

A robust digital predistortion algorithm for power amplifiers

Yun Bu^{a)}, Tian-Qian Li, and Yong-Qiang Chen

*School of Electrical and Information Engineering, Xihua University,
Cheng Du, 610039, China*

a) buyun_xihua@sina.cn

Abstract: The Volterra series and polynomial based indirect learning digital predistortion methods suffer from the noise at the power amplifiers output. We propose an improvement using two robust non-polynomial functions to substitute the higher items, and utilizing the robust maximum correntropy criterion to replace the minima square error criterion. The convergence of the method is proved and simulations show its robust performance for noisy signals.

Keywords: digital predistorter, power amplifier, polynomial model, maximum correntropy

Classification: Microwave and millimeter wave devices, circuits, and systems

References

- [1] A. Zhu, J. C. Pedro and T. J. Brazil: IEEE Trans. Microw. Theory Tech. **54** [12] (2006) 4323.
- [2] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt and J. Pastalan: IEEE Trans. Signal Process. **54** [10] (2006) 3852.
- [3] M. Rawat, K. Rawat and F. M. Ghannouchi: IEEE Trans. Microw. Theory Tech. **58** [1] (2010) 95.
- [4] O. Gorbachov: Microwaves & RF **39** [8] (2000) 137.
- [5] S. Oualid Hammi: IEEE Trans. Microw. Theory Tech. **55** [6] (2007) 1382.
- [6] Y.-J. Liu, W. Chen, J. Zhou, B.-H. Zhou, F. M. Ghannouchi and Y.-N. Liu: IEEE Trans. Microw. Theory Tech. **60** [11] (2012) 3559.
- [7] B. Chen and J. C. Principe: IEEE Signal Process. Lett. **19** [8] (2012) 491.

1 Introduction

The limited spectrum in modern wireless communication systems motivates wide applications of efficient modulated techniques, such as orthogonal frequency division multiplexing (OFDM), which has dramatically fluctuant envelopes and is sensitive to nonlinearity. Therefore linearization techniques are introduced to correct the distortion causing mainly by power amplifiers (PA), which must work close to saturation to achieve high-power efficiency. As a popular technique, digital predistortion (DPD) method employs a predistorter (PD) before the PA, to compensate the nonlinear effects. Traditional DPD models include Volterra series [1],

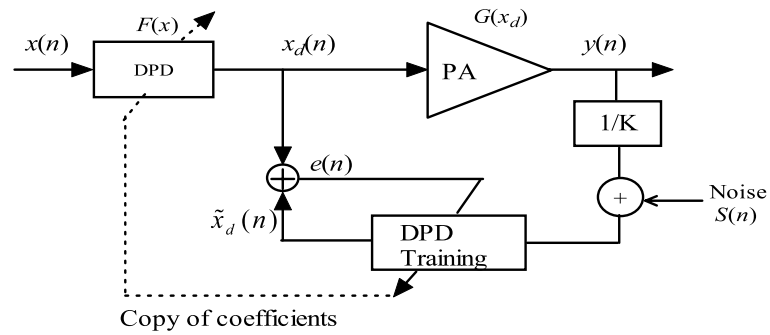


Fig. 1. Indirect learning architecture for DPD

polynomial models [2], lookup table, and neural networks [3]. An indirect learning DPD is shown in Fig. 1.

As shown in Fig. 1, in DPD training, the attenuated PA output signal $y(n)$ feeds into DPD, and the instantaneous PA input signal $x(n)$ is the desired DPD output signal, namely $x_d(n)=x(n)$. The process is to achieve the DPD model. $s(n)$ denotes external noise introduced in feedback path. After DPD model convergences, it is inserted before PA, and $x(n)$ becomes DPD input signal. Let $G(\cdot)$ and $F(\cdot)$ represent PA and PD model, the linearization is to make output

$$y(n) = G(F(x(n))) = Kx(n) \quad (1)$$

where $x(n)$ is transmitted symbols, K denotes the desired PA gain.

However, noise will cause the DPD coefficients dispersion. Noise includes PA output noise and external noise $s(n)$. The former mainly consists of PA low inherent noise and intermodulation distortion (IMD) products [4]. External noise mainly comes from electromagnetic interference and measurement errors in feedback path, like I/Q imbalance and DC-offset in quadrature demodulation (QDEMOD), A/D noise and so on. The frequency dependent noise or error like IMD and I/Q imbalance error can be easy modeled using Volterra series or polynomial for it dependents on input signals, then they can be compensated by DPD, as can be shown in the latter simulations. But the other noise like A/D noise has not a similar mathematic express with DPD model, and then it will not be eliminated but will accumulate by iteration. To reduce coefficients dispersion, O. Hammi *et al* proposed normalizing DPD input samples [5] to attempt higher order polynomials, but noise was disregarded. Y. J. Liu designed a modified least squares method [6] to enhance the DPD robustness, however which needs an extra measurement to achieve the response of feedback path to the PA output $y(n)$. In this article, we improve the algorithm by introducing robust basis functions and a new optimization criterion.

2 Using non-polynomial in DPD system

The popular Volterra series and polynomial DPD are sensitive to noise, because of their high order polynomial derivatives, $f'(x) = px^{p-1}$, where $f(x)$ is a P -order polynomial. To enhance the DPD robustness, we propose a hybrid method using non-polynomial functions with bounded derivates as substitutes for higher order polynomials:

$$\begin{aligned} f_1(x) &= \tanh(x) \\ f_2(x) &= \exp(-x^2/2) \end{aligned} \quad (2)$$

In independent component analysis, the two functions have been used as substitutes of higher order cumulants, whose counterparts are just the higher-order polynomial. The robustness DPD method can be expressed as:

$$\begin{aligned} x_d(n) &= \sum_{k=0}^3 \sum_{m=0}^M \omega_{k,m} x(n-m) |x(n-m)|^k \\ &\quad + \sum_{m=0}^M \sum_{i=1}^2 \omega'_{m,i} f_i(x(n-m)) = \omega^H(n) u(n) \\ \omega(n) &= [\omega_{0,0}, \omega_{0,1}, \dots, \omega_{3,M}, \omega'_{0,1}, \omega'_{0,2}, \dots, \omega'_{M,2}]^T \\ u(n) &= [x(n), x(n-1), \dots, x(n-M) |x(n)|^3, \\ &\quad f_1(x(n)), f_2(x(n)), \dots, f_2(x(n-M))]^T \end{aligned} \quad (3)$$

where $x(n)$ and $x_d(n)$ are the DPD input and output signal, respectively, M is the memory depth, $\omega(n)$ and $u(n)$ represent DPD coefficient vector and input vector, $f_i(x)$ is defined as (2). The input signal that feed into non-polynomial functions will be normalized as $x(n)/\max |x|$. In training, input signal $x(n)$ is replaced by $y(n)$.

3 Maximum correntropy criterion (MCC)

In DPD training, the popular minimum square error (MSE) is a second-order statistics that is often not competent to non-Gaussian distribution. Then we use maximum correntropy criterion [7] to cope the DPD complex noise case:

$$V_\sigma(X, X_d) = E[\kappa_\sigma(X - X_d)] = E\left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{|e|^2}{2\sigma^2}\right)\right] \quad (4)$$

where X and X_d are two random variables, $\kappa_\sigma(\cdot)$ is a positive definite kernel, here using Gaussian kernel with the width being σ , $E(\cdot)$ denotes the expectation operator, and error is defined as $e = X - X_d$. When $e=0$, correntropy reaches its maximum value. Hence the strategy is named as maximum correntropy criterion (MCC). To achieve fast convergence, we need to derive the MCC recursive least squares (RLS). Derivative the MCC cost function and let it equals to zero to calculate the model coefficients:

$$\begin{aligned} \omega(n) &= \left\{ \sum_{i=0}^n \lambda^{n-i} u(i) u^H(i) \kappa_\sigma(|e(i)|) \right\}^{-1} \left\{ \sum_{i=0}^n \lambda^{n-i} u(i) x_d^*(i) \kappa_\sigma(|e(i)|) \right\} \\ &= R_u^{-1}(n) r_{ud}(n) \end{aligned} \quad (5)$$

where H indicates the conjugate transposition operation. Let input vector be $u_t(n) = \sqrt{\kappa_\sigma(e(n))} \cdot u(n)$, the MCC RLS becomes a standard RLS. If $e(n)$ is small, $u_t(n) \approx \kappa_\sigma(e(n)) \cdot u(n)$ and the Kalman gain of MCC RLS can be approximately expressed as:

$$k(n) = \frac{P(n-1)u(n)\sqrt{\kappa_\sigma(|e(n)|)}}{(\lambda + \kappa_\sigma(|e(n)|))u^H(n)P(n-1)u(n)} \quad (6)$$

$$\approx \frac{P(n-1)u(n)\kappa_\sigma(|e(n)|)}{\lambda + u^H(n)P(n-1)u(n)}$$

Then we can get the simplified MCC-RLS criterion (SMCC-RLS) update rule:

$$\omega(n) = \omega(n-1) + \frac{P(n-1)u(n)}{\lambda + u^H(n)P(n-1)u(n)} \kappa_\sigma(|e(n)|) e^*(n) \quad (7)$$

4 Convergence of the hybrid DPD

MCC and MSE show evident relation when correnentropy is expanded as [7]:

$$\kappa_\sigma(|e|^2) = 1/(\sqrt{2\pi}\sigma) \sum_{n=0}^{\infty} (-1)^n / (2^n n!) E\{|e(n)|^{2n} / \sigma^{2n}\} \quad (8)$$

The equation shows that crrenentropy consists of all even moments. If kernel width is set to a large value, higher order even moments will be repressed. Further more, when error is small, the second-order statistics is dominant, then it is reasonable to analysis the convergence of the hybrid method for noisy signal utilizing the well known RLS convergence conclusion:

$$E\{\varepsilon^H(n)\varepsilon(n)\} = \frac{\sigma_0^2}{n} \text{Tr}(R_u^{-1}) = \frac{\sigma_0^2}{n} \sum_{i=1}^M \frac{1}{\lambda_i} \quad (9)$$

where σ_0^2 is the variance of system error, λ_k is the eigenvalue of R_u , and coefficient error $\varepsilon(n) = \omega_o - \omega(n)$ is the deviation between the optimum coefficient ω_o and the n -th estimated one $\omega(n)$.

For convenience, we only considered a real input signal and let forget factor $\lambda = 1$. Besides this, noise $s(n)$ is assumed independent and with zero mean value and σ_1^2 variance. Then deteriorated input signal is $\tilde{u}(n) = F(x(n) + s(n))$, where $F(x)$ represents basis function set, and the input autocorrelation matrix becomes:

$$\begin{aligned} R_{\tilde{u}} &= E\{\tilde{u}(n)\tilde{u}^T(n)\} \\ &\approx E\{(F(x(n) + s(n))s^T(n))(F'(x(n) + s(n))s^T(n))^T\} \quad (10) \\ &= E\{u(n)u^T(n)\} + E\{F'(x(n))s^T(n)s(n)F'^T(x(n))\} \\ &= E\{u(n)u^T(n)\} + \sigma_1^2 I \cdot E\{F'(x(n))F'^T(x(n))\} \end{aligned}$$

where I denotes the identity matrix. From the equation, we can achieve:

$$E\{\tilde{\varepsilon}^H(n)\tilde{\varepsilon}(n)\} = \frac{\sigma_0^2}{n} \text{Tr}(R_{\tilde{u}}^{-1}) = \frac{\sigma_0^2}{n} \sum \frac{1}{\lambda_i + f_j'^2(x)\sigma_1^2} \quad (11)$$

where $f_j'(x)$ represents the derivative of the j -th basis function.

Because the minimum derivate value of polynomial is less than that of non-polynomial, the minimum eigenvalue of polynomial-based $R_{\tilde{u}}$ is also smaller. This means square coefficients error of polynomial method is larger than that of the proposed hybrid method, namely the latter performs more robust.

5 Simulations and conclusion

To test the proposed algorithm, we constructed a hybrid DPD with 5 memory depth in Matlab. We also employed a 5 order polynomial DPD using RLS criterion as a control study. The PA is a 200 W base station power amplifier (nxp LDMOS) with OFDM signal, and its input and output data is measured and provided by Mathworks in “Amplifier With DPD Block” of Matlab.

The first 10000 samples were used for training and the learning curves were obtained by averaging 100 trials. Fig. 2 is the simulation results for noiseless signal (without external noise $s(n)$), where (a) shows the hybrid method has lower error and is more stability, and (b) illuminates PA output noise caused by nonlinearity has been reduced since IMD can be modeled easily. But external noise $s(n)$ is complex from different sources and we even can't exactly know its statistic distribution for electromagnetic interference may be time varying. So the simulations were conducted to different distributed noises, uniform noise, Gaussian noise and Laplace noise. In simulation, noise was injected into the PA output signals from a low level to a high level. We found that the performances of different DPD change slightly and they were very close to each other when the SNR is higher than 36 dB. If SNR becomes smaller, the performance curves obviously deviated from each other. The following experiments were conducted under 32 dB SNR. Uniform noise is used to simulate the case A/D error is mainly, and Gaussian noise is for the sum of independent identically distribution noises, as shown in Fig. 3 and Fig. 4. The methods can achieve their convergences soon, but the hybrid approach (red) shows more stable with lower mean absolute errors and fluctuate within narrower ranges. For Laplace noise, the DPD performance is very similar with Fig. 3 and Fig. 4. But there still are some obvious distinctions to different noises, as listed in table I.

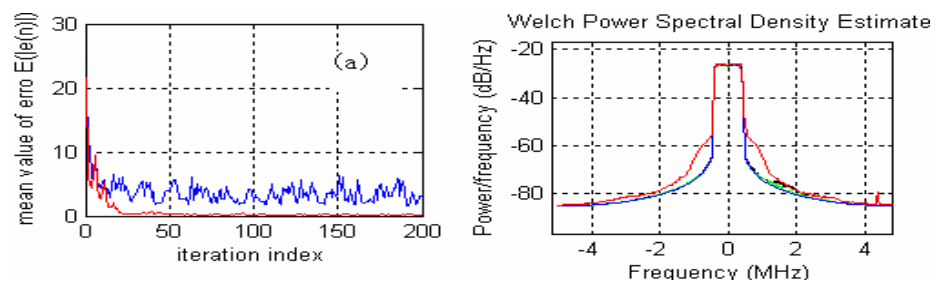


Fig. 2. DPD for noiseless signal (a) Convergence performances of RLS (blue) and SMCC-RLS (red) for noiseless signal. (b) PSD curves, without DPD (red), RLS DPD (black), SMCC-RLS (green) and OFDM source (blue)

From the simulations, we can see that in noiseless case, power spectral density (PSD) of our algorithm is almost coincident with that of the OFDM signal source, and also a little better than that of the polynomial DPD. When noise is introduced, PSD curves raise and deviate from the ideal one. The hybrid method shows more robust under different noises conditions for the method can repress non-Gaussian noise.

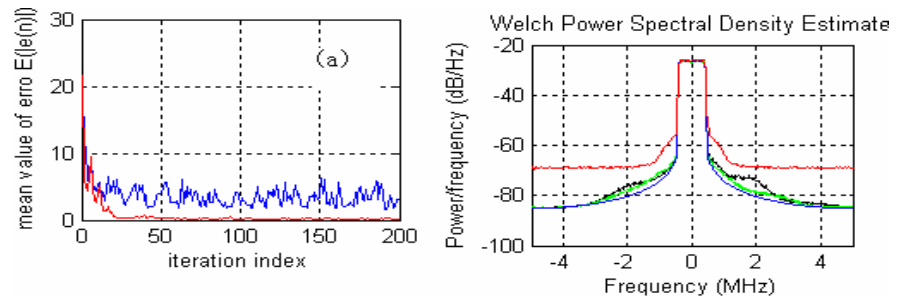


Fig. 3. DPD for uniform noise signal (a) Convergence performances of RLS (blue) and SMCC-RLS (red) for uniform noise. (b) PSD curves, without DPD (red), RLS DPD (black), SMCC-RLS (green) and OFDM source (blue)

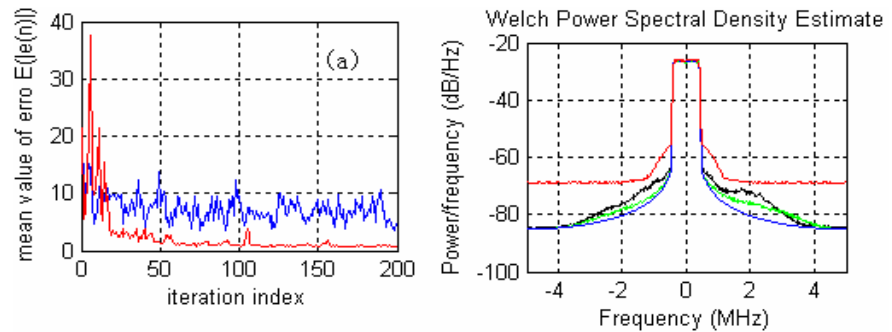


Fig. 4. DPD for Gaussian noise (a) Convergence performances of RLS (blue) and SMCC-RLS (red) for Gaussian noise. (b) PSD curves, without DPD (red), RLS DPD (black), SMCC-RLS (green) and OFDM source (blue)

Table I. Summary of Simulation results

	Polynomial-based RLS DPD		Hybrid SMCC-RLS DPD	
	Mean(e)	Std(e)	Mean(e)	Std(e)
Uniform noise	5.96	1.55	0.82	0.19
Gaussian noise	6.52	1.63	0.79	0.16
Laplace noise	6.97	1.67	0.78	0.11
Noiseless	3.29	1.06	0.19	0.03

Acknowledgments

This work is partially supported by the Education Bureau of Si Chuan Province through Grant No. 12ZB132, and supported by XiHua University through Grants No. Z1120944, No. Z1120942 and No. Z1120943.