BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF RADIO ELECTRONICS

DIGITAL PREDISTORTERS WITH LOW-COMPLEXITY ADAPTATION

ADAPTACE DIGITÁLNÍCH PŘEDZKRESLOVAČŮ S NÍZKOU SLOŽITOSTÍ

DOCTORAL THESIS
DISERTAČNÍ PRÁCE

AUTHOR
AUTOR PRÁCE

Jan Král

ADVISORS
VEDOUCÍ PRÁCE

doc. dr. Ing. Tomáš Götthans
prof. Ing. Roman Maršálek, Ph.D.

BRNO 2022

ABSTRACT

Modern communication systems often require digital predistorters (DPDs), advanced signal-processing units, to satisfy stringent demands on transmitter linearity and efficiency. Nevertheless, DPD significantly increases the hardware and computational complexity of transmitters, which leads to increased power consumption and expenses. Therefore, we propose methods to achieve lower hardware and computational complexity of DPD adaptation. The principle of real-valued feedback samples allows for saving one of two originally-needed feedback analogue-to-digital converters (ADCs), which implies reduced transmitter complexity and power consumption. Furthermore, the hardware and computational complexity can be reduced if the feedback samples for the DPD adaptation are undersampled and carefully selected. The proposed techniques select samples based on histograms and can reduce the required number of feedback samples to a few tens. The provided analyses show approximately 400-times reduced computational complexity achieved by the sample selection and 40-times reduced power consumption of the undersampling feedback ADCs. The real-valued feedback, its undersampling, and sample selection constitute fundamental principles of the proposed DPD adaptation with a level-crossing ADC, which is realised by a simple comparator. Replacing the conventional ADCs with a comparator significantly reduces the design complexity and power consumption. All the proposed and described techniques are accompanied by simulations, usually confirmed by measurements on real hardware, and compared with state-of-the-art methods. The final discussion analyses the limitations, usability and advantages of the proposed techniques. It shows that reducing complexity might not be universally applicable and all the design constraints and specifications must be carefully assessed.

KEYWORDS

Digital predistortion, power amplifier linearisation, low-complexity predistorter adaptation, feedback undersampling, histogram-based sample selection, comparator-based level-crossing analogue-to-digital converter.

ABSTRAKT

Moderní komunikační systémy často vyžadují digitální předzkreslovače (DPD), jednotky pokročilého zpracování signálu, ke splnění požadavků na linearitu a účinnost vysílačů. Nicméně DPD výrazně zvyšují hardwarovou a výpočetní složitost vysílačů, což vede ke zvýšení spotřeby a nákladů. Z toho důvodu představujeme metody k dosažení nižší hardwarové a výpočetní složitosti adaptace DPD. Metoda zpětnovazebních vzorků s reálnou hodnotou vyžaduje pouze jeden ze dvou původně potřebných zpětnovazebních analogově-digitálních převodníků (ADC), což vede ke snížení složitosti vysílače a spotřeby energie. Hardwarovou a výpočetní složitost lze více snížit, pokud je zpětnovazební signál pro adaptaci DPD podvzorkován a jsou vybrány pouze vhodné vzorky. Navrhované techniky vybírají vzorky dle histogramu a mohou snížit potřebný počet vzorků zpětné vazby na několik desítek. Provedené analýzy ukazují přibližně 400násobné snížení výpočetní náročnosti dosažené výběrem vzorků a 40násobné snížení spotřeby energie díky podvzorkování zpětné vazby DPD. Metody zpětné vazby s reálnými hodnotami, její podvzorkování a výběr vzorků tvoří základ navrhované adaptace DPD využívající ADC s detekcí jedné úrovně, který je realizován jednoduchým komparátorem. Nahrazení běžných zpětnovazebních ADC jednoduchým komparátorem výrazně snižuje složitost návrhu a spotřebu energie. Všechny navržené a popsané techniky jsou doprovozeny simulacemi, obvykle i měřením na reálném hardwaru a porovnány s dostupnými metodami. Závěrečný rozbor řeší použitelnost a omezení jednotlivých metod. Ukazuje se, že snížení složitosti nemusí být univerzálně dosažitelné a je třeba pečlivě posoudit všechna specifika daného návrhu.

KLÍČOVÁ SLOVA

Digitální předzkreslovač, linearizace výkonových zesilovačů, adaptace předzkreslovače s nízkou složitostí, podvzorkování zpětné vazby, výběr vzorků pomocí histogramu, analogově digitální převodník s detekcí jedné úrovně s komparátorem.

BIBLIOGRAPHIC CITATION

KRÁL, Jan. *Digital Predistorters with Low-Complexity Adaptation*. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Radio Electronics, 2022, 161 p. Doctoral thesis. Advised by doc. dr. Ing. Tomáš Götthans and prof. Ing. Roman Maršálek, Ph.D.

Author's Declaration

Author:	Jan Král
Author's ID:	125500
Publication type:	Doctoral thesis
Academic year:	2021/22
Topic:	Digital Predistorters with Low-Complexity Adaptation

I declare that I have written this paper independently, under the guidance of my advisors and using exclusively the technical references and other sources of information cited in the paper and listed in the comprehensive bibliography at the end of the paper.

As the author, I furthermore declare that, with respect to the creation of this paper, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll. of the Czech Republic, Section 2, Head VI, Part 4.

In Brno, 17 March 2022.

Jan Král

ACKNOWLEDGEMENT

First of all, I would like to thank kindly my advisors, Tomáš Götthans and Roman Maršálek, for their exceptional encouragement, highly sophisticated research ideas, and valuable tips when I was stuck and lost in my nonworking Matlab simulations, as well as their thorough revision and improvement of our papers and this thesis. Thank you for shielding me from the administrative stuff. I could not fully acknowledge this before, but I know now how painful this administrative work is. I am very glad that both of you have looked after me during my doctoral studies so that I could benefit from the combination of your extraordinary qualities. Without you, this thesis would be unachievable. Thank you also very much for being great friendly colleagues.

I am really glad that we could collaborate with Professor Markus Rupp and his group at Technical University Wien. I would like to thank them for this enriching collaboration, for their expert advice and comments. Professor Rupp, thank you for reviewing our papers and moving them to the next level. Also, thank you for your patience with us.

I must remember and acknowledge our former student Lukáš Jagla for designing the feedback module with a real comparator in the feedback for the digital predistorter adaptation and building a measurement setup with this module as part of his Master's thesis. Thank you, Lukáš. Without your work, our results would be much less valuable.

My gratitude also belongs to all my colleagues at the Department of Radio Electronics for creating a nice working environment. In particular, I would like to thank my friends: Michal Harvánek for being a supportive colleague and sports partner; Martin Pospíšil for knowing everything and discussing all my various theoretical and practical questions – generally known as "Ask Martin" service; Michal Kubíček, Aleš Povalač and Vojtěch Derbek for being excellent supportive project partners; Aslihan Kartci for creating an international office; Ondřej Domanský for his crazy ideas; and Lukáš Langhammer for being an example in publication activities.

I would like to express my thanks to Zbyněk Raida and Tomáš Kratochvíl, the current and former heads of the Department of Radio Electronics, for all the support and for letting me become a part of the team.

Last but not least, I would like to thank the reviewers of this thesis for their time and valuable comments.

In Brno, 17 March 2022.

Jan Král

VĚNOVÁNÍ

Tuto práci bych chtěl věnovat našim synům, Lukášovi a Davidovi. Chtěl bych, abyste věděli, že tahle práce pro mě byla zatím nejtežší a zároveň i jedna z nejdelších životních zkoušek, které jsem si vytyčil. Největší obtíží pro mě nebyla odborná náročnost práce, byť řešená problematika není úplně jednoduchá na pochopení, ale hlavně jsem bojoval s nalezením motivace tuto práci dokončit – protože je to cíl, který vlastně k ničemu nepotřebuji. Nicméně jsem nechtěl jen tak to vzdát a zahodit úsilí, které jsem posledních šest let této práci věnoval, což mě donutilo k tomu ji nakonec dokončit. Tímto bych vám chtěl předat životní moudro, že někdy je potřeba hodně úsilí, zejména toho psychického, zatnout zuby, vydžet a výsledek se nakonec dostaví. Dovolil bych si zde citovat svého středoškolského profesora fyziky, který mě toho hodně naučil a říkal: „Podívejte se, komunisté sebrali lidem majetky, sebrali jim svobodu, ale vzdělání, vzdělání vám nikdo nikdy nesebere.“ Budete-li mít tu možnost – a vězte, že se budeme maximálně snažit, abyste tu možnost měli – potom se vzdělávejte, protože jednak je to zábava a za druhé se budete stejně vzdělávat celý život v různých směrech.

Chtěl bych zde moc poděkovat svojí ženě Jolče za všechnu podporu. Vím, že je to občas dost časově náročné skloubit všechny moje aktivity dohromady. Joli, jsem Ti moc vděčný, že jsem nikdy nemusel vysvětlovat, proč někdy chodím z práce tak pozdě nebo proč jsem občas chtěl chodit do práce i o víkendu. Moc děkuji za Tvoje pochopení.

Poděkovat bych chtěl také svým rodičům a bráchovi Tomovi, že mi poskytli podporu a prostor účastnit se všelijakých středoškolských aktivit – fyzikálních, matematických a programovacích olympiád a vytváření klukovských robotů – vím, že jsem v těchto dobách nemusel tolik uklízet jako jindy a mnoho dalšího. Děkuji vám také za bezstarostný vysokoškolský život, bez něho bych nemohl dělat to, co jsem chtěl.

V Brně, dne 17.03.2022.

Honza

Contents

List of Figures

List of Tables

Introduction

Modern communication systems must provide ever-increasing data throughput. This demand is usually satisfied by increased communication bandwidth and by developing more spectrum-efficient modulation schemes. The modern spectrum-efficient modulations put stringent demands on transmitters and receivers, especially on their hardware imperfections. In the currently developed *fifth generation* (5G) and *sixth generation* (6G) systems, one of the main concerns is the linearity of *radio frequency* (RF) *power amplifiers* (PAs) in transmitters [1], because the nonlinear distortion leads to a degraded constellation diagram resulting in higher bit error rate and to the spectrum regrowth, undesired transmission and potential interference in regions close to the communication channel. The linear PAs, however, suffer from low power efficiency. They are usually physically larger and require powerful cooling systems, which naturally implies their higher manufacturing expenses. Therefore, designers tend to prioritise their power efficiency at the expense of their linearity and overall transmitter linearity is achieved by linearisation techniques.

One of the linearisation techniques is digital predistortion. The *digital predistorter* (DPD) modifies a signal going to the nonlinear PA by artificial nonlinearity with complementary characteristics to cancel the PA nonlinearity. The digital predistortion is currently the most promising linearisation technique with the highest achievable linearisation performance [2]. However, the DPDs, especially their adaptation, are very computationally demanding and often require complex feedback circuits to track changing PA characteristics.

The higher computational and hardware complexity represents one of the current research challenges. Even though a lot of research has been conducted in this direction, e.g. [3–5], there is still plenty of space for improvements. In this thesis, therefore, we have oriented to low-complexity methods for DPDs. The goal of this thesis is to extend state-of-the-art knowledge of low-complexity methods for PA linearisation to make DPDs generally more applicable, less expensive and more efficient. We introduce three key methods to lower the computational and hardware complexity of the DPD adaptation:

- real-valued feedback,
- feedback sample selection,
- feedback with a *level-crossing analogue-to-digital converter* (LC-ADC).

For the conventional PA adaptation, the PA output is usually down converted by an *in-phase and quadrature* (IQ) mixer and both in-phase and quadrature signals are sampled by two *analogue-to-digital converters* (ADCs). The real-valued feedback method enables avoiding one of the two ADCs which reduces the hardware complexity, saves power and lowers transmitter costs.

Usually the feedback ADCs are required to continually sample the PA output at a sampling frequency which covers three to five times of the communication channel bandwidth [6, 7]. The feedback sample selection naturally allows for feedback undersampling, which significantly decreases the power consumption of the feedback ADCs. Additionally, if the feedback samples are carefully selected, only a few samples are required for successful DPD adaptation. Consequently, calculations with lower dimension matrices lead to a significant reduction of computational complexity. Our simulations and analyses show that the computational complexity can be reduced up to 400 times compared with the conventional approach.

The LC-ADC in feedback replaces two conventional feedback ADCs by a simple comparator. The principle of this method is based on combining principles of the real-valued feedback samples and feedback sample selection. The PA output is not continuously sampled and cannot be fully recovered in the digital processing, but rather the comparator detects when the PA output is crossing a set reference voltage. Knowledge of time instants of the comparator output edges and the set reference voltage is sufficient for the DPD adaptation. Our measurements indicate performance of DPD adaptation with the proposed feedback comparable to DPD adaptation with the conventional feedback with two ADCs. The advantages of the LC-ADC are lower power consumption, costs and footprint than of the equivalent feedback solution with conventional ADCs.

This thesis is structured as follows. Chapter 1 introduces theoretical and historical basics of PA linearisation. Mathematical models of PAs and DPDs, signal modulations, and performance metrics, which are used throughout this thesis, are defined here as well. Chapter 2 reviews the recent advances of PA linearisation, especially of the DPDs. We identified several research directions and focused particularly at low-complexity predistortion methods. In Chapter 3, DPD adaptation with real-valued feedback samples by different DPD architectures is described. The linearisation performance of the DPD adapted by the real-valued feedback samples and the conventional approach are simulated and compared. Feedback sample selection methods are analysed in Chapter 4. We present a general principle of sample selection methods for DPD adaptation in different architectures and later we propose various sample selection methods. We analyse their computational complexity reduction and simulated linearisation performance. Additionally, we confirmed the simulation results by measurements. Chapter 5 describes principles of DPD adaptation with the LC-ADC in the feedback, and measurements of its performance on three different hardware setups. Finally, Chapter 6 summarises the proposed low-complexity approaches and discusses their limitations, usability, and advantages.

1 Theoretical Introduction

1.1 Power Amplifier Basics

Amplifiers are electronic components that increase signal strength in terms of voltage, current, or power. The ideal amplifier output $y(t)$ is fully proportional to its input $x(t)$ and, without limitations, follows the equation [8, 9]

$$y(t) = G\,x(t), \tag{1.1}$$

where G is a constant amplifier gain. The schematic symbol of an ideal amplifier is depicted in Fig. 1.1.

The described linear model is sufficient when the output signal of the amplifier is relatively small and the amplifier works within very limited operating conditions. However, in reality, these conditions are usually met only for low-power amplifiers working with no or negligible load. In contrast, high-power amplifiers (simply PAs) usually suffer from nonlinearity and memory effects.

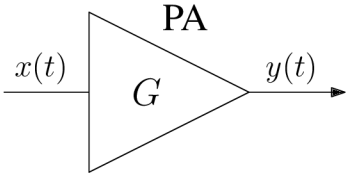


Fig. 1.1: Schematic symbol of an ideal amplifier with gain G.

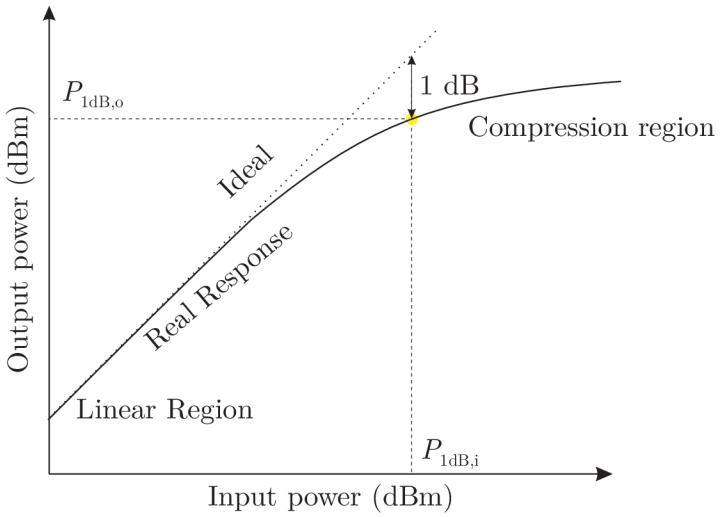


Fig. 1.2: The compression point of 1 dB in output-input power characteristics.

1.1.1 Compression Point of 1 dB

As the input signal level increases and the amplifier output approaches the supply voltage, the amplifier output cannot follow the linear relation (1.1) anymore and saturates. As a result, the output signal level is lower than the ideal output level, as shown in Fig 1.2. The point, where the actual output level is 1 dB lower than the ideal output level, is called the 1-dB-compression point. It is described by its output power level $P_{\mathrm{1dB,o}}$ and, correspondingly, by its input power level $P_{\mathrm{1dB,i}}$. The 1-dB-compression point is the simplest indicator of the nonlinearity of an amplifier. [8, 9]

1.1.2 AM/AM and AM/PM Characteristics

A more advanced method of describing the RF amplifier is by its *amplitude/amplitude* (AM/AM) and *amplitude/phase* (AM/PM) characteristics. Let us assume that the amplifier input signal is given as

$$x_{\mathrm{c}}(t) = x(t)e^{\mathrm{j}2\pi f_{\mathrm{c}}t}, \tag{1.2}$$

where $x(t)$ is a complex envelope and f_{c} is the carrier frequency . The power amplifier output can be described as

$$y(t) = F_{\mathrm{AM}}(|x(t)|) \cdot e^{\mathrm{j}\arg\{x(t)\}+\mathrm{j}F_{\mathrm{PM}}(|x(t)|)}e^{\mathrm{j}2\pi f_{\mathrm{c}}t}, \tag{1.3}$$

where $F_{\mathrm{AM}}(|x(t)|)$ represents the AM/AM characteristics and $F_{\mathrm{PM}}(|x(t)|)$ represents the AM/PM characteristics. The operators $|\cdot|$ and $\arg\{\cdot\}$ signify the magnitude and the angle of the given complex parameters, respectively. [8]

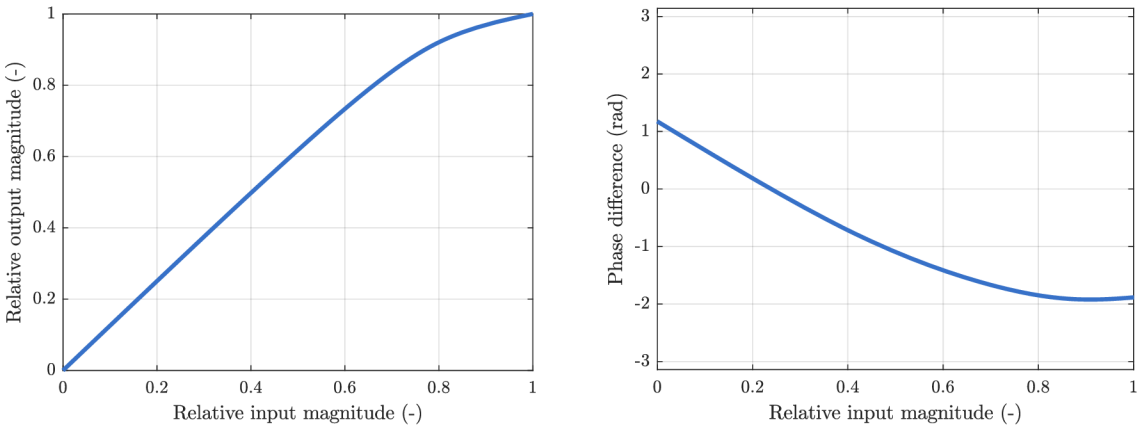
In other words, the AM/AM characteristic defines the envelope magnitude of the amplifier output with respect to the envelope magnitude of the amplifier input. This can be seen as an amplitude distortion. The AM/PM characteristic defines the added angle to the complex envelope phase with respect to the input complex envelope magnitude. This can therefore be seen as a phase distortion. The examples of AM/AM and AM/PM characteristics are plotted in Fig. 1.3.

1.1.3 Intermodulation Distortion

Intermodulation occurs when two (or generally multiple) signals or tones of different frequencies are input into a nonlinear component. Let us assume that the input signal $x(t)$ consists of two tones with frequencies f_1, f_2 and amplitudes A. It can be described as [9]

$$x(t) = A\cos(2\pi f_1 t) + A\cos(2\pi f_2 t) = 2A\ \cos(2\pi f_{\mathrm{d}} t)\ \cos(2\pi f_{\mathrm{p}} t), \tag{1.4}$$

where $f_{\mathrm{d}} = \frac{f_2 - f_1}{2}$ and $f_{\mathrm{p}} = \frac{f_1 + f_2}{2}$.



(a) AM/AM characteristic. (b) AM/PM characteristic.

Fig. 1.3: An example of PA characteristics.

Assuming the amplifier has no memory effects, the output $y(t)$ is given by the actual input $x(t)$ at time t and does not depend on its history. When considering the amplitude distortion only, the nonlinearity can be modelled by general polynomial series and the amplifier output can be written as

$$y(t) = a_1 x(t) + a_2 x^2(t) + \ldots + a_k x^k(t) + \ldots + a_K x^K(t), \tag{1.5}$$

where k is the nonlinearity order and K is the maximum nonlinearity order. By incorporating (1.4) into (1.5), we get

$$y(t) = \sum_{k=1}^{K} 2A\, a_k \, \cos^k(2\pi f_\mathrm{d} t)\, \cos^k(2\pi f_\mathrm{p} t), \tag{1.6}$$

which can be further expanded based on the binomial theorem, the exponential definition of the cosine, and the substitution $u = 2\pi f_\mathrm{d} t$, $v = 2\pi f_\mathrm{p} t$ as [9]

$$\begin{aligned}
x^n(t) &= 2^n A^n \left(\frac{e^{\mathrm{j}u} + e^{-\mathrm{j}u} + e^{\mathrm{j}v} + e^{-\mathrm{j}v}}{2} \right)^n = \left| \begin{array}{l} c = e^{\mathrm{j}u} \\ d = e^{\mathrm{j}v} \end{array} \right| = \\
&= A^n \left(c + c^{-1} + d + d^{-1} \right)^n = A^n \sum_{k=0}^{n} \binom{n}{k} (c + c^{-1})^{n-k} (d + d^{-1})^k = \\
&= A^n \sum_{k=0}^{n} \binom{n}{k} \left[\sum_{i=0}^{n-k} \binom{n-k}{i} c^{n-k-i}\, c^{-i} \right] \left[\sum_{h=0}^{k} \binom{k}{h} d^{k-h}\, d^{-h} \right] = \\
&= A^n \sum_{k=0}^{n} \sum_{i=0}^{n-k} \sum_{h=0}^{k} \binom{n}{k} \binom{k}{h} \binom{n-k}{i} c^{n-k-2i}\, d^{k-2h} = \\
&= A^n \sum_{k=0}^{n} \sum_{i=0}^{n-k} \sum_{h=0}^{k} \binom{n}{k} \binom{k}{h} \binom{n-k}{i} e^{\mathrm{j}u(n-k-2i) + \mathrm{j}v(k-2h)}.
\end{aligned} \tag{1.7}$$

When looking only at complex harmonic components, by back substituting for u, v, f_d, and f_p we get an intermodulation product at frequency $f_\mathrm{im}[n,k,i,h]$ given by eq. (1.7) as

$$\begin{aligned} f_\mathrm{im}[n,k,i,h] &= (n-k-2i)f_\mathrm{d} + (k-2h)f_\mathrm{p} = \\ &= \frac{(n-k-2i)(f_2-f_1)+(k-2h)(f_1+f_2)}{2} = \\ &= \left(\frac{n}{2}-i-h\right)f_2 - \left(\frac{n}{2}-k-i-h\right)f_1. \end{aligned} \tag{1.8}$$

As the frequency of the intermodulation product is a combination of frequencies f_1 and f_2, we define the frequency $f_\mathrm{im}[m]$ of the m-th order intermodulation product as

$$f_\mathrm{im}[m] = p\,f_1 + q\,f_2, \tag{1.9}$$

where p, q is any combination of positive or negative integers which comply with the relation $m = |p| + |q|$. Even-order *intermodulation distortion* (IMD) products are far away from the original signals. They are often ignored, as they are filtered out by analogue filters anyway. However, odd-order IMD products are problematic as they are too close to the original signals and cannot be filtered by analogue filters. Dominant IMD products closest to the original signals are of the 3$^{\mathrm{rd}}$ and 5$^{\mathrm{th}}$ orders (Fig. 1.4). [8,9]

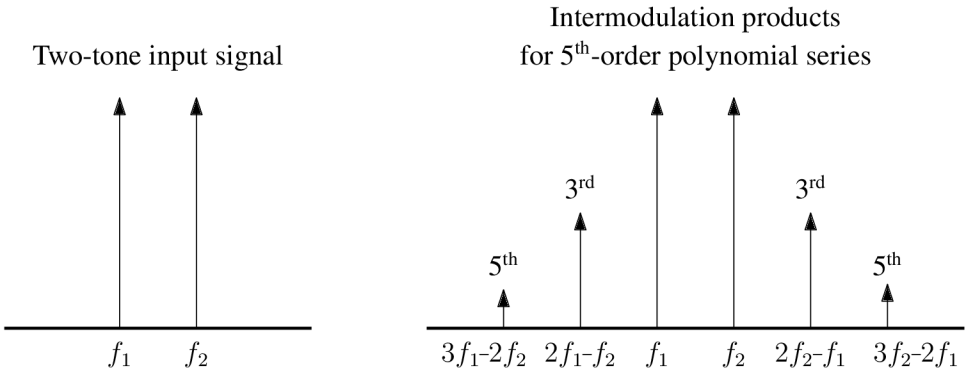


Fig. 1.4: IMD of two-tone input (left) generates intermodulation products in the vicinity of the original tones (right).

The intermodulation effect has been analysed by the two-tone test. This test was extensively employed in history to test the linearity, when most of the communication signals were amplitude or frequency modulated and narrow-band. Nowadays, the two-tone test importance has been lowered as the character of communication signals has changed significantly, and the test might not provide sufficient information about the impact of the distortion caused by a tested amplifier on a signal with a modern modulation scheme.

1.1.4 Intercept Point

The intercept point of the m-th order (IPm) can be defined in two ways. Firstly, it can be defined based on the harmonics of a single tone input signal that appear due to nonlinear distortion at the m-th multiple of the original signal frequency. The other definition is based on m-th order intermodulation products of two closely spaced input tones. The second variant is mostly used in practice, as it does not require the wide bandwidth of the tested amplifier. Please note that these two definitions differ by approximately 9.5 dB ($20 \log_{10} 3$). [8, 9]

The intercept point is graphically obtained (Fig. 1.5) from the AM/AM plot with logarithmic scales. The linear parts of the output power of the wanted amplified original signal and the intermodulation distortion (or harmonics) of m-th order are extrapolated. In logarithmic scales, the function x^m is translated into a straight line with slope of m which means that the slope of the wanted (linearly amplified) component is 1 and e.g. the 3$^{\text{rd}}$ order IMD has slope of 3. The intercept point of m-th order is given by the intersection of extrapolated lines. In other words, the intercept point is a point for which the power of the m-th order intermodulation product is equal to the linearly amplified component power at the output.

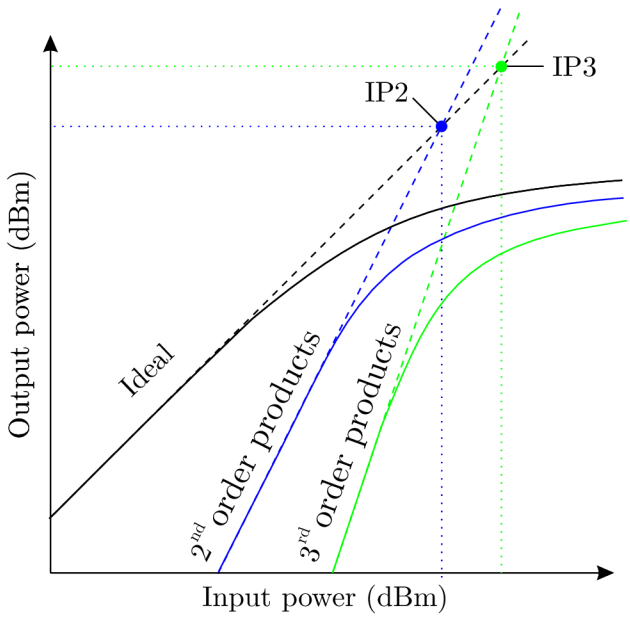


Fig. 1.5: A graphical representation of the intercept point in AM/AM plots.

Similar to the compression point of 1 dB, the concept of the intercept point is an inappropriate way to characterise amplifiers in present communication systems where the signals are different from narrow band, almost single tone, signals.

1.1.5 Power Amplifier Efficiency

Power amplifier efficiency is an essential parameter of power amplifiers. It denotes how much of the spent power is delivered to the PA output. The lower the efficiency is, the higher the power consumption the PA has, considering a constant output power. This is the primary concern in battery-supplied devices. Another aspect is power dissipation, since the PA with lower efficiency requires better cooling.

To objectively evaluate PA efficiency, *power-added efficiency* (PAE) has been introduced. It is defined as [8]

$$\eta_{\mathrm{PAE}} = \frac{P_y - P_x}{P_{\mathrm{DC}}}, \tag{1.10}$$

where P_y and P_x are output and input signal powers, respectively, and P_{DC} is the PA power consumption.

Fig. 1.6 depicts a typical contradiction between PA efficiency and linearity. If the PA operating point is set far from saturation, the PA exhibits good linearity, but the efficiency is low. Vice versa, if the operating point is set close to saturation, the efficiency is high, but the linearity is inferior. A common approach is to set the PA back-off (operating point) such that the maximum level of the amplified signal is still in the region with decent linearity.

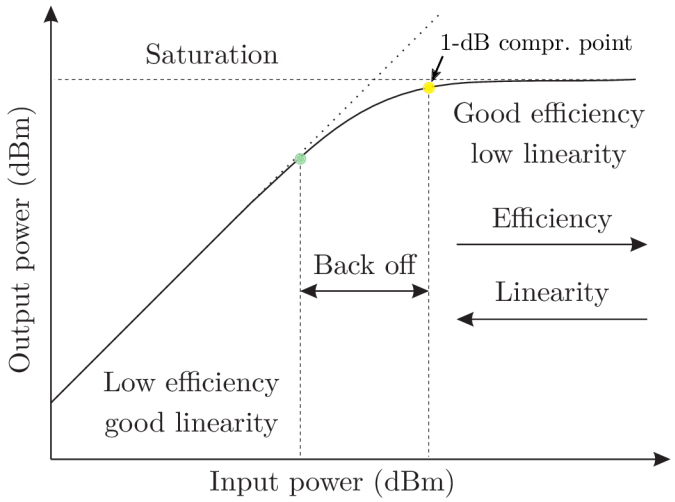


Fig. 1.6: The trade-off between PA efficiency and linearity in AM/AM characteristics.

1.2 Basic Principles of Power Amplifier Linearisation

The early ideas of amplifier linearisation are quite old already. The first references are even older than the discovery of the transistor effect in 1956. Bendel in [10] showed the linearisation of an amplifier by negative feedback, quite interestingly already in the baseband. Leypold [11] demonstrated negative feedback for the phase distortion correction a few years later. Their techniques were, of course, purely analogue and designed for the electron tubes.

The basic concepts of linearisation are simple. The transmitted signal or the PA parameters are modified to obtain a desired output signal. The PA parameters are modified mainly in the analogue domain, and this type of linearisation is often referred to as nonlinearity compensation. For example, the amplifier gain can be increased for higher magnitudes of the input signal. On the other hand, the modification of the transmitted signal before entering the PA is called predistortion. Predistortion can be achieved in both the analogue or digital domain.

Ideally, the predistorter is set to cancel out the distortion of the transmitter nonlinearity. Its characteristics can be seen as a nonlinearity inverse, if the inversion exists. This ideal case is depicted in Fig. 1.7.

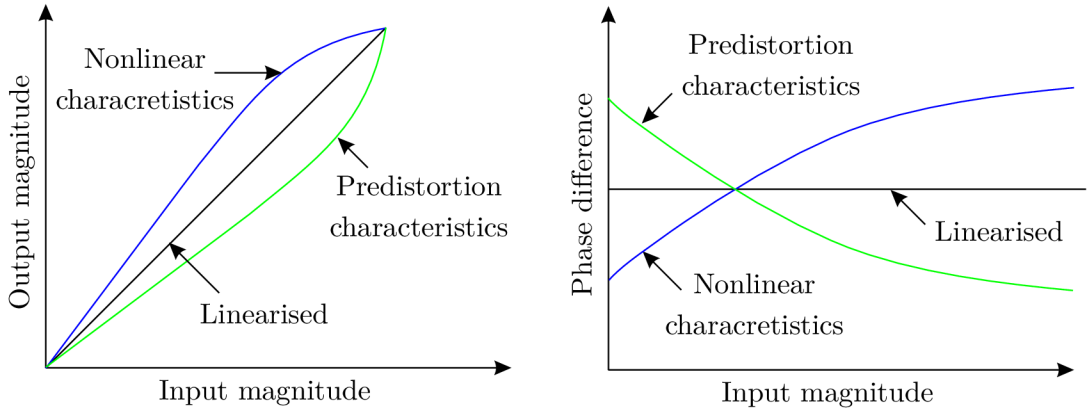


Fig. 1.7: Nonlinearity and ideally-inverted predistortion characteristics in AM/AM and AM/PM plots.

The motivation behind the predistortion is an increased operating range of amplifiers. The predistortion allows for a reduction in PA back-off and, as the operation moves closer to saturation, an increase in PA efficiency without compromising system linearity. This effect is shown in Fig. 1.8. Consequently, the PA dimensions and power rating can be decreased, heat sinks can be miniaturised, and smaller power supplies are required, which naturally results in resource savings and price reduction.

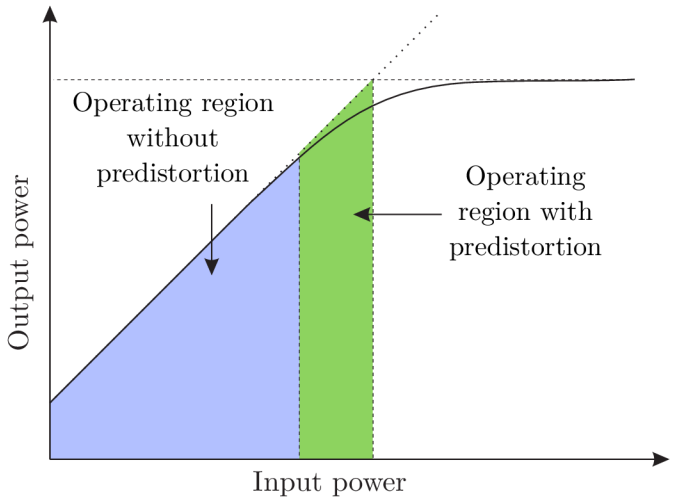


Fig. 1.8: Increased PA operating range by employing predistortion.

Predistortion can be introduced at any point in the system preceding the PA. Based on the point where the predistortion is applied, we distinguish between baseband, intermediate frequency, and RF predistorters. Later, throughout this thesis, we focus exclusively on the baseband DPDs.

The first references to the baseband DPDs were published by Nagata [12] and Cavers [13]. Both employed adaptive predistorters with feedback. The adaptive predistorter can follow changes in nonlinear PA characteristics over time. These changes occur mainly due to temperature changes and component ageing. The concept of the DPD with feedback is shown in Fig. 1.9.

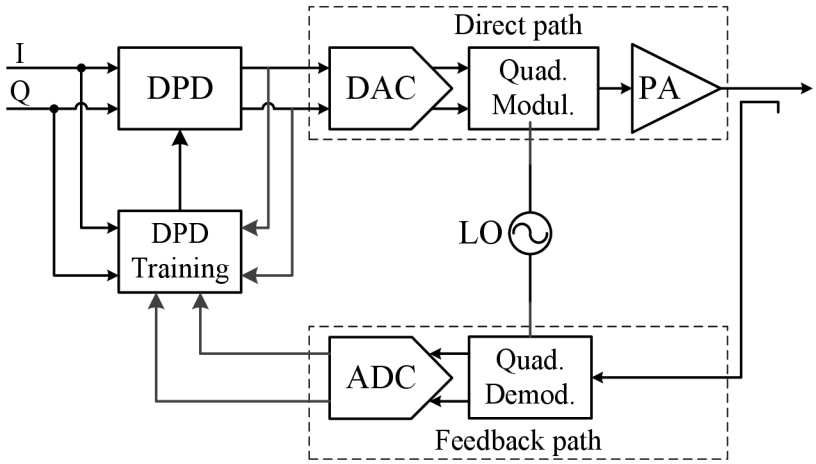


Fig. 1.9: A block diagram of a simplified real system with a digital baseband predistorter.

The PA output signal is taken as feedback by an RF coupler. It is usually down-converted by an IQ demodulator and sampled by an ADC; hence, it can be digitally processed. The digital processing identifies the PA or DPD parameters; a few strategies to identify the PA or DPD parameters are presented later in Section 1.4. Subsequently, the transmitted signal is predistorted by the DPD with already identified parameters, converted into the analogue domain by the *digital-to-analogue converter* (DAC), up-converted by an IQ modulator, and amplified.

Please note that the DPD can be employed in a system with a transmitter of any architecture (direct conversion, heterodyne, superheterodyne, etc.). The depicted architecture is just an example, and the main principle remains the same over different architectures.

1.3 Employed Models

In this section, we introduce readers to the mathematical models of PAs and DPDs employed in this thesis. All the presented models describe the baseband behaviour of the modelled component.

1.3.1 Memory Polynomial Model

The *memory polynomial* (MP) model is a basic mathematical model for modelling PAs and DPDs. It was firstly introduced by Kim and Konstantinou in [14] as the evolution of a simple polynomial baseband model without memory. The model is simple and, generally, it achieves limited linearisation performance, compared to other more complex models.

The discrete baseband output y of the MP model is given as [14]

$$y[n] = \sum_{k=1}^{K} \sum_{q=0}^{Q} b_{k,q} x[n-q] |x[n-q]|^{k-1}, \tag{1.11}$$

where x is the MP model input, $b_{k,q}$ is the coefficient of the MP model, and K and Q represent the maximum nonlinearity order and memory length, respectively. The product $x[n-q]\,|x[n-q]|^{k-1}$ is often called a basis waveform or a basis function. Example basis functions for the MP model are shown in Fig. 1.10. We denote the basis function as

$$\phi_{k,q}^{(x)}[n] = x[n-q]\,|x[n-q]|^{k-1}. \tag{1.12}$$

The input samples x, model coefficients $b_{k,q}$, and the basis waveforms $\phi_{k,q}^{(x)}[n]$ can be arranged into vectors and matrices

$$\begin{aligned}
\boldsymbol{\phi}_{k,q}^{(x)} &= \left[\phi_{k,q}^{(x)}[0] \quad \phi_{k,q}^{(x)}[1] \quad \dots \quad \phi_{k,q}^{(x)}[N-1]\right]^T, \\
\boldsymbol{x} &= \left[x[0] \quad x[1] \quad \dots \quad x[N-1]\right]^T, \\
\boldsymbol{y} &= \left[y[0] \quad y[1] \quad \dots \quad y[N-1]\right]^T, \\
\boldsymbol{b} &= \left[b_{1,0} \quad b_{1,1} \quad \dots \quad b_{1,Q} \quad b_{2,0} \quad \dots \quad b_{K,Q}\right]^T, \\
\boldsymbol{U}_x &= \left[\boldsymbol{\phi}_{1,0}^{(x)} \quad \boldsymbol{\phi}_{1,1}^{(x)} \quad \dots \quad \boldsymbol{\phi}_{1,Q}^{(x)} \quad \boldsymbol{\phi}_{2,0}^{(x)} \quad \dots \quad \boldsymbol{\phi}_{K,Q}^{(x)}\right]^T,
\end{aligned} \tag{1.13}$$

where $\boldsymbol{b}$ is a column vector with $P = K(Q+1)$ rows, and the size of the matrix $\boldsymbol{U}_x$ is $N \times K(Q+1)$. Equation (1.11) can be rewritten into the matrix form

$$\boldsymbol{y} = \boldsymbol{U}_x \boldsymbol{b}. \tag{1.14}$$

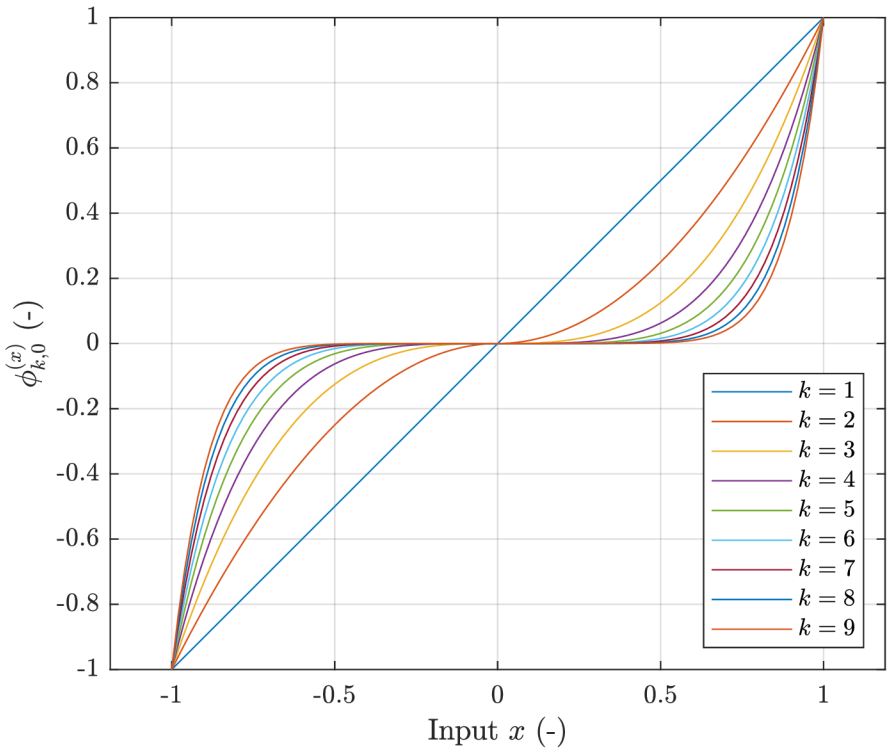


Fig. 1.10: MP model basis without memory effects ($q = 0$) for the real-valued model input ($x_\mathrm{i}(t) = 0$).

1.3.2 Simplified 2nd-Order Dynamic Deviation Reduction-Based Volterra Model

The *simplified 2nd-order dynamic deviation reduction-based Volterra* (DDR2) model is a popular model employed for modelling PAs and DPDs. Guan and Zhu presented it for the first time in [15]. The DDR2 model is based on the Volterra series [16, 17] which is a nonlinear model similar to the Taylor series [18], but it adds the ability to capture memory effects. The full Volterra series is, however, impractically complex for modelling PAs and DPDs. Therefore, the DDR2 model includes only certain selected terms from the full Volterra series.

The DDR2 model has been defined as [15]

$$\begin{aligned} y[n] = &\sum_{k=0}^{\frac{K-1}{2}} \sum_{q=0}^{Q} b_{0,k,q} \, |x[n]|^{2k} \, x[n-q] + \\ &+ \sum_{k=1}^{\frac{K-1}{2}} \sum_{q=1}^{Q} b_{1,k,q} \, |x[n]|^{2(k-1)} \, x^2[n] \, x^*[n-q] + \\ &+ \sum_{k=1}^{\frac{K-1}{2}} \sum_{q=1}^{Q} b_{2,k,q} \, |x[n]|^{2(k-1)} \, x[n] \, |x[n-q]|^2 + \\ &+ \sum_{k=1}^{\frac{K-1}{2}} \sum_{q=1}^{Q} b_{3,k,q} \, |x[n]|^{2(k-1)} \, x^*[n] \, x^2[n-q], \end{aligned} \tag{1.15}$$

where $b_{0,k,q}$, $b_{1,k,q}$, $b_{2,k,q}$, $b_{3,k,q}$ are the model coefficients. We can denote the basis function as

$$\phi_{i,k,q}^{(x)}[n] = \begin{cases} |x[n]|^{2k} \, x[n-q] & \text{if } i = 0, \\ |x[n]|^{2(k-1)} \, x^2[n] \, x^*[n-q] & \text{if } i = 1, \\ |x[n]|^{2(k-1)} \, x[n] \, |x[n-q]|^2 & \text{if } i = 2, \\ |x[n]|^{2(k-1)} \, x^*[n] \, x^2[n-q] & \text{if } i = 3. \end{cases} \tag{1.16}$$

The model coefficients $b_{i,k,q}$ and the basis waveforms $\phi_{k,q}[n]$ can be arranged into vectors and a matrix similarly as in (1.13)

$$\begin{aligned} \boldsymbol{\phi}_{i,k,q}^{(x)} &= \begin{bmatrix} \phi_{i,k,q}^{(x)}[0] & \phi_{i,k,q}^{(x)}[1] & \dots & \phi_{i,k,q}^{(x)}[N-1] \end{bmatrix}^T, \\ \boldsymbol{b} &= \begin{bmatrix} b_{0,1,0} & b_{0,1,1} & \dots & b_{0,1,Q} & b_{0,2,0} & \dots & b_{3,K,Q} \end{bmatrix}^T, \\ \boldsymbol{U}_x &= \begin{bmatrix} \boldsymbol{\phi}_{0,1,0}^{(x)} & \boldsymbol{\phi}_{0,1,1}^{(x)} & \dots & \boldsymbol{\phi}_{0,1,Q}^{(x)} & \boldsymbol{\phi}_{0,2,0}^{(x)} & \dots & \boldsymbol{\phi}_{3,K,Q}^{(x)} \end{bmatrix}^T, \end{aligned} \tag{1.17}$$

where $\boldsymbol{b}$ is a column vector with $P = K(2Q + \frac{1}{2}) - Q + \frac{1}{2}$ rows, and the size of the matrix $\boldsymbol{U}_x$ is $N \times K(2Q + \frac{1}{2}) - Q + \frac{1}{2}$. The matrix form of the model equation is the same as (1.14).

1.3.3 Generalised Memory Polynomial Model

The *generalised memory polynomial* (GMP) model is another popular model employed for modelling PAs and DPDs. It was presented by Morgan et al. in [19]. The GMP model is also derived from the Volterra series [16, 17]. In contrast to the DDR2 model, the GMP model selects different terms from the Volterra series and is defined more universally. The selected terms do not follow a predefined order, but they can be selected sparsely. The GMP model is one of the most complex models employed in this thesis. Naturally, it requires the highest number of coefficients and, due to our observation, it can achieve the best linearisation performance compared with the other presented models.

The model can be defined as [19]

$$\begin{aligned} y[n] = &\sum_{k\in K_a}\sum_{l\in L_a} b_{0,k,l}\, x[n-l]\,|x[n-l]|^k + \\ &+ \sum_{k\in K_b}\sum_{l\in L_b}\sum_{m\in M_b} b_{1,k,l,m}\, x[n-l]\,|x[n-l-m]|^k + \\ &+ \sum_{k\in K_c}\sum_{l\in L_c}\sum_{m\in M_c} b_{2,k,l,m}\, x[n-l]\,|x[n-l+m]|^k, \end{aligned} \tag{1.18}$$

where $b_{0,k,l}$, $b_{1,k,l,m}$, and $b_{2,k,l,m}$ are the PA model coefficients; K_a and L_a are the index arrays for aligned signal and envelope (memory polynomial); K_b, L_b, and M_b are the index arrays for signal and lagging envelope; and K_c, L_c, and M_c are index arrays for signal and leading envelope. We can denote its basis function as

$$\phi^{(x)}_{i,k,l,m}[n] = \begin{cases} x[n-l]\,|x[n-l]|^k & \text{if } i=0, \\ x[n-l]\,|x[n-l-m]|^k & \text{if } i=1, \\ x[n-l]\,|x[n-l+m]|^k & \text{if } i=2. \end{cases} \tag{1.19}$$

The model coefficients $b_{0,k,l}$ and $b_{i,k,l,m}$, and the basis functions $\phi_{k,q}[n]$ can be arranged into vectors and a matrix similarly as in (1.17)

$$\begin{aligned} \boldsymbol{\phi}^{(x)}_{i,k,l,m} &= \left[\phi^{(x)}_{i,k,l,m}[0] \quad \phi^{(x)}_{i,k,l,m}[1] \quad \dots \quad \phi^{(x)}_{i,k,l,m}[N-1]\right]^T, \\ \boldsymbol{b} &= \left[b_{0,k_0,l_0} \quad b_{0,k_0,l_1} \quad \dots \quad b_{0,k_0,l_{\overline{\overline{L_a}}}} \quad b_{0,k_1,l_0} \quad \dots \quad b_{3,k_{\overline{\overline{K_c}}},l_{\overline{\overline{L_c}}},m_{\overline{\overline{M_c}}}}\right]^T, \\ \boldsymbol{U}_x &= \left[\boldsymbol{\phi}^{(x)}_{0,k_0,l_0,0} \quad \boldsymbol{\phi}^{(x)}_{0,k_0,l_1} \quad \dots \quad \boldsymbol{\phi}^{(x)}_{0,k_0,l_{\overline{\overline{L_a}}},0} \quad \boldsymbol{\phi}^{(x)}_{0,k_1,l_0,0} \quad \dots \quad \boldsymbol{\phi}^{(x)}_{3,k_{\overline{\overline{K_c}}},l_{\overline{\overline{L_c}}},m_{\overline{\overline{M_c}}}}\right]^T, \end{aligned} \tag{1.20}$$

where $\overline{\overline{(.)}}$ is the cardinality (number of elements) of a given set, $\boldsymbol{b}$ is a column vector with $P = \overline{\overline{K_a}}\cdot\overline{\overline{L_a}} + \overline{\overline{K_b}}\cdot\overline{\overline{L_b}}\cdot\overline{\overline{M_b}} + \overline{\overline{K_c}}\cdot\overline{\overline{L_c}}\cdot\overline{\overline{M_c}}$ rows, and the size of the matrix $\boldsymbol{U}_x$ is $N \times P$. The matrix form of the model equation is the same as (1.14).

1.4 DPD Adaptation Strategies

In this section, we present basic DPD adaptation strategies, also called architectures, specifically those architectures required for understanding the proposed methods or employed for comparison. This section does not represent a comprehensive elaboration of all known DPD adaptation strategies.

1.4.1 Direct Learning Architecture

The *direct learning architecture* (DLA) [20] is a DPD adaptation strategy which determines the DPD coefficients directly from the PA baseband output $y(t)$ and the DPD input $z(t)$. Its typical topology is shown in Fig. 1.11.

Let us assume that $\mathcal{P}(\cdot)$ is a nonlinear transfer function of the PA and $F_{pre}(\cdot)$ is a nonlinear function of the predistorter. The predistorter function should ideally cancel out the PA nonlinearity, resulting in the overall system being linear. This condition can be mathematically written as [9]

$$\mathcal{P}(F_{\text{pre}}(z)) = G_0 z, \tag{1.21}$$

where G_0 is intended gain. If the inversion of $\mathcal{P}(\cdot)$ exists, the predistortion function is given as

$$F_{\text{pre}}(z) = \mathcal{P}^{-1}(G_0 z). \tag{1.22}$$

The $\mathcal{P}(\cdot)$ is nonlinear, therefore, solving equation (1.22) is a nonlinear problem. To obtain its solution, we can employ optimisation techniques and minimise a criterion function $J(\epsilon)$ with the error signal $\epsilon(t)$ defined as [9]

$$\epsilon(t) = \frac{y(t)}{G_0} - z(t) = \frac{\mathcal{P}(F_{\text{pre}}(z))}{G_0} - z(t). \tag{1.23}$$

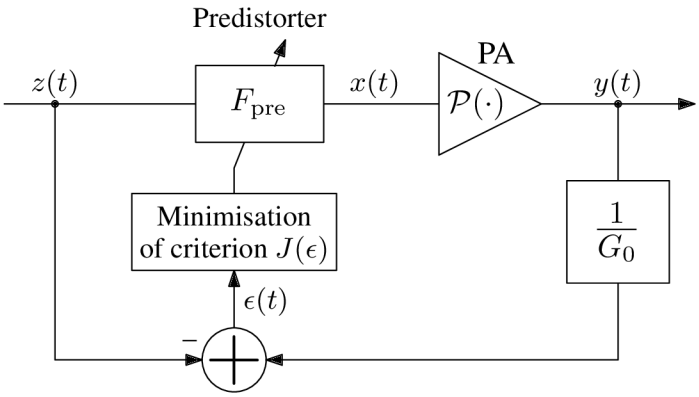


Fig. 1.11: The principle of DPD adaptation by the direct learning architecture.

A typical approach to obtaining the solution is based on solving the nonlinear function numerically using Newton's method. Let us assume that the predistortion function is prescribed by the nonlinear model stated in Sec. 1.3 with a matrix form in equation (1.14). Throughout this thesis, the iterative damped Newton's method is employed to solve the DPD coefficients. The coefficients are updated as [21]

$$\hat{\boldsymbol{b}} = \boldsymbol{b}' - \mu \boldsymbol{e}, \tag{1.24}$$

where vector $\hat{\boldsymbol{b}}$ is a new solution of the DPD coefficients, vector $\boldsymbol{b}'$ is the current vector of the DPD coefficients, and μ is the iteration step size. Vector $\boldsymbol{e}$ is the coefficient error vector. It has the same dimensions as the coefficient vector $\boldsymbol{b}'$ and is given as the *least squares* (LS) solution of

$$\boldsymbol{\Delta} \approx \boldsymbol{U}_z \boldsymbol{e}, \tag{1.25}$$

where $\boldsymbol{\Delta} = \boldsymbol{z} - \boldsymbol{y}$. The solution of $\boldsymbol{e}$ with a pseudoinverse of matrix $\boldsymbol{U}_z$ can be written as

$$\boldsymbol{e} = (\boldsymbol{U}_z^H \boldsymbol{U}_z)^{-1} \boldsymbol{U}_z^H \boldsymbol{\Delta}. \tag{1.26}$$

The final equation of the coefficient update is given by incorporating equation (1.26) into (1.24) as

$$\hat{\boldsymbol{b}} = \boldsymbol{b}' - \mu (\boldsymbol{U}_z^H \boldsymbol{U}_z)^{-1} \boldsymbol{U}_z^H (\boldsymbol{z} - \boldsymbol{y}). \tag{1.27}$$

1.4.2 Indirect Learning Architecture

The *indirect learning architecture* (ILA), contrary to the DLA, solves the DPD model in a noniterative way. The problem of solving the predistorter coefficients is reformulated and the ILA solves the postdistorter coefficients instead of the predistorter coefficients. The employed topology is depicted in Fig. 1.12. We minimise a criterion function $J(e)$ with the error signal $e(t)$ which is a difference between the postdistorted PA output and the PA input [9, 22]

$$\epsilon(t) = F_{\text{post}}\left(\frac{y}{G_0}\right) - x(t), \tag{1.28}$$

where $F_{\text{post}}(\cdot)$ is a nonlinear function of the postdistorter. If the employed model is linear in its unknown parameters, the solution of its coefficients is a linear problem. All models employed in this thesis, and most DPD models, comply with this condition. Note that both signals $y(t)$ and $x(t)$ must be known.

By substituting signal x with the inverse function $\mathcal{P}^{-1}(y)$ in equation (1.28) and assuming the error signal ϵ to be zero, which corresponds to the ideally adapted postdistorter function $F_{\text{post}}(.)$, we get

$$F_{\text{post}}\left(\frac{y}{G_0}\right) = \mathcal{P}^{-1}(y) = \mathcal{P}^{-1}(G_0 z). \tag{1.29}$$

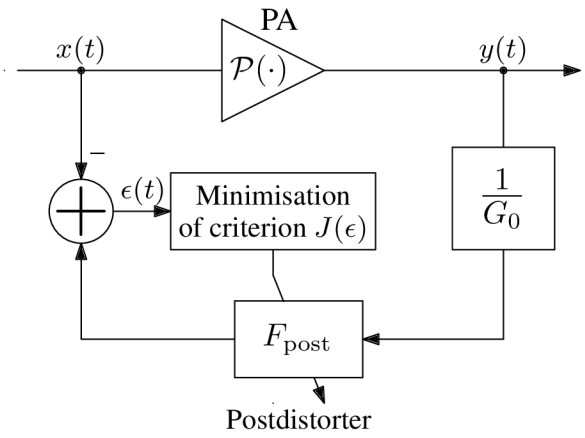


Fig. 1.12: The principle of DPD adaptation by the indirect learning architecture.

If we compare the right-hand side of this equation with equation (1.22), we can see that both equations are equivalent. Consequently, we can state that the solved postdistorter model can be employed as the predistorter and we can set

$$F_{\text{pre}}(\cdot) = F_{\text{post}}(\cdot). \tag{1.30}$$

The fully derived evolution of the predistorter function from the postdistorter function can be found in [23].

The solution of the postdistorter can be written in the form of a matrix equation. Assuming that the postdistorter model is described in the matrix form (1.14), the solution of a typical overdetermined system of equations is given by linear algebra as [24]

$$\boldsymbol{b'} = (\boldsymbol{U}_y^H \boldsymbol{U}_y)^{-1} \boldsymbol{U}_y^H \boldsymbol{x}, \tag{1.31}$$

where $(\boldsymbol{U}_y^H \boldsymbol{U}_y)^{-1} \boldsymbol{U}_y^H$ is a pseudoinverse of matrix $\boldsymbol{U}_y$. Please note that for the postdistorter model, signal $y(t)$ is the input and goes into matrix $\boldsymbol{U}_y$, while signal $x(t)$ is the desired postdistorter output.

1.4.3 Indirect Learning Architecture with Forward Model Estimation

It has been known [20, 25, 26] that the predistorter solution provided by the ILA is biassed and that the ILA could be suboptimal. This phenomenon is related to the present noise in the measured PA output $y(t)$ which samples form the system matrix $\boldsymbol{U}_y$. This problem does not exist for the DLA, because matrix $\boldsymbol{U}_z$ is created by noise-free samples of signal $z(t)$.

Morgan [25] and Landin [26] solved the problem with the biassed solution by identifying the PA forward model followed by an internal calculation of the noise-free output of the forward PA model and its usage for the regular ILA calculation. Hereinafter, we will refer to this method as the *forward model indirect learning architecture* (FM-ILA). Its block diagram is depicted in Fig. 1.13. The main drawback of FM-ILA is the increased computational complexity as the forward PA model and its output is calculated in addition to the calculation of the regular ILA.

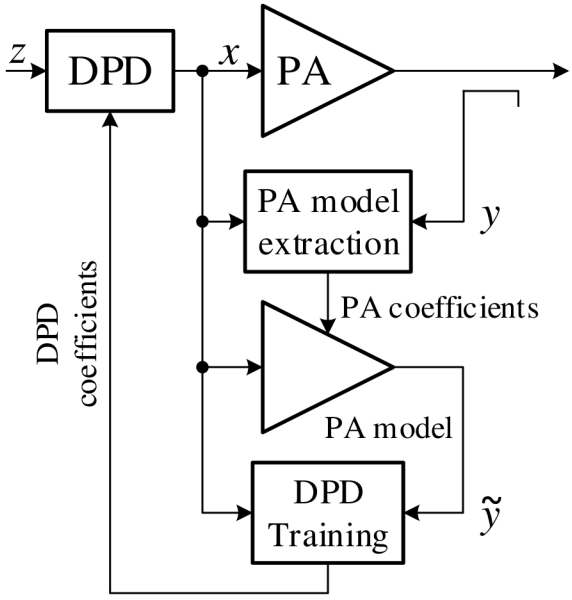


Fig. 1.13: The block diagram of the DPD adapted by the FM-ILA.

The FM-ILA procedure can be described mathematically as solving the forward PA model coefficients as

$$\boldsymbol{b} = (\boldsymbol{U}_x^H \boldsymbol{U}_x)^{-1} \boldsymbol{U}_x^H \boldsymbol{y}. \tag{1.32}$$

The forward PA model output $\tilde{\boldsymbol{y}}$ is calculated as

$$\tilde{\boldsymbol{y}} = \boldsymbol{U}_x \, \boldsymbol{b}, \tag{1.33}$$

and the postdistorter coefficients $\boldsymbol{b}'$ are solved using the forward PA model output $\tilde{\boldsymbol{y}}$ and the PA input $\boldsymbol{x}$ as

$$\boldsymbol{b}' = (\boldsymbol{U}_{\tilde{y}}^H \boldsymbol{U}_{\tilde{y}})^{-1} \boldsymbol{U}_{\tilde{y}}^H \boldsymbol{x}. \tag{1.34}$$

The calculated postdistorter coefficients are used as the predistorter coefficients in the next iterations.

1.5 Employed Modulations

In this section, we introduce the basics of modulation schemes employed for test signals, hereinafter in this thesis. This brief overview serves only as a reference for the performed simulations and experiments. A comprehensive elaboration would be out of the scope of this thesis.

1.5.1 Orthogonal Frequency Division Multiplexing

orthogonal frequency-division multiplexing (OFDM) is a multicarrier modulation scheme that dates back to 1966 and respective works [27–29]. Recently, multicarrier modulations have become very popular due to their favourable properties and the possibility of realistic and effective *fast Fourier transform* (FFT) and *inverse fast Fourier transform* (IFFT) implementations. The multicarrier modulations divide a given frequency channel into several subchannels, known as subcarriers. These subcarriers spread over a narrow band only, which allows for the effective equalisation of a radio channel, as one-tap equalisers can be employed to equalise subcarrier by subcarrier.

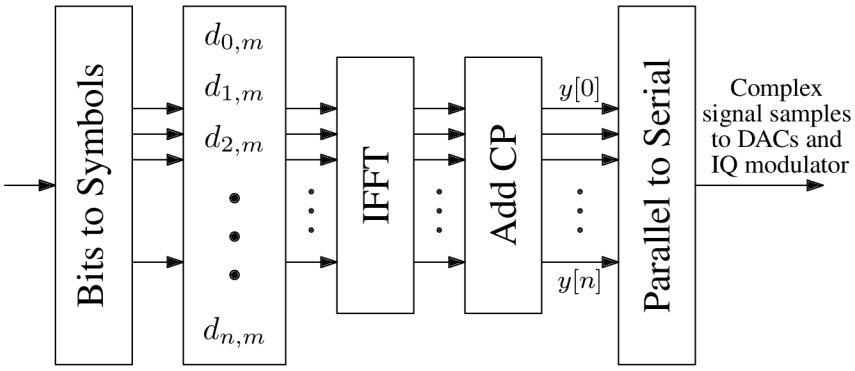


Fig. 1.14: A simplified block diagram of an OFDM modulator. Data bits are mapped into data symbols from a predefined alphabet in the first block and the CP is added in the fourth block.

A simple block diagram of an OFDM modulator is shown in Fig. 1.14. The OFDM-modulated signal $y(t)$ can be mathematically described as [30]

$$y(t) = \sum_{m=-\infty}^{\infty} \sum_{k \in \mathcal{K}} g_{k,m}(t)\, d_{k,m}, \tag{1.35}$$

where $d_{k,m}$ is a transmitted complex symbol at the m-th time position and k-th subcarrier, and $\mathcal{K}$ is a set of the active subcarriers. Signal $g_{k,m}(t)$ describes a modulation

pulse shape which is defined as [30]

$$g_{k,m}(t) = \rho(t - m\Delta T)e^{\mathrm{j}2\pi kF(t-m\Delta T)}, \tag{1.36}$$

where F is a frequency spacing between subcarriers, ΔT is a time spacing between transmitted symbols and function $\rho(t)$ describes the modulation prototype filter. The OFDM modulation has the prototype filter defined as [30]

$$\rho(t) = \begin{cases} 1 & \text{if } t \in [0, \Delta T] \\ 0 & \text{otherwise,} \end{cases} \tag{1.37}$$

which is a rectangular pulse with the width of $\Delta T = 1/F$. The final equation for the OFDM modulation can be obtained by incorporating (1.36) into (1.35) as

$$y(t) = \sum_{m=-\infty}^{\infty} \sum_{k \in \mathcal{K}} d_{k,m}\rho(t - m\Delta T)e^{\mathrm{j}2\pi kF(t-m\Delta T)}. \tag{1.38}$$

1.5.2 Filtered Orthogonal Frequency Division Multiplexing

The *filtered orthogonal frequency-division multiplexing* (F-OFDM) modulation is an evolution of the OFDM. The original OFDM suffers from high out-of-band emissions caused by the rectangular-pulse filtering. The problem of out-of-band emissions has been addressed and solved by the F-OFDM in exchange for dropping strict orthogonality of subcarriers.

The mathematical description of an F-OFDM modulated signal can be defined similarly to the OFDM and equation (1.35), employing the prototype filter. The F-OFDM prototype filter for this thesis has been designed as proposed in [31], [32]. The filter coefficients $f_B[n]$ are given as

$$f_B[n] = \frac{p_B[n]\, w[n]}{\sum\limits_n p_B[n]\, w[n]}, \tag{1.39}$$

where $p_B[n]$ is the sinc function and $w[n]$ is the window function. The sinc function is defined as

$$p_B[n] = \mathrm{sinc}\left((W + 2\,\Delta W)\frac{n}{Z}\right), \tag{1.40}$$

where Z is the FFT size, W is the number of assigned data subcarriers, and ΔW is the tone-offset, where $2\Delta W$ is the difference between the desired pass band and the designed sinc filter pass band [32]. The window function is defined as

$$w[n] = \left(0.5\left(1 + \cos\left(\frac{2\pi n}{L-1}\right)\right)\right)^{0.6}, \tag{1.41}$$

where $-\left\lfloor \frac{L}{2} \right\rfloor \leq n \leq \left\lfloor \frac{L}{2} \right\rfloor$ and L is the number of filter taps.

Discrete Fourier Transform Precoding

discrete Fourier transform (DFT) precoding is a well-known technique that reduces the *peak-to-average power ratio* (PAPR) of the OFDM modulated signals before their transmission. The precoding principal and properties have been described in [33–35]. The basic principle is quite simple: a DFT precoding block is added before the F-OFDM (alternatively OFDM) modulator, which results in a similar PAPR to single carrier modulations, while the beneficent properties of the OFDM modulation scheme are preserved. A simplified block diagram of the modulator with DFT precoding can be seen in Fig. 1.15.

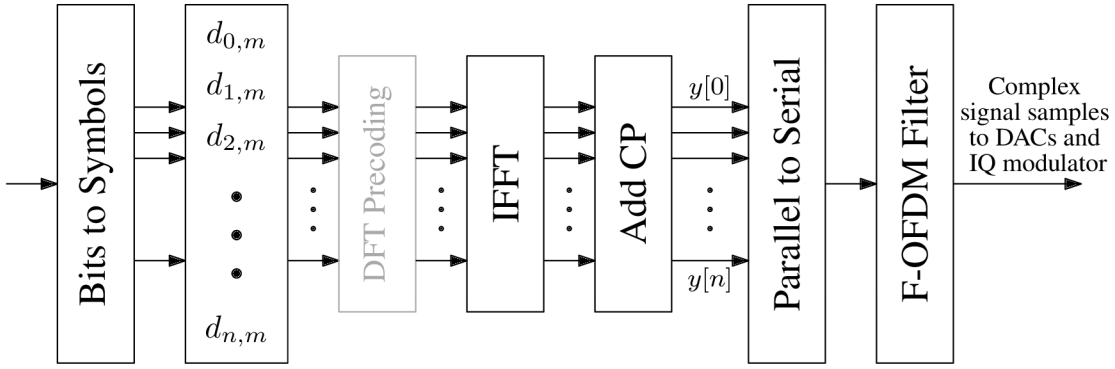


Fig. 1.15: A simplified block diagram of an F-OFDM modulator with an optional DFT precoding.

1.5.3 Filter Bank Multicarrier

The *filter bank multicarrier* (FBMC) modulation is based on transmission through a bank of filters. This approach was first presented in [27, 28]. The idea is to divide a given band into several subbands, similar to the OFDM subcarriers, by bandpass filters. The filter design is a crucial task and it needs to fulfil contradictory requirements, such as symbol density, out-of-band emissions, and orthogonality, have to be balanced [36–38].

We have employed the prototype filter and principles defined by the PHYDYAS project [39]. The impulse response of the filter is defined as [39–42]

$$\rho(t) = \begin{cases} 1 + 2\sum\limits_{i=1}^{\mathcal{O}-1} H_i \cos\left(\frac{2\pi t}{\mathcal{O}T_0}\right) & \text{if } -\frac{\mathcal{O}T}{2} < t \leq \frac{\mathcal{O}T}{2}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.42}$$

where $\mathcal{O}$ is the overlapping factor, T_0 represents a time-scaling parameter and depends on the desired subcarrier spacing (or time spacing), and the numerical

coefficients H_i are defined by the PHYDYAS project [39, 41] and are stated in Tab. 1.1. Our modulator and demodulator are based on the publicly available Matlab code [43].

Tab. 1.1: Numerical coefficients for the PHYDYAS filter.

$\mathcal{O}$	H_0	H_1	H_2	H_3
2	1	$\sqrt{2}/2$	–	–
3	1	0.911438	0.411438	–
4	1	0.971960	$\sqrt{2}/2$	0.235147

1.6 System Performance Metrics

To analyse the DPD performance, we observe these metrics: the *adjacent channel power ratio* (ACPR), *error vector magnitude* (EVM), and the *normalised mean square error* (NMSE).

1.6.1 Adjacent Channel Power Ratio

The ACPR qualifies the unwanted spectrum regrowth caused mainly by the nonlinearity of the transmitter. This metric shows how much the transmitter would interfere in other communications ongoing in channels close to the transmitter's channel. It is defined as the power in the adjacent channel over the power in the main channel [44, 45] (see Fig. 1.16)

$$\mathrm{ACPR} = 10\log_{10}\frac{\int_{f_{\mathrm{adj1}}}^{f_{\mathrm{adj2}}} P_y(f)\,df}{\int_{f_{\mathrm{m1}}}^{f_{\mathrm{m2}}} P_y(f)\,df}, \tag{1.43}$$

where $P_y(f)$ is the signal power spectral density, f_{adj1}, f_{adj2} are bound frequencies of the adjacent channel and f_{m1}, f_{m2} are bound frequencies of the main channel. The difference $f_{\mathrm{m2}} - f_{\mathrm{m2}}$ denotes the main channel bandwidth B. Adjacent channel frequencies are set in the same way as the adjacent channel bandwidth $f_{\mathrm{adj2}} - f_{\mathrm{adj1}} = B$. The ACPR is evaluated for the lower (left hand side) and higher (right hand side) adjacent channels separately. The adjacent channel can also be defined with an offset, e.g., $f_{\mathrm{adj1}} = f_{\mathrm{m2}} + 0.1B$ for the lower adjacent channel and $f_{\mathrm{adj2}} = f_{\mathrm{m1}} - 0.1B$ for the higher adjacent channel.

Throughout this thesis, when we present the ACPR results, the ACPR is averaged for the lower and higher adjacent channels as

$$\mathrm{ACPR(dB)} = 10\log_{10}\frac{1}{2}\left(10^{\frac{\mathrm{ACPR_L}}{10}} + 10^{\frac{\mathrm{ACPR_H}}{10}}\right), \tag{1.44}$$

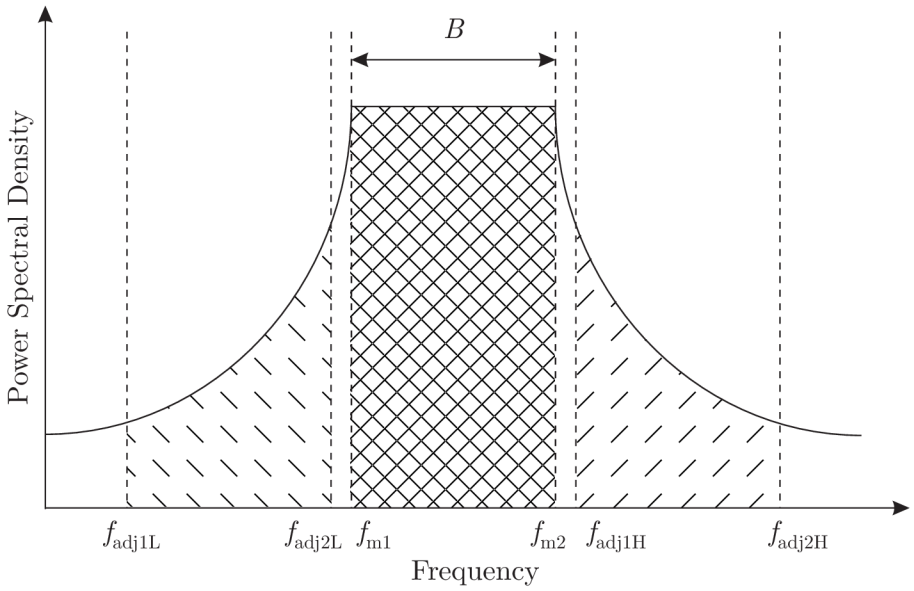


Fig. 1.16: The ACPR bounds depicted in the spectrum.

where $\mathrm{ACPR_L}$ and $\mathrm{ACPR_H}$ are the ACPRs for the lower and higher adjacent channels, respectively. Similarly to (1.44), throughout this thesis, the average ACPR value is calculated as

$$\mathrm{ACPR_{avg}(dB)} = 10\log_{10}\frac{1}{N}\sum_{n=0}^{N-1} 10^{\frac{\mathrm{ACPR}[n]}{10}}, \tag{1.45}$$

where N is the number of averaged values.

1.6.2 Error Vector Magnitude

The EVM evaluates the in-band signal distortion [46, 47]. In contrast to the other presented metrics, the EVM requires the signal to be demodulated and it qualifies errors in the demodulated symbols. The EVM is defined as the *root mean square* (RMS) of the error vectors which connect the reference symbols with the demodulated symbols (see Fig. 1.17) over the normalisation reference expressed in percentages [46, 47]

$$\mathrm{EVM}(\%) = \frac{\sqrt{\frac{1}{N_S}\sum_{n=0}^{N_S-1}\left(I_{\mathrm{err}}[n]^2 + Q_{\mathrm{err}}[n]^2\right)}}{A_{\mathrm{ref}}} \cdot 100\%, \tag{1.46}$$

where n is the symbol index, N_S is the number of demodulated symbols, $I_{\mathrm{err}}[n] = I_{\mathrm{ref}}[n] - I_{\mathrm{demod}}[n]$ is the in-phase component of the error vector, $Q_{\mathrm{err}}[n] = Q_{\mathrm{ref}}[n] - Q_{\mathrm{demod}}[n]$ is the quadrature component of the error vector, and A_{ref} is the EVM normalisation reference.

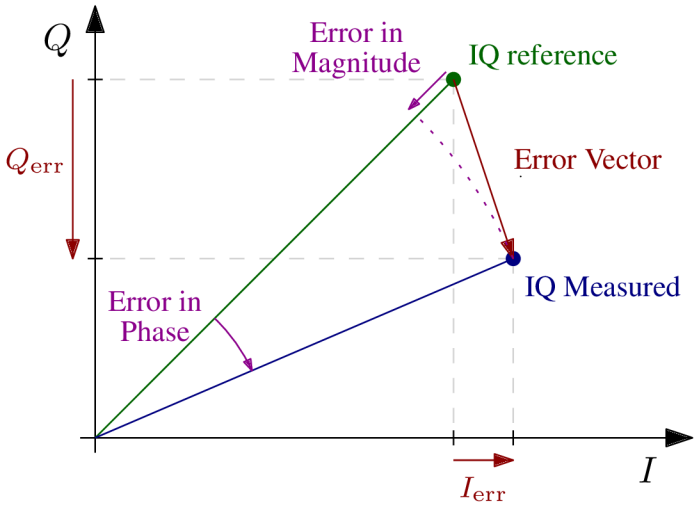


Fig. 1.17: A graphical representation of the received-symbol error vector for calculating the EVM.

There are two possible values for the EVM normalisation reference: the constellation maximum or the reference RMS. EVM normalisation reference as the constellation maximum can be formally written as

$$A_{\text{ref}} = \max_{n} \sqrt{I_{\text{ref}}[n]^2 + Q_{\text{ref}}[n]^2}, \tag{1.47}$$

or as the reference RMS, it can be written as

$$A_{\text{ref}} = \frac{1}{N_S} \sum_{n=0}^{N_S-1} \sqrt{I_{\text{ref}}[n]^2 + Q_{\text{ref}}[n]^2}. \tag{1.48}$$

Throughout this thesis, we strictly show the EVM metric normalised to the reference RMS. Considering its definition, an EVM average is calculated as a root-mean-square value

$$\text{EVM}_{\text{avg}}(\%) = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} |\text{EVM}[n]|^2}, \tag{1.49}$$

where N is the number of averaged values.

1.6.3 Normalised Mean Square Error

The NMSE is used to qualify the difference between ideal and actual signals in at certain point in a system, usually at the output of the transmitter. The NMSE is defined as the mean power of the error signal divided by the mean power of the reference signal x and it is expressed in dB as [48, 49]

$$\text{NMSE(dB)} = 10\log_{10}\frac{\sum\limits_{n}|y[n]-x[n]|^2}{\sum\limits_{n}|x[n]|^2}, \tag{1.50}$$

where y is the actual signal and x is the reference signal. Similarly to the ACPR average in (1.45), the NMSE is averaged as

$$\text{NMSE}_{\text{avg}}\text{(dB)} = 10\log_{10}\frac{1}{N}\sum_{n=0}^{N-1}10^{\frac{\text{NMSE}[n]}{10}}, \tag{1.51}$$

where N is the number of averaged values.

2 State-of-the-Art

In this chapter, we present a summary of the research in the field of predistorters, and especially of digital predistorters, in recent years. We do not intend to provide a comprehensive overview of all papers from the whole history of predistortion. A very brief outline of the history can be found in Sec. 1.2 and more can be found in the referenced review papers in Sec. 2.1.

In recent publications in the field of predistortion, we have identified the following research directions:

- adaptation methods,
- analogue and hybrid predistorters,
- analysis,
- antenna arrays and *multiple-input and multiple-output* (MIMO) systems,
- Doherty power amplifiers,
- hardware implementation,
- low-complexity methods,
- machine learning,
- mathematical models.

Furthermore, the papers about low-complexity predistortion can be divided based on their focus:

- computational complexity,
- model pruning and sizing,
- one-bit feedback samples,
- real-valued feedback samples,
- subband and multiband predistortion,
- and undersampling methods.

In the rest of this chapter, we briefly introduce a few papers for each research direction that received the most interest. More extensively, we described the papers aiming at low-complexity methods for digital predistortion, especially the papers on which this thesis is based.

2.1 Review Papers

Katz et al. [50] presented an overview of PA linearisation techniques such as feed-forward linearisation, predistortion, mitigating memory effects, DPDs, *look-up tables* (LUTs), and approaches to DPD adaptation. At the end, the authors outlined the challenges in the linearisation field.

An extensive overview was presented by Fager et al. in [1]. They focused especially on antenna phased arrays and MIMO systems in 5G communications.

Namely, they covered the following areas: PA behavioural modelling, PA-antenna interactions, power efficiency and linearisation in MIMO transmitters and hybrid or phased-array transmitters, and DPD adaptation with few observation receivers.

Gilabert et al. [51] clearly explained several solutions for the identification of DPD parameters. They covered adaptation strategies such as least mean squares, recursive LS, QR decomposition-based recursive LS, simultaneous perturbation stochastic approximation, *principal component analysis* (PCA), partial LS, and others.

Borel et al. [2] compared and evaluated the most popular linearisation techniques, including feedforward, feedback, and predistortion. They provided a comprehensive comparison of state-of-the-art methods and their performance. The authors clearly stated the advantages and disadvantages of the various linearisation methods and predicted potential future directions, bottlenecks, and obstacles. Finally, they considered a hybrid of analogue and digital predistorters as a promising linearisation architecture.

Singh and Malik [52] presented another review on PA linearity and efficiency. The paper covered feedback, feedforward, and predistortion as linearisation techniques and outphasing, Doherty, and envelope-tracking PAs as efficiency enhancement techniques. The authors described the differences between these techniques and provided a comparison.

Desgreys et al. [53] briefly reviewed predistortion techniques for wideband PAs. Wood in [54] presented challenges posed by wider bandwidth and lower power consumption and offered some potential solutions. Ahmad et al. [55] reviewed machine-learning-based approaches in communications and partially covered applications of machine learning in DPDs. Liu et al. [56] presented a comprehensive review of linearisation techniques targeting Doherty PAs and massive MIMO systems. Gupta et al. [57] presented a review generally oriented towards nonlinear system identification.

2.2 General Research Directions in Predistortion

2.2.1 Adaptation Methods

Researchers still aim at improving the adaptation methods of the predistorter coefficients. Chani-Cahuana et al. [58] proposed an adaptation method based on the *iterative learning control* (ILC), a well-established control theory technique. The proposed method iteratively finds an unknown PA input signal for a desired PA output signal. By knowing the required PA input for the desired PA output, the predistorter coefficients can be easily found. The authors reported improved resistance to noise in the feedback compared with conventional predistorters. Similarly, Schoukens et al. [59] presented a modified technique based on the ILC.

Lopez-Bueno et al. and Pham et al. [60, 61] described an estimation and adaptation method based on the adaptive PCA. Their method leads to an orthogonal transformed matrix which reduces the number of coefficients. Due to its orthogonality, new coefficients can be estimated independently. The proposed adaptive PCA method can find the minimum number of necessary coefficients to meet a specific linearity level.

Belabad et al. [62] proposed a genetic algorithm optimisation method to identify coefficients of the Wiener model of the PA. According to this model, they proposed the Hammerstein-model based DPD.

Rodrigues et al. [63] presented an adaptation of the DPD based on different scalar measurements taken from the feedback. They specifically proposed measurement of the ACPR and spectral mask margin. The DPD model coefficient estimation was reduced to a generic numerical optimisation problem. Additionally, they proposed an orthogonalisation algorithm, resulting in a faster convergence time and lower output power variations across the iterations.

Le Duc et al. [64] showed an adaptive DPD with a cascade of the adaptive ILA and static DLA with a linearly interpolated LUT. The static DLA compensated for the time-invariant part of the nonlinear distortion while the adaptive ILA compensated for the variable nonlinear distortion which was caused by temperature changes and PA ageing.

Le et al. [65] proposed a DPD technique based on adaptive ILA with recursive prediction error minimisation. They designed the forgetting factor to be variant in time; a lower forgetting factor at the beginning ensures fast convergence, while a higher forgetting factor at the end ensures higher immunity against noise and more consistent results in the steady state.

2.2.2 Analogue and Hybrid Predistorters

The main motivations and arguments for analogue predistorters are the computation complexity of DPDs and, consequently, their higher power consumption. The possibilities of DPDs are, however, far beyond the limits of their analogue counterparts. Some researchers have tried combining the advantages of both analogue and digital predistorters and came up with the concept of hybrid and digitally assisted analogue predistorters.

Cai et al. [66] presented a two-transistor-based RF *analogue predistorter* (APD). Their predistorter was based on the similar nonlinearity of the two transistors. However, the predistorter missed adaptability and had to be adjusted for the main PA. The authors extended their work in [67] by adding a realisation of the predistorter

and its measurements to the originally simulated-only results. They reported similar or better predistortion performance with the simpler predistorter compared to competitive works.

Braithwaite [68] proposed a hybrid analogue and digital predistorter for dual-band RF transmitters. Each subband was compensated separately in the digital and analogue predistorters. The DPD was intended for compensating close intermodulation products, whereas the APD should compensate for intermodulation products at multiples of the carrier frequency.

The combination of the analogue and digital predistorter was also employed by Tomé et al. in [69]. They employed the APD to cancel long-term memory effects in GaN *high-electron-mobility transistor* (HEMT) PAs caused by the electron trapping phenomena. They reported improvement in the intermodulation product level by 6.8 dB for the combination of the APD and the DPD compared to the DPD alone.

In [70], the authors proposed a tunable reflective predistorter based on a parallel Schottky diode power-dependent impedance matching network. They reported an improvement of 6.3 dB in the power of the third-order intermodulation products achieved by their predistorter compared to the conventional APDs.

Finally, we mention the paper [71] written by Gumber and Rawat where they proposed a new control scheme for controlling the gain and phase of the predistorted signal. In addition, they compared their method with many state-of-the-art analogue and hybrid predistorters.

2.2.3 Analysis

Cheaito et al. [72, 73] derived an analytical expression for the EVM of the nonlinear PA output. They modelled PA nonlinearity with the clipping effect and PA memory with the MP model. Additionally, they analysed the predistortion complexity with respect to EVM constraints.

Liu et al. [74, 75] studied the influence of reduced ADC dynamic range on DPDs. They also included analyses of the effect of gain and delay mismatch on predistortion performance in [75, 76].

The impact of the normalisation gain of DPDs on the linearisation performance was studied by Wang et al. [77] and Jardin and Baudoin [78] . They showed that adjusting the normalisation gain can improve the PAE, EVM, or the ACPR.

Xiao [79] studied the long-term stability of closed-loop adaptation algorithms. She proposed a modified closed-loop adaptation algorithm achieving the long-term stability which the conventional closed-loop algorithms suffer from.

Chani-Cahuana [49] derived a closed-form expression for the minimum NMSE that could be obtained in systems with predistorters. She based her study on an ideal

scenario where all the distortions introduced by the PA are perfectly compensated for and the linearisation is limited only by the noise added to the measured PA output. Naturally, this limits the DPD adaptation and the linearisation performance.

Finally, Wang et al. [77, 80] investigated the effects of crest factor reduction and digital predistortion complexity on the linearity and power efficiency of PAs. The authors proposed an algorithm to trace an approximate Pareto front of the two criteria, the ACPR and the number of DPD coefficients. Their algorithm helps to analyse the influence of the PA operating point on its efficiency and the required number of DPD coefficients for given ACPR or NMSE requirements.

2.2.4 Antenna Arrays and MIMO Systems

Researchers were interested in predistorters for antenna arrays and MIMO systems, because these systems usually contain many PAs to be linearised. There are many approaches to linearising a bunch of PAs. Liu et al. [81] proposed a linearisation method based on the feedback from a single PA to linearise beamforming MIMO systems. Tervo et al. [82] introduced single over-the-air feedback for nonlinear phased arrays. The authors extended their work in [83] by adding the calibration of the feedback path and considering multiple DPD adaptation strategies. A single-input single-output DPD was also presented by Ng et al. [84, 85]. Their architecture aimed at millimetre-wave beamforming arrays.

Abdelhafiz et al. [86] proposed an augmented crossover MP model for linearisation in MIMO systems in the presence of linear and nonlinear crosstalk. The proposed model was an improvement of the crossover MP model [87]. Abdelaziz et al. reduced the complexity of DPDs by reduced-bandwidth filtered basis functions and estimated the model coefficients based on decorrelation-based closed-loop processing in [88]. The authors extended their work in [89] where they presented a DPD technique utilising the decorrelation-based learning rule in conjunction with a single combined output signal of the individual PAs for hybrid MIMO transmitters.

Wang et al. [90] focused on the linearisation of MIMO transmitters using a real-time single-channel over-the-air data acquisition loop. The proposed technique identified the nonlinear behaviour of all PAs, as well as their combined signals in the far field. They reported excellent linearisation performance with low-complexity hardware.

Zanen et al. [91] analysed the power consumption of different DPD architectures for MIMO systems and antenna arrays and noticed that the power consumption of complex DPDs themselves, in combination with wideband signals, can be higher than the overall consumption of the PAs.

2.2.5 Doherty Power Amplifiers

Doherty PAs are based on two parallel PAs. One PA, called a carrier amplifier, is biassed for Class-A/B or Class-B operation, while the other one, called a peaking amplifier, is biassed for Class-C operation. The peaking amplifier amplifies the peaks in the signal which are normally clipped by the carrier amplifier. The main challenges with the Doherty PAs are splitting the input for the two parallel PAs and connecting the PA outputs via an output power combiner. Two variants of the Doherty PAs exist. There are single-input Doherty PAs [92] and digitally-driven Doherty PAs with two separate inputs, one for the carrier amplifier and the other one for the peaking amplifier [93].

Ghosh and Rawat [94] proposed a digitally-driven Doherty PA with an analytically designed combiner network and employed the DPD for its linearisation. They reported a drain efficiency of 45%–56%.

Chang et al. [95] presented an aliasing-free multilevel pulse modulation architecture for Doherty PA implementation accompanied by a simple memoryless DPD. They reported the improved linearisation performance in ACPR from -36 dB of a conventionally linearised Doherty PA to -41 dB of their linearised architecture.

Masood et al. [96] showed a digitally assisted Doherty PA with a segmented DPD architecture. According to their paper, the improved adjustment of the phase delay between the carrier and peaking PA and the thresholding of the peaking PA result in higher efficiency but degrade the linearity. The proposed segmented DPD architecture is reported to improve linearisation performance by $\approx$ 11 dB in terms of the ACPR compared with the conventional DPD.

Peng et al. [97] proposed an *adaptive signal separation algorithm* (ASSA) for dual-input Doherty PAs to minimise the mismatch between the signal separation functions and the intended operation which is the main source of nonlinearity. They reported linearisation improvement of the ACPR from -43 dB without ASSA to -53 dB with ASSA for test signals with 10-MHz bandwidth.

2.2.6 Hardware Implementation

The researchers also focused on obstacles to implementing predistorters into real hardware, mainly in *field programmable gate arrays* (FPGAs). They optimised the predistorter topology in [98] and modified the decomposed vector rotation-based behavioural model in [99]. Huang et al. [100] proposed a new approach for a parallel processing DPD implementation and achieved a linearisation bandwidth of up to 2.4 GHz with an FPGA running at a 300-MHz clock.

Other researchers were interested also in implementing LUT-based predistorters. Ren [101] proposed the interpolation of the LUT values by the Talyor series to achieve more accurate linearisation. Molina et al. [102] implemented the LUT-based DPD with linear interpolation and extrapolation and extended their work in [103] where they introduced a 2-D LUT based predistorter with bilinear interpolation for concurrent dual-band predistortion.

2.2.7 Machine Learning

Machine learning in the field of PA predistortion has experienced a boom recently. Bhuyan and Sarma [104] proposed an artificial *neural network* (NN) structure called the real-valued focused time delay NN for modelling PA predistorter. The proposed structure splits the baseband complex signal into real and imaginary samples and feeds these samples into delay shift registers of a certain length. All values from the registers are taken as inputs for the NN. This structure was more thoroughly analysed and compared with the conventional methods [105].

Wang et al. [106] continued development of the NN predistorters and introduced an augmented real-valued time-delay NN for joint mitigation of PA nonlinearity and modulator imperfections. They added the magnitudes of baseband complex samples and their powers as inputs to the NN. The authors reported improved mitigation of imperfections by 3–4 dB in terms of the ACPR compared to the real-valued focused time delay NNs.

Wu et al. [107] improved the real-valued time delay NN predistorter by adding direct connections from input to output. They called the proposed NN the residual real-valued time-delay NN, because, in fact, the NN learns only the PA nonlinearity behaviour. The authors reported a lower complexity of the proposed NN compared to the original NN.

Jaraut et al. [108] proposed a composite NN DPD for MIMO systems. They focused on compensating for the PA nonlinearity, IQ imbalances, crosstalk, and the *direct current* (DC) offset. Li et al. in [109] extended the NN predistortion for MIMO systems by the ILC and focused on compensating for nonlinear crosstalk.

Cai et al. [110] proposed a model based on a time-delay *support vector regression* (SVR). The authors reported that the SVR-based method can obtain the optimal model in a short time compared to the conventional NN models. Their SVR model improved modelling accuracy compared to the conventional Volterra-based models and to other NN-based models. Xu et al. [111] extended the SVR-based model by adding the magnitude and phase of complex baseband samples, which reportedly improved the linearisation performance.

Sun et al. [112] presented a link between the PA memory and the memory of bidirectional long short-term memory NNs and based a predistorter on it. Their paper showed that the proposed predistorter can achieve successful linearisation; however, it lacks thorough comparison with the conventional and other NN methods.

Zhang et al. [113] proposed a vector decomposed time-delay NN and an augmented vector decomposed time-delay NN. The authors claimed that the proposed NNs comply more with the physical characteristics of PAs than conventional NNs. Their proposed models take signal sample magnitudes only and recover the phase in the phase recovery layer. The augmented version added the p-th order magnitude inputs. Li et al. [114] applied the idea of vector decomposed NNs and presented a vector decomposed long short-term memory NN and its simplified variant. The authors showed that the proposed model outperformed other DPD variants, but with much more variable parameters. Tanio et al. [115] proposed an envelope time-delay NN DPD, quite similar to the vector decomposed time-delay NNs, and a pruning strategy. They demonstrated that their NN can outperform vector-decomposed NNs while using far fewer floating point operations.

Tripathy et al. [116] presented a deep NN predistorter with Swish [117] and Sigmoid-weighted linear unit [118] activation functions instead of the conventional *rectified linear units* (ReLU) activation function.

Hu et al. [119] employed a deep NN for signal recovery from the measured undersampled and band-limited signal in the feedback for adaptation of DPDs. Sun et al. [120] presented an adaptive strategy for deep NN DPDs.

2.2.8 Mathematical Models

Abdelrahman et al. [121] extended the conventional MP model by adding weighting functions. Their model is suitable for PA characteristics that exhibit strong memory effects at low input power levels and mild memory effects at high input power levels. In such cases, the proposed weighted model can improve linearisation performance while preserving the number of model coefficients.

Li et al. [122] redesigned the cross terms of the *decomposed vector rotation* (DVR) model and presented a simplified model. They showed that the proposed model improves the modelling performance and reduces the implementation complexity.

Zhai et al. [123] presented a new canonical piecewise-linear function-based model for band-limited DPDs with a structure similar to the dynamic deviation reduction Volterra series model. Their model was simpler as it contained no higher order terms and no finite impulse response filters. The authors reported linearisation performance similar to the original model.

Luo et al. [124] proposed a dual-input canonical piecewise-linear function-based model for DPDs of multi-antenna transmitters. They reported that the proposed model could mitigate the higher-order nonlinearity of the antenna crosstalk.

2.2.9 Other Works

In this section, we include selected papers that topically did not fit into the identified research directions and we consider them interesting to be mentioned in this thesis. Papers [125, 126] were focused on over-the-air feedback for DPD adaptation specifically aiming at MIMO and multi-antenna systems.

Barradas et al. [127] proposed a DPD specifically aiming at linearisation of long-term memory effects in GaN HEMTs attributed to the electron trapping phenomena. Florian et al. [128] addressed a similar problem and proposed a prepulsing technique for characterisation and measurements of GaN-based PAs.

Li et al. [129] introduced a Cartesian error feedback architecture. Their linearisation was based on a typical feedback architecture where the error signal was created by subtracting the desired signal from the PA output at the carrier frequency. For this purpose, the feedforward signal was modulated at the carrier frequency of the original complex IQ signal.

2.3 Low-Complexity Methods for DPD

2.3.1 Computational Complexity

Zhai et al. [130] simplified the DVR model [131] and presented a similar model without vector rotations. Their results show that the new model provides similar linearisation performance as the original DVR model while decreasing the computational complexity.

Wang et al. [132, 133] replaced the computationally-intensive matrix inversion $(\boldsymbol{U}_z^H \boldsymbol{U}_z)^{-1}$ in (1.27) by a precalculated inversion of the correlation matrix of the basis functions. The authors reported the same linearisation results as for the traditional approach. Additionally, they proposed an extension for dual-band applications [133].

Kelly and Zhu [134,135] proposed a low-complexity stochastic optimisation-based DPD adaptation. They directly measured the loss function for little varying DPD coefficients and by its change (its approximate derivation) they decided the update direction.

2.3.2 Model Pruning and Sizing

Gilabert et al. [136] proposed an iterative search algorithm, called the less relevant basis removal, for the model order reduction and compared it with other three reduction techniques, based on the compressed sensing, PCA, and partial LS. All the compared techniques lead to approximately the same reduction factor; however, the PCA and partial LS need the fewest iterations.

Li et al. [137] showed a model pruning method, called the regularised sparsity adaptive matching pursuit to prune MP-based models. Their method combines the adaptability of the adaptive sparsity matching pursuit [138] and the regularity of the regularised *orthogonal matching pursuit* (OMP) [139].

Wang et al. [140,141] presented a technique for optimal GMP-model sizing based on the Hill-Climbing algorithm. Their optimisation criteria are the number of coefficients and the NMSE of the model output. Additionally, they presented a single criterion combining the number of coefficients and the NMSE. The authors extended their work in [142] by a size-determining algorithm for multistage sparse models.

Peng et al. [143] presented a sparse Bayesian learning algorithm. The proposed algorithm decreased the number of model coefficients and reduced the required number of samplings.

Peng et al. [144] proposed a sparse parameter identification algorithm. They selected kernels one by one according to the projection of the residual vector onto the kernel. The authors reported that the proposed method achieves the performance of the OMP and the regularised matching pursuit while having much lower computational complexity.

Abdelhafiz et al. [145] presented a swarm optimisation with the Akaike information criterion [146] for determining the dimensions of PA models. They applied the optimisation to the GMP model and compared its results with those of simulated annealing. The obtained results showed similar performance for all the compared methods.

Becerra et al. [147] proposed a method for kernel selection based on the OMP and the Gram-Schmidt orthogonalisation. Compared with the OMP, they added one extra orthogonalisation into the algorithm, which improved the performance. They compared the results of the proposed algorithm with those of the original OMP, PCA, and the compressed-sampling matching pursuit [148]. They reported that the proposed doubly OMP achieved the best pruning performance compared with the other methods. In [149], the authors added subspace pursuit pruning and presented an overview of all the pruning methods; and in [150], they reduced the computational complexity of the doubly OMP by avoiding the calculation of the pseudoinverse matrix.

2.3.3 Real-Valued Feedback Samples

Chani-Cahuana et al. [3] proposed a DPD architecture with real-valued feedback samples. They estimated the PA forward model similarly to the procedure described in Sec. 3.2. Afterwards, they employed the same procedure for the ILC-based DPD adaptation, originally described in [58]. The authors reported the same linearisation performance of the proposed ILC with real-valued feedback samples as that of the original ILC DPD.

Guan et al. [5] presented a DLA DPD with one undersampling ADC in the feedback path. Basically, they proposed the same idea as Chani-Cahuana et al. in [3], but employed the DLA instead of the ILC and added undersampling. Both papers seem to be independent. Guan et al. reported, quite surprisingly, a higher linearisation performance of their real-valued undersampled DLA DPD than that of the conventional full-speed complex-valued DPD.

Zhang et al. [151] derived the same formula as Guan et al. [5] in a little different way and without the undersampling.

2.3.4 One-Bit and Signed-Based Feedback Samples

Wang et al. [4, 152] proposed DPD adaptation based on observing the sign of the PA output error signal. They observed the sign employing the comparators in the feedback path and, in principle, these comparators could be understood as 1-bit ADCs. The simplified diagram of their DPD architecture is depicted in Fig. 2.1.

Wang et al. modified the DLA in equation (1.27) and derived its form with the sign. We can summarise their derivation as follows: The difference $\boldsymbol{\Delta} = \boldsymbol{z} - \boldsymbol{y}$ in the original can be split into real and imaginary parts. The vector elements are

$$\Delta[n] = \Delta_{\mathrm{r}}[n] + j \cdot \Delta_{\mathrm{i}}[n], \tag{2.1}$$

where $(\cdot)_{\mathrm{r}}$ and $(\cdot)_{\mathrm{i}}$ are real and imaginary parts, respectively. By employing the fact that an arbitrary real number can be written as a multiplication of its sign and its absolute value $a = \mathrm{sign}(a) \cdot |a|$, equation (2.1) can be reformulated as

$$\Delta[n] = \mathrm{sign}(\Delta_{\mathrm{r}}[n]) \cdot |\Delta_{\mathrm{r}}[n]| + j \cdot \mathrm{sign}(\Delta_{\mathrm{i}}[n]) \cdot |\Delta_{\mathrm{i}}[n]|. \tag{2.2}$$

Since both absolute values, $|\Delta_{\mathrm{r}}[n]|$ and $|\Delta_{\mathrm{i}}[n]|$, tend to be very small and they are decreased through the DPD adaptation process, they can be replaced by an update constant μ_0 [4]

$$\Delta[n] = \mu_0 \,\mathrm{sign}(\Delta_{\mathrm{r}}[n]) + j \cdot \mu_0 \,\mathrm{sign}(\Delta_{\mathrm{i}}[n]) = \mu_0 \cdot \mathrm{sign}(\Delta[n]), \tag{2.3}$$

where $\mathrm{sign}(\Delta[n])$ calculates the signs of real and imaginary parts of $\Delta[n]$ separately.

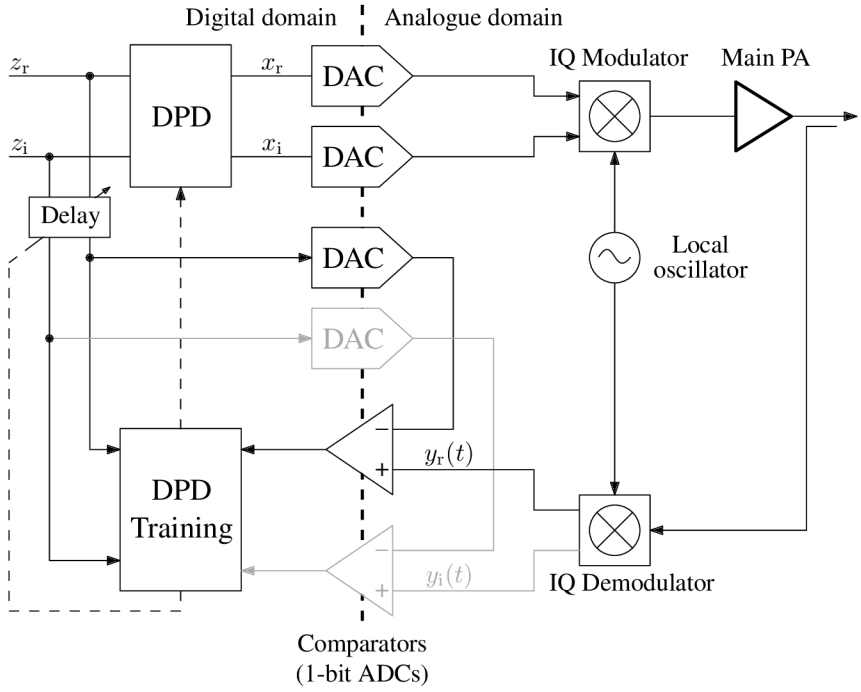


Fig. 2.1: The simplified DPD architecture with comparators in the feedback path as proposed by Wang et al. in [4, 152]. Grey-coloured parts are not required if real-valued feedback principle [153] is employed.

The final formula of the modified DLA can be stated as

$$\hat{\boldsymbol{b}} = \boldsymbol{b}' - \hat{\mu}(\boldsymbol{U}_z^H \boldsymbol{U}_z)^{-1} \boldsymbol{U}_z^H \cdot \operatorname{sign}(\boldsymbol{\Delta}), \tag{2.4}$$

where $\hat{\mu} = \mu\mu_0$ and the sign$(\cdot)$ function on the vector is defined as element-wise function.

Zhang et al. [153] adapted the real-valued feedback principle (in this thesis presented in Chapter 3) to the feedback with two comparators and simplified the architecture by employing one comparator only.

One can notice that the proposed principle in Fig. 2.1 requires two additional high-speed DACs (or one DAC if real-valued feedback is implemented [153]) which naturally complicates the design and increases the system power consumption.

Guan et al. [154] proposed a low-complexity approach for DPD adaptation employing a signed regressor algorithm (sign-based calculation) of the DLA. Although their algorithm requires complex-valued feedback samples, we consider it to be related to this topic and have included it here in this section.

Pascual Campo et al. [155] presented a comprehensive overview of sign-based algorithms for DPD adaptation. The authors compared the complexities and performance of individual algorithms as well as combinations of them.

2.3.5 Subband and Multiband Predistortion

Abdelaziz et al. [156,157] and Tarver et al. [158] presented a low-complexity DPD for subband linearisation for spectrally noncontiguous transmissions. Furthermore, they developed a decorrelation-based parameter learning solution. The authors reported more than 10 times the reduced computational complexity of the proposed DPD compared with the conventional ones.

Mkadem et al. [159] expounded a complexity-reduced GMP model for linearisation of multiband PAs. The authors compared the proposed models with dual-input and triple-input DPD models and reported similar linearisation performance with a reduced number of coefficients.

Zhang and Chen in [160] presented the DPD adapted by multiple band-limited feedback signals. Their solution estimated the PA forward model based on the incomplete feedback signals and, by employing the forward model, they finally extracted the coefficients of the predistorter. Zhang et al. [161] implemented the dual-band PA linearisation employing reduced LUTs.

Jaraut et al. [162] proposed an independent component analysis method for the DPD models employed in carrier aggregation scenarios. The authors reported reduced requirements on the bit resolution of calculations and on the FPGA memory. Additionally, the proposed method led to improved numerical stability of the DPD solution.

2.3.6 Undersampling Methods

Huang et al. [163] proposed the DPD adaptation with undersampled feedback. They employed the DLA and kept every n-th feedback sample. They noted that the feedback must be sampled with sufficient bandwidth. Wang et al. [164] presented a similar approach with the undersampled feedback, but they estimated the PA forward model and employed this model to adapt the DPD by the ILA. Beltagy et al. [165] extended the work [163] by deriving the DPD adaptation formulas for the undersampled feedback sampling at an intermediate frequency. The authors claimed the intermediate frequency sampling could be beneficial as it does not suffer from IQ imbalances. The adapted DPD achieved similar performance to its full-rate equivalent.

Prata et al. [166] presented feedback loops employing RF subsampling ADCs to improve concurrent dual-band transmitter linearisation. The aliasing between upper and lower bands could naturally occur which the authors compensated for based on statistical approximated nonoverlapped multisines. The authors reported similar performance to the other subsampling techniques but with a lower hardware complexity.

Chung et al. [167] proposed methods for joint mitigation of hardware imperfections, especially IQ imbalances and nonlinear distortion caused by the IQ modulator and PA. They estimated both the imbalances and nonlinear distortion based on observing the undersampled feedback.

Li et al. [168] focused on time-interleaved DPDs to reduce their sampling rates. The authors proposed a few architectures with time-interleaved DPDs and solved the aliasing effect. Their proposed low-speed DPD was naturally narrow band and compensated only for the distortion close to the communication band. In that region, the authors reported performance similar to that of the full high-speed DPDs.

3 Real-Valued Feedback

3.1 Introduction

In this chapter, we describe a simplification of the feedback circuitry for the DPD adaptation by employing only the in-phase or quadrature output of the IQ down-converting mixer. The in-phase feedback is depicted in a block diagram in Fig. 3.1. The main motivation is to avoid one feedback ADC, which is one of the main contributors to the total system power consumption. Alongside, the feedback complexity is reduced. Under certain circumstances, the DPD adaptation with the real-valued feedback samples might be less susceptible to IQ imbalances [A9].

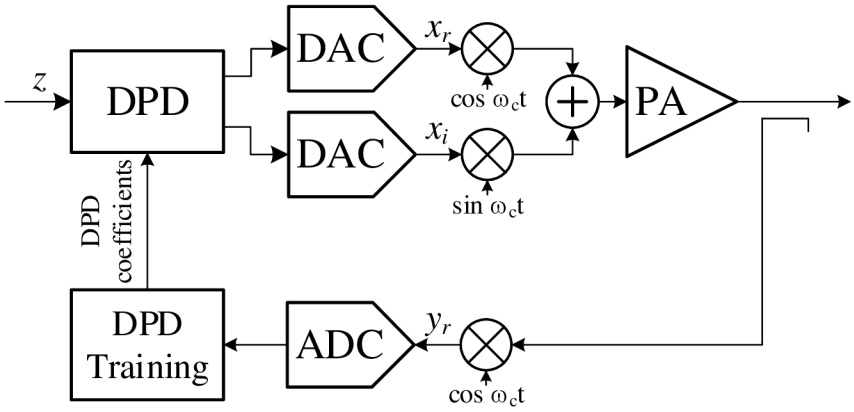


Fig. 3.1: The system diagram for the DPD with real-valued feedback.

Our contribution has been presented in the paper [A3] and is the natural evolution of methods described by Chani-Cahuana et al. in [3] and by Guan et al. in [5]. In this thesis, both works are described in Sec. 2.3.3. All Matlab source codes related to this chapter are provided at `www.github.com/jankralx/rfm_ila`.

We must mention papers by Yu et al. [169] and Zhang et al. [170] which proposed a similar idea to our paper [A3] presented here. Both works were presented later in the same year, independently of our work.

3.2 Forward Model Estimation

For simplicity, we start our derivation of the real-valued feedback with the MP model. However, the procedure is applicable to most of the currently well-known PA or DPD models. We can rewrite equation (1.14) with the real and imaginary parts, denoted as $(\cdot)_\mathrm{r}$ and $(\cdot)_\mathrm{i}$, respectively, as

$$\boldsymbol{y}_\mathrm{r} + j\boldsymbol{y}_\mathrm{i} = (\boldsymbol{U}_{x\mathrm{r}} + j\boldsymbol{U}_{x\mathrm{i}})(\boldsymbol{b}_\mathrm{r} + j\boldsymbol{b}_\mathrm{i}). \tag{3.1}$$

Expanding the multiplication on the right-hand side yields

$$\boldsymbol{y}_{\mathrm{r}} + j\boldsymbol{y}_{\mathrm{i}} = \boldsymbol{U}_{x\mathrm{r}}\boldsymbol{b}_{\mathrm{r}} + j\boldsymbol{U}_{x\mathrm{i}}\boldsymbol{b}_{\mathrm{r}} + j\boldsymbol{U}_{x\mathrm{r}}\boldsymbol{b}_{\mathrm{i}} - \boldsymbol{U}_{x\mathrm{i}}\boldsymbol{b}_{\mathrm{i}}, \tag{3.2}$$

which can be split into two systems of equations, one for the real part of $\boldsymbol{y}$ and the other one for the imaginary part of $\boldsymbol{y}$

$$\boldsymbol{y}_{\mathrm{r}} = \boldsymbol{U}_{x\mathrm{r}}\boldsymbol{b}_{\mathrm{r}} - \boldsymbol{U}_{x\mathrm{i}}\boldsymbol{b}_{\mathrm{i}}, \tag{3.3}$$

$$\boldsymbol{y}_{\mathrm{i}} = \boldsymbol{U}_{x\mathrm{i}}\boldsymbol{b}_{\mathrm{r}} + \boldsymbol{U}_{x\mathrm{r}}\boldsymbol{b}_{\mathrm{i}}. \tag{3.4}$$

To obtain the PA coefficient vector $\boldsymbol{b}$, it is sufficient to solve only one of the two systems of equations in (3.3) and (3.4). Advantageously, each system of equations requires only the real or imaginary part of the feedback signal $\boldsymbol{y}$. In real hardware, the real part of the feedback signal is equivalent to the in-phase of the feedback quadrature mixer and similarly, the imaginary part is equivalent to the quadrature output.

Matrix $\boldsymbol{U}_x$ is fully known, as it consists of samples given by the transmitted signal x. Therefore, we can obtain $\boldsymbol{b}$ as the LS solution of eq. (3.3) with the real feedback samples

$$\begin{bmatrix}\boldsymbol{b}_{\mathrm{r}}\\ \boldsymbol{b}_{\mathrm{i}}\end{bmatrix} = (\boldsymbol{A}_x^H\boldsymbol{A}_x)^{-1}\boldsymbol{A}_x^H\boldsymbol{y}_{\mathrm{r}}, \tag{3.5}$$

or as the LS solution of eq. (3.4) with the imaginary feedback samples

$$\begin{bmatrix}\boldsymbol{b}_{\mathrm{r}}\\ \boldsymbol{b}_{\mathrm{i}}\end{bmatrix} = (\boldsymbol{B}_x^H\boldsymbol{B}_x)^{-1}\boldsymbol{B}_x^H\boldsymbol{y}_{\mathrm{i}}, \tag{3.6}$$

where we have substituted the real and imaginary matrix $\boldsymbol{U}_x$ by matrices $\boldsymbol{A}_x$ and $\boldsymbol{B}_x$, defined as

$$\boldsymbol{A}_x = \begin{bmatrix}\boldsymbol{U}_{x\mathrm{r}} & -\boldsymbol{U}_{x\mathrm{i}}\end{bmatrix}, \tag{3.7}$$

$$\boldsymbol{B}_x = \begin{bmatrix}\boldsymbol{U}_{x\mathrm{i}} & \boldsymbol{U}_{x\mathrm{r}}\end{bmatrix}. \tag{3.8}$$

In this way, we have solved the coefficients of the PA forward model. However, these coefficients cannot be directly recomputed into the DPD coefficients. One of the methods allowing for calculating the DPD coefficients employing the known forward model is the method developed by Morgan et al. [25] and Landin et al. [26].

3.3 Indirect Learning Architecture

Direct derivation of the ILA with real-valued feedback cannot be achieved. The ILA takes the feedback samples and puts them into matrix $\boldsymbol{U}_y$ on the right-hand side of

the system of equations and the PA input on the left-hand side. In fact, the PA input and output are swapped compared with the PA forward model. However, we can benefit from [25, 26] and employ the FM-ILA with real-valued feedback.

The equations with the real-valued feedback are similar to the original complex feedback equations described in Sec. 1.4.3 with PA model coefficients, originally calculated using equation (1.32), and calculated with real feedback samples by eq. (3.5). Hereinafter, we will refer to the proposed method as the *real-valued forward-model indirect learning architecture* (R-FM-ILA).

3.4 Direct Learning Architecture

Let us continue with the development started in Sec. 1.4.1. We can split eq. (1.25) into the real and imaginary parts as

$$\boldsymbol{\Delta}_{\mathrm{r}} + \mathrm{j}\boldsymbol{\Delta}_{\mathrm{i}} = (\boldsymbol{U}_{z\mathrm{r}} + \mathrm{j}\boldsymbol{U}_{z\mathrm{i}})(\boldsymbol{e}_{\mathrm{r}} + \mathrm{j}\boldsymbol{e}_{\mathrm{i}}). \tag{3.9}$$

By expanding the multiplication on the right-hand side, we get two equation systems (similarly as for the forward model output as described by eq. (3.3) and (3.4)

$$\boldsymbol{\Delta}_{\mathrm{r}} = \boldsymbol{U}_{z\mathrm{r}}\boldsymbol{e}_{\mathrm{r}} - \boldsymbol{U}_{z\mathrm{i}}\boldsymbol{e}_{\mathrm{i}}, \tag{3.10}$$

$$\boldsymbol{\Delta}_{\mathrm{i}} = \boldsymbol{U}_{z\mathrm{i}}\boldsymbol{e}_{\mathrm{r}} + \boldsymbol{U}_{z\mathrm{r}}\boldsymbol{e}_{\mathrm{i}}. \tag{3.11}$$

Vector $\boldsymbol{e}$ can be solved from either the system of equations (3.10) requiring only the real part of the error vector $\boldsymbol{\Delta}$ (consequently only the real feedback samples are required) as

$$\begin{bmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_i \end{bmatrix} = (\boldsymbol{A}_z^H \boldsymbol{A}_z)^{-1} \boldsymbol{A}_z^H \boldsymbol{\Delta}_r \tag{3.12}$$

or the system of equations (3.11) requiring only the imaginary feedback samples as

$$\begin{bmatrix} \boldsymbol{e}_r \\ \boldsymbol{e}_i \end{bmatrix} = (\boldsymbol{B}_z^H \boldsymbol{B}_z)^{-1} \boldsymbol{B}_z^H \boldsymbol{\Delta}_i, \tag{3.13}$$

where $\boldsymbol{A}_z$ was defined in eq. (3.7) and $\boldsymbol{B}_z$ in eq. (3.8). Back substitution of vector $\boldsymbol{e}$ into eq. (1.24) yields the solution for the updated DPD coefficients

$$\begin{bmatrix} \hat{\boldsymbol{b}}_r \\ \hat{\boldsymbol{b}}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}'_r \\ \boldsymbol{b}'_i \end{bmatrix} - \mu(\boldsymbol{A}_z^H \boldsymbol{A}_z)^{-1} \boldsymbol{A}_z^H (\boldsymbol{z}_r - \boldsymbol{y}_r). \tag{3.14}$$

Hereinafter, the DLA with real feedback samples is referred to as the *real-valued direct learning architecture* (R-DLA).

3.5 Simulation

We have implemented and simulated DPD architectures with real-valued feedback (R-DLA, R-FM-ILA) to evaluate their linearisation performance. We have performed simulations for conventional architectures with complex feedback (ILA, DLA, FM-ILA) and for the system without the DPD to compare all of them with the recently introduced R-FM-ILA and R-DLA.

For all the architectures, the PA has been modelled by the MP model (eq. 1.12) with $K = 7, Q = 1$. We have extracted the PA model coefficients from measurements of a real PA. The PA has been excited with an FBMC signal with 1024 subcarriers, 18 frames in each iteration. The sampling frequency has been set six times higher than the main channel bandwidth B.

We have simulated all the DPD architectures iteratively. This means that the DPD coefficients have been trained on a signal different from the signal used for evaluating the DPD linearisation performance. In this way, the evaluation is more objective, as it represents more closely a real transmitter with the DPD.

For all the architectures, one iteration consists of:

1. generating the random-data FBMC signal z (same signal for all the DPD architectures),
2. predistorting the desired signal z with current DPD coefficients $\boldsymbol{b}'$ yielding the PA input x,
3. adjusting the signal amplitude to set the desired PA output power,
4. calculating the PA model output y,
5. evaluating the linearisation performance,
6. and calculating the new DPD coefficients $\hat{\boldsymbol{b}}$.

The important step is setting the PA reference gain. The maximum signal amplitude in the system without DPD was set such that the PA provided the same output power in the communication channel as the system with the ILA. The detailed procedure for setting the PA reference gain is described in the work of Jardin and Baudoin [77,78]. The calculation of the new DPD coefficients (last step) varies based on the DPD architecture.

The procedure for the ILA consists of:

1. adjusting the maximum Euclidean norm of signal x and y such as $||x|| \leq 1$, $||y|| \leq 1$ for the DPD coefficient calculation,
2. solving coefficients $\hat{\boldsymbol{b}}$ of the postdistorter as described by (1.31),
3. using the calculated coefficients of the post-distorter as the DPD coefficients in the next iteration.

The procedure for the DLA consists of:

1. adjusting the maximum Euclidean norm of z and y as $||z|| \leq 1$, $||y|| \leq 1$ for updating the DPD coefficient,
2. updating the DPD coefficients due to the desired signal z and feedback y as described by eq. (1.27).

One can achieve the fastest convergence if the initial DPD coefficients are set close to the optimum. We have suggested the optimum DPD coefficients be unknown in the first iteration. Generally, we have achieved a reasonable convergence speed with the initial coefficients set to $\boldsymbol{b}' = \begin{bmatrix} 0.5 & 0 & 0 & \dots \end{bmatrix}^T$. This procedure is slightly modified for the R-DLA as described in Sec. 3.4.

Finally, the procedure for FM-ILA consists of

1. adjusting the maximum Euclidean norm of x and y as $||x|| \leq 1$, $||y|| \leq 1$ for the following calculations,
2. estimating the forward PA model coefficients,
3. calculating the forward model output,
4. solving coefficients $\hat{\boldsymbol{b}}$ of the post-distorter,
5. employing the calculated postdistorter coefficients $\hat{\boldsymbol{b}}$ as the DPD coefficients in the next iteration.

This procedure is slightly modified for the R-FM-ILA as described in Sec. 3.3.

3.6 Simulation Results

The linearisation performance has been evaluated based on the NMSE and ACPR. We have simulated all architectures with the DPD modelled by the MP model with the maximum nonlinearity order of $K = 7$ and the maximum memory length of $Q = 3$. In the simulation, all the architectures have been computed with 220 iterations. The evaluated metrics were averaged over the last 200 iterations, which corresponds to the range where the DLA and R-DLA converged.

Fig. 3.2 shows the AM/AM characteristics for the R-FM-ILA with the trained DPD coefficients. The black circles represent the characteristics of the PA obtained from the measurement of the feedback signal, and the orange dots represent the characteristics of the estimated forward PA model. One can see that the PA shows a certain memory effect which is mostly compensated for by the predistorter.

The frequency spectra of the PA output are shown in Fig. 3.3. The spectra for all the DPD architectures are almost the same and well improved compared to the spectrum for the system without the DPD.

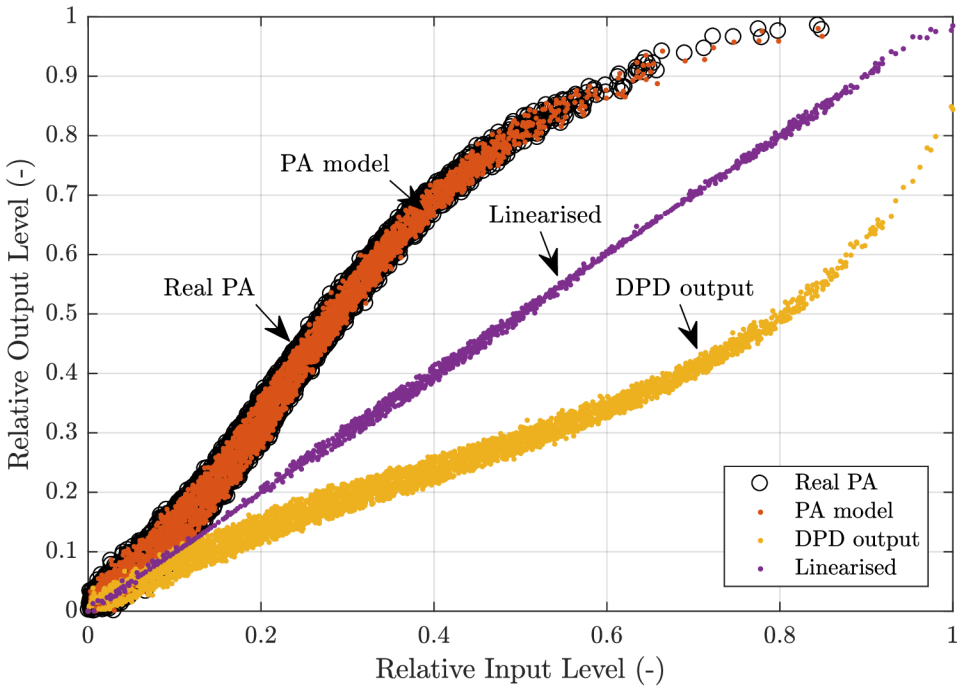


Fig. 3.2: The AM/AM characteristics with the R-FM-ILA DPD.

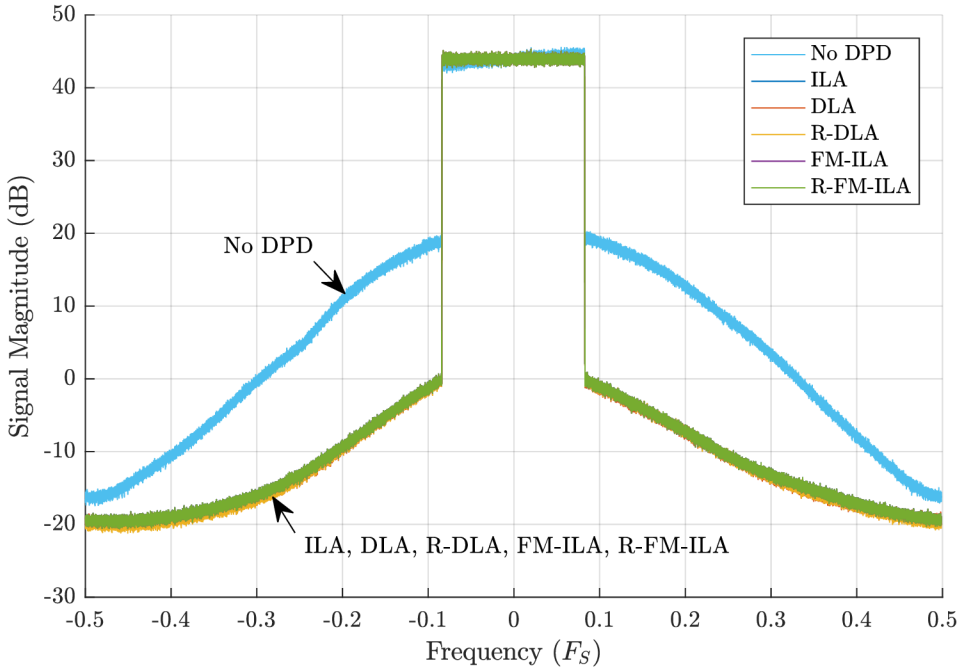


Fig. 3.3: The frequency spectra of the PA output for all simulated architectures. The frequency axis is normalised to the sampling frequency F_s.

The evolution of the NMSE in a few first iteration cycles is depicted in Fig. 3.4. At the beginning of the simulation, ILA, FM-ILA, and R-FM-ILA are not trained yet and therefore provide the same NMSE as the system without the DPD. On the other hand, DLA and R-DLA start with a lower output power due to the first DPD coefficient being set to 0.5 which results in a lower NMSE. In the second iteration, ILA, FM-ILA and R-FM-ILA achieve almost the optimum predistortion. The output power of the system without the DPD is lowered to be the same as for the system with ILA. Note the peak in the NMSE for DLA and R-DLA which the optimisation process goes through to achieve the solution. The convergence speed of the DLA and R-DLA highly depends on the initial coefficients, step size μ, and the desired signal z. Higher convergence speeds could be achieved by increasing the step size μ, but the convergence probability would be lower.

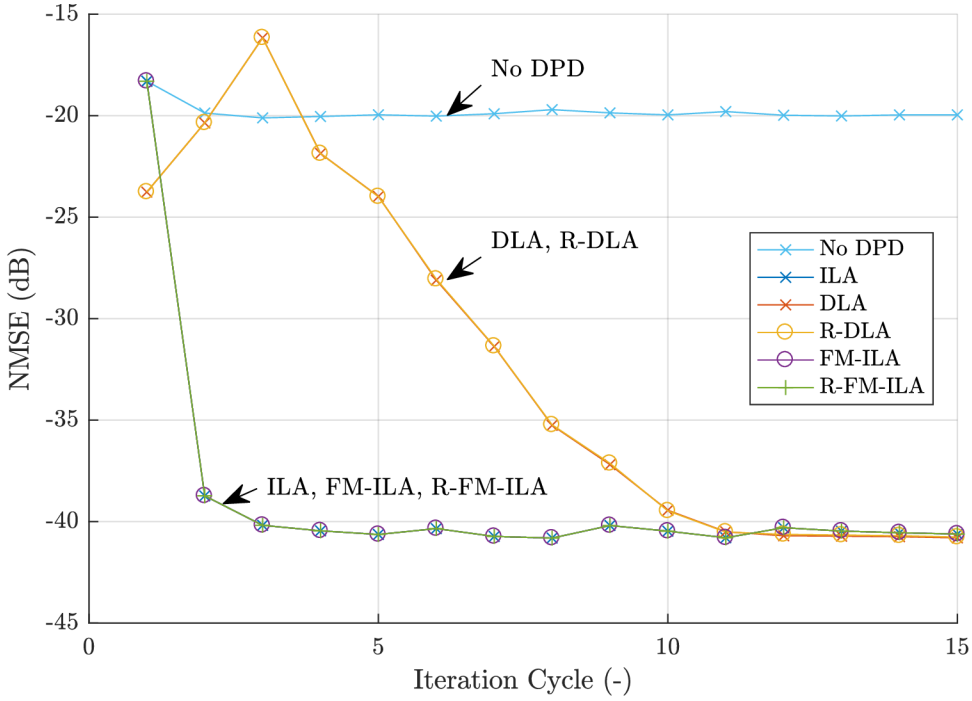


Fig. 3.4: Evolution of the NMSE in a few first iteration cycles.

The detailed comparison of linearisation performance for all the architectures is given in Tab. 3.1. All the ILA-based systems provide very similar linearisation performance based on the evaluated metrics. Again, the DLA and R-DLA provide very similar results, but they are slightly better than the ILA-based systems.

Tab. 3.1: The linearisation performance of systems with real-valued feedback compared with conventional architectures.

DPD architecture	NMSE (dB)	ACPR-1st (dB)	ACPR-2nd (dB)
No DPD	-19.9	-29.6	-46.5
ILA	-40.5	-49.6	-60.3
DLA	-40.8	-49.8	-60.7
R-DLA	-40.7	-49.7	-60.7
FM-ILA	-40.5	-49.6	-60.3
R-FM-ILA	-40.5	-49.6	-60.3

3.7 Conclusion

In this chapter, we have presented a DPD architecture with real-valued feedback samples and employing the forward PA model estimation. The presented method has been verified and compared to state-of-the-art DPD methods in simulations. We have shown that the R-FM-ILA can achieve the same linearisation performance as its complex variant and the ILA DPD and very similar results as complex DLA and R-DLA. The main advantage of the presented R-FM-ILA over the DLA and R-DLA is noniterative calculation and the related unnecessary setting of the initial solution and step size. Additionally, it allows the employment of a simple RF mixer and one ADC instead of a quadrature mixer with two ADCs which decreases power consumption, system complexity, and the transmitter price.

4 DPD Adaptation with Sample Selection

4.1 Introduction

In this chapter, we extend methods for undersampling and sample selection [5, 163, 164, 171–176] and we propose methods for the selection of samples for DPD adaptation allowing computational complexity reduction. The block diagram for DPD adaptation with sample selection is shown in Fig. 4.1.

The proposed methods are: a method based on the identification of important samples using QR decomposition [177], a gradient-based sampling method [178], and two histogram-based methods. The first histogram method equalises the histogram of signal magnitudes to ensure evenly sampled PA characteristics while the second one optimises a histogram optimised by a genetic algorithm, which respects both

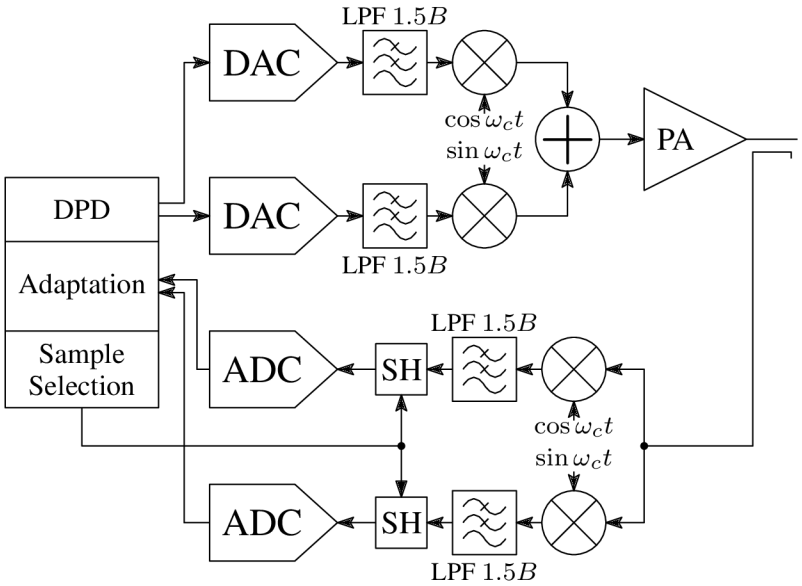


Fig. 4.1: Linearisation of an RF PA with the DPD using an optional sample selection method. In the direct path, DACs are accompanied by reconstruction *low-pass filters* (LPFs) with bandwidth $1.5B$, where B is the desired channel bandwidth. In the feedback, there are LPFs with bandwidth $> 1.5B$ to limit the noise bandwidth, optional *sample and hold* (SH) circuits and ADCs. Note that the resulting bandwidth of the forward and feedback paths is $3B$ due to IQ sampling. Although the depicted SH circuits as discrete components are one possible approach, a more practical implementation would employ ADCs with integrated SH circuits and with periodic and equidistant sampling, performing sample selection from a sample buffer in the digital domain.

the transmitted signal statistical properties and the specific PA characteristics. All these methods are analysed with respect to the conventional DPD and to the simple undersampling method [5, 174, 175]. Moreover, we provide a detailed analysis of the computational complexity of these methods and show how significantly some of them can reduce the required computational resources. The performance of the proposed sample selection methods is shown in simulations and eventually verified by measurements.

The contributions of this chapter have been originally presented in our paper [A2] and include

- a proposal of methods for sample selection for DPD adaptation compatible with the conventional DPD-enabled transmitter architectures, especially the histogram-based method with the histogram optimised by a genetic algorithm, which highly reduces the computational complexity while preserving the linearisation capabilities,
- a comparison of the computational complexity of the proposed methods,
- and all related Matlab source codes available at `www.github.com/jankralx/dpd_sample_selection`.

The rest of this chapter is organised as follows. We introduce the theory behind the sample selection for DPD adaptation in Sec. 4.2. Sample selection methods are proposed and discussed in Sec. 4.5. We show how the proposed sample selection methods reduce the computational and hardware resources needed for DPD adaptation in Sec. 4.6 and Sec. 4.6.5. Finally, the simulations are described in Sec. 4.7 and the measurements in Sec. 4.8.

4.2 Basic Principles of Sample Selection

In this chapter, we start our description of the basic principles of undersampling in DPD systems on the problem of PA-model identification and later we develop the basic idea also for the ILA, FM-ILA, and the DLA. Let us assume that the PA is modelled by the MP baseband model. We have chosen the MP model for its formal simplicity, but all the presented concepts can be generalised for any model which is linear in its unknown parameters. To identify the unknown coefficients $b_{k,q}$ of the PA model, described by equation (1.11), the conventional methods construct a system of equations by taking consecutive input and output baseband samples of the PA and inserting them into the system of equations (1.14).

Let us supplement the theory of the undersampling in DPD systems with a fictitious example of the MP model with the maximum nonlinearity order $K = 2$ and the maximum memory length $Q = 1$, which have $P = K \cdot (Q+1) = 4$ coefficients.

For completeness, we start the example with the conventional method. The system of equations in the vector form can be constructed as

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \\ y[3] \\ y[4] \\ y[5] \\ y[6] \\ y[7] \end{bmatrix} = \begin{bmatrix} x[0] & x[-1] & x[0]|x[0]| & x[-1]|x[-1]| \\ x[1] & x[0] & x[1]|x[1]| & x[0]|x[0]| \\ x[2] & x[1] & x[2]|x[2]| & x[1]|x[1]| \\ x[3] & x[2] & x[3]|x[3]| & x[2]|x[2]| \\ x[4] & x[3] & x[4]|x[4]| & x[3]|x[3]| \\ x[5] & x[4] & x[5]|x[5]| & x[4]|x[4]| \\ x[6] & x[5] & x[6]|x[6]| & x[5]|x[5]| \\ x[7] & x[6] & x[7]|x[7]| & x[6]|x[6]| \end{bmatrix} \begin{bmatrix} b_{0,0} \\ b_{0,1} \\ b_{1,0} \\ b_{1,1} \end{bmatrix}. \tag{4.1}$$

Here we would like to point out that the conventional methods take consecutive samples of the PA input $x[n]$ and the PA output $y[n]$ to construct the system of equations. Practically, to solve the PA model coefficients, the system of equations has to be overdetermined to mitigate the imperfections of the PA input generation and of the PA output measurement. One can figure out that if there are no imperfections, every arbitrary combination of P rows from (4.1) leads to the solution of the model coefficients. We can benefit from a similar principle with the overdetermined system and eliminate arbitrary rows in the system of equations. The solution of the system is not changed significantly until the system contains a sufficient number of equations. Practically, the number of equations can be drastically reduced if correct samples (equations) are selected, as we will demonstrate later. The reduced system of equations could look like

$$\begin{bmatrix} \cancel{y[0]} \\ y[1] \\ y[2] \\ \cancel{y[3]} \\ \cancel{y[4]} \\ y[5] \\ y[6] \\ y[7] \end{bmatrix} = \begin{bmatrix} \cancel{x[0]} & \cancel{x[-1]} & \cancel{x[0]|x[0]|} & \cancel{x[-1]|x[-1]|} \\ x[1] & x[0] & x[1]|x[1]| & x[0]|x[0]| \\ x[2] & x[1] & x[2]|x[2]| & x[1]|x[1]| \\ \cancel{x[3]} & \cancel{x[2]} & \cancel{x[3]|x[3]|} & \cancel{x[2]|x[2]|} \\ \cancel{x[4]} & \cancel{x[3]} & \cancel{x[4]|x[4]|} & \cancel{x[3]|x[3]|} \\ x[5] & x[4] & x[5]|x[5]| & x[4]|x[4]| \\ x[6] & x[5] & x[6]|x[6]| & x[5]|x[5]| \\ x[7] & x[6] & x[7]|x[7]| & x[6]|x[6]| \end{bmatrix} \begin{bmatrix} b_{0,0} \\ b_{0,1} \\ b_{1,0} \\ b_{1,1} \end{bmatrix}. \tag{4.2}$$

Advantageously, the signal on the left-hand side of (4.2), in this case signal $y[n]$, can be arbitrarily undersampled. This allows for reducing the hardware requirements, e.g., the sampling rate of the feedback ADCs. It is important to note that the Nyquist-Shannon sampling theorem [179, 180] is not violated by reducing the rows as long as we do not need to fully recover the signal waveform. One needs to realise that filtering and other standard signal processing methods are not achievable if the signal is undersampled.

Formally, we can construct the system of equations by taking the signal samples $y[n_i]$ with arbitrary sample indices $n_i \in \mathbb{N}; i \in \{1, 2, ..., N\}$ which satisfy the condition $n_i \neq n_\kappa; \forall i, \kappa : i \neq \kappa$. The reduced equation system can be expressed in the matrix form by arranging the output samples, model coefficients and basis functions into vectors, i.e.,

$$\begin{aligned} \boldsymbol{\phi}_{k,q}^{(x)} &= \begin{bmatrix} \phi_{k,q}^{(x)}[n_1] & \phi_{k,q}^{(x)}[n_2] & \dots & \phi_{k,q}^{(x)}[n_N] \end{bmatrix}^T, \\ \boldsymbol{y} &= \begin{bmatrix} y[n_1] & y[n_2] & \dots & y[n_N] \end{bmatrix}^T, \\ \boldsymbol{b} &= \begin{bmatrix} b_{1,0} & b_{1,1} & \dots & b_{1,Q} & b_{2,0} & \dots & b_{P,Q} \end{bmatrix}^T, \\ \boldsymbol{U}_x &= \begin{bmatrix} \boldsymbol{\phi}_{1,0}^{(x)} & \boldsymbol{\phi}_{1,1}^{(x)} & \dots & \boldsymbol{\phi}_{1,Q}^{(x)} & \boldsymbol{\phi}_{2,0}^{(x)} & \dots & \boldsymbol{\phi}_{K,Q}^{(x)} \end{bmatrix}, \end{aligned} \tag{4.3}$$

where $\boldsymbol{b}$ is a column vector with P rows, and the size of the matrix $\boldsymbol{U}_x$ is $N \times P$. The system of equations is then expressed as

$$\boldsymbol{y} = \boldsymbol{U}_x \boldsymbol{b}. \tag{4.4}$$

By solving (4.4), we obtain the model coefficients $\boldsymbol{b}$ and hence identify the PA model. Typically, we choose $N \gg P$ to sufficiently sample the whole characteristics of the PA with diverse signal samples and to mitigate the influence of noise and other imperfections of a real system. The solution can be obtained as a projection of $\boldsymbol{b}$ into the column space of $\boldsymbol{U}_x$, in other words it is the LS solution which minimises the difference between the observed PA output and the model output

$$\boldsymbol{b} = (\boldsymbol{U}_x^H \boldsymbol{U}_x)^{-1} \boldsymbol{U}_x^H \boldsymbol{y}. \tag{4.5}$$

We can also imagine the above procedure such that we first take consecutive samples of the PA output, build a conventional overdetermined system of equations and before solving it, we leave out some arbitrary rows from matrix $\boldsymbol{U}_x$ and the equivalent samples from vector $\boldsymbol{y}$. We would like to emphasise that the reduction of the system of equations preserves the memory modelling, because equation (1.11) is still fully valid. Please note that only the PA output $y[n]$, on the left-hand side of the system in (4.4), can be arbitrarily undersampled. The PA input $x[n]$ has to be known, to cover at least the memory modelling of the PA to build matrix $\boldsymbol{U}_x$.

Hereinafter, the described method is referred to as the *sample selection method* (SSM). In the following sections, we show some possible ways of employing SSMs in the DPD identification process.

4.3 Sample Selection and ILA

In the ILA, the measured feedback samples fill the matrix $\boldsymbol{U}_y$. Consequently, the advantage of arbitrarily undersampled feedback disappears and the usage of SSM is

limited. However, we can benefit from the solution of the suboptimal ILA provided by Morgan [25] and Landin [26] (described in Sec. 1.4.3) which first estimates the forward PA model. The proposed SSM can be employed for the PA model estimation and the DPD coefficients are calculated by the ILA in the standard way. The block diagram of the method is depicted in Fig. 4.2.

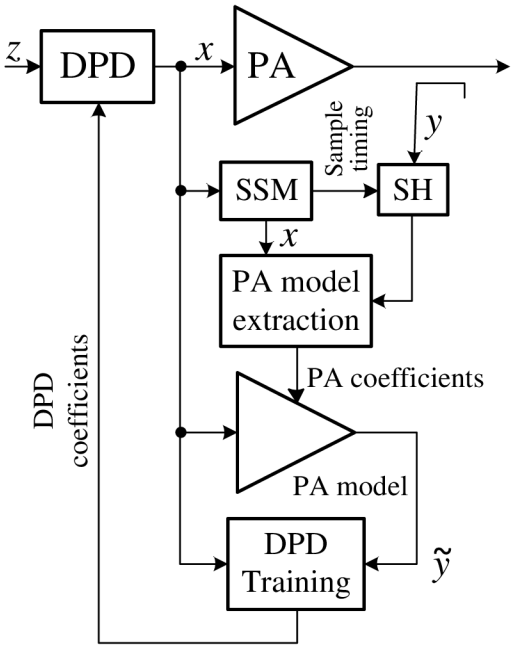


Fig. 4.2: Block diagram of a transmitter with a baseband PA model linearised by DPD with FM-ILA. The PA output is sampled by a *sample and hold* (SH) circuit at times driven by SSM.

4.4 Sample Selection and DLA

We assume that equation (1.25) is originally overdetermined, therefore we can omit arbitrary rows of matrix $\boldsymbol{U}_z$ and the equivalent samples of vector $\boldsymbol{\Delta}$, similarly as in Sec. 4.2 for the PA model estimation. The reduced vector and matrix can be denoted as

$$\begin{aligned} \boldsymbol{z} &= \begin{bmatrix} z[n_1] & z[n_2] & \dots & z[n_S] \end{bmatrix}^T, \\ \boldsymbol{U}_z &= \begin{bmatrix} \boldsymbol{\phi}_{1,0}^{(z)} & \boldsymbol{\phi}_{1,1}^{(z)} & \dots & \boldsymbol{\phi}_{1,Q}^{(z)} & \boldsymbol{\phi}_{2,0}^{(z)} & \dots & \boldsymbol{\phi}_{K,Q}^{(z)} \end{bmatrix}. \end{aligned} \tag{4.6}$$

The notation of the final solution employing SSM does not differ from the conventional DLA in eq. (1.27) and can be written as

$$\hat{\boldsymbol{b}} = \boldsymbol{b}' - \mu(\boldsymbol{U}_z^H \boldsymbol{U}_z)^{-1} \boldsymbol{U}_z^H (\boldsymbol{z} - \boldsymbol{y}). \tag{4.7}$$

The block diagram of DLA is depicted in Fig. 4.3. Please note that SSMs allow for arbitrary undersampling of only the PA output y. The desired scaled PA output z must be sufficiently known to construct the matrix $\boldsymbol{U}_z$. This usually does not represent any complication, because the desired scaled PA output is also the DPD input. In the rest of this chapter, we focus our analysis and findings primarily on the DLA, because we find the DLA more efficient and usable than the FM-ILA.

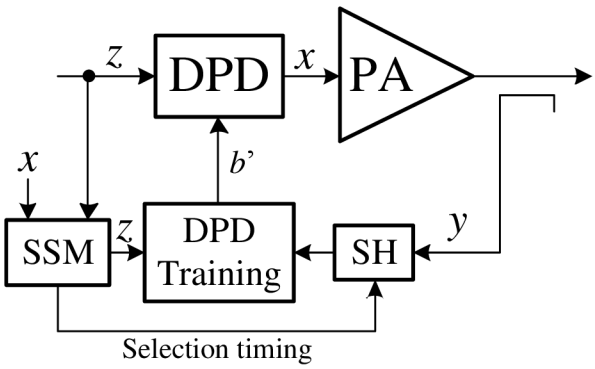


Fig. 4.3: Block diagram of the transmitter with a baseband PA model linearised by the DPD with the DLA. The PA output is sampled by the SH circuit at times driven by SSM based on the PA input samples x.

4.5 Sample Selection Methods

The best DPD adaptation can be achieved if the observation errors of the input and output are uncorrelated [181]. However, the adjacent samples used by the conventional DPD are not independent, and therefore the observation errors are correlated. To minimise the observation errors, the conventional methods without sample selection require a high number of samples N, usually $N > 1000$ [181]. For a small N, the subsequent samples cause the system of equations to be ill-conditioned. Furthermore, a limited number of subsequent samples cannot cover the statistical properties of the transmitted signal.

We show that the proposed SSM does not suffer from the mentioned drawbacks for small N if the samples $n_1, n_2, ..., n_N$ are selected carefully. The problem for the sample selection method can be defined as the selection of N samples from all samples which were acquired by the feedback ADCs. The number of all acquired samples is N_0 and corresponds to the acquisition time and hence to the required update rate of DPD coefficients. The number of selected samples is naturally limited by the condition $N \leq N_0$.

4.5.1 Undersampling

One possible approach to SSM is a simple undersampling of the feedback signal. It can be seen as an untargeted sample selection and henceforth it is referred to as the *undersampling-based sample selection* (US). Similar approaches were presented in [5, 174,175]. Although this method is very simple and does not consider the statistics of the transmitted signal, it can improve the conditioning of the system of equations, as it takes distant samples which have the potential to be more independent. Therefore, this simple approach can be sufficient for less demanding applications.

4.5.2 Sample Selection Based on QR Decomposition

The problem of selecting the samples $n_1, n_2, ..., n_N$ for the calculation of new DPD coefficients can be solved analytically using QR decomposition with column pivoting [177], hereinafter referred to as *QR-decomposition-based sample selection* (QRS). First we construct the matrix $\boldsymbol{U}_z$ using all the samples from the feedback. Afterwards we perform the QR decomposition with column pivoting of the transposed matrix $\boldsymbol{U}_z$, which is defined as

$$\boldsymbol{U}_z^T \boldsymbol{E} = \boldsymbol{Q}\boldsymbol{R}, \tag{4.8}$$

where $\boldsymbol{R}$ is an $N_0 \times P$ upper triangular matrix, $\boldsymbol{Q}$ is an $N_0 \times N_0$ unitary matrix, and $\boldsymbol{E}$ is an $N_0 \times N_0$ permutation matrix. The first N columns of the permutation matrix identify the N most important rows of matrix $\boldsymbol{U}_z$ for the solution of new DPD coefficients. Formally, we can write that the sample indexes of selected samples are given as

$$\left[n_1, n_2, ..., n_N\right] = \left[1, 2, ..., N_0\right] \boldsymbol{E}_{1:N}, \tag{4.9}$$

where $\boldsymbol{E}_{1:N}$ is a submatrix of matrix $\boldsymbol{E}$ consisting of its first N columns.

The drawback of this method is that it improves the conditioning of the pseudoinverse of matrix $\boldsymbol{U}_z$, but this does not guarantee improved DPD linearisation performance, because this method does not respect vector $\boldsymbol{\Delta}$ in the calculation of the DPD coefficients in (4.7) nor the PA characteristics.

Moreover, the practical usability of this method for real-time DPD adaptation is limited, because it requires performing the QR decomposition of the full matrix $\boldsymbol{U}_z$, which in principle already solves the system of equations using all feedback samples. Potential usage is limited to cases where a fast QR decomposition with low precision [182] is applied to identify a few important samples and thus requires fewer computation resources and high precision calculation is performed by only the selected samples.

4.5.3 Gradient Sampling

Another method for sample selection can be the *gradient-based sampling* (GS) [178], henceforth referred to as *GS-based sample selection* (GSS). GS is a representative of the methods for solving least squares problems of a large sample size. For DLA, we calculate the gradient values g_i for all rows of matrix $\boldsymbol{U}_z$ and all samples of vector $\boldsymbol{\Delta}$ [178]

$$g_i = ||\boldsymbol{u}_i(\Delta_i - \boldsymbol{u}_i\boldsymbol{e}_0)||, \tag{4.10}$$

where $\boldsymbol{u}_i$ is the i-th row of matrix $\boldsymbol{U}_z$, Δ_i is the i-th element of vector $\boldsymbol{\Delta}$, and $\boldsymbol{e}_0$ is a given pilot estimate (good guess) for solving the coefficient error vector $\boldsymbol{e}$. Since $\boldsymbol{e}$ is ideally a zero vector in the converged state, we set $\boldsymbol{e}_0$ as a zero vector and thus we can simplify the calculation of gradient values to

$$g_i = ||\boldsymbol{u}_i\Delta_i||. \tag{4.11}$$

The probability that the i-th feedback sample and the equivalent row of matrix $\boldsymbol{U}_z$ are taken into the calculation of the DPD adaptation in (4.7) is calculated as [178]

$$p_i = \frac{Ng_i}{\sum\limits_{i=1}^{N_0} g_i}. \tag{4.12}$$

Based on the calculated probabilities, N samples $n_1, n_2, ..., n_N$ are selected for the calculation of new DPD coefficients.

4.5.4 Histogram-Based SSMs

Even though both QRS and GSS are analytically based, they do not consider a priori knowledge of the PA characteristics nor signal statistics. To overcome this disadvantage, we propose an SSM based on the histogram of signal magnitudes of the PA input x. Let us make the following notations: J is the number of histogram bins. The target bin counts are denoted d_j with $j \in 1, 2, ..., J$. The lower and upper bin boundaries are θ_{j-1} and θ_j respectively (see Fig. 4.4). D_j is a set of selected indices n_i whose samples belong to the j-th bin, defined as

$$D_j = \{n_i : \theta_{j-1} < |x[n_i]| < \theta_j\}, \tag{4.13}$$

and $\overline{\overline{D_j}}$ is the cardinality (number of elements) of set D_j.

Samples for DPD adaptation are selected randomly such that the j-th histogram bin count reaches the target bin count d_j, i.e., cardinality $\overline{\overline{D_j}}$ is equal to d_j. At the same time, samples should be selected such that they are distant in time and hence the matrix conditioning is improved. Formally, the selected indices need to satisfy the condition

$$|n_i - n_\kappa| \geq \gamma;\ \forall i, \kappa : i \neq \kappa, \tag{4.14}$$

where γ is the minimum sample time distance and is set such that it ensures the independence of selected samples.

Even though the papers [183,184] have proposed histogram methods with slightly different usage, the condition of a minimum sample distance expressed by (4.14) has not, to the author's best knowledge, been proposed for the histogram-based sample selection methods.

Evenly Distributed Histogram

The *evenly distributed histogram* (EDH) is a possible setting for histogram-based sampling. If N is divisible by J, all target bin counts are set to the same constant $d_j = \frac{N}{J}$, otherwise they are selected randomly from the values $d_j \in \{\lfloor \frac{N}{J} \rfloor; \lfloor \frac{N}{J} \rfloor + 1\}$ such as $\sum d_j = N$. The motivation for making the histogram evenly distributed is to cover the whole AM/AM characteristics of PA, as shown in Fig. 4.4. Unfortunately, this simple approach of setting target bin counts does not respect either the shape of the AM/AM characteristics of the used PA nor the statistical properties of the transmitted signal (Fig. 4.4) which results in a DPD model with similar modelling capabilities in all regions of AM/AM characteristics, even those not frequently utilised due to the high PAPR of the transmitted signal. In other words, if we consider the limited number of points for DPD adaptation, selecting points from highly nonlinear regions and high signal probability provides more information for DPD adaptation than selecting points from linear regions and low signal probability.

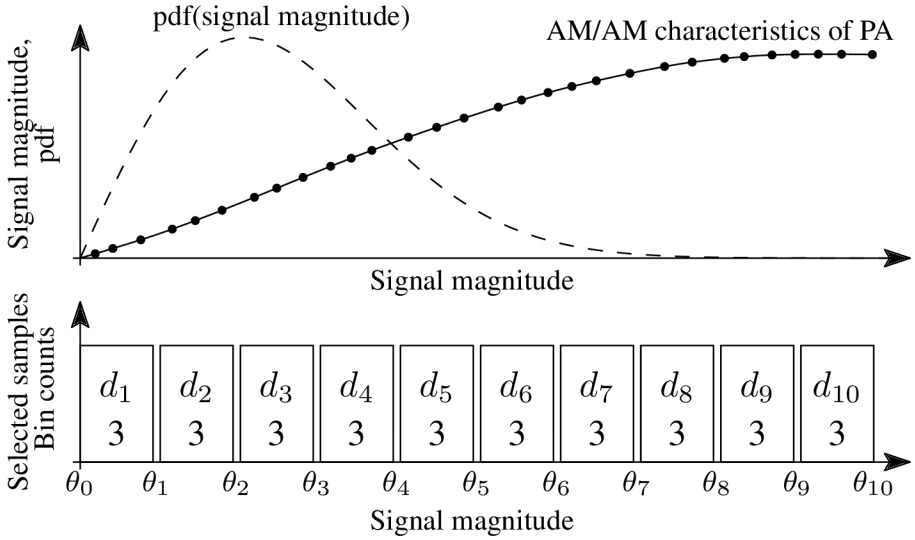


Fig. 4.4: Principle of sample selection by EDH with $J = 10$ related to the probability density function (pdf) of signal magnitude and the AM/AM characteristics of the PA. The depicted $N = 30$ selected samples are evenly distributed over the whole AM/AM characteristics of the PA, irrespective of the signal statistics nor the shape of the AM/AM characteristics.

Genetically Optimised Histogram

The *genetically optimised histogram* (GOH) is proposed to suppress EDH imperfections. An optimised histogram can respect the AM/AM characteristics of the PA and the statistics of the transmitted signal, as depicted in Fig. 4.5. We show that it is possible to set the target bin count such that a chosen criterion is optimised. In general, the histogram bin counts should be set such that GOH selects samples from regions with strong nonlinearity and high signal amplitude probability.

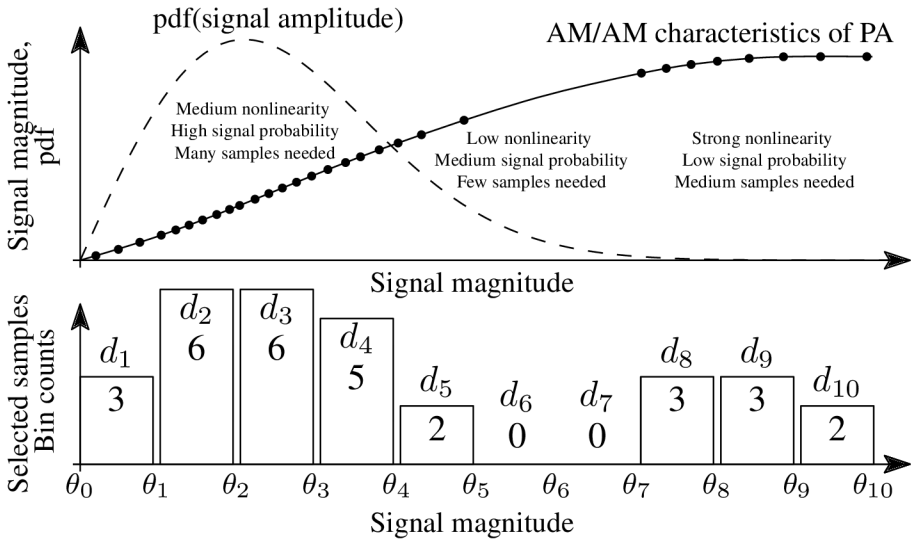


Fig. 4.5: Principle of the sample selection by GOH with $J = 10$ related to the probability density function (pdf) of signal magnitude and the AM/AM characteristics of the PA. The depicted $N = 30$ selected samples are placed in regions with strong nonlinearity and/or high signal probability.

We optimise target histogram bin counts by the genetic algorithm [185] to minimise the NMSE of the PA output with respect to the desired output. We have chosen the genetic algorithm optimisation for the integer histogram bin counts, because the genetic algorithm does integer optimisation by default. We assume that the histogram is optimised once for a specific type of PA and the transmit signals, just to set the target bin counts which are afterwards applied for DPD adaptation. Therefore, higher computational costs of optimisation calculation can be neglected, although it can take several hours to optimise the histogram. Therefore, the optimisation is most likely unfeasible in real time in a transmitter, and we consider precalculated histograms only. However, we have not aimed to make the optimisation faster, and this might be a focus in future research.

The application of GOH can be limited in cases of changing working conditions. The PA characteristics change due to temperature changes and ageing. However,

due to our experience, temperature changes cause mainly a change of the PA gain and only small changes in the shape of the AM/AM characteristics (usually within 10%). A small margin in the number of selected samples N should ensure GOH insensitivity to these small changes. However, extensive analysis should always be done for the specific application.

Another potential usage limitation arises when the communication system employs signals with significantly changing statistics of the signal magnitudes. If a single precalculated histogram does not lead to sufficient linearisation performance, then we suggest optimising the histogram for all different types of signals and applying the corresponding precalculated histogram in hardware.

In both cases, it is always possible to apply EDH which does not require histogram optimisation. The simulations with the results in Sec. 4.7.6 show that EDH leads to only a slightly higher computation complexity or slightly worse linearisation performance compared with GOH. GOH and EDH are further discussed in Sec. 4.7.4 and Sec. 4.7.5.

4.6 Reduction of Computational Complexity of DPD Adaptation

A few properly selected samples N significantly reduces the computational complexity of DPD adaptation. We evaluate the computational complexity of DPD adaptation with P coefficients and N samples with respect to the required number of real-valued multiplications $O_{\otimes}(N, P)$ and real-valued additions $O_{\oplus}(N, P)$. The following analysis does not include any optimisation and does not exploit the properties of the PA model, and the presented numbers of required multiplications and additions can therefore be seen as the upper bounds. We assume that a single complex multiplication requires four real-valued multiplications and three real-valued additions; a single complex addition requires two real-valued additions.

We split the analysis into three parts:

1. the calculation of the DPD model represented by the matrix $\boldsymbol{U}_z$,
2. calculation of the coefficient error vector $\boldsymbol{e}$,
3. and the coefficient update.

The calculation of $\boldsymbol{U}_z$ for the MP model in equation (1.14) requires $O_{U\otimes}(N, P)$ real-valued multiplications and none of the real-valued additions. Assuming that the calculation of $\phi_{k,q}^{(x)}[n]$ is reused for the calculation of $\phi_{k+1,q}^{(x)}[n]$, the number of multiplications can be expressed as

$$O_{U\otimes}(N, R) \approx 2NR. \tag{4.15}$$

The calculation of coefficient error vector $\boldsymbol{e}$ in equation (1.27) requires $O_{\boldsymbol{e}\hat{\otimes}}(N,P)$ complex multiplications: $\frac{1}{2}NP(P+1)$ for the $\boldsymbol{U}_z^H\boldsymbol{U}_z$ matrix multiplication because the resulting matrix is of size $P \times P$ and is symmetrical, $O_{I\hat{\otimes}}(P) \approx P^3$ for the inversion of the symmetrical $P \times P$ complex matrix $(\boldsymbol{U}_z^H\boldsymbol{U}_z)$, NP for $(\boldsymbol{U}_z^H(\boldsymbol{z}-\boldsymbol{y}))$, and P^2 for the multiplication of the matrix inverse $(\boldsymbol{U}_z^H\boldsymbol{U}_z)^{-1}$ and vector $\boldsymbol{U}_z^H(\boldsymbol{z}-\boldsymbol{y})$. The calculation of coefficient correction $\boldsymbol{e}$ in equation (1.26) requires $O_{\boldsymbol{e}\hat{\oplus}}(N,P)$ complex additions: $\frac{1}{2}(N-1)P(P+1)$ additions for $(\boldsymbol{U}_z^H\boldsymbol{U}_z)$ calculation, $O_{I\hat{\oplus}}(P) \approx P^3$ additions for the inversion of the symmetrical $P \times P$ complex matrix, N additions for subtraction $(\boldsymbol{z}-\boldsymbol{y})$, $(N-1)P$ for $\boldsymbol{U}_z^H(\boldsymbol{z}-\boldsymbol{y})$, and $(P-1)P$ additions for the multiplication of the matrix inverse $(\boldsymbol{U}_z^H\boldsymbol{U}_z)^{-1}$ and vector $(\boldsymbol{U}_z^H(\boldsymbol{z}-\boldsymbol{y}))$. Complexity can be expressed as

$$O_{\boldsymbol{e}\hat{\otimes}}(N,P) \approx \frac{1}{2}NP(P+3) + P^2 + P^3, \tag{4.16}$$

$$O_{\boldsymbol{e}\hat{\oplus}}(N,P) \approx \frac{1}{2}(N-1)P(P+3) + P^3 + N + (P-1)P. \tag{4.17}$$

The calculation of a single DLA iteration in equation (1.27) additionally requires $2P$ real-valued multiplications and additions. The total number of real-valued multiplications and additions can therefore be estimated as

$$\begin{aligned} O_{\otimes}(N,P) = {} & O_{U\otimes}(N,P) + 4O_{\boldsymbol{e}\hat{\otimes}}(N,P) + 2P \approx \\ & 2NP(P+4) + 4P^2 + 4P^3 + 2P, \end{aligned} \tag{4.18}$$

$$\begin{aligned} O_{\oplus}(N,P) = {} & 2O_{\boldsymbol{e}\hat{\oplus}}(N,P) + 3O_{\boldsymbol{e}\hat{\otimes}}(N,P) + 2P \approx \\ & \left(\frac{5}{2}NP - P\right)(P+3) + 5P^2 + 5P^3 + 2N. \end{aligned} \tag{4.19}$$

4.6.1 Additional Complexity of Histogram-Based SSMs

In terms of resources, histogram-based SSMs are very simple, because they only require storing and incrementing the actual bin counts to select the samples. Hence, additional computational resources required by these methods can be neglected.

4.6.2 Additional Complexity of GSS

GSS additionally requires calculating the matrix $\boldsymbol{U}_z$ for all N_0 samples which requires $O_{U\otimes}(N_0-N,P) \approx 2(N_0-N)P$ real-valued multiplications for the MP model. In (4.11) the calculation of $\boldsymbol{u}_i\Delta_i$ needs P complex multiplications while the calculation of the l_2-norm of a complex vector needs $5P$ real-valued multiplications and $4P-1$ real-valued additions, and this equation needs to be calculated N_0 times.

Additional computational complexity to implement GSS is therefore

$$O_{\otimes}^{\mathrm{GS}}(N, P, N_0) \approx 11N_0P - 2NP, \tag{4.20}$$

$$O_{\oplus}^{\mathrm{GS}}(N, P, N_0) \approx N_0(7P + 1). \tag{4.21}$$

4.6.3 Additional Complexity of QRS

QRS also requires calculating the matrix $\boldsymbol{U}_z$ for all N_0 samples and QR decomposition requires $33N_0 + 1$ complex multiplications and $8N_0$ complex additions [182]. Additional computational complexity to implement QRS is therefore

$$O_{\otimes}^{\mathrm{QR}}(N, P, N_0) \approx 132N_0 + 2(N_0 - N)P + 4, \tag{4.22}$$

$$O_{\oplus}^{\mathrm{QR}}(N, P, N_0) \approx 107N_0 + 3. \tag{4.23}$$

4.6.4 Comparison of SSM Complexity

In Tab. 4.1 we provide a comparison of the computational complexity for the above SSMs in terms of the number of real-valued multiplications and additions. In all cases there are $N_0 = 20\,000$ samples collected from the feedback. The complexity is evaluated for the MP model with $K = 7$, and $Q = 1$ which leads to $P = 14$ DPD coefficients. We can conclude that QRS is computationally the most demanding as it performs the inversion of a large matrix $\boldsymbol{U}_z$ constructed from all N_0 samples. The least complex methods are histogram-based SSMs whose complexity is the same as for conventional DPD adaptation using a block of continuous feedback samples. However, the conventional methods provide poor linearisation capabilities for a few samples N as we show in Sec. 4.7.6. Although QRS and GSS have higher computational complexity for the same number of selected samples N than the conventional DPD, the computational complexity reduction is achieved for them as QRS and GSS allow predistortion with lower N.

Although we have neglected memory requirements in our analysis, please note that QRS and GSS require much more memory to store the whole $N_0 \times P$ matrix $\boldsymbol{U}_z$ and the $N_0 \times 1$ vector $\boldsymbol{\Delta}$ than histogram-based SSMs, which only need to store the reduced $N \times P$ matrix $\boldsymbol{U}_z$ and the $N \times 1$ vector $\boldsymbol{\Delta}$.

4.6.5 Reduction of Hardware Resources

Since SSM allows undersampling the PA output, it allows for ADCs with decreased sampling frequency in the feedback. However, the ADC analogue input bandwidth and the speed of the ADC sample-and-hold circuit have to be sufficient to cover the desired signal bandwidth including close intermodulation products. In common

Tab. 4.1: Comparison of Computational Complexity

		Conv. DPD	QRS	GSS	Hist. SSM
$P = 14$	$O_{\otimes}$	DNC	$3.2 \cdot 10^6$	$3.1 \cdot 10^6$	$22 \cdot 10^3$
$N = 20$	$O_{\oplus}$	DNC	$2.2 \cdot 10^6$	$2.0 \cdot 10^6$	$26 \cdot 10^3$
$P = 14$	$O_{\otimes}$	DNC	$3.3 \cdot 10^6$	$3.1 \cdot 10^6$	$62 \cdot 10^3$
$N = 100$	$O_{\oplus}$	DNC	$2.2 \cdot 10^6$	$2.1 \cdot 10^6$	$74 \cdot 10^3$
$P = 14$	$O_{\otimes}$	$5.1 \cdot 10^6$	$8.0 \cdot 10^6$	$7.8 \cdot 10^6$	$5.1 \cdot 10^6$
$N = 10^4$	$O_{\oplus}$	$6.0 \cdot 10^6$	$8.1 \cdot 10^6$	$7.9 \cdot 10^6$	$6.0 \cdot 10^6$

Conventional DPD does not converge (DNC) for a low number of selected samples N. The complexity in the table is calculated for $N_0 = 20\,000$ and the MP model with $K = 7$, and $Q = 1$.

state of the art, it is required that the feedback bandwidth be three to five times the desired channel bandwidth [6], [7]. ADCs with high analogue input bandwidth are commonly employed in current communication systems either for Sub-Nyquist sampling or in interleaved ADCs.

The main advantages of the decreased sampling frequency of the feedback ADCs are the lower power consumption, the decreased system complexity, and the price. We present these parameters for the example transmitter from Fig. 4.1. We consider two cases: in one case the conventional DPD and in the other case an undersampled ADC in the feedback. The communication signal bandwidth was chosen to be 300 MHz, which implies a required feedback bandwidth of at least 900 MHz. The conventional DPD requires two ADCs with the sampling frequency $F_{\mathrm{S}} \geq 900$MSps. DPD with SSM requires two ADCs for IQ sampling with the input bandwidth $B_{\mathrm{in}} \geq$ 450 MHz and arbitrary F_{S}. For this example we have selected ADCs AD9690 [186] with $F_{\mathrm{S}} = 1$ GSps for the conventional DPD and AD9629 [187] with $B_{\mathrm{in}} = 700$ MHz and $F_{\mathrm{S}} = 20$ MSps for DPD with SSM, both labelled as low power ADCs. Tab. 4.2 gives a comparison of a transmitter with the conventional DPD and a transmitter with the undersampled feedback allowed by the proposed SSM. The values were taken from ADC data sheets [186], [187]. The comparison of system parameters in Tab. 4.2 does not include the power consumption and price of auxiliary components, e.g. clock generators, buffers, filters, etc.

Tab. 4.2: Comparison of System Parameters

	Conventional DPD 2× AD9690	DPD with SSM 2× AD9629
F_S	1000 MSps	20 MSps
Max. B_in	500 MHz	700 MHz
ENOB	10.5 bits	11 bits
Digital interface	JESD204B	Parallel
Power consumption	$\approx$ 4 W	$\approx$ 90 mW
Price	$\approx$ 720 USD	$\approx$ 20 USD

4.7 Simulations

We simulate the DPD performance with the proposed SSM described in Sec. 4.5. Additionally, all simulations include the conventional DPD without SSM as a reference. For all methods, the DLA has been used. Since it is iterative, each simulation run consists of $M = 80$ iterations. At the beginning of each iteration the transmit signal $\boldsymbol{z}$ is generated with random data symbols and used for all simulated SSMs. Afterwards, the following steps bound to the specific SSM are executed:

1. signal $\boldsymbol{z}$ is predistorted using coefficients $\boldsymbol{b}'$ to obtain the PA input $\boldsymbol{x}$,
2. signal $\boldsymbol{x}$ is sent through the PA model to get the PA output $\boldsymbol{y}$,
3. new DPD coefficients $\hat{\boldsymbol{b}}$ are calculated,
4. the linearisation performance metrics are evaluated,
5. and the DPD coefficients are updated $\hat{\boldsymbol{b}} \rightarrow \boldsymbol{b}'$.

For the first iteration, the DPD coefficients $\boldsymbol{b}'$ are set to $[0.5, 0, \ldots, 0]^T$. The step size parameter μ was set to 0.1 as a decent trade-off between the convergence probability and the convergence speed. The whole process is repeated 1000 times to evaluate the 95% confidence intervals of the results.

4.7.1 PA Models for Simulations

In simulations we employ the PA model based on GMP [19] (see Sec. 1.3.3). We have chosen GMP, because it can be sufficiently complex to accurately model the state-of-the-art PAs. Moreover, for realistic simulation results it is necessary to apply the PA model of higher complexity than the DPD model.

We include models of the following PAs in our simulations:

PA1: Two-stage PA in class AB designed for Satcom applications at 1625 MHz, with monolithic PA GALI-24+ from Mini-Circuits [188] in the first stage and GaN SiC HEMT TGF2965 from Qorvo [189] in the second stage. The first stage power supply voltage was set to 8 V, the second stage power supply voltage was set to 32 V, and the gate voltage was set to ensure that the drain quiescent current was 20 mA. The total gain of both stages is approximately 30 dB.

PA2: Class-F power amplifier designed for Satcom applications at 1625 MHz, with output power $P_{3\text{dB}} = 43\,\text{dBm}$, based on the GaN SiC HEMT T2G6003028 from Qorvo [190]. The PA gain is approximately 12 dB, the power supply voltage 28 V, the gate voltage -3.03 V. Together with this PA, the 5-W amplifier Minicircuits ZHL-5W-2G-S+ [191] was used as a predriver.

PA3: Block amplifier ADL5610 [192].

All PA models were extracted from measurements at a centre frequency of 1600 MHz, excited by a test signal with the bandwidth $B \approx 6$ MHz, further described in Sec. 4.7.3. Index arrays of PA models were set to ensure sufficient fidelity of the models. The particular indexes were set to $K_a = \{0, 1, 2, 3, 4, 5\}$, $L_a = \{0, 1, 2, 3\}$, $K_b = \{2, 4\}$, $L_b = \{0\}$, $M_b = \{1\}$, $K_c = \{2, 4\}$, $L_c = \{0\}$, $M_c = \{1\}$. The PA model coefficients can be found in the source codes provided.

4.7.2 DPD Model

As a model of digital predistorter we have chosen the DDR2 model [15] (described in Sec. 1.3.2), because the simple MP model did not achieve sufficient linearisation results for the tested amplifiers. For the following simulations, we have chosen the DPD order to be $K = 7$ and $Q = 1$, which yields $P = 17$ DPD coefficients.

4.7.3 Test Signal

We have chosen the F-OFDM signal waveform with the *64-state quadrature amplitude modulation* (64-QAM) as the internal modulation to demonstrate the linearisation capabilities of the proposed SSM for DPD adaptation. The F-OFDM has been designed as described in Sec. 1.5.2.

In each iteration, we generate the F-OFDM signal with 12 frames, each frame with 68 resource blocks, a block size of 12 subcarriers, resulting in $W = 12 \cdot 68 = 816$. The FFT size is set to $Z = 4096$ and the filter length is $L = 2049$. The sampling frequency is limited by the measurement setup and is set to $F_\text{S} = 30\,\text{MHz}$. These parameters provide the test signal vector with the channel bandwidth $B \approx F_\text{S}/5 \approx 6\,\text{MHz}$ and $N_0 = 52\,064$ samples. The signal mean power was set constant during

the DPD adaptation such that the mean PA output power over the last 20 iterations is equal for all the methods.

4.7.4 Settings Specific to Histogram-Based Methods

We have set the number J of histogram bins for EDH and GOH based on the simulation presented in Fig. 4.6, where we have analysed the influence of the number of bins on achievable NMSE, using DPD with EDH. We can see that for DPD adaptation the sufficient number of histogram bins J is 5. We set $J = 10$ to ensure a sufficient margin, as we expect a very low computational resource allocation for histogram-based SSMs.

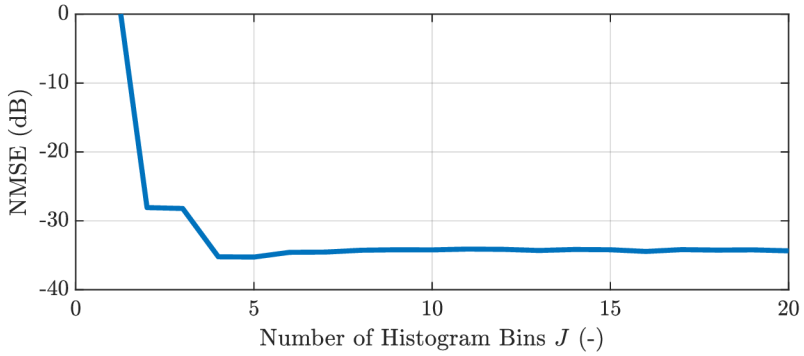


Fig. 4.6: Simulated influence of the number of histogram bins on achievable NMSE using DPD with EDH and $N = 30$.

4.7.5 Histogram Optimisation by Genetic Algorithm

We have optimised histogram target bin counts d_j by the genetic algorithm [185] for each simulated number of selected samples N. Optimised bin counts are integers from the interval $[0, N]$ and need to fulfil the condition $\sum d_j = N$. We have set the optimisation parameters as follows: the population size 100, the maximum number of generations 20, the population fraction at the next generation created by crossover 0.8, the probability of mutation 1%. We have defined the objective function as an average of the NMSE results over 10 runs. Each run consists of 80 iterations and the NMSE results for averaging are taken only from the last 20 iterations.

4.7.6 Simulation Results

Hereinafter we present detailed simulation results for the model PA1 in Fig. 4.7, Fig. 4.8, and Fig. 4.9. The simulation results for the other models, PA2 and PA3, are provided to verify SSMs in a condensed form in Tab. 4.3 and Tab. 4.4.

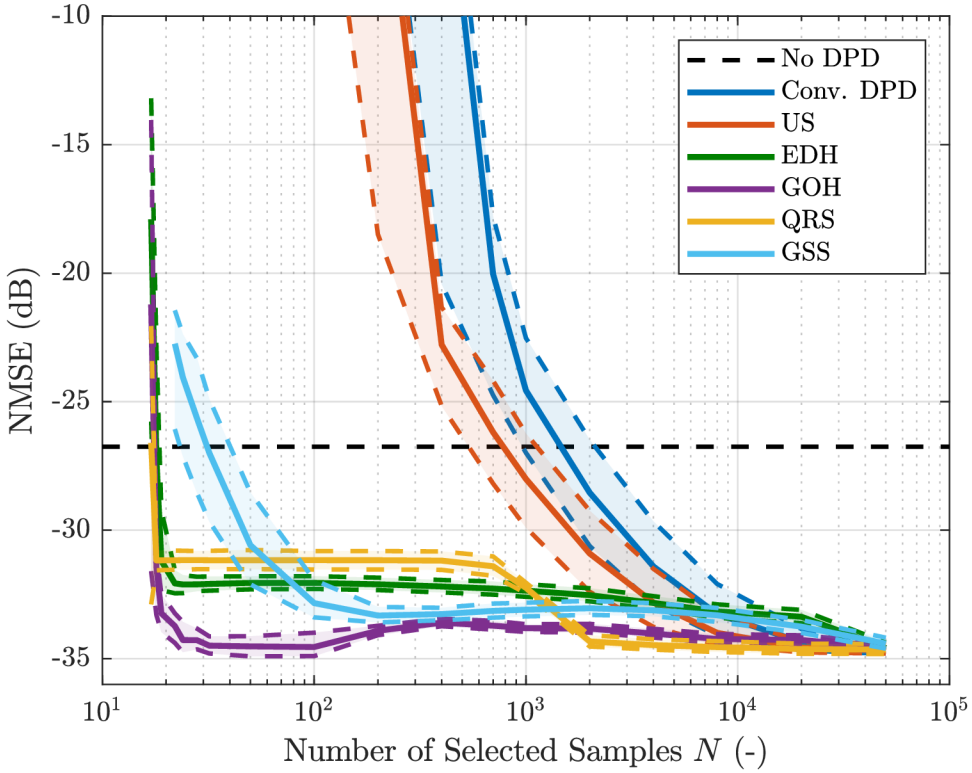


Fig. 4.7: Simulation results of NMSE as a function of the number of selected samples N with 95% confidence intervals depicted by coloured dashed lines for the model PA1. The black dashed line represents the simulated NMSE of the PA output without DPD.

The relationship between NMSE and the number of selected samples N is depicted in Fig. 4.7. We can observe that the conventional DPD starts to improve the transmitter linearity when more than 1300 samples are selected for DPD adaptation. US follows the behaviour of the conventional DPD but with slightly less needed samples. QRS starts to improve linearity of the output signal from 18 selected samples, but up to 2000 samples it provides suboptimal results. This is mainly due to improving problem conditioning but considering neither the signal statistics nor the observed feedback samples, as has been discussed above. Please recall that we have 17 DPD coefficients and taking only 18 samples is almost equivalent to solving a fully determined system.

GSS shows good linearisation performance for more than 100 selected samples. We suppose this is caused by choosing samples on a random basis with respect to the probabilities calculated by GS. For a few selected samples there is no margin for selecting unimportant samples. EDH starts to linearise from 19 selected samples and slightly improves with an increasing number of selected samples. For $N < 100$, EDH outperforms GS and for $N > 100$, EDH provides the NMSE less than 0.8 dB higher than GS.

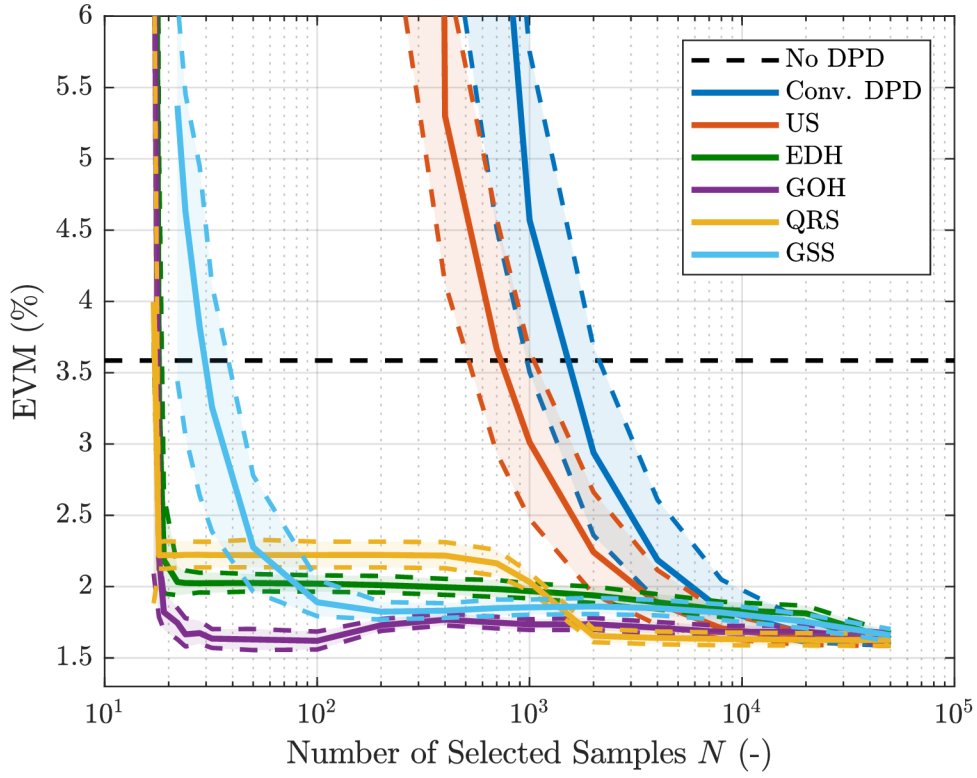


Fig. 4.8: Simulation results of the EVM as a function of the number of selected samples N with 95% confidence intervals depicted by coloured dashed lines for the model PA1. The black dashed line represents the simulated EVM of the PA output without DPD. The EVM of the generated signal is 1.1% due to the nonorthogonality caused by the inherent F-OFDM filtering.

In the region of up to $N < 2000$ samples, GOH outperforms all other methods. For $N > 2000$, GOH is very close to QRS. Please note that this difference is negligible, about 0.5 dB in terms of NMSE. For GOH, in the region from 100 samples, we can see the effect of a larger space in which the genetic algorithm looks for the optimal histogram. In this region, the NMSE is up to 1 dB worse than for $N = 100$ or $N = 50\,000$. This effect is caused by the fixed number of maximum generations and the population size of the genetic algorithm, set irrespective of the number of selected samples N.

EVM as a function of the number of selected samples N is depicted in Fig. 4.8. The simulated EVM results agree with the NMSE results shown in Fig. 4.7. The ACPR results are presented in Fig. 4.9 and agree similarly with the NMSE results. One can notice that for PA1 the linearisation performance in terms of the ACPR improvement is not impressive. This is likely caused by specific characteristics of PA1 and the DPD model. The key point is that the proposed methods achieve the maximum linearisation performance of the conventional DPD. The higher ACPR improvement is achieved for PA3 with results in Tab. 4.4, as described below.

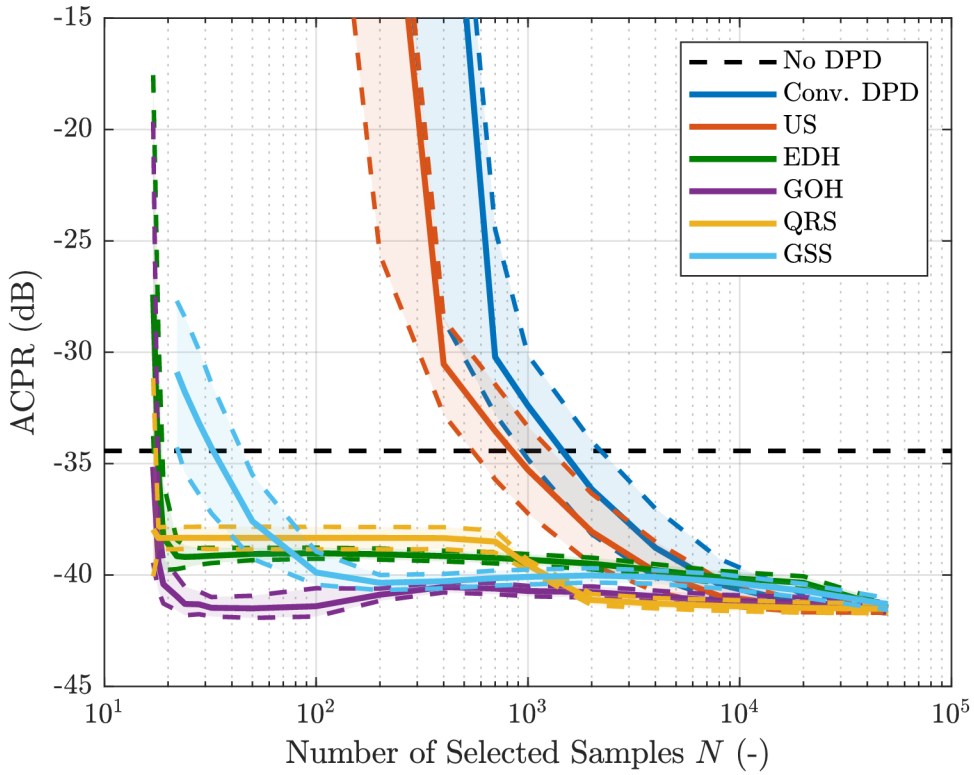


Fig. 4.9: Simulation results of ACPR as a function of the number of selected samples N with 95% confidence intervals depicted by coloured dashed lines for the model PA1. The black dashed line represents the simulated ACPR of the PA output without DPD.

Summary results for PA2 and PA3 are presented in Tab. 4.3 and Tab. 4.4 respectively. Although EDH shows good linearisation capabilities for a small number of samples, GOH outperforms other SSMs for a few selected samples in all cases and is close to the other methods for a high number of samples.

Model PA2 with results in Tab. 4.3 is highly nonlinear, as PA2 is designed in class F, and we can see that the conventional DPD and the simple undersampling do not converge for a small number of samples $N < 10\,000$. On the other hand, the model PA3 with results shown in Tab. 4.4 is less nonlinear and the conventional DPD and US provide a decent linearisation for $N = 1000$ selected samples. In Tab. 4.4 we can notice that the PA nonlinearity causes mainly leakage into adjacent channels while the signal in the main channel remains undistorted, as illustrated by EVM.

Tab. 4.3: Simulation Results for Model PA2

Metric	N	Conv. DPD	US	EDH	GOH
NMSE (dB)	22	DNC	DNC	-15.4	-20.4
	40	DNC	DNC	-19.8	-20.8
	10 000	-14.0	-18.0	-20.8	-21.1
	20 000	-20.6	-21.0	-20.9	-21.1
EVM (%)	22	DNC	DNC	9.8	4.4
	40	DNC	DNC	4.9	4.2
	10 000	13.2	7.2	4.1	3.9
	20 000	4.2	4.0	4.0	3.9
ACPR (dB)	22	DNC	DNC	-22.5	-26.8
	40	DNC	DNC	-26.2	-27.2
	10 000	-21.0	-26.4	-27.1	-27.2
	20 000	-27.0	-27.2	-27.2	-27.3

Conventional DPD and US *do not converge* (DNC) for a low number of selected samples N. Simulated metrics without DPD: NMSE = -13.9 dB, EVM = 13.2%, and ACPR = -21.8 dB. The EVM of the generated signal is 1.1% due to the nonorthogonality caused by the inherent F-OFDM filtering.

4.8 Measurements

In addition to simulations for all models PA1, PA2, and PA3, we evaluate the SSMs for PA1 in measurements. Our measurement setup (Fig. 4.10) employs the vector signal generator Rohde & Schwarz SMU200A to generate the input signal of PA1. The PA1 output is connected to the real time spectrum analyser Rohde & Schwarz FSVR through a high-power RF attenuator with an attenuation of 10 dB and a maximum dissipated power of 50 W followed by two smaller RF attenuators of 10 dB each. We use the digital multimeter Keysight 34461A as an Ampere metre to set the drain quiescent current of the PA1 second-stage transistor.

We have performed the measurements with the same settings as for simulations. The only change is that for measurements we perform 80 iterations only once, and hence the values presented are not averaged.

Tab. 4.4: Simulation Results for Model PA3

Metric	N	Conv. DPD	US	EDH	GOH
NMSE (dB)	19	DNC	DNC	DNC	-45.8
	22	DNC	DNC	-46.3	-46.7
	1000	-44.8	-46.0	-47.0	-47.0
	4000	-46.7	-46.6	-47.0	-47.1
EVM (%)	19	DNC	DNC	DNC	1.15
	22	DNC	DNC	1.15	1.15
	1000	1.17	1.16	1.15	1.15
	4000	1.95	1.15	1.15	1.15
ACPR (dB)	19	DNC	DNC	DNC	-58.2
	22	DNC	DNC	-60.4	-61.8
	1000	-56.5	-59.5	-64.4	-64.9
	4000	-62.4	-61.8	-64.8	-65.0

Conventional DPD and US *do not converge* (DNC) for a low number of selected samples N. Simulated metrics without DPD: NMSE = -33.3 dB, EVM = 1.5%, and ACPR = -42.6 dB. The EVM of the generated signal is 1.1% due to the nonorthogonality caused by the inherent F-OFDM filtering.

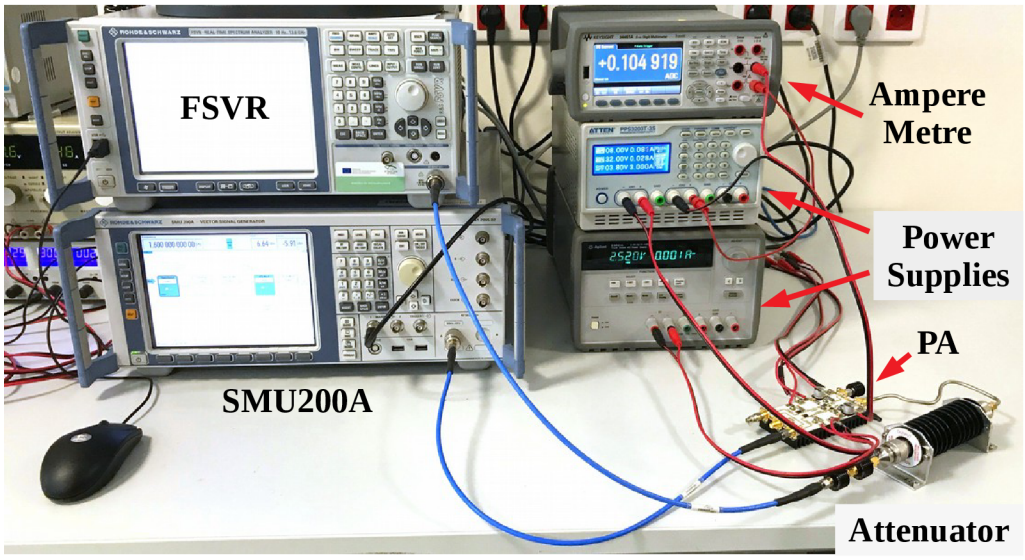


Fig. 4.10: Measurement setup with PA1, the vector signal generator SMU200A, and the real time spectrum analyser FSVR.

In Fig. 4.11 we present the measured AM/AM characteristics of PA1, DPD and the whole linearised transmitter. The characteristics of DPD and the linearised transmitter are depicted for DPD adapted using GOH with $N = 24$ selected samples. Please note that the selected samples are not distributed evenly over the input magnitudes, but rather concentrated in the region with lower magnitudes to cover the region with higher signal occurrence probability and then in the region with higher magnitudes to cover the most nonlinear region of PA.

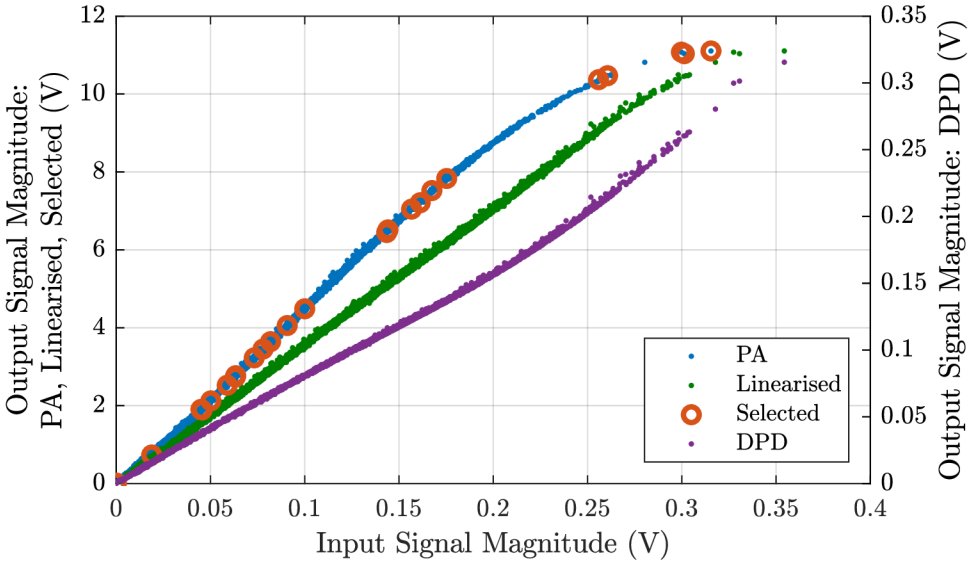


Fig. 4.11: Measured AM/AM characteristics of PA1, DPD, and the whole linearised transmitter. The depicted $N = 24$ samples were selected by GOH and used for DPD adaptation.

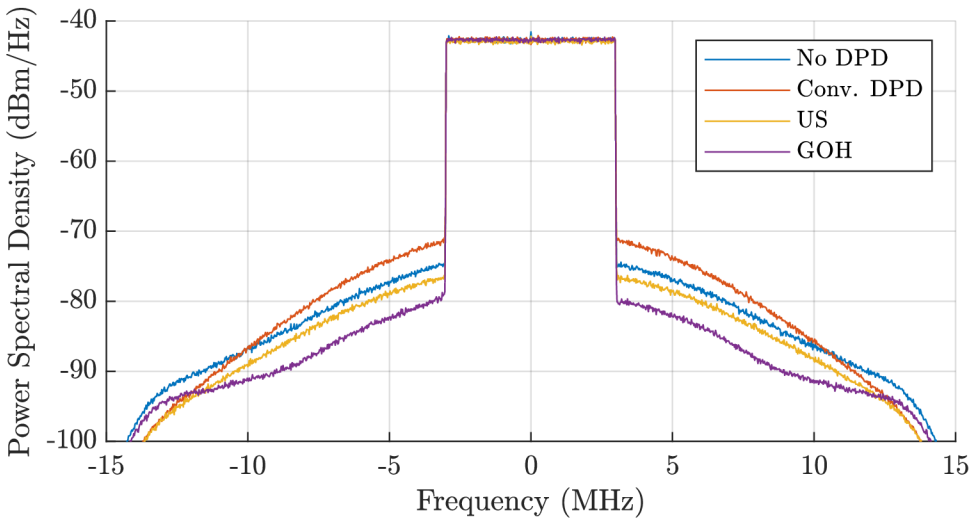


Fig. 4.12: Measured power spectral density of PA1 output without DPD and with DPD adapted by different SSMs for the number of selected samples $N = 400$.

The measured *power spectral density* (PSD) of the PA1 output linearised with DPD adapted using different techniques is depicted in Fig. 4.12. The PSD of the PA1 output without DPD is shown as a reference. The results are shown for $N = 400$ selected samples. Please note that for $N = 400$ the conventional DPD causes higher out-of-band emissions than PA1 without DPD. US slightly improves the out-of-band emissions, while GOH in this case provides the best linearisation.

The measured DPD performance in terms of NMSE, EVM, and ACPR is shown in Fig. 4.13, Fig. 4.14, Fig. 4.15 respectively. We can see that the conventional DPD and US in real measurements work slightly better than in simulations. This could indicate that the extracted PA model is more difficult to linearise by these methods than the real PA. The trend for these methods follows the simulation results, and we can see a spread of measured points from this trend. This spread complies with the wider 95% confidence intervals shown in the simulation results. The measured NMSE for GOH is around 1 dB worse with respect to the average NMSE in the simulations. This increase could probably be caused by the measurement noise which is neglected in simulations. Measurement results in terms of EVM and ACPR follow the NMSE measurement results which exactly complies with the simulations.

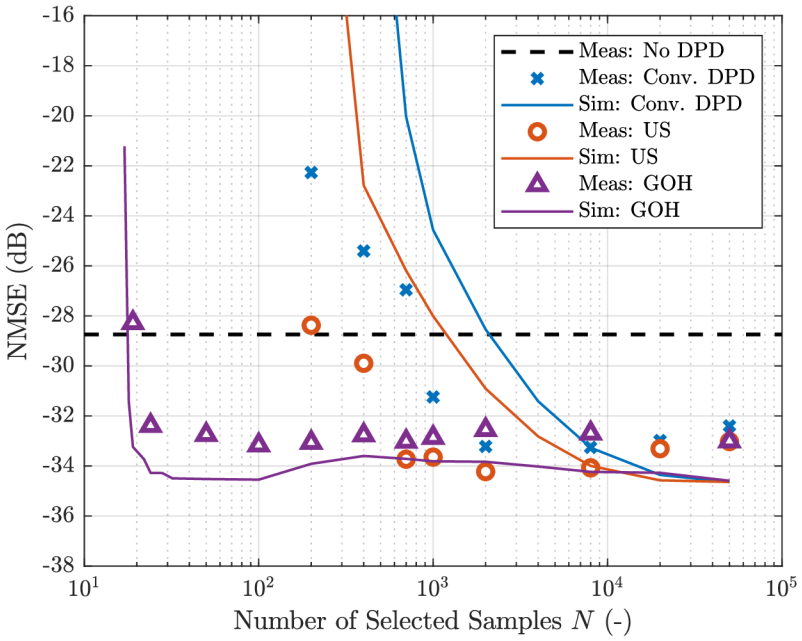


Fig. 4.13: Measurement results of NMSE as a function of the number of selected samples N compared to the simulation results. The black dashed line represents the measured NMSE of the PA1 output without DPD.

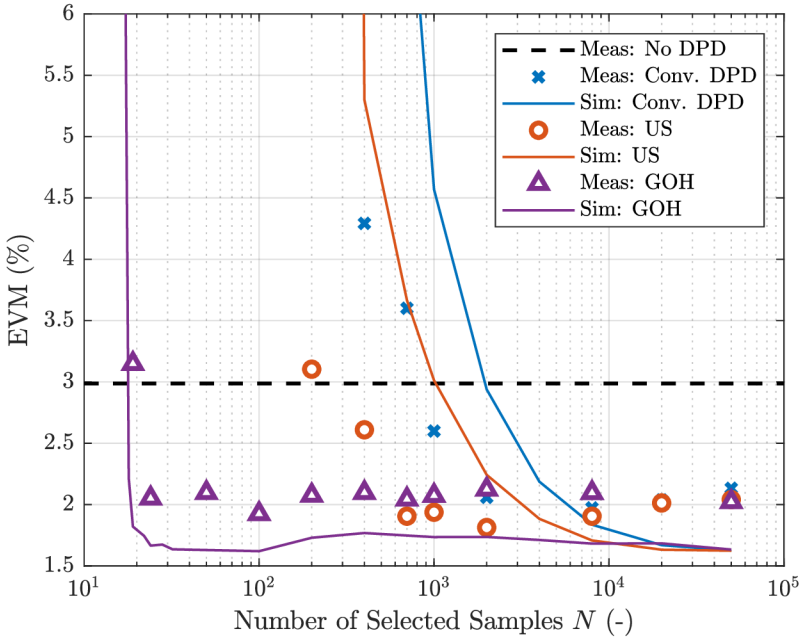


Fig. 4.14: Measurement results of EVM as a function of the number of selected samples N compared to the simulation results. The black dashed line represents the measured EVM of the PA1 output without DPD. The EVM of the generated signal is 1.1% due to the nonorthogonality caused by the inherent F-OFDM filtering.

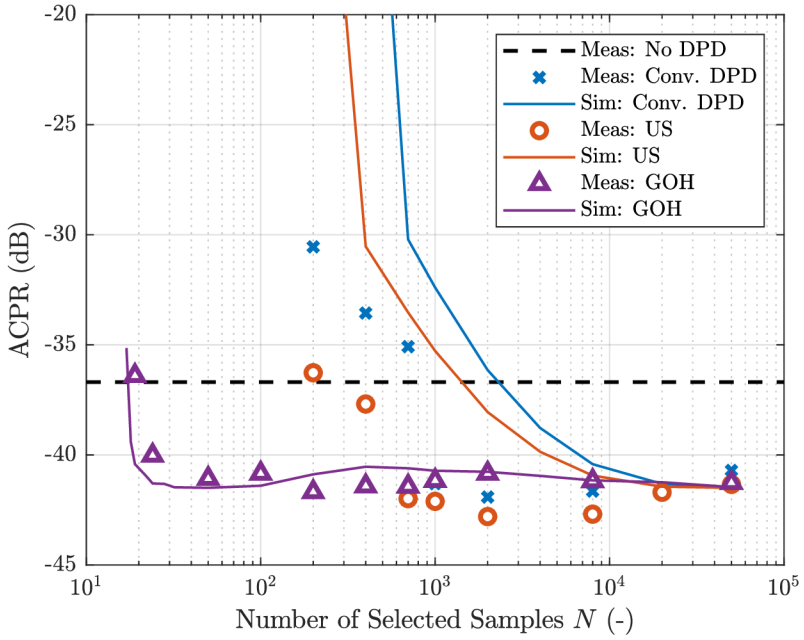


Fig. 4.15: Measurement results of ACPR as a function of the number of selected samples N compared to the simulation results. The black dashed line represents the simulated ACPR of the PA1 output without DPD.

4.9 Conclusion

The number of samples for DPD adaptation directly influences the computational complexity of DPD adaptation. In this chapter we have proposed sample selection methods for DPD adaptation with the intention to minimise the required number of samples for DPD adaptation and thus to minimise the computational complexity of DPD adaptation. We have shown that the proposed GOH outperforms other sample selection methods in terms of linearisation capabilities. For a very small number of selected samples, GOH provides a linearisation performance equivalent to the maximum achievable linearisation performance of the conventional DPD. Simulations on the model of a two-stage power amplifier (PA1) designed for Satcom applications have revealed that GOH can achieve a sufficient linearisation performance already for $N = 24$ selected samples while the conventional DPD achieves the equivalent performance only for $N \geq 10^4$. Since the computational complexity is linear with respect to the number of required samples, this indicates a 400-times improvement over the conventional DPD in terms of computational complexity.

5 DPD Adaptation with Level-Crossing ADC

5.1 Introduction

In this chapter, we propose a novel method for predistorter adaptation with an LC-ADC [193] based on a comparator and an *low-speed digital-to-analogue converter* (LSDAC) which replaces conventional ADCs in the feedback path. The block diagram of the proposed architecture is depicted in Fig. 5.1. In the feedback, there is an RF mixer or an IQ mixer with a single connected output followed by an LPF providing the in-phase feedback signal $y_\mathrm{r}(t)$. Signal $y_\mathrm{r}(t)$ is compared with a reference voltage $r(t)$ generated by the LSDAC. The output of the comparator is connected to a circuit for the time extraction of the comparator output edges. Employing the LC-ADC significantly reduces the system power consumption, system complexity and price.

The idea of a comparator in the DPD feedback has already been presented by Wang et al. [4, 152], and Zhang et al. [153]. Their methods have been analysed in Sec. 2.3.4. However, our architecture highly differs from their approach in the following aspects:

- It does not require additional high-speed DACs and, consequently, the proposed architecture reduces the power consumption more significantly.
- It is not sensitive to gain mismatch and is only partially sensitive to delay mismatch [76, 194].

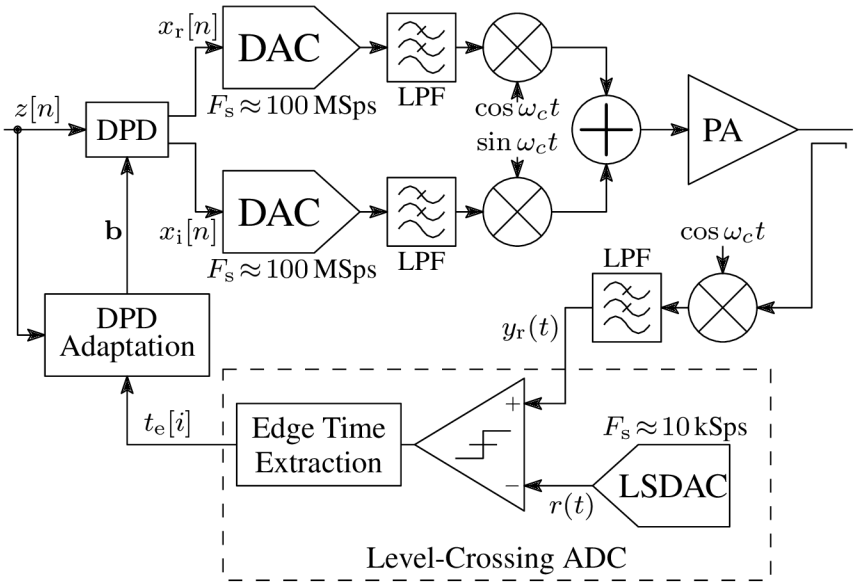


Fig. 5.1: Linearisation of a PA by the DPD with a level-crossing ADC. The edge time extraction circuit provides the time stamps $t_\mathrm{e}[i]$ of edges at the comparator output.

The rest of this chapter is organised as follows. The second section introduces the basic principles of the proposed DPD adaptation approach with a comparator in the feedback. In the third and fourth sections, we discuss the adoption of the proposed principles into the ILA and the DLA. System implementation details are presented in the fifth section. The sixth and seventh sections present the measurement procedures and the achieved results employing the simulated comparator in the feedback. On the contrary, the two final sections describe a custom-made feedback module with a hardware LC-ADC and a real comparator and the linearisation results obtained with this module.

5.2 Basic Principles of DPD with LC-ADC

Let us continue with the idea of the sample selection presented in Sec. 4.2, where the system of equations for solving the PA model coefficients contains samples sampled at arbitrary discrete times. The key idea allowing for mathematical derivation of DPD adaptation by the feedback with the LC-ADC, which is constituted by a simple comparator, is moving from the discrete time samples into the domain of continuous time samples. The following notations are required: The continuous-time signals $x(t)$, $y(t)$, $z(t)$ represent the PA input, PA output, and the desired scaled PA output in the baseband, respectively. Their corresponding discrete-time equivalents are $x[n]$, $y[n]$, $z[n]$ with n being the sample index. The transition between continuous and discrete time representation can be written as $x[n] = x(nT)$, where T is the sampling period of the discrete-time signals.

We can easily rewrite equation (1.12) of the MP model into the continuous-time domain as

$$y(t) = \sum_{k=1}^{K} \sum_{q=0}^{Q} b_{k,q} x(t - qT)|x(t - qT)|^{k-1}. \tag{5.1}$$

The system of equations for solving the PA coefficients consists of N equations. Assuming that signal $x(t)$ is sampled at distinct arbitrary time instants $t_1, t_2, \ldots, t_N$, we can arrange the system of equations as

$$\begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_N) \end{bmatrix} = \begin{bmatrix} x(t_1) & \ldots & x(t_1)|x(t_1)| & x(t_1 - T)|x(t_1 - T)| & \ldots & x(t_1 - QT)|x(t_1 - QT)|^{K-1} \\ x(t_2) & \ldots & x(t_2)|x(t_2)| & x(t_2 - T)|x(t_2 - T)| & \ldots & x(t_2 - QT)|x(t_2 - QT)|^{K-1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x(t_N) & \ldots & x(t_N)|x(t_N)| & x(t_N - T)|x(t_N - T)| & \ldots & x(t_N - QT)|x(t_N - QT)|^{K-1} \end{bmatrix} \begin{bmatrix} b_{1,0} \\ \vdots \\ b_{2,0} \\ b_{2,1} \\ \vdots \\ b_{K,Q} \end{bmatrix}. \tag{5.2}$$

On the left-hand side, this system of equations has a vector of complex feedback samples $y(t_i)$. The LC-ADC with a comparator, however, provides only information

about the level crossing of the in-phase or quadrature signal of the feedback. Even if we installed comparators in both IQ branches, we would not know the exact level of the second signal at the precise moment the first signal crossed the set threshold. Consequently, we need to employ the real-valued feedback, presented in Chapter 3, to overcome this obstacle. The derivation is straightforward. We can split the real and imaginary parts, similarly to equations (3.1) and (3.2), and rewrite equation (5.2) in the form of equation (3.3), which yields

$$\begin{aligned}\begin{bmatrix} y_\mathrm{r}(t_1) \\ y_\mathrm{r}(t_2) \\ \vdots \\ y_\mathrm{r}(t_N) \end{bmatrix} = &\begin{bmatrix} x(t_1) & \dots & x(t_1-QT)|x(t_1-QT)|^{K-1} \\ x(t_2) & \dots & x(t_2-QT)|x(t_2-QT)|^{K-1} \\ \vdots & \ddots & \vdots \\ x(t_N) & \dots & x(t_N-QT)|x(t_N-QT)|^{K-1} \end{bmatrix}_\mathrm{r} \begin{bmatrix} b_{1,0} \\ \vdots \\ b_{K,Q} \end{bmatrix}_\mathrm{r} - \\ - &\begin{bmatrix} x(t_1) & \dots & x(t_1-QT)|x(t_1-QT)|^{K-1} \\ x(t_2) & \dots & x(t_2-QT)|x(t_2-QT)|^{K-1} \\ \vdots & \ddots & \vdots \\ x(t_N) & \dots & x(t_N-QT)|x(t_N-QT)|^{K-1} \end{bmatrix}_\mathrm{i} \begin{bmatrix} b_{1,0} \\ \vdots \\ b_{K,Q} \end{bmatrix}_\mathrm{i} .\end{aligned} \tag{5.3}$$

Considering the DPD feedback with the LC-ADC from Fig. 5.1, the signal $y_\mathrm{r}(t)$ is known only for $t = t_\mathrm{e}[i]$ when the feedback signal equals the set threshold $r(t)$. Time stamps $t_\mathrm{e}[i]$ are the instantaneous times of transitions of signal $y_\mathrm{r}(t)$ over the reference voltage $r(t)$. In other words, by knowing the time stamps $t_\mathrm{e}[i]$ and the reference voltage $r(t)$, we can recover values of signal y_r at the time instants $t_\mathrm{e}[i]$. Therefore, we can set $t_i = t_\mathrm{e}[i]$ and $y_\mathrm{r}(t_i) = r(t_i)$.

The equation system can be expressed in the matrix form by arranging the output samples, model coefficients and the basis functions into vectors, i.e.

$$\begin{aligned} \boldsymbol{\phi}_{k,q}^{(x)} &= \begin{bmatrix} \phi_{k,q}^{(x)}(t_1) & \phi_{k,q}^{(x)}(t_2) & \dots & \phi_{k,q}^{(x)}(t_N) \end{bmatrix}^T , \\ \boldsymbol{r} &= \begin{bmatrix} r(t_1) & r(t_2) & \dots & r(t_N) \end{bmatrix}^T = \\ &= \begin{bmatrix} y_\mathrm{r}(t_1) & y_\mathrm{r}(t_2) & \dots & y_\mathrm{r}(t_N) \end{bmatrix}^T , \\ \boldsymbol{b} &= \begin{bmatrix} b_{1,0} & b_{1,1} & \dots & b_{1,Q} & b_{2,0} & \dots & b_{K,Q} \end{bmatrix}^T , \\ \boldsymbol{U}_x &= \begin{bmatrix} \boldsymbol{\phi}_{1,0}^{(x)} & \boldsymbol{\phi}_{1,1}^{(x)} & \dots & \boldsymbol{\phi}_{1,Q}^{(x)} & \boldsymbol{\phi}_{2,0}^{(x)} & \dots & \boldsymbol{\phi}_{K,Q}^{(x)} \end{bmatrix} , \end{aligned} \tag{5.4}$$

where $\boldsymbol{b}$ is a column vector with P rows, and the size of matrix $\boldsymbol{U}_x$ is $N \times P$. The basis functions in vector $\boldsymbol{\phi}_{k,q}^{(x)}$ and matrix $\boldsymbol{U}_x$ require the knowledge of signal values $x(t_i - nT)$ at times $t_i = t_\mathrm{e}[i]$ while signal $x(t)$ is practically sampled at times nT. As signal $x(t)$ is the PA input, it is fully known and, therefore, the signal values $x(t_\mathrm{e}[i] - nT)$ can be calculated by interpolation.

The LS solution of the PA model coefficients in the matrix form is expressed as

$$\begin{bmatrix} \boldsymbol{b}_{\mathrm{r}} \\ \boldsymbol{b}_{\mathrm{i}} \end{bmatrix} = (\boldsymbol{A}_x^H \boldsymbol{A}_x)^{-1} \boldsymbol{A}_x^H \boldsymbol{r}, \tag{5.5}$$

where

$$\boldsymbol{A}_x = \begin{bmatrix} \boldsymbol{U}_{x\mathrm{r}} & -\boldsymbol{U}_{x\mathrm{i}} \end{bmatrix}. \tag{5.6}$$

Please note that the described method has the capability of fully estimating the PA memory effects. Moreover, the output of the LC-ADC can be undersampled (Chap. 4), and therefore, the edge time extraction circuit need not register all the comparator output edges.

5.3 ILA Employing LC-ADC

The direct application of the LC-ADC for the DPD with ILA is not achievable for the same reasons as for the real-valued feedback (Sec. 3.3) and sample selection (Sec. 4.3). We can simply state that the methods by Morgan [25] and Landin [26] are required to employ ILA for the DPD adaptation with the LC-ADC in the feedback. However, further in this thesis, we will explore exclusively DLA as it generally achieves better linearisation performance.

5.4 DLA Employing LC-ADC

We start our derivation of the DLA with the LC-ADC from the basic iteration step

$$\hat{\boldsymbol{b}} = \boldsymbol{b}' - \mu \boldsymbol{e}. \tag{5.7}$$

Considering the real-valued feedback, vector $\boldsymbol{e}$ can be solved by equation (3.13) where vector $\boldsymbol{\Delta}_{\mathrm{r}}$ is defined as $\boldsymbol{\Delta}_{\mathrm{r}} = \boldsymbol{z}_{\mathrm{r}} - \boldsymbol{y}_{\mathrm{r}}$. For our derivation, assuming the above definitions, vector $\boldsymbol{y}_{\mathrm{r}}$ is equal to vector $\boldsymbol{r}$ and we need to redefine vector $\boldsymbol{z}_{\mathrm{r}}$ as

$$\boldsymbol{z}_{\mathrm{r}} = \begin{bmatrix} z_{\mathrm{r}}(t_1) & z_{\mathrm{r}}(t_2) & \dots & z_{\mathrm{r}}(t_N) \end{bmatrix}^T. \tag{5.8}$$

The solution of the updated DPD coefficients can be obtained similarly to equation (3.14) incorporating the above definitions as

$$\begin{bmatrix} \hat{\boldsymbol{b}}_r \\ \hat{\boldsymbol{b}}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}'_r \\ \boldsymbol{b}'_i \end{bmatrix} - \mu (\boldsymbol{A}_z^H \boldsymbol{A}_z)^{-1} \boldsymbol{A}_z^H (\boldsymbol{z}_r - \boldsymbol{r}), \tag{5.9}$$

where

$$\boldsymbol{A}_z = \begin{bmatrix} \boldsymbol{U}_{z\mathrm{r}} & -\boldsymbol{U}_{z\mathrm{i}} \end{bmatrix}. \tag{5.10}$$

5.5 System Implementation Details

Tab. 5.1 shows a system comparison of the proposed architecture with the conventional DPD and DPD with a comparator from [4]. The stated power consumption P_{pwr} does not include components common for all the architectures, i.e. modulators and demodulators, and does not reflect the computational complexity of different adaptation algorithms. The power consumption of the proposed architecture does not include the circuit for the edge time extraction described below. We can see that for DACs with $F_{\text{s}} = 500\,\text{MSps}$ the total power P_{pwr} of the proposed architecture is only half of the power of the DPD from [4] and more than four times smaller than the power of the conventional DPD.

Tab. 5.1: Comparison of the system parameters for the proposed and conventional DPD.

	Convent.	**DPD [4]**	**Proposed**	
	DAC			
Model	AD9779	AD9779	AD9779	AD9136
Quantity	1	2	1	1
F_{s} (MSps)	500	500	500	**2 000**
$\approx P_{\text{pwr}}$	0.6 W	1.2 W	0.6 W	1.45 W
	ADC	**Comparator**		
Model	AD9684	ADCMP553	ADCMP553	ADCMP573
Quantity	1	2	1	1
F_{s} (MSps)	500	500	750 MHz*	**8 GHz***
$\approx P_{\text{pwr}}$	2.2 W	0.12 W	0.06 W	0.2 W
$\approx \sum \boldsymbol{P_{\text{pwr}}}$	**2.8 W**	**1.32 W**	**0.66 W**	**1.65 W**

* Sampling frequency is not applicable, the equivalent input bandwidth of the comparator is stated instead.

5.5.1 Edge Time Extraction

The most straightforward method for extracting the comparator output edge time stamps is to sample the comparator output with a digital *flip-flop* (FF) at a very fast clock frequency. In current FPGAs, regular input pins can be sampled with

clock frequencies of $f_{\mathrm{clk}} \leq 1250\,\mathrm{MHz}$ [195]. It should be noted that the high clock frequency is required only for the input deserialiser and it does not necessarily imply a high operational frequency of the FPGA core. Consequently, in this case the maximum uncertainty of the edge time would be the clock period $T_{\mathrm{clk}} = 1/f_{\mathrm{clk}} > 0.8\,\mathrm{ns}$.

Another possible approach to edge time extraction is depicted in Fig. 5.2a with a simplified time diagram shown in Fig. 5.2b. In the proposed circuit, two FFs sample the comparator output with and without a certain delay t_{d}. The edge detection is ensured only if the delayed edge arrives at input D1 after the active clock edge plus the FF hold time t_{h} and the undelayed edge arrives at input D2 before the active clock edge minus the FF setup time t_{su}. This results in $\mathrm{Q2} = 1$ and $\mathrm{Q1} = 0$ indicating a rising edge, or $\mathrm{Q2} = 0$ and $\mathrm{Q1} = 1$ indicating a falling edge. It can be shown that the maximum time uncertainty of the edge is given as $t_{\mathrm{d}} + t_{\mathrm{su}} + t_{\mathrm{h}}$. For example, the digital FF NB7V52M [196] has $t_{\mathrm{su}} \leq 40\,\mathrm{ps}$ and $t_{\mathrm{h}} \leq 50\,\mathrm{ps}$, which yields the maximum edge-time uncertainty $t_{\mathrm{d}} + 90\,\mathrm{ps}$. The probability that the time stamp is extracted can be approximated as $(t_{\mathrm{d}} + 0.5t_{\mathrm{su}} + 0.5t_{\mathrm{h}})/T_{\mathrm{clk}}$.

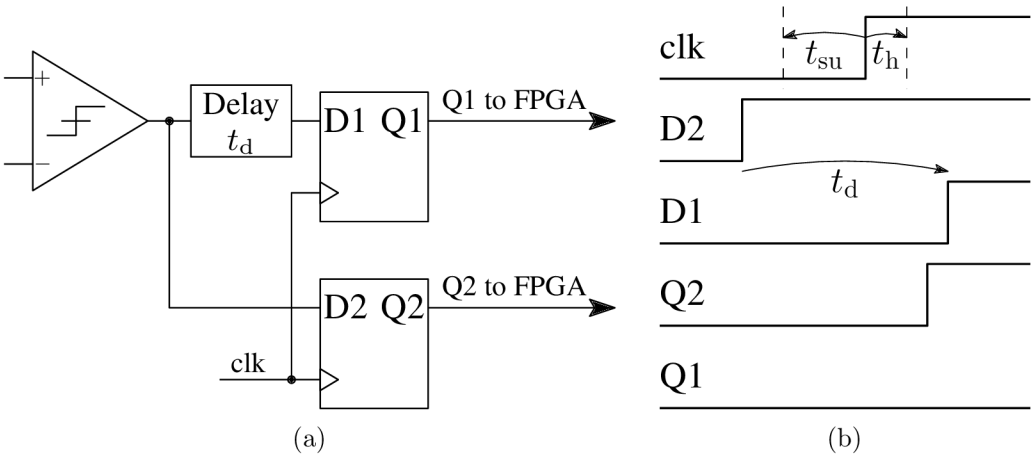


Fig. 5.2: (a) A possible practical circuit implementation for edge time extraction. (b) Waveforms explaining the function of the circuit.

5.6 Measurements

The linearisation performance of the DPD adapted by a comparator in the feedback was evaluated at two distinct setups. The first setup was composed of a laboratory RF generator, a real-time spectrum analyser, and a PA supplemented by an RF attenuator. Even though, its usable bandwidth for our experiment is limited only to $\approx$ 5 MHz, the experiment is valid and can prove the proposed concept. Subsequently, we have also performed measurements with another, wide-band, setup which provides bandwidth of up to 4096 MHz. This wide-band setup has certain

imperfections which cannot be fully compensated for in our laboratory environment and, therefore, the linearisation performance might be degraded. In both setups, we have replaced the feedback comparator with its model created in Matlab. All details, for both narrow-band and wide-band measurements, can be found in Matlab source codes available at `www.github.com/jankralx/comparator_dpd`.

5.6.1 Narrow-Band Measurements

The narrow-band measurement setup is depicted in Fig. 5.3. The vector signal generator, Rohde & Schwarz SMU200A, generates the input signal for the PA. The PA output is attenuated by 40 dB and connected to the real time spectrum analyser Rohde & Schwarz FSVR. The model of an ideal comparator is applied on the recorded IQ signal in Matlab. The Ampere metre, digital multimeter Keysight 34461A, serves to set the drain quiescent current of the PA second-stage transistor.

For the narrow-band measurements, the linearised PA is a two-stage PA in class AB presented in Sec. 4.7.1 as PA1. The drain quiescent current for this measurement was set to 15 mA. The measurements were performed with F-OFDM signals with the 64-QAM as the internal modulation.

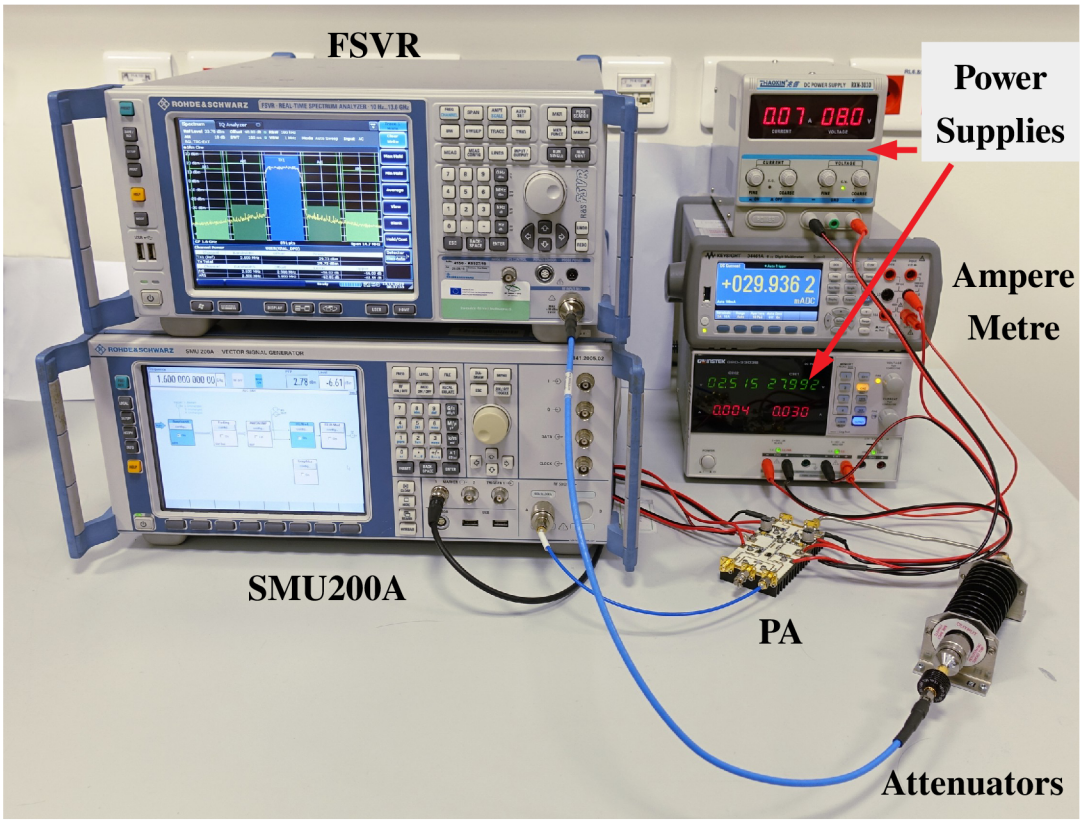


Fig. 5.3: Photograph of the narrow-band measurement setup.

In each iteration, we generate the F-OFDM signal with 64 frames, each frame with 114 resource blocks, a block size of 6 subcarriers, resulting in $W = 6 \cdot 114 = 684$. The FFT size is set to $Z = 8192$ and the filter length is $L = 4097$. The sampling frequency is limited by the measurement setup and is set to $F_\mathrm{s} = 30\,\mathrm{MHz}$. The DFT precoding is employed to reduce the signal PAPR. The signal generator with these settings provides the test signal vector with the channel bandwidth $B \approx 2.5\,\mathrm{MHz}$ and $N_0 = 532\,992$ samples.

DPD for the PA linearisation employs the MP model with $K = 7$ and $Q = 3$, which yields $P = 28$ DPD coefficients. The DLA has driven the DPD adaptation, the iteration step was set to $\mu = 0.05$, and 50 iterations were performed during the measurements. The reference voltage $r(t)$ was swept in equidistant steps over the whole range of the in-phase feedback amplitudes. The step size was set to 7.5% of the full scale.

5.6.2 Wide-Band Measurements

The block diagram of the wide-band measurement setup is depicted in Fig. 5.4 and the real setup is captured in photographs in Fig. 5.5 and Fig. 5.6. The setup is based on the development board ZCU111 [197] with the Zynq Ultrascale+ *radio frequency system-on-chip* (RFSoC) by Xilinx. The RFSoC on ZCU111 provides up to eight RF ADCs with a sampling frequency $\leq 4\,096$ MHz and a nominal resolution of 12 bits, and up to eight RF DACs with the sampling frequency $\leq 6\,554$ MHz and the nominal resolution of 14 bits. For transmitting, two RF DACs generate the baseband IQ signal which is frequency-shifted by IQ mixer HMC8191 [198] directly to the carrier frequency. The carrier frequency is set to $F_c = 9$ GHz and is generated by the RF generator SMF100A [199] from Rohde Schwarz. The signal from the

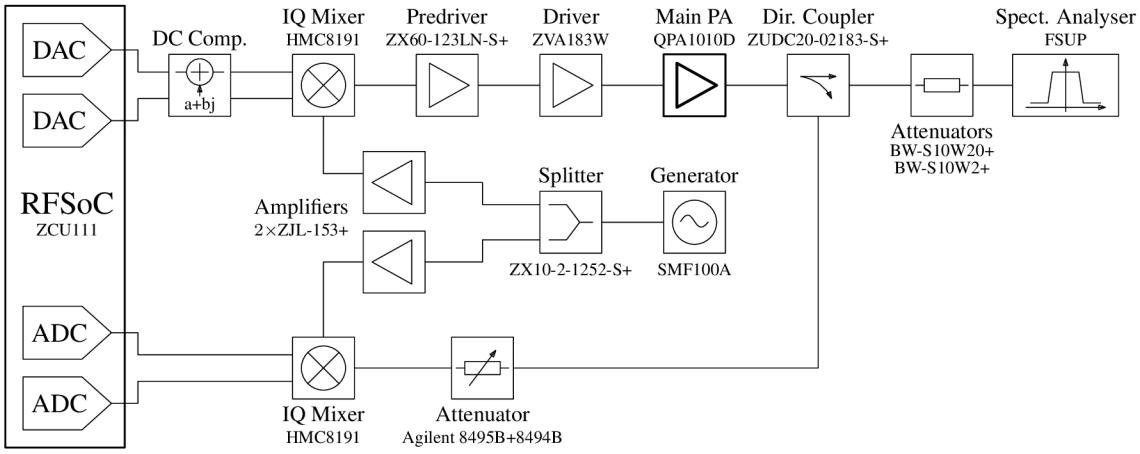


Fig. 5.4: Block diagram of the wide-band measurement setup.

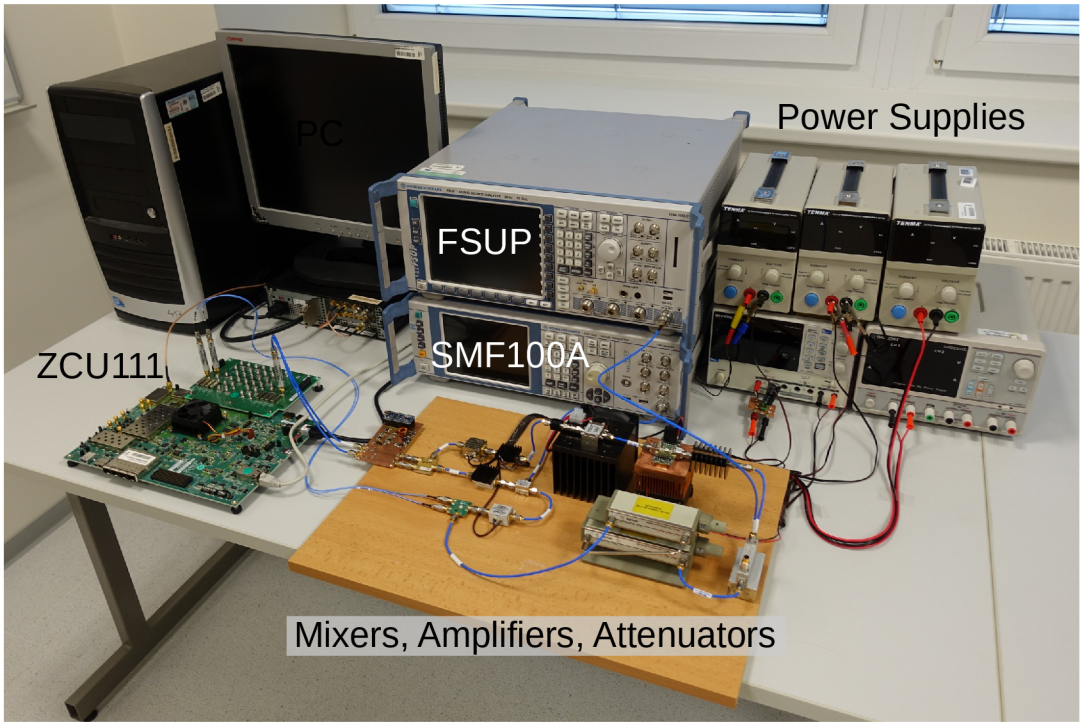


Fig. 5.5: Photograph of the wide-band measurement setup.

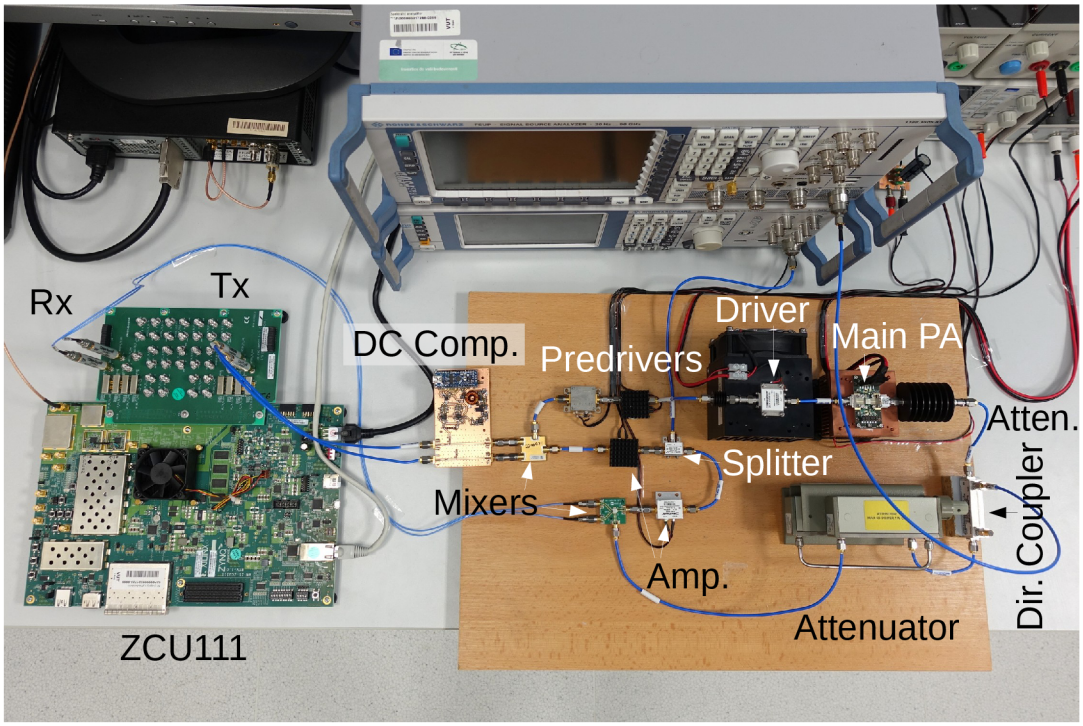


Fig. 5.6: Detailed view on the wide-band measurement setup.

carrier-frequency generator is split to ensure the same frequency for IQ mixers in the direct and feedback paths and is amplified by two amplifiers ZJL-153+ [200] to achieve a sufficient signal level for the IQ mixers. The direct-path IQ mixer is followed by a series of amplifiers which drive the main PA close to the saturation point. The output of the main PA goes through a directional coupler and high-power attenuators to the spectrum analyser FSUP [201] from Rohde Schwarz. The directional coupler probes the main PA output for the feedback path. The feedback signal level is adjusted by two variable attenuators Agilent 8495B and 8494B [202], down-converted by IQ mixer HMC8191, and converted into a digital baseband IQ signal by two RFSoC ADCs.

The main PA is an X-band GaN-on-SiC high power amplifier QPA1010 which operates in the frequency range from 7.9 GHz to 11 GHz and typically provides 15-W saturated output power with power-added efficiency of 38% and large-signal gain of 18 dB [203]. Hereinafter, we will denote this PA as PA4. The test signals were the same as for the narrow-band measurements, except for the sampling frequency set to $F_\mathrm{s} = 2048\,\mathrm{MHz}$, number of resource blocks set to 333 and the number of frames set to 16. These settings result in the test signal with the channel bandwidth $B \approx 500\,\mathrm{MHz}$ and $N_0 = 136\,320$ samples.

The DPD for the PA linearisation employs the DDR2 model with $K = 9$ and $Q = 9$, which yields $P = 121$ DPD coefficients. The DLA has driven the DPD adaptation, the iteration step was set to $\mu = 0.2$, and 50 iterations were performed during the measurements. The reference voltage $r(t)$ was swept in the same way as for the narrow-band measurements.

5.7 Measurement Results

5.7.1 Results of Narrow-Band Measurements

The narrow-band measurement results of the ACPR, EVM, and NMSE evolution over the adaptation iterations are shown in Fig. 5.7, Fig. 5.8, Fig. 5.9, respectively. We can see that the results of the proposed DPD adaptation with the comparator in the feedback fairly follow the conventional DPD adaptation results. For both predistorters, the adaptation reaches a steady state in approximately 20 iterations. The horizontal dashed black line depicts the metrics of the transmitter without the DPD for the same output power in the main channel. The measured PSD for the proposed and conventional DPD compared to the PA output PSD without DPD is shown in Fig. 5.10.

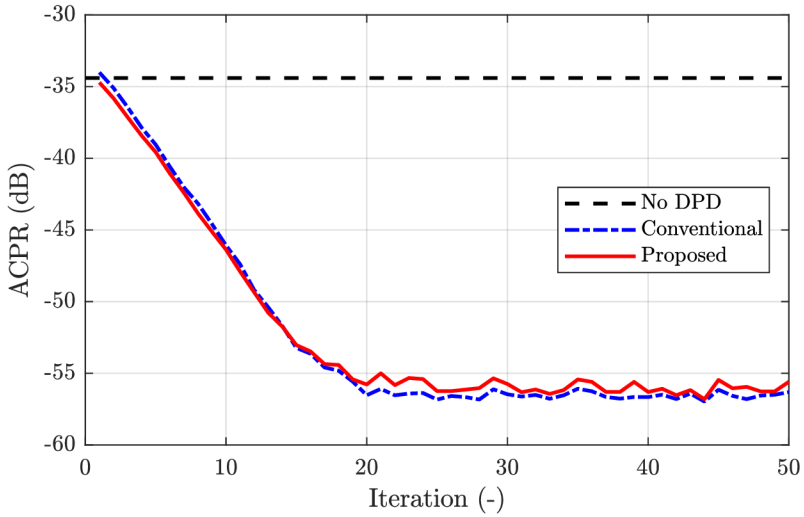


Fig. 5.7: Evolution of the ACPR during the DPD adaptation with the narrow-band setup.

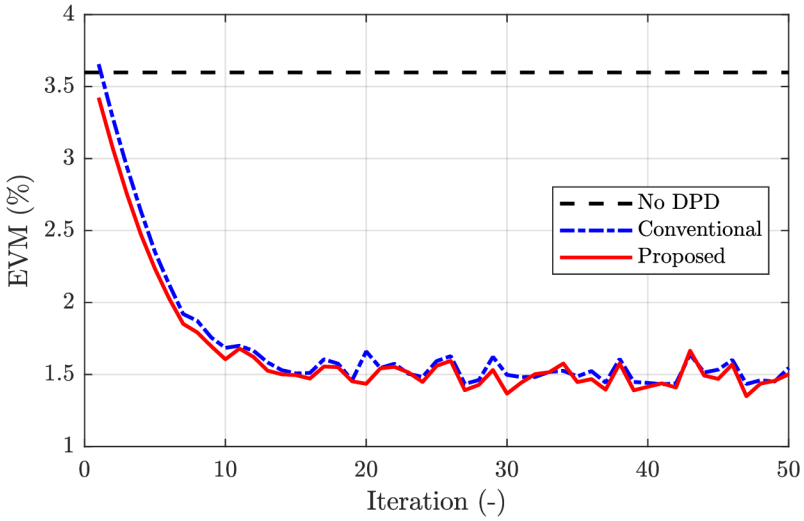


Fig. 5.8: Evolution of the EVM during the DPD adaptation with the narrow-band setup.

Tab. 5.2 shows the performance of adapted DPD in terms of NMSE, EVM, and ACPR. The results are averaged over the last 10 iterations. The proposed DPD provides comparable linearisation performance to the conventional DPD. It is to be noted that the conventional DPD performs the adaptation with the whole test signal vector containing 532 992 samples, whereas the proposed DPD adaptation is limited by the occurrences of the signal transitions over the reference voltage $r(t)$ resulting in the DPD adaptation with approximately $4000 - 5000$ samples.

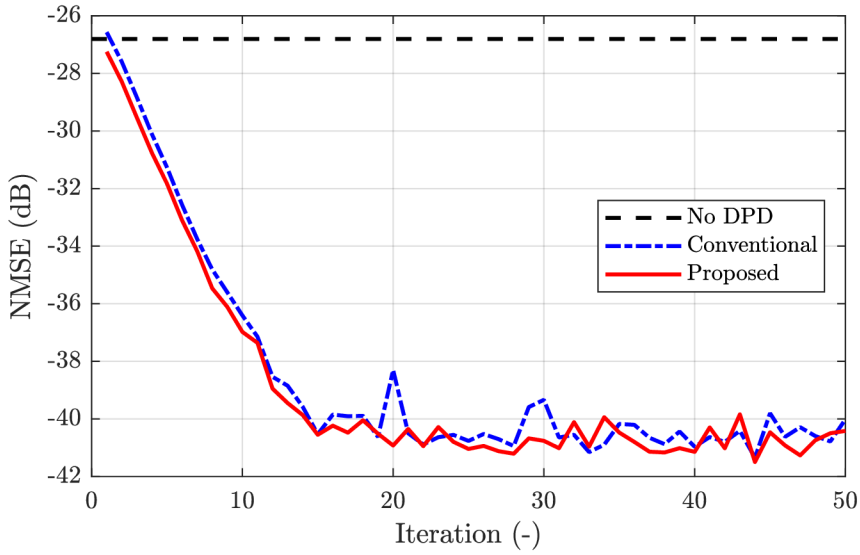


Fig. 5.9: Evolution of the NMSE during the DPD adaptation with the narrow-band setup.

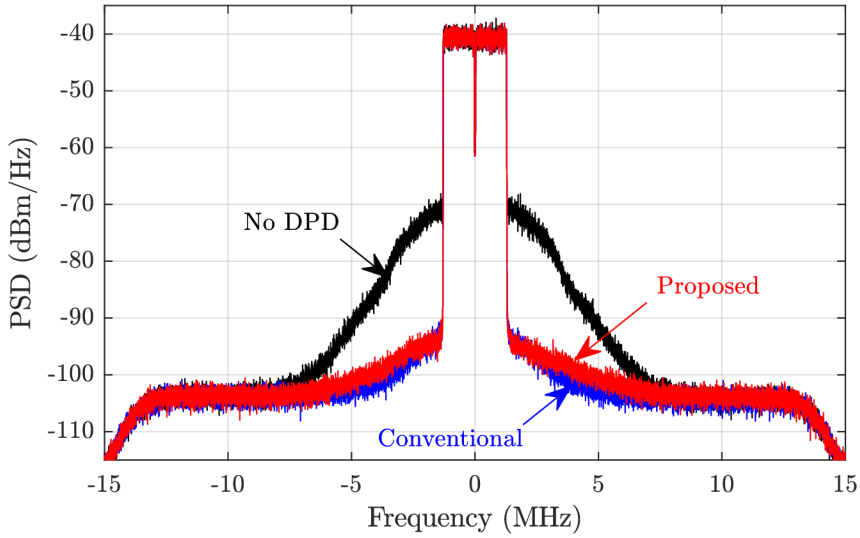


Fig. 5.10: Comparison of the measured power spectral density of the PA output without DPD and with the conventional and proposed DPD with the narrow-band setup.

5.7.2 Results of Wide-Band Measurements

Fig. 5.11 and Fig. 5.12 depict the AM/AM and AM/PM characteristics of the PA with the wide-band setup, adapted DPD, and linearised transmitter. The PA under test exhibits strong nonlinearity and even the smaller signal magnitudes are compressed. The memory effects of the PA widen the AM/AM characteristics of the

Tab. 5.2: Narrow-band measurement results of NMSE, EVM, and ACPR for the DPD with the proposed LC-ADC and conventional feedback.

	P_{main} (dBm)	NMSE (dB)	EVM (%)	ACPR (dB)
Without DPD	23.2	-26.8	3.60	-34.4
Conventional	23.3	-40.5	1.51	-56.6
LC-ADC	23.3	-40.7	1.48	-56.1

EVM of the generated signal is 1.25% due to the nonorthogonality caused by the inherent F-OFDM filtering.

PA and we can see that the employed DPD could not fully compensate for them. We believe that uncompensated memory effects are of a long-term character and are beyond the maximum memory length of the employed DPD. The straight-forward option of increasing maximum memory length would be very impractical, because due to high sampling frequency, capturing the long-term memory effects would drastically increase the number of DPD coefficients. The AM/PM characteristic of the PA is almost flat and its contribution to the overall PA nonlinearity is negligible.

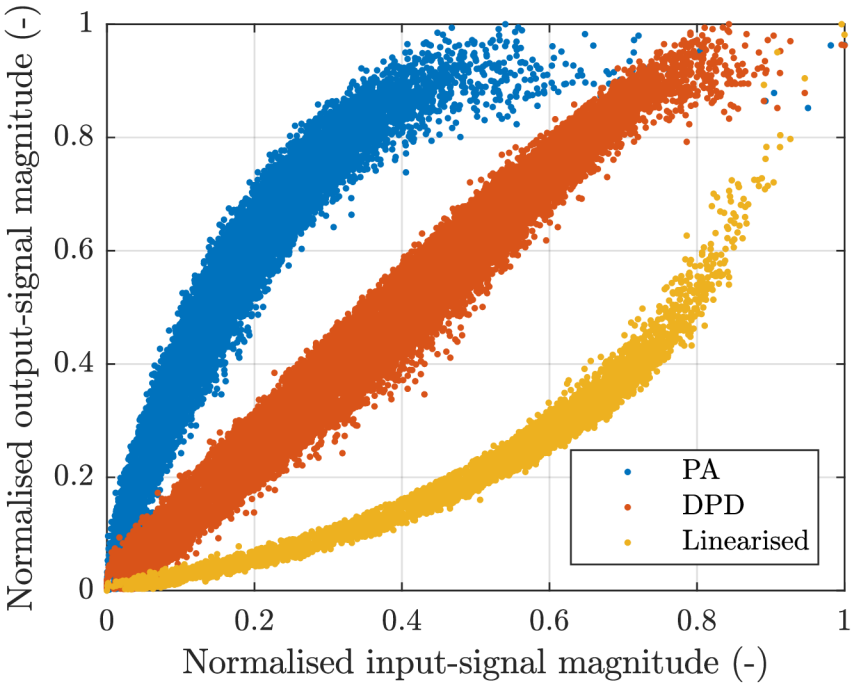


Fig. 5.11: Measured AM/AM characteristics of the PA, DPD and of the linearised transmitter.

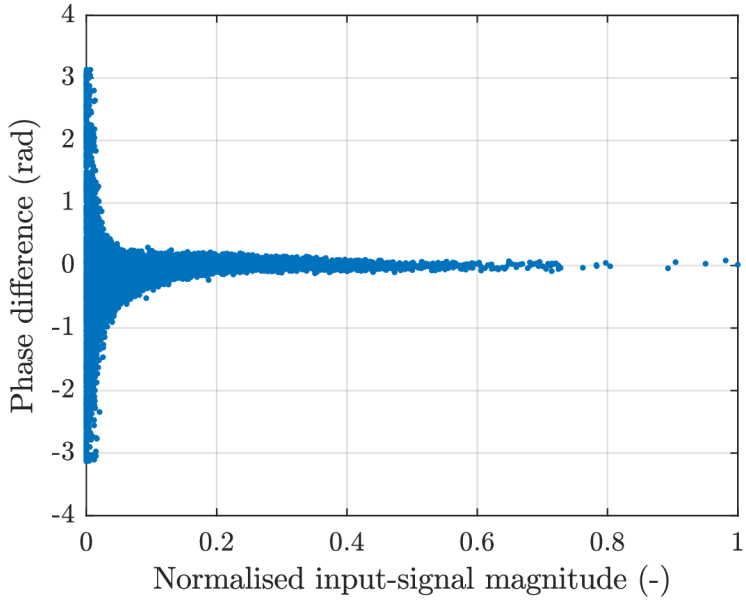


Fig. 5.12: Measured AM/PM characteristics of the PA. (The AM/PM characteristics of the DPD and of the linearised transmitter are not plotted, because all the characteristics overlap.)

The evolution of ACPR, EVM, NMSE during the DPD adaptation is shown in Fig. 5.13, Fig. 5.14, Fig. 5.15, respectively. The proposed adaptation with a comparator in the feedback achieves slightly higher (worse) ACPR and NMSE than the conventional DPD adaptation. Both adaptation methods achieve a similar EVM of the transmitted signal. Both predistorters reach a steady state in approximately 15 iterations. The horizontal dashed black lines depict the metrics of the PA without DPD for the same output power in the main channel. The measured PSD for the proposed and conventional DPD compared to the PA output PSD without DPD is shown in Fig. 5.16.

The achieved linearisation performance metrics are summarised in Tab. 5.3. The results are averaged over the last 10 iterations. Again, the conventional DPD adaptation calculates with the whole test signal vector, whereas the proposed DPD adaptation is limited by the occurrences of the signal transitions over the reference voltage $r(t)$. Due to that the linearisation performance of the DPD with a comparator might be slightly limited.

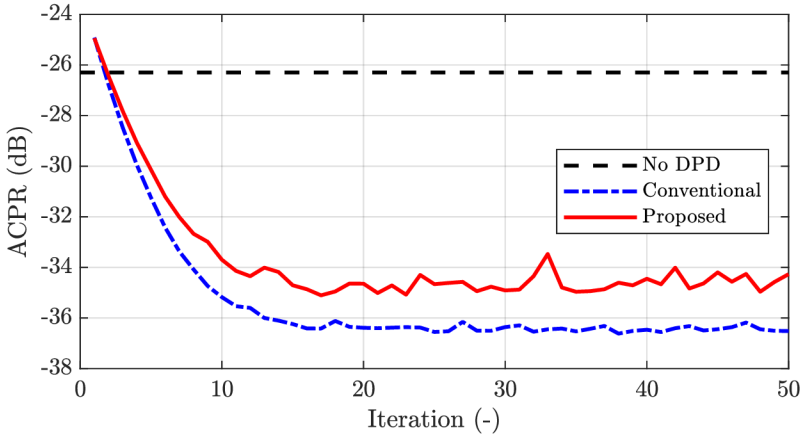


Fig. 5.13: Evolution of the ACPR during the DPD adaptation with the wide-band setup.

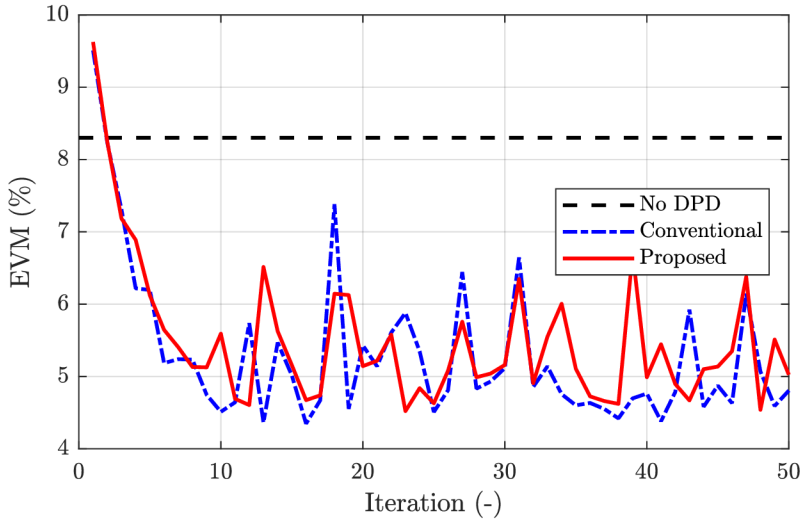


Fig. 5.14: Evolution of the EVM during the DPD adaptation with the wide-band setup.

5.8 Measurement Setup for DPD Adaptation with a Hardware LC-ADC

To prove the proposed concept of the DPD adaptation with an LC-ADC in the feedback, we have built up a measurement setup, which is depicted in a block diagram in Fig. 5.17 and in the photograph in Fig. 5.18. The setup is based on the Spartan-3A DSP Starter Board [204], EXP High-Speed DAC Converter Module from Avnet (with dual-channel DAC DAC5682Z, hereinafter denoted also as the DAC module) [205], our custom-designed feedback module with a real hardware LC-ADC, two RF signal generators for generating the carrier signal and reference clock, and a real-time spectrum analyser FSVR from Rohde-Schwarz [206].

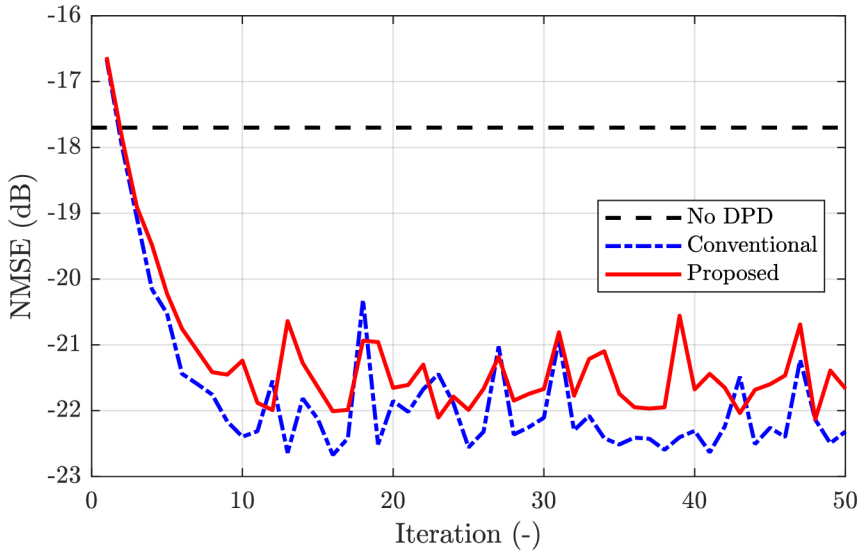


Fig. 5.15: Evolution of the NMSE during the DPD adaptation with the wide-band setup.

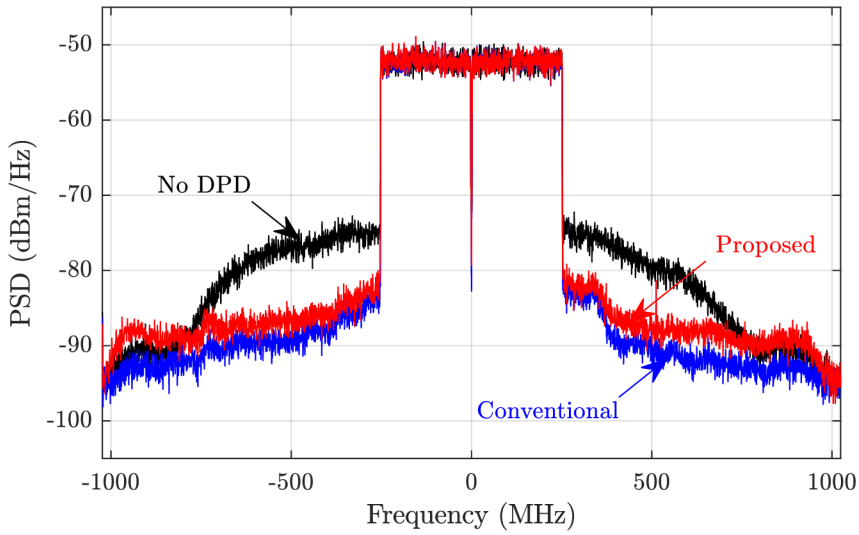


Fig. 5.16: Comparison of the measured power spectral density of the PA output without DPD and with the conventional and proposed DPD with the wide-band setup.

The simplified block diagram of the DAC module is depicted in Fig. 5.19. It contains the dual-channel DAC DAC5682Z [207], reconstruction LPFs, and up-converting IQ mixer TRF3703 [208]. The DAC sampling clock is fed from the feedback module and the carrier frequency signal from an external RF generator. The output of the mixer is connected to the PA.

Tab. 5.3: Wide-band measurement results of NMSE, EVM, and ACPR for the DPD with the proposed LC-ADC and conventional feedback.

	P_{main} (dBm)	NMSE (dB)	EVM (%)	ACPR (dB)
Without DPD	34.8	-17.7	8.31	-26.3
Conventional	34.8	-22.2	5.01	-36.4
LC-ADC	34.9	-21.6	5.23	-34.5

EVM of the generated signal is 1.25% due to the nonorthogonality caused by the inherent F-OFDM filtering.

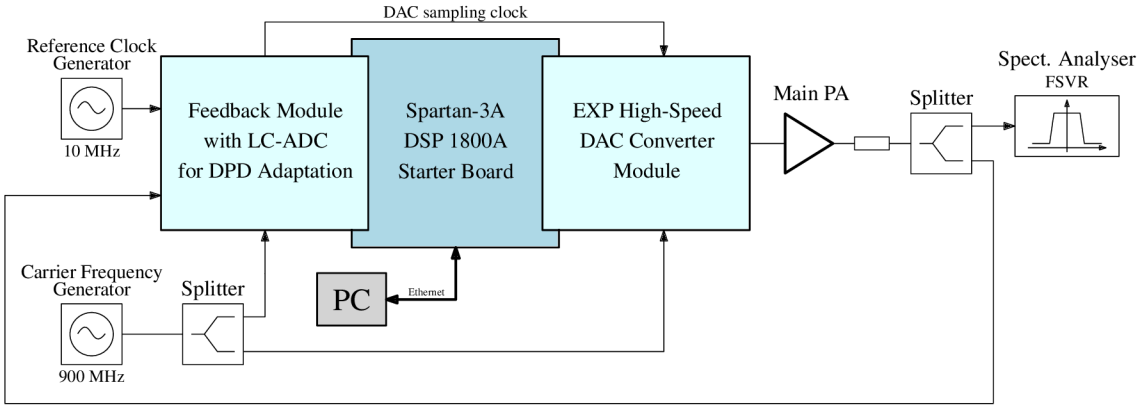


Fig. 5.17: Block diagram of the measurement setup for the DPD adaptation with a hardware LC-ADC.

We had to design and build our own feedback module with a real hardware LC-ADC. The module contains the comparator, an edge-time extraction circuit as described in Sec. 5.5.1 and in Fig. 5.2a, feedback down-converting IQ mixer, and the clock distribution system. The block diagram of the designed module is shown in Fig. 5.20. The realised feedback module is captured in photographs in Fig. 5.21. The feedback signal from the main PA is down-converted to the baseband by IQ mixer TRF371125 [209] and its real differential output is converted to a single-ended signal by an active balun. We denote this signal $\hat{y}_{\text{r}}(t)$, ideally it would be a baseband PA output, however, the down-converting mixer and active balun add certain distortion, which is noted by the hat. A high-speed comparator ADCMP582 [210] compares the feedback signal $\hat{y}_{\text{r}}(t)$ with a generated reference voltage $r(t)$. The comparator output is sampled by two FFs NB7V52 [196]. One FF is connected directly to the

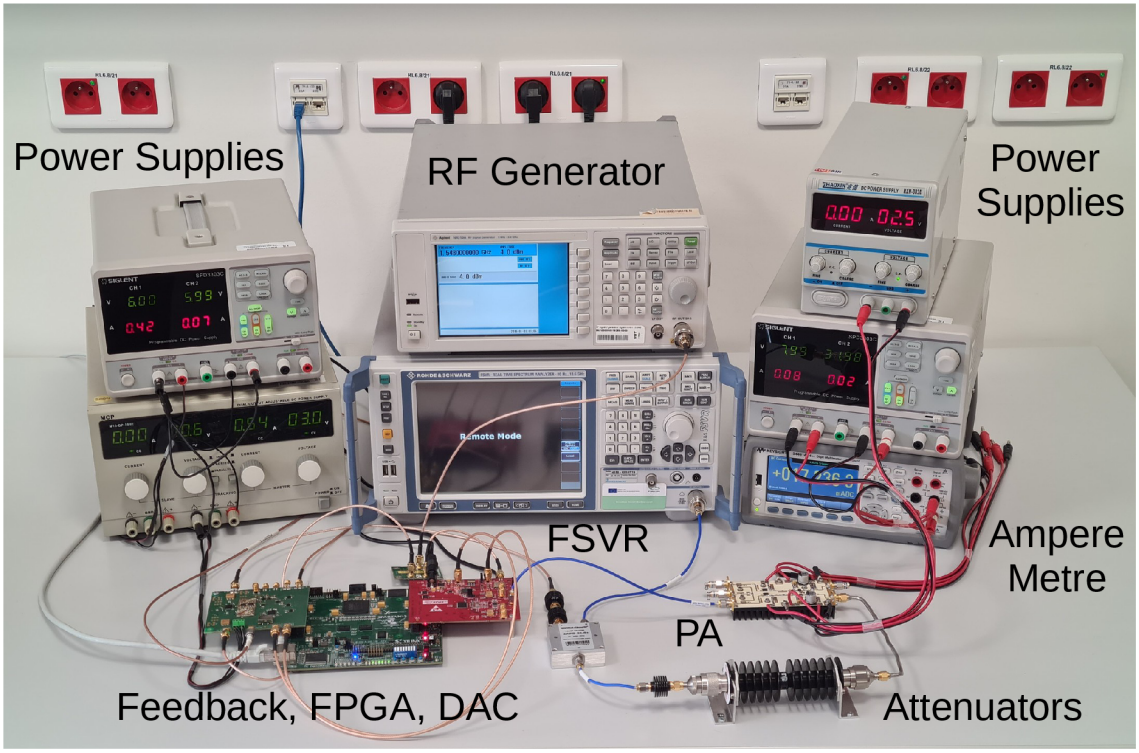


Fig. 5.18: Photograph of the measurement setup for the DPD adaptation with a hardware LC-ADC. Transmission signal is generated by the DAC module and the FPGA and amplified by the main PA. The PA output is attenuated and the splitter divides the signal for the real-time spectrum analyser FSVR and the feedback module with the comparator. The RF generator generates the carrier frequency.

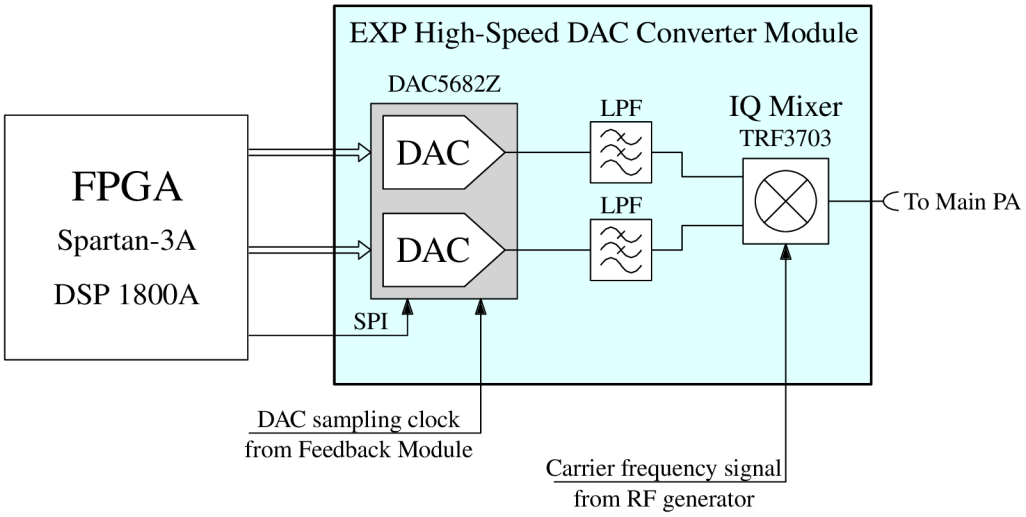


Fig. 5.19: Simplified block diagram of the high-speed DAC converter module.

comparator output, the other one is connected with a certain small delay τ created by *printed circuit board* (PCB) traces. The FF outputs are converted from the *current-mode logic* (CML) standard to the *low-voltage differential signaling* (LVDS) standard and connected to the FPGA for further processing. The clocks for the FPGA, DAC sampling, and the feedback FFs are generated in a jitter cleaner and clock distributor LMK04133 [211]. This is the only clock source for the system (other clock sources are not employed or are bypassed) which ensures the clock coherency and minimises problems with clock-domain signal crossing. More details about the hardware design of the feedback module can be found in [212].

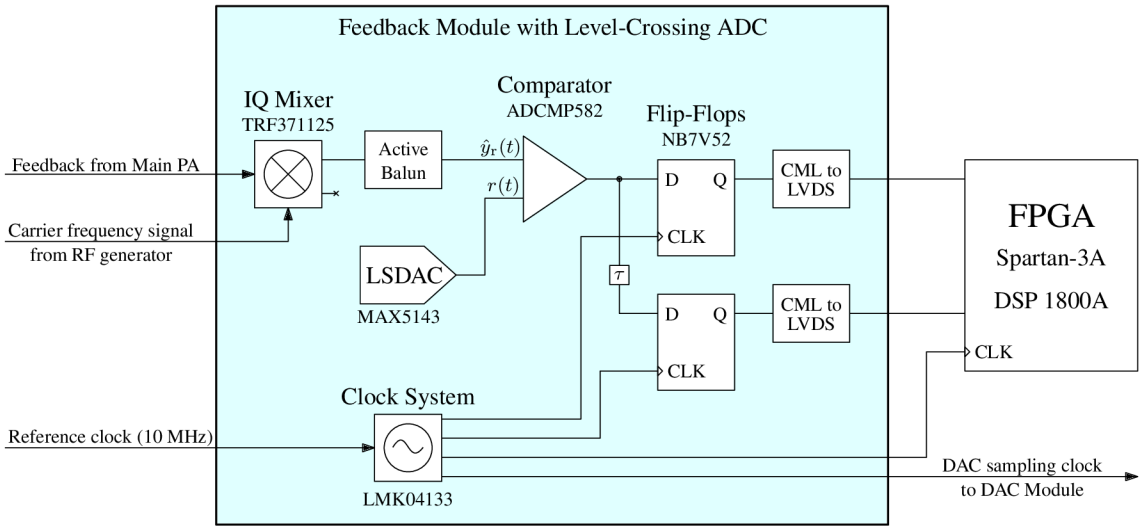
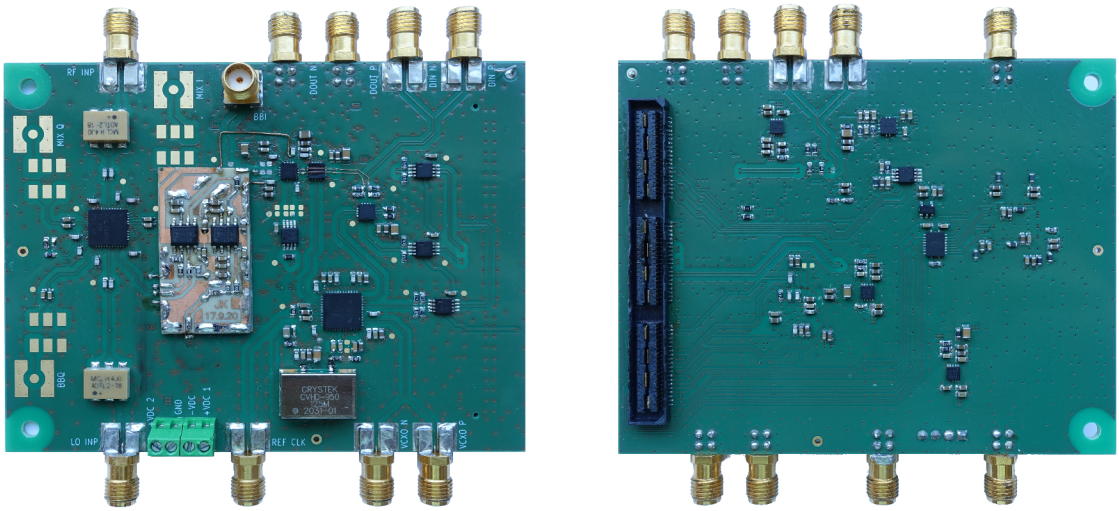


Fig. 5.20: Block diagram of the feedback module with LC-ADC.



(a) Top view. (b) Bottom view.

Fig. 5.21: Photographs of the realised the feedback module.

5.8.1 Operating Procedures

The measurement setup operating procedures are executed in two places. Low-level operations related to the hardware of the DAC module and feedback module are managed by the FPGA firmware which contains a soft core processor Microblaze [213] for hardware initialisation and the FPGA logic dedicated to high-performance operations such the signal transmission and the feedback acquisition. High-level operations related to the signal generation, DPD adaptation and predistortion are managed by Matlab on a connected *personal computer* (PC).

The measurement setup, after its start-up, has to be initialised. The FPGA processor takes these steps to initialise the hardware:

- initialising the clock generating system with LMK041333 on the feedback module,
- bypassing clock generator CDCM7005 on DAC module,
- initialising the main DAC DAC5682Z on DAC module,
- initialising the down-converting IQ mixer TRF371125 on the feedback module,
- initialising Ethernet modules,
- and starting the main application server.

The main application server receives and processes commands from the PC. It is based on the lwIP stack and *user datagram protocol* (UDP) and implements a simple custom-designed protocol which ensures reliable data delivery. We have also experimented with *transmission control protocol* (TCP) which would inherently ensure the communication reliability, but we have found the soft processor Microblaze to be underperforming for sufficient high-throughput TCP communication.

The typical single measurement with the setup from the Matlab perspective consists of

- opening a UDP socket,
- setting the number of samples to be transmitted and the operating mode,
- sending the LSDAC waveform – the waveform of the reference voltage $r(t)$,
- sending the samples of signal $x(t)$ which is up-converted and fed into the main PA,
- sending a request for the measurement results,
- obtaining the measurement results from the FPGA,
- compensating for the setup delay,
- and applying the calibration compensations.

The application server in Microblaze receives the LSDAC waveform and stores it in a dedicated fast on-chip *block random-access memory* (BRAM). The BRAM is read during the measurement by the FPGA logic and sent to the LSDAC to create the desired reference voltage waveform $r(t)$. The main-DAC samples cannot be re-

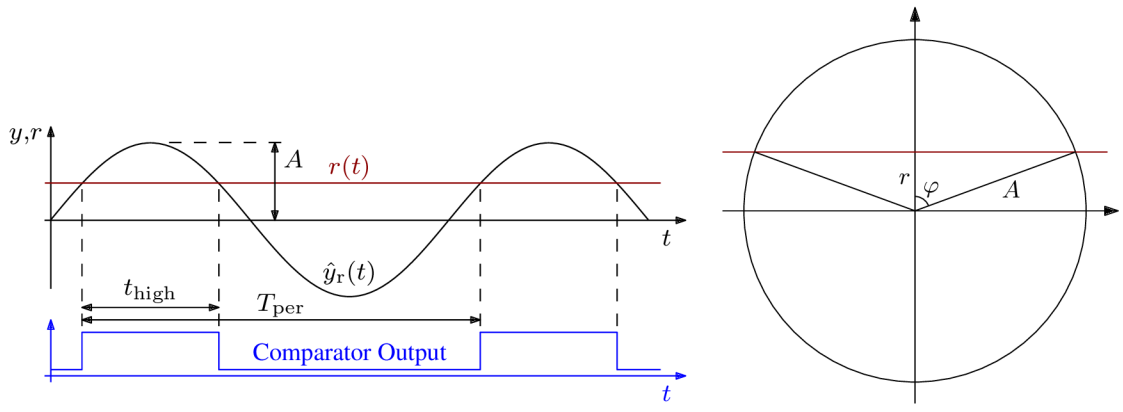
ceived via Ethernet and sent to the DAC instantly. Therefore, the application server receives the samples and stores them in a dedicated space of the external *double data rate* (DDR) *synchronous dynamic random-access memory* (SDRAM). When all the DAC samples are received, the application server triggers the measurement. The FPGA logic starts reading DAC samples from the DDR SDRAM and transmits them via the main DAC and an up-converting IQ mixer to the PA. Simultaneously, the FPGA logic monitors the output of the two FFs on the feedback module, and when it detects an edge, it saves the edge timestamp and polarity into a dedicated BRAM. When the measurement is finished, the application server reads the edge timestamps and polarities and sends them to Matlab for further processing.

5.8.2 Reference Voltage Amplitude Calibration

To calibrate the reference voltage set by the LSDAC, we have proposed a method based on measuring the duty cycle of the comparator output expecting a periodic sinewave test signal as its input. If the test signal is sufficiently low to not be distorted by the PA, the comparator input $\hat{y}_\mathrm{r}(t)$ can be considered fully known as it is approximately equal to the transmitted signal, i.e. $\hat{y}_\mathrm{r}(t) \approx y_\mathrm{r}(t) \approx x_\mathrm{r}(t)$.

The principle of the method is shown in Fig. 5.22. The method can be seen in the following way: The known periodic sinewave signal $\hat{y}_\mathrm{r}(t)$ crosses the unknown reference voltage $r(t)$ and generates a square wave at the comparator output (Fig. 5.22a). The duty cycle of the square wave is determined by a ratio of the reference voltage $r(t)$ and the amplitude A of the test sinewave signal.

The test signal period is preferred to be long, because in that case the small time errors of the edge time extraction circuit can be neglected. The described method is independent of the delay calibration, as measuring the duty cycle requires no information about the system delays.



(a) Illustration in the time. (b) Explanation in the unit circle.

Fig. 5.22: Principle of reference voltage calibration by measuring the duty cycle for a sinewave test signal.

Formally, we can make the following derivations: Let us expect that the real-part feedback is a scaled sinewave signal

$$\hat{y}_{\mathrm{r}}(t) = A \sin\left(2\pi\, T_{\mathrm{per}}^{-1}\, t\right), \tag{5.11}$$

where T_{per} is its period and A is its amplitude. The set reference voltage $r(t)$ is constant during the measurement, thus $r(t) = r$. We can calculate the reference voltage r due to Fig. 5.22b as

$$r = A\cos(\varphi), \tag{5.12}$$

and we know that

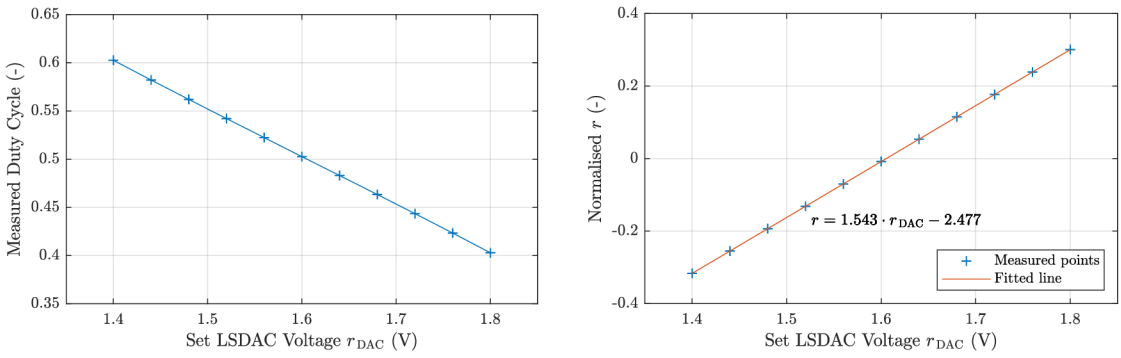
$$\frac{2\varphi}{2\pi} = \frac{t_{\mathrm{high}}}{T_{\mathrm{per}}}, \tag{5.13}$$

which yields the final solution of r to be

$$r = A\cos\left(\frac{t_{\mathrm{high}}}{T_{\mathrm{per}}}\,\pi\right). \tag{5.14}$$

Fig. 5.23 shows the measured duty cycle $\frac{t_{\mathrm{high}}}{T_{\mathrm{per}}}$ of the comparator output and the normalised reference voltage r as functions of the set LSDAC voltage r_{DAC}. The reference voltage r is normalised with respect to the test signal amplitude A. We can observe that the reference voltage r is a linear function of the set LSDAC voltage r_{DAC}. We can fit a line into the measured points; its equation determines the relation between the reference voltage r and the set LSDAC voltage r_{DAC} as

$$r = (1.543 \cdot r_{\mathrm{DAC}} - 2.477) \cdot A. \tag{5.15}$$



(a) Measured duty cycle.

(b) Normalised reference voltage r.

Fig. 5.23: Measured duty cycle and the reference voltage as functions of the set LSDAC voltage.

5.8.3 Edge-Detection Delay Calibration

In real hardware, DACs, up-converting and down-converting mixers, PA itself, acquisition circuit in the feedback, and the signal propagation cause a certain delay between the transmitted digital baseband signal $x(t)$ and the respective observed feedback signal $y(t)$. Practically, this delay is often considered constant and has to be measured once to be compensated for before the DPD adaptation.

We have proposed a measurement method for the delay between the digital transmitter and the feedback receiver in two steps. The first step is a coarse delay measurement, when the edge acquisition circuit works in coarse mode where it detects all edges with higher time uncertainty. This mode allows us to measure the approximate delay and perform a fine delay measurement. The signal waveform transmitted for the coarse delay measurement is depicted in Fig. 5.24. The coarse delay is computed by fitting a transmitted signal to the recovered received signal by finding the LS optimum.

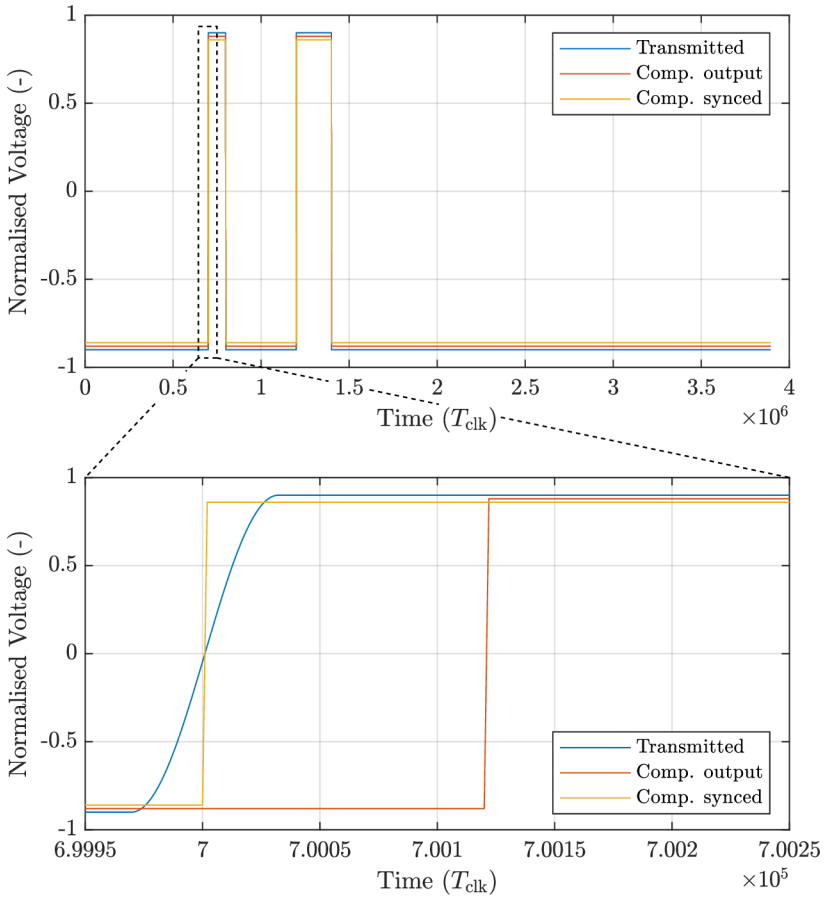


Fig. 5.24: The transmitted signal waveform and received comparator output before and after synchronisation during the coarse delay measurements.

The second step is a fine delay measurement which measures the delay with a resolution lower than the receiver clock period T_{clk}. During this measurement, the edge acquisition circuit works with precise edge time detection, as proposed in Fig. 5.2a. The transmitter generates a signal waveform containing artificial edges at time instants $nT_{\mathrm{clk}} + t_{\mathrm{f}}$, where t_{f} is a subperiod delay $t_{\mathrm{f}} < T_{\mathrm{clk}}$ between the nT_{clk} time instant and the interpolated transmitted signal crosses the set reference voltage $r(t)$. The edge subperiod delay is explained in Fig. 5.25. The edge acquisition circuit in the precise edge time detection mode does not register all edges, but only edges which belong within a certain t_{f} interval. Due to the operation of the edge acquisition circuit, edge detection cannot be ensured at the boundaries of the t_{f} interval. Consequently, we employ a signal waveform with a delay of t_{f} swept in steps in interval $[-0.5\,T_{\mathrm{clk}}, 0.5\,T_{\mathrm{clk}}]$ with a step size equal to $0.01T_{\mathrm{clk}}$. Several edges are generated for each step which should increase the edge-detection probability at the boundaries of the interval of interest. A practical example of the detected edges by fine delay measurement is illustrated in Fig. 5.26. The results can be processed as the histogram shown in Fig. 5.27.

The depicted t_{f} interval of edge detection defines the interval of the fine delay between the transmitter and the feedback receiver. For the DPD adaptation, we reduce the delay interval with a single-value delay represented by the interval centre. The overall measured delay is employed for the delay compensation before the DPD adaptation. Shifting the feedback signal $y(t)$ in time is not achievable, as we only know the feedback signal in particular discrete time instants t_i and, generally, signal $y(t)$ cannot be fully recovered from samples $y(t_i)$. Instead, the delay compensation is achieved by time shifting the transmitted signal $x(t)$ which is fully known. In our measurement setup, the delay is constant after the setup powers on, therefore, we can measure it once during the initialisation phase.

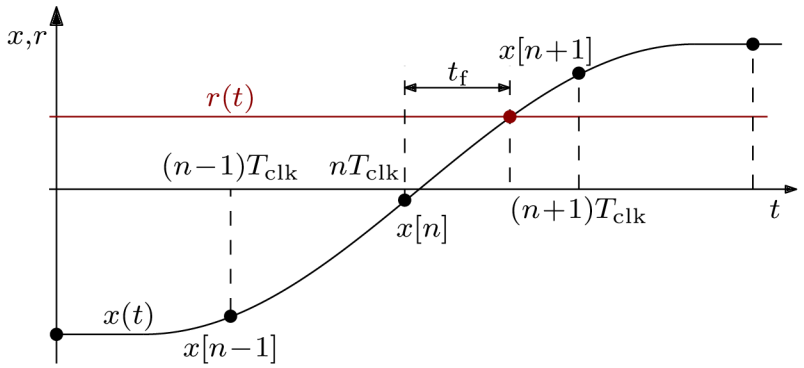


Fig. 5.25: The graphical explanation of the edge subperiod delay. The signal samples $x[n]$ form the reconstructed signal edge $x(t)$ which intersects the reference voltage $r(t)$ at the set time instant $nT_{\mathrm{clk}} + t_{\mathrm{f}}$.

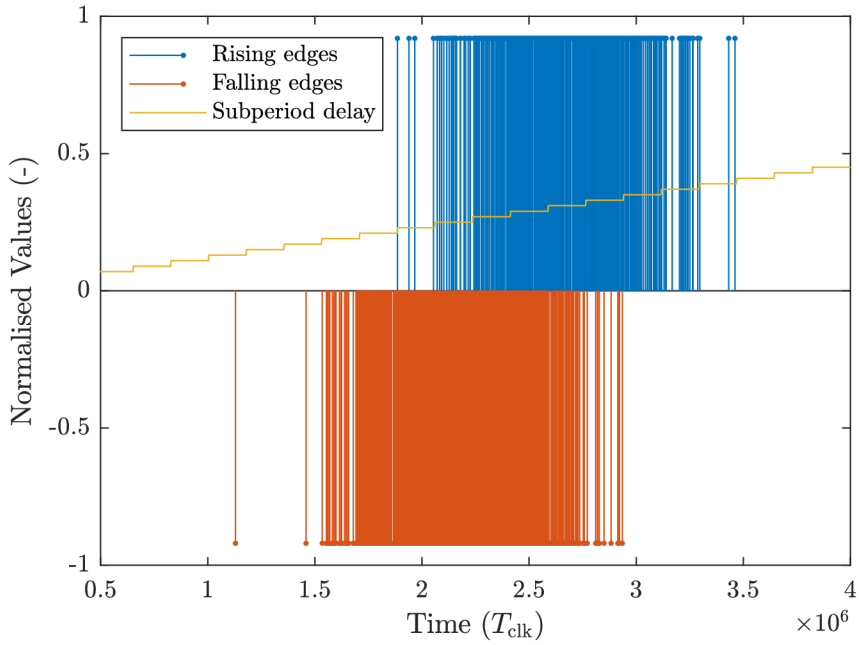


Fig. 5.26: Rising and falling edges at the comparator output based on the variable generated subperiod delay. The subperiod delay has been normalised to T_{clk}.

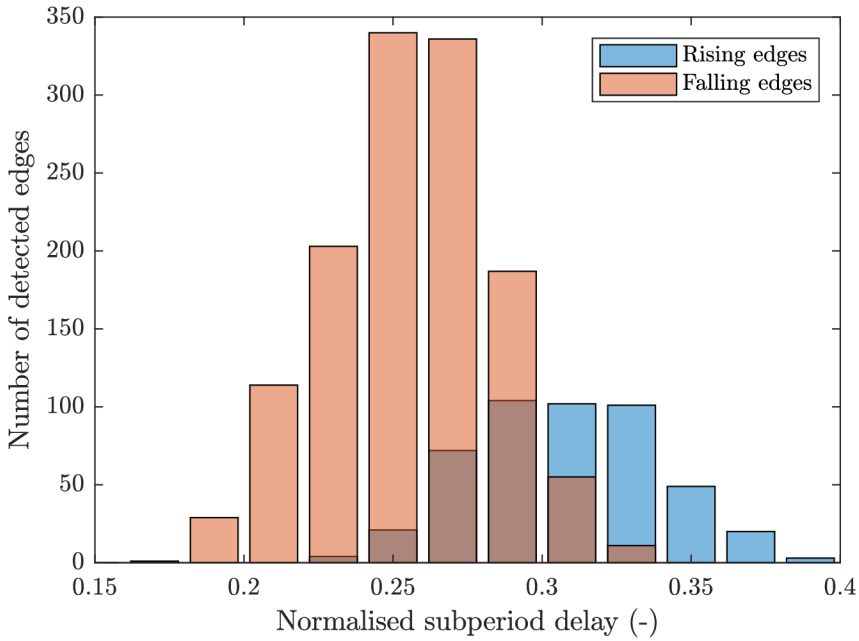


Fig. 5.27: Histogram of the detected rising and falling edges at the comparator output as the function of the generated subperiod delay. The subperiod delay has been normalised to T_{clk}.

5.8.4 Optional LS Compensations

The test measurements with the amplitude and the delay compensated for as described above have shown a sub-optimal performance of the feedback receiver. An analysis showed biased errors in the amplitude and delay calibration. Therefore, we have decided to tune the delay and the amplitude based on the measurement with a real transmission signal. We have employed a two-parameter optimisation with the *mean square error* (MSE) as the optimisation criterion.

Although, by employing the described optional LS compensations, we have achieved a lower MSE for the feedback-sample acquisition, the real DPD adaptation attempts showed that this optional compensation decreased the linearisation performance. We believe that the introduced LS compensation well compensated for the average residual delay mismatch and the amplitude offset for the individual comparison levels, but at the same time, it introduced very small delay mismatches and amplitude offsets which were different for the individual comparison levels. Gotthans et al. [194] and Liu et al. [76] showed that the DPD with memory modelling ($M > 0$) can intrinsically compensate for a small feedback path delay. Naturally, the DPD can compensate for linear gain in the feedback. Due to these findings, we have not employed the proposed optional LS compensation for the DPD adaptation.

5.8.5 Potential Future Improvement

Although, the presented a simple circuit for edge-time extraction with two FFs works sufficiently well for our measurement setup; its inherent time uncertainty caused by setup and hold times of the FFs is the main limiting factor for higher achievable bandwidth. One could overcome these limitations in the future by employing a more sophisticated method for edge-time extraction. We would suggest employing a time-to-digital converter. Mattada and Guhilot [214] presented a comprehensive review of various time-to-digital converters. Some of these converters achieve time resolution being better than 1 ps and, simultaneously, very low power consumption of a few milliwatts.

5.9 Measurement Results for DPD Adaptation with a Hardware LC-ADC

The measurement results presented in this chapter have been obtained employing the hardware introduced in Sec. 5.8. The method labelled as conventional obtains the feedback samples from the spectrum analyser FSVR, while the proposed method employs the feedback comparator. For all these measurements, the linearised PA is

the two-stage PA in class AB presented in Sec. 4.7.1 as PA1. The drain quiescent current for this measurement was set to 15 mA. The measurements were performed with F-OFDM signals with the 64-QAM as the internal modulation.

In each iteration, we generate the F-OFDM signal with 384 frames, each frame with 10 resource blocks, a block size of 6 subcarriers, resulting in $W = 6 \cdot 10 = 60$. The FFT size is set to $Z = 4096$ and the filter length is $L = 2049$. The sampling frequency is limited by the measurement setup and is set to $F_\mathrm{s} = 125\,\mathrm{MHz}$. Note that the sampling frequency of the spectrum analyser FSVR is set to 25 MHz and the obtained signal is upsampled five times for the processing. DFT precoding is employed to reduce the signal PAPR. The signal generator with these settings provides the test signal vector with the channel bandwidth $B \approx 1.9\,\mathrm{MHz}$ and $N_0 = 1\,602\,560$ samples.

The DPD for the PA linearisation employs the MP model with $K = 3$ and $Q = 1$, which yields $P = 6$ DPD coefficients. The DLA has driven the DPD adaptation, the iteration step was set to $\mu = 0.1$, and 30 iterations were performed during the measurements. The reference voltage $r(t)$ was set to -0.1 and to 0.05 of the full scale of the in-phase feedback amplitudes and the two values were changed in the middle of the measurement iteration.

Fig. 5.28, Fig. 5.29, and Fig. 5.30 show the evolution of ACPR, EVM, NMSE, respectively, in the iterations throughout the measurement. The horizontal black dashed lines depict the individual metrics without the DPD for the same main-channel power. Fig. 5.31 shows the comparison of the resulting spectra for the conventional DPD and the proposed feedback comparator with the spectrum of the PA output without the linearisation.

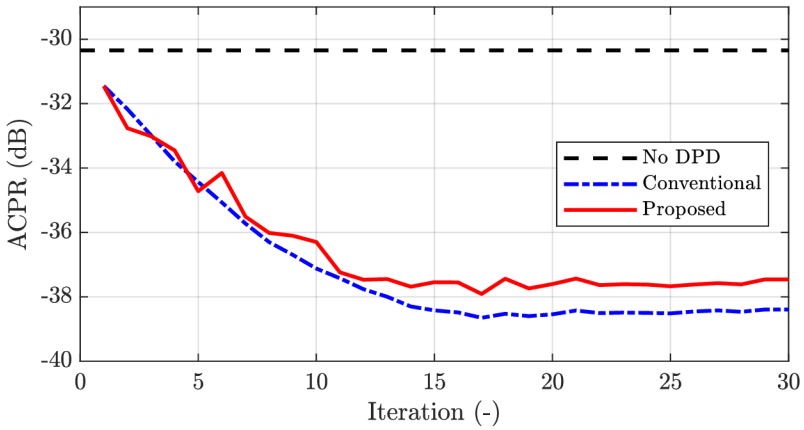


Fig. 5.28: Evolution of the ACPR during the DPD adaptation with the setup with a hardware LC-ADC.

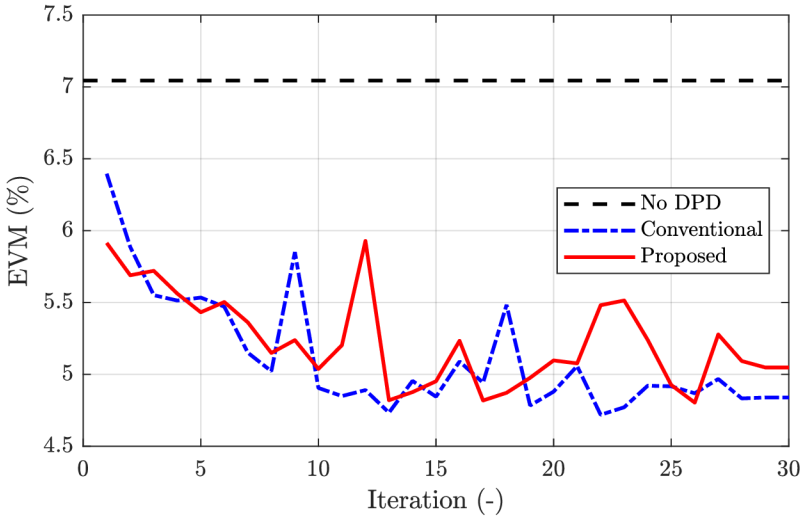


Fig. 5.29: Evolution of the EVM during the DPD adaptation with the setup with a hardware LC-ADC.

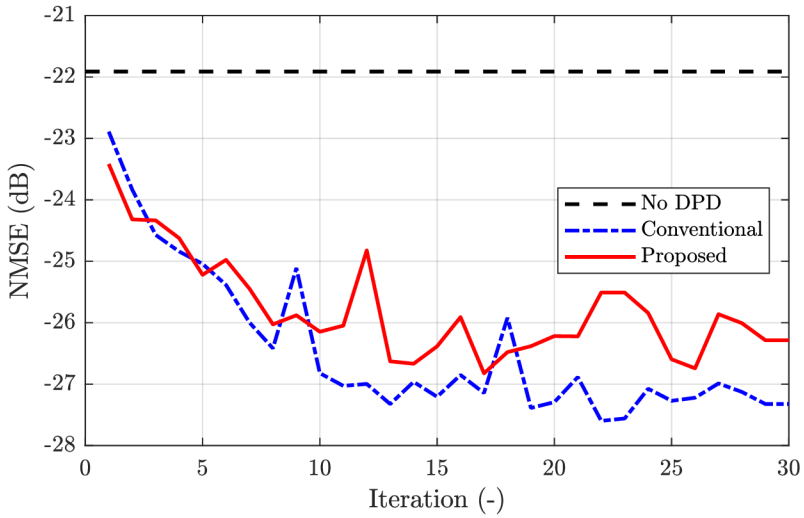


Fig. 5.30: Evolution of the NMSE during the DPD adaptation with the setup with a hardware LC-ADC.

Tab. 5.4 summarises the DPD linearisation performance. The presented results are averaged over the last ten iterations. We can observe that the proposed method with the feedback comparator did not achieve the linearisation performance of the conventional method. This could be caused by the limited number of points acquired by the comparator and used for the DPD adaptation. The comparator acquired approx. 100 samples while the conventional method employed all 1 602 560 samples. This limitation is, however, the limitation of our setup rather than the limitation of the proposed method.

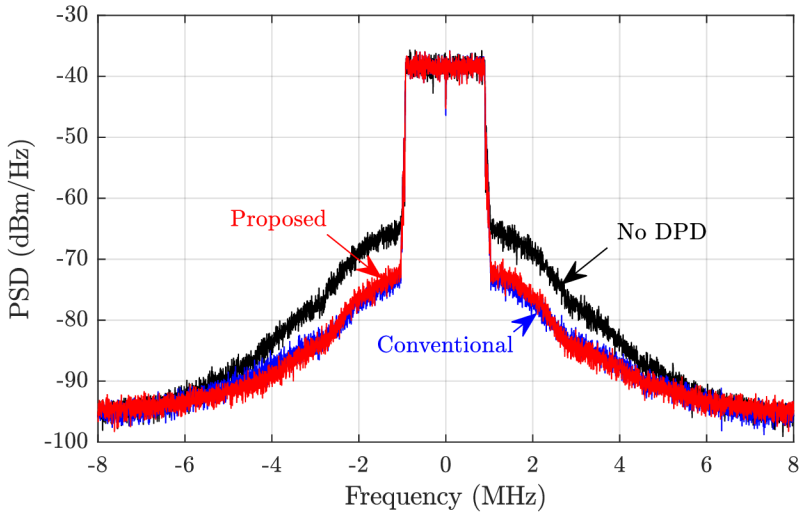


Fig. 5.31: Comparison of the measured power spectral density of the PA output without DPD and with the DPD adapted by the conventional and proposed methods on the setup with a hardware LC-ADC.

Tab. 5.4: Measurement results of NMSE, EVM, and ACPR for the DPD with the proposed LC-ADC and conventional feedback on the setup with the hardware LC-ADC.

	P_{main} (dBm)	NMSE (dB)	EVM (%)	ACPR (dB)
Without DPD	24.4	-21.9	7.0	-30.4
Conventional	24.3	-27.2	4.9	-38.5
LC-ADC	24.2	-26.0	5.2	-37.6

EVM of the generated signal is 3.1% due to the nonorthogonality caused by the inherent F-OFDM filtering.

5.10 Conclusion

In this chapter, we have proposed a novel method for DPD adaptation with the LC-ADC replacing the conventional ADCs. We have demonstrated the principle of the DPD adaptation with level-crossing detection by a comparator in the feedback. We have shown that the proposed architecture reduces power consumption by approximately 75% compared with the conventional DPD architectures and by approximately 50% compared with the DPD from [4]. The linearisation performance has been evaluated by three measurements, where the proposed DPD achieved com-

parable linearisation performance to the conventional DPD architectures. The results achieved with the narrow-band measurement setup indicate an improvement of 22 dB in ACPR and 2.1% in EVM by both the conventional and proposed architectures. We have achieved linearisation of a signal with 500 MHz bandwidth with the wide-band setup with the improvement of 10 dB in ACPR and 3.3% in EVM for the conventional architecture and 8 dB in ACPR and 3.1 % in EVM for the proposed architecture with the LC-ADC. For the last measurement, we designed our feedback module with a real comparator and incorporated it into the measurement setup for the predistortion. Despite all the technical limitations of this setup, the adapted predistorter achieved an improvement of 8 dB in ACPR and 2.1% in EVM for the conventional architecture and 7 dB in ACPR and 1.8% in EVM for the proposed architecture with the designed feedback module.

6 Conclusion

Digital predistortion is considered to be the most advanced and best performing linearisation technique. However, at the same time, it is one of the most complex and computationally demanding techniques. The required extra computational resources and feedback circuits for DPD adaptability represent additional expenses for implementing DPDs and the main limiting factors for their wide spread use across various applications. Although a lot of research has been conducted in this field, there is still plenty of room for improvements in this area. Therefore, we have oriented this thesis toward low-complexity methods for DPD adaptation. The main contributions of this thesis are three key methods to lower the complexity of DPDs:

- real-valued feedback,
- feedback sample selection,
- feedback with an LC-ADC.

The adaptation with real-valued feedback samples enables saving one of two conventionally-employed feedback ADCs. We have shown that the proposed method achieves the same linearisation performance as the conventional approaches and reduces the power consumption and additional expenses on the feedback circuit. Moreover, if real-valued feedback is employed, DPD adaptation might be less susceptible to IQ imbalances of the feedback mixer [A9]. The real-valued feedback principle is generally applicable without limitations. Even already realised transmitters with conventional feedback circuits could turn off one feedback ADC, reduce the feedback-circuit power consumption and benefit from improved immunity to feedback IQ imbalances.

The feedback sample selection has been intrinsically based on the undersampling feedback methods. We have shown that only a few feedback samples are required for successful DPD adaptation if the samples are carefully selected. The limited number of required collected samples decreases the size of vectors and matrices entering the calculations and, hence, reduces the computational complexity of DPD adaptation. We have proposed several methods for feedback sample selection, two of them have been driven by respecting a predefined histogram. The proposed histogram-based methods respect both nonlinear PA characteristics and statistical properties of the transmitted signal. In our simulations and measurements, they have achieved the highest reduction of the required number of feedback samples and, consequently, the highest computational complexity reduction. Even the undersampling feedback alone can lead to a significant reduction ($\approx$ 40 times) of the feedback ADC power consumption. The sample selection additionally reduces the computational complexity of DPD adaptation. The performed simulations indicate a 400-time reduction in computation complexity in the number of required multiplications and additions.

The feedback sample selection and undersampling principles are generally applicable and, even more, both principles can be potentially combined with the real-valued feedback principle to reduce even more power consumption and hardware complexity. The feedback sample selection can be applied without any limitations in already-realised transmitters, because its implementation requires only firmware modification. In certain cases, the undersampling could be potentially implemented in already-realised transmitters to reduce the feedback ADC power consumption. To achieve this, the hardware would need to allow for a changing ADC sampling clock frequency. Of course, maximum benefits can be exploited if a new transmitter design is adjusted for all these methods.

The proposed feedback with an LC-ADC replaces a conventional feedback ADC with a simple comparator complemented by a low-speed DAC. Its real implementation in the hardware requires a different approach for time and amplitude calibrations. For this purpose, we have designed and tested a method based on duty cycle measurements for amplitude calibration which does not require time synchronisation. We have synchronised the time in two steps: first coarsely with sample resolution and later finely with subsample resolution. In both steps, we transmitted an arbitrary signal with edges at different time positions and calculated the signal delay by fitting the transmitted signal with the observed comparator output. In all the conducted measurements, the DPD with LC-ADC adaptation achieved performance similar to that of the conventional DPD. A system comparison example has shown that the proposed LC-ADC feedback can significantly reduce the feedback power consumption (≈ 36 times) or can achieve higher linearisation bandwidth with unchanged power consumption.

The usability of the LC-ADC for DPD adaptation is practically limited to special designs. We expect the LC-ADC could replace conventional ADCs in highly integrated feedback circuits. These integrated designs could benefit from the smaller footprint and lower power consumption of a simple comparator than that of complex high-speed ADCs. Additionally, the presented limitations and required compensations and calibrations could be more easily achievable as the circuit parameters can be better controlled on a chip than in a discrete realisation. The LC-ADC concept, therefore, currently seems to be impractically applicable to designs with discrete feedback circuits and is not suitable for already-realised transmitters, in contrast to the two previously presented methods. Even though the LC-ADC concept is not generally applicable, we demonstrated its funcionality by the first proof-of-concept implementation and we consider it the most advanced and interesting technique presented in this thesis. We believe its main ideas will be employed or improved in the future.

6.1 Potential Future Challenges

Recent review papers [1,2] emphasised the importance of digital predistortion for the 5G or even 6G high-speed wireless communications and identified some potential future challenges. These primarily included improving DPD linearisation performance in general, lowering power consumption demands for DPD linearisation by utilising hybrids of analogue and digital predistorters, effective predistortion for phased array antennas and massive MIMO systems, and increasing linearised bandwidth. In addition, due to our experience, the power consumption of the auxiliary circuits required for the DPD adaptation might be reduced more. One possible approach, which we would like to analyse in the future, is the possibility of avoiding the feedback down-converting mixer which is usually a very power demanding component, especially in wideband applications. Another potential of reducing digital predistortion complexity could be completely avoiding classic feedback and adapting to the PA nonlinearity changes based on a simpler input, e.g. PA temperature. Even though this approach seems to be straightforward, due to our best knowledge, the current state-of-the-art research lacks a comprehensive study of temperature and ageing effects on PA linearity changes and, therefore, it is hard to predict the abilities and performance of simpler DPDs without full adaptability.

Although there are still unanswered questions and many potential challenges to be addressed in the future, we believe that this dissertation thesis satisfies its goal and extends the current state-of-the-art knowledge in the field of digital predistortion aiming at low-complexity methods. We hope this thesis will supplement the published papers and provide a different view on the presented topics and will be one of the starting points for young researchers working on low-complexity digital predistortion.

References

[1] Christian Fager, Thomas Eriksson, Filipe Barradas, Katharina Hausmair, Telmo Cunha, and Jose Carlos Pedro. Linearity and Efficiency in 5G Transmitters: New Techniques for Analyzing Efficiency, Linearity, and Linearization in a 5G Active Antenna Transmitter Context. *IEEE Microwave Magazine*, 20(5):35–49, May 2019. doi:10.1109/MMM.2019.2898020.

[2] Andzej Borel, Vaidotas Barzdenas, and Aleksandr Vasjanov. Linearization as a Solution for Power Amplifier Imperfections: A Review of Methods. *Electronics*, 10(9):1073, May 2021. doi:10.3390/electronics10091073.

[3] Jessica Chani-Cahuana, Mustafa Özen, Christian Fager, and Thomas Eriksson. Digital Predistortion Parameter Identification for RF Power Amplifiers Using Real-Valued Output Data. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 64(10):1227–1231, Oct. 2017. doi:10.1109/TCSII.2017.2686004.

[4] Haoyu Wang, Gang Li, Chongbin Zhou, Wei Tao, Falin Liu, and Anding Zhu. 1-bit Observation for Direct-Learning-Based Digital Predistortion of RF Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 65(7):2465–2475, Jul. 2017. doi:10.1109/TMTT.2016.2642945.

[5] N. Guan, N. Wu, and H. Wang. Digital Predistortion of Wideband Power Amplifier With Single Undersampling ADC. *IEEE Microwave and Wireless Components Letters*, 27(11):1016–1018, Nov. 2017. doi:10.1109/LMWC.2017.2750059.

[6] F. M. Ghannouchi and O. Hammi. Behavioral Modeling and Predistortion. *IEEE Microwave Magazine*, 10(7):52–64, Dec. 2009. doi:10.1109/MMM.2009.934516.

[7] J. Zhao, Y. Liu, C. Yu, J. Yu, and S. Li. A Modified Band-Limited Digital Predistortion Technique for Broadband Power Amplifiers. *IEEE Communications Letters*, 20(9):1800–1803, Sep. 2016. doi:10.1109/LCOMM.2016.2585652.

[8] M. K. Kazimierczuk. *RF Power Amplifiers*. John Wiley & Sons, 2015.

[9] Tomas Gotthans. *Advanced methods for analyzing nonlinear dynamical systems*. PhD thesis, Université Paris-Est, 2014.

[10] H. Bendel. Amplifier Linearized by Negative Feedback, Jun. 1940. US Patent 2,203,484. URL: https://www.google.com/patents/US2203484.

[11] D. Leypold, H. Leysieffer, and H. Holzwarth. Transmitter Linearized by Negative Feedback, Nov. 1959. US Patent 2,912,570. URL: https://www.google.com/patents/US2912570.

[12] Y. Nagata. Linear Amplification Technique for Digital Mobile Communications. In *IEEE Vehicular Technology Conference*, pages 159–164 vol.1, May 1989. doi:10.1109/VETEC.1989.40066.

[13] J. K. Cavers. Amplifier Linearization Using a Digital Predistorter with Fast Adaptation and Low Memory Requirements. *IEEE Transactions on Vehicular Technology*, 39(4):374–382, Nov. 1990. doi:10.1109/25.61359.

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

[15] L. Guan and A. Zhu. Simplified Dynamic Deviation Reduction-Based Volterra Model for Doherty Power Amplifiers. In *Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits*, pages 1–4, Apr. 2011. doi:10.1109/INMMIC.2011.5773325.

[16] Vito Volterra. *Sopra le funzioni che dipendono da altre funzioni.* Tipografia della R. Accademia dei Lincei, 1887.

[17] Vito Volterra and Edmund Taylor Whittaker. *Theory of functionals and of integral and integro-differential equations.* Dover publications, 1959.

[18] Brook Taylor. *Methodus incrementorum directa et inversa.* Innys, 1717.

[19] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan. A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers. *IEEE Transactions on Signal Processing*, 54(10):3852–3860, Oct. 2006. doi:10.1109/TSP.2006.879264.

[20] H. Paaso and A. Mammela. Comparison of Direct Learning and Indirect Learning Predistortion Architectures. In *IEEE International Symposium on Wireless Communication Systems*, pages 309–313, Oct. 2008. doi:10.1109/ISWCS.2008.4726067.

[21] Z. Yu, H. Xie, and E. Zhu. Wide-band Linear Characteristic Compensation for DPD Systems with Direct Learning. In *Asia-Pacific Microwave Conference Proceedings (APMC)*, pages 790–793, Nov. 2013. doi:10.1109/APMC.2013.6694933.

[22] Roman Marsalek. *Contributions to the power amplifier linearization using digital baseband adaptive predistortion.* PhD thesis, Université de Marne la Vallée, 2003.

[23] R. Marsalek, P. Jardin, and G. Baudoin. From Post-Distortion to Pre-Distortion for Power Amplifiers Linearization. *IEEE Communications Letters*, 7(7):308–310, Jul. 2003. doi:10.1109/LCOMM.2003.814714.

[24] Lloyd N Trefethen and David Bau III. *Numerical Linear Algebra*, volume 50. Siam, 1997.

[25] D. R. Morgan, Zhengxiang Ma, and Lei Ding. Reducing Measurement Noise Effects in Digital Predistortion of RF Power Amplifiers. In *IEEE International Conference on Communications (ICC)*, volume 4, pages 2436–2439, May 2003. doi:10.1109/ICC.2003.1204371.

[26] P. N. Landin, A. E. Mayer, and T. Eriksson. MILA - a noise mitigation technique for RF power amplifier linearization. In *IEEE International Multi-Conference on Systems, Signals Devices (SSD14)*, pages 1–4, Feb. 2014. doi:10.1109/SSD.2014.6808874.

[27] Robert W. Chang. Synthesis of Band-Limited Orthogonal Signals for Multi-channel Data Transmission. *The Bell System Technical Journal*, 45(10):1775–1796, Dec. 1966. doi:10.1002/j.1538-7305.1966.tb02435.x.

[28] B. Saltzberg. Performance of an Efficient Parallel Data Transmission System. *IEEE Transactions on Communication Technology*, 15(6):805–811, Dec. 1967. doi:10.1109/TCOM.1967.1089674.

[29] S. Weinstein and P. Ebert. Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform. *IEEE Transactions on Communication Technology*, 19(5):628–634, Oct. 1971. doi:10.1109/TCOM.1971.1090705.

[30] Ramjee Prasad. *OFDM for Wireless Communications Systems.* Artech House, 2004.

[31] J. Abdoli, M. Jia, and J. Ma. Filtered OFDM: A New Waveform for Future Wireless Systems. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 66–70, Jun. 2015. doi:10.1109/SPAWC.2015.7227001.

[32] Huawei, HiSilicon. f-OFDM scheme and filter design. R1-165425, 3GPP TSG RAN WG1 Meeting #85, Nanjing, China, May 2016.

[33] Hyung G. Myung, Junsung Lim, and David J. Goodman. Single Carrier FDMA for Uplink Wireless Transmission. *IEEE Vehicular Technology Magazine*, 1(3):30–38, 2006. doi:10.1109/MVT.2006.307304.

[34] Mohamed A Aboul-Dahab, AAA Esam, and Ahmad A Elhaseeb. PAPR Reduction Based on DFT Precoding for OFDM Signals. *International Journal of Future Computer and Communication*, 2(4):325, 2013.

[35] David D. Falconer. Linear Precoding of OFDMA Signals to Minimize Their Instantaneous Power Variance. *IEEE Transactions on Communications*, 59(4):1154–1162, 2011. doi:10.1109/TCOMM.2011.11.100042.

[36] Yeo Hun Yun, Chanhong Kim, Kyeongyeon Kim, Zuleita Ho, Byunghwan Lee, and Ji-Yun Seol. A New Waveform Enabling Enhanced QAM-FBMC Systems. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 116–120, Jun. 2015. doi:10.1109/SPAWC.2015.7227011.

[37] Hyungju Nam, Moonchang Choi, Seongbae Han, Chanhong Kim, Sooyong Choi, and Daesik Hong. A New Filter-Bank Multicarrier System With Two Prototype Filters for QAM Symbols Transmission and Reception. *IEEE Transactions on Wireless Communications*, 15(9):5998–6009, Sep. 2016. doi:10.1109/TWC.2016.2575839.

[38] Chanhong Kim, Yeo Hun Yun, Kyeongyeon Kim, and Ji-Yun Seol. Introduction to QAM-FBMC: From Waveform Optimization to System Design. *IEEE Communications Magazine*, 54(11):66–73, Nov. 2016. doi:10.1109/MCOM.2016.1600384CM.

[39] Maurice Bellanger, D Le Ruyet, D Roviras, M Terré, J Nossek, L Baltar, Q Bai, D Waldhauser, M Renfors, T Ihalainen, et al. FBMC physical layer: a primer. *PHYDYAS*, 25(4):7–10, Jan. 2010.

[40] M.G. Bellanger. Specification and Design of a Prototype Filter for Filter Bank Based Multicarrier Transmission. In *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings*, volume 4, pages 2417–2420 vol.4, May 2001. doi:10.1109/ICASSP.2001.940488.

[41] Ali Jasim Ramadhan. Implementation of 5G FBMC PHYDYAS Prototype Filter. *International Journal of Applied Engineering Research*, 12(23):13476–13481, 2017.

[42] Ronald Nissel. *Filter Bank Multicarrier Modulation for Future Wireless Systems*. PhD thesis, Faculty of Electrical Engineering and Information Technology, TU Wien, Dec. 2017.

[43] Michel. FBMC Modulation / Demodulation [online]. Matlab Central File Exchange, Jan. 2014. Cited 19 Nov. 2021. URL: https://www.mathworks.com/matlabcentral/fileexchange/45100-fbmc-modulation-demodulation.

[44] David Ballo and Joe Gorin. Adjacent channel power measurements in the digital wireless era. *Microwave Journal*, 37(7):74–83, Jul. 1994.

[45] J.F. Sevic and J. Staudinger. Simulation of Power Amplifier Adjacent-Channel Power Ratio for Digital Wireless Communication Systems. In *IEEE Vehicular Technology Conference. Technology in Motion*, volume 2, pages 681–685 vol.2, May 1997. doi:10.1109/VETEC.1997.600415.

[46] R. Hassun, M. Flaherty, R. Matreci, and M. Taylor. Effective Evaluation of Link Quality Using Error Vector Magnitude Techniques. In *Proceedings of Wireless Communications Conference*, pages 89–94, Aug. 1997. doi:10.1109/WCC.1997.622254.

[47] Keysight Technologies. EVM (Digital Demod) [onl.]. Cited 19 Nov. 2021. URL: https://rfmw.em.keysight.com/wireless/helpfiles/89600b/webhelp/subsystems/digdemod/Content/digdemod_symtblerrdata_evm.htm.

[48] Peter Händel. Understanding Normalized Mean Squared Error in Power Amplifier Linearization. *IEEE Microwave and Wireless Components Letters*, 28(11):1047–1049, Nov. 2018. doi:10.1109/LMWC.2018.2869299.

[49] Jessica Chani-Cahuana, Christian Fager, and Thomas Eriksson. Lower Bound for the Normalized Mean Square Error in Power Amplifier Linearization. *IEEE Microwave and Wireless Components Letters*, 28(5):425–427, May 2018. doi:10.1109/LMWC.2018.2817021.

[50] Allen Katz, John Wood, and Daniel Chokola. The Evolution of PA Linearization: From Classic Feedforward and Feedback Through Analog and Digital Predistortion. *IEEE Microwave Magazine*, 17(2):32–40, Feb. 2016. doi:10.1109/MMM.2015.2498079.

[51] Pere L. Gilabert, R. Neil Braithwaite, and Gabriel Montoro. Beyond the Moore-Penrose Inverse: Strategies for the Estimation of Digital Predistortion Linearization Parameters. *IEEE Microwave Magazine*, 21(12):34–46, Dec. 2020. doi:10.1109/MMM.2020.3023220.

[52] Sukwinder Singh and Jagannath Malik. Review of Efficiency Enhancement Techniques and Linearization Techniques for Power Amplifier. *International Journal of Circuit Theory and Applications*, 49(3):762–777, Mar. 2021. doi:10.1002/cta.2956.

[53] Patricia Desgreys, Venkata Narasimha Manyam, Kelly Tchambake, Dang-Kièn Germain Pham, and Chadi Jabbour. Wideband Power Amplifier Predistortion: Trends, Challenges and Solutions. In *IEEE International Conference on ASIC (ASICON)*, pages 100–103, Oct. 2017. doi:10.1109/ASICON.2017.8252421.

[54] John Wood. Digital Pre-Distortion of RF Power Amplifiers. In *IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, pages 1–3, Jan. 2017. doi:10.1109/PAWR.2017.7875557.

[55] Ijaz Ahmad, Shariar Shahabuddin, Hassan Malik, Erkki Harjula, Teemu Leppänen, Lauri Lovén, Antti Anttonen, Ali Hassan Sodhro, Muhammad Mahtab Alam, Markku Juntti, Antti Ylä-Jääski, Thilo Sauter, Andrei Gurtov, Mika Ylianttila, and Jukka Riekki. Machine Learning Meets Communication Networks: Current Trends and Future Challenges. *IEEE Access*, 8:223418–223460, 2020. doi:10.1109/ACCESS.2020.3041765.

[56] Xin Liu, Guan-sheng Lv, De-han Wang, Wen-hua Chen, and Fadhel M. Ghannouchi. Energy-efficient power amplifiers and linearization techniques for massive MIMO transmitters: a review. *Frontiers of Information Technology & Electronic Engineering*, 21(1):72–96, Jan. 2020. doi:10.1631/FITEE.1900467.

[57] Saurav Gupta, Ajit Kumar Sahoo, and Upendra Kumar Sahoo. Volterra and Wiener Model Based Temporally and Spatio-Temporally Coupled Nonlinear System Identification: A Synthesized Review. *IETE Technical Review*, 38(3):303–327, May 2021. doi:10.1080/02564602.2020.1732233.

[58] Jessica Chani-Cahuana, Per Niklas Landin, Christian Fager, and Thomas Eriksson. Iterative Learning Control for RF Power Amplifier Linearization. *IEEE Transactions on Microwave Theory and Techniques*, 64(9):2778–2789, Sep. 2016. doi:10.1109/TMTT.2016.2588483.

[59] Maarten Schoukens, Jules Hammenecker, and Adam Cooman. Obtaining the Preinverse of a Power Amplifier Using Iterative Learning Control. *IEEE Transactions on Microwave Theory and Techniques*, 65(11):4266–4273, Nov. 2017. doi:10.1109/TMTT.2017.2694822.

[60] David Lopez-Bueno, Quynh Anh Pham, Gabriel Montoro, and Pere L. Gilabert. Independent Digital Predistortion Parameters Estimation Using Adaptive Principal Component Analysis. *IEEE Transactions on Microwave Theory and Techniques*, 66(12):5771–5779, Dec. 2018. Patent Number: 2 Place: Piscataway Publisher: IEEE-Inst Electrical Electronics Engineers Inc. WOS:000453528100039. doi:10.1109/TMTT.2018.2870420.

[61] Quynh Anh Pham, David Lopez-Bueno, Gabriel Montoro, and Pere L. Gilabert. Adaptive Principal Component Analysis for Online Reduced Order Parameter Extraction in PA Behavioral Modeling and DPD Linearization. In *IEEE/MTT-S International Microwave Symposium - IMS*, pages 160–163, New York, 2018. IEEE. ISSN: 0149-645X WOS:000451173600043.

[62] Ahmad Rahati Belabad, Saeed Sharifian, and Seyed Ahmad Motamedi. An Accurate Digital Baseband Predistorter Design for Linearization of RF Power Amplifiers by a Genetic Algorithm Based Hammerstein Structure. *Analog Integrated Circuits and Signal Processing*, 95(2):231–247, May 2018. doi:10.1007/s10470-018-1173-x.

[63] Henry Douglas Rodrigues, Tales Cleber Pimenta, Rausley Adriano Amaral de Souza, and Luciano Leonel Mendes. Orthogonal Scalar Feedback Digital Pre-Distortion Linearization. *IEEE Transactions on Broadcasting*, 64(2):319–330, Jun. 2018. doi:10.1109/TBC.2017.2755261.

[64] Han Le Duc, Bruno Feuvrie, Matthieu Pastore, and Yide Wang. An Adaptive Cascaded ILA- and DLA-Based Digital Predistorter for Linearizing an RF Power Amplifier. *IEEE Transactions on Circuits and Systems I-Regular Papers*, 66(3):1031–1041, Mar. 2019. doi:10.1109/TCSI.2018.2872465.

[65] Duc Han Le, Van-Phuc Hoang, Minh Hong Nguyen, Hien M. Nguyen, and Duc Minh Nguyen. Linearization of RF Power Amplifiers in Wideband Communication Systems by Adaptive Indirect Learning Using RPEM Algorithm. *Mobile Networks and Applications*, 25(5):1988–1997, Oct. 2020. doi:10.1007/s11036-020-01545-z.

[66] Qi Cai, Wenquan Che, and Kaixue Ma. A Linear GaN Power Amplifier Using Novel Transistor Based Analog Predistortion Method. In *IEEE MTT-S International Microwave Workshop Series on Advanced Materials and Processes for RF and THz Applications (IMWS-AMP)*, pages 1–4, Jul. 2016. doi:10.1109/IMWS-AMP.2016.7588314.

[67] Qi Cai, Wenquan Che, Kaixue Ma, and Mi Zhang. A Simplified Transistor-Based Analog Predistorter for a GaN Power Amplifier. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(3):326–330, Mar. 2018. doi:10.1109/TCSII.2017.2735022.

[68] R. Neil Braithwaite. Using a Cascade of Digital and Analog Predistortion to Linearize a Dual-Band RF Transmitter. In *IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, pages 77–80, Jan. 2017. doi:10.1109/PAWR.2017.7875578.

[69] Pedro M. Tomé, Filipe M. Barradas, Telmo R. Cunha, and José Carlos Pedro. Hybrid Analog/Digital Linearization of GaN HEMT-Based Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 67(1):288–294, Jan. 2019. doi:10.1109/TMTT.2018.2880911.

[70] Hailin Deng, Dalong Lv, Yi Zhang, Dewei Zhang, and Dongfang Zhou. Compact Analog Predistorter With Shape Tuning Capability Using Power-Dependent Impedance Matching Network. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 67(10):1705–1709, Oct. 2020. doi:10.1109/TCSII.2019.2945062.

[71] Karan Gumber and Meenakshi Rawat. Analogue Predistortion Lineariser Control Schemes for Ultra-Broadband Signal Transmission in 5G Transmitters. *Iet Microwaves Antennas & Propagation*, 14(8):718–727, Jul. 2020. doi:10.1049/iet-map.2019.0253.

[72] Ali Cheaito, Jean-Francois Helard, Matthieu Crussiere, and Yves Louet. EVM Derivation of Multicarrier Signals to Determine the Operating Point of the Power Amplifier Considering Clipping and Predistortion. *Eurasip Journal on Wireless Communications and Networking*, page 281, Dec. 2016. doi:10.1186/s13638-016-0771-5.

[73] Ali Cheaito, Matthieu Crussiere, Jean-Francois Helard, and Yves Louet. Quantifying the Memory Effects of Power Amplifiers: EVM Closed-Form Derivations of Multicarrier Signals. *IEEE Wireless Communications Letters*, 6(1):34–37, Feb. 2017. doi:10.1109/LWC.2016.2627037.

[74] Ying Liu, Xin Quan, Shihai Shao, and Youxi Tang. Digital Predistortion Architecture with Reduced ADC Dynamic Range. *Electronics Letters*, 52(6):435–436, Mar. 2016. doi:10.1049/el.2015.4174.

[75] Ying Liu, Chuang Huang, Xin Quan, Patrick Roblin, Wensheng Pan, and Youxi Tang. Novel Linearization Architecture with Limited ADC Dynamic

Range for Green Power Amplifiers. *IEEE Journal on Selected Areas in Communications*, 34(12):3902–3914, Dec. 2016. doi:10.1109/JSAC.2016.2600415.

[76] Ying Liu, Xin Quan, Wensheng Pan, Shihai Shao, and Youxi Tang. Performance Analysis of Direct-Learning Digital Predistortion With Loop Delay Mismatch in Wideband Transmitters. *IEEE Transactions on Vehicular Technology*, 65(9):7078–7089, Sep. 2016. doi:10.1109/TVT.2015.2496188.

[77] Siqi Wang, Mazen Abi Hussein, Olivier Venard, and Geneviève Baudoin. Impact of the Normalization Gain of Digital Predistortion on Linearization Performance and Power Added Efficiency of the Linearized Power Amplifier. In *European Microwave Conference (EuMC)*, pages 1050–1053, Oct. 2017. doi:10.23919/EuMC.2017.8231026.

[78] P. Jardin and G. Baudoin. Filter Lookup Table Method for Power Amplifier Linearization. *IEEE Transactions on Vehicular Technology*, 56(3):1076–1087, May 2007. doi:10.1109/TVT.2007.895566.

[79] Yue-Yu Xiao. Stability Analysis of Closed-loop Algorithms for Cancelling Nonlinear Memory Effects. *IEEE Transactions on Signal Processing*, 65(6):1477–1485, Mar. 2017. doi:10.1109/TSP.2016.2631462.

[80] Siqi Wang, Morgan Roger, and Caroline Lelandais-Perrault. Impacts of Crest Factor Reduction and Digital Predistortion on Linearity and Power Efficiency of Power Amplifiers. *IEEE Transactions on Circuits and Systems Ii-Express Briefs*, 66(3):407–411, Mar. 2019. doi:10.1109/TCSII.2018.2855084.

[81] Ling Liu, Wenhua Chen, Liyong Ma, and He Sun. Single-PA-Feedback Digital Predistortion for Beamforming MIMO Transmitter. In *IEEE International Conference on Microwave and Millimeter Wave Technology (ICMMT)*, volume 2, pages 573–575, Jun. 2016. doi:10.1109/ICMMT.2016.7762371.

[82] Nuutti Tervo, Janne Aikio, Tommi Tuovinen, Timo Rahkonen, and Aarno Parssinen. Digital Predistortion of Amplitude Varying Phased Array Utilising Over-the-Air Combining. In *IEEE MTT-S International Microwave Symposium (IMS)*, pages 1165–1168, Jun. 2017. doi:10.1109/MWSYM.2017.8058809.

[83] Nuutti Tervo, Bilal Khan, Olli Kursu, Janne P. Aikio, Markku Jokinen, Marko E. Leinonen, Markku Juntti, Timo Rahkonen, and Aarno Pärssinen. Digital Predistortion of Phased-Array Transmitter With Shared Feedback and

Far-Field Calibration. *IEEE Transactions on Microwave Theory and Techniques*, 69(1):1000–1015, Jan. 2021. doi:10.1109/TMTT.2020.3038193.

[84] Eric Ng, Yehia Beltagy, Patrick Mitran, and Slim Boumaiza. Single-Input Single-Output Digital Predistortion of Power Amplifier Arrays in Millimeter Wave RF Beamforming Transmitters. In *IEEE/MTT-S International Microwave Symposium - IMS*, pages 481–484, Jun. 2018. doi:10.1109/MWSYM.2018.8439680.

[85] Eric Ng, Yehia Beltagy, Giovanni Scarlato, Ahmed Ben Ayed, Patrick Mitran, and Slim Boumaiza. Digital Predistortion of Millimeter-Wave RF Beamforming Arrays Using Low Number of Steering Angle-Dependent Coefficient Sets. *IEEE Transactions on Microwave Theory and Techniques*, 67(11):4479–4492, Nov. 2019. doi:10.1109/TMTT.2019.2924893.

[86] Abubaker Abdelhafiz, Laleh Behjat, Fadhel M. Ghannouchi, Mohamed Helaoui, and Oualid Hammi. A High-Performance Complexity Reduced Behavioral Model and Digital Predistorter for MIMO Systems With Crosstalk. *IEEE Transactions on Communications*, 64(5):1996–2004, May 2016. doi:10.1109/TCOMM.2016.2545654.

[87] Seyed Aidin Bassam, Mohamed Helaoui, and Fadhel M. Ghannouchi. Crossover Digital Predistorter for the Compensation of Crosstalk and Nonlinearity in MIMO Transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 57(5):1119–1128, May 2009. doi:10.1109/TMTT.2009.2017258.

[88] Mahmoud Abdelaziz, Lauri Anttila, and Mikko E. Valkama. Reduced-Complexity Digital Predistortion for Massive MIMO. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6478–6482, Mar. 2017. doi:10.1109/ICASSP.2017.7953404.

[89] Mahmoud Abdelaziz, Lauri Anttila, Alberto Brihuega, Fredrik Tufvesson, and Mikko Valkama. Digital Predistortion for Hybrid MIMO Transmitters. *IEEE Journal of Selected Topics in Signal Processing*, 12(3):445–454, Jun. 2018. doi:10.1109/JSTSP.2018.2824981.

[90] Xiaoyu Wang, Yue Li, Chao Yu, Wei Hong, and Anding Zhu. Digital Predistortion of 5G Massive MIMO Wireless Transmitters Based on Indirect Identification of Power Amplifier Behavior With OTA Tests. *IEEE Transactions on Microwave Theory and Techniques*, 68(1):315–327, Jan. 2020. doi:10.1109/TMTT.2019.2944828.

[91] Joep Zanen, Eric Klumperink, and Bram Nauta. Power Efficiency Model for MIMO Transmitters Including Memory Polynomial Digital Predistortion. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 68(4):1183–1187, Apr. 2021. doi:10.1109/TCSII.2020.3028493.

[92] O. Hammi, S. Bousnina, and F.M. Ghannouchi. A Linearized Doherty Amplifier Using Complex Baseband Digital Predistortion Driven by CDMA Signals. In *Proceedings. IEEE Radio and Wireless Conference (IEEE Cat. No.04TH8746)*, pages 435–438, Sep. 2004. doi:10.1109/RAWCON.2004.1389170.

[93] Ramzi Darraji, Fadhel M. Ghannouchi, and Oualid Hammi. A Dual-Input Digitally Driven Doherty Amplifier Architecture for Performance Enhancement of Doherty Transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 59(5):1284–1293, May 2011. doi:10.1109/TMTT.2011.2106137.

[94] Sujata Ghosh and Karun Rawat. Hybrid Analog/Digital Continuous Class B/J Mode for Broadband Doherty Power Amplifiers. *IEEE Access*, 7:74986–74995, 2019. doi:10.1109/ACCESS.2019.2920487.

[95] Chun-Wei Chang, Shu-Chen Lin, Jau-Horng Chen, and Joseph Staudinger. A Multilevel Pulse-Modulated Polar Transmitter Based on a Doherty Power Amplifier and Memoryless Digital Predistortion. *IEEE Microwave and Wireless Components Letters*, 28(10):933–935, Oct. 2018. doi:10.1109/LMWC.2018.2864875.

[96] Mir Masood, Peter Rashev, and J. Stevenson Kenney. Digitally-assisted Doherty Power Amplifier: Efficiency Enhancement and Linearity Improvement. In *IEEE/MTT-S International Microwave Symposium - IMS*, pages 801–804, Jun. 2018. doi:10.1109/MWSYM.2018.8439376.

[97] Jun Peng, Songbai He, Weimin Shi, Tingting Yao, Jiayan Wu, and Jinchen Wang. Adaptive Signal Separation for Dual-Input Doherty Power Amplifier. *IEEE Transactions on Microwave Theory and Techniques*, 68(1):121–131, Jan. 2020. doi:10.1109/TMTT.2019.2947486.

[98] Chak-Fong Cheang, Pui-In Mak, and Rui P. Martins. A Hardware-Efficient Feedback Polynomial Topology for DPD Linearization of Power Amplifiers: Theory and FPGA Validation. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 65(9):2889–2902, Sep. 2018. doi:10.1109/TCSI.2017.2788082.

[99] Wenhui Cao and Anding Zhu. A Modified Decomposed Vector Rotation-Based Behavioral Model With Efficient Hardware Implementation for Digital Predistortion of RF Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 65(7):2443–2452, Jul. 2017. doi:10.1109/TMTT.2016.2640318.

[100] Hai Huang, Jingjing Xia, and Slim Boumaiza. Novel Parallel-Processing-Based Hardware Implementation of Baseband Digital Predistorters for Linearizing Wideband 5G Transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 68(9):4066–4076, Sep. 2020. doi:10.1109/TMTT.2020.2993236.

[101] Jijun Ren. FPGA Implementation of Adaptive Digital Pre-Distorter with Improving Accuracy of Lookup Table by Taylor Series Method. *Electronics Letters*, 54(15):959–961, Jul. 2018. doi:10.1049/el.2018.1082.

[102] Albert Molina, Kannan Rajamani, and Kamran Azadet. Digital Predistortion Using Lookup Tables With Linear Interpolation and Extrapolation: Direct Least Squares Coefficient Adaptation. *IEEE Transactions on Microwave Theory and Techniques*, 65(3):980–987, Mar. 2017. doi:10.1109/TMTT.2016.2627562.

[103] Albert Molina, Kannan Rajamani, and Kamran Azadet. Concurrent Dual-Band Digital Predistortion Using 2-D Lookup Tables With Bilinear Interpolation and Extrapolation: Direct Least Squares Coefficient Adaptation. *IEEE Transactions on Microwave Theory and Techniques*, 65(4):1381–1393, Apr. 2017. doi:10.1109/TMTT.2016.2634001.

[104] Manasjyoti Bhuyan and Kandarpa Kumar Sarma. Learning Aided Behavioral Modeling and Adaptive Digital Predistortion Design for Nonlinear Power Amplifier. *IEEE Sensors Journal*, 16(16):6167–6174, Aug. 2016. doi:10.1109/JSEN.2016.2579261.

[105] Omar Z. Alngar, Walid S. El-Deeb, El-Sayed, and M. El-Rabaie. On-Line Predistortion Algorithm for Nonlinear Power Amplifiers with Memory Effects Based on Real-Valued Time-Delay Neural Network. In *National Radio Science Conference (NRSC)*, pages 338–344, Mar. 2018. doi:10.1109/NRSC.2018.8354388.

[106] Dongming Wang, Mohsin Aziz, Mohamed Helaoui, and Fadhel M. Ghannouchi. Augmented Real-Valued Time-Delay Neural Network for Compensation of Distortions and Impairments in Wireless Transmitters. *IEEE Transactions on Neural Networks and Learning Systems*, 30(1):242–254, Jan. 2019. doi:10.1109/TNNLS.2018.2838039.

[107] Yibo Wu, Ulf Gustavsson, Alexandre Graell Amat, and Henk Wymeersch. Residual Neural Networks for Digital Predistortion. In *IEEE Global Communications Conference (globecom)*, New York, 2020. IEEE. ISSN: 2334-0983 WOS:000668970501118. doi:10.1109/GLOBECOM42002.2020.9322327.

[108] Praveen Jaraut, Meenakshi Rawat, and Fadhel M. Ghannouchi. Composite Neural Network Digital Predistortion Model for Joint Mitigation of Crosstalk, I/Q Imbalance, Nonlinearity in MIMO Transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 66(11):5011–5020, Nov. 2018. doi:10.1109/TMTT.2018.2869602.

[109] Kenan Li, Ning Guan, and Hua Wang. Iterative Learning Control Assisted Neural Network for Digital Predistortion of MIMO Power Amplifier. In *IEEE Vehicular Technology Conference (VTC Spring)*, pages 1–5, Jun. 2018. doi:10.1109/VTCSpring.2018.8417491.

[110] Jialin Cai, Chao Yu, Lingling Sun, Shichang Chen, and Justin B. King. Dynamic Behavioral Modeling of RF Power Amplifier Based on Time-Delay Support Vector Regression. *IEEE Transactions on Microwave Theory and Techniques*, 67(2):533–543, Feb. 2019. doi:10.1109/TMTT.2018.2884414.

[111] Jin Xu, Weiliang Jiang, Linhua Ma, Mingyu Li, Zhiqiang Yu, and Zhen Geng. Augmented Time-Delay Twin Support Vector Regression-Based Behavioral Modeling for Digital Predistortion of RF Power Amplifier. *IEEE Access*, 7:59832–59843, 2019. doi:10.1109/ACCESS.2019.2915281.

[112] Jinlong Sun, Wenjuan Shi, Zhutian Yang, Jie Yang, and Guan Gui. Behavioral Modeling and Linearization of Wideband RF Power Amplifiers Using BiLSTM Networks for 5G Wireless Systems. *IEEE Transactions on Vehicular Technology*, 68(11):10348–10356, Nov. 2019. doi:10.1109/TVT.2019.2925562.

[113] Yikang Zhang, Yue Li, Falin Liu, and Anding Zhu. Vector Decomposition Based Time-Delay Neural Network Behavioral Model for Digital Predistortion of RF Power Amplifiers. *IEEE Access*, 7:91559–91568, 2019. doi:10.1109/ACCESS.2019.2927875.

[114] Hongmin Li, Yikang Zhang, Gang Li, and Falin Liu. Vector Decomposed Long Short-Term Memory Model for Behavioral Modeling and Digital Predistortion for Wideband RF Power Amplifiers. *IEEE Access*, 8:63780–63789, 2020. doi:10.1109/ACCESS.2020.2984682.

[115] Masaaki Tanio, Naoto Ishii, and Norifumi Kamiya. Efficient Digital Predistortion Using Sparse Neural Network. *IEEE Access*, 8:117841–117852, 2020. doi:10.1109/ACCESS.2020.3005146.

[116] Girish Chandra Tripathi, Meenakshi Rawat, and Karun Rawat. Swish Activation Based Deep Neural Network Predistorter for RF-PA. In *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, pages 1239–1242, Oct. 2019. doi:10.1109/TENCON.2019.8929500.

[117] Prajit Ramachandran, Barret Zoph, and Quoc V. Le. Swish: a Self-Gated Activation Function. *arXiv:1710.05941*, Oct. 2017.

[118] Stefan Elfwing, Eiji Uchibe, and Kenji Doya. Sigmoid-Weighted Linear Units for Neural Network Function Approximation in Reinforcement Learning. *arXiv:1702.03118*, Nov. 2017.

[119] Xin Hu, Zhijun Liu, Weidong Wang, Mohamed Helaoui, and Fadhel M. Ghannouchi. Low-Feedback Sampling Rate Digital Predistortion Using Deep Neural Network for Wideband Wireless Transmitters. *IEEE Transactions on Communications*, 68(4):2621–2633, Apr. 2020. doi:10.1109/TCOMM.2020.2966718.

[120] Jinlong Sun, Juan Wang, Liang Guo, Jie Yang, and Guan Gui. Adaptive Deep Learning Aided Digital Predistorter Considering Dynamic Envelope. *IEEE Transactions on Vehicular Technology*, 69(4):4487–4491, Apr. 2020. doi:10.1109/TVT.2020.2974506.

[121] Abdalla E. Abdelrahman, Oualid Hammi, Andrew K. Kwan, Azzedine Zerguine, and Fadhel M. Ghannouchi. A Novel Weighted Memory Polynomial for Behavioral Modeling and Digital Predistortion of Nonlinear Wireless Transmitters. *IEEE Transactions on Industrial Electronics*, 63(3):1745–1753, Mar. 2016. doi:10.1109/TIE.2015.2494040.

[122] Yue Li, Wenhui Cao, and Anding Zhu. Instantaneous Sample Indexed Magnitude-Selective Affine Function-Based Behavioral Model for Digital Predistortion of RF Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 66(11):5000–5010, Nov. 2018. doi:10.1109/TMTT.2018.2855134.

[123] Jianfeng Zhai, Yang Li, Chao Yu, Lei Zhang, Jianyi Zhou, and Wei Hong. A Band-Limited Canonical Piecewise-Linear Function-Based Behavioral Model for Wideband Power Amplifiers. *IEEE Microwave and Wireless Components Letters*, 27(11):1022–1024, Nov. 2017. doi:10.1109/LMWC.2017.2750092.

[124] Qing Luo, Chao Yu, and Xiao-Wei Zhu. A Dual-Input Canonical Piecewise-Linear Function-Based Model for Digital Predistortion of Multi-Antenna Transmitters. In *IEEE/MTT-S International Microwave Symposium - IMS*, pages 559–562, Jun. 2018. doi:10.1109/MWSYM.2018.8439236.

[125] Xin Liu, Wenhua Chen, Long Chen, Fadhel M. Ghannouchi, and Zhenghe Feng. Linearization for Hybrid Beamforming Array Utilizing Embedded Over-the-Air Diversity Feedbacks. *IEEE Transactions on Microwave Theory and Techniques*, 67(12):5235–5248, Dec. 2019. doi:10.1109/TMTT.2019.2944821.

[126] Qing Luo, Xiao-Wei Zhu, Chao Yu, and Wei Hong. Single-Receiver Over-the-Air Digital Predistortion for Massive MIMO Transmitters With Antenna Crosstalk. *IEEE Transactions on Microwave Theory and Techniques*, 68(1):301–315, Jan. 2020. doi:10.1109/TMTT.2019.2943287.

[127] Filipe M. Barradas, Luís C. Nunes, Telmo R. Cunha, Pedro M. Lavrador, Pedro M. Cabral, and José C. Pedro. Compensation of Long-Term Memory Effects on GaN HEMT-Based Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 65(9):3379–3388, Sep. 2017. doi:10.1109/TMTT.2017.2671368.

[128] Corrado Florian, Tommaso Cappello, Alberto Santarelli, Daniel Niessen, Fabio Filicori, and Zoya Popović. A Prepulsing Technique for the Characterization of GaN Power Amplifiers With Dynamic Supply Under Controlled Thermal and Trapping States. *IEEE Transactions on Microwave Theory and Techniques*, 65(12):5046–5062, Dec. 2017. doi:10.1109/TMTT.2017.2723003.

[129] Jinbo Li, Zhiwei Xu, Wei Hong, and Qun Jane Gu. A Cartesian Error Feedback Architecture. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 65(3):1133–1142, Mar. 2018. doi:10.1109/TCSI.2017.2761394.

[130] Jianfeng Zhai, Lei Zhang, Zhiqiang Yu, Jianyi Zhou, and Wei Hong. A Modified Canonical Piecewise-Linear Function-Based Behavioral Model for Wideband Power Amplifiers. *IEEE Microwave and Wireless Components Letters*, 26(3):195–197, Mar. 2016. doi:10.1109/LMWC.2016.2524512.

[131] Anding Zhu. Decomposed vector rotation-based behavioral modeling for digital predistortion of rf power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 63(2):737–744, Feb. 2015. doi:10.1109/TMTT.2014.2387853.

[132] Zonghao Wang, Wenhua Chen, and Gongzhe Su. Low Computational Complexity Digital Pre-Distortion for Broadband Power Amplifiers. In *Asia-Pacific Microwave Conference (APMC)*, volume 1, pages 1–3, Dec. 2015. doi:10.1109/APMC.2015.7411820.

[133] Zonghao Wang, Wenhua Chen, Gongzhe Su, Fadhel M. Ghannouchi, Zhenghe Feng, and Yuanan Liu. Low Computational Complexity Digital Predistortion Based on Direct Learning With Covariance Matrix. *IEEE Transactions on Microwave Theory and Techniques*, 65(11):4274–4284, Nov. 2017. doi:10.1109/TMTT.2017.2690290.

[134] Noel Kelly and Anding Zhu. Low-complexity stochastic optimization-based model extraction for digital predistortion of rf power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 64(5):1373–1382, May 2016. doi:10.1109/TMTT.2016.2547383.

[135] Noel Kelly and Anding Zhu. Direct Error-Searching SPSA-Based Model Extraction for Digital Predistortion of RF Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 66(3):1512–1523, Mar. 2018. doi:10.1109/TMTT.2017.2748128.

[136] Pere L. Gilabert, Gabriel Montoro, Teng Wang, M. Nieves Ruiz, and Jose A. Garcia. Comparison of Model Order Reduction Techniques for Digital Predistortion of Power Amplifiers. In *European Microwave Conference (EuMC)*, pages 182–185, New York, 2016. IEEE. ISSN: 2325-0305 WOS:000393581100047.

[137] Mingyu Li, Zhenxing Yang, Zhongming Zhang, Ruoyu Li, Qing Dong, and Shigetoshi Nakatake. Sparsity Adaptive Estimation of Memory Polynomial Based Models for Power Amplifier Behavioral Modeling. *IEEE Microwave and Wireless Components Letters*, 26(5):370–372, May 2016. doi:10.1109/LMWC.2016.2549024.

[138] Honglin Wu and Shu Wang. Adaptive sparsity matching pursuit algorithm for sparse reconstruction. *IEEE Signal Processing Letters*, 19(8):471–474, Aug. 2012. doi:10.1109/LSP.2012.2188793.

[139] Deanna Needell and Roman Vershynin. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):310–316, Apr. 2010. doi:10.1109/JSTSP.2010.2042412.

[140] Siqi Wang, Mazen Abi Hussein, Olivier Venard, and Geneviève Baudoin. Optimal Sizing of Generalized Memory Polynomial Model Structure Based on Hill-Climbing Heuristic. In *European Microwave Conference (EuMC)*, pages 190–193, Oct. 2016. doi:10.1109/EuMC.2016.7824310.

[141] Siqi Wang, Mazen Abi Hussein, Olivier Venard, and Geneviève Baudoin. A Novel Algorithm for Determining the Structure of Digital Predistortion Models. *IEEE Transactions on Vehicular Technology*, 67(8):7326–7340, Aug. 2018. doi:10.1109/TVT.2018.2833283.

[142] Siqi Wang, Mazen Abi Hussein, Olivier Venard, and Geneviève Baudoin. Optimal Sizing of Two-Stage Cascaded Sparse Memory Polynomial Model for High Power Amplifiers Linearization. *IEEE Transactions on Microwave Theory and Techniques*, 66(9):3958–3965, Sep. 2018. doi:10.1109/TMTT.2018.2838126.

[143] Jun Peng, Songbai He, Bingwen Wang, Zhijiang Dai, and Jingzhou Pang. Digital Predistortion for Power Amplifier Based on Sparse Bayesian Learning. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 63(9):828–832, Sep. 2016. doi:10.1109/TCSII.2016.2534718.

[144] Jun Peng, Songbai He, Zhijiang Dai, and Bingwen Wang. A Simplified Sparse Parameter Identification Algorithm Suitable for Power Amplifier Behavioral Modeling. *IEEE Microwave and Wireless Components Letters*, 27(3):290–292, Mar. 2017. doi:10.1109/LMWC.2017.2661682.

[145] A. Abdelhafiz, L. Behjat, and F. M. Ghannouchi. Generalized Memory Polynomial Model Dimension Selection Using Particle Swarm Optimization. *IEEE Microwave and Wireless Components Letters*, 28(2):96–98, Feb. 2018. doi:10.1109/LMWC.2017.2783847.

[146] P. Stoica and Y. Selen. Model-order selection: A review of information criterion rules. *IEEE Signal Processing Magazine*, 21(4):36–47, Jul. 2004. doi:10.1109/MSP.2004.1311138.

[147] Juan A. Becerra, María J. Madero-Ayora, Javier Reina-Tosina, Carlos Crespo-Cadenas, Javier García-Frías, and Gonzalo Arce. A Doubly Orthogonal Matching Pursuit Algorithm for Sparse Predistortion of Power Amplifiers. *IEEE Microwave and Wireless Components Letters*, 28(8):726–728, Aug. 2018. doi:10.1109/LMWC.2018.2845947.

[148] Abubaker Abdelhafiz, Andrew Kwan, Oualid Hammi, and Fadhel M. Ghannouchi. Digital predistortion of lte-a power amplifiers using compressed-sampling-based unstructured pruning of volterra series. *IEEE Transactions*

on Microwave Theory and Techniques, 62(11):2583–2593, Nov. 2014. doi:10.1109/TMTT.2014.2360845.

[149] Juan A. Becerra, Maria J. Madero-Ayora, and Carlos Crespo-Cadenas. Comparative Analysis of Greedy Pursuits for the Order Reduction of Wideband Digital Predistorters. *IEEE Transactions on Microwave Theory and Techniques*, 67(9):3575–3585, Sep. 2019. doi:10.1109/TMTT.2019.2928290.

[150] Juan A. Becerra, Maria Jose Madero Ayora, Javier Reina-Tosina, and Carlos Crespo-Cadenas. Sparse Identification of Volterra Models for Power Amplifiers Without Pseudoinverse Computation. *IEEE Transactions on Microwave Theory and Techniques*, 68(11):4570–4578, Nov. 2020. doi:10.1109/TMTT.2020.3016967.

[151] Qian Zhang, Wenhua Chen, and Zhenghe Feng. Reduced Cost Digital Predistortion Only With In-Phase Feedback Signal. *IEEE Microwave and Wireless Components Letters*, 28(3):257–259, Mar. 2018. doi:10.1109/LMWC.2018.2797541.

[152] Haoyu Wang, Gang Li, Yikang Zhang, Falin Liu, and Anding Zhu. Forward modeling assisted 1-bit data acquisition based model extraction for digital predistortion of RF power amplifiers. In *IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, pages 59–62, Jan. 2017. doi:10.1109/PAWR.2017.7875573.

[153] Yikang Zhang, Haoyu Wang, Gang Li, and Falin Liu. One-bit in-phase observation for direct learning-based digital predistortion with modified frequency-domain delay estimation and alignment. *International Journal of RF and Microwave Computer-Aided Engineering*, 27(9):e21149, 2017. doi:10.1002/mmce.21149.

[154] Ning Guan, Nan Wu, and Hua Wang. Model Identification for Digital Predistortion of Power Amplifier With Signed Regressor Algorithm. *IEEE Microwave and Wireless Components Letters*, 28(10):921–923, Oct. 2018. doi:10.1109/LMWC.2018.2860790.

[155] Pablo Pascual Campo, Vesa Lampu, Lauri Anttila, Alberto Brihuega, Markus Allén, Yan Guo, and Mikko Valkama. Closed-Loop Sign Algorithms for Low-Complexity Digital Predistortion: Methods and Performance. *IEEE Transactions on Microwave Theory and Techniques*, 69(1):1048–1062, Jan. 2021. doi:10.1109/TMTT.2020.3038316.

[156] Mahmoud Abdelaziz, Lauri Anttila, Chance Tarver, Kaipeng Li, Joseph R. Cavallaro, and Mikko Valkama. Low-Complexity Subband Digital Predistortion for Spurious Emission Suppression in Noncontiguous Spectrum Access. *IEEE Transactions on Microwave Theory and Techniques*, 64(11):3501–3517, Nov. 2016. doi:10.1109/TMTT.2016.2602208.

[157] Mahmoud Abdelaziz, Lauri Anttila, Adnan Kiayani, and Mikko Valkama. Decorrelation-Based Concurrent Digital Predistortion With a Single Feedback Path. *IEEE Transactions on Microwave Theory and Techniques*, 66(1):280–293, Jan. 2018. doi:10.1109/TMTT.2017.2706688.

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

[159] Farouk Mkadem, Anik Islam, and Slim Boumaiza. Multi-Band Complexity-Reduced Generalized-Memory-Polynomial Power-Amplifier Digital Predistortion. *IEEE Transactions on Microwave Theory and Techniques*, 64(6):1763–1774, Jun. 2016. doi:10.1109/TMTT.2016.2561279.

[160] Qian Zhang and Wenhua Chen. Digital Predistortion for 5G Wideband Power Amplifiers Using Multiple Band-Limited Feedback Signals. In *ARFTG Microwave Measurement Conference (ARFTG)*, pages 1–4, Jun. 2017. doi:10.1109/ARFTG.2017.8000818.

[161] Qian Zhang, Wenhua Chen, and Zhenghe Feng. Concurrent Dual-Band Digital Predistortion Implemented with Reduced Look-up-Tables. *Electronics Letters*, 53(12):802–803, Jun. 2017. doi:10.1049/el.2017.0787.

[162] Praveen Jaraut, Meenakshi Rawat, and Patrick Roblin. Digital Predistortion Technique for Low Resource Consumption Using Carrier Aggregated 4G/5G Signals. *Iet Microwaves Antennas & Propagation*, 13(2):197–207, Feb. 2019. doi:10.1049/iet-map.2018.5608.

[163] Hai Huang, Patrick Mitran, and Slim Boumaiza. Digital Predistortion Function Synthesis Using Undersampled Feedback Signal. *IEEE Microwave and Wireless Components Letters*, 26(10):855–857, Oct. 2016. doi:10.1109/LMWC.2016.2605500.

[164] Zonghao Wang, Wenhua Chen, Gongzhe Su, Fadhel M. Ghannouchi, Zhenghe Feng, and Yuanan Liu. Low Feedback Sampling Rate Digital Predistortion

for Wideband Wireless Transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 64(11):3528–3539, Nov. 2016. doi:10.1109/TMTT.2016.2602216.

[165] Yehia Beltagy, Patrick Mitran, and Slim Boumaiza. Direct Learning Algorithm for Digital Predistortion Training Using Sub-Nyquist Intermediate Frequency Feedback Signal. *IEEE Transactions on Microwave Theory and Techniques*, 67(1):267–277, Jan. 2019. doi:10.1109/TMTT.2018.2873331.

[166] André Prata, Diogo C. Ribeiro, Pedro Miguel Cruz, Arnaldo S. R. Oliveira, and Nuno Borges Carvalho. RF Subsampling Feedback Loop Technique for Concurrent Dual-Band PA Linearization. *IEEE Transactions on Microwave Theory and Techniques*, 64(12):4174–4182, Dec. 2016. doi:10.1109/TMTT.2016.2608893.

[167] Arthur Chung, Marwen Ben Rejeb, Yehia Beltagy, Ali M. Darwish, H. Alfred Hung, and Slim Boumaiza. IQ Imbalance Compensation and Digital Predistortion for Millimeter-Wave Transmitters Using Reduced Sampling Rate Observations. *IEEE Transactions on Microwave Theory and Techniques*, 66(7):3433–3442, Jul. 2018. doi:10.1109/TMTT.2018.2817224.

[168] Yue Li, Xiaoyu Wang, and Anding Zhu. Sampling Rate Reduction for Digital Predistortion of Broadband RF Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 68(3):1054–1064, Mar. 2020. doi:10.1109/TMTT.2019.2944813.

[169] Cuiping Yu, Quan Tang, and Yuanan Liu. A Novel Indirect Learning Digital Predistortion Architecture Only with In-Phase Component. In *Asia-Pacific Microwave Conference (APMC)*, pages 995–997, Nov. 2018. doi:10.23919/APMC.2018.8617557.

[170] Yikang Zhang, Gang Li, Hongmin Li, Wen Qiao, and Falin Liu. In-Phase or Quadrature Observation for Indirect Learning Architecture Digital Predistortion Method Based on Forward Modeling. In *Asia-Pacific Microwave Conference (APMC)*, pages 938–940, Nov. 2018. doi:10.23919/APMC.2018.8617577.

[171] Y. Liu, J. J. Yan, H. T. Dabag, and P. M. Asbeck. Novel Technique for Wideband Digital Predistortion of Power Amplifiers With an Under-Sampling ADC. *IEEE Transactions on Microwave Theory and Techniques*, 62(11):2604–2617, Nov. 2014. doi:10.1109/TMTT.2014.2360398.

[172] L. Zhang and Y. Feng. An Improved Digital Predistortion in Wideband Wireless Transmitters Using an Under-Sampled Feedback Loop. *IEEE Communications Letters*, 20(5):910–913, May 2016. doi:10.1109/LCOMM.2016.2546257.

[173] Z. Wang, S. Ibrahim, H. Su, and R. Farrell. Generalised Digital Predistortion of RF Power Amplifiers with Low-Rate Feedback Signal. In *European Microwave Conference (EuMC)*, pages 831–834, Oct. 2016. doi:10.1109/EuMC.2016.7824472.

[174] Z. Wang, L. Guan, and R. Farrell. Compact Undersampled Digital Predistortion for Flexible Single-chain Multi-band RF Transmitter. In *IEEE MTT-S International Microwave Symposium (IMS)*, pages 1542–1545, Jun. 2017. doi:10.1109/MWSYM.2017.8058923.

[175] Z. Wang, L. Guan, and R. Farrell. Undersampling Observation-Based Compact Digital Predistortion for Single-Chain Multiband and Wideband Direct-to-RF Transmitter. *IEEE Transactions on Microwave Theory and Techniques*, 65(12):5274–5283, Dec. 2017. doi:10.1109/TMTT.2017.2758366.

[176] Q. Zhang, W. Chen, Z. Wang, and G. Su. A Single Feedback Architecture for Dual-Band Digital Predistortion With Under-Sampling Technique. In *IEEE MTT-S International Wireless Symposium (IWS)*, pages 1–4, Mar. 2016. doi:10.1109/IEEE-IWS.2016.7585462.

[177] Tony F. Chan. Rank Revealing QR Factorizations. *Linear Algebra and its Applications*, 88-89:67 – 82, 1987. doi:https://doi.org/10.1016/0024-3795(87)90103-0.

[178] Rong Zhu. Gradient-Based Sampling: An Adaptive Importance Sampling for Least-squares. In *Conference on Neural Information Processing Systems (NIPS 2016)*, Dec. 2016.

[179] H. Nyquist. Certain Topics in Telegraph Transmission Theory. *Transactions of the American Institute of Electrical Engineers*, 47(2):617–644, Apr. 1928. doi:10.1109/T-AIEE.1928.5055024.

[180] C.E. Shannon. Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1):10–21, Jan. 1949. doi:10.1109/JRPROC.1949.232969.

[181] L. Guan and A. Zhu. Optimized Low-Complexity Implementation of Least Squares Based Model Extraction for Digital Predistortion of RF Power Amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 60(3):594–603, Mar. 2012. doi:10.1109/TMTT.2011.2182656.

[182] A. A. Rontogiannis and S. Theodoridis. New Fast Inverse QR Least Squares Adaptive Algorithms. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 1412–1415 vol.2, May 1995. doi:10.1109/ICASSP.1995.480506.

[183] T. Wang, P. L. Gilabert, and G. Montoro. Under-Sampling Effects and Computational Cost Reduction in RF Power Amplifier Behavioral Modeling. In *European Microwave Integrated Circuits Conference (EuMIC)*, pages 57–60, Sep. 2015. doi:10.1109/EuMIC.2015.7345067.

[184] Z. Wang, J. Dooley, K. Finnerty, and R. Farrell. Selection of Compressed Training Data for RF Power Amplifier Behavioral Modeling. In *European Microwave Integrated Circuits Conference (EuMIC)*, pages 53–56, Sep. 2015. doi:10.1109/EuMIC.2015.7345066.

[185] Oliver Kramer. *Genetic Algorithm Essentials*, volume 679 of *Studies in Computational Intelligence*. Springer, 2017.

[186] Analog Devices, Inc. *Data Sheet AD9690*, May 2017. Rev. B.

[187] Analog Devices, Inc. *Data Sheet AD9629*, Feb. 2017. Rev. B.

[188] Mini-Circuits. *Data Sheet Gali 24+*. Rev. B.

[189] Qorvo, Inc. *Data Sheet TGF2965-SM*, May 2019. Rev. B.

[190] Qorvo, Inc. *Data Sheet T2G6003028-FL*, May 2021. Rev. C.

[191] Mini-Circuits. *Data Sheet ZHL-5W-2G-S+, ZHL-5W-2GX-S+*. Rev. H.

[192] Analog Devices, Inc. *Data Sheet ADL5610*, Dec. 2015. Rev. B.

[193] Y. Hou, J. Qu, Z. Tian, M. Atef, K. Yousef, Y. Lian, and G. Wang. A 61-nW Level-Crossing ADC With Adaptive Sampling for Biomedical Applications. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 66(1):56–60, Jan. 2019. doi:10.1109/TCSII.2018.2841037.

[194] T. Gotthans, G. Baudoin, and A. Mbaye. Influence of Delay Mismatch on Digital Predistortion for Power Amplifiers. In *Proceedings of the 20th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES*, pages 490–493, Jun. 2013.

[195] Xilinx, Inc. *Virtex UltraScale+ FPGA Data Sheet: DC and AC Switching Characteristics*, Nov. 2019. DS923 (v1.14).

[196] Semiconductor Components Industries, LLC. *Data Sheet NB7V52M*, Jul. 2014. Rev. 4.

[197] Xilinx, Inc. *ZCU111 Board User Guide*, Oct. 2018. UG1271 (v1.2).

[198] Analog Devices, Inc. *Data Sheet HMC8191*, Aug. 2019. Rev. C.

[199] Rohde & Schwarz GmbH & Co. KG. *SMF100A Microwave Signal Generator – Specifications*, May 2019. Data Sheet | Version 06.00.

[200] Analog Devices, Inc. *Coaxial Wideband Amplifier ZJL-153+*. REV. OR.

[201] Rohde & Schwarz GmbH & Co. KG. *FSUP Signal Source Analyzer – Specifications*, May 2011. Data Sheet | Version 06.03.

[202] Keysight Technologies. *Keysight 8494/95/96A/B Operating and Service Manual*, Oct. 2021. Edition 7.

[203] Qorvo, Inc. *Data Sheet QPA1010*, May 2021. Rev. H.

[204] Xilinx, Inc. *Spartan-3A DSP Starter Platform User Guide*, Jan. 2009. UG454 (v1.1).

[205] Xilinx, Inc. *EXP High-Speed Dititал-to-Analog Converter Module – User Guide*, Sep. 2009. Rev. C.

[206] Rohde & Schwarz GmbH & Co. KG. *FSVR Real-Time Spectrum Analyzer – Specifications*, Jul. 2021. Data Sheet | Version 04.00.

[207] Texas Instruments Incroporated. *Data Sheet DAC5682Z*, Jan. 2015. Rev. F.

[208] Texas Instruments Incroporated. *Data Sheet TRF370315, TRF370333*, May 2011. Rev. J.

[209] Texas Instruments Incroporated. *Data Sheet TRF371125*, Dec. 2010. Rev. B.

[210] Analog Devices, Inc. *Data Sheet ADCMP580/ADCMP581/ADCMP582*, Apr. 2016. Rev. B.

[211] Texas Instruments Incroporated. *Data Sheet for LMK04100, LMK04101, LMK04102, LMK04110, LMK04111, LMK04131, LMK04133*, Nov. 2012. Rev. B.

[212] Lukas Jagla. Adaptation of digital predistorter to linearize amplifiers using comparator. Master's thesis, Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Radio Electronics, 2020.

[213] Xilinx, Inc. *MicroBlaze Micro Controller System v3.0 – LogiCORE IP Product Guide*, Jul. 2021. PG116.

[214] Mahantesh P. Mattada and Hansraj Guhilot. Time-to-digital converters—A comprehensive review. *International Journal of Circuit Theory and Applications*, 49(3):778–800, Mar. 2021. doi:10.1002/cta.2936.

Author's Bibliography

[A1] Jan Kral, Tomas Gotthans, Roman Marsalek, Lukas Jagla, Michal Harvanek, Martin Pospisil, and Markus Rupp. Digital Predistorter Adaptation with a Level-Crossing Analogue-to-Digital Converter [Submitted]. 2022.

[A2] Jan Kral, Tomas Gotthans, Roman Marsalek, Michal Harvanek, and Markus Rupp. On Feedback Sample Selection Methods Allowing Lightweight Digital Predistorter Adaptation. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 67(6):1976–1988, Jun. 2020. doi:10.1109/TCSI.2020.2975532.

[A3] Jan Kral, Tomas Gothans, Roman Marsalek, and Michal Harvanek. Digital Predistorter with Real-Valued Feedback Employing Forward Model Estimation. In *International Conference on Telecommunications (ICT)*, pages 471–475, Jun. 2018. doi:10.1109/ICT.2018.8464937.

[A4] Jan Kral, Tomas Gotthans, and Michal Harvanek. Analytical Method of Fractional Sample Period Synchronisation for Digital Predistortion Systems. In *International Conference Radioelektronika (RADIOELEKTRONIKA)*, pages 1–5, Apr. 2017. doi:10.1109/RADIOELEK.2017.7937603.

[A5] Jan Kral and Tomas Gothans. Direct Learning Architecture For Digital Predistortion with Real-Valued Feedback. In *Conference STUDENT EEICT 2018*, pages 332–336, 2018.

[A6] Jan Kral and Tomas Gothans. Evaluation of Influence of Anti-Aliasing and Reconstruction Filters on Digital Predistortion. In *Conference STUDENT EEICT 2017*, pages 405–409, 2017.

[A7] Michal Harvanek, Roman Marsalek, Jan Kral, Tomas Gotthans, Jiri Blumenstein, Martin Pospisil, and Markus Rupp. Adjacent Channel Interference Cancellation in FDM Transmissions. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 67(12):5417–5428, Dec. 2020. doi:10.1109/TCSI.2020.2995350.

[A8] Tomáš Götthans, Roman Maršálek, Jan Král, and Tomáš Urbanec. Linearity and Efficiency Enhancement Techniques for Satellite Communications. In *IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, pages 54–57, Jan. 2022. doi:10.1109/PAWR53092.2022.9719692.

[A9] Tomas Gotthans, Roman Maršálek, Martin Pospíšil, Tomáš Urbanec, Jan Král, and Jiří Blumenstein. On the Lower I/Q Imbalance Sensitivity Using Real-Valued Feedback of Digital Predistortion. In *IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, pages 79–82, Jan. 2018. doi:10.1109/PAWR.2018.8310073.

[A10] Michal Harvanek, Roman Marsalek, Tomas Gotthans, Jan Kral, and Martin Pospisil. Influence of Transceiver Imperfections to Digital Pre-Distortion Linearization — Application to 5G Waveforms. In *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 238–242, Dec. 2017. doi:10.1109/ISSPIT.2017.8388648.

Abbreviations

5G	fifth generation
6G	sixth generation
64-QAM	64-state quadrature amplitude modulation
ACPR	adjacent channel power ratio
ADC	analogue-to-digital converter
AM/AM	amplitude/amplitude
AM/PM	amplitude/phase
APD	analogue predistorter
ASSA	adaptive signal separation algorithm
BRAM	block random-access memory
CML	current-mode logic
CP	cyclic prefix
DAC	digital-to-analogue converter
DC	direct current
DDR	double data rate
DDR2	simplified 2^{nd}-order dynamic deviation reduction-based Volterra
DFT	discrete Fourier transform
DLA	direct learning architecture
DNC	do not converge
DPD	digital predistorter
DRAM	dynamic random access memory
DSP	digital signal processing
DVR	decomposed vector rotation
EDH	evenly distributed histogram

EVM	error vector magnitude
F-OFDM	filtered orthogonal frequency-division multiplexing
FBMC	filter bank multicarrier
FF	flip-flop
FFT	fast Fourier transform
FM-ILA	forward model indirect learning architecture
FPGA	field programmable gate array
GMP	generalised memory polynomial
GOH	genetically optimised histogram
GS	gradient-based sampling
GSS	GS-based sample selection
HEMT	high-electron-mobility transistor
IFFT	inverse fast Fourier transform
ILA	indirect learning architecture
ILC	iterative learning control
IMD	intermodulation distortion
IQ	in-phase and quadrature
LC-ADC	level-crossing analogue-to-digital converter
LPF	low-pass filter
LS	least squares
LSDAC	low-speed digital-to-analogue converter
LUT	look-up table
LVDS	low-voltage differential signaling
MIMO	multiple-input and multiple-output
MP	memory polynomial

MSE mean square error

NMSE normalised mean square error

NN neural network

OFDM orthogonal frequency-division multiplexing

OMP orthogonal matching pursuit

PA power amplifier

PAE power-added efficiency

PAPR peak-to-average power ratio

PC personal computer

PCA principal component analysis

PCB printed circuit board

pdf probability density function

PSD power spectral density

QRS QR-decomposition-based sample selection

R-DLA real-valued direct learning architecture

R-FM-ILA real-valued forward-model indirect learning architecture

RAM random-access Memory

ReLU rectified linear units

RF radio frequency

RFSoC radio frequency system-on-chip

RMS root mean square

SDRAM synchronous dynamic random-access memory

SH sample and hold

SH sample and hold

SSM sample selection method

SVR	support vector regression
TCP	transmission control protocol
UDP	user datagram protocol
US	undersampling-based sample selection

Symbols

j	complex unit
$\arg\{\cdot\}$	angle of a complex number
$\lvert\cdot\rvert$	magnitude of a complex number
$(\cdot)_\mathrm{r}$	real part of a complex number
$(\cdot)_\mathrm{i}$	imaginary part of a complex number
$(\cdot)^*$	complex conjugate
$(\cdot)^T$	transposed matrix
$(\cdot)^H$	Hermitian matrix
$\lVert\cdot\rVert$	l_2-norm
$\lfloor\cdot\rfloor$	number rounded down towards zero
$\overline{\overline{D_j}}$	cardinality (number of elements) of set D_j
$\mathcal{K}$	set of active OFDM subcarriers
$\mathcal{O}$	overlapping factor in FBMC
$\mathcal{P}(\cdot)$	nonlinear transfer function of the PA
ΔT	time spacing between transmitted OFDM symbols
ΔW	tone offset for the F-OFDM modulation
$2\Delta W$	difference between the desired passband and the designed sinc filter passband
γ	minimum sample time distance in SSMs
$\epsilon(t)$	error signal
η_PAE	power added efficiency
θ_{j-1}	lower boundary of the j-th bin in histogram-based SSMs
θ_j	higher boundary of the j-th bin in histogram-based SSMs
μ	iteration step size
$\rho(t)$	modulation prototype filter
$\phi_{k,q}^{(x)}(\cdot)$	basis function with k nonlinearity order and q memory index created for signal x
$\boldsymbol{b}$	vector of PA model coefficients
$\boldsymbol{b}'$	vector of DPD coefficients
$\hat{\boldsymbol{b}}$	vector of updated DPD coefficients
$\boldsymbol{e}$	vector of coefficient errors
$\tilde{\boldsymbol{y}}$	vector of forward PA model output samples
$\hat{y}_\mathrm{r}(t)$	real part of the PA output including imperfections of down-converting mixer and active balun
A	constant amplitude
B	channel bandwidth
C_DAC	set LSDAC code

d_j	target count for the j-th bin in histogram-based SSMs
$d_{k,m}$	transmitted complex OFDM symbol at the m-th time position and the k-th subcarrier
D_j	set of selected indices n_i whose samples belong to the j-th bin
$f_{\mathrm{adj1}}, f_{\mathrm{adj2}}$	bound frequencies of the adjacent channel for ACPR calculation
$f_B[n]$	prototype-filter coefficients for the F-OFDM modulation
f_{c}	carrier frequency
f_{clk}	clock frequency
$f_{\mathrm{m1}}, f_{\mathrm{m2}}$	bound frequencies of the main channel for ACPR calculation
F	frequency spacing between OFDM subcarriers
$F_{\mathrm{AM}}(\cdot)$	function of AM/AM characteristics
$F_{\mathrm{PM}}(\cdot)$	function of AM/PM characteristics
$F_{\mathrm{post}}(\cdot)$	nonlinear function of postdistorter
$F_{\mathrm{pre}}(\cdot)$	nonlinear function of predistorter
F_{s}	sampling frequency
$g_{k,m}(t)$	modulation pulse at the m-th time position and the k-th subcarrier
G	linear PA gain
H_i	numerical coefficients for PHYDYAS filter
J	number of histogram bins in histogram-based SSMs
k	general index, mainly nonlinearity order index of a DPD or PA model
K	maximum nonlinearity order of a DPD or PA model
K_a	GMP index array for nonlinearity order of aligned signal and envelope
K_b	GMP index array for nonlinearity order of signal and lagging envelope
K_c	GMP index array for nonlinearity order of signal and leading envelope
L	number of taps of the prototype filter for the F-OFDM modulation
L_a	GMP index array for memory back shift of aligned signal and envelope
L_b	GMP index array for memory back shift of signal and lagging envelope
L_c	GMP index array for memory back shift of signal and leading envelope
M_b	GMP index array for memory forward shift of signal and lagging envelope
M_c	GMP index array for memory forward shift of signal and leading envelope
m	general index, mainly memory order index of a DPD or PA model
n	sample index
N	number of selected samples or number of averaged values
N_0	number of acquired samples
N_S	number of demodulated symbols
$p_B[n]$	sinc function for the F-OFDM modulation
P	number of model coefficients

P_{DC}	power consumed from power supply rails
P_{pwr}	power consumption
P_x	amplifier input-signal power
P_y	amplifier output-signal power
q	memory order index of a DPD or PA model
Q	maximum memory length of a DPD or PA model
$r(t)$	reference voltage signal
r_{DAC}	LSDAC output voltage
t	time
t_{d}	time of delay
t_{f}	subperiod delay of generated edges
t_{h}	hold time of digital FF
t_{high}	time of the comparator output in the high state
t_{su}	setup time of digital FF
T	sampling period of discrete-time signals
T_0	time-scaling parameter in FBMC
T_{clk}	clock period
T_{per}	period of the test sinewave signal
$w[n]$	window function for the F-OFDM modulation
W	number of assigned data subcarriers for the F-OFDM modulation
x	PA input
y	PA output
z	desired PA output and usually DPD input
Z	FFT length for the F-OFDM modulation