Journal of Applied Dynamic Systems and Control

Vol.7, No.1, 2024: 1-64

Address of Editorial office:
Department of Engineering, Aliabad Katoul Branch, Islamic Azad University, Aliabad Katoul , Iran
Tel:0098-17-34239511
Fax:0098-17-34235141
Website:https://sanad.iau.ir/journal/jadsc

Journal of Applied Dynamic Systems and Control, Vol. 7, No.1, 2024: 1-64

Contents:

Using Magic Square Chaotic Algorithm and DNA for Evolutionary-based Image Encryption Operators

Mahdi Tahbaz[1],Hossein Shirgahi[2*], Mohammad Reza Yamaghani[1]

Abstract – The development of digital technologies has improved the transfer of data over the Inte rnet in recent years. Image encryption is a technique to ensure security in information transfers. The current paper presents an evolutionary model on the basis of a hybridization of DNA biomolecule o perators and the LS2 Map chaos function for encryption of image. The model proposed here include s three stages. In the initial stage, the MSC (Magic Square Chaotic) algorithm and a secret key are utilized with the SHA-256 algorithm to determine the initiating the LS2 Map function value, which i s then employed to manipulate the pixels of the image. Then, DNA biomolecule operators and the c haos function are used for propagation. Additionally, the previous stages process is iterated with the starting population of the genetic algorithm in the third stage. Afterward, the optimization is carried out through genetic algorithm operators. The results indicate that the introduced model is superior to other rivals. Furthermore, as for the high level of entropy obtained, the model exhibits strong resista nce to common attacks.

Keywords: LS2 map, DNA operators, Magic square algorithm, Image encryption, Genetic algorithm, hash function.

1. Introduction

Every day, individuals use audio, images, and video to communicated through the media Specifically, information contained in images can be easily accessed via wireless networks and the Internet. Thus, ensuring the security of digital image content is essential to prevent illegal viewing, copying, or editing [23, 27]. There are several methods and algorithms proposed for image encryption [3]. Cryptography can be non-textual or textual [21]. Given the data volume (to many pixels), encryption of image is different significantly from texts. As a result, text and image protection are completely different and the methods for encrypting texts may not be used for encryption of images [4,18]. The systems designed for image encryption mostly used chaotic and turbulent functions because of the distinctive characteristics of chaotic maps, like random chaotic sequence values, excessive sensitivity of initial values, and simplicity of the maps. Such methods essentially aim to divide pixels of image into smaller units (bits) and encryption is carried out based on operators like XOR. Therefore, the encryption cannot be recognized unless the encryption key is given. In addition, the proposed algorithms feature unique specifications. That is, some of them have a higher sensitivity to the starting condition, while others have sensitivity to other parameters [5, 10-13]. The systems for cryptography rely on diffusion and permutation of pixels and creating chaotic patterns. Furthermore, because of the encryption keys with large spaces, chaotic systems can stand exhaustive search attacks. In general, researchers primarily utilize complex chaotic systems or integrate novel methods of encryption and available chaotic systems to improve the security of cryptographic algorithms for image encryption [24-26]. There are two different methods to encrypt images including diffusion and permutation. In the latter, pixels position is changed using matrix transformation or chaotic sequences like checkerboard transformations. In these algorithms keep the value of pixels unchanged while the pixel position is altered. Therefore, these algorithms are vulnerable to statistical analysis. The value of pixels is changed by diffusion phase based on a chaotic sequence. In comparison to permutation technique, diffusion provides a higher level of security; however, its efficiency is lower in terms of cryptographic impact. Therefore, these two techniques are combined to enhance effectiveness and security of cryptography [2,16].

Recently, the properties and structural complexity of biological molecules such as RNA and DNA have been extensively discussed in the literature. There is a higher

1 Department of Computer Engineering, Lahijan Branch, Islamic Azad University, Lahijan, Iran.

Email: mahdi_tahbaz@liau.ac.ir, o_yamaghani@liau.ac.ir

2* Corresponding Author : Department of Computer Engineering, Jouybar Branch, Islamic Azad University, Jouybar, Iran.

Email: h.shirgahi@jouybariau.ac.ir

Received: 08.08.2023 ; Accepted:06.03.2024

efficiency in biomolecule encryption techniques for emerging security applications with several levels as for security and performance. As a result, a mixture of DNA sequences and chaotic systems is used for encryption of images with a higher level of security. Thus, cryptography based on DNA is a desirable supplement to the standard mathematical cryptography [11, 14, 17]. There have been many studies on utilizing evolutionary algorithms and chaotic functions in encryption of images. The chaotic functions are mostly utilized to generation starting encryption images; while evolutionary algorithm and general solutions are used to achieve a higher level of quality of solutions [6, 19]. Abbasi et al. [1, 7] encrypted images using a novel evolutionary chaotic model using the specifications of biological molecules. The introduced method demonstrated strong reliability against normal attacks thanks to 256-bit concealed key and incorporating the chaotic function's dependency on the key to produce sequences that are quasi-random, along with the integration of evolutionary algorithms and biomolecule operators.

A new model was introduced that utilizes genetic algorithms to increase entropy values and attenuates correlation coefficients in image pixels by employing DNA operators. The goal is to enhance the diffusion criterion through mixing magical square algorithm and the LS2 Map chaos function, ultimately increasing the level of turbulence to address drawbacks of image encryption. The remaining of the article is arranged as follows: the fundamental concepts of chaotic functions, the magic square algorithm, biological molecules, hash functions, and genetic algorithms are introduced in Section two; Section 3 provides a more detailed focus on the model. The results of simulations for the proposed model are given in Section 4 along with comparing them with those of previous models. Eventually, section five concludes the paper with conclusion remarks.

2. Basic Concepts

The basics of chaotic functions, magic square algorithms, biological molecules, hash functions, and genetic algorithms are discussed as follows:

2.1 Chaos Function

The term "chaos" explains the complicated function of dynamic systems. It indeed is a behavior exhibited by these systems, offering rapid and highly secure methods for steganography and encryption. It boasts unique benefits, including sensitivity to initial parameters and conditions, unpredictability, and random behavior [16,26].

2.1.1. Logistic-Sinusoidal Mapping

As one of the main mappings in chaos theory, the logistic mapping (Eq.1) is characterized by complex chaotic behavior and simple equations in its output. In addition, the sine mapping (Eq. 2) has a highly similar behavior to the logistic mapping. Failure to generation consistently chaotic output in specific intervals is one issue with these maps, and they also exhibit inefficient chaotic behavior within their chaotic interval. A novel approach known as LS Map was introduced to deal with this drawback (Eq. 3). By incorporating an additional operation, altering the output using the mod operation, and giving it back to the interval, the authors enhanced the chaotic performance of the two maps [0, 1]. Still, according to the results, we do not have a uniform mapping. Hence, the mapping is not a suitable option for encryption, given that the maximum intervals of distribution are easily determined through a simple statistical analysis. As illustrated Figure 1, it is possible to have a uniform output utilizing XOR operator in series (Eq. 4). The output of this result is further combined with two sinusoidal and logistic maps and then constrained to the interval [0,1] (LS2 Map) [8, 28].

$$X\_(n+1) = r.X\_n\,(1 - X\_n) \quad (1)$$

$$(X\_(n+1) = (4 - r)\ .sin\,(\pi.X\_n)) / 4 \quad (2)$$

$$X\_(n+1) = (r.X\_n\,(1 - X\_n) + ((4 - r)\ .\sin\,(\pi.X\_n)) / 4) mod\ 1 \quad (3)$$

Where $r \in (0..4]$

$$X\_(n+1) = (((r.X_n(1 - X_n) \oplus ((4 - r)\ .\sin\,(\pi.X\_n)))) / 4) + (r.X\_n\,(1 - X\_n) + ((4 - r)\ .\sin\,(\pi.X\_n)) / 4)) mod\ 1 \quad (4)$$

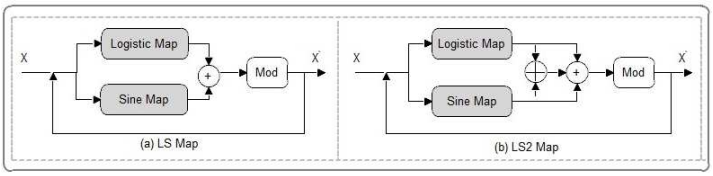


Fig. 1. (a) LS Map and (b) LS2 Map

2.2 Biological Molecules

Three categories of biological molecules are used by all forms of complex life on Earth to carry out essential functions. Proteins have various function in cells like providing structure, acting as chemical transporters, and promoting catalytic functions. Meanwhile, DNA and RNA contain genetic information, which is transferred to next generations [17,31].

2.2.1. DNA

DNA (Deoxyribonucleic Acid) is the chemical compound that contains all the genetic information and hereditary traits of living organisms. This molecule contains two very long strands that coil and form a double-helix structure. DNA is found in all cells of living organisms and is passed from parent cells to offspring. Based on Figure 2, nitrogenous organic bases have a circular structure and exist in four forms within the DNA molecule: guanine (G), cytosine (C), adenine (A), and thymine (T).

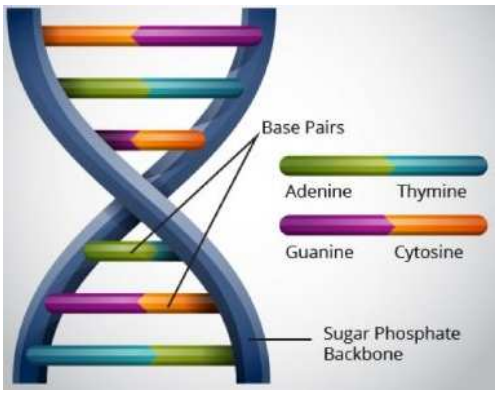


Fig. 2. Structure of DNA molecules

2.2.2. Image encryption and decoding by DNA

Adelman conducted the first DNA computational experiment in 1967. Furthermore, he developed a novel method in the field of molecular computing to address combinatorial problems related to information principles. The emergence of DNA computing marked the formation of a new field that utilized DNA sequences as carriers of information and as implementation tools.

Bases A and T on one hand and G and C on the other hand are complementary to each other. As 0 and 1 are complementary, 00 and 11 will also be complementary. Similarly, 01 and 10 are complementary and have the same relationship. There are 24 different encryption modes achievable utilizing the four combinations. Nevertheless, considering the complementary nature of the A-T and G-C base pair rules, only the four modes presented in Table 1 will be valid [17, 20].

As shown in Table 1, each element in a row or column is repeated only once. Various operators, like subtraction, addition, XNOR, and XOR, are utilized in image encryption according to Table 2.

As illustrated in Table 1, each element in a row or column is repeated only once. Different operators including addition and subtraction, XOR and XNOR are used in image encryption based on Table 2.

Table 1. Rules of DNA sequences operators for decoding and encryption

	A	T	C	G
Rule 1	00	11	10	01
Rule 2	00	11	01	10
Rule 3	11	00	10	01
Rule 4	11	00	01	10
Rule 5	10	01	00	11
Rule 6	01	10	00	11
Rule 7	10	01	11	00
Rule 8	01	10	11	00

Table 2. XOR, Subtraction, addition, and XNOR operations for DNA operators.

	+				-				XOR				XNOR			
	A	G	C	T	A	G	C	T	A	G	C	T	A	G	C	T
A	A	G	C	T	A	T	C	G	A	G	C	T	T	C	G	A
G	G	C	T	A	G	A	T	C	G	A	T	C	C	T	A	G
C	C	T	A	G	C	G	A	T	C	T	A	G	G	A	T	C
T	T	A	G	C	T	C	G	A	T	C	G	A	A	G	C	T

2.3 Magic Square Algorithm

There are n rows and n columns in a magic square, containing n^2 cells filled with unique natural numbers (see Eqs. 5 and 6).

$$1 + 2 + 3 + \cdots + n^2 = n.M \qquad (5)$$

$$\sum_{i=1}^{n^2} i = n.M. \qquad (6)$$

M refers to the total numbers on columns, rows, or principal diagonal of the magic square of order n, so that the total value of figures in the magic square is equal n*M (Eq. 7).

$$n.M = \frac{n^2(n^2+1)}{2} \qquad (7)$$

The total value of all columns, rows, and diagonals of a magic square is fixed. The figures in each cell of a magic square of order n includes all figures between 1 to n^2. The magic square is known as n^2 order magic square. Here, the total of the columns, rows, and diagonals of this square is calculated based on Equation 8.

$$M = \frac{n(n^2+1)}{2} \qquad (8)$$

The 8x8 magic square are used in the model proposed here. Thus the total value of each column, row, and

diagonal of the magic square is equal to eight, which equals (8^2 + 1)/2 = 260. The magic square 8x8 arrangement is: first, counting is started from the cell in the top left corner, and the numbers given to the non-diagonal cells are added to them (black numbers in Figure 3). Afterward, the count in the next row at the end of each row is restarted from the cells on the left. Then, after going through all non-diagonal cells, counting begins at the right side of the row below and the number dedicated to the diagonal cells are inserted (figures in blue and red color in Figure 3). Then, we start counting at the right to left, and with each row completed, it restarts in the top row from the right side cell.

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64

→

64	2	3	61	60	6	7	57
9	55	54	12	13	51	50	16
17	47	46	20	21	43	42	24
40	26	27	37	36	30	31	33
32	34	35	29	28	38	39	25
41	23	22	44	45	19	18	48
49	15	14	52	53	11	10	56
8	58	59	5	4	62	63	1

Fig. 3. The 8*8 magic square left (before moving) and right (after moving)

2.4. Hash function

A hash is a mathematical function that converts any input value to another compressed value. The input of the hash function is a value of unknown length; however, the output is always of a constant length. Hash functions are widely-used and are used in almost all information security applications [8].

2.4.1. Encryption hash function

Cryptographic hash function with its own unique features is used for security and authentication applications. A cryptographic hash function is:

1- Definite and Constant

Each specified input is converted to a fixed output by the hash function. In this process, it does not matter how many times or when the hash function affects the input, in any case the output is fixed. If this is not the case and the value of the hash function is different in each effect, data retrieval will be impossible.

2 - Calculation speed

Calculation speed of the input hash is a key factor for the efficiency of the hash function.

3- One-way function

The cryptographic hash function is a one-way relationship, it means that the output can be easily calculated by hashing. However, an output corresponding to an input is a barrier to obtaining the input.

2.4.2. Secure Hash Algorithm (SHA-256)

The SHA-256 algorithm is characterized by one of the most secure hash functions and is one of the branches of SHA-2 created by the National Security Agency in 2001 as a replacement for SHA-1. SHA-256 is a 256-bit encrypted hash function, with three features which make the algorithm highly secure:

a) Restructuring raw data from a hash value is almost not possible. An attack to create raw data requires 2^{256} attempts!

b) It is very unlikely that two messages with the same amount of hashes (so-called collisions) are possible. With 2^{256} possible hash values (which is more than the total number of known atoms in the universe), the probability that two numbers are the same is infinitely unimaginable.

c) A trivial change in the original data alters the hash value so that it is unclear whether the new hash value was taken from the same data – i.e. avalanche effect.

2.5. Genetic Algorithm

As a specialized type of evolutionary algorithm, a genetic algorithm employs biological techniques like mutation and inheritance. To achieve the best formula to predict or match the patterns, genetic algorithms (GA) actually utilize the idea of natural selection as introduced by Darwin. They are frequently a viable choice for regression-based prediction methods. In artificial intelligence, a GA is a method for programming that employs genetic evolution for solving problems. The problem involves inputs that are transformed into solutions through a simulated process of genetic evolution. Subsequently, the Fitness Function assesses the solutions as candidates. Furthermore, with an exit condition, the algorithm will terminate. The GA is typically an iteration-based algorithm in which most components are selected through random processes [19].

3. Proposed Model

The introduced model contains four phases. At the beginning, a 256-bit secret key is processed based on the SHA-256 algorithm to obtain the starting value for the LS2map function. Then, in the next phase (permutation), the encryption process uses chaos sequence and magic algorithm to enhance security. Afterward, in the diffusion phase, the chaotic sequence and the DNA sequence operators are utilized to alter the gray surface on the image pixels. In the fourth phase, based on a genetic algorithm, the optimization process is performed following the encryption.

Phase 1: Secret key generation

To improve the security against attacks, A 256-bit key

was utilized. To create pseudo-random figures using the LS2map function, we needed a starting value (x0), obtained using a 256-bit key. Using the input data (image) the SHA-256 algorithm produces a unique 256-bit hash key, through this, a direct link is created that connects the input data and the generated key (image). In other word, a unique identifier is assigned to each input data.

As depicted in Figure 4, to generate secrete key, the proposed algorithm involves the initial conversion of the input image to a 64-character hexadecimal figure based on the SHA-256. Subsequently, to improve the level of security, using SHA-256, the same 64-character string is further changed into a new 64-character hexadecimal number. Afterward, the new secret key is turned into binary form. For instance, with an image of Rayan, it is first converted to "7d23f64f7759bfe39acc056b0bc7c85aa4d46bf5fefee15b54c416526b9d02cc" using the SHA-256 algorithm.

Then, it is further converted to "6445a9943240fe983d058158a2b07e4bf3aad8b842914c18ba9f9b48cd9930dc" before being transformed into binary form. The obtained key id demonstrated as "bi," where "bi" indicates the ith bit of the key. The 256 bits are divided into 32 8-bit blocks (Eq. 9).

$$\text{Key}=\{K\_1,K\_2,\ldots,K\_32\} \tag{9}$$

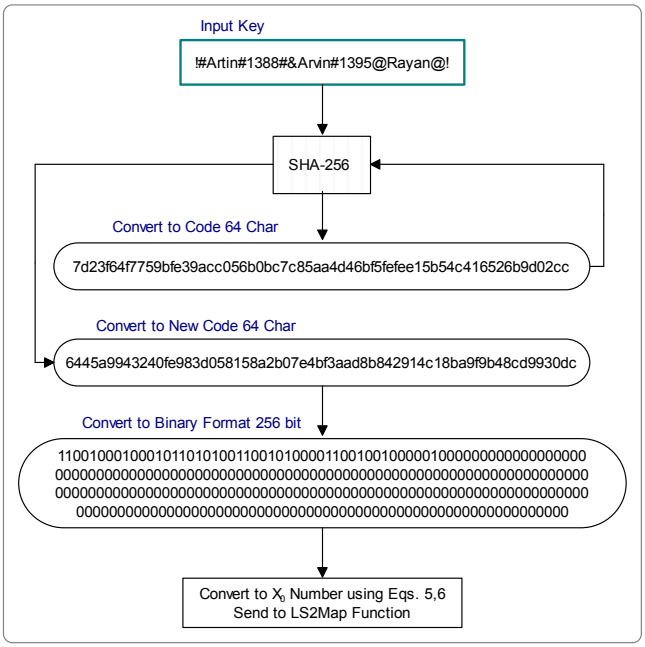


Fig.4. Generating the secret key

Formula 10 is employed to obtain x0. Which ⊕ represents the XOR operator.

$$X_0 = \left. \begin{matrix} \left(\sum_{i=1}^{8} K_i + (K_{17\oplus} \ldots \oplus K_{24})\right) + \\ \left(\sum_{i=9}^{16} K_i + (K_{25\oplus} \ldots \oplus K_{32})\right) \end{matrix} \middle/ 2^{12} \right. \tag{10}$$

Phase 2: permutation

Step 1. Level one blocking:

This step involves dividing the main image into blocks (64 to 64×64 pixel blocks for 512×512 images and 64 to 32×32 pixel blocks for 256×256 images).

Step 2: Move the level one blocks based on the magic square algorithm.

Section 3-2. Figure 6 displays the changes in the blocks of the original image.

Step 3. Level 2 Blocking:

The blocks dispositioned from step 2 are chosen through the LS2map chaos (Eq. 3) to achieve a random number (X_i) in the [1..64] based Eq. 11. The blocks are then is divided 8x8 blocks (64 to 8×8 pixel blocks 512x512 images and 16 to 8x8 pixel blocks 256x256 images).

$$\text{Select\_Block}=\text{Round}(X_i\times 64)+1 \tag{11}$$

Step 4. Blocks selected based on the rate of permutation:

A random number (X_i) is generated within the range of 4 to 8 for 256 × 256 images, and within the range of 16 to 32 for the same image size. The LS2 map function is utilized for selecting blocks for the rate of image pixel permutation (Permute_Rate).

$$\text{Select\_B\_256}=\text{Round}(X_{i+1} \times 8)+4 \tag{12}$$

$$\text{Select\_B\_512}=\text{Round}(X_{i+1} \times 32)+16 \tag{13}$$

Step 5. Moving the second level blocks using the magic square algorithm:

To rearrange the pixels inside the selected blocks, we used the magic square algorithm (see section 3-2).

Step 6: Repeat steps 4 and 5 for all selected blocks, based on the number of hash rates.

The difference lies in using the values of X_i and X_{i+1} for every image, starting from the last value created by the image earlier.

Phase 3: Diffusion

Step 1. Convert the image into a one-dimensional (1D) array:

As shown in Figure 4, at this stage, in order to convert the image obtained from the previous stage (permutation) into a 1D array, first, Row i+1 of the image is placed at the end of Row i, and this action is repeated for all the rows of the image. The 2D $N\times M$ array is converted into a 1D array with $N\times M$ elements in the following form:

$\{P_1, P_2, \ldots P_{N\times M}\}$. In addition, Equations 14 and 15 are used to calculate the number of the row and column of Pixel (P_i) in the 2D array.

$$\text{Row} = \left\lceil \frac{P_i}{N} \right\rceil \qquad (14)$$

$$Column = \begin{cases} (P_i \ mod \ M) & if \ (P_i \ mod \ M) \neq 0 \\ M & if \ (P_i \ mod \ M) = 0 \end{cases} \qquad (15)$$

Step 2. Conversion into a DNA sequence:

At this stage, the five sequential numbers [X_i, ..., X_{i+4}] of the chaotic function "LS2 map" are used to encrypt both pixels, in the following way: X_i is used to generate a random number in the interval [1..8] based on Equation 16 in order to select DNA rules from Table 1, X_{i+1} and X_{i+2} to generate a random number in the interval [1..256] based on Equation 17 in order to select two bits from the key to select the biomolecules operators (XOR, ADD), and X_{i+3} and X_{i+4} to generate two random numbers in the interval [0..255] based on Equation 18 in order to select two numbers from the chaotic function in order to apply the selected operator.

$$\text{Select}_{\text{Rule}} = \text{Round}(X_i \times 7) + 1 \qquad (16)$$

$$\text{Select}_{\text{Operator}} = \text{Round}(X_{i+1} \text{ and } X_{i+2} \times 255) + 1 \qquad (17)$$

$$\text{Select}_{\text{Number}} = \text{Round}(X_{i+3} \text{ and } X_{i+4} \times 255) \qquad (18)$$

Step 3. Replacement of pixels:

At each stage, two pixels are selected from the 1D array and converted into a DNA sequence using Equation 12 and Table 1.

As shown in Figure 6, two low-value nucleotides from the selected pixels are replaced by one another.

The two numbers, selected from the chaotic function based on Equation 14, are converted into a DNA sequence using Equation 12 and Table 1.

At this stage, the XOR operator is used to the two bits selected from the secret key. If the result of the XOR operator on the two bits selected from the key is zero, the XOR operator will be applied to the two pixels selected from the 1D array and the two numbers selected from the chaotic function. Otherwise, (if the result of the XOR operator on the two bits selected from the secret key is 1,) the XNOR operator will be applied to the two pixels selected from the 1D array and the two numbers selected from the chaotic function. The process is repeated for all the pixels of the image.

Step 4. The step three output will again be turned into a binary form based on Table 1, upon completion of the encryption process and then into a decimal form, and then into a 2D array using Equations 14 and 15. Figure 5 shows an example of the encryption process of two pixels selected from the image in the diffusion phase.

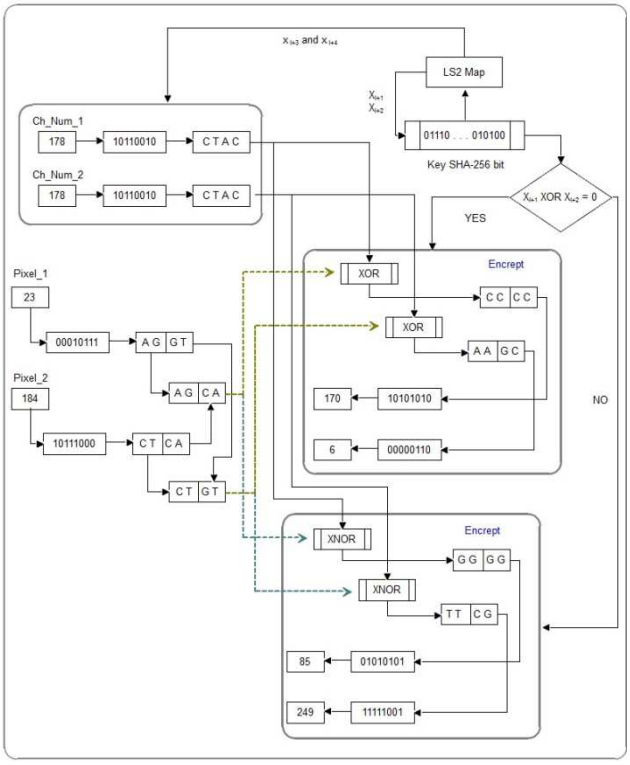


Fig. 5. The process of encryption two selected pixels of an image at the diffusion step

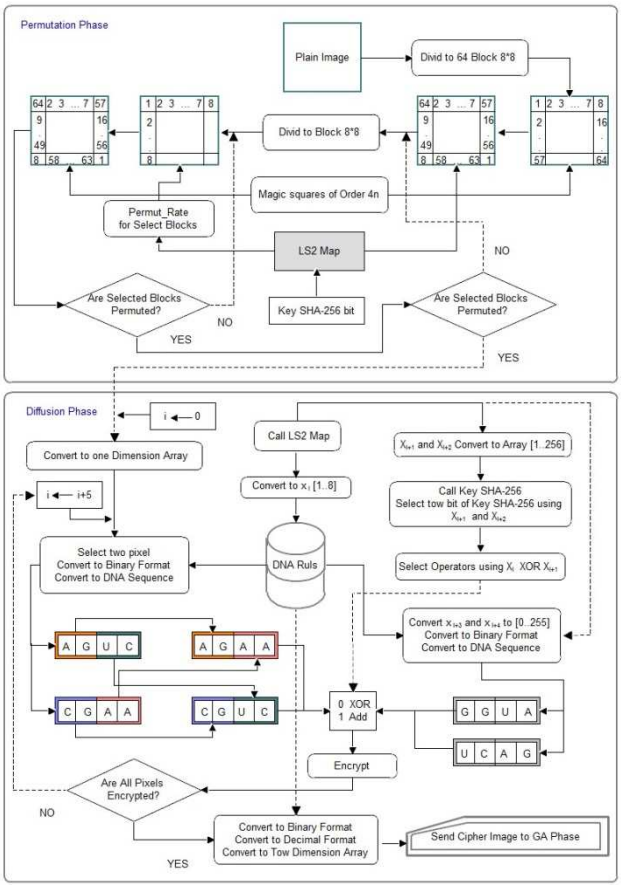


Fig. 6. Proposed Model Encryption Process

Phase 4: GA

Following diffusion and permutation process, and based on the genetic algorithm rules outlined earlier, the optimization takes place for achieved an image encrypted with greater resistance in comparison to the original one, taking into account the criterion of entropy criterion for the evaluation function:

a) Creation of the initial population:

Using 100 images, the starting population is formed and used in the diffusion. The distinction lies in using the chaotic function values, which are obtained from the chaotic function for each image. This process starts with the final value given by the image in the last step. In addition, each image is assigned with three integers depending on the duration of encryption.

b) Selection operator:

The operator indicates the chromosome that can survival in the next generation. That is, individuals with a superior chromosome have a higher survival chance. This actually happens in the process of evolution. Still, in some cases, using a mutation operator or crossover operator to a wrong chromosome may lead to the correct chromosome. Here, the top-quality image with the maximum entropy are adopted based on a roulette wheel to undergo crossover operator and mutations operator.

c) Crossover operator:

After encryption (as depicted in Figure 7), images are categorized into four equal sections following selection and arranged in a line side by side.

Here, a figure between 1 and 3 is selected through a random function.

When one is the obtained random number, the one-point crossover method is utilized. When two is the number randomly generated, the two-point crossover method is selected; and when three is the number randomly generated, the multi-point crossover method is employed. Moreover, with crossover operator employed on the image, three numbers are inserted in the three-byte encryption table. For the decryption and determine the images and operators utilized for image encryption this is essential.

d) Mutation operator:

Following the crossover operator, the mutation happens prior to the final evaluation in the search space. Actually, the mutation randomly finds the proper chromosome for change. Here, a mutation happens by random selection of a new encryption image to replace the original image using the operator.

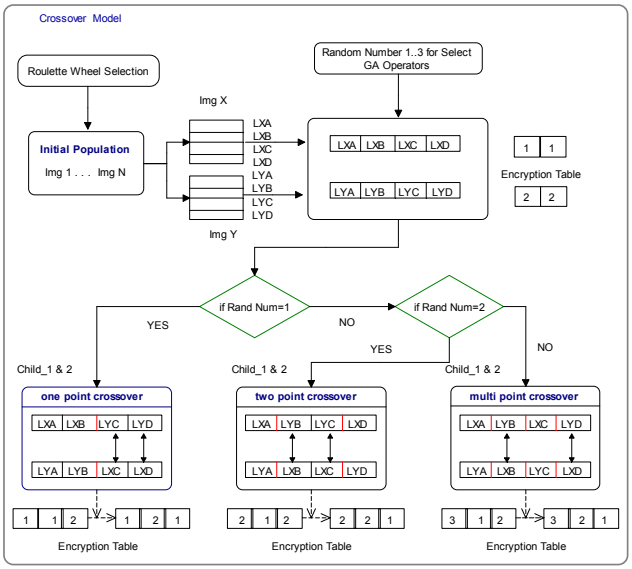


Fig. 7. Crossover of the proposed model

4. Simulation Results

Here, the performance and security of the introduced model are analyzed and evaluated through brute-force analysis, statistical analysis, information entropy analysis, differential attacks, and convergence analysis. The evaluations were carried out in MATLAB software (2.50 GHz Dual Core, 4GB memory, and Windows 7). The analysis was performed on different image with of 256 × 256 and 512 × 512 dimensions, which were developed from the USC-SIPI database based on Figure 8.

Fig. 8. Test image

4.1. Statistical Analysis

To be an efficient algorithm for encryption, the output must be resistive to statistical attacks. We generated the histogram of image to carry out statistical analysis. Additionally, using images retrieved the USC-SIPI database we analyzed adjacent vertical, horizontal, and diagonal pixels correlation.

4.1.1. Histogram Analysis

The pixel elements distribution in an image is illustrated

by image histogram. The algorithm of encryption must resist statistical attacks. Actually, with an increase in uniformity of the histography of the image of the encryption algorithm, the rate of statistical attacks decreases. The proposed model conducted analyses on the results of various experiments based on histogram analysis to check the statistical distributions. Figure 9 depicts the histograms of the images before employed and after employing the proposed model. Clearly, the histogram of the encrypted image appears uniform and is not identical to the histogram of the original image. This prevents hackers from accessing valuable information, reducing the likelihood of successful statistical attacks. This illustrates the effectiveness of the model proposed here.

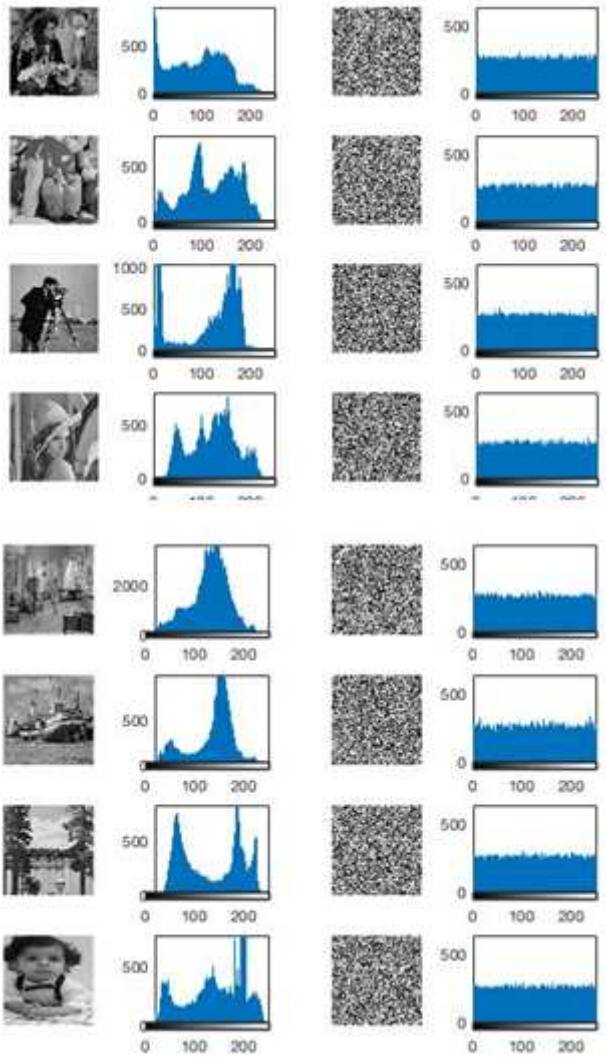


Fig. 9. From left to right in each row, simple image, simple image histogram, encrypted image and encrypted image histogram

4.1.2. Analysis of Correlation Coefficients

Because of the interdependence of adjacent pixels in digital images, a robust cryptographic techniques requires generating encrypted images that minimize the vertical, diagonal, and horizontal correlations among the pixels. As a statistical index, the correlation coefficient is utilized to measure the two variables relationship, returning a number within the range of -1 to 1. With a number closer to zero, the variations of the two variables become more independent, and as the figure merges to -1 or 1, the interdependence of variations increases. The coefficient of correlation is obtained as follows [7].

$$r_{xy}=\frac{|\mathrm{cov}(x,y)|}{\sqrt{D(x)}\times\sqrt{D(y)}} \qquad (19)$$

$$E(x)=\frac{1}{N}\sum_{i=1}^{N}x_i \qquad (20)$$

$$D(x)=\frac{1}{N}\sum_{i=1}^{N}(x_i-E(x))^2 \qquad (21)$$

$$\mathrm{cov}(x,y)=\frac{1}{N}\sum_{i=1}^{N}(x_i-E(x))(y_i-E(y)) \qquad (22)$$

where, Y and X represent the brightness of two pixels that are next to each other. In terms of similarity of pixel, the two adjacent pixels become more similar. When plot points are closer to the original diameter. The objective of cryptography to minimize the similarity of the encrypted images to decrease the likelihood of unauthorized image access by analyzing the resemblance among pixels.

The correlation of 8192 pairs of adjacent horizontal, vertical, and diagonal pixels extracted from the database of the image using Equation 19 is listed in Table 2.

Table 3. The coefficient of correlation of adjacent pixels (vertical, diagonal, and horizontal)

Correlation of pixels					
		Cameraman	Forest	Lena	Rayan
256 × 256	Horizontal	**-0.0098**	**-0.0096**	**-0.0178**	**-0.0106**
	Vertical	**-0.0003**	**0.0005**	**0.0005**	**-0.0107**
	Diagonal	**0.0085**	**0.0122**	**-0.0101**	**-0.0029**
512 × 512	Horizontal	**0.0130**	**-0.0191**	**-0.0045**	**0.0062**
	Vertical	**0.0101**	**0.0123**	**0.0008**	**-0.0057**
	Diagonal	**-0.0211**	**0.0118**	**0.0013**	**0.0010**
		peppers	Living room	Boat	Painter
256 × 256	Horizontal	**0.0043**	**0.0064**	**0.0090**	**0.0021**
	Vertical	**-0.0048**	**-0.0193**	**0.0174**	**0.0131**
	Diagonal	**0.0074**	**-0.0125**	**0.0025**	**-0.0019**
512 × 512	Horizontal	**0.0147**	**0.0040**	**0.0023**	**-0.0094**
	Vertical	**-0.0016**	**0.0168**	**-0.0045**	**0.0047**
	Diagonal	**0.0087**	**-0.0062**	**0.0062**	**0.0022**

The distribution of gray surface for the two diagonally, horizontally, and vertically adjacent pixels of the Lena image is shown in Figure 10. As depicted in Figure 10, the pixels are uniformly distributed and independence of the successive pixels is confirmed. Therefore, the findings

indicate that the original image differs from the encrypted image in terms of adjacent pixels.

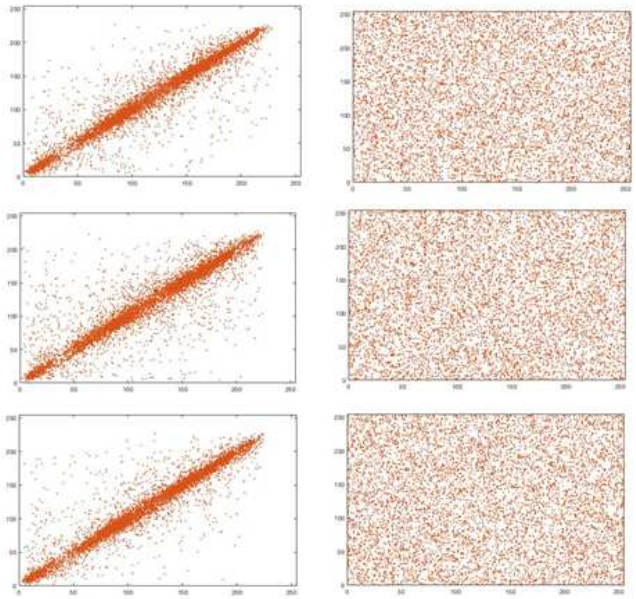


Fig. 10. Distribution of gray surface for two adjacent pixels (top to bottom) vertically, diagonally, and horizontally on the Lena image prior to (left diagram) and following encryption (right diagram)

4.2. Information Entropy

Shannon entropy, also known as information entropy, is a prominent characteristic of randomness. It is a mathematical theory for storing and transferring data, which was developed by Claude E. Shannon in 1949. Equation 23 is a well-known formulas for calculating entropy [1].

$$H(s)=\sum_{i=0}^{2^M-1} P(s_i)\log_2\frac{1}{P(s_i)} \qquad (23)$$

Where N refers to the number of gray levels in the image (256 in 8-bit images) and indicates the occurrence chance of the i-th gray level. In completely random images, this figure is eight, which is an ideal figure. The proximity of these figures 8 in the model proposed here indicates a negligible leakage of information in the image decoding. Therefore, the encryption model proposed here has a highly resistive to entropy attack. The encrypted images entropy is given in Table 4.

Table 4. Encrypted images entropy on image database

Table 4. Encrypted images entropy on image database

Image Size	Cameraman	Forest	Lena	Rayan
256 × 256	**7.9977**	**7.9976**	**7.9976**	**7.9970**
512 × 512	**7.9993**	**7.9994**	**7.9993**	**7.9993**
Image Size	Peppers	Living room	Boat	Painter
256 × 256	**7.9977**	**7.9976**	**7.9976**	**7.9970**
512 × 512	**7.9993**	**7.9994**	**7.9993**	**7.9993**

4.3. Brute-Force Attack

The sensitivity of the secret key and its state space size are examined to prevent brute-force attacks.

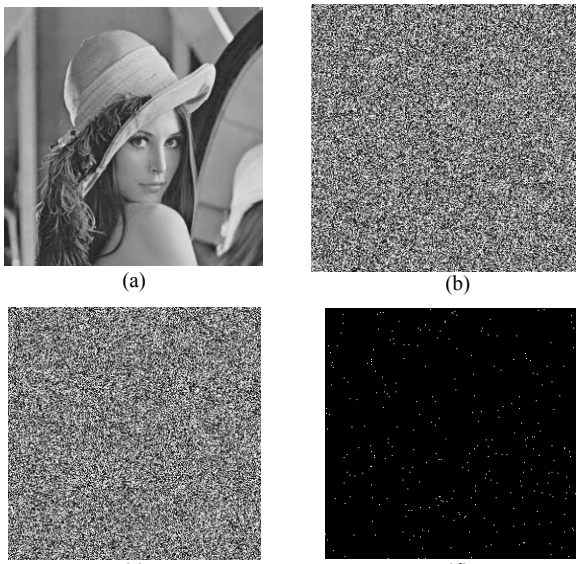
4.3.1. Secret Key Sensitivity Analysis

A main aspect of security of a robust encryption algorithm is the password sensitivity. The sensitivity level of the chaotic systems to the initial state and parameters is high. With a group of possible parameter states, selecting two closely related initial parameters can take the system into two diverse ways. Thus, chaotic systems are suitable candidates for use in cryptography. Therefore and to assess the proposed model sensitivity to the encryption key (Fig. 11), the Lena image (256 * 256) in Figure 11a was encrypted through a 256-bit hidden key "!#Artin#1388#&Arvin#1395@Rayan@!". (11b). Then, it is encrypted once more using the key "!#Artin#1388#&Brvin#1395@Rayan@!""; still, by randomly changing a bit (11 c). Figure 11d (highlighting the differences between images 11b and 11c) clearly illustrates that the two images are encoded with white pixels at the same gray level. As pictured in Figure 11d, the model has sensitivity to slightest changes in initial key, which ensures the model's resilience to a comprehensive search attack. Table 4 lists the findings of the analysis of the secret key applied to the database images.

4.3.2. Analysis of Secret Key Mode Space

The total keys usable in cryptography is given by the key state space. With a bigger key state space, the performance of encryption increases along with the resistance to cryptographic attacks like unrestrained attacks.

Here, a 256-bit key is employed as the initial value for the ls2map chaotic function, resulting in a key state space of 2^{256} states. Due to the big state space, the model proposed here demonstrated good resistance to brute force attacks.



(a) (b) (c) (d)

Fig. 11. a. image of Lena, b. encrypted image with a 256-bit secret key, c. encrypted image with the same key and by changing one bit, d. image differences between 10b and 10c

Table 5. Sensitivity to the secret key in the model on the image database

Image Size	Cameraman	Forest	Lena	Rayan
256 × 256	**99.58%**	**99.60%**	**99.61%**	**99.77%**
512 × 512	**99.62%**	**99.60%**	**99.62%**	**99.63%**
Image Size	Peppers	Living room	Boat	Painter
256 × 256	**99.62%**	**99.59%**	**99.66%**	**99.63%**
512 × 512	**99.62%**	**99.61%**	**99.61%**	**99.61%**

4.4. Differential Attack

In image encryption, the use of a differential attack is a significant and necessary criterion for evaluating the security level. In practice, a Differential Attack aim is to indicate if altering a pixel in the original image can influence image after encryption.

4.4.1. Evaluation of NPCR and UACI Benchmark

The influence of altering a single pixel of the original image after encryption can be assessed by two criteria: NPCR and UACI. NPCR is given as the pixel change rate in the image encrypted based on the change of one pixel in the original image. Additionally, UACI is the average of such changes, determined as follows [30].

$$NPCR=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} D(i,j)}{M\times N}\times 100\% \qquad (24)$$

$$UACI=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} |C_1(i,j)-C_2(i,j)|}{255\times M\times N}\times 100\% \qquad (25)$$

Where W refers to the width of the images and H refers to length of the images; additionally, I1 and I2 are two images encrypted on the basis of two images with one pixel difference; and D is obtained as follows:

$$D(i,j)=\begin{cases}1 & \text{if } C_1(i,j)=C_2(i,j)\\ 0 & \text{otherwise}\end{cases} \qquad (26)$$

The values based on Tables 6 and 7 clearly indicate that the model also has an acceptable resistance to differential attacks, meaning that a single pixel change in the input image will result in a significant variation in the output.

Table 6. NPCR NPCR outputs of the model on the images

Image Size		Cameraman	Forest	Lena	Rayan
256×256	Average	0.995962	0.996145	0.995962	0.996104
	Best	0.996078	0.996323	0.996094	0.996338
	Average	0.996174	0.996175	0.996038	0.996183
	Best	0.996197	0.996269	0.996098	0.996288
Image Size		peppers	Living room	Boat	Painter
256×256	Average	0.996084	0.996099	0.996043	0.996175
	Best	0.996292	0.996231	0.996216	0.996175
	Average	0.996086	0.996062	0.996091	0.996188
	Best	0.996208	0.996294	0.996223	0.996201

Table 7. UACI results of proposed model on image database

Image Size		Cameraman	Forest	Lena	Rayan
256×256	Average	0.333867	0.333893	0.333888	0.334350
	Best	0.334233	0.335576	0.335081	0.335352
	Average	0.334577	0.334666	0.334754	0.334694
	Best	0.335111	0.334830	0.335243	0.335185
Image Size		peppers	Living room	Boat	Painter
256×256	Average	0.334140	0.334459	0.335302	0.334343
	Best	0.334990	0.335960	0.336555	0.335406
	Average	0.334736	0.333551	0.334434	0.334265
	Best	0.335001	0.334656	0.334794	0.335435

4.5. Comparisons

Table 8 shows the model performance in comparison to [23, 27, 30, 1, 19, 6, 7, 28]. The value of entropy of the model can be improved significantly by incorporating the LS2 Map chaos function and integrating it with DNA biomolecule operators, and utilizing genetic algorithms. As listed in Table 8, the model demonstrates superior performance for the boat image compared to other

references and is competitive with a few references in terms of alternative criteria.

Table 8. The proposed model function compared to other methods

	Entropy	Correlation Coefficient			NPCR (%)	UACI (%)
		V	H	D		
Ref. [23]	7.9965	0.0074	0.0044	0.0081	0.995890	0.335002
Ref. [27]	7.9976	0.0015	0.0007	-0.0007	0.996200	0.336900
Ref. [30]	7.9986	0.0104	0.0167	0.0074	0.971029	0.328471
Ref. [1]	7.9992	0.0103	-0.0069	-0.0011	0.996240	0.334408
Ref. [19]	7.9986	0.0167	0.0104	0.0074	0.971029	0.328471
Ref. [6]	--	0.0215	0.0154	0.0081	--	--
Ref. [7]	7.9993	0.0408	0.0123	0.0040	0.996129	0.333711
Ref. [28]	7.9965	0.0074	0.0044	0.0081	0.995890	0.335002
Proposed Model	7.9994	-0.0045	0.0023	0.0062	0.996294	0.334656

5. Conclusion

An evolutionary model based on a combination of DNA biomolecule operators and the LS2 Map chaos function was proposed for image encryption. The chaos function dependence on the 256-bit secret key and the MSC algorithm in for permutation, combined with the utilization of DNA biomolecule operators and genetic algorithms, contribute to the advantages of the proposed model. These advantages include high entropy value, encrypted image histogram uniformity, a significant decline in the adjacent pixels correlation in the image of encryption, and a large key mode space. All these factors collectively support the high efficiency of the presented model in encrypted applications. Moreover, the model has demonstrated strong resilience against different attack types such as code deciphering attacks, statistical attacks, and brute-force attacks.

References

[1] Abbasi, A. A., Mazinani, M., & Hosseini, R. (2020). Chaotic evolutionary-based image ncryption using RNA codons and amino acid truth table. Optics & Laser Technology, 132, 106465.

[2] Enayatifar, R., Abdullah, A. H., Isnin, I. F., Altameem, A., & Lee, M. (2017). Image encryption using a synchronous permutation-diffusion technique. Optics and Lasers in Engineering, 90, 146-154.

[3] Gulshan, K., Praveen, P., Rahul, S., & Kumar, R. M. (2016). Chaotic image encryption technique based on IDEA and discrete wavelet transformation. Indian J. Sci. Technol, 9(15), 71871.

[4] Hu, T., Liu, Y., Gong, L.-H., & Ouyang, C.-J. (2017). An image encryption scheme combining chaos with cycle operation for DNA sequences. Nonlinear Dynamics, 87(1), 51-66.

[5] Ibrahim, Y., Khalifa, F., Mohamed, M. A., & Samrah, A. S. (2020). A New Image Encryption Scheme Based on Hybrid Chaotic Maps. Complexity, 2020.

[6] Abdullah, A. H., Enayatifar, R., & Lee, M. (2012). A hybrid genetic algorithm and chaotic function model for image encryption. AEU-International Journal of Electronics and Communications, 66(10), 806-816.

[7] Abbasi, A. A., Mazinani, M., & Hosseini, R. (2020). Evolutionary-based image encryption using biomolecules operators and non-coupled map lattice. Optik, 219, 164949.

[8] Chenaghlu, M. A., & Khasmakhi, N. N. A fast and secure keyed hash function based on coupled chaotic maps for crypto-currencies.

[9] Abdullah, D., Rahim, R., Siahaan, A. P. U., Ulva, A. F., Fitri, Z., Malahayati, M., & Harun, H. (2018). Super-encryption cryptography with IDEA and WAKE algorithm. Paper presented at the J. Phys. Conf. Ser.

[10] Luo, Y., Yu, J., Lai, W., & Liu, L. (2019). A novel chaotic image encryption algorithm based on improved baker map and logistic map. Multimedia Tools and Applications, 78(15), 22023-

[11] Al-Mashhadi, H. M., & Abduljaleel, I. Q. (2017). Color image encryption using chaotic maps, triangular scrambling, with DNA sequences. Paper presented at the 2017 International Conference on Current Research in Computer Science and Information Technology (ICCIT).

[12] Alabaichi, A. M. (2016). Color image encryption using 3D chaotic map with AES key dependent S-Box. International Journal of Computer Science and Network Security (IJCSNS), 16(10), 105-115.

[13] Ye, G., Pan, C., Huang, X., & Mei, Q. (2018). An efficient pixel-level chaotic image encryption algorithm. Nonlinear Dynamics, 94(1), 745-756.

[14] Yin, Q., Cao, B., Li, X., Wang, B., Zhang, Q., & Wei, X. (2020). An Intelligent Optimization Algorithm for Constructing a DNA Storage Code: NOL-HHO.International journal of molecular sciences,212191.

[15] Belazi, A., Abd El-Latif, A. A., Diaconu, A.-V., Rhouma, R., & Belghith, S. (2017). Chaos-based partial image encryption scheme based on linear fractional and lifting wavelet transforms. Optics and Lasers in Engineering, 88, 37-50.

[16] Li, Y., Wang, C., & Chen, H. (2017). A hyper-chaos-based image encryption algorithm using pixel-level permutation and bit-level permutation. Optics and Lasers in Engineering, 90, 238-246.

[17] Chai, X., Chen, Y., & Broyde, L. (2017). A novel chaos-based image encryption algorithm using DNA sequence operations. Optics and Lasers in Engineering, 88, 197-213.

[18] El-Zoghdy, S. F., Nada, Y. A., & Abdo, A. (2011). How good is the DES algorithm in image ciphering? International Journal of Advanced Networking and Applications, 2(5), 796-803.

[19] Enayatifar, R., Abdullah, A. H., & Isnin, I. F. (2014). Chaos-based image encryption using a hybrid genetic algorithm and a DNA sequence. Optics and Lasers in Engineering, 56, 83-93.

[20] Wang, X., Wang, Y., Zhu, X., & Luo, C. (2020). A novel chaotic algorithm for image encryption utilizing one-time pad based on pixel level and DNA level. Optics and Lasers in Engineering, 125, 105851.

[21] Joshy, A., Baby, K. A., Padma, S., & Fasila, K. (2017). Text to image encryption technique using RGB substitution and AES. Paper presented at the 2017 International Conference on Inventive Computing and Informatics (ICICI).

[22] Wang, Y., Lei, P., Yang, H., & Cao, H. (2015). Security analysis on a color image encryption based on DNA encoding and chaos map. Computers & Electrical Engineering, 46, 433-446. 22043.

[23] Tsafack, N., Sankar, S., Abd-El-Atty, B., Kengne, J., Jithin, K. C., Belazi, A., ... & Abd El-Latif, A. A. (2020). A new chaotic map with dynamic analysis and encryption application in internet of health things. IEEE Access, 8, 137731-137744.

[24] Som, S., Mitra, A., Palit, S., & Chaudhuri, B. (2019). A selective bitplane image encryption scheme using chaotic maps. Multimedia Tools and Applications, 78(8), 10373-10400.

[25] ur Rehman, A., Liao, X., Ashraf, R., Ullah, S., & Wang, H. (2018). A color image encryption technique using exclusive-OR with DNA complementary rules based on chaos theory and SHA-2. Optik, 159, 348-367.

[26] Wang, X., Feng, L., Li, R., & Zhang, F. (2019). A fast image encryption algorithm based on non-adjacent dynamically coupled map lattice model. Nonlinear Dynamics, 95(4), 2797-2824.

[27] Nestor, T., De Dieu, N. J., Jacques, K., Yves, E. J., Iliyasu, A. M., El-Latif, A., & Ahmed, A. (2020). A multidimensional hyperjerk oscillator: Dynamics analysis, analogue and embedded systems implementation, and its application as a cryptosystem. Sensors, 20(1), 83.

[28] Maqsood Mahmuda, Atta-ur-Rahmanb, Malrey Leec, Jae-Young Choi, Evolutionary-based image encryption using RNA codons truth table, Optics and Laser Technology 121 (2020) 105818.

[29] Yadollahi, M., Enayatifar, R., Nematzadeh, H., Lee, M., & Choi, J.-Y. (2020). A novel image security technique based on nucleic acid concepts. Journal of Information Security and Applications, 53, 102505.

[30] Nematzadeh, H., Enayatifar, R., Yadollahi, M., Lee, M., & Jeong, G. (2020). Binary search tree image encryption with DNA. Optik, 202, 163505.

[31] Tahbaz, M., Shirgahi, H., Yamaghani, M.R. (2023). Evolutionary-based image encryption using Magic Square Chaotic algorithm and RNA codons truth table. Multimedia Tools and Applications, 83, 503–526. https://doi.org/10.1007/s11042-023-15677-3

A Novel Method for Optimal Sensor and Actuator Placements: The "Infinite Value Algorithm"

Mahdi Zavar[1],Mohammad Shahraeini[2*], Alireza Safa[3],Niki Manouchehri[4]

Abstract–Cyber-physical systems rely heavily on the functionality of sensors and actuators to operate effectively. Sensors transfer measurements from the physical part to the control and cyber part and while actuators that play the role of applying control signals to the physical part. The placement of sensors and actuators is an important issue for ensuring system's observability and controllability. This placement should be done in way that the system remains controllable and observable with fewest number of sensors and actuators.. Minimizing the number of sensors and actuators has a significant effect on cost and energy reduction. In this study, a new approach is introduced based on the existing paths and non-existent paths between the states of a system. In the proposed approach, the non-existent paths are defined as infinite paths. The best nodes are selected as the location of sensors and actuators based on the infinite paths and their numbers.The numerical simulations illustrate that the Infinite Value Algorithm has performed better in the placements in this way that it has consumed the unique solution in less time compared the Genetic Algorithms.

Keywords:Infinite value algorithm, Sensor placement, Actuator placement, Observability, Controllability.

1. Introduction

Cyber-physical systems (CPSs) integrate physical components with cyber capabilities, revolutionizing industries like electrical infrastructure and communication systems, to create a networked environment [1]. These systems also incorporate the concept of system dynamics, which means that physics and dynamics are intertwined. CPSs and networks form a pervasive computing platform that underlies many modern technologies. They are created by integrating and interconnecting physical equipment and systems from an engineering perspective and different algorithms, enabling them to influence each other. Since CPSs aim to control the dynamics of physical systems, they can be considered dynamic systems. Examples include medical devices [2, 3], communication peripherals [4], intelligent vehicles [5], transportation networks [6, 7], power generation networks [8, 9], and water distribution systems [10]. CPS is closely related to buzzwords such as the Internet of Things, Industry 4.0, Industrial Internet, and Machine-to-Machine. All of these reflect a vision of modern technology that connects the physical world with the world of information.

Graph-based algorithms can be used to analyze and model switched-mode systems [11]. A recent study proposed a new approach using graphs to model switching dynamical systems, capturing interactions between objects, and learning both intra-object and inter-object mode-switching behavior. This method has been successfully applied to model various real-world systems, from crowds of people to groups of immune cells and swarms. Additionally, graph models are gaining popularity in various fields such as social networks, knowledge graphs, and protein-protein interaction networks. Furthermore, a recent work explores the potential of graph models in the biomedical field and introduces a new framework for creating medical records as graphs while maintaining privacy [12].

1.1 CPS Challenges

CPSs pose numerous challenges, with modeling being the most crucial. Modeling helps to analyze how a system works, determine necessary criteria, and tackle other

1 Department of Electrical Engineering, Golestan University, Gorgan, Iran.
Email:m.zavar400@stu.gu.ac.ir

2* CorrespondingAuthor :Department of Electrical Engineering, Golestan University, Gorgan, Iran.
Email:m.shahr@gu.ac.ir

3Department of Electrical Engineering, Golestan University, Gorgan, Iran.
Email:a.safa@gu.ac.ir

4 Department of Electrical Engineering, Golestan University, Gorgan, Iran.
Email: n.manouchehri400@stu.gu.ac.ir

Received: 2024.01.12; Accepted: 2024.04.03

challenges confidently [13]. Other challenges include ensuring security, managing complexity, scalability, integration, and resilience of the system, which become more critical as the dimensions and size of the systems increase [14-18]. Another challenge is observability and controllability, which has recently gained attention from researchers. Observability and controllability allow for establishing a connection between the physical and cyber parts of the system through a series of equipment.

To address this challenge and control and observe a CPS, sensors and actuators must be placed correctly. Leitold and his colleagues have proposed an approach based on Tarjan's algorithm, which locates sensors by finding strongly connected components (SCC) in the system's graph. However, this approach has four significant flaws. Firstly, finding SCCs differs depending on the starting location of the algorithm. Secondly, the location of sensors in SCCs is not unique. Thirdly, the number of sensors is more than required for system observability. Finally, this approach is only used for placing sensors [19].

Using optimization algorithms such as the Genetic Algorithms (GAs) is another way to place sensors and actuators. However, these algorithms require defining the cost function and necessary conditions for execution, leading to many iterations that take a lot of time to execute. Additionally, the final answer is not unique for large systems. Moreover, these algorithms must be executed twice and separately for the placement of sensors and actuators, further complicating the placement process and increasing execution time [20, 21].

1.2Motivation and Contribution

The presented approach utilizes a graph-based method that leverages the system's structural model and adjacency matrix to identify both existing and non-existing paths between system states. While previous methods have limitations, the Infinite Value Algorithm (IVA) addresses these issues. The IVA offers several advancements, such as:

i. The simultaneous placement of sensors and actuators
ii. Achieving observability and controllability with a minimal number of sensors and actuators
iii. Unique results for sensor and actuator placement
iv. Significantly short time to reach the answers

1.3Paper Structure

To provide context for the proposed approach, Section 2 defines key concepts and terms used in the research, including state space model (SSM) and its two sets of equations: state equations and output equations. Section 3 outlines the proposed approach and provides pseudocode for its implementation. To improve the approach's effectiveness, Section 4 briefly discusses the use of GAs and related settings. To evaluate the effectiveness of the proposed approach, Section 5 presents simulations of two systems with 10 and 100 states, created using Erdős–Rényi random graphs. These simulations are evaluated by IVA and GAs, highlighting the advantages of IVA over previous methods. Finally, Section 6 presents conclusions based on the simulations.

2. Preliminaries

2.1 State Space Model

A SSM is a method for describing the behavior of a dynamic system using variables known as states. These states represent the internal state of the system, even if they cannot be observed directly, but can be inferred from available measurements. SSM consists of two sets of equations: state equations and output equations.

State equations illustrate how states change over time and are often expressed as first-order differential equations. They can be represented using matrices and vectors. For a linear time-invariant system with n state variables and m inputs, the state equations can be expressed as:

$$\dot{x}(t) = \mathbf{A}\, x(t) + \mathbf{B}\, u(t) \tag{1}$$

where$\mathbf{A}$ with $n \times n$ dimensions and $\mathbf{B}$ with $n \times m$ dimensions are state matrix and output matrix, respectively. Vector x with dimensions $n \times 1$, u with dimensions $m \times 1$ and $\dot{x}$ with dimensions $n \times 1$ are respectively referred to as state vector, input vector, and state vector derivative.

Output equations describe the relationship between observed system outputs and states. They specify how states and inputs affect outputs and are usually presented as algebraic equations. Assuming that the system has p outputs, the output equations can be written in the form:

$$y(t) = \mathbf{C}x(t) + \mathbf{D}u(t) \tag{2}$$

where$\mathbf{C}$ with dimensions $p \times n$ and $\mathbf{D}$ with dimensions $p \times m$ are the output matrix and transmission matrix, respectively [22, 23].

2.2Graph Theory

Graph theory is a significant branch of mathematics that deals with graphs and mathematical structures used to represent relationships between objects. This theory has wide applications for modeling distributed systems, large-scale systems, and multi-agent systems. Graph theory focuses on researching the structure and properties of graphs, including their connectivity, and developing algorithms to solve computational problems related to graphs. In recent years, graph theory has become increasingly important in the analysis of complex networks, such as social media platforms and biological networks. In general, graph theory is a valuable tool that presents the analysis of complex systems along with the characteristics of dynamic systems in a simpler and more comprehensible way. Since the stability of CPSs is of great importance, two concepts of controllability and visibility are defined for systems [24, 25].

In a graph, there are two main components: nodes (vertices) and edges. Nodes represent objects or entities, while edges represent connections between them. The mathematical notation of a graph is defined as:

$$\mathcal{G} = (\mathrm{N}, \mathrm{E}) \begin{cases} \mathrm{N} \coloneqq \text{Nodes} \\ \mathrm{E} \coloneqq \text{Edges} \end{cases} \tag{3}$$

A directed graph $\mathcal{G}$ is a graph whose edges define the connection between nodes one-way. In the sense that if there is an edge from node i to node$j$, it does not mean that there is an edge from node j to node i.

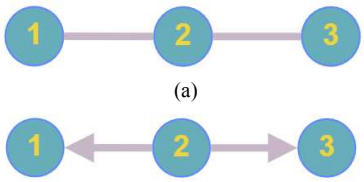


Fig. 1.(a) Undirected graph, (b) Directed Graph or digraph

A path is a sequence of connected edges that allows one to move from one node to another in a graph. The length of a path is determined by the number of edges it covers. A simple path does not repeat vertices or edges, except for the initial and final vertices. In Fig. 1(a), the path is formed by two edges. Figure 2 shows a path between 4 nodes, from node 1 to node 4 with three edges.

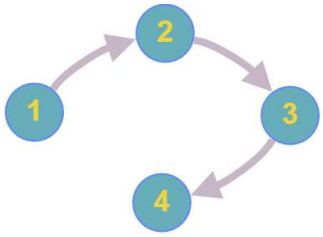


Fig. 2.A directed path with 4 nodes

2.3Adjacency Matrix

Each graph is drawn with a matrix called adjacency matrix. This matrix shows the relationship between the states of a CPS and is a square matrix whose dimensions are equal to the number of states or system nodes [26-28].

The adjacency matrix denoted by $\mathcal{A}$ in (4) is an$n \times n$ matrix where each row and column corresponds to a vertex. The value of a_{ij} in the matrix indicates whether there is an edge between vertices i and j. For a directed graph, if there is a directed edge from node i to node j, then a_{ij} has a value of one. If there is no directed edge from vertex i to vertex j, a_{ij} is set to zero. It is important to note that the diagonal entry a_{ii} in the adjacency matrix indicates whether there is a self-loop at node i or not. It can be set to a specified value if self-loops are allowed, or zero otherwise.

$$\mathcal{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{4}$$

The transpose of the adjacency matrix represents the state matrix of the system in the state space model as follows:

$$\mathcal{A} = \mathbf{A}^{\mathrm{T}} \tag{5}$$

2.4Controllability and Observability

Studying the observability and controllability of systems is one of the main applications of graph theory. For this purpose, a graph refers to the desired system in which the vertices correspond to the states, inputs, and outputs, and the edges express the relationships between the vertices. The graph can be assumed to be directed and weighted or without them, and the characteristics of dynamic systems can be informed using adjacency matrices. The main pillars of the systems are observability and controllability. By examining these two criteria, sensors and actuators can be placed in the systems [29, 30].

One of the widely used methods in evaluating the controllability and observability of CPSs is the use of

observability and controllability matrices and the Kalman rank check. Equations (6) and (7) show these matrices and the Kalman rank check of the system.

$$\mathcal{C} = [\mathbf{B} \quad \mathbf{AB} \quad \mathbf{A}^2\mathbf{B} \quad \cdots \quad \mathbf{A}^{n-1}\mathbf{B}] \tag{6}$$

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \tag{7}$$

A necessary and sufficient condition for a CPS to be controllable and observable is that the rank of the controllability matrix $\mathcal{C}$ and the observability matrix $\mathcal{O}$ is equal to the number of states of the system, i.e. the dimensions of the state matrix **A** or the adjacency matrix $\mathcal{A}$. These conditions ensure that any state variable can be affected by appropriate choices of input and output vectors.

However,structuralcontrollability and observability occur based on the direction of the edges. In structural controllability, if an actuator is placed on a node that is connected to another node by a directed edge (output edge), both nodes are controlled. This process also transpires for observability with the input edge.

2.5 Structured Model of CPS

As mentioned earlier, modeling is one of the important challenges of CPS. Structural modeling is a new approach that has been used in recent years. The basic concept of the structured model is to store non-zero information in the state space matrices of the system. Zeros remain and non-zero entries are replaced by free parameters. Therefore, the state space matrices of the system as well as the adjacency matrix have only zero and one members. A system with this feature is known as a structured system, and in this way, modeling is called a structured model [31, 32].

This construct consists of several components, including certain zero-valued variables that do not have a direct effect on other variables. Some inputs, such as one, are constant and can help extract state variables effortlessly from other variables. In addition, the inputs that interact and communicate with each other are dependent on each other based on the definition of system parameters through algebraic equations. The symbol$\mathcal{A}$, which was previously used to represent the adjacency matrix, is also used as a structured model in the form of zeros and ones:

$$a_{ii}, a_{ij} := 0 \text{ or } 1 \tag{8}$$

To build a structuredmodel, you can create a directed graph with nodes that represent variables. These nodes include input nodes, state variables, and output nodes. If there are non-zero parameters for these variables in the state space equations, the associated nodes are connected by an edge.

By combining the techniques of structuredmodeling and graph theory, one can benefit from several advantages. First, for the extraction of additional structured knowledge by analyzing physical laws and decomposing the overall system into subsystems. Second, it provides the advantage of creating a clear visual representation of the structure,which is easier to understand. Third, it enables the possible of examine features that are purely structurally relevant without considering unknown parameter values. Typically, these unknown parameters refer to physical parameters in functions. Finally, it minimizes the computational load, making CPS manageable, especially for large systems.

3. ProposedApproach: Infinite Value Algorithm

This algorithm is defined based on edges and paths between nodes so that the system becomes observable with the fewest number of sensors and controllable with the fewest number of actuators. In fact, in graph $\mathcal{G}$, a graph built based on the adjacency matrix, the i-th node with a directed edge has a path to the j-th node, by placing a sensor on the j-th node, both nodes can be viewed structurally. By examining the Kalman rank of the system confirms the correctness of this concept. The same concept is realized in the $\mathcal{G}$ graph transpose, i.e. the graph implemented based on the system state matrix, for system actuators and structural controllability. The IVA is named for the way it works, which treats the number of non-existent paths between nodes as an infinite value.

Algorithm 1. Shortest Path

ShortestPath():
Require:adj
Ensure: Sorted nodes paths
1.**for**each node**do**
2. $sp \leftarrow$ totalPath(adj)
3. $NoInfinite \leftarrow$ infPath(sp)
4. **end for**
5. $sortNoInf \leftarrow$ Sort($NoInfinite$)
return: $sortNoInf$

The way IVA works is that it first determines the shortest paths between each node and other nodes. If a node is not

connected to another node, it means that there is no input edge in the sensor placement and no output edge to that node in the actuator placement. In this algorithm, non-existent paths are specified with infinite value. After determining the shortest paths, the number of infinite paths is estimated and sorted in ascending order based on the node number. Then, if a node has no infinite path, that node is known as the sensor and operator location. If the number of nodes with zero infinite paths is more than one, then the node with the highest number of paths with the rest of the nodes is selected. If the minimum number of infinite paths is not zero, the same number of sensors and actuators are added.

Algorithm 2. Infinite Value Algorithm

Require: adj
Ensure: $driverNodes, SensorNodes$
1. $[NoNodes, NoEdges] \leftarrow$ findNum (adj)
2. Call ShortestPath(adj)
3. $driverNoInf \leftarrow sortNoInf(1)$
4. **while** $driverNoInf = 0$ **do**
5.**for**each node**do**
6.**if***i*th node hasn't path to jth node **then**
7.**for**each node**do**
8.**if***k*th node has path to jth node**then**
9. $NoPathToInf \leftarrow [k, shortestPath(k, j)]$
10.**end if**
11. **end for**
12. **for**each node**do**
13. **if***k*th node hasn't path to jth node**then**
14. $NoPathToInf \leftarrow [j, shortestPath(j, j)]$
15. **end if**
16.**end for**
17.**for**each $NoPathToInf$ **do**
18. $bestDriver \leftarrow$ selectBest($NoPathToInf$)
19.$minInf \leftarrow$ selectMinInf($bestDriver$)
20.**end for**
21.Set new $driverNodes$ from $minInf$
22.**end if**
23.**end for**
24. $driverNoInf \leftarrow driverNoInf - 1$
19. **end while**
20.The same process is done for sensor placement
return: $driverNodes, SensorNodes$

Equation (9) expresses the number of sensors and actuators based on infinite paths. If the number of infinite paths is non-zero in the lowest case, you should consider the paths between the nodes. If the first node and other sensor nodes are specified, the path between them should be checked, and if there is a round edge path between two nodes, choosing one of them is sufficient for the system's observability.This procedure is also true for the actuator placement.

$$\text{if}\begin{cases}\text{number of}\infty\ = 0\ \Rightarrow \text{Number of sensors or actuators}\ = 1\\ \text{number of}\infty = k\ \Rightarrow \text{Maximum number of sensors or actuators} = k+1\end{cases} \quad (9)$$

The main issue in sensor and actuator placement is determining what it means to minimize a matrix-valued function. Doubtlessly, the suitability of a sensor and actuator configuration depends on the evaluation criteria. Within the literature, various cost functions have been proposed, including maximizing the trace, the determinant, the rank, or the minimum eigenvalue of the observability and its dual, the controllability Gramian. For example, the authors in [33-36] used the observability Gramian and its dual, the controllability Gramian for sensor and actuator placement. More precisely, the observability of a deterministic linear system is given by the observability Gramian, which is defined as:

$$\mathrm{W}_0 = \int_{t_0}^{t_f} e^{\mathbf{A}^{\mathrm{T}}t}\mathbf{C}^{\mathrm{T}}\mathbf{C}e^{\mathbf{A}t}\,dt \quad (10)$$

It is proven that maximizing the observability Gramian will in some sense minimize the estimation error. Generally, the calculation of Gramian is computationally intractable. One intuitive way to handle this issue is to assume that systems are stable. The state transition matrix $e^{\mathbf{A}t}$ comprises decaying exponentials for stable systems, so a finite positive definite limit of the observability Gramian always exists and can be calculated by solving a Lyapunov equation in (11).

$$\mathbf{A}^{\mathrm{T}}\mathrm{W}_{\mathrm{o}} + \mathbf{C}^{\mathrm{T}}\mathbf{C} + \mathrm{W}_{\mathrm{o}}\mathbf{A} = 0 \quad (11)$$

However, this assumption restricts the applicability of these methods only for stable systems. In this light, we use the Kalman rank-check-based criterion to minimize for considering both stable and unstable linear systems.

4. Genetic Algorithm

A genetic algorithm is an evolutionary algorithm that is formulated based on biological methods such as mutation, inheritance, selection principles, etc. Figure 3 shows the

flowchart related to GA. This algorithm is used for prediction and mathematical modeling. Designed to replicate specific natural evolutionary processes, GA has become a highly effective stochastic search technique that relies on the principles of natural selection and genetics. A GA starts with a set of arbitrary solutions called a population, where each individual is coded as a chromosome that proposes a solution to the problem. These chromosomes evolve through several iterations or generations. During each generation, chromosomes are evaluated based on specific fitness criteria. As the process proceeds over several generations, the algorithm converges to the best chromosome, which represents the optimal solution [37].

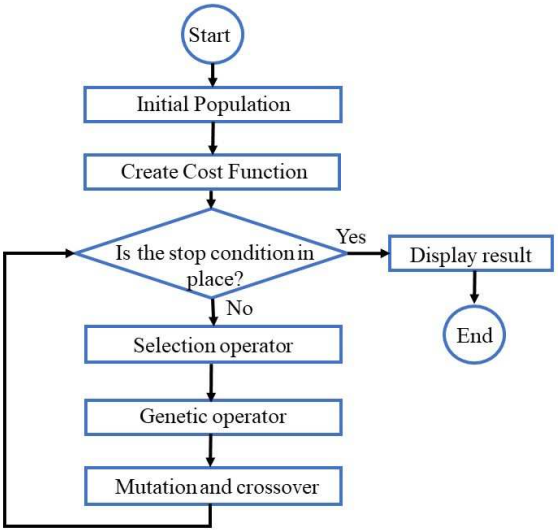


Fig. 3.A simple flowchart of GA

$$\text{Cost} = \text{observability or controllability} \tag{12}$$

The cost function considered for GA is system observability and controllability, which are performed in two separate steps.In (12) cost functions related to sensor and actuator placement are defined.Chromosome coding is binary and the number of genes is equal to the number of system states. The probability of crossoverand mutationequals 60% and 20%, respectively. The selection function is "TournamentSelection" and the population equals 25 individuals. The stop condition is defined as the number of executions equal to 200 iterations.

$$\begin{aligned} &\text{Minimize: Cost} = \text{Not}(OC) + Penalty \\ &OC = \text{Obsrvability or Controllability} \\ &Penalty = \frac{\text{Number of } Sensors}{\text{Number of } State} \end{aligned} \tag{13}$$

Another change has been made in the GA condition and cost function so that the number of sensors and actuators can be minimized. It can be seen in (13) cost functions and theirs conditions. However, the stopping condition this time is finding the minimum number of sensors and actuators to guarantee observability and controllability.

5. Simulation and Analysis

In this section, for 10-state and 100-state systems, the directed Erdős–Rényirandom graph model is used to compare IVA and GA. We have followed the approach presented in [38] to produce random directed graphs. All simulations and results are done in MATLAB software[1].The GA has been executed 10 times in such a way as to confirm the controllability and observability of the system.

5.1 The 10-state System

The first simulation is a system with 10 nodes, or a system that has 10 states in its state space model. The nodes or states of the 10-state system and their connections are expressed based on the adjacency matrix (14).

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{14}$$

As noted earlier, from the adjacency matrix, it can be understood that each node has an output edge and an input edge. Therefore, according to (14), the graph of the system can be drawn and its connections can be observed. Figure 4 shows the graph associated with this 10×10 adjacency matrix, where blue circles refer to nodes and black arrows refer to edges.

[1]System info: Core(TM) i3 CPU, M370 @ 2.40GHz, RAM 4.00 GB

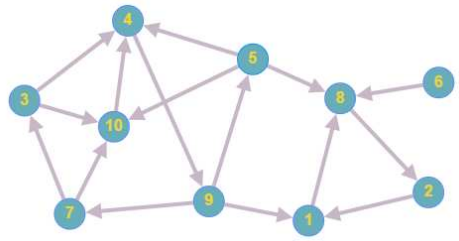


Fig. 4. Graph representation for 10-state system

In the first step, using the proximity matrix and Algorithm 1, it is possible to determine the existing paths and the infinite paths between the nodes. This tsakcan be easily in small dimensions of the drawing graph, but it becomes very difficult and sometimes impossible in large dimention.Therefore, utilizing the adjacency matrix will greatly assist in carrying out and implementing this algorithm. Table 1 and Table 2 show the paths between the nodes for the sensors positioning based on the proximity matrix and for the actuators positioning according to the state matrix. The sum of the infinite paths for each node is calculated and presented in ascending order in Table 2. Based on the IVA, the node with the fewest infinite paths is selected as candidate for both sensor and actuator. The system becomes observable with a sensor due to its structure, which allows access to all nodes of the system through one node.

Table 1. Existent paths and infinite paths in 10-state system for sensor placement

Node	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	0	1	3	2	3	3	4	2	1	3
x_2	2	0	5	4	2	2	6	1	3	5
x_3	∞	∞	0	3	4	∞	1	∞	2	4
x_4	∞	∞	1	0	1	∞	2	∞	2	1
x_5	∞	∞	3	2	0	∞	4	∞	1	3
x_6	∞	∞	∞	∞	∞	0	∞	∞	∞	∞
x_7	∞	∞	3	2	3	∞	0	∞	1	3
x_8	1	2	4	3	1	1	5	0	2	4
x_9	∞	∞	2	1	2	∞	3	∞	0	2
x_{10}	∞	∞	1	3	1	∞	1	∞	2	0

Table 2. Existent paths and infinite paths in 10-state system for actuator placement

Node	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	0	2	∞	∞	∞	∞	∞	1	∞	∞
x_2	1	0	∞	∞	∞	∞	∞	2	∞	∞
x_3	3	5	0	1	3	∞	3	4	2	1
x_4	2	4	3	0	2	∞	2	3	1	3
x_5	3	2	4	1	0	∞	3	1	2	1
x_6	3	2	∞	∞	∞	0	∞	1	∞	∞
x_7	4	6	1	2	4	∞	0	5	3	1
x_8	2	1	∞	∞	∞	∞	∞	0	∞	∞
x_9	1	3	2	2	1	∞	1	2	0	2
x_{10}	3	5	4	1	3	∞	3	4	2	0

According to Algorithm 1, Table 1 and Table 2, Table 3 is formed. Based on this table, 3 nodes are candidates for placing sensors and 6 nodes for actuators. But one must be chosen among them for sensor placement and two nodes be chosen for actuator placement. For this purpose, the sum of the shortest paths available for each node is calculated. In other words, the node that is connected to the largest node with the largest edge is selected. This table shows the number of infinite paths and the sum of the shortest paths available for placing sensors and actuators.

Now you can choose sensors and actuators according to Table 3. The number of infinite paths means that one node has no path to multiple nodes. In controllability, it means not controlling the node and in observability it means not seeing the desired node. In this table for actuators and sensors, the minimum number of infinite paths is zero and one, so this system is observable and controllable with one sensor and twoactuators. In the positioning of the sensor betweenx_1, $x_2$and x_8, node x_2 and the first positioning of the actuator betweenx_3, x_4, x_5, x_7, $x_9$andx_{10}, nodex_7 includes the most paths, so they are chosen as the right place for the placement of the sensor and the actuator.After that, to select the location of the second actuator, it should be determined which x_7 does not have a path to.then by referring to Table 2, another node that has a path to it should be selected. If several nodes have paths to that node, the node that has the highest sum of paths will be selected. Here, referring to Table 2 and the row related tox_7, it can be seen that this node does not have a path tox_6 and there is no other node that has a path to x_6. Therefore, $x_6$itself is also selected as the location of the actuator.

Table 3. Sum of the infinite and existent paths for 10-state system

Node	Number of Infinite Paths		Sum of Existent Paths	
	Sensors	Actuators	Sensors	Actuators
x_1	0	7	22	3
x_2	0	7	30	3
x_3	4	1	14	22
x_4	4	1	7	20
x_5	4	1	13	17
x_6	9	6	0	6
x_7	4	1	12	26
x_8	0	7	23	3
x_9	4	1	10	14
x_{10}	4	1	8	25

Figure 5 shows the comparison of the execution time of IVA and GAin 10runs. It should be noted that in IVA sensor and actuator placement are determined at the same time, so the displayed execution time is the sum of the execution time of the actuator and sensor placement. GA_{oc-sen}andGA_{oc-act} mean sensor placement and actuator placement respectively by GA along with (12). Also,

$GA_{min-sen}$ and $GA_{min-act}$ refer to sensor placement and actuator placement by GA along with (13). As can be seen, the running times of GA_{oc} and GA_{min} each alone are significantly different from IVA. However, Fig. 6 shows the execution time of GA_{oc} and GA_{min} as the total execution time of sensor placement and actuator placement. Based on this figure, it can be seen that the execution time of GAs is very different from the execution time of the proposed IVA.

Figure 7 and Fig. 8 respectively show the output results of the algorithms during 10 separate runs for sensor and actuator placements. The IVA had a unique answer, while the GA_oc and GA_min had different answers in the output. As mentioned earlier, one of the disadvantages of GA is the creation of the cost function and its necessary conditions. This has caused the failure to define the necessary conditions for sensor and actuator placement, resulting lack of a unique output. In addition, in Fig. 8, the performance of GA_{oc} in the ninth run is specified and it has selected one more actuator that is not minimized.

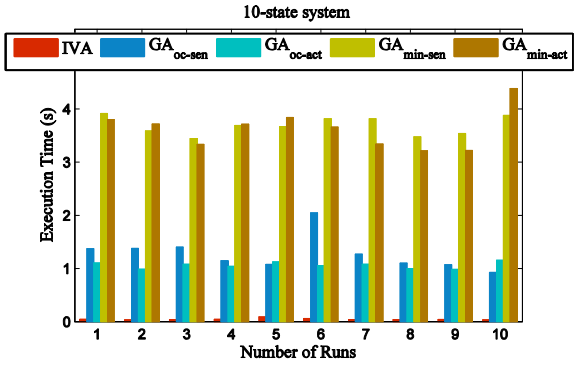


Fig. 5. Comparison of execution time of IVA and two separate cost functions ofGA_{oc} and GA_{min} for sensor and actuator in 10-state system

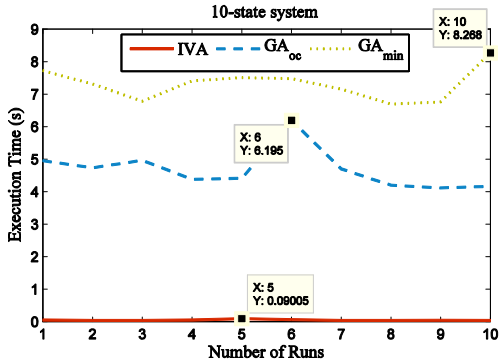


Fig. 6. Comparison of execution time of IVA, GA_{oc} and GA_{min} for sensor and actuator placement in 10-state system

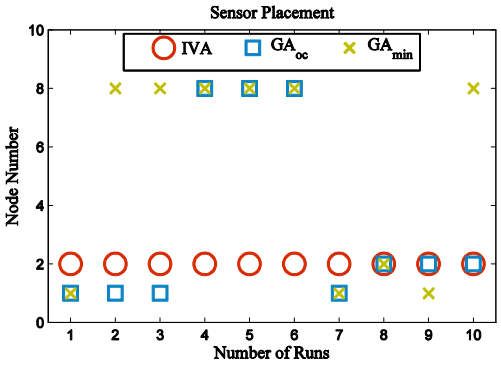


Fig. 7. Comparsionof sensorplacements in 10- state system

Table 4 clearly compares the performance of these three algorithms for placements. As can be seen, if the number of sensors and actuators is considered almost equal, their execution time is very defferent. Approximately, the execution time of $GA_{oc}$was 100 times and the execution time of $GA_{min}$was 150 times that of IVA.

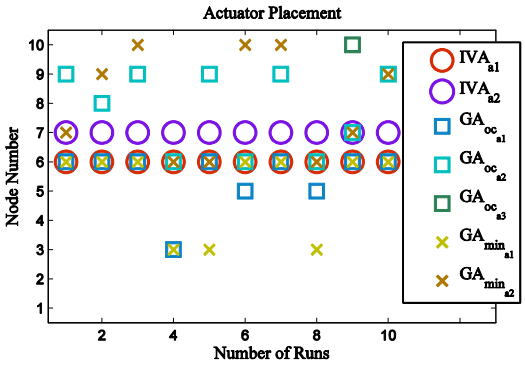


Fig. 8. Comparsionof actuatorplacements in 10- state system

Table 4. Compare algorithm results in 10-state system

IVA	GA_{min}	GA_{oc}		
1	1	1	Avg. Number of Sensors	
2	2	2.1	Avg. Number of Actuators	
0.048	3.684	2.122	Sensors	Avg. Execution Time
	3.624	2.558	Actuators	
0.048	7.308	4.680	Avg. Total Execution Time	

5.3The 100-state System

Now a larger system with 100 nodes or states is considered, on which the same processes as before are performed. The related graph is shown in Fig. 9. Now, using the tables related to IVA, you can start placing the sensor and actuator. The calculation of the shortest existing paths and non-existing paths or infinite paths between nodes is done in the first step and can be seen in Table 5.

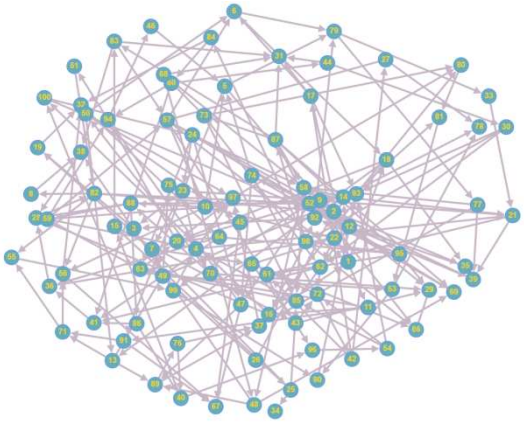


Fig. 9. Graph representation for 100-state system

As can be seen in Table 5, one node for the sensor and 8 nodes for the actuators are identified as having the fewest infinite paths. For positioning the sensors, only x_{78}has the fewest infinite paths, and for positioning the actuators, x_{84} has the most paths by 590. Therefore, these two nodes are selected as the first sensor and actuator nodes, but according to (9), because each of these nodes has an infinite path, another node for the sensor and another node for the actuator must be selected. To determine the other sensors and actuators, it should be checked which node is the infinite path of the nodes.

Table 5. Sum of the paths for 100-state system

Placement	Node	Number of Infinite Paths	Sum of Existent Paths
Sensors	x_{78}	5	172
Actoators	x_{30}	1	439
	x_{44}	1	530
	x_{60}	1	525
	x_{68}	1	421
	x_{73}	1	438
	x_{84}	1	590
	x_{87}	1	371
	x_{90}	1	15

Table 6 shows the nodes that do not have a path to x_{78}. Among them, two nodes x_{66} and x_{29}have a round trip path, so by choosing one of them as a sensor location, the other node will also be seen. Considering the total number of paths of each node in Table 6, $x_{29}$is placed in more paths than x_{66}, so $x_{29}$is chosen as the location of the sensor.In addition, for the other nodes listed in Table 6, there is no path to other nodes, so those nodes themselves are selected as the location of the sensors.

Table 6. Existent paths and infinite paths in 100-state system for sensor placement

Node	x_{29}	x_{31}	x_{34}	x_{66}	x_{75}	Sum of Existent Paths
x_{29}	0	∞	∞	1	∞	357
x_{31}	∞	0	∞	∞	∞	583
x_{34}	∞	∞	0	∞	∞	742
x_{66}	1	∞	∞	0	∞	351
x_{75}	∞	∞	∞	∞	0	519
x_{78}	∞	∞	∞	∞	∞	172

To select the second actuator node, the node to which x_{84}does not have a path must be specified first. Node x_{26} is the mentioned node and is not related to any other node, so it is chosen as the location of the actuator.

Table 7. Existent paths and infinite paths in 100-state system for actuator placement

Node	x_{26}	x_{84}	Sum of Existent Paths
x_{26}	0	∞	560
x_{84}	∞	∞	590

Figure10 and Fig. 11 show the execution time of IVA and the execution time of GAs. Figure 10 is the comparison of IVA,$GA_{oc}$and $GA_{min}$for sensor and actuator placement with two separate cost functions, and Fig.11 is the comparison of IVA, $GA_{oc}$and GA_{min}in general.

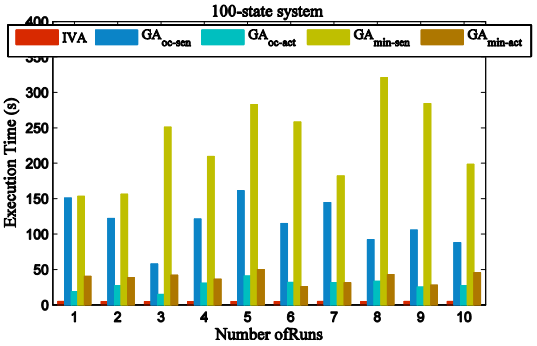


Fig. 10. Comparison of execution time of IVA and two separate cost functions of GAs for sensor and actuator in 100-state system

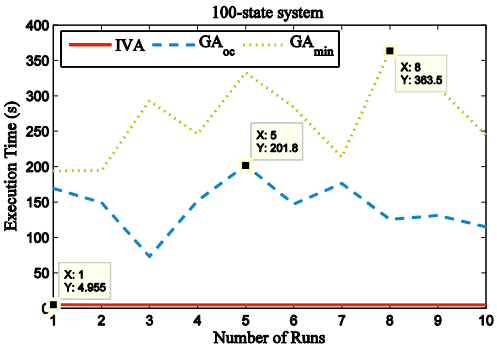


Fig. 11. Comparison of execution time of IVA and GAs for placements in 100-state system

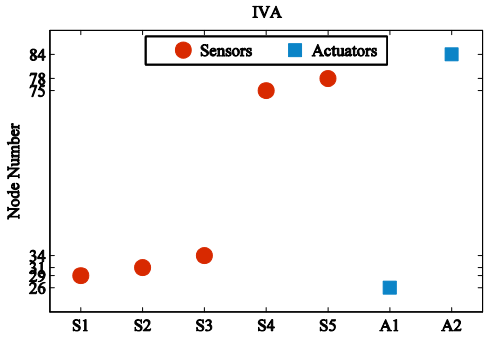


Fig. 12. Sensor and actuator placements in 100- state system by IVA

Figure 12 is related to the placement of the sensor and actuator in IVA, which can be observed and controlled by checking the Kalman rank of this 100-state system with 5 sensors and 2 actuators. Additionally, Fig. 13 displays the number of sensors and actuators in this system for 10 different runs of GA_{oc}. It is evident that the minimization of the number of sensors and actuators did not occur in this method. Furthermore, Fig. 14 illustrates the sensor and actuator placement by GA_{min}. However, this algorithm was unable to minimize the number of sensors in its fifth run within 200 iterations.

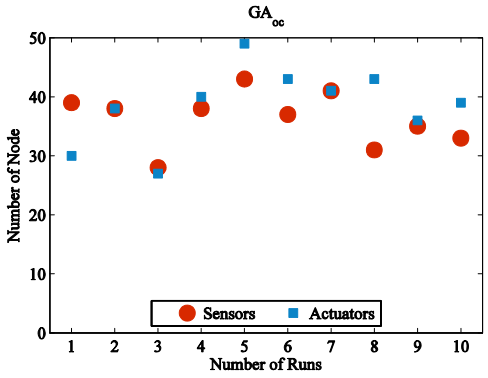


Fig. 13. Number of Sensors and actuators in 100- state system by GA_{oc}

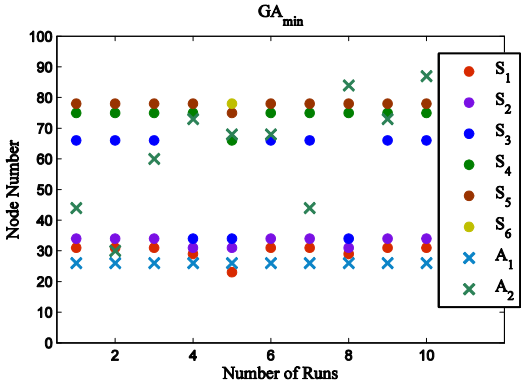


Fig. 14. Sensor and actuator placements in 100- state system by GA_{min}

Another comparison between the algorithms for the 100-state system is shown in Table 8. Here, the GA_{oc}considers a very large number of nodes for sensors and actuators, while in $GA_{min}$and IVA, their number is very small. On the other hand, the difference in execution time of GAs has been significantly higher than IVA. The total execution time of $GA_{min}$was more than 1.5 times that of $GA_{oc}$and more than 50 times that of IVA. Also, the total execution time of $GA_{oc}$was more than 30 times that of IVA.

Table 8. Results of proposed algotithm for 100-state system

IVA	GA_{min}	GA_{oc}		
5	5.1	36.30	Avg. Number of Sensors	
2	2	38.60	Avg. Number of Actuators	
4.783	229.668	115.828	Sensors	Avg. Execution Time
	38.154	28.136	Actuators	
4.783	267.822	143.964	Total Avg.Execution Time	

5. Conclusion

In this study, a new approach introduced,which simultaneously and in parallel advances the placement of sensors and actuators. This graph-based algorithm is based on the paths between nodes, particularly the paths that are introduced as infinite paths. After addressing the basic and necessary definitions, IVA is presented and a comparison with GAs has been made in the placements to evaluate it. Additionally, two methods have been used for GA, one of which only works with system controllability and observability as GA_{oc} , and the other determines the minimum number of sensors and actuators required to obtain controllability and observability as GA_{min}.

The simulation results indicate three main and fundamental advantages of the IVA algorithm, which can be an important reason for its use. The most important factor in IVA is the simultaneous sensor and actuator placement. This process prevents the algorithm from repeating itself and both placements occur with a single execution. This is done separately in the GA optimization algorithms. Another advantage is that IVA guarantees the uniqueness of the solution, while GA_{min}, although it is designed to find the minimum value, was not able to minimize in 200 iterations. The third benefitof IVA is the much lower execution time compared to GAs.The aforementioned advantages have been shown in simulations related to 10-state and 100-state systems. However, execution time will take a significant amount in large-scale systems, and the time difference between these two algorithms (i.e. the proposed one and GAs) will be more visible. In general, IVA as a graph-based method has performed better than GAsas optimization-based method and can be a suitable option for sensor and

actuator placement.

In future research, there will be a focus on combining IVA and centrality measures for placements. This will be particularly useful when infinite paths in IVA are equal, as the use of centrality measures will allow for more efficient selection of placement candidates. Additionally, important real systems will be studied to determine the impact of various factors and to assess the effectiveness of the proposed algorithm. Furthermore, future work should extend the proposed methodology to incorporate more sophisticated models. In particular, we optimized sensor and actuator placement with the assumption that the understudied models are linear time invariant systems. However, real-world systems are often influenced by time varying and nonlinear processes. Neglecting these dynamics may impair the optimality of the sensor and actuator placement. Further work should also investigate ways to extend our proposed sensor and actuator placement algorithm to time-varying and nonlinear systems.

References

[1] Bakirtzis, Georgios, Christina Vasilakopoulou, and Cody H Fleming. "Compositional Cyber-Physical Systems Modeling." arXiv preprint arXiv:2101.10484 (2021).

[2] Ho, Nicholas, Pooi-Mun Wong, Ngoc-Son Hoang, Dun-Kai Koh, Matthew Chin Heng Chua, and Chee-Kong Chui. "Cps-Based Manufacturing Workcell for the Production of Hybrid Medical Devices." Journal of Ambient Intelligence and Humanized Computing 12 (2021): 10865-79.

[3] Tavčar, Jože, Jože Duhovnik, and Imre Horváth. "From Validation of Medical Devices Towards Validation of Adaptive Cyber-Physical Systems." Journal of Integrated Design and Process Science 23, no. 1 (2019): 37-59.

[4] Castillo-Martínez, Diego Hilario, Adolfo Josué Rodríguez-Rodríguez, Adrian Soto, Alberto Berrueta, David Tomás Vargas-Requena, Ignacio R Matias, Pablo Sanchis, Alfredo Ursúa, and Wenceslao Eduardo Rodríguez-Rodríguez. "Design and on-Field Validation of an Embedded System for Monitoring Second-Life Electric Vehicle Lithium-Ion Batteries." Sensors 22, no. 17 (2022): 6376.

[5] Hu, Zhongxu, Shanhe Lou, Yang Xing, Xiao Wang, Dongpu Cao, and Chen Lv. "Review and Perspectives on Driver Digital Twin and Its Enabling Technologies for Intelligent Vehicles." IEEE Transactions on Intelligent Vehicles (2022).

[6] Li, Ning, Haiyi Sun, Xin Jing, and Zhongtang Chen. "Dynamic Modeling and Aperiodically Intermittent Strategy for Adaptive Finite-Time Synchronization Control of the Multi-Weighted Complex Transportation Networks with Multiple Delays." Chinese Physics B 30, no. 9 (2021): 090507.

[7] Wu, Jianping, and Dongping Fang. "Role of Cps in Smart Cities." Cyber-Physical Systems in the Built Environment (2020): 255-72.

[8] Cui, Yi, Feifei Bai, Tapan Saha, and Jalil Yaghoobi. "Authenticating Source Information of Distribution Synchrophasors at Intra-State Locations for Cyber-Physical Resilient Power Networks." International Journal of Electrical Power & Energy Systems 139 (2022): 108009.

[9] Khan, Izhar Ahmed, Dechang Pi, Nasrullah Khan, Zaheer Ullah Khan, Yasir Hussain, Asif Nawaz, and Farman Ali. "A Privacy-Conserving Framework Based Intrusion Detection Method for Detecting and Recognizing Malicious Behaviours in Cyber-Physical Power Networks." Applied Intelligence (2021): 1-16.

[10] Cui, Xinyue. "Cyber-Physical System (Cps) Architecture for Real-Time Water Sustainability Management in Manufacturing Industry." Procedia CIRP 99 (2021): 543-48.

[11] Liu, Yongtuo, Sara Magliacane, Miltiadis Kofinas, and Efstratios Gavves. "Graph Switching Dynamical Systems." arXiv preprint arXiv:2306.00370 (2023).

[12] Vazirgiannis, Michalis. "Gnns and Graph Generative Models for Biomedical Applications." Proceedings of the ACM Web Conference 2023, Austin, TX, USA, Association for Computing Machinery, 2023.

[13] Adamos, Konstantinos, George Stergiopoulos, Michalis Karamousadakis, and Dimitris Gritzalis. "Enhancing Attack Resilience of Cyber-Physical Systems through State Dependency Graph Models." International Journal of Information Security (2023): 1-12.

[14] Bodkhe, Umesh, Dhyey Mehta, Sudeep Tanwar, Pronaya Bhattacharya, Pradeep Kumar Singh, and Wei-Chiang Hong. "A Survey on Decentralized Consensus Mechanisms for Cyber Physical Systems." IEEE Access 8 (2020): 54371-401.

[15] Mittal, Saurabh, and Andreas Tolk. Complexity Challenges in Cyber Physical Systems: Using Modeling and Simulation (M&S) to Support Intelligence, Adaptation and Autonomy. John Wiley & Sons, 2019.

[16] Abbas, Arbab Waseem, and Safdar Nawaz Khan Marwat. "Scalable Emulated Framework for Iot

Devices in Smart Logistics Based Cyber-Physical Systems: Bonded Coverage and Connectivity Analysis." IEEE Access 8 (2020): 138350-72.

[17] Fataliyev, Tahmasib Kh, and Shakir A Mehdiyev. "Integration of Cyber-Physical Systems in E-Science Environment: State-of-the-Art, Problems and Effective Solutions." International Journal of Modern Education and Computer Science 11, no. 9 (2019): 35.

[18] Colabianchi, Silvia, Francesco Costantino, Giulio Di Gravio, Fabio Nonino, and Riccardo Patriarca. "Discussing Resilience in the Context of Cyber Physical Systems." Computers & Industrial Engineering 160 (2021): 107534.

[19] Leitold, Dániel, Ágnes Vathy-Fogarassy, and János Abonyi. "Network-Based Observability and Controllability Analysis of Dynamical Systems: The Nocad Toolbox." F1000Research 8 (2019).

[20] Civera, Marco, Marica Leonarda Pecorelli, Rosario Ceravolo, Cecilia Surace, and Luca Zanotti Fragonara. "A Multi□Objective Genetic Algorithm Strategy for Robust Optimal Sensor Placement." Computer□Aided Civil and Infrastructure Engineering 36, no. 9 (2021): 1185-202.

[21] Dhuri, KD, and P Seshu. "Multi-Objective Optimization of Piezo Actuator Placement and Sizing Using Genetic Algorithm." Journal of sound and vibration 323, no. 3-5 (2009): 495-514.

[22] Peng, Yuhuai, Alireza Jolfaei, Qiaozhi Hua, Wen-Long Shang, and Keping Yu. "Real-Time Transmission Optimization for Edge Computing in Industrial Cyber-Physical Systems." IEEE Transactions on Industrial Informatics 18, no. 12 (2022): 9292-301.

[23] Zhou, Xin, Xiaodong Gou, Tingting Huang, and Shunkun Yang. "Review on Testing of Cyber Physical Systems: Methods and Testbeds." IEEE Access 6 (2018): 52179-94.

[24] Tutte, William Thomas. Graph Theory. Vol. 21: Cambridge university press, 2001.

[25] Shahraeini, Mohammad, Panayiotis Kotzanikolaou, and Mehrab Nasrolahi. "Communication Resilience for Smart Grids Based on Dependence Graphs and Eigenspectral Analysis." IEEE Systems Journal 16, no. 4 (2022): 6558-68.

[26] Bhunia, Pintu, Santanu Bag, and Kallol Paul. "Bounds for Eigenvalues of the Adjacency Matrix of a Graph." Journal of Interdisciplinary Mathematics 22, no. 4 (2019): 415-31.

[27] Shahraeini, Mohammad, Shahla Khormali, and Ahad Alvandi. "Optimal Pmu Placement Considering Reliability of Measurement System in Smart Grids." Paper presented at the 2022 12th International Conference on Computer and Knowledge Engineering (ICCKE), 2022.

[28] Xie, Jun, Qiguang Miao, Ruyi Liu, Wentian Xin, Lei Tang, Sheng Zhong, and Xuesong Gao. "Attention Adjacency Matrix Based Graph Convolutional Networks for Skeleton-Based Action Recognition." Neurocomputing 440 (2021): 230-39.

[29] Jungers, Raphael M, Atreyee Kundu, and WPMH Heemels. "Observability and Controllability Analysis of Linear Systems Subject to Data Losses." IEEE Transactions on Automatic Control 63, no. 10 (2017): 3361-76.

[30] Yan, Jiayuan, Bin Hu, Zhi-Hong Guan, Tao Li, and Ding-Xue Zhang. "On Controllability and Observability of a Class of Fractional-Order Switched Systems with Impulse." Nonlinear Analysis: Hybrid Systems 50 (2023): 101378.

[31] Chen, Chen, Ruiyue Peng, Lei Ying, and Hanghang Tong. "Fast Connectivity Minimization on Large-Scale Networks." ACM Transactions on Knowledge Discovery from Data (TKDD) 15, no. 3 (2021): 1-25.

[32] Sun, Wen, Junxia Guan, Jinhu Lü, Zhigang Zheng, Xinghuo Yu, and Shihua Chen. "Synchronization of the Networked System with Continuous and Impulsive Hybrid Communications." IEEE Transactions on Neural Networks and Learning Systems 31, no. 3 (2019): 960-71.

[33] Bopardikar, Shaunak D. "A Randomized Approach to Sensor Placement with Observability Assurance." Automatica 123 (2021): 109340. https://doi.org/https://doi.org/10.1016/j.automatica.2020.109340.

[34] Manohar, Krithika, J Nathan Kutz, and Steven L Brunton. "Optimal Sensor and Actuator Selection Using Balanced Model Reduction." IEEE Transactions on Automatic Control 67, no. 4 (2021): 2108-15. https://doi.org/https://doi.org/10.1109/TAC.2021.3082502.

[35] Takahashi, Shun, Yasuo Sasaki, Takayuki Nagata, Keigo Yamada, Kumi Nakai, Yuji Saito, and Taku Nonomura. "Sensor Selection by Greedy Method for Linear Dynamical Systems: Comparative Study on Fisher-Information-Matrix, Observability-Gramian and Kalman-Filter-Based Indices." IEEE Access (2023). https://doi.org/https://doi.org/10.1109/ACCESS.2023.3291415.

[36] Yamada, Keigo, Yasuo Sasaki, Takayuki Nagata, Kumi

Nakai, Daisuke Tsubakino, and Taku Nonomura. "Efficient Sensor Node Selection for Observability Gramian Optimization." Sensors 23, no. 13 (2023): 5961. https://doi.org/https://doi.org/10.3390/s23135961.

[37] Shirajuddin, Talhah Mohamad, Nur Shazwani Muhammad, and Jazuri Abdullah. "Optimization Problems in Water Distribution Systems Using Non-Dominated Sorting Genetic Algorithm Ii: An Overview." Ain Shams Engineering Journal 14, no. 4 (2023): 101932.

[38] Shahraeini, Mohammad. “Modified Erdős–Rényi Random Graph Model for Generating Synthetic Power Grids.” IEEE Systems Journal. Institute of Electrical and Electronics Engineers (IEEE), 2023. https://doi.org/10.1109/jsyst.2023.3339664.

Disturbance Rejection Nonlinear Control of Flexible Joint Robot Arm

Hadi Hasanpour[1], Esmaeil Alibeiki[2*], SeyyedMostafa Ghadami[3]

Abstract- In this article, we investigate different control methods and control input disturbance removal in the two-link flexible robot. For this purpose, we use the linear feedback control method to control the trajectory tracking of the robot's anglewith a nonlinear model by reducing the effects of disturbances in the control signal. The main contribiution of this paper is disturbance rejection that effect on performance of robot. In the end, by analyzing and comparing the obtained results, the best method is selected to control and eliminate the control input disturbance in the flexible two-link robot. To analyze the results of this research, MATLAB software and the system simulation are performed and according to the desired trajectory for the robot arms, the optimum control input for the system is obtained. Based on the results, the proposed control structure controls the model property without overshooting and by considering the saturation limit of input control.

Keywords: Flexible robot – Nonlinear model-Disturbance rejection

1. Introduction

In recent years, fixed-time control algorithms have become increasingly popular in robotics systems due to their fast convergence rates and ability to estimate convergence time without relying on initial states [1-3]. Compared to traditional finite-time control methods, fixed-time control algorithms possess an advantage in convergence time estimation, making them more useful for real-time applications [3].Various fixed-time control schemes have been developed for systems with unknown uncertainties and multi-agent systems. For example, a distributed fixed-time controller has been designed for an order integrator system with unknown uncertainties caused by inherent dynamics, guaranteeing the convergence of tracking errors in finite time [4]. Nonlinear fixed-time consensus schemes have also been constructed to guarantee the fixed-time consensus problem for a multi-agent system with undirected communication graphs [4].Neural network control methods have also become increasingly popular in recent years to deal with uncertainties and nonlinearities in robotics systems [5]. These methods include fault-tolerant controllers, approximation control schemes, cooperative adaptive control schemes, and fuzzy neural network control methods.For example, a neural networks-based fault-tolerant controller has been constructed to estimate actuator failures and uncertain dynamics [5]. Approximation control schemes based on neural networks have been applied to handle model parametric uncertainties and unknown disturbances. The radial basis function neural network control method has been proposed to handle the problems of the unknown dynamic model of coordinated dual arms robot and the saturation nonlinearity of the motor. Cooperative adaptive control schemes have been investigated to deal with nonlinear systems, while fuzzy neural network control methods have been proposed for nonlinear systems [6].However, the transient performances of these methods are not always satisfying, and the accuracy of the models for robotic systems can be difficult to obtain. Additionally, the methods proposed in the literature often rely on the assumption that all parameters of the robotic systems are known, which may not be practical in real-world applications.As a result, the development of a neural networks-based fixed-time control method for a robot with uncertainties and input

1 Department of Electrical Engineering, Aliabad Katoul Branch, Islamic Azad University, Aliabad Katoul, Iran. Email: hadihasanpor@gmail.com

2* Corresponding Author : Department of Computer Engineering, Aliabad katoul Branch, Islamic Azad University, Aliabad Katoul, Iran. Email:esmail_alibeiki@aliabadiau.ac.ir

3 Department of Electrical Engineering, Aliabad Katoul Branch, Islamic Azad University, Aliabad Katoul, Iran. Email: ghadami@aliabadiau.ac.ir

Received: 2023.07.23; Accepted: 2024.03.12

deadzone is significant [7]. This approach could provide better transient performance while also addressing uncertainties and input dead zones, which are common challenges in robotics systems.Two-joint flexible robots have many potential applications that can benefit from their unique combination of flexibility and controllability. Here are some examples [8]:

Medical robotics: Two-joint flexible robots can be used in medical applications such as minimally invasive surgery and endoscopy. The flexibility of the robot can allow it to navigate through tight spaces and around organs, while the controllability of the system can enable precise manipulation of surgical tools.

Manufacturing: Two-joint flexible robots can be used in manufacturing processes that require delicate handling of objects, such as assembling small electronic components or handling fragile materials. The flexibility of the robot can enable it to adapt to different shapes and sizes of objects, while the controllability of the system can ensure precise positioning and manipulation.

Inspection and maintenance: Two-joint flexible robots can be used for the inspection and maintenance of structures such as pipelines, bridges, and buildings. The flexibility of the robot can enable it to navigate through complex environments, while the controllability of the system can ensure precise positioning and inspection.

Entertainment: Two-joint flexible robots can be used in entertainment applications such as animatronics and robot performances. The flexibility of the robot can enable it to move in a lifelike manner, while the controllability of the system can enable precise and coordinated movement.

Education and research: Two-joint flexible robots can be used in educational settings to teach robotics and control theory, as well as in research settings to study the dynamics and control of flexible robotic systems.

Control methods for two-joint flexible robots typically involve a combination of feedforward and feedback control techniques to achieve accurate and robust control of the system [9]. Feedforward control is used to compensate for the dynamic effects of the flexible link, while feedback control is used to track desired trajectories and reject disturbances.Some recent research papers on control methods for two-joint flexible robots have focused on using advanced control techniques such as adaptive control, model predictive control, and sliding mode control [10]. These methods can improve the performance of the system while also providing robustness to uncertainties and disturbances.Other papers have explored the use of machine learning techniques, such as reinforcement learning and neural networks, to learn control policies for two-joint flexible robots. These methods can be effective in learning complex control strategies that are difficult to design manually.Overall, the field of control methods for two-joint flexible robots is an active area of research, and many new approaches and techniques are being developed and tested.Two-joint flexible robots are a type of robotic system that typically consists of two links connected by a flexible joint. The flexible joint can deform under external loads or forces, which can cause the robot to behave in a non-linear and unpredictable manner. This makes control of the system challenging, as the dynamics of the flexible joint can be difficult to model accurately.To address this challenge, researchers have developed a variety of control methods for two-joint flexible robots. One common approach is to use a combination of feedforward and feedback control techniques. Feedforward control involves estimating the dynamic effects of the flexible joint and compensating for them in the control signal. Feedback control uses sensors to measure the position, velocity, and acceleration of the robot, and adjusts the control signal to track desired trajectories and reject disturbances. Adaptive control is another technique that has been used for two-joint flexible robots [11]. Adaptive control involves continuously updating the control parameters based on feedback from the system, which can improve the accuracy and robustness of the control. Model predictive control is another advanced control technique that has been applied to two-joint flexible robots, which involves predicting the future behavior of the system and optimizing the control signal accordingly [12].

In this paper, the proposed controller aims to reject input control disturbance and to follow the desired path for the links model with two flexible joints. Numerical simulations are done in Matlabsoftwareand the proposed controller is designed using linearization input-output feedback with a disturbance rejection block [13]. A simplified-adaptive controller with proper adjustment of the obtained constant values is presented, and the simulation results for the performance of the tracking systems are presented [14-15].Our research innovations are a multi-controller design-based disturbance rejection structure and the use of a linearization input-output feedback method to achieve the desired path of the links. disturbance rejection from the robot system is very important because its presence can disturb the overall performance of the system and also weaken the

effectiveness of the controller. Therefore, it is necessary to provide a solution for disturbance rejection.

The rest of the paper is organized as follows. In the second section, the model of two flexible jointsis presented. In the third section, the proposed controller is designed. In the fourth section, the the nonlinear model is simulated in MATLAB software and based on the results, the performance of the method is evaluated. In the fifth section, according to the evaluations performed, the conclusion of the paper is explained.

2. Flexible Robot

The model is used to predict the behavior of the robot, including its motion, position, velocity, and acceleration, as well as the internal stresses and strains in the flexible link.The mathematical model typically includes a set of differential equations that describe the dynamics of the system. These equations take into account the forces and torques acting on the robot, as well as the elastic properties of the flexible link. The model can be used to simulate the behavior of the robot under different conditions, such as different input commands or external disturbances.One approach to developing a mathematical model of a two-joint flexible robot is to use a Lagrangian formulation, which describes the system's dynamics in terms of its kinetic and potential energies. This approach can lead to a set of coupled differential equations that describe the motion of the robot and the deformation of the flexible link.Another approach is to use a finite element method, which discretizes the flexible link into a set of smaller elements and solves the behavior of each element separately. This approach can be more computationally intensive, but it can provide more detailed information about the internal stresses and strains in the flexible link.The two-joint flexible robot structure is shown in Figure 1 with one motor attached to each arm. In this robot the arm movement is elastic and its effect is in the form of a spring.

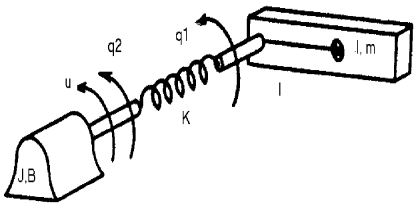


Fig 1. Structure of FJR robot with a single link

The dynamic equation of the FJR robot for the n-joint state is in the form of the following relationships, which include the joint rotation relationships and the motor rotation angle of the joints:

$$M\left(q_1\right)\ddot{q}_1 + C\left(q_1.\dot{q}_1\right)\dot{q}_1 + G\left(q_1\right) = K\left(q_1 - q_2\right) \quad (1)$$

$$J\ddot{q}_2 + B\dot{q}_1 + K\left(q_2 - q_1\right) = t \quad (2)$$

In the above relation, q1 and q2 are the vectors of the joint angle and the motor rotation angle, respectively. The matrix M,G, C, and Bare the general components of inertia, the Coriolis vector, the gravity vector, and the actuator damping component, respectively. K Indicates the coefficient of elastic force. Also in the motor equation, the motor inertia matrix(J), the drive attenuation matrix, and the matrix coefficient of articulation are considered. In single robot mode, the above equations will be in scalar values.

3. Disturbance rejection method

Based on the results in [14], a motion control design based on a continuous single-output (SISO) control system using a progressive control system is proposed. In this system, only one perturbation is applied to the primary plant input. In some cases, such as flexible systems controlled by full closed-loop position control systems, a single disturbance model is not sufficient if physical factors that are not present in the plant model affect the output of small and large loops. To explain this point, consider a flexible joint that is modeled as a two-mass system, depending on Figure 2, where Jmand Dm are investigated.

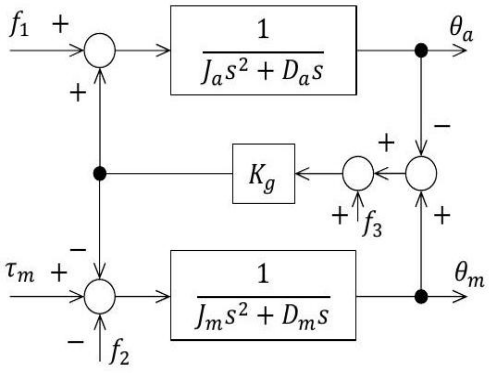


Fig. 2. Two-mass system with physical disturbances.

moment of inertia and viscous damping factor of the motor side; J_a and D_a are the moment of inertia and viscous damping factor of the arm side; K_g is the spring factor of the elastic component in the gear; f_1 is the physical disturbance torque applied at the arm side including the interference force generated from the other axis movement and the external force; f_2 is the physical disturbance torque applied at the motor side such as nonlinear friction; f_3 is the physical disturbance applied within the elastic component such as the gear's angular transmission error. The input/output relationship of this system is expressed as follows:

$$\theta_m = G_1(s)\left(\tau_m - G_{f_1,m}(s)f_1 - G_{f_2,m}(s)f_2 - G_{f_3,m}(s)f_3\right), \theta_a = G_2(s)\left(\theta_m - G_{f_1,a}(s)f_1 - G_{f_2,a}(s)f_2 - G_{f_3,a}(s)f_3\right) \quad (3)$$

where the transfer functions are computed as:

$$G_1(s) = \frac{P_{n,m}(s)}{P_d(s)}, \quad G_2(s) = \frac{P_{n,a}(s)}{P_{n,m}(s)} \quad (4)$$

$$G_{f_1,m}(s) = -\frac{K_g}{J_a s^2 + D_a s + K_g}, \quad G_{f_2,m}(s) = 1 \quad (5)$$

$$G_{f_3,m}(s) = \frac{K_g(J_a s^2 + D_a s)}{J_a s^2 + D_a s + K_g} \quad (6)$$

$$G_{f_1,a}(s) = -\frac{1}{K_g}, \quad G_{f_2,a}(s) = 0, \quad G_{f_3,a}(s) = -1 \quad (7)$$

$$P_{n,a}(a) = K_g \quad (8)$$

$$P_{n,m}(s) = J_a s^2 + D_a s + K_g, \quad (9)$$

$$P_d(s) = J_m J_a s^4 + (J_m D_a + J_a D_m)s^3 + (D_m D_a + K_g J_m + K_g J_a)s^2 + K_g(D_m + D_a)s \quad (10)$$

The Eqs. (3) and (4) can be described by the block diagram shown in Fig. 3, where:

$$d_1 = G_{f_1,m}(s)f_1 + G_{f_2,m}(s)f_2 + G_{f_3,m}(s)f_3 \quad (11)$$

$$d_2 = G_{f_1,a}(s)f_1 + G_{f_2,a}(s)f_2 + G_{f_3,a}(s)f_3 \quad (12)$$

To capture the impact of physical disturbances on both the inner and outer loop outputs of a fully closed control system, the double-disturbance model shown in Figure 3 must be examined. This is particularly crucial for multi-joint flexible robots, as the joints' simultaneous motion generates significant interference force [10], causing f_1 to become substantial and d_2 to become non-negligible. In this model, d_1 and d_2 represent the overall effects of physical disturbances on the minor and major loops, respectively, of a cascade positioning control system.

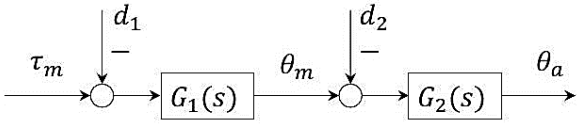


Fig. 3. Double-disturbance model

The controlled output θ_a can be then expressed as follows:

$$\begin{aligned}\theta_a = & \frac{C_p(s)C_v(s)G_1(s)G_2(s)}{1+C_v(s)G_1(s)s+C_p(s)C_v(s)G_1(s)G_2(s)} r \\ & + \frac{C_v(s)G_1(s)G_2(s)}{1+C_v(s)G_1(s)s+C_p(s)C_v(s)G_1(s)G_2(s)} c_\omega \\ & + \frac{G_1(s)G_2(s)}{1+C_v(s)G_1(s)s+C_p(s)C_v(s)G_1(s)G_2(s)} c_\tau \\ & - \frac{G_1(s)G_2(s)}{1+C_v(s)G_1(s)s+C_p(s)C_v(s)G_1(s)G_2(s)} d_1 \\ & - \frac{G_2(s)+C_v(s)G_1(s)sG_2(s)}{1+C_v(s)G_1(s)s+C_p(s)C_v(s)G_1(s)G_2(s)} d_2\end{aligned} \quad (13)$$

From the above transfer functions, the effect of disturbances on the controlled output can be eliminated by using either one of two following sets of compensation inputs:

- Set 1 :

$$c_\omega = \frac{1+C_v(s)G_1(s)s}{C_v(s)G_1(s)} d_2, \quad c_\tau = d_1 \quad (14)$$

- Set 2 :

$$c_\omega = sd_2, \quad c_\tau = d_1 + G_1^{-1}(s)d_2. \quad (15)$$

Although the first set of equations in (14) is a more intuitive solution, as it uses c_τ to compensate for d_1 in the minor loop and c_ω to compensate for d_2 in the major loop, the second set of equations in (15) should be chosen due to its independence from feedback controllers.

(19)

4. Simulation

The purpose of this paper is to track the desired sinusoidal path as well as disturbance rejection at both

angles of the robot joint. The minimum disturbance effect on the input control signal and the elimination of this effect in the output signal is the aim of this paper. For this purpose, we analyze the behavior of the system in a linear model and the designed controller is used to achieve the desired path track based on [13]. In this simulation, the robot parameter is given in Table 1.

Table 1:The Parameters of the FJR Two-Joint Robot Model

Unit	Value	Parameter
Nm/rad	100	K_1
Nm/rad	100	K_2
Nms/rad	0.9	B_2
Nms/rad	0.9	B_2
Kg	0.5	m_1
Kg	0.5	m_2
M	1	L_1
M	1	L_2

In this simulation, the optimal path for the first and second joint angles is 0.25cos(t). In this simulation, the sampling rate is 0.01 s and the initial values of the angles of the robot are set to zero. This simulation is performed in 9 seconds. The simulation results with the conditions expressed in the form of two outputs q1 and q2 are flexible robots and the control inputs in the form of τ1 and τ2 are shown in Figures 2 and 3. The sinusoidal disturbance is added at 4.9s with attitude 0.5 N.m.

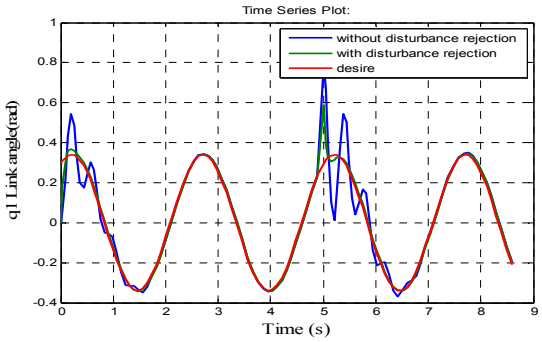


Fig.4 Angle Related to Dynamics q_{11} and q_{12}

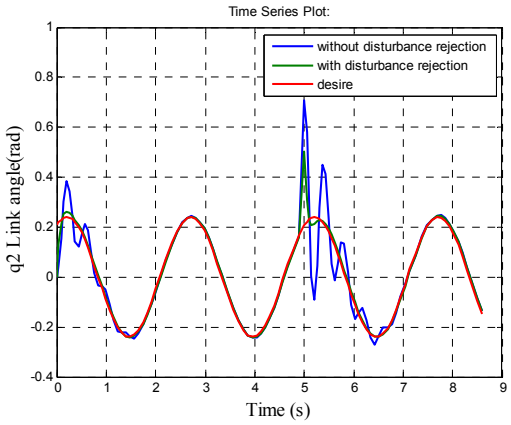


Fig.5 The related angles q_{21} and q_{22}

In this section, the adaptive backstepping controller design is discussed in [15].The simulation results of input control for the controller are shown in Figures 4 and 5.

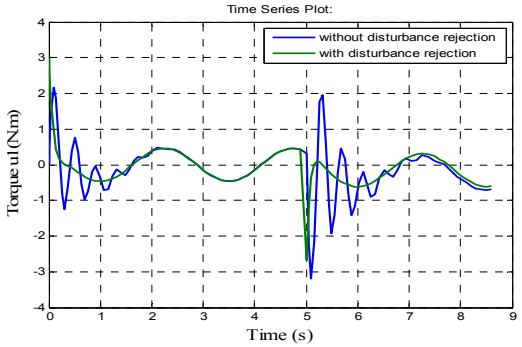


Fig. 6. Input controller of the first joint

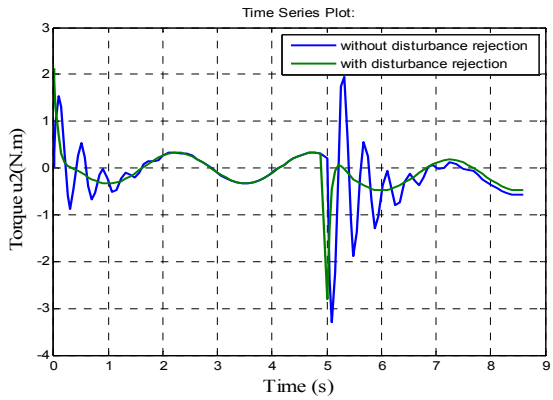


Fig.7. Input controller of second joint

According to the results, the first and second joint angular followed the desired path, and the sinusoidal disturbanceby considering the proposed method is rejected in minimum time(0.1s). Based on these results,with input disturbance in this model, the controller acts properly.

5. Conclusion

In this paper, a two-degree-of-freedom flexible robot system is modeled and the control law is designed based on the disturbance rejection. The nonlinear model is considered a linear model with feedback linearization and based on multiple inputs of input controller and disturbance effects, a controller designed to reject disturbance. By using thebackstepping control method in [15], the control input is designed.The proposed control law is simulated and analyzed in a flexible two-joint robot in MATLAB software. The simulation results show the ability of the proposed control law to reject the disturbance effect quickly and track the desired path in a minimum time.According to the method proposed in this article for flexible joint robots, this method can be used for systems with high coupling that face disturbances. Therefore, for future works, this method can be used for similar systems despite the disturbance. Another suggestion for researchers is to use a disturbance observer for the model, which can be used to observe the disturbance and reduce or eliminate its effect accordingly.

Reference

[1] L. Sun, W. Yin, M. Wang and J. Liu, "Position control for flexible joint robot based on online gravity compensation with vibration suppression", IEEE Trans. Ind. Electron., vol. 65, no. 6, pp. 4840-4848, Jun. 2018.

[2] R. Dubay, M. Hassan, C. Li and M. Charest, "Finite element based model predictive control for active vibration suppression of a one-link flexible manipulator", ISA Trans., vol. 53, no. 5, pp. 1609-19, Sep. 2014.

[3] Y. Pan, H. Wang, X. Li and H. Yu, "Adaptive command-filtered backstepping control of robot arms with compliant actuators", IEEE Trans. Control Syst. Technol., vol. 26, no. 3, pp. 1149-1156, May 2018.

[4] B. Du, S. Wu, S. Han and S. mei Cui, "Application of linear active disturbance rejection controller for sensorless control of internal permanent-magnet synchronous motor", IEEE Trans. Ind. Electron., vol. 63, no. 5, pp. 3019-3027, May 2016.

[5] X. Liu, F. Zhao, S. S. Ge, Y. Wu and X. Mei, "End-effector force estimation for flexible-joint robots with global friction approximation using neural networks", IEEE Trans. Ind. Inform., vol. 15, no. 3, pp. 1730-1741, Mar. 2019.

[6] Z. Cui, W. Li, X. Zhang, P. W. Y. Chiu and Z. Li, "Accelerated dual neural network controller for visual servoing of flexible endoscopic robot with tracking error joint motion and RCM constraints", IEEE Trans. Ind. Electron, vol. 69, no. 9, pp. 9246-9257, Sep. 2022.

[7] E. Spyrakos-Papastavridis and J. S. Dai, "Minimally model-based trajectory tracking and variable impedance control of flexible-joint robots", IEEE Trans. Ind. Electron., vol. 68, no. 7, pp. 6031-6041, Jul. 2021.

[8] V. Azimi, T. Nguyen, M. Sharifi, S. A. Fakoorian and D. Simon, "Robust ground reaction force estimation and control of lower-limb prostheses: Theory and simulation", IEEE Trans. Syst. Man Cybern.: Syst., vol. 50, no. 8, pp. 3024-3035, Aug. 2020.

[9] Z. Yan, X. Lai, Q. Meng and M. Wu, "A novel robust control method for motion control of uncertain single-link flexible-joint manipulator", IEEE Trans. Syst. Man Cybern.: Syst., vol. 51, no. 3, pp. 1671-1678, Mar. 2021.

[10] K. Rsetam, Z. Cao and Z. Man, "Cascaded-extended-state-observer-based sliding-mode control for underactuated flexible joint robot", IEEE Trans. Ind. Electron., vol. 67, no. 12, pp. 10822-10832, Dec. 2020.

[11] W. Sun, S. Su, J. Xia and V. Nguyen, "Adaptive fuzzy tracking control of flexible-joint robots with full-state constraints", IEEE Trans. Syst. Man Cybern.: Syst., vol. 49, no. 11, pp. 2201-2209, Nov. 2019.

[12] H. Hasanpour, E. Alibeiki, &, S. M. Ghadami. Flexible Robot Arm Control Using Adaptive Control Structure. Journal of Applied Dynamic Systems and Control, 5(1), 5-10, 2022.

[13] Y. Pan, X. Li and H. Yu, "Efficient PID tracking control of robotic manipulators driven by compliant actuators", IEEE Trans. Control Syst. Technol., vol. 27, no. 2, pp. 915-922, Mar. 2019.

[14] J. Kim and E. A. Croft, "Full-state tracking control for flexible joint robots with singular perturbation techniques", IEEE Trans. Control Syst. Technol., vol. 27, no. 1, pp. 63-73, Jan. 2019.

[15] H. Wang, Q. Zhang, Z. Sun, X. Tang and I.-M. Chen, "Continuous terminal sliding-mode control for FJR subject to matched/mismatched disturbances", IEEE Trans. Cybern., vol. 52, no. 10, pp. 10479-10489, Oct. 2022.

Multiple-input Single-output Nonlinear System Identification using Bezier- Bernstein Polynomials with Noise Cancellation

Mohammad Jahani Moghaddam[1*]

Abstract- This article deals with an identification method for the fractional multiple-input single-output model. It is considered the Hammerstein model to separate dynamic linear and static nonlinear behaviors. Which Bezier-Bernstein polynomials are used to approximate the nonlinear functions and the fractional order transfer function is applied to estimate the linear part. A hybrid identification method based on a modified evolutionary algorithm and a recursive classic method is presented. As an advantage, this method can also correctly identify the system in the presence of output noise. A photovoltaic experimental system and a numerical example are used to illustrate the efficiency and performance of the proposed scheme.

Keywords: Nonlinear System Identification, Multi Input- Single Output Hammerstein Model, Bezier-Bernste in Polynomial, Modified Genetic Algorithm

1. Introduction

The utilization of fractional calculations has been observed in a variety of fields in science and technology, including image processing [1,2], signal processing [3], mechanics [4], path planning and tracking [5], control theory and applications [6], and real physical systems modeling and identification [7-12], etc. This supports the notion that real plants and processes inherently possess some degree of fractionality [6], thus requiring fractional-order models to yield more accurate and satisfactory identification results compared to integer models [7]. However, for systems with wide operating areas, nonlinear models such as neural networks, Volterra series, or block-oriented models should be considered [13].

The Hammerstein model is a sequential structure comprising a static nonlinear subsystem and a dynamic linear block. It has been widely utilized to identify various nonlinear systems across different engineering problems, including chemical processes [14], DC/DC converters [15], model-based controller design [16], speech signal processing [17], electric drives [18], electrically stimulated muscles [19], RF transmitters [20], and more.

The different literatures presented various identification algorithms for Hammerstein models that correspond to different applications [13, 21-26]. The various methods differ in how they represent memory less non-linear static parts and optimization algorithms used to approximate unknown parameters. A particular approach for identifying the Hammerstein model with commensurate fractional order transfer function as the linear part is proposed in [27], assuming the differential order is a known parameter. In [28], an iterative algorithm is presented to identify the continuous fractional order of the Hammerstein system in ARX and OE forms. This method applies only to nonlinear systems with quasi-linear properties and identifies two categories of linear and non-linear unknowns through least squares-based estimation in two steps. The P-type order learning algorithm is used iteratively to estimate the order of differentiation. Moreover, [29] proposes identifying the fractional order Wiener-Hammerstein model but faces issues such as initial guesses, mathematical complications, and making a large number of data assumptions. Reference [30] suggests identifying the fractional Hammerstein model by minimizing output error using the LM algorithm. Additionally, [31] uses Particle Swarm Optimization (PSO) to minimize output error.

One possibility to represent the static part is Bezier-Bernstein polynomials that are commonly utilized in computer design and graphics [32, 33]. In [32], an identification method based on the Bezier-Bernstein approximation single-input single-output (SISO) Hammerstein systems is proposed. However, this approach has several disadvantages that hinder its use in real-world nonlinear process identification, such as computational complexity and inaccurate estimation due to non-linear coefficient estimation using the Gauss-Newton algorithm, especially in noisy data. Ref. [34] considers the Hammerstein model in a multiple-input-single-output (MISO) structure and estimates both linear and non-linear parameters using the Levenberg-Marquardt (LM)

1* Corresponding author: Department of Electrical Engineering, Langarud Branch, Islamic Azad University, Langarud, Iran.

Email: Jahani.iaul@yahoo.com

Received: 2024.01.30; Accepted: 2024.04.30

algorithm. However, this article only focuses on estimating outputs of non-linear static parts. Both of these articles are based on integer Hammerstein model.

Prior methods of identification have not been able to adequately capture the nonlinearities present in the system. This issue stems mainly from utilizing a predetermined structure such as known-order polynomials, multi-segment piecewise linear forms, etc., for describing the nonlinear behavior of the system. In practical scenarios, it is often difficult to describe the characteristics of many plants using these predefined structures. The primary objective of this research paper is to propose a novel identification scheme for MISO fractional order Hammerstein models. Unlike previous approaches, we consider the form of nonlinear functions to be unknown and estimate them during the identification process. We use Bezier-Bernstein polynomials to estimate the static nonlinear component for this purpose. It has been demonstrated that among all polynomial functions utilized in function approximation, Bernstein basis is the most stable and best conditioned [35]. This provides an additional advantage to our application. Furthermore, we employ the Fractional Order Transfer Function (FOTF) to achieve a more precise estimation of the linear dynamic component.

The other sections of this article are prepared as follow: Section 2 provides a summary of fractional order transfer function mathematics background. Section 3 gives a brief introduction to Bezier-Bernstein polynomials. Section 4 contains my proposed identification method. Section 5 includes the numerical results and finally, the conclusion is explained in detail.

2. FRACTIONAL ORDER TRANSFER FUNCTION

A fractional differential equation (1) represents a dynamical continuous-time fractional system:

$$\begin{aligned} a_n D^{\alpha_n} y(t) + a_{n-1} D^{\alpha_{n-1}} y(t) + \cdots + a_0 D^{\alpha_0} y(t) = \\ = b_m D^{\beta_m} u(t) + b_{m-1} D^{\beta_{m-1}} u(t) + \cdots + b_0 D^{\beta_0} u(t) \end{aligned} \quad (1)$$

where the coefficients and fractional orders are denoted by $a_0, a_1, ..., a_n, b_0, b_1, ..., b_m \in R$ and $\alpha_0, \alpha_1, ..., \alpha_n, \beta_0, \beta_1, ..., \beta_m \in R$, respectively. The Riemann and Liouville arbitrary order differentiation is introduced as [12]:

$$D^{\alpha} = \left(\frac{d}{dt}\right)^{\alpha} \quad (2)$$

The Grünwald-Letnikov (GL) definition is used in this paper for simplicity of calculations. For $\alpha, a \in R, \alpha > 0$, GL formula of fractional derivative of order α (3), uses sampling step h and operating limits *a* and *t* of ${}_aD_t^{\alpha} f(t)$ [12]:

$${}_aD_t^{\alpha} f(t) = \lim_{h \to 0} \frac{1}{h^{\alpha}} \sum_{j=0}^{\left[\frac{t-a}{h}\right]} (-1)^j \binom{\alpha}{j} f(t - jh) \quad (3)$$

The binomial coefficients are calculated using the relationship between the factorial and the Euler gamma function [12]:

$$\binom{\alpha}{j} = \frac{\alpha!}{j!(\alpha - j)!} = \frac{\Gamma(\alpha + 1)}{\Gamma(j + 1)\Gamma(\alpha - j + 1)} \quad (4)$$

where the Euler Gamma function $\Gamma(.)$ is defined as [12]:

$$\Gamma(n) = \int_0^{\infty} t^{n-1} e^{-t} dt \quad (5)$$

Due to the existence of long memory behavior in fractional operator, Newton's binomial $\binom{\alpha}{j}$ rate of convergence to zero with j will be very slow, and in accordance with the principle of short memory for real implementation, an approximation of equation (2) using only recently past values of f(t) is defined as the simplest solution for the simulation of fractional systems in the time domain [36].

$$D^{\alpha} f(t) = \frac{1}{h^{\alpha}} \sum_{j=0}^{N} (-1)^j \binom{\alpha}{j} f(t - jh) \quad (6)$$

where $N = [T/h]$ is the approximation addends number [36].

The FOTF corresponding to Equation (1) is:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_m s^{\beta_m} + b_{m-1} s^{\beta_{m-1}} + \cdots + b_0 s^{\beta_0}}{a_n s^{\alpha_n} + a_{n-1} s^{\alpha_{n-1}} + \cdots + a_0 s^{\alpha_0}} \quad (7)$$

3. BEZIER- BERNSTEIN POLYNOMIALS

The expanding of $[x + (1-x)]^d$ is formed the univariate Bernstein polynomial basis functions $B_{i,d}(x)$ defined as [33]:

$$B_{i,d}(x) = \binom{d}{i} x^i (1-x)^{d-i} \quad (8)$$

where *i* and *d* are nonnegative integers $i \le d$ over the region $x \in [0,1]$. The number of the single-variable Bernstein polynomials of order *d* is d + 1. Also, Bernstein polynomials are computed recursively:

$$B_{i,d}(x) = (1-x) B_{i,d-1}(x) + x B_{i-1,d-1}(x) \quad (9)$$

Bezier-Bernstein polynomials over the region [0,1] are defined as:

$$B_{i,n} f(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} f\left(\frac{i}{n}\right), \quad n=1,2,.... \quad (10)$$

The Bezier-Bernstein polynomials over the [a, b] domain, which a, b are arbitrary numbers, are defined as:

$$B_{i,n}(t)=\binom{n}{i}\frac{(t-a)^{i}(b-t)^{n-i}}{(b-a)^{n}},\ \text{i=0,1,2,....,n} \quad (11)$$

4. THE IDENTIFICATION ALGORITHM

4.1. The Problem Statement

The fractional MISO Hammerstein model is constructed from a series of the nonlinear gain functions and FOTF as the dynamic part (Fig. 1) and is generally specified by:

$$y(t)=-\sum_{i=1}^{n}a_i D^{\alpha_i}y(t)+\sum_{j=0}^{m}b_j D^{\beta_j}(w_1(t)+...+w_k(t)) \quad (12)$$

where $w_k(t), k=1,2,...,r$ are the nonlinear gain functions outputs that entered to dynamic subsystem. The input and output delays for the dynamic part are denoted by n, m . In this paper, the nonlinear blocks are modelled as:

$$w_k(t)=f_k\left(u_k(t)\right)=\sum_{i=0}^{d}\delta_{ki}B_{ki,d}(u_k(t)) \quad (13)$$

where y(t) represents the system output, and $u_1,u_2,...,u_r$ represents the inputs. The Bezier-Bernstein polynomial associated to *k*th inputs are represented by $B_{ki,d}(u_k(t)), i=0,...,d, k=1,...,r$ and the corresponding weights are represented by $\delta_{ki}, i=0,1,\ldots,d, k=1,...,r$. Moreover, the dynamic part gain is determined by coefficients $a_1,...,a_n,b_0,...,b_m\in R$ and fractional orders $\alpha_1,...,\alpha_n,\beta_0,...,\beta_m\in R$ as:

$$G(s)=\frac{B(s)}{A(s)}=\frac{b_m s^{\beta_m}+\cdots+b_1 s^{\beta_1}+b_0 s^{\beta_0}}{1+a_n s^{\alpha_n}+\cdots+a_1 s^{\alpha_1}} \quad (14)$$

Based on equations (12), (13), the relationship between input and output can be expressed using the following equation:

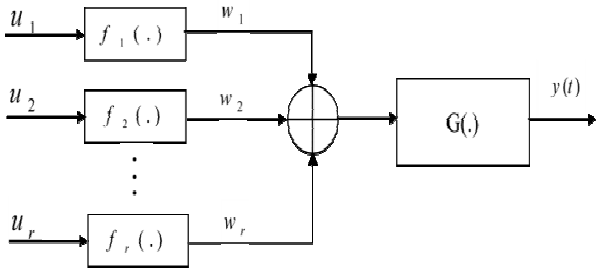


Fig. 1. The MISO Hammerstein model

$$y(t)=-\sum_{i=1}^{n}a_i D^{\alpha_i}y(t)+\sum_{j=0}^{m}b_j D^{\beta_j}(\sum_{k=0}^{d}\left(\delta_{1k}B_{1k,d}(u_1(t))+...\right.$$
$$\left.\ldots+\delta_{rk}B_{rk,d}(u_r(t))\right)) \quad (15)$$

Equation (15) is rewritten using the GL definition of fractional difference-integral as shown in equations (16) to (18) [36]:

$$y(k+1)=-\sum_{i=1}^{n}a_i' Y_i(k)+$$
$$+\sum_{j=0}^{m}\left(\sum_{g=0}^{d}b'_{1jg}(F_{1jg}(k))+...+\sum_{g=0}^{d}b'_{rjg}(F_{rjg}(k))\right) \quad (16)$$

$$a_i'=\frac{\frac{a_i}{h^{\alpha_i}}}{1+\sum_{k=1}^{n}\frac{a_k}{h^{\alpha_k}}},\quad b'_{wjg}=\frac{\frac{b_j(\delta w)_g}{h^{\beta_j}}}{1+\sum_{k=1}^{n}\frac{a_k}{h^{\alpha_k}}}, \quad (17)$$

$$1\le \text{i}\le \text{n},\ 1\le w\le r,\ 0\le g\le \text{d},\ 1\le j\le m$$

$$Y_i(k)=\sum_{j=1}^{N}(-1)^j\binom{\alpha_i}{j}y(k+1-j),$$

$$\text{F}_{\text{w}jg}(k)=\sum_{i=0}^{N}(-1)^i\binom{\beta_i}{j}\left(B_{wg,d}(u_w(k))\right), \quad (18)$$
$$,1\le w\le r, 0\le j\le d$$

By considering the actual inputs (*u(t)*) and measured outputs ($y^*(t)=y(t)+p(t)$), and defining Gaussian random noise $p(t)$ with zero mean and variance σ^2 , the equation (16) can be further rewritten as equation (19) [36].

$$y^*(k+1)=-\sum_{i=1}^{n}a_i' Y_i^*(k)+$$
$$+\sum_{j=0}^{m}\left(\sum_{g=0}^{d}b'_{1jg}(F_{1jg}(k))+...+\sum_{g=0}^{d}b'_{rjg}(Fr_{jg}(k))\right)+ \quad (19)$$
$$+e(k+1)$$

where:

$$Y_i^*(k)=\sum_{j=1}^{N}(-1)^j\binom{\alpha_i}{j}y^*(k+1-j) \quad (20)$$

$$e(k+1)=p(k+1)+\sum_{i=1}^{n}a_i'\sum_{j=1}^{N}(-1)^j\binom{\alpha_i}{j}p(k+1-j) \quad (21)$$

The unknown parameters in this model include the fractional orders $\alpha_1,...,\alpha_n,\beta_0,...,\beta_m$ and the coefficients of the transfer function $a_1,...,a_n,b_0,...,b_m$ and the weights of Bezier- Bernstein polynomials

$\delta_{ki}, i = 0,1,\ldots,d, k = 1,...,r$. To obtain these unknown parameters, this paper proposes the Modified Genetic Algorithm (MGA) [37] to identify the fractional orders α_i, β_j and provide an initial estimation for other unknown parameters a_i, b_j, δ_k . A recursive optimization algorithm is then used to update these other unknowns.

The cost function for this approach is the mean of squared output errors, as given in equation (22), where $y(k)$ is the real plant output and $y_{est}(k,\theta)$ is the estimated output, and *p* is the number of data samples.

$$e(\theta) = mean(y(k) - y_{est}(k,\theta))^2, k = 0,1,\ldots,p \quad (22)$$

Once the fractional orders are obtained, the recursive algorithm is applied to update the unknown coefficients and weights (a_i, b_j, δ_k) accordingly.

4.3. RLS Optimization Approach

The standard RLS algorithm requires an initial set of value samples to obtain the unique solution of the unknown parameter vector θ , which is determined by the dimension of the regressor ϕ . Equation (19) in the RLS form can be expressed as equation (23) which is linear with respect to the coefficients, where θ includes unknown parameters and $\phi(k)$ is the known vectors in *k*th step as shown in equations (24) and (25).

$$y(k+1) = \theta^T \phi(k) + e(k+1) \quad (23)$$

$$\theta = \left[a_1', \ldots, a_n', b_{100}', \ldots, b_{10d}', \ldots, b_{1m0}', \ldots, b_{1md}', \ldots, b_{r00}', \ldots, b_{r0d}', \ldots, b_{rm0}', \ldots, b_{rmd}' \right]^T \quad (24)$$

$$\phi(k) = \left[-Y_1^*(k), \ldots, -Y_n^*(k), F_{100}(k), \ldots, F_{10d}(k), \ldots, F_{1m0}(k), \ldots, F_{1md}(k), \ldots, F_{r00}(k), \ldots, F_{r0d}(k), \ldots, F_{rm0}(k), \ldots, F_{rmd}(k) \right]^T \quad (25)$$

The LS method provides a solution as equation (26) if the matrix $\left[\sum_{i=1}^{k} \phi(i-1)\phi^T(i-1) \right]^{-1}$ exists, but the matrix ϕ may become poorly conditioned or singular, making it difficult to calculate its inverse. To avoid this issue, MGA is used to obtain the required RLS initial values without needing to compute any inverses.

$$\hat{\theta}_k = \left[\sum_{i=1}^{k} \phi(i-1)\phi^T(i-1) \right]^{-1} \sum_{i=1}^{k} \phi(i-1) y(i) \quad (26)$$

Equation (27) is the recursive version of equation (26) and can be extended to online identification applications [36, 38, 39], where optimization results are updated in each iteration using new measured input/output.

$$\begin{cases} \hat{\theta}_{k+1} = \hat{\theta}_k + G_k \phi(k) \varepsilon(k+1) \\ G_{k+1} = G_k - \dfrac{G_k \phi(k) \phi^T(k) G_k}{1 + \phi^T(k) G_k \phi(k)} \\ \varepsilon(k+1) = \dfrac{y(k+1) - \hat{\theta}_k^T \phi(k)}{1 + \phi^T(k) G_k \phi(k)} \end{cases} \quad (27)$$

The adaptation gain matrix G for an initial value is generally selected according to equation (28) [36].

$$G_0 = \frac{1}{\gamma} I; \quad 0 < \gamma \ll 1 \quad (28)$$

The coefficients $a_i', i = 1,...,n$ are obtained from $\hat{\theta}$ directly and the assumption $b_0 = 1$ is used to calculate the values of $b_j \left(\delta_w\right)_g$, $1 \le w \le r$, $0 \le g \le d$, $1 \le j \le m$.

MGA estimates fractional orders ($\alpha_1,...,\alpha_n, \beta_0,...,\beta_m$) and produces initial estimations of other unknown parameters including $a_i', i = 1,...,n$ and $b_{w\,jg}', 1 \le w \le r, 0 \le g \le d, 1 \le j \le m$.

Where the multiplication of b_j s and $\left(\delta_w\right)_g$ s results in the unknown parameter beta $b_{w\,jg}'$ s. These estimations are frequently updated by RLS. The $b_j, 1 \le j \le m$ coefficients and $b_{w\,jg}'$ s are then used to calculate the weights of Bezier Bernstein polynomials $\left(\delta_w\right)_g, 1 \le w \le r, 0 \le g \le d$. If the stop conditions are satisfied, the identification process is complete. If not, RLS is applied again to reduce the estimation error. The identification process will continue until the desired accuracy is achieved.

5. SIMULATION RESULTS

To verify the effectiveness of the proposed approach, two examples are presented. The FOTF's general form is defined as the dynamic subsystem by equation (29).

$$G(s) = \frac{b_1 s^{\beta_1} + s^{\beta_0}}{a_2 s^{\alpha_2} + a_1 s^{\alpha_1} + a_0} \quad (29)$$

5.1. Example 1

In the first example, the real Photo-Voltaic (PV) module MF120EC3 is from Pusat Tenaga Malaysia (Malaysia Energy Centre, MEC) is taken into consideration [40].The inputs are Solar Irradiance (Ir) and Cell Temperature (CT), while the output is DC Current (DCC).The transfer function of the MISO PV system is expressed in equations (30) and (31) [40]. The Tustin estimation operator is used to obtain

the continuous versions of the transfer functions given by equations (32) and (33).

$$TF_{CT/DCC} = = \frac{.01358z^4 - .02393z^3 - .001419z^2 + .01159z}{z^4 - 1.148z^3 + 0.251z^2 - .01503z + .0381} \quad (30)$$

$$TF_{Ir/DCC} = = \frac{.05644z^4 - .06703z^3 - .002245z^2 + .00039z}{z^4 - 1.148z^3 + .251z^2 - .01503z + .0381} \quad (31)$$

$$TF_{CT/DCC} = = \frac{.009992s^4 + .1022s^3 + .1375s^2 - .05455s - .001168}{s^4 + 4.986s^3 + 9.341s^2 + 5.16s + .8226} \quad (32)$$

$$TF_{Ir/DCC} = = \frac{.04928s^4 + .2941s^3 + .5597s^2 + .2966s - .0812}{s^4 + 4.986s^3 + 9.341s^2 + 5.16s + .8226} \quad (33)$$

The proposed method involves generating a simulated MISO Hammerstein model using equations (32) and (33) and adding Gaussian noise $p(t) \in N(0,\sigma^2)$ with two various $\sigma^2 = 0.0001$ and $\sigma^2 = 0.01$ to the output. This method effectively cancels out the output noise. The inputs *CT* and *Ir* have uniform distributions within the ranges of $[20,80]$ and $[0,1000]$, respectively. To achieve optimal accuracy, MGA was utilized to determine the appropriate degree for the Bernstein basis function, starting from 3. Eventually, the polynomial degree was set to d = 5, using Equation (10) for the Bernstein polynomials, resulting in the generation of a sequence of regressors $B_{i,n}(u(t)), i = 0,1,...,5$ .This generates a sequence of regressors and Bezier-Bernstein polynomial basis functions for each input, with 5 knots set for each input. For the first input, we set 5 knots at [20, 35, 50, 65, 80], and for the second input, we selected another set of 5 knots at [0, 250, 500, 750, 1000]. These knots generated 5 Bezier-Bernstein polynomial basis functions for each input that are depicted in Figs. 2- 5.

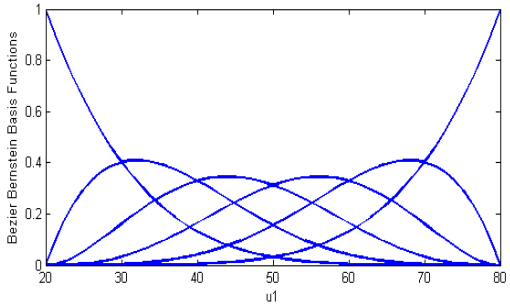


Fig. 2. Six Bezier–Bernstein polynomial basis functions constructed in Example 1 over the first input data u1

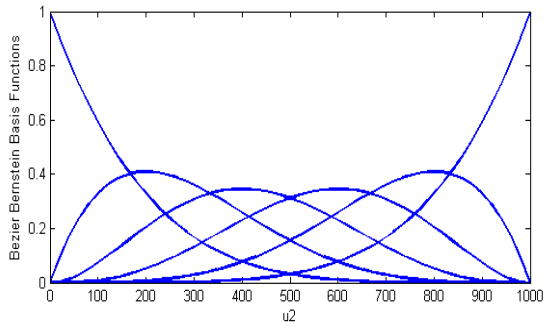


Fig. 3. Six Bezier–Bernstein polynomial basis functions constructed in Example 1 over the second input data u2.

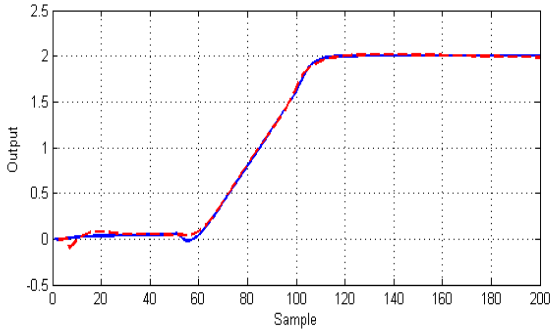


Fig. 4. The actual (blue line) and the estimated (red dashed line) outputs- PV Module ($\sigma^2 = 0.0001$)

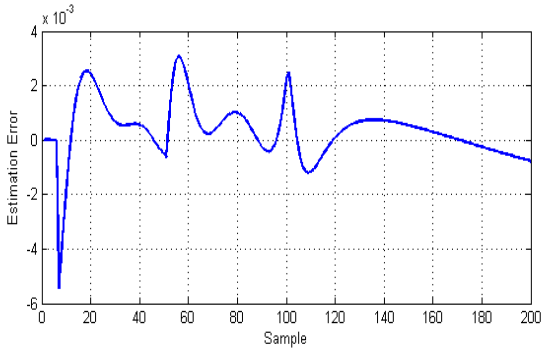


Fig. 5. The output estimation error- Example 1 ($\sigma^2 = 0.0001$)

The algorithm was iterated 200 times, reaching a Mean Squared Errors (MSE) value of for the clean output and for the noisy output. The estimated outputs and related errors are shown in Figs. 6, 7 for low noise and Figs. 9, 10 for high noise cases, demonstrating improved accuracy compared to reference [33]. The coefficients estimation using RLS is illustrated in Figs. 8, 11. The identified FOTFs are presented as (34) for low output noise and (35) for high output noise.

The algorithm was iterated over 200 cycles, resulting in a Mean Squared Error (MSE) value of 1.1857×10^{-6} for the low noise case and 2.4803×10^{-5} for the high noise measurements i.e., the noisy case. The estimated outputs and their corresponding estimation errors can be seen in Figs. 6, 7 and Figs. 9, 10 for the low and high noise cases, respectively. These results demonstrate an improvement in accuracy compared to the reference [33]. The coefficient

estimation process using RLS is shown in Figs. 8, 11 for both noise levels. The identified FOTFs for the low and high output noise cases are represented by Equations (34) and (35), respectively.

$$G(s)=\frac{s^{1.3461}-0.6257s^{0.0439}}{s^{1.2176}+2.5644s^{0.8157}+0.5815} \quad (34)$$

$$G(s)=\frac{2.4708s^{1.6082}+s^{0.0048}}{2.6007s^{0.9297}+s^{0.6486}+1.3365} \quad (35)$$

Tables 1 and 2 showcase a comparative analysis between fractional order and integer order transfer functions at two different levels of noise. Additionally, Table 2 compares the performance of MGA, classic GA, and PSO. The estimated outputs and corresponding errors using an integer order transfer function can be observed in Figs. 12-15. To prove the robustness of our proposed approach against input uncertainties, we introduced Gaussian noise $p(t)\in N(0,\sigma^2)$ with a standard deviation of $\sigma^2=0.01$ to the inputs of the Hammerstein model. Table 2 also presents the simulation results of this case, supporting our assertion that FOTF and MGA lead to superior identification accuracy.

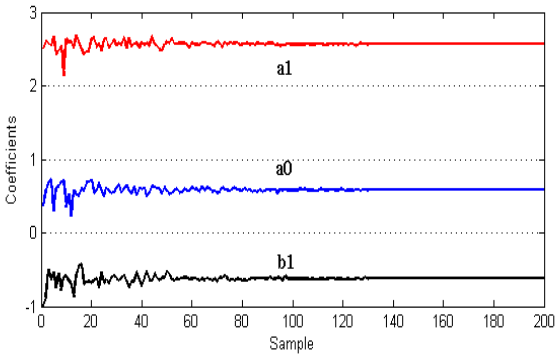


Fig. 6. RLS coefficients estimation process-PV Module ($\sigma^2=0.0001$)

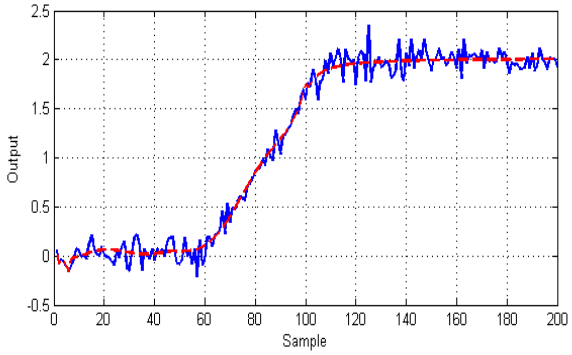


Fig. 7. The actual (blue line) and the estimated (red dashed line) outputs-PV Module ($\sigma^2=0.01$)

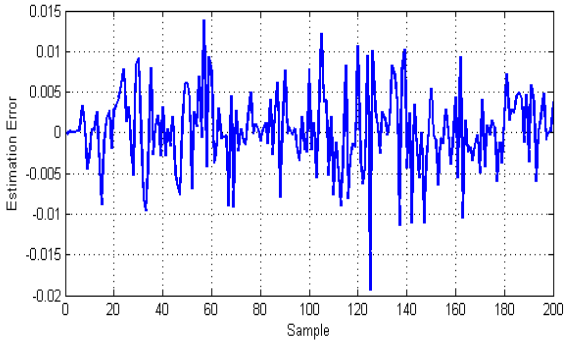


Fig. 8. The output estimation error-PV Module ($\sigma^2=0.01$)

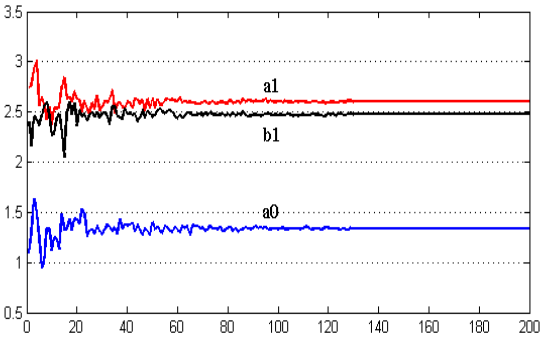


Fig. 9. RLS coefficients estimation process-PV Module ($\sigma^2=0.01$)

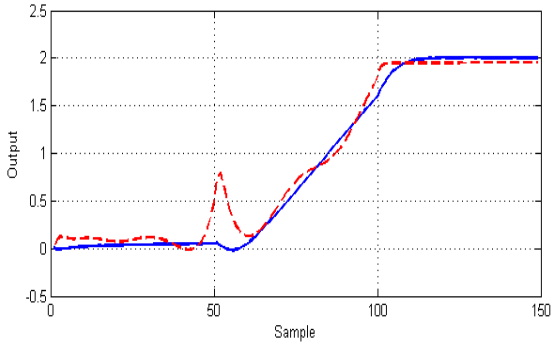


Fig. 10. The actual (blue line) and the estimated (red dashed line)outputs-integer order transfer function- PV Module($\sigma^2=0.0001$)

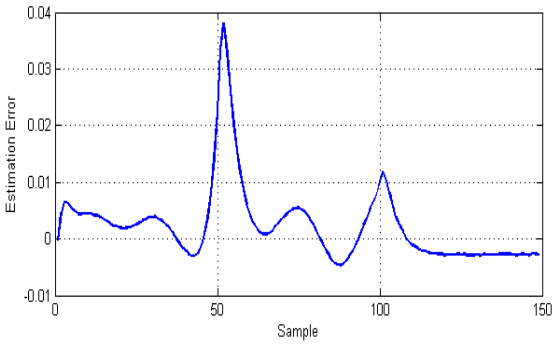


Fig. 11. The output estimation error-integer order transfer function -PV Module($\sigma^2=0.0001$)

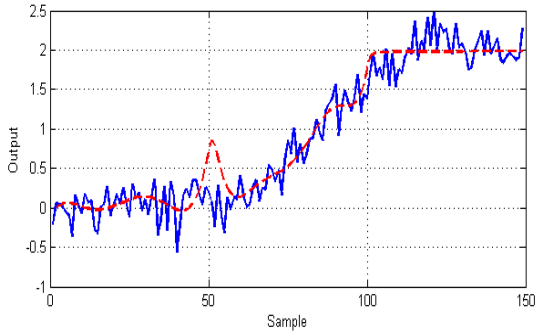


Fig. 12. The actual (blue line) and the estimated (red dashed line)outputs- integer order transfer function- PV Module ($\sigma^2 = 0.01$)

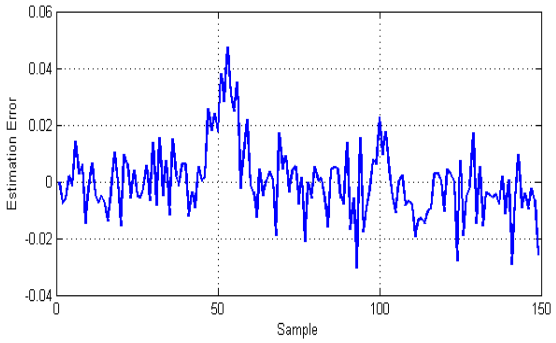


Fig. 13.The output estimation error- integer order transfer function -PV Module($\sigma^2 = 0.01$)

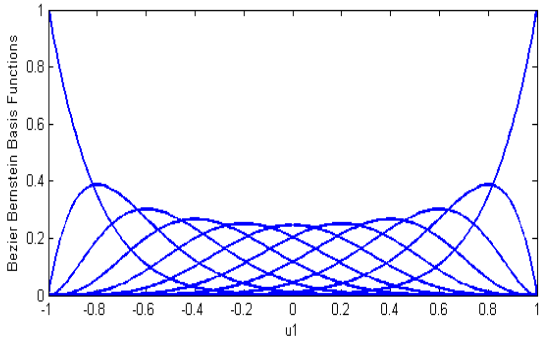


Fig. 14. Eleven Bezier–Bernstein polynomial basis functions constructed in Example 2 over the first input data u1

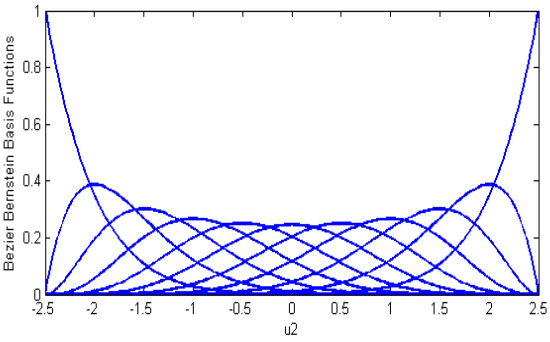


Fig. 15. Eleven Bezier–Bernstein polynomial basis functions constructed in Example 2 over the second input data u2

Table 1. The estimation error comparison between two different linear parts in PV Module ($\sigma^2 = 0.0001$).

Approach	MSE	RMS
Integer Order Transfer Function	4.1182×10^{-5}	0.0073
Fractional Order Transfer Function	1.1857×10^{-6}	0.0018

Table 2. The estimation error comparison between different optimization approaches and with two different linear parts in PV Module ($\sigma^2 = 0.01$).

Approach			MSE	RMS
PSO & RLS			$1.3886\times10^{-}$	0.011723
GA & RLS			1.2601×10^{-4}	0.011637
MGA & RLS	Integer Order Transfer Function		1.3777×10^{-4}	0.012102
	Fractional Order Transfer Function	Without Input Noise	$2.4803\times10^{-}$	0.004305
		With Input Noise	$5.9813\times10^{-}$	0.008000

5.2. Example 2

The Steam-Water Heat Exchanger (SWHE) nonlinear gain function [34] is used as a numerical example in the second case. Eq. (36) describes the SWHE which has been studied previously in [41], whereas $f_2(u)$ in Eq. (37) represents its nonlinear gain function. To create a MISO system, $f_1(u)$ is incorporated into the system model. The function $sat(x, x_-, x_+)$ is the saturation function with respective left and right breaking points in x_- and x_+ [34].

$$y(t) = 1.608y(t-1) - 0.6385y(t-2) + 0.3f_1(u_1(t)) + \\ + 0.207f_2(u_2(t-1)) - 0.1764f_2(u_2(t-2)) + p(t) \quad (36)$$

$$f_1(u) = sat(u, -0.6, 0.6)$$

$$f_2(u) = -31.549u + 41.732u^2 - 24.201u^3 + 68.634u^4 \quad (37)$$

Gaussian noise $p(t) \in N(0, \sigma^2)$ with different $\sigma^2 = 0.0001$ and $\sigma^2 = 0.01$ values is added to the Hammerstein model to demonstrate the proposed method's ability to eliminate output noise. The intervals [-1, 1] and [-2.5, 2.5] set the input limits for u_1 and u_2, respectively. In order to achieve accurate results, MGA is implemented to

determine the appropriate degree of Bernstein basis functions. Consequently, 11 knots are selected for each input, producing 11 Bezier-Bernstein polynomial basis functions for each input as demonstrated in Figs. 16 and 17. The 11 knots for the first input was set as [-1, -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8, 1] and a set of 11 knots for the second input were selected as [-2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5]. Finally, one hundred input/output data samples are generated using Eqs. (36) and (37).

And as a result, aset of In the case of $\sigma^2 = 0.0001$, MSE after 100 iteration cycles became 1.8646×10^{-5}. The estimation of dynamic subsystem is obtained as:

$$G(s) = \frac{0.2880s^{0.5594} + s^{0.8616}}{3.9933s^{2.3921} + 4.7773s^{0.4164} + 2.4674} \quad (38)$$

In the case of a rather noisy data, i.e. $\sigma^2 = 0.01$, after 100 iterations, the MSE value was equaled to 5.3781×10^{-5}. The dynamic transfer function is obtained as:

$$G(s) = \frac{4.6881s^{0.3307} + s^{0.0526}}{2.6694s^{3.4542} + 4.5400s^{0.5089} + 2.1056} \quad (39)$$

Figures 16-19 display the estimation accuracy of nonlinear static parts and corresponding errors. Additionally, Figures 21 and 23 provide a visual representation of outputs approximated with low and high noise levels, respectively, with their related errors depicted in Figures 22 and 24. Furthermore, Figures 20 and 25 illustrate coefficients estimation using RLS for two distinct amounts of noise.

For this particular experiment, a comparison was carried out between fractional order and integer order transfer functions as linear dynamic parts. Tables 3 and 4 present results for two different noise levels, with estimated outputs and corresponding errors depicted in Figures 26-29. The optimization method employed significantly improved estimation accuracy compared to the GA and PSO methods, as showcased in Table 4. Moreover, the proposed method's robustness against input uncertainty was demonstrated by introducing Gaussian noise $p(t) \in N(0,\sigma^2)$ with standard deviation $\sigma^2 = 0.01$ to each model input, as shown in Table 4. Finally, Tables 3 and 4 highlight the superior identification accuracy achieved using FOTF and MGA.

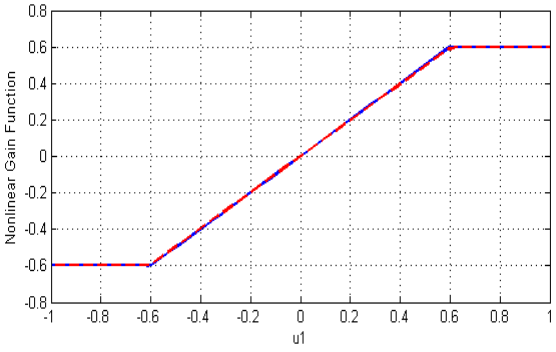


Fig. 16. The actual (blue line) and the estimated (red dashed line)nonlinear gain functions associated with the first input in low noise condition(SWHE)

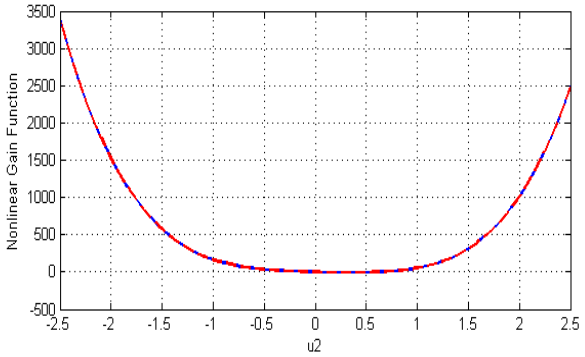


Fig. 17. The actual (blue line) and the estimated (red dashed line) nonlinear gain functions associated with the second input in low noise condition (SWHE)

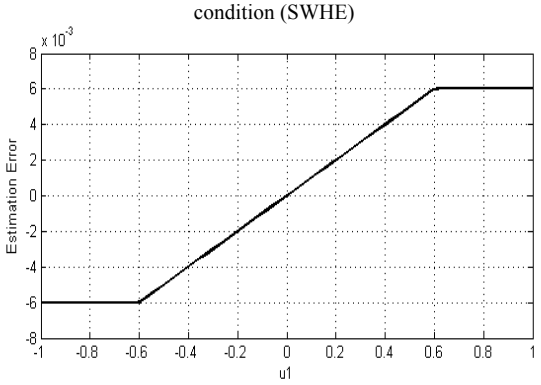


Fig. 18. Corresponding estimation errors regarding the first nonlinear function (SWHE)

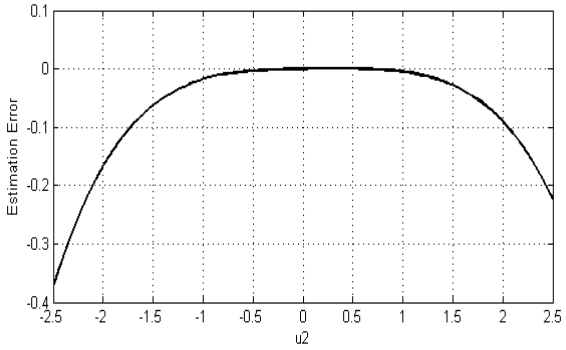


Fig. 19. Corresponding estimation errors regarding the second nonlinear function (SWHE)

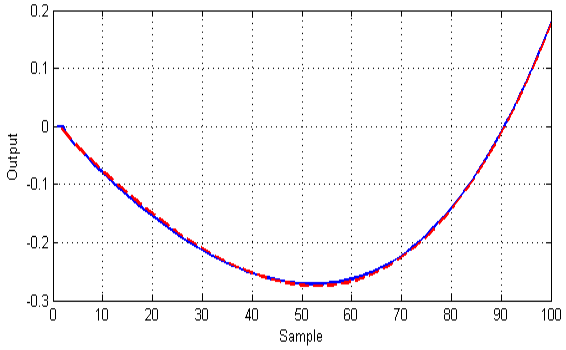


Fig. 20. The actual output (blue line) and the estimated output (red dashed line)-SWHE ($\sigma^2 = 0.0001$)

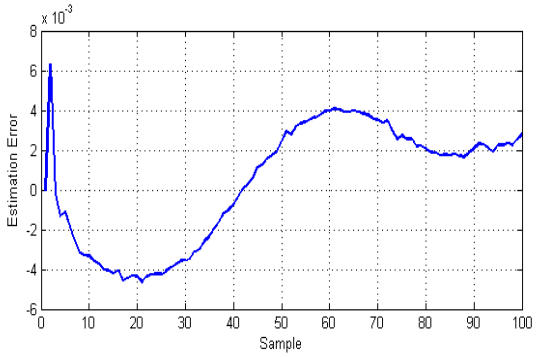


Fig. 21. The output estimation error- SWHE ($\sigma^2 = 0.0001$)

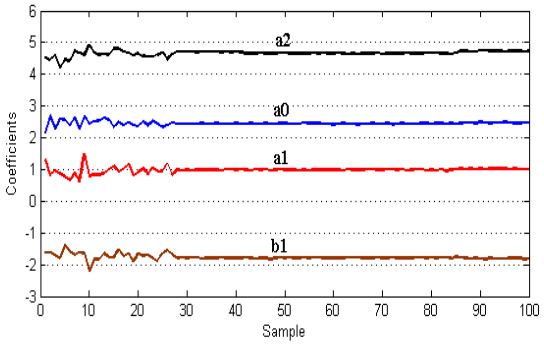


Fig. 22. RLS coefficients estimation regarding SWHE ($\sigma^2 = 0.0001$)

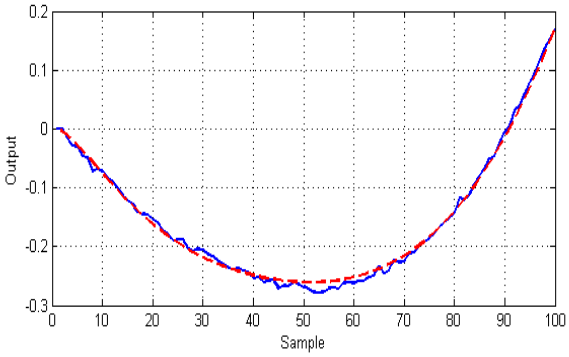


Fig. 23. The actual (blue line) and the estimated (red dashed line) outputs regarding SWHE ($\sigma^2 = 0.01$)

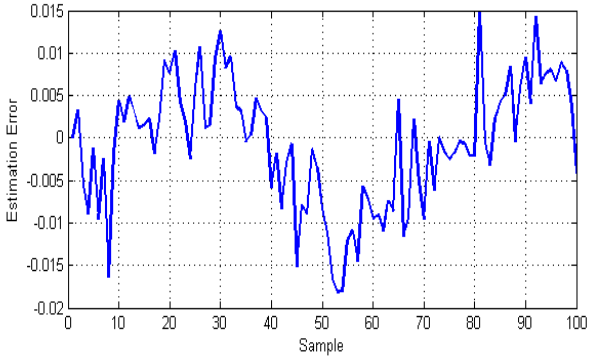


Fig. 24. The output estimation error-SWHE ($\sigma^2 = 0.01$)

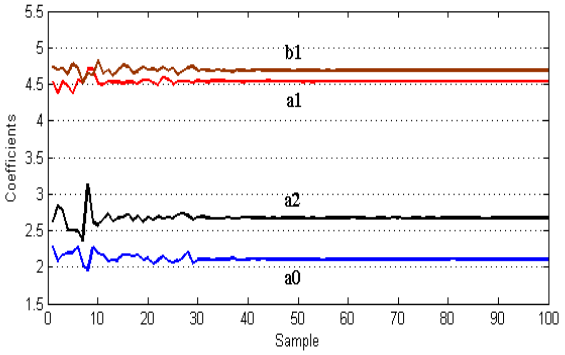


Fig. 25. RLS coefficients estimation regarding SWHE ($\sigma^2 = 0.01$)

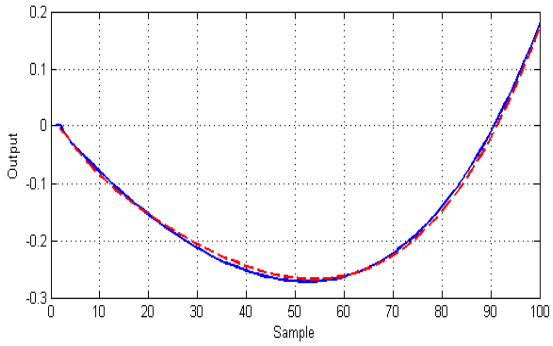


Fig. 26. The actual (blue line) and the estimated (red dashed line) outputs-integer order transfer function-SWHE ($\sigma^2 = 0.0001$)

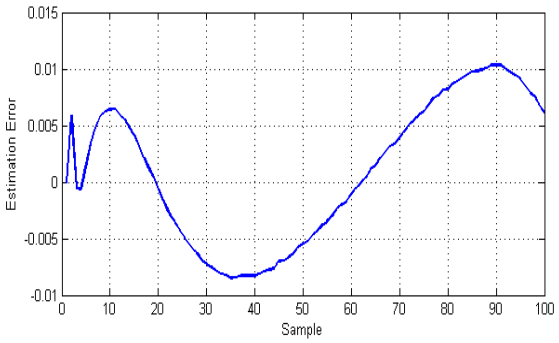


Fig. 27. The output estimation error- integer order transfer function-SWHE($\sigma^2 = 0.0001$)

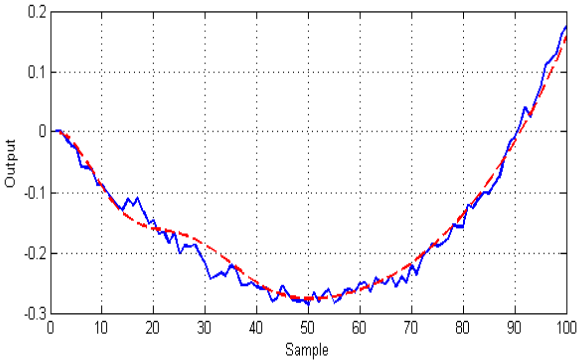


Fig. 28. The actual (blue line) and the estimated (red dashed line) outputs-integer order transfer function-SWHE($\sigma^2 = 0.01$)

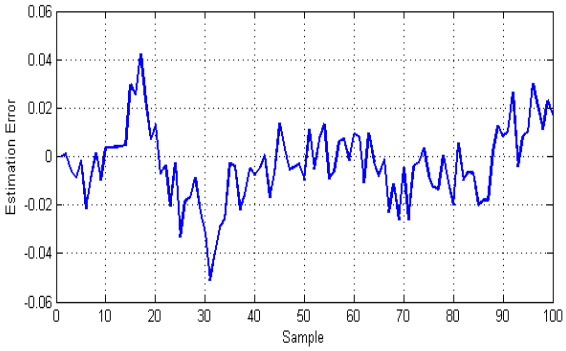


Fig. 29. The output estimation error- integer order transfer function-SWHE ($\sigma^2 = 0.01$)

Table 3. The estimation errorcomparison between two different linear parts in SWHE ($\sigma^2 = 0.0001$).

Approach	MSE	RMS
Integer Order Transfer Function	3.5895×10^{-4}	0.0104
Fractional Order Transfer Function	1.8646×10^{-5}	0.0019

Table 4. The output estimation errorcomparison between two different optimization approaches and with two different linear parts in SWHE ($\sigma^2 = 0.01$).

Approach			MSE	RMS
PSO & RLS			3.2069×10^{-4}	0.0151
GA & RLS			1.7230×10^{-4}	0.0122
MGA & RLS	Integer Order Transfer Function		5.7775×10^{-4}	0.0157
	Fractional Order Transfer Function	Without Input Noise	5.3781×10^{-5}	0.0054
		With Input Noise	8.2548×10^{-5}	0.0088

6. CONCLUSIONS

The study presented a novel approach for the identification of Multiple-Input Single-Output (MISO) fractional order Hammerstein models using Bezier-Bernstein polynomials with noise cancellation. Through comprehensive simulation results and analysis on the presented PV module and SWHE case studies, we have demonstrated the effectiveness and robustness of the proposed method in accurately capturing the complex nonlinear behaviors of the systems while mitigating the impact of output noise. The achieved results showcase superior identification performance and noise cancellation capabilities, laying the foundation for advancements in nonlinear system identification methodologies.

Looking ahead, future research endeavors could extend this method to more diverse and complex practical systems, exploring its applicability in real-world scenarios. Furthermore, the application of this method in other domains such as control systems, robotics, and industrial processes presents an exciting avenue for further investigation. Additionally, research efforts could focus on refining the algorithm's efficiency and extending its capabilities to address broader classes of nonlinear systems, thus contributing to a more comprehensive understanding and utilization of fractional order models in practical engineering applications

Statements & Declarations

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability Statement

The examples data are available in references [34, 40].The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

Study concept and design; analysis and interpretation of data**;** drafting of the manuscript**;** critical revision of the manuscript for important intellectual content; and statistical analysis are all done by the corresponding author.

REFERENCES

[1] Birs, I. R., Muresan, C. I., Folea, S., & Prodan, O. (2016). A comparison between integer and fractional order pd controllers for vibration suppression. Applied Mathematics and Nonlinear Sciences, 1(1), 273-282.

[2] Yang Q, Chen D, Zhao T, Chen Y. Fractional calculus in image processing: a review. Fractional Calculus and Applied Analysis. 2016 Oct 1;19(5):1222-49.

[3] Sheng H, Chen Y, Qiu T. An Overview of Fractional Processes and Fractional-Order Signal Processing Techniques. In Fractional Processes and Fractional-Order Signal Processing 2012 (pp. 31-46). Springer London. ISBN: 978-1-4471-2232-6

[4] Baleanu D, Golmankhaneh AK, Nigmatullin R, Golmankhaneh AK. Fractional newtonian mechanics. Central European Journal of Physics. 2010 Feb 1;8(1):120-125.

[5] Yousfi N, Melchior P, Rekik C, Derbel N, Oustaloup A. Path tracking design by fractional prefilter using a combined QFT/H∞ design for TDOF uncertain feedback systems. J. Appl. Nonlinear Dyn. 2012;1(3):239-61.

[6] Monje CA, Vinagre BM, Feliu V, Chen Y. Tuning and auto-tuning of fractional order controllers for industry applications. Control engineering practice. 2008 Jul 31;16(7):798-812.

[7] Gabano JD, Poinot T, Kanoun H. Identification of a thermal system using continuous linear parameter-varying fractional modelling. IET Control Theory & Applications. 2011 May 5;5(7):889-99.
[8] Cois O, Oustaloup A, Battaglia E, Battaglia JL. Non integer model from modal decomposition for time domain system identification. IFAC Proceedings Volumes. 2000 Jun 30;33(15):989-94.
[9] Poinot T, Trigeassou JC. Identification of fractional systems using an output-error technique. Nonlinear Dynamics. 2004 Dec 1;38(1):133-54.
[10] Lin J, Poinot T, Li ST, Trigeassou JC. Identification of non-integer-order systems in frequency domain. Journal Control Theory Appl.2008;25:517-20.
[11] Valerio D, Sa da Costa J. Identifying digital and fractional transfer functions from a frequency response. International Journal of Control. 2011 Mar 1;84(3):445-57.
[12] Petras I. Fractional-order nonlinear systems: modeling, analysis and simulation. Springer Science & Business Media; 2011 May 30. ISBN: 978-3-642-18100-9.
[13] Ivanov DV. Identification discrete fractional order Hammerstein systems. In Control and Communications (SIBCON), 2015 International Siberian Conference on 2015 May 21 (pp. 1-4). IEEE.
[14] Rattanawaorahirunkul R, Sanposh P, Panjapornpon C. Nonlinear system identification of pH process using Hammerstein-Wiener model. InElectronics, Information, and Communications (ICEIC), 2016 International Conference on 2016 Jan 27 (pp. 1-4). IEEE.
[15] Alonge F, D'Ippolito F, Raimondi FM, Tumminaro S. Nonlinear modeling of DC/DC converters using the Hammerstein's approach. IEEE transactions on power electronics. 2007 Jul;22(4):1210-21.
[16] Huo HB, Zhu XJ, Hu WQ, Tu HY, Li J, Yang J. Nonlinear model predictive control of SOFC based on a Hammerstein model. Journal of Power Sources. 2008 Oct 15;185(1):338-44.
[17] Turunen J, Tanttu JT, Loula P. Hammerstein model for speech coding. EURASIP Journal on Applied Signal Processing. 2003 Jan 1;2003:1238-49.
[18] Balestrino A, Landi A, Ould-Zmirli M, Sani L. Automatic nonlinear auto-tuning method for Hammerstein modeling of electrical drives. IEEE Transactions on Industrial Electronics. 2001 Jun;48(3):645-55.
[19] Le F, Markovsky I, Freeman CT, Rogers E. Recursive identification of Hammerstein systems with application to electrically stimulated muscle. Control Engineering Practice. 2012 Apr 30;20(4):386-96.
[20] Taringou F, Hammi O, Srinivasan B, Malhame R, Ghannouchi F M, Behaviour modelling of wideband RF transmitters using Hammerstein–Wiener models, IET circuits, devices & systems, 2010, 4(4), 282--290.
[21] Bai EW, Fu M. A blind approach to Hammerstein model identification. IEEE Transactions on Signal Processing. 2002 Jul;50(7):1610-9.
[22]Hasiewicz Z, Mzyk G. Hammerstein system identification by non-parametric instrumental variables. International Journal of Control. 2009 Mar 1;82(3):440-55.
[23] Greblicki W. Stochastic approximation in nonparametric identification of Hammerstein systems. IEEE Transactions on Automatic Control. 2002 Nov;47(11):1800-10.
[24] Chen H F, Pathwise convergence of recursive identification algorithms for Hammerstein systems, IEEE Transactions on Automatic Control, 2004, 49(10), 1641--1649.
[25] Goethals I, Pelckmans K, Suykens JA, De Moor B. Subspace identification of Hammerstein systems using least squares support vector machines. IEEE Transactions on Automatic Control. 2005 Oct;50(10):1509-19.
[26] Mete S, Ozer S, Zorlu H. System identification using Hammerstein model optimized with differential evolution algorithm. AEU-International Journal of Electronics and Communications. 2016 Dec 31;70(12):1667-75.
[27] Mohamed AO, Malti R, Olivier CO, Oustaloup A. System Identification Using Fractional Hammerstein Models. IFAC Proceedings Volumes. 2002 Dec 31;35(1):265-9.
[28] Zhao Y, Li Y, Chen Y. Complete parametric identification of fractional order Hammerstein systems. InFractional Differentiation and Its Applications (ICFDA), 2014 International Conference on 2014 Jun 23 (pp. 1-6). IEEE.
[29] Vanbeylen L. A fractional approach to identify Wiener–Hammerstein systems. Automatica. 2014 Mar 31;50(3):903-9.
[30] Hammar K, Djamah T, Bettayeb M. Fractional Hammerstein system identification using polynomial non-linear state space model. InControl, Engineering & Information Technology (CEIT), 2015 3rd International Conference on 2015 May 25 (pp. 1-6). IEEE.
[31] Hammar K, Djamah T, Bettayeb M. Fractional hammerstein system identification using particle swarm optimization. InModelling, Identification and Control (ICMIC), 2015 7th International Conference on 2015 Dec 18 (pp. 1-6). IEEE.
[32] Hong X, Mitchell RJ. Hammerstein model identification algorithm using Bezier–Bernstein approximation. IET Control Theory & Applications. 2007 Jul 1;1(4):1149-59.
[33] Farin G. Curves and surfaces for computer-aided geometric design: a practical guide. Elsevier; 2014 Jun 28.
[34] Ahmadi M, Mojallali H. Identification of multiple-input single-output Hammerstein models using Bezier curves and Bernstein polynomials. Applied Mathematical Modelling. 2011 Apr 30;35(4):1969-82.
[35] Farouki R, Goodman T. On the optimal stability of the Bernstein basis. Mathematics of Computation of the American Mathematical Society. 1996;65(216):1553-66.
[36] Jahani Moghaddam, M., Mojallali, H., &Teshnehlab, M. (2018). Recursive identification of multiple-input single-output fractional-order Hammerstein model with time delay. Applied Soft Computing, 70, 486-500.
[37] Jahani Moghaddam, M., Mojallali, H., &Teshnehlab, M. (2018). A multiple-input–single-output fractional-order Hammerstein model identification based on modified

neural network. Mathematical Methods in the Applied Sciences, 41(16), 6252-6271.
[38] Sundari S, Nachiappan A. Online Identification Using RLS Algorithm and Kaczmarz's Projection Algorithm for a Bioreactor Process. International Journal of Engineering and Computer Science. 2014;3:7974-8.
[39] Vahidi A, Stefanopoulou A, Peng H. Recursive least squares with forgetting for online estimation of vehicle mass and road grade: theory and experiments. Vehicle System Dynamics. 2005 Jan 1;43(1):31-55.
[40] Hussain MN, Omar AM, Samat AA. Identification of multiple input-single output (miso) model for mppt of photovoltaic system. InControl System, Computing and Engineering (ICCSCE), 2011 IEEE International Conference on 2011 Nov 25 (pp. 49-53). IEEE.
[41] Eskinat E, Johnson SH, Luyben WL. Use of Hammerstein models in identification of nonlinear systems. AIChE Journal. 1991 Feb 1;37(2):255-68.

Passivity-based Design of Controller and Observer for a Class of Nonlinear Systems with Application to Hepatitis B Disease

Shaghayegh Gorji[1] , Ahmad Fakharian[2*], Rezvan Abbasi[1]

Abstract: In this paper, a strictly passive formulation has been developed to design a passive state-observer for both time-invariant and time-varying Lipschitz nonlinear systems. During this formulation, a convergence and strictly passive state-observer is provided to have passive closed-loop system. Some definitions and charts are defined here for time-invariant and time-varying systems in different scenarios. A new interconnection between passivity of subsystems and passivity/stability of the closed-loop system has been introduced from a different point of view. All definitions are organized based on the systematic method called "virtually Euler-Lagrange" form of passivation. Utilizing this form and theses definitions, make the design process simpler and straightforward, while, some conditions of design will be released due to using these definitions. The designed controller/observer has been applied to control the hepatitis B virus infection disease. The reliability of the proposed definitions are examined by using MATLAB/SIMULINK, while, the results demonstrate the ability and power of this novel approach.

Keywords: passivity-based design, state-observer, adaptive control, Lipschitz nonlinear systems, virtually Euler-Lagrange, hepatitis B virus.

1. Introduction

Passivity-based method is a powerful and well-known nonlinear tool to design and analysis of systems. Passive systems are a class of dynamical systems where the exchanging energy with the environment plays a major role in them. Besides, the feature of "passivity" is preserved in feedback connection [1]. The concept of passivity-based control (PBC) has been established for the first time in [2]. A systematic method for passivation of non-Euler-Lagrange systems is proposed in [3]. Passivity-based observer and controller have been designed in some PHD dissertations [4,5]. This well-established technique has been used to control/observe in a wide range of systems, including biologic [3,6], robot manipulators [7,8], real-time systems [9] induction motors [10,11], swash-mass helicopter and pendulum [12,13], and so on.

The field of observer design has been activated for many years. Typically, observers are designed to estimate the system states based on the available measurements, which are called state-observer (SO). Passivity concepts for state-observer design is pursued in [4,14]. For the set-valued Lur'e systems, a convergence and passive SO is designed in [15]. Observer-based robust control has been used for multi-agent systems in [16]. For a class of Lipschitz nonlinear systems, some important issues are stated for observer design in papers [17-18].Also, for one-sided Lipschitz systems some nonlinear observer is designed [19]. Some authors study on the interval observer design [20-22].

The hepatitis B virus (HBV) viral disease has always been a serious infectious disease all around the world. This dangerous virus is almost classified into two phases, acute and chronic. The chronic phase of HBV causes the liver cirrhosis and liver cancer to progress [23]. In this phase, most of the sick people must carry on drug consumption up to the end of life [24]. Authors in [3], [25] and [26], respectively apply adaptive passivity-based control (APBC), Kalman-filter and Lyapunov-based methods in order to treat infected body. Some adaptive approaches are employed in [27,28], while in [29], a robust method is applied to handle the parametric and non-parametric uncertainties in HBV model. In this paper, the Lyapunov direct method (based on [30]) has been exerted as the compared method. Besides, due to the importance of controlling human diseases, many authors have worked in various biological fields, such as Ebola [31] and Corona virus[32].

According to availability of variable states and system parameters, four scenarios will be appeared. Some definitions are presented based on papers [3] for time-invariant and time-varying nonlinear Lipschitz systems. These definitions have been provided to analyze and design

1 Department of Electrical Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran

2* **Corresponding author:** Department of Electrical Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran
Email: Ahmad.Fakharian@qiau.ac.ir

Received: 2024.03.24; Accepted: 2024.04.17

control systems in the sense of passivity. Using these definitions, link the passivity of several subsystems to the stability of total closed-loop system in a new point of view. In this regard, some charts are structured here to pursue the issue in the best way. All of the definitions are shaped based on virtually Euler-Lagrange form (VEL-form) of passivation in order to possess the advantages of this form. During using these definitions, one will face into less complexity and more speed in passivity-based analysis and control. The capability of proposed definitions is confirmed by applying them on the HBV viral disease model.

This paper is structured as follows. Some definitions are presented in Section 2 for both time-varying and time-invariant systems. Moreover, two charts are provided to better follow up the design process. The under-study time-invariant HBV model has been investigated in Section 3. While, in Section 4, the controller/observer is designed for HBV model in different scenarios for the passivity-based method and compared method. The simulation results have been depicted for 1000-days drug therapy in Section 5, and finally, Section 6 has been dedicated to the conclusion of the paper.

2. Definitions

Nonlinear Time-Varying Systems:

Consider the nonlinear time-varying system

$$\begin{cases} \dot{X} = f(t, X, \varpi); \ t \in \mathbb{R}, & X \in \mathbb{R}^n, & \varpi \in \\ Y = CX; & Y \in \mathbb{R}^p \end{cases} \quad (1)$$

where f is piecewise continuous in t and Lipchitz (at least locally)in (X, ϖ) and $f(t, 0, 0) = 0; \ \forall t \geq 0$. Rewrite (1)into the following form

$$\begin{cases} \dot{X} = AX + \Psi(t, X, \varpi) \\ Y = CX \end{cases}, \quad (2)$$

where A is a constant matrix and Ψ is a Lipschitz nonlinear term (at least locally).

Definition 1: If system(2)with the mapping $\varpi(X) \rightarrow Y$,is strictly passive (SP)in the sense of VEL-form, by choosing a suitable function $\varpi = -\phi(Y)$, s.t. ϕ is any locally Lipchitz function and $\phi(0) = 0, Y^T\phi(Y) > 0; \ \forall Y \neq 0$, the origin of system (2)will be globally uniformly asymptotically stable [3].

For adaptive case, rewrite (2) in the following form

$$\begin{cases} \dot{X} = A_1 X + \Psi(t, X, \varpi) \ ; \ t \in \mathbb{R}, \varpi \in \mathbb{R}^m, X \in \mathbb{R}^n \\ Y = CX \end{cases}, \quad (3)$$

with

$$\Psi(t, X, \varpi) = \Psi'(t, X, \varpi) + A_2 X. \quad (4)$$

Let Ψ' be a nonlinear Lipschitz term which is piecewise continuous in$t$, and $\Psi'(t, 0,0) = 0$.Suppose that$A_1$and A_2 contain the fix and unknown system parameters, respectively.

Definition2: If system(2)with the mapping $\varpi(X) \rightarrow Y$,is strictly passive (SP)in the sense of VEL-form, and there exists an estimation mechanism as

$$\dot{\hat{\theta}} = \Gamma\left(\hat{\Phi}^T \mathrm{Y} + \hat{K}\right), \quad (5)$$

with a radially unbounded (RU) and positive definite (PD)storage function; therefore, the closed-loop system (3)will be output strictly passive(OSP),and the sufficient condition to asymptotically stabilizing the origin $X = 0$ is that $\dot{X} \in L_\infty$. So, the origin of system (3) with the mapping $\varpi(X, \hat{\theta}) \rightarrow Y$, will be globally uniformly asymptotically stabilized by $\varpi = -\phi(Y)$, s.t. ϕ is any locally Lipchitz function, where $\phi(0) = 0$ and $Y^T\phi(Y) > 0; \ \forall Y \neq 0$. While, it is guaranteed that $\hat{\theta}$and $\tilde{\theta}$remain bounded($\hat{\theta}, \tilde{\theta} \in L_\infty$).

Definition3: Consider system (2)with Lipschitz constant$\varrho \geq 0$, i.e.,

$$\left\|\Psi(t, X, \varpi) - \Psi\left(t, \hat{X}, \varpi\right)\right\| \leq \varrho\left\|X - \hat{X}\right\|, \quad (6)$$

and, with the following state-observer

$$\dot{\hat{X}} = A\hat{X} + \Psi\left(t, \hat{X}, \varpi\right) + L\left[Y - C\hat{X}\right]. \quad (7)$$

This SO is asymptotically convergence and strictly passive, if for the LMI

$$\begin{bmatrix} (A - LC)^T P + P(A - LC) + I + \frac{1}{2} C^T C & P \\ P & -\varrho^{-2} I \end{bmatrix} \prec 0, \quad (8)$$

there exists a PD symmetric solution P [6].

Definition4: If system (2) with the mapping $\varpi(X) \rightarrow Y$,is strictly passive (SP)in the sense of VEL-form, and there exists a strictly passive SO as (7) with a RU and PD storage function, therefore, the closed-loop system given by

$$\begin{cases} \dot{X} = AX + \Psi\left(t, X, \varpi(\hat{X})\right); \ t \in \mathbb{R}, \varpi \in \mathbb{R}^m, X \in \mathbb{R}^n \\ \dot{\hat{X}} = A\hat{X} + \Psi\left(t, \hat{X}, \varpi(\hat{X})\right) + L[Y - \mathbb{Y}]; \ Y \in \mathbb{R}^m, \hat{X} \in \\ \mathbb{Y} = C\hat{X}; \quad \mathbb{Y} \in \mathbb{R}^m \end{cases} \quad (9)$$

with the mapping $\varpi(\hat{X}) \rightarrow \mathbb{Y}$, is strictly passive. By defining$\chi = [X \quad \hat{X}]^T$,the origin $(\chi = 0)$ can be globally uniformly asymptotically stabilized by $\varpi = -\phi(\mathbb{Y})$, s.t. ϕ is any locally Lipchitz function, where $\phi(0) = 0$ and $\mathbb{Y}^T\phi(\mathbb{Y}) > 0; \ \forall \mathbb{Y} \neq 0$ [**Error! Bookmark not defined.**].

Definition5: If system (2)with the mapping $\varpi(X) \rightarrow Y$, is strictly passive (SP)in the sense of VEL-form, and there exists a strictly passive SO as

$$\dot{\hat{X}} = A_1\hat{X} + \Psi\left(t, \hat{X}, \varpi\right) + L[Y - \mathbb{Y}], \quad (10)$$

with a RU and PD storage function; and an estimation machine as

$$\dot{\hat{\theta}} = \Gamma\left(\hat{\Phi}^T \mathbb{Y} + \hat{K}\right), \tag{11}$$

with a RU and PD storage function; therefore, the closed-loop system $\varpi(\hat{X}, \hat{\theta}) \to \mathbb{Y}$given by

$$\begin{cases} \dot{X} = A_1 X + \Psi\big(t, X, \varpi(\hat{X}, \hat{\theta})\big);\ t \in \mathbb{R}, \varpi \in \mathbb{R}^m, \\ \dot{\hat{X}} = A_1 \hat{X} + \Psi\big(t, \hat{X}, \varpi(\hat{X}, \hat{\theta})\big) + L[Y - \mathbb{Y}];\ Y \in \mathbb{R}^m, \\ \mathbb{Y} = C\hat{X};\quad \mathbb{Y} \in \mathbb{R}^m \end{cases} \tag{12}$$

will be output strictly passive (OSP). By defining$\chi = [X \quad \hat{X}]^T$, the sufficient condition to asymptotically stabilizing the origin $\chi = 0$ is that $\dot{X} \in L_\infty$. So, the origin will be globally uniformly asymptotically stabilized by $\varpi = -\phi(\mathbb{Y})$, s.t. ϕ is any locally Lipchitz function, where $\phi(0) = 0$ and $\mathbb{Y}^T\phi(\mathbb{Y}) > 0;\ \forall\, \mathbb{Y} \neq 0$. While, it is guaranteed that$\hat{\theta}$and $\tilde{\theta}$ will be bounded$(\hat{\theta}, \tilde{\theta} \in L_\infty)$ [6].

Nonlinear Time-Invariant Systems:

Consider the nonlinear time-invariant system

$$\begin{cases} \dot{X} = f(X, \varpi);\ X \in \mathbb{R}^n, \quad \varpi \in \mathbb{R}^p \\ Y = CX;\ \ Y \in \mathbb{R}^p \end{cases}, \tag{13}$$

where f is Lipchitz (at least locally) in (X, ϖ) and $f(0,0) = 0$. Rewrite (13) into the form

$$\begin{cases} \dot{X} = AX + \Psi(X, \varpi) \\ Y = CX \end{cases}, \tag{14}$$

where A is a constant matrix and Ψ is a Lipschitz nonlinear term (at least locally).

Definition6: If system(14)with the mapping $\varpi(X) \to Y$,is strictly passive (SP)in the sense of VEL-form, by choosing a suitable function $\varpi = -\phi(Y)$, s.t. ϕ is any locally Lipchitz function and $\phi(0) = 0, Y^T\phi(Y) > 0;\ \forall\, Y \neq 0$, the origin of the system (14)will be globally asymptotically stable.

For adaptive case, rewrite (14) in the form of

$$\begin{cases} \dot{X} = A_1 X + \Psi(X, \varpi)\ ;\ \varpi \in \mathbb{R}^m, X \in \mathbb{R}^n \\ Y = CX \end{cases}, \tag{15}$$

with

$$\Psi(X, \varpi) = \Psi'(X, \varpi) + A_2 X. \tag{16}$$

where Ψ' is nonlinear Lipschitz term and $\Psi'(0,0) = 0$. Suppose A_1 and A_2 as the known and unknown system parameters, respectively.

Definition 7: If system(14)with the mapping $\varpi(X) \to Y$,is strictly passive (SP)in the sense of VEL-form, and there exists an estimation mechanism as

$$\dot{\hat{\theta}} = \Gamma\left(\hat{\Phi}^T \mathrm{Y} + \hat{K}\right), \tag{17}$$

with a RU and PD storage function; therefore, the closed-loop system with the mapping $\varpi(X, \hat{\theta}) \to Y$will be output strictly passive(OSP).The sufficient condition to asymptotically stabilizing the origin $X = 0$ is that $\dot{X} \in L_\infty$. So, the origin will be globally uniformly asymptotically stabilized by $\varpi = -\phi(Y)$, s.t. ϕ is any locally Lipchitz function, where $\phi(0) = 0$ and $Y^T\phi(Y) > 0;\ \forall\, Y \neq 0$. While, it is guaranteed that $\hat{\theta}$and $\tilde{\theta}$remain bounded.

Definition 8: Consider system (14)with Lipschitz constant$\varrho \geq 0$, i.e.,

$$\big\|\Psi(X, \varpi) - \Psi\big(\hat{X}, \varpi\big)\big\| \leq \varrho\big\|X - \hat{X}\big\|, \tag{18}$$

and, with the following state-observer

$$\dot{\hat{X}} = A\hat{X} + \Psi\big(\hat{X}, \varpi\big) + L\big[Y - C\hat{X}\big]. \tag{19}$$

This SO is asymptotically convergence and strictly passive, if for the LMI

$$\begin{bmatrix} (A - LC)^T P + P(A - LC) + I + \frac{1}{2}C^T C & P \\ P & -\varrho^{-2} I \end{bmatrix} \prec 0, \tag{20}$$

there exists a PD symmetric solution P.

Definition 9: If system (14) with the mapping $\varpi(X) \to Y$,is strictly passive (SP)in the sense of VEL-form, and there exists a strictly passive SO as (19) with a RU and PD storage function, therefore, the closed-loop system given by

$$\begin{cases} \dot{X} = AX + \Psi\big(X, \varpi(\hat{X})\big);\ \varpi \in \mathbb{R}^m, X \in \mathbb{R}^n \\ \dot{\hat{X}} = A\hat{X} + \Psi\big(\hat{X}, \varpi(\hat{X})\big) + L[Y - \mathbb{Y}];\ \hat{X} \in \mathbb{R}^n, \\ \mathbb{Y} = C\hat{X};\quad \mathbb{Y} \in \mathbb{R}^m, \quad Y \in \mathbb{R}^m \end{cases} \tag{21}$$

with the mapping $\varpi(\hat{X}) \to \mathbb{Y}$, is strictly passive. By defining$\chi = [X \quad \hat{X}]^T$,the origin $(\chi = 0)$ can be globally asymptotically stabilized by $\varpi = -\phi(\mathbb{Y})$, s.t. ϕ is any locally Lipchitz function, where $\phi(0) = 0$and $\mathbb{Y}^T\phi(\mathbb{Y}) > 0;\ \forall\, \mathbb{Y} \neq 0$.

Definition 10: If system (14)with the mapping $\varpi(X) \to Y$, is strictly passive (SP)in the sense of VEL-form, and there exists a strictly passive SO as

$$\dot{\hat{X}} = A_1\hat{X} + \Psi\big(\hat{X}, \varpi\big) + L[Y - \mathbb{Y}], \tag{22}$$

with a RU and PD storage function; and an estimation machine as

$$\dot{\hat{\theta}} = \Gamma\left(\hat{\Phi}^T \mathbb{Y} + \hat{K}\right), \tag{23}$$

with a RU and PD storage function; therefore, the closed-loop system

$$\begin{cases} \dot{X} = A_1 X + \Psi\big(X, \varpi(\hat{X}, \hat{\theta})\big);\ \varpi \in \mathbb{R}^m, X \in \mathbb{R}^n \\ \dot{\hat{X}} = A_1\hat{X} + \Psi\big(\hat{X}, \varpi(\hat{X}, \hat{\theta})\big) + L[Y - \mathbb{Y}];\ \hat{X} \in \mathbb{R}^n, \\ \mathbb{Y} = C\hat{X};\quad \mathbb{Y} \in \mathbb{R}^m, Y \in \mathbb{R}^m \end{cases} \tag{24}$$

with the mapping $\varpi(\hat{X},\hat{\theta}) \rightarrow \mathbb{Y}$,will be output strictly passive (OSP). By defining $\chi = [X \quad \hat{X}]^T$, the sufficient condition to asymptotically stabilizing the origin $(\chi = 0)$ is that $\dot{\chi}$ is bounded. So, the origin can be globally asymptotically stabilized by $\varpi = -\phi(\mathbb{Y})$, s.t. ϕ is any locally Lipchitz function, where $\phi(0) = 0$and $\mathbb{Y}^T\phi(\mathbb{Y}) > 0;\ \forall\, \mathbb{Y} \neq 0$. While, it is guaranteed that $\hat{\theta}$ and $\tilde{\theta}$ will be bounded.

Remark 1: The VEL-form is a systematic formulation for system passivation which has been proposed in details in paper [3].In this formulation,ϖ is called virtual input,and thereis a relationship between $\varpi$and actual input$u$ in all definitions here. Foraffine-in-control systems one will have$u = a + b\varpi$,where,

- $a$is continuously differentiable in X(and piecewise continuous in t for TV systems), in the state space or work space.
- $b$is nonsingular in X(and piecewise continuous in t for TV systems), in the state space or work space.

Remark 2: These definitions are presented according to the VEL-form and theorems in [3] and [6].All the related proofs are omitted here due to limits on the paper volume.

For better follow-up of the issue, one can pursue the charts below. Different categories of nonlinear systems have been illustrated in Fig. **1**. After finding a suitable u, i.e. the input which forces the system to be passive, refer to Fig. **2** and track the process related to your current scenario.

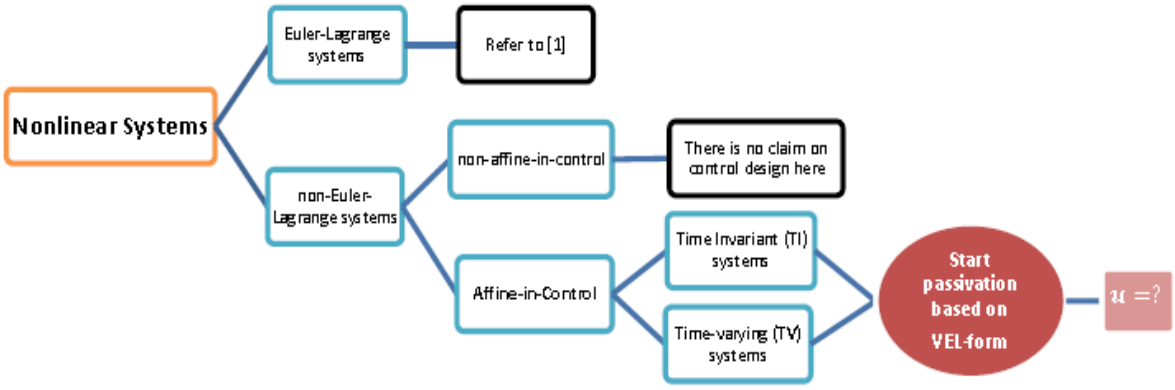


Fig. 1: A classification for nonlinear systems to use VEL-form of passivation

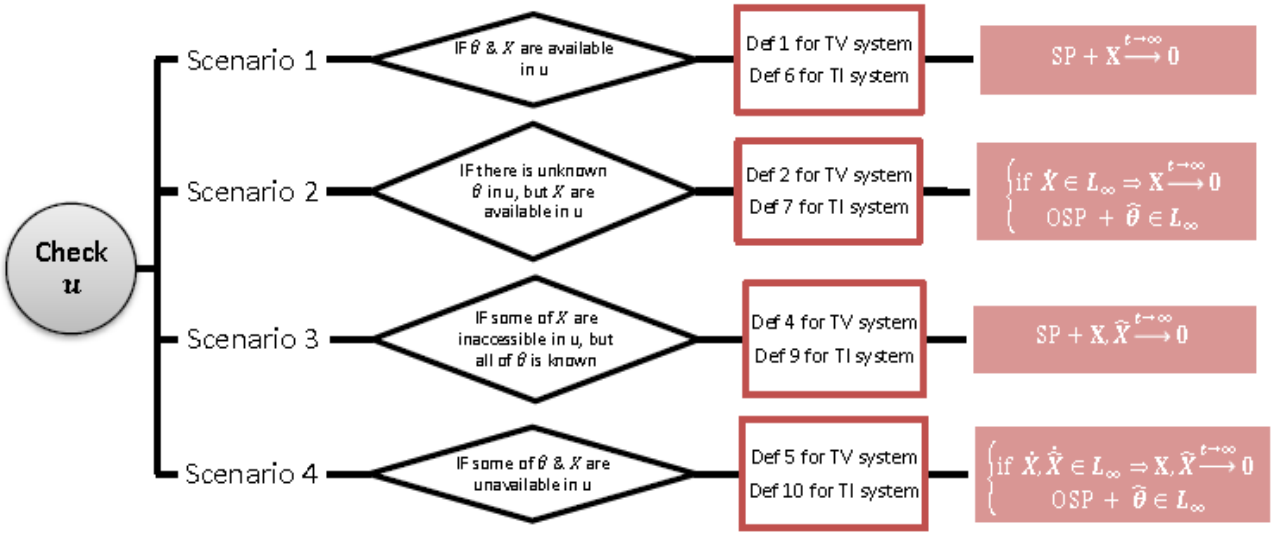


Fig. 2: Different scenarios in passivity-based control/ observe in the sense of VEL-form; SP: strictly passive, OSP: output strictly passive, TV: time-varying, TI: time-invariant

3. HBV disease model

The time-invariant HBV infection model is considered as [33-34].

$$\begin{cases} \dot{x} = \lambda - dx - \beta xv \\ \dot{y} = \beta xv - \delta y \\ \dot{v} = (1-\mu u)py - cv \end{cases}, \quad (25)$$

where x, y and v are respectively uninfected hepatocytes (healthy cells), infected hepatocytes (sick cells) and free viruses in blood. The system parameters and initial conditions are listed in Table 1.Control input$u$has been exerted to infection treatment in the infected body, whereis bounded in [0, 1].The untreated HBV model (without drug usage) is given by

$$\begin{cases} \dot{x} = \lambda - dx - \beta xv \\ \dot{y} = \beta xv - \delta y \\ \dot{v} = py - cv \end{cases}, \quad (26)$$

which has two equilibrium points, infection-free point (unstable) and endemic point (stable) as

$$X^*_{inf-free} = \begin{bmatrix} \frac{\lambda}{d} \\ 0 \\ 0 \end{bmatrix}, \quad X^*_{endemic} = \begin{bmatrix} \frac{c\delta}{p\beta} \\ \frac{\lambda}{\delta} - \frac{dc}{p\beta} \\ \frac{\lambda p}{c\delta} - \frac{d}{\beta} \end{bmatrix}. \quad (27)$$

In this study, all investigations will be done based on the infection-free point. The shifted model around the infection-free equilibrium point will be presented as

$$\begin{cases} \dot{\bar{x}} = -d\bar{x} - \beta v(\bar{x} + x^*) \\ \dot{y} = \beta v(\bar{x} + x^*) - \delta y \\ \dot{v} = (1-\mu u)py - cv \end{cases}, \quad Y = v, \quad (28)$$

where, $\bar{x} = x - x^*$ and $x^* = \frac{\lambda}{d}$, and Y states system output vector.The most significant criterion in observing the infection status is HBV DNA serum level which is related directly to the free virus quantity (v). HBV DNA is distinguished by some serological tests such as RT-PCR, therefore, the signal v can be selected as measurable system output.

4. Controller/ observer design

In this section, all scenarios are investigated for passivity-based approach and compared method (Lyapunov direct method).

Scenario 1: Here, the controller is designed for nominal model in the absence of state-observer.

-Passivity based design: For the system (28) one can find the following matrices by using the VEL method [**Error! Bookmark not defined.**]as

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, N = \begin{bmatrix} d & 0 & \beta(\bar{x} + x^*) \\ 0 & \delta & -\beta(\bar{x} + x^*) \\ 0 & -p & c \end{bmatrix},$$
$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, K'_d = \begin{bmatrix} d & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & c \end{bmatrix}.$$

According to [**Error! Bookmark not defined.**], the storage function and its time-derivative will be

$$V = \frac{1}{2} X^T D X, \qquad \dot{V} \le -X^T K'_d X + \varpi^T Y, \quad (29)$$

with the virtual input

$$\varpi = \beta(\bar{x} + x^*)(y - \bar{x}) + py(1 - \mu u). \quad (30)$$

Thus, the mapping $\varpi(X) \to Y$ is SP and according to Definition 6, a suitable choose for ϖ can be $\varpi = -\alpha v$.Finally, it concludes

$$u = \frac{1}{p\mu y}[\beta(\bar{x} + x^*)(y - \bar{x}) + py + \alpha v]. \quad (31)$$

With the control input (31), the origin of (28)will be globally asymptotically stable. In order to avoid making the paper longer, it is neglected to show $\Psi$is Lipschitzand $\Psi(0,0) = 0$in the HBV model based on general form of (14).The process of this proof could be similar to [3].

-Lyapunov based design: For compared method, the Lyapunov function can be selected as

$$H = \frac{1}{2} X^T X, \quad (32)$$

and, taking time derivative of H yields

$$\dot{H} \le -d\bar{x}^2 - \delta y^2 - cv^2 + \beta(y - \bar{x})(\bar{x} + x^*)v + pvy - \mu pyvu.$$

By selecting the control input as

$$u = \frac{1}{p\mu y}[\beta(\bar{x} + x^*)(y - \bar{x}) + py], \quad (33)$$

one has

$$\dot{H} \le -X^T \begin{bmatrix} d & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & c \end{bmatrix} X,$$

whered, δ and c are PD constants, so,$\dot{H}$ is ND. Since the Lyapunov function (32)is RU, the origin of (28)will be globally asymptotically stable.

Scenario 2: The controller is designed for adaptive model in the absence of state-observer, while, the unknown system parameters are considered as$\theta = [\beta \quad \mu]$. Thus, the estimation error of unknown parameters is $\tilde{\theta} = [\tilde{\beta} \quad \tilde{\mu}]$, where $\tilde{\theta} = \theta - \hat{\theta}$ and $\hat{\theta}$ states the estimation vector of parameters.

-Passivity based design: Based on (15), (16) and (28),

$$\Psi' = \begin{bmatrix} -\beta\bar{x}v \\ \beta\bar{x}v \\ -\mu puy \end{bmatrix}, \quad A_1 = \begin{bmatrix} -d & 0 & 0 \\ 0 & -\delta & 0 \\ 0 & p & -c \end{bmatrix},$$
$$A_2 = \begin{bmatrix} 0 & 0 & -\beta x^* \\ 0 & 0 & \beta x^* \\ 0 & 0 & 0 \end{bmatrix}.$$

For HBV model, the adaptation mechanism is given as

$$\dot{\hat{\theta}} = \Gamma \begin{bmatrix} (\bar{x} + x^*)(y - \bar{x}) \\ -upy \end{bmatrix} v, \tag{34}$$

with a RU and PD storage function as $V_\theta = \frac{1}{2}\tilde{\theta}^T\Gamma^{-1}\tilde{\theta}$, where Γ is a free PD gain matrix. According to (23), one has

$$\widehat{\Phi} = [(\bar{x} + x^*)(y - \bar{x}) \quad 0], \ \widehat{K} = [0 \quad -upyv]^T.$$

On the other hand, Scenario 1 reveals that the nominal system $(\varpi(X) \to Y)$ is SP with (30).Therefore, based on Definition 7, if $\dot{X} \in L_\infty$, the origin of adaptive model will be stabilized. Since all states and control input in (28)are considered finite, thus, $\dot{X}$ is bounded. Thereupon, by choosing $\varpi = \alpha v$, the origin $X = 0$ is globally asymptotically stable by using the adaptive control input as

$$u_{(X,\hat{\theta})} = \frac{1}{p\hat{\mu}y}[\hat{\beta}(\bar{x} + x^*)(y - \bar{x}) + py + \alpha v]. \tag{35}$$

-Lyapunov based design: For compared method, one can choose the Lyapunov function for adaptive closed-loop system as

$$H_{Acl(X,\tilde{\theta})} = \frac{1}{2}X^TX + \frac{1}{2}\tilde{\theta}^T\Gamma^{-1}\tilde{\theta}. \tag{36}$$

By select the following control input

$$u_{(X,\hat{\theta})} = \frac{1}{p\hat{\mu}y}[\hat{\beta}(\bar{x} + x^*)(y - \bar{x}) + py], \tag{37}$$

and the adaptation law similar to (34), $\dot{H}_{Acl}$ will benegative semi-definite (NSD). Based on previous section, $\dot{X} \in L_\infty$. Accordingly, $X \to 0$asymptotically, while, $\tilde{\theta}$only can remain bounded.

Scenario 3: The nominal system in the presence of SO has been studied here. Suppose that all using parameters in control inputs are available, while, some of using states are unavailable.

-Passivity based design: Consider the SO (19) with a RU and PD storage function $V_{ob} = \tilde{X}^TP\tilde{X}$, where $\tilde{X}$ is defined as estimation error of states and $\tilde{X} = X - \hat{X}$. Let define $L = L_0\text{ones}(3,1)$; $L_0 > 0$. By setting $L_0 = 200$ a solution $P = P^T > 0$ is found for LMI (20), while, $A - LC$ is Hurwitz. Thus,based on Definition 8, one will have a strictly passive SO for HBV model. According to Definition 9, since the mapping $\varpi(X) \to Y$ is SP,the closed-loop HBV model will be SP with

$$\varpi_{(\hat{X})} = \beta(\hat{\bar{x}} + x^*)(\hat{y} - \hat{\bar{x}}) + p\hat{y}(1 - \mu u).$$

With suitable choice of $\varpi = \alpha\tan^{-1}(\hat{v})$, the origin will be globally asymptotically stabilized by the following control input

$$u_{(\hat{X})} = \frac{1}{p\mu\hat{y}}[\beta(\hat{\bar{x}} + x^*)(\hat{y} - \hat{\bar{x}}) + p\hat{y} + \alpha\tan^{-1}(\hat{v})]. \tag{38}$$

-Lyapunov based design: Consider the Lyapunov function as

$$H_{cl(\hat{X},\tilde{X})} = \frac{1}{2}\hat{X}^T\hat{X} + \tilde{X}^TP\tilde{X}, \tag{39}$$

and the following LMI instead of (20)

$$\begin{bmatrix} (A - LC)^TP + P(A - LC) + I & P \\ P & -\varrho_0{}^{-2}I \end{bmatrix} < 0.$$

Choosing $L_0 = 200$ leads to have a suitable solution $P$for this LMI.Now, by selecting the following control input

$$u_{(\hat{X})} = \frac{1}{p\mu\hat{y}}[\beta(\hat{\bar{x}} + x^*)(\hat{y} - \hat{\bar{x}}) + p\hat{y}], \tag{40}$$

$\dot{H}_{cl}$will beND and the origin of the system is globally asymptotically stable.

Scenario 4: In this case, the adaptive model in the presence of SO has been investigated. In other words, there are some inaccessible parameters and variable states in the control inputs which need to be estimated.

-Passivity based design: By setting $L_0 = 300$,a strictly passive SO is achieved for adaptive HBV model. On the other hand, the mapping $\varpi(X) \to Y$ is SP, and the passive adaptive law is as

$$\dot{\hat{\theta}} = \Gamma \begin{bmatrix} (\hat{\bar{x}} + x^*)(\hat{y} - \hat{\bar{x}}) \\ -up\hat{y} \end{bmatrix} \hat{v}, \tag{41}$$

where Γ is a free PD gain matrix.According to Definition 10, the closed-loop HBV model will be SP with

$$\varpi_{(\hat{X},\hat{\theta})} = \hat{\beta}(\hat{\bar{x}} + x^*)(\hat{y} - \hat{\bar{x}}) + p\hat{y}(1 - \hat{\mu}u).$$

Now, if $\dot{X}$ and $\dot{\hat{X}}$ are bounded, the origin can be stabilized. Since all states and estimation signals and control input are considered finite, therefore, $\dot{X}, \dot{\hat{X}} \in L_\infty$and with suitable choice of $\varpi = \alpha\tan^{-1}(\hat{v})$, the origin will be globally asymptotically stabilized by the following control input

$$u_{(\hat{X},\hat{\theta})} = \frac{1}{p\hat{\mu}\hat{y}}[\hat{\beta}(\hat{\bar{x}} + x^*)(\hat{y} - \hat{\bar{x}}) + p\hat{y} + \alpha\tan^{-1}(\hat{v})]. \tag{42}$$

-Lyapunov based design: Consider the Lyapunov function as

$$H_{Acl(\hat{X},\tilde{X},\tilde{\theta})} = \frac{1}{2}\hat{X}^T\hat{X} + \tilde{X}^TP\tilde{X} + \frac{1}{2}\tilde{\theta}^T\Gamma^{-1}\tilde{\theta}, \tag{43}$$

Choosing $L_0 = 300$ leads to have a convergence SO. Now, by selecting the following control input

$$u_{(\hat{X},\hat{\theta})} = \frac{1}{p\hat{\mu}\hat{y}}[\hat{\beta}(\hat{\bar{x}} + x^*)(\hat{y} - \hat{\bar{x}}) + p\hat{y}], \tag{44}$$

$\dot{H}_{Acl}$ will beNSD. Since $\dot{X}, \dot{\hat{X}} \in L_\infty$, thus, $\dot{X}$ and $\dot{\hat{X}}$ are

uniformly continuous. Finally, it leads to have $X \xrightarrow{t\to\infty} 0and\tilde{\theta} \in L_\infty$.

5. Simulation results

In the first step, according to the open-loop system (26), Fig. **3**illustrates the behavior of the HBV dynamics in the absence of drug therapy in 5000 days.

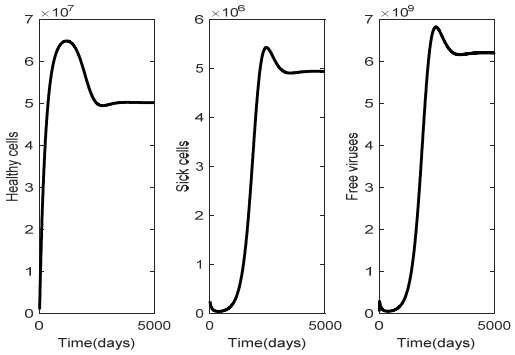


Fig. 3: Open-loop states in untreated infected body during 5000 days (without drug consumption)

In the following, the closed-loop HBV model has been simulated for both control approaches in four different scenarios during 1000-days drug therapy. To have a better comprehension of preference of the proposed method, Table 2 is provided to show the comparison criteria (norm 2) for some signals in both methods.

Scenario 1: Themost important signals are the number of free viruses and drug dosage which are depicted in Fig. **4**. For PBC approach set $\alpha = 50$ in (31).

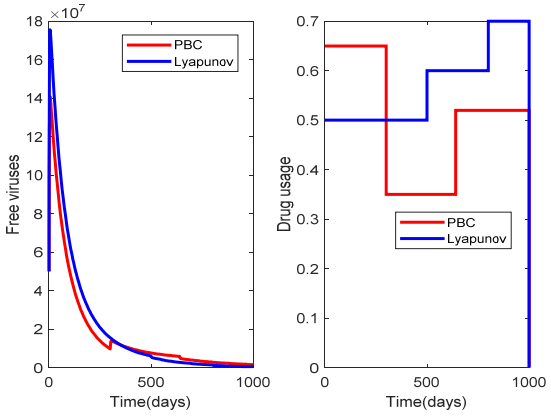


Fig. 4: Number of free viruses $(\boldsymbol{v})$ and drug dosage $(\boldsymbol{u})$ for PBC and Lyapunovdirect method during 1000-day drug therapy

Scenario 2: Here, the signals v and u have been simulated in Fig. **5** for APBC and adaptive Lyapunov direct method based on (35) and (37).Besides, the parametric estimation errors, i.e.$\tilde{\beta}$and $\tilde{\mu}$, are demonstrated in Fig. **6** by setting

$$\Gamma = \mathrm{diag}([10^{-20}, 10^{-15}]),$$

in(34)for both methods, and $\alpha = 1000$ in (35).

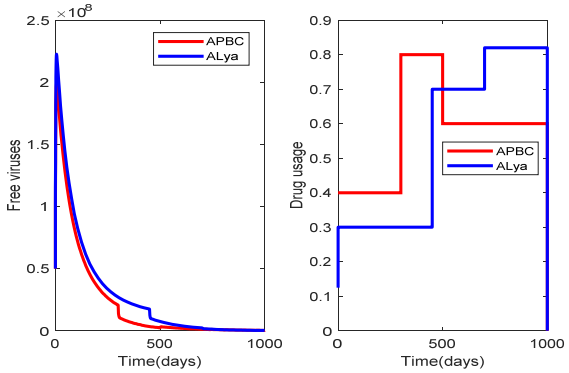


Fig. 5:Number of free viruses $(\boldsymbol{v})$ and drug dosage $(\boldsymbol{u})$ for APBC and adaptive Lyapunov direct method during 1000-day drug therapy

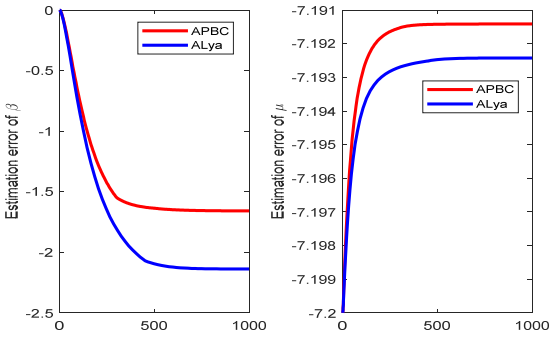


Fig. 6: Estimation error of unknown system parameters for APBC and adaptive Lyapunov direct method during 1000-day drug therapy

Scenario 3: In this case, set$\alpha = 500$ in (38). According to the control inputs in (38) and (40) free viruses and its estimation error are displayed in Figure 7 for PBC and Lyapunov method in the presence of SO. Fig. **8** depicts the drug usage in 1000 days. For more information, one can refer to Table 2.

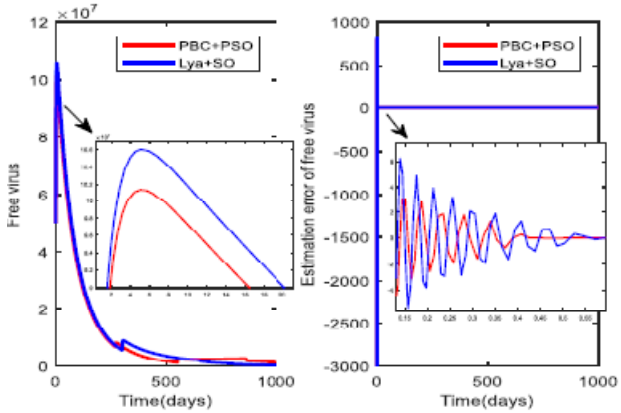


Fig. 7:Number of free viruses $(\boldsymbol{v})$ and estimation error of free viruses $(\tilde{\boldsymbol{v}})$ for PBC/PSO and Lyapunov direct method/SOduring 1000-day drug therapy

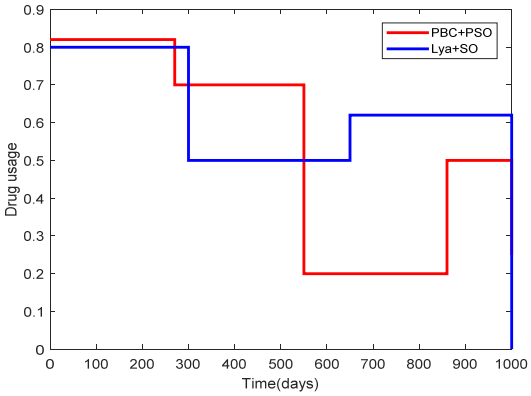


Fig. 8: Drug dosage in PBC/PSO and Lyapunov direct method/SO during 1000-daydrug therapy

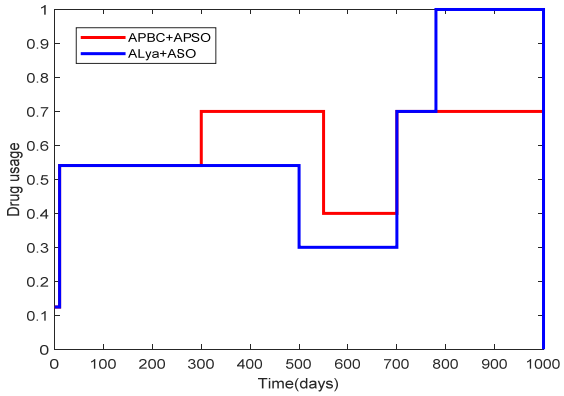


Fig.10: Drug dosage in APBC/APSO and adaptive Lyapunov direct method/ASO during 1000-day drug therapy

Scenario 4: For the last scenario, the status of free viruses and its error of estimation have been indicated in Fig. 9. In addition, one can follow the drug consumption for treated body in Fig.10 based on (42) and (44). The estimation errors of unknown parameters are depicted inFig. 11for both control methods as well. Set $\alpha = 100$ in (42) and

$$\Gamma = \mathrm{diag}([10^{-23}, 10^{-18}]),$$

in(41) for both control approaches.

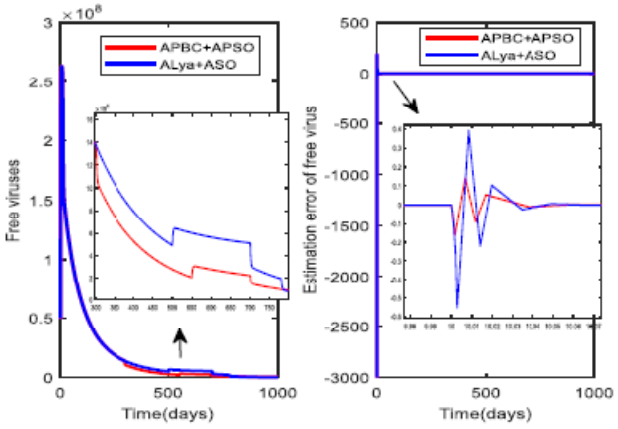


Fig. 9: Number of free viruses $(\tilde{v})$ and estimation error of free viruses $(\tilde{v})$ for APBC/APSO and adaptive Lyapunov direct method/ASOduring 1000-day drug therapy

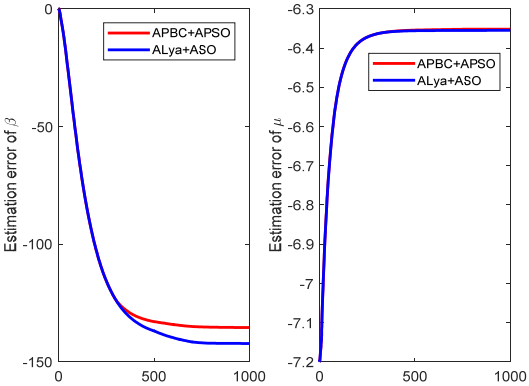


Fig. 11: Estimation error of unknown system parameters for APBC/APSO and adaptive Lyapunov direct method/ASO during 1000-day drug therapy

6.Conclusion

In this paper, a formulation based on LMI is provided to have a passive SO for time-varying and time-invariant Lipschitz nonlinear systems. According to the definitions, passivity and stability of the closed-loop can be connected to passivity of subsystems in a different way for both time-varying and time-invariant systems. A passivity-based controller/ observer is designed based on VEL-form and some definitions, which offer significant advantages, e.g. decreasing design complexities and accelerating the analysis and design process. Different scenarios are categorized in a char to guide the designer. These contributions have been validated on the HBV disease model, while, the capability of the proposed approach is confirmed based on the simulation results. As stated in the definitions, the parametric estimation errors will only be guaranteed to remain bounded in all adaptive cases, and as is clear from the results, the main goal of control has been achieved during this situation. From simulations and Table 2, it is revealed that the proposed method has more acceptable performance rather than the compared method. Norm 2 of drug input in the proposed approach is less in all scenarios. Moreover, it is obvious that the status of free viruses in passivity-based method is in a better condition.

Table 1: Parameters and initial conditions for HBV model [33]

Parameters	*Definitions*	*Numeric values*
λ	Constant production rate of new target cells (healthy hepatocytes)	2.5251×10^5
d	Death rate of healthy hepatocytes	0.0038
β	Infection rate of healthy hepatocytes	1.981×10^{-13}
δ	Death rate of Infected cells(sick hepatocytes)	0.0125
μ	Efficacy of drug usage	0.8
p	Production rate of free viruses by sick hepatocytes	842.0948
c	Death rate of free viruses	0.67
$x(0)$	Initial number of uninfected hepatocytes	10^6
$y(0)$	Initial number of infected hepatocytes	2.5×10^5
$v(0)$	Initial number of free viruses	0.5×10^8

Table 2: Comparison of the control approaches in four different scenarios during 1000-day drug therapy

Scenario	*Control Method*	*Explanations*	*Norm 2*		
			$\sqrt{\int_{t=0}^{1000} v^2}$	$\sqrt{\int_{t=0}^{1000} \widetilde{v}^2}$	$\sqrt{\int_{t=0}^{1000} u^2}$
1	PBC	Nominal case without state-observer	10^9	-	16.3
	Lyapunov		1.28×10^9	-	18.2
2	APBC	Adaptive case without state-observer	1.47×10^9	-	18.8
	Adaptive Lyapunov		1.69×10^9	-	19.1
3	PBC/ PSO	Nominal case with state-observer	7.1×10^8	150	19.1
	Lyapunov/ SO		7.5×10^8	163	20.3
4	APBC/ APSO	Adaptive case with state-observer	1.36×10^9	210	19.5
	ALyapunov/ ASO		1.37×10^9	220	20.6

References

[[1] R. Ortega, et al, *Passivity-based Control of Euler-Lagrange Systems*, London: Springer-Verlag, 1998.

[2] R. Ortega and M.Spong, "Adaptive motion control of rigid robots: a tutorial", *Automatica*,vol. 25, pp. 877-888, 1989.

[3] S. Gorji, A. Fakharian and R.Abbasi,"Virtually Euler-Lagrange form for feedback passivation of a wide range of non-Euler-Lagrange systems: Adaptive Passivity-Based Control of Hepatitis B Disease", *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering,* vol. 237, pp. 1313-1330, 2023.

[4] H.Shim,"*A Passivity-based Nonlinear Observer and a Semi-global Separation Principle"*, PhD Thesis, Seoul National University, 2000.

[5] TE.Pare ,"*Passivity-based analysis and control of nonlinear systems"*, PhD Thesis, Stanford University, 2000.

[6] S. Gorji, A. Fakharian and R.Abbasi ,"Passive observer-based controller design for a class of Lipschitz nonlinear time-varying systems with application to Hepatitis B disease", *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering,*2024,doi:10.1177/09596518241236925.

[7] M. Ali and CK.Alexander,"Robust tracking control of a robot manipulator using a passivity-based extended-state observer approach",*IET Cyber-syst Robot*, vol. 1, pp. 63-71, 2019.

[8]F. Bouakrif, D. Boukhetala and F.Boudjema,"Passivity-based controller-observer for robot manipulators",*3rd International Conference on Information and Communication Technologies: From Theory to Applications*, Damascus, Syria, 7-11 April, 2008.

[9] LDH. Peiris, JL. Du Bois and AR. Plummer, "Passivity-based adaptive delay compensation for real-time hybrid tests",*Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 235, pp.427-432, 2020.

[10] R. Salim, et al,"Sensorless passivity based control for

induction motor via an adaptive Observer",*ISA Transactions,* vol. 84, pp. 118-127, 2019.

[11] Y. Liu, et al, "Passivity Control of Induction Motors based on Adaptive Observer Design",*3rd IFAC International Conference on Intelligent Control and Automation Science*, Chengdu, China, 2-4 Sep. 2013, pp. 176-181.

[12] B. Salamat and AM.Tonello,"Energy Based Control of a Swash Mass Helicopter through Decoupling Change of Coordinates",*IEEE Access,* vol. 8, pp. 77449-77458, 2020.

[13] B. Salamat and AM.Tonello,"A Swash Mass Pendulum with Passivity-Based Control",*IEEE Robotics and Automation Letters,* vol. 6, pp. 199-206, 2021.

[14] L. Tiitinen, M. Hinkkanen and L.Harnefors,"Stable and Passive Observer-Based V/Hz Control for Induction Motors",*IEEE Energy Conversion Congress and Exposition*, Detroit, MI, USA, 9-13 Oct. 2022, pp. 1-8.

[15] B. Brogliato, R. Lozano, et al,*Dissipative systems analysis and control*, 3rd ed. Switzerland AG: Springer Nature, 2020.

[16] M. Salimifard and HA.Talebi,"Robust output feedback fault-tolerant control of non-linear multi-agent systems based on wavelet neural networks",*IET Control Theory & Applications*, vol. 11, pp. 3004-3015, 2017.

[17] R. Rajamani and YM.Cho,"Existence and design of observers for nonlinear systems: relation to distance to unobservability"*, International Journal of Control*, vol. 69, pp. 717-731, 1998.

[18] R.Rajamani,"Observers for Lipschitz Nonlinear Systems",*IEEE Transaction on Automatic Control,* vol. 43, pp. 397-401, 1998.

[19] Y. Dong, W. Liu and S.Liang,"Nonlinear observer design for one-sided Lipschitz systems with time-varying delay and uncertainties",*International Journal of Robust and Nonlinear Control,* vol. 27, pp.1974-1998, 2017.

[20] T. Raïssi, D. Efimov and A.Zolghadri,"Interval state estimation for a class of nonlinear systems",*IEEE Trans. Automatic Control,* vol. 57, pp. 260-265, 2012.

[21] T. Raïssi, G. Videau and A.Zolghadri,"Interval observers design for consistency checks of nonlinear continuous-time systems",*Automatica,* vol. 46, pp. 518-527, 2010.

[22] D. Efimov, et al,"Interval State Observer for Nonlinear Time Varying Systems",*Automatica*, vol. 49, pp. 200-205, 2013.

[23] S. Manzoor, et al,"Hepatitis B virus therapy: What's the future holding for us?",*World J. Gastroenterol,*vol. 21, pp. 12558-12575, 2015.

[24] World Health Organization.*Online Referencing*, http://www.who.int.

[25] J.Jang,"Feedback control of an HBV model based on ensemble kalman filter and differential evolution", *Mathematical Biosciences and Engineering,* vol. 15, pp. 667-691, 2018.

[26] Y. Ding and WJ.Wang,"Control strategy design for the anti-HBV mathematical model", *IET Syst. Biol,*vol. 13, pp. 186-193, 2019.

[27] O. Aghajanzadeh, et al,"Robust adaptive Lyapunov-based control of hepatitis B infection", *IET Syst. Biol,* vol.12, pp. 62-67, 2018.

[28] O. Aghajanzadeh, et al,"Nonlinear Adaptive Control Method for Treatment of Uncertain Hepatitis B Virus Infection",*Biomedical Signal Processing and Control,*vol. 38, pp. 174-181, 2017.

[29] O. Aghajanzadeh, M. Sharifi andA.Falsafi,"Robust control strategy for HBV treatment: Considering parametric and nonparametric uncertainties",*Control Applications for Biomedical Engineering Systems,*vol. 5, pp. 127-147, 2020.

[30] JJE. Slotin and W.Li,*Applied nonlinear control,* Englewood Cliffs, NJ: Prentice Hall, 1991.

[31] R. Shadi, F. Bouzari Liavoli and A.Fakharian,"Nonlinear Sub-Optimal Controller for Ebola Virus Disease: State-Dependent Riccati Equation Approach",*7th International Conference on Control, Instrumentation and Automation*, Tabriz, Iran, 23-24 Feb. 2021, pp. 1-6.

[32] R. Shadi, A. Fakharian and H. Khaloozadeh,"Transmission Dynamics and Optimal Control Strategy to Mitigate the Spread of Novel Coronavirus: The Case of Iran",*The 30th Iranian Conference on Electrical Engineering (ICEE*), Tehran, Iran, 17-19 May. 2022, pp. 356-361.

[33] M. Sheikhan and SA. Ghoreishi, "Antiviral therapy using a fuzzy controller optimized by modified evolutionary algorithms: a comparative study", *Neural Computing and Applications*, vol. 23, pp. 1801-1813, 2013.

[34] MA. Nowak, et al, "Viral dynamics in hepatitis B virus infection", *Proceedings of the National Academy of* Sciences, vol. 93, pp. 4398-4402, 1996.

Multipath Cascaded Single Stage Distributed Power Detector; Analysis and Design

Nader Javadifar[1*], Yaghoob Mohammadmoradi[2], Atila skandarnezhad[3]

Abstract – In this paper, a novel technique to improve the input dynamic range and bandwidth of the microwave power detectors (PDs) simultaneously is presented, which utilizes the piecewise linear approximation in conjunction with the distributed structure to achieve the goals. This is an efficient method that requires less number of active devices and therefore saves more power consumption and active chip area, rather than the previous conventional methods. The analysis of circuit based on the transmission line theory is discussed and the transfer function is extracted mathematically for the proposed model. Moreover, a transistor level design is performed using a 0.15μm PHEMT GaAs technology for 24GHz applications. The post layout simulation results are presented.

Keywords: Distributed circuit; Logarithmic amplifier; Microwave; Power detector; Transmission line.

1. Introduction

Many of applications such as wireless communication systems and test & measurement equipment need to detect radio frequency (RF) signals at very weak levels. Signal level can be detected by generating an output proportional to the power of input signal.

On the other hand, the power level of RF signals in many cases varies in the wide range of multi octave in decibel (dB). For that reason, the linear amplifiers are inefficient to detect such signals, because they have a limited dynamic range.

An alternative approach has been used is the logarithmic amplifiers (commonly known as log-amps). The log-amps are divided into the two main categories: The first, those that utilize the inherent exponential characteristic of the semiconductor devices such as p-n junction diodes [1], bipolar transistors [2], [3] and MOSFET transistors in weak inversion region [4-6]. However, they don't have a wide dynamic range usually. Furthermore, the operation of these devices is restricted at higher frequencies, because of the modification of their fundamental characteristic equation by parasitic elements [7]. To extend the total input dynamic range, the second type of log-amps employs more complex configuration with multiple parallel paths. The successive detection logarithmic amplifier (SDLA) is the most famous configuration, which uses this approach to realize the approximated logarithmic function (Fig. 1) [8-13].

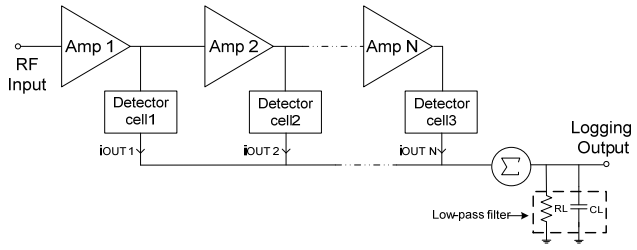


Fig.1. General block diagram of SDLA

The SDLA is conventionally made up of a number of limiting-amplifier stages in a cascaded manner. Each amplifier drives a detector cell (or rectifier cell) to convert its RF input signal to a DC voltage proportional to the amplitude of the RF input. Sum of the detector cells output forms a desired logarithmic transfer function by using the piecewise linear approximation method [14]. The more number of stages in this circuit, results in a wider dynamic

1* Corresponding Author: Department of Electrical Engineering , Aliabad Katoul Branch,Islamic Azad University, Aliabad Katou, Iran.
Email: javadifar@aliabadiau.ac.ir
2 Department of Physics , Aliabad Katoul Branch,Islamic Azad University, Aliabad Katou, Iran.Email: mohammadmoradi@aliabadiau.ac.ir
3 Department of Electrical Engineering , Aliabad Katoul Branch,Islamic Azad University, Aliabad Katou, Iran.
Email: eskandarnejad@aliabadiau.ac.ir
Received: 2024.01.13 ; Accepted:2024.04.13

range. However, this decreases the total bandwidth of circuit as below [15]:

$$F_T = F_S\sqrt{2^{1/N}-1} \quad (1)$$

Where F_T indicates the total bandwidth of the cascaded circuit with N gain stages and F_S is the bandwidth of a single gain stage. Therefore, a very strict trade-off between the total dynamic range and frequency of operation exists in this topology, where improving one proportionally degrades the other.

In this paper, a microwave PD based on the SDLA structure in combination with distributed circuit theory is proposed to overcome this challenge. Section II, describes the principle of PD design based on the distributed circuit concept. The analysis and design of proposed distributed PD is presented in section III. The post layout simulation results are shown in section IV, and finally a conclusion is given in section V.

2. Distributed PD concept and structure

The term 'distributed' has been used quite the opposite of the 'lumped' notion in many of literature. However, for the purposes of this paper, a distributed system is defined to be a system with multiple parallel signal paths and devices, cooperating in harmony to achieve a desired task [16]. On the basis of this definition, the distributed circuit configuration is in contrast to the conventional serial cascaded systems, which have a single signal path. In addition, the physical size of the circuit compared with the signal wavelength has no importance in this view.

As mentioned in the preceding section, the total bandwidth of an analog cascaded circuit block decreases by increasing the number of cascaded stages. This stems from this fact that the more number of stages (or sections), imposes additional poles to the circuit. In distributed systems, however, the core of a given circuit is placed in the multiple parallel paths between the two artificial transmission lines, so that the parasitic capacitances may be absorbed into the line's LC sections. The examples of such circuits has been presented previously are distributed amplifier (DA) [17-20], distributed voltage controlled oscillator (DVCO) [21] and distributed active transformer (DAT) power amplifier [22].

Generalize this idea to logarithmic detector topologies such as SDLA, necessitates exploitation of the two transmission lines (including on-chip passive lumped inductors) that the core of PD circuit is embedded between them and through which the RF input signal can travel without deterioration of the bandwidth. The parasitic capacitances will be merged with the lumped inductors in this condition and the artificial transmission lines are constructed. This technique allows the core of PD is repeated in a cascaded structure to extend the total dynamic range without sacrificing the bandwidth.

Fig. 2 shows the simplified schematic of the proposed PD core, which is suggested based on the described idea. This comprises a single stage DA as an amplifier stage followed by a simple common gate MOS as a detector cell.

The parasitic capacitances in connection with the amplifier stage are absorbed into the two artificial transmission lines, which we called them input line and inter-stage line, respectively. The detector cell parasitic capacitances at its input node are also absorbed into the inter-stage line. Because of the desired signal at the PD output is the DC portion, there is no concern associated with the parasitic capacitances at the output tips of the detector cell.

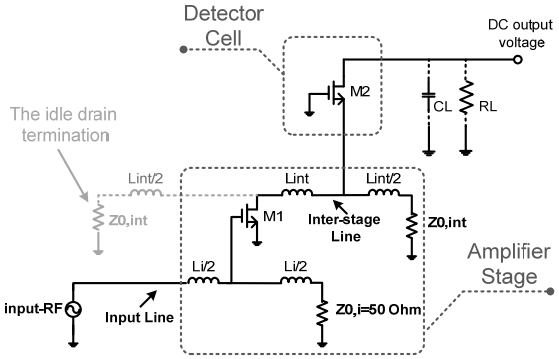


Fig. 2. Simplified schematic of proposed PD core

The power detection investigation by using the above PD core will be described in the next section. Moreover, we will exploit the proposed PD core in a SDLA structure to extend the total dynamic range. Extraction of the logarithmic function for the SDLA will be presented subsequently.

3. Circuit analysis and design

3.1. RF to DC conversion for the proposed PD core

The power detection process is often performed in three steps: V/I (voltage to current) conversion, RF to DC conversion and finally I/V conversion [4], [23-24]. A voltage may be converted to current easily by using a MOS

transistor through its transconductance. Moreover, the RF to DC conversion can be realized by using a single MOS transistor in one of the common-source, common-drain or common-gate forms, like those we have proposed as a detector cell in Fig. 2. We will prove this claim in the continuation.

As seen from Fig. 2, the amplifier stage in the proposed PD core is a single stage DA, comprises an active device (M1) and the two LC ladders at its input and output as a transmission line. The detector cell is a simple common gate transistor M2, which produces a DC voltage proportional to the power of amplified RF signal at the drain of M1. The parasitic capacitances of M1 and M2 in conjunction with the lumped inductors of the input and inter-stage lines make the two artificial transmission lines. Contrary to conventional DA structure, the idle drain line termination (marked with dotted line in this figure) has been eliminated in this configuration. Therefore, the voltage swing across the output drain of M1 is increased, which consequently enhances the amplifier's gain.

In general, the transfer characteristic curve of a linear amplifier may be shown as Fig. 3. As seen, a linear amplifier has the two operation modes. If the input signal amplitude is less than a threshold value of $V_{i,sat}$, then it works in amplifying (or linear) mode and if the amplitude is larger than this value, the amplifier is forced to operate in limiting mode.

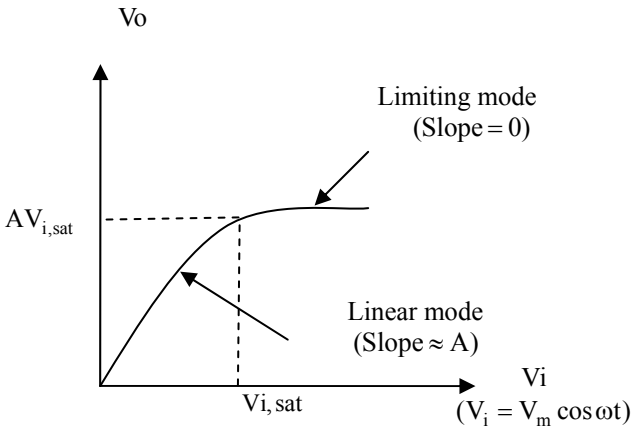


Fig. 3. General characteristic curve of a linear amplifier

By assumption that the single stage DA in Fig. 2 operates in linear mode, its output is a sine wave proportional to the RF input signal. In this condition, the detector transistor M2 meets a sine wave at its input.

The instantaneous current at the output of detector cell, now is given by:

$$i_{OUT} = K(V_{GS_2} + v_{gs_2} - V_t)^2 = K(V_{GS_2} - V_t)^2 + K[v_{gs_2}{}^2 + 2(V_{GS_2} - V_t)v_{gs_2}] \quad (2)$$

where the V_{GS_2} and v_{gs_2} indicate the DC and AC portion of the gate to source voltage of M2, respectively and V_t is the threshold voltage. From above, the instantaneous current of detector cell can be divided into the two individual parts, we name them i_{out1} and $i_{out\,2}$. The first part, $i_{out1} = K(V_{GS_2} - V_t)^2$, is a DC portion due to the biasing voltage and just affects the DC offset of the output current. The second part, $i_{out\,2} = K[v_{gs_2}{}^2 + 2(V_{GS_2} - V_t)v_{gs_2}]$, originates from the RF signal at the input of M2.

By considering the RF input signal of the detector cell transistor M2 as $v_{i,rf}$ and from the simplified schematic of Fig. 2, we have:

$$v_{gs_2} = -v_{i,rf} \quad (3)$$

Substituting (3) into the second part of equation (2), gives:

$$i_{out\,2} = K[v_{i,rf}{}^2 - 2(V_{GS_2} - V_t)v_{i,rf}] \quad (4)$$

By considering the RF received signal at the gate of M1 equal to $v_{g1} = V_m \cos \omega t$, magnitude of amplified signal at its output is given by:

$$|v_{i,rf}| = |v_{d1}| = |I_{d1}|Z_0 = g_{m1}Z_{0,int}|v_{g1}| = g_{m1}Z_{0,int}V_m \cos \omega t \quad (5)$$

The artificial transmission lines have been supposed lossless in above equation. From (4) and (5), we have:

$$|i_{out2}| = \frac{K}{2}(g_{m1}Z_{0,\text{int}}V_m)^2 + \frac{K}{2}(g_{m1}Z_{0,\text{int}}V_m)^2 \cos 2\omega t - 2K(V_{GS_2} - V_t)g_{m1}Z_{0,\text{int}}V_m \cos \omega t \quad (6)$$

A low-pass RC filter at the output eliminates the AC terms and so a constant DC current remains as:

$$i_{out,dc} = \frac{K}{2}\left(g_{m1}Z_{0,int}\right)^2 V_m{}^2 \qquad (7)$$

This proves a MOS transistor is capable to convert a part of its RF input signal to a DC value, proportional to the RF input power. If the amplifier stage operates in limiting mode, its output is a constant saturated voltage of $AV_{i,sat}$ and therefore there is no RF signal at the input of detector cell (i.e., $v_{i,rf} = 0$). Accordingly, the detector cell adds just a constant DC current to output, independent of an RF input signal amplitude.

3.2.Multipath cascaded single stage distributed power detector

The PD core proposed in previous section was designed based on the distributed circuit theory, especially to operate at microwave frequencies using minimum active devices. Nevertheless, it has a limited dynamic range and has not sufficient sensitivity to detect very week RF signals. Accordingly, a proper enhancement must be done for this structure to increase the dynamic range using the logarithmic topologies.

At first glance, it maybe seems that achieving a logarithmic transfer function is possible by developing the proposed PD core in a parallel summation schematic of Fig. 4(a). Indeed, this structure is the well-known topology of DA that each section drives a detector cell at its output. The main drawback concerned with this structure is the additive gain nature of the DA. Assuming that the transmission lines are lossless, the voltage gain at the last section of the DA drain line can be estimated by [25]:

$$A_V \approx \frac{1}{2} N.g_m.Z_{0,int} \qquad (8)$$

As seen, no logarithmic function can be derived from this equation.

As an alternative, we can use the proposed PD core in a multipath cascaded single stage distributed power detector (CSSDPD) configuration of Fig. 4(b), which is inspired from the SDLA configuration. The complete circuit schematic along with the biasing network has been indicated in this figure.

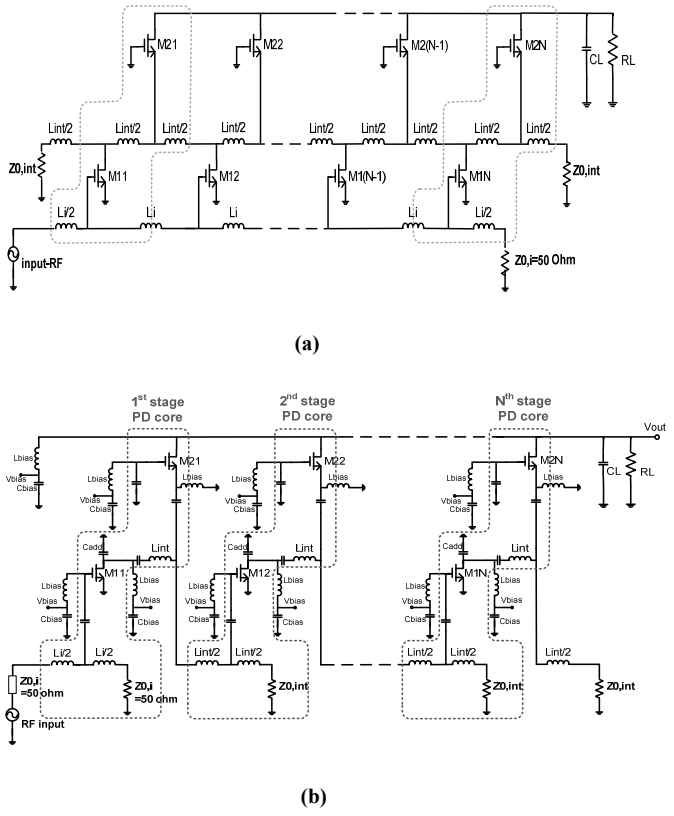


Fig. 4. a) Extension of the proposed PD core with additive gain nature; b) A complete schematic of the multipath CSSDPD with multiplicative gain nature.

Cascade of single stage distributed amplifier stages takes advantage of multiplicative gain feature [26]. This arises from the fact that the RF input signal is multiplied by the gain of each amplifier stage and so emerges multiplicatively at the input of subsequent stages. Summing the detected signal of each core through the multiple parallel paths successively approximates the desired logarithmic transfer function at the output node [14].

There are two artificial transmission lines in this configuration. One of them is the input line of the first stage PD core, we called it 'input line'. Another is called 'inter-stage line', which includes all of the other transmission lines. Only the input line must be terminated to 50Ω. There is no necessity for the inter-stage transmission lines to be matched to 50Ω. This gives the more degree of freedom to design this circuit and allows the inter-stage lines to have a larger characteristic impedance, which means attaining a given gain requirement by using the less number of sections compared with the conventional DA structure. Additionally, this consumes lower power and occupies lower chip area. Unlike the usual structure of DA, the phase velocity equalization is not mandatory in this

circuit, also.

Another problem related with the DA is the lossy nature of transmission lines, which results from the low quality of on-chip inductors. Accordingly, the losses in the long artificial transmission lines may be greater than the amount of signal amplification. This provides extra restrictions on this circuit, from the viewpoint of maximum number of sections. It is in contrary to achieve a large dynamic range. Cascade of single stage DA (in Fig. 4(b)) can solve this problem, also. Because the transmission lines are not continuous in this topology and the line of each section is separated from the others. However, an additional capacitor (C_{add}) is probably required in each inter-stage line of CSSDPD, to equalize the LC sections capacitance.

The bandwidth of proposed CSSDPD is restricted here by minimum cut-off frequency of the two transmission lines. Supposing all sections are similar, the cut-off frequency of two transmission lines and their characteristic impedance can be expressed as:

$$f_{c,i} = \frac{1}{\pi\sqrt{L_i C_{gs_{1j}}}}$$

$$Z_{0,i} = \sqrt{\frac{L_i}{C_{gs_{1j}}}} = 50^{\Omega}$$

$$f_{c,\text{int}} = \frac{1}{\pi\sqrt{L_{\text{int}} C g s_{1j}}} = \frac{1}{\pi\sqrt{L_{\text{int}}\left(C g s_{2k} + C_{ds_{2k}}\right)}} = \frac{1}{\pi\sqrt{2L_{\text{int}} \times \left(C_{ds_{1j}} + C_{add}\right)}}$$

$$Z_{0,\text{int}} = \sqrt{\frac{L_{\text{int}}}{C_{gs_{1j}}}} = \sqrt{\frac{L_{\text{int}}}{C_{gs_{2k}} + C_{ds_{2k}}}} = \sqrt{\frac{L_{\text{int}}/2}{C_{ds_{1j}} + C_{add}}} \tag{9}$$

A mathematical analysis to derive a logarithmic transfer function for this circuit is provided in next section.

3.3. Transfer function extraction

From the Fig. 4(b), by considering the RF input signal as $v_i = V_m \cos\omega t$ and the voltage gain of amplifier stages in linear mode as $A = g_{m_{11}} Z_{0,\text{int}}$, the voltage at drain tips of amplification transistors are given as:

$$\begin{aligned}
\left|v_{d_{11}}\right| &= \left|I_{d_{11}} Z_{0,\text{int}}\right| = \left|g_{m_{11}} v_{g_{11}} Z_{0,\text{int}}\right| = \frac{1}{2} A.\left|v_i\right| \\
\left|v_{d_{12}}\right| &= \left|I_{d_{12}} Z_{0,\text{int}}\right| = \left|g_{m_{11}} v_{g_{12}} Z_{0,\text{int}}\right| = \frac{1}{2} A^2.\left|v_i\right| \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
\left|v_{d_{1N}}\right| &= \left|I_{d_{1N}} Z_{0,\text{int}}\right| = \left|g_{m_{11}} v_{g_{1N}} Z_{0,\text{int}}\right| = \frac{1}{2} A^N.\left|v_i\right|
\end{aligned} \tag{10}$$

Where $g_{m_{11}}$ the transconductance of the amplifier is stages and $v_{g_{1N}}$ is the voltage at the gate of Nth amplification transistor. Multiplicative gain can be obviously observed from these equations.

The gate voltage of each amplifier may be expressed as:

$$\begin{aligned}
\left|v_{g_{11}}\right| &= \frac{1}{2}\left|v_i\right| \\
\left|v_{g_{12}}\right| &= \left|v_{d_{11}}\right| = \frac{1}{2} A.\left|v_i\right| \\
&\vdots \qquad\qquad \vdots \\
\left|v_{g_{1N}}\right| &= \left|v_{d_{1(N-1)}}\right| = \frac{1}{2} A^{N-1}.\left|v_i\right|
\end{aligned} \tag{11}$$

On the other hand, the DC converted current at the output of jth path may be rewritten by generalizing the equation (7) as following:

$$i_{out,dc_j} = \frac{K}{2}\left(g_{m1} Z_{0,\text{int}}\right)^2 \left|v_{g_{1j}}\right|_{\max}^{\ 2} = \frac{K}{2} A^2 \left|v_{g_{1j}}\right|_{\max}^{\ 2} = \frac{K}{2} A^{2j}\left(\frac{V_m}{2}\right)^2, \qquad j = 1,2,...,N \tag{12}$$

In above, $\left|v_{g_{1j}}\right|_{\max}$ is the peak of voltage magnitude at the gate of amplifying transistor in jth section.

For a multipath CSSDPD with N stages, when all amplifier stages are in linear mode, the converted DC current of each path can be written as below:

$$\begin{aligned} i_{out,dc_1} &= \frac{K}{2}A^2\left(\frac{V_m}{2}\right)^2 \\ i_{out,dc_2} &= \frac{K}{2}A^4\left(\frac{V_m}{2}\right)^2 \\ &\vdots \qquad\qquad \vdots \\ i_{out,dc_N} &= \frac{K}{2}A^{2N}\left(\frac{V_m}{2}\right)^2 \end{aligned} \tag{13}$$

In limiting mode also, the amplifier stages output has a saturated value of $AV_{i,sat}$ and the output of paths in this case considered as a constant value of C.

By boosting the RF input signal in the cascaded structure, some of the amplifier stages may be in linear mode and some others in limiting mode. As a result, the converted DC currents summation of N parallel paths at output node may have $(N+1)$ situations:

$$\begin{aligned} &I_{O-S1} = NC && V_{i,sat} \le V_m \\ &I_{O-S2} = i_{out,dc_1} + (N-1)C && \frac{V_{i,sat}}{A} \le V_m < V_{i,sat} \\ &\vdots \qquad\qquad \vdots \\ &I_{O-SJ} = i_{out,dc_1} + \cdots + i_{out,dc_{(J-1)}} + (N-J+1)C && \frac{V_{i,sat}}{A^{J-1}} \le V_m < \frac{V_{i,sat}}{A^{J-2}} \\ &\vdots \qquad\qquad \vdots \\ &I_{O-S(N+1)} = i_{out,dc_1} + \cdots + i_{out,dc_{(N-1)}} + i_{out,dc_N} && \frac{V_{i,sat}}{A^{N}} \le V_m < \frac{V_{i,sat}}{A^{N-1}} \end{aligned} \tag{14}$$

In above, I_{O-SJ} indicates the converted DC output current at Jth situation, when $(J-1)$ stages operate in linear mode and $(N-J+1)$ ones in limiting mode.

From (14), given the RF input voltage is at its lower boundary at Jth situation, i.e. $V_m = \frac{V_{i,sat}}{A^{J-1}} = \frac{AV_{i,sat}}{A^J}$, we have:

$$J = \log_A\left(\frac{A.V_{i,sat}}{V_m}\right) \tag{15}$$

Replacing (13) in (14) for the Jth situation gives us:

$$\begin{aligned} I_{O-SJ} &= \frac{K}{2}A^2\left(\frac{V_m}{2}\right)^2 + \cdots + \frac{K}{2}A^{2(J-1)}\left(\frac{V_m}{2}\right)^2 + (N-J+1)C \\ &= \frac{K}{2}\left(\frac{V_m}{2}\right)^2 \times A^2 \times \frac{A^{2(J-1)}-1}{A^2-1} + (N-J+1)C \end{aligned} \tag{16}$$

The above equation by assumption that the gain of linear mode (A) is sufficiently greater than 1, can be approximated as:

$$I_{O-SJ} \approx \frac{K}{2}\left(\frac{V_m}{2}\right)^2 \times A^{2(J-1)} + (N-J+1)C \tag{17}$$

Substituting (15) in above and using straightforward calculations, the DC output current at Jth situation is given by:

$$\begin{aligned} I_{O-SJ} &\approx C\left[\frac{K.V_{i,sat}{}^2}{8C} + N + 1 + \log_A\left(\frac{V_m}{A.V_{i,sat}}\right)\right] \\ &\approx \frac{C}{\log_{10} A} \times \log_{10}\left(\frac{V_m}{\frac{V_{i,sat}}{A^{\left(\frac{K.V_{i,sat}{}^2}{8C}+N\right)}}}\right) \end{aligned} \tag{18}$$

The derived equation shows a logarithmic relationship between the output current of CSSDPD and the amplitude of RF input signal explicitly. The term $\frac{C}{\log_{10} A}$ is a constant value which indicates the slope of logarithmic function. The term $\frac{V_{i,sat}}{A^{\left(\frac{K.V_{i,sat}{}^2}{8C}+N\right)}}$ is reversely proportional to the number of sections (N) and will decrease by increasing N. This determines the beginning point of the logarithmic function, so that when the RF input signal amplitude (V_m) reaches this value, I_{O-SJ} will be zero. Accordingly, the more number of sections cause to decrease the minimum detectable input signal and increases the sensitivity.

4. Simulation results, layout and comparison

Over the microwave frequencies, the atmospheric loss meets a peak near the centre of 24GHz band $(24.05-24.25\text{GHz})$, which mainly occurs due to the absorption by water vapor molecules. Consequently, detecting the RF signals at this band encounters a challenge especially for those applications trying to cover distances of more than a few kilometers. This necessitates the design of power detectors with very high sensitivity for this band.

Here, the simulation results of a logarithmic power detector for 24GHz band based on the twelve stage CSSDPD circuit by using the $0.15\mu\text{m}$ PHEMT GaAs technology is presented.

4.1. Simulation results of PD core

Fig. 5(a) depicts the DC portion of output voltage for the proposed PD core versus the RF input amplitude sweep from -1.6V to $+1.6\text{V}$. The curve shown in this figure, belongs to an even function of equation (7) and confirms that the output characteristic of the proposed PD core is similar to those which can be achieved by a full-wave rectifier. Accordingly, the proposed PD core can do the desired task with just one active device, which is much lower than those are required by a full-wave rectifier in conventional SDLA configuration [8-10].

Fig. 5(b) indicates the transient response of output voltage versus the RF input amplitude variations from 0 to 1.6V. This shows the DC portion of output voltage increases by increasing the RF input amplitude and proves the RF input signal is converted to a DC voltage proportionally by the proposed PD core. The ripples are seen in this figure can be alleviated by applying a larger load capacitor (C_L) in the output RC filter (shown in Fig. 2).

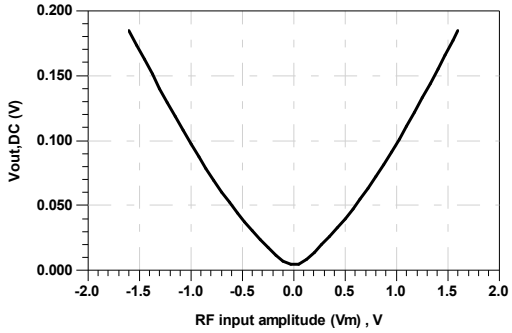


Fig.5(a). The DC component of output voltage as a function of input amplitude

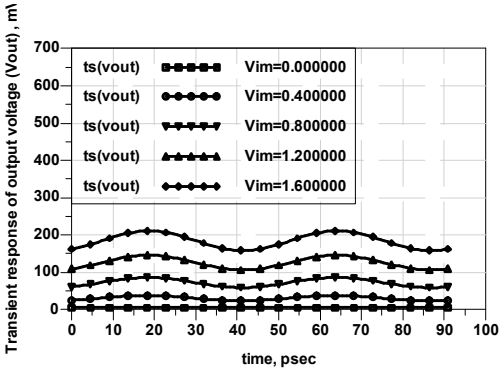


Fig. 5(b). Transient response of the proposed PD core for $\text{C}_\text{L}=105\text{fF}$

4.2. Simulation results for the twelve stage multipath CSSDPD

Fig. 6 depicts the DC output voltage of the twelve stage CSSDPD versus the amplitude of RF input. The logarithmic behavior is obvious in this figure.

Moreover, the transfer characteristic of proposed power detector and its detection error is shown in Fig.7. Indeed, this figure is the same as Fig. 6, except that the horizontal axis is in logarithmic form. As seen, the total dynamic range is 110dBm from -94 to $+16\text{dBm}$ and the maximum detection error over the dynamic range is $\pm 1.5\text{dBm}$. Minimum detectable power is -94dBm, which shows an ultra-high sensitivity.

Fig. 8 shows the total power gain (S_{21}) and the input return loss (S_{11}) in dB. As shown, the power gain is very flat $(108.1\pm 0.1\text{dB})$ over the frequency band. Moreover, S_{11} is less than -20dB over the operating frequency range, which boasts ease of input matching for the cascaded single stage DA topology.

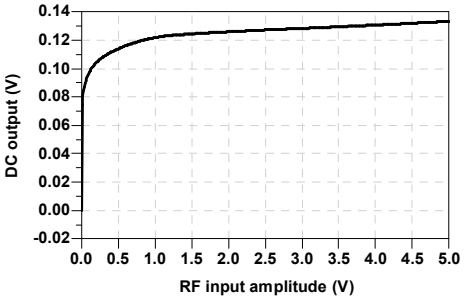


Fig. 6. DC output voltage of CSSDPD versus the RF input amplitude

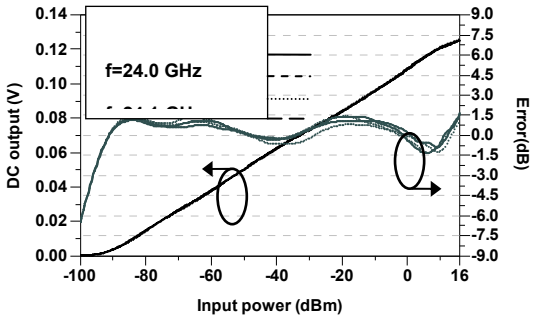


Fig.7. Transfer characteristic of the twelve stage CSSDPD and its error

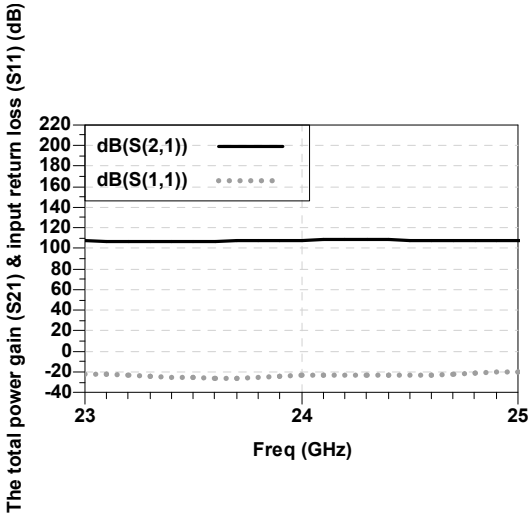


Fig.8. S-parameters of twelve stage CSSDPD

The layout of the designed power detector is shown in Fig.9. The total chip area including twelve stages, supplies nets and pads is $1.5\text{mm} \times 3.6\text{mm}$. This is $1\text{mm} \times 0.32\text{mm}$ for each PD core. The total power consumption is 194mW from $\pm 3\text{V}$ power supplies.

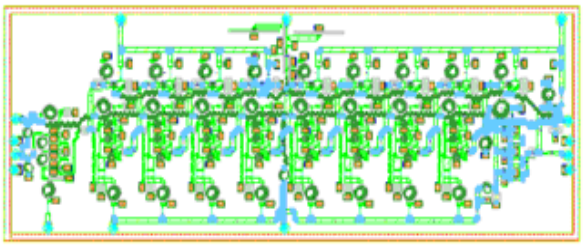


Fig. 9. Layout of the designed multipath CSSDPD

Table 1 summarizes the proposed power detector characteristics and compares this work with others.

Table1. Comparison of logarithmic power detectors

Ref. no.	Freq. (GHz)	Dynamic range (dB)	Logging error (dB)	Power cons. (mW)	Supply voltage (V)	Topology &Technology
[8]	DC to 8	40 60	±1	70 150	2.8	(SDLA) 180nm CMOS **(measured)**
[9]	1	65 74	±2 ±3	24	1.2	(SDLA) 65nm CMOS **(simulated)**
[10]	2 to 18	33	±1	690	-5	(SDLA) InP DHBT **(measured)**
[12]	0.5 to 1.2	60	±2.5	700	NA	(SDLA) GaAs **(measured)**
[13]	0.5 to 1.5	60	±1	400	±5	(SDLA) GaAs HBT **(measured)**
This work	24 to 24.3	95 110	±1 ±1.5	194	±3	(SDLA) 0.15 μm GaAs PHEMT **(simulated)**

5. Conclusion

In this study, the theory of a linear-in-dB microwave power detector with novel structure was analyzed and mathematically discussed. The PD core configuration exploited here is ultra wideband and requires the minimum number of transistors. A significant progress has been made in extending the dynamic range at frequencies of more than $20GHz$, while keeping the power consumption as low as possible. The total dynamic range of $110dBm$ with the minimum detectable power of $-94dBm$ has been achieved. The detection error over the total dynamic range is less than $\pm 1.5dB$. It is also less than $\pm 1dB$ from $-94dBm$ to $+1dBm$. The total power consumption of designed power detector is $194mW$.

REFERENCES

[1] Huggins, R. W. "Analytical fit of the transfer function of a logarithmic electrometer and correction for ambient temperature variations," Review of Scientific Instruments, 1973, Vol. 44, no.3, pp. 297–300.

[2] Meyer, R.: "Low-power monolithic RF peak detector analysis," IEEE J. Solid-State Circuits, 1995, Vol. 30, no.1, pp. 65–67.

[3] Wilson, B., & Al-Gahtani, M. "Improved logarithmic converter based on a transconductance feedback amplifier," In the 2001 IEEE international symposium on circuits and systems ISCAS, Vol. 1, pp. 651–654.

[4] Rami, S., Tuni, W., and Eisenstadt, W.R. 'Millimeter wave MOSFET amplitude detector', IEEE Silicon Monolithic Integrated Circuits in RF Systems (SiRF), 2010, pp. 84–87

[5] Huang, C., & Chakrabartty, S. "Current-input current-output CMOS logarithmic amplifier based on translinear Ohm's law," Electronics Letters, 2011, Vol. 47, no.7, pp. 433–434.

[6] K. A. Townsend and J. W. Haslett, "A wideband power detection system optimized for the UWB spectrum", IEEE J. Solid-State Circuits, vol. 44, no. 2, pp. 371–381, Feb. 2009.

[7] Barber, W. L., & Brown, E. R., "A true logarithmic amplifier for radar IF applications," IEEE Journal of Solid-State Circuits, 1980, Vol. 15, no.3, pp. 291–295.

[8] Wu, J., Hsu, K., Lai, W., To, C., Chen, S., Tang, C., and Juang, Y.: "A linear-in-dB radio-frequency power detector," IEEE Microwave Symposium Digest (MTT), 2011, pp. 1-4.

[9] K. Kiela, M. Jurgo, and R. Navickas, "Design of a linear-in-dB power detector in 65nm CMOS technology", *ELEKTRONIKA IR ELEKTROTECHNIKA*, 2013, Vol. 19, no. 10, pp. 91-94.

[10] Haynes, M.: "Wideband Monolithic SDLA Design Using InP DHBT Technology," IET Seminar on RF and Microwave IC Design, 2008, pp. 1-6.

[11] Kenneth A. Townsend, James W. Haslett, "A Wideband Power Detection System Optimized for the UWB Spectrum," IEEE Journal of Solid-State Circuits, Vol. 44, No. 2, February, 2009.

[12] R. Michels, N. Scheinberg, J. Gluck, "An L-Band Temperature Compensated Ultra Low Power Successive Detection Logarithmic Amplifier," IEEE MTT-S International Microwave Symposium Digest, pp. 541-4, 1989.

[13] G. M. Gorman, A. K. Oki, E. M. Mrozek, J. B. Camou, D. K. Umemoto, and M. E. Kim, "A GaAs HBT Monolithic Logarithmic IF (0.5 TO 1.5 GHz) Amplifier with 60 dB Dynamic Range and 400 mW Power Consumption," IEEE MTT-S Digest, vol.2, pp. 537-540, 1989.

[14] R. S. Hughes, Logarithmic Amplification with Application to Radar and EW. Dedham, Artech House, 1986.

[15] P.-C. Huang, Y.-H. Chen, and C.-K. Wang, "A 2-V 10.7-MHz CMOS limiting amplifier/RSSI," *IEEE J. of Solid-State Circuits*, vol. 35, no. 10, pp. 1474–1480, Oct. 2000.

[16] A. Hajimiri, "Distributed Integrated Circuits: An Alternative Approach to High-Frequency Design," IEEE Communication Magazine, 2002, Vol. 40, no. 2, pp. 168-173.

[17] Kambiz Moez and Mohamed Elmasry, "A Low-Noise CMOS Distributed Amplifier for Ultra- Wideband Applications," IEEE Transactions on Circuits and Systems II, Vol. 55, no. 2, pp 126-130, Feb. 2008.

[18] A. Ghadiri and K. Moez, "Compact Transformer-Based Distributed Amplifier for UWB Systems," IEEE Transactions on Circuits and Systems - II: Express Briefs, vol. 58, no. 5, pp. 259-263, May 2011.

[19] Wai-Kai Chen, "Theory and design of distributed amplifiers," International Journal of Electronics, vol.26, no. 5, pp. 405-421, 1969.

[20] Kenji Kumabe and Hiroshi Kanbe, "GaAs travelling-wave amplifier," International Journal of Electronics, vol. 58, no. 4, pp. 587-611, 1985.

[21] H. Wu and A. Hajimiri, "Silicon-Based Distributed Voltage Controlled Oscillator," IEEE J. Solid-State Circuits, Mar. 2001.

[22] Aoki, I.; Kee, S.D.; Rutledge, D.B.; Hajimiri, A., "Fully integrated CMOS power amplifier design using the distributed active-transformer architecture," J. Solid-State Circuits, vol.37, no.3, pp.371-383, Mar. 2002.

[23] Valdes-Garcia, A., Venkatasubramanian, R., Srinivasan, R., Silva-Martinez, J., & Sinencio E. S. "A CMOS RF rms detector for built in testing of wireless transceivers," 23rd IEEE VLSI Test Symposium, doi:10.1109/VTS.2005.8.

[24] Yin, Q., Eisenstadt, W. R., Fox, R. M., & Zhang, T. (2005). "A translinear rms detector for embedded test of RF ICs," IEEE Trans. On Instrumentation and Measurement, doi:10.1109/TIM.2005.855105.

[25] H. Wu and A. Hajimiri, "Silicon-Based Distributed Voltage Controlled Oscillator," IEEE J. Solid-State Circuits, Mar. 2001.

[26] Liang, J. Y. and C. S. Aitchison, "Gain Performance of Cascade of Single Stage Distributed Amplifiers," Electronics Letters, Vol. 31, No. 15, July 1995, pp.1260-1261.