Journal Pre-proofs

Regular paper

Optimization of Digital Predistortion Models for RF Power Amplifiers Using a Modified Differential Evolution Algorithm

Sanjika Devi R V, Robin Kalyan, K.R Bindu, Dhanesh G. Kurup

PII: S1434-8411(20)30304-6
DOI: https://doi.org/10.1016/j.aeue.2020.153323
Reference: AEUE 153323

To appear in: *International Journal of Electronics and Communications*

Received Date: 6 February 2020
Accepted Date: 18 June 2020

Please cite this article as: S.D. R V, R. Kalyan, K.R Bindu, D.G. Kurup, Optimization of Digital Predistortion Models for RF Power Amplifiers Using a Modified Differential Evolution Algorithm, *International Journal of Electronics and Communications* (2020), doi: https://doi.org/10.1016/j.aeue.2020.153323

Optimization of Digital Predistortion Models for RF Power Amplifiers Using a Modified Differential Evolution Algorithm

Sanjika Devi R V*[1], Robin Kalyan[2], K R Bindu[2] and Dhanesh G. Kurup[1]

[1]Department of Electronics and Communication Engineering,

Amrita School of Engineering, Bengaluru,

Amrita Vishwa Vidyapeetham,

India.

*Email: r_sanjika@blr.amrita.edu

[2]Communication Systems Group, U R Rao Satellite Centre,

Indian Space Research Organization (ISRO), Bengaluru,

India.

Abstract

This paper, investigates and presents the optimal parameter identification of digital pre-distortion (DPD) models for radio frequency power amplifiers (RF PAs) using a modified differential evolution (MDE) based optimization algorithm. Compared to the conventional exhaustive search method which is computationally intensive, our proposed approach enables the identification of a best-fit DPD model from a combinatorially large model space in a short time. In addition, applying information criteria based objective functions in the optimization process enables us to achieve sparse selection of dynamical models, which balances the model accuracy and model complexity. Experimental validation on a GaN based class AB power amplifier illustrates that, our proposed approach was able to accurately identify complexity reduced optimal DPD models without compromising the modeling accuracy.

Keywords

RF power amplifier, digital predistortion, modified differential evolution algorithm.

I. Introduction

The proliferation of high data-rate wireless communication systems [1] with large operational distances, has increased the requirement for power during signal transmission [2]. RF power amplifier fulfills this requirement of providing adequate power for signal transmission. However, power amplifiers consume major part of the available power [3] which causes problems especially to battery operated wireless devices. Therefore, for achieving high power efficiency, power amplifiers has to be operated in the non-linear region. However, operating power amplifiers in nonlinear region heavily distorts the signal causing degradation in the bit error rate (BER) performance of the system, thereby affecting the quality of service. Hence, designing power amplifiers with high power efficiency and acceptable signal distortions is a major research area [4][5].

Over the past few years, data-driven modeling techniques are finding tremendous applications in various fields of engineering [6]-[8]. Behavioral modeling of PAs, which is essentially a data-driven modeling technique, thus enables us to derive accurate mathematical description of the PA from its simulations or measurements [9]. Once RF PAs are modeled accurately, the input baseband signal corrected with inverse model of the PA known as digital predistortion (DPD) technique, will enable us to linearize the PA characteristics for increased power efficiency [10].

The PA behavioral models reported in the literature can be classified into two main categories, namely the non-parametric and parametric based models. The non-parametric based PA models essentially provides a functional mapping between the measured input and outputs with no prior knowledge of the underlying nonlinear dynamics of the system. On the other hand, parametric models provides a meaningful insight into the physical behavior of the PA as a nonlinear dynamical system. Typical parametric based PA models include memory polynomials (MP) [11], generalized memory polynomials (GMP) [12], dynamic-deviation-reduction Volterra method [13] and different variants of these three models such as PLUME, augmented complexity-reduced GMP and Weighted MPs [14]-[16]. On the other hand, Artificial Neural Networks (ANN) [17]-[20] and Support Vector Regression (SVR) [21] are some of the non-parametric models used for modeling RF PAs. Compared to the non-parametric models which are adopted when the system model structure is unknown, the parametric models provide an accurate model structure whose parameters needs to be identified. However, the behavioral modeling and DPD techniques based on parametric based models, is challenging due to the fact that, complexity of model which depends on number of model coefficients is unknown. Also, parsimonious representation of PAs to model the complex nonlinear characteristics of the PA with minimal model coefficients is also a major challenge.

GMP model has excellent modeling accuracy when the model parameters are accurately identified [22]. However, one of the major challenge in GMP model is the selection of the best GMP model out of an enormous number of combinatorial models. Indeed, identifying the parameters of GMP model generally requires an intractable search through all possible combinatorial models. Hence, model selection process becomes computationally intractable even for a modest number of nonlinear order and memory depth. In such cases, applying optimization techniques significantly reduces the computational time. Particle swarm optimization (PSO), Hill Climbing algorithm and Genetic algorithms (GA) [23]-[28] are some of the most widely used optimization techniques for deriving optimal PA and DPD models. However, Differential Evolution (DE) algorithm [29] was found to be less sensitive to parameter changes, efficient, converges faster and robust than PSO and GA for numerical optimizations over a large and diverse set of numerical benchmark functions [31]. The DE is a population based meta-heuristic algorithm and is regarded as one of the best stochastic optimization method for solving many real-time engineering problems [32]-[44] due to its ability to find globally optimum solution for any complex problems. A survey of the state-of-the-art DE algorithm [45]-[46] reveals the remarkable performance of DE in optimizing multi-dimensional, multi-modal and multi-objective optimization problems. Hence, DE can be regarded as an excellent choice for parameter optimization of the GMP model.

In this paper, an enhanced version of the DE algorithm namely, the modified differential evolution algorithm (MDE) proposed in [47] is used to determine the optimal size of GMP model for the DPD. We also propose a weighted information criteria based cost function for determining the optimal GMP model which enables us to balance both the modeling accuracy as well as model complexity. It is to be noted that, the presented approach is not only restricted to GMP models, but can also be applied to any parametric based PA and DPD models.

This article is organized as follows: Section II discusses the GMP model, the modified differential evolution algorithm and the different objective functions used in the optimization process, Section III presents the experimental results and discussions on the identification of optimal models for DPDs using the proposed approach and finally Section IV marks the conclusion.

II. Optimal GMP Model Selection Using MDE Algorithm And Information Criteria

A. The GMP Model

The GMP model [12], used in this paper for model identification of DPD is an improved version of the widely used memory polynomial model. It is to be noted that, this model contains eight integer parameters representing the non-linearity orders and memory depths of the PA. For the digital pre-distortion (DPD), the widely used indirect learning architecture [48] is adopted, where, the post-inverse model of the PA is first identified using the GMP model as given by

$$\begin{aligned}\tilde{x}(n) &= \sum_{k=0}^{K_x}\sum_{p=0}^{P_x}\alpha_{k,p} z(n-p)|z(n-p)|^k \\ &+ \sum_{k=1}^{K_y}\sum_{p=0}^{P_y}\sum_{q=1}^{Q_y}\beta_{k,p,q} z(n-p)|z(n-p-q)|^k \\ &+ \sum_{k=1}^{K_z}\sum_{p=0}^{P_z}\sum_{q=1}^{Q_z}\gamma_{k,p,q} z(n-p)|z(n-p+q)|^k\end{aligned} \tag{1}$$

where, $\tilde{x}(n)$ is the modeled output of the post-inverse PA model and $z(n) = y(n)/G$ is the normalised output complex envelope of the PA baseband signal, $[K_x,\ P_x]$ represents the nonlinear order and memory depth for the diagonal terms, similar to the memory polynomial (MP) model [11]. Additional parameters used in the GMP model are $[K_y,\ Q_y,\ P_y]$ and $[K_z,\ Q_z,\ P_z]$ which represents the size of lagging envelope terms and leading envelope terms, respectively. The model coefficients of the GMP model $[\alpha_{k,p},\ \beta_{k,p,q},\ \gamma_{k,p,q}]$ can be determined using least squares techniques. The identified post-inverse PA model is then used as the pre-inverse DPD model for linearizing the PA as shown in Fig. 1. As can be seen from (1), the model complexity of the GMP model depends on the eight integer valued parameters described above, resulting in total N_c coefficients given by,

$$N_c = (K_x+1)(P_x+1) + K_y(P_y+1)Q_y + K_z(P_z+1)Q_z \tag{2}$$

Therefore, from (1) and (2), we can conclude that the number of floating point operations (FLOPs) performed for calculating the model output is given by [22],

$$N_f = 10 + 2(K_x+1) + 2K_yQ_y + 2K_zQ_z + 8N_c - 2 \tag{3}$$

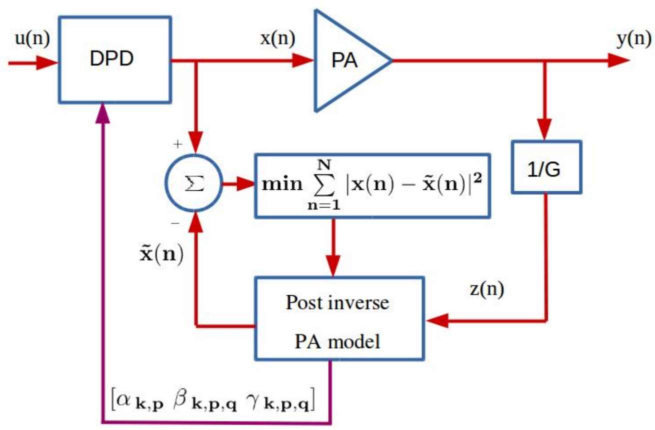

Fig. 1. DPD based on the indirect learning architecture.

From (2) and (3), it can be inferred that, the model complexity increases with the nonlinear order and memory depth of the GMP model. Also, determining the accurate GMP model by sweeping the integer parameters over the entire range could be computationally intensive and can also result in model over-fitting. Hence, to address the problem of identifying optimum model coefficients, we propose an global optimization method namely the Modified Differential Evolution (MDE) [47], which is an improved version of the classical Differential Evolution (DE) [29]. The MDE algorithm determines the optimal parameters of the DPD models by iteratively refining a parent population over J iterations, where each member in j^{th} iteration is a set of model parameters uniformly distributed in the parameter space given by,

$$\bar{\bar{P}}^{(j)} = [\,\bar{p}_1^{(j)}\ \bar{p}_2^{(j)}\ .....\ \bar{p}_{N_p}^{(j)}\,],\ \ j = [\,0\,:\,(J-1)\,] \tag{4}$$

where, N_p is the size of parent population and $\bar{p}_i^{(j)}$, $i = [1 : N_p]$ is described by the eight integer parameters of the GMP model of the DPD as,

$$\bar{p}_i^{(j)} = [\,K_x^{(j)}\ P_x^{(j)}\ K_y^{(j)}\ P_y^{(j)}\ Q_y^{(j)}\ K_z^{(j)}\ P_z^{(j)}\ Q_z^{(j)}\,] \tag{5}$$

When the best parent in terms of the objective function meets the optimization criteria or the number of iterations reaches the maximum specified iterations J, the optimization is terminated. The original DE algorithm uses a mutation strategy to generate a trial child population of same N_p members in each iteration. When the fitness of each child member in terms of the objective

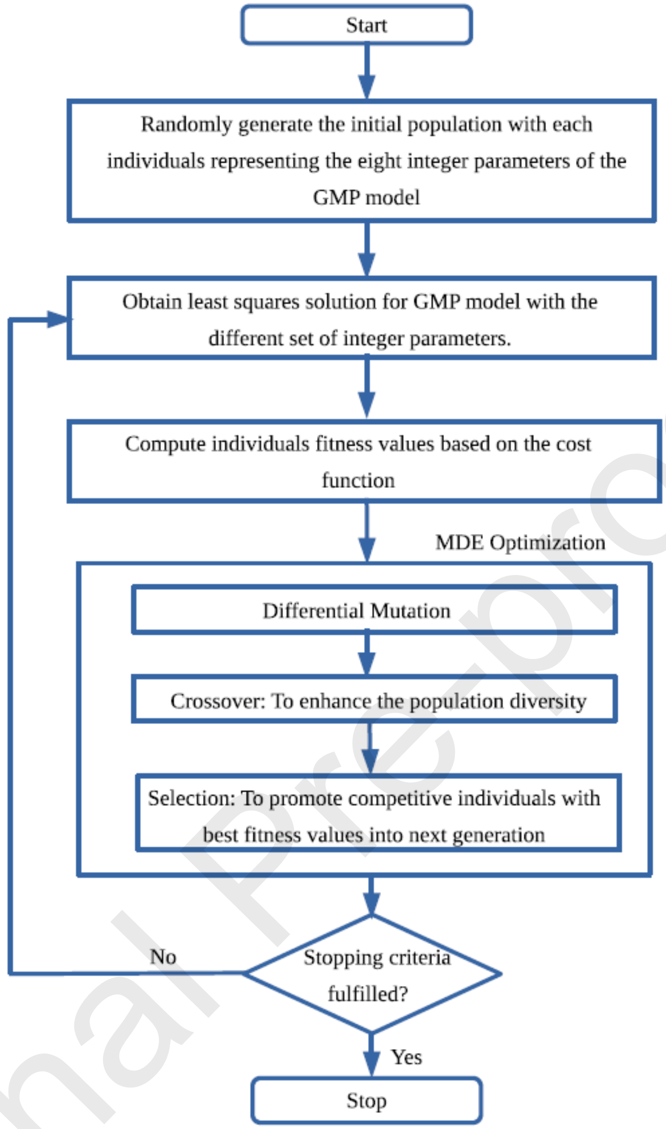

Fig. 2. The work-flow of the proposed method for parameter optimization of the GMP model using MDE algorithm.

function is better than the corresponding parent member in the current iteration, the parent will be replaced by the child member yielding a new parent population for next iteration. Since its introduction, many flavors of the original DE algorithm have been proposed to improve the performance of the basic DE algorithm [46]. The improved DE algorithm namely MDE used in this paper is based on competitive strategies [47], where M different mutation strategies compete in terms of fitness to generate each child member in the child population. Thus, any bias in terms of strategy is removed in the mutation process, whereby faster convergence is achieved. The implementation details of MDE algorithm can be found in [47]. Fig. 2 illustrates the work-flow of the proposed method for parameter optimization of the GMP model using the MDE algorithm.

B. Objective Function for Optimization of GMP Parameters

The objective function defines the search criteria for the parameter optimization of GMP model. During the optimization process, the MDE algorithm searches for the values of the individual model parameters by minimizing the value of the objective function.

The most commonly used objective function for RF PA behavioral modeling or DPD modeling is the normalized mean squared error (NMSE). NMSE is a measure of how good the modeled output resembles the actual output and is given by,

$$\mathrm{NMSE_{dB}} = 10\,\log_{10}\frac{\sum\limits_{n=1}^{N}|x(n)-\tilde{x}(n)|^2}{\sum\limits_{n=1}^{N}|x(n)|^2} \tag{6}$$

where $x(n)$ and $\tilde{x}(n)$ are the PA input and the output of the post-inverse PA model, respectively. However, using NMSE as the objective function in the optimization process leads to more complex models, as there is no constraint on the model complexity. We address this well known problem in the next section using information criteria based cost function.

C. Application of Information Criteria based Cost Function

Application of information criteria (IC) based cost functions enables us to balance the model complexity with model accuracy. The information criteria based cost function accomplishes this goal by the selection of parameters which are most important for the model specification [50]. Akaike Information Criteria (AIC) [51] and Bayesian Information Criteria (BIC) [52] are some of the popular and widely used information criteria tools for model selection. Using information criteria as the objective function in the optimization process helps in reducing the number of model parameters or the model complexity without compromising the modeling accuracy. The scores attained by AIC and BIC serves as the statistical criterion for choosing the best fit GMP model, and is given by,

$$\text{AIC} = N\,ln(\sigma^2) + 2\,N_c \tag{7}$$

$$\text{BIC} = N\,ln(\sigma^2) + ln(N)\,N_c \tag{8}$$

where, N_c is the number of model coefficients, N is the number of samples used for model identification and σ^2 is the variance of the modeling error between the modeled output and the actual output. From (7) and (8) we can conclude that, the information criterion based cost functions penalizes the GMP models that have large number of coefficients and those which fails to accurately model the characteristics by assigning a higher AIC/BIC score. The proposed approach uses a weighted information criterion based cost function which has a form similar to (7) and (8) but with a weighted penalty term and hence offers more flexibility in model selection. The weighted-information criteria, η_{ic} is given by,

$$\eta_{\text{ic}} = \eta\,ln(\sigma^2) + N_c \tag{9}$$

where, η is a weighting factor. The weighting factor η can be used to control the number of model coefficients for achieving trade-off between the modeling accuracy and complexity. From (9), we can infer that, higher the value of η, better will be the modeling accuracy. The following section presents the results of the proposed approach in determining optimal DPD model for PA linearization and experimental results.

III. RESULTS

A. Measurement Setup

The proposed MDE optimization based model identification for the DPD, is validated for its modeling accuracy using the RF WebLab measurement setup provided by the Chalmers University and National Instruments [55]. The RF WebLab is a remotely accessible PA measurement system which incorporates a Vector Signal Transceiver (PXIe-5646R) and a nonlinear 6W GaN PA (Cree CGH40006-TB) supplied by a DC power source module, PXI-4130. The Vector Signal Transceiver incorporates a Vector Signal Generator (VSG) as well as a Vector Signal Analyser (VSA) into a single PXI Express module and has an instantaneous bandwidth of 200 MHz. The RF weblab measurement setup used for PA characterization and DPD validation is shown in

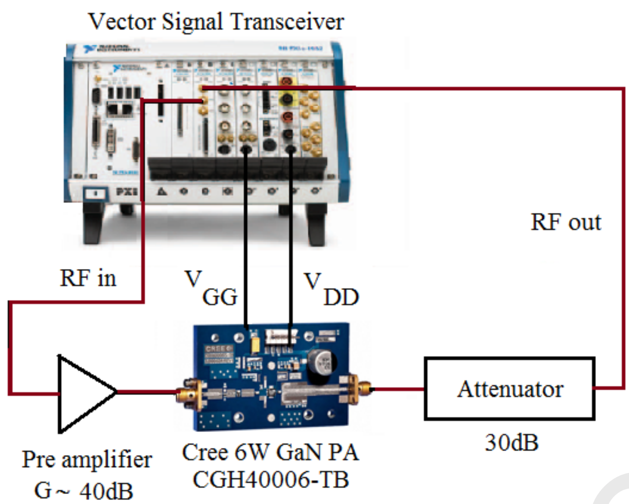

Fig. 3. RF Weblab Measurement Setup

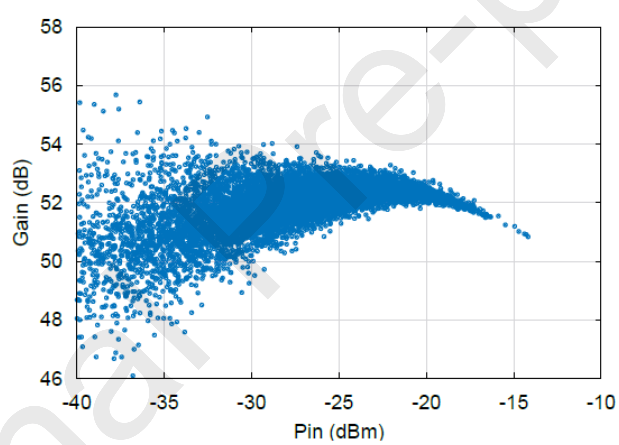

Fig. 4. AM/AM distortion characteristics of the measured PA

Fig. 3. A more detailed description of the experimental setup and the access details are given in [56].

The GaN PA under test is excited by an input 2.14 GHz carrier modulated by baseband signal with randomly generated in-phase and quadrature components, band-limited to 20 MHz. The input signal is generated using the Keysight Advanced Design System (ADS) tool and then send to the RF Weblab for PA measurement. The input signal has an rms power of -22 dBm and a PAPR of 10.8 dB. It is to be noted that, the signal generated by the Vector Signal Transceiver

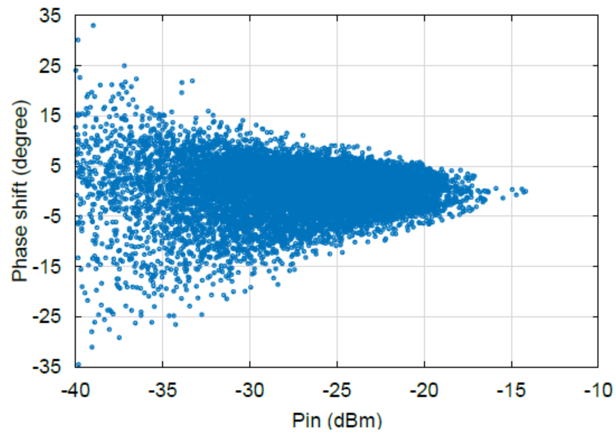

Fig. 5. AM/PM distortion characteristics of the measured PA

(VST) is amplified by a linear driver amplifier having a gain of $\approx 40dB$ before it is amplified by the GaN PA. The measured rms output power of the signal is 30.25 dBm and the peak power of the output is 36.59 dBm. The measured AM/AM and AM/PM characteristics of the PA under test are shown in fig. 4 and fig. 5, respectively. From the figures we can conclude that, the PA exhibits significant distortions in the low power region due to the memory effects of the PA and in the high power region due to the nonlinearity of PA.

B. Modeling Performance of DPD Models

In this section, the proposed modified differential evolution algorithm is applied to determine the optimal GMP model for the DPD. The accurate time delay between the baseband input and output signal is estimated and suitably compensated using the cross correlation technique. The time aligned baseband signals are then used to derive the optimal DPD models. The behavioral modeling and the identification of optimal DPD models using MDE is carried out using the C++ based open source library, ITPP [57]. For extracting the DPD models, 12000 (12 K) samples of the baseband input and output signals are used. The models are validated using another 12000 (12 K) samples taken from a different time interval.

The optimization variables for the MDE algorithm are the eight integer parameters of the GMP model given in (1). The performance of the MDE algorithm depends on the mutation strategy adopted, the population size N_P and the associated control parameters used in the optimization

TABLE I. MDE SETTINGS

Parameters	***Value***
Dimension, D	8
Population size, N_P	20
Population Initialization	Random
Weighting Factor, F	0.5
Crossover Rate, p_{cr}	0.7
Maximum number of iterations, J	20

process [47]. Therefore, appropriate choice of the optimization parameters are crucial to avoid premature convergence and stagnation at local optimum. In order to balance the reliability and speed of the optimization process in determining the global optimum, a population size of 20 was chosen. The parameters of the MDE algorithm used for optimizing the GMP model are shown in Table I. Fine tuning of control parameters such as the weighting factor F and the crossover rate p_{cr} are crucial for the algorithm to obtain the desired solutions. The weighting factor F plays a major role in controlling the accuracy and exploration capability of the algorithm. In the proposed MDE algorithm, F is made adaptive which increases the accuracy and exploration capability in the search process. In [29], it is said that, F = 0.5 or 0.6 would be a good initial choice. Hence, in the proposed work the initial value of F is chosen as 0.5. As the generation evolves, the value of F adapts itself to determine the global optimum from the search space. Similarly, a good initial choice for the the crossover rate p_{cr}, which represents the probability of mixing between the trial and target vectors is said to be between 0.3 and 0.9 [45]. Hence, in the proposed work the crossover rate is chosen as 0.7. To ensure consistency, the optimization process for the GMP model extraction was repeated for 30 independent trials and the best solution is selected.

The parameters which contribute to the computational complexity of the algorithm are: the population size N_P, number of generations J, number of competitive strategies M used in the MDE and the computational complexity of the fitness function to be optimized. The overall complexity of the algorithm alone is at the most $O(N_P\, J\, M)$ excluding the complexity of the fitness function evaluation which depends on the application domain. However, in practical cases we incorporate break points while running the algorithm, for example, if the RMSE of the

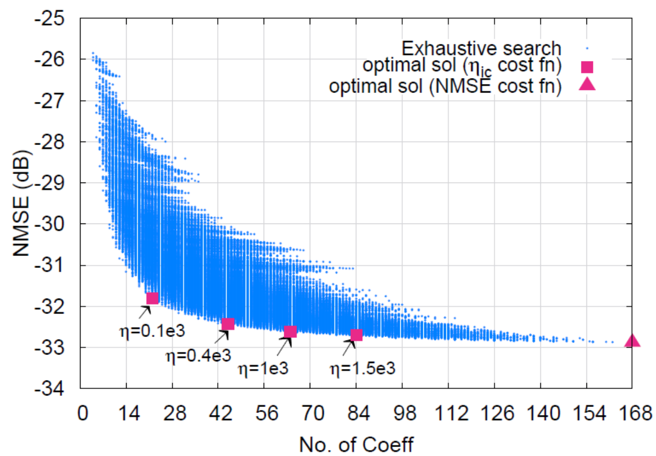

Fig. 6. All possible solutions of the GMP model using exhaustive search vs. the best optimal solution obtained using MDE optimization with cost functions based on NMSE and the proposed η_{ic} with different η values.

TABLE II. DESIGN VARIABLES OF THE GMP MODEL AND ITS CORRESPONDING SEARCH SPACE

Integer Parameters of GMP model								**Total number of possible**
K_x	Q_x	K_y	P_y	Q_y	K_z	P_z	Q_z	**combinatorial models**
$[0,7]$	$[0,5]$	$[1,5]$	$[0,3]$	$[1,3]$	$[1,5]$	$[0,3]$	$[1,3]$	172800

optimization function reaches a certain value, the algorithm will stop the computation. Hence, in practical applications the computational complexity is much less than $O(N_P\, J\, M)$. However, the most important advantage of the algorithm is its capability of finding combination of parameters which spans a large parametric space with multiple global minima and maxima with excellent accuracy. The large execution time of the algorithm is not at all a problem, due to the fact that, the non-linear system identification of the PA and DPD is not performed in real time. In addition, only the accuracy of the system identification is paramount for digital per-distorter design. Once the parameters of the generalised memory polynomial is identified using MDE, the model can be programmed in embedded hardware such as FPGA.

The design variables of the GMP model and its corresponding search space are tabulated in Table II. An exhaustive search technique was initially used to determine the best GMP model

TABLE III. PERFORMANCE COMPARISON OF BASIC DE STRATEGIES AND MDE WITH NMSE BASED COST FUNCTION FOR OPTIMAL DPD MODEL IDENTIFICATION

DE	**NMSE(dB)**			
Strategies	**Best**	**Worst**	**Mean**	**Std. dev.**
DE/rand [29]	-32.852	-32.622	-32.754	0.00824
DE/best [30]	-32.867	-32.579	-32.753	0.00614
MDE	$\mathbf{-32.883}$	-32.801	$\mathbf{-32.854}$	0.01977

TABLE IV. COMPARISON OF MDE PERFORMANCE WITH THE DIFFERENT OBJECTIVE FUNCTIONS

Objective Function	**NMSE(dB)**	**No. of model coeff, N_c**	**No. of FLOPs, N_f**	**Computational Time (Mins)**
NMSE	-32.883	168	1428	4.23
AIC	-32.856	144	1228	3.89
BIC	-32.815	114	1214	3.46
η_{ic} ($\eta = 1.5e3$)	-32.698	84	720	2.62

for the candidate space defined in Table II. It was found that, the GMP model with an NMSE of -32.883 dB has the best modeling accuracy using the exhaustive search process. It took $\approx$ 6 hours to determine the solution using the regular exhaustive search technique. On the other hand, MDE optimization technique enables us to significantly reduce the computational time to $3 - 5$ minutes, depending on the cost function, see Table IV. The computational time is calculated based on the model extraction process performed on an Intel CORE i7-4760 CPU having a clock speed of 3.6 GHz. The performance of MDE is first compared with the basic DE strategies widely used in the literature such as "DE/rand" [29] and "DE/best" [30]. The optimal DPD models obtained with the basic DE strategies and the MDE algorithm for the NMSE based cost function are summarized in Table III. It is to be noted that, the mean NMSE is obtained by averaging the results over 30 trials. It can be inferred from Table III that, compared to the basic DE strategies, MDE determines the best DPD model with the minimum NMSE of -32.883 dB. This is due to the fact that, MDE uses a competitive strategy, where M different mutation strategies compete in terms of fitness in each generation to generate the solution. Due to the superior performance of MDE in determining optimal DPD model, we apply MDE to determine the optimal DPD model with the other information criteria based cost functions.

The total number of possible combinatorial models are 1,72,800 for the search space defined in Table II. Fig. 6 depicts all the possible combinatorial solutions of the GMP model obtained using the exhaustive search technique and the final best optimal solution obtained after the MDE optimization process with NMSE as cost function and the proposed η_{ic} with different η values as the cost function. From Fig. 6, we can infer that, when NMSE based cost function is used in the optimization process, the MDE chooses the GMP model having the lowest NMSE of -32.883 dB as the optimal solution, without considering the model complexity in the model selection process. The optimal model chosen using NMSE based cost function has 168 model coefficients. On the other hand, when the proposed η_{ic} based cost function is used in the optimization process, the MDE determines the GMP model which balances model accuracy with complexity. The free parameter in the η_{ic} cost function, namely the weighting factor η has a significant influence in extracting complexity reduced DPD models. It can be inferred from Fig. 6 that, η can be increased in fixed steps, until the optimal DPD model achieves a balance between model complexity and goodness-of-fit. From Fig. 6, we can conclude that, for an $\eta = 1.5e3$, the proposed η_{ic} based cost function, gives almost similar modeling performance with an NMSE of -32.698 dB but with only half the number of coefficients compared to when NMSE is used as the cost function. Hence, the proposed η_{ic} cost function helps in avoiding model over fitting in the model selection process. However, the limitation of the proposed method is that, the weighting factor η has to be tuned, until the extracted DPD model achieves a balance between complexity and accuracy.

In order to ensure fair comparison between the methods, the standard performance metric NMSE, widely used in the literature is used to assess the performance of the DPD models. Table IV compares the performance of the MDE algorithm in terms of the modeling accuracy (NMSE) [14]-[16], the model complexity and the computational time taken for determining the optimal solution with different cost functions in the optimization process. It can be inferred from Table IV that, compared to the other cost functions such as NMSE, AIC and BIC, the proposed weighted information criteria based cost function helps in identifying complexity reduced models in the model order selection process. This is because, the standard least squares technique [11]-[16] used for extracting the coefficients of DPD models requires larger number of training data

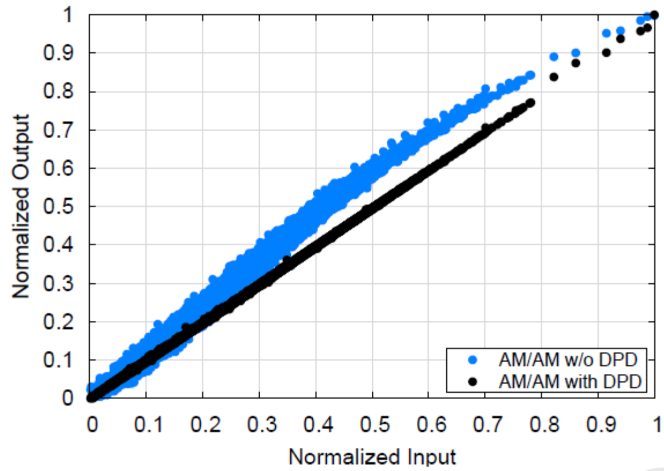

Fig. 7. Linearized AM/AM characteristics of the measured PA

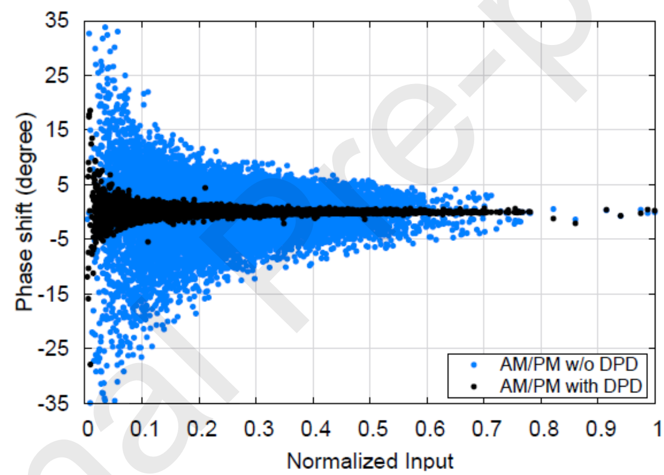

Fig. 8. Linearized AM/PM characteristics of the measured PA

samples in the model extraction process, so that, the model can predict the behavior reasonably well for the samples outside the training data. In the proposed work, 12000 (12K) baseband samples are used for extracting the DPD models using least squares technique. However, for larger number of data samples N, when AIC is used as a model order selection tool, it tends to overestimate the model order due to the factor $N\,ln(\sigma^2)$ in the first term of (7). The probability of overestimating the model order, increases with increase in the sample size [53]. Asymptotically, as the sample size $N \to \infty$, AIC considers only the model accuracy in the model selection process and does not consider its complexity [54]. From Table IV it can be inferred that, the AIC based cost function chooses a model, which is only 14.3% less complex than the one chosen by the NMSE based cost function. Therefore, AIC as an objective function is not an ideal choice for extracting complexity reduced DPD models for a larger sample size N. On the other hand, the tendency of BIC based cost function to overestimate the model order is comparatively less than AIC due to the factor $ln(N)$ in the second term of (8). For models having similar accuracy and different model complexity, the factor $ln(N)$ when multiplied with the number of coefficients N_c increases the information criteria score of the more complex models. Subsequently, the more complex models will not be selected in the optimization process, where the model order selection is formulated as a cost minimization problem. However, the model selection can still be improved by using a weighting factor in the cost function instead of the fixed sample size, N. Hence, the proposed work uses a weighting factor η in the cost function, which can be tuned in the model selection process. This ensures that, over modelling is avoided and helps in deriving optimal DPD models which balances model complexity and accuracy. From Table IV, we can infer that, compared to the other cost functions such as NMSE, AIC and BIC, the proposed η_{ic} based cost function in the optimization process performs best as it substantially reduces the model complexity without significantly compromising the model accuracy. It is to be noted that, compromise in NMSE performance of approximately 0.2dB for the derived DPD model enables us to reduce the number of model coefficients from 168 to 84, which is significant. Also, there is a significant reduction in the number of FLOPs from 1428 to 720, when η_{ic} based cost function is used. Hence, the proposed weighted information criteria based cost function and MDE optimization enables us to derive optimum DPD models balancing model complexity and model accuracy. In addition, it can also be inferred from Table IV that, deriving optimal GMP model using the proposed approach requires only a fraction of the time as required for determining the best model using exhaustive search technique.

C. Linearization performance

The performance of the derived DPD model in linearizing the PA behavior is investigated in this section. The eight integer parameters of the GMP model determined by the MDE algorithm with $\eta_{ic}\,(\eta = 1.5e3)$ based cost function are used in the DPD model for linearizing the PA. The AM/AM and the AM/PM characteristics of the measured PA and the linearized PA using the optimal DPD model are shown in Fig. 7 and 8, respectively. From the figures we can infer that, both the memory effects and the non-linearities due to PA imperfections are corrected by the derived optimal DPD model. The normalized power spectral density (PSD) of the PA without and with linearization using the DPD models derived using different cost functions in the optimization process are shown in Fig. 9. It is apparent from Fig. 9 that, the DPD model obtained using the proposed η_{ic} cost function achieves a similar linearization performance, while requiring only 50% and 41.6% of the coefficients, as compared to the DPD models obtained using NMSE and AIC based cost functions, respectively. Hence, we can conclude that, the optimal DPD model obtained using the proposed method reduces the spectral regrowth in the sidebands and achieves excellent linearization performance with reduced number of coefficients, compared to NMSE and AIC based cost functions, respectively.

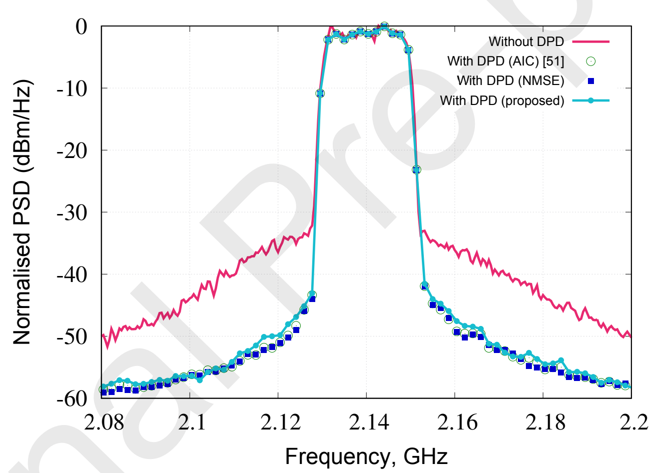

Fig. 9. Normalized Power Spectral Density of the PA without and with linearization using the DPD models obtained with different cost functions in the optimization process.

IV. Conclusion

The digital predistortion of radio frequency power amplifiers (RF PA) using the modified differential evolution (MDE) algorithm, introduced in this paper, enabled us to derive optimum DPD models. It was found that, MDE based optimization technique for determining the DPD models of the PA, saves lot of computer resources and computational time. Further, when we incorporate the weighted information criteria as objective function in the optimization, accurate

DPD models which balances model accuracy and model complexity can be achieved. Hence, the proposed search criteria and MDE based parameter optimization of the GMP model enables an expedited method for deriving optimal DPD models.

REFERENCES

[1] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, Shu Sun, and Shuai Nie, "Millimeter-Wave Enhanced Local Area Systems: A High-Data-Rate Approach for Future Wireless Networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152-1163, Jun. 2014.

[2] S. Buzzi, Chih-Lin I, T. E. Klein, H. Vincent Poor, Chenyang Yang and A. Zappone, "A Survey of Energy-Efficient Techniques for 5G Networks and Challenges Ahead," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 697-709, Apr. 2016.

[3] H. Bogucka, A. Conti, "Degrees of Freedom for Energy Savings in Practical Adaptive Wireless Systems," *IEEE Commun. Mag.*, pp. 38-44, Jun. 2011.

[4] Jingon Joung, Chin Keong Ho and Sumei Sun, "Spectral Efficiency and Energy Efficiency of OFDM Systems: Impact of Power Amplifiers and Countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 208-220, Feb. 2014.

[5] Jingon Joung, Chin Keong Ho, Koichi Adachi and Sumei Sun, "A Survey on Power-Amplifier-Centric Techniques for Spectrum and Energy Efficient Wireless Communications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 315-333, First quarter 2015.

[6] S. L. Brunton, J. L. Proctor and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Nat. Acad. Sci.*, vol. 113, no. 15, pp. 3932-3937, 2016.

[7] N. M. Mangan, S. L. Brunton, J. L. Proctor and J. N. Kutz "Inferring biological networks by sparse identification of nonlinear dynamics," *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, vol. 2, no. 1, pp. 52-63, Jun. 2016.

[8] Neethu Mohan, K. P. Soman, S. Sachin Kumar, "A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model," *Applied Energy*, vol. 232, pp. 229-244, 2018.

[9] Lei Guan and Anding Zhu, "Green Communications," *IEEE Microw. Mag.*, vol. 15, no.7, pp. 84-99, Nov./Dec. 2014.

[10] F. M. Ghannouchi and O. Hammi, "Behavioral Modeling and Predistortion," *IEEE Microw. Mag.*, vol. 10, no.7, pp. 52-64, Dec. 2009.

[11] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159-165, Jan. 2004.

[12] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt and J. Pastalan, "A generalized memory polynomial for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852-3860, Oct. 2006.

[13] A. Zhu, J. C. Pedro and T. J. Brazil, "Dynamic deviation reduction-based Volterra behavioral modeling of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.,* vol. 54, no. 12, pp. 4323-4332, Dec. 2006.

[14] M. Younes, O. Hammi, A. Kwan and F. M. Ghannouchi, "An accurate complexity-reduced 'PLUME' model for behavioral modeling and digital predistortion of RF power amplifiers," *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1397-1405, Apr. 2011.

[15] Y. J. Liu, J. Zhou, W. Chen and B. H. Zhou "A robust augmented complexity-reduced generalized memory Polynomial for wideband RF power amplifiers," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2389-2401, May. 2014.

[16] A. E. Abdelrahman, O. Hammi, Andrew K. Kwan, A. Zerguine, F. M. Ghannouchi, “A novel weighted memory Polynomial for Behavioral Modeling and Digital Predistortion of Nonlinear Wireless Transmitters,” *IEEE Trans. Ind. Electron.*, vol. 63, no. 3, pp. 1745-1753, Mar. 2016.

[17] M. Isaksson, D. Wisell and D. Rönnow, “Wide-band dynamic modeling of power amplifiers using Radial-Basis function neural networks,” *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 11, Nov. 2005.

[18] Meenakshi Rawat and F. M. Ghannouchi, “Distributed Spatiotemporal neural network for nonlinear dynamic transmitter modeling and adaptive digital predistortion,” *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, Mar. 2012.

[19] Ana Anastasijevica, D. Cojaa, N. Neskovica, A. Neskovica, D. Budimir, “Joint power amplifier and I/Q modulator impairments modelling and compensation for LTE transmitters using artificial neural networks,” *Int. J. Electron. Commun. (AEÜ),* 69 (2015) pp. 529-538.

[20] R. V. S. Devi and D. G. Kurup, “Behavioral Modeling of RF Power Amplifiers for Designing Energy Efficient Wireless Systems,” *Proc. Int. Conf. Wireless Commun. Signal Process. Netw. (WISPNET)* pp.1994-1998, Mar. 2017.

[21] J Cai, Chao Y, L Sun, S Chen and J B King, “Dynamic behavioral modeling of RF Power Amplifier based on time-delay support vector regression,” *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 2, pp. 533-543, Feb. 2019.

[22] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson and C. Fager, “A Comparative Analysis of the Complexity/Accuracy Tradeoff in Power Amplifier Behavioral Models,” *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 6, pp. 1510-1520, Jun. 2010.

[23] Gholamreza Karimi and Ali Lotfi, “ An analog/digital predistorter using particle swarm optimization for RF power amplifiers,” *Int. J. Electron. Commun. (AEÜ),* 67 (2013), pp. 723-728.

[24] A. Abdelhafiz, O. Hammi, A. Zerguine, A. Al-Awami and F. M. Ghannouchi, “A PSO based memory polynomial predistorter with embedded dimension estimation,” *IEEE Trans. Broadcast.*, vol. 59, no. 4, pp. 665-673, Dec. 2013.

[25] A. Abdelhafiz, L. Behjat and F.M. Ghannouchi, “Generalized memory polynomial model dimension selection using Particle Swarm Optimization,” *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 2, pp. 96 - 98, 2018.

[26] S. R. Bhavanam, Sanjika Devi R V, Sriram M and D. G. Kurup, “Information criteria based optimal structure identification of RF power amplifier models,” *Journal of Intelligent and Fuzzy Systems,* 36, pp. 2137-2145, 2019.

[27] Siqi Wang, Mazen Abi Hussein, Olivier Venard and Genevieve Baudoin, “A novel algorithm for determining the structure of digital predistortion models”, *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7326-7340, Aug. 2018.

[28] C. Ciflikli and A. C. Yapici, “Genetic algorithm optimization of a hybrid analog/digital predistorter for RF power amplifiers,” *Analog Integrated Circuits and Signal Processing,* vol. 52, no. 1, pp. 25-30, 2007.

[29] R. Storn and K. Price, “Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces,” *J. Global optim.,* vol. 11, no. 4, pp. 341-359, 1997.

[30] R. Storn and K. Price, “Minimizing the real functions of the ICEC’96 contest by differential evolution,” *Proceedings of IEEE International Conference on Evolutionary Computation,* pp. 842-844, Nagoya, Japan, 1996.

[31] Vesterstrom J, Thomsen R, “A comparative study of differential evolution, particle swarm optimization and evolutionary algorithms for numerical benchmark problems. In:Evolutionary Computation, 2004.” *CEC2004. Congress on. IEEE,* vol. 2, pp. 1980-1987, 2004.

[32] Kotha Srinivasa Reddy and Subhendu Kumar Sahoo, “An approach for FIR filter coefficient optimization using differential evolution algorithm,” *Int. J. Electron. Commun. (AEÜ),* 69 (2015), pp. 101-108.

[33] D. Mandal, A. Chatterjee, T. Mondal and A. K. Bhattacharjee, "Shaped beam synthesis of concentric ring array antenna using differential evolution algorithm," *Int. J. Electron. Commun. (AEÜ),* 69 (2015), pp. 1025-1031.

[34] Selcuk Mete, Saban Ozer and Hasan Zorlu, "System identification using Hammerstein model optimized with differential evolution algorithm," *Int. J. Electron. Commun. (AEÜ),* 70 (2016), pp. 1667-1675.

[35] Hasan Zorlu, "Optimization of weighted myriad filters with differential evolution algorithm," *Int. J. Electron. Commun. (AEÜ),* 77 (2017), pp. 1-9.

[36] Yuan Mao, S. Niu and Yun Yang, "Differential evolution-based multiobjective optimization of the electrical continuously variable transmission system," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2080-2089, Mar. 2018.

[37] Burhanettin Durmus, "Optimal components selection for active filter design with average differential evolution algorithm," *Int. J. Electron. Commun. (AEÜ),* 94 (2018), pp. 293-302.

[38] Jie Hu, Min Wu, Xin Chen, Sheng Du, Pan Zhang, Weihua Cao and Jinhua She, "A multilevel prediction model of carbon efficiency based on the differential evolution algorithm for the iron ore sintering process," *IEEE Trans. Ind. Electron.*, vol. 65, no. 11, pp. 8778-8787, Nov. 2018.

[39] W. Zhao, A. Ma, J. Ji, X. Chen and T. Yao, "Multiobjective optimization of a double-side linear vernier PM motor using response surface method and Differential Evolution," *IEEE Trans. Ind. Electron.*, vol. 67, no. 1, pp. 80-90, Jan. 2020.

[40] V. Gonzalez, C. A. Monje, S. Garrido, L. Moreno and C. Balaguer, "Coverage mission for UAVs using differential evolution and fast marching square methods," *IEEE Aerospace and Electronic Systems Magazine,* vol. 35, no. 2, pp. 18-29, Feb. 2020.

[41] L. A. C. Najarro, I. Song and K. Kim, "Differential evolution with opposition and redirection for source localization using RSS measurements in wireless sensor networks," *IEEE Trans. Autom. Sci. Eng.*, Mar. 2020, http://dx.doi.org/10.1109/TASE.2020.2975287

[42] W. Gong, Y. Wang, Z. Cai and L. Wang, "Finding multiple roots of nonlinear equation systems via a repulsion-based adaptive differential evolution," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 4, pp. 1499-1513, Apr. 2020.

[43] L. Liu, C. Guo, Y. Tu, H. Mei and L. Wang, "Differential evolution fitting-based optical step-phase thermography for micrometer thickness measurement of atmospheric corrosion layer," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5213-5222, Aug. 2020.

[44] Judhisthir Dash, Bivas Dam and Rajkishore Swain, "Design and implementation of sharp edge FIR filters using hybrid differential evolution particle swarm optimization," *Int. J. Electron. Commun. (AEÜ),* 114 (2020).

[45] S. Das and P. N. Suganthan, "Differential evolution: a survey of the state-of-the-art," *IEEE Trans. Evol. Comput.*, vol. 15, no. 1, pp. 4-31, 2011.

[46] S. Das, S. S. Mullick and P. N. Suganthan, "Recent advances in differential evolution – An updated survey," *Swarm Evol. Comput.*, vol. 27, pp. 1-30, 2016.

[47] D. G. Kurup, M. Himdi and A. Rydberg, "Synthesis of uniform amplitude unequally spaced antenna arrays using the differential evolution algorithm," *IEEE Trans. Antennas Propag.*, vol. 51, no. 9, pp. 2210-2215, 2003.

[48] C. Eun and E. J. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 223-227, 1997.

[49] R. Mallipeddi, P. N. Suganthan, Q. K. Pan and M. F. Tasgetiren, "Differential evolution algorithm with ensemble of parameters and mutation strategies," *Applied Soft Computing,* vol. 11, pp. 1679-1696, 2011.

[50] P. Stoica and Y. Selen, “Model-order selection: A review of Information criterion rules,” *IEEE Signal Process. Mag.,* vol. 21, no. 4 pp. 36-47, 2004.

[51] H. Bozdogan, “Akaike’s information criterion and recent developments in information complexity,” *Journal of Mathematical Psychology,* vol. 44, no. 1, pp. 62-91, 2000.

[52] G. Schwarz, “Estimating the dimension of a model,” *The Annals of Statistics,* vol. 6, no. 2, pp. 461-464, 1973.

[53] A. P. Liavas and P. A. Regalia, “On the behavior of information theoretic criteria for model order selection,” *IEEE Trans. Signal Process.,* vol. 49, no. 8, pp. 1689-1695, 2001.

[54] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust. Speech Signal Process.,* vol. 33, pp. 387-392, 1985.

[55] P. N. Landin, S. Gustafsson, C. Fager and T. Eriksson “WebLab: A Web-based setup for PA digital predistortion and characterization [application notes],” *IEEE Microw. Mag.*, vol. 16, no. 1 pp. 138-140, Feb. 2015.

[56] (Sep. 2019). *Weblab Measurement Setup*. [Online]. Available: http://dpdcompetition.com/rfweblab/measurement-setup/

[57] C++ Signal Processing and Communication Library (ITPP). [Online]. Available: http://itpp.sourceforge.net/4.3.1/

Declaration of interests -None