



دانشگاه سمنان



Semnan University

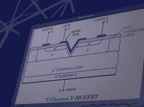
Journal of Modeling & Simulation in Electrical & Electronics Engineering (MSEEE)

8

Volume 2, Number 2, August 2022

- | | | | |
|--|-------|---|-------|
| ❖ Design Processes Linear Permanent Magnet Electrical Vernier Machines For Future
Research Directions: A Review | 29-36 | ❖ Copy-Move Forgery Detection Using Fast Retina Keypoint (FREAK) Descriptor | 1-7 |
| Hamid Yaghobi; Mohammad Reza Mohammadzadeh Moghaddam | | Ardeshir Ghasemi Yegane; kourosh Kiani; Razieh Rastgoo | |
| ❖ Coordinated Management of EVs Charging Station with a Wide Presence of Renewable
Energy Sources | 37-47 | ❖ Wideband Balun-LNA Employing gm-Boosting Feedback Technique and CBLD Circuit
for Digital Televisions Tuner Application | 9-15 |
| Morteza Shamani; Peyman Talebi; Asghar Akbari Foroud | | Nima Rahimzadeh; Pejman Rezaei | |
| ❖ Cancer Model Simulation in Simulink Environment for Educational Purposes of Cancer
Drug Dose Control | 49-60 | ❖ Improvement of Mesh Simplification Using Normal Vector Diversity | 17-22 |
| Ali Maleki; Matin Kordbacheh | | Masoud Ebadi; kourosh Kiani; Razieh Rastgoo | |
| | | ❖ Implementation of a Fourth-Order Compact Quasi-Elliptic Substrate Integrated
Waveguide Filter in C-Band | 23-27 |
| | | Narges kiani; Majid Afsahi; Farzad Tavakkol Hamedani; Pejman Rezaei | |

Online ISSN 2821-0786



Copy-Move Forgery Detection Using Fast Retina Keypoint (FREAK) Descriptor

Ardeshir Ghasemi Yegane¹, Kouros Kaini^{1*}, and Razieh Rastgoo¹

Abstract— Image forgery, the manipulation of an image to hide some meaningful or helpful information, is widely used to manage the large amount of information being exchanged in the form of images. There are different forms of image forgery, and copy-move forgery is the most common form of it. The copy-move forgery is easy to perform yet challenging to detect due to the similarity between the original part of the image and the copied part. In this paper, we employ a keypoint descriptor inspired by the human visual system, namely the FREAK (Fast Retina Keypoint) descriptor, for robust copy-move forgery detection. This method uses the advantages of FREAK descriptor such as fast computing and low memory load compared to SIFT, SURF, and BRISK. Finally, geometric transformation parameters are extracted and discussed. Results confirm promising results in the case of image post-processing operations such as adding noise, illumination change, and geometric transformations such as rotation and scaling.

Index Terms— Copy-move forgery detection, Fast Retina Keypoint (FREAK), Keypoint descriptor.

I. INTRODUCTION

Nowadays, a large amount of information is being exchanged in the form of images/videos. Different methods in Computer Vision are used to process this information [1-9]. However, the manipulation of digital images has been straightforward because of the existence of powerful computers, advanced editing software packages, and high-resolution imaging devices [10]. Developing software, tools, and digital image processing techniques facilitated the use of image forgery techniques that are not simply detectable. To deal with this problem, digital forensics has been introduced, which

are techniques to detect any type of forgery in digital images. Applications of forensics are in the fields of a court of law, criminal investigations, insurance, scientific claims, medical imaging, and so on [1,11-12]. Fig. 1 shows a typical image and its forged version.

Tampering methods are divided into three main classes: copy-move, splicing, and retouching. In copy-move tampering, one or more patches of an image are copied and moved to another location in the same image to duplicate some object or hide some other scenes of that image. In splicing forgeries, a part of one image is copied and moved to another image. In the case of retouching, some techniques are used to make some changes to images. Copy-move is one type of forgeries for which many detection methods have been proposed to solve. All forgery detection methods may be divided into two main categories: block-based and keypoint-based methods. In the block-based methods, the query image is divided into some overlapping blocks that differ from each other only by one row or one column. For each block, some features are extracted and then sorted. Neighbor feature vectors in the sorted matrix rely on features extracted from similar blocks. In keypoint-based methods, some important points of the image are detected, and features of these keypoints are extracted. After matching extracted features, duplicated patches can be found. Some forgery detection methods rely on keypoints, such as SIFT (Scale Invariant Feature Transform), SURF (Speeded Up Robust Features), BRISK (Binary Robust Invariant Scalable Keypoints), and the proposed keypoint descriptor, Fast Retina Keypoint (FREAK), introduced in [13].

The rest of this paper is organized as follows: In section II, the related works on the copy-move forgery detection field are introduced. In section III, the proposed method is introduced

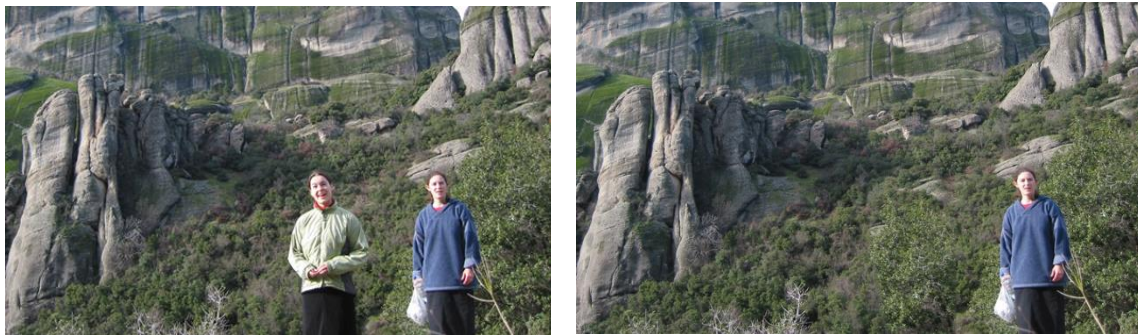


Fig. 1. (Left) A typical image, (Right) The tampered version of it.

step by step. Section IV demonstrates the experimental results of the proposed method. Finally, the conclusions are drawn in section V.

II. RELATED WORKS

Many methods have been presented in the literature to overcome copy-move attacks. Some good studies can be found in [14-18]. The main idea of all methods proposed to solve the copy-move problem is that both copied and pasted patches have similarities in some properties that can be used as features. Using these features and searching for similar features may lead to discovering similar patches. A good forgery detection method should be able to detect duplicated patches. In addition, a good method must be stable against geometric transformations such as translation, rotation, and scaling as well as post-processing manipulations such as noise addition, JPEG compression, blurring, and so on. Almost the first attempt to solve the copy-move forgeries was proposed in [19]. It is a block-based approach that uses DCT as a feature extraction method and lexicographic as a sorting algorithm. Using a lexicographic sorting algorithm instead of the nearest neighbor algorithm was done to improve match finding and decrease running time. Several statistical methods for detecting copy-move attacks were proposed in [20, 21]. In [22], a method for reducing dimensions by applying Principle Component Analysis (PCA) to small, fixed-size blocks was proposed. This model is robust against additive noise and JPEG compression. In [23], four components—R, G, B, and Y—were used in blocks to obtain the energy distribution of luminance along with four different directions. Another method was proposed in [24] by using the singular value decomposition (SVD) for feature vector dimensionality reduction along with a discrete wavelet transform (DWT) for duplication detection and lexicographic sorting. This model is suitable for image compression and edge processing. A method for detecting duplicated patches was proposed in [25] that uses blur moment invariants, PCA, and a k-d tree. While this model uses 24 blur invariants up to the seventh order, this method suffers from high computational time. The method described in [26] uses Zernike moments that are rotation-invariant. Another method introduced in [27] uses a log-polar map for extracting descriptors into a 1-D vector and is invariant to reflection, rotation, and scaling.

Other methods extract descriptors using a relatively recent feature known as the Local Binary Pattern (LBP). LBP has some texture-related applications in image processing, such as texture categorization and copy-move forgery detection.

Local visual features such as SIFT [28], SURF [29], BRISK [30], BRIEF [31], and so on, are widely used in different image processing areas such as object recognition, object matching, and image retrieval because of their robustness to some geometric transformations such as translation, rotation, and scaling. In [32], a novel method using SIFT was introduced that is stable against changes in illumination, rotation, and scaling to detect duplicated patches. A method based on SIFT was

proposed in [33] that estimates the geometric transformation parameters. This model shows a good true positive rate (TPR) in detecting duplicated regions. In this paper, we use a fast and robust descriptor, namely the FREAK descriptor. FREAK uses a comparison between pixel intensities and makes a binary string. Using this binary descriptor, copy-move forgery detection is done, and duplicated patches are discovered.

III. PROPOSED METHOD

In this paper, we propose a solution in the field of copy-move forgery detection based on extracting key point features using the FREAK descriptor. FREAK is a strong keypoint descriptor inspired by the human visual system, namely the retina. Fig. 2 depicts the proposed method's procedure. This approach works with grayscale images, so if the query image is not in grayscale format, it should be converted to grayscale. The proposed approach is separated into four major steps: (1) recognizing keypoints, (2) extracting keypoint descriptors, (3) matching extracted features to obtain forged areas, and (4) determining transform parameters using the RANSAC method.

A. Detecting keypoints

FREAK extracts keypoint descriptors from some pre-specified keypoints but it does not detect keypoints. Thus, we need some methods to detect robust keypoints to pass to the FREAK algorithm. Someone may select some random points but the result might not be satisfying. In the case of FREAK, we can use some keypoints detection methods such as corner points as keypoints. Other options may be using SURF or BRISK keypoints. There are some kinds of corner point detector algorithms in the literature [34, 35]. One of the most useful methods is the Harris algorithm that was introduced in [34].

The combination of the Harris corner detector as a keypoint detector and FREAK as a keypoint descriptor leads to good results. The other useful corner detector is FAST, which was introduced in [36]. In this paper, we use a combination of Harris corners and SURF points as keypoints in the proposed model. Some other keypoints, such as FAST, SURF, and BRISK keypoints, are used to make some comparisons.

B. Extracting keypoint descriptors

FREAK is a method to extract keypoint descriptors in the same way as the retina does in the human visual system. FREAK uses a pattern similar to that of retinal ganglion cells in the retina. This pattern is circular, and the density of the central points is higher than the peripheral ones. This grid pattern is illustrated in Fig. 3. Only 512 pairs of this grid are used and the rest are ignored. FREAK is a binary descriptor of size 512 bits, which means it makes a comparison between the intensities of each point of every pair in the pattern. Each bit is calculated by thresholding the difference between one pair's intensities that have been smoothed using a Gaussian kernel.

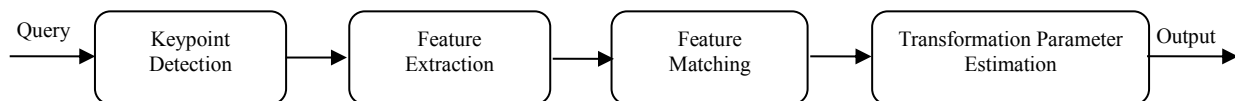


Fig. 2. A schematization of the proposed method.

Equation (1) demonstrates this concept:

$$F = \sum_{0 \leq a < N} 2^a T(P_a) \quad (1)$$

Where P_a is a pair of points, N is the size of descriptor and T is thresholding function defined as:

$$T(P_a) = \begin{cases} 1 & \text{if } (I(P_a^{r_1}) - I(P_a^{r_2})) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $I(P_a^{r_1})$ and $I(P_a^{r_2})$ are smoothed intensities of two points in the pair P_a .

For each keypoint detected in the previous step, we extract a FREAK descriptor as described above. Now we have some binary descriptors, each of which is related to one keypoint in the image.

C. Matching features

In the copy-move forgery attacks, the copied patch has essentially similar characteristics to the original patch. Therefore, two similar patches in the same image have similar features extracted by the same method. So, matching features in all FREAK descriptors extracted from an image can reveal similar patches in the image.

With a query image, a set of n keypoints $P = \{p_1, \dots, p_n\}$ are extracted. For each keypoint, the corresponding descriptor is also extracted. All of these descriptors form a set of n descriptors $D = \{d_1, \dots, d_n\}$. Given an individual feature d_i , the goal is to find the nearest neighbor feature in the feature space of n-1 remaining features. As we use the FREAK binary descriptor, the used distance function should be Hamming distance. The distance must be lower than a global threshold; T1 (matching threshold), to accept two features as similar. This method does not perform well because of some ambiguous matches that may be made. So, we use the procedure introduced in [27]. Assume that d_1 is the nearest neighbor and d_2 is the second nearest neighbor. Then, two features are tagged as similar, if the ratio between d_1 and d_2 is lower than threshold T2 (matching ratio). It means that the candidate features are matched only if the following constraint is fulfilled:

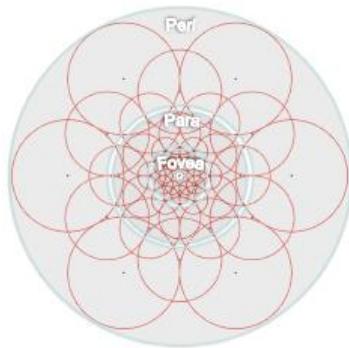


Fig. 3. Illustration of FREAK sampling pattern introduced in [1]

$$d_1/d_2 < T_2 \text{ where } T_2 \in (0,1) \quad (3)$$

Using this procedure, ambiguous matches will be rejected. The two thresholds mentioned above play key roles in the number of discovered matches. In both cases, increasing the

threshold leads to finding more matches. Both thresholds are in the interval of (0,1). In the case of both T1 and T2, increasing the threshold leads to an increment in discovered matched pairs but may also increase the false matches. Therefore, a tradeoff is needed.

Matching features using this method is called the "nearest neighbor ratio" method. This method is an instance of a lazy algorithm. Thus, it is a time-consuming method because of its many distance computing operations. If there are n feature vectors, the algorithm has an order of $O(n^2)$. Increasing the number of features leads to increase in the needed time for computing all distances in the order of two. So using fewer features makes the computing timeless. After matching all features, there may be some unwanted matched pairs. These pairs are outlier pairs and should be rejected. For this reason, an instance of the RANSAC algorithm is needed to discover the transformation parameters and reject the outliers. This step is described in the next section.

D. Determining geometric transform parameters

RANSAC (random sample consensus) is an iterative algorithm for estimating model parameters when the data is contaminated by unwanted data. RANSAC was first introduced by Fischler and Bolles in [30] in 1981 as a method to estimate the parameters of a model having a set of data contaminated by some unwanted data named as outliers. The RANSAC algorithm has some modifications. Despite these modifications, it is essentially made of two main steps that are repeated iteratively. These steps are known as "hypothesize and test framework."

In the hypothesize step, a minimum sample set is randomly selected from the input data, and the model parameters are computed using these samples.

In the test framework step, the algorithm verifies which elements of the entire dataset fit in the model obtained in the previous step and inserts these elements into the model. The next iteration will be done with a new model.

These two steps run iteratively. RANSAC terminates when it cannot find a better model than the one found. In the copy-move forgery field, a patch of an image is copied and moved to another place in the same image. This transformation can be modeled as a geometric transformation. A geometric transformation is made of three main simple transformations: translation, rotation, and scaling. Each of these main transformations has its own transform matrix and parameters. For a given point (x,y) in a 2D plane, the translation, rotation, and scaling matrices are as below, respectively:

$$T = \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (4)$$

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (5)$$

$$S = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \quad (6)$$

The combination of these base transformation matrices can be written as a single matrix as follows:

$$H = \begin{bmatrix} A & T \\ 0^T & 1 \end{bmatrix} \quad (7)$$

The vector T is a translation vector as (4). Matrix A is the combination of rotation and scaling matrices in (5) and (6). For

any point (x,y) , the transformation point is computed as:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (8)$$

In the case of copy-move attacks, the copied patch is usually transformed using a transformation according to (8). Using the RANSAC algorithm, only inliers are selected to form a model, and outlier points will be rejected. Also, transformation parameters can be obtained. Thus, only keypoints that have similar features and the same transformation parameters will remain.

In this paper, we use the RANSAC algorithm two times. For the first time, after matching features, RANSAC is applied to matched features to find some inlier points and their corresponding transformation parameters. There may be some matched pairs that can fit in that transformation, but their transformation directions are opposite and therefore are rejected in the RANSAC algorithm. We use the KNN(K-Nearest Neighbors) algorithm with $k = 1$ overall matched pairs to add these pairs to the discovered transformation. So, all matched points are classified into two classes based on their Euclidean distance from two center points of previously found inliers. Consequently, another instance of the RANSAC algorithm is applied to these two classes, and final inlier points and transformation parameters are achieved. This process makes the results better. An image will be tagged as “forged” if the number of final inlier points is at least four.

Fig. 4 demonstrates the step-by-step output of the proposed method over the query image mentioned in fig. 1 with $T1=0.6$ and $T2=0.8$. First, all keypoints and their corresponding descriptors are extracted (Fig 4-a). After matching extracted descriptors, only those keypoints that have a match in the rest of keypoints are remain (Fig 4-b). Matched keypoints can be shown as some connected pairs (Fig 4-c). In this step, there may be some matched outlier keypoints. After applying the RANSAC algorithm, the keypoint pairs that fit in a geometric transformation are extracted (Fig 4-d). Geometric transformation parameters are obtained in this step.

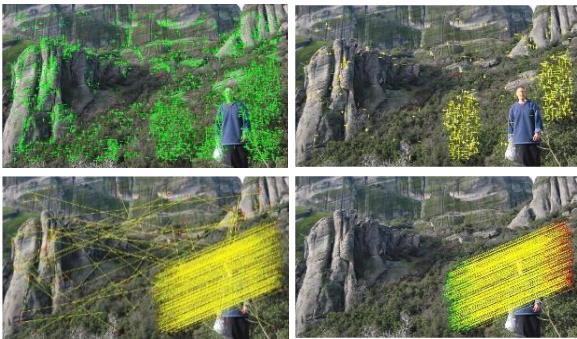


Fig. 4. Step-by-step output of the proposed method for a query image including four steps: First row – left: Detected keypoints, First row – right: Matched keypoints, Second row – left: Matched keypoint pairs, Second row – right: Matched keypoint pairs after applying RANSAC algorithm.

IV. EXPERIMENTAL RESULTS

In this section, we evaluate the proposed method over some test images. The used test dataset is MICC-F220, which was introduced in [27]. To make a comparison between different

keypoint detector algorithms, we use some factors, such as the average number of detected keypoints and the average number of matched keypoints. This comparison can be found in Table I. To evaluate the proposed method against other methods, we define the True Positive Rate (TPR) and False Positive Rate (FPR) metrics. These metrics are defined as follows:

$$TPR = \frac{TP}{N_f} \quad (9)$$

$$FPR = \frac{FP}{N_o} \quad (10)$$

Where TP (True Positive) is the number of images tagged as forged being forged, N_f is the number of all forged images in the dataset, FP (False Positive) is the number of images tagged as forged being original and N_o is the number of all original images in the dataset.

Table II shows the results of various amounts for the Matching Threshold and Matching Ratio parameters when the detection method is a combination of Harris corners and SURF points. In this table, we compare TPR, FPR, and the average time needed for any single image as metrics. It is obvious that the best result is obtained when the matching threshold and



Fig.5. Some images with various post processing operations and their corresponding outputs.

Matching Ratio are 0.6 and 0.8, respectively. Table III shows the values used for parameters in the proposed model.

In Table IV, we compare four kinds of keypoint detection methods when the Matching Threshold and Matching Ratio are 0.6 and 0.8, respectively. We can see that the combination of Harris corners and SURF points has the best results.

Fig. 5 shows four typical images that have been tampered with by various post-processing operations such as rotation, scaling, illumination change, and a combination of them and their corresponding outputs. We can see that the duplicated area is easily found in all types of tampering.

Multiple tampering attacks can also be easily handled by the proposed method. Using an iterative operation, after finding one duplicated region, one set of found points should be removed from all found matches, and the same process must be done on the remaining points until no more patches can be found.

V. CONCLUSIONS

In this paper, we discussed copy-move forgery detection methods. A new method to discover image forgery detection based on the FREAK descriptor was proposed. This method combines Harris corners and SURF points and uses them as keypoints. Then FREAK descriptor is extracted for that keypoints and the matching phase is done. Using the proposed algorithm, duplicated patches are appeared, and transformation parameters are obtained. According to experimental results, the highest performance is obtained in the case of using Harris and SURF points as keypoints and matching threshold and matching ratio with values of 0.6 and 0.8, respectively. The main advantage of the proposed method is low running time and computational load. Therefore, this method can be used in cases where hardware specifications are low and the low run time is considered such as in smartphones. In addition, this method is invariant to some post-processing operations such as scaling, rotation, adding noise, illumination change, etc. In future works, we would like to employ deep learning-based models in the field.

TABLE I
Comparison Between Different Detection Methods on Detected Keypoints and Matched Points

Detection Method	Average Number of detected keypoints	Average Number of matched points
Harris	1218.47	6.95
FAST	1317.26	6.91
BRISK	598.40	2.76
Harris + SURF	2098.15	10.07

TABLE II
Comparison Between Different Matching Thresholds and Matching Ratio Amounts

(Matching Threshold, Matching Ratio)	True Positive Rate (%)	False Positive Rate (%)	Average Time(s)
(0.3, 0.6)	60	2.72	0.58
(0.3, 0.7)	78.18	7.27	0.79
(0.4, 0.7)	77.27	7.27	0.78
(0.5, 0.7)	77.27	9.09	0.79
(0.3, 0.8)	89.09	13.64	1.45
(0.4, 0.8)	89.09	14.54	1.49
(0.5, 0.8)	84.54	10.90	1.47
(0.6, 0.8)	91.82	8.18	1.45
(0.7, 0.8)	88.18	13.63	1.45
(0.8, 0.8)	89.09	11.81	1.36
(0.3, 0.9)	79.09	24.55	2.03
(0.4, 0.9)	82.72	27.27	1.99
(0.5, 0.9)	82.72	30.90	2.03
(0.6, 0.9)	76.36	31.81	2.03
(0.7, 0.9)	80.90	31.81	2.05

TABLE III
The Values Used for Model Parameters

Parameter	Value	Parameter	Value
Matching Threshold	0.3, 0.4, 0.5, 0.6	Matching Ratio	0.6, 0.7, 0.8, 0.9
T1	0.6	T2	0.8

TABLE IV
Comparison Between Different Detection Methods on TPR, FPR, and Average Time

Matching Threshold = 0.6, Matching Ratio = 0.8			
Detection Method	True Positive Rate	False Positive Rate	Average Time
Harris	75.45	4.5	0.93
FAST	64.54	8.18	0.93
BRISK	71.81	1.81	0.80
Harris + SURF	91.82	8.18	1.45

VI. REFERENCES

- [1] Diwan, A., Sharma, R., Roy, A.K., Mitra, S.K. (2021). Keypoint-based comprehensive copy-move forgery detection. *IET Image Process* 15, pp. 1298–1309.
- [2] Gan, Y., Zhong, J., Vong, Ch. (2022). A Novel Copy-Move Forgery Detection Algorithm via Feature Label Matching and Hierarchical Segmentation Filtering. *Information Processing & Management* 59, No. 102783.
- [3] Li, F.F. (2022). Machine Learning in Computer Vision, <https://www.cs.princeton.edu>.
- [4] Ouyang, J., Liu, Y., Liao, M. (2017). Copy-move forgery detection based on deep learning. 2017 10th International Congress on Image and Signal Processing, Bio-Medical Engineering and Informatics (CISP-BMEI), Shanghai, China.
- [5] Rodriguez-Ortega, Y., Ballesteros, D.M., Renza, D. (2021). Copy-Move Forgery Detection (CMFD) Using Deep Learning for Image and Video Forensics, *J Imaging*. Vol. 7, no. 3.
- [6] Liu, Y., Xia, Ch., Zhu, X., Xu, Sh. (2020). Two-Stage Copy-Move Forgery Detection with Self Deep Matching and Proposal Superglue, arXiv: 2012.08697.
- [7] Zainal Abidin, A.B., Abdul Majid, H.B., Samah, A.B.A., Hashim, H.B. (2019). Copy-Move Image Forgery Detection Using Deep Learning Methods: A Review, 2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS), Johor Bahru, Malaysia.
- [8] Kiani, K., Rezaeeraad, S., Rastgoo, R. (2021). HMM-based Face Recognition Using SVD and Half of the Face Image, *Journal of Modeling & Simulation in Electrical & Electronics Engineering (MSEEE)*, vol. 1, no. 2.
- [9] Majidi, N., Kiani, K., Rastgoo, R. (2020). A deep model for super-resolution enhancement from a single image, *Journal of AI and Data Mining*, vol. 8, no. 4, pp. 451-460.
- [10] Youssef, B., Atta, E. (2016). Image Forgery Detection using FREAK Binary Descriptor and Level Set Segmentation. *International Journal of Scientific & Engineering Research* 7.
- [11] Sridevi, M., Aishwarya, S., Bokadia, D. (2019). Parallel Image Forgery Detection Using FREAK Descriptor. *Information and Communication Technology for Intelligent Systems* 107, pp. 619-630.
- [12] Khudhair, Z.N., Mohamed, F., Kadhim, K.A. (2021). A Review on Copy-Move Image Forgery Detection Techniques. *Journal of Physics: Conference Series*.
- [13] Alahi, A., Ortiz, R. & Vandergheynst, P. (2012). FREAK: Fast Retina Keypoint. *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 510–517.
- [14] Warif, N et al. (2016). Copy-move forgery detection: Survey, challenges, and future directions. *Journal of Network and Computer Applications*, vol. 75, pp. 259–278.
- [15] Birajdar, G.K. & Mankar, V.H. (2013). Digital image forgery detection using passive techniques: A survey. *Digital Investigation*, Vol. 10, no. 3, pp. 226-245.
- [16] Sridevi M., Mala C. & Sanyam S. (2012). Comparative Study of Image Forgery and Copy-Move Techniques. In: Wyld D., Zizka J., Nagamalai D. (eds). *Advances in Computer Science, Engineering & Applications. Advances in Intelligent and Soft Computing*, vol. 166. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-30157-5_71.
- [17] Meena, K.B. & Tyagi, V. (2019) Image Forgery Detection: Survey and Future Directions. In: Shukla R., Agrawal J., Sharma S., Singh Tomer G. (eds) *Data, Engineering, and Applications*. Springer, Singapore. https://doi.org/10.1007/978-981-13-6351-1_14.
- [18] Al_Azrak, F.M., Sedik, A., Dessowky, M.I. et al. (2020). An efficient method for image forgery detection based on trigonometric transforms and deep learning. *Multimed Tools Appl*, vol. 79, pp. 18221–18243. <https://doi.org/10.1007/s11042-019-08162-3>.
- [19] Fridrich, J., Soukal, D. & Lukas, J. (2003). Detection of copy-move forgery in the digital images. *Proceeding of DFRWS*, Cleveland, OH.
- [20] Farid, H. (2003). A picture tells a thousand lies. *New Scientist*, 6 Sep 2003.
- [21] Popescu, A. & Farid, H. (2005). Exposing Digital Forgeries in Color Filter Array Interpolated Images. *IEEE Trans. on signal processing*, vol. 53, no. 10, pp. 3948-3959.
- [22] Popescu, A. & Farid, H. (2004). Exposing digital forgeries by detecting duplicated image regions. *Dartmouth College, Computer Science, Tech. Rep. TR2004-515*.
- [23] Luo, W., Huang, J. & Qiu, G. (2006). Robust detection of region-duplication forgery in digital image. *Proc. of ICPR*, Washington, USA.
- [24] Li, G., Wu, Q., Tu, D. & Sun, S.J. (2007). A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD. *Proc. of IEEE ICME*, Beijing, China.
- [25] Mahdian, B & Saic, S. (2007). Detection of copy-move forgery using a method based on blur moment invariants. *Forensic Science International*, vol. 171, no. 2-3, pp. 180–189.
- [26] Ryu, S.J., Lee, M.-J., Lee, H.-K. (2010). Detection of copy-rotate-move forgery using Zernike moments. *Proc. of International Workshop on Information Hiding*, Calgary, Canada.
- [27] Bravo, S. & Nandi, A.K. (2011). Automated detection and localization of duplicated regions affected by reflection, rotation, and scaling in image forensics. *Signal Processing*, vol. 91, pp. 1759–70.
- [28] D. Lowe. (1999). Object recognition from local scale-invariant features., 1999. *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, pp. 1150–1157.
- [29] Bay, H., Tuytelaars, T. & Van Gool. T. (2006). SURF: Speeded up robust features. *ECCV*, Graz, Austria, pp. 404–417.
- [30] Leutenegger, S., Chli, M. & Siegwart, R. (2011). Brisk: Binary robust invariant scalable keypoints. *2011 International Conference on Computer Vision*,

Barcelona, Spain.

- [31] Calonder, M., Lepetit, V., Strecha, C. & Fua, P. (2010). Brief: Binary robust independent elementary features. ECCV, Heraklion, Crete, Greece, pp. 778–792.
- [32] Huang, H., Guo, W. & Zhang, Y. (2008). Detection of copy-move forgery in digital images using SIFT algorithm. Proc. of IEEE Pacific-Asia Workshop on Computational Intell. and Industrial Application, Wuhan, China.
- [33] Irene, L., Caldelli, R., Del, A. & Serra, G. (2011). A SIFT-based forensic method for copy-move attack detection and transformation recovery. IEEE Trans Inf Forensics Security, vol. 6, pp. 1099–110.
- [34] Harris, C. & Stephens, M. (1988). A combined corner and edge detector. In Alvey vision conference, Manchester, UK, vol. 15, pp. 147-151.
- [35] Rosten, E. & Drummond, T. (2006). Machine learning for high-speed corner detection. ECCV, Graz, Austria, vol. 1, pp. 430–443.
- [36] Fischler, M. & Bolles, R. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, vol. 24, no. 6, pp. 381–395.

Wideband Balun-LNA Employing g_m -Boosting Feedback Technique and CBLD Circuit for Digital Televisions Tuner Application

Nima Rahimzadeh¹ and Pejman Rezaei²

Abstract-- A balun Low Noise Amplifier (Balun-LNA) with technique of g_m -boosting feedback and a modified current bleeding (CBLD) circuit is proposed for application in the tuner of digital television (DTV) and other wideband radio and microwave receivers. Using the technique of g_m -boosting feedback causes input impedance matching to not just depend on the CG transistor, and input impedance matching is satisfied by the transconductance of CG and CS transistors. Therefore, the transconductance of the CG transistor increases to boost the differential voltage gain of Balun-LNA and decrease its NF. Also, a modified current bleeding circuit is used in the CS stage in order to make the CS transistor have higher transconductance and its output current be identical to the output current of the CG stage. To compensate for having identical output currents, symmetrical loads are used in differential output so that they cause the gain and phase balance at the differential output. This Balun-LNA is built on 90-nm CMOS technology and operates in the digital television frequency band of 48 to 864 MHz. This Balun-LNA achieves a maximum differential voltage gain of 24 dB, an input return loss of less than -10 dB, and a minimum NF of 5 dB. This Balun-LNA works at 2.8 v nominal supply voltage and consumes the power of 2.5 mW.

Index Terms-- Balanced output, Balun-LNA, Low power consumption, Noise Cancelling, Symmetrical loads, Digital television tuner.

I. INTRODUCTION

In recent years, wideband radio and microwave receivers, especially digital television receivers, have been more popular because wireless communication devices must support several features in one IC [1]–[9]. Nowadays, digital televisions are used for different purposes and in many places. Digital television manufacturers compete to produce digital televisions with better quality in displaying video and pictures on digital television screens and better quality in playing audio through digital television speakers. In order to develop the performance of displaying pictures and videos on screen and audio in the speakers of digital television, the low-noise amplifier must be modified. Reducing noise, increasing voltage gain, improving linearity due to higher linear performance of tuner IC next circuits, improving input impedance matching, and lowering

LNA power consumption, which are the most important features of designing a LNA for a tuner of digital televisions, are important challenges of modifying LNA for application tuner of digital televisions.

In digital television, first the electromagnetic waves of digital video broadcasting (DVB) are received with the antenna of digital television, and then they are transformed into electrical signals and are transferred to the tuner of digital television. A tuner is an implement with a metal shield which is located at the beginning of the mainboard of digital televisions and has an IC by the name of tuner IC. The Low Noise Amplifier is the first and most important block of the tuner of digital television that has an important effect on the performance quality of digital television in playing pictures, videos, and audio [10]–[15].

There are several challenges in designing a low-noise amplifier for the tuner IC of digital televisions. The digital television frequency band ranges from 48 to 860 MHz, which is very wide. The tuner IC of digital television should handle the signals with broadband and should have good input impedance matching, sufficient gain, high linearity, and low NF over the desired frequency bandwidth. In order to endure strong analogue and digital interference in the terrestrial environment and avoid multiple distortions by hundreds of broadcast channels in the cable environment, the LNA should have a high third-order input referred intercept point (IIP3) and a high second-order input referred intercept point (IIP2) [16] [20].

For the application of a digital television tuner, a single-ended input LNA is usually preferred [21]. Furthermore, differential LNAs are used to achieve higher SNR because differential signal processing has a high common mode rejection ratio, power supply rejection ratio, and low second order distortion [22]–[23]. For wideband applications like the tuner of digital television, the passive transformer like the Balun circuit is too bulky to be integrated on a chip, and its insertion loss degrades the NF and the sensitivity. Hence, it is desirable to adopt the S-to-D LNA, combining the Balun and LNA functionality (as an active Balun-LNA) into a single integrated circuit for application of a digital television tuner. However, the involved antenna is usually single-ended.

1: N. Rahimzadeh is with the Department of Communications, Faculty of Electrical and Computer Engineering, Semnan University, Semnan, Iran.

2: P. Rezaei is with the Department of Communications, Faculty of Electrical

and Computer Engineering, Semnan University, Semnan, Iran.

Corresponding author: prezaei@semnan.ac.ir

The digital television tuner IC must process the electrical signals with a wideband input frequency range of 48 to 864 MHz. Also, digital television tuner ICs and other radio and microwave receivers have several blocks like low noise amplifiers, mixers, oscillators, and other circuits that all of them have very linear and low power consumption performance characteristics. For this reason, the tuner IC of digital televisions needs a wideband low noise amplifier which has good input impedance matching, high differential voltage gain, low noise factor, and high linear and low power consumption performance characteristics in the desired frequency bandwidth [24].

Usually, all of these characteristics do not happen in one low-noise amplifier. But like many other low-noise amplifiers, this proposed Balun-LNA has an acceptable tradeoff between different characteristics like voltage gain, NF and power consumption. In this paper we propose a balun low-noise amplifier for high third order input intercept point and high second order input intercept point, high differential voltage gain, low noise factor, and low power consumption, which is one of the most important characteristics in low-noise amplifiers in wideband receivers, especially for tuners of digital television and for other wideband wireless radio and microwave receivers too.

II. THE TOPOLOGY AND PROPERTIES OF TYPICAL BALUN-LNAs

Single-ended input low noise amplifiers are typically addressed, as are single-ended output low noise amplifiers [25-27]. However, because differential signal processing has high common mode signal reflection and low second order distortion, inductorless differential low noise amplifiers have higher SNR [28-33]. For wideband applications like the tuner of digital television, integrating the passive transformer onto a chip is difficult. Accordingly, the single-input to single-output low noise amplifiers is used so that this technique integrates the Balun and low noise amplifier into one chip.

In recent years of study, several active Balun-LNAs with the performance of low noise, high differential voltage gain, and high linearity have been introduced. The technique of noise cancellation is one of the most satisfactory ways of designing a Balun-LNA with low noise and high gain that has been discussed. Fig. 1. (a) Shows the most basic Balun-LNA. In this circuit, CS and CG transistors are used in two different stages so that the source of the CG transistor is connected to the gate of the CG transistor and the input voltage.

In this figure, the CG transistor satisfies input impedance matching and the CS transistor cancels noise and distortion generated by the CG transistor because they appear in the form of common mode signals at differential outputs and are cancelled by the process of differential signals at the output of the LNA [34]. But this basic Balun-LNA can't afford high voltage gain and low power consumption with a low noise factor. Decreasing NF of CG-CS Balun-LNA requires that noise of the CS transistor decreases, because NF in this CG-CS Balun-LNA is approximately limited by the CS transistor. With increasing size and the transconductance of the CS transistor, the noise of it reaches a limited amount.

Fig. 1. (b) shows this designed structure. In Fig. 1. (b), the transconductance (g_m) of the CS transistor is increased N times

compared to the CG transistor, and the load resistor of the CS transistor is divided N times compared to the load resistor of CG [35]. In this circuit, since the transconductance of the CS transistor is increased N times, the current of the CS stage is increased N times too. In order to have identical current at the differential outputs of the CS and CG stages, the output resistor of the CS stage is divided N times. It causes the nonsymmetrical output load at the differential output. This condition leads to imbalance and mismatch in gain and phase of differential output, which increases second-order distortion and third-order distortion and makes the noise cancellation less perfect.

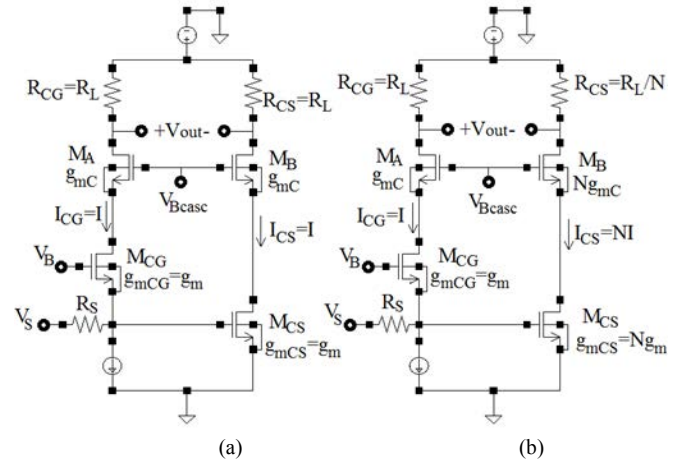


Fig. 1 CG-CS Balun LNAs: (a) conventional topology with $g_{mCS}=g_{mCG}$, $R_{CS}=R_{CG}$ and $I_{CS}=I_{CS}$ (b) practical topology with $g_{mCS}=N g_{mCG}$, $R_{CS}=R_{CG}/N$ and $I_{CS}=N I_{CG}$.

Thus, the designed circuit does not have an acceptable noise and linear coefficient, which is one of the most important features of low-noise amplifiers in digital television tuners. For this reason, we must solve this essential problem.

Another CGCS-based Balun-LNA that has been presented uses feedback between the CS and CG stages to increase CG transconductance while lowering power consumption [36]. The gain of the CS transistor is reused for this feedback to increase the transconductance of the CG transistor. A Balun-LNA is used with balanced loads in another paper [37]. A modified current bleeding is used in this structure to allow the Balun-LNA to use symmetrical currents at differential output. The performance of this structure is as follows: by increasing the transconductance of the CS transistor N times, the current of the CS transistor increases N times, and $N-1$ times of this current is consumed in the transistor of the current bleeding circuit, while the current of the CS stage at differential output remains constant.

As a result, this structure of Balun-LNA can have differential output with symmetrical loads and with any gain and phase unbalancing. In another LNA, feedback is used to reduce power consumption, so that in this structure, the local feedback increases the transconductance of the CG transistor [38]. In this Balun-LNA, a current bleeding circuit is also used for boosting the transconductance of the CS transistor and improving the voltage gain and reducing power consumption. A Balun low noise amplifier which has three inverters at its output is presented in another structure [39].

We consider the transconductance of the M_{BLD} transistor to be $(N-1)$ times the transconductance of the cascode transistors of M_A and M_B . In addition, we consider the resistive load of R_{BLD} to be $R_{L1}/(N-1)$. we multiply N times the transconductance of the CS transistor and therefore we multiply N times its current, the current amount of $(N-1)I$ goes through the M_{BLD} transistor and the current of I goes through the cascode transistor of M_B . As a result, the current and resistive load of the CS stage at the differential output remain constant and equal to the current and resistive load of the CG stage at the differential output. On the other hand, in the conventional Balun-LNA circuit, because input impedance matching is just determined by the CG transistor, the transconductance of the CG transistor and the output current of the CG stage remain fixed.

We use a local feedback technique in this proposed structure to increase the transconductance of the CG transistor and thus the current of the CG stage's output. This feedback multiplies the transconductance of the CG transistor by the transconductance of the CG transistor, increasing the amount of transconductance of the CG stage and the load current of the CG stage.

Also, as it will be shown by this feedback, the input impedance matching of the designed Balun-LNA does not just depend on the transconductance of the CG transistor, and the input impedance matching is determined by the transconductance of the CS transistor too. Therefore, we can change the transconductance of the CG transistor. Thus, by employing the technique of g_m boosting feedback and a modified current bleeding circuit, this proposed Balun-LNA achieves symmetrical resistive loads and current at the differential output.

b. Voltage Gain

The voltage gain from V_{in} to the output of the CS transistor stage (inverting output) is obtained as:

$$\frac{V_{out+}}{V_{in}} = g_{mCG}(1 + g_{mCS}R_{CS})R_{CG} \quad (1)$$

Also, the voltage gain of the proposed Balun-LNA from V_{in} to the output of the CG transistor (noninverting output) is obtained as this equation. :

$$\frac{V_{out-}}{V_{in}} = -g_{mCS}R_{CS} \quad (2)$$

As a result, the proposed Balun-LNA's differential voltage gain is calculated as:

$$A_V = g_{mCG}(1 + g_{mCS}R_{CS})R_{CG} + g_{mCS}R_{CS} \quad (3)$$

Which g_{mCG} is the transconductance of the CG transistor, R_{CG} is the load resistor of the CG transistor, g_{mCS} is the transconductance of the CS transistor, and R_{CS} is the load resistor of the CS transistor. According to equation (3), it can be seen that the differential voltage gain of the designed Balun-LNA has a direct relationship with the transconductance of CG and CS transistors and resistive loads of R_{CS} and R_{CG} . The differential voltage gain of the designed Balun-LNA increases

if the transconductance of CG and CS transistors and load resistors of R_{CS} and R_{CG} are boosted.

c. Input Impedance Matching

Input impedance matching of the proposed Balun-LNA is obtained by the following equation:

$$R_S = \frac{1}{g_{mCG}(1+g_{mCS}R_{CS})} \quad (4)$$

As it can be seen in the equation of (4), by applying feedback, the transconductance of the CS transistor is multiplied by the transconductance of the CG transistor by a coefficient of $(1+g_{mCS}R_{CS})$. Thus, the input impedance matching of this proposed Balun-LNA is not just determined by the transconductance of the CG transistor, and the transconductance of the CS transistor is another parameter in determining the input impedance matching. As a result, the transconductance of the CS transistor determines the impedance matching of this proposed Balun-LNA, and the amount of the transconductance of the CG transistor can be changed to boost the differential voltage gain, reduce the noise factor, and improve the power consumption of the proposed Balun-LNA.

d. Noise Analysis

In this proposed Balun-LNA, all noise sources are uncorrelated to each other. The generated noise of cascode transistors is negligible because most of their noise is rotated in their loops and doesn't appear at the output. By using the noise factor equation, because the generated noise by R_{BLD} is small enough to be negligible, it can be excluded from the noise factor [37], [28], [43]. Thus, almost the noise factor of Balun-LNA is limited by CS and CBLD transistors and resistive loads, and can be expressed as:

$$NF \cong 1 + \frac{\gamma}{g_{mCG}R_S} + \frac{\gamma g_{mCS}}{g_{mCG} \left(\frac{1}{1 + g_{mCS} \left(\frac{r_o}{N} \parallel Z_{INBLD} \right)} \right)^2} + \frac{(R_{CG}+R_{CS})(1+g_{mCG}(1+g_{mCS}R_{CS})R_S)^2}{R_S A_V^2} \quad (5)$$

Which in this equation, Z_{INBLD} is equal to:

$$Z_{INBLD} \cong \frac{R_{BLD} + r_{oBLD}}{1 + g_{mBLD}r_{oBLD}} = \frac{1}{g_{mCG} - 1} \frac{R_L}{1 + g_{mCS}} \quad (6)$$

R_S is the resistive of input source which has been aligned to 50Ω .

As can be seen in equation (5), by boosting the transconductance of CG and CS transistors, the NF of this proposed Balun-LNA reduces, besides the boosting of differential voltage gain.

IV. SIMULATION RESULTS

In 90-nm CMOS technology, the proposed Balun-LNA is designed and simulated. Resistors R1 and R2 are biased by the CG and CS transistors, while R3 and R4 are biased by the cascode transistors MA and MB, as well as the modified current

bleeding transistor MBLD. The output load resistors $R_{CS} = R_{CG}$ are designed to be equal to 3 k Ω . In this designed balun-LNA, the amount of $N = 5$ is considered. Furthermore, the output load resistor of the MBLD is equal to $R_{BLD} = R_L/(N-1) = 750$. The current of output loads is equal to $I = 0.9$ mA. The transconductance of CS and CG transistors is equal to 12 mS. This proposed Balun-LNA works in the operating frequency band of digital television, namely 48 to 864 MHz. Also, the proposed Balun-LNA works at a supply voltage of 2.8 v, which is the supply voltage of most digital television tuners. The power consumption of this proposed Balun-LNA is equal to 2.5 mW, which is very low compared to other LNAs. One of the most important achievements in this Balun-LNA is its low power consumption, which makes it suitable for tuners of digital televisions and other wideband radio and microwave receivers.

In Fig. 3, the curve of the noise factor (NF) of this proposed Balun-LNA is displayed. As it can be seen in this figure, the amount of NF in the whole operating frequency band of the proposed Balun-LNA, which is the operating frequency band of the tuner of digital televisions, is about 5 dB. The minimum amount of obtained NF is equal to 5 dB and is at a frequency of 400 MHz.

In Fig. 4, the curve of input return loss (S_{11}) of the proposed Balun-LNA can be seen. As it can be seen in Fig. 4., the proposed Balun-LNA has an input return loss of much less than -10 dB in the whole frequency band, which is suitable for tuners of digital televisions and other wideband applications.

In Fig. 5., the differential voltage gain of the designed Balun-LNA for application of the tuner of digital television is displayed. As it can be seen in Fig. 5., the differential voltage gain of the proposed Balun-LNA is changed between 23.8 and 24 dB, in a frequency range of 48 to 864 MHz. The maximum obtained voltage gain is until the frequency of 400 MHz, which is equal to 24 dB.

The table I shows the performance summary and comparison between the designed and proposed Balun-LNA with several other Balun-LNAs. As it can be seen in the table I, this designed Balun-LNA has higher differential voltage gain compared to previous structures and has an input return loss of less than -10 dB, which are essential parameters of a low-noise amplifier for wideband applications, especially for the tuners of digital television. Furthermore, the proposed Balun-LNA has a power consumption of 2.5 mW, which is significantly lower than that of many previous works. Having low power consumption is the most important advantage of this designed Balun-LNA. Reducing power consumption is one of the important characteristics, so that should be considered as a major characteristic in the design of the low noise amplifier in the tuner IC of digital television. The noise factor of this designed and proposed Balun-LNA is a little higher than other structures, but it is suitable for the application of tuners of digital television. Also, the proposed Balun-LNA has the performance of both balun and symmetrical loads at the output. The presence of symmetrical loads causes the gain and phase balance at the differential output, which is necessary for the tuner of digital televisions and other wideband applications.

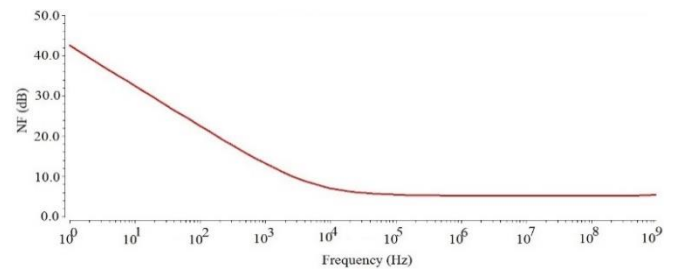


Fig. 3. NF of the proposed Balun-LNA.

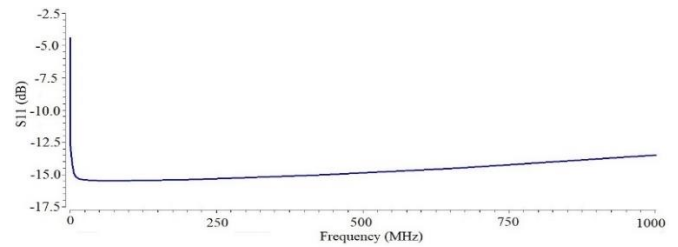


Fig. 4. S_{11} of proposed Balun-LNA.

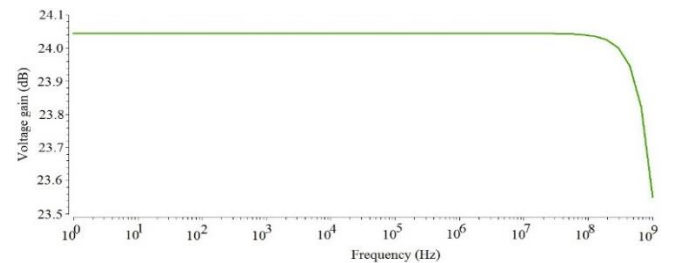


Fig. 5. Voltage gain of proposed Balun-LNA.

V. CONCLUSION

In 90-nm CMOS technology, a balun low noise amplifier employing gm-boosting feedback technique and current bleeding circuit for application of digital television tuner was designed and proposed. In this Balun-LNA, a circuit of modified current bleeding was used to boost the transconductance of the CS transistor. Also, a technique of g_m -boosting feedback was used for boosting the transconductance of the CG transistor. The proposed Balun-LNA works in the frequency band of digital television, namely 48 to 864 MHz, and achieves a maximum differential voltage gain of 24 dB at the frequency of 400 MHz. Also, input return loss of much less than -10 dB and a minimum NF of 5 dB are achieved in this proposed Balun-LNA. This proposed Balun-LNA has a balun function and symmetrical loads at the differential output. Using the symmetrical loads causes the gain and phase balance at the differential output to be symmetrical. This designed and presented Balun-LNA has achieved a power consumption of 2.5 mW, which is very low and suitable for tuners of digital television.

ACKNOWLEDGMENT

The authors acknowledge the Semnan University staff for their beneficial and professional help. The authors would like to thank the journal editor and reviewers for their valuable comments.

TABLE I

Comparison of proposed Balun-LNA and previous works.

Ref.	CMOS Process	Frequency range (GHz)	S11 (dB)	Voltage gain (dB)	NF (dB)	Symmetric load	Power consumption (mW)	Supply voltage (v)
[1]	180 nm	0.05-0.86	<-8.5	17-18	2.5-3	YES	30	1.2
[2]	130 nm	0.2-3.8	<-10	16-19	2.8-3.4	YES	5.7	1
[4]	28 nm	1-6.2	<-10	48.2	3.4-4.2	YES	22.2	1.8
[11]	130 nm	0.01-1	<-10	14.5	2.5	NO	3.6	1.2
[24]	40 nm	1-11	<-10	14-17	3.5-5.5	YES	9	1.2
[30]	28 nm	0.02-2	<-10	18.5	2.5-3.5	NO	4.1	1
[37]	65 nm	0.05-1	<-10	24-30	2.2-3.3	YES	19.8	2.2
[38]	65 nm	0.05-1.3	<-10	24-27.5	2.3-3	YES	5.7	1
[43]	130 nm	0.1-1	<-10	14	2.4	YES	2.7	1.2
[44]	180 nm	0.13-0.93	<-10	16.6-19.6	3.6-5	YES	3	1.8
This work	90 nm	0.04-0.864	<-10	24	5	YES	2.5	2.8

VI. REFERENCES

- [1] D. Im, I. Nam and K. Lee. A CMOS active feedback balun-LNA with high IIP2 for wideband digital TV receivers. *IEEE Trans. Microw. Theory Techn.* 58(12), pp. 3566-3578, 2010.
- [2] H. Wang, L. Zhang and Zh. Yu. A wideband inductorless LNA with local feedback and noise cancelling for low-power low-voltage applications. *IEEE Trans. Circ. Syst. -I: Regular Papers.* 57(8), 2010.
- [3] D. Im, O. Lee and I. Nam. A TV Receiver Front-End with Linearized LNA and Current-Summing Harmonic Rejection Mixer. *IEEE Trans. Circuits Syst. II.* 64(3), pp. 269-273, 2017.
- [4] B. Guo, D. Prevedelli, R. Castello, and D. Manstretta, A 0.008 mm² 1-6.2 GHz Receiver Front-End with Inverter-Base Shunt-Feedback Balun-LNA. *IEEE Radio Frequency Integrated Circuits Symposium*, 2020.
- [5] Bevilacqua, A. and Niknejad, A. M. An Ultra-Wideband CMOS Low-Noise Amplifier for 3.1 - 10.6-GHz Wireless Receivers," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2259-2267, 2004.
- [6] R. Bagheri, et al. An 800-MHz-6-GHz software-defined wireless receiver in 90-nm CMOS. *IEEE J. Solid-State Circ.* 41(12), pp. 2260-2262, 2006.
- [7] M. Hotti, et al. An IIP2 Calibration Techniques for Direct Conversion Receivers. *IEEE Int. Symp. Circuits and Syst.*, pp. 257-258, 2004.
- [8] C. Svensson. The Blocker Challenge When Implementing Software Defined Radio Receiver RF Frontends. *Analog Integr. Circuits Signal Processing*, pp. 81-82, 2010.
- [9] J. Ryyanen, et al. A Dual-Band RF Front-End for WCDMA and GSM Applications. *IEEE J. Solid-State Circuits.* 36(8), pp. 1198-1199, 2001.
- [10] P. I. Mak and R. P. Martins. A 0.46-mm² 4-Db NF unified receiver front-end for full-band mobile TV in 65-nm CMOS. *IEEE J. Solid-State Circ.* 46(9), 2011.
- [11] S. Jang, S. Park and D. Im, A PVT Insensitive Noise Canceling Balun-LNA for TV Receiver Application. *Progress in Electromagnetics Research Symposium-Fall (PIERS-FALL)*. pp.1726-1728, 2017.
- [12] D. Mastantuono, and D. Manstretta, A Low-Noise Active Balun with IM2 Cancellation for Multiband Portable DVB-H Receivers. *IEEE Int. Solid-State Circuits Conf.* pp. 215-217, 2009.
- [13] D. Im, I. Nam and S.S. Song. A CMOS resistive feedback single to differential low noise amplifier with multiple-tuner-outputs for a digital TV tuner. *IEEE Radio Freq. Integr. Circuits symp.* pp. 555-558, 2009.
- [14] S. Jin, T.Y. Oh, K.T. Hong, H.T. Kim and B. Kim. Wide-Band CMOS Loop-Through Amplifier for Cable TV Tuner. *IEEE Radio Freq. Integr. Circuits Symp.* pp. 215-218, 2008.
- [15] K. Kwon, H.T. Kim and K. Lee, A 50-300-MHz highly linear and low-noise CMOS Gm-C filter adopting multiple gated transistors for digital TV tuner ICs., *Trans. Microw. Theory Techn.* 57(2), pp. 306-312, 2009.
- [16] N. Vitee, H. Ramiah, P. I. Mak, and R. P. Martins. A 3.15-mW +16.0-dBm IIP3 22-dB CG inductively source degenerated balun-LNA mixer with integrated transformer-based gate inductor and IM2 injection technique. *IEEE Trans. Very Large Scale Integration Syst.* 28(3), pp. 700-710, 2020.
- [17] W. H. Chen, et al. A highly linear broadband CMOS LNA employing noise and distortion cancellation. *IEEE J. Solid-State Circ.* 43(5), pp. 1164-1173, 2008.
- [18] M. El-Nozahi, et al. A CMOS low-noise amplifier with reconfigurable input matching network. *IEEE Trans. Microw. Theory Techn.* 57(5), pp. 1054-1059, 2009.
- [19] H. Koizumi, et al. A GaAs single balanced mixer MMIC with built-in active balun for personal communication systems. *IEEE Microw. Millimeter-Wav. Monolithic Circ. Symp.* pp. 77-79, 1995.
- [20] H.-K. Chen, et al. 5.8-GHz merged LNA-mixer with on-chip balun. *Microw. Opt. Technol. Lett.* 48(3), pp. 508-511, 2006.
- [21] M. Ch. Kuo, Ch. N. Kuo and T. Ch. Chueh. Wideband LNA compatible for differential and single-ended inputs. *IEEE Microw. Wirel. Compon. Lett.* 19(7), pp. 482-484, 2009.
- [22] M. A. Martins, et al. A single-to-differential LNA topology with robust output gain-phase balancing against balun imbalance. *IEEE Int. Sym. Circ. Syst. (ISCAS)*. pp. 289-291, 2011.
- [23] S. Joo, et al. A 3-to-5 GHz UWB LNA with a low-power balanced active balun. *IEEE Radio Freq. Integr. Circ. Symp.* pp. 303-306, 2009.
- [24] Z. Liu, et al. A 0.061-mm² 1-11GHz noise-cancelling low-noise amplifier employing active feedforward with simultaneous current and noise reduction. *IEEE Trans. Microw. Theory Techn.* 69(6), pp. 3093-3104, 2021.
- [25] S. Masih, P. Rezaei and M. Panahi. Compact chip-resistor loaded active integrated patch antenna for ISM band applications. *Wirel. Pers. Commun.* 97(4), pp. 5738-5746, 2017.
- [26] A. Valizadeh, P. Rezaei and A.A. Orouji, Design of reconfigurable active integrated microstrip antenna with switchable low-noise amplifier/power amplifier performance for wireless local area network and WiMAX applications. *IET Microw. Antennas Propag.* 9(9), pp. 874-881, 2015.
- [27] A. Valizadeh, P. Rezaei, and A.A. Orouji, A new design of dual-port active integrated antenna for 2.4/5.2 GHz WLAN applications. *Prog. Electromag. Res. B.* 58, pp. 86-94, 2014.

- [28] N. Rahimzadeh and P. Rezaei. A balun-LNA employing positive feedback and modifies current-bleeding technique and symmetrical loads for tuner of digital televisions. 3rd Iranian Conf. Microelectron. Tarbiat Modares University, Dec. 2021 (in Persian).
- [29] N. Rahimzadeh and P. Rezaei. Decreasing power consumption of BLNA with balanced output with gm boosting feedback for application of tuner of digital televisions. 6th Conf. Elec. and Comput. Eng. Tech., Tafresh University, March 2022. (in Persian).
- [30] A. Bozorg and R.B. Staszewski. A 20 MHz-2 GHz inductorless two-fold noise-canceling low-noise amplifier in 28-nm CMOS. *IEEE Trans. Circ. Syst. I: Regular Papers*, vol. 69(1), pp. 42-50, 2022.
- [31] Sh. Tiwari, and J. Mukherjee, An inductorless noise cancelling wideband balun LNA with dual shunt feedback and current reuse. *IEEE 62nd Int. Midwest Symp. Circ. Syst.* pp. 432-435, 2019.
- [32] T. Chang, et al. ESD-protected wideband CMOS LNAs using modified resistive feedback techniques with chip-on-board packaging," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 8, pp. 1817-1819, 2008.
- [33] H. Ma, et al. Novel active differential phase splitters in RFIC for wireless applications. *IEEE Radio Freq. Integr. Circ. Symp.*, pp. 51-52, June 1998.
- [34] F. Bruccoleri, EA and M. Klumperink. Wide-Band CMOS Low-Noise Amplifier Exploiting Thermal Noise Canceling. *IEEE J. Solid-State Circ.* 39(2), pp. 275-277, 2004.
- [35] S.C. Blaakmeer, et al. Wideband balun-LNA with simultaneous output balancing, noise-canceling and distortion-canceling. *IEEE J. Solid-State Circ.* 43(6), pp. 1341-1347, 2008.
- [36] J. Kim and J. Silva-Martinez, Wideband Inductorless Balun-LNA Employing Feedback for Low-Power Low-Voltage Applications. *Trans. Microw. Theory Techn.* 60(9), 2012.
- [37] S. Kim and K. Kwon. (2019, Feb.). A 50-MHz–1-ghz 2.3-dB NF Noise-cancelling Balun-LNA Employing a Modified Current-Bleeding Technique and Balanced loads," *IEEE Trans. Circuits Syst. I: Regular Papers*. 66(2), pp. 546–554.
- [38] S. Kim and K. Kwon. Broadband balun-LNA employing local feedback gm-boosting technique and balanced loads for low-power low-voltage applications. *Trans. Circ. Syst. I.* 67(12), pp. 4631-4640, 2020.
- [39] J. Yi, Ch. Liu and J. Sh. Chen. A wideband inductorless single-to-differential LNA in 0.18 μ m CMOS technology for digital TV receivers. *IEEE Microw. Wirel. Compon. Lett.* 24(7), 2014.
- [40] S. Blaakmeer, E. Klumperink, D. Leenaerts and B. Nautra, A Wideband Balun LNA I/Q-Mixer Combination in 65nm CMOS. *IEEE Int. Solid-State Circ. Conf.* pp. 326-328, 2008.
- [41] B. ShirMohammadi and M. Yavari. A low power wideband balun-LNA employing local feedback, modified current-bleeding technique, and balanced loads. 28th Iranian Conf. Elec. Eng. (ICEE), Aug. 2020.
- [42] D. Manstretta. A broadband low-power low-noise active balun with second-order distortion cancellation," *IEEE J. Solid-State Circ.* 47(2), 2012.
- [43] D. Kim, et al. A broadband PVT-insensitive all-nMOS noise-cancelling balun-LNA for subgigahertz wireless communication applications. *IEEE Microw. Wirel. Compon. Lett.* 31(2), pp. 165-167, 2021.
- [44] Sh. Tiwari and J. Mukherjee. An inductorless wideband gm-boosted balun LNA with nMOS-pMOS configuration and capability coupled loads for sub-GHz IOT applications. *IEEE Trans. Circ. Syst. II: Express Briefs.* 68(10), pp. 3204-3207, 2021.
- [45] L. Liu, Zh. Lu, K. Zhang, Zh. Ren, A. Hu and X. Zou. Wideband balun-LNA exploiting noise cancellation and gm compensation technique. *Electron. Lett.* 52(8), pp. 673-674, 2016.
- [46] J. Borremans, et al. Low-area active-feedback low-noise amplifier design in scaled digital CMOS. *IEEE J. Solid-State.* 43(11), pp. 2422-30, 2008.
- [47] J. Kim, S. Hoyos and J. Silva-Martinez. Wideband common-gate CMOS LNA employing dual negative feedback with simultaneous noise, gain, and bandwidth optimization. *Trans. Microw. Theory Techn.* 58(9), 2010.
- [48] B. Afshar, and A.M. Niknejad, X/Ku band CMOS LNA design techniques. *IEEE Custom Integr. Circ. Conf.* pp. 389-390, 2006.
- [49] S. Woo, et al. A 3.6 mW differential common-gate CMOS LNA with positive-negative feedback. *IEEE Int. Solid-State Circ. Conf.*, pp. 217-219, Feb. 2009.

Improvement of Mesh Simplification Using Normal Vector Diversity

Masoud Ebadi¹, Kourosh Kiani² and Razieh Rastgoo³

Abstract-- 3D mesh simplification is an important challenge in various fields. While different simplification methods have been proposed in recent years, the focus has shifted to keeping properties such as ridges and valleys along with mesh simplification. While most of the proposed models have used curvature, some challenges exist, such as the computational complexity and sensitivity to the neighborhood size. The latter can be solved by averaging several neighborhoods. This paper proposes a simple yet fast method with less sensitivity to the neighborhood size. To this end, we use the normal vector and the parameters of a probability distribution of its variations to detect the elevations, depressions (geometrical changes), and curve parts. We combine this method with the Quadric Error Metric (QEM) method to produce a hybrid method for 3D mesh simplification, preserving its elevations and depressions. Evaluation results show that our method has a lower error than the other methods.

Index Terms-- 3D Mesh simplification, Normal vector, Curvature, Improvement, Sensitivity.

I. INTRODUCTION

Nowadays, 3D models are used in various fields, such as virtual reality (VR) [1], heterogeneous materials [2], and Infrastructural Work [3]. Different high complexity methods are used to provide 3D models in each field. For example, in some computer games, many complicated mathematical calculations are necessary to have an accurate collision detection of some objects. These calculations can be able to perform on less detailed models. Therefore, a suitable simplification method is required to reduce the model complexity. In general, simplification methods decrease the calculations, memory usage, and volume of the transmitted data while keeping the quality of the results. Different researchers have proposed different simplification methods [4-28]. These methods can be categorized into two general groups: sensitive to feature and insensitive to feature methods. While the sensitive feature methods consider special priorities for different features of the model, the insensitive feature methods act equally on all parts of the model without special attention to

different features. The latter only considers an error measurement function to control the operation and keeps the simplified model similar to the reference. These methods maintain special properties, like curvature, by directing and controlling operations during simplification. The basic algorithm of both categories is similar, but the first category models act somewhat intelligently to increase or decrease the level of simplification. One of the solutions to control the level of simplification is the detection of the important regions and saliency in two-dimensional images, which generally state that regions with lots of changes in light intensity are more important for the human visual system. However, the light intensity is substituted by the geometrical data in 3D models. So, regions with lots of geometrical changes are more important from the point of view of the human visual system. This way, finding the correlation of several features with regions selected by humans is used to find the important regions in 3D models [18]. The curvature feature is the most important among other features. This feature is usually used for determining geometrical changes. In line with this, in this paper, we investigate the curvature problem and propose a simplification method for determining the regions of elevation and depression by using the normal vector of the vertex.

The rest of this paper is organized as follows: In section 2, normal diversity expression is introduced. Related work and the proposed method are presented in sections 3 and 4, respectively. Section 5 demonstrates the experimental results of the proposed method. Finally, the conclusions are drawn in section 6.

II. NORMAL DIVERSITY EXPRESSION

We aim to search the regions that are more visually important and assist in reducing the damage to the sensitive mesh regions. There are several methods to detect the considered regions. The most common way among them is based on the curvature [16], in which the amount of geometrical changes in a region is the criterion to measure the region's importance. This procedure comes from 2D image processing, where more light intensity changes may be more important to

1: M. Ebadi is with the Electrical and Computer Engineering Faculty Semnan University, Semnan, Iran

2: K. Kiani is with the Faculty of Electrical and Computer Engineering, Semnan University, Semnan, Iran.

3: R. Rastgoo is with the Electrical and Computer Engineering Faculty Semnan University, Semnan, Iran

Corresponding author: kourosh.kiani@semnan.ac.ir

the eyes. This way, we first study the curvature-based methods and the corresponding concepts.

The curvature is one of the geometrical properties representing a deviation from a flat surface in a specific direction. This is a local feature and varies for different points of the object. From a mathematical point of view, for a 2D curve $F(t)$ (that is, in parametric form), the curvature of the point $K(t)$ is equal to the ratio of angle changes of the normal vector of the curve (between the considered point and a point in its vicinity) to over the curve length between two points (Fig. 1). It can be calculated as (1):

$$F(t) = g(t) * \vec{i} + h(t) * \vec{j} \quad (1)$$

$$T(t) = \frac{F'(t)}{|F'(t)|}$$

$$K(t) = \frac{T'(t)}{|F'(t)|}$$

Where $g(t)$ and $h(t)$ are the parametric functions of i and j . $F'(t)$ is the parametric derivation, $T(t)$ is the tangent vector on the curve, and $T'(t)$ is the derivation of the curve. For 3D surfaces, the surface is cut in one direction, and the curvature is calculated for the resulting curve. The results are obtained in different numbers in different directions. The smallest and biggest numbers are called min and max curvature, respectively.

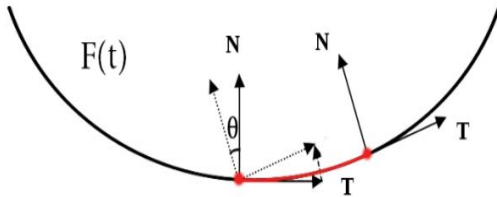


Fig. 1. The curve of $F(t)$ and angle changes of a normal vector.

Two common methods are used to calculate curvature in the mesh where space is discrete: The first group includes methods that initially estimate the surface using a continuous function, then using the second derivation to compute the curvature. The second group estimates the second derivation and curvature directly without the continuous function estimation and using the geometrical data of the mesh such as face normal vector or face angle. One of the advantages of the second group is high performance. However, they also have larger errors. Curvature calculation needs a neighborhood for each point. The curvature determines the deflection amount of the surface at each point, but we are searching for elevation and depression in an area. Averaging in a ring is usually used to achieve this result. The size of the ring is important. A smaller ring gives more details, while more local and larger rings result in more global information. Determining the appropriate ring size also affects the correct determination of saliency. So, in addition to the

neighborhood of the curvature calculation, the averaging also needs the neighborhood rings.

III. RELATED WORK

One well-known proposed work in mesh saliency is [16] where the mean curvature is calculated for each point using a neighborhood $(N(v, \sigma))$. Then, a weighted mean for each vertex is calculated in another neighborhood $(G(\vartheta, \sigma))$ to reduce the effect of the size of $N(v, \sigma)$. The weights are calculated based on a Gaussian neighborhood centered on the considered vertex. After that, two neighborhoods $(G(\vartheta, \sigma))$ with a little difference in size are selected and the curvature is subtracted to calculate the saliency. The variance or width of this Gaussian neighborhood affects saliency results. The larger neighborhood results in more continuous regions as shown in Fig. 2. Finally, to solve this problem, saliency is calculated in different sizes and after normalizing the results, they are added to each other as shown in Fig. 2. Therefore, determining the elevation and depression regions using curvature is both costly and time-consuming and also very sensitive to the size of neighborhoods at different steps. Changes in the neighborhood size can cause serious differences in the results. To tackle these challenges and consider the curvature behavior, we propose a method using a normal vector of vertices to reduce the challenges and improve the results. The proposed method attempts to determine regions with lots of geometrical changes at a low cost and without complex calculations. These regions should be continuous as possible and the results must be approximately consistent, with minor parameter changes that can be used in different meshes. Vertex normal vectors can provide the necessary solution.

The normal vector is one of the vertex properties in the mesh. Its most important application is facing angle identification for lighting in computer graphics. As shown in Fig. 3, the normal vector has several directions in the regions with curvature. This property helps to use a normal vector instead of curvature to identify the change areas. This will tackle the challenge of the computational complexity of the curvature. However, the challenge of sensitivity to the neighborhood size will remain. The curvature computes the amount of deflection that is strong or weak according to the speed of the surface change. It seems this causes curvature changes by changing the neighborhood because there are different changes in the mesh. However, we do not care about the intensity of changes for simplification. As Fig. 3 shows, the normal vectors have different directions where there are geometrical changes. As with other natural phenomena, these changes clearly have a distribution probability. We don't know this distribution, but a general distribution can be supposed. Most natural phenomena follow a normal or Gaussian distribution. Thus, we assumed that the distribution is Gaussian. Usually, there are different changes in the mesh surface, so the pattern of direction change is not the same everywhere. We are going to estimate the mean and

variance of direction changes statistically. Therefore, we calculate the mean and variance for each vertex in a neighborhood, including some rings, with the following equations:

$$m(v) = \frac{\sum_{i=1}^N n_i}{N} \quad (2)$$

$$s(v) = \frac{\sum_{i=1}^N (n_i - m(v))^2}{N} \quad (3)$$

where $m(v)$ is the mean of normal vectors in the neighborhood of v and $s(v)$ is the normal vector variance around the mean vector. N is the number of vertices in the neighborhood, and n_i is the normal vector of the i -th vertex. It should be noted that the calculations are done separately in each dimension of 3D space. We aim to identify the direction of changes. As shown in Fig. 3, where the direction changes, there is diversity, and the intensity of the diversity represents the number of changes. Therefore, $s(v)$ is directly proportional to the direction of changes. Each dimension of $s(v)$ shows the diversity in that dimension on the mesh. So, we can use the magnitude of $s(v)$ as the criterion to detect the number of geometrical changes. Using the normal vector and the probability distribution of the normal vector direction changes (because of less sensitivity to the surface changes) to find the regions with geometrical changes that lead to more continuous output without using the complex method as [16]. Fig. 4 shows the differences between the proposed model and the model of [16]. It should be noticed that normal vector information of vertices usually is stored with mesh files, so it is not required to calculate it. If there was no mesh data, calculating a normal vector is not somewhat difficult, and because of the elevation and depression, the finding procedure is based on the probability distribution. Changing the neighborhood size doesn't change the results until the pattern of the direction change does not vary.

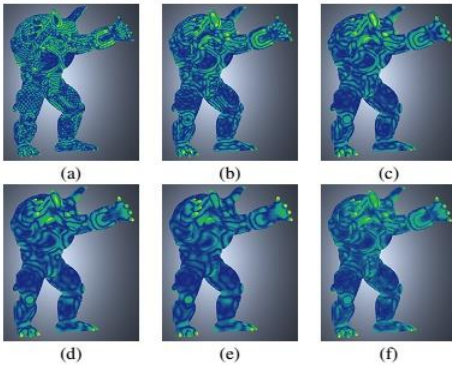


Fig. 3. Result of saliency detection [16] Results have been changed by scale changes.

IV. THE PROPOSED METHOD

The basic purpose of this work is mesh simplification. As mentioned before, we aim to decrease the changes in the

important areas during mesh simplification. Since human eyes are more sensitive to areas with changes, we tend to change these areas less than the others. The way to identify changing areas has been explained in section II. In the previous steps, the obtained weight for each vertex represents the geometrical changes around the vertex. Now, we are going to use this weight for simplification.

Furthermore, an error measurement or cost function in most

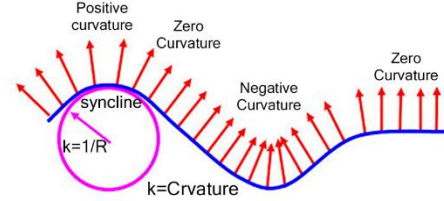


Fig. 2. The position of the normal vector at various locations of curve.

simplification algorithms controls the element selection for elimination and keeps the criterion small. Using the calculated weight in the error measurement of such algorithms allows more attention to the important areas. We combined our work with the method introduced in [4] First, we explain this method and combine our weight with it. The method in [4] is based on edge contraction. More concretely, in this method, vertex v is replaced by an edge (v_1, v_2) (Fig. 5). This algorithm follows the edges that lead to the minimum changes in the mesh. To this end, the cost is calculated for each contraction.

The v could be anywhere between v_1 and v_2 . Each point has a different cost. The minimum cost for v is selected by [4] The cost is the summation of the squared distances between v and the planes that meet v_1 and v_2 . Here, the plane means the planes created by using the normal vector of each triangle met by the vertex and the coordinate of the vertex. The cost is updated after each contraction. After sorting, an edge with a minimum cost is removed, causing a minimum change in the shape at each step. This action has high computational complexity due to the duplicate calculations of distances. To deal with this challenge, [4] proposes a method, including a mathematical proof, which is led to the squared distance from vertex v to a plane p , as shown in (4).

$$d(v) = v k_p v^T \quad (4)$$

where k_p is a matrix based on the coefficients of the plane equation, p and v are the considered points, and T is the transpose sign. Initially, k_p is calculated for each triangle; then, the summation of k_p corresponds to meeting triangles for each vertex calculated (Q). To calculate the distance from v to meet planes with v_1 and v_2 , the specified Q are used as follows:

$$C(v) = v(Q_1 + Q_2)v^T \quad (5)$$

$$Q_i = \sum_{p \in \text{planes of } v_i} k_p$$

where $C(v)$ is the cost of replacing vertex v to edge (v_1, v_2) and others, as in (4). In [4] for new vertex v , Q equals Q_1+Q_2 , and recalculation is not required. Now, if we want to apply the considered weight to each vertex, we need to multiply the weight by (4), which needs the changing (4) and (5) to (6) and (7).

$$d(v) = w_v(v k_p v^T) = v(w_v k_p)v^T \quad (6)$$

$$C(v) = v(wQ_1 + wQ_2)v^T \quad (7)$$

$$wQ_i = w_v \sum_{p \in \text{planes of } v_i} (k_p)$$

where w_v is our weight, other parameters are the same as (4) and (5). This way, we engaged the procedure's importance in high change areas. To create more differences between the areas, we also add a nonlinear transformation as follows:

$$W_{v=(\omega_v)^\gamma} \quad (8)$$

where ω_v is the calculated weight in the last step and γ is an exponent to create nonlinear differences between low and high change areas. Fig. 6 shows an example of output where QEM had done the simplification with and without the calculated weights.

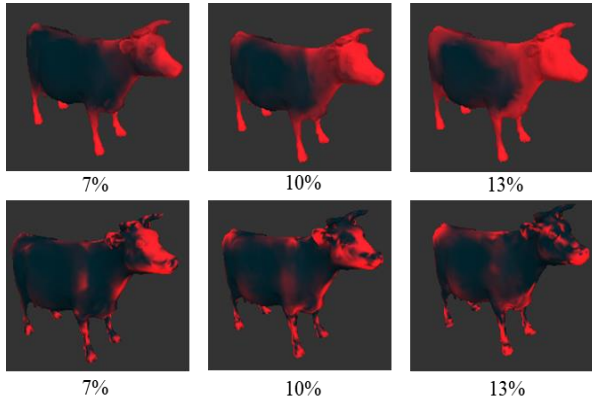


Fig. 4. Top row- Identification of normal diversity for different neighborhoods. Bottom row- mean curvature results for different neighborhoods. The size of the neighborhood is determined based on the bounding box diameter. More red means more geometric.

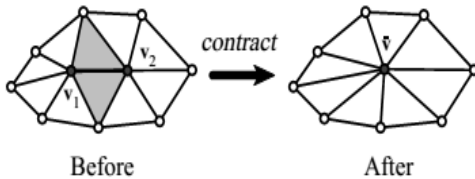


Fig. 5. Edge contraction.

V. EXPERIMENTAL RESULTS

In this section, several experiments are done, and the results are analyzed. First, an experiment is done to find the appropriate value of γ and neighborhood size (d). After determining the appropriate values of these parameters, our results and QEM

results are compared. Finally, the outputs of some well-known models are shown and compared. We performed these experiments with C++ and used the official implementation of QEM with a dataset provided by Princeton University for segmentation evaluation, including 380 models in 19 different classes. The average of the mean curvature is used as the comparison criterion.

A. Determining γ and d

Both parameters, γ , and d , are directly related to the output; therefore, we changed each in an interval. By each change, the outputs for 100 identical models are obtained. The difference between the average mean curvature before and after simplification is considered the error of combination (γ , d). The neighborhood is considered based on a percentage of the diameter of the bounding box (Fig. 6). We change γ in [1 - 9] and d in [4-14] intervals. The best state is (5, 11), but it is not stable. Thus, to find an appropriate γ and d , each dimension's average is calculated separately according to the values of other dimensions (to understand error changes based on each parameter). The results are shown in Fig. 7. The minimum error occurs at 5 and 11 for γ and d , respectively, consistent with the previous results.

B. Comparison with QEM

The results of this work are compared with the results of the QEM. In other words, this is a comparison between the QEM with and without a normal diversity. The γ and d are set to the results of 4.1. For comparison, the level of simplification is changed from 10% to 80% by step 10%. Each method is tested with 100 identical models. The error criterion is the average of mean curvature. Fig. 7 shows the results. It is observed that our method has a lower error.

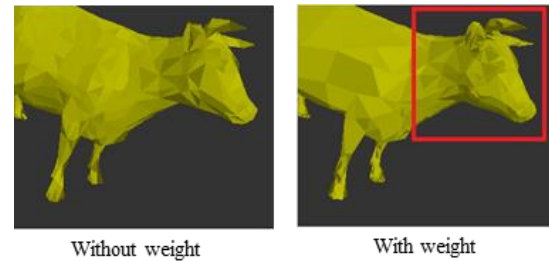


Fig. 6. Differences of output with and without normal diversity weight. Number of initial triangles: 5084 rate of simplification: 70%, retained triangle: 1074. The specific area of the right side has more density.

C. Output

In this section, we show our output for several models. The output of normal diversity is shown in Fig. 8, and the simplification output in Fig. 9 shows that the simplification level is not the same everywhere when using the weight of normal diversity. In the flat areas, there is more simplification level, and the number of triangles is less while their size is larger. In contrast, there are more triangles and smaller ones in the areas with more geometrical changes, causing a more accurate initial shape.

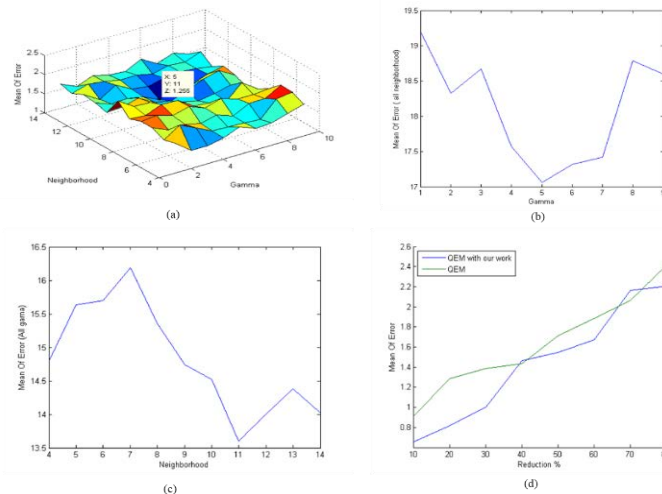


Fig. 7. (a) Error change according to d and γ , (b) The mean of error of all neighborhoods based on gamma, (c) The mean of error of all Gammas based on neighborhood size, (d) The error of the two methods comparison.

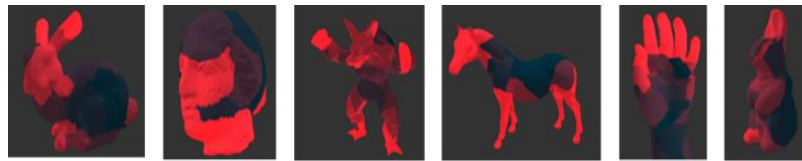


Fig. 8. Differences of output with and without normal diversity weight. Number of initial triangles: 5084 rate of simplification: 70%, retained triangle: 1074. The specific area of the right side has more density.

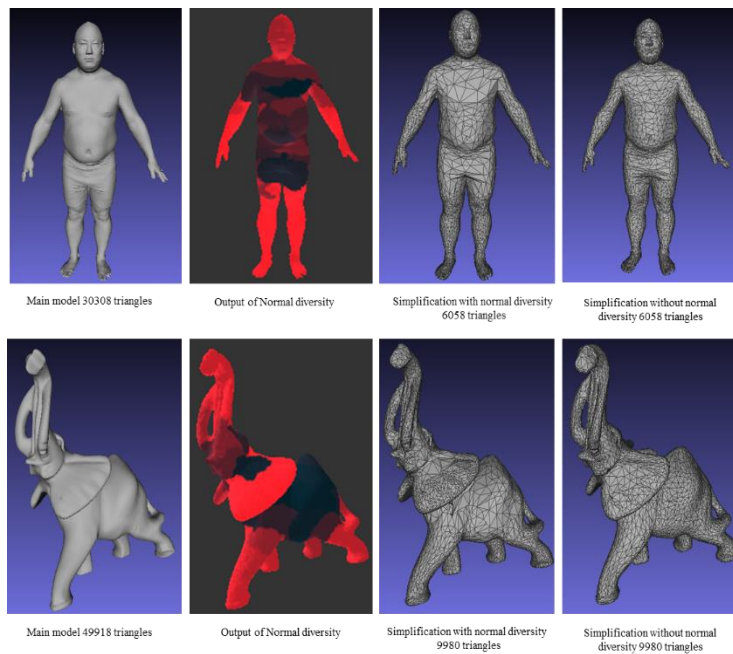


Fig. 9. Simplification output- normal diversity caused more simplification in the flat regions and less in the areas with geometrical changes. Simplification level is 80%.

VI. CONCLUSION

In this paper, we proposed a method to detect the elevations and depressions in the 3D mesh by using a normal vertex vector based on its probability distribution function, which does not have the problem of using curvature. We used the mean and variance of the normal distribution corresponding to the normal vector direction to detect the geometrical changes. Unlike the curvature, which usually suffers from high computational complexity and high sensitivity to the size of its neighborhoods, our method is faster and less sensitive, and its results are

consistent with each other when neighborhood size is changed. We combined our results as a weight for each vertex with QEM. To this end, the calculated weight was multiplied by the initial errors of vertices. The simplification results show that the proposed method maintained the elevations, depressions (geometrical changes), and curve parts better than the general QEM. We used the average mean curvature as the error measure of elevations and depression changes. In future works, we would like to employ deep learning-based models in the field.

VII. REFERENCES

- [1] Yuk Ming Tang, H.L. Ho, 3D Modeling and Computer Graphics in Virtual Reality, Mixed Reality and Three-Dimensional Computer Graphics, 2020, 10.5772/intechopen.91443.
- [2] Swantje Bargmann, Benjamin Klusemann, Jürgen Markmann, Jan Eike Schnabel, Konrad Schneider, Celal Soyarslan, Jana Wilmers, Generation of 3D representative volume elements for heterogeneous materials: A review, *Progress in Materials Science*, Volume 96, July 2018, Pages 322-384.
- [3] Mauro Mazzei and Davide Quaroni, Development of a 3D WebGIS Application for the Visualization of Seismic Risk on Infrastructural Work, *Int. J. Geo-Inf.* 2022, 11(1), 22; <https://doi.org/10.3390/ijgi11010022>.
- [4] Rumeng Lv, Xiaobing Chen, and Bingying Zhang, A simplified algorithm for 3D mesh model considering the influence of edge features, *Journal of Physics: Conference Series*, 2021.
- [5] Guangyou Zhou; Shangda Yuan; Sumei Luo, Mesh Simplification Algorithm Based on the Quadratic Error Metric and Triangle Collapse, *IEEE Access*, Vol. 8, pp. 196341–196350, 2020.
- [6] Schroeder, William J., Jonathan A. Zarge, and William E. Lorensen. "Decimation of triangle meshes." *ACM Siggraph Computer Graphics*. Vol. 26. No. 2. ACM, 1992.
- [7] Renze, Kevin J., and James H. Oliver. "Generalized unstructured decimation [computer graphics]." *Computer Graphics and Applications*, IEEE 16.6 (1996): 24-32.
- [8] Hoppe, Hugues. "Progressive meshes." *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996.
- [9] Garland, Michael, and Paul S. Heckbert. "Surface simplification using quadric error metrics." *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 1997.
- [10] Cao, Weiqun, Hujun Bao, and Qunsheng Peng. "An algorithm for LOD by merging near coplanar faces based on gauss sphere." *Journal of Computer Science and Technology* 16.5 (2001): 450-457.
- [11] DeHaemer Jr, Michael J., and Michael J. Zyda. "Simplification of objects rendered by polygonal approximations." *Computers & Graphics* 15.2 (1991): 175-184.
- [12] Hamann, Bernd. "A data reduction scheme for triangulated surfaces." *Computer aided geometric design* 11.2 (1994): 197-214.
- [13] Schaefer, Scott, and Joe Warren. "Adaptive vertex clustering using octrees." *Geometric Design and Computing* 2.5 (2003).
- [14] Bayik, Tolga, and Mehmet B. Akhan. "3d object database simplification using a vertex clustering algorithm." University of West Bohemia, Plzen, Czech Republic. 1999.
- [15] Turk, Greg. "Re-tiling polygonal surfaces." *ACM SIGGRAPH Computer Graphics* 26.2 (1992): 55-64.
- [16] Campomanes-Álvarez, B. Rosario, Sergio Damas, and Oscar Cordon. "Mesh simplification for 3D modeling using evolutionary multi-objective optimization." *Evolutionary Computation (CEC)*, 2012 IEEE Congress on. IEEE, 2012.
- [17] Huang, Hui-Ling, and Shinn-Ying Ho. "Mesh optimization for surface approximation using an efficient coarse-to-fine evolutionary algorithm." *Pattern Recognition* 36.5 (2003): 1065-1081.
- [18] Álvarez, Rafael, et al. "A mesh optimization algorithm based on neural networks." *Information Sciences* 177.23 (2007): 5347-5364.
- [19] Hoppe, Hugues. "New quadric metric for simplifying meshes with appearance attributes." *Proceedings of the conference on Visualization'99: celebrating ten years*. IEEE Computer Society Press, 1999.
- [20] Wei, Jin, and Yu Lou. "Feature preserving mesh simplification using feature sensitive metric." *Journal of Computer Science and Technology* 25.3 (2010): 595-605.
- [21] Lee, Chang Ha, Amitabh Varshney, and David W. Jacobs. "Mesh saliency." *ACM Transactions on Graphics (TOG)*. Vol. 24. No. 3. ACM, 2005.
- [22] Wu, Jinliang, et al. "Mesh saliency with global rarity." *Graphical Models* 75.5 (2013): 255-264.
- [23] Chen, Xiaobai, et al. "Schelling points on 3D surface meshes." *ACM Transactions on Graphics (TOG)* 31.4 (2012): 29
- [24] Hongle Li and SeongKi Kim, A Novel Mesh Simplification Method Based on Vertex Removal Using Surface Angle, *International Journal of Engineering Research and Technology*. Volume 12, Number 8 (2019), pp. 1313-1320.
- [25] Guangyou Zhou, Shangda Yuan, and Sumei Luo, Mesh Simplification Algorithm Based on the Quadratic Error Metric and Triangle Collapse, *IEEE Access*, 2020. 10.1109/ACCESS.2020.3034075.
- [26] Rolandos Alexandros Potamias, Stylianos Ploumpis, Stefanos Zafeiriou, Neural Mesh Simplification, *CVPR*, pp. 18583- 18592, 2022.
- [27] Marie-Julie Rakotosaona, Paul Guerrero, Noam Aigerman, Niloy J Mitra, and Maks Ovsjanikov. Learning delaunay surface elements for mesh reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22–31, 2021.
- [28] Thibault Lescoat, Hsueh-Ti Derek Liu, Jean-Marc Thiery, Alec Jacobson, Tamy Boubekeur, and Maks Ovsjanikov. Spectral mesh simplification. *Computer Graphics Forum*, 39(2):315–324, 2020.

Implementation of a Fourth-Order Compact Quasi-Elliptic Substrate Integrated Waveguide Filter in C-Band

Narges kiani^{1,*}, Majid Afsahi², Farzad Tavakkol Hamedani³ and Pejman Rezaei³

Abstract— Substrate Integrated Waveguide (SIW) technology by application of the planar construction process obtains a catchy means for integrating planar and nonplanar circuits. Further, it is hard to realize the negative coupling needed to create a compact quasi-elliptic bandpass filter based on a single-layered SIW structure. The presented work proposes a specific planar and negative coupling configuration that provides two transmission zeroes at 5.03 GHz and 6.26 GHz. This article presents a fourth-order quasi-elliptic filter. The proposed filter is also wide-band. This structure is implemented in SIW knowledge. The SIW filter has a central frequency of 5.5 GHz. The perfect bandwidth is 0.7 GHz. It is realized on a single-layer substrate from Rogers Ro4003. The Thickness of the substrate is considered as 0.508 mm. The measured outcomes of this filter, which show an excellent selectivity, and a low insertion loss of about 1.9 dB, agree suitably with simulation results. The designed filter's innovation provides these features: compact, low cost, wide-band, good selectivity, low insertion loss, and agreement between simulation and fabrication.

Index Terms— Substrate Integrated Waveguide (SIW), Quasi-Elliptic Filter, Band Pass Filter (BPF), Electric Coupling, Magnetic Coupling.

I. INTRODUCTION

The rapid growth of communications services and devices happened lately [1, 2]. SIW technology has attracted a lot of consideration because of its benefits such as: low-cost fabrication, low loss, high-quality factor (Q-factor), high power handling capability, and easy integration. In comparison to multilayered SIW, single-layer SIW structures have more stuff. They include filters [3–18], antennas [19–21], and so on. The quasi-elliptic filters are trendy. Among the benefits of these filters, excellent out-of-band rejection is emphasized. Electric coupling (negative or cross-coupling) and magnetic coupling (positive or direct coupling) is applied in the structure of these types of SIW filters. In general, finite transmission zeroes (TZs) can be observed in the quasi-elliptic filter structures. However, mutual coupling between non-adjacent resonators is required to bring TZs from infinite to finite situations. It has been confirmed that both magnetic and electrical coupling is needed to generate TZs at limited frequency positions.

Furthermore, they obtain excellent selectivity in a cross-coupled filter. Generally, direct coupling is performed by using

a magnetic post-wall window (positive coupling) in single-layered SIW filters. Usually, it is not easy to achieve negative coupling. BPF microstrip filters typically have a lower Q-factor compared to SIW filters. This is the basic weakness of BPF microstrip filters compared to SIW filters [22, 23]. There are electromagnetic excitations at the interface between metal and dielectric. These excitations are called surface plasmon polaritons (SPPs). Various devices based on SPPs have been reported. We can mention plasmonic filters among the tools that are very useful in this field. Unlike microwave filters, they cannot be manufactured, which is plasmonic filters' main weakness [24, 25].

In [3], a dual-mode SIW doublet bandpass filter was investigated using a CSRR. A SIW transversal filter using mixed source and load coupling was realized in [4]. In [5], a wide stopband SIW filter with corner cavities was explained. The multilayer technique was used to realize compact SIW filters in [6]. Analysis of SIW dual-band and broad stop-band filters were elaborated in [7]. A SIW filter based on a parallel-coupled microstrip line (PCML) reported [8]. A design system for SIW-EBG filters was reported in [9]. In [10], a study of SIW and HMSIW BPFs was presented by applying periodic configurations. Implementation of dielectric SIW filter by using two variant substrates proposed [11]. In [12], the performance of two different topologies of SIW BPFs was displayed in the S-band. The wide SIW BPFs were designed using stepped impedance structures [13]. A circular SIW BPF with a broad tuning range was studied in [14]. In [15], a compact SIW BPF was implemented by the compact microstrip resonant cell method. An investigation of the factors of the slow wave phenomenon on two species of miniaturized SIW BPFs was deduced in [16]. The access to H-plane SIW filters was introduced using propagative and evanescent modes in [17].

This paper presents a fourth-order SIW broadband filter. Both magnetic and electric couplings are used in the structure of this filter. The purpose of providing electric coupling is to reach an excellent coefficient and wide coupling stability range. The designed filter's innovation provides these features: compact, low cost, wide-band, good selectivity, low insertion loss, and agreement between simulation and fabrication. Simulations are implemented in CST Microwave Studio software [26].

1- Shariati Technical and Vocational College, Tehran, Iran.

2- Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran.

3- Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran.

Corresponding author: n.kiani@semnan.ac.ir

In section 2, the proposed filter coupling structure is presented, along with a detailed description of its coupling matrix. Section 3 describes the design features of the fourth-order bandpass filter. The simulation results of the designed structure are reported in Section 4. Finally, the fifth section presents the test and measurement results which are in suitable compromise with the simulation outcomes.

II. COUPLING MATRIX

The design of the fourth-order SIW filter is illustrated in Fig. 1.

The coupling diagram of the designed filter is remarkable in Fig. 2.

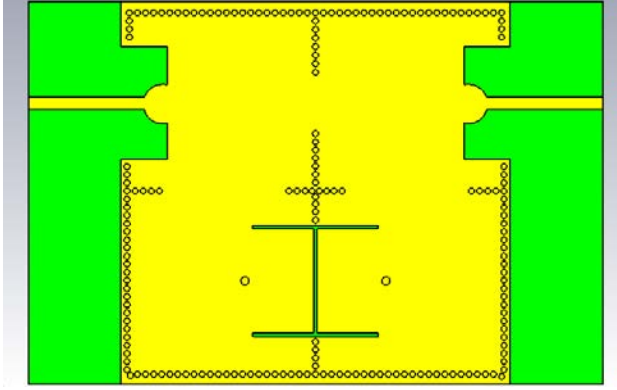


Fig. 1. Configuration of the proposed structure.

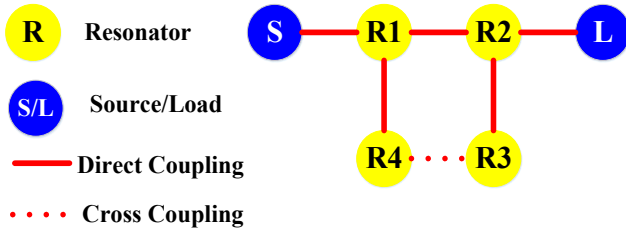


Fig. 2. Coupling diagram.

In general, the coupling matrix is derived from Eq. 1, which is as follows [27, 28]:

$$K_{ij} = K_{ji} = \frac{f_j^2 - f_i^2}{f_j^2 + f_i^2} \quad (1)$$

Applying Eq.1 to the proposed SIW filter structure yields this result. The cases $k > 0$ and $k < 0$, are for magnetic coupling (direct coupling) and electric coupling (cross-coupling), respectively.

$$K_{ij} = \begin{bmatrix} +0.00 & +0.25 & +0.00 & +0.16 \\ +0.25 & +0.00 & +0.17 & +0.00 \\ +0.00 & +0.17 & +0.00 & -0.23 \\ +0.16 & +0.00 & -0.23 & +0.00 \end{bmatrix} \quad (2)$$

The K_{34} and K_{43} coupling coefficients are the same, and both are negative. The negative sign in the coupling coefficient displays TZs before and after the passband. These TZs are located at 5.03 and 6.256 GHz. Their position appears in the form of design results in future sections. The center frequency (f_0) of the SIW filter is located at 5.5686 GHz. The designed

SIW filter bandwidth (BW) is 0.7 GHz. This is how the generalized coupling matrix [29] is obtained, which is evident in Eq.3.

$$M_{ij} = K_{ij} \times \frac{f_0}{BW} = K_{ij} \times \frac{5.5}{0.7} = K_{ij} \times 7.9 \quad (3)$$

The generalized coupling matrix is given using Eq.3, matrix two, and the substitution of the values obtained from the filter frequency characteristic.

$$M_{ij} = \begin{bmatrix} +0.0 & +1.1 & +0.0 & +0.0 & +0.0 & +0.0 \\ +1.1 & +0.0 & +1.9 & +0.0 & +1.3 & +0.0 \\ +0.0 & +1.9 & +0.0 & +1.4 & +0.0 & +1.1 \\ +0.0 & +0.0 & +1.4 & +0.0 & -1.8 & +0.0 \\ +0.0 & +1.3 & +0.0 & -1.8 & +0.0 & +0.0 \\ +0.0 & +0.0 & +1.1 & +0.0 & +0.0 & +0.0 \end{bmatrix} \quad (4)$$

In the next step, the dimensions of the square cavity are approximated. To achieve this goal, the iris dimensions need to be modified. Such design adjustments are useful and appropriate to optimize coupling coefficients. In general, it can be said that this general rule applies to full-wave simulations of coupled cavities. The final step in designing is to adjust the dimensions of the cavities. The goal of implementing these proceedings is to repair the resonance movement created by the iris structure [27, 28]. The quality factor (Q_{ext}) is determined by Eq.5.

$$Q_{external} = \frac{f_0}{\Delta f_{\mp 90}} = 64.8 \quad (5)$$

In this Equation, the parameter $\Delta f_{\mp 90}$ is in the sense of bandwidth where the phase variation is ± 90 degrees [29].

III. FILTER DESIGN

The configuration of the proposed filter structure in CST software with the required design dimensions is illustrated in Fig. 3.

The SIW filter dimension values are listed in Table I.

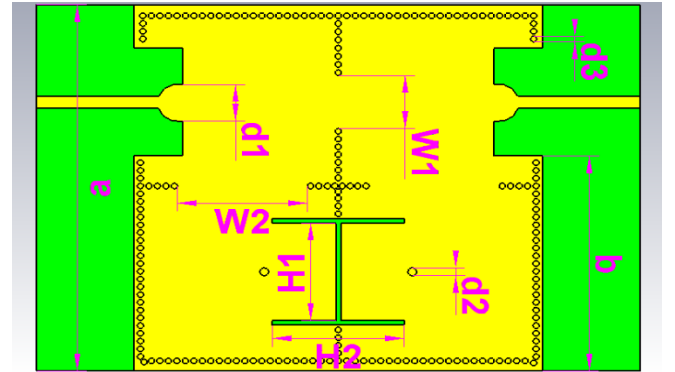




Fig. 3. Dimensions and configuration of the designed filter.

TABLE I
Designed Dimensions of the SIW Filter.

Designed Dimensions (in mm)				
a	b	W_1	W_2	
33	20	5	11	
d_1	d_2	d_3	H_1	H_2
4	0.7	0.5	9	11

The benefits of using fourth-order filters to achieve isolation are very high. The entire structure is implemented in a standard PCB process. The substrate is a Rogers RO4003. The substrate thickness is selected as 0.508 mm. The dielectric constant (ϵ_r) of this substrate is assumed to be 3.38. The value of loss tangent ($\tan \delta$) is also considered 0.0027. All sidewall vias should be drilled at a diameter of 0.5 mm. The central vias should be drilled at a diameter of 0.7 mm. The segmented filter structure is shown in Fig. 4.

PEC	
Ro4003	

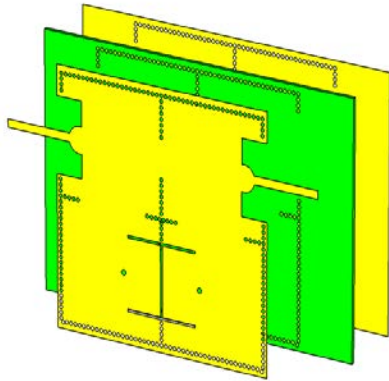


Fig. 4. Configuration of the proposed structure.

IV. RESULTS OF THE DESIGN

The S-parameter parameter frequency response (S_{11}) of the designed SIW fourth-order structure is evident in Fig. 5. According to Fig.5, the central frequency of the designed SIW fourth-order filter is set at 5.5 GHz. The bandwidth of the designed SIW filter is approximately 0.7 GHz. The filter poles are at 5.226, 5.5255, and 5.6726 GHz, respectively. Return loss is reported at about 30.2 dB. The insertion loss is about 1.9 dB.

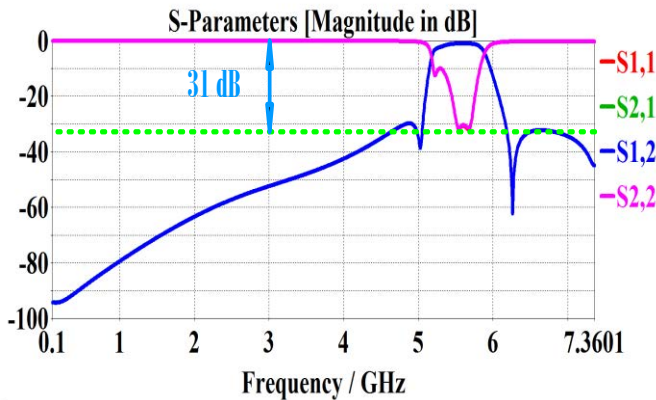


Fig. 5. Wideband frequency response.

Fig. 5, which shows the wideband frequency response, indicates that the 31-dB rejected level until 7.4 GHz. The distribution of Forth order SIW filter E- field lines is illustrated in Fig. 6.

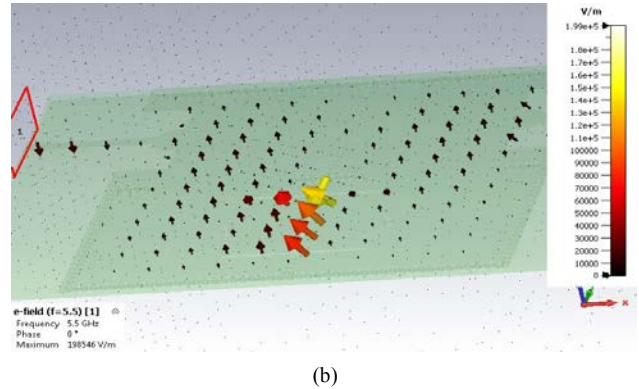
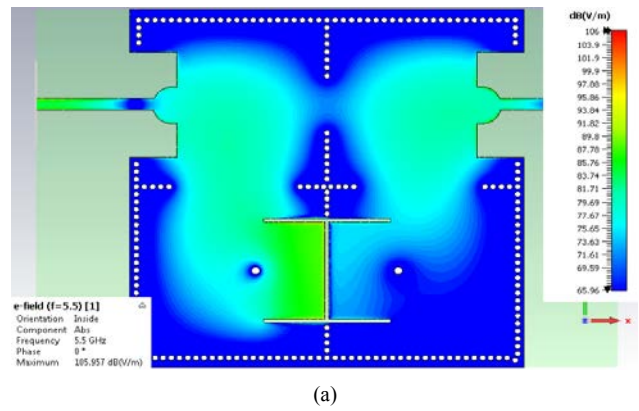


Fig. 6. E-field distribution in central frequency ($f_0 = 5.5 \text{ GHz}$).

The H- field distribution diagram of the proposed filter structure is presented in Fig. 7.

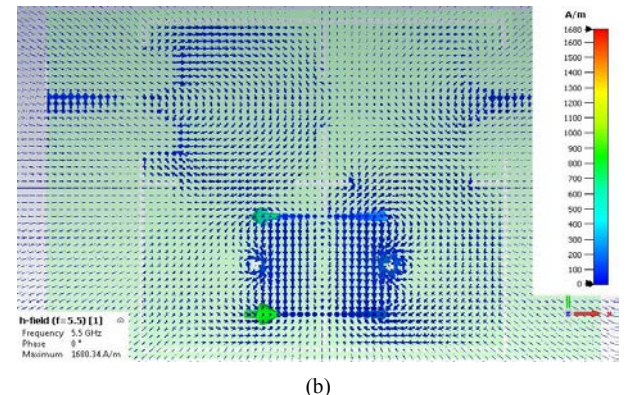
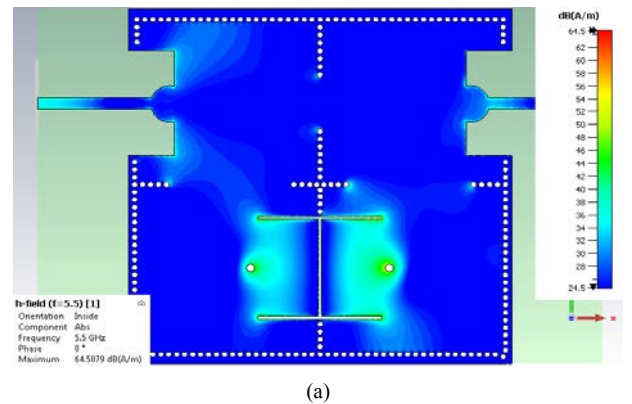


Fig. 7. H-field distribution in central frequency ($f_0 = 5.5 \text{ GHz}$).

In Fig. 8, the distribution diagram of the surface current of the quasi-elliptic filter is shown.

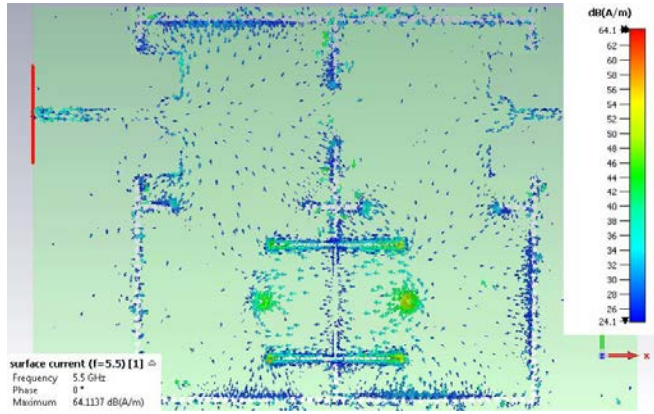


Fig.8. Distribution diagram of the surface current in central frequency ($f_0 = 5.5 \text{ GHz}$).

The VSWR curve is shown in Fig. 9 and provides excellent results.

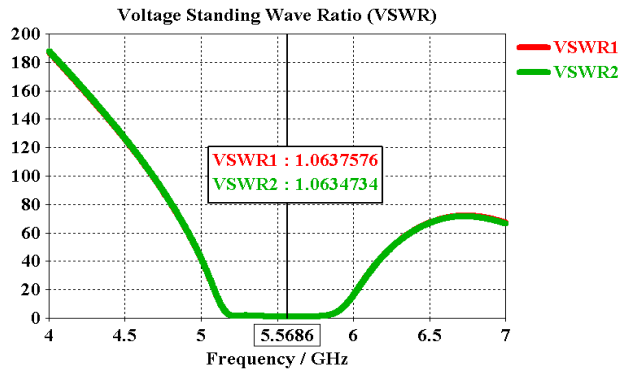


Fig. 9. VSWR of the proposed SIW filter.

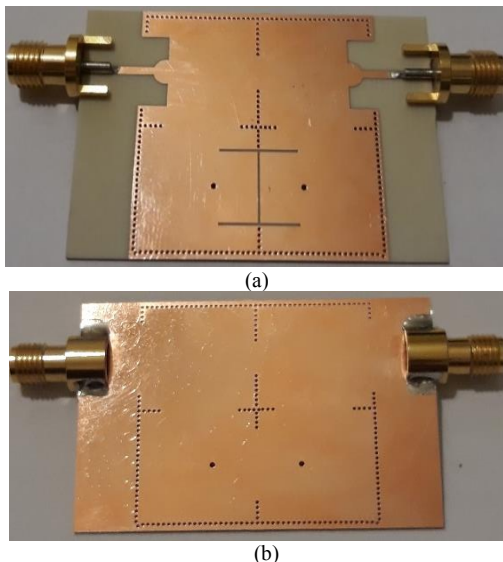


Fig. 10. The form of the fabricated fourth-order SIW filter, (a) Top view, (b) Bottom view.

V. FABRICATION RESULTS

The prototype of the fabricated SIW fourth-order filter is displayed in Fig. 10.

The results of the frequency response curve measurement of the fourth-order SIW filter structure by the Network Analyzer are shown in Fig. 11.

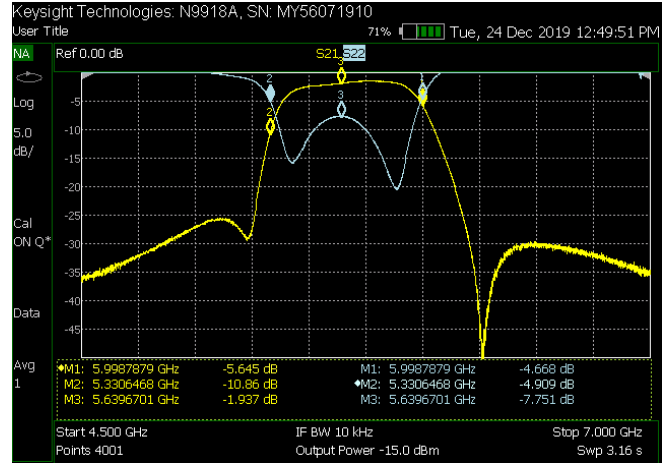


Fig. 11. The measurement results of the fabricated SIW filter.

As seen in Fig. 11, there is a cooperative agreement between the measurement results and the simulation, confirming of each.

A comparison of the performance of the designed filter structure and similar presented structures in other articles are shown in Table II.

TABLE II

Comparison of the Designed Filter with Other Articles in This Field.

References	BW_{3dB} (GHz)	f_0 (GHz)	Filter Order	IL (dB)	Number of TZs	FBW	Rejection	Size $(\lambda_g)^2$
[3]	0.2	4.99	3	1.5	2	5	Null	1.25
[4]	0.3	3.6	3	1.35	2	9.7	$20/1.99f_0$	0.63
[5]	0.55	20	4	2.78	1	2.75	$50/1.49$	2.53
[6]	0.6	13.2	2	1.5	1	4.55	$20/2.3f_0$	0.47
[7]	0.7/0.5	10	3	2.91	1	1.51	$25/2.1f_0$	1.86
[8]	-	10	2	0.8	2	1	-	3
Presented work	0.7	5.5	4	1.9	2	12.73	$31/1.4f_0$	0.65

VI. CONCLUSION

In this proposed paper, a compact fourth-order quasi-elliptic filter has been studied. Various electric and magnetic couplings have been used to implement the structure of this filter. Studies of filter coupling configuration have been fully reported. SIW technology implements this filter. The advantages of using SIW in this structure were low loss, low cost, and easy construction. The quasi-elliptic filter has been designed in the C band. The central frequency of the designed SIW fourth-order filter was set at 5.5 GHz. The bandwidth of the designed SIW filter was approximately 0.7 GHz. The filter poles were at 5.226, 5.5255, and 5.6726 GHz, respectively. Return loss was reported at

about 30.2 dB. The insertion loss was about 1.9 dB. There was a good agreement between the fabrication results and the simulation.

VII. ACKNOWLEDGMENT

The authors would like to thank the support from Semnan University. The authors thank the journal editor and reviewers for their valuable comments.

VIII. REFERENCES

- [1] Kiani N, Afsahi M. Design and fabrication of a compact SIW diplexer in C-band. *Iranian J. Electr. Electron. Eng.* 2019;15 (2):189–194.
- [2] Kiani N, Tavakkol Hamedani F, Rezaei P. Implementation of a reconfigurable miniaturized graphene-based SIW antenna for THz applications. *Micr. Na.* 2022.
- [3] Wu LS, Zhou XL, Wei QF, Yin W. An extended doublet substrate integrated waveguide (SIW) bandpass filter with a complementary split ring resonator (CSRR). *IEEE Microw. Wireless Compon. Lett.* 2009;19 (12): 777–779.
- [4] Shen W, Yin WY, Sun XW. Compact substrate integrated waveguide transversal filter with microstrip dual-mode resonator. *J. Electromagn. Waves. Appl.* 2010;24 (14): 1887–1896.
- [5] Zhu F, Hong W, Chen JX, Wu K. Wide stopband substrate integrated waveguide filter using corner cavities. *Electron. Lett.* 2013;49 (1): 50–52.
- [6] Jia D, Feng Q, Xiang Q, Wu K. Multilayer substrate integrated waveguide (SIW) filters with higher-order mode suppression. *IEEE Microw. Wireless Compon. Lett.* 2016;26 (9): 678–680.
- [7] Zhou K, Zhou CX, Wu W. Resonance characteristics of the substrate-integrated rectangular cavity and their applications to dual-band and wide-stopband bandpass filter design. *IEEE Trans. Microw. Theory Techn.* 2017;65 (5): 1511–1524.
- [8] Kiani S, Rezaei P, Karami M, R.A. Sadeghzadeh. Substrate integrated waveguide quasi-elliptic bandpass filter with parallel coupled microstrip resonator. *Electron. Lett.* 2018; 54: 667–668.
- [9] Simsek S, Rezaeieh SA. A design method for substrate integrated waveguide electromagnetic bandgap (SIW-EBG) filters. *AEU Int. J. Electron. Commun.* 2013;67:981–983.
- [10] Moitra S, Bhowmik PS. Modeling and analysis of Substrate Integrated Waveguide (SIW) and half-mode SIW (HMSIW) band-pass filter using reactive longitudinal periodic structures. *AEU Int. J. Electron. Commun.* 2016;70:1593–1600.
- [11] Aghayari H, Nourinia J, Ghobadi C, Mohammadi B. Realization of dielectric-loaded waveguide filter with substrate integrated waveguide technique based on the incorporation of two substrates with different relative permittivity. *AEU Int. J. Electron. Commun.*2018;86:17–24.
- [12] Martínez J, Coves Á, Bronchalo E, San Blas ÁA, Bozzi M. Band-pass filters based on periodic structures in SIW technology. *AEU Int. J. Electron. Commun.* 2019;112:152942.
- [13] Máximo-Gutiérrez C, Hinojosa J, Alvarez-Melcon A. Design of wide band-pass substrate integrated waveguide (SIW) filters based on stepped impedances. *AEU Int. J. Electron. Commun.*2019;100:1–8.
- [14] Wang X, Zhang D, Liu Q, Deng H, Lv D. Tunable bandpass filter with a wide tuning range of center frequency and bandwidth based on circular SIW. *AEU Int. J. Electron. Commun.*2020;114:153002.
- [15] Huang L, Wu W, Zhang X, Lu H, Zhou Y, Yuan N. A novel compact and high-performance bandpass filter based on SIW and CMRC techniques. *AEU Int. J. Electron. Commun.*2017;82:420–425.
- [16] Parameswaran A, Raghavan APS. Miniaturizing SIW filters with slow-wave technique. *AEU Int. J. Electron. Commun.*2018;84:360–365.
- [17] Aghayari H, Nourinia J, Ghobadi C. Incorporated substrate integrated waveguide filters in propagative and evanescent mode: realization and comparison. *AEU Int. J. Electron. Commun.*2018;91:150–159.
- [18] Meiguni JS, Ghobadi Rad A. WLAN substrate integrated waveguide filter with novel negative coupling structure. *J. Model. Sim. Electr. Electron. Eng.* 2015;1 (2):15–18.
- [19] Amn-e-Elahi A, Rezaei P. SIW corporate-feed network for circular polarization slot array antenna. *Wireless Pers. Commun.* 2020;111: 2129–2136.
- [20] Meiguni JS, Pommerenke A. Theory and experiment of UWB archimedean conformal spiral antennas. *IEEE Trans. Antennas Propag.* 2019;67 (10): 6371–6377.
- [21] Meiguni JS, Kamyab M, Hosseinbeig A. Theory and experiment of spherical aperture-coupled antennas. *IEEE Trans. Antennas Propag.* 2013;61 (5): 2397–2403.
- [22] Wang B, Cappelli MA. A tunable microwave plasma photonic crystal filter. *Appl. Phys. Lett.* 2015; 107(17): 171107.
- [23] Pal B, Mandal MK, Dwari S. Varactor tuned dual-band bandpass filter with independently tunable band positions. *IEEE Microw. Wireless Compon. Lett.* 2019; 29(4): 255–257.
- [24] Zhu H, Abbosh A. Compact tunable bandpass filter with a wide tuning range of center frequency and bandwidth using coupled lines and short-ended stubs. *IET Microw. Antennas Propag.* 2016; 10(8): 863–870.
- [25] Xiao JK, Su XB, Wang HX, Ma JG. Compact microstrip balanced bandpass filter with adjustable transmission zeros. *Electron. Lett.* 2019; 55(4): 212–214.
- [26] CST GmbH, Germany, CST STUDIO SUITE Ver. 2015 – User’s Manual, Dec. 2015. www.cst.com.
- [27] Cameron RJ. General coupling matrix synthesis methods for Chebyshev filtering functions. *IEEE Trans. Microw. Theory Tech.* 1999; 47 (4): 433–442.
- [28] Cameron RJ. Advanced coupling matrix synthesis techniques for microwave filters. *IEEE Trans. Microw. Theory Tech.* 2003; 51 (1): 1–10.
- [29] Hong JS, Lancaster MJ. *Microstrip filter for RF/microwave applications.* New York: Wiley, 2001: 235–272.

Design Processes Linear Permanent Magnet Electrical Vernier Machines For Future Research Directions: A Review

Hamid Yaghobi^{1*} and Mohammad Reza Mohammadzadeh Moghaddam²

Abstract— In this study, the technology of Linear Permanent Magnet Vernier Electric Machines (LPMEVMs) is reviewed. Since the introduction of the first LPMVM designs, many designs have been proposed, and many researchers have come up with different ideas. The new LPMEVM technology has attracted the attention of many researchers, so with the efforts of researchers, it can be used in industrial applications of this machine and will be welcomed by craftsmen shortly. In the following, this article examines the work done in the field of LPMVM. In this study, the articles presented in this field are generally reviewed, the new technologies presented, and the structures have been analyzed by researchers and mentioned in the articles. Detailed more precisely on the structures presented and designs made for different parts of the machine, including the mover and translator, which include fixed and moving parts in the LPMV machine, the shape of the grooves and teeth, and the placement of permanent magnets and coils, will be discussed in this paper. This research discusses the designs presented in this LPMV machine and paves the future research path.

Index Terms— Linear Vernier machine, Permanent magnet, Winding, Core.

I. INTRODUCTION

Linear permanent magnet Vernier machines are among those that have attracted special attention in linear motion applications due to their outstanding features. In describing the features, papers in this field state that the LPMVM machine produces high torque at low speeds and that its performance base is based on magnetic gearbox technology. Due to the problems with mechanical gearboxes, such as parts failure and maintenance of technology, Vernier machines have been able to solve this need without friction with each other and by using magnetic relations. Research and development in magnetic gearboxes in recent years show the importance of this issue in Fig. 1 (the x-axis is the years, and the y-axis is the number of publications). LPMVM can involve linear, conventional, multi-harmonic, and modular LPMV primary permanent magnet machines [1], [2], [3]. That is the leading technology in the electrical linear machine industry, mainly functioning based on the magnetic gearbox. Machines are generally composed of

both stationary and moving parts. The components of conventional machines include: (i) **iron core**, (ii) **north permanent magnet**, (iii) **south permanent magnet**, and **coil conductors** (iv). For example, a linear permanent magnet Vernier machine (FP-LPMVM) in a fractional pole-pair linear force is transferred by a coiled actuator. The fixed parts include the said core and magnets [4].

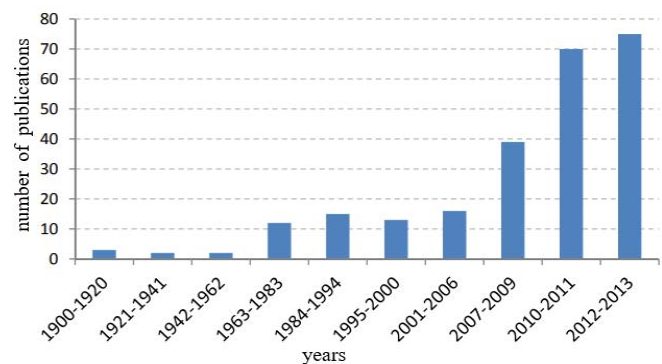


Fig. 1 - Number of publication magnetic gears (1900-2013) [2].

The histogram in Fig. 2 depicts studies on linear Vernier machines and research done by researchers and published in conferences and journals (IEEE) (the x-axis is the years, and the y-axis is the number of publications). This histogram shows researchers' growth, research, and development of linear Vernier machines in recent years.

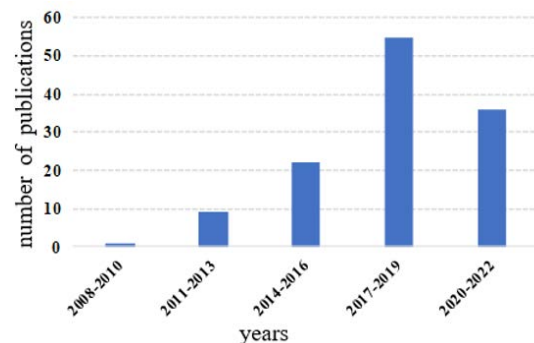


Fig. 2. Number of publication LVMs (2008-2022).

1- Hamid Yaghobi is Faculty of Electrical and Computer Engineering, Semnan University, Semnan, Iran.

2- Mohammad Reza Mohammadzadeh Moghaddam is Faculty of Electrical and Computer Engineering, Semnan University, Semnan, Iran.
Corresponding author: yaghobi@semnan.ac.ir

Rotary Vernier machines were first introduced more than a century ago [5], and numerous designs have since been presented. Along with this advancement, linear Vernier machines have begun to gain popularity among industrialists and researchers. As a result, the following section will go over the presented plans and the growth and developments that have occurred to express the current state of this technology, as well as review the articles on linear Vernier machines to cause the future development of this technology.

II. DESIGN PATTERN LPMVM

Linear Vernier machines can have a variety of mover and stator structures that aim to create linear motion at the machine's output and to create linear motion. Fig. 2 shows the stator of the linear machine.

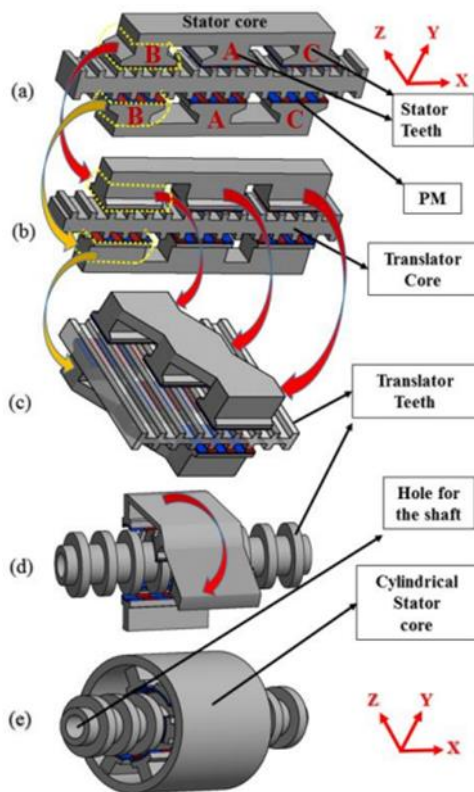


Fig.3. Development of VHM tubular technology from the flattish VHM [6].

The linear flat machine (Fig. 3a) served as the foundation for the most recent Vernier machine design technologies. Changing the tooth deformation in the stator by 90 degrees (Fig. 3b) led to the emergence of a new design (Fig. 3c). Figure 3d shows another advancement in the design of such machines.

The next step was to change the translator core. The cylindrical stator structure led to the linear motion of the mover (Fig. 3e), which was based on Fig. 3a of the linear flat machine. LPMs can have a variety of stators, movers, and toothed shape designs, but they all use linear motion [6], [7].

Magnetic materials in electric machines are classified into two types: (a) soft magnetic materials such as stator core laminations and (b) hard magnetic materials such as permanent magnets. The properties of these materials have a great impact

on the performance of the linear machine. Many parameters can effectively produce a magnetic field in a linear Vernier machine. However, the most important ones are the material of permanent magnet magnets, the teeth' shapes, the coils' design, the amount of copper, and the type of placement.

It directly affects the core's magnetization and the Vernier machine's saturation points [6], [8]. Studies show that increasing the number of phases from three to six improves fault tolerance, reliability, and the sinusoidal voltage waveform of the machine [9].

III. RECENT ADVANCES IN LINEAR VERNIER MACHINES

In recent years, a lot of research has been on Vernier machines, perhaps due to the PM feature and properties such as high energy density, which can create high torque at low speeds, and on the other hand, provide new topologies and efficiency. In addition, new skills and abilities have led to the expression of new concepts.

A. Application of permanent magnet materials

The management of the use of PMs in the design has played a very important role in the final cost, and according to the different types of arrays and the variety of materials and shapes available, the best PM selections according to the design of machines.

In 1977, flux changes in air distances, the power factor of the Vernier machine, and its reliability were studied [10]. The Vernier machine has a higher torque than conventional flux modulation machines based on producing this amount of high torque due to the presence of pairs of magnetic poles in the machine design [11].

PM thickness significantly affects flux density and permeability and affects the machine's performance [12]. In some designs, the linear machine uses non-magnetic space between the magnet and the stator, improving the machine's efficiency and reducing the amount of leakage flux [13]. The placement of the PMs in the stator is critical to the operation of the machine. Figure 4 presents a new design that can be placed in (i) **Linear shape PM**, (ii) **V-shaped PM**, and (iii) **Linear V-shaped PM**, designed at a certain distance and angle in addition to a combination of poles with a particular magnetic orientation.

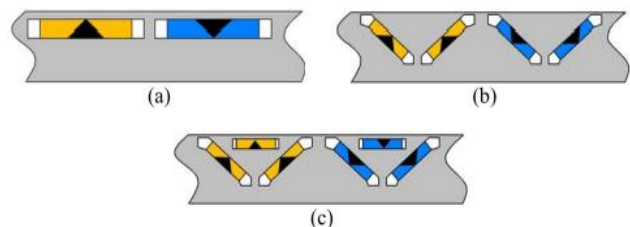


Fig.4. Three PM design types. (a) Linear shape. (b) V-shape. (c) Linear V-shape [16].

Among the new PM topologies, researchers' attention to the proposed Halbach PM arrays: where all kinds of placement in the actuator, efficiency, leakage rate, amount of flux in the air distance, ripple, and efficiency can have a significant impact on the machine's performance.

Fig. 5. shows the Halbach arrays for high-force density

Vernier machines, and the application of PMs directly impacts the final price of the machine [14], [15], [16], and [17].

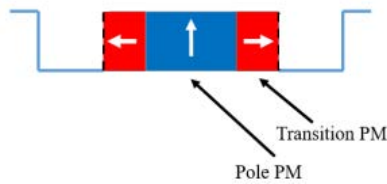


Fig. 5. Halbach PM array [18].

Fig. 6. Displays the other three PM arrays that researchers are most familiar with, in addition to the Halbach arrays. One of the factors in reducing the cost of linear machines is using fewer PMs in machine design [18].

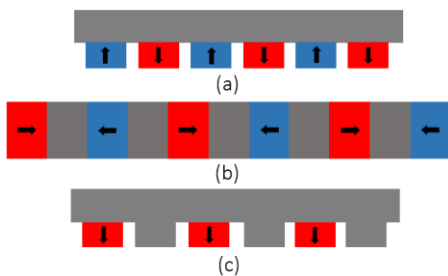


Fig. 6. Configurations (a) surface-mounted [16], (b) flux concentrator [13], (c) consequent pole [18].

The following results were obtained by comparing three designs (a) **surface-mounted** (SPM-LVM), (b) **Halbach** (SPM-LVM with Halbach), (c) **V-shape** (IPM-LVM) in LPMVMs with the same and no-load flux distribution in Fig 7. characteristics condition. In Fig. 8. the results of no-load LVPM flux linkage analysis and EMF analysis.

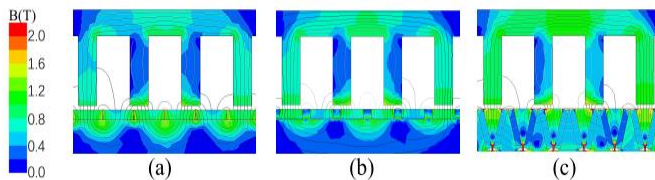


Fig. 7. No-load LVPM flux distribution three design [16].

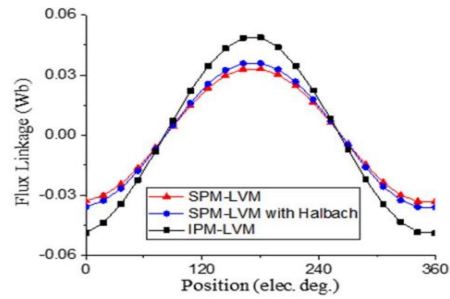
B. low and High-temperature superconductor (LTS &HTS)

One of the effective ways to get more current density from stator winding is to use high-temperature, and low-temperature superconductors (LTS and HTS) designs that can significantly improve the performance and efficiency of the machine, thus reducing the effect of the coil wire's end windings. Fig. 9 shows the proposed DCLVM-LTS machine design with a 40% improvement in current density using LTS [19].

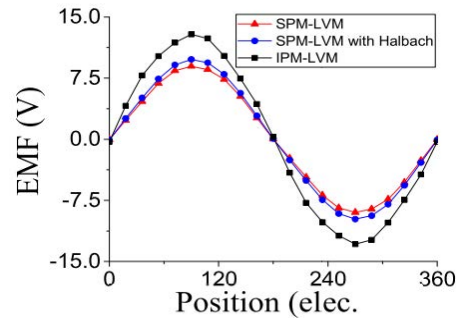
Chaojie Shi and all colleagues in the line generator use HTS technology to receive more power and increase the power factor from 0.38 to 0.57 for DS-HTS-LVG, resulting in about 36% higher power density [20].

A high-temperature superconductor (HTS) can reduce the flux leakage in the LPMV machine and improve efficiency. The higher power in a linear machine and the equivalent power

received from a larger machine can be obtained from a smaller machine [21].



(a)



(b)

Fig. 8. Comparison three designs (a) Comparison no-load LVPM flux linkage (b) Comparison LVPM of EMF [16].

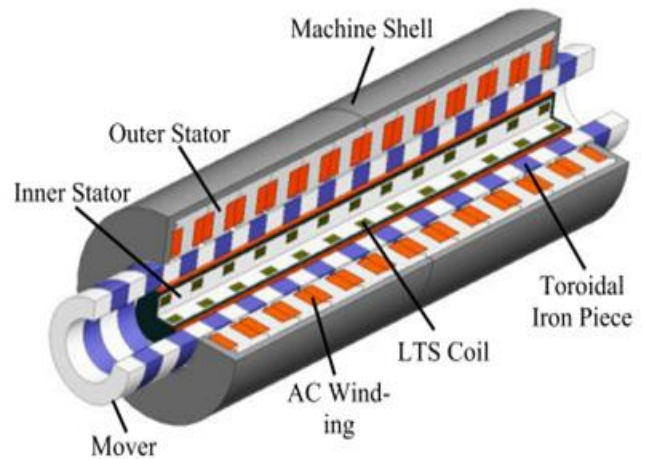


Fig. 9. View from DCLVM-LTS technology [19].

One of the applications of LPMVM with a lighter weight is the sea wave energy conversion (WEC) into electricity. A machine with less weight in the same situation would have a higher speed and, as a result, produce more electricity in generator mode and achieve more efficiency in generating energy from sea waves [22].

Considering HTS in a similar design, the LSPMV with HTS has more capabilities. Considering HTS in a similar design, the LSPMV with HTS has more capabilities. Compared to LSPMV with the LSPMV-HTS, more output values can be received from HTS in Fig. 10. [23].

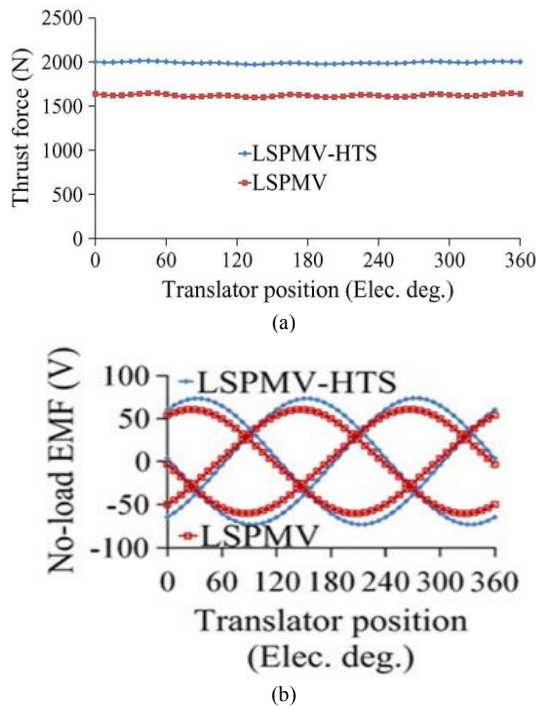


Fig. 10. In (a) the thrust force of LSPMV and LSPMV-HTS (b) and the no-load EMF of LSPMV and LSPMV-HTS [23].

Most operating costs can be covered by the cooling system and HTS, which improve thrust performance while also raising the machine's overall cost. Consideration is given to the trade-off between cost and thrust performance [24]. The use of direct drive WEC (DD-WEC) line generators with HTS technology can extract higher power and, similarly, reduce winding losses [25], [26].

The DD-WEC extraction with HTS-DSTVM topology, located on the coil in the outer layer of the stator, increases the thrust and reduces the leakage flux and end-flux lines [27]. Fig. 11. shows that a DC field excitation coil can increase power multiplication, and an HTS high-temperature superconductor can improve stability [28].

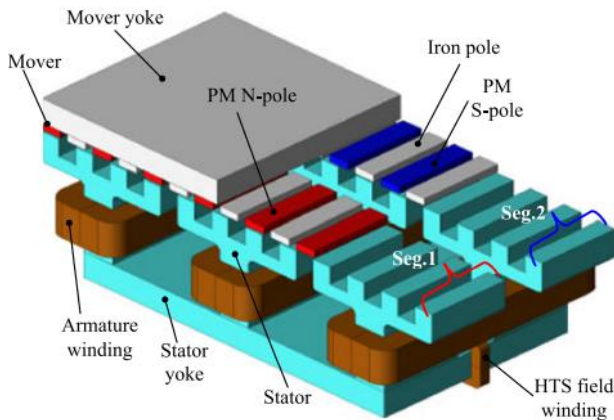


Fig. 11. A new LVPMM using HTS by DC field excitation [28].

In the double stator structure of the HTS-DSSLVM type, the application of HTS can improve the power factor by reducing the inductance and, thus, reducing the leakage flux in the

machine [29], resulting in less electrical loss and higher torque stability [30]. In LPMV, the combination of permanent-magnet arrays can improve flux density and reduce leakage flux [31]. Here, the structures frequently introduced in different studies have been compiled to provide a better analysis. Also, the used parts were discussed in detail to identify the different structures. The collected data is summarized in Table I.

TABLE I
an Overview of the Issues Under Analysis

An Important Structure Of Lvpmm	Highly Repetitive Content	References	Number Of References
PM structure	<ul style="list-style-type: none"> • Dimensions • Arrangement • Shape of placement • Number • PM distribution 	[1], [11], [12], [13], [16], [18], [37], [38], [50].	9
Core & Tooth structure	<ul style="list-style-type: none"> • Core with Flux barrier • slot / poles • Number of teeth • Tooth dimensions 	[4], [8], [31], [36], [47], [48].	6
Machine structure	<ul style="list-style-type: none"> • Several phases • Structure double • Combined structures • Tubular structures 	[3], [6], [7], [14], [15], [19], [22], [27], [42], [45], [49].	11
winding structure	<ul style="list-style-type: none"> • Coil shape • How to place the coil • Connections between coils 	[9], [46], [51].	3
Conductor structure	<ul style="list-style-type: none"> • HTS • Superconducting 	[20], [21], [23], [24], [25], [26], [28], [29], [30].	9
Controller structure	<ul style="list-style-type: none"> • Control methods 	[17], [32], [33], [35], [39], [40], [41], [43], [44].	9

IV. WINDING, STATOR, AND MOVER DESIGN

From the design point of view, the three main parts of Vernier machines are discussed: **1) patterns of winding** and **2) mover and stator design**. Before that, the given general information in Table 1 is presented to get acquainted with the designers' ideas. It has hoped that the information collected will be beneficial and open doors for the advancement of future researchers.

A. Patterns of winding

The winding design is influenced by the design of other parts of the LVPMM machine. The design of the coils and winding has been subjected to the design of the core of the machine. It can be done according to the design, and the behavior of the winding must be done carefully enough because it has a direct relationship with the performance and behavior of the machine.

According to the basic principles of the voltage of the input terminals of the PM machine, papers that have been in this field [23] define the 3-phase voltage relationship:

$$\begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} = \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} R + \frac{d}{dt} \left\{ \begin{bmatrix} L_a & M_{ab} & M_{ac} \\ M_{ba} & L_b & M_{bc} \\ M_{ca} & M_{cb} & L_c \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} \varphi_{PMa} \\ \varphi_{PMb} \\ \varphi_{PMc} \end{bmatrix} \right\} \quad (1)$$

Where i_c , i_b and i_a , are the currents, resistance the winding is the R , and M_{ab} , M_{ac} , M_{ba} , M_{bc} , M_{ca} and M_{cb} are mutual inductances, L_c , L_b and L_a are self-inductances, φ_{PMc} , φ_{PMb} and φ_{PMa} are PM flux linkage.

The most important relationship is given (2) in a Vernier machine

$$P_{pm} = N_s \pm P \quad (2)$$

In this relationship, P_{pm} the PM pairs are the mover and mean the number of poles of the winding and the number of the stator flux teeth. Power factor equation (3) without considering the resistance [27].

$$PF = \frac{1}{\sqrt{1 + \left(\frac{L_s I}{\varphi_m}\right)^2}} \quad (3)$$

The windings' design is directly related to the machine's output parameters. Fig. 12. shows three nodes of the machine design for a better understanding [8].

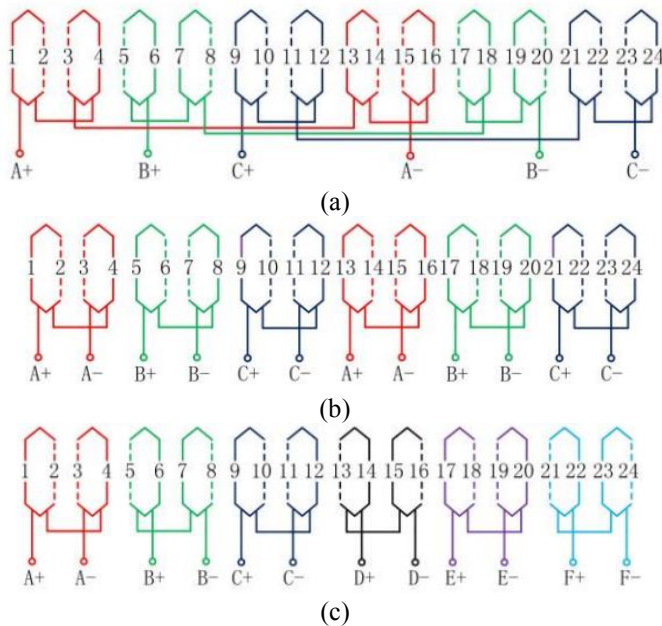


Fig. 12. Three structures winding. (a) Machine three-phase (b) Machine of dual three-phase (c) Machine six phase [8].

Vernier machines have direct-drive systems, including the power supply, switches, and logic arrangements that improve engine control [32]. Vernier machine is a direct drive; if it is

conventionally controlled, it is controlled I_c by a three-phase and a sinusoidal waveform [33].

$$I_a = I_{rms} \sin(\omega t) \quad (4)$$

$$I_b = I_{rms} \sin\left(\omega t - \frac{2\pi}{3}\right) \quad (5)$$

$$I_c = I_{rms} \sin\left(\omega t + \frac{2\pi}{3}\right) \quad (6)$$

Since linear machines have a beginning and an end, several studies have been conducted on their initial and end effects (Fig. 13). The system's performance was affected by these machines because the flux rate in LPMVM's end effect is different [34].

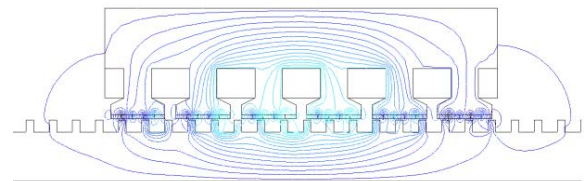


Fig. 13. End effect flux-line of LVPM [34].

B. Mover and stator design in LPMVM

According to the research done in the design of the LPMVM machine, the operation conditions are a very important point and have been one of the important points in the design of the mover and stator. The design of the mover and stator depends on the conditions in which the LPMVM operates so that it can have the best performance.

There is a lot of research on the structure of linear Vernier machines, which can be referred to as double stator and double mover structures in Fig. 14. [35].

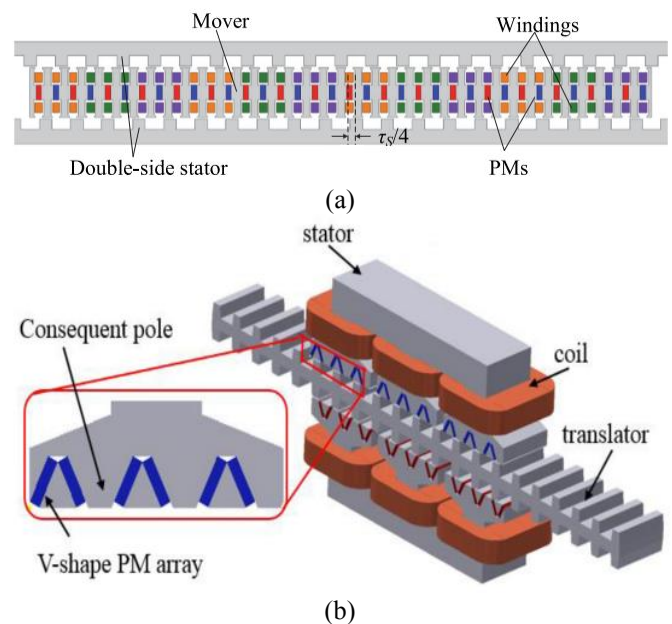


Fig. 14. (a) Double mover design [35], (b) Double stator design [37].

Of course, researchers have proposed many designs, such as cylindrical linear Vernier machines or the use of different PM arrays in the design.

Linear Vernier machines can be used to (i) **produce compressive**, (ii) **ascending or descending**, and (iii) **tensile forces** [34].

Changing the three phases in a linear machine can be asymmetrically reduced using a different coil (copper) and core (iron) structural design. The design of a three-phase linear machine (3MS) is based on three separate parts with non-magnetic spacers, such as aluminum. The places of the phases are ABC, CAB, and BCA, respectively; each coil with the same phase is in the other part of the series machine (Fig. 15), leading to asymmetric reduction, reduced ripple, and better control. Various designs involving putting magnets in the primary and secondary parts can help alter the thrust density of the machine [36].

The diameter of the magnet, type of magnet, and placement of magnets can considerably influence the function of the electric machine. Moreover, they affect another critical parameter, the air distance, which changes the air density flux, affecting the force density and, thus, the power factor in the machine [37].

Fig. 1. presents a new design that has been able to provide a special arrangement for the coils and a special arrangement for placing the coils in the DFM-CP-LPM and with two flux barriers made of aluminum, the machine is divided into three parts, including module 1, module 2, module 3, able to have an almost symmetrical magnetic circuit.

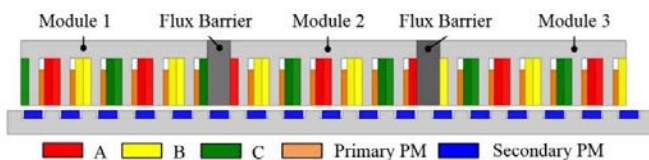


Fig. 15. Configuration of the DFM-CP-LPM [36].

V. FAULT TOLERANCE AND RELIABILITY

Reliability is a recognized feature of linear Vernier machines in industrial applications. One way researchers have proposed is to use more than three phases. In the event of an error in one phase of a multi-phase machine and being in a situation where the machine must continue to operate, the machine can be forced to operate, which has increased the reliability of multi-phase machines [8], [9].

Many researchers have researched linear Vernier machine drives. Vernier machines are direct drives. These drives have been gradually inspired by the drives of conventional permanent magnet electric machines. Sensors can provide accurate parameter data to ensure better machine control to get the phase's location, speed, voltage, and current and measure the force [38]. With the advantages of linear machines in electric trains, the need for location, and the increased cost of implementing position sensors, the article [39] introduced a drive without a position sensor for this project.

Control with two inverters is a design including two voltage power supplies. The first is a DC power source, and the second is a capacitor responsible for supplying reactive power and correcting reactive power, suitable for heavy-load conditions [40]. The controller with a floating capacitor in OEW-LVPM can operate under open fault conditions in open circuit conditions [41]. In linear Vernier generator strategic applications, such as wave energy due to voltage and current

tolerances, the output power from the generator terminals is rectified and converted to DC (Fig. 16). The DC is converted to AC by the inverter to output the voltage, and current waveform with the ripple fluctuations lower (the sea wave is not always uniform) power factor and the capacitor improve it [42]. An optimally designed controller can significantly reduce the amount of THD and extract maximum power from the system [43]. The DTFC controller in the LVPM machine is error-resistant [44].

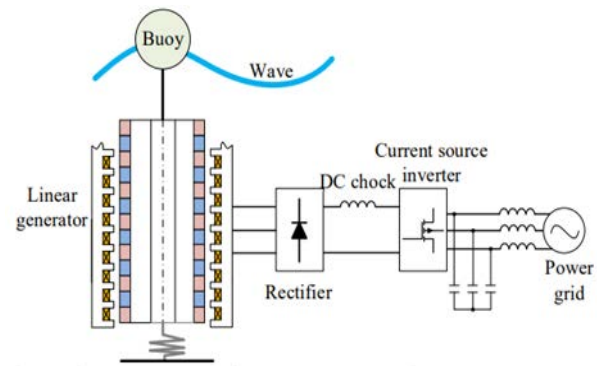


Fig. 16. LPMVG technology for wave power [42].

VI. APPLICATION OF LINEAR VERNIER MACHINES

The linear Vernier machine has been designed based on a magnetic gearbox. This machine has unique features, making it one of the choices in the industry for linear movement. With recent advances, the linear Vernier machine is a good choice in linear applications.

Vernier machines are low-speed and have a high driving force density. Depending on the structure of the Vernier machine, it can have multiple industrial applications. The structure of LPPMV machines provides more thrust and less toothed force than LVH machines, making them more suitable for low-speed applications [45].

The high torque originates from various features of this machine, including the ability to stretch the linear movement of objects, lift heavy loads, generate a compressive force, etc.

- **Wave energy:** Vernier linear machine is suitable for converting the energy of waves in the sea or ocean into electrical energy, ~ 0.5 m/s, and >1 MW;
- **Free-piston generator:** The application of the machine is linear, $10\sim 20$ m/s and ~ 10 kW.
- **Railroad transportation:** It has been used in the traction system of rail transport and inner-city trains.

The use of the linear device largely depends on the device's design. The number of slots/poles affects the performance and application of the machine [46], [47].

The amount of leakage flux in PMLV differs in various conditions: the highest amount of leakage flux is in the machine teeth, and its amount varies according to the machine's location [48].

The yoke thickness in the design of a DSTWLPMVM results in a tendency to create a small thrust wave for the machine, although it imposes a restraining force on the end tooth, which can be reduced by design optimization [49]. In the LVPM machine, the coils and the machine's structure affect thrust density, despite the high thrust density provided by the Halbach array [50], [51].

VII. CONCLUSION

An overview of what has been said about the various methods and designs that have been expressed so far. Due to the many efforts made by researchers developing the trend of this machine, the attention toward the machine is increasing. Located and see more enthusiasts than in the past due to the developments. Many designs have been proposed for this machine so far with development approach, performance improvement, power factor, economical designs, efficiency increase, accuracy increase, drive performance improvement, power density, development plans, and efficiency, and it is hoped that this research can be used to effective development steps will help future researchers. In this article, the most critical comprehensive plans have been done to help dynamic researchers in this field with a review approach.

VIII. REFERENCES

- [1] Guohai Liu, Huan Zhong, Liang Xu, and Wenxiang Zhao, "Analysis and Evaluation of a Linear Primary Permanent Magnet Vernier Machine With Multi Harmonics," *IEEE Transactions on Industrial Electronics*, vol. 68, no.1, pp.1983-1993, 2021.
- [2] P.M. Tlali, R-J. Wang, S. Gerber, "Magnetic Gear Technologies: A Review," *IEEE International Conference on Electrical Machines (ICEM)*, 20 November 2014.
- [3] Wenxiang Zhao, Tian Yao, Liang Xu, Xu Chen, Xinxin Song, "Multi-Objective Optimization Design of a Modular Linear Permanent-Magnet Vernier Machine by Combined Approximation Models and Differential Evolution," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 6, pp. 4634 - 4645, 2021.
- [4] Chaojie Shi, Ronghai Qu, Dawei Li, Xiang Ren, Yuting Gao, Zhi Chen, "Analysis of the Fractional Pole-Pair Linear PM Vernier Machine for Force Ripple Reduction," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 6, pp. 4748 - 4759, 2021.
- [5] C. H. Lee, "Vernier Motor and Its Design," *IEEE Transactions on Power Apparatus and Systems*, vol. 82, Issue: 66, June 1963.
- [6] Nick J. Baker, Mohammad A. H. Raihan, Ahmed A. Almoray a, "A Cylindrical Linear Permanent Magnet Vernier Hybrid Machine for Wave Energy," *IEEE Transactions on Energy Conversion*, Vol. 34, no. 2, pp. 691 - 700, 2019.
- [7] Mohammad A. H. Raihan, Nick J. Baker, Kristopher J. Smith, Ahmed A. Almoraya, "Development and Testing of a Novel Cylindrical Permanent Magnet Linear Generator," *IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 3668 - 3678, 2020.
- [8] Kangkang Du, Wenxiang Zhao, Liang Xu, Jinghua Ji, "Design of a New Fault-Tolerant Linear Permanent-Magnet Vernier Machine," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 1, no. 2, pp. 172 - 181, 2020.
- [9] Anqi Ma, Wenxiang Zhao, Liang Xu, Jinghua Ji, Fangfang Bian, "Influence of Armature Windings Pole Numbers on Performances of Linear Permanent-Magnet Vernier Machines," *IEEE Transactions on Transportation Electrification*, Vol 5, no. 2, pp. 385 - 394, 2019.
- [10] D.J. Rhodes, "Assessment of vernier motor design using generalised machine concepts," *IEEE Transactions on Power Apparatus and Systems*, vol. 96, no. 4 ,pp. 1346 - 1352, 1977.
- [11] Ziyi Liang, Yuting Gao, Dawei Li, Ronghai Qu, "Design of a novel dual flux modulation machine with consequent-pole spoke-array permanent magnets in both stator and rotor," in *2018 IEEE, CES Transactions on Electrical Machines and Systems*, vol. 2, n o. 1, pp. 73 - 81, March 2018
- [12] Myounghyun Choi, Byungtaek Kim, "Calculation of PM Vernier Motors Using an Improved Air-Gap Permeance Function," *IEEE Transactions on Magnetics*, vol. 55, no. 6, 2019.
- [13] Alireza Nematsaber; Jawad Faiz, "A Novel Linear Stator-PM Vernier Machine With Spoke-Type Magnets," *IEEE Transactions on Magnetics*, vol. 54, no.11, 2018.
- [14] Xuhui Zhu; Jinghua Ji; Liang Xu; Mei Kang, "Design and Analysis of Dual-Stator PM Vernier Linear Machine With PMs Surface-Mounted on the Mover," *IEEE Transactions on Applied Superconductivity*, vol.28, no. 3, 2017.
- [15] Yi Du, K. T. Chau, Ming Cheng, Ying Fan, Yubin Wang, Wei Hua, Zheng Wang, "Design and Analysis of Linear Stator Permanent Magnet Vernier Machines," *IEEE Transactions on Magnetics*, vol. 47, no. 10, 2011.
- [16] Chaojie Shi, Ronghai Qu, Yuting Gao, Dawei Li, Libing Jing, You Zhou, "Design and Analysis of an Interior Permanent Magnet Linear Vernier Machine," *IEEE Transactions on Magnetics*, vol. 54, no. 11, 2018.
- [17] Zhi Chen, Wubin Kong, Ronghai Qu, Dawei Li, Chaojie Shi, "Symmetrical Components Observer for LPMM Back EMF Based on Negative-Sequence Current Injection," *IEEE Transactions on Industrial Electronics*, vol. 69 no. 2, pp. 1346 - 1355, Feb. 2022.
- [18] Nick J. Baker, Mohammad A. H. Raihan, Ahmed A. Almoraya, Joseph W. Burchell, Markus A. Mueller, "Evaluating Alternative Linear Vernier Hybrid Machine Topologies for Integration Into Wave Energy Converters," *IEEE Transactions on Energy Conversion*, vol. 33, no. 4, pp. 2007 - 2017, 2018.
- [19] Chaojie Shi, Ronghai Qu, Yuting Gao, You Zhou, "Design and Analysis of a Novel Dual-Stator Cylindrical Linear Vernier Machine With LTS Field Windings," *IEEE Transactions on Applied Superconductivity*, vol .29, no. 5, 2019.
- [20] You Zhou, Yuting Gao, Ronghai Qu, Yi Cheng, Chaojie Shi, "A Novel Dual-Stator HTS Linear Vernier Generator for Direct Drive Marine Wave Energy Conversion," *IEEE Transactions on Applied Superconductivity*, vol .29, no. 5, 2019.
- [21] Yi Du, K. T. Chau, Ming Cheng, Ying Fan, Wenxiang Zhao, Fuhua Li, "A Linear Stator Permanent Magnet Vernier HTS Machine for Wave Energy Conversion," *IEEE Transactions on Applied Superconductivity*, vol. 23 ,no.1, 2012.
- [22] Omar Farrok, Md. Rabiul Islam, Md. Rafiqul Islam Sheikh, Youguang Guo, Jianguo G. Zhu, "A Split Translator Secondary Stator Permanent Magnet Linear Generator for Oceanic Wave Energy Conversion," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 9, pp. 7600 - 7608, 2018.
- [23] Feng Xiao; Yi Du; Yubin Wang; Mu Chen; T. W. Ching; Xianxing Liu, "Modeling and Analysis of a Linear Stator Permanent-Magnet Vernier HTS Machine," *IEEE Transactions on Applied Superconductivity*, vol. 25, no. 3, June 2015.
- [24] Chaojie Shi, Ronghai Qu, Baoquan Kou, Dawei Li, Yuting Gao, You Zhou, "A Novel HTS Flux-Reversal Linear Permanent Magnet Machine With a Lower Number of Mover Teeth and Higher Thrust Density," *IEEE Transactions on Applied Superconductivity*, Vol. 28, No. 3, April 2018.
- [25] Lei Huang, Jing Liu, Haitao Yu, Ronghai Qu, Hao Chen, Haiyang Fang, "Winding Configuration and Performance Investigations of a Tubular Superconducting Flux-Switching Linear Generator," *IEEE Transactions on Applied Superconductivity*, vol. 25, no. 3, June 2015.
- [26] Lei Huang, Minqiang Hu, Jing Liu, Haitao Yu, Cengcang Zeng, and Zhongxian Chen, "Electromagnetic Design of a 10-kW-Class Flux-Switching Linear Superconducting Hybrid Excitation Generator for Wave Energy Conversion," *IEEE Transactions on Applied Superconductivity*, vol. 27, no. 4, June 2017.
- [27] Noman Baloch, Salman Khaliq, Byung-Il Kwon, "A High Force Density HTS Tubular Vernier Machine," *IEEE Transactions on Magnetic*, vol.53, no.11, Nov.2017.
- [28] Wenlong Li; T. W. Ching; K. T. Chau, "A New Linear Vernier Permanent-Magnet Machine Using High-Temperature Superconducting DC Field Excitation," in *2017 IEEE Transactions on Applied Superconductivity*, vol .27, no. 4, 2017.
- [29] Noman Baloch, Salman Khaliq, Byung-Il Kwon, "HTS Dual-Stator Spoke-Type Linear Vernier Machine for Leakage Flux Reduction," *IEEE Transactions on Magnetics*, vol. 53, no. 11, 2017.
- [30] Jianguo Li, K. T. Chau, "Design and Analysis of an HTS Vernier PM Machine," *IEEE Transactions on Applied Superconductivity*, vol. 20, no. 3, pp. 1055 - 1059, 2010.
- [31] Jinghua Ji, Wenxiang Zhao, Zhuoya Fang, Jianxing Zhao, Jihong Zhu, "A Novel Linear Permanent-Magnet Vernier Machine With Improved Force Performance," *IEEE Transactions on Magnetics*, vol. 51, no. 8, Aug. 2015.
- [32] Jinghua Ji, Yang Jiang, Wenxiang Zhao, Qian Chen, Anchen Yang, "Sensorless Control of Linear Vernier Permanent-Magnet Motor Based on Improved Mover Flux Observer," *IEEE Transactions on Power Electronics*, vol.35, n o.4, pp. 3869 - 3877, 2019.
- [33] Zhi Chen, Wubin Kong, Chaojie Shi, Ronghai Qu, Vincent Fedida, "A Novel Negative-Sequence Current Injection Method for Reducing LVPMM Electromagnetic Thrust Ripple Considering Static

- Longitudinal End Effect,” *IEEE Transactions on Industrial Electronics*, vol. 68, no. 2, pp. 1108 - 1117, 2021.
- [34] Imanol Eiguren, Gaizka Almandoz, Aritz Egea, Gaizka Ugalde, Ana Julia Escalada, “Linear Machines for Long Stroke Applications—A Review,” *IEEE Access*, vol. 8, pp. 3960 - 3979, 2019.
- [35] Jinghua Ji, Rui Xue, Wenxiang Zhao, Tao Tao, Linsen Huang, “Simplified Three-Vector-Based Model Predictive Thrust Force Control With Cascaded Optimization Process for a Double-Side Linear Vernier Permanent Magnet Motor,” *IEEE Transactions on Power Electronics*, vol. 35, no. 10, pp. 10681 - 10689, 2020.
- [36] You Zhou, Ronghai Qu, Yuting Gao, Chaojie Shi, Cong Wang, “Modeling and Analyzing a Novel Dual-Flux-Modulation Consequent-Pole Linear Permanent-Magnet Machine,” *IEEE Transactions on Magnetics*, vol. 55, no. 7, 2019.
- [37] Ahmed A. Almor aya, Nick J. Baker, Kristopher J. Smith, Mohammad A. H. Raihan, “Design and Analysis of a Flux-Concentrated Linear Vernier Hybrid Machine With Consequent Poles,” *IEEE Transactions on Industry Applications*, Vol.55, No. 5, pp. 4595 - 4604, 2019.
- [38] Wenxiang Zhao, Junqiang Zheng, Jiabin Wang, Guohai Liu, Jianxing Zhao, Zhuoya Fang, “Design and Analysis of a Linear Permanent-Magnet Vernier Machine With Improved Force Density,” *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2072 - 2082, 2016.
- [39] Wenxiang Zhao, Shuai Jiao, Qian Chen, Dezhi Xu, Jinghua Ji, “Sensorless Control of a Linear Permanent-Magnet Motor Based on an Improved Disturbance Observer,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 12, pp. 9291 - 9300, Dec 2010.
- [40] Wenxiang Zhao, Zhonghua Chen, Dezhi Xu, Jinghua Ji, Peng Zhao, “Unity Power Factor Fault-Tolerant Control of Linear Permanent-Magnet Vernier Motor Fed by a Floating Bridge Multilevel Inverter With Switch Fault,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 11, pp. 9113 - 9123, 2018.
- [41] Wenxiang Zhao, Peng Zhao, Dezhi Xu, Zhonghua Chen, Jihong Zhu, “Hybrid Modulation Fault-Tolerant Control of Open-End Windings Linear Vernier Permanent-Magnet Motor With Floating Capacitor Inverter,” *IEEE Transactions on Power Electronics*, vol. 34, no. 3, pp. 2563 - 2572, March 2019.
- [42] T. W. Ching; K. T. Chau; Wenlong Li, “Power Factor Improvement of a Linear Vernier Permanent-Magnet Machine Using Auxiliary DC Field Excitation,” *IEEE Transactions on Magnetics*, vol. 52, no. 7, July 2016.
- [43] Dezhi Xu; Wenxiang Zhao; Hongyu Tang; Xinxin Song; Rui Xue, “Three-Vector-Based Model Predictive Current Control With Zero-Sequence Current Suppression for Open-Winding LPMVM Drives,” *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 225 - 236, Jan. 2021.
- [44] Wenxiang Zhao, Binyu Wu, Qian Chen, Jihong Zhu, “Fault-Tolerant Direct Thrust Force Control for a Dual Inverter Fed Open-End Winding Linear Vernier Permanent-Magnet Motor Using Improved SVPWM,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 9, pp. 7458 - 7467, Sept. 2018.
- [45] Yi Du, Ming Cheng, K. T. Chau, Xianxing Liu, Feng Xiao, Wenxiang Zhao, Kai Shi, Lihong Mo, “Comparison of Linear Primary Permanent Magnet Vernier Machine and Linear Vernier Hybrid Machine,” *IEEE Transactions on Magnetics*, Vol. 50, no. 11, 2014.
- [46] Fan Wu, Ayman M. EL-Refai, “Permanent Magnet Vernier Machines: A Review,” in IEEE 2018 XIII International Conference on Electrical Machines (ICEM), 2018.
- [47] Li Fang; Dawei Li; Dawei Li; Chaojie Shi; Ronghai Qu, “Design and Analysis of Fractional Pole-Pair Linear Permanent Magnet Machine,” *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 2021.
- [48] Wenlong Li, K. T. Chau, Chunhua Liu, Shuang Gao, Diyun Wu, “Analysis of Tooth-Tip Flux Leakage in Surface-Mounted Permanent Magnet Linear Vernier Machines,” *IEEE Transactions on Magnetics*, vol. 49, no. 7, pp. 3949 - 3952, 2013.
- [49] He Zhang, Baoquan Kou, Z. Q. Zhu, Ronghai Qu, Jun Luo, Yi Shao, “Thrust Ripple Analysis on Toroidal-Winding Linear Permanent Magnet Vernier Machine,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 12, 2018.
- [50] Chaojie Shi, Dawei Li, Ronghai Qu, He Zhang, Yuting Gao, Yongsheng Huo, “A Novel Linear Permanent Magnet Vernier Machine With Consequent-Pole Permanent Magnets and Halbach Permanent Magnet Arrays,” *IEEE Transactions on Magnetics*, vol. 53, no. 11, 2017.
- [51] Yuting Gao, Ronghai Qu, Dawei Li, Fengxiang Chen, “Force Ripple Minimization of a Linear Vernier Permanent Magnet Machine for Direct-Drive Servo Applications,” *IEEE Transactions on Magnetics*, vol. 53, no. 6, 2017.

Coordinated Management of EVs Charging Station with a Wide Presence of Renewable Energy Sources

Morteza Shamani^{1*}, Peyman Talebi² and Asghar Akbari Foroud³

Abstract-- Renewable energy generations depend on environmental conditions, and often don't coincide with energy consumptions. In addition, uncontrolled charging of electric vehicles (EVs) leads to technical problems in the grid and economic inefficiency for EVs owners. In this paper, a bi-level energy management algorithm is presented for a distribution network consisting of EVs and renewable energy generations at high penetration level. Energy management at the distribution network level is done centrally by distribution system operator (DSO). In the second level, the parking operator is responsible for providing EVs energy. The parking operator uses the difference in energy prices between the time steps to determine the level of power exchange between EVs and the distribution network in order to reduce the cost of providing energy for EVs in the parking lot. The proposed algorithm coordinates the charging and discharging of EVs with the network conditions and the output of renewable generations while provides financial benefits for vehicle owners. The results indicate the proposed bi-level energy management has been able to use the freedom of the EVs to balance the power of the network and support renewable generations as well as increase the financial benefits of EVs owners.

Index Terms-- EVs, Distributed power generation, Financial benefits of vehicle owners, Energy management

I. INTRODUCTION

Power generation and energy management are some of the most important issues in the policies of today's societies. Several structures have been proposed to provide clean, permanent and low-cost energy. Clean energy is provided by increasing usage of solar, wind, sea waves and etc. On the consumption-side, clean energy consumptions such as electric vehicles (EVs) are replacing fossil fuel vehicles. These EVs are dependent on the power grid to provide power.

Due to dependence on environmental conditions, renewable energy resources are available just for a few hours a day. Also, these resources often do not necessarily coincide with energy consumption, causing imbalances between electricity production and consumption. If renewable power generation exceeds the power consumption, it will lead to overvoltage in the network, followed by a reversal of current power. Conversely, during peak energy consumption periods when the grid does not generate enough renewable power, the

grid voltage drops. Hence, increasing the penetration of renewable energy sources is only possible if it is supported by energy consumption [1].

EVs are increasing as new energy consumptions. The presence of a large number of plug-in EVs in the network and the lack of correct management of these EVs can cause problems for the network. The most important of these problems is the creation of a new peak in the distribution network [2].

The plug-in EVs store the required energy for travel in the batteries embedded in the EVs. The battery capacity of every plug-in electric vehicle is not enough to be managed by distribution system operator (DSO). While the aggregation of these EVs in charge parking can create significant capacity for charging or discharging energy, this capacity is sufficient to be used as an energy management program. Therefore, in order to power balancing in the network, DSO can determine policies for EVs parking lots. DSO's policies for power balancing, along with energy pricing policies, can reinforce motivation of owners of EVs to participate in this direction [3].

Several studies have been conducted on the simultaneous control of renewable energy generations and EVs at the distribution network level. Article [4] for the planning of daily energy resources, coordinated with dynamic electricity prices for EVs, presents a stochastic model for solving the challenges of demand and renewable resources. Reference [5] investigates the economic benefits of coordinated control of distributed energy sources and EVs in smart micro grids operation. Impact of electric vehicle customer response to time-of-use rates on distribution power grids is presented in reference [6]. The authors of [3] have investigated the related topics for integrating the EVs into smart grids, as well as integrating renewable energy generations into EVs. Studies present different strategies for charging the EVs in the distribution network are [7,8]. Optimization, control and charging management of EVs fleet in intelligent networks are discussed in reference [9]. In this study, the effects of EVs managed charging on the performance of the transmission system, distribution, local generation, especially renewable energy generations, as well as the profits of EV owners, have been investigated. In [10] the authors study the combination of EVs and wind power sources for peak modification. Reference [11] provides the formulation

1. Tehran Province DISCO, Tehran, Iran.

2. CRM/BOS Engineer, Behpardaz Hamrah Samaneh Aval (BEHSA), Tehran, Iran.

3. Faculty of Electrical & Computer Engineering, Semnan University, Semnan, Iran.

Corresponding author: mortezashamani@yahoo.com

of a hybrid model for optimal allocation of resources for EVs charging stations, renewable energy sources, and energy storage systems in distribution networks.

The reliability of the distribution network consists of renewable generations and EVs was investigated in [12] and it has been proven that increasing the bidirectional power exchanges between the EVs and the network increases the reliability of the distribution network.

Providing the managed power for EVs can reduce daily load fluctuations. Researchers in [13] determine the final capacity of parking lots by performing optimal power flow to balance energy consumption and production, and maintenance the network voltage level within the allowed range. In [14], an innovative and repetitive method is used to provide required energy for EVs by distributed generations. The strategy of this method creates a virtual price, based on the difference between energy consumption and production. The authors in [15] used the concept of filling low load hours to balance energy consumption and production.

Creating coordination between renewable energy generations and EVs as controllable energy consumptions increase the penetration of renewable energy. Researchers in [16] used a dynamic approach to maximize usage of renewable generations to charge EVs. The basis of determining the power levels of EVs is their state of charge (SOC). The authors [17] have used mixed integer linear programming to reduce usage of conventional power generation and increasing penetration of renewable generations. Paper [18] assesses the economic and environmental impacts of providing renewable energy for charging EVs.

Many factors effect in reducing costs of distribution network. Using low-cost power generations and providing energy at low-cost intervals will reduce network costs. Researchers in [19] used extra power of EVs as reserve in the distribution network. The amount of reserve in the grid is based on the difference between energy consumption and generation. In [20], by changing the distribution network structure through reconfiguration, selects the most appropriate structure in terms of energy providing cost for the network. In [21], a mixed-integer non-linear method was used to reduce the cost of network energy. This methodology was presented by the market agent and EVs agent and by making changes in the price of energy.

Unmanaged increasing of penetration level of distributed power generation and EVs (as energy consumptions), cause intensifying the imbalance between energy consumption and production, resulting in the deviation of voltage levels in the network. Researchers in [22] used optimal power flow to determine the level of parking power exchanges under the allowed conditions for network voltage. In addition, the fuzzy management method was used to prioritize EVs and select the most suitable vehicle for charging.

Among the major setbacks in the distribution network energy management formed from renewable energy and EVs is the uncertainty about the behavior of EVs and local generations. Different methods for compensating of the uncertainty inherent in the behavior of EVs and distributed power generations were

investigated in [23,24]. Monte Carlo method was proposed to compensate for the uncertainty in the behavior of EVs and renewable generations in [25]. In addition, energy storage for integration of renewable energy and EVs was also used in this reference. Researchers in [26] compensated for uncertainty in the behavior of EVs and renewable generations by expanding power flow at 12-hour time steps. The authors in [27] used the probabilistic method to estimate time to start charging and the time taken to charge of EVs to compensate for the uncertainty in EVs behavior. Paper [28] presented a method for developing interaction between EV parking lots and the distribution system operator in the energy and reserve markets, which considers the uncertainty of load and wind power.

EVs charging / discharging management at the parking lots level provide an opportunity to increase the benefits of EVs. In this regard, the issues and problems of the EVs energy supply by parking lots were thoroughly investigated by the researchers in [29,30]. Researchers in [31] have presented a model for increasing the financial benefits of EV parking lots as a multi-energy system. In this reference, the concept of exchanging power between vehicle and parking lot instead of power exchange between vehicle and network is used. Authors of [32] maximized the number of charged EVs with the lowest cost, using demand response methods.

One of the issues that has been neglected in the studies, is the use of EVs in the parking lot as an electrical energy storage. For this reason, EVs can be used to transfer power between peak hours of local generation and peak energy consumption. EVs, in addition to the controllable capability for providing the required energy, can be used like energy storage due to the freedom of action in terms of available battery capacity (as compared to the required battery capacity) and parking time. Especially when EVs can be integrated into parking lots. In addition, the other issues that have not been addressed are creating dependencies in the profits of EVs to support renewable power generation. In this case, any vehicle that benefits from renewable energy generation to provide its required energy or transfers generated power of renewable energy generation to peak hours of consumption will earn more benefits. In this context, creating an infrastructure that can optimize the benefits of network assets is critical for renewable energy generations alongside the benefits of EVs.

In this paper, an algorithm based on bi-level energy management is proposed to control the combination of renewable energy generation and EVs under high penetration. At the first level, DSO will calculate the tariff energy price as day-ahead, based on the behavior of renewable generations, in order to create coherence between controllable consumptions (EVs) and renewable energy generations. In addition, DSO will change parking lots design capacity by performing optimal power flow, with the goal of reducing the energy providing cost and maximizing usage of renewable generations, as well as satisfying network operation constraints. In the second level, the parking lot operator will determine the ability of EVs to earn the financial benefits, using the price difference in the time intervals that EVs are in the parking lot. Then, by performing linear optimization based on the financial benefits of EVs

(depending on difference in energy price), the level of vehicle power exchange with the network will be determined at any time. The innovations of this study are summarized as follows.

- 1- In this method, due to the available excess capacity of EVs batteries, the parking of EVs is used as energy storage in order to transfer as possible as the energy surplus of renewable energy generations to peak hours of consumptions.
- 2- To create coordination between controllable consumptions (EVs) and renewable energy generations, tariff energy price has been created based on the behavior of renewable generations. In this case, the benefits of controllable energy consumptions will depend on the exploitation of renewable generations. Additionally, viewpoint from ahead time intervals will be available to the parking operator by the tariff price.
- 3- In the proposed method, the operation of DSO and parking operator is real-time and is based on hourly electricity production and load balancing. This will cover uncertainties in the behavior of renewable energy generations and owners of EVs.
- 4- The profitability of EVs is improved in two ways. A) Power transmission from the peak hours of renewable generation to peak hours of consumption. B) Providing energy for EVs at time intervals with low price (the peak of renewable generations).

The paper is organized as follows. In Section II, the energy management mechanism is presented at the distribution network level. The energy management of EVs in parking lots is described in Section III. The results of simulation studies are presented in Section IV and the conclusions are presented in Section V.

II. ENERGY MANAGEMENT AT THE DISTRIBUTION NETWORK

For proper operation of a distribution network consisting of distributed power generation and EVs at high penetration level, a parking capacity management approach is presented with regard to energy price. This method is based on the optimal power flow. In this context, the energy tariff rate will be determined in such a way that energy price will be low for EVs at peak hours of renewable power generations and high at peak hours of energy consumption. This will encourage EVs to supply their energy at peak hours of renewable generations. In addition, based on the imbalance between local power generations and energy consumptions, parking lots capacity will be managed in real time market.

The proposed energy management algorithm is performed at one hour's steps. For this purpose, all parameters are assumed to be constant over a one-hour period. In addition, it is assumed that the initial capacity and location of parking lots of EVs and local power generation are already designed and known.

A. Formation of tariff price

The energy tariff rate for controllable consumptions (EVs) will be calculated by DSO as day-ahead. Electricity tariffs are determined based on forecasts of renewable energy production

and energy consumption. For this purpose, the average 24-hour deployment cost per kWh of energy required for electric vehicles is taken as a reference price. Then, according to the difference between the output of renewable generations and energy consumptions, changes in the reference price of energy will be formed (see Fig.1).

The average 24-hours cost of energy needed for EVs can be calculated based on the electricity market rate or the expired cost of energy in the distribution network. The cost of energy in the distribution network is cheaper than energy market rate; because renewable energy generations cover a part of the network's energy consumptions. In this regard, to calculate the reference price of energy, the expired cost of energy in the distribution network is used.

$$\pi_{avg,N} = \frac{\sum_{t=1}^{24} \left[\left(\sum_{j=1}^N P_l^{t,j} \pm \sum_{j=1}^N P_p^{t,j} - \sum_{j=1}^N P_{DG}^{t,j} \right) \pi^t \right] - \left(\sum_{j=1}^N P_{LDG}^{t,j} \right) \pi_{LDG}^t}{\sum_{t=1}^{24} \left(\sum_{j=1}^N P_l^{t,j} \pm \sum_{j=1}^N P_p^{t,j} - \sum_{j=1}^N P_{DG}^{t,j} - \sum_{j=1}^N P_{LDG}^{t,j} \right)} \quad (1)$$

$$\pi_{avg,M} = \frac{\sum_{t=1}^{24} \left(\sum_{j=1}^N P_p^{t,j} \right) \pi^t}{\sum_{t=1}^{24} \sum_{j=1}^N P_p^{t,j}} \quad (2)$$

$$\pi_{avg} = \beta \times \pi_{avg,N} + (1 - \beta) \times \pi_{avg,M} \quad (3)$$

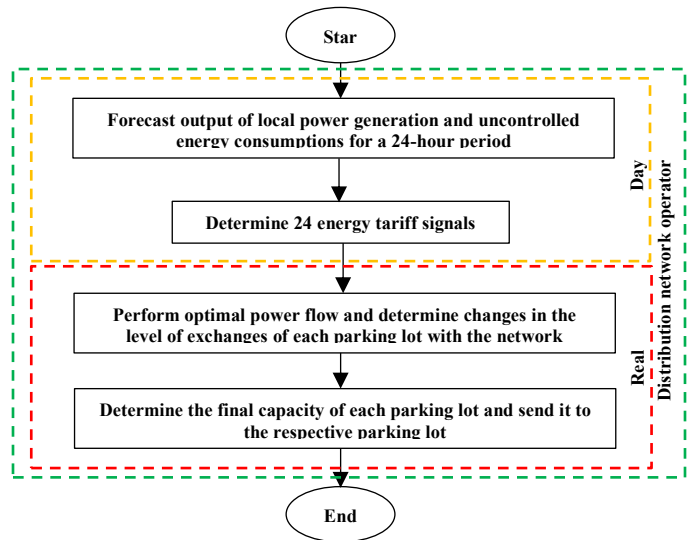


Fig.1. Flowchart of the DSO actions

Equation (1) shows the average cost per kWh of energy in the distribution network. The first term of this equation shows the cost of energy exchanged with the upstream network and the second term shows the cost of energy from non-renewable local production. According to this, the power exchanged with the upstream network is settled at the price of π^t , and the cost of producing non-renewable power is equal to π_{LDG}^t . As mentioned, distributed generation units cover part of the network load, and therefore, in this equation, the amount of electricity production of these units is shown with the opposite sign of the load. The parking lots can also be connected to the network as charging or discharging at any time and have the sign corresponding to electricity production or consumption. In equation (2), the average cost per kWh of energy for parking lots is calculated. In this regard, parking lots can have signs corresponding to electricity production or consumption at any

time. In equation (3), the average 24-hour cost of supplying energy needed for EVs is calculated based on the variation in parking design capacity. By changing the β parameter, it is possible to increase the effectiveness of parking lots in the production cost.

Parking lots tariff price rate is based on difference in the output of renewable generations and energy consumptions, as expressed in (4). If renewable energy generations are more than energy consumptions, the cost of energy will be lower for parking lots. In contrast, with increasing energy consumptions, the energy price for parking lots is rising, so parking lots can earn revenue by selling (discharging) energy to the network.

$$\pi_p^t = \pi_{avg} + \gamma \left[\frac{[(x \cdot \sum_{j=1}^N P_l^{t,j} - \sum_{j=1}^N P_{DG}^{t,j})]}{\sum_{j=1}^N P_p^{t,j}} \cdot \pi_{avg} \right] \quad (4)$$

In this equation, coefficient γ is effect of the price of the parking tariff from difference between renewable generation and energy consumption. The value of this parameter is at the discretion of the system operator to reflect impact of renewable resources in supplying the load.

B. Parking lots capacity management

The DSO manages capacity of parking lots in real time market based on information of energy consumptions and power generations for each time step. Renewable energy generations will be used at the maximum available capacity. In addition, parking lots will operate at nearest value to their design capacity. DSOs change parking lot design capacity with the goal of reducing the cost of providing energy for targeted consumption and maintaining grid constraints based on optimal power flow. The imbalance between power consumption and generation in the studied distribution network will be exchanged with the upstream network. In the event of violation of the voltage constraint or grid lines loading limitations, the surplus or deficit will be exchanged with parking lots and will change their design capacity.

As mentioned, the objective function is minimizing the cost of providing energy of covered consumptions:

$$\min \left\{ \sum_{j=1}^N \left((P_p^{t,j} \cdot \pi_p^t) + (P_{LDG}^{t,j} \cdot \pi_{LDG}^t) \right) + \left(\sum_{j=1}^N (P_p^{t,j} + P_{DG}^{t,j}) - \sum_{j=1}^N (P_l^{t,j}) \right) \cdot \pi^t \right\} \quad (5)$$

Given the following constraints:

$$V_{min} \leq V_m \leq V_{max} \quad (6)$$

$$p_{ij}^2 + q_{ij}^2 \leq S_{ij,max}^2 \quad (7)$$

$$P_{p,min} \leq P_p^{t,j} \leq P_{p,max} \quad (8)$$

$$S_{DG} \leq S_{DG,max} \quad (9)$$

$$p_i = V_i \sum_{k=1}^N Y_{ij} V_j \cos(\delta_i - \delta_j - \theta_{ij}) \quad (10)$$

$$q_i = V_i \sum_{k=1}^N Y_{ij} V_j \sin(\delta_i - \delta_j - \theta_{ij}) \quad (11)$$

$$p_{ij} = V_i V_j (G_{ij} \sin \delta_{ij} + B_{ij} \cos \delta_{ij}) - G_{ij} V_i^2 \quad (12)$$

$$q_{ij} = V_i V_j (G_{ij} \cos \delta_{ij} - B_{ij} \sin \delta_{ij}) + B_{ij} V_i^2 \quad (13)$$

The first part of equation (5) represents the cost of energy supplied through parking lots and distributed generation sources. The second part of this equation also shows settlement of the power imbalance with the real time market price of energy.

Constraints (6), (7), (8), and (9) indicate network voltage limitation, lines loading limits, capacity limitation for parking lots, and limitation in local power production, respectively. The active and reactive injection power to the buses in (10) and (11), respectively, and the active and reactive power of the lines in (12) and (13) are visible.

C. Modeling distributed generations

The sources of distributed power generation studied in this paper include photovoltaic and wind power generations as renewable power generation [33], and the power generation sources of micro-turbine and fuel cell as non-renewable power generation [34,35].

$$P_{pv}^{t,j} = \eta_g n A_m G_t \quad (14)$$

Where η_g is the instantaneous PV generator efficiency, A_m the area of a single module used in a system (m^2), G_t the global irradiance incident on the titled plane (W/m^2) and N is the number of modules.

$$P_{wt}^{t,j} = \begin{cases} 0 & v_w^t < v_{min} \text{ or } v_w^t > v_{max} \\ \frac{v_w^t - v_{min}}{v_r - v_{min}} & v_{min} \leq v_w^t < v_r \\ P_r & v_r \leq v_w^t < v_{max} \end{cases} \quad (15)$$

$$P_{DG}^{t,j} = P_{pv}^{t,j} + P_{wt}^{t,j} \quad (16)$$

$$\pi_{LDG}(t) = \left(U^t / P_{LDG}^t \right) + b \cdot P_{LDG}^t \quad (17)$$

$$P_{LDG}^t - P_{LDG}^{t-1} \leq RU \quad (18)$$

$$P_{LDG}^{t-1} - P_{LDG}^t \leq RD \quad (19)$$

$$P_{LDG,min}^t \leq P_{LDG}^t \leq P_{LDG,max}^t \quad (20)$$

The equations (14) and (15) respectively represent the output of photovoltaic power generation and dependence on the solar radiation level and the output of wind power generation, depending on velocity of the wind. Equation (16) shows distributed renewable energy generations at time t and bus j is equal to wind and photovoltaic power generation. Equation (17) expresses the cost of local non-renewable electricity production based on the output power and amount of start-up cost per kilowatt of production. The constraints (18 to 20) are correlated with the rate of power growth, power loss and local generation capacity.

III. ENERGY MANAGEMENT AT THE EVS PARKING LOT LEVEL

The ideal conditions for EVs owners are to provide their energy at the cheapest time steps of vehicle presence in the parking lot. In addition, owners of EVs can earn revenue by

doing charge and discharge managed at the other time steps of their presence in the parking lot. Therefore, if the average price of energy at the effective rate of return is taken as the base price, buying energy at a lower price and selling energy at a higher price than the base price increases the economic benefits of electric vehicle owners. Also, the larger the time step price difference from the base price, the more profitable the EV. Thus, by analyzing the state of charge of electric vehicles and analyzing the price of energy charges, the degree of profitability of electric vehicles is determined (see Fig.2).

In the proposed method, the percentage (half) of the financial benefits of EVs is allocated to the owner of the parking lot. On the other hand, due to limited capacity of parking lots, the benefits of all EVs can't be provided. Therefore, the owner of the parking lot is trying to support a vehicle that is more profitable.

A. Determine the level of power exchange between EVs and the network through parking lot

The level of power exchange between EVs and the network (through parking) is optimized by linear programming. The purpose of determining the level of power exchange between EVs with the network based on the financial profitability of EVs is to maximize the revenue of set of EVs inside parking lot.

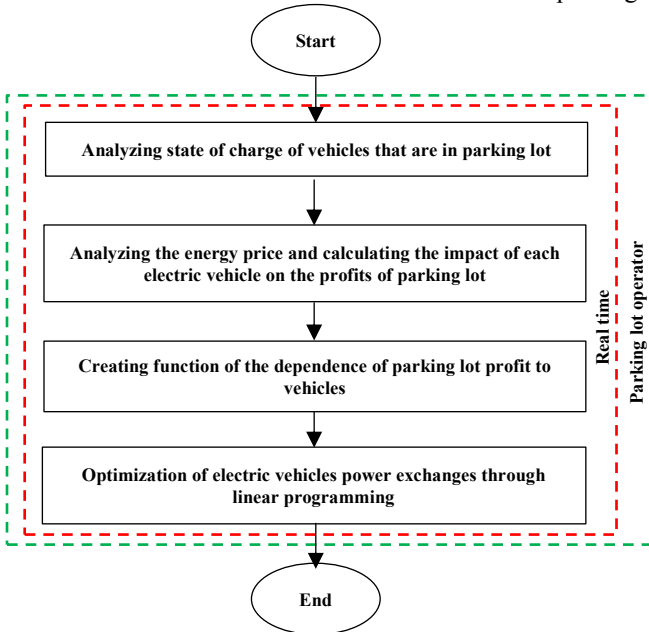


Fig.2. Flowchart of the parking lot operator behavior

$$\max\{\sum_{e=1}^n [M_c^{e,t} \cdot P_{EV}^{e,t}] - K[P_P - \sum_{e=1}^n P_{EV}^{e,t}]\} \quad (21)$$

Constraints of linear optimization method will also be managed with regard to EVs constraints.

$$P_{EV,min}^{e,t} \leq P_{EV}^{e,t} \leq P_{EV,max}^{e,t} \quad (22)$$

Equation (21) expresses the goal of the parking operator in two parts. The first part is related to the financial benefits of EVs inside the parking lot. The second part is guaranteed to maintain parking capacity within the allowed range. K is a great value which puts the greatest benefit of parking in subject to

achievement the parking design capacity.

The analysis of the state of charge of EVs shows the amount of power required or the excess energy stored in the battery of EVs for the parking operator. In this regard, final state of charge of EVs is considered as a reference for state of charge. Based on the difference in state of charge of each vehicle from its reference value, the number of required steps one-hour is calculated to provide the vehicle's final energy according to equation (23).

$$T_m^{e,t} = \frac{(soc_{fin}^e - soc^{e,t}) \times E_{batt}^e}{P_{charger}^e} \quad (23)$$

The parking lot operator calculates the profitability of existing EVs in the parking lot by analyzing energy price. In this regard, according to the results of the analysis of the state of charge of EVs, the cheapest time steps for the providing of final energy of vehicle are predicted and the average price of other steps will be considered as a reference price. The price deviation of each step compared to the price reference indicates the amount of profitability of the electric vehicle at that time step.

The average energy price in time steps and the charge motive function of the vehicle are presented using equations (24) and (25).

$$C_{avg}^{e,t} = \frac{\sum_l \pi_p^l}{t_{ex}^{e,t} - T_m^{e,t}}, l \in A \quad (24)$$

$$M_c^{e,t} = C_{avg}^{e,t} - \pi_p^t \quad (25)$$

B. Modeling of EVs

The most important parameters in modeling the behavior of EVs are battery capacity, state of charge of battery at any time step, and charger power. In this regard, the technical constraints and limitations below are important [36].

$$(1 - D_{plag}^{e,t})(P_{ch}^{e,t} + P_{dis}^{e,t}) = 0 \quad (26)$$

$$P_{ch}^{e,t} \times P_{dis}^{e,t} = 0 \quad (27)$$

$$0 \leq Q_{ch}^{e,t} \leq (soc_{max}^e - soc^{e,t}) \cdot E_{batt}^e \quad (28)$$

$$0 \leq Q_{dis}^{e,t} \leq (soc^{e,t} - soc_{min}^e) \cdot E_{batt}^e \quad (29)$$

$$P_{ch}^{e,t} \leq P_{charger}^e \quad (30)$$

$$P_{dis}^{e,t} \leq P_{charger}^e \quad (31)$$

$$soc^{e,t_{ex}} \geq soc_{fin}^e \quad (32)$$

$$soc^{e,t+1} = soc^{e,t} + \left(\frac{P_{EV}^{e,t}}{E_{batt}^e}\right) \quad (33)$$

$$P_{EV}^{e,t} = P_{ch}^{e,t} + P_{dis}^{e,t} \quad (34)$$

$$P_p^{j,t} = \sum_e P_{EV}^{e,t} \quad (35)$$

Based on (26), EVs have the ability to exchange with the network only if they are connected to the network. Furthermore,

EVs do not have the ability to simultaneously charge and discharge, which is modeled (27). The available capacity for charging and discharging EVs is modeled in (28) and (29). Equations (30) and (31) model the charger's limitation and (32) represents a guarantee for the providing energy required for EVs. The charging state of EVs at the end of each step is calculated according to (33). Equation (34) also indicates the power exchanged with each EV at each node. It should be noted that equations (26) to (34) are valid for all nodes. Equation (35) also shows the total power exchanged with each parking lot at any time.

IV. SIMULATIONS AND NUMERICAL STUDIES

In this section, the proposed bi-level energy management method will be tested. In this study, simulation is presented using GAMS optimization software tested on the standard IEEE 33-bus network [38]. In this regard, in order to verify the validity of the proposed method, the behavior of DSO has been investigated in six states. These states, which depend on the level of penetration of distributed power generation and EVs, are shown in Table I. The second part of the study was done in a sample parking lot and the parking operator's behavior was investigated in three states. In the first and second modes, the ability to change the parking capacity is about 1/4 and 1/8 of parking design capacity, respectively. In the third case, parking constraints are not considered and parking operator behavior is only in order to meet the needs of EVs.

TABLE I
Penetration Level of Local Generation and Evs in Study Scenarios

Study of DSO behavior	Study states					
	1	2	3	4	5	6
Penetration level of local generation	66%	66%	66%	33%	33%	33%
Penetration level of EVs	0%	30%	50%	0%	30%	50%

A. Input data

The initial information required to study the behavior of the DSO under the proposed method of energy management involves predicting the behavior of distributed generations, energy price, and energy consumptions. This information is presented in Fig.3 [28, 39].

Initial information on fuel cell and microturbine energy sources is also provided in Table II [34, 35]. Additionally, the allowed voltage deviation is considered to be 0.05 per unit [40]. In this study, half of local power productions are allocated to photovoltaic and 40% to wind resources, and the rest to non-renewable productions. In addition, four parking spaces are located at buss 9, 15, 27 and 32 with a design capacity of 500 KWh and a 25% changeability capability.

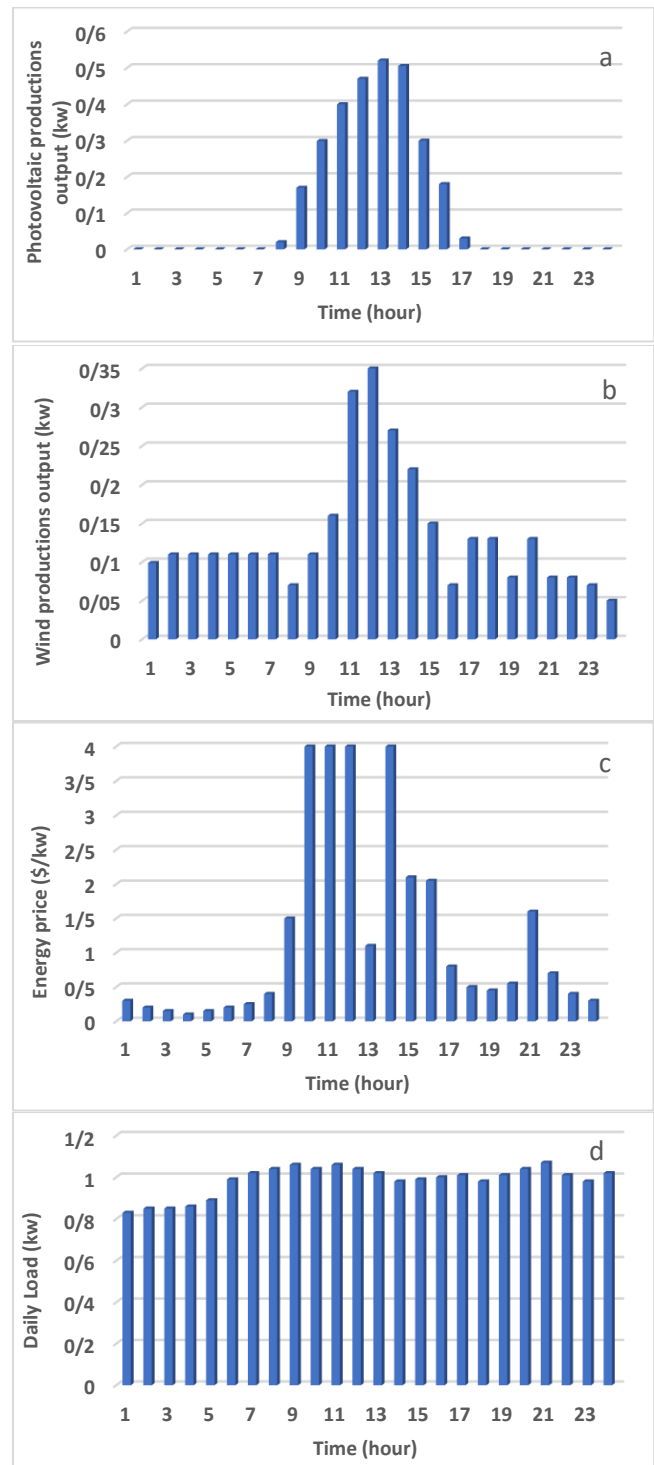


Fig.3. a) Output of photovoltaic power generation. b) output of wind power generation. c) Energy prices at one hour's steps. d) predicted daily load figure

TABLE II
Information of Non-Renewable Local Power Generations

Type	a	b	P_{min}	P_{max}	MUT	MDT	RU	RD
Micro turbine	0.015	0.01	150	700	3	3	350	350
Fuel cell	0.02	0.01	50	300	1	1	150	150

The initial information for EVs is presented in Table III. The EVs in this study use batteries with a capacity between 10 and 30 kWh [41]. In addition, the charger power used for EVs is up to 2.5 kWh. This power is related to parking lots with a voltage level of 208 to 240 V and a current of about 10 A [27]. On the other hand, for the purpose of increasing the battery life of EVs and preventing high discharge, as well as guaranteeing the availability of EVs in an emergency, the minimum state of charge is 30% of the vehicle's battery capacity.

TABLE III
Information of Evs Studied in Sample Parking

Type	Input time(h)	Exit time(h)	Input SOC (%)	Exit SOC (%)	Battery capacity (kWh)	Number
vehicle fleet 1	1	6	30	80	18	260
vehicle fleet 2	2	9	45	85	21	240
vehicle fleet 3	7	13	45	90	18	165
vehicle fleet 4	8	16	30	85	18	150
vehicle fleet 5	13	21	30	80	20	185
vehicle fleet 6	19	24	40	85	24	240

In this study, the fleet of EVs was examined in three categories. The first category of EVs enter the parking lot before the peak of renewable generations. The second and third categories are present in the parking lot during and after the peak of renewables generations. Each of these three categories, according to the input SOC, is grouped into two groups with input SOC and without input SOC.

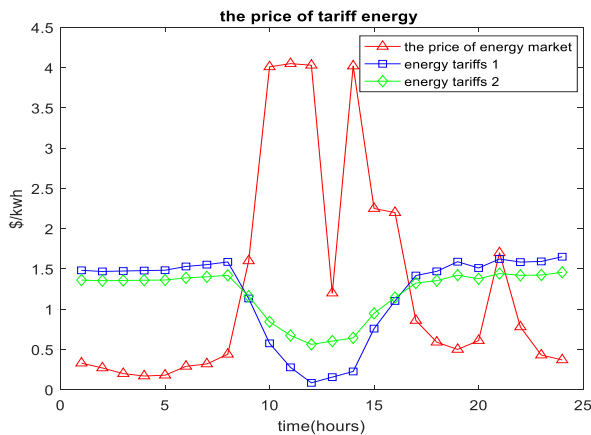


Fig.4. Tariff energy price provided by DSO

B. Results of optimization of DSO behavior

DSO calculates the energy price in the first step. Fig. 4 shows the tariff decision results. On the other hand, as can be seen from the figure, parking lots have low energy costs during 10:00 to 15:00 (the peak of renewable energy generation), and expensive at the peak of energy consumption. In addition, the energy tariff 1 in the peak hours of renewable energy generations is cheaper than the energy tariff 2. Because energy

tariff 1 is related to the penetration level of 66% of local generations and energy tariff 2 is related to the penetration of 33% of local generations.

The bus voltage level of the distribution network at peak hours of energy consumption (at 21:00 o'clock) and the peak of renewable generation (at 13:00 o'clock) is shown in Fig.5 and Fig.6.

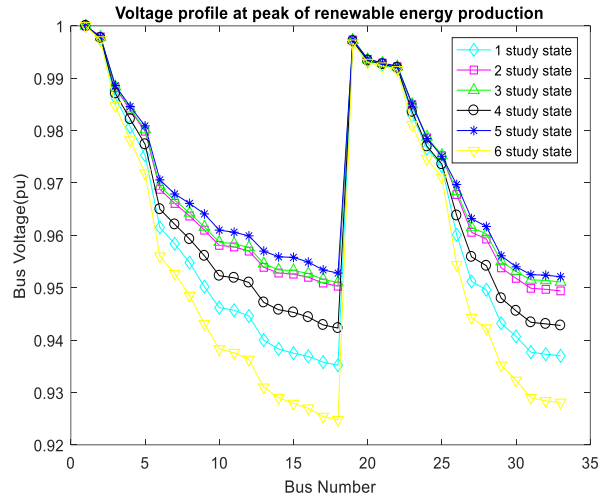


Fig.5. Network voltage level at the peak of renewable power generation

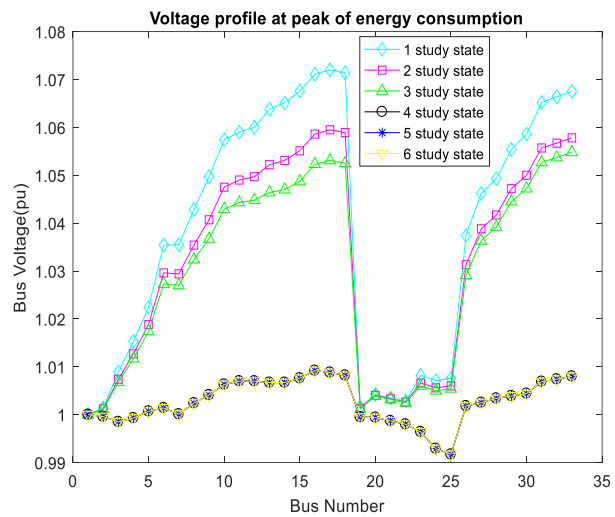


Fig.6. Network voltage level at peak power consumption

By reflecting to Fig. 5, in State 3, while penetration coefficient of EVs and renewable energy sources is simultaneously at its highest value, the voltage level is in the best way within the permissible range. From Fig.6, it can be seen that with increase of penetration of EVs (from state 1 to 3), the voltage profile of the buses is in a better range. This shows that the increasing penetration of EVs will help to improve the voltage profile in all buses.

As presented in Fig.7 an imbalance between local production and energy consumption exists. When the wind and photovoltaic power generation are at their highest level (10:00-15:00), the power generation is more than of consumption, on the other hand, during the peak consumption hours (18:00-23:00), the amount of consumption is more than production, and this imbalance must be compensated at a later stage.

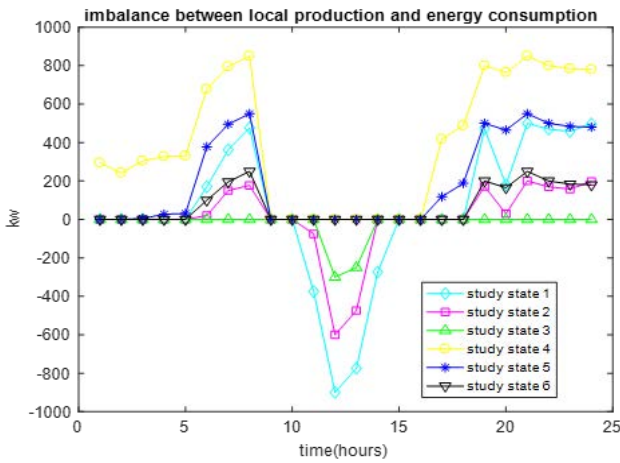


Fig.7. Imbalance between local generations

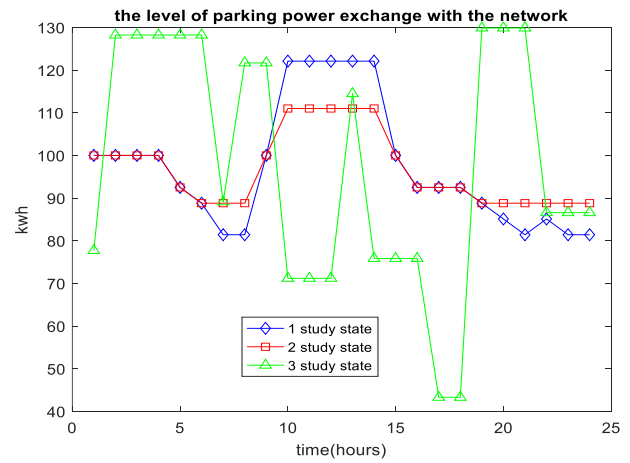


Fig.9. Power exchange level of parking with the network

The results of optimal power flow manage the level of power exchange (capacity) of the parking lots. This is presented in Fig.8. As it is clear from this figure, the level of energy exchange of parking lots decreases during peak consumption hours (18:00-24:00) or when renewable energy generation decreases (5:00-8:00). Instead, in the peak state of production of renewable energy resources (10:00-15:00) the energy exchange level of the parking lots has increased and reached 650 KW in the highest state.

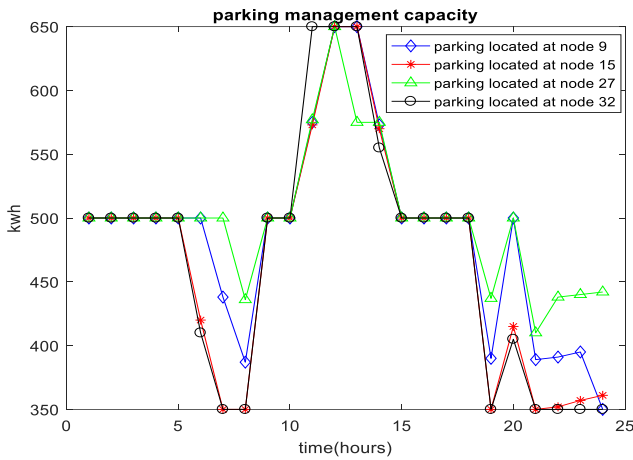


Fig.8. Parking lots managed capacity

At some times steps, parking design capacity is unchanged. In these intervals, local generation and energy consumption are at a level. In other words, energy consumptions are provided locally, or the supply of energy from the upstream network will not violate the network conditions.

The proposed model could also capitalize on the positive potential of connecting electric vehicles to the grid to increase the use of renewable energy.

C. Results of optimization of parking operator behavior

The power exchange level between parking and network for the three study states in Fig.9 is investigated. State 3, which shows the level of power needs of vehicle, expresses that the needs of the EVs aren't matched to the parking capacity.

The cumulative profit of the parking operator for the two study states 1 and 2 is compared in Fig.10. This is the result of the difference of parking cost in two states than the parking cost for state 3 (Elimination of the power needs of EVs). As it is known, parking is in the first state more profitable.

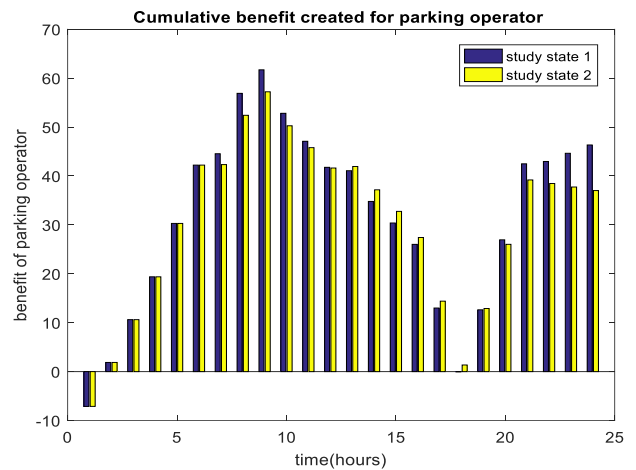


Fig.10. Cumulative profit created for parking operator

The results of the management program of charging/discharging EVs based on linear optimization are shown in Table IV. The fleet of EVs that were present in the parking lot before the peak of renewable generations, with the release of available battery capacity at expensive intervals, provided their energy at the peak of renewable generations.

For example, during 2:00-4:00, when renewable energy generation is low due to the lack of solar energy, in fleet 2, EVs benefit by discharging their batteries and selling power to the grid. Then during 6:00-9:00, when renewable energy generation increases, EVs receive energy from the grid and charge their batteries. The fleets of EVs that were in parking lots after peak renewable energy generation increased their financial benefits by transferring power from peak hours of renewable energy generation to peak consumption hours. In the meantime, the fleet of EVs that provided energy at the peak of renewable generations made the balance between power consumption and production.

TABLE IV
Results for Scheduling of Evs State Of Charge

hour	SOC of vehicle1 (%)	SOC of vehicle2 (%)	SOC of vehicle3 (%)	SOC of vehicle4 (%)	SOC of vehicle5 (%)	SOC of vehicle6 (%)
1	40.71					
2	54.6	42.07				
3	68.49	39.14				
4	82.38	36.21				
5	81.73	45.97				
6	80.68	55.73				
7		67.63	39.75			
8		79.53	30	35.07		
9		85	42	30		
10			54	37.33		
11			66	44.67		
12			78	52		
13			90	52	35.28	
14				63	42.36	
15				74	47.94	
16				85	53.02	
17					65.52	
18					78.02	
19					78.02	48.01
20					78.02	55.71
21					80	62.13
22						69.82
23						77.51
24						85.21

In order to better understand the performance of electric vehicle battery charging process in the presence of renewable energy sources, Fig.11 is presented.

In Fig.11, the SOC of EVs batteries for fleet 3 and fleet 4 are shown during the hours of high penetration of renewable energy sources. These two fleets enter the parking lot before the extensive renewable energy generation and leave the parking lot after its reduction. As seen in the performances of these two fleets, with the increase in wind and photovoltaic energy generation, EV owners charge their batteries more and increase their SOC. In fact, these EVs have supported the increase in renewable energy generation by managing their SOC.

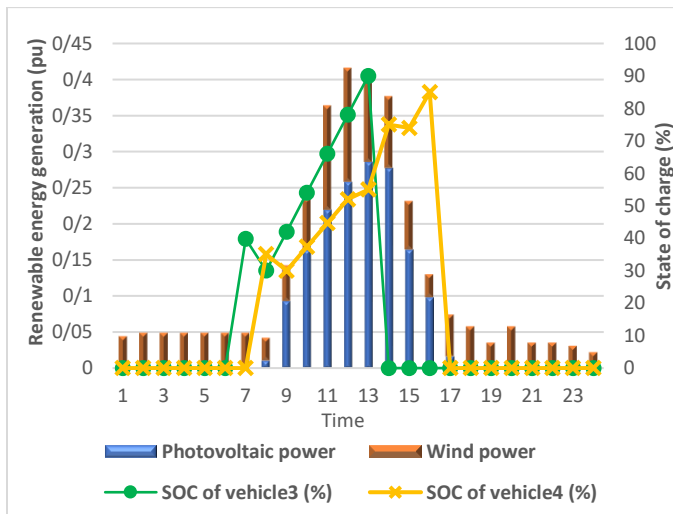


Fig.11. SOC of vehicle fleet 3 and vehicle fleet 4

V. CONCLUSION

This study sought to assess the impact of using energy prices as a basis for establishing coordination among EVs and the output of renewable energy sources in two stages. In addition, the freedom of EVs has been used to offset the imbalance between energy consumption and production. As seen from the simulation results, the profit of electric vehicles has increased and the grid voltage profile has also improved. It was observed that by using the presented method, the capacities of EVs parking lots are well managed. The proposed model could also capitalize on the positive potential of connecting EVs to the electricity distribution network to increase the use of renewable energy.

NOMENCLATURE

Indexes

- A Index of time steps that considered for earning revenue of EV
- β Variability coefficient of design capacity of parking lots
- e Index of EVs
- i, j Index of Buses
- n Index related to number of EVs available in the parking lot
- N Total number of buses in network
- t Index of Time
- t_{ex} Index related to the time which the EV exits from the parking lot
- x Percentage of daily consumptions that local productions are considered to meet them
- γ Impact of tariff price from diversion between renewable production and energy consumption

Parameters

- a, b Coefficient of cost function of non-renewable productions
- δ_i Voltage angle at bus i
- δ_{ij} The angle difference between the voltages of buses i and j
- E_{batt}^e Battery capacity of EV
- $M_c^{e,t}$ The charge motive function of the vehicle e based on the energy price at time t
- MDT Minimum downtime
- MUT Minimum uptime
- π^t Price of energy market
- π_l^t Contractual energy price between DSO and uncontrollable consumptions
- π_{LDG}^t Energy price of non-renewable productions
- $P_{charger}^e$ Charger power of EV e
- $P_{EV,max}^{e,t}$ Maximum level of vehicle power exchange with the network
- $P_{EV,min}^{e,t}$ Minimum level of vehicle power exchange with the network
- $P_{LDG,max}$ Maximum output of non-renewable energy production
- $P_{LDG,min}$ Minimum output of non-renewable energy production

$P_l^{t,j}$	Power consumption of uncontrollable loads at time t and bus j
$Q_{ch}^{e,t}$	Capacity available for charging EV at time t
$Q_{dis}^{e,t}$	Capacity available for discharging EV at time t
RD	Power downward rate for non-renewable productions
RU	Power upward rate for non-renewable productions
$S_{ij,max}$	Maximum power of the line between buss i and j
$S_{DG,max}$	Maximum of distributed generation output
soc_{fin}^e	Final state of charge of EV e
soc_{max}^e	Maximum of state of charge of EV e
soc_{min}^e	Minimum of state of charge of EV e
$T_m^{e,t}$	The time it takes to prepare the vehicle e for the first zero at time t
θ_{ij}	Angle of Y_{ij}
U	Start-up cost of non-renewable productions
V_m	Voltage of bus m
V_{max}	Maximum network voltage level
V_{min}	Minimum network voltage level
Y_{ij}	Admittance of line between buses i and j

Variables

$C_{avg}^{e,t}$	Average energy price in time steps related to index A
$D_{plag}^{e,t}$	A binary variable indicating that vehicle e be in the parking lot at a time step t
π_{avg}	The average 24-hour cost of providing energy of EVs
π_p^t	Tariff energy price for parking lots
$P_{ch}^{e,t}$	Transferred power from network to EV at time t
$P_{dis}^{e,t}$	Transferred power from EV e to the network at time t
$P_{DG}^{t,j}$	Output of distributed renewable energy generations at time t and bus j
$P_{EV}^{e,t}$	Power exchange level of vehicle e with the network at time t
p_i	Injected active power to bus i
p_{ij}	Active power of the line between buss i and j
$P_{LDG}^{t,j}$	Output of non-renewable energy generation at time t and bus j
$P_p^{t,j}$	Level of power exchanges based on the tariff price between the network and parking lot at time t and bus j
P_{pv}^t	Photovoltaic power generation output
P_{wt}^t	Wind power generation output
q_i	Injected reactive power to bus i
q_{ij}	Reactive power of the line between buss i and j
$soc^{e,t}$	State of charge of EV at time t

VI. REFERENCES

- [1] A. Mandev, P. Plötz, F. Sprei, G. Tal, "Empirical charging behavior of plug-in hybrid electric vehicles," *Applied Energy*, Vol. 321, September, 2022, 119293.
- [2] F. Mwasilu, J. John Justo, E.-K. Kim, T. Duc Do and J.-W. Jung, "Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration," *Renewable and Sustainable Energy Reviews*, Vol.34, pp. 501-516, June, 2014.
- [3] C. Roe, A. P. Meliopoulos, J. Meisel and T. Overbye, "Power System Level Impacts of Plug-In Hybrid EVs Using Simulation Data," *IEEE Energy 2030*, Atlanta, GA USA, November, 2008.
- [4] J. Soares, M. A. Fotouhi Ghazvini, N. Borges, Z. Vale, "Dynamic electricity pricing for EVs using stochastic programming," *Energy*, Vol. 122, pp. 111-127, March, 2017.
- [5] A. G. Anastasiadis, G. P. Kondylis, G. A. Vokas, S. A. Konstantinopoulos, C-T. Salame, A. Polyzakis, K. Tsatsakis, "Economic benefits from the coordinated control of Distributed Energy Resources and different Charging Technologies of EVs in a Smart Microgrid," *Energy Procedia*, Vol. 119, pp. 417-425, July, 2017.
- [6] C.B. Jones, W. Vining, M. Lave, T. Haines, C. Neuman, J. Bennett, D.R. Scoffield, "Impact of Electric Vehicle customer response to Time-of-Use rates on distribution power grids," *Energy Reports*, Vol. 8, pp. 8225-8235, November, 2022.
- [7] L. Luo, P. He, S. Zhou, G. Lou, B. Fang, P. Wang, "Optimal scheduling strategy of EVs considering the limitation of battery state switching times," *Energy Reports*, Vol. 8, Supplement 5, pp. 918-927, August, 2022.
- [8] K. Gorgani Firouzjah, "Profit-based electric vehicle charging scheduling: Comparison with different strategies and impact assessment on distribution networks," *International Journal of Electrical Power & Energy Systems*, Vol. 138, 107977, June, 2022.
- [9] Z.J. MaxShen, B. Feng, C. Mao, L. Ran, "Optimization models for electric vehicle service operations: A literature review," *Transportation Research Part B: Methodological*, Vol. 128, pp. 462-477, October, 2019.
- [10] A. G. Anastasiadis, G. A. Vokas, S. A. Konstantinopoulos, G. P. Kondylis, T. Khalilof, A. Polyzakis, K. Tsatsakis, "Wind Generation and EVs coordination in Microgrids for Peak Shaving purposes," *Energy Procedia*, Vol. 119, pp. 407-416, July, 2017.
- [11] S. M. Kandil, H E.Z. Farag, M F. Shaaban, M. Zaki El-Sharafy, "A combined resource allocation framework for PEVs charging stations, renewable energy resources and distributed energy storage systems," *Energy*, Vol. 143, pp. 961-972, January, 2018.
- [12] S. Shafiq, A.T. Al-awami, "Reliability and Economic Assessment of Renewable Micro-Grid with V2G EVs Coordination," *IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies*, Amman, Jordan, December, 2015.
- [13] T. Mao, W. H. Lau, C. Shum, H. Chung, K. F. Tsang, N.C. F. Tse, "A new schedule-controlled strategy for charging large number of Vs with load shifting and voltage regulation," *IEEE PES Asia-Pacific power and Energy Engineering Conference (APPEEC)*, Brisbane, QLD, Australia, November, 2015.
- [14] E.L. Karfopoulos, N.D. Hatziazyriou, "Distributed Coordination of EVs Providing V2G Services," *IEEE Transactions on Power Systems*, Vol. 31, No. 1, pp. 329-338, February, 2016.
- [15] E. Xydias, C. Marmaras, L. M. Cipcigan, "A multi-agent based scheduling algorithm for adaptive EVs charging," *Appl. Energy* 177, pp. 354-365, September, 2016.
- [16] F. Khoucha, M. Benbouzid, Y. Amirat, A. Kheloui, "Integrated Energy Management of a Plug-in Electric Vehicle in Residential Distribution Systems with Renewables," *IEEE 24th International Symposium on Industrial Electronics (ISIE)*, pp. 717-722, June, 2015.
- [17] A. Schuller, C. M. Flath, S. Gottwalt, "Quantifying load flexibility of EVs for renewable energy integration," *Appl. Energy*, Vol. 151, pp. 335-344, August, 2015.
- [18] A. Alsharif, C. Wei Tan, R. Ayop, A. Dobi, K. Yiew Lau, "A comprehensive review of energy management strategy in Vehicle-to-Grid technology integrated with renewable energy sources," *Sustainable Energy Technologies and Assessments*, Vol. 47, 101439, October, 2021.
- [19] A. Zakariazadeh, S. Jadid, P. Siano, "Integrated operation of EVs and renewable generation in a smart distribution system," *ENERGY Convers. Manag.*, vol. 89, pp. 99-110, January, 2015.
- [20] A. Kavousi-fard and A. Khodaei, "Efficient integration of plug-in EVs via reconfigurable microgrids," *Energy*, Vol. 111, pp. 653-663, September, 2016.
- [21] J. Hu, H. Morais, M. Lind, H. W. Bindner, "Multi-agent based modeling for electric vehicle integration in a distribution network operation," *Electr. Power Syst. Res.*, Vol. 136, pp. 341-351, July, 2016.
- [22] E. Akhavan-rezai, S. Member, M. F. Shaaban, S. Member, S. Member, F. Karray, S. Member, "Online Intelligent Demand Management of Plug-In EVs in Future Smart Parking Lots," *IEEE Systems Journal*, Vol. 10, No. 2, pp. 483-494, April, 2016.

- [23] J. Zhou, Y. Zhang, Z. Li, R. Zhu, A. zeman, "Stochastic scheduling of a power grid in the presence of EVs, RESS, and risk index with a developed lightning search algorithm," *Journal of Cleaner Production*, Vol. 364, 132473, September, 2022.
- [24] J. Wu, Y. Liu, X. Chen, C. Wang, W. Li, "Data-driven adjustable robust Day-ahead economic dispatch strategy considering uncertainties of wind power generation and electric vehicles," *International Journal of Electrical Power & Energy Systems*, Vol. 138, 107898, June, 2022.
- [25] M. S. Elnozahy, T. K. Abdel-galil, M. M. A. Salama, "Probabilistic ESS sizing and scheduling for improved integration of PHEVs and PV systems in residential distribution systems," *Electr. Power Syst. Res.*, Vol. 125, pp. 55–66, August, 2015.
- [26] A. O. Connell, S. Member, D. Flynn, S. Member, A. Keane, "Rolling Multi-Period Optimization to Control Electric Vehicle Charging in Distribution Networks," *IEEE Transactions on Power Systems*, Vol. 29, No. 1, pp. 340–348, January, 2014.
- [27] H. Kamankesh, V. G. Agelidis, A. Kavousi-fard, "Optimal scheduling of renewable micro-grids considering plug-in hybrid electric vehicle charging demand," *Energy*, Vol. 100, pp. 285–297, April, 2016.
- [28] S. Aghajani, M. Kalantar. "Operational scheduling of EVs parking lot integrated with renewable generation based on bilevel programming approach," *Energy*, Vol. 139, pp. 422-432, November, 2017.
- [29] J. C. Mukherjee and A. Gupta, "A Review of Charge Scheduling of EVs in Smart Grid," *IEEE Systems Journal*, Vol. 9, No. 4, pp. 1541-1553, October, 2015.
- [30] M. Kuran, A. C. Viana, L. Iannone, D. Kofman, G. Mermoud, J. p. Vasseur, "A Smart Parking Lot Management System for Scheduling the Recharging of EVs," *IEEE Transactions on Smart Grid*, Vol. 6, No. 6, pp. 2942-2953, November, 2015.
- [31] M. Yazdani-Damavandi, M. P. Moghaddam, M.R. Haghifam, M. Shafiekhah, J. P. S. Catalão, "Modeling Operational Behavior of Plug-in EVs' Parking Lot in Multienergy Systems," *IEEE Transactions on Smart Grid*, pp. 1-12, January, 2016.
- [32] L. Yao, W. H. Lim and T. S. Tsai, "A Real-Time Charging Scheme for Demand Response in Electric Vehicle Parking Station," *IEEE Transactions on Smart Grid*, pp. 1-11, January, 2017.
- [33] J. H. Lim, "Optimal Combination and Sizing of a New and Renewable Hybrid Generation System," *Future Generation Communication and Networking*, Vol. 5, No. 2, December, 2013.
- [34] W. Gu, Z. Wu, R. Bo, W. Liu, G. Zhou, W. Chen and Z. Wu, "Modeling, planning and optimal energy management of combined cooling, heating and power microgrid: A review," *International Journal of Electrical Power & Energy Systems*, Vol. 54, pp. 26-37, January, 2014.
- [35] S. Diaf, D. Diaf, M. Belhamel, M. Haddad, A. Louche, "A methodology for optimal sizing of autonomous hybrid PV/wind system," *Energy Policy*, Vol. 35, No. 11, pp. 5708-5718, November, 2007.
- [36] T. Leonor, M. J. Humberto, "Power demand impacts of the charging of EVs on the power distribution network in a residential area," in *Proceedings of the 2011 3rd International Youth Conference on Energetic (IYCE)*, pp. 1-6, July 2011.
- [37] J. Tant, S. Member, F. Geth, S. Member, D. Six, P. Tant, J. Driesen, S. Member, "Multiobjective Battery Storage to Improve PV Integration in Residential Distribution Grids," *IEEE Transactions on Sustainable Energy*, Vol. 4, No. 1, pp. 182–191, January, 2013.
- [38] R. Christie, "power system test case archive," [online]. Available: <http://www.ee.washington.edu/research/pstea>.
- [39] P. Mesari'c, S. Krajcar, "Home demand side management integrated with EVs and renewable energy sources," *Energy and Buildings*, Vol. 108, pp. 1-9, December, 2015.
- [40] *Requirements for Voltage Characteristics in Public Distribution Systems Standard EN 50160*, 2014.
- [41] G. Li and X. P. Zhang, "Modeling of Plug-in Hybrid Electric Vehicle Charging Demand in Probabilistic Power Flow Calculations," *IEEE Transactions on Smart Grid*, Vol. 3, No. 1, pp. 492-499, March, 2012.

Cancer Model Simulation in Simulink Environment for Educational Purposes of Cancer Drug Dose Control

Matin Kordbacheh¹ and Ali Maleki^{2*}

Abstract— Cancer is one of the leading causes of death worldwide and the third cause of death in Iran after cardiovascular diseases and driving incidents. Therefore, having models that can describe and explain the process of cancer treatment is vital. One novel method for cancer treatment is to combine chemotherapy and immunotherapy. The purpose of immunotherapy is to enhance the body's immune system and reduce chemotherapy's side effects. Different models have been presented. The significance of the models is to understand the effect of each element on the treatment process. This article concerns the implementation of chemotherapy-immunotherapy cancer treatment in the Simulink toolbox of Matlab. Following the implementation of the model, the impact of chemotherapy, immunotherapy, and the combined technique on cancer cell growth is assessed. Simulating with Simulink allows for a graphical explanation as well as the ability to explain the hierarchy and effect of each piece. As a result, it may be advantageous for educational purposes. Simulink Toolboxes such as optimization, signal processing, and system identification allow us to analyze and control the cancer model.

Index Terms— Tumor cell model, chemotherapy, immunotherapy, Simulink, Simulation

I. INTRODUCTION

It is common to categorize cancers based on the cell tissue from which they have been generated. Generally, cancers are categorized into five categories. These classes are carcinoma, sarcoma, lymphoma, leukemia, and blastoma. Cancers occur due to the uncontrolled proliferation of cells and could exist in all parts of the body, from hard tissue such as bones to nervous tissue.

Most cancers do not have a particular symptom and most symptoms, do not appear in the same way from patient to patient. Many factors, such as substances containing arsenic, air pollutants, and some viruses, are involved in the existence of this disease. Cancer is a disease that disrupts the cellular disposition and results in interference with key and vital genes. This molecular disorder has a detrimental effect on cellular division and results in a lack of cellular differentiation [1, 2, 3]. There are various mathematical models regarding cancer therapy processes in the literature. Pillis and his colleagues [4]

have suggested a model that involves Cancer cells, three types of defensive cells and two drug concentrations. In a continuation on his efforts Pillis has utilized empirical data to enhance the values of the parameters in his model [5]. Kim and colleagues proposed a model with cancer cells, regulatory cells, natural killer cells, and two types of immune cells (CD4T cells & CD8T cells). There are two types of cancer genes worth mentioning. The first are oncogene genes, which have a positive effect on tumor formation. The second type of gene is the tumor suppression gene, which has a negative effect on tumor formation. Cancer can be cured in a variety of ways. Radiotherapy is one of the most common ways that cancer cells are targeted by X-ray. Surgery is another common treatment option for immature tumors (in clinical terms, have not yet metastasized). Chemotherapy is a method of cancer treatment that uses various chemical substances in order to eliminate cancer cells. Although it may have unfavorable side effects, it has proven to be very effective in the majority of cases. Using chemotherapy may have multiple purposes, such as cancer treatment without the presence of other treatment methods or utilizing chemotherapy after specific therapies, such as surgery, to assist in the treatment process. Other purposes may include reducing pain and preparing the patient for further stages of the overall therapy. Different chemotherapy drugs have different side effects, but the key thing to note is that the injected drugs also have a negative effect on the immune system in addition to their positive effect on eliminating cancer cells. In order to neutralize this negative effect, immunotherapy is used. Immunotherapy utilizes the immune system of the individual to recover from cancer. The goal of immunotherapy is to boost the immune system of the patient against cancer cells by increasing its efficiency. Immunotherapy has three subcategories. Immune response modifiers, monoclonal antibodies, and vaccines. Immune response modifiers target the interleukins, which also include IL-2. Monoclonal antibodies target certain cancerous antigens. Monoclonal antibodies can differentiate between healthy and cancerous cells. Vaccines, which are produced by cancerous cells, assist the immune system in the identification of cancer cells [7]. Modeling is the correct portrayal of a system's behavior. We can examine and simulate the system under investigation using a model. Because the chosen model

1-the Electrical and computer engineering faculty, Semnan university, Semnan, Iran

2- the Biomedical engineering department, Semnan university, Semnan, Iran
Corresponding author: amaleki@semnan.ac.ir

is heavily dependent on our demands, the system model is chosen depending on our control objectives. One of the most well-versed and well-rounded models is that of Pinho, which was published in 2013[8]. The model presented in [9] is based on the model of [7] which has more detail for cellular interactions and chemo-immunotherapy representation. Therefore, the assumptions of both papers are similar. In [10], a more improved and enhanced model is presented that utilizes interleukins to model cancer growth and its treatment. Naderi [11] has designed a controller to regulate the injected drug dosages based on the model presented in [6]. Sotolongo-Costa [12] proposed a model to represent the interaction dynamics of lymphocytes and tumor cells. They use cytokines alone as their periodical immunotherapy treatment and explain the dynamics by examining the interaction between the immune system and hostile cells. Robertson-Tessi [13] has also proposed a model for the interaction of tumor cells and the immune system. In order to govern the equations and parameters, experimental and clinical outcomes are incorporated. This model includes three types of T cells: dendritic cells, tumor cells, and plasma cells. The main purpose of this model is to find optimal antigenicity for maximum immune system response. It is stated that as antigenicity increases, immune cell access to tumor cells will be limited due to an increase in immunosuppression. Roberto A. Ku-Carrillo [14] has evaluated the effect of obesity on cancer growth and the immune system response. He does so by consolidating four differential equations containing the density of Immune cells, density of Cancer cells and the density of normal cells and the density of Fat. Denise Kirschner [15] utilizes a four-dimensional model that illustrates the dynamics between tumor cells, effector cells, and IL-2. Svetlana Bunimovich-Mendrazitsky [16] presents a system of nonlinear ODE's to Bladder Cancer and utilizes mouse and human in

vitro data to verify the model. In regards to experimental and clinical studies, Alexandre Boissonnas [17] has evaluated the effect of cytotoxic CD8+ T-cells on the elimination of tumor cells through in vivo imaging. It is shown that the CD8+ T- cells migrate, and when they come into contact with tumor cells, their associated antigen is dispersed. It is seen that CD8+ T- cells continue their movement if they come across dead tumor cells. By including multiple state variables and evaluating the dynamical uncertainty of the system, Kim presents a complex model for cancer. This model represents the cancer treatment process by having four classifications, with three of those classes being related to immune system cells and one relating to cancer cells. The amount of drug concentration in the bloodstream is also considered. Model [6] includes a system of differential equations containing eleven ordinary differential equations (ODEs). We have used model [6] as the basis of our study. The purpose of this paper is to implement Kim's comprehensive model into the Simulink toolbox of MATLAB, thereby simulating and evaluating the model in different circumstances and in a more thorough manner. Because of the high level of abstraction, Simulink simulation allows for a graphical description as well as the ability to illustrate the hierarchy and effect of each element. As a result, it can be beneficial for educational purposes.

II. RESEARCH METHODOLOGY

The model used in this article is a hybrid of chemotherapy and immunotherapy. This model, which is the one presented in [6], is comprehensive and has great potential for simulating different treatment methods. This model has eleven ODEs. The state variables are introduced in Table I. In addition, the differential equations are presented.

TABLE I
State Variables and Their Definition

variable	type	Definition	Unit
T	Cancer cell	cancer cell	number
N	Immune system	natural immune response (Natural Killer cell-NK)	number
E_T		Adaptive Immune response(CD8+T)	number
D_U		Unlicensed dendritic cells	number
D_L		Licensed dendritic cells	number
H_T		CD4+T cells	number
I_2		IL-2	IU/L
G_T	Immune	Regulatory cells	number
S	Suppressive	TGF- β	IU/L
I_10		IL-10	IU/L
M	Chemotherapy	Chemotherapy drug concentration	IU/L

$$\frac{dT}{dt} = aT(1 - bT) - c_1TN - \frac{dTE_T}{E_T + eT} \frac{1}{\left(1 + \left(\frac{G_T}{E_T}\right)\right)\left(1 + \left(\frac{S}{S_1}\right)\right)} - k_T(1 - e^{-M})T \quad (1)$$

$$\frac{dN}{dt} = b_1 - d_NN - c_2TN + \frac{P_NNI_2}{q_N + I_2} - k_N(1 - e^{-M})N \quad (2)$$

$$\frac{dE_T}{dt} = \frac{\alpha_1 I_2 D_L m_E}{\left(1 + \left(\frac{S}{S_1}\right)\right)(i_1 + I_2)(d_1 + D_L)} - c_3 E_T T - K_{E_T}(1 - e^{-M})E_T + w_1 u_E(t) \quad (3)$$

$$\frac{dD_U}{dt} = \frac{pT}{\left(1 + \left(\frac{I_{10}}{i_2}\right)\right)\left(\left(1 + \left(\frac{G_T}{g_2}\right)\right)\right)} - \frac{\gamma_1 D_U}{1 + \left(\frac{D_U}{m_H}\right)} - d_{D_U}D_U - K_{D_U}(1 - e^{-M})D_U \quad (4)$$

$$\frac{dD_L}{dt} = \frac{\gamma_1 D_U}{1 + \left(\frac{D_U}{m_H}\right)} - d_{D_L}D_L \quad (5)$$

$$\frac{dH_T}{dt} = \frac{\alpha_2 I_2 (D_L + D_U) m_H}{\left(1 + \left(\frac{S}{S_2}\right)\right)(i_1 + I_2)(d_1 + D_L + D_U)} - \frac{\gamma_2 H_T S}{S + s_3} - d_{H_T}H_T - k_{H_T}(1 - e^{-M})H_T \quad (6)$$

$$\frac{dG_T}{dt} = \frac{\gamma_2 H_T S}{S + s_3} + \frac{\alpha_3 I_2 D_L m_G}{(i_1 + I_2)(d_1 + D_L)} - d_{G_T}G_T - K_{G_T}(1 - e^{-M})G_T \quad (7)$$

$$\frac{dS}{dt} = p_1 G_T + p_2 T - d_s S \quad (8)$$

$$\frac{dI_2}{dt} = \frac{\alpha_4 H_T}{\left(1 + \left(\frac{I_{10}}{i_3}\right)\right)\left(\left(1 + \left(\frac{S}{S_4}\right)\right)\right)} - d_{I_2}I_2 + w_2 u_{I_2}(t) \quad (9)$$

$$\frac{dI_{10}}{dt} = p_3 G_T + p_4 T - d_{I_{10}}I_{10} \quad (10)$$

$$\frac{dM}{dt} = -d_M M + w_3 u_M(t) \quad (11)$$

The treatment process and eliminating cancer cells is as follow: Natural killer cells naturally attempt to oppose cancer cells. To enhance the treatment process to eliminate cancer cells, chemotherapy and immunotherapy are utilized. CD4 T cells and CD8 T cells are produced during chemotherapy and immunotherapy drug injection. CD4T & CD8T cells secrete IL-2, which reinforces one another. In addition, dendritic cells (DC) assist CD4T, CD8T, and NK cells in identifying cancer cells. DC cells are divided into two categories: licensed and unlicensed. The crucial element to remember is that unlicensed dendritic cells are incapable of detecting cancer cells. Another important point to consider is the unfavorable consequences of chemotherapy, which result in the creation of regulatory cells. Regulatory cells reduce the effectiveness of the immune system. As illustrated in Fig. 1, these cells release IL-10 and

TGF-B, resulting in a decrease in the efficacy of the immune system. Table II displays the model parameters and the amount of each parameter. Table III also shows the model's initial settings.

Fig. 1 shows the implementation of the model in the Simulink environment. It can be seen that the model has three control inputs for cancer control. The goal in cancer control is to eliminate as many cancer cells as possible by utilizing the least amount of Chemo and immune therapies. In Fig. 2, the next implemented level in Simulink is presented, which shows eleven state variables in the frame of eleven distinct blocks. The relationships between the blocks are also visible. Fig. 3-Fig. 13 shows the details of the differential equations for each state variable at the next level.

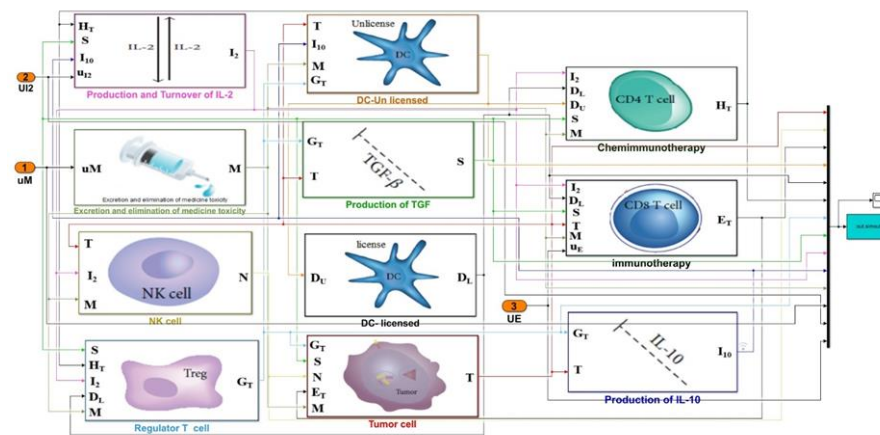


Fig. 2. Schematic of the block diagram of the implemented cancer model in the Simulink environment (this is the lower level to Figure 1) each block represents a state variable of each model. Number of cancer cells (T), number of Natural Killer cells (N), number of CD8T cells (E), number of CD4T cells (H), the amount of IL-2 discharge (I2), number of regulatory cells T (G), the amount of B-TGF discharge (S), the amount of IL-10 discharge (I10), number of licensed dendritic cells (DL), number of unsilenced dendritic cells (DU) and the amount of drug concentration in the blood stream (M) are represented.

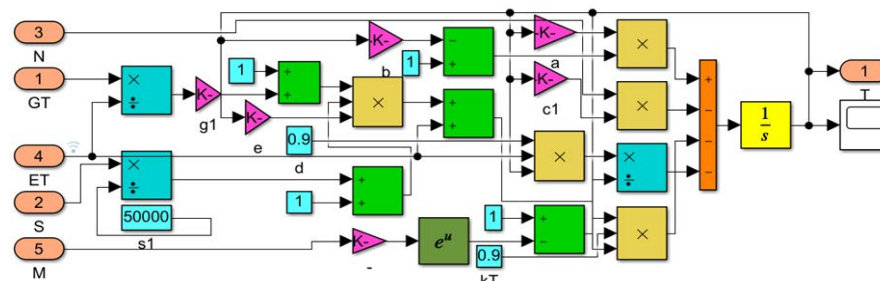


Fig. 3. Implementation of the differential equation of the “number of cancer cell” state variables in the third Level of the Simulink environment.

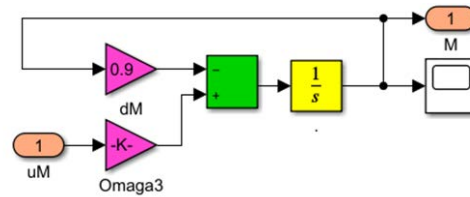


Fig. 4. Implementation of the differential equation of the “Chemotherapy drug concentration” state variable in the Simulink environment.

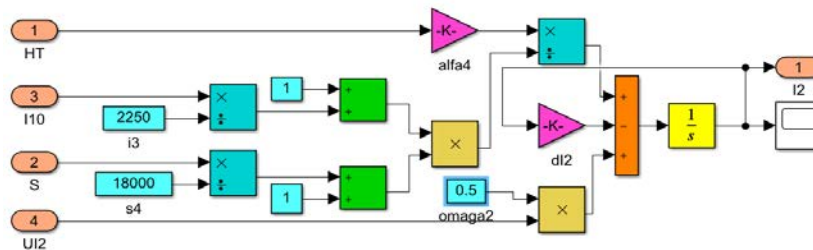


Fig. 5. Implementation of the differential equation of the amount of discharge of IL-2 in the Simulink environment.

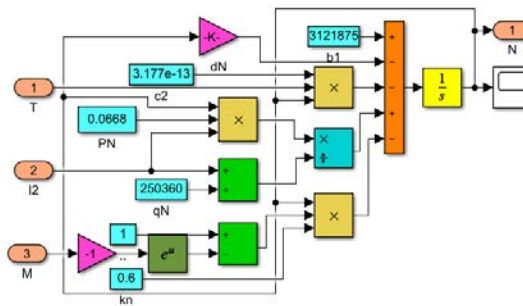


Fig. 6. Implementation of the differential equation of “number of Natural Killer cells” in the Simulink environment.

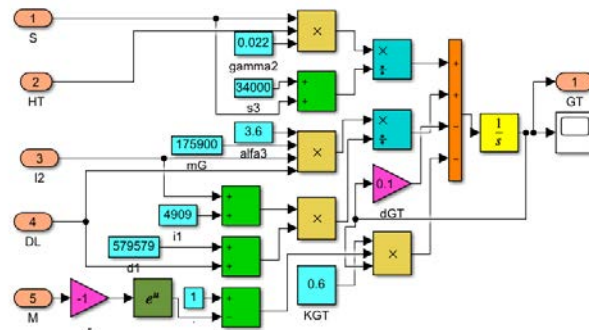


Fig. 7. Implementation of the differential equation of “number of Regulatory T cells” in the Simulink environment

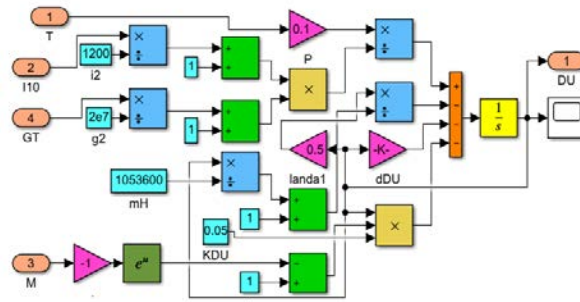


Fig. 8. Implementation of the differential equation of “Unlicensed Dendritic cells” in the Simulink environment.

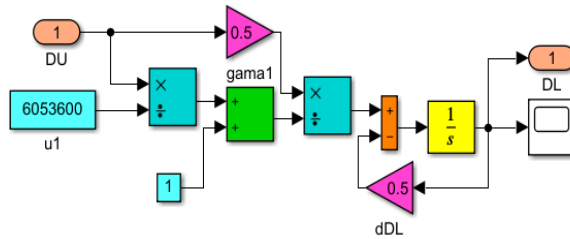


Fig. 9. Implementation of the differential equation of “licensed Dendritic cells” in the Simulink environment.

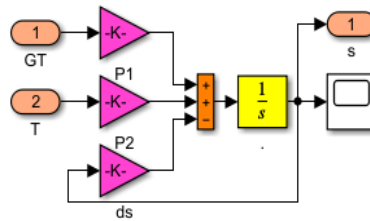


Fig. 10. Implementation of the differential equation of “TGF-β” discharge in the Simulink environment.

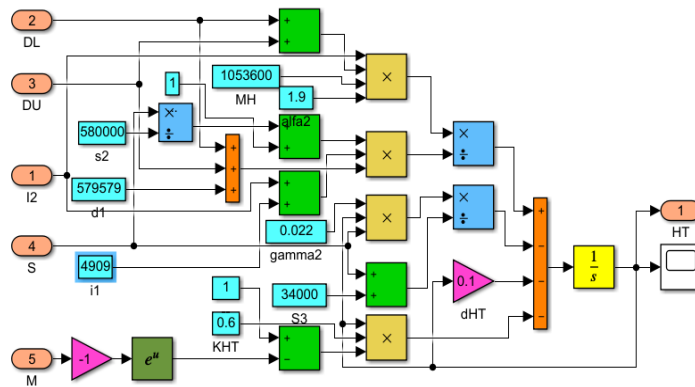


Fig. 11. Implementation of the differential equation of “number of CD4T cells” in the Simulink environment.

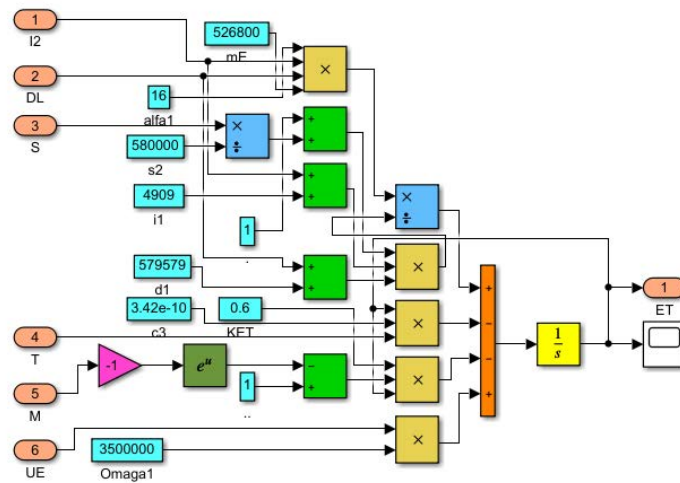


Fig. 12. Implementation of the differential equation of “number of CD8T cells” in the Simulink environment.

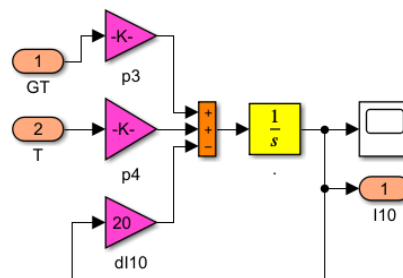


Fig. 13. Implementation of the differential equation of the amount of discharge of IL-10 in the Simulink environment.

III. SIMULATION RESULTS

By simulating the implemented model, the dynamical changes of the eleven state variables of the model are evaluated. Figure 14 depicts the simulation results for the events of no therapeutic intervention and only chemotherapy over the course of fifty days. The results show that the number of cancer cells in the absence of medicine becomes ten times larger than its original size in the span of ten days. Only chemotherapy with the recommended drug dosage reduces the number of cancer cells in that time frame. Figure 15 shows the effect of the changes in drug dosage (u_M) for chemotherapy on eliminating cancer cells. With the increase in chemotherapy drug dosage, the rate of cancer cell elimination increases. This effect is such that in the absence of u_M the number of cancer cells increases and, after a while, remains constant, but with the increase in drug dosage, the uphill trend of the number of cancer cells is reduced. The reduction of the number of cancer cells is reduced in such a way that at $u_M=1e-7$ the ultimate number of cancer cells decreases and by further increasing the amount of dosage. This downhill trend continues until cancer cells do not have the opportunity to increase from the very beginning. This is evident at $u_M=50e-7$ Figure 16 shows the effect of Immunotherapy drug dosage on the number of cancer cells. With the increase of the drug dosage of immunotherapy, the rate at which the number of cancer cells increases also experiences a decline such that before $u_E=10$ the effect of the Immunotherapy drug dosage is

not apparent, and the number of cancer cells reaches their maximum after a few days. Even so, after $u_E=10$ as we increase the amount of drug dosage not only does the rate of cancer cell growth decline but also the Maximum number of cancer cells also decreases. As shown in Figure 16 as u_E reaches the amount of 10.9 cancer cell growth is limited and eventually reaches zero. By increasing the amount of drug dosage further and reaching $u_E=50$, one can see the decrease in cancer cells from the very beginning of the treatment. In Figure 17, three different therapeutic methods with determined amounts of drug dosages are compared to each other and the state of no medicine in relation to the efficacy of cancer cell elimination. Immunotherapy has a significant effect on cancer cell elimination in such a way that in only fifteen days it stops Figure 16 shows the effect of Immunotherapy drug dosage on the number of cancer cells. With the increase of the drug dosage of immunotherapy, the rate at which the number of cancer cells increases also experiences a decline such that before $u_E=10$ the effect of the Immunotherapy drug dosage is not apparent, and the number of cancer cells reaches their maximum after a few days. Even so, after $u_E=10$ as we increase the amount of drug dosage not only does the rate of cancer cell growth decline but also the Maximum number of cancer cells also decreases. As shown in Figure 16 as u_E reaches the amount of 10.9 cancer cell growth is limited and eventually reaches zero. By increasing the amount of drug dosage further and reaching $u_E=50$, one can

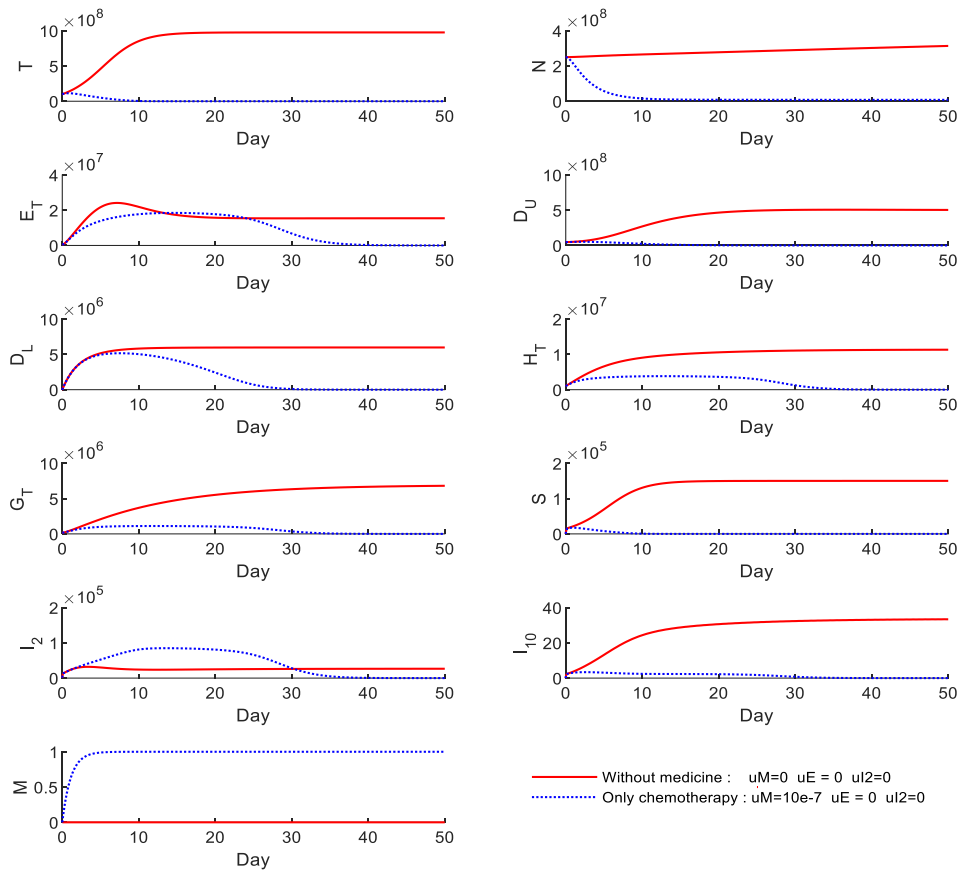


Fig. 14. State variable changes in the case of no medicine and only chemotherapy

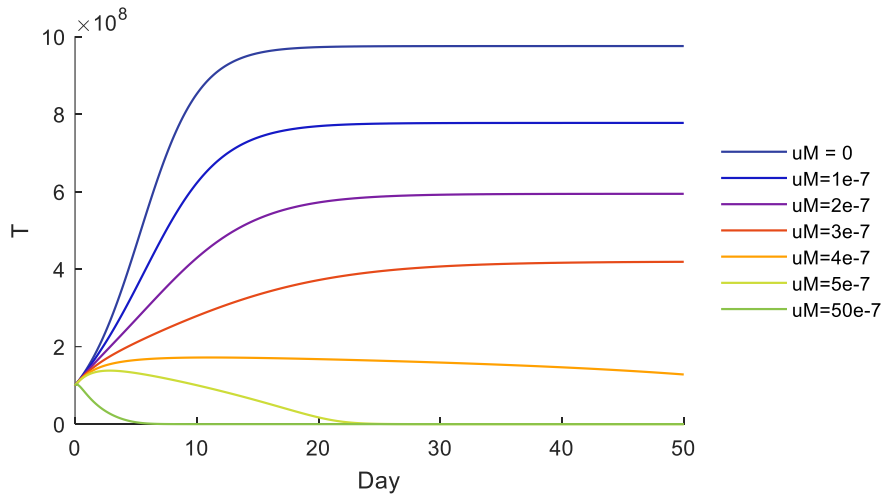


Fig. 15. The effect of Chemotherapy drug dosage (uM) on cancer cells elimination

see the decrease in cancer cells from the very beginning of the treatment. In Figure 17, three different therapeutic methods with determined amounts of drug dosages are compared to each other and the state of no medicine in relation to the efficacy of cancer cell elimination. Immunotherapy has a significant effect on cancer cell elimination in such a way that in only fifteen days it stops cancer cell growth and compared to no medicine the maximum amount of cancer cells is much lower. In immunotherapy, the number of cancer cells converges to zero only after 35 days. chemotherapy is even much more effective,

that is it stops cancer cell growth very quickly and in a short span of time and causes a decrease in the number of cancer cells. The time needed for chemotherapy to eliminate cancer cells is quite long, such that after fifty days of treatment, there are still a significant number of cancer cells. Combining chemotherapy and immunotherapy with the same amount of dosage will give better results in the elimination of cancer cells. In the combined chemo-immunotherapy treatment, cancer cell growth is prevented from the very beginning and quickly results in cancer cell elimination, such that after ten days the number

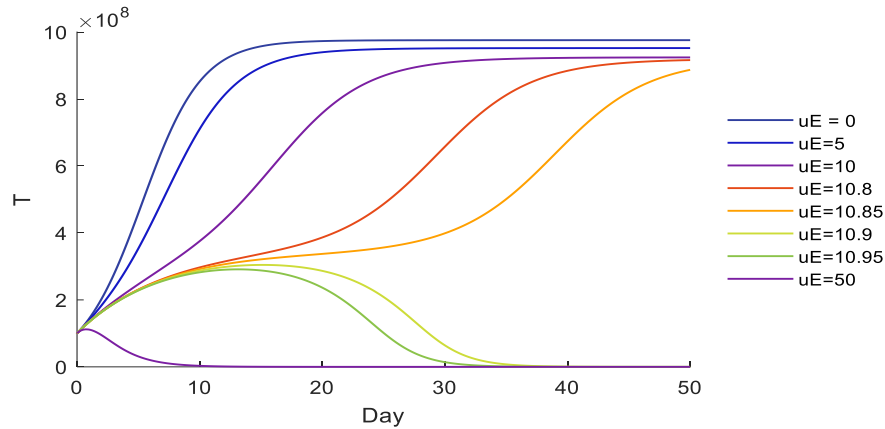


Fig. 16. The effect of changes in Immunotherapy drug dosage (uE) on cancer cell elimination

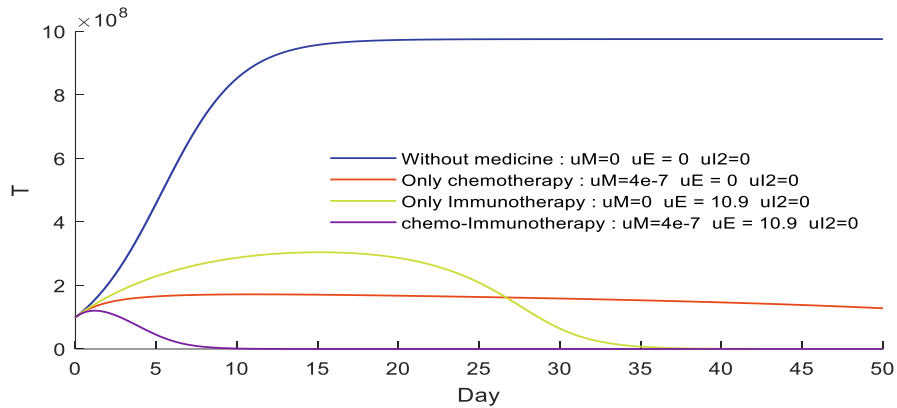


Fig. 17. Comparing the efficacy of three different treatment methods (only Chemotherapy, Only Immunotherapy, Chemo-Immunotherapy) in contrast to without medicine on cancer cell elimination

of cancer cells has converged to zero. The key point to note here is that these results are based on a particular amount of drug dosage. Therefore, utilizing different drug dosages may have different results.

The drug dosage change of Chemotherapy and Immunotherapy on different state variables is evaluated. Figure 18 shows the change in “the number of regulatory T cells” based on the increase of Immunotherapy drug dosage, against the amount of time (days) spent from the beginning of the treatment. In the absence of the Immunotherapy drug, “the number of regulatory T cells” reaches their maximum number of 10×10^8 in ten days. With the increase in the immunotherapy drug dosage, the maximum number of regulatory T cells decreases slightly, simultaneously the number of days it takes to reach this maximum increases. This trend can be seen up to $uE=10.85$ where it takes fifty days for the regulatory T cells to reach their maximum number. From this dosage on and as we increase the dosage, a strong downhill trend is apparent in which the number of regulatory T cells will eventually converge to zero. The

amount of dosage at $uE=50$ has a more desirable effect compared to other dosages, such that from the start, the increase in the number of regulatory T cells is prevented and converges to zero after 10 days. Figure 19 illustrates the change in the number of regulatory T cells based on the increase in chemotherapy drug dosage and the amount of time (days) spent since the beginning of the treatment. In the absence of the immunotherapy drug, “the number of regulatory T cells” increases. With the increase in drug dosage, the rate of increase in the number of regulatory T cells declines, such that by injecting a small amount of chemotherapy $uM=1e-7$, the number of regulatory T cells decreases from the 7×10^6 reach in the absence of chemotherapy to 4×10^6 . By continuing to increase the chemotherapy drug dosage, the increase in the number of regulatory T cells also continues to decline. In the maximum amount of drug dosage, $uM=50e-7$, the increase of the number of regulatory T cells is very limited and converges to a small value after thirty days.

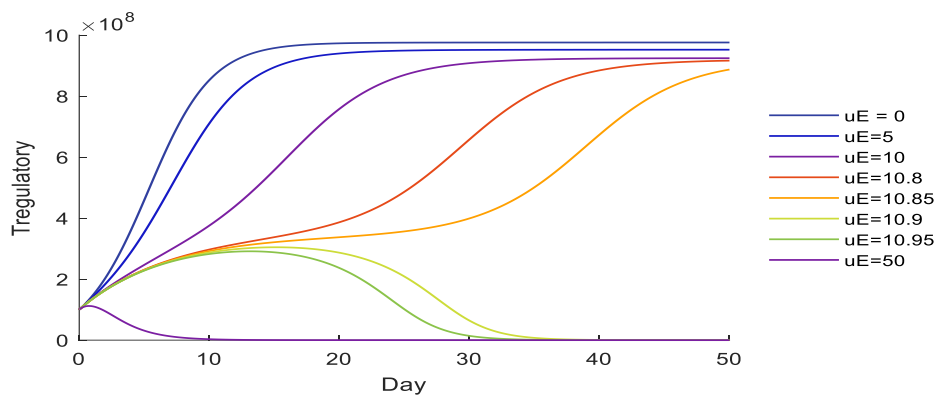


Fig .18. The change of the number of regulatory T cells based on the increase of the Immunotherapy drug dosage against time.

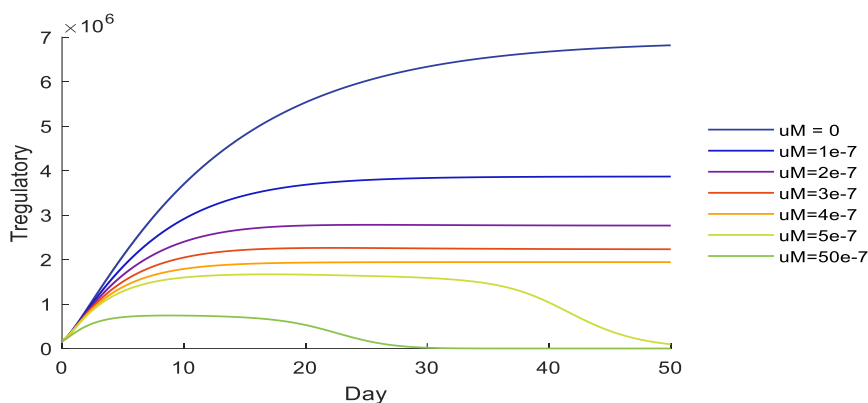


Fig .19. The change of the number of regulatory T cells based on the increase of the Chemotherapy drug dosage against time.

Figure 20 illustrates the change in CD4T cells which decreases with the increase in the chemotherapy drug dosage as the uphill trend of the number of CD4T cells decreases. In the absence of any drug dosage and after fifty days, the number of CD4T cells will reach approximately to 12×10^6 . By injecting a dosage of $uM=1e-7$ of the chemotherapy drug the number of CD4T cells will reach than 8×10^6 after fifty days. In addition, by increasing of the drug dosage to $uM=50e-7$ the maximum of number of CD4T cells will only be 2×10^6 and will converge to zero after thirty days. Figure 21 shows the change in CD8T cells. The number of CD8T cells increases with the increase in the chemotherapy drug dosage, such that in the absence of the drug dosage, after some fluctuation, the number of CD8T cells converges to 1.5×10^7 and by increasing the drug dosage, this convergence will be at a higher number. At $uM=4e-7$ the

number of CD8T cells will not even have a fluctuation phase and will continuously increase. However, by increasing the Chemotherapy drug dosage further, the uphill trend suddenly stops and the number of CD8T cells will start decreasing such that at $uM=5e-7$ after thirty days, the number of CD8T cells will reach its maximum value, but will immediately have a strong downhill trend and after fifty days will have an approximate 0.5×10^7 cells. Eventually, at $M=50e-7$ the upward trend only takes ten days, and in only thirty days, the CD8T cells will converge to zero. This abrupt alternation in the number of CD8T cells with the increase of chemotherapy might be interpreted as one of the side effects of chemotherapy meaning, by further increasing the amount of chemotherapy from a certain dosage, a decrease and eventual elimination of other influential cells is expected.

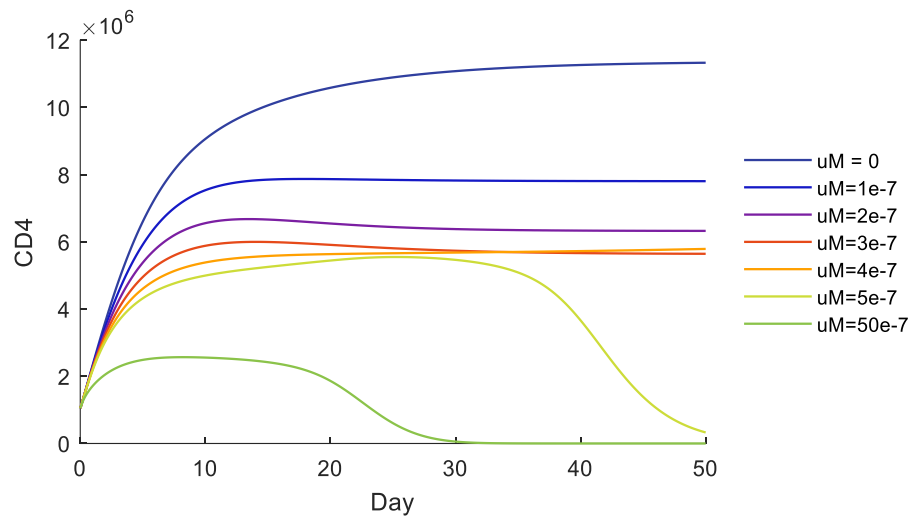


Fig. 20. The change of the number of CD4T cells based on the increase of the Chemotherapy drug dosage against time

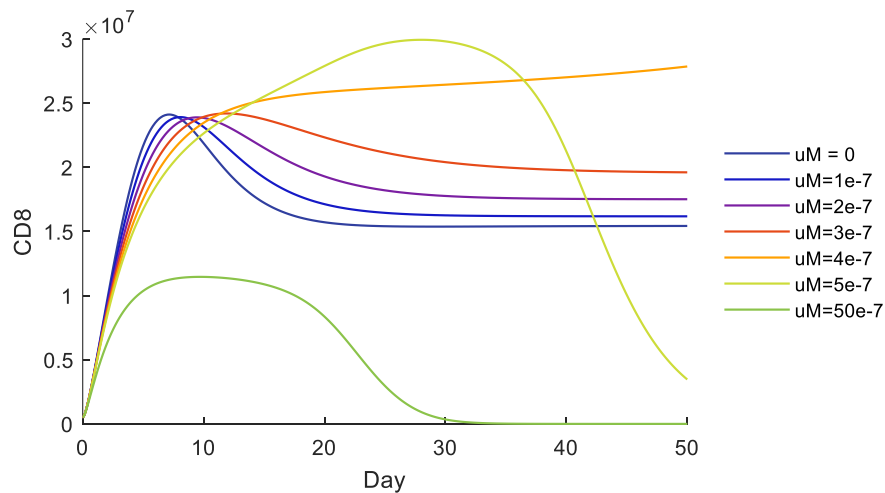


Fig. 21. Change in the CD8T cells, based on the increase of the Chemotherapy drug dosage against time

IV. DISCUSSION

A model of chemo-immunotherapy was simulated. Simulating this model in the graphical environment of Simulink provides us with valuable information regarding the evaluation of cancer growth and the treatment process. Due to the graphical description of the model in Simulink, understanding each element and their relations to one another is much simpler, and one can easily observe the effects of changes in each element on the model as a whole. The description of the model is done in a hierarchical manner, which makes it possible for individuals with different levels of knowledge and experience to connect with the model.

By simulating the cancer model in different circumstances, including solely chemotherapy, solely immunotherapy, and the combination method, it is seen that the combination method is much more effective. In future works, the efficacy of different controllers with the purpose of reducing cancer cells in the least amount of time and with the least amount of Chemo-Immunotherapy will be studied. From an educational standpoint, the knowledge acquired from the simulation can be utilized for other cancer models and other dynamical models in the fields of engineering and medical sciences.

V. REFERENCES

- [1] Sonnenschein C and Soto AM, "Theories of carcinogenesis: an emerging perspective", *Seminars in cancer biology*, 8(5): 372-377, 2008 Oct 1.
- [2] Pakin DM, "The global health burden of infection-associated cancers in the years 2002", *International Journal of Cancer*; 118(12):3030-44, 2006.
- [3] Beir VI, "Health risks from exposure to low levels of ionizing radiation", *The National Academies report in brief*. 2005.
- [4] De Pillis LG and Radunskaya A, "A mathematical model of immune response to tumor invasion", *Computational fluid and solid mechanics*, 1661-1668, 2003.
- [5] De Pillis L and Renee Fister K and Gu W and Collins C and Daub M and Gross D and Moore J and Preskill B, "Mathematical model creation for cancer chemo-immunotherapy", *Computational and Mathematical Methods in Medicine*, 10(3):165-84, 2009.
- [6] Kim KS and Cho G and Jung IH, "Optimal treatment strategy for a tumor model under immune suppression"

- Computational and mathematical methods in medicine, 2014.
- [7] De Pillis LG and Gu W and Radunskaya AE, “Mixed immunotherapy and chemotherapy of tumors: modeling, applications and biological interpretations”, *Journal of theoretical biology*; 238(4):841-62, 2006.
- [8] Pinho ST and Bacelar FS and Andrade RF and Freedman HI, “A mathematical model for the effect of anti-angiogenic therapy in the treatment of cancer tumours by chemotherapy”, *Nonlinear Analysis: Real World Applications*; 14(1):815-828, 2013.
- [9] Adam JA and Bellomo N. “A survey of models for tumor-immune system dynamics”, Springer Science & Business Media; 2012.
- [10] Usman A and Cunningham C, “Application of the mathematical model of tumor-immune interactions for IL-2 Adoptive Immunotherapy to studies on patients with Metastatic Melanoma or Renal Cell Cancer”, *Rose-Hulman Undergraduate Mathematics Journal*, 6(2):9, 2005.
- [11] Naderi H and Mehrabi M and Ahmadian MT, “Adaptive fuzzy controller design of drug dosage using optimal trajectories in a chemoimmunotherapy cancer treatment model”, *Informatics in Medicine Unlocked*; 27:100782. 2021.
- [12] otolongo-Costa O, Molina LM, Perez DR, Antoranz JC, Reyes MC. Behavior of tumors under nonstationary therapy. *Physica D: Nonlinear Phenomena.*;178(3-4):242-53, 2003 Apr 15.
- [13] Robertson-Tessi M, El-Kareh A, Goriely A. A mathematical model of tumor-immune interactions. *Journal of theoretical biology*; 294:56-73. 2012 Feb 7.
- [14] Roberto A. Ku-Carrillo, Sandra E. Delgadillo, B.M. Chen-Charpentier, A mathematical model for the effect of obesity on cancer growth and on the immune system response, *Applied Mathematical Modelling* (2015)
- [15] Kirschner D, Panetta JC. Modeling immunotherapy of the tumor-immune interaction. *Journal of mathematical biology*; 37(3):235-52. 1998 Sep.
- [16] Bunimovich-Mendrazitsky S, Shochat E, Stone L. Mathematical model of BCG immunotherapy in superficial bladder cancer. *Bulletin of Mathematical Biology*;69(6):1847-70.2007 Aug.
- [17] Boissonnas A, Fetler L, Zeelenberg IS, Hugues S, Amigorena S. In vivo imaging of cytotoxic T cell infiltration and elimination of a solid tumor. *The Journal of experimental medicine*; 204(2):345-56, 2007 Feb 19.