DIGITAL BASEBAND PREDISTORTION OF NONLINEAR POWER AMPLIFIERS USING ORTHOGONAL POLYNOMIALS

Raviv Raich, Hua Qian, and G. Tong Zhou *

Georgia Institute of Techonology
School of Electrical and Computer Engineering
Atlanta, GA 30332-0250, USA

ABSTRACT

The polynomial model is commonly used in predistorter design. However, the conventional polynomial model exhibits numerical instabilities when high-order terms are included. In this paper, we introduce a novel set of orthogonal polynomial basis functions for predistorter modeling. Theoretically, the conventional and the orthogonal polynomial models are "equivalent" and thus should have the same performance. In practice, however, the two approaches can perform quite differently in the presence of quantization noise and with finite precision processing. Simulation results show that the orthogonal polynomials can alleviate the numerical instability problem associated with the conventional polynomials and generally yield better predistortion linearization performance.

1 INTRODUCTION

Power amplifier (PA) is a major source of nonlinearity in a communication system. To increase efficiency, PAs are often driven into their nonlinear region, thus causing spectral regrowth (broadening) as well as in-band distortion. PA linearization is often necessary to suppress spectral regrowth, contain adjacent channel interference, and to reduce bit error rate (BER).

Among all linearization techniques, digital baseband predistortion is one of the most cost effective. A predistorter, which (ideally) has the inverse characteristic of the PA, is used to compensate for the nonlinearity in the PA. To linearize a memoryless nonlinear PA, the polynomial model is a common choice and is widely used in predistorter modeling [1, Sec. 3.3]. In practice, however, the polynomial model may experience numerical difficulties when high-order polynomial terms are included [1, p. 86]. Volterra series and certain special cases of the Volterra series, for example, the Hammerstein model [2] and the memory polynomial model [3], have been proposed for predistorter design that includes memory effects.

To the best of our knowledge, [4, 5] are the only published results on orthogonal polynomials for predistorter design. Our approach is different and has the following advantages: (i) Our orthogonal polynomial basis functions are expressed in closed form (non-iterative), and the coefficients are free of round-off errors. (ii) Our basis functions are pre-determined and can be implemented with little demand on the computation resources. In [4, 5], the basis functions are calculated online and iteratively, thus requiring much more computational power. (iii) Our basis set consists of both even and odd-order terms whereas that of [4, 5] allows odd-powered series only. Moreover, our basis function expressions are for generally complex-valued baseband data; their application to nonlinear systems with memory is also prescribed.

In Section 2, we describe the conventional polynomial model and point out its deficiencies. In Section 3, we derive novel orthogonal polynomial basis functions which are numerically more stable than the conventional one. In Section 4, simulation results are presented to illustrate the benefits of the orthogonal polynomials. Finally, conclusions are drawn in Section 5.

2 THE POLYNOMIAL MODEL

We consider the indirect learning architecture [6] as shown in Fig. 1. The baseband predistorter input is denoted by

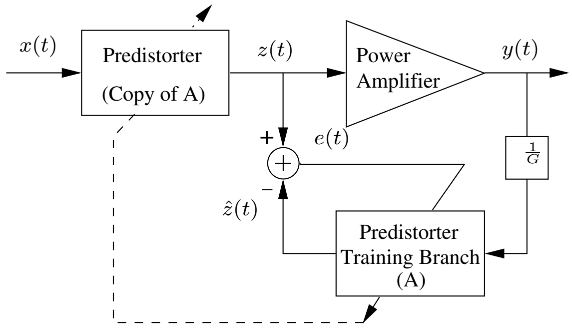

Figure 1. Indirect learning architecture.

$x(t)$, the baseband predistorter output/PA input is denoted by $z(t)$, and the baseband PA output is denoted by $y(t)$. The feedback path labeled "Predistorter Training Branch" (block A) has $y(t)/G$ as its input, where G is the intended gain of the PA, and $\hat{z}(t)$ is its output. The actual predistorter (copy of A) is an exact copy of the predistorter training branch. When $y(t) = Gx(t)$, the error $e(t) = z(t) - \hat{z}(t)$ is 0. To reduce the error between $y(t)$ and $Gx(t)$, we choose the predistorter parameters that minimizes the error $e(t)$. The benefit of the indirect learning architecture is that, instead of assuming a model for the PA, estimating the PA parameters and then constructing its inverse, we can go directly after the predistorter[1].

For the predistorter training branch, if an all-order (even and odd) polynomial model [7] is adopted, we have,

$$\hat{z}(t) = \sum_{k=1}^{K} a_k \left|G^{-1}y(t)\right|^{k-1} G^{-1}y(t). \qquad (1)$$

Define the polynomial basis function $\phi_k(x) = |x|^{k-1}x$, Eq. (1) becomes

*This work was supported in part by the National Science Foundation grant MIP 9703312, the State of Georgia's Yamacraw Initiative, and Danam USA Incorporated.

[1]The term "indirect learning" seems counter-intuitive here, since the predistorter is learned directly; it is the PA characteristics that are learned indirectly.

$$\hat{z}(t) = \sum_{k=1}^{K} a_k \ \phi_k(G^{-1}y(t)). \qquad (2)$$

Based on a set of PA input/output measurements, $\mathbf{z} = [z(t_1), \ldots, z(t_N)]^T$ and $\mathbf{y} = [y(t_1), \ldots, y(t_N)]^T$, a simple least-squares (LS) estimator is obtained for the coefficients, $\mathbf{a} = [a_1, \ldots, a_K]^T$:

$$\hat{\mathbf{a}}_{LS} = \left(\mathbf{Y}^H\mathbf{Y}\right)^{-1}\mathbf{Y}^H\mathbf{z}, \qquad (3)$$

where $\mathbf{Y} = [\boldsymbol{\phi}_1(G^{-1}\mathbf{y}) \ \boldsymbol{\phi}_2(G^{-1}\mathbf{y}) \ \ldots \ \boldsymbol{\phi}_K(G^{-1}\mathbf{y})]$, and for $1 \le k \le K$, $\boldsymbol{\phi}_k(\mathbf{y}) = [\phi_k(y(t_1)), \ldots, \phi_k(y(t_N))]^T$. Once the coefficients $\hat{\mathbf{a}}_{LS}$ are found, they are plugged into the predistorter. This procedure can be repeated iteratively. Such iterative procedure enables the predistorter to adapt to a slowly time-varying PA.

The inversion of the matrix $\mathbf{Y}^H\mathbf{Y}$ in (3) can experience a numerical instability problem. Let us assume for simplicity, $G = 1$. Consider $y(t_1), \ldots, y(t_N)$ identically distributed whose magnitude $|y(t_i)|$ is uniformly distributed in $[0, 1]$. It can be shown that $E[\frac{1}{N}\mathbf{Y}^H\mathbf{Y}]_{kl} = \frac{1}{k+l+1}$ (See the Appendix). This matrix is known as a segment of the generalized Hilbert matrix with $p = 2$ [8], which is ill-conditioned. The condition number (norm 2) of a matrix is defined as $\rho = \left|\frac{\lambda_{\max}}{\lambda_{\min}}\right|$, where $\lambda_{\max}$ and $\lambda_{\min}$ are its maximum and minimum eigenvalues, respectively. It can be used to predict the numerical stability associated with matrix inversion. In general, when the condition number is much larger than 1, the numerical error involved in inverting the matrix can be significant.

3 ORTHOGONAL POLYNOMIALS

To alleviate the numerical instability problem associated with the basis set $\boldsymbol{\Phi} = [\phi_1(x), \phi_2(x), \ldots, \phi_K(x)]^T$ in (2), we consider orthogonal polynomials. To derive a set of orthogonal polynomial basis, $\boldsymbol{\Psi} = [\psi_1(x), \psi_2(x), \ldots, \psi_K(x)]^T$, which spans the same space as $\boldsymbol{\Phi}$, we consider the following requirements:

1. Orthogonality: any two different basis functions, $\psi_k(x)$ and $\psi_l(x)$, are orthogonal; i.e., $\langle\psi_k(x), \psi_l(x)\rangle = 0$ for $k \ne l$, and $\langle\psi_k(x), \psi_k(x)\rangle = d_k > 0$. The inner product $\langle\psi_k(x), \psi_l(x)\rangle$ is discussed in detail in the Appendix.
2. Form of the polynomial basis: we consider polynomial basis $\psi_k(x) = \sum_{l=1}^{k} B_{kl} \ \phi_l(x) = \sum_{l=1}^{k} B_{kl}|x|^{l-1} x$, where B_{kl} is a generally complex-valued coefficient. Note that $\psi_k(x)$ has order k.

Therefore, we seek a lower triangular matrix $\mathbf{B}$, with $[\mathbf{B}]_{kl} = B_{kl}$ for $k \ge l$ and 0 otherwise, to construct the basis $\boldsymbol{\Psi} = \mathbf{B}\boldsymbol{\Phi}$ satisfying:

$$\langle\boldsymbol{\Psi}, \boldsymbol{\Psi}\rangle = \mathbf{B}\langle\boldsymbol{\Phi}, \boldsymbol{\Phi}\rangle\mathbf{B}^H = \mathrm{diag}(d_1, \ldots, d_K), \qquad (4)$$

where $\langle\boldsymbol{\Psi}, \boldsymbol{\Psi}\rangle_{kl} = \langle\psi_k(x), \psi_l(x)\rangle$, and $\mathrm{diag}(d_1, \ldots, d_K)$ denotes a diagonal matrix with $[\mathrm{diag}(d_1, \ldots, d_K)]_{kk} = d_k$. As we show in the Appendix, the (k, l)th element of $\langle\boldsymbol{\Phi}, \boldsymbol{\Phi}\rangle$ is $E[r^{k+l}]$; i.e., the $(k + l)$th-order moment of $r = |x|$. For example, when $r \sim U[0, 1]$; i.e., r is uniformly distributed in $[0, 1]$, $\langle\boldsymbol{\Phi}, \boldsymbol{\Phi}\rangle_{kl} = \frac{1}{1+k+l}$. Given a probability density function (PDF) of r, the orthogonal polynomial basis construction problem becomes finding the lower triangular matrix $\mathbf{B}$ such that $\mathbf{B}\langle\boldsymbol{\Phi}, \boldsymbol{\Phi}\rangle\mathbf{B}^H$ is diagonal. Therefore in theory, orthogonal polynomials are tied to the PDF of the signal amplitude.

When $r \sim U[0, 1]$ and with the requirement that the squared norm of the basis to be preserved; i.e., $\|\psi_k(x)\|^2 = \|\phi_k(x)\|^2 = \frac{1}{2k+1}$, we show in [9] that the matrix $\mathbf{B}$ that solves this problem is

$$B_{kl} = (-1)^{l+k}\frac{(k + l)!}{(l - 1)!(l + 1)!(k - l)!}. \qquad (5)$$

Therefore, the kth-order orthogonal polynomial basis function for the $|x| \sim U[0, 1]$ distribution is

$$\psi_k(x) = \sum_{l=1}^{k}\frac{(-1)^{l+k}(k + l)!}{(l - 1)!(l + 1)!(k - l)!}|x|^{l-1}x. \qquad (6)$$

Table 1 shows the first 5 such orthogonal polynomials. If x were real-valued, applying the same procedure to orthogonalize the basis $\{1, x, x^2, \ldots, x^K\}$ (i.e., adding the x^0 term) would yield the shifted Legendre polynomials [10]. Although the construction of orthogonal basis is often an iterative procedure, we were able to obtain novel, *closed form* expression (6) for complex-valued $x(t)$.

Table 1. **Orthogonal polynomial basis functions $\psi_k(x)$ for $1 \le k \le 5$.**

$\psi_1(x) = x$
$\psi_2(x) = 4\|x\|x - 3x$
$\psi_3(x) = 15\|x\|^2x - 20\|x\|x + 6x$
$\psi_4(x) = 56\|x\|^3x - 105\|x\|^2x + 60\|x\|x - 10x$
$\psi_5(x) = 210\|x\|^4x - 504\|x\|^3x + 420\|x\|^2x - 140\|x\|x + 15x$

The orthogonal polynomial predistorter in the training branch is given by:

$$\hat{z}(t) = \sum_{k=1}^{K} c_k \ \psi_k(G^{-1}y(t)). \qquad (7)$$

Therefore, the LS estimator for the coefficients $\mathbf{c} = [c_1, c_2, \ldots, c_K]^T$ is

$$\hat{\mathbf{c}}_{LS} = \left(\mathbf{M}^H\mathbf{M}\right)^{-1}\mathbf{M}^H\mathbf{z}, \qquad (8)$$

where $\mathbf{M} = [\boldsymbol{\psi}_1(G^{-1}\mathbf{y}) \ \boldsymbol{\psi}_2(G^{-1}\mathbf{y}) \ \ldots \ \boldsymbol{\psi}_K(G^{-1}\mathbf{y})]$, and for $1 \le k \le K$, $\boldsymbol{\psi}_k(\mathbf{y}) = [\psi_k(y(t_1)), \ldots, \psi_k(y(t_N))]^T$.

Since $\psi_k(x)$ is a linear combination of $\{\phi_l(x)\}_{l=1}^{k}$, models (7) and (2) are equivalent, in theory. However, in practice, sampling the input and output of a PA using a finite precision A/D converter may introduce error to the samples. Furthermore, since obtaining the LS estimates of the predistorter coefficients requires a matrix inversion (see (3)), the digital signal processor (DSP) precision may impact the accuracy of the resulting matrix inverse. To avoid numerical problems due to quantization and finite precision calculations in the DSP, the eigenvalue spread should be minimized. In the case where $y(t)$ is uniformly distributed in amplitude between 0 and 1, the condition number, ρ, for the matrix $\mathbf{Y}^H\mathbf{Y}$ (3), increases with the matrix dimension exponentially. In contrast, when orthogonal polynomials are used; i.e., when matrix $\mathbf{M}^H\mathbf{M}$ (8) is used, the condition number is only $\rho = \frac{2K+1}{3}$ when $|x(t)|$ is uniformly distributed in $[0, 1]$.

The set of orthogonal polynomial basis (6) derived earlier, can be extended to built a predistorter with memory.

In the memory case, we assume uniform sampling with the sampling period of T. The memory polynomial predistorter model, which has the form,

$$\hat{z}[n] = \sum_{k=1}^{K}\sum_{q=0}^{Q-1} a_{kq} \ \phi_k\left(G^{-1}y[n - q]\right), \qquad (9)$$

where $y[n] = y(t_n) = y(nT)$ and $z[n] = z(t_n) = z(nT)$, is shown to be a robust predistorter model for PAs with memory [3], where K is the highest polynomial order, and $Q-1$ is the maximum delay. The coefficients a_{kq} can be obtained using least squared method.

We propose to construct an orthogonal memory polynomial predistorter as follows:

$$\hat{z}[n] = \sum_{k=1}^{K} \sum_{q=0}^{Q-1} a_{kq} \ \psi_k \left(G^{-1} y[n-q] \right), \qquad (10)$$

where $\psi_k(\cdot)$ is the orthogonal polynomial basis function (6). Although orthogonality still holds for $\psi_k(G^{-1}y[n-q_1])$ and $\psi_l(G^{-1}y[n-q_2])$ when $q_1 = q_2$, it doesn't hold for $q_1 \neq q_2$ due to the correlation introduced by the common terms. However, we still expect that the predistorter (10) exhibits better numerical stability as compared to (9). Result for the performance of the memory orthogonal polynomials are presented in [9].

4 NUMERICAL EXAMPLES

In this section, we would like to explore the benefits of applying orthogonal polynomials to the predistorter design.

4.1 Robustness of Orthogonal Polynomials

The simulation environment is C, floating point data with 64-bit precision accuracy. Although the set of orthogonal basis functions (6) (see also Table 1) are derived assuming that the input amplitude, $r = |x|$, is uniformly distributed, we would like to show that such basis functions are beneficial even if r is not uniformly distributed. Let us consider the truncated Rayleigh distribution,

$$f_r(r) = \begin{cases} \frac{1}{1-e^{-\frac{1}{2\sigma^2}}} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, & 0 \le r \le 1, \\ 0, & \text{otherwise}, \end{cases} \qquad (11)$$

and the truncated exponential distribution,

$$f_r(r) = \begin{cases} \frac{1}{1-e^{-\frac{1}{\lambda}}} \frac{1}{\lambda} e^{-\frac{r}{\lambda}}, & 0 \le r \le 1, \\ 0, & \text{otherwise}. \end{cases} \qquad (12)$$

Consider as examples, five specific distributions of r: (i) r is uniformly distributed between 0 and 1; (ii) r is truncated Rayleigh distributed (c.f. (11)) with $\sigma^2 = 0.1086$; (iii) r is truncated Rayleigh distributed (c.f. (11)) with $\sigma^2 = 0.5$; (iv) r is truncated exponentially distributed (c.f. (12)) with $\lambda = 0.2127$; (v) r is truncated exponentially distributed (c.f. (12)) with $\lambda = 1$.

Fig. 2 shows the condition number for the matrix $\mathbf{Y}^H\mathbf{Y}$ (which is an empirical estimate of $N\langle \mathbf{\Phi}, \mathbf{\Phi} \rangle$) for each of the PDFs. We note that, for each PDF, the condition number grows exponentially with the polynomial order K. Fig. 3 shows the condition number for the matrix $\mathbf{M}^H\mathbf{M}$ (which is an empirical estimate of $N\langle \mathbf{\Psi}, \mathbf{\Psi} \rangle$) for each of the PDFs, which increases at a much slower rate as K increases and are within 100 for the cases tested. The low condition number will ensure good numerical stability when finite precision computation of the predistorter coefficients is carried out.

4.2 Predistorter design

The simulation environment in this section is C, floating point data with 32-bit precision accuracy. An example is given to demonstrate how the numerical problems associated with estimating the predistorter coefficients affect the performance of the predistorter in terms of spectral regrowth suppression. We utilize the system shown in Fig. 1 to perform predistortion linearization. The predistorter's input, $x(t)$, is a three carrier UMTS signal. Both conventional polynomials and orthogonal polynomials are considered for the construction of the predistorter with orders $K = 7$.

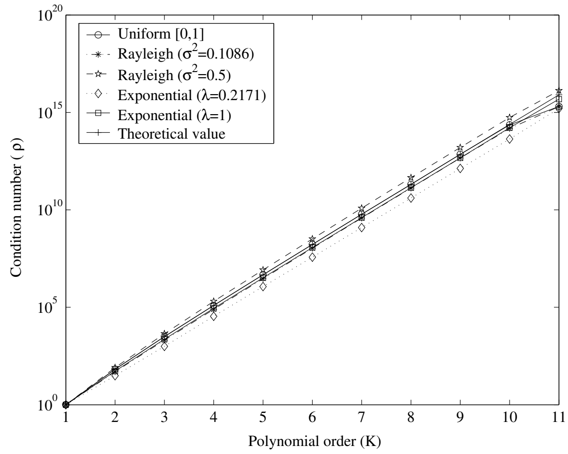

Figure 2. Condition number of $\mathbf{Y}^H\mathbf{Y}$ for the PDFs given in (i)-(v).

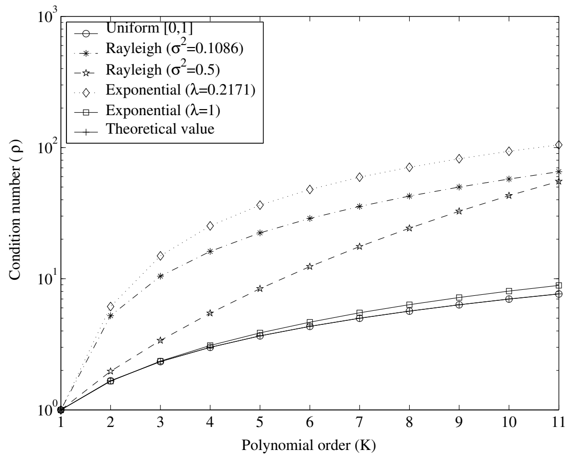

Figure 3. Condition number of $\mathbf{M}^H\mathbf{M}$ for the PDFs given in (i)-(v).

The PA has the following input/output relationship (arctan model):

$$y(t) = \left(\alpha_1 \tan^{-1}(\beta_1 |z(t)|) + \alpha_2 \tan^{-1}(\beta_2 |z(t)|) \right) e^{j\angle z(t)}, \ (13)$$

where $\alpha_1 = 8.0034 - j4.6116$, $\alpha_2 = -3.7717 + j12.0376$, $\beta_1 = 2.2690$, and $\beta_2 = 0.8234$. This PA model fits well measured data from an actual class AB PA. The intended linear gain is set to $G = 7$.

Fig. 4 shows the PSD at the output of the PA for the conventional polynomial predistorter with a polynomial order $K = 7$. The PSD is presented for iterations 15, 18, and 21, respectively and shows no sign of convergence. In contrast, Fig. 5 shows the PSD at the output of the PA for the orthogonal polynomial predistorter with the same order $K = 7$. In this case, the predistorter shows stability and effectiveness.

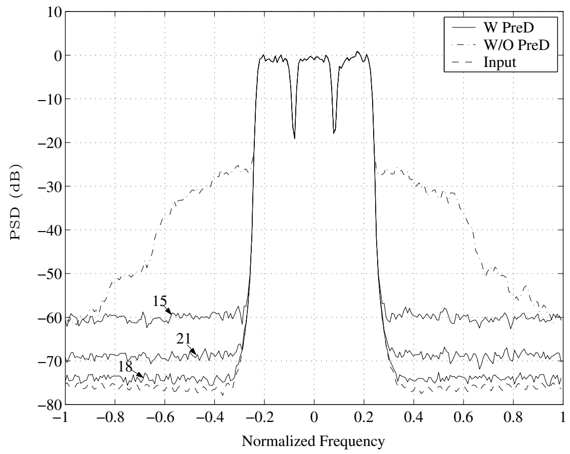

Figure 4. Conventional polynomial predistorter example with $K = 7$. Dash dotted line shows the PA output PSD without predistortion; solid lines show the PA output PSD with predistortion (results are shown for iteration numbers 15, 18 and 21); dashed line shows the PA input PSD. For easy visual comparison, output PSDs are normalized with respect to the input PSD. The predistorter did not converge, revealing a numerical instability problem.

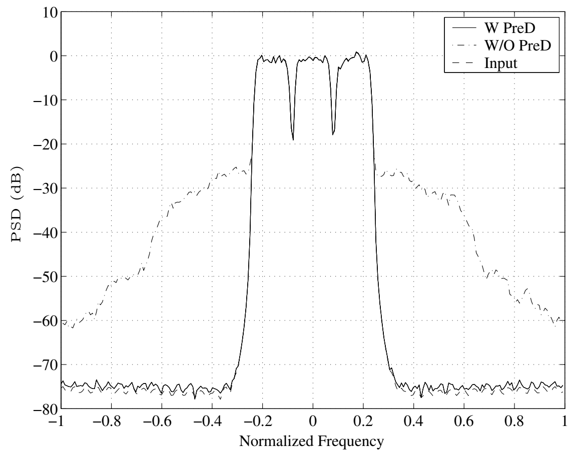

Figure 5. Orthogonal polynomial predistorter example with $K = 7$. Dash dotted line shows the PA output PSD without predistortion; solid lines show the PA output PSD with predistortion (results are shown for iteration numbers 15, 18 and 21); dashed line shows the PA input PSD. For easy visual comparison, output PSDs are normalized with respect to the input PSD. The predistorter converged and could fully suppress spectral regrowth.

5 CONCLUSION

In this paper, the benefits of using the orthogonal polynomials as opposed to conventional polynomials are explored. Closed-form expression for the orthogonal polynomial basis is derived. We demonstrated via simulations the numerical instability problem when high order conventional polynomials are used. In terms of predistortion linearization performance, the spectral regrowth suppression pattern exhibits instability when a high-order, conventional polynomial predistorter is used. This instability problem appears when the matrix inverted in the LS estimator of the polynomial coefficients, is ill-conditioned. The problem can be resolved by using an orthogonal polynomial predistorter. Extension to an orthogonal polynomial predistorter with memory is also described.

REFERENCES

[1] S. C. Cripps, *Advanced Techniques in RF Power Amplifier Design.* Norwood, MA: Artech House, 2002.

[2] L. Ding, R. Raich, and G. T. Zhou, "A Hammerstein linearization design based on the indirect learning architecture," in *Proc. IEEE Conf. Acoust. Speech Sig. Process.*, vol. 3, (Orlando, FL), pp. 2689–2692, May 2002.

[3] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "Memory polynomial predistorter based on the indirect learning architecture," in *Proc. IEEE Global Telecommunications Conference*, (Taipei, Taiwan), Nov 2002.

[4] P. Midya, "Polynomial predistortion linearizing device, method, phone and base station." United States Patent No. 6,236,837, July 30, 1998.

[5] P. Midya and J. Grosspietsch, "Scalar cost function based predistortion linearizing device, method, phone and basestation." United States Patent No. 6,240,278, July 30, 1998.

[6] C. Eun and E. J. Powers, "A predistorter design for a memory-less nonlinearity preceded by a dynamic linear system," in *Proc. Global Telecommunications Conference*, vol. 1, (Singapore), pp. 152–156, November 1995.

[7] L. Ding and G. Zhou, "Effects of even-order nonlinear terms on predistortion," in *Proc. 10th IEEE DSP Workshop (DSP'2002)*, (Pine Mountain, GA), Oct 2002.

[8] R. B. Smith, "Two theorems on inverse of finite segments of the generalized Hilbert matrix," *Mathematical Tables and Other Aids to Computation*, vol. 13, pp. 41–43, January 1959.

[9] R. Raich, H. Qian, and G. T. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design." submitted to *IEEE Trans. on Vehicular Technology*, Jan. 2003.

[10] W. Gautschi, "On the preceding paper "A Legendre polynomial integral" by James L. Blue," *Mathematics of Computation*, vol. 33, pp. 742–743, April 1979.

APPENDIX - INNER PRODUCT OF THE POLYNOMIAL BASIS

Assume x is a complex-valued random variable with a probability density function $f_{r,\theta}(r,\theta)$, where $r = |x|$ and $\theta = \angle x$. Let $\psi_k(x)$ and $\psi_l(x)$ be functions of the random variable x. The inner product of the two functions is defined as:

$$\begin{aligned}\langle \psi_k(x), \psi_l(x)\rangle &\triangleq E_x[\psi_k(x)\psi_l^*(x)] \\ &= \iint_{\theta,r} \psi_k(x)\psi_l^*(x) f_{r,\theta}(r,\theta)\ dr\ d\theta, \quad (14)\end{aligned}$$

where $[\cdot]^*$ denotes complex conjugation.

For the conventional polynomial basis, $\phi_k(x) = |x|^{k-1}x$, the inner product of $\phi_k(x)$ and $\phi_l(x)$ is

$$\langle \phi_k(x), \phi_l(x)\rangle = E[|x|^{k+l}] = E[r^{k+l}],$$

which depends only on the PDF of $r = |x|$.

As a special case, consider the magnitude $r = |x|$ uniformly distributed in $[0,1]$; i.e., the PDF

$$f_r(r) = \begin{cases} 1, & r \le 1, \\ 0, & r > 1. \end{cases} \quad (15)$$

It follows easily that $\langle \phi_k(x), \phi_l(x)\rangle = \frac{1}{1+k+l}$.