Characterization and Linearization of 5G mm-Wave Power Amplifiers Using Reduced-Sampling Observation Receiver

by

Yehia Beltagy

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2017

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Yehia Beltagy

Abstract

In this work, an advanced calibration routine is proposed to determine the frequency response of a feedback receiver over a targeted linearization bandwidth, when only sub-Nyquist (aliased) samples are available. A new approach is then devised to apply a direct learning algorithm along with the proposed receiver calibration routine, and thus linearize a millimeter wave power amplifier (PA), driven by a modulated signal, using digital pre-distortion (DPD) with a reduced feedback sampling rate. The proposed new calibration routine and DPD approach are successfully applied to linearize a PA under test, operating at 24 GHz and driven by single carrier 16-QAM and carrier aggregated LTE signals of 200MHz modulation bandwidth using a feedback receiver with sampling rates of 2 Gsps, 1 Gsps and 500 Msps. Adjacent channel power ratio of about 49 dBc and normalized mean square error of about 2% are obtained at the linearized PA output using the three sampling rates.

Acknowledgements

I would like to thank every one who was evolved in successfully finishing this Masters. I see my self very lucky to have the chance to meet sincere people who shaped me in a way that made the completion of this Master possible. Starting from my family that grew me up with passion to science, and ending with the people I met during my graduate studies here in Waterloo. I would like to thank my teachers and mentors who taught me in an early stage of my life how to approach difficult problems. I want to thank my professors who taught me the engineering basics and inspired me to start my academic career journey. I want to thank all my friends and family who helped me to make a seamless move to this new country. Finally, I want to thank my god, Allah, who gave me the chance to meet these people, and the resources to learn from them.

Table of Contents

vi

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation for Wide-Band Low-Power Digital Pre-distortion

Global mobile traffic is expected to reach 62ZB/month by 2020, compared to 885EB/month in 2012[1][3]. This continuous growth in traffic is not only due to the support for higher data rates, but also due to the increase in the number of devices accessing the network[4]. It is expected that the number of devices connected to the IP networks will be three times as high as the global population by 2021[6]. This support for higher data rates and dense connectivity will allow for many emerging applications: IoT, vehicle to vehicle communication, augmented realty, etc. These different applications require, in general, enhanced mobile broadband, ultra-reliable and low latency communication, or massive machine type communication[4]. These requirements establish the new communication requirements for 5G. Although 5G is still amorphous, it is agreed on that the above mentioned applications will require several hundred of MHz up to at least 1GHz of signals modulation bandwidths. Moreover, having this wide bandwidth as one contiguous bandwidth would enhance data delivery efficiency and ease the complexity of hardware implementation[4]. Therefore, mm-wave unlicensed bands at 30 and 60 GHz are investigated for these wide bandwidth applications. Furthermore, having small cell size base stations will allow the mobile network to improve in both capacity and coverage, especially at mm-wave frequencies as the path-loss is anticipated to be high[5].

---

[1]ZB: zettabytt; ZB = 1000^4GB; trillion gigabyte. EB: exabyte; ZB = 1000EB.

These small cells base station power amplifiers (PAs) will have the same PA linearity-efficiency trade-off as conventional PAs used in today's communication systems [17]. In detail, PAs can either work in an efficient region while exhibiting nonlinear behavior—introducing unwanted distortion—, or in contrast, in an inefficient but linear region. Since most of the base-station DC power is dissipated to operate and cool the PA [13], designing efficient PAs is of great importance to reduce such wasted power. Digital pre-distortion (DPD) is one solution to resolve this linearity-efficiency trade-off, such that the PA is operated at the efficient, yet, nonlinear region. Then the DPD mitigates the PA nonlinear effects. Designing the DPD for 5G mm-wave PAs deployed in small cells must incorporate the wide-band anticipated signals.

1.2   Problem Statement

DPD solutions operate a PA in the efficient nonlinear region. Then the DPD tries to mitigate the PA nonlinear behavior. The DPD is only useful when the absolute power saved by operating the PA in the nonlinear region—as a result of allowing high PA efficiency—is larger than the power consumed by the added DPD extra circuitry. These two powers relate differently to the RF PA output power. The saved absolute power is, inherently, directly related to the RF output power, i.e., a DPD linearizing a 40dBm PA saves more power than a 30dBm one, given both PAs operate at the same high PA efficiency after applying the DPD. In contrast, the overhead consumed power to operate the added DPD is independent of the PA output RF power. In other words, for the previously mentioned 30 and 40dBm case, the DPD circuitry consumes the same power, even though the absolute saved power by the DPD in the two cases are not the same. Collectively, DPD solutions are advantageous for high power transmitters, since DPD can achieve higher system efficiency[2].

Unfortunately, small cell base stations are known for having low RF output power, since they have small coverage areas. Therefore, the DPD added circuity must consume a small amount of power, to achieve an improved system efficiency. Knowing small cell base-stations for 5G applications are anticipated to transmit wide band signals, the DPD added circuity have to be wide-band enough to process these signals. However, operating these wide-band extra circuitry consumes more power. In particular, the transmitter-observation-receiver analog-to-digital-converter (TOR ADC) consumes most of DPD added circuity power. The ADC sampling speed is one key parameter that governs the ADC

---

[2]System efficiency incorporated the efficiency of PA and the DPD. For PA efficiency, the DC consumed power only accounts for the DC consumed power by the PA. In contrast, for system efficiency, the DC consumed power accounts for the PA and the DPD circuity consumed power as well

consumed power. This sampling speed should be high enough to capture all the PA non-linear output, which conventionally, has a bandwidth that is five times the original linear signal modulation bandwidth. Thus, this ADC with high sampling speed would consume substantiation amount of power, especially for the wide-band signals scenario. This is challenging since, for small cell base-stations, we have small margin of absolute saved power after utilizing the DPD.

1.3   Thesis Outline

This thesis investigates different methods to reduce the ADC sampling speed without affecting the wide-band DPD linearization capacity. Having a TOR that employs a low sampling ADC reduces the overall power consumption of the DPD circuitry. This thesis work is not the first attempt to reduce the DPD hardware consumed power by utilizing low sampling speed ADCs. Nevertheless, we tried to show a practical implementation for one of reduced sampling speed DPD ideas that was previously proposed.

After the first chapter, motivation, chapter 2 establishes the DPD theory and the important previous similar attempts to reduce the ADC sampling speed. By comparing these different approaches and their validation results, we draw a few conclusions that will guide our proposed sub-sampling method. In addition, it is evident from this literature review that an important practical implementation aspect was ignored in all of the presented work: receiver calibration using sub-sampled data. Hence the next chapters, in turn, fills this "practical implementation" gap by proposing a new calibration routine that is suitable for sub-sampling DPD[3] applications.

Chapter 3 proposes different techniques to tackle the sub-rate receiver calibration problem from a practical point of view. First, this chapter reviews the theoretical solution adopted by this thesis work. Then, this chapter highlights the missing practicability in that approach by introducing the receiver calibration problem. As a result, this chapter demonstrates a new receiver calibration routine that is suitable for the sub-sampling DPD application. This calibration routine requires a small modification to the conventional DPD system diagram. This chapter also discusses few assumptions regarding this modification on the DPD system that affects the DPD convergence behavior. Hence, this chapter recommends some settings for the direct algorithm used to have better convergence speed.

---

[3]Sub-sampling DPD and sub-rate DPD are used interchangeably in this thesis, which means a DPD that only uses reduced sampling speed data.

Chapter 4 demonstrates a DPD measurements to validate the previously proposed sub-sampling calibration and DPD techniques. The measurements setup used utilizes real sub-sampling TOR hardware, i.e., the TOR ADC runs at the reduced sampling speed. Worth mentioning that this type of measurement is missing from most of the previously presented solution in literature. Chapter four also provides a practical routine to identify the TOR and the transmitter linear nonidealities using sub-sampled data. Using this routine, the ADC sampling frequency is relaxed to a very low sampling speed value (below Nyquist); hence, the consumed ADC power is reduced. We showed in this chapter a linearization results for wide-band LTE and 16-QAM signals using different reduced sampling speeds.

Finally, chapter 5 provides the future work needed to complement this work. We still want to conduct a further study on the practical limit of the lowest reduced sampling speed (highest sub-rate factor) that can be used without affecting the DPD linearization capacity. We think that this limit can be constrained by the accuracy of time-alignment algorithms that only employ sub-sampled data. Hence, a better way of time aligning the signals in sub-sampling DPD is still needed to be devised.

Chapter 2

Background and Theory

This chapter motivates the need for low-power wide-band digital pre-distortion solutions. DPD is one solution for the PA linearity/efficiency trade-off dilemma that has been widely adopted in the past years. In brief, the DPD block should synthesize the PA inverse behavior, thus, the cascade of the DPD and the PA yields one linear block. To understand the different DPD design aspects, this chapter reviews the PA different non-linear effects.

DPD systems consist of three main blocks: DPD engine, estimator, and transmitter observation receiver (TOR). These modules are becoming more challenging to design and implement as communication standards are evolving, in particular, as the signal bandwidth and the peak power to average power ration (PAPR) increase. Out of those three DPD components, the TOR consumes most of the overhead power. In fact, the two other DPD blocks, engine and estimator, only require extra digital signal processing power, without the need of adding extra hardware.

Most of the TOR power, and the DPD system, is consumed in the TOR ADC. This ADC power is a linear function of the sampling speed. Hence, as the signal bandwidth increases, the ADC sampling speed and power increase accordingly. Therefore, there have been multiple attempts in the literature to reduce the ADC sampling speed without affecting the DPD linearization performance.

After establishing the bases of PA non-linearity and the wide-band DPD challenges, this chapter provides a review on the different approaches proposed to reduce the ADC sampling speed; since it is of great importance for wide-band DPD power consumption.

2.1 Power Amplifier as an Important Nonlinear Block

PAs are power gain blocks used to boost the signal power before it is transmitted over the air. This means PAs usually generate large amounts of power to allow for wider wireless signal coverage. The specifications of a PA that are related to this thesis scope are: power gain, maximum output power, and efficiency.

- **Power gain** G (2.1). It is understood from this equation that this is the *linear* power gain, since the gain is independent of the input power level. For nonlinear analysis, output and input powers can no longer be related using such simple relationship.

$$G = \frac{P_{out}}{P_{in}} \tag{2.1}$$

- **Output power** P_{1dB} (2.2). This the the output power at which the power gain is smaller than the linear gain by $1dB$. It is usually preferred not to operate the PA beyond this point, since the PA is very nonlinear.

$$P_{1dB} = P_{out} \quad | \quad G(P_{in})[dB] = G[dB] - 1 \tag{2.2}$$

- **Efficiency** η. Quantifies how the PA efficiently transforms the input power into an RF output power. There are two known definitions for PA efficiency: drain efficiency (2.3), and power added efficiency (2.4).

$$\eta_{DE} = \frac{P_{out}}{P_{DC}} \tag{2.3}$$

$$\eta_{PAE} = \frac{P_{out} - P_{in}}{P_{DC}} \tag{2.4}$$

  Any DC power that is not transformed into RF output power is lost in the form of heat. This heat, in turn, is needed to be properly handled using a suitable cooling system, which requires another extra consumed power to operate. Consequently, the PA efficiency is of great importance to reduce the total lost power in a base station transmitter.

For PAs, it is usually the case that efficiency trades off with linearity, Fig. 2.1 [14]. That is, efficiency is usually high for high output power, however, the PA is nonlinear in that region. This can be seen from Fig. 2.1: for high input power levels, the output

6
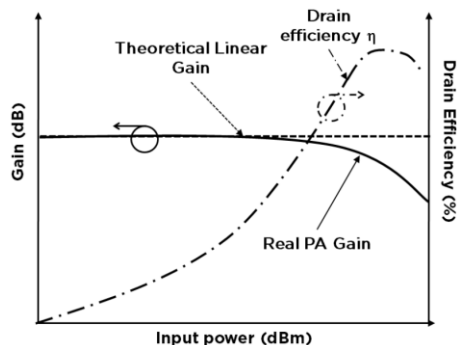
Figure 2.1: Typical linearlity efficiency trade off.

power compresses and the gain decreases. In DPD-like solutions, the PA is operated at the high efficient region, and DPD, in turn, mitigates the nonlinear effects. So, DPD does not affect the PA efficiency; it only allows for nonlinear region operation while reducing the PA nonlinear effects. This is possible since DPD tries to model the inverse function of the PA. To understand the different DPD models, the different PA nonlinear effects must be first quantified.

2.1.1 PA Nonlinearity

The formal definition of a linear block, gain block of gain $g(\cdot)$ for example, is the following

$$g(A + B) = g(A) + g(B) \tag{2.5}$$

In other words, the gain of that block is independent of the input drive level. For nonlinear PA, the gain value changes with input power level. Having a gain that is not constant, and a function of the input power level means the output cannot be expressed as a linear function of the input, as in (2.1). There are many causes of power amplifier nonlinearity. The most pronounced one is output power saturation, which shows up as a form of gain reduction for higher input power levels.

Since the gain is now a function of the input signal, input signal characteristics should be defined. In general, the majority of signal characteristics are dominated by three main quantities: average power, bandwidth and peak to average ratio (PAPR). Therefore, the PA behavior is different with the change of any the above mentioned input signal characteristics. Since average power and bandwidth are widely known parameters, only PAPR is reviewed next.

7

2.1.2 PAPR

PAPR quantifies the variation of signal envelope power. Having high PAPR signal means the signal envelop—instantaneous envelop power—has large variations. Hence, the amount of amplification for each of those instantaneous input power is different. In other words, for nonlinear amplifiers, the gain is input power dependent.

PAPR can be defined in two different ways, both should lead to the same value. Each one of these two methods use different signal for the analysis: either baseband (2.7) or RF signal (2.8). By RF we mean the signal that is generated by up-converting the baseband signal to the high carrier frequency, f_c. Baseband signal is a complex valued signal centered around the zero frequency. While RF is a real up-converted signal that is symmetric around the zero frequency. The two signals can be related by:

$$x_{RF}(t) = \Re(\widetilde{x}_{BB}(t) \cdot e^{jw_c t}) \tag{2.6}$$

$$PAPR(\widetilde{x}_{BB}) = \frac{P_{Peak}}{P_{Avg}} = \frac{\max(|\widetilde{x}_{BB}|^2)}{mean(|\widetilde{x}_{BB}|^2)} \tag{2.7}$$

$$PAPR(x_{RF}(t)) = \frac{P_{Peak}}{P_{Avg}} = \frac{\max\limits_{-\infty < to < \infty} \frac{1}{T_o} \int\limits_{T_o}^{To+to} x_{RF}^2 dt}{\lim\limits_{T \to \infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x_{RF}^2 dt} \tag{2.8}$$

A more practical method of determining the PAPR is by finding the CCDF of the time domain signal. Theoretically, as the signal frame time increases, the higher the chance of having higher peak power in this lengthy frame time, hence the larger the PAPR. This means the PAPR is frame time dependent, which is impractical. In contrast, the maximum powers in (2.7) or (2.8) can be defined as the power that only 0.1% of waveform signal samples exceeds. This statistical definition eliminates the increase in PAPR due to the outlier samples.

There are multiple reasons why signal PAPR might increase. For single carrier signals, modulation type plays the most important role in specifying the signal PAPR. Some modulation schemes do not have any envelop variations, such as PSK, 4-QAM. To achieve higher data rates, more information (bits) are encoded in one constellation point, like 64-QAM. This high index encoding leads to envelop power variation, in other words, higher PAPR values. In addition, carrier aggregated signals have, in general, larger PAPR values.

With the constant demand for higher data rate communications, it is usually the case that signals PAPR tend to increases, which adds more burden to the DPD. The higher the

PAPR and the wider the bandwidth of the PA input signal, the more PA non-linearities are stimulated. Knowing that DPD is the inverse of the PA, we need to first accurately measure and quantify the PA non-linear behavior. This is important to efficiently design the different DPD blocks. AM/AM, AM/PM are a visual tools used generally for quantifying the PA nonlinearities.

2.1.3 AM/AM, AM/PM

In the AM/AM, the instantaneous output power of each data point is plotted versus the instantaneous input power. For PA that is compressed (saturated), the curve shows gain drop as the input power signal level increases. Same curve can be plotted for the signal output phase, which is AM/PM. In AM/PM the phase response behavior of the PA is plotted versus input power. For low input power levels, the phase response should be constant. In contrast, as the input power increases, phase response starts to deviate from the linear value. Fig. 2.2 shows a typical AM/AM and AM/PM graphs for a Doherty power amplifier driven by a single tone sine wave. For a wide-band signal inputs, these graphs tend to have a different profile due to memory effects, which is explained next.

2.1.4 Memory Effects

For power amplifiers with memory, the output signal does not only depend on the instantaneous input signal, but on the L previous input samples as well. This shows as spectrum variations in the frequency domain, a non-flatness in the PA output spectrum. For AM/AM, and AM/PM curves, memory shows as a spreading of points around the traditional CW curve. This is, for certain input power, we can have multiple possible output powers, each corresponds to a different previous samples sequence[1], Fig. 2.2.

For memory effects, there are linear memory effects, where the frequency response explained before is not a function of the input drive level. On the other hand, nonlinear frequency response is when this frequency response is a function of the input power drive.

---

[1]To show why memory effects cause spectrum variations, assume, for simplicity, a linear system with an impulse response $h(t)$. With memory, $h(t)$ has non-zeros values for $t > 0$, since the output depends on the previous inputs. This shows as a non-constant frequency repose. In contrast, for memoryless case, $h(t)$ only has non-zero value for $t = 0$, which shows as a flat frequency response.
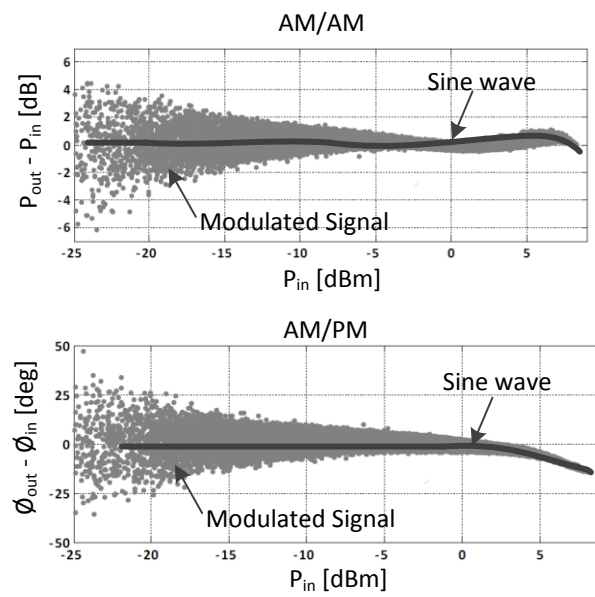
Figure 2.2: AM/AM, AM/PM for PA driven by single tone signal (CW) or modulated signal.

2.1.5   Effects of PA Distortion

As was mentioned before, operating the PA in the nonlinear region is preferred from the efficiency point of view. However, nonlinear PAs distort the signal; this distortion shows in two main ways. First, spectrum regrowth, which shows as out of band distortion. This out of band distortion can be observed by degradation in ACPR/ACLR. For any transmission device, certain transmission mask is required by the different regulation authorities before being certified for production.

The second type of distortion caused by non-linear PA is the deterioration in EVM/NMSE, which is dominated by the in-band distortion. EVM is measured for the demodulated I/Q signal; NMSE is computed for the raw time domain signal before demodulation, equation[2.9]. In this thesis, NMSE is used as a metric for the in-band distortion.

$$NMSE/EVM(\widetilde{x}_{1BB}, \widetilde{x}_{2BB}) = \sqrt{\frac{(I_1 - I_2)^2 + (Q_1 - Q_2)^2}{I_1{}^2 + Q_1{}^2}} \qquad (2.9)$$

As communication standards evolve, higher PAPR and wider bandwidth signals are used. High PAPR pushes the PA into a more nonlinear regions and stimulate more of its non-linearities. DPD is an efficient solution that has been adapted to mitigate these nonlinear effects by linearizing the PA.

2.2   PA Linearization using Digital Pre-distortion DPD

DPD applies the inverse function of the PA, such that the cascade of DPD and PA yield one linear block. For example, if PA introduces gain compression, hence a reduction in the PA output signal PAPR, DPD should introduce gain expansion which increases the signal PAPR. The cascade of the PA and DPD, compression and expansion, results in one linear block, Fig. 2.3, 2.4, and 2.5 show the PAPR of the different signals in the DPD system before and after convergence[2]. Worth noting that in these figures, we assume that the average power of all the signals is 0 dBm.

DPD is a behavioral modeling problem, where DPD tries to model the PA inverse function. In DPD-like solutions, we operate the PA in the high efficient, yet nonlinear

---

[2]From Fig. 2.5, the PAPR of the pre-distorted signal $\tilde{u}$ is 4dB larger than the original signal $\tilde{x}$. 4dB is considered a big increase to correct for 1dB of PA compression (from Fig. 2.3). There are crest factor reduction techniques to limit this PAPR increase in the predistorted signal. Crest factor reduction (CFR) is out side the scope of this thesis [16].

region. Then, DPD tries to mitigate the consequent nonlinear effects. DPD systems usually consist of three main components:

1. **Engine**, where DPD applies the inverse PA function. This inverse function is conventionally applied in digital domain. DPD behavioral models can be either polynomials based, or neural-networks based. [26] shows a comparative overview of these different models.

2. **Coefficients Estimator**, where the DPD coefficients are estimated. Two main estimators architectures are used: direct and indirect learning. [25, 10, 11] show a comparison between those two approaches.

3. **Transmitter observation receiver (TOR)**, an extra hardware block that receives the PA output to be used in coefficients identification.

2.2.1   Wide-Band Signals as a Challenge for DPD

As communication standards are evolving to support higher data rates, the modulation bandwidth of communication signals increases. This increase in bandwidth adds more burden to the DPD to correct for the PA non-linearities. In particular, PA used inside small cell base stations transmitting wide band signals can be very challenging for DPD. Concretely, wide-band signal linearization can be challenging for the following reasons:

1. **Complex DPD Models.** To support higher data rates, signals PAPR and bandwidth tend to increase. This means the DPD models should be complex enough to accommodate the newly excited non-linearities and linear/non-linear memory effects. With higher model complexity, more numerical instability in coefficients estimation is often observed, mainly due to over-fitting. Another disadvantage for increasing model complexity is the need for more digital signal processing power for the different optimization algorithms to identify these coefficients.

2. **DPD Circuit Power Consumption.** The power consumed by the DPD added circuitry reduces the total system efficiency. Thus, these circuitry should be carefully designed for low power consumption. Low power DPD is of great importance for small-cell base station as the PA RF power, and consequently the absolute power saved by DPD, is relatively low. The amount power consumed by the DPD circuity should be less than the amount of power saved by allowing the PA to operate in the
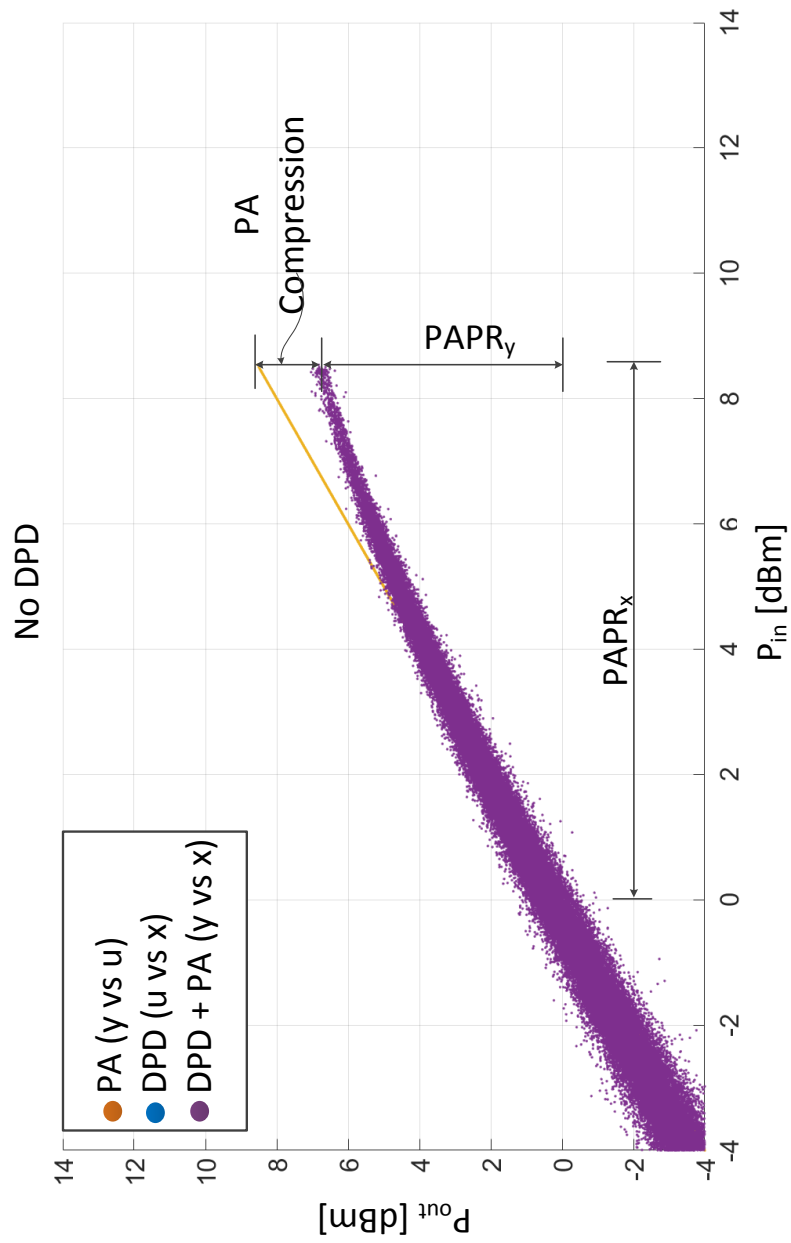
Figure 2.3: PAPR profile of $\tilde{x}$ and $\tilde{y}$ before DPD.
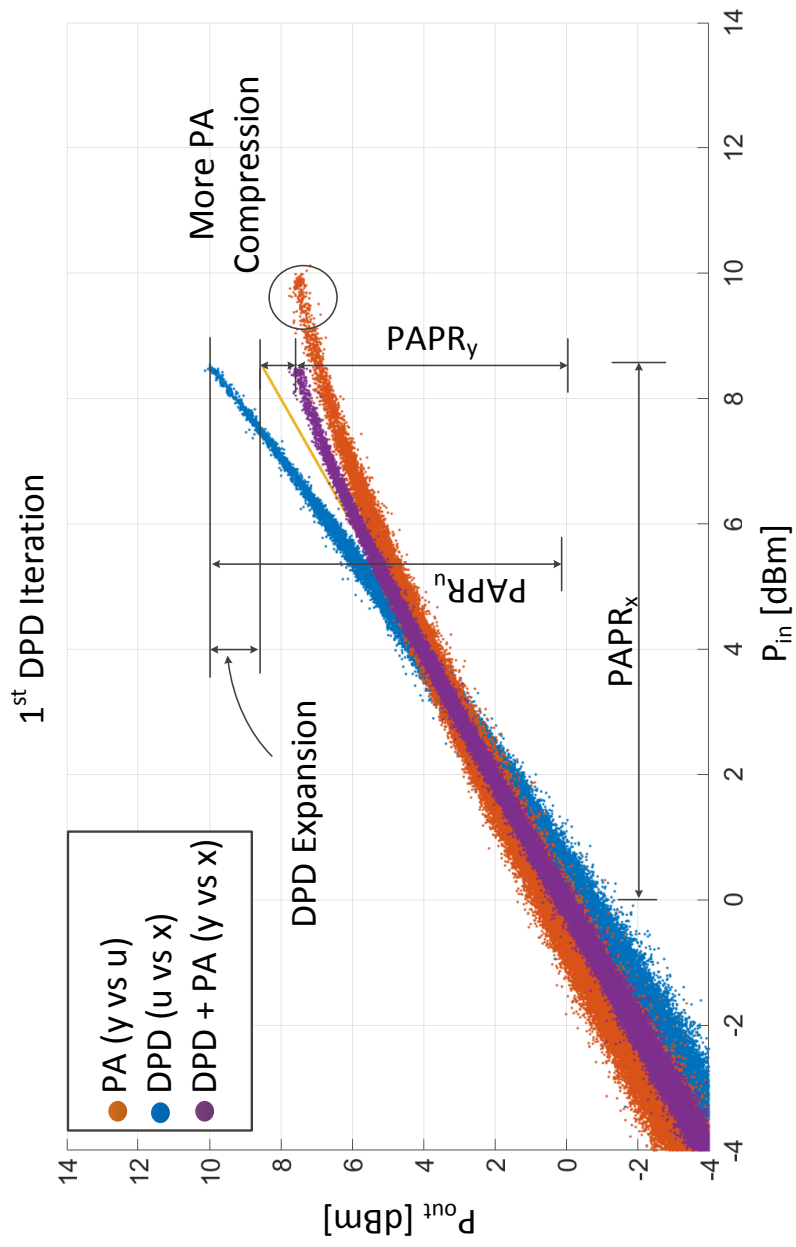
Figure 2.4: PAPR profile of the different signals, $\tilde{x}$, $\tilde{u}$, and $\tilde{y}$, in the DPD system after the first iteration.
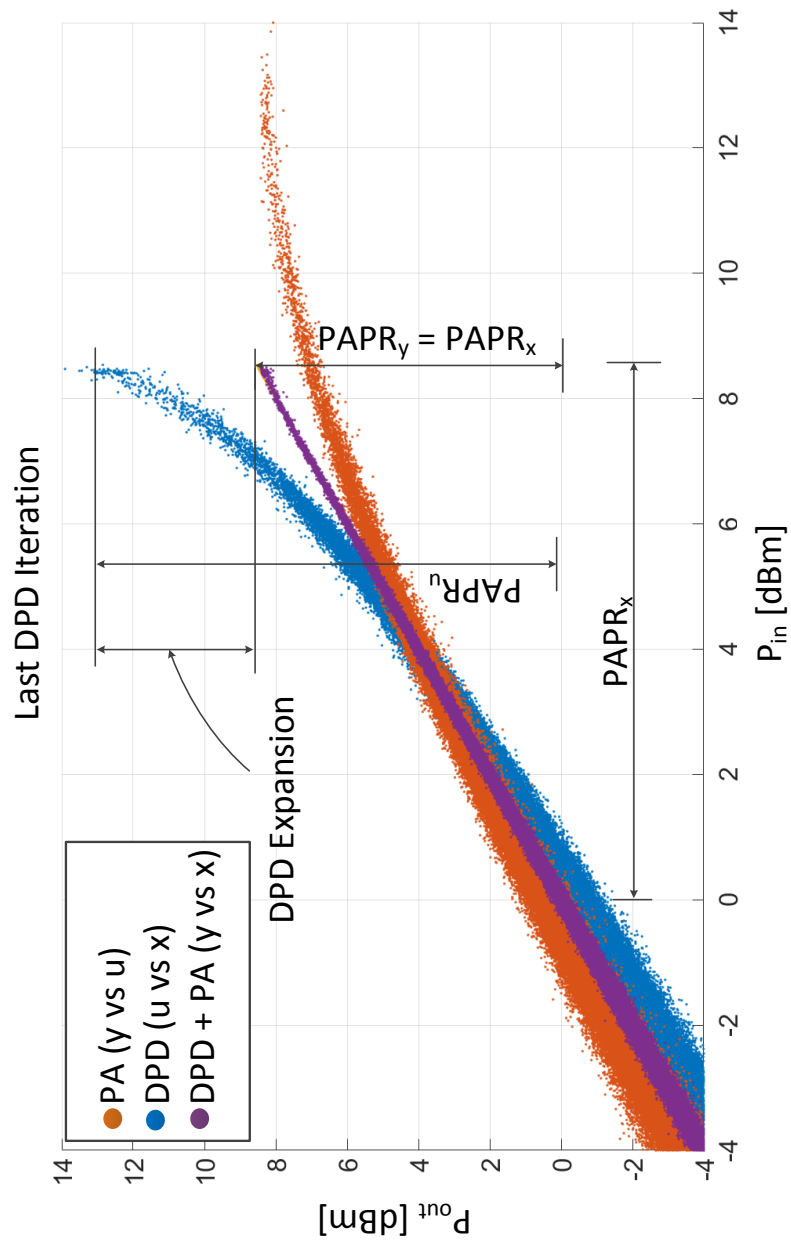
Figure 2.5: PAPR profile of the different signals, $\tilde{x}$, $\tilde{u}$, and $\tilde{y}$, in the DPD system after after DPD convergence.
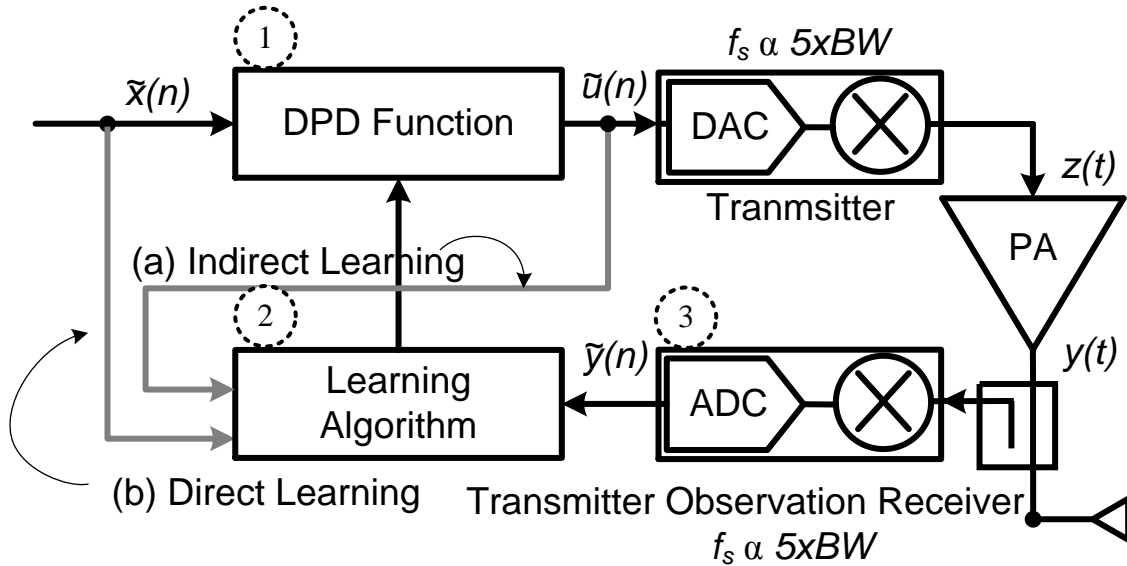
Figure 2.6: DPD system block diagram.

non-linear region. Unfortunately, for these small-cells transmitting wide band signals, no clear method exists that to lower the consumed power in the DPD circuitry. In other words, the DPD circuity consumption is not necessary scalable with the PA absolute operating power.

TOR power consumption is one example for this DPD added circuitry power consumption problem; as the PA signal bandwidth increases, the analogy circuity used inside the TOR should be all wide-band enough to capture this wide-band signal. This is translated to more power consumption, hence lower over all system efficiency. Knowing that the PA nonlinear output signal bandwidth usually occupies a bandwidth which is five time the original signal bandwidth, the ADC used inside the TOR should be running at a sampling speed that is high enough to capture all this nonlinear output. This high sampling speed requires big amount of power overhead that severely degrades the overall system efficiency. Indeed this ADC power consumption is the same regardless of the PA output power is 30 or 40dBm.

3. **TOR Calibration.** In DPD, the estimator attempts to minimizing the error between the original input signal, $\tilde{x}(n)$ and the PA output signal capture by the TOR, $\tilde{y}(n)$. In other words, as the DPD is converging, the lease square error between the signals that the estimator sees is minimal. However, for wide-band TOR, it is in-

evitable to have some variation in the receiver frequency response. As a result, as the DPD algorithm converges, the PA output signal will not approach the needed $\tilde{x}(n)$ signal. Instead, $y(t)$ converges to a $\tilde{x}(n) * h_{Rx}^{-1}(n)$, where h_{Rx}^{-1} is the impulse response that represents the inverse of the nonidealities of the TOR. This is the case since the signal estimator sees is $\tilde{x}(n) * h_{Rx}^{-1}(n)$, as appose to $\tilde{x}(n)$. Although this is a wrong solution, it provides the minimum error for the DPD estimator.

TOR design and calibration is the main scope of this thesis. One solution to reduce the TOR consumed power is to reduce the ADC sampling speed. There have been multiple attempts to run the DPD algorithm using under-sampled data. Next section presents a detailed review on these approaches. It will be evident that non of the proposed solution in the literature provided a complete practical DPD solution using sub-sampling ADC.

2.3   Literature Review on Reduced Sampling Rate DPD

There have been multiple attempts to reduce the sampling speed of the TOR ADC. In all these attempts, a band limited version of the PA wide-band output signal, either aliased or filtered, was used in DPD coefficient estimation. Since the signal bandwidth now is reduced, the ADC sampling speed can be reduced accordingly. These attempts can be categorized into two main groups.

1. **Sub-sampling DPD**. In these approaches, the DPD is trained without restoring the full bandwidth signal. DPD coefficients are estimated by only using the incomplete information of the sub-rated signal [24, 10, 32].

2. **Full-rate signal restoration**. In these approaches, signal processing techniques are proposed to restore the full rate wide-band PA output signal from the narrow band signal. The novelty in these approaches is how to construct the full bandwidth signal with minimum error. Once the signal in restored, normal DPD estimation algorithm is applied [30, 21, 22].

The following section reviews separately each group of these approaches. Then, comments are provided on each approach regarding important implementation issues like time alignment, validation measurements conducted, and final results. Finally, after this review, we introduce the work of this thesis as a better solution than most of the reviewed approaches.

2.3.1 Sub-Sampling DPD

In these approaches, the DPD model was trained using a filtered version of the PA output signal. Having a filtered signal using a bandpass filter means some of the signal frequency components are rejected. Hence, the information embedded in these components are completely lost. Having a smaller bandwidth signal reduces the ADC sampling-rate compared to the conventional DPD.

The most famous approach to reduce the TOR ADC sampling ADC was proposed in [32]. Chao *et al* proposed the idea of band-limited DPD, Fig. 2.7a. They modified the DPD model so as to linearize the PA output signal within certain preselected bandwidth. Any nonlinear components outside this bandwidth are rejected using a RF bandpass filter. As a result of that filter, the PA is needed to be linear within a certain bandwidth. To do so, the basis had to be filtered during coefficients estimation for DPD training, and during DPD application as well.

In contrast, authors in [10] and [24] tried to linearize the PA output across the whole output signal bandwidth. They tried to do this while capturing only a band-limited version of the signal. In [10], Braithwaite applied basis filtering only during coefficients estimation, Fig. 2.7b. This means after coefficients estimation and while applying the DPD, the basis was not filtered. He showed mathematical analysis based on superposition principle to support his approach. On the other hand, Farouk [24] did not apply any kind of basis filtering, Fig. 2.7c. He used the PA band-limited signal, as it is, for DPD coefficients training without any basis filtering during estimation or DPD application. In these two approaches the DPD was expected to extrapolate and linearize the PA signal at frequencies that were filtered, i.e., were not available during coefficients estimation.

These attempts demonstrated some gain in reducing the ADC sampling speed. Chao *et al* could linearize a 100MHz signal within a 140MHz of bandwidth. Braithwaite and Farouk both reduced the ADC sampling speed by a factor of 2 compared to the conventional DPD, while achieving full-bandwidth linearization. They could not reduce, however, the sampling speed substantially because they relied on behavioral model extrapolation capabilities. To further clarify their work shortcomings, next sections provides more details about the low level implementation. It will be evident that non of the proposed solutions demonstrated full linearization capabilities similar to conventional DPD.

Band-Limited DPD [32]

In [32], Chao *et al* proposed the idea of band-limited predistortion, such that the DPD will only linearize the PA output signal within certain pre-selected bandwidth. They suggested
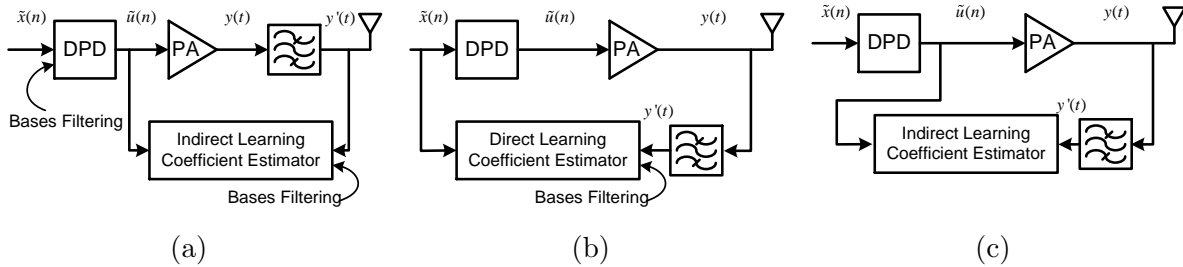
Figure 2.7: Sub-sampling DPD literature review. (a) Bases filtering during estimation and DPD application. (b) Bases filtering during estimation only. (c) No bases filtering.

to use a bandpass filter to reject all the nonlinear outputs outside the passband of this filter, Fig. 2.8. Besides adding this extra hardware, the basis function of whatever model that will be used must also be band-limited to the same bandpass filter bandwidth. This is necessary to match the bandwidth of the training (and linearization signal) to the basis waveforms bandwidth. This basis filtering can be shown in Fig. 2.9b, in contrast to Fig. 2.9a, and mathematically using 2.10 and 2.11, where L is a linear bandpass filter operator with a bandwidth that matches the added bandpass filter, and y' is the filtered output signal.

$$\text{Coefficients estimation: } \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ \psi_q(\tilde{y}'(n)) \Rightarrow \tilde{u}'(n) = \sum_{q=1}^{Q} a_q' \ L\{\psi_q(\tilde{y}'(n))\} \quad (2.10)$$

$$\text{DPD application: } \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ \psi_q(\tilde{x}(n)) \Rightarrow \tilde{u}'(n) = \sum_{q=1}^{Q} a_q' \ L\{\psi_q(\tilde{x}(n))\} \qquad (2.11)$$

By applying these two techniques, the authors had only to measure, and linearize, the PA output signal within that bandpass filter bandwidth. Hence, the ADC sampling rate now can be less than the one used in conventional DPD, Fig. 2.10a. One key drawback for this technique is the necessity to add the extra bandpass filter. This filter is not easily designed at high RF frequencies; moreover, it will be bulky, and with some insertion loss that will reduce the total PA efficiency. For these reasons the authors could not themselves add this bandpass filter. They instead introduced filtering in the digital domain. So, they measured with hardware that supports full-rate sampling; then they digitally filtered the data to the intended bandpass filter bandwidth. Finally, they used this signal in training

19
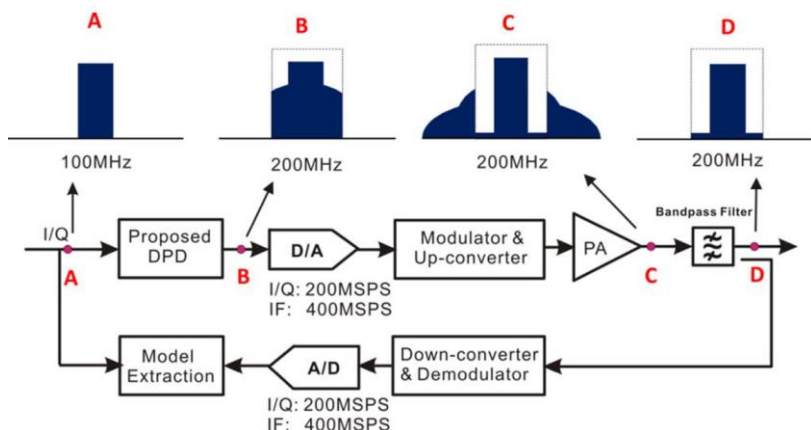
Figure 2.8: Band-limited DPD [32].

as a verification of their idea. Using this technique, wide-band signal linearization of modulation bandwidth of 100MHz signal was demonstrated. They linearized the signal within 140MHz of bandwidth, Fig. 2.10b.

Direct Learning Wide-band DPD [10]

For the second approach to reduce the ADC sampling rate, Braithwaite used direct learning and superposition principles to devise a wide-band pre-distorter. This pre-distorter can linearize wide-band signals by only capturing a narrow band filtered version of the same signal for coefficients estimation. Filtering here is done explicitly using an IF bandpass filter in the receiver, or implicitly by having a narrow band TOR.

In this approach, the bases waveforms were only filtered in the estimator, while solving for the coefficients. Once these coefficients were found, the DPD function was applied using the same nonlinear basis, without any filtering. This assumes that the DPD module is performing an extrapolation, i.e., the DPD module can still generate a correct output at frequency bands that were not available during coefficient estimation. The author justification of this approach based on the superposition principle and direct learning, which both will be explained next.

In direct learning, the error signal (difference between $\tilde{x}$ and $\tilde{y}$) is expressed as a weighted sum of nonlinear basis, where the weights are the error in coefficients, Δa_i, Equation 2.12. If a linear operation is applied to both hand side of the equation, mathematically speaking, the unknown (error coefficients Δa_i) will not change. For sub-sampling

20
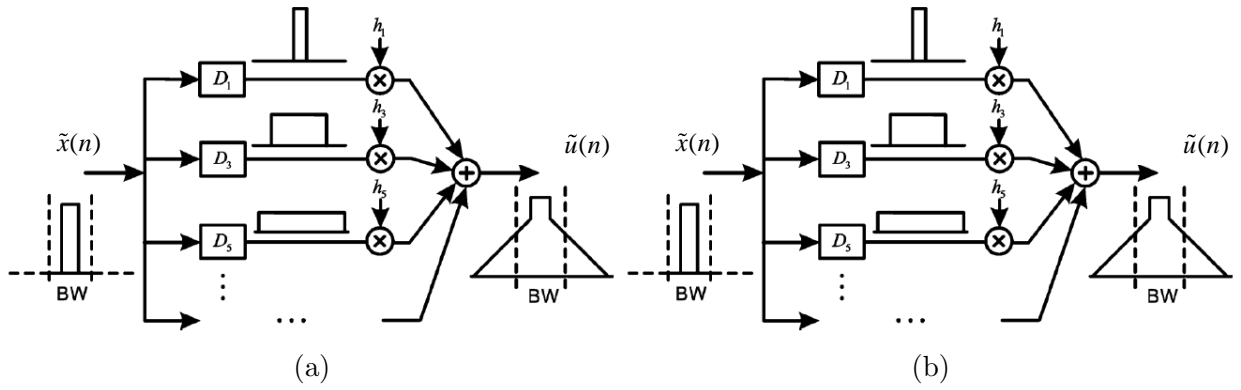
Figure 2.9: (a) conventional Volterra DPD model (b) band-limited propossed Volterra model [32].
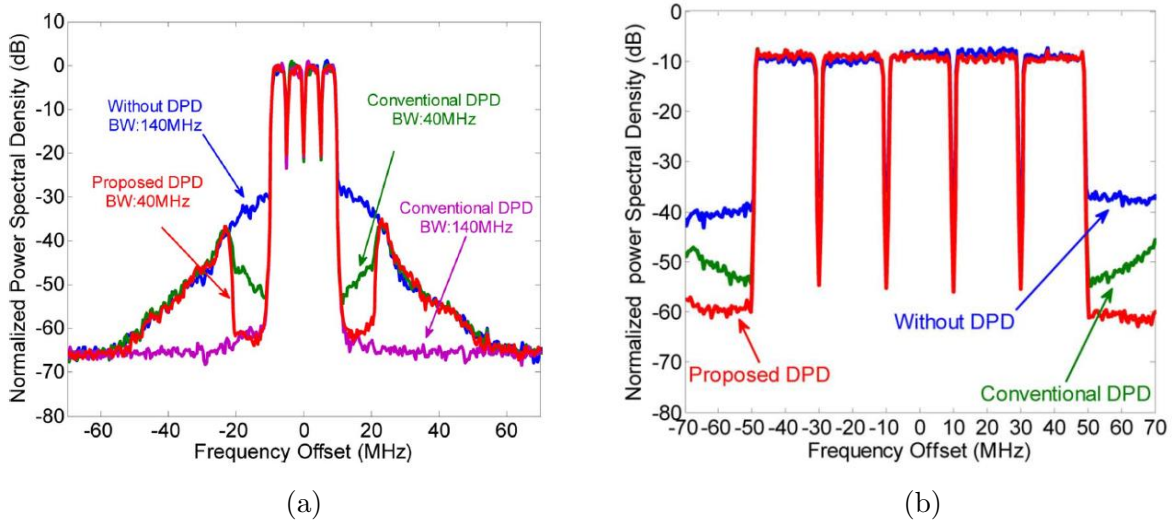
Figure 2.10: Band-limited DPD meassurments results (a) 10MHz (b) 100MHz [32].

DPD purposes, this linear operation was chosen to be a bandpass filter L with a bandwidth of the TOR bandwidth, or the bandpass filter bandwidth used. Based on that, the coefficient error values found form the equation with filtered bases should be the same as the values of the conventional DPD (if filtering was not included). As a result, these values can be used for wide-band signal linearization, because the coefficients are identical to the ones found in conventional DPD.

$$\text{Coefficients estimation: } \tilde{e}(n) = \sum_{q=1}^{Q} \Delta a_q \ \psi_q(\tilde{x}(n)) \Rightarrow L\{\tilde{e}(n)\} = \sum_{q=1}^{Q} \Delta a_q \ L\{\psi_q(\tilde{x}(n))\} \tag{2.12}$$

$$\text{DPD application (no change): } \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ \psi_q(\tilde{x}(n)) \Rightarrow \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ \psi_q(\tilde{x}(n)) \tag{2.13}$$

The main drawback of this approach is the following. Knowing that DPD is a behavioral modeling problem, it is fairly known that behavioral models have very poor extrapolation capabilities, which is one form of over-fitting. By over-fitting here I mean the estimated model will only perform well within the range of frequencies that was used during training, which is not the case in this approach. That is why in this work a new technique was investigated to reduce this degradation. These techniques can be understood as ways of regularization. Equivalently, it can be also understood as a method to reduce the bias during estimation for only reducing the measured band error. The author used two methods to reduce this bias: first method was to limit the number of basis in the DPD model using rank reduction techniques. Braithwaite used a standard method based on choosing the dominant eigenvalues [18][20][12]. The second method, which will be explained in the next paragraph, was to use a notch filter to reduce the in-band signal power.

In this notch filter technique, the in-band signal power was reduced using a notch filter bandstop filter. This extra filtering operation was applied in the digital domain, and only applied to the signal during coefficients estimation. This notch filter will reduce the amplitude of the modulation signal used during training. Without this filter, the optimization algorithm will try to reduce the in-band error while giving the out-of-band errors smaller weight. This occurs because in a typical DPD scenario, there is at least 20dB difference between the in and out-of-band errors. It was also suggested in this work that this notch filter technique can be also used to mitigate the in-band errors due to TOR impairments. This was achieved because the estimator now will not be heavily affected by any in-band linear distortion caused by the frequency response of the TOR.

It worth mentioning that in this work measurements, when the author wanted to demonstrate how a narrow band receiver can be used without affecting the linearization capabili-
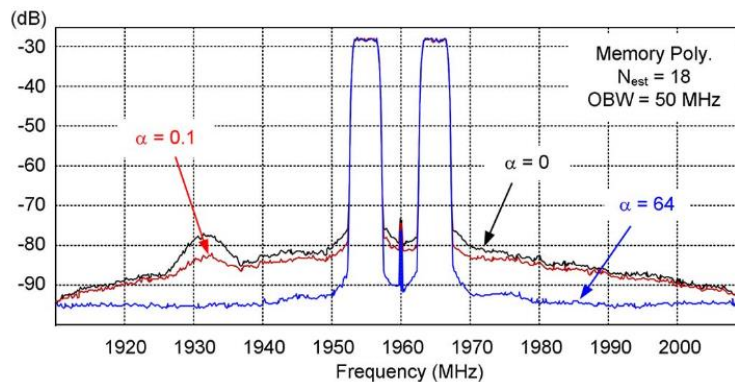
Figure 2.11: Open Loop wide band DPD [10].

ties, a wide band TOR was used. After measuring the signal with this wide band receiver, the IF bandpass filtering was introduced in digital domain. The filtered signal is then used in DPD training. The author showed linearization results of a dual band WCDM signal with a modulation bandwidth of 15MHz using an observation receiver of 50MHz of bandwidth, Fig. 2.11. It can be seen from these results that within the 50MHz of observation bandwidth, most of the nonlinearity is already included, hence, these results do not show a real demonstration of the proposed sub-sampling. For these reasons, the DPD extrapolation capabilities is questionable.

Indirect Learning Wide-band DPD [24]

In this work [24], Farouk tried to linearize the wide-band PA output signal without capturing the full bandwidth. He proposed a technique to achieve good linearization performance while capturing a bandwidth which is 2 - 2.5 times the input signal bandwidth. This is still considered a reduction in the ADC sampling speed from 5 times to only 2.5.

Farouk's approach can be considered similar to Braithwaite. Both of them applied filtering in the receiver path, without applying any filtering in the DPD application path. The main difference is Braithwaite filtered the PA signal and the basis, while farouk filtered the PA signal only without any basis filtering, Equation 2.14, 2.15; filtering here can be due to the addition of a bandpass filter, or due to having a narrow band TOR. Another less important difference is that Braithwaite used direct learning, while Farouk used indirect learning.

$$\text{Coefficients estimation: } \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ \psi_q(\tilde{y}(n)) \Rightarrow \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ \psi_q(\tilde{y}'(n)) \qquad (2.14)$$

$$\text{DPD application (no change): } \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ \psi_q(\tilde{x}(n)) \Rightarrow \tilde{u}(n) = \sum_{q=1}^{Q} a_q \ L\{\psi_q(\tilde{x}(n))\}$$
$$(2.15)$$

It was clear from my comments before that these kind of approaches has a huge tendency towards over-fitting; authors are trying to linearize a wide-band signal while providing the the estimator a band limited version of the signal. Since the estimator does not have any information about the filtered bands, it is expected DPD *will not* linearize these bands. Farouk and Braithwaite both used different techniques to increase the extrapolation capabilities of their models. Specifically, Farouk stated two requirements for this approach to work:

- Model order should be chosen carefully. If a complex model is used in DPD modeling, it will fail to linearize the out-of band components, however, excellent in-band linearization capabilities are often achievable. This is a very clear sign of over-fitting.

- Increasing the receiver bandwidth. Farouk could not reduce the observation bandwidth below 2 -2.5 times the input signal bandwidth. Although this might be better that the conventional DPD, it still a limitation on the TOR ADC sampling frequency.

Using these two techniques, Farouk could linearize a multi-carrier modulated signal of 60MHz using an TOR with an observation bandwidth of 150MHz, Fig. 2.12. It worth mentioning that in his measurements, signal filtering was applied in digital domain.

From reviewing the three previous approaches, it was evident non of them could achieve similar results to the conventional DPD. Chao *et al.* only linearized a specific band; then they relied on a bandpass filter that it hard to implement. Moreover, Braithwaite and Farouk both applied signal filtering while expecting the DPD will extrapolate. It was shown from their results they could not achieve substantial reduction on the ADC sampling speed. Next section will review the other family of approaches. It will be clear that using aliased signal for DPD training is, in general, better than filtered signals.
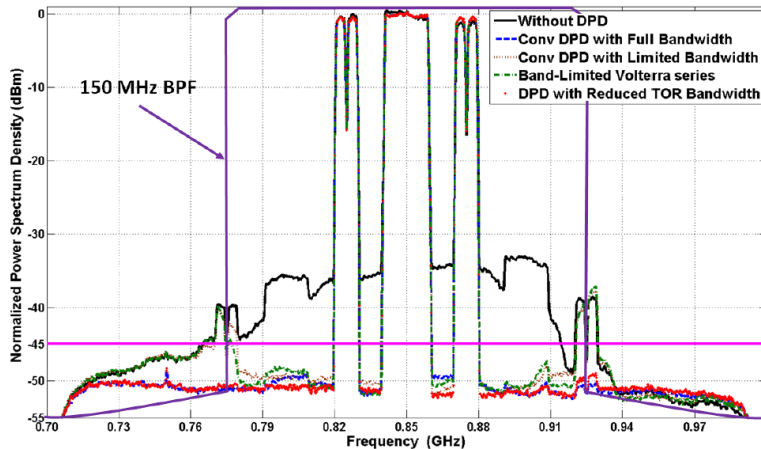
Figure 2.12: Indirect learning DPD linearlzation [24].

2.3.2 Full-rate Signal Restoration

In this second group of approaches to reduce the ADC sampling speed, no modification on the DPD algorithm were proposed. Instead, wide-band signal restoration techniques were devised. In [30], authors tried to train a PA forward model first using the sub-rated aliased data. This model is used to generate a full rate data that will be then used for DPD training. From the next discussion it will be shown that to train a PA model, no need to have access to the full rate data, however, subrated data can be enough. In this approach, in contrast to the previous ones, the signal bandwidth was reduced by reducing the signal sampling speed, and not by filtering some of the signal components. Reducing the signal sampling speed means the signal is now aliased. In aliasing, unlike signal filtering, the high frequency components can be still measured, but in an aliased form. This means some of the signal information at these high frequencies can be still obtained and are not completely lost like the filtering case. Both approaches (aliasing and filtering) will reduce the ADC sampling speed, but in a different way.

In [21] authors proposed a simple iterative method to reconstruct the full rate signal using an aliased version of the signal. Their method also required digital signal processing overhead. Besides that, for practical implementation the authors mentioned another signal transmitter hardware is required, which reduces the practicability of their approach.

Finally In [22], the authors proposed an spectrum extrapolation method. This method uses a filtered signal to predict the signal spectrum at the filtered frequencies. The extrap-

olated signal can be then used for normal DPD training. Each one of those approaches will be discussed to show more of its implementation details.

It will be evident from the next review that aliasing based solutions outperform filtered based ones. For that particular reason, spectrum extrapolation technique did not achieve promising results. On the other hand, authors in [30] showed huge gain in sampling speed reduction. However, they did not show in gain for a real measurements system that uses low sampling speed hardware. In contrast, authors in [21] showed a real implementation of their approach; they could linearize a PA with ADCs that run at the intended reduced sampling speed.

Full-Rate Signal Restoration through PA Forward Modeling [30]

In this work [30] Zonghao *et al.* tackled the under-sampled DPD problem in two steps; they had first to devise a forward model for the PA; then they trained the DPD engine, Fig. 2.13. For the first part they only used under-sampled output data for the forward model training. This was possible since they still had access to the full rate PA input data. The idea they used was elegant and fairly common, which will be explained later. Once they had the forward model of the PA, they could use it to generate a full rate output signal, and use this signal in a normal indirect learning scenario.

For PA forward modeling using under-sampled data, the output $\tilde{y}(n)$ (baseband equivalent of $y(t)$) is expressed as weighted sum of nonlinear basis:

$$\tilde{y}(n) = \sum_{q=1}^{Q} a_q \, \psi_q(\tilde{u}(n)) \tag{2.16}$$

This equation has Q unknowns: the forward model coefficients a_i. To find these coefficients, a system of linear equations is constructed using measurements data points. Usually, the number of points is larger than the number of unknowns; hence, the system is overdetermined. A standard way to write these equations using matrix notations can be as follows:

$$
\begin{bmatrix}
\tilde{y}(n) \\
\tilde{y}(n+1) \\
\tilde{y}(n+2) \\
\vdots \\
\tilde{y}(n+N)
\end{bmatrix}
=
\begin{bmatrix}
\psi_1(\tilde{\mathbf{U}}(\mathbf{n})) & \psi_2(\tilde{\mathbf{U}}(\mathbf{n})) & \ldots & \psi_Q(\tilde{\mathbf{U}}(\mathbf{n})) \\
\psi_1(\tilde{\mathbf{U}}(\mathbf{n+1})) & \psi_2(\tilde{\mathbf{U}}(\mathbf{n+1})) & \ldots & \psi_Q(\tilde{\mathbf{U}}(\mathbf{n+1})) \\
\psi_1(\tilde{\mathbf{U}}(\mathbf{n+2})) & \psi_2(\tilde{\mathbf{U}}(\mathbf{n+2})) & \ldots & \psi_Q(\tilde{\mathbf{U}}(\mathbf{n+2})) \\
\vdots & \vdots & \ddots & \vdots \\
\psi_1(\tilde{\mathbf{U}}(\mathbf{n+N})) & \psi_2(\tilde{\mathbf{U}}(\mathbf{n+N})) & \ldots & \psi_Q(\tilde{\mathbf{U}}(\mathbf{n+N}))
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2 \\
a_3 \\
\vdots \\
a_Q
\end{bmatrix}
\tag{2.17}
$$

$$\tilde{\mathbf{U}}(\mathbf{n}) = \begin{bmatrix} \tilde{u}(n) & \tilde{u}(n-1) & \ldots & \tilde{u}(n-l) \end{bmatrix}; \, l : \text{ memory depth} \tag{2.18}$$

$$\mathbf{Y}_{Nx1} = \mathbf{\Psi}_{NxQ} \ \mathbf{a}_{Qx1} \tag{2.19}$$

Lease square algorithms can find the set of coefficients that minimize the square error between the predicted and measured PA output. For example, LSE can find the best set of coefficients by:

$$\mathbf{a} = (\mathbf{\Psi^H \Psi})^{-1} \mathbf{\Psi^H Y} \tag{2.20}$$

The idea behind using under-sampled data is like the following: since the system of linear equation is overdetermined, some of the rows in 2.17 (data points) can be discarded without affecting the final coefficients solution. In particular, if we discarded each other row, the final solution will not change. Since discarding each other row is equivalent to ignoring each other measured point; it is also similar to say the sampling rate of the output signal is reduced by factor of 2. It worth mentioning that the input signal $\tilde{u}(n)$ is still sampled at the full rate. This is necessary since the PA forward model requires $\tilde{u}(n-1)$, $\tilde{u}(n-2)$, ... etc. In other words, even though we discarded some of the rows, the content of the kept rows should not change.

The main difference between this approach and the previous ones is that the authors are using an aliased signal for forward model training. Although the down-sampled signal does not has all the wide-band signal information, however, aliasing has some information of the wide-band signal. This given an impression that aliasing based solution are more likely to achieve better results than filtered bases solutions. And indeed this was the case, in this work, authors could linearize a 40MHz LTE signal using 200, 20, 5 and 2.5Msps. 2.5Msps is equivalent to reducing the sampling speed by a factor of 80 than conventional DPD, Fig 2.14.

For their measurements, authors had to use measurement setup that runs at the full sampling rate; then they down-sampled the signals digitally and use it for training. Although this was presented as a proof of concept for the under-sampling DPD, it did not show a real implementation using a slow ADCs. It also worth mentioning that authors proposed an iterative time alignment algorithm for under-sampled data alignment. It was also mentioned, however, that this technique is very processing intensive. As a result, the sup-sampling speed is limited to 20MHz, which is equivalent to sub-rate of 10.

Under-Sampling Restoration DPD [21]

Similar to the previous approach [30], the authors in this work proposed an unconventional way to restore the full-rate PA output signal from an under-sampled aliased signal. To do so, authors tried to model the aliasing distortion effect as a simple linear operation. Once
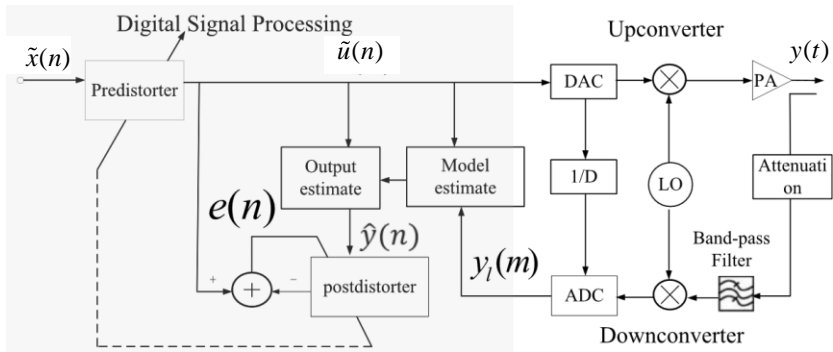
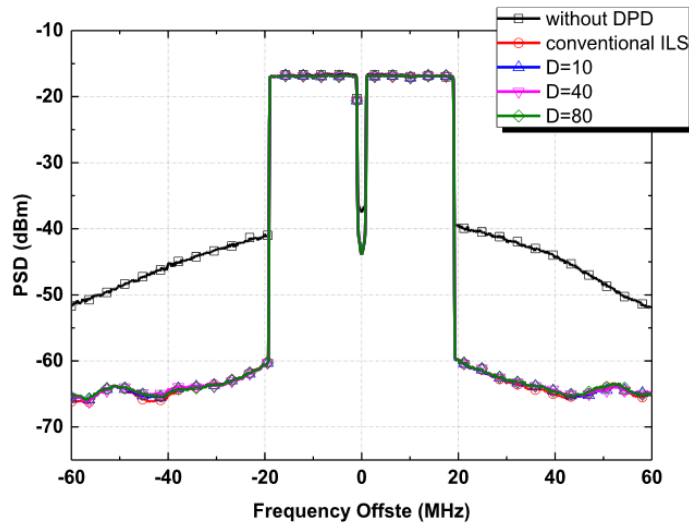Figure 2.13: Proposed DPD architecture [30].

Figure 2.14: 40MHz LTE linearization results using subsampling by D factor [30].

this operation is modeled, they could apply the inverse to the aliased data to restore the full rate signal, that can be then used for DPD training.

At the first iteration, the PA was bypassed, Fig. 2.15. The DAC output signal $y_a(t)$ was measured with an under-sampled ADC; an ADC with a sampling rate less than the Nyquist rate. This under-sampled signal was then up-sampled to the original signal sampling rate. This step is necessary to allow the comparison between the two signals. A linear function was devised to relate the two signals, $y_f(n)$ and $y_a(n)$, together 2.21. This G transfer function is a sequence of complex numbers with the same length as the signal itself. It worth mentioning that up till now $y_a(n)$ is the same as $x(n)$. Since in a more practical situation $y_a(n)$ will be distorted version of $x(n)$, hence, the G function will need to be re-found as $y_a(n)$ changes.

$$G(n) = \frac{y_f(n)}{y_a(n)} \qquad (2.21)$$

Next step is to connect the PA back to the loop and measure the PA $y_a(t)$. Although we do not have access the the full-rate PA output data, but using the previously found G and the under-sampled measured signal $y_f(t)$, the full rate signal can be found using the following:

$$y_a(n) = \frac{y_f(n)}{G(n)} \qquad (2.22)$$

Once the full rate data is found, the DPD engine can be now trained using indirect learning, and a new pre-distorted signal $u(n)$ can be found. Next step is to find a better estimate for the receiver transfer function, G. To do so, first, PA forward model is trained using the most recent measurements. Then, having this model and the new pre-distorted signal $u(n)$, the full rate output can be predicted. This output is fed to the TOR by bypassing the PA again. Having this full rate predicted output and the corresponding under-sampled measured signal, a new G transfer function can be devised using 2.21.

Using these techniques, the authors showed linearization results for 10 and 20MHz LTE signals with a reduction in sampling speed by 3-5 factor. In their measurements they used hardware that runs at the reduced sampling speed. These results shows a *real* implementation of *sub-sampling* DPD. The main disadvantage of this technique is the need for another signal generator in the transmitter, Fig. 2.15. Another disadvantage is the need to re-train the TOR transfer function for each DPD iteration. This will add huge processing complexity to the total algorithm.
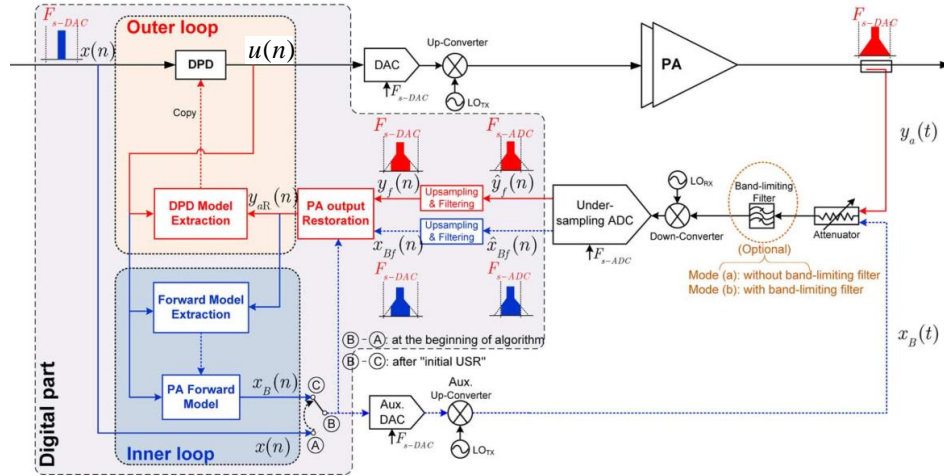
Figure 2.15: Block digtam of USR-DPD [30].

<center>(a)</center>

<center>(b)</center>
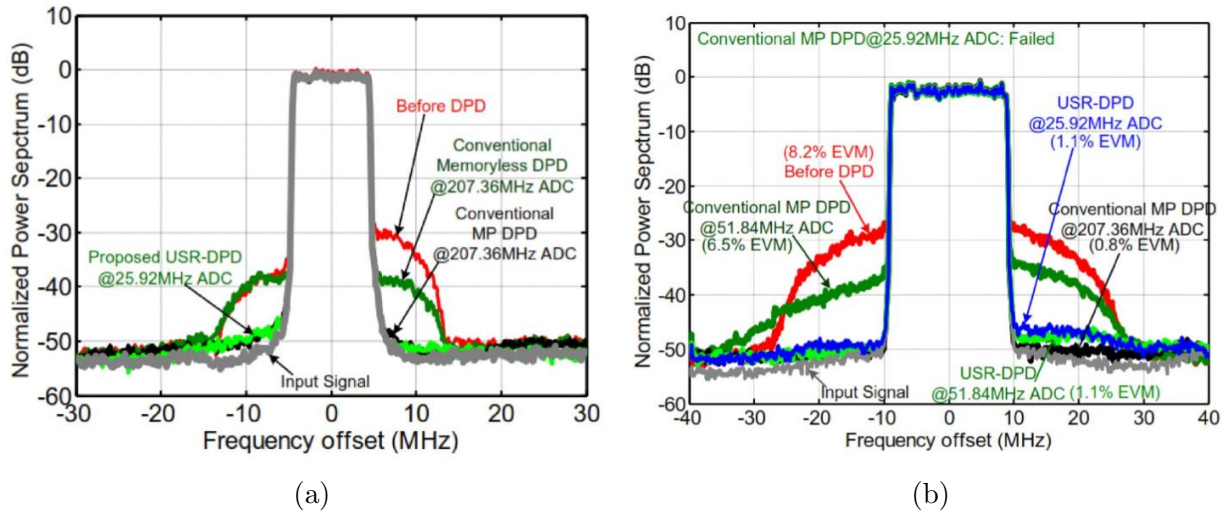
Figure 2.16: Linearization results (a)10MHz (b) 20MHz [21].
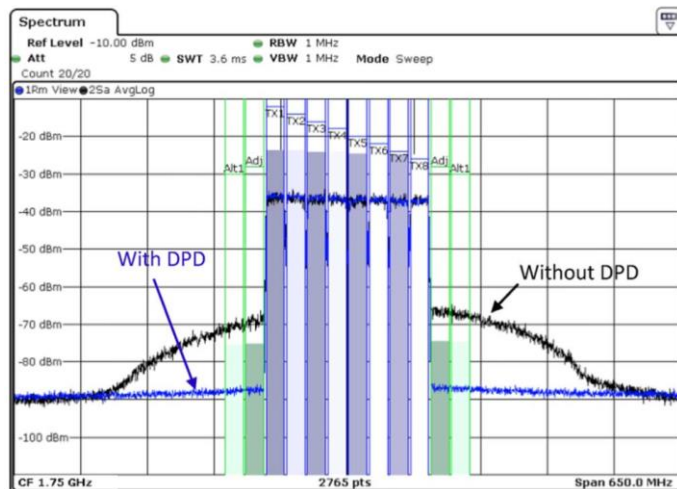
Figure 2.17: Linearizion results for 160MHz signal using spectrum extrapolation [22].

Spectrum Extrapolation [22]

In [22], the authors tried to construct the wide-band signal before DPD training. In particular, they used a spectrum extrapolation technique for signal reconstruction, and direct learning for DPD training.

In this approach, to reduce the ADC sampling rate, the signal spectrum was first filtered to a bandwidth that is less than the ADC Nyquist sampling rate. This means no information is available for the filtered frequency components. Such approaches usually over-fits and simple DPD models must be used. In this work for example, authors used a MP model with 6 coefficients to linearize a 140MHz signal.

For measurements, authors introduced down-sampling in digital domain; all measurements were done using hardware that runs at full sampling rate. The data is then down-sampled in digital domain and used for DPD training. In this work they could achieve notable reduction in sampling rate without degrading the linearization capabilities. For example, they linearized 160MHz LTE signal while measuring 160MHz of PA output signal with 200Msps ADC Fig.

31

2.3.3 Summary of Literature Review

The final linearization results are summarized in table 2.1. Theses results can be interpreted in the following way:

- It is evident form [10], [24], and [22] that filtered based solutions demonstrated poor ADC soampling rate reduction factor. The reason behind that was explained in the previous sections; DPD bahavioural models exhibit poor extrapolation capabilities, and they should be only used where they were trained.

- Although both [21] and [30] used aliasing based approaches, there is a huge discrepancy between each approach gain in ADC sampling speed reduction. This can be due to:

  - [30] used measurements hardware that runs at the full rate, and then introduced down-sampling in the digital domain. In contrary, [21] used hardware that runs at the reduced sampling speed. By doing do, distortion due to operating the circuits near DC can be avoided.

  - [21] approach in restoring the full rate signal laked any theoretical support. Modeling aliasing using simple linear operation is questionable. On the other hand, [30] seemed to be theoretically solid. In a sense authors provided detailed analysis on the propagation of modeling error from forward modeling to DPD training. From this analysis some insight on the DPD training performance were gained that guided the validation measurement process.

- From the previous point, it is clear now that implementing sub-sampling DPD using low speed ADCs is very challenging problem. This gives the impression that sub-sampling DPD is a measurements problem rather than a conventional modeling one. For this particular reason, big attention should be paid to the measurement procedure and calibration techniques.

All this review can be used to guide this thesis work. For the before mentioned reasons a novel sub-sampling was applied that tackles the sub-sampling DPD from a practical point of view. At the beginning, an aliased signal will be used for DPD training in a direct learning scenario. A similar idea to one used in [30] will be exploited to directly train the DPD model without having to do any forward modeling. The next and important step is to build this algorithm using a hardware running at the reduced sampling speed. This setup can be used as a *real* demonstration of the proposed algorithm. To do so, a

calibration routine was devised to compensate for receiver and transmitter linear distortion. At the end, good linearization results are achieved for a 160, 200 and 320MHz signal using full-rate, sub-rate of 2 and 4.

2.4 Conclusion

In this chapter, the PA non-linearity problem was introduced. The effect of this nonlinearity of efficiency, EVM and ACPR distortion, was also discussed. To mitigate these non-linear effect, the DPD solution were devised. Although DPD is a widely adopted and mature solution, some DPD challenges are showing up every day as communication standards are evolving. In particular, as the signal bandwidth is increasing, new DPD models should be devise, better TOR should be designed, and lower power hardware circuits should be used for the TOR. Out of all these circuits, the ADC sampling is the most important factor that dictates where the majority TOR power is dissipated.

To reduce the ADC power, and the sampling speed accordingly, multiple solutions were proposed in the literature and reviewed in this chapter. Out of this review it was evident that practical implementation of sub-sampling DPD was missing in the literature. For that reason, a new DPD algorithm is proposed with some emphasis on the practical implementation aspects: using hardware running at reduced sampling speed while paying big attention to TOR calibration.

Table 2.1: Literature Review Measurements Summary.

Sub-sampling approach	BW [MHz]	F_s [Msps]	$F_s/$ BW [Msps/MHz]	Band-limiting approach Filtered vs Aliased	Down-sampling technique Real vs Digital	DPD estimator Direct vs Indirect
Band-limited DPD [32]	100	140	-	-	-	Indirect
Direct Leaning Wide Band DPD [10]	15	50	3⅓	Filtered	Digital	Direct
Indirect Leaning Wide Band DPD [24]	60	150	2.5	Filtered	Digital	Indirect
Signal Restoration through PA forward Modeling [30]	40	2.5	¹⁄₁₆	**Aliased**	Digital	Indirect
Under-Sampling Restoration DPD [21]	20	25	1.25	**Aliased**	**True**	Indirect
Spectrum Extrapolation [22]	**160**	200	1.25	Filtered	Digital	Direct

Chapter 3

Proposed Sub-sampling DPD

The previous literature review has established that sub-sampling methods that are based on aliasing achieve, in general, better linearization results than those based on filtering methods. Unfortunately, the two sub-sampling DPD approaches that used aliased PA output signal required heavy digital signal processing overhead, either for PA forward modeling [30] or for feedback path alias modeling [21]. Furthermore, and aside from sub-sampling algorithm computation complexity of the sub-sampling algorithm, neither of the two sub-sampling DPD approaches using aliased PA output signal research presented any implementation using hardware that runs at the reduced sampling speed. Collectively, out of the six sub-sampling attempts reviewed, only one used sub-sampling hardware [21].

To solve the computation complexity for the aliased-based methods, Hai *et al* [19] proposed a sub-sampling DPD based on direct learning. Hai's method does not require any extra processing overhead since the DPD training is done using the sub-sampled data directly. Consequently, Hai's approach has been determined to be suitable for this thesis.

Regarding the hardware challenge, while Hai's approach provides a good theory for sub-sampling DPD, it does not show measurement results using hardware running at the reduced sampling speed. Hence, to demonstrate the practicability of Hai's approach, and sub-sampling DPD in general, this chapter provides a discussion regarding the practical implementation of this approach. It will be shown that the sub-sampling algorithms can be implemented if the TOR IF circuitry can work very close to DC.

This chapter is organized as follows: the first section reviews the theoretical details of Hai's approach. Next section first shows the measurements setup used in Hai's approach. Then introduces the method to do similar measurements using sub-rate hardware. Finally, we discuss the receiver calibration problem using sub-rate hardware.

3.1 Digital Pre-Distortion Function Synthesis Using Under-Sampled Feedback Signal [19]

This section reviews the theory of the work that is adopted by this thesis. Hai *et al* [19] devised a sub-sampling DPD that did not require any extra digital processing power. Their work was based on direct learning algorithm where the DPD coefficient has to be updated iteratively. In the direct learning algorithm, the error signal is modeled using the nonlinear bases, where error here is the difference between the PA output and the required linear output. Since Hai's work is about under-sampling DPD, they have to modify the DPD formulation to be suitable for aliased data. As a result, they modeled the *aliased* under-sampled error signal using the same non-linear bases. Finally, they achieved similar linearization results to the conventional DPD system for up to 80MHz of modulation bandwidth.

3.1.1 Direct Learning

Since Hai's work is based on direct learning, it is helpful to first review direct learning to show how it is suitable for under-sampling applications. For direct learning [11], the DPD coefficients are updated iteratively. At each iteration, the error signal is modeled as a weighted sum of nonlinear bases, where the weights are the error in coefficients (3.1)[1]. This error equation has Q unknowns, the delta coefficients. To solve for these unknowns, Q linearly independent equations are needed. In a typical DPD scenario, N data points are collected where $N >> Q$. Since the number of equations is larger than the number of unknowns, the system of N linear equations is overdetermined (3.2). To find the set of delta coefficients that minimize the least square error between the measured and real error (3.6), LSE algorithms (3.5) can be used. Once these delta coefficients are found, they are used to update the DPD coefficients (3.7). Finally, after the coefficients are updated, the pre-distorted signal can be found and applied to the PA input (3.8).

$$\hat{e}(n) = \tilde{y}(n) - \tilde{x}(n) = \sum_{q=1}^{Q} \Delta a_q \ \psi_q(\tilde{x}(n)) \tag{3.1}$$

---

[1]Although modeling the error signal using the same DPD coefficients is counterintuitive, a formal proof of this equation can be found here: [33, 15].

$$\begin{bmatrix} \tilde{e}(n) \\ \tilde{e}(n+1) \\ \tilde{e}(n+2) \\ \vdots \\ \tilde{e}(n+N) \end{bmatrix} = \begin{bmatrix} \psi_1(\mathbf{X(n)}) & \psi_2(\mathbf{X(n)}) & \ldots & \psi_Q(\mathbf{X(n)}) \\ \psi_1(\mathbf{X(n+1)}) & \psi_2(\mathbf{X(n+1)}) & \ldots & \psi_Q(\mathbf{X(n+1)}) \\ \psi_1(\mathbf{X(n+2)}) & \psi_2(\mathbf{X(n+2)}) & \ldots & \psi_Q(\mathbf{X(n+2)}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{X(n+N)}) & \psi_2(\mathbf{X(n+N)}) & \ldots & \psi_Q(\mathbf{X(n+N)}) \end{bmatrix} \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \\ \Delta a_3 \\ \vdots \\ \Delta a_Q \end{bmatrix} \quad (3.2)$$

$$\mathbf{X(n)} = \begin{bmatrix} \tilde{x}(n) & \tilde{x}(n-1) & \ldots & \tilde{x}(n-l) \end{bmatrix} ; \; l: \quad \text{memory depth} \quad (3.3)$$

$$\mathbf{E}_{Nx1} = \mathbf{\Psi}_{NxQ} \; \mathbf{\Delta a}_{Qx1} \quad (3.4)$$

$$\mathbf{\Delta a} = (\mathbf{\Psi^H \Psi})^{-1} \mathbf{\Psi^H E} \quad (3.5)$$

$$LSE = \min_{\Delta a} \|\hat{e} - \tilde{e}\|^2 = \min_{\Delta a} \left\| \sum_{q=1}^{Q} \Delta a_q \; \psi_q(\tilde{u}(n)) - \tilde{e} \right\|^2 \quad (3.6)$$

$$a_q^t = a_q^{t-1} - \eta \; \Delta a_q \quad (3.7)$$

$$\tilde{u}(n) = \sum_{q=1}^{Q} \Delta a_q \cdot \psi_q(\tilde{x}(n)) \quad (3.8)$$

Furthermore, for indirect learning to achieve good results, some constraints regarding the input and output signals sampling frequencies must be met. In particular both the input, $\tilde{x}$, and output, $\tilde{y}$, should be sampled at a sampling frequency which—according to the conventional rule of thumb—is five times the input modulation bandwidth.

Input and output sampling frequencies govern the minimum time differences between different entry of the direct learning data matrix (3.2). In this matrix, each row corresponds to a different measurement data point, where the time difference between each successive data point is dedicated by the *output* signal sampling frequency. For example, the time difference between the $\tilde{e}(n)$ and $\tilde{e}(n+1)$ is T_{out}, where T_{out} is $1/F_{Sout}$, F_{Sout} is the output $\tilde{y}$ sampling frequency. On the other hand, the bases of each row dedicates the inputs, $\tilde{x}$ and $\tilde{u}$, sampling frequency. In a conventional DPD system, the sampling frequency of the PA input, $\tilde{u}$, is also five times the modulation bandwidth of the original input $\tilde{x}$. This corresponds to a minimum time difference between the bases of T_{in}, where T_{in} is $1/F_{Sin}$. For example, to model the linear memory of a PA, the bases used are usually in the form of: $\tilde{x}(n)$, $\tilde{x}(n-1)$, ... $\tilde{x}(n-L)$. The time difference between each successive bases, between $\tilde{x}(n)$ and $\tilde{x}(n-1)$ for example, is T_{in}.

In summary, in a conventional DPD system, both the PA input and output are sampled at a sampling frequency which is five times the original signal bandwidth. These two sampling frequencies show up in the direct learning matrix (3.2) in two different ways:

- Input sampling shows up as the minimum time difference between two successive bases of the same row.

- Output sampling shows up as the minimum time difference between two successive outputs, which is equivalent to two successive rows.

Since direct learning is introduced and the the sampling requirements are discussed, the next sub-section explains the theory of the DPD approach used in Hai's work. We adopt the same theory for the work of this thesis.

3.1.2   Sub-Sampling Direct Learning Digital Pre-distortion [19]

Hai used the same direct learning algorithm explained before. In addition, and to introduce sub-sampling, he has to make a small modification to (3.2). For sub-sampling, he only has access to the sub-sampled output data. For example, if output data sub-sampled by a factor of 2, this means, compared to the conventional DPD case, we discarded every other output data point (3.9). Since we have an overdetermined system of equation, having a system with some of the equations discarded is still solvable. In other words, the reduced system of equation (3.9) should yield the same solution as the one with no data points discarded (3.2).

Since reducing the size of the data set used in coefficients estimation might affect the performance of the optimization algorithm, that is discarding half of the data points might deteriorate the DPD overall performance. To solve this problem, longer frame time can be used to construct the matrix in (3.9).

$$
\begin{bmatrix}
\tilde{e}(n) \\
\tilde{e}(n+2) \\
\tilde{e}(n+4) \\
\vdots \\
\tilde{e}(n+2N)
\end{bmatrix}
=
\begin{bmatrix}
\psi_1(\mathbf{X(n)}) & \psi_2(\mathbf{X(n)}) & \ldots & \psi_Q(\mathbf{X(n)}) \\
\psi_1(\mathbf{X(n+2)}) & \psi_2(\mathbf{X(n+2)}) & \ldots & \psi_Q(\mathbf{X(n+2)}) \\
\psi_1(\mathbf{X(n+4)}) & \psi_2(\mathbf{X(n+4)}) & \ldots & \psi_Q(\mathbf{X(n+4)}) \\
\vdots & \vdots & \ddots & \vdots \\
\psi_1(\mathbf{X(n+2N)}) & \psi_2(\mathbf{X(n+2N)}) & \ldots & \psi_Q(\mathbf{X(n+2N)})
\end{bmatrix}
\begin{bmatrix}
\Delta a_1 \\
\Delta a_2 \\
\Delta a_3 \\
\vdots \\
\Delta a_Q
\end{bmatrix}
\quad (3.9)
$$

Worth mentioning is that the input signals, $\tilde{x}$ and $\tilde{u}$, are still sampled at the conventional sampling rate. This shows up in (3.9) as the content of the retained rows should be identical to the ones in (3.2). As a result, indirect learning is not a suitable method be used here. The reason for that can be seen from the indirect learning formulation. For indirect

38

learning, the PA input $\tilde{u}$ is modeled as weighted sum of non-linear bases, $\psi(\mathbf{Y(n)})$. This can be shown in (3.10) in a matrix form. If the output sampling frequency changed, the time difference between the output sample changes. knowing the content of each row is a function of the output signal and output signal sampling, changing the output sampling changes the content of each rows. Hence, indirect learning can not be used here.

$$
\begin{bmatrix} \tilde{u}(n) \\ \tilde{u}(n+1) \\ \tilde{u}(n+2) \\ \vdots \\ \tilde{u}(n+N) \end{bmatrix} = \begin{bmatrix} \psi_1(\tilde{\mathbf{Y}}(\mathbf{n})) & \psi_2(\tilde{\mathbf{Y}}(\mathbf{n})) & \dots & \psi_Q(\tilde{\mathbf{Y}}(\mathbf{n})) \\ \psi_1(\tilde{\mathbf{Y}}(\mathbf{n+1})) & \psi_2(\tilde{\mathbf{Y}}(\mathbf{n+1})) & \dots & \psi_Q(\tilde{\mathbf{Y}}(\mathbf{n+1})) \\ \psi_1(\tilde{\mathbf{Y}}(\mathbf{n+2})) & \psi_2(\tilde{\mathbf{Y}}(\mathbf{n+2})) & \dots & \psi_Q(\tilde{\mathbf{Y}}(\mathbf{n+2})) \\ \vdots & \vdots & \ddots & \vdots \\ i\psi_1(\tilde{\mathbf{Y}}(\mathbf{n+N})) & \psi_2(\tilde{\mathbf{Y}}(\mathbf{n+N})) & \dots & \psi_Q(\tilde{\mathbf{Y}}(\mathbf{n+N})) \end{bmatrix} \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \\ \Delta a_3 \\ \vdots \\ \Delta a_Q \end{bmatrix} \quad (3.10)
$$

3.1.3 Comments on Validation Measurements Results

In Hai's work [19], he showed linearization results for 20 and 80MHz, with sub-rate factor of 5, which is sampling rate of 20 and 80Msps respectively, Fig. 3.1a and Fig. 3.1b. For these measurements, as a proof of sub-sampling concept, he used hardware running at the full sampling speed. Then, he introduced sub-sampling on the digital domain.

Although these measurements showed validation results for sub-sampling DPD, it lacked the practical proof that sub-sampling can be implemented in a real scenario. This is the case because Hai did not demonstrate his algorithm can work using hardware that runs at the reduced sampling speed. This kind of demonstration, using sub-sampled hardware, shows the real benefits of using sub-sampling DPD algorithms. Besides, this demonstration highlights the practical challenges of implementing sub-sampling DPD; receiver calibration for example. The next section shows how the hardware setup should be adjusted to allow for sub-sampling DPD implementation.

3.2 Hardware Implementation of Sub-Sampling Digital Pre-distortion

This section shows the method to generate the sub-sampled signal using hardware running at the reduced sampling speed. First we review the conventional DPD TOR receiver with emphasis on receiver calibration. Then we discus the modifications on this conventional DPD if the hardware is only operating at the reduced sampling speed.
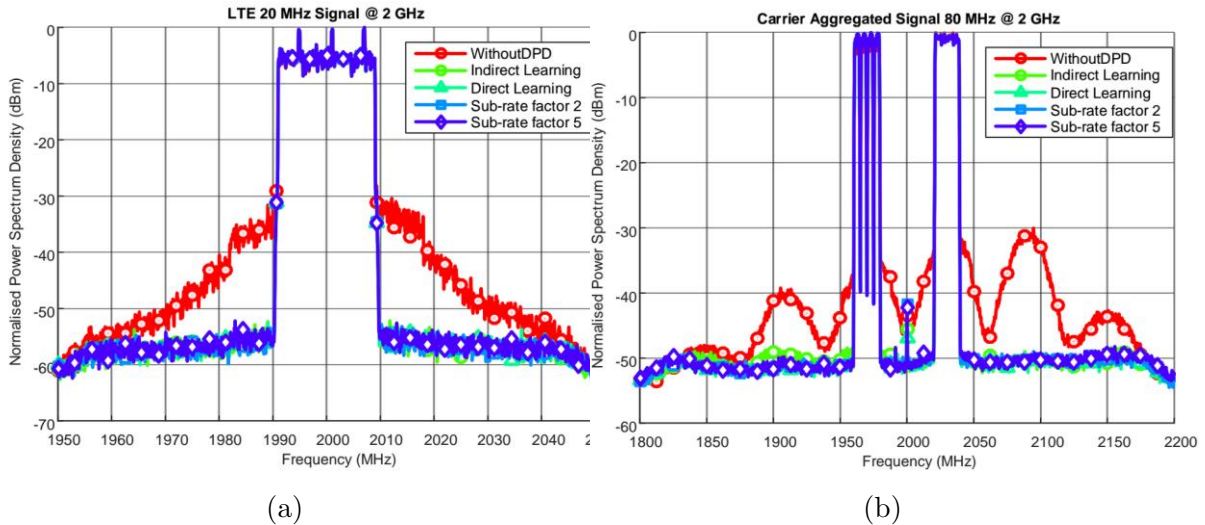
Figure 3.1: Linearization results (a) LTE 20MHz (b) Carrier aggregated 80MHz [19] .

3.2.1   Full-Rate Hardware for Sub-Rate Digital Pre-distortion

Most of the work presented in the literature did not show real implementation of sub-sampling DPD using hardware (receiver) running at the reduced sampling speed. They, instead, measured the PA output signal with a hardware that runs at the full sampling speed, and sub-sampling was later introduced in the digital domain. As a result, the PA output signal is captured and transformed to the digital domain without being distorted by aliasing or image overlap.

An example of a block diagram for a typical measurements setup (receiver) that uses a super heterodyne-like receiver is depicted in Fig. 3.2a. At those receivers, the down-conversion from RF to the baseband is done in two steps, where the intermediate step is in the IF stage. The first down-conversion step, from RF to IF, is done in the RF domain using real hardware. Then the data is captured with an ADC that has high sampling speed. This sampling speed should be proportional to the received signal bandwidth, $TLBW$, where $TLBW$ is the targeted linerization bandwidth. According to the rule of thumb, TLBW is equivalent to five times the input signal modulation bandwidth. Once the data are in the digital domain, the second down-conversion step, from IF to the baseband (zero frequency), is done using digital signal processing techniques. The TLBW affects the choice of F_{lo} and F_{sADC}[2] for the following reasons:

---

[2]In this thesis, F_S or f_s are used for *real* sampling speed. Besides, F_S^c or f_s^c are used for *complex*

- F_{IF}, and accordingly F_{lo}, is chosen high enough to avoid any distortion due to image overlap after down-conversion to IF frequency. In other words, $F_{IF} > TLBW/2$.

- $F_{s\ ADC}$, the ADC sampling speed, should be high enough to avoid any distortion due to aliasing. Hence, $F_{s\ ADC}/2 > F_{IF} + TLBW/2$

By relating this measurement setup to DPD, or in particular to the construction of the reduced bases matrix (3.9), the way this setup constructs the reduced bases matrix can be understood in two different, but equivalent, ways:

1. The easy way to understand this, was demonstrated before in section 3.1.2, is to first construct the full rate bases matrix (3.2) using the full-rate signal. Then, discard some of the rows of this matrix according to the intended sub-sampling ratio (3.9).

2. Equivalently, we could have generated the same reduced bases matrix (3.9) by using only the aliased signal in constructing the bases matrix. By aliasing signal we mean here the signal that is generated by the TOR after the digital down-sampling step. This method generates the reduced bases matrix, with the rows automatically discarded.

In contrast, the next approach, the one that generates only sub-rated data without any digital down-sampling, uses the sub-sampled data directly from the TOR ADC. Hence, this approach is equivalent to generating the reduced matrix directly (3.9). The next presented approach can be considered as a real implementation of a real sub-sampling DPD using a sub-sampling hardware.

3.2.2   Sub-Rate Hardware for Sub-Rate Digital Pre-distortion

The objective of this section is to show how to generate the sub-rated data from the measurements hardware without any digital down-sampling. The previous section established that constructing the bases matrix from the aliased data automatically generates the reduced version of the bases matrix (3.9). As a result, this thesis work is about generating a similar aliased signal using the sub-sampling hardware. In other words, if the hardware that runs at the reduced sampling speed generates an aliased signal that is identical to the aliased signal resulting from the digital down-sampling, we can construct the reduced bases

sampling frequencies, where $F_S^c = F_S/2$. Therefore, for full-rate $f_{sADC}/2 > TLBW$, equivalently $f_{sADC}^c > TLBW$.
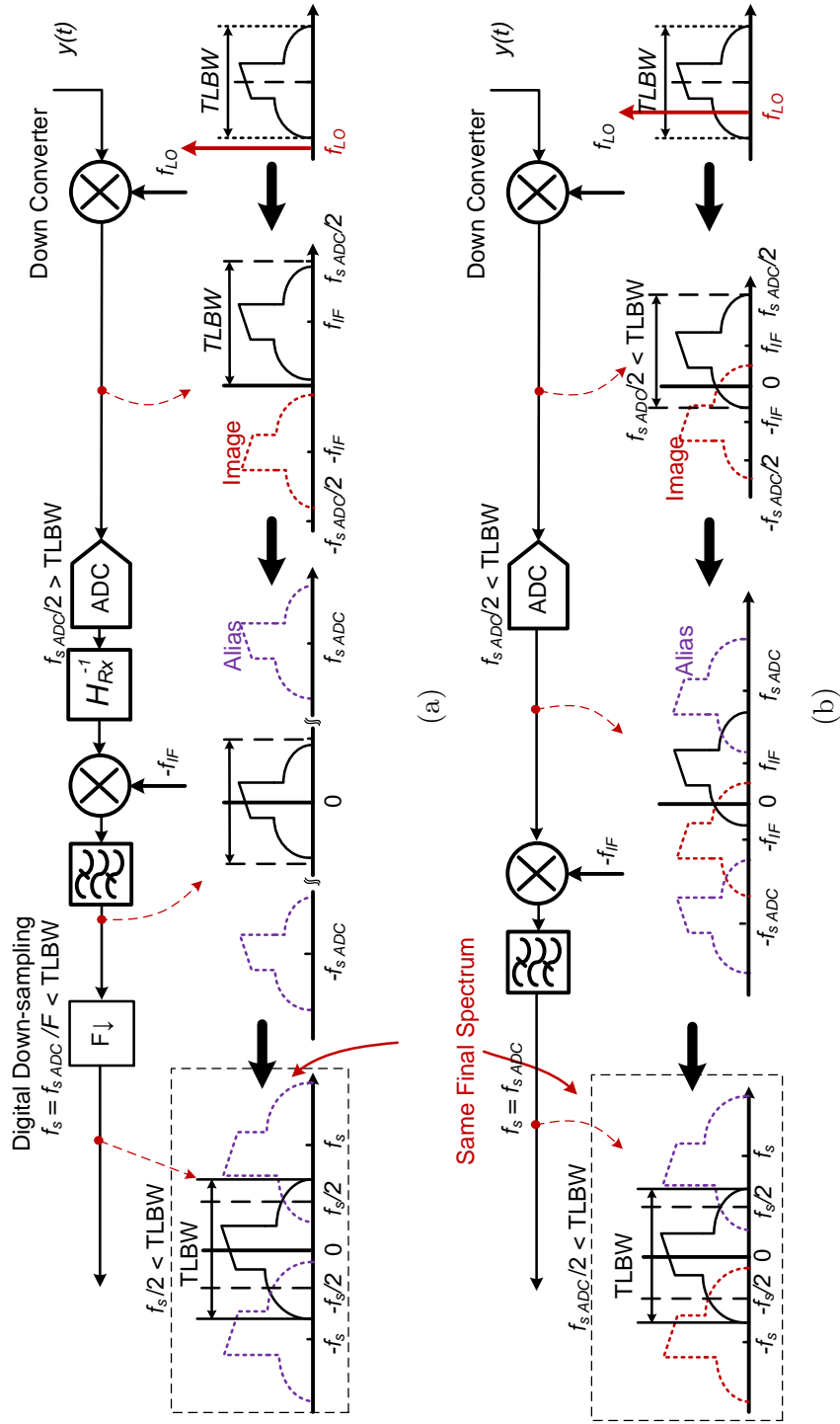
Figure 3.2: Sub-sampling DPD using: (a) Full-rate hardware. (b) Sub-rate hardware.

42

matrix automatically. This automatic construction is possible because the two aliased signals are the same. So, the link between the two approaches is to generate, using hardware that runs at the reduced sampling speed, an aliased signal that is identical to the aliased signal generated before using digital down-sampling.

In the newly proposed setup (Fig. 3.2b), the hardware ADC is running at the reduced sampling speed, where this ADC speed is less than the signal bandwidth, $FsADC/4 < F_{IF} + TLBW/2$. Having such sampling speed means aliasing distortion is inevitable. To generate the needed aliased spectrum, we chose the the new IF frequency to be exactly $FsADC/4$. By doing so, seen from Fig. 3.2b, the final spectrum is identical to the one generated digital down-sampling approach. It is noticed from this figure that the new distortion caused by image overlap at the left hand side of the IF signal spectrum is identical to the distortion caused by aliasing at the previous *digital down-sampling* method, Fig. 3.2a.

Since the new aliased signal is now identical to the one generated using full-rate hardware, we use it in normal DPD bases matrix construction. This bases matrix is exactly in the form of (3.9), and as a result, can be used directly for DPD coefficients estimation.

To generate this aliased signal using sub-sampling hardware, we unintentionally created another problem related to calibration. By comparing the two block diagrams in Fig. 3.2a and Fig. 3.2b, it can be seen that the receiver calibration block is missing for the sub-rate case. This occurs because identifying the TOR transfer function and applying receiver calibration in the sub-rate case is not a straight forward task. To understand how to perform these two actions, the next section first reviews how calibration is applied in the full-rate case. Furthermore, we demonstrate calibration examples using the *half-rate* (sub-rate of 2) case.

3.3   Receiver Calibration

As the input signal bandwidth increases, the non-linearly distorted version of the same signal, and accordingly TLBW, increases. To capture this wide-band signal, we must use a receiver of the same bandwidth, or preferably larger. This wide-band receiver should not exhibits lots of frequency variation across its bandwidth. However, for large signal bandwidths, having these frequency variations is inevitable. Consequently, more burden is added to the receiver calibration, given that these variations on the receiver frequency response are not very severe.

Receiver calibration can be normally applied on the received data in the digital domain. Receiver calibration [3] in this context is meant to correct for the receiver linear distortion, i.e., receiver frequency response or receiver frequency in-flatness. There are other receiver calibration techniques, that are not part of this thesis work, corrects for the receiver non-linear distortion behavior. This non-linear calibration is meant to extend the equipment SFDR, rather than a frequency response simple correction [27]. Receiver calibration involves two different routine:

1. **Receiver transfer function (frequency response) identification**. This step is performed only once to find the receiver transfer function, H_{Rx}^{-1}.

2. **Receiver transfer function application**. Once the receiver transfer function is found, either in time or frequency domain, it is applied on the received data. Calibration is applied just directly after receiving the data before any other processing.

3.3.1   Full-Rate Receiver Calibration

One way to identify the receiver transfer function is to use a calibrated transmitter. However, to get this calibrated receiver, a calibrated transmitter must also exists beforehand. Since in a real measurement setup scenario we do not have access to neither calibrated receiver nor transmitter, we can not separate the receiver errors from the transmitter. In summary, to calibrate a receiver, a calibrate transmitter is needed. Similarly, to calibrate a transmitter, a calibrate receiver is needed.

Some research ideas have been proposed to resolve this calibration dilemma [31, 7]. These ideas depend on different signal processing tricks to separate the transmitter from the receiver error terms. In contrast to the signal processing based solution, and much simpler too, we can use a reference ideal source to calibrate the receiver, then use this calibrated receiver to calibrate the transmitter. A commonly known ideal source, and used widely for this particular application, is the comb generator. Comb generators are an ultra broad-band sources used for this thesis work. Comb generators are found to be very useful in general for receiver calibration, and in particular for sub-rate receiver calibration applications, as we explain in the next chapter.

So, in summary, for full-rate setup, the receiver and transmitter calibration can be done in the following way: first use a calibrated source, comb generator, to calibrate the receiver, Fig. 3.3. Then, use this calibrated receiver to calibrate the transmitter, Fig.

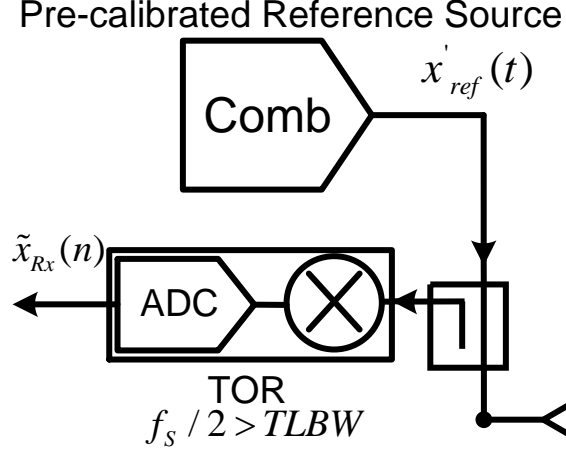[3]cal and calibration can be used interchangeably

Figure 3.3: Full-rate receiver cal using comb generator.

3.4. The outcome of this process is the identification the individual transfer function of both receiver (3.11) and transmitter (3.13), $H_{Rx}^{-1} \, H_{Tx}^{-1}$. Worth mentioning that receiver and transmitter transfer functions relate two baseband equivalent quantities. For that reason, the baseband equivalent of the RF signal is used in those transfer function calculations.

$$H_{Rx}^{-1} = \frac{\tilde{x}_{ref}(n)}{\tilde{x}_{rx}(n)} \tag{3.11}$$

$$x_{ref}(t) = \Re(\tilde{x}_{ref}(n) \cdot e^{jw_c t}) \tag{3.12}$$

$$H_{Tx}^{-1} = \frac{\tilde{x}_{tx}(n)}{\tilde{x}_{rx}(n) \cdot H_{Rx}^{-1}} \tag{3.13}$$

Once these two transfer function are well identified, the second step is to apply them on the transmitted and received data. In particular, transmitter calibration should be applied on the transmitter data just before it is being transmitted. Similarly, receiver calibration should be applied on the received data right after they are received. Collectively, and after the application of the calibration functions, the transmitter and receiver hardware can be now considered ideal, Fig. 3.5.
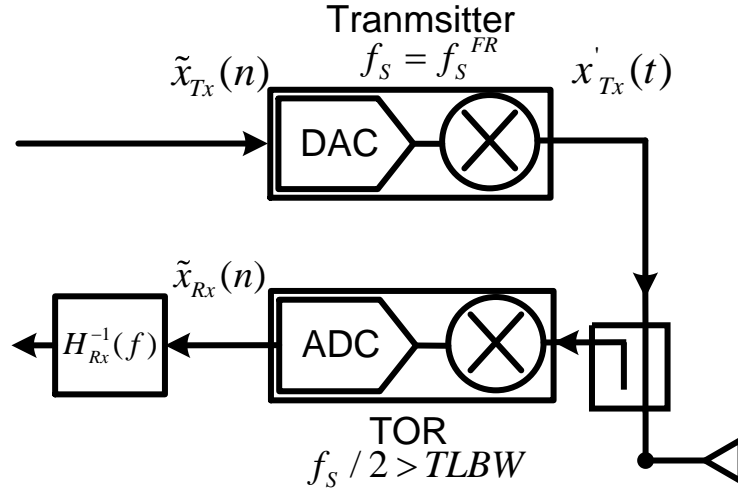
Figure 3.4: Transmitter calibration using full-rate receiver.

3.3.2 Sub-Rate Receiver Calibration

The previous section establish the proper background to understand the sub-rate receiver calibration problem. Inspired form the last section, receiver calibration problem is a two part problem: receiver transfer function identification, and receiver transfer function application. For the next particular section, we only investigate the transfer function *application* part. We will discuss the identification problem, however, in the next chapter. The identification problem is more of a characterization issue and involves low practical measurements details. For this reasons, we find it is better to explored the identification problem in the hardware measurements chapter.

Sub-rate Receiver Application for DPD Applications

From Fig. 3.5 and (3.11) we can observe that to successfully apply the receiver calibration, the sampling frequency of the received signal should be above the Nyquist rate, so as to avoid any aliasing distortion. An output sampling rate below Nyquist rate yields an overlap between some of the frequency contents due to *aliasing*. This aliasing prevents the correct application of the receiver cal. This happens because those overlapped frequency components should be calibrated (or corrected) using different parts of the receiver transfer function, Fig. 3.6. Since the overlapped components can not be separated, the
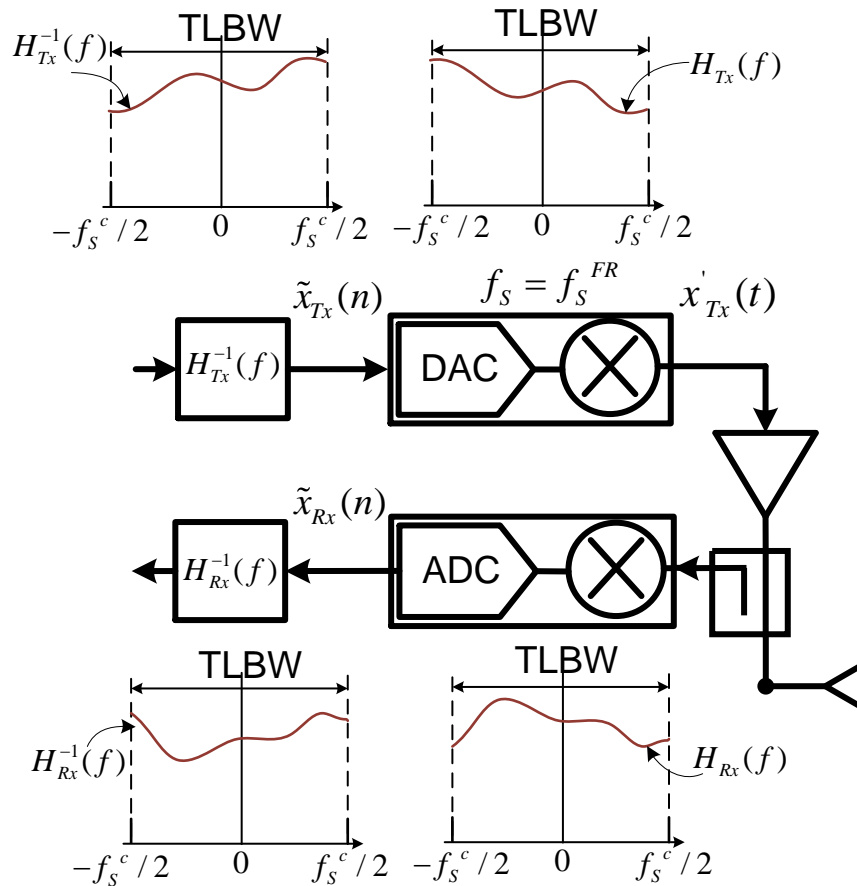
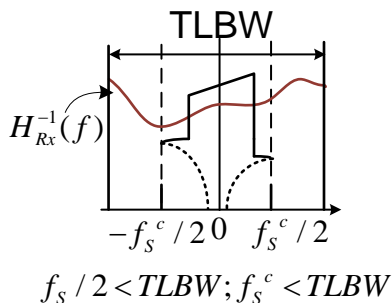Figure 3.5: Transmitter and receiver calibration using full-rate receiver.

Figure 3.6: Sub-rate receive calibration problem.

receiver calibration routine must be modified. Such problem only arises in the case of using sub-sampling hardware, where we only have access to the sub-sampling data. That is why receiver calibration was not addressed in any of the previously reviewed methods for sub-sampling DPD in the literature review. For those sub-sampling DPD that was demonstrated on full-rata hardware, they could apply receiver calibration on the full-rate data, before the digital down-sampling step, where no aliasing distortion has yet been accrued, Fig. 3.2a.

For this receiver calibration problem, we modify the DPD system by moving the receiver calibration from the receiver part to the transmitter, 3.7a. At the transmitter, the data is sampled at the full-rate sampling speed. This mean the receiver calibration can be normally applied on the original input $\tilde{x}(n)$. From mathematical point of view, this movement of blocks can be only valid for linear systems, i.e., $g(f(x)) = f(g(x))$ iff both g and f are both linear function. As a result, for this move to be valid, every thing between the old place of the receiver calibration block to the new receiver calibration place, every thing should be linear. Unfortunately, this is not the case for our DPD problem, because the DPD and PA blocks are both non-linear functions. Luckily, we can consider the cascade of the DPD and PA blocks to be one linear block.

One of the drawbacks of the proposed approach is that during DPD training the PA output is always distorted by the linear receiver calibration function. This is why after the DPD is successfully trained, the receiver calibration block is bypassed, Fig. 3.7b. This step is necessary to assure that the PA output signal is not linearly distorted by the receiver calibration block.
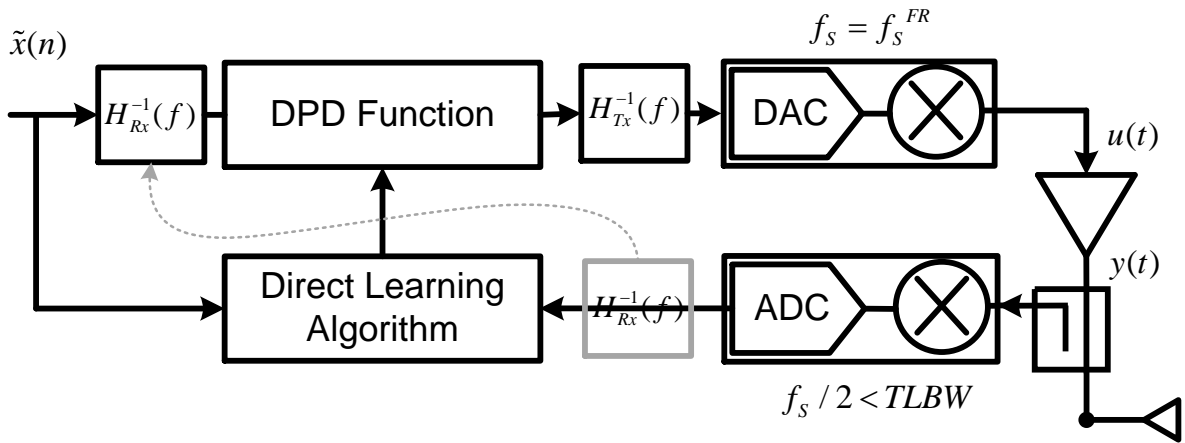
(a) DPD identification.
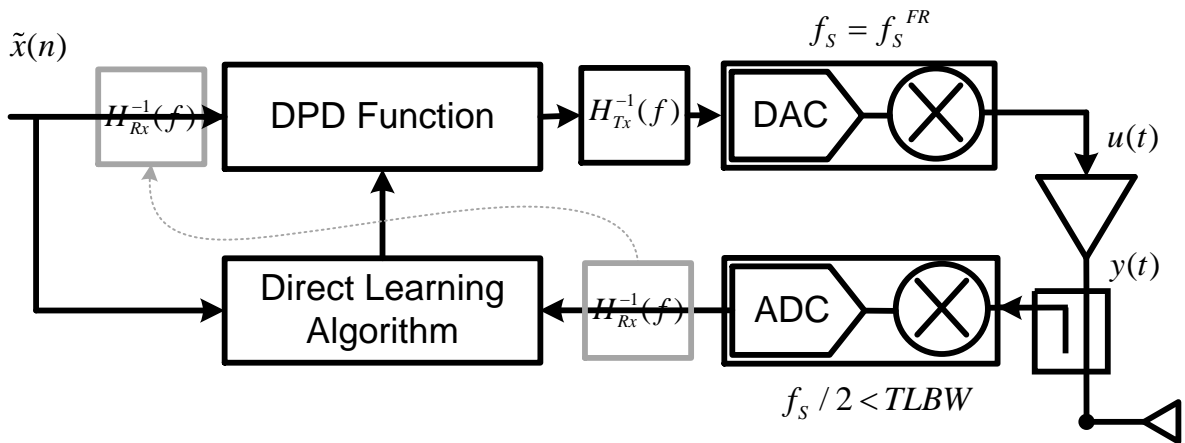
(b) DPD application.

Figure 3.7: Sub-sampling DPD using sub-sampled hardware.

Comments on Convergence

Although this proposed solution may seem intuitive, we have to address few subtle details regarding the convergence here:

- For the first iteration, where the DPD function is still far from the real inverse of the PA, the assumption made before, regarding the linear nature of the cascade of the DPD and PA, is not valid. This definitely affects the performance of the DPD algorithm. However, if the DPD converged, i.e., the DPD now is exactly the inverse of the PA, the receiver calibration block movement is valid.

- The receiver calibration transfer function is defined over the whole TLBW. In the sub-rate scenario, receiver calibration is applied to the input signal before it is distorted by the DPD, $\tilde{x}(n)$, at the transmitter. This signal has a bandwidth which is 5 times less than the TLBW which means that the frequency bands within TLBW and outside the input signal modulation bandwidth are not calibrated. This occurs because when the input signal $\tilde{x}$, at the transmitter, is multiplied by the inverse of the receiver transfer function, some of the calibrated frequencies bands have zeros spectrum content. This happens as a result there is not signal power at those frequencies. Consequently, those frequency bands are not calibrated. At convergence, nevertheless, the PA output, $\tilde{y}(n)$, has a bandwidth similar to the input signal $\tilde{x}(n)$. So, at convergence, calibrating the signal output signal within this small bandwidth is sufficient for correct DPD operation.

From the previous two mentioned points we conclude that the proposed calibration does not yield the best results for the first couple of DPD iterations where the PA output is still a wide-band signal. As the DPD algorithm is converging, the two assumptions made before are becoming more valid, and the final solution should achieve similar results compared to the conventional DPD. However, the converging speed of the algorithm will be slow, since the first couple of iterations do not provide the optimum descent towards the global minima. Another drawback of the previously mentioned points is that DPD algorithm is prone to divergence; again, because the first couple of iterations might steer the gradient descent towards a bad solution. To avoid this, the step size in the direct learning coefficients update equation (3.7) should be picked to be small, especially at the first couple of iterations[4]. This might slow down the convergence speed compared to the conventional DPD, but it lowers the chance of divergence.

---

[4]It is interesting to point that this choice of step size is the opposite to the conventional optimization case. Conventionally, the step size is picked to be decreasing with iterations.

By applying these techniques, the sub-rate DPD algorithm can be practically implemented using sub-rate hardware. Our proposed modifications can be summarized as follows:

- **Receiver IF frequency**. Modify the receiver IF frequency to be exactly $F_{sADC}/4$, assuming super-heterodyne receivers.

- **Receiver calibration application**. During DPD training, we apply receiver calibration at the transmitter. Once the DPD is trained, we bypass receiver calibration.

- **Coefficients update parameter**. This parameter should be chosen to be small to lower the chance of DPD divergence.

3.4   Conclusion

This chapter presented a hardware implementation based on the sub-sampling DPD algorithm theory proposed by Hai *et al* [19]. The advanced method in this chapter uses only hardware that runs at the reduced sampling speed.

We showed that the only modification to the hardware is the choice of the receiver IF frequency. However, for this practical implementation, a calibration problem showed up. As a result, this chapter also demonstrated how to apply receiver calibration for sub-rate hardwares. In detail, we slightly modified the DPD system block digram by moving the receiver calibration to the transmitter. It was determined that this move should only be applied during DPD training. To allow for this block movement, we highlighted some assumptions regarding the different blocks in the DPD system. Since, these assumptions are not always true, especially at the first couple of DPD iterations, the DPD divergence speed is affected. As a result, we provided few comments on the convergence of the DPD training algorithm's speed.

The next chapter shows the real hardware implementation of the ideas we have proposed. Next chapter shows linearization measurements results that demonstrate the validity of the proposed method. In the first part of the next chapter, we present the identification of the receiver and transmitter transfer functions for calibration. Finally, we show the linearization measurements results for different sub-rate values. Importantly, all the measurements are done using hardware that runs at the reduced sampling speed, without any digital down-sampling involved.

Chapter 4

Practical Implementation of Sub-sampling DPD

This chapter demonstrates a practical implementation of the sub-sampling DPD using the receiver calibration technique proposed in the previous chapter. First, we introduce the measurement hardware setup used. To calibrate this hardware, we utilized a multi-tone comb generator. Besides the previously mentioned techniques related to applying the receiver calibration on the sub-sampled data, this chapter shows how to identify the receiver transfer function using only sub-sampled data as well. Moreover, we also present a vector transmitter calibration. This transmitter calibration involves identifying the transmitter transfer function by measuring only sub-sampled signals.

To demonstrate sub-sampling using real hardware, our measurement setup employs interleaved ADCs. These interleaved ADCs, used in the TOR, are exploited to synthesize different sampling speeds; in other words, different sub-rate ratios.

Using all the above mentioned techniques, we test a real sub-sampling DPD on this setup. We show that sub-sampling DPD is capable of achieving linearization results similar to those of the conventional DPD. Hence, we demonstrate a successful linerization of 320MHz, 200MHz bandwidth signal using different sampling speeds.

4.1 Transmitter Receiver Transfer Function Identification using Sub-sampled Data

This section demonstrates a novel method to calibrate a wide-band setup using ADCs running at a sampling speed much below Nyquist. By calibrating a measurement setup here we mean to find the receiver and the transmitter transfer functions, H_{Rx}^{-1}, H_{Tx}^{-1} respectively. For the transmitter we include the lump of all the hardware components that transform the ideal digital data (bits) to the RF signal. Similarly, the receiver refers to all the hardware components that down-convert the PA output RF signal to IF, and then to the digital domain. Those transmitter and receiver transfer functions represent the lumped frequency response effect of all the components used inside.

The system calibration involves first calibrating the receiver using a multi-tone reference source, i.e., comb generator. This calibrated receiver is used, in turn, to calibrate the transmitter. As described in the next sections, we linearize a TLBW of 1GHz. If the receiver ADC is running at full sampling speed (full rate), the ADC sampling speed is 2Gsps, assuming a super heterodyne receiver. Similarly, the ADC can run at a sampling speed of 1Gsps (sub-rate of 2, half rate), or at a sampling speed of 500Msps (sub-rate of 4, quarter rate). All the next measurements are completed around 24GHz, which is our RF center frequency f_c. The next section provides numerical examples for receiver calibration with an ADC running at a sub-rate of two. Nevertheless, the same techniques can be applied for any further sub-sampling value.

4.1.1 Wide-band Receiver Calibration Using Sub-Nyquist Signal

The problem that we are trying to solve here is how to measure a wide band reference signal using a sub-sampling TOR. This sub-sampling TOR has an ADC that runs at a sampling speed that is much lower than the Nyquist rate, i.e., $F_{s\ ADC}/2 < TLBW$. Having this low sampling speed results in distortion due to the image and alias overlap with the in-band signal[1].

In general, once these frequency bands are folded over each other, they are no-longer separable. Nevertheless, since the comb signal is periodic, we can exploit this periodicity

---

[1]Although image and alias distortion are both due to the same thing (aliasing), nevertheless, in this thesis we are using the two terms to differentiate between at which side of the signal spectrum the overlap is happening. In particular, image overlap distortion happens at the left hand side of the IF signal spectrum, and alias overlap distortion happens at the right hand side of the IF spectrum in Fig. 3.2a

to avoid such overlap. Periodic signals are known to have a discrete frequency spectrum. That is, the frequency spectrum is sparse with plenty of empty spots inside. Although the folding is inevitable, we can still apply some techniques to control it. If we carefully controlled the folding such that the folded tones will only fall in the spectrum of empty spots, we can separate the different tones and restore the original signal.

We calibrate the receiver by feeding the comb generator signal to the receiver input. Comb signal is periodic and discrete in frequency. Specifically, it is a reference multi-tone signal with Δf spacing (usually Δf is 10MHz), where those tones have a known magnitude and phase. The comb generator used in this work has an ultra wide-band output, i.e., from 10MHz all the way up to 67GHz. Within this wide band signal, we are only interested in calibrating a TLBW of 1GHz around a certain center frequency, $f_c = 24$GHz, Fig. 4.1.

Since the comb signal frequency band that we are interested at has a bandwidth of TLBW (which is larger than the ADC sampling speed), it experiences a distortion after down-conversion due to image and alias folding. The folded tones will overlap with the in-band tones resulting in an unwanted distortion. We should not discard these folded tones, however, we should measure them since they have the receiver transfer function information at the image and alias bands within the TLBW, as can be seen in Fig. 4.1. Hence, these tones must be separated from the in-band ones and measured individually. The next two sections demonstrate the techniques to separate those tones. The first section shows how to measure the image band, and the next one shows how to measure the alias band.

Image Tones Separation

The image problem is a well-known RF problem. If the RF signal (before down-conversion) has any frequency contents at frequencies of equal distance to the LO, this frequency part is called an image. The down-conversion mixer down-converts the image signal to the same frequency band as the intended signal, which causes distortion due to overlap between the two down-converted signal, Fig. 4.2a.

For example, in our case (for half rate) the $TLBW$ is 1GHz, F_{sADC} is 1Gsps, f_{IF} is 250MHz, and the comb tones frequency spacing Δf is 10MHz. We can also assume lower LO frequency injection, i.e., $f_{LO} = F_{RF} - f_{IF}$. For this particular case, image overlapping happens as the following: RF comb tones at $f_{LO} + i \, \Delta f$ and $f_{LO} - i \, \Delta f$ both be are down-converted to the same IF frequency, $i \, \Delta f$. Hence, the two down-converted signals overlap, 4.2a. Concretely, tones at 23.73GHz and 23.77GHz are down-converted to the same frequency, 20MHz.
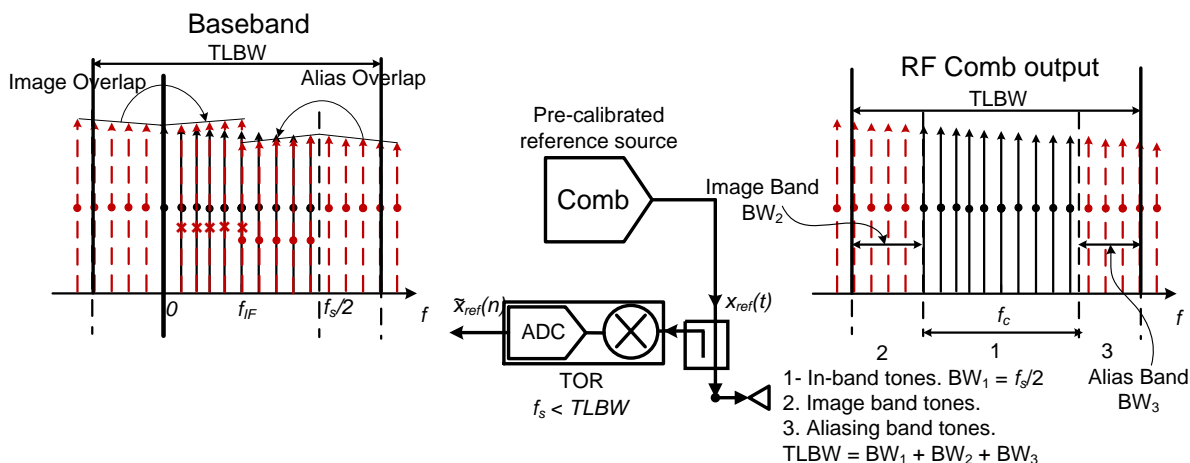
Figure 4.1: Receiver calibration using comb generator and sub-Nyquist sampling ADC.

The conventional solution to the image problem overlap is to use an RF filter for rejecting this image frequency band before down-conversion. However, in our particular case, our goal is not only to measure the in-band tones, like a conventional RF down-converter. Moreover, we want to measure the image band as well. Another conventional solution is to use IQ demodulators. These demodulators can separate the in-band signal from the image. However, they require twice the receiver hardware resources, and a difficult calibration routine for any mismatch between the I and Q path, especially for a wide-band operation.

The solution we propose here is very simple. We adjust the LO such that $F_{IF} \neq i \cdot \Delta f/2$. For example, instead of having f_{IF} and f_{LO} of 250MHz and 23.75GHz in the half-rate case, we can use $250 - 2.5$MHz and 23.75GHz $+ 2.5$MHz. Using the same previous example, now tones at 23.7325GHz and 23.7725GHz are down-converted to 17.5 and 22.5MHz respectively; hence they can be separated and measured individually, Fig. 4.2b. This separation is possible because we break the symmetry around the f_{LO} frequency. Once those image tones are measured, we can find the receiver transfer function at the image band.

Alias Tones Separation

Regarding the aliasing tones, tones at the aliased band fold back to the in-band tones after digital sampling. For example, in the previous case after we fixed the image tones overlap
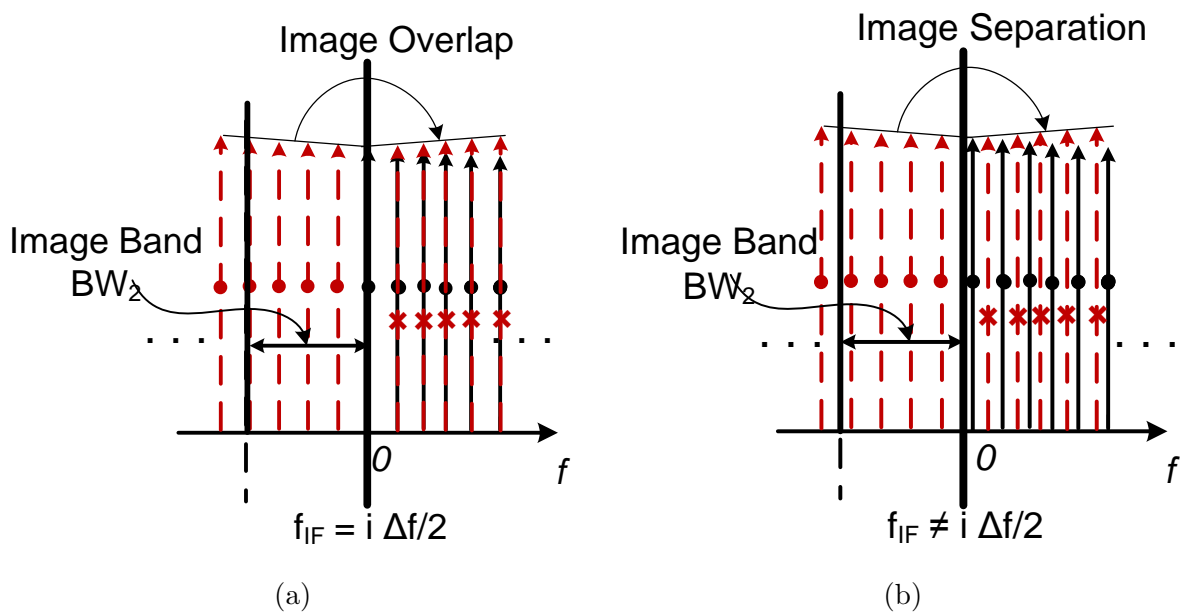
55

Figure 4.2: Down-converted image tones. (a) Image tones overlapping. (b) Image tones separation.

problem, the down-converted tones have frequencies of [**2.5 12.5 22.5** $\cdots$ **482.5 492.5 502.5 512.5MHz** $\cdots$] for the in-band, and [**7.5 17.5 27.5** $\cdots$ **487.5 497.5 507.5 517.5MHz** $\cdots$] for the image. It can be seen that those tones are symmetric around $F_{sADC}/2 = 500$GHz. Hence, tones at the aliased band, which are down-converted to frequencies above 500MHz, fold back on the in-band tones, causing distortion. Concretely, 497.5 and 502.5MHz tones both appear as they are at 497.5MHz, 4.3a. The conventional solution to this problem, which we can not use in our case, is the use of anti-aliasing filter. Anti-aliasing filters reject all frequency contents at bands higher than $F_{sADC}/2$, however we can not use this solution since we want to measure these aliased tones.

It can interpreted from the previous example that the aliasing folding problem is happening due to having a symmetry in tones frequencies around $F_{sADC}/2$. If we break this symmetry, aliased tones no longer overlap with the in-band ones. For example, if $F_{sADC}/2 \neq i \cdot \Delta f/4$, the tones at the right and left of $F_{sADC}/2$ no longer have an equal distance from $F_{sADC}/2$. Hence, the aliased tones will fold on an empty spot in the frequency and will not overlap with any other tone. For example, let $F_{sADC} = 998$Msps instead of 1000Msps, i.e., $F_{sADC}/2 = 499$Msps. This small change in the sampling frequency beaks the symmetry of tones frequencies around F_{sADC}. Consequently, the tone at 502.5MHz mentioned above folds on 493.5MHz. This 493.5MHz frequency is still empty in our case, hence, we can separate and measure those aliased tones and find the receiver transfer function at the aliased band, Fig. 4.3b.

Using the two previously techniques, we can calibrate a wide band receiver using an ADC running at a very low sampling speed. Generally, we adjust the f_{IF} and, hence, the f_{LO} to separate the image, and adjust the F_{sADC} to separate the alias. These separations are possible by exploiting the discrete nature of the reference signal. Since we now have a calibrated receiver, we can use it to calibrate our transmitter. But before moving to the transmitter calibration routine, a few assumptions regarding the proposed receiver calibration routine that were implicitly made, must be mentioned.

Sub-sampling Receiver Calibration Assumptions

The following assumptions all result from one important aspect: the calibration conditions are different from the validation ones. Concretely, for the sub-rate of two case, we used an f_{IF}, f_{LO}, and F_{sADC} of 257.5MHz, 23.7525GHz, and 998Msps to separate the overlapped tones during calibration. However, when we do the validation measurement (DPD), we will change those frequencies back to 250, 23.75GHz and 1Gsps respectively.

As a reminder, these validation frequency values are necessary to generate the intended
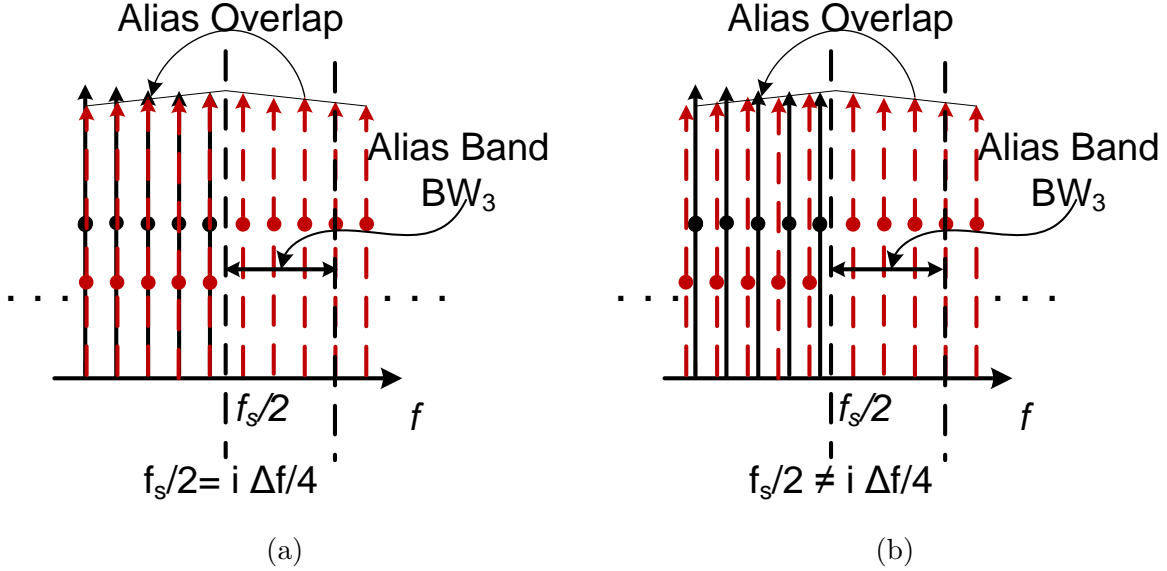
Figure 4.3: Down-converted aliased tones. (a) Aliased tones overlapping. (b) Aliased tones separation.

aliased signal as we explained at the previous chapter 3.2.2. Theses small discrepancies in frequency values, between calibration and validation, add some error in the calibration identification. Therefore, we state the next few assumptions to define the conditions by which this error is minimized. If any of these assumptions is not true for a specific application, the aforementioned routine should not be used.

1. **Conditioning circuits bandwidth**. For the receiver used, all the RF and IF circuity must have a wide bandwidth of TLBW. Therefore, the proposed technique only reduces the ADC sampling speed. The wide analog bandwidth requirement for the circuity is important since we do not want to filter the image and the alias bands.

2. **Effect of changing the f_{LO} and f_{IF}**. The change of F_{LO} at calibration to the value that will be used in the validation measurements (DPD) will affect the accuracy of the calibration routine. We mentioned before that the proposed calibration routine finds the transfer function of the lumped effect of all the circuity at the receiver. These circuity can be divided into two subgroups: RF and IF. RF circuity are the ones that operate near our RF center frequency, 24GHz in our case, and IF circuity are the ones that operate near our IF frequency, from DC to 500MHz in our case. Fig 4.4 depicts that the mixer can be used to separate the IF from the RF ones.

The way receiver calibration is done is to find the receiver transfer function on a specific frequency grid. Then, we interpolate to find the transfer function at all the in between frequencies. In particular, during calibration the in band RF frequency grid is $[\cdots \mathbf{23.98\ 23.99\ 24\ 24.01\ 24.02GHz} \cdots]$, and the IF grid is $[\cdots \mathbf{227.5\ 237.5\ 247.5\ 257.5\ 267.5MHz} \cdots]$. While doing the DPD measurements, the RF frequency grid is the same, however, the IF grid is different: $[\cdots \mathbf{230\ 240\ 250\ 260\ 270MHz} \cdots]$. having this 2.5MHz mismatch between the two grids is a source of error. Nevertheless we can assume that the RF circuity response is not changing within this 2.5MHz difference. This assumption is valid since the amount of frequency variations around high frequencies is negligible.

3. **Effect of changing F_{sADC}.** Since the ADC sampling speed is slightly modified at calibration that the value that will be used while doing the actual measurements, this can introduce some error. The we assume that the ADC response should not change with this small modification in the sampling frequency.

4. **No reflection or mismatch correction**. The proposed routine is a vector single port calibration, i.e., we had only to capture the voltage at the ADC input. In contrast, a 2 port vector calibration is often required to fully calibrate the linear distortion. There are different methods for 2-port vector calibration (TRL, SOLT, SOLR, LRM) [28], but they require a 2 port receiver, i.e., a coupler and another receiver to measure the reflected wave.

   Single port calibration, like the one we propose, can be still valid under a certain assumption, reflected wave is very small such that its effect can be ignore. This assumption is valid if the mismatch at the different components ports is very small, typically below -10dB. If this is not valid, the proposed routine must be extended to account for mismatch (reflection). However, we did not have to account this in our work.

By meeting these assumptions, the proposed receiver calibration routine can be used. The next step in to calibrate the transmitter using this calibrated receiver. To show a real implementation, we had to modify the calibration routine to allow for sub-sampling calibration.
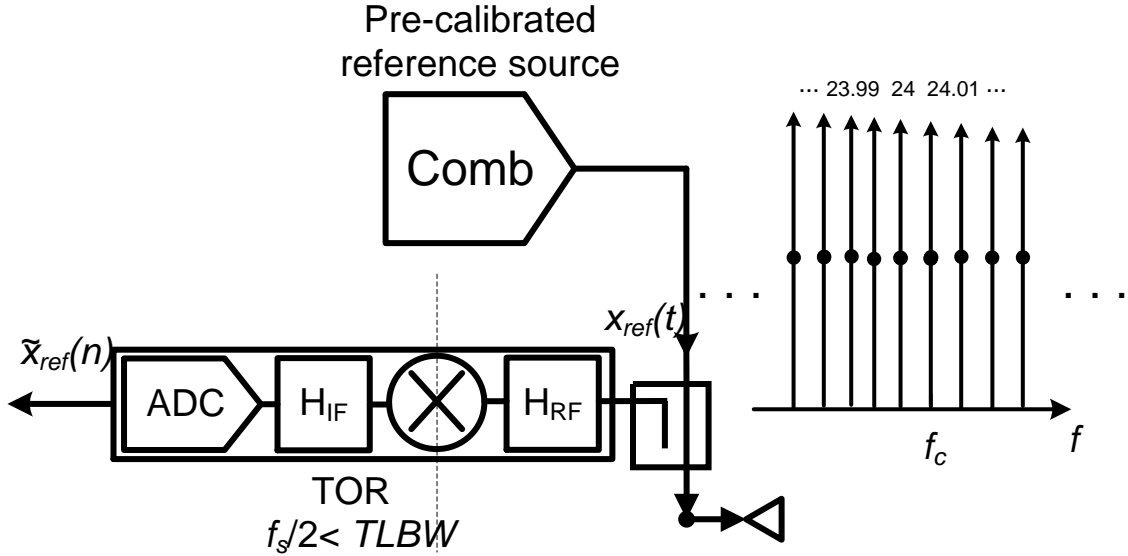
Figure 4.4: Receiver RF and IF transfer functions.

4.1.2 Wide-band Transmitter Calibration using Sub-Nyquist Signal

For our particular application, the transmitters we utilize wide band signal generators. That is, the sampling speed of the signal generator is high enough to allow for a normal generation of a wide band signal of bandwidth TLBW. So, our work is not about sub-sampling transmitters, our work is more focused towards the receivers side. However, we still have to calibrate this wide-band transmitter using a narrow band receiver.

We introduced a transmitter calibration routine using a calibrated receiver in previous chapter, 3.3.1. A multi-tone signal is sent form the transmitter and measured with this calibrated receiver. By measuring the distortion at these tones, we can find the inverse of the transmitter transfer function, H_{Tx}^{-1}. In a conventional transmitter calibration, a tones spacing of 10Mhz can be used. Those tones, when they are up-converted to the RF frequency, will have a frequency grid that is similar to the comb generator tones: [**23**.5 $\cdots$ **23.98 23.99 24 24.01 24.02** $\cdots$ **24.5GHz**]. One possible method to measure these tones is to use the same previously mentioned comb generator technique: change the f_{LO} and F_{sADC}. Nevertheless, we can use another approach that does not require any modifications in the above mentioned frequency parameters; hence, we do not

have to make any assumptions regarding our hardware. This other approach relies on generating the multi-tone signal with different frequency spacing. Such control on the frequency spacing was not possible in the comb generator case.

If we generated the tones such that the up-converted tones has the following pattern [**23**.**502** $\cdots$ **23**.**982 23**.**992 24**.**02 24**.**012 24**.**022** $\cdots$ **24**.**492GHz**]. This pattern breaks the symmetry around the f_{LO} of 23.75GHz and f_{sADC} of 1Gsps automatically. Hence, allowing for image and alias tones separation. For the DPD measurements, for the sub-rate of 2, we chose f_{LO} and f_{IF} to be 23.75GHz and 500Mhz respectively. This means the down-converted frequency grid for the in-band tones will be: [**2 12** $\cdots$ **242 252 262** $\cdots$ **492MHz**], and for the image ones is: [**8 18** $\cdots$ **258MHz**]. The aliased tones are down-converted to [**502 512** $\cdots$ **742MHz**], which are folded back on [**498 488** $\cdots$ **258MHz**].

The careful reader will notice that if a larger TLBW is needed, this technique fails, since it can be seen from the previous example that the image and aliased tones are now overlapping at the last tone 458MHz. In other words, if the TLBW is 1.2GHz, offsetting the generated tones by 2MHz, like the previous example, is not enough 4.5a.

To solve this problem, we can change the frequency spacing of the generated tones such that no overlap happens. For example, knowing that the tone 742MHz at the aliased band are folded at 258MHz,causes such over lap. We can, while generating these multi-tones, shift this tone, and all the next tones by 1MHz, i.e., the new tones sequence after down-conversion is [**2 12** $\cdots$ **732 742 753 763 773MHz**]. The 753MHz tone now is folded back on 247MHz, so no overlapping happens, 4.5b.

The higher the sub-rate ratio, the more folding happens, the more attention should be paid to where to apply these sudden changes in the tone spacing. Using those three techniques, any sub-ratio can be achieved, theoretically. Collectively, those three axis of freedom (f_{LO}, f_{sADC}, and generated tones spacing) can allow for a *substantial* decrease in the ADC sampling speed.

4.2   Wide-band Sub-sampling Linearization Results

This section shows the PA linerization results using sub-sampling DPD. First, we introduce the setup that allowed for real sub-sampling DPD implementation. In particular, this setup utilizes an interleaved ADCs that can synthesize multiple sampling speed: 2Gsps, 1Gsps, and 500Msps. The obtained results shown that sub-sampling DPD is capable of achieving
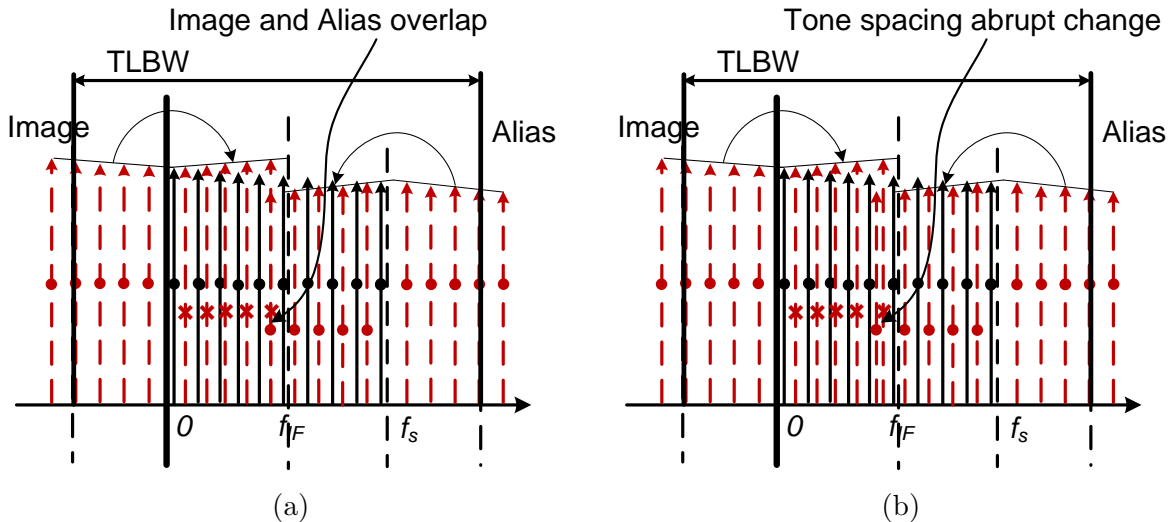
Figure 4.5: Transmitter tones overlapping. (a) Transmitted tones without tones spacing modifications. (b) Transmitted tones with tones spacing modifications.

similar linearization results, compared to conventional full-rate DPD. Signals with 320 and 200MHz modulation bandwidth are used to validate the proposed methods.

4.2.1 DPD Setup Calibration

DPD measurement setup is, simply, a transmitter and a receiver. The transmitter generates the pre-distorted RF signal that is fed to the PA. Likewise, the receiver measures the PA output RF signal. The receiver we use must be a vector receiver, i.e., it demodulates the RF output signal to a complex base-band signal using a down-converter and a ADC. Besides this vector receiver, it is commonly to use another scaler receiver to measure the PA output RF signal [2]. This scaler receiver is important to assure having a linear signal at the output of the PA rather than at the output of the vector receiver. Therefore, we use a spectrum analyzer to measure the ACPR of the PA output RF signal .

The vector receiver we use in this measurement setup mimics the TOR in the DPD block digram. For that receiver, and to synthesize multiple sampling frequencies, we use

---

[2]Scaler receiver only measures power across frequencies, in contrast to vector receivers that measure complex signal (magnitude and phase). Power measurement can be useful for multiple application, ACPR and IMD measurements for example. Main advantage of spectrum analyzer is they can easily achieve wide-band

interleaved ADCs. Specifically, we used the M9708A digitizer from Keysight [2]. This digitizer consists of 8 channels each with 1Gsps sampling speed. To synthesize different sampling speed, we can use ADCs interleaving [8]. For instance, this digitizer allows for interleaving each two channels to synthesize a combined sampling speed of 2Gsps. In addition, each one of the 8 channels already consists of two ADCs interleaved together. Therefore, we can also synthesize a sampling speed of 500Msps by usign only one of these two interleaved ADCs. In summary, using these interleaved ADCs we can test our setup using three different sampling speeds, 2Gsps, 1Gsps and 500Msps. Since our maximum sampling speed is 2Gsps, we limited our TLBW to 1GHz.

The final measurement setup used is shown in Fig. 4.6. For the transmitter part We use a M8190A 8Gsps/14bits AWG from keysight, with an IF output centered around 2GHz. To up-convert the baseband signal from 2GHz to 24GHz, we use two stage up-converters; first stage up-converts the signal to 5GHz, and the second to 24GHz. For the receiver part, we used single stage down-converter to down-convert the signal to the receiver IF frequency. Finally, the digitizer would convert the signal to the digital domain for further processing.

Designing such setup with many components is not a straight forward task. Filters, attenuators, and amplifiers have to be carefully added to assure proper operation. Filters are generally used at the transmitter to reject the different images generated from the sub-sequent stages. Attenuators and amplifiers are used to adjust the signal power level according to the best operating conditions for the different components. With the addition of these extra components, receiver and transmitter transfer function have more variations, and in this case calibration is of great importance. Consequently, the previously mentioned techniques for calibrating the transmitter and the receiver must all applied before doing any DPD measurements. For 1GHz of multi-tone signals, an EVM below 1% could be obtained after calibration.

4.2.2 Linearization Measurements Results

After calibrating the setup, we add the PA. We use the 83018A [1] class AB PA operating at P_{1dB} of 17dBm. This PA is driven by 200 and 320MHz signal. These signals can be either 16QAM with PAPR of 6dB or carrier aggregated LTE signal with 8.4dB PAPR. We use a memory polynomial DPD with $N = 7$, and $M = 3$. We linearized this PA with and without applying the receiver calibration, and using different sub-sampling rate values.

Fig. 4.7a and 4.7b depict the DPD results for the 200MHz case. It can be seen that sub-rate linearization results are similar to the full-rates case. From the same figures, it can
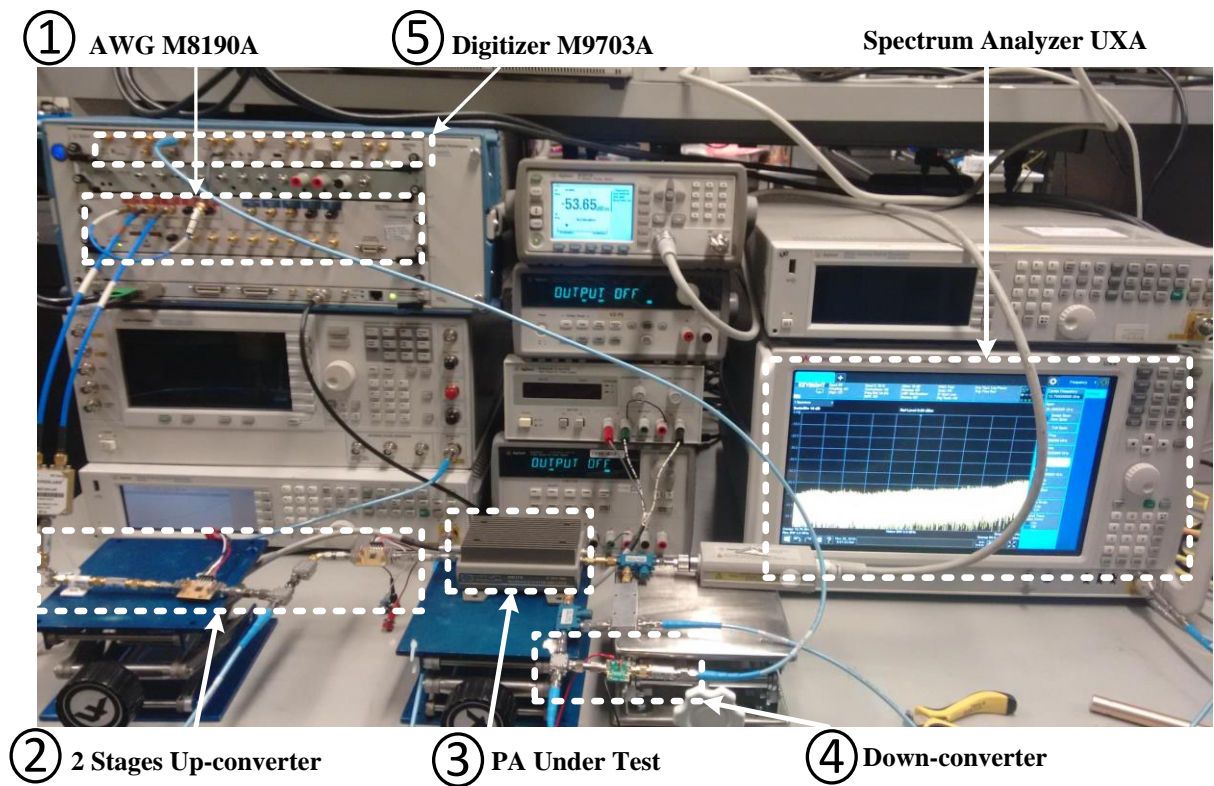
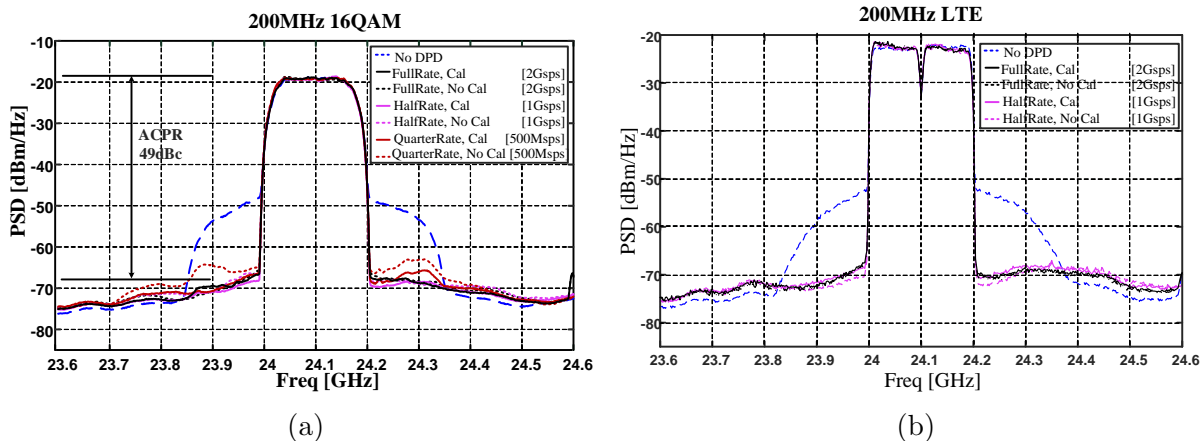Figure 4.6: Real Sub-sampling measurements Setup.

Figure 4.7: DPD linerization results for 200MHz (a) 16QAM. (b) LTE.

Table 4.1: Proposed sub-sampling DPD linearization results.

Signal	PAPR [dB]	Full-Rate		Half-Rate		Quarter-Rate	
		DPD+ Rx Cal	No DPD+ Rx Cal	DPD+ RX Cal	DPD+No Rx Cal	DPD+ RX Cal	DPD+No Rx Cal
200MHz 16QAM	6	2.2/51	7.4/29	2.5/51	3.6/47	2/49	3.4/43
200MHz LTE	8.4	2/46	9.7/29	2.2/45	3.8/45	2/45	4/43
320MHz 16AQM	6	2.7/48	7.6/29	2.8/47	5.3/40	N/A	N/A
320MHz LTE	8.4	2.6/44	9.7/27	2.2/45	3.8/45	N/A	N/A

be concluded that calibration has important effect on DPD. Without calibration applied, the DPD algorithm usually diverges, or, in best scenario, converges to a worst solution.

Similar results are obtained for the 300MHz case, Fig. 4.8a and 4.8b. For both LTE and 16QAM, sub-sampling DPD could achieve similar linearization results, compared to the full-rate. It is evident that calibration plays an important rule in algorithm convergence. Finally, table 4.1 summarizes the linearization results for the 4 different signals. It can be seen that all of them met the LTE spectrum emission mask, 45dBc.
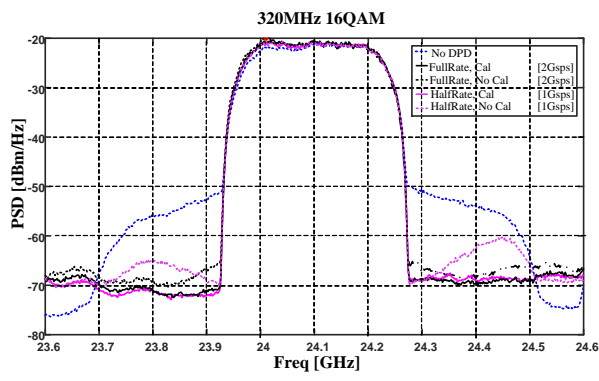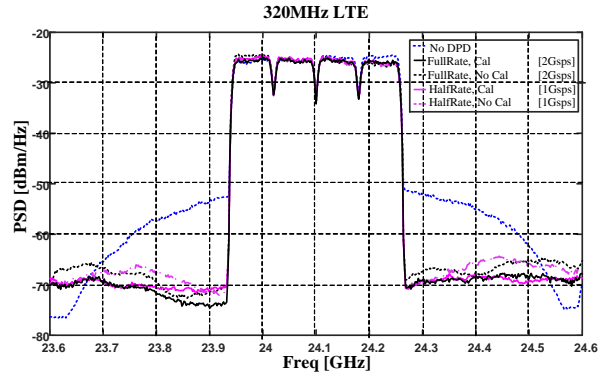
Figure 4.8: DPD linerization results for 320MHz (a) 16QAM. (b) LTE.

Chapter 5

Conclusion and Future Work

This thesis aims to reduce the DPD added circuitry power consumption to enable efficient linearization for PAs employed in small-cell basestations driven by wide-band modulated signals. There is a growing interest to deploy small-cell basestations for some of the 5G anticipated use cases, since they achieve improved mobile network capacity and coverage. Chapter 1 explains how the DPD extra circuitry power consumption contributes to the overall system efficiency. For the small-cell basestations, PA outputs relatively small amount of power; hence, the absolute power saved from improving the PA efficiency is low. On the other hand, the TOR ADC power consumption is high, since the base-station is transmitting wide-band signals. Furthermore, this consumed power in independent from the PA output power, i.e., the ADC consumed power does not scale down with the PA output RF power.

In this work, we investigate a method to reduce the TOR consumed power by reducing the ADC sampling speed, since sampling speed is one key parameter that dictates the ADC power consumption. Although multiple researchers have proposed multiple ideas for reduced sampling speed DPD, they paid very little attention to the practical implementability of their different approaches. It was evident from chapter 2 that the sub-sampled receivers calibration issue was barely discussed. Furthermore, receiver calibration will have a big impact on the DPD performance, especially for the wide-band signal scenarios we are addressing here.

As a result, we propose a modified DPD settings that allow for applying receiver calibration on the sub-sampled data. We suggest to move the receiver calibration from the receiver side to the transmitter side where the full-sampled data is still available. This move is only valid under a certain set of assumptions that are not always necessary cor-

rect, especially at the first training iteration. Thus, we had to perform the DPD training part carefully—using small update step size—to avoid algorithm divergence.

Besides, we demonstrate receiver and transmitter calibration routine to identify the receiver and transmitter transfer functions using sub-sampled ADCs. Identifying these transfer functions using sub-sampled received data is one another practical aspect that we tackle. Using this proposed routine, we prove that receive sampling speed can be reduced to a very low sampling rate, much below the conventional Nyquist rate.

Finally, we use all these proposed methods to apply DPD on 200 and 320MHz, LTE and 16QAM. For both bandwidths, we demonstrate a sub-sampling linearization capacity similar to the conventional full-rate DPD. It was evident form the measurement results that receiver calibration plays an important role in the DPD performance. The measurements presented show that when we do not apply receiver calibration, the DPD usually diverges, or in best case achieves worst linearization performance (in terms of EVM and ACPR). For all the tested measurement cases, we achieve and EVM below 3% and ACPR better than 45dBc using different reduced sampling speed. Our best results are an EVM of 2% and ACPR of 49dBC for 200MHz 16QAM signal using a real sampling speed of 500Msps.

5.1   Future Work

A theoretical limit on the minimum sampling speed possible must be devised. In this work we showed DPD using sub-rate factor of 2 and 4. However, further reduction in the sampling speed has yet to be investigated. We expect that two main factors will be affecting the DPD performance with low sampling speed values (high sub-rate factor): time alignment and hardware distortion near DC. Time alignment using under-sampled data might not yield accurate results. Hence, we want to investigate a two step time alignment: coarse using frequency domain approach, and then fine time alignment using cross-correlation. Regarding hardware distortion near DC, for high sub-rate factors, the receiver IF frequency moves closer to the DC. Near DC, coupling capacitors introduces distortion that is hard to calibrate.

It was evident from the measurements results that when we increase the model memory order, DPD diverges; therefore, we had to use simple low order DPD models. As a result, we still want to investigate different regularization techniques that might reduce the divergence behavior of the DPD algorithm.

Finally, this work aims to reduce the TOR ADC consumed power by reducing the sampling speed. Another factor that reduces the ADC power is the number of quantization

bits [29]. ADC power, in general, grows exponentially with the number of bit (in contrast to the sampling speed, where the power in linearly related to the sampling speed). Thus, reducing the number of ADC bits has a substantial impact on reducing the ADC power. Similar to this work, the DPD and calibration aspect using low-bits data must be investigated.

References

[1] 83018a microwave system amplifier, 2 ghz to 26.5 gh. Avilable: `http://www.keysight.com/en/pd-1000001781%3Aepsg%3Apro-pn-83018A/microwave-system-amplifier-2-ghz-to-265-ghz?cc=US&lc=eng` [Online].

[2] M9703a axie 12-bit high-speed digitizer/wideband digital receiver. Avilable: `http://www.keysight.com/en/pd-2031289-pn-M9703A/Wideband-Digital-Receiver?cc=US&lc=eng` [Online].

[3] IMT traffic estimates for the years 2020 to 2030. report itu-r m.2370-0. Avilable: `https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf` [Online], Jul 2015.

[4] IMT vision framework and overall objectives of the future development of imt for 2020 and beyond. recommendation itu-r m.2083-0. Avilable: `https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf` [Online], Sep 2015.

[5] Technical feasibility of imt in bands above 6 ghz. report itu-r m.2376-0. Avilable: `https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2376-2015-PDF-E.pdf` [Online], Jul 2015.

[6] Cisco visual networking index: Forecast and methodology, 20162021. Avilable: `http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf` [Online], Jun 2017.

[7] S. A. Bassam, S. Boumaiza, and F. M. Ghannouchi. Block-wise estimation of and compensation for i/q imbalance in direct-conversion transmitters. *IEEE Transactions on Signal Processing*, 57(12):4970–4973, Dec 2009.

[8] W. C. Black and D. A. Hodges. Time interleaved converter arrays. *IEEE Journal of Solid-State Circuits*, 15(6):1022–1029, Dec 1980.

[9] W. C. Black and D. A. Hodges. Time interleaved converter arrays. *IEEE Journal of Solid-State Circuits*, Sep 2015.

[10] R. N. Braithwaite. Closed-loop digital predistortion (dpd) using an observation path with limited bandwidth. *IEEE Transactions on Microwave Theory and Techniques*, 63(2):726–736, Feb 2015.

[11] R. N. Braithwaite. A comparison of indirect learning and closed loop estimators used in digital predistortion of power amplifiers. In *2015 IEEE MTT-S International Microwave Symposium*, pages 1–4, May 2015.

[12] R Neil Braithwaite. General principles and design overview of digital predistortion. *Digital Processing for Front End in Wireless Communication and Broadcasting*, pages 143–191, 2011.

[13] T. Chen, H. Kim, and Y. Yang. Energy efficiency metrics for green wireless communications. In *2010 International Conference on Wireless Communications Signal Processing (WCSP)*, pages 1–6, Oct 2010.

[14] Steve C Cripps. *RF power amplifiers for wireless communications.* Artech House, 2006.

[15] L. Ding, F. Mujica, and Z. Yang. Digital predistortion using direct learning with reduced bandwidth feedback. In *2013 IEEE MTT-S International Microwave Symposium Digest (MTT)*, pages 1–3, June 2013.

[16] B. Fehri, S. Boumaiza, and E. Sich. Crest factor reduction of inter-band multi-standard carrier aggregated signals. *IEEE Transactions on Microwave Theory and Techniques*, 62(12):3286–3297, Dec 2014.

[17] F. M. Ghannouchi and O. Hammi. Behavioral modeling and predistortion. *IEEE Microwave Magazine*, 10(7):52–64, Dec 2009.

[18] P. L. Gilabert, G. Montoro, D. Lpez, N. Bartzoudis, E. Bertran, M. Payar, and A. Hourtane. Order reduction of wideband digital predistorters using principal component analysis. In *2013 IEEE MTT-S International Microwave Symposium Digest (MTT)*, pages 1–7, June 2013.

[19] H. Huang, P. Mitran, and S. Boumaiza. Digital predistortion function synthesis using undersampled feedback signal. *IEEE Microwave and Wireless Components Letters*, 26(10):855–857, Oct 2016.

[20] Ian Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.

[21] 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.

[22] Y. Ma, Y. Yamao, Y. Akaiwa, and K. Ishibashi. Wideband digital predistortion using spectral extrapolation of band-limited feedback signal. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 61(7):2088–2097, July 2014.

[23] F. Mkadem and S. Boumaiza. Nonlinear system behavioral modeling using reduced transmitter observation receiver bandwidth. In *83rd ARFTG Microwave Measurement Conference*, pages 1–3, June 2014.

[24] Mkadem, Farouk. *Behavioral Modeling and Digital Predistortion of Wide- and Multi-Band Transmitter Systems*. PhD thesis, 2014.

[25] H. Paaso and A. Mammela. Comparison of direct learning and indirect learning predistortion architectures. In *2008 IEEE International Symposium on Wireless Communication Systems*, pages 309–313, Oct 2008.

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

[27] M. Ben Rejeb, J. M. Moreau, J. Dunsmore, J. P. Teyssier, and S. Boumaiza. Rf receiver characterization and spurious-free dynamic range enhancement for pim/weakly nonlinear device measurements. *IEEE Transactions on Microwave Theory and Techniques*, 64(11):4030–4038, Nov 2016.

[28] Valeria Teppati, Andrea Ferrero, and Mohamed Sayed. *Modern RF and microwave measurement techniques*. Cambridge University Press, 2013.

[29] H. Wang, G. Li, C. Zhou, W. Tao, F. Liu, and A. 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, July 2017.

[30] Z. Wang, W. Chen, G. Su, F. M. Ghannouchi, Z. Feng, and Y. Liu. Low feedback sampling rate digital predistortion for wideband wireless transmitters. *IEEE Transactions on Microwave Theory and Techniques*, 64(11):3528–3539, Nov 2016.

[31] David Wisell and Peter Händel. *Simultaneous measurement of transmitter and receiver amplitude and phase ripple.* Högskolan i Gävle, 2009.

[32] C. Yu, L. Guan, E. Zhu, and A. Zhu. Band-limited volterra series-based digital predistortion for wideband rf power amplifiers. *IEEE Transactions on Microwave Theory and Techniques*, 60(12):4198–4208, Dec 2012.

[33] D. Zhou and V. E. DeBrunner. Novel adaptive nonlinear predistorters based on the direct learning algorithm. *IEEE Transactions on Signal Processing*, 55(1):120–133, Jan 2007.