摘要——文章提出将增广实值时延项引入到深度神经网络结构中的数字预失真，并研究了采用ELU激活函数来代替Sigmoid函数以避免神经网络训练时的梯度消失问题，加快网络的训练速度。同时，采用增广实值时延的输入结构，将输入信号的包络引入网络的输入端，可以使所提出的模型产生更丰富的基函数，以提升建模精度并减轻PAs失真。所提出的模型在以80MHz与100MHz的大宽带信号驱动的高度非线性的SLCG功率放大器上进行了大量的实验，证明它的非线性补偿与模拟缺陷缓解能力明显优于DDR与单隐层RVTDFNN，邻道泄露比优化了大约4-5dB，归一化均方误差优化了大约3dB；同时，相较于单隐层的RVTDFNN与ARVTDNN，所提出的模型随着驱动信号带宽的增加具有更优的鲁棒性。

*关键词——*功率放大器，数字预失真, 神经网络, ARVTDNN, 无线发射机，ELU激活函数。

II. INTRODUCTION

在无线发射机系统中，针对功率放大器的线性化是一直备受关注的重要问题。在第五代通信系统中，功率放大器所要处理的大带宽信号往往都具体高PAPR特点，而这一类信号对于功率放大器的非线性特征更加敏感，进而导致严重的失真与效率降低。目前，已有非常多的基于多项式模型的DPD技术被提出，如Volterra、MP、DDR[1]等。近年来，随着人工神经网络的发展，许多研究者开始将目光转向将神经网络运用于DPD技术当中，并取得丰富的成果。

关注于输入I/Q信号与其过去时间的值的RVTDFNN[2][3]结构，是目前主流的基于神经网络的数字预失真的基础。而将输入信号的包络引入神经网络的输入端，构成了所谓的ARVTDFNN[4]结构，相较于RVTDFNN，它更够产生更丰富的基函数，即具有更高的建模精度。最近，深度神经网络在图像识别、分类等领域表现出的优异性能也引起了许多PA建模研究者的关注。[3]-[5]中讨论了将神经网络运用于DPD中的效果，但其使用的都是渐层的神经网络，而有关采用深层神经网络结构的DPD的报道较少。

本文提出了一个采用增广实值时延深度神经网络结构的数字预失真，利用ELU激活函数来克服深层神经网络训练的梯度消失问题，同时在输入层引入包络项来加强网络的建模精度。实验结果表明，与传统的DDR-DPD、单隐层RVTDFNN-DPD和ARVTDFNN-DPD相比，所提出的结构能够表现出更优越的性能，同时能够减少计算复杂度。



Figure 1

II. THE PROPOSED DPD

RVTDFNN是在全连接前向神经网络的基础上考虑记忆效应所提出的一种动态神经网络结构，这样的网络可以有效的模拟具有强记忆效应的PA[3]。ARVTDNN是在前者的基础上，将输入信号的包络项的非线性样本引入其输入端，这样能够产生更为丰富的基函数集，更能表现PA的完整行为特征，从而提升建模精度[4]。图2展示了所提出的ARVTD-DNN，是在RVTDFNN的基础上，将输入信号的幅度及其非线性version引入网络输入端，并采用多个隐藏层的深层网络结构，以ELU激活函数代替Sigmoid激活函数。该模型输入端由实值输入正交信号与以及其包络的当前与过去的值组成，输出端有正交信号与组成。

1. ARVTD-DNN

在所提出的模型当中，训练过程的任意时刻，网络输入端的样本序列为一个向量，包括与的现在和过去时间的值：



其中，表示输入样本序列的记忆深度。

，表示输入信号的幅值，表示网络输入端引入输入信号包络的非线性阶数。延迟响应是通过使用作为延迟运算符来实现的，其中是单位延迟运算符，将其用时，会产生其延迟信号。抽头延迟线会存储来自先前时间步长的值。

该网络的训练过程与RVTDFNN的训练过程一致，每一层中单个神经元的输出由下列步骤给出：

1. 第层第个神经元加法器输出（表示上一层神经元数量，表示连接上一层第个神经元与第层第个神经元的突触权值）：



其中表示输入的第个样本，即第0层表示输入层。

1. 第层第个神经元激活函数输出（函数表示）：



1. 在输出层的I/Q分量表示为：（网络层数为L，神经元数量D=2）



1. 在神经网络输出端计算成本函数：



其中表示样本长度，与表示目标分量，而与表示网络输出分量。

本文中所使用的反向传播算法为1维Levenberg-Marquart[6]算法，目的是使得成本函数最小化。网络训练过程中，在前向传播中，计算每一次迭代中的成本函数；在反向传播中，利用LM算法计算，使得成本函数尽可能小的最新参数（突触权重与偏置）。

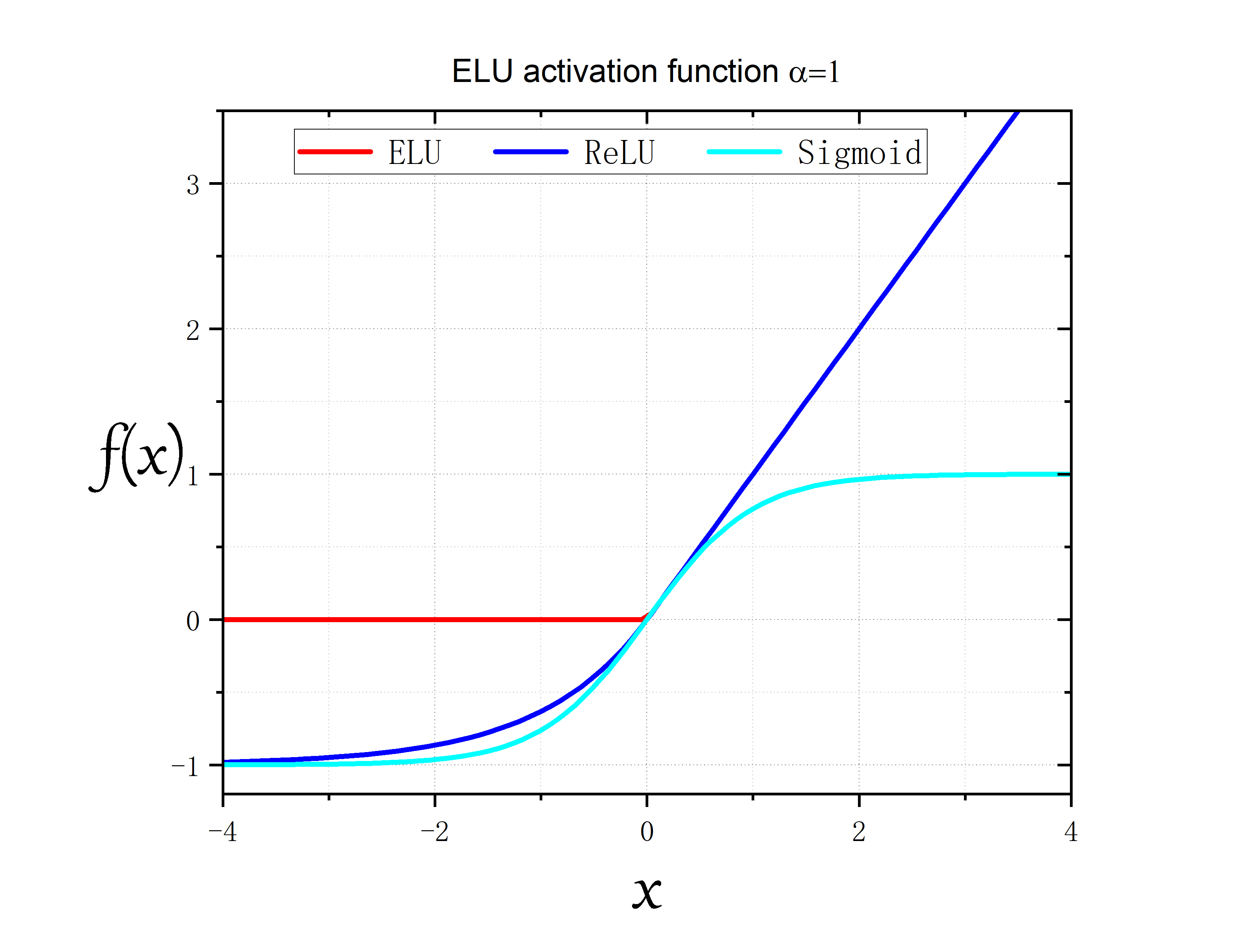
1. Activiation

在这些文献当中，所提到的神经网络中神经元的激活函数均采用的是Sigmoid函数（如图2所示）数学上等于双曲正切，其数学公式如下：



但是S型函数，在处理深度神经网络会出现梯度消失问题，即在Sigmoid函数自变量取值为极大的正值或负值时，函数会饱和，对输入微小的改变变得不敏感，然后导致反向传播过程中参数的梯度收敛到零的情况[7]。在[8]这篇文章中，使用了ReLU激活函数代替Sigmoid激活函数来解决上述问题，并得到更好的补偿性能。但是，在PA的逆向建模过程中，为了完成输入和输出信号之间的时间对准(可以使用互相关技术实现[9])与加快神经网络的训练速度，我们将用于建模的信号（input与output）进行平均功率归一化处理。在测试过程中，我们发现这些训练数据用于采用ReLU作为激活函数的深度神经网络中会造成比较严重的“神经元坏死现象”[10]，原因是进行了归一化后的训练样本中存在大量的负值部分，由于这些负值所引起的负值自变量会造成对于的ReLU神经元“关闭”，最终导致网络的训练失败。所以，本文引入ELU[11]来替代Sigmoid函数，既能克服梯度消失问题，也能避免采用ReLU激活函数时神经元坏死问题。

ELU激活函数（如图2所示）的数学公式如下，其中为超参数，一般取1。



**Figure 2**

III. EXPERIMENTAL RESULTS & DISCUSSION

为了测量和实验验证的目的，所提出的ARVTD-DNN-DPD在MATLAB平台上实现，并与两种高效DPD分别在以宽带信号驱动的真实PA上进行了线性化性能对比，它们分别是基于传统多项式模型的MP-DPD与DDR-DPD，基于单隐层的RVTDFNN-DPD，所使用的PA是具有记忆效应的高度非线性的SLCG PA，其饱和输出功率为37dB，漏极效率为35%。四种DPD均采用间接学习结构。图3展示了测量设置与预失真实现流程。



Figure 3

提出的预失真器所生成的预失真信号上传至信号发生器中，进行调制与上变频，经过驱动级放大后馈入PA，PA的输出信号通过频谱仪进行下变频与采样，获得输出的复数调制信号的I/Q分量。然后将输入与输出的复信号I/Q分量通过MATLAB代码实现时间调整（对准），最终得到PA的非线性特性。这些非线性特性在MATLAB软件平台上由DPD结构获取，并进行逆向建模，以生成预失真器。

RVTDFNN-DPD采用单隐层，根据文章[10]使用25个神经元，时延抽头数选择4。所提出的ARVTD-DNN-DPD的输入端包含三阶包络项，时延抽头数选择4，并隐藏层数L=9，并采用ELU作为激活函数。用于DPD的样本数量为9000。

输出层均含有两个神经元，采用线性激活函数。

为了验证所提出DPD结构的鲁棒性，实验测量了在两种不同的大带宽信号驱动下的具有严重非线性特性的SLCG PA的性能表现，输入信号分别为带宽为40MHz与80MHz的单载波256QAM信号。

应用于SLCG PA的100MHz信号下DPD结果如 表1与图4所示，表I中的NMSE表明了所采用的DPD模型对PA的建模精度，NMSE的公式如下：

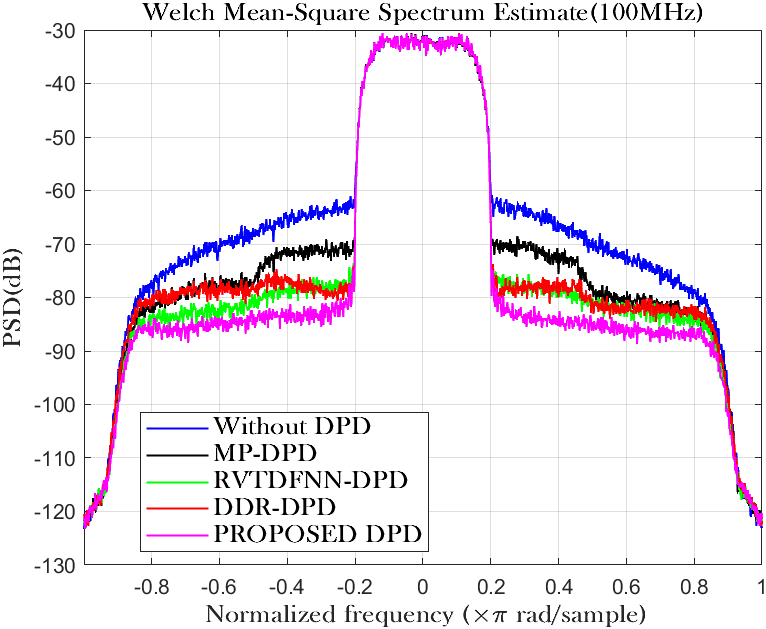


其中表示期望结果，表示测量结果，表示样本点的个数。

图5表现了上述几种DPD在不同带宽信号下的线性化性能，可以发现，所提出的结构相较于传统多项式模型中最优的DDR-DPD能够更好地适应大带宽信号情况下线性化要求。

IV. CONCLUSION

本文提出了ARVTD-DNN-DPD模型，并采用ELU激活函数来代替DNN中的ReLU激活函数，以避免神经元坏死现象。使用SLCG PA进行的实验证明，所提出的DPD方法可以实现很好的补偿效果，并且在大宽带下仍能保持其优良的线性化性能。具有3-4dB的ACLR抑制，2-4dB的ACPR抑制效果，建模精度NMSE也有2-4dB的提升。与性能最优秀的基于传统多项式模型DDR-DPD相比较，所提出的DPD具有更强的鲁棒性，在具有严重的非线性与记忆效应的SLCG PA实验中，在采用80MHz大带宽信号驱动下，DDR-DPD失去了其线性化能力，而所提出的DPD仍然能够进行有效的线性化。



**Figure 4 图中表现了使用不同DPD的SLCG PA线性化情况**

TABLE 1

PERFORMANCE OF VARIOUS MODELS IN TERMS OF NMSE, OPTIMAL

PARAMETERS, AND TOTAL NUMBER OF NEURONS

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BW of Signal** | **Model** | **Layers** | **Number of** **Neurons** | **ACLR (dBc)** | | **ACPR (dBc)** | | **NMSE (dB)** | **EVM (%)** |
| **Low** | **High** | **Low** | **High** |
| **80MHz** | w/o DPD | 0 | 0 | -31.12 | -32.75 | -32.57 | -33.88 |  | 3.5 |
| MP | 0 | 0 | -38.55 | -40.39 | -40.01 | -41.19 | -30.16 | 1.03 |
| RVTDFNN | 2 | 27 | -43.92 | -44.53 | -44.61 | -44.95 | -35.82 | 1.04 |
| DDR | 0 | 0 | -44.92 | -45.91 | -46.49 | -46.69 | -37.57 | 0.79 |
| PROPOSED | 9 | 82 | -48.3 | -47.97 | -49.7 | -50.25 | -38.59 | 0.8 |
| **100MHz** | w/o DPD | 0 | 0 | -31.12 | -32.75 | -32.57 | -33.88 |  | 3.5 |
| MP | 0 | 0 | -38.47 | -38.28 | -39.26 | -39.45 | -30.3 | 1.19 |
| RVTDFNN | 2 | 17 | -45.28 | -44.96 | -45.21 | -45.58 | -35.97 | 0.85 |
| DDR | 0 | 0 | -44.61 | -45.2 | -44.95 | -45.42 | -36.9 | 0.94 |
| PROPOSED | 9 | 82 | -50.71 | -49.52 | -50.76 | -51.37 | -39.17 | 0.6 |

V. REFERENCES

[1] Ghannouchi, Fadhel M. Behavioral Modelling and Predistortion of Wideband Wireless Transmitters. Hoboken: John Wiley & Sons Inc., 2015. Print.

[2] T. Liu, S. Boumaiza, and F. Ghannouchi, “Dynamic behavioral modeling of 3G power amplifiers using real-valued time-delay neural networks,” IEEE Trans. Microw. Theory Tech., vol. 52, no. 3, pp.1025–1033, Mar. 2004.

[3] Rawat, M., K. Rawat, and F.M. Ghannouchi. “Adaptive Digital Predistortion of Wireless Power Amplifiers/Transmitters Using Dynamic Real-Valued Focused Time-Delay Line Neural Networks.” IEEE transactions on microwave theory and techniques 58.1 (2010): 95–104. Web.

[4] Wang, Dongming et al. “Augmented Real-Valued Time-Delay Neural Network for Compensation of Distortions and Impairments in Wireless Transmitters.” IEEE transaction on neural networks and learning systems 30.1 (2019): 242–254. Web.

[5] B. E. Watkins et al., “Neural network based adaptive predistortion for the linearization of nonlinear RF amplifiers,” Military Commun. Conf. 1995, pp. 145-149, San Diego, USA, November 1995.

[6] Hagan, M.T., and M.B. Menhaj. “Training Feedforward Networks with the Marquardt Algorithm.” IEEE transactions on neural networks 5.6 (1994): 989–993. Web.

[7] S. Haykin, Neural Networks: A Comprehensive Foundation. UpperSaddle River, NJ: Prentice-Hall, 1999.

[8] Hongyo, Reina, Yoshimasa Egashira, and Keiichi Yamaguchi. “Deep Neural Network Based Predistorter with ReLU Activation for Doherty Power Amplifiers.” 2018 Asia-Pacific Microwave Conference (APMC). IEICE, 2018. 959–961. Web.

[9] M. S. O. Alink, E. A. M. Klumperink, M. C. M. Soer, A. B. J. Kokkeler, G. J. M. Smit, and B. Nauta, “A CMOS-compatible spectrum analyzer for cognitive radio exploiting crosscorrelation to improve linearity and noise performance,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 59, no. 3, pp. 479–492, Mar. 2012.

[10] Lu Lu,Yeonjong Shin,Yanhui Su George E. Karniadakis.(2019).Dying ReLU and Initialization: Theory and Numerical Examples..CoRR

[11] Djork-Arné Clevert,Thomas Unterthiner Sepp Hochreiter.(2015).Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)..CoRR