# DIGITAL PREDISTORTION FOR HYBRID PRECODING ARCHITECTURE IN MILLIMETER-WAVE MASSIVE MIMO SYSTEMS

Han Yan and Danijela Cabric

Electrical Engineering Department University of California, Los Angeles Email: yhaddint@ucla.edu, danijela@ee.ucla.edu

#### ABSTRACT

Millimeter-wave (mmWave) systems require massive antennas at both transmitter and receiver to reach desirable link budget. Analog-digital hybrid beamforming is a promising architecture since it significantly reduces the hardware cost while approaches the performance of digital beamforming. However, the nonlinear power amplifier (PA) introduces intermodulation interference and degrades the spectral efficiency. In this work, we use memory polynomial model for modelling mmWave PA and study the inter-beam and inter-symbol interference in hybrid precoding architecture. We also propose a digital predistortion (DPD) algorithm to mitigate interference. The proposed DPD reduces interference power by an order of magnitude and compared with system without predistortion, the proposed system can achieve up to 60% spectral efficiency.

*Index Terms*— Millimeter-wave, phased array, beamforming, predistortion, power amplifier

## 1. INTRODUCTION

Millimeter-wave (mmWave) communication is a promising technology for the future 5G cellular network [1]. Due to the total 10 GHz spectrum to be assigned for mobile systems in the 28, 38, and 72 GHz band [2], 5G networks can support 100 times higher data rate than the current 4G. As shown in both theory and prototypes, mmWave system requires beamforming (BF) with large antenna arrays at both base station and user equipments (UE) to combat severe propagation loss in mmWave band [3, 4]. The massive size of antenna array makes traditional digital beamforming less attractive due to its high cost. The analog-digital hybrid beamforming architecture gains attentions since it potentially achieves optimal capacity in mmWave band while maintaining reasonable hardware cost and power consumption [5].

Many works studied the the hybrid precoding and combining algorithm with assumptions of ideal hardwares [5, 6]. Only few considered the impact of hardware imperfect to hybrid beamforming. In [7] the impact of finite phase shifting resolution in hybrid beamforming is studied. In [8] the impact



**Fig. 1**. The analog-digital hybrid precoding architecture in the BS. Mathematical symbols are marked as blue.

of phase shifting uncertainty is studied and robust precoding algorithm is presented. In [9] the digital predistortion is used to compensate nonlinear effect of mmWave power amplifiers, but the hybrid beamforming architecture has not been considered.

In this work, we study the impact of power amplifier (PA) nonlinearity on hybrid precoding and propose a digital predistortion (DPD) algorithm. The DPD reduces interference power by an order of magnitude and achieves up to 60% higher spectral efficiency at mmWave BS. The hybrid precoding architecture and millimeter-wave power amplifier model are introduced in Section II. In Section III, we study the interference that arises due to PA nonlinearity. We present the DPD algorithm in Section IV followed by numerical results in Section V. We conclude this work in Section VI.

# 2. SYSTEM MODEL

In this section, we introduce the analog-digital hybrid beamforming system and nonlinear model of millimeter-wave power amplifier (PA).

Consider a system where one mmWave BS transmits M data streams towards M UEs. The transmitter uses analogdigital hybrid precoding architecture as shown in Fig. 1. The M data streams are precoded in the DSP, up-converted to radio frequency (RF), precoded by RF phase shifters, and then amplified by  $N_t$  PAs ( $N_t \gg M$ ) in the antenna array. The precoded signal is expressed as

$$\mathbf{y}[n] = \mathbf{F}_{\mathsf{RF}} \mathbf{F}_{\mathsf{BB}} \mathbf{s}[n]. \tag{1}$$

where  $\mathbf{s}[n] \in \mathbb{C}^M$  and  $\mathbf{y}[n] \in \mathbb{C}^{N_t}$  are the symbols vector before and after precoding, respectively, n is the time index, and  $\mathbf{F}_{BB} \in \mathbb{C}^{M \times M}$  and  $\mathbf{F}_{BB} \in \mathbb{C}^{N_t \times M}$  are the digital and analog precoding matrix, respectively. Since the analog precoding procedure is implemented via phase shifters, each element in the  $\mathbf{F}_{RF}$  has unit amplitude.

Each precoded symbol is amplified by a PA before transmitted via the corresponding antenna. We consider a baseband memory polynomial model (MPM) with linear and cubic terms to model the behavior of mmWave PA [10], and the output at the PA is expressed as

$$z_k[n] = \sum_{d=0}^{D-1} \left( \beta_d^{(1)} y_k[n-d] + \beta_d^{(3)} y_k[n-d] \left| y_k[n-d] \right|^2 \right),$$
(2)

where  $y_k[n]$ , the  $k^{th}$  element of  $\mathbf{y}[n]$ , is the input symbols of  $k^{th}$  PA at instance n and  $z_k[n]$  is the corresponding output.  $\beta_d^{(1)}$  and  $\beta_d^{(3)}$  are the linear and cubic parameter of the PA at delay d. They are assumed to be identical for all PAs and are known to the BS. D is the memory depth of the model.

We assume each mmWave UE is equipped with  $N_r$  antennas and has only one RF chain for the cost concern. Thus the UE uses analog combining. Components in the combining vector  $\mathbf{w}_m$  have unit amplitude. The combined symbol at UE m is expressed as

$$r_m[n] = \mathbf{v}_m^H \mathbf{H}_m \mathbf{z}[n] + \mathbf{v}_m^H \mathbf{q}[n], \qquad (3)$$

where  $\mathbf{z}[n] = [z_1[n], \dots, z_M[n]]^T$  represents the transmitted symbols at BS arrays,  $\mathbf{v}_m$  is the combining vector,  $\mathbf{H}_m$  is the channel matrix between BS and UE m, and  $\mathbf{q}[n] \sim C\mathcal{N}(0, \sigma_q^2 \mathbf{I})$  is the noise.

Since mmWave channel has limited scattering [3], the narrowband channel matrix has low rank and is expressed as

$$\mathbf{H}_{m} = \frac{1}{\sqrt{N_{t}N_{r}}} \sum_{l=1}^{L} \alpha_{m,l} \mathbf{a}_{r}(\phi_{m,l}) \mathbf{a}_{t}(\theta_{m,l})^{H}, \qquad (4)$$

where *L* represents the number of propagation paths,  $\alpha_{m,l} \in \mathbb{C}$  is the propagation loss in the descending order, i.e.,  $|\alpha_{m,1}| \geq |\alpha_{m,2}| \cdots \geq |\alpha_{m,L}|$ , and  $\mathbf{a}_t(\theta_{m,l})$  and  $\mathbf{a}_r(\phi_{m,l})$  are angular channel response vectors at the BS and UE. We consider uniform linear arrays in both BS and UEs and antenna spacing is half of the wavelength. The channel response vectors are expressed as

$$\mathbf{a}_{r}(\phi_{m,l}) \triangleq \left[1, e^{j\pi \sin(\phi_{m,l})}, \cdots, e^{j\pi N_{r} \sin(\phi_{m,l})}\right]^{T} \text{ and } \\ \mathbf{a}_{t}(\theta_{m,l}) \triangleq \left[1, e^{j\pi \sin(\theta_{m,l})}, \cdots, e^{j\pi N_{t} \sin(\theta_{m,l})}\right]^{T},$$

where  $j = \sqrt{-1}$  is the unit imaginary number,  $\theta_{m,l}$  and  $\phi_{m,l}$  are the angle of departure (AOD) and angle of arrival (AOA) corresponding to propagation path l, respectively.  $\phi_{m,l}, \theta_{m,l} \in [-\pi/2, \pi/2], \forall m, l.$ 

We use the achievable rate of M data stream,  $C = \sum_{m=1}^{M} \log_2(1 + \text{SINR}_m)$ , as the performance metric, where  $\text{SINR}_m$  represents the signal to interference and noise ratio of  $r_m[n]$ .

#### 3. ANALYSIS OF INTERFERENCE FROM NONLINEAR PA

In this section, we study the interference at UE introduced by the nonlinear PA. Assuming CSI is known to both BS and UE, and transmission angular response are orthogonal.

$$\mathbf{a}(\theta_{m_1,l_1})^H \mathbf{a}(\theta_{m_2,l_2}) = \begin{cases} N_t, & \text{if } l_1 = l_2, m_1 = m_2 \\ 0, & \text{otherwise.} \end{cases}$$

Such assumption ignores the sidelobe of each transmission beam and is valid for sufficiently large  $N_t$ . As such, the optimal analog precoder at BS is  $\mathbf{F}_{RF} = [\mathbf{a}_t(\theta_{m,1}), \cdots, \mathbf{a}_t(\theta_{M,1})]$ is optimal. The digital precoder is water-filling based power allocation of each data stream. The optimal combining weight at UE is  $\mathbf{v}_m = \mathbf{a}_r(\phi_{m,1})$ . We use  $\theta_m$  and  $\alpha_m$  for  $\theta_{m,1}$  and  $\alpha_m$  for clarity.

Under the optimal transmitter beamforming, (2) becomes

$$z_{k}[n] = \sum_{d=0}^{D-1} \sum_{m=1}^{M} \beta_{d}^{(1)} \bar{s}_{m}[n-d] e^{j(k-1)\pi \sin(\theta_{m})} + \sum_{d=0}^{D-1} \sum_{m_{1}=1}^{M} \sum_{m_{2}=1}^{M} \sum_{m_{3}=1}^{M} \left\{ \beta_{d}^{(3)} \bar{s}_{m_{1}}[n-d] \bar{s}_{m_{2}}[n-d] \cdot \bar{s}_{m_{3}}^{*}[n-d] e^{j(k-1)\pi \left[ \sin(\theta_{m_{1}}) + \sin(\theta_{m_{2}}) - \sin(\theta_{m_{3}}) \right]} \right\}, \quad (5)$$

where  $\bar{s}_m[n]$  represents the power scaled symbol  $s_m[n]$ , which is the  $m^{th}$  element of  $s_n$ .

According to (5), the nonlinear PAs not only steer beams at original angles  $\theta_m$  as desired, but also generate spurious beams at new angles. In this work, we ignore the incidence where spurious beam overlaps with original beam, i.e.,  $\sin(\theta_{m_1}) + \sin(\theta_{m_2}) - \sin(\theta_{m_3}) \neq \sin(\theta_m)$  except  $m_1 = m, m_2 = m_3$  or  $m_2 = m, m_1 = m_3$ . As such, the combined symbol by combining vector  $\mathbf{v}_m$  is expressed as

$$r_m[n] = \alpha_m \left( \bar{s}_m[n] + \mathrm{ISI}_m[n] + \mathrm{IBI}_m[n] \right) + \mathbf{v}_m^H \mathbf{q}[n], \quad (6)$$

where  $\beta_0^{(1)} \bar{s}_m[n]$  is the desired symbol in  $m^{th}$  stream. The inter-beam interference (IBI), expressed as

$$\operatorname{IBI}_{m}[n] = 2 \sum_{d=0}^{D-1} \sum_{j=1, j \neq m}^{M} \beta_{d}^{(3)} \bar{s}_{m}[n-d] \left| \bar{s}_{j}[n-d] \right|^{2}, \quad (7)$$



**Fig. 2**. The proposed DPD algorithm in mmWave BS with hybrid precoding. Mathematical symbols are marked as blue.

is the multiuser interference from symbols in other beam. The inter-symbol interference (ISI), expressed as

$$ISI_{m}[n] = (\beta_{0}^{(1)} - 1)\bar{s}_{m}[n] + \sum_{d=1}^{D-1} \beta_{d}^{(1)}\bar{s}_{m}[n-d] + \sum_{d=0}^{D-1} \beta_{d}^{(3)}\bar{s}_{m}[n-d] \big| \bar{s}_{m}[n-d] \big|^{2}, \qquad (8)$$

comes from past symbols of the same data stream due to the memory effect of PA.  $\mathbf{v}_m^H \mathbf{q}[n]$  is the noise. The arise of interference degrades SINR and such impact cannot be mitigated in the UE. We propose a digital compensation algorithm as a solution in the next section.

## 4. PROPOSED PREDISTORTION ALGORITHM

In this section, we propose a DPD algorithm at the BS in order to mitigate interference introduced from nonlinear PA.

The DPD is presented in Fig. 2 as signal flow illustration and **Algorithm 1**. Following [10, 11], we use memory polynomial for the DPD algorithm in system with hybrid precoding architecture. The algorithm firstly uses power scaled symbol  $\overline{\mathbf{s}}[n]$  to formulate an extended vector  $\mathbf{u}[n]$  as follows

$$\mathbf{u}[n] = \left[\bar{\mathbf{s}}^{T}[n], \bar{\mathbf{s}}^{T}[n] \otimes \bar{\mathbf{s}}^{T}[n] \otimes \bar{\mathbf{s}}^{H}[n], \cdots, \bar{\mathbf{s}}^{T}[n-D+1], \\ \bar{\mathbf{s}}^{T}[n-D+1] \otimes \bar{\mathbf{s}}^{T}[n-D+1] \otimes \bar{\mathbf{s}}^{H}[n-D+1] \right]^{T},$$
(9)

where  $\otimes$  represents kronecker product. Note that  $\mathbf{u}[n] \in \mathbb{C}^{D(M+M^3)}$ . A linear predortortion of  $\mathbf{u}[n]$  is used to formulate predistorted symbols  $\mathbf{x}[n]$  as output of DPD. It is expressed as

$$\mathbf{x}[n] = \mathbf{W}^H \mathbf{u}[n] \tag{10}$$

where  $\mathbf{W} \in \mathbb{C}^{D(M+M^3) \times M}$  is the weight.

The system uses a training phase to adaptively learn the optimal weight  $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_M]$  to minimize ISI and IBI. The training phase,  $\mathbf{F}_{BB}$  and  $\mathbf{F}_{RF}$  are set the same as in data transmission, and training symbols  $\mathbf{s}_n$  is used. Besides, an alternative feedback loop is used, which consists of phase

shifter network  $\mathbf{F}_{RF}^{H}$ , to allow the system get the an estimation of received symbols at UE,  $\hat{\mathbf{r}}[n] = \mathbf{F}_{RF}^{H} \mathbf{z}[n]$ . Each of its component is expressed as

$$\hat{r}_{m}[n] = \sum_{d=0}^{D-1} \left( \beta_{d}^{(1)} x_{m}[n-d] + \beta_{d}^{(3)} x_{m}[n-d] \big| x_{m}[n-d] \big|^{2} + 2 \sum_{j=1, j \neq m}^{M} \beta_{d}^{(3)} x_{m}[n-d] \big| x_{j}[n-d] \big|^{2} \right).$$
(11)

Note that (11) is a noiseless version of (6) with  $\bar{s}_m[n]$  replaced by  $x_m[n]$ ,  $m^{th}$  component of  $\mathbf{x}[n]$ .

The training procedure uses least-mean-square (LMS) filter to reduce mean square of interference symbol  $e_m[n] \triangleq$  $ISI_m[n] + IBI_m[n]$ , which can also be evaluated by taking difference between ideally amplified symbols  $\mathbf{s}[n]$  and  $\hat{\mathbf{r}}[n]$ . The instantaneous knowledge of  $\sum_{m=1}^{M} ||e_m[n]||^2$  gives the following LMS updating equation

$$\mathbf{w}_{m}[n+1] = \mathbf{w}_{m}[n] - \mu \nabla_{\mathbf{w}_{m}[n]} \|\overline{\mathbf{s}}[n] - \hat{\mathbf{r}}[n]\|^{2}$$
$$= \mathbf{w}_{m}[n] + \mu \sum_{k=1}^{M} \left( e_{k}^{*}[n] \sum_{d=0}^{D-1} \sum_{j=1}^{M} \frac{\partial \hat{r}_{k}[n]}{\partial x_{j}[n-d]} \frac{\partial x_{j}[n-d]}{\partial \mathbf{w}_{m}[n]} \right).$$
(12)

where  $\mathbf{w}_n^{(m)}$  is the DPD weight for  $m^{th}$  stream at instance n.  $\mu$  is the stepsize of LMS updating. According to (11), the derivative  $\partial r_k[n]/\partial x_l[n-d]$  is

$$\frac{\partial \hat{r}_k[n]}{\partial x_j[n-d]} = \begin{cases} \beta_d^{(1)} + 2\beta_d^{(3)} \sum_{k=1}^M |x_k[n-d]|^2, & k=j, \\ 2\beta_d^{(3)} x_k[n-d] x_j^*[n-d], & k\neq j. \end{cases}$$

The derivative  $\partial x_j[n-d]/\partial \mathbf{w}_m[n]$  is  $\mathbf{u}[n-d]$  when l=m, and a zero vector otherwise.

Algorithm 1 LMS-based DPD for hybrid precoding

**Input:** Digitally precoded symbols  $\bar{\mathbf{s}}[n]$ , first N as training symbols

**Output:** Predistorted symbols  $\mathbf{x}_n$ 

- 1: Initialize  $\mathbf{w}_m[1] = [1, \mathbf{0}_{D(M+M^3)-1}]^T, \forall m$
- 2: for n = 1 : N do
- 3: Transmit  $\mathbf{x}[n]$  as output of DPD training via (9) and (10), receive feedback symbol  $\hat{\mathbf{r}}[n]$
- 4: Use symbol error vector  $\mathbf{s}[n] \hat{\mathbf{r}}[n]$  to update  $\mathbf{w}_m[n+1]$  for all m via (12)
- 5: end for
- 6: Use  $\hat{\mathbf{W}} = [\mathbf{w}_1[N], \cdots, \mathbf{w}_M[N]]$  as  $\mathbf{W}$ , and predistort data symbols  $\bar{\mathbf{s}}[n], n > N$  via (9) and (10)

#### 5. NUMERICAL RESULTS

In this section, we present numerical results of the proposed DPD in a mmWave system using hybrid precoding.



**Fig. 3**. The learning curve of LMS-based DPD weight training in mmWave BS.

For simulating the DPD training performance, we use Gaussian random sequence  $s_m[n]$  in both training and data transmission. Truncation is made for peak-to-average power ratio being 5 dB. To the best knowledge of the authors, there is not work reporting behavior model for mmWave PA. As a compromise, we use parameters of microwave PA reported in [11]. We choose memory depth D = 2 for both PA modeling and PDP algorithm.

The learning curve and mean square error  $\sum_{m=1}^{M} ||e_m[n]||^2$  (MSE) performance of LMS-based DPD training is presented in Fig. 3. Results are averaged over 1000 realizations for smoothness. Note that for different number of RF chain M, the convergence speed is different. We manually select stepsize for fast convergence. As observed, the MSE converges to same level for different M. The convergence can be reached in a few hundreds of samples and in convergence the MSE is an order of magnitude smaller than pre-compensation scenario (N = 0).

The achievable capacity C is presented in Fig. 4 for different SNR and number of RF chains M. The SNR values for all M UEs  $N_t N_r \alpha_m \mathbb{E} || \bar{s}_m [n ||^2) / \sigma_q^2$  are set to be identical. In serving UE with the low SNR (SNR = 10 dB), even for system with ideal PAs increasing M provide little improvements in terms of spectral efficiency. Because adding extra RF chains (beams) splits transmitted power, it results in SINR decreases for all UE. Although PA nonlinearity introduces extra degradation and can be compensated by the proposed DPD, the difference is small. In serving UEs with high SNR (SNR = 25 dB), the impact of IBI and ISI is significant, especially for large M. If interference remains uncompensated, a clear performance loss can be observed. The proposed DPD reduces interference power by around 10 dB. Admittedly, certain per-



**Fig. 4**. Spectral efficiency of mmWave BS using hybrid precoding with varying number of RF chain.

formance gap remains between ideal system and the one using DPD due to the existence of residual interference. Once adopting the proposed DPD, the spectral efficiency increases by 60% for system equipped with large number of RF chain operating in high SNR region (M = 8 and 25 dB SNR).

## 6. CONCLUSION AND FUTURE WORKS

Although hybrid precoding architecture approaches optimal channel capacity, the interference caused by nonlinear power amplifier potentially degrades its performance. We studied the inter-symbol and inter-beam interference using memory polynomial model of mmWave power amplifier. We proposed a linear digital predistortion algorithm that reduces the interference. It uses least-mean-square filter to adaptively estimate optimal compensation weights. The DPD reduces interference by an order of magnitude and the post-compensation spectral efficiency increases by up to 60%.

In the future, study of DPD for hybrid precoding BS with non-orthogonal transmission beams will be considered. A scenario where UEs also use hybrid receiving beamforming is another important extension.

## 7. REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [2] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk,

P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152–1163, June 2014.

- [3] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, May 2013.
- [4] W. Hong, K. H. Baek, Y. Lee, Y. Kim, and S. T. Ko, "Study and prototyping of practically large-scale mmwave antenna systems for 5g cellular devices," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 63–69, Sept. 2014.
- [5] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [6] R. W. Heath, N. Gonzlez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave mimo systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [7] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial ue discovery in millimeter-wave mimo systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.
- [8] S. H. Wu, L. K. Chiu, K. Y. Lin, and T. H. Chang, "Robust hybrid beamforming with phased antenna arrays for downlink sdma in indoor 60 ghz channels," vol. 12, no. 9, pp. 4542–4557, Sept. 2013.
- [9] P. Y. Wu, Y. Liu, B. Hanafi, H. Dabag, P. Asbeck, and J. Buckwalter, "A 45-ghz si/sige 256-qam transmitter with digital predistortion," in 2015 IEEE MTT-S International Microwave Symposium, May 2015, pp. 1–3.
- [10] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417– 1418, Nov. 2001.
- [11] Lei Ding, G. T. Zhou, D. R. Morgan, Zhengxiang Ma, J. S. Kenney, Jaehyeong Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, Jan. 2004.