Comparison between Direct and Indirect Learnings for the Digital Pre-distortion of Concurrent Dual-band Power Amplifiers

Luis Schuartz
Group of Integrated Circuits and Systems (GICS)
Federal University of Paraná
Curitiba, Paraná
luisschuartz@ufpr.br

Artur T. Hara
Group of Integrated Circuits and Systems (GICS)
Federal University of Paraná
Curitiba, Paraná
arturhara@gmail.com

André A. Mariano
Group of Integrated Circuits and Systems (GICS)
Federal University of Paraná
Curitiba, Paraná
mariano@ufpr.br

Bernardo Leite
Group of Integrated Circuits and Systems (GICS)
Federal University of Paraná
Curitiba, Paraná
leite@eletrica.ufpr.br

Eduardo G. Lima
Group of Integrated Circuits and Systems (GICS)
Federal University of Paraná
Curitiba, Paraná
elima@eletrica.ufpr.br

ABSTRACT

Current radio-communication systems demand high linearity and high efficiency. The digital baseband pre-distorter (DPD) is a cost-effective solution to guarantee the required linearity without compromising the efficiency. In the design of a DPD for a single band power amplifier (PA), the position of the inverse system is exchanged during the identification procedure to avoid the necessity of a PA model within a cumbersome closed-loop process. However, in a practical environment where only an approximation to the inverse is achieved, the linearization capability is affected by shifting the post-inverse placed after the PA to a pre-inverse located before the PA. In DPD intended for concurrent dual-band PAs, an additional advantage of such approach is that the post-inverse identifications for each band are completely independent of each other. This work performs a comparative analysis between two learning architectures applied to the linearization of two concurrent dual-band PAs stimulated by 2.4 GHz Wi-Fi and 3.5 GHz LTE signals. For the first PA, an exact PA model is known and the replacement of a post-inverse to a pre-inverse produces only negligible degradation in linearity. For the second PA, only an approximate PA model is available and the accuracy of such PA model produces a major impact on the linearization capability than the shifting of the inverse.

SBCCI '19, August 26–30, 2019, Sao Paulo, Brazil

ACM ISBN 978-1-4503-6844-5/19/08…$15.00
https://doi.org/10.1145/3338852.3339842

KEYWORDS

Concurrent dual-band, Digital pre-distortion, Efficiency, Linearity, Power amplifier.

ACM Reference Format:
Luis Schuartz, Artur T. Hara, André A. Mariano, Bernardo Leite, and Eduardo G. Lima. 2019. Comparison between Direct and Indirect Learnings for the Digital Pre-distortion of Concurrent Dual-band Power Amplifiers. In *32nd Symposium on Integrated Circuits and Systems Design (SBCCI '19), August 26–30, 2019, Sao Paulo, Brazil.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3338852.3339842

1 INTRODUCTION

In current radio communication systems, high demand for data rate in a scarce spectrum imposes high linearity of transmitters [14]. In addition, the multitude of wireless applications as Wi-Fi and LTE puts toward communication systems that operate in broadband and concurrent multi-mode operations [10].

By one hand, vast possibilities of communications impose a wide-band operation and high linearity. By other hand, communication systems need to be efficient [5, 6, 10, 14]. In the transmitter system, the power amplifier (PA) is the highest power consumer. However, PA is more efficient in the highly nonlinear region [4], which generates a trade-off between linearity and efficiency. In current communications, linearity is mandatory. Consequently, a linearizer is indispensable. The inclusion of a digital baseband pre-distorter (DPD) is the most adequate and low-cost technique that improves PA efficiency without compromising the linearity. The DPD applies the PA inverse characteristic [4, 6, 13, 14].

Let F and G represent two systems. If G is exactly the inverse system of F, e.g. $G = F^{-1}$, then the resultant system from their cascade connection is a system in which the output and input signals are the same, independently of the specific locations of F and G in the cascade connection. However, if G is only approximately the inverse system of F, the particular position of each block in the cascade connection will have an influence in the resultant system.

The DPD parameter can be performed by a direct learning of the pre-inverse or by an indirect learning in which the parameters from a post-inverse are first identified and then copied to a pre-inverse of same topology. As a consequence, the accuracy of indirect learning can be compromised because the disposition of the blocks in the cascade connection is exchanged during training and linearization stages. Previous works have investigated the possible degradation in linearization capability by adopting a training procedure which exchanges the positions of the PA and DPD blocks. However, the investigation was limited to single band PAs [2, 3, 6, 12]. The main contribution of this work is to extend such investigation for concurrent dual-band PAs, motivated by the larger set of phenomena responsible for the generation of distortions in these circuits. In a 2D feed-forward model the outputs are individually formulated as functions of both input signals. Because of the absence of feedback, the parameter identification of the last block in the cascade connection, which is the post-inverse in the indirect learning, can be done separately by each band. Contrarily, the parameter identification of the first block in the cascade connection, which is the pre-inverse in the direct learning, must compulsorily include both bands at a single step, even in presence of a 2D feed-forward model. Such dependency between bands occurs because each band will affect the PA output at both bands. In other words, this work aims to investigate the degradation in linearity capability by performing the identification of two separated post-inverses, one for each band, instead of the unique identification of two related pre-inverses. The latter is a situation that has no counterpart in the single band case.

This paper is divided as follows: Section 2 resumes DPD concepts, Section 3 details direct and indirect learnings; Section 4 reports simulation results and Section 5 summarizes the conclusions.

2 DIGITAL BASEBAND PRE-DISTORTION CONCEPTS

The PA amplifies a radio-frequency signal, at GHz range for the systems considered in this work, expressed by a time-varying amplitude multiplied by the cosine of a phase composed by a random constant and a time-varying component. The time-varying amplitude and phase compose the complex baseband envelope signal at MHz range. PA modifies the envelope signal and can be expressed by a baseband equivalent behavioral model [11]. The DPD implements an inverse PA characteristic modifying the envelope signal in the baseband.

In what concerns the PA behavioral model, Volterra-based series have as main advantages the accuracy and linearity in their parameters. The polynomial model with memory is one example of such series [7, 8]. For concurrent dual-band PA, DPD could apply completely independent inverse characteristics for each channel or just one inverse characteristic to linearize the two bands simultaneously. The first approach has the performance compromised because interferences between bands are not considered and the second one is impractical because the required sampling frequency is too high. Thus, a two-dimensional model, in which each band is individually modeled and interferences between channels are considered, is mandatory to correctly linearize concurrent dual-band PAs [6]. The two-dimensional improved memory polynomial (2D-IMP) complies with such requirements. Been $x_1(n)$ and $x_2(n)$ the input samples of channels 1 and 2 measured at instant n, respectively, the output samples $y_1(n)$ and $y_2(n)$ of channels 1 and 2, respectively, are given by:

$$y_1(n) = \sum_{p=0}^{P-1}\sum_{m=0}^{M} h_{m,p}^{(1)} x_1(n-m)|x_1(n-m)|^p$$
$$\sum_{q=0}^{Q-1}\sum_{p=1}^{P-1}\sum_{m=0}^{M} h_{q,m,p}^{(1)} x_1(n-m)|x_1(n-m)|^q|x_2(n-m)|^p, \quad (1)$$

and,

$$y_2(n) = \sum_{p=0}^{P-1}\sum_{m=0}^{M} h_{m,p}^{(2)} x_2(n-m)|x_2(n-m)|^p$$
$$\sum_{q=0}^{Q-1}\sum_{p=1}^{P-1}\sum_{m=0}^{M} h_{q,m,p}^{(2)} x_2(n-m)|x_2(n-m)|^q|x_1(n-m)|^p, \quad (2)$$

where P and Q are the polynomial truncation orders, M is the memory length and h the coefficient. If Q is set equal to P, then the 2D-IMP is exactly equal to the 2D memory polynomial of [1]. Observe that each output from (1) or (2) can be rewritten as:

$$y(n) = X(n)H, \quad (3)$$

where $X(n)$ is the vector of inputs at present n and past $n-m$ (with m ranging from 1 to M) sample times, and H the vector formed by the coefficients h. The model is linear in its parameters and the least squares (LS) method can be applied to its parameter extraction [8].

3 DIRECT AND INDIRECT LEARNING ARCHITECTURES

When (1) and (2) are employed for pre-distortion purposes, to obtain the optimal values for the coefficient vector H, direct and indirect learning architectures are available.

3.1 Indirect learning architecture

Indirect learning is an architecture extracted from [12] that indirectly computes the DPD parameters by copying the coefficients previously extracted from a post-distorter (PoD) of the same topology. Fig. 1 shows the baseband equivalent of indirect learning architecture [3, 12]. In the indirect learning, the PoD input and output signals are known. The signal applied as PoD input is the signal measured at PA output, whereas the desired PoD output signal is the signal applied at PA input. Hence, indirect learning does not require any PA model because the PA measurements are done offline.

Indirect learning architecture can be performed several times. However, two iterations are able to achieve acceptable accuracies [3]. At the first iteration, the DPD does not apply distortion, and $x'(n)$ is identical to $x(n)$. The PA output signal is divided by the gain g to remove the PA small-signal gain. The PoD is then adjusted to minimize the error signal $e(n)$. Observe that, at first iteration, $y(n)/g \neq x(n)$ due to the nonlinearities of the PA. At the second iteration, the DPD applies the PA inverse behavior and $x'(n) \neq x(n)$. Again, the PoD coefficient extraction is performed. Observe that,

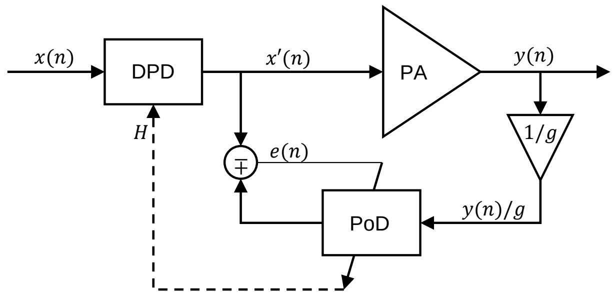

Figure 1: Baseband equivalent of indirect learning.

at the second iteration, $y(n)/g \approx x(n)$, which is a scenario closer to the DPD application and, as a consequence, a more accurate PA inverse behavior is expected in comparison with the first iteration.

3.2 Direct learning architecture

Direct learning is an architecture extracted from [2], which proposes an iterative direct DPD parameter extraction by a nonlinear function to be minimized. Fig. 2 shows the baseband equivalent of direct learning architecture [2, 3]. In the direct learning, the desired pre-inverse output signal is not known. The signal applied as pre-inverse input is the undistorted envelope signal, whereas the desired PA output signal is a linear replica of the DPD input signal. Hence, direct learning does require a PA model because the PA is part of the closed-loop system to be optimized.

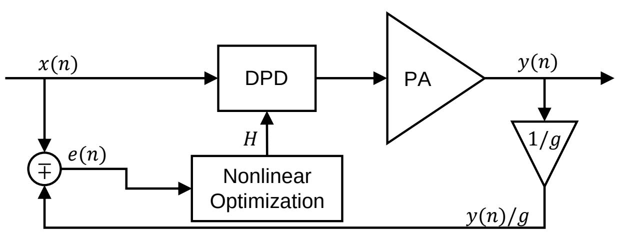

Figure 2: Baseband equivalent of direct learning.

Direct learning applies nonlinear optimization to adjust the coefficients. As consequence, initial values for H are demanded, besides several iterations for convergence [2, 3, 6]. By one hand, direct learning allows DPD identification to be more accurate. By the other hand, the complexity of nonlinear optimization is the most negative aspect of this approach [3, 6].

4 SIMULATION RESULTS

The direct and indirect learnings are applied to the identification of DPDs for two concurrent dual-band PAs using Matlab software. The first PA is a behavioral model adopting the 2D-IMP equation of (1) with the truncation factors set to $M = 3$, $P = 8$ and $Q = 8$. The second PA is a post-layout netlist from a broadband class AB circuit, suitable for integration at 130 nm CMOS and designed to operate from 2 to 7 GHz. Both PAs are concurrently stimulated by a Wi-Fi signal at 2.4 GHz, using IEEE 802.11n with 64-QAM and 20 MHz bandwidth, and a Long-Term Evolution (LTE) envelope at 3.5 GHz with 64-QAM and 20 MHz bandwidth. The 2D-IMP of (1) is used as the DPD model in all cases. To provide a fair comparison, the same values for the truncation factors $M = 3$, $P = 8$ and $Q = 8$ are adopted in both learnings. The truncation factors are set to provide the best compromise between number or parameters and modeling accuracy. Four complex-valued time-domain sequences of the same length must be collected from the PA prior to each iteration of the indirect learning. Two iterations of indirect learning are performed. In the first iteration, the PA is fed with undistorted envelopes. In the second iteration, pre-distorted sequences are inputted to the PA. A huge computational cost is demanded if the post-layout circuit schematic is directly employed as the PA model inside the optimization routine. A more viable solution, which is adopted in this work, is to replace the post-layout circuit netlist by a behavioral model based on the 2D-IMP of (1). The results from the first iteration of the indirect learning are employed as initial estimations for the parameter set during the direct learning.

In what concerns the first PA, the normalized mean square error (NMSE) [9] is computed for both learnings. The NMSE computes the error signal defined as the difference between the output and input signals of the cascade connection given by the DPD followed by the PA. For indirect learning, NMSE values of −48.1 dB and −48.8 dB are evaluated for Wi-Fi and LTE bands, respectively. For direct learning, NMSE values of −50.1 dB and −49.5 dB are evaluated for Wi-Fi and LTE bands, respectively. Since there is no PA modeling error in this case, such degradations of 2.0 dB and 0.7 dB for Wi-Fi and LTE bands, respectively, are attributed to the displacement from post-inverse to pre-inverse adopted by the indirect learning. Figs. 3 and 4 show the output amplitude *vs.* input amplitude (AM-AM) transfer behavior of the cascade (DPD+PA) applying direct and indirect learnings. Very similar plots are observed for the two learnings and both of them provide highly linearized systems. Figs. 5 and 6 show the power spectral densities (PSDs) at PA outputs. Very reduced levels of distortions are noticeable in all PSD plots, although a slightly superior performance is reported to the direct learning.

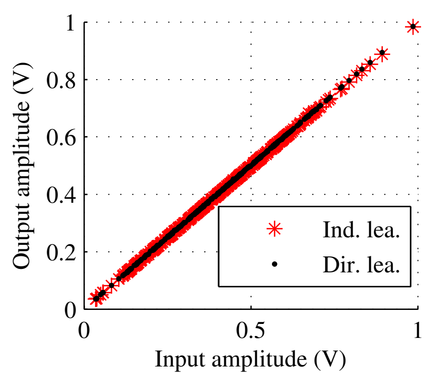

Figure 3: Baseband equivalent AM-AM conversions for Wi-Fi at 2.4 GHz using the first PA.

In what concerns the second PA, Table 1 reports the DC power consumption (P_{DC}), output mean power (P_{out}), efficiency (η) and error vector magnitude (EVM) for the cases without DPD and with

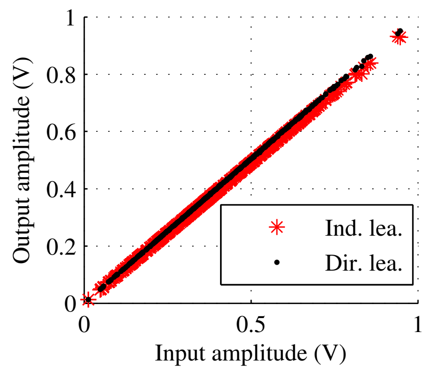

Figure 4: Baseband equivalent AM-AM conversions for LTE at 3.5 GHz using the first PA.

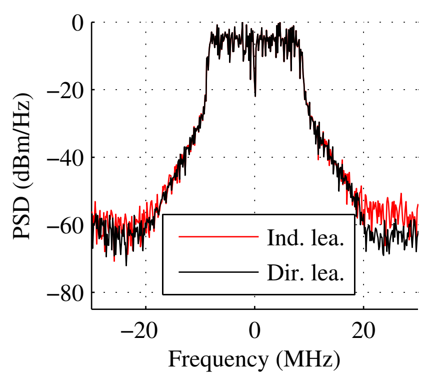

Figure 5: Baseband equivalent PSD of PA output signals for Wi-Fi at 2.4 GHz using the first PA.

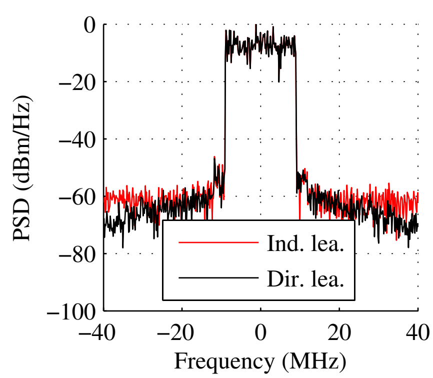

Figure 6: Baseband equivalent PSD of PA output signals for LTE at 3.5 GHz using the first PA.

DPD designed with direct and indirect learnings. P_{DC} is calculated by the sum of the mean powers of all the DC sources. In this scenario of similar P_{out}, a superior improvement in linearity (estimated by EVM) is achieved by the inclusion of the DPD trained by the indirect learning in comparison with the DPD trained by the direct learning. The performances of the DPDs trained by indirect and direct learnings are also compared in Figs. 7 to 10. The AM-AM plots from Fig. 7 are very close to the ideal unitary line, certifying that highly linear curves are achieved when the Wi-Fi band is linearized by both DPDs. However, from the AM-AM plots of Fig. 8, lower distortions at higher amplitudes are observed if indirect learning is adopted instead of the direct one. Whereas the PSDs shown in Fig. 9 are very close, noticeable differences are illustrated in Fig. 10, in which much lower distortion levels at adjacent channels are exhibited by the indirect learning. The superior performance of the indirect learning in comparison with the direct learning indicates that the errors related to the displacement from post-inverse to pre-inverse adopted by the indirect learning are much lower than the errors associated to the PA modeling, since the latter only degrade the performance of the direct learning.

Table 1: Linearization results.

Case	P_{DC} (mW)	P_{out} (mW)	η (%)	EVM (%) Wi-Fi	EVM (%) LTE
Indirect learning Linearization	828.0	60.4	7.3	0.9	1.0
Direct learning Linearization	828.1	60.1	7.3	1.9	3.2
Without Linearization	827.0	60.4	7.3	8.3	10.0

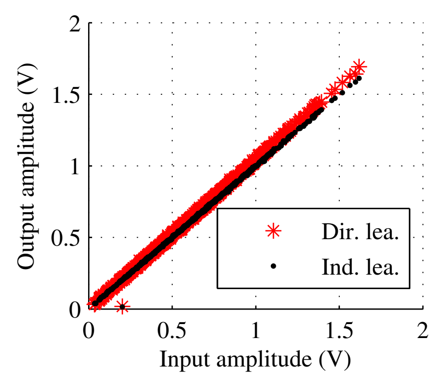

Figure 7: Baseband equivalent AM-AM conversions for Wi-Fi at 2.4 GHz using the second PA.

In what concerns the DPD practical implementation, its power consumption is mostly related to the real-time processing of a known model, since the direct and indirect learnings are performed off-line with adaptation rates several orders lower than the real-time processing [15].

5 CONCLUSIONS

This paper reports results of comparison between direct and indirect learnings in two scenarios of concurrent dual-band PA linearization. At the first scenario, where the PA was represented by a completely

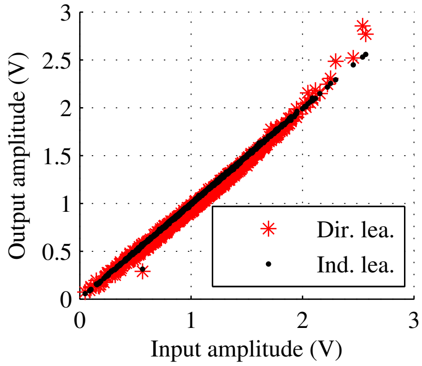

Figure 8: Baseband equivalent AM-AM conversions for LTE at 3.5 GHz using the second PA.

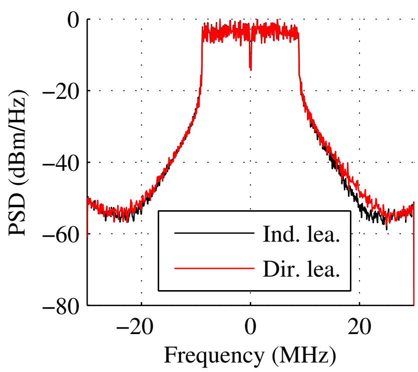

Figure 9: Baseband equivalent PSD of PA output signals for Wi-Fi at 2.4 GHz using the second PA.

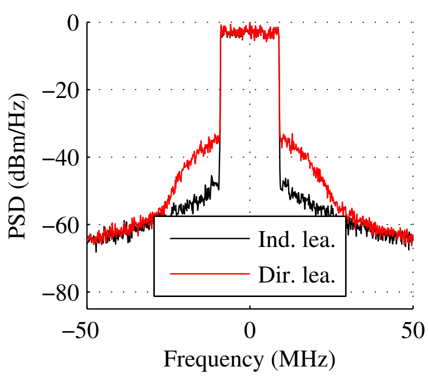

Figure 10: Baseband equivalent PSD of PA output signals for LTE at 3.5 GHz using the second PA.

known behavioral model, direct learning architecture was able to allow more linearity in transmitter, gaining about 2 dB in NMSE metric. For the second scenario, a post-layout circuit-level description of a CMOS 130 nm PA is linearized. Because the direct learning required an approximated PA model, such modeling error resulted in better linearization applying indirect learning, quantified by a difference of 1 p.p. for Wi-Fi and 2 p.p. for LTE in EVM metric.

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and by National Council for Scientific and Technological Development (CNPq).

REFERENCES

[1] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi. 2011. 2-D Digital Predistortion (2-D-DPD) Architecture for Concurrent Dual-Band Transmitters. *IEEE Transactions on Microwave Theory and Techniques* 59, 10 (Oct 2011), 2547–2553. https://doi.org/10.1109/TMTT.2011.2163802

[2] G. Baudoin and P. Jardin. 2001. Adaptive polynomial pre-distortion for linearization of power amplifiers in wireless communications and WLAN. In *EUROCON'2001. International Conference on Trends in Communications. Technical Program, Proceedings (Cat. No.01EX439)*, Vol. 1. 157–160 vol.1. https://doi.org/10.1109/EURCON.2001.937787

[3] J. H. Chavez, C. L. R. Machado, L. H. A. Lolis, and E. G. Lima. 2018. Optimal Parameter Identification for Look-up Table based Band-limited Memory Polynomial Model using Direct and Indirect Learnings. 13, 2 (2018), 1–6. https://doi.org/10.29292/jics.v13i2.36 Exported from https://app.dimensions.ai on 2019/03/30.

[4] Steve C Cripps. 2016. *RF power amplifiers for wireless communications* (2 ed.). Artech House, Norwood, MA.

[5] M. Dardaillon, C. Jabbour, and V. P. Srini. 2015. Adaptive digital pre-distortion for future wireless transmitters. In *2015 IEEE International Conference on Electronics, Circuits, and Systems (ICECS)*. 332–335. https://doi.org/10.1109/ICECS.2015.7440316

[6] N. Kelly, W. Cao, and A. Zhu. 2017. Preparing Linearity and Efficiency for 5G: Digital Predistortion for Dual-Band Doherty Power Amplifiers with Mixed-Mode Carrier Aggregation. *IEEE Microwave Magazine* 18, 1 (Jan 2017), 76–84. https://doi.org/10.1109/MMM.2016.2616185

[7] J. Kim and K. Konstantinou. 2001. Digital predistortion of wideband signals based on power amplifier model with memory. *Electronics Letters* 37, 23 (Nov 2001), 1417–1418. https://doi.org/10.1049/el:20010940

[8] V John Mathews and Giovanni Sicuranza. 2000. *Polynomial signal processing*. John Wiley & Sons, Inc.

[9] M. S. Muha, C. J. Clark, A. A. Moulthrop, and C. P. Silva. 1999. Validation of power amplifier nonlinear block models. In *1999 IEEE MTT-S International Microwave Symposium Digest (Cat. No.99CH36282)*, Vol. 2. 759–762 vol.2. https://doi.org/10.1109/MWSYM.1999.779870

[10] G. Naik, J. Liu, and J. J. Park. 2018. Coexistence of Wireless Technologies in the 5 GHz Bands: A Survey of Existing Solutions and a Roadmap for Future Research. *IEEE Communications Surveys Tutorials* 20, 3 (thirdquarter 2018), 1777–1798. https://doi.org/10.1109/COMST.2018.2815585

[11] J. C. Pedro and S. A. Maas. 2005. A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches. *IEEE Transactions on Microwave Theory and Techniques* 53, 4 (April 2005), 1150–1163. https://doi.org/10.1109/TMTT.2005.845723

[12] E. J. Powers. 1997. A new Volterra predistorter based on the indirect learning architecture. *IEEE Transactions on Signal Processing* 45, 1 (Jan 1997), 223–227. https://doi.org/10.1109/78.552219

[13] D. Raychaudhuri and N. B. Mandayam. 2012. Frontiers of Wireless and Mobile Communications. *Proc. IEEE* 100, 4 (April 2012), 824–840. https://doi.org/10.1109/JPROC.2011.2182095

[14] Chance Tarver, Mahmoud Abdelaziz, Lauri Anttila, Mikko Valkama, and Joseph R. Cavallaro. 2018. Low-complexity, Multi Sub-band Digital Predistortion. *Journal of Signal Processing Systems* 90, 10 (01 Oct 2018), 1495–1505. https://doi.org/10.1007/s11265-017-1303-1

[15] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager. 2010. A Comparative Analysis of the Complexity/Accuracy Tradeoff in Power Amplifier Behavioral Models. *IEEE Transactions on Microwave Theory and Techniques* 58, 6 (June 2010), 1510–1520. https://doi.org/10.1109/TMTT.2010.2047920