

SeMicro-PR 2018

Configuração em aritmética de vírgula fixa para síntese em FPGA dos modelos MP, EMP e CMEMP

Luis Schuartz, Isabella F. Wosniack, Felipe A. Schoulten, Émeli L. C. Silveira, Sibilla B. L. França, Eduardo G. Lima
Grupo de Concepção de Circuitos e Sistemas Integrados
Universidade Federal do Paraná, Curitiba, Brasil
luisschuartz@ufpr.br

Resumo—Os sistemas de comunicação sem fio atuais, para suprir a alta demanda de transferência de dados cumprindo com as normas de comunicação e garantir maior rendimento, devem ser lineares e eficientes simultaneamente. Para manter linearidade e eficiência o uso de um pré-distorcedor digital (DPD), conectado em cascata com o amplificador de potência (PA), é uma solução viável e de baixo custo. Contudo, o DPD deve ser capaz de caracterizar inversamente a característica de não linearidade e memória do PA cujos modelos devem ser escolhidos para garantir o melhor compromisso entre complexidade e acurácia. O objetivo deste trabalho é comparar os requisitos necessários para implementação em hardware demandados pelos seguintes modelos simplificados das séries de Volterra: polinômio de memória (MP), polinômio de memória de envoltória (EMP) e polinômio de memória e de memória de envoltória combinado (CMEMP), aplicados a um estudo de caso modelando um PA LDMOS classe AB. Resultados de simulação apresentam que o CMEMP demanda mais custo de hardware comparado aos demais, no entanto garante melhor acurácia de modelagem, enquanto o modelo MP necessita de menor quantidade de recursos lógicos para obter a saída em uma mesma quantidade de ciclos de clock que o modelo CMEMP.

Palavras-chave: Amplificador de potência, eficiência, linearização, LUT, vírgula flutuante, VHDL.

I. INTRODUÇÃO

Nos sistemas de comunicação sem fio atuais, devido aos esquemas de modulação em amplitude e fase, a linearidade e a eficiência são requisitos fundamentais para cumprir com a demanda de tráfego de dados e otimizar o uso de baterias [1]. No transmissor, o amplificador de potência (PA) é o componente que exige maior atenção nestes aspectos devido à sua característica de linearidade vs. rendimento: em regiões de maior eficiência, a linearidade é comprometida, e em regiões de maior linearidade, a eficiência é comprometida [2]. Para resolver tal situação, o uso de um pré-distorcedor digital (DPD), em banda base, conectado em cascata com o PA é uma

solução de baixo custo que permite manter linearidade e eficiência no sistema simultaneamente [2]-[4].

O DPD aplica ao sinal a ser modulado uma característica de distorção inversa a do PA, de tal forma que a característica de transferência da cascata (DPD + PA) seja linear. Para realizar este procedimento, modelos polinomiais com memória, especificamente as séries de Volterra [5], são a técnica mais difundida na literatura para representar as características de memória e não linearidade dos PAs. Modelos polinomiais possuem as vantagens de linearidade entre seus parâmetros e acurácia, em contrapartida, o número de coeficientes cresce exponencialmente com o aumento de não linearidade e memória [6]. Para resolver tal situação, simplificações das séries de Volterra permitem boa relação entre acurácia e complexidade (número de coeficientes) [7], [8]. Além da complexidade do modelo, outras questões que necessitam atenção especial são a conversão de vírgula flutuante para vírgula fixa e a implementação de funções não lineares por meio de tabelas de busca (LUTs) [9], [10], que se resumem a duas situações: número mínimo de bits para precisão numérica das operações básicas e número de bits para endereçamento das LUTs. Portanto, o objetivo desta pesquisa é apresentar a melhor configuração da pré-distorção digital, para três modelos da literatura aplicados a um estudo de caso, garantindo a melhor relação entre acurácia e demanda computacional para o processamento digital.

Este artigo está dividido conforme segue: a Seção II revisa conceitos aplicados à pré-distorção digital e aos modelos polinomiais; a Seção III apresenta a implementação de funções não lineares por meio de tabelas de busca; a Seção IV apresenta a implementação em vírgula fixa; a Seção V detalha os resultados aplicados ao estudo de caso; e a Seção VI contém as conclusões.

II. CONCEITOS E MODELOS APLICADOS À PRÉ-DISTORÇÃO DIGITAL

Um sinal modulado em amplitude e fase pode ser descrito como uma amplitude em função do tempo

multiplicada por uma função periódica de frequência específica de fase com uma parcela variante no tempo e uma constante aleatória conforme:

$$s(t) = r(t) \cos(2\pi f_c t + \phi(t) + \theta), \quad (1)$$

onde $s(t)$ é o sinal modulado, f_c a frequência de portadora, θ é a fase aleatória da portadora, $r(t)$ e $\phi(t)$ formam a envoltória complexa, em banda base, descrita como:

$$x(t) = r(t)e^{j\phi(t)}. \quad (2)$$

A taxa de variação de $x(t)$, na faixa de MHz, é muito inferior à taxa de variação de $s(t)$, em ordem de GHz [6]. A pré-distorção digital atua sobre a envoltória e , sendo um processo digital, atua sobre a informação discretizada amostrada em taxa nT_s , onde n é um valor inteiro e T_s é o período de amostragem do sinal em banda base. Para o DPD, basta que o período de amostragem seja constante, e a atuação é sobre a envoltória complexa do instante n . Consequentemente, o DPD, por meio de um modelo polinomial modifica a envoltória do sinal inserido no PA de tal forma que a envoltória do sinal de saída do PA é linear em relação ao sinal de entrada do DPD. Logo, os processos digitais e modelos são apresentados apenas como função da amostra no instante n . Na sequência estão apresentados os modelos polinomiais, derivados das séries de Volterra, capazes de modelar o comportamento de não linearidade e memória de um PA, a serem avaliados.

A. Polinômio de memória (MP)

O polinômio de memória (MP) é uma simplificação das séries de Volterra que considera apenas as componentes unidimensionais [7]. O MP possui a vantagem de ser linear em seus parâmetros, os quais podem ser obtidos por um processo de aprendizagem linear, como os mínimos quadrados (LS), que é o algoritmo de melhor custo-benefício entre demanda computacional e acurácia [3]. O modelo MP é descrito pela equação:

$$y(n) = \sum_{p=0}^{P-1} \sum_{m=0}^M h_{m,p} x(n-m) |x(n-m)|^p, \quad (3)$$

onde x é a entrada, y a saída, h o coeficiente e M e P os fatores de truncamento de memória e não linearidade, respectivamente.

B. Polinômio de memória de envoltória (EMP)

Outro modelo da literatura é o polinômio de memória de envoltória (EMP). A principal diferença em relação ao MP é a utilização apenas da amplitude das amostras passadas para modelagem da saída no instante atual. O EMP é escrito conforme:

$$y(n) = x(n) \sum_{p=0}^{P-1} \sum_{m=0}^M h_{m,p} |x(n-m)|^p. \quad (4)$$

Por ser linear em seus parâmetros, a extração de coeficientes pode ser realizada com LS.

C. Polinômio de memória e de memória de envoltória combinado (CMEMP)

O polinômio de memória e de memória de envoltória combinado (CMEMP), proposto em [11], faz uma combinação entre os modelos MP e EMP representado conforme:

$$y(n) = \left[\sum_{m=0}^M a_m x(n-m) \right] \left[\sum_{p=0}^{P-1} \sum_{m=0}^M b_{m,p} |x(n-m)|^p \right], \quad (5)$$

onde a e b são os coeficientes do modelo. O CMEMP combina termos de diferentes instantes de tempo, ponderados pelos coeficientes. O CMEMP não é linear em seus parâmetros, os quais devem ser extraídos por algoritmos de solução não linear. A técnica mais adequada é o algoritmo de mínimos quadrados não lineares que ajusta iterativamente os coeficientes reduzindo o erro de modelagem.

III. APLICAÇÃO DE LUTS PARA PROCESSAMENTO NÃO LINEAR

Os modelos apresentados na Seção II apresentam processamento não linear de ordem definida pelo fator de truncamento P . O número de operações básicas necessárias para obter estes termos cresce exponencialmente com P , aumentando o tempo de processamento e consumo de potência do DPD. Uma das técnicas propostas na literatura e utilizada em [9]-[11] é a substituição da função não linear por uma interpolação linear entre pontos pré-calculados dispostos em uma tabela de busca (LUT). Cada LUT implementa uma função não linear cuja entrada é o quadrado do valor absoluto. Genericamente, para cada instante de memória (m), a função equivalente da LUT é representada por:

$$f_{NL}(in) = \sum_{p=0}^{P-1} h_{p,m} (\sqrt{in})^p, \quad (6)$$

onde $in = |x(n-m)|^2$. Ao criar uma função do módulo ao quadrado da entrada o processo digital é simplificado.

As LUTs implementam uma discretização da função não linear em $Q+1$ pontos, onde, a saída é a interpolação linear entre os pontos discretos. Cada LUT armazena os coeficientes angular (α) e linear (β) da interpolação linear, ou seja, para cada entrada da LUT (in), a saída (out) é:

$$out(n) = \alpha \cdot in(n) + \beta. \quad (7)$$

Os coeficientes angular e linear são obtidos por:

$$\alpha = \frac{y_2 - y_1}{x_2 - x_1}, \quad (8)$$

$$\beta = y_1 - x_1 \frac{y_2 - y_1}{x_2 - x_1}, \quad (9)$$

respectivamente, onde x_1, x_2, y_1 e y_2 são pontos discretos da função não linear (6). Cada LUT contém Q endereços, onde Q é uma potência de 2 para determinado número de bits de endereçamento, ou seja, $Q = 2^{(\text{bits de endereçamento})}$.

IV. CONVERSÃO DE VÍRGULA FLUTUANTE PARA VÍRGULA FIXA

Sistemas digitais exigem a conversão da vírgula flutuante para vírgula fixa. Para realizar esta operação, primeiramente os valores devem possuir uma normalização conhecida, a exemplo de -1 a 1. O número máximo então é convertido para um valor inteiro que pode ser escrito por uma sequência binária de até $2^{(\text{número de bits de precisão})}$. Todo processo digital faz uso de operações básicas de soma e multiplicação, podendo atingir valores maiores que o máximo da normalização. Para isso, alguns bits extras devem ser considerados para evitar *overflow*, assim como mais um bit de sinal [10], [12]. Especificamente para a operação de multiplicação, que dobra a quantidade de bits, apenas metade dos bits é mantida. Em particular, apenas os bits de precisão mais significativos e os bits extras menos significativos são mantidos.

Cada procedimento realizado exige uma operação extra para a implementação do DPD em um sistema. A implementação adequada reduz o tempo e consumo de processamento digital.

V. ESTUDO DE CASO APLICANDO OS MODELOS DA LITERATURA

Os 3 modelos (MP, EMP e CMEMP) são implementados em um estudo comparativo, procurando modelar acuradamente um PA e verificando os requisitos necessários para implementação em *hardware*.

A. Configuração dos testes

Testes são realizados com dados de um PA LDMOS classe AB alimentado com uma portadora de 2 GHz modulada por uma envoltória WCDMA de 3,84 MHz, medidos através de um analisador vetorial de sinais (VSA) Agilent MXA N9020A, com frequência de amostragem de 30,72 MHz, contendo 26180 amostras para extração de coeficientes e 8501 amostras para validação. A modelagem e escolha dos parâmetros são realizados no *software Matlab*. Os procedimentos de distorção utilizando vírgula fixa são processados em linguagem VHDL, simulando uma FPGA Virtex5 LX50T.

A seguinte sequência de procedimentos é realizada. Primeiramente, todos os modelos são identificados usando

vírgula flutuante e representação polinomial. Os fatores de truncamento de memória e não linearidade são escolhidos. O melhor resultado de modelagem é obtido por meio do erro quadrático médio normalizado (NMSE) conforme [13]:

$$NMSE = 10 \log \left\{ \frac{\sum_{k=1}^N \left[\left(y_{I,k}^{des} - y_{I,k}^{cal} \right)^2 + \left(y_{Q,k}^{des} - y_{Q,k}^{cal} \right)^2 \right]}{\sum_{k=1}^N \left[\left(y_{I,k}^{des} \right)^2 + \left(y_{Q,k}^{des} \right)^2 \right]} \right\}, \quad (10)$$

onde $(.)^{des}$ indica os valores desejados, $(.)^{cal}$ os valores calculados, $(.)_I$ a componente em fase, $(.)_Q$ a componente em quadratura e N o número total de amostras medidas.

Na sequência, ainda usando vírgula flutuante, polinômios são substituídos por LUTs com interpolação linear, verificando para cada caso o número de bits de endereçamento necessário para manter a acurácia do modelo. Então o código é convertido para vírgula fixa e é determinada a quantidade mínima de bits de precisão para manter valor de NMSE próximo ao valor obtido com vírgula flutuante.

Definida a quantidade de bits para endereçamento das LUTs e escolhido o número de bits de precisão, o código é programado em VHDL e simulado para determinar os recursos lógicos necessários para sua implementação.

B. Resultados obtidos

Os três modelos descritos na Seção II são implementados no *Matlab*. Os fatores de truncamento são escolhidos para que o NMSE atinja valor mínimo. Os dados de extração e validação possuem pouca memória e todo o sistema é implementado com uma amostra de atraso ($M = 1$). O grau de não linearidade ($P = 7$) é escolhido de forma que, ao aumentar-se o valor escolhido de P , a variação de NMSE é desprezível. Para todos os casos, as LUTs são endereçadas com 5 bits, apresentando um total de 32 endereços chegando muito próximo ao melhor resultado conforme apresenta a figura 1, com NMSE em função do número de bits de endereçamento para as LUTs. A tabela 1 resume os resultados obtidos aplicando vírgula flutuante e vírgula fixa para diferentes valores de bits de precisão e NMSEs de validação na modelagem inversa do PA.

TABELA 1. RESULTADOS DE NMSE OBTIDOS PARA OS DIFERENTES MODELOS APLICANDO VÍRGULA FLUTUANTE E VÍRGULA FIXA.

Modelo	NMSE [dB]			
	Vírgula Flutuante	Vírgula fixa 14 bits	Vírgula fixa 15 bits	Vírgula fixa 16 bits
MP	-37,6	-36,0	-38,1	---
EMP	-36,5	-36,1	-36,6	---
CMEMP	-39,8	---	-38,7	-41,0

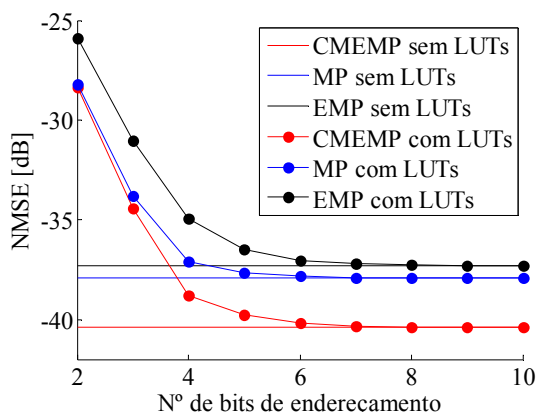


Fig. 1 NMSE em função do número de bits de endereçamento das LUTs para os modelos MP, EMP e CMEMP e precisão de vírgula flutuante.

O código é implementado em VHDL e simulado, necessitando os recursos para cada caso dispostos na tabela 2. São valores absolutos de recursos necessários para um processador digital dedicado à pré-distorção digital.

TABELA 2. RECURSOS NECESSÁRIOS PARA IMPLEMENTAÇÃO DOS MODELOS EM VHDL.

Modelo	Quantidade de recursos lógicos necessários			
	Ciclos de clock	Flip-Flops	Slice LUTs	DSP48
MP	11	122	195	16
EMP	12	190	162	12
CMEMP	11	130	361	22

VI. CONCLUSÃO

Este artigo apresenta uma comparação entre os modelos MP, EMP e CMEMP da literatura em um estudo de caso aplicando diferentes etapas de implementação em VHDL. CMEMP apresenta resultados de melhor acurácia comparados ao MP e EMP. No entanto, a necessidade de um método de solução não linear, a quantidade de bits de precisão e necessidade de *hardware* são os fatores mais agravantes. Em contrapartida, o modelo MP demanda menor custo de *hardware*, conseguindo atingir o resultado em uma mesma quantidade de ciclos de *clock* que o modelo CMEMP, necessitando de 15 bits de precisão e tendo um bit extra para o sinal. A quantidade de recursos lógicos da FPGA para o modelo MP é de 122 *flip-flops*, 195 *slice LUTs* e 16 unidades DSP48 dedicadas para multiplicações exclusivas da FPGA Vixtex5 LX50T.

AGRADECIMENTOS

Os autores agradecem ao suporte financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível

Superior, da Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Estado do Paraná, do Fundo Nacional de Desenvolvimento da Educação e Programa de Iniciação Científica modalidade UFPR/Tesouro Nacional.

REFERÊNCIAS

- [1] M. Dardaillon, C. Jabbour, e V. P. Srin, "Adaptive digital pre-distortion for future wireless transmitters," em IEEE International Conference on Electronics, Circuits, and Systems, 2015, pp. 332-335.
- [2] S. Cripps, RF Power Amplifiers for Wireless Communications, 2ª edição. Norwood, MA: Artech House, 2006.
- [3] L. Schuartz e E. G. Lima, "Comparison among Algorithms for the Identification of Adaptive Memory Polynomial Predistorter Models", em 30º Simpósio Sul de Microeletrônica, Maio 2015, pp. 1-4.
- [4] D. Raychaudhuri e N.B. Mandayam, "Frontiers of Wireless and Mobile Communications," Proc. IEEE, vol.100, pp.824-840, Abril 2012.
- [5] V. Mathews e G. Sicuranza, Polynomial Signal Processing. New York: Wiley, 2000.
- [6] J. C. Pedro e S. A. Maas, "A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches," IEEE Trans. Microw. Theory Tech., vol. 53, n.º 4, pp. 1150-1163, Abril 2005.
- [7] J. Kim e K. Konstantinou. "Digital predistortion of wideband signals based on power amplifier model with memory." Electronics Letters 37.23 (2001): 1.
- [8] O. Hammi, F. M. Ghannouchi e B. Vassilakis. "A compact envelope-memory polynomial for RF transmitters modeling with application to baseband and rf-digital predistortion." IEEE Microwave and Wireless Components Letters, 18(5):359-361, 2008
- [9] I. F. Wosniack, E. L. C. Silveira, F. A. Schoulten, S. B. L. França, E. G. Lima, "Fixed-point VHDL Description of a Look-up Table Based Combined Memory and Envelope Memory Polynomial Model", em Simpósio Sul de Microeletrônica (SIM), Curitiba, 2018.
- [10] F. A. Schoulten, I. F. Wosniack, F. I. Yasuda, S. B. L. França e E. G. Lima, "Fixed-point Arithmetic Description of Limiting and Filtering for Peak-to-average Power Ratio Reduction", em Simpósio Sul de Microeletrônica (SIM), Curitiba, 2018.
- [11] C. L. R. Machado, E. G. Lima, "A Combined Memory and Envelope-Memory Polynomial Model for RF Power Amplifiers", em XXII Iberchip Workshop, Bariloche, Argentina, 2017.
- [12] J. C. L. Pereira, F. A. Schoulten, C. P. Mizerkowski, S. B. L. França, E. G. Lima, "Fixed-point Arithmetic Architecture of a Physicallymeaningful Perceptron for Digital Pre-distorters", em 18th Microelectronics Students Forum (SFORUM), Bento Gonçalves, 2018.
- [13] M. S. Muha, C. J. Clark, A. Moulthrop, and C. P. Silva, "Validation of power amplifier nonlinear block models," em IEEE MTT-S Int. Microwave Symp. Dig., Anaheim, CA, Jun. 1999, pp. 759-762.