Title: Design of a neural network-based VCO with high linearity and wide tuning range

Author(s): Guo, Rui; Qian, Kun; Wei, Jinping; Chen, Tupei; Liu, Yanchen; Kong, Deyu; Wang, J. J.; Wu, Yuancong; Hu, S. G.; Yu, Qi; Liu, Yang


Date: 2019

URL: http://hdl.handle.net/10220/48929

Rights: © 2019 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Design of a Neural Network-Based VCO With High Linearity and Wide Tuning Range

RUI GUO, KUN QIAN, JINPING WEI, TUPEI CHEN, (Member, IEEE), YANCHEN LIU, DEYU KONG, J. J. WANG, YUANCONG WU, S. G. HU, QI YU, (Member, IEEE), AND YANG LIU, (Member, IEEE)

State Key Laboratory of Electronic Thin Films and Integrated Devices, University of Electronic Science and Technology of China, Chengdu 610054, China
School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

Corresponding author: Yang Liu (yliu1975@uestc.edu.cn)

This work was supported in part by the NSFC under Project 61774028 and Project 61771097, and in part by the Fundamental Research Funds for the China Central Universities under Project ZYGX2016Z007.

ABSTRACT In this paper, a 2 GHz LC-VCO with neural network (Multilayer Perceptron) has been designed in a 0.13 μm CMOS technology. With the integrated neural network, the linearity and tuning range of the LC-VCO has been substantially improved. Compared to a conventional VCO design, the proposed technique can improve the linearity by selecting optimized bias voltages obtained from the output of the neuron network. The result shows that the tuning nonlinearity of the proposed VCO is further optimized from 0.335% to 0.254%.

INDEX TERMS LC-VCO, neural network, MLP, varactor tank, linearity.

I. INTRODUCTION

In recent years, RF circuits and systems are making rapid progress due to the push from wireless communication [1]–[4]. As the most crucial part of a phase-locked loop for current RF circuit, VCO has been studied widely. Comparing with the ring oscillator, the LC-VCO has a much lower phase noise than ring oscillator. In the past years, for the LC-VCO, phase noise, low power consumption, and wide tuning range have taken much attention. There have been fewer efforts concentrating on high tuning linearity. The VCO gain Kvco as an extremely significant parameter affects phase noise, frequency lock range, and stability of the loop. In the past research, many techniques have been published, including varactor banks in LC tank [5], differently-biased varactor banks [6]. Linear MOS varactor bank or varactor array is gradually applied in practice widely to provide constant Kvco. Then the main structure of varactor in the LC-VCO circuit is gradually fixed in many types of research [7]–[10]. Though all of these methods and strategies can effectively optimize the linearity of LC-VCO, it is hard to further optimize the linearity without the extra circuit. A new bias voltage selection algorithm is needed to solve the problem. Artificial Neural Network (ANN), which is inspired by biological neural networks, is based on the integration of artificial neurons and synapses. The strong parallel processing power and brain-like behavior of the neural network can also be used in IC design. Reference [11] proposed an ANN to improve the FOM of ADC. Reference [12] realized a new Power-Efficient Transmitter based on ANN to classify the input data. In [13], RF Power Amplifier was also enlightened by ANN. And [14] utilized the brain-like behavior of ANN to finish the gesture recognition function.

In this paper, a varactor resonant tank composed of four identical MOS varactor pairs is used and optimized by the neural network algorithm. When the optimized value of the input layer is assigned to the bias voltage of the varactor, the nonlinearity of the VCO is finally reduced to less than 0.254%. The neural network algorithm can also linearize the tuning curve of a VCO significantly and maintain a wide linear tuning range without the extra circuit.

II. DESIGN OF MLP-BASED VCO

When the bias voltage and the linearity of the VCO are taken as input and output separately, optimizing the linearity of VCO could be treated as a regression problem. Considering that the Multilayer Perceptron (MLP) algorithm takes advantage in solving the regression problem [15], it is used in this design. The MLP consists of one input layer, at least one hidden layer and one output layer. Each layer includes...
multiple neurons which act as the fundamental computation units. The forward propagation which is based on the current inputs and layer weights generate output signals. The loss function is obtained by the label of the training dataset and the output of the forward propagation. In the backward propagation (BP), the layer weights are updated with reducing the loss \[16, 17\]. Then the bias voltage selection algorithm has been carried out in two steps and shown in Fig. 1(a).

In this work, the MLP consists of one input layer, one output layer, and three hidden layers, where there are 32 and 64 and 32 nodes in the first and the second and the third hidden layers, respectively, as shown in Fig. 1(b).

Corresponding to the different bias voltage VA, VB, VC and VD, four neurons are selected as the input layer. The linearity of the predicted curve could be described by one neuron at the output layer. The number of layers and neurons are gradually confirmed during the training process.

Traditionally, there are always three or four same set of varactor arrays in LC-VCO circuit as shown in Fig. 2. As the Fig. 3 shown, if only one bias voltage is utilized in the VCO, the linearity of the tuning curve at 0V, 0.6V and 1.2V can’t be maintained at same time. More bias voltages are needed to optimize the linearity in the whole range at the same time. In this design, four bias voltages (VA, VB, VC, VD) are chosen to improve the linearity of the VCO. The range of the control voltage is the same as the supply voltage of 1.2 V.

Then the sampling tuning curves are obtained when the bias voltages are set as VA, VB, VC, and VD ∈ \{0, 0.15, 0.3, 0.45, 0.6, 0.75, 0.9, 1.05, 1.2\}. When VA, VB, VC, and VD are assigned with all above values, 6561 (i.e. \(9^4\)) curves are acquired as samples to train the MLP, as shown in Fig. 4.

The linearity is used as the Label for MLP training in supervised learning in this design. The label can be obtained as

\[
grad_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad i = 1, 2, \ldots, n - 1
\]

\[
grad_{n-1} = \frac{y_n - y_{n-1}}{x_n - x_{n-1}}
\]

\[
E(\text{grad}) = \frac{1}{n-1} \sum_{i=1}^{n-1} \|grad_i - E(\text{grad})\|^2.
\]

In (1)-(4), \((x_i, y_i)\) are the coordinates of the \(i\)th sample in the tuning curve, and \(n\) is the number of samples in one curve. Then the linearity of each curve in the sampling could be described in the form of Label.
Once the specific structure of MLP and the target have been confirmed, the training procedure is started. Rectified Linear Unit (ReLU) [18] is used as the activation function in this design. By utilizing the mini-batch gradient descent technique, all of these curves are divided into batches. The number of curves in each batch is \( m (m = 64) \). Then the loss and error rate (\( \text{ER} \)) can be obtained as

\[
\text{Loss} = \frac{1}{m} \sum_{i=1}^{m} (\text{predict}_i - \text{Label}_i)^2, \tag{5}
\]

\[
\text{ER} = \frac{|\text{predict} - \text{label}|}{\text{label}}. \tag{6}
\]

By using the back-propagation algorithm with gradient descent [19]–[21], the MLP network was trained as follows:

\[
W^{(i)} := W^{(i)} - \alpha \frac{\partial \text{Loss}}{\partial W^{(i)}}, \tag{7}
\]

\[
b^{(i)} := b^{(i)} - \alpha \frac{\partial \text{Loss}}{\partial b^{(i)}}, \tag{8}
\]

where \( \alpha \) is the learning rate and \( i \) represents the serial number of the layer.

Error Rate is used to describe the learning ability of MLP. If the Error Rate is less than 10\%, it means MLP has learned the relationship between the varactor and bias voltage. In other word, MLP can predict the tuning curve of the VCO accurately.

In the training process, learning rate reduction is used to optimize the MLP. There is a total of 200 epochs in the whole training process. The forward and backward propagation processes are recycled until the early stopping is reported. Learning rate is reduced to one-tenth at the 150th epoch from 0.005 to 0.0005. Adam optimizer [22] is used as the optimizer. The early stopping strategy is also used to avoid overfitting in the training process.

When the structure and weights of the MLP have been confirmed, the input bias voltage would be updated to optimize the linearity of the VCO. Compared with the previous training process, the only difference of training locates in that the gradient of inputs is updated as shown in Fig. 5. Although the gradient of the MLP is calculated for transmitting the gradient, the weights and bias matrix of the MLP remain unchanged. Then the initial learning rate is set as 0.00005.

The loss in this training process can be obtained as

\[
\text{Loss} = |y|^\frac{5}{4}, \tag{9}
\]

where \( y \) is the output of the MLP. With the descending of the loss function, the bias voltage is optimized, and the linearity of the VCO is further improved.

### III. RESULTS AND DISCUSSIONS

To verify the learning and optimization ability of the MLP, the characteristics with the highest linearity in the samples is used as a reference for comparison, as shown in Fig. 6. More varactor pairs are used to optimize the linearity of the VCO effectively. In this design, the center frequency and gain of the VCO are set as 2GHz and \( K_{\text{vco}} = 140\text{MHz} \), respectively, for the training dataset.

R-squared (\( R^2 \)) [23] is used to evaluate the fitting result for the model. An \( R^2 = 1 \) indicates that the regression perfectly fits the data. When the four bias voltages are assigned with \([0 0.15 0.6 1.05]\), this F-V curve reaches the best linearity when \( R^2 = 0.99986 \) in samples. The nonlinearity can be defined with non-linear error \( \delta \) and calculated as 0.335\% by equation

\[
\delta = \Delta Y \max / Y \times 100\%, \tag{10}
\]

\[
\Delta Y = Y_{3\text{rd degree polynomial}} - Y_{1\text{st degree polynomial}}, \tag{11}
\]

where \( Y \) is the frequency of the VCO.

In the training process, the error rate and loss are used to describe the learning ability of the MLP as shown in Fig. 7.
Error rate and loss reduce gradually and finally tend to be stable. At the 150th epoch, early stopping strategy is used to reduce the error rate further [24]. When the error rate is less than 10%, the MLP can be used to update the input bias voltage in the second step.

When the input voltages have been updated by the MLP within [0.6234, 1.082, 0.0122, 0.165], the optimizing VCO tuning curve is obtained as shown in Fig. 8. The updated curve has high linearity indicators for $R^2 = 0.99989$ and nonlinearity = 0.247%. As shown in the figure, comparing with the highest linearity tuning curve in the training sample, the optimal bias voltage algorithm can further optimize the linearity of VCO.

Considering the fact that the bias voltages cannot be obtained as accurate as the updated input voltage, the bias voltages are selected from [0.62, 1.08, 0, 0.165]. Then the indicator ($R^2$) can still be maintained and the nonlinearity is reduced to 0.254%. The difference in tuning curves between the in-application bias voltage and an optimized bias voltage is shown in Fig. 9. As can be observed in the figure, the actual bias voltage can keep linearity well.

### IV. CONCLUSION

This paper presents a VCO with MLP neural network that can improve the linearity. The VCO is biased with different voltages to generate the training data for the MLP network. The linearity of each curve of the sample can be extracted in the form of Label. By optimizing the bias voltage with the MLP, the proposed MLP-based VCO can improve the linearity without extra circuit cost. Simulation results show that the MLP network can optimize the linearity for a nonlinearity of 0.254% automatically.

### REFERENCES


RUI GUO received the B.S. degree in microelectronics from the University of Electronic Science and Technology of China, Chengdu, China, where he is currently pursuing the Ph.D. degree in microelectronics. His current research interests include frequency synthesizer design and artificial neural network algorithm.

KUN QIAN received the B.S. degree in integrated circuit design and integrated system from the University of Electronic Science and Technology of China, Chengdu, China, where he is currently pursuing the master’s degree. His current research interests include hardware neural network accelerator and high-performance computing.

JINPING WEI received the B.S. degree in microelectronics from the University of Electronic Science and Technology of China, Chengdu, China, where he is currently pursuing the master’s degree. His current research interests include deep learning, machine learning, and their applications.

TUPEI CHEN has been a Faculty Member with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, since 2000. He has authored or coauthored more than 280 peer-reviewed SCI journal papers, more than 130 conference presentations, one book, and six book chapters. He has filed ten U.S. patents and 12 technology disclosures. His research interests include the applications of nanoscale materials in electronic devices, nanophotonic devices and memory devices, memristor and its application in artificial neural networks, Si photonics, and metal oxide thin films and their applications in flexible / transparent devices.

JINYAN LIU received the B.S. degree in microelectronics from the University of Electronic Science and Technology of China, Chengdu, China, where he is currently pursuing the Ph.D. degree in microelectronics. His current research interests include RF circuits design, VLSI design, and artificial neural networks.

DEYU KONG received the B.S. degree in microelectronics from the University of Electronic Science and Technology of China, Chengdu, China, where he is currently pursuing the Ph.D. degree in microelectronics. His current research interests include artificial neural network algorithm, and intelligent medical systems.

J. J. WANG received the B.S. degree in microelectronics from the University of Electronic Science and Technology of China, Chengdu, China, where he is currently pursuing the Ph.D. degree. His current research interests include thin-film transistor, nonvolatile memory devices, and their applications in artificial intelligence.
YUANCONG WU received the B.S. degree in microelectronics. He is currently pursuing the Ph.D. degree with the University of Electronic Science and Technology of China, Chengdu, China. His current research interests include artificial intelligence chips and systems.

S. G. HU received the Ph.D. degree in microelectronics from the University of Electronic Science and Technology of China, Chengdu, China, where he has been an Associate Professor, since 2016. His current research interests include neuromorphic devices, chips, and systems.

QI YU received the Ph.D. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2010, where he is currently a Professor and the Vice Dean of the School of Microelectronics and Solid-State Electronics.

YANG LIU received the B.S. degree in microelectronics from Jilin University, Changchun, China, in 1998, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2005. In 2008, he joined the School of Microelectronics, University of Electronic Science and Technology of China, Chengdu, China, as a Full Professor. He has authored or coauthored more than 130 peer-reviewed journal papers and more than 100 conference papers. His current research interests include memristor neural network systems and neuromorphic ICs. He was a recipient of the Prestigious Singapore Millennium Foundation Fellowship, in 2006.

***