytical chemistry* 409.29 (2017): 6771-6780.

[12] Arneth, Borros M. "Clinical significance of measuring prostate-specific antigen." *Laboratory Medicine* 40.8 (2009): 487-491.

[13] Abdolkader, Tarek M., et al. "ISFET pH-sensor sensitivity extraction using conventional MOSFET simulation tools." International Journal of Chemical Engineering and Applications 6.5 (2015): 346.

[14] Sant, William, et al. "On-line monitoring of urea using enzymatic field-effect transistors." Sensors and Actuators B: Chemical 160.1 (2011): 59-64.

[15] Cazalé, Arnaud, et al. "Study of field-effect transistors for the sodium ion detection using fluoropolysiloxane-based sensitive layers." Sensors and Actuators B: Chemical 177 (2013): 515-521.

[16] Sinha, Soumendu, et al. "Fabrication, characterization and electrochemical simulation of AlN-gate ISFET pH sensor." Journal of Materials Science: Materials in Electronics 30.7 (2019): 7163-7174.

[17] Dutta, Jiten Ch. "Ion sensitive field-effect transistor for applications in bioelectronic sensors: A research review." 2012 2nd National Conference on Computational Intelligence and Signal Processing (CISP). IEEE, 2012

[18] Muangsuwan, Wannaporn, et al. "Development of an immunoFET biosensor for the detection of biotinylated PCR product." Heliyon 2.10 (2016): e00188.

[19] . Madec, Morgan, et al. "Environment for modeling and simulation of biosystems, biosensors, and lab-on-chips." IEEE Transactions on Electron Devices 66.1 (2018): 34-43.

[20] Jiang, Yuting, et al. "Transient electroosmotic slip flow of fractional Oldroyd-B fluids." Microfluidics and Nanofluidics 21.1 (2017): 7.

[21] Bazant, Martin Z., and Todd M. Squires. "Induced-charge electrokinetic phenomena: theory and microfluidic applications." Physical Review Letters 92.6 (2004): 066101.

[22] Hu, Guoqing, Yali Gao, and Dongqing Li. "Modeling micropatterned antigen-antibody binding kinetics in a microfluidic chip." Biosensors and Bioelectronics 22.7 (2007): 1403-1409.

Fuzzy Logic Timing Control for Standard Crossroad Lights

Mohammad Hasan Shojaeefard [1], Morteza Mollajafari [2]* and Majid Talebi[3]

***Abstract*— Fuzzy logic can be arranged concerning practical experiences and blended with conventional control techniques. Even though it is not the replacement of conventional control methods in many cases, a fuzzy control system eases the implementation and design process. It has also been put in use in other matters like traffic control. Increasing vehicles and insufficiency of passages capacity have led to widespread traffic emergence. While it is very difficult to widen existing roads, optimizing traffic like control schemas is still possible. This paper assumes a common four-directional crossroad where vehicles can move in a bidirectional way from each direction. A Sugeno fuzzy logic set of rules is presented to regulate the timing schedule of green lights for the crossroad concerning the vehicle accumulation at each line.**

Index Terms—Fuzzy logic, Crossroad, Traffic lights control

I. Introduction

Traffic problem is a major inconvenience in most urban areas around the world. This transportation issue can affect the economy, slow development, reduce production, increase cost, and distort every aspect of social life. This phenomenon causes increasing vehicles, insufficient roads and highways, and unfit traffic lights control strategies. All of these factors can create traffic congestion in the intersection, but a traditional traffic light system is one of the major factors. Traffic signals are common features of urban areas worldwide; their main purpose is to improve the traffic qualities at the intersection, maximizing the capacity at the intersection and minimizing the delays by controlling the number of vehicles that can pass at each timing phase [1]. A fuzzy logic controller allows linguistic and inexact traffic data to be applied in controlling the signal timings. A fuzzy control system is a rule-based control system characterized by expressing an expert's control rules using a fuzzy theory and determining a control command by a fuzzy inference program.

The idea of a fuzzy traffic signal controller is about modeling a control strategy based on expert human knowledge [2, 3]. In a conventional traffic light controller, the traffic lights change at a constant cycle time which is not the optimal solution when the density of passing vehicles varies with time. Obviously, it would be more feasible and sensible to pass more cars at the green interval if fewer cars were waiting behind the red lights at other lines [4, 5]. The fuzzy logic theory is introduced for the traffic controller to adopt a green interval response based on dynamic traffic load inputs instead of default pre-set timing. In this paper, we proposed a new fuzzy traffic light control system that can effectively handle confusing traffic situations with interference and a long queue of vehicles waiting at the red light. It can adaptively manage green phase lengths according to the traffic frequency and waiting for a queue. The rest of the paper is organized as follows. In the next section, some previous and close works on this problem are presented and discussed. In Section 3, the proposed method is formulated, and its algorithm is presented in detail. The simulation results are provided and discussed in Section 4, and finally, the paper is concluded in section 5.

II. Previous Research

Many researchers have already researched fuzzy logic control methods; in [6], researchers have established fuzzy rule base and expert systems to control various traffic situations, which began to apply fuzzy control algorithms in traffic management. Based on this in [7], the authors designed a multi-level fuzzy controller that may even alter the fuzzy logic upon severe condition changes [8]. In [9], researchers applied fuzzy logic to control two adjacent intersections with the same situations. This controller determined the extension or termination of the green signal based on the upstream traffic. In [10], a fully distributed system with local cooperative controllers to self-organizing traffic signal control is applied in a multi-intersection network.

In [11], a two-layer fuzzy control algorithm for traffic control of the network is proposed, which is supposed to have a large traffic flow and a high possibility of congestion. This work is also reviewed and improved in [12]. This paper indicated better performance of fuzzy-based controllers than traditional traffic signal controls, specifically during uneven and heavy traffic conditions.

Researchers in [13] have presented a hybrid algorithm that combines Fuzzy Logic Controller and Genetic Algorithms. The Genetic Algorithm has been used to adapt the decision rules of Fuzzy Logic Controllers (FLC). Others have also done some research on the finite-time adaptive fuzzy tracking control issue [14, 15]. Others proposed an optimal general type-2 fuzzy controller for urban traffic networks [16]. The general type-2 fuzzy logic sets and the modified backtracking search algorithm

1- Iran University of Science and Technology
2- Iran University of Science and Technology
3- Iran University of Science and Technology
Corresponding author: mollajafari@iust.ac.ir

techniques control the traffic signal scheduling and phase. Aiming at the coordination and dynamic uncertainty problem in arterial traffic, other scholars offered a type-2 fuzzy coordination arterial traffic control method [17]. Aiming at the situation that the timing cycle changes with time, Mo Hong used the time-varying universe and parallel system theory to study the dynamic linguistic trajectory of the timing evolution process [18,19].

The traditional control strategy of traffic lights mostly uses a fixed timing method, whose cycle, green time and phase structure are invariable [20,21]. But all researchers agree that it is hard to achieve optimal control for the variable and complicated traffic flow [22–25]. It is imperative to optimize the phase structure and select the appropriate timing method to adapt to the changing traffic flow [26-27].

For the fuzzy logic system, Q-learning and neural networks commonly used in traffic control, scholars have compared their control effect in traffic control [28]. The results show that fuzzy logic has better performance than other methods. Thus in the current work, the fuzzy control under time-varying conditions is used to realize an adaptive timing table for a traffic signal. Compared with the traditional fixed timing order, the results show that a rationally tuned adaptive approach will improve the traffic qualities under asymmetric traffic flow.

In addition, Cang Zhou in [26] has used the Mamdani-type of the fuzzy logic system, a timing project with phase optimization based on the time-varying universe proposed for the situation where the traffic control is unreasonable. Firstly, the cycling universe is determined by the base word combination of traffic flow in each lane. Secondly, the key traffic volume and average maximum queue length are inputs to determine each phase's green time. Thirdly, when the phase structure does not match the traffic flow, the phase structure is optimized by adding additional phases or merging phases. Finally, VISSIM and MATLAB simulation software compares the original and optimization timing projects proposed in this paper. The delay increases the capacity during peak hours.

In [29], a predictive controller is proposed for urban traffic in which state-space dynamics are used to estimate the number of vehicles at an isolated intersection and its queue length. The simulation results show that traffic volumes and car crashes are reduced by controlling traffic on an urban road using model predictive control.

The authors of the previous research proposed a new stable TS (Takagi–Sugeno) fuzzy controller for urban traffic [30]. Their approach formulates the state-space dynamics for the vehicle's average waiting time at an isolated intersection and the length of queues. Then, a fuzzy intelligent controller is used to control the lights based on the queue length. The results show that the performance of the control and its stability are proved using the Lyapunov theorem.

[31] developed a method based on the fuzzy factors of the intensity of pedestrian flow and flow discontinuity to improve the accuracy of modeling and predicting the traffic capacity of intersections. The results are analytical, and no simulation results are provided.

The main contribution of this paper is using the fuzzy logic in predicting, optimizing and reducing the traffic and the number of stopped cars at crossroads. Meanwhile, in many recent projects, artificial smart or intelligent vision systems have been used. The results obtained according to the fuzzy inference output chart are clear: the number of stopped cars increased, whether in total, in total mean, or even in separate mode. Moreover, in the proposed approach, each car's waiting time is less than those of smart modes and system controls.

III. Current Work

The field of operation is presumed to be a common crossroad with four directions of passages, as shown in Fig. 1. In each direction, vehicles can either enter or leave the crossroad.

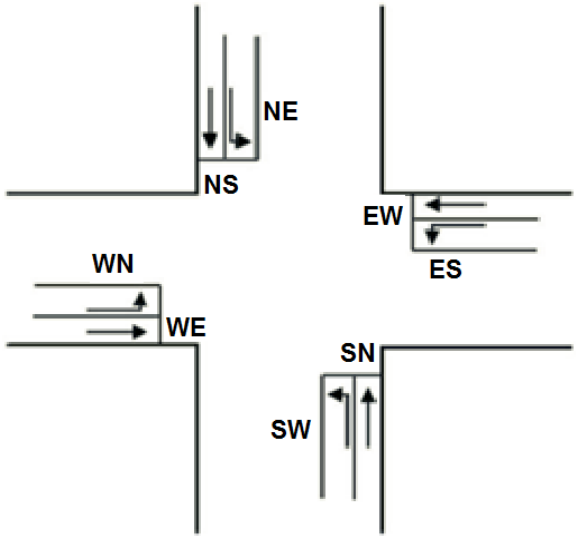


Fig. 1. The important directions of motion at this crossroad are considered from West to East (WN), West to North (WN), and similarly NS, NE, EW, ES, SN, and SW. Vehicles that leave the area or turn to the right don't have any deterministic effect on the traffic situation here.

Obviously, motions with the most queue should be given the most time. According to the practical data provided in [26] (driven from standard Chinese urban traffic scales), a total timing sequence of traffic lights in such a crossroad should be between 80-220[1] seconds, and vehicles queued in the line should count less than 100 for each direction and 600 for the whole quarter. It is assumed that the vehicle numbers are countably utilizing modern intelligent surveillance monitoring systems presented at each important passage.

Thus assuming the total cars in the quarter as our first input independent variable called TC, this parameter is a subset of Low (TC<200), Medium (200<TC<400) and High (TC>400) traffic density. Total cycle duration (TCD) is a dependent variable on the TC. According to the Fuzzy Inference System Act, it should be considered as three areas for the total car waiting time. Since the plan, the total period of stopping at the crossroads of 80 to 220 seconds can be reached in three-level

[1] Quantities may vary slightly considering a necessary tolerance

"low", "moderate," and "long" divided time intervals and also According to international law, the cars do not stop for the right-turn, so it is not considered in this article. The first set of fuzzy rules are now set as below:

TABLE I
Sugeno Fuzzy Logic Rules for TCD Concerning the TC

Rule 1	If TC is low, then TCD is short (80-120)	TCD=80+0.2*TC
Rule 2	If TC is medium then TCD is mediocre (120-160)	TCD=100+0.15*TC
Rule 3	If TC is high then TCD is long (160-220)	TCD=40+0.3*TC

The TCD parameter naturally contains all the crossroad's red, yellow and green light times. Considering the existing 8 lines, when the TCD is set to be 80 seconds, and the car queues at all lines are equal, it means that each line has 10 seconds of green light, 20 seconds of yellow and 50 seconds of red light.

After determining the cycle duration, it is now turned to dividing this duration between 8 lines. The direction with the highest Local Queue Density (LQD) should be allocated with the longest Green Light Duration, and the sum of all eight GLDs should not pass the authorized TCD. Here we take a simple solution to assign green time to each line according to its queued vehicle quantity. If TC_i is the number of waiting for cars at each line i, a ratio can be considered as a priority right like (1, 2):

$$P_i = \frac{TC_i}{TC} \quad (1)$$

Obviously:

$$\sum_{i=1}^{8} P_i = \frac{\sum_{i=1}^{8} TC_i}{TC} = \frac{TC}{TC} = 1 \quad (2)$$

Each Pi can be a real number between 0-1. The line with the highest priority can be let first to green light.

The time quota from the TCD for each car to pass among the whole TC can be assumed as:

$$Ct = \frac{TCD}{TC} \quad (3)$$

And finally, each line green timeshare can be calculated as:

$$GT_i = C_t * TC_i \quad (4)$$

And

$$\sum_{i=1}^{8} GT_i = TCD \quad (5)$$

However, this approach will cause an allocation time shortage when too few or too many vehicles are in line. For example, when there are less than ten cars in a line, the assigned GT may run below 3 seconds, which is insufficient to pass. Therefore, another set of fuzzy rules is introduced to compensate for this matter:

TABLE II
Sugeno Fuzzy Logic Rules to Assign Suitable GT for Each Line

Rule 1	If TCi is low (TCi<30), then GTi is short	$GT_i = A_i + C_t * TC_i$
Rule 2	If TCi is medium, then GTi is mediocre	$GT_i = B_i + C_t * TC_i$
Rule 3	If TCi is high (TCi>60), then GTi is long	$GT_i = C_i + C_t * TC_i$

For an instant, it seems that constants $A_i$between3-5, B_i between 5-10 and C_i between 10-15 seconds would suit well; however, these constants can be further optimized in practice noticing that the sum of all GTs should not pass 1.2*TCD. Consequently:

$$\sum_{i=1}^{8} A_i < 0.2 * TCD \quad (7)$$

The algorithm will re-assign A, B, and C for each line to adhere to the above condition. Beginning from the upper limit for each one and reducing it if necessary. Considering these, a better set of fuzzy rules for this stage can be written below.

TABLE III
Accurate Sugeno Fuzzy Logic Rules to Assign Suitable GT for Each Line

Rule 1	If TCi is low (TCi<33), then GTi is short	$GT_i = 3 + \frac{TC_i}{15} + C_t * TC_i$
Rule 2	If TCi is medium, then GTi is mediocre	$GT_i = 5 + \frac{TC_i}{12} + C_t * TC_i$
Rule 3	If TCi is high (TCi>66), then GTi is long	$GT_i = 10 + \frac{TC_i}{20} + C_t * TC_i$

This set of fuzzy rules must be applied for all GTs of all directions. After recalculating the GTs, the real cycling duration (RCD) will be the sum of all GTs. This may differ from the predicted TCD.

IV. Simulation and Results

While the proposed algorithm is not sensitive to daily cycles, its application will still be in daily life. Thus for the simulation, a 24 hours cycle is assumed, but the quantities of incoming cars are assumed as independent toward daily hours. At the end of each TC cycle, the TCi numbers are re-extracted the set of rules and parameters are updated.

The TC (and subsequently TC_i) is a random independent input variable that varies between 0-600. The TCD parameter is first considered constant and then recalculated with our fuzzy rules. Pi, Ct, GTi and AWT must be calculated for both cases. The simulation is repeated each time that TCD cycling time is passed and subtracted from the daily time duration in seconds.

In a sample simulation, the total number of 409 cycles was

allocated (this number depends on the total passing cars as a random input variable and our fuzzy rules). The results are as follows.

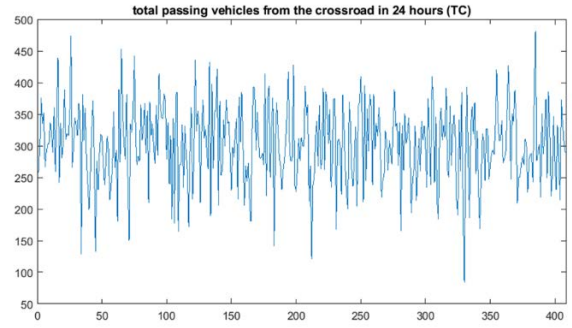


Fig. 2. The total quantities of passing cars per cycle were between 75 and 475

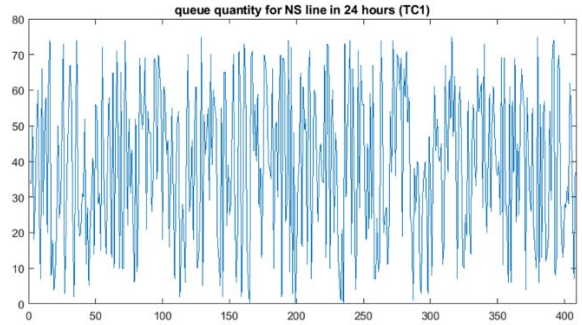


Fig. 3. The queued cars in the North to South line at each cycle

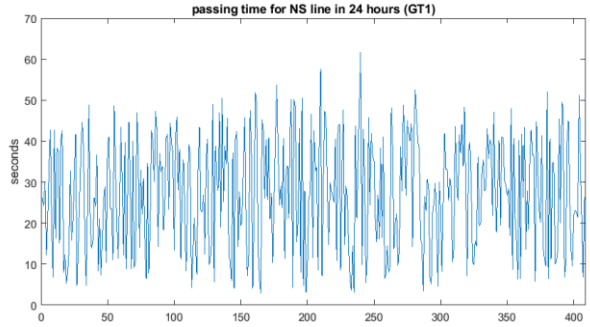


Fig. 4. The green light time for NS passage at each cycle has always been always been between 3-62 seconds

Each line's average green light time was about 30 seconds.

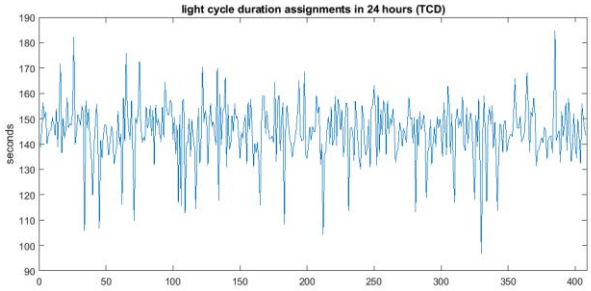


Fig. 5. Each cycle duration is in seconds. The threshold of 220 seconds has not been violated, and cycles were usually less than 3 minutes lengthened.

Fig. 6. is the most important parameter on the topic since it shows the time wasted at the crossroad for the vehicles. One may claim that a constant timing cycle, for example, at the length of 2 minutes, could be shorter on most occasions. Still, there are no guarantees that this constant timing of 15 seconds for each line to pass could suffice for all the vehicles in a queue to pass, and the line could be accumulated with incoming cycles. This will lead to an unpredictable situation of long queues which may be delayed up to many hours while the current fuzzy logic-based solution ensures that all the vehicles will receive enough time to pass and the crossroad will be vacant at the end of each cycle from the previous cycle's passengers even though it sometimes has violated the threshold of the maximum recommended light cycling time.

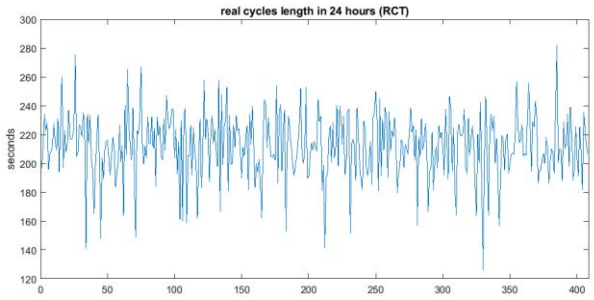


Fig. 6. The real length of each passing cycle per 24 hours

TABLE IV
Parametric Comparison Between Methods (seconds).

	Current method	The method presented in [25]	The method presented in [26]
average green light time	30	45	40
Average cycling time	200	240	220

This article is based on a fuzzy model of an urban traffic network designed for a single intersection. The model's state variables are the length of queues and the average waiting time for the vehicles. Moreover, the vehicle's average waiting time at an isolated intersection and the length of queues are considered controller inputs. Additionally, the effectiveness of the suggested controller is verified by simulation results. The percentage of improvement with attention to simulation results using the proposed method decreases the vehicles in each intersection phase to the fixed time Control, as shown in Figures 7 and 5.

Moreover, the presented method does not bind the situation to any specific time over the day-night conditions like early morning traffic. The current approach can quickly adapt itself to any unpredicted traffic conditions. However, interested researchers can still work on further optimizing the fuzzy rules set presented here. As seen in Figure 8, the proposed can effectively reduce the delay.

TABLE V
Vehicle Queue Length Performance Results from Fuzzy Intelligent and Fixed Time Control

Queue times	Fixed-time Controls [25]	Fuzzy Intelligent Control [Proposed]	Total Improvement Percentage
7:30 – 8:30	85	20	76.47
10:30 – 11:30	30	8	73.33
15:30 – 16:30	45	15	66.67
18:30 – 19:30	320	67	79.06
SUM	480	110	77.08

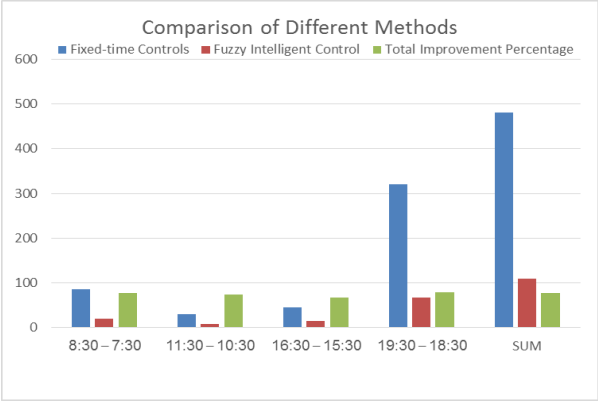
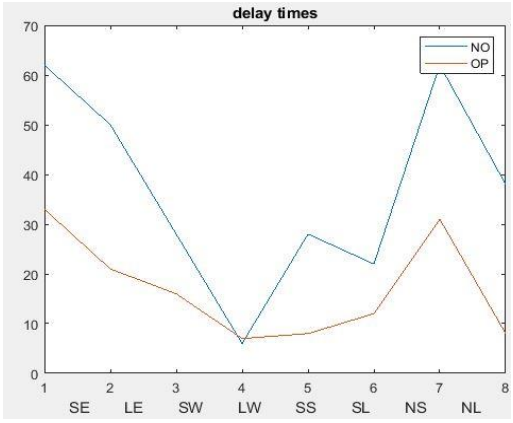
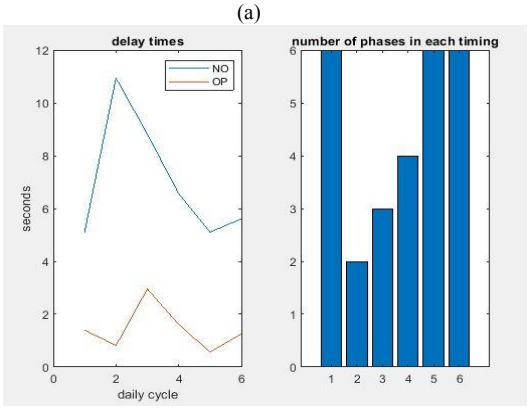


Fig. 7. Comparison of results in the length of the queue of vehicles using a fuzzy model.



(a)



(b)

Fig. 8. (a) Comparison of results in delay times between the method with and without optimization, (b) impact of the number of phases in each timing on the delay times.

The novelty of the proposed method using fuzzy logic in predicting, optimizing, and reducing traffic and stopping cars at crossroads. Meanwhile, in many recent projects, artificial smart or intelligent vision systems have been used. The results obtained according to the fuzzy inference output chart are quite clear: the number of car stops, whether in total and in total mean and even in a separate mode, each car's waiting time is less than smart mode and system controls. In addition, the presented method does not bind the situation to any specific time over the day-night conditions like early morning traffic. The current approach can quickly adapt itself to any unpredicted traffic conditions. However, interested researchers can still work on further optimizing the fuzzy rules set presented here.

V. Conclusion

This paper proposes a fuzzy logic predicting method that optimizes traffic and stopped cars at crossroads. The fuzzy inference output chart of the proposed method clearly shows that the amount of stopped cars and their waiting time at crossroads is less than those of intelligent vision-based controlling systems. Moreover, the presented method does not bind the situation to any specific time over the day-night conditions like early morning traffic. The current approach can quickly adapt itself to any unpredicted traffic conditions. In the future, we will work on further optimizing the fuzzy rules set presented here.

VI. Biographies

Prof. **Mohammad Hassan Shojaeefard** was born in 1950 in Jahrom, Iran. He completed his B.Sc. degree in mechanical engineering from the Iran University of Science and Technology (IUST) in 1977. He also successfully attained an M.Sc. degree in mechanical engineering from the University of Birmingham, UK, in 1983. Prof. Shojaeefard has held a Ph.D. in mechanical engineering from the University of Birmingham, the UK, since 1987. Now, he is the Professor and Head of Automotive Eng. Research Center (AERC) & Automotive Eng. Dept. (AED) (IUST). The Specific topics of interest of Prof. Shojaeefard are fluid dynamics, internal combustion engines and automotive engineering.

Morteza Mollajafari received his BS in electrical engineering from the Electrical Engineering Department at K.N. Toosi University of Technology in 2008; and his MSc. and a Ph.D. degree in electrical engineering from Iran University of Science and Technology (IUST) in 2011 and 2017, respectively. He is an assistant professor in the automotive engineering department at the IUST. His research areas include parallel and distributed computing systems, especially Cloud Computing, developing scheduling algorithms, Evolutionary Algorithms and autonomous vehicular systems.

Majid Talebi received BS degree from the Azad University of Bojnord, Bojnord, Iran, in 2006, where he is currently pursuing the master's degree in Digital Electronics Engineering with the school of Automotive Engineering, Iran University of Science and Technology. His research interests include machine learning, data mining, deep learning, machine vision, and optimization.

VII. References

[1] Rustem, S., Regina, N., "Traffic safety system management", *Transportat. Res. Procedia.* 36, 676–681 (2018)

[2] Mohan, D., “Traffic safety: Rights and obligations”, *Accidents Analysis and Prevent.* 128, 159–163 (2019)
[3] Li, Z.H., Cao, Q., Zhao, Y.H., Tao, P.F., Zhou, R., “Krill herd algorithm for signal optimization of cooperative control with traffic supply and demand”, *IEEE Access*. 7, 10776–10786 (2019)
[4] Dotoli, M., Fanti, M.P., Meloni, C., "A signal timing plan formulation for urban traffic control", *Control. Eng. Pract*. 42(3), 1297–1311 (2016)
[5] Akcelik, R., “Traffic signals: capacity and timing analysis”, *Transportat. Res*. 5(6), 505–505 (2011)
[6] Webster, F.V., “Traffic signal settings”, *Technical Report. HM, Stationery Office* (2018)
[7] Pappis, C.P., Mamdani, E.H., “A fuzzy logic controller for a traffic junction”, *IEEE. Transact*. Syst. Man. Cybernet. 7(10), 707–717 (2017)
[8] Chen, H., Chen, S.F., “A method for real-time traffic fuzzy control of a single intersection”, Informat. Control. 36(3), 227–233 (2017)
[9] Nakatsuyama, M., Nagahashi, H., Nishizuka, N., “Fuzzy logic phase controller for traffic junctions in the one-way arterial road”, *In IFAC proceedings series*. 2865C2870 (2015)
[10] Chiu, S., Chand, S., “Self-organizing traffic control via fuzzy logic”, In: Proceedings of the 32nd IEEE conference on decision and control. 2, 1897C1902 (2013)
[11] Zhang, W.B., Wu, B.Z., Liu, W.J., “Anti-congestion fuzzy algorithm for traffic control of a class of traffic networks”, In IEEE international conference on granular computing. 124 (2007)
[12] Rahman, S.M., Ratrout, N.T., “Review of the fuzzy logic-based approach in traffic signal control: prospects in Saudi Arabia”, *J. Transportat. Syst. Eng. In format. Technol*. 9, 5870 (2009)
[13] Odeh, S.M., Mora, A.M., Moreno, M.N., Merelo, J.J., “A hybrid fuzzy genetic algorithm for an adaptive traffic signal system”, *Adv. Fuzzy. Syst*., (2015).
[14] Sun, W., Wu, Y.Q., Sun, ZY, “Command filter-based finite-time adaptive fuzzy control for uncertain nonlinear systems with prescribed performance”, *IEEE. Transact*. Fuzzy. Syst. 10(1109), 2967295 (2020)
[15] Sun, W., Su, S.F., Wu, Y.Q., Xia, J.W., “A novel adaptive fuzzy control for output constrained stochastic non-strict feedback nonlinear systems”, *IEEE. Transact. Fuzzy. Syst*. 10(1109), 2969909 (2020).
[16] Khooban, M.H., Vafamand, N., Liaghat, A., Dragicevic, T., “An optimal general type-2 fuzzy controller for Urban Traffic Network”, *ISA. Transact*. 66, 335–343 (2017).
[17] Bi, Y.R., Lu, X.B., Sun, Z., Srinivasan, D., Sun, ZX, “Optimal type-2 fuzzy system for arterial traffic signal control”, *IEEE. Transact. Intell. Transportat*. Syst. 19(9), 3009–3027 (2018)
[18] Li, R.M., Jiang, C.Y., Zhu, F.H., Chen, X.L., “Traffic flow data forecasting based on interval type-2 fuzzy sets theory”, *IEEE/CAA J. Autom. Sinica*. 3(2), 141–148 (2016)
[19] Cao, X.L., Mo, H., Zhu, F.H., “Fuzzy control of timing for traffic lights based on the time-varying universe”, *Measure. Control. Technol*. 38(11), 115–120 (2019).
[20] Mo, H., Hao, X.X., Zheng, H.B., Liu, Z.Z., Wen, D., “Linguistic dynamic analysis of traffic flow based on social media case study”, *IEEE Transact. Intell. Transport. Syst.* 17(9), 2668–2676 (2016).
[21] Henrique, D., Norian, M., Furio, D., “Genetic algorithm-based traffic lights timing optimization and routes definition using Petri net model of urban traffic flow*”, IFAC. Proceed*. Vol. 47(3), 11326–11331 (2014).
[22] Talab, H.S., Mohammadkhani, H., Haddadnia, J., “Controlling multi variable traffic light timing in an isolated intersection using a novel fuzzy algorithm”, *Journal of Intelligent Fuzzy Systems* 25(1), 103–116 (2013).
[23] Maythem, K.A., Mohd, N.K., Madzlan, N., Brahim, BS, Marwan, A., “High accuracy traffic light controller for increasing the given green time utilization”, *Comput. Electri. Eng*. 41, 40–51 (2015).
[24] Shahhoseini, H. S., Saleh Kandzi, E., Mollajafari, M. “Nonflat surface level pyramid: a high connectivity multidimensional interconnection network,” *The Journal of Supercomputing*, 67(1), 31-46, (2014).
[25] Araghi, S., Khosravi, A., Creighton, D., “A review on computational intelligence methods for controlling traffic signal timing”, *Exp. Syst. Appl*. 42(3), 1538–1550 (2015).
[26] Cang Zhou, “Fuzzy Control under Time-Varying Universe and Phase Optimization in Traffic Lights (ICSSE 2020)”, *Int. J. Fuzzy Syst*. (2021).
[26] Mollajafari, M., Shojaeefard, M. H. “TC3PoP: a time-cost compromised workflow scheduling heuristic customized for cloud environments,” *Cluster Computing*, 24(3), 2639-2656, (2021).
[27] Mollajafari, M., Shahhoseini, H.S. “A repair-less genetic algorithm for scheduling tasks onto dynamically reconfigurable hardware.” *International Review on Computers and Software* 6.2, 206-212, (2011).
[28] Shojaeefard, M., Mollajafari, M., Mousavitabar, S., Khordehbinan, M., Hosseinalibeiki, H., "A TSP-based nested clustering approach to solve multi-depot heterogeneous fleet routing problem," Rev. int. métodos numér. cálc. diseño ing, 38 (1), pp. 1-14, (2022).
[29] Jafari, S., Shahbazi, Z., & Byun, Y. C. “Improving the Performance of Single-Intersection Urban Traffic Networks Based on a Model Predictive Controller,” Sustainability, 13(10), 5630, (2021).
[30] Jafari, S., Shahbazi, Z., & Byun, Y. C.. “Traffic Control Prediction Design Based on Fuzzy Logic and Lyapunov Approaches to Improve the Performance of Road Intersection. Processes," 9(12), 2205, (2021).
[31] Shepelev, V. D., Glushkov, A. I., Makarova, I. V., & Boyko, A. “Clustering Urban Transport Network Junctions Using Convolutional Neural Networks and Fuzzy Logic Methods,” Transportation Research Procedia, 62, 581-588, (2022).

Evaluation of Ground Systems Performance on Rail Potential and Stray Current in Tehran Railway System

Ebrahim Zare Juybari[1]* and Ahmad Gholami[2]

Abstract— **DC Stray currents represent serious problems for any electrified railway system. This paper describes a complete simulation model for evaluating rail potential and stray current for the Tehran multi-train DC railway system based on different earthing systems. In the present paper, a complete analysis of the results is presented. The simulation results clearly indicate that earthing system arrangements and other parameters influence stray currents and rail potential. The simulation results show the performance of three ground system types: grounded, ungrounded, and diode-grounded systems on rail potential and stray current.**

Index Terms— DC Railway System, Multi-Train Simulation, Rail Potential, Stray Current, Monte Carlo Simulation

I. Introduction

Most railways systems use the running rails as the return conductor for traction current to reduce the wiring cost. The disadvantages of such an arrangement are those of rail potential and stray current problems [1, 2]:

- Rail potential may rise or fall above a certain threshold and can be hazardous in the forms of touch or step voltages. Rail potential should not exceed the critical value of correlative standards to avoid endangering the person's safety.
- Stray currents can dramatically create or accelerate the electrochemical corrosion of underground metallic utility pipes, power and telephone cables and the earthing grids laid in the vicinity of the DC railway system.

Therefore, controlling rail potential and stray current is important in all DC railway system designs. Rail potential and stray current are influenced by factors such as the conductance per unit length between running rails and earth, rail resistance, distance between traction substations, line voltage, weather, position and loads condition. Among them, the earthing strategies profoundly influence both rail potential and stray current [3, 4].

The present paper deals with modeling and simulation of a multi-train DC railway system to determine rail potential and stray current values in different structures (earthing strategies) for different conditions of the Tehran metro system. The metro is used with a direct current power supply for short distances in urban transportation. An alternating current power supply is suitable for distances of more than 30 km between two stations, and for distances over 200 km, magnetic levitation trains are used [5, 6]. Stray currents that crawl through rails into the earth can flow through building foundations, pipes, and other underground metallic installations with less resistivity concerning the soil to reach the negative terminal of the substation. The points at which current enters/leaves the metallic structure are called the cathodic/anodic regions [7- 9]. The current leaving the anodic region causes the corrosion phenomenon on the metallic structure. Various methods have already been used for stray current modeling in railway transit systems. In most of the literature and simulations, such as in [10- 12], the metro system is presented as resistive grids, and the mathematical equations are given based on the traction substation current, rail track resistance, the distance between the train and substation, the resistance between the rail and earth, and stray current. In summary, the contributions of the study are as follows:

• Analysis and comparison of different types of metro ground systems by TPS and ENS simulation;

• Presenting and comparing the effect of longitudinal resistance of running rails conductivity of rails to ground on stray current and rail potential;

The rest of the paper is organized as follows. Section 2 applied the Block diagram of train motion equations and presented mathematical and analytical relationships. Section 3 Speed and power simulations are performed for the DC metro line in Tehran. In Section 4, the simulation results obtained from the software are studied and analyzed.

II. Modeling and Simulation of DC Electrified Railway System

Simulation of the railway system is complicated for its multiple nonlinear and time-varying equalities. Because it is composed of many components, the relation between them changes in time. Anyway, the Simulation of the DC Electrified Railway System was conducted in the following three stages:

• Train performance simulation
• Electric network simulation

1- Electrical Engineering Faculty, Imam Mohammad Bagher, Technical and Vocational University, Mazandaran, Iran

2- Electrical Engineering Faculty, University of Science and Technology, Tehran, Iran
Corresponding author: ebrahim.zare.juybari@gmail.com

• Rail potential and current ground simulation

2.1 Train Performance Simulation

2.1.1 Train Motion Equation (TME)

Train motion can be best described by Lomonossoff's equation [13, 14]:

$$(M + DM)\frac{dv}{dt} = TE - (a + bv + cv^2) - Mg(\alpha - \alpha') \quad (1)$$

Where M is the mass of the train, DM is the dynamic mass of the train (typically 5%-10% of M), v is train speed, TE is tractive effort, "a,b,c" are running resistance coefficients of the train, g is gravity, α is a slope at the current position, and α' is the curvature at current position converted to equivalent slope according to the following:

$$\alpha' = \frac{\mu(L+G)}{2\times R}\times 10^6 \quad (2)$$

Where G is track gauge (m), L is axial length (m), R is the radius of curvature (m), and μ is wheel-rail adhesion coefficient. The adhesion coefficient is empirical and can be approximated by Curtius Kniffler's formula for "dry rail" conditions [13, 14].

$$\mu = 0.161 + \frac{7.5}{44 + 3.6v} \quad (3)$$

The running resistance coefficients depend on the number of axles, total train length, and train cross-section area. The tractive effort of electric locomotives depends on various factors but mainly on the weight and speed of the train. For metro operations with a high passenger density, the variation in train weight can significantly affect train performance and energy consumption. With a time-based (time of day) passenger profile for each station, the weight of trains can be altered [14, 15].

Usually, the locomotive tractive effort determines as a function of speed and line voltage. Therefore, for each time snapshot, the output mechanical power of the train (motors) is obtained by equation (4). The total input electrical power (P_{tot}) of the locomotive measured at the pantograph is calculated by equation (5). The auxiliary power, electrical losses and mechanical losses, as well as all other losses, are included in P_{tot} [14, 15, 16]:

$$P(v) = TE(v)\times v \quad (4)$$

$$P_{tot} = \frac{P(v)}{\eta_t} \quad (5)$$

$$I(v) = P_{tot}(v)/V_{dc}(v) \quad (6)$$

Where $P(v)$ the tractive power, η_t the total power efficiency; $V_{dc}(v)$ the supply voltage of the train,the train current.and $I(v)$

Train motion equations (*TME*) are then solved using the built-in integration methods equipped with the algebraic loop solver to account for train speed, position, and line voltage effects. The Block diagram of single train dynamics is shown in Fig. 1. It consists of three major modules, i.e., the Trains Performance Simulator (*TPS*), the Electric Network Simulator (*ENS*), and the Train Movement Simulator (*TMS*). The railway system's operation timetable (T-V* profile) is applied to define each train's location and speed profile along the main line for each time snapshot. In reality, the *T-V** profile corresponds to either the speed code received from the track signaling system or the speed demand set by the driver during operation. The speed control module produces the desired *TE / BE* to achieve the desired speed. The desired *TE* demand is only subject to acceleration and jerk limits [12, 14].

The actual speed and position are computed using the train motion equation (*TME*). Traction power (*TE*) or braking power (*BE*) is obtained by equation (4). If trains use an inverter propulsion system, then the efficiency characteristics are used to calculate the electric power demand. Similarly, train regeneration power output can be obtained from efficiency characteristics [12, 14].

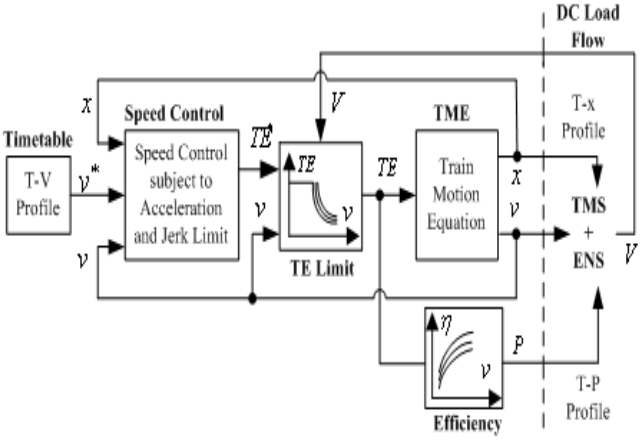


Fig. 1. Block diagram of train motion equations as related to *TPS*, *TMS*, and *ENS*

2.1.2 Monte Carlo Simulation of Operation Timetable Delays

The Monte Carlo method is a stochastic simulation technique implemented to solve a category of problems dealing with the variables of random nature [17]. Timetable delays per train are considered as the random variable in this study. Obviously, the probability that a train leaves the station with delay is much more than that it leaves at the scheduled time (T_{sch}). Therefore, triangular or normal distribution for this variable (the departure time) is a proper choice. The simulator output will include the number of trains of each line (headway), timetable delays per train, delays per trip, total and average delays and average and maximum speed. This information is available as a "per train", "per day" or summary output.

2.1.3 The AC/ DC Power Supply System Model

The power supply system is one of the most important facilities in electric railway systems. It provides the necessary energy for the operation of trains and the load at stations and depots. The Bulk Supply Substations (BSS) connect the railway power supply system to the national supply at several locations and step the high-voltage (63 kV) down to medium-voltage (20 kV). Then, it distributes a 20 kV supply to the traction supply

substation (TSS), the Station Supply Substations (SSS), and the depots. In TSS**,** 20 kV voltage is stepped down to low-voltage (592V AC) and (12 pulses) rectifiers are used to convert the AC power supply to 750 (780) V DC power source for train sets [18].

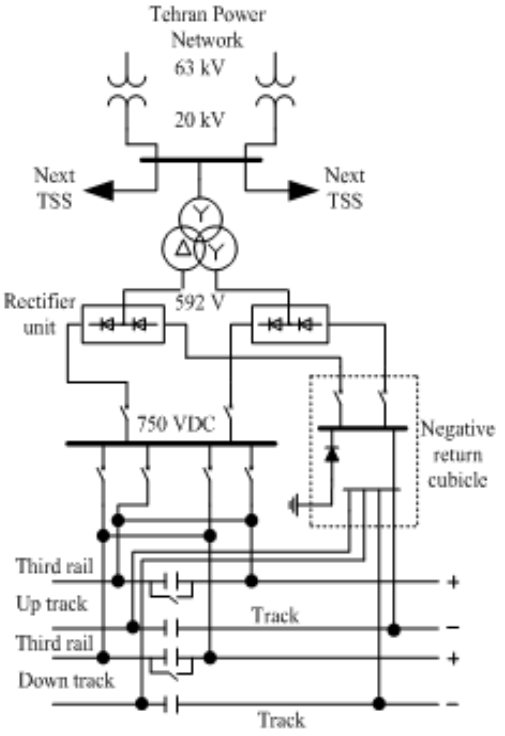


Fig. 2. Power supply system of propulsion traction for Tehran railway system

As shown in Fig. 2, the positive bus of TSS is connected to the third rail and the negative bus to the running rails. To prevent cathode corrosion phenomena due to the stray current, the negative return of DC is collected by the running rail and connected to the negative terminal of the rectifiers in TSS.
The train sets, operated along the main lines, pick up DC power from the third rail for traction motors to drive the train set.

2.1.4 Traction Motor Model

For each train set, several units of motors (DC motors or three-phase induction motors with VVVF[1] inverters) are used to provide propulsion power. An equivalent motor is used to represent several traction motors in each train. The motor has an equivalent rating (voltage, power, etc.).

2.2 Electrical Network Simulation

Having obtained the location and power consumption data for each train, the equivalent electrical network for each snapshot or specific time can be configured, as shown in Fig. 3. Each train set is treated as a load bus in the load flow analysis, and the number of train sets to be put into operation for each metro line is determined by scheduled headway. The impedance matrix of the DC network is updated according to the actual locations of all train sets. There are several methods to solve a railway traction power network. Then the voltage drop or power flow can be evaluated by electrical network simulation (ENS), which leads to the current flow into or out of the train, depending on the train's operation mode (i.e., acceleration, constant-speed, or deceleration). The current outflow from each traction substation (TSS) was also evaluated. Then, the rail potential and current ground simulation were followed by simulating each substation as a current source and each train as a current sink or a current source, depending on the train's operation mode [17, 18].

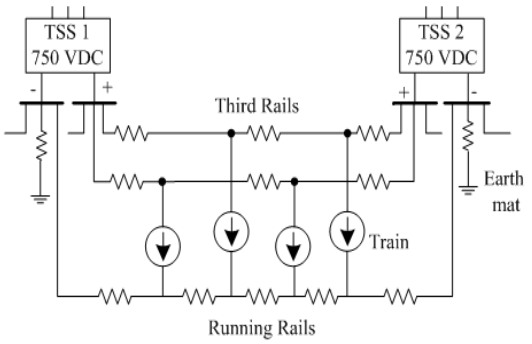


Fig. 3. Multi-train and double-tracks model of DC electrified railway at a snapshot or specific time

2.3 Rail Potential and Ground Current Simulation

The rail potential and current ground simulation were followed by simulating substations, trains, running rails, collection mats (if any), and substation grounding.

2.3.1 Running Rail Model

The rail potential and current ground equations can be obtained by a simple model based on DC transmission line equations under stationary conditions. These equations describe the system represented in Figure 4, constituted by a linear rail embedded in the soil. Where R_R is DC rail resistance per unit length $[\Omega/m]$, G is DC rail to ground conductance per unit length [*S/m*].
Rail-to-earth conductance can vary depending on the type of construction, amount of dirt on rail fasteners, temperature, ballast condition, and ambient moisture. If assumed that the rail is a cylindrical conductor with a radius $r_0 (<< d)$ embedded in the ground, then the conductance per unit length can be calculated as [19, 20]:

$$G = 2\pi/[\rho.\ln(2d/r_0)] \quad (7)$$

Where ρ is soil resistance $[\Omega.m]$, and d is rail length [*m*].

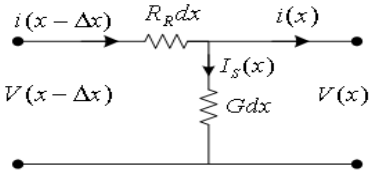


Fig. 4. DC transmission line model

2.3.2 Rail Potential and Stray Current Modeling for Multi Trains System

[1] - Variable Voltage and Variable Frequency

If rail resistance and the rail-to-earth conductance are assumed constant and homogeneous, then rail current $i_i(x)$, rail potential $v_i(x)$ and stray currentsection the inside $Is_i(x)$ $P_iP_{(i+1)}$ can be calculated as follows [4, 6]:

$$i_i(x) = c_{(2i-1)}e^{\gamma x} + c_{2i}e^{-\gamma x} \tag{8}$$

$$v_i(x) = -R_0(c_{(2i-1)}e^{\gamma x} - c_{2i}e^{-\gamma x}) \tag{9}$$

$$Is_i(x) = Gv_i(x) \tag{10}$$

$$1 \le i \le n+m$$

Where:

γ : Propagation constant (m^{-1}) = $\sqrt{GR_R}$

R_0 : Characteristic resistance of the rail conductor earth system (Ω^{-1}) = $\sqrt{R_R/G}$

$c_{(2i-1)}, c_{2i}$: Constants are decided according to specific boundary conditions (as shown in Figure 5) and using Kirchhoff's current and voltage laws.

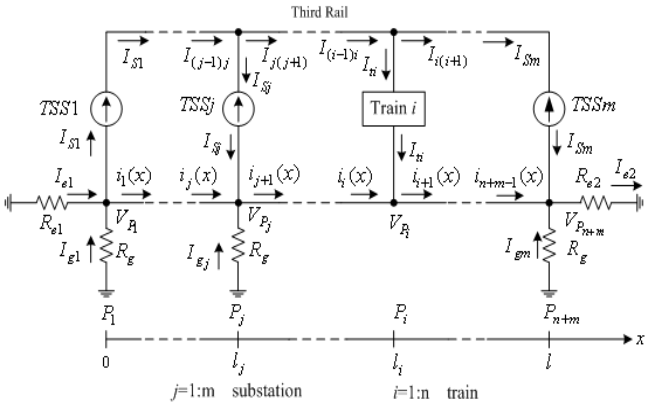


Fig. 5. Simple DC-feed multi-train rail system model at a snapshot

2.3.3 Grounding Strategies Modeling

The analytical solutions for each grounding scheme (including ungrounded, solidly grounded, and diode grounded negative return circuit) can be obtained and are useful in demonstrating the general form of rail potential and stray currents.

The analysis model of the solidly grounded traction schemes is similar to the ungrounded traction scheme, except the negative bus at each substation is directly grounded. Thus, the rail potential and stray current derivation are similar to the ungrounded traction scheme. The rail potential and stray current equations at each section are exactly the same, except for the constants (in the rail current and rail potential equations).

The analytical models of the rail potential and stray current in the diode-grounded scheme are the combined results of the ungrounded and solidly grounded schemes. The diode circuit in the diode-grounded scheme allows current to flow from the grounding mat at substations to the negative bus when a certain threshold voltage is reached. The flowchart for computations of the rail potential and stray current distribution in the diode-grounded scheme is described in reference [20, 21].

III. Tehran DC Metro Line Simulation

The performance of rail line 1 of the Tehran DC railway system is simulated to analyze rail potential and stray current. The effects of grounding strategies on rail potential and stray currents are simulated. For simplicity, we use the data from a section of line 1 (between station Q1 (Mirdamad) and station G1 (Panzdah Khordad)) for our simulation. As shown in Figure 6, this section of rail line 1 is a double-track line with a route length of 8.72km, 11 passenger stations and 5 rectifier transformer substations. The train sets, which are operated along the rail line, pick up the DC power from the third rail for their four units of 132 kW DC traction motors that drive the train set.

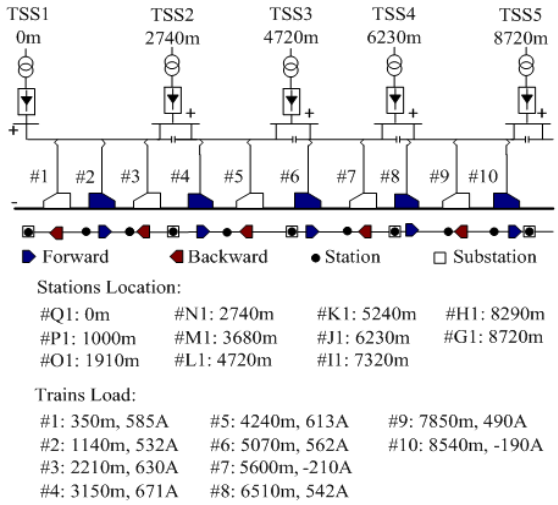


Fig. 6. The location and current consumption for the individual trains are based on an instantaneous "snapshot."

The system parameters for line 1 of the Tehran metro system are as follows: The running rail resistance is $0.035\,\Omega/km$, the rail to ground conductance is $0.02\ S/km$, and the earth resistance for the negative side of a substation grounded via the diode is $0.5\,\Omega$. The maximum slope is 5%, and the minimum radius of curvature is 250 m. The car-technical specifications are in Table I. The total run time of the train from Q1 to G1 (up-truck) is about 14 min, including 25sec of dwell time at each station. The minimum headway is designed to be 2 min [22].

TABLE I
Urban Cars-Technical Specifications of Tehran Metro Network

Number of the car per train (loco. + wagon)	7 (1+6)
Train length (m)	135
Normal static weight/ dynamic weight (tons)	341 / 363
Track gauge (m)	1.435
Front area (m2)	9.6
Number of axles	4
Starting acceleration (m/s2)	0.7
Braking deceleration (m/s2)	1.0
Max. speed (km/h)	80
Feeding voltage (Min. / Max)	750VDC (550/ 900)
Motorized car power (kW)	4*132
Auxiliary power (kW)	45

The Distance-Weight Profile for the up track under normal load is shown in Fig. 7. The Time-Speed profile for a single train in the up and down tracks (for a perfect cycle of journey) is shown

in Fig. 8, and the corresponding Time-Distance and Time-Power profiles are shown respectively in Fig. 9 and 10.

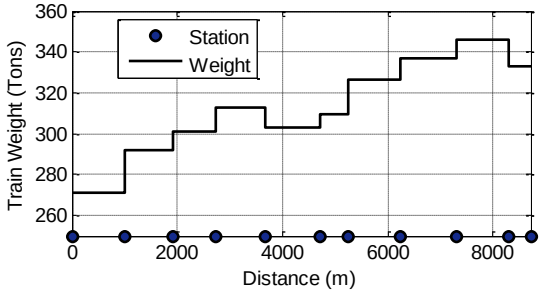


Fig. 7. Distance-Weight Profile for Up track (Normal load)

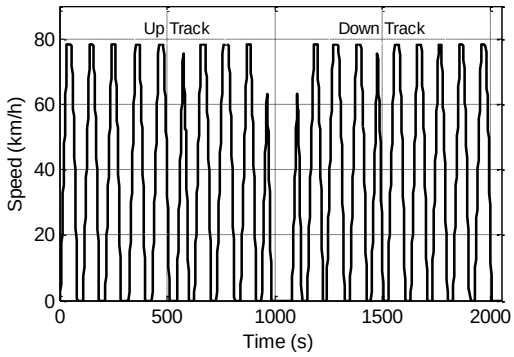


Fig. 8. Time-Speed Profile for a single train in Up and Down tracks

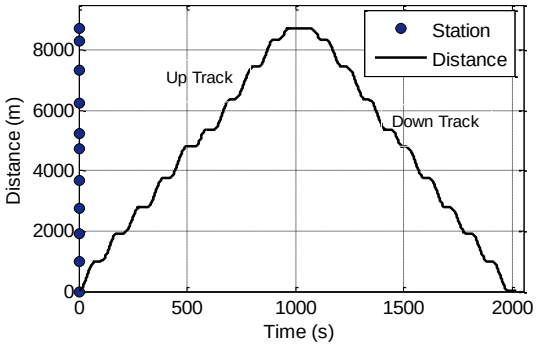


Fig. 9. Time-Distance Profile for Up and Down tracks

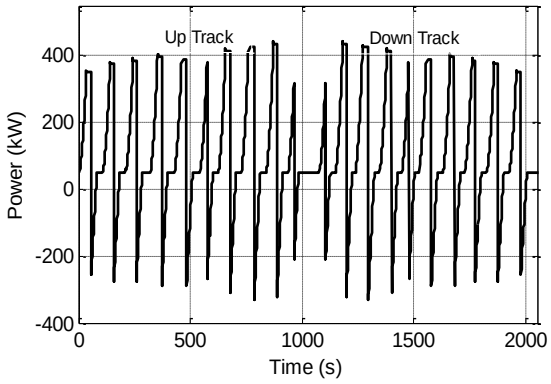


Fig. 10. Time-Power Profile for a single train in the Up and Down tracks

IV. Simulation Results and Analysis

The simulation aimed to investigate the rail potential and stray current characteristics for different earthing strategies assumed for the Line 1 of the Tehran metro system. Fig. 11, 12 and 13 depict one of the typical simulation results. This case simulated the rail potential at a typical snapshot for grounded, ungrounded, and diode-grounded systems. The locations and current consumptions for the individual trains are based on an instantaneous "snapshot" of the system (this would be determined from TPS and ENS simulation).

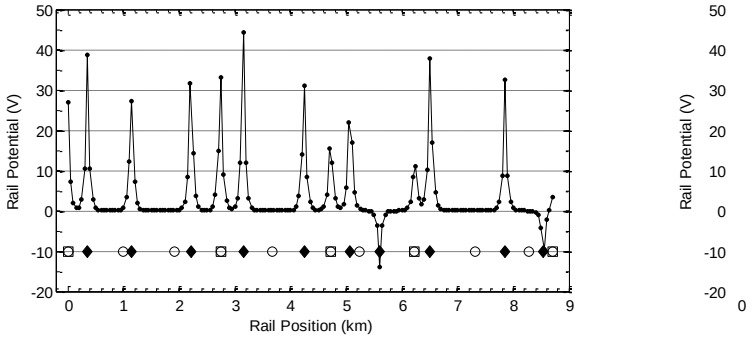


Fig. 11. Rail potential against the rail position at a typical snapshot for a grounded system

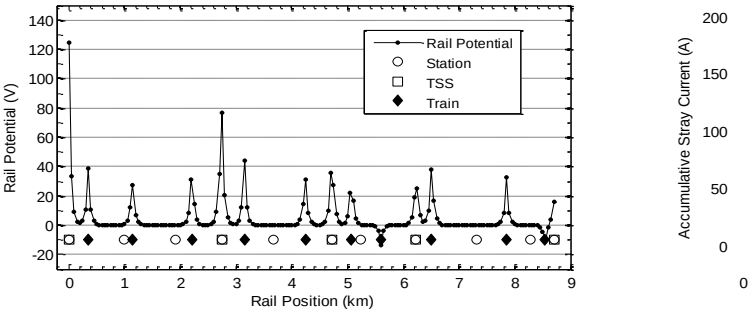


Fig. 12. Rail potential against the rail position at a typical snapshot for an ungrounded system

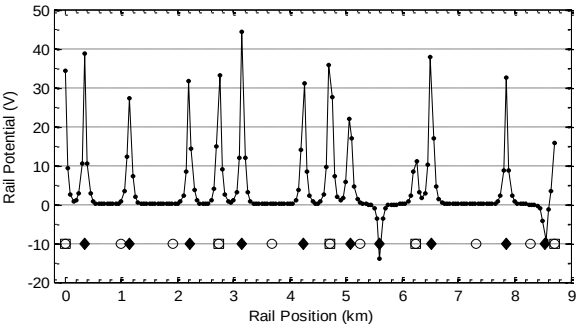


Fig. 13. Rail potential against the rail position at a typical snapshot for a diode-grounded system
(Diodes in TSS #1, 2 and 3 are turned on)

The results show that for the grounded system in some places, the rail potential increases to 45 V; for the diode-grounded system, this rail potential is up to 47 V, but for the non-grounded system in some places, the rail potential increases to 123 V.

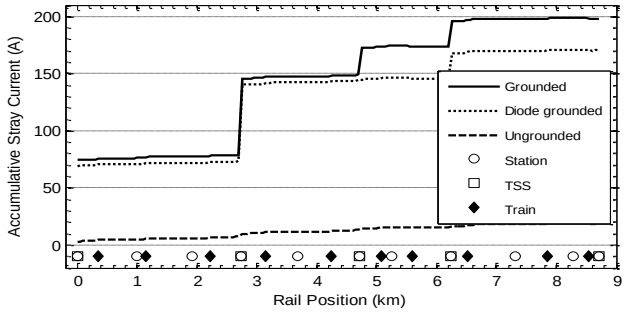


Fig. 14. Cumulative stray current against the rail position at the defined snapshot

Fig. 14 represents the accumulative stray current at a typical snapshot in grounded, diode-grounded, and ungrounded systems.

As seen in Figure 14, the accumulative stray current in solidly grounded systems is at least ten times larger than the accumulative stray current in an ungrounded system. Also, the accumulative stray current obtained in the diode-ground scheme is small compared with the result obtained in the solidly grounded scheme. Since the diode-grounded scheme represents a compromise between a solidly grounded and ungrounded traction scheme, the rail potential and stray current of the diode-grounded system are then between the other two grounding systems. The diode-grounded scheme for the DC railway system mainly depended on the number of substations having their drainage diodes turned on. According to the results of Figures 13 and 14, it can be seen that the non-grounded system, the diode-grounded system, and the grounded system have the best performance in reducing the stray current, respectively. Still, in the case of increasing the rail potential, this arrangement is quite the opposite between them.

The simulation results show that the maximum rail potentials occur in the location of trains and traction substations (TSSs). Tables II and III show that the rail potential and accumulative stray current can vary considerably with variations in assumed rail resistance, rail-to-earth leakage resistance, and substation grounding resistance (R_g).

TABLE II

Systems Grounding Versus Touch Potential and Stray Current ($R_g = 0.5\Omega$, $dx = 50m$)

Rail Data		Maximum Rail Potential (V)		Maximum Accumulative Stray Current (A)	
$R_r(\Omega/km)$	$G(S/km)$	Grounded	Ungrounded	Grounded	Ungrounded
0.035	0.04	31.4	88.2	163.6	20.0
	0.02	44.4	124.7	198.5	19.0
	0.01	62.8	176.5	235.9	18.4
0.02	0.04	23.7	66.7	159.7	19.2
	0.02	33.6	94.3	193.2	18.5
	0.01	47.6	133.9	231.0	18.1

TABLE III

Systems Grounding Versus Touch Potential and Stray Current

($R_r = 0.035\Omega/km$, $G = 0.02S/km$, $dx = 50m$)

R_g (Ω)	Maximum Rail Potential (V)		Maximum Accumulative Stray Current (A)	
	Grounded	Diode-grounded	Grounded	Diode-grounded
0.5	44.4	44.4	198.5	170.6
1.0	46.3	53.7	158.2	130.6
5.0	67.0	98.6	91.9	55.6
10.0	72.2	110.2	79.0	39.0

Generally, the most important form of stray current reduction is maintaining high rail-to-earth resistance along the lines while minimizing line resistance by cross bonding parallel track and cables and maximizing rail cross-sections.

Using running rails with larger cross-sections and higher conductivity, which have lower longitudinal resistance, reduces the maximum Rail Potential and the maximum Accumulative Stray Current.

In addition, according to equation (6), if a higher voltage (such as 1500V DC) is used in a power traction system, the current flowing in the return conductors will be smaller**.** Therefore, stray current and rail potential will be smaller.

V. Conclusion

This paper presents a computer simulation model for analyzing rail potential and stray current in a multi-train DC railway system. Based on the simulation results, the following concluding remarks have been drawn:

- Grounded system results in very high stray currents with reasonably low rail voltages.
- Ungrounded system results in very high rail potentials with low stray currents.
- Diode-grounded system results in high rail voltages and stray currents compared to an ungrounded system.
- The stray current and rail potential depend on headway, the number of trains in operation, and the length of line (distance between traction substations).
- The changes of the diode-grounded system into a partially ungrounded system can effectively reduce the stray current.
- The isolation of two cross-junction lines by disconnecting the impedance bond at the tie line or the cross-junction can also effectively reduce the total amount of stray current and the peak of rail potential.
- If higher voltages are used in the power traction system, then the stray current and rail potential will be reduced**.**

Generally, each grounding scheme provides an essential traction design for rail systems concerning rail potential and stray current. The tasks of controlling rail potential and stray

current are conflicting; therefore, a balance has to be struck between the two.

VI. References

[1] M. Niasati., A. Gholami., 'Overview of stray current control in DC railway systems', IEEE International Conference on Railway Engineering (Hong Kong), Dec. 2008.

[2] Cotton, I., and eds., 'Stray current control in DC Mass Transit Systems'. IEEE Transactions On Vehicular Technology, Vol. 54, No. 2, March 2005, pp.722 – 730.

[3] Lee, C.H., and Wang, H.M., 'Effects of grounding schemes on rail potential and stray currents in Taipei Rail Transit Systems'. IEE Proc. Electr. Power Appl., vol. 148, No 2, 2001, pp. 148-154.

[4] Ku, B.Y., Lee, C.Y., Yen, K.H., Yang, J.J., and Lin, C.F., 'Touch potential and stray current of diode grounded traction systems'. Proc. 19th Symp. Electrical Power Engineering, Taipei, Taiwan, 1998, pp. 677–681.

[5] Hao Xue., Xiaofeng Yang., Yuhao Zhou., Trillion Q. Zheng., 'Multi-interval DC Traction System Simulator for Stray Current and Rail Potential Distribution' IEEE Energy Conversion Congress and Exposition (ECCE), 2018, Portland, USA

[6] Marek Siranec., Michal Regula., Alena Otcenasova., Juraj Altus., 'Measurement and Analysis of Stray Currents' IEEE 20th International Scientific Conference on Electric Power Engineering (EPE), 2019, Czech Republic

[7] Chen Zhiguang., Zhang Xuyan., Qin Chaokui., 'Field testing and analysis of stray current interference on gas pipeline from metro maintenance base' IEEE Conference on Power, Energy, Control and Transmission Systems (ICPECTS), 2018, Chennai, India

[8] George Marulli., Ian Cook., 'Earthing and Negative Return Systems in the Melbourne DC Railway' Down to Earth Conference (DTEC), IEEE, 2018, Melbourne, Australia

[9] Guifu Du., Jun Wang., Xingxing Jiang., Dongliang Zhang., Longyue Yang., Yihua Hu., 'Evaluation of Rail Potential and Stray Current with Dynamic Traction Networks in Multitrain Subway Systems' IEEE Transaction on Transportation Electrication. 2020. Pages: 784 - 796

[10] Sheng Lin., Qi Zhou., Xiaohong., Mingjie., Amin Wang., 'Infinitesimal Method Based Calculation of Metro Stray Current in Multiple Power Supply Sections' IEEE Access. 2020, Pages: 96581 - 96591

[11] Ebrahim Zare Juybari., Reza Keypour., Mohsen Niasati., 'Voltage Distribution Indices Method to Analyze the Performance of Various Structures of Stray Current Collectors in DC Transit Lines' IET Electrical Systems in Transportation. May 2021.

[12] Sheng Lin., Yang Huangfu., Qi Zhou and Aimin Wang., 'Evaluation and Analysis Model of Stray Current in the Metro Depot' IEEE Transaction on Transportation Electrication. 2020, Pages: 1780 – 1794.

[13] M. Niasati, A. Gholami, 'Evaluation of rail potential control devices performance for control of rail potential of DC electrified railway systems', IEEE International Conference on Railway Engineering (Hong Kong), Dec. 2008.

[14] Salman Aatif, Haitao Hu, Fezan Rafiq, Zhengyou He, 'of rail potential and stray current in MVDC railway electrification system', Springer (Railway Engineering Science journal), July 2021, volume 29, pages 394–407.

[15] Ade Ogunsola, Leonardo Sandrolini, Andrea Mariscotti, 'Evaluation of Stray Current From a DC-Electrified Railway With Integrated Electric–Electromechanical Modeling and Traffic Simulation', IEEE Transactions on Industry, Nov. 2015. Volume: 51, Issue: 6, 5431 – 5441.

[16] Aydin Zaboli; Behrooz Vahidi; Sasan Yousefi; Mohammad Mahdi Hosseini-Biyouki, 'Evaluation and Control of Stray Current in DC-Electrified Railway Systems', IEEE Transactions on Vehicular Technology, Vol: 66, Issue: 2, Feb. 2017. 974 – 980.

[17] Chengtao Wang, Wei Li, Yuqiao Wang, Shaoyi Xu, Mengbao Fan, 'Stray Current Distributing Model in the Subway System: A review and outlook', International Journal of Electrochemical Science, Dec. 2017. 13(2): 1700-1727.

[18] Tzeng, Y. S., 'DC RAIL—A power system simulator for DC electrified railways. Proc. World Metro Symp., Taipei, Taiwan, 2002, pp. 332– 337.

[19] Chang, C.S., Khambadkone, A., and Xu, Z., 'Modeling and simulation of DC traction system with VSI-fed induction motor drive train using PSB/MATLAB'. IEEE PEDS 2001, Indonesia, 2001, pp. 881–885.

[20] Lee, C.H., 'Evaluation of the maximum potential rise in Taipei rail transit systems. Power Delivery, IEEE Transactions on Volume 20, Issue 2, April 2005, pp. 1379 – 1384.

[21] 'The Effects of Moisture and Ballast Contact on Electrical Isolation of Running Rail on Wood Ties and Ballasted Track'. Deleuw Cather and Company, July 1982.

[22] Yu, J.G., 'The effects of earthing strategies on rail potential and stray currents in DC transit railways'. Proc. Int. Conf. (Conf. Publ. No. 453) Developments in Mass Transit Systems, 1998, pp. 303–309.