Additive Neural Network Based Static and Dynamic Distortion Modeling for Prior-Knowledge-Free Nyquist ADC Characterization

Danfeng Zhai, Peizhe Li, Jiushan Zhang, Chixiao Chen, Fan Ye, Junyan Ren
State Key Laboratory of ASIC and System
Department of Microelectronics, Fudan University
Shanghai 201203, China
Email: junyanren@fudan.edu.cn

2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS) | 978-1-6654-2461-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/MWSCAS47672.2021.9531763

***Abstract*—This paper presents a prior-knowledge free modeling method for Nyquist ADCs. Current ADC modeling methods mainly base on known circuit implementation and non-idealities, thus hard to recover non-linear static and dynamic distortions. The proposed method adopts an additive neural network with binary inputs to achieve a data driven, prior-knowledge free modeling method. Both static and dynamic distortions are modeled by two separate sub-network. Also, a batch generation scheme is used to enhance the noise insensitivity, facilitating small sample training, when only simulation results are available. The proposed methods are validated by three typical non-ideal ADC designs, including a SAR ADC with capacitor mismatch, an ultra-high speed ADC with NMOS sampling switch, and a SAR ADC with a bandwidth limited reference source. All the non-linearity and FFT spectrum plots show the proposing model can accurately model both static and dynamic distortion with less than 1dB spur mismatch.**

***Index Terms*—ADC modeling, neural network, high-order harmonic, prior-knowledge free,**

I. Introduction

The rapid increase in analog-to-digital converters' (ADCs) energy efficiency drives the higher integration trend of analog and mixed-signal circuits into system-on-chips (SoC). Monolithic commercial ADC dies are nearly extinct recently. Instead, ADC-oriented systems are playing important roles in many applications, such as wireline/wireless communication [1] and accurate biomedical sensors [2]. Meanwhile, it is of great interest how the non-idealities of ADCs affect the entire systems' performance. With the increasing system complexity, there is no inevitability between the ADC distortion and system mitigation. It is hard to precisely predict these effects by a single formula.

One straightforward way to evaluate such an effect is to substitute the ADC circuit into the entire system simulation. However, it is time-consuming and impractical. Therefore, the need for general-usage and accurate ADC modeling methods is growing. Thanks to the recent computing power boom, the model complexity could be relieved as long as it is much simpler than the transistor-based model. Besides, ADC modeling can also help develop new calibration schemes [3].

Previous ADC modeling schemes highly depend on prior-knowledge, especially ADC topologies. Linear approximation of capacitor mismatch and the amplifier's gain is the most common non-idealities to model [4]. Symbolic tools are adopted to model active RC-integrators with finite gain-bandwidth in [5], where closed-form distortion formulas can be derived. For high-speed scenarios, linear interleaving models with sampling skews are also established in [6]. ADC performances are obtained by simulating the combination of non-ideal ADC sub-blocks.

Although featuring simplicity, these prior-knowledge-based models suffer from the following issues. First, the model is usually first-order, which is not sufficient. For example, the single-channel ADC model for time interleaving cases only includes gain and offset errors in [6], while higher-order harmonic distortions and bandwidth mismatches are neglected. However, these non-idealities may mitigate ADC performances. Second, certain signal-dependent dynamic distortions cannot be recovered by the modeling schemes mentioned above. A Volterra series based dynamic model is proposed in [7], which combines static non-linear error model and dynamic linear error model. But it still not characterizes signal-derivative-independent dynamic errors.

In the paper, we propose a data-driven neural network (NN) based prior-knowledge-free modeling method that characterizes both static and dynamic distortion. The rest of the paper is organized as follows. Section II reviews common ADC distortion sources. The proposed additive binary input neural network is detailed in Section III. The effectiveness of the model is verified by three different cases in Section IV. Section V concludes the paper.

II. Distortion Characterization in Nyquist ADCs

Common distortion sources in Nyquist ADCs can be divided into two categories, static errors, and dynamic ones. They need to be characterized differently.

Static Distortion mainly refers to fixed amplitude-dependent distortions as long as ADC chips are fabricated. Common linear distortion models can characterize comparator offset, passive device mismatches, and finite gain of operational amplifiers. These distortion sources normally can be directly mapped to physical nature. Therefore, the distortion extraction

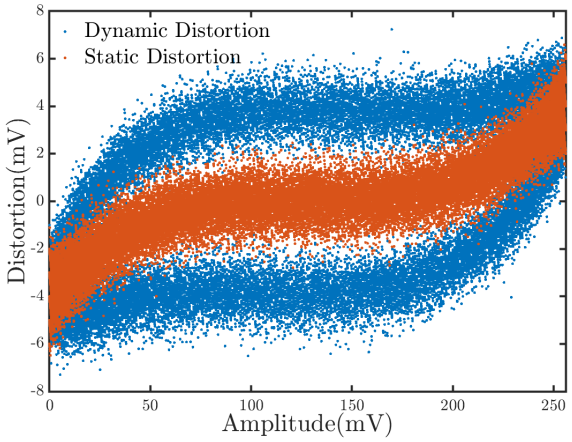


Fig. 1. Static vs. Dynamic in terms of INL

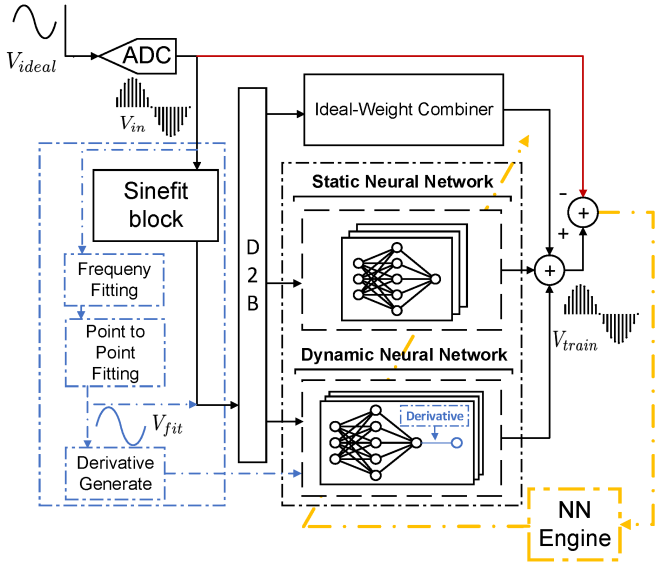


Fig. 2. The proposed prior-knowledge-free ADC modeling method

based on such models can be not only easily simulated but also fast and accurate [8]. Other static distortions are complicated. Non-linearity of amplifiers requires high-order polynomial modeling or Taylor approximation. Recent low power design trends using open-loop dynamic amplifiers highly depends on these models, but requires a very complicated statistical derivation to extract precise parameters [9].

Dynamic Distortion, in contrast, varies even input amplitude is identical. A typical dynamic distortion source is sampling skew occurred in multi-channel time-interleaved ADCs (TIADC). The errors due to sampling skew depend on both input slope and the skew [10], among which the only skew can be regarded as signal independent fixed errors. Filters with frequency domain response also contribute to dynamic distortions. Fortunately, they are normally employed by continuous-time noise-shaping ADCs, and rare in Nyquist ADCs, which is not within the scope of the paper.

Distortions are usually characterized by histogram integral non-linear (INL) testing where inputs are clipping low-frequency sinusoidal waves. However, dynamic errors are often ignored when using this method. To investigate the dynamic issues, high-frequency waves are desired, which significantly differentiates the static/low frequency in terms of INL. In order to distinguish the two differences, we use the method of subtracting ADC output from ADC ideal input to obtain the distortions and then match ADC ideal input and distortions to plot the figure.

Figure 1 illustrates the INL diagram obtained by low and high frequency respectively. For low-frequency distortion where only static distortion is concerned, the results can fit into one curve shown in red. On the other hand, this one curve is split into two for high-frequency input with dynamic distortions, illustrated as blue points. A detailed analysis shows that the two curves represent the INL of positive and negative slope separately.

III. Proposed Modeling Scheme

A. Prior-knowledge-free Modeling using NN

In this paper, a data-driven prior-knowledge-free model is proposed to precisely characterize both static and dynamic errors. Inspired by the recent success of machine learning and artificial intelligence, the proposed method establishes a neural work trained by the gradient descent algorithm. The network is illustrated in Fig. 2. The input of the modeling engine is the real ADC outputs with sinusoidal inputs, from either simulation or testing results.

The entire modeling process includes the following steps. The first is to build an ideal input by sine-fit. Accurate values of the sine amplitude, frequency, and initial phase are obtained by the Nelder-Mead Simplex method [11]. Afterward, the outputs of the sine-fit are quantized and transformed into a bit-wise binary vector. Assuming that an N-bit ADC is modeled, the sine-fit outputs are quantized to $N+1$-bit numbers, and then each number would be transformed to an $N+1$-element binary vector. This vector serves as the NN input. The NN block consisted of both static and dynamic distortion NN, which will be detailed in the next subsection. All NN weights are updated by a training engine. The train engine takes the differences between NN outputs and the real ADC outputs as errors. Note that the process does not utilize any ADC design information to modeling, except for the resolution bit.

B. Additive NN with Binary Inputs

The NN consists of three sub-networks, ideal weight combiner (IWC), static, and dynamic distortion network. The IWC reconstructs the binary $N+1$ vector into a real number, i.e., $\Sigma_i D_i 2^i$. The IWC weights are always radix-2 and fixed. The weights of the other two sub-network are updated by the training engine. The static network is a typical shallow NN similar to [3], modeling amplitude-dependent non-linearity.

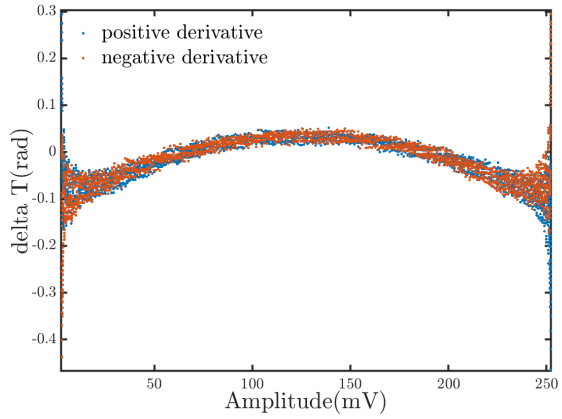


Fig. 3. Delta T with different derivative

On the other hand, the dynamic network first predicts an amplitude-dependent time shift ΔT by a static similar NN and then multiplies the time shift with the wave derivative. The product stands for the first-order approximation of the voltage errors due to dynamic distortion. The wave derivative can be also obtained by the sine fit process. The results of the three networks are combined with a direct summation. To avoid complex chain-rule derivative for cascading multi-layer networks, the proposed NN adopts a parallel structure instead.

In practice, static distortions are more obvious than dynamic ones. To ease the convergence and avoid the local minimum, we would like to first train the static sub-network and then freezes it but train the dynamic one. A sign of correct convergence is the two INL curves with high-frequency inputs approach and finally overlap, shown in Fig. 3.

C. Noise Insensitivity Enhancement

The main challenge of the original static network is signal noise sensitivity. Tens of millions of samples are required to average these random statistical effects in [3]. However, this amount is infeasible for simulation result based modeling. If weights are updated after each sample, this noise would distract the correct convergence. Each real ADC output can be represented by,

$$D_{\text{real}} = D_{\text{ideal}} + \Delta\epsilon + N_{\text{Qn}} + N_{\text{Wh}}, \tag{1}$$

where $\Delta\epsilon$ is the distorted error, and N_{Qn} and N_{Wh} represents quantization and white noise respectively. For signal-to-noise ratio (SNR) dominated ADCs, N_{Wh} is likely greater than $\Delta\epsilon$. To eliminate these effects, a multi-sample-based batch is introduced to average those white noise effects, such as thermal noise. In other words, the mathematical expectation is applied to (1), and thus,

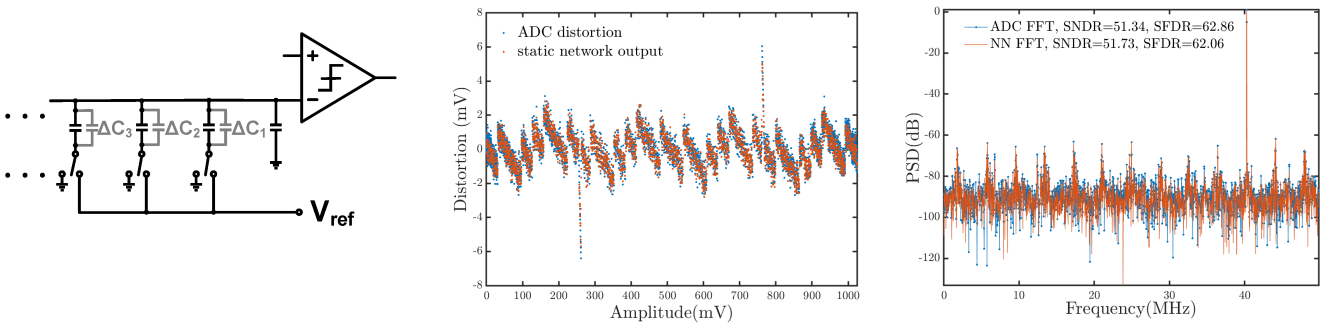
$$E(D_{\text{real}}) = E(D_{\text{ideal}}) + E(\Delta\epsilon). \tag{2}$$

Note that batches with a randomly selected sample might have zero $E(\Delta\epsilon)$. Thanks to the differential nature of ADCs, a positive $\Delta\epsilon$ is likely to be canceled by a negative one with complementary input amplitude. Therefore, each batch, corresponding to once weight update, is supposed to demonstrate a similar $\Delta\epsilon$. One batch in the proposed method contains real outputs of close amplitude, to ensure the correct convergence. A re-order queue buffer is adopted to achieve the batch generation.

IV. Experiment Results

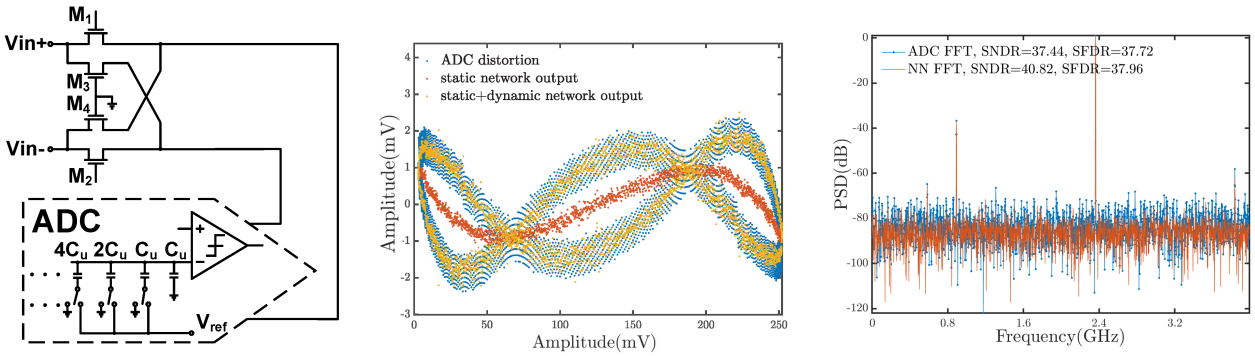
To verify the effectiveness of the proposed method, we apply it to model three practical ADCs with different distortion sources. The static network has 2 layers, between which the first layer input is a binary $N+1$-element vector. The output of the first layer is sent to the second layer after the rectified linear unit (ReLU). The first layer has 128 neurons. The second layer is a fully connected network, which merges the 128 output into one value. The dynamic network has 3 layers, and the first two layers are the same as the static one. The third layer is an element-wise multiplier, whose inputs are the derivative of the current input and the output of the previous layer. Note that the derivative is known during the sine fit phase.

All ADC prototypes are designed by 65nm CMOS technology. Each design is simulated with sinusoidal input, generating about 4400 samples. Among these samples, 4140 samples are used for training, while the rest 256 samples for testing. Comparisons between the transistor-level circuit simulation



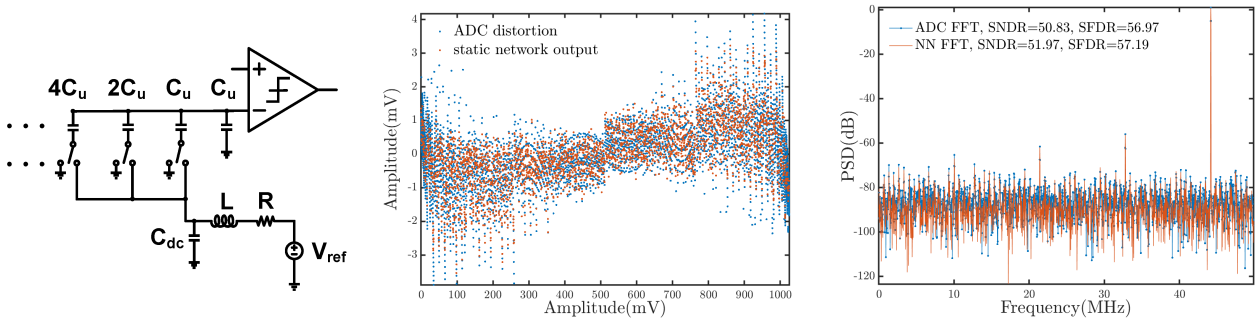
(a) Topology with non-linear characterization (b) Non-linearity plot vs. Modeled distortion (c) FFT spectrum

Fig. 4. Modeling case A: capacitor mismatch for a binary weighted SAR ADC



(a) NMOS sampling switch (b) Non-linearity plot vs. Modeled distortion (c) FFT spectrum

Fig. 5. Modeling case B: NMOS sampling switch distortion for a binary weighted SAR ADC



(a) Bandwidth limited Reference (b) Non-linearity plot vs. Modeled distortion (c) FFT spectrum

Fig. 6. Modeling case C: The influence of Bandwidth limited Reference on SAR ADC

and the proposed modeling method are made. Both the non-linearity and spectrum plot for each case is investigated.

A. Design A: SAR ADC with Capacitor Mismatch

The first case is the successive approximation register (SAR) ADC with a capacitive DAC (CDAC) mismatch. Recently, the capacitors are custom designed with a small area. The mismatch among capacitors affects achievable linearity, and thus resolution. As shown in Fig. 4(a), a binary-weighted capacitor array with a 2.3%-σ mismatch would bring wrong charge redistribution and offset voltage of the comparator.

Fig. 4(b) demonstrates the INL modeling effectiveness. The plot shows a good match between the two sets of curves. Missing codes are also extracted. The spectrum plot shows similar results.

B. Design B: High-Speed ADCs using NMOS sampling switch

Though widely used, bootstrapped switches are rarely employed by high-speed ADCs, because their latency performance will decrease rapidly when both signal and sampling frequency are above few Giga hertz. NMOS sampling transistors are more commonly used [12]. However, the on-resistance of the sampling switch, and the time constant is signal and derivative dependent. Figure 5(a) illustrates the schematic of high-speed ADCs using NMOS sampling switches with neutralization. The differential input signals were connected with NMOS switch transistors M_1 and M_2. The cross-coupled transistors M_3 and M_4 are used to neutralize the feedthrough effect of M_1 and M_2 at hold mode.

Due to the fixed gate voltage of switch transistors (M_1M_2), the impedance of them varies with the input voltage, because of their gate-source voltage variation. Moreover, the caused distortion is not only amplitude-dependent but also dynamic. The proposed modeling method tracks both static and dynamic non-linearity, shown in Fig. 5(b). The spectrum result is also close to the real results, except that the noise floor is slightly less.

C. Design C: SAR ADCs with Bandwidth limited Reference

The constant reference voltage is critical for high-speed SAR ADCs. However, inductive bonding wires or active buffer can limit the reference voltage response in practical. The limit causes incomplete voltage settling during quantization, thus mitigating ADC resolutions. In Fig.6(a), an off-chip reference voltage source V_{ref} connects to the bottom plate of capacitors in a SAR ADC through a bonding wire. The bonding wire can be modeled as an inductor L and a resistor R, and to relax the negative effect of the bonding wire, a decoupling capacitor C_{dc} is applied.

The incomplete settling results in a complicated non-linearity, which can be well addressed in Fig. 6(b). The spectrum in Fig. 6(c) result is also close to the actual ones.

V. Conclusion

This paper presents a prior-knowledge free modeling method for Nyquist ADCs. Two separate neural networks are proposed to model a different type of distortions in ADC. A new batch generation scheme is proposed to solve the small data set question. The method can cover various static and dynamic distortions at one time without too many settings. The availability of this method can be proved by the above three typical non-ideal ADC designs.

References

[1] Y. Lin, M. Keel, A. Faust, A. Xu, N. R. Shanbhag, E. Rosenbaum, and A. C. Singer, "A study of ber-optimal adc-based receiver for serial links," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 5, pp. 693–704, 2016.

[2] C. Kim, S. Joshi, H. Courellis, J. Wang, C. Miller, and G. Cauwenberghs, "A 92db dynamic range sub-μvrms-noise 0.8μ w/ch neural-recording adc array with predictive digital autoranging," in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, 2018, pp. 470–472.

[3] T. Zhang, Y. Cao, S. Zhang, C. Chen, F. Ye, and J. Ren, "Machine Learning Based Prior-Knowledge-Free Calibration for Split Pipelined-SAR ADCs with Open-Loop Amplifiers Achieving 93.7-dB SFDR," *ESSCIRC 2019 - IEEE 45th European Solid State Circuits Conference*, pp. 189–192, 2019.

[4] G. M. Salgado, D. O'Hare, and I. O'Connell, "Modeling and analysis of error feedback noise-shaping sar adcs," in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2020, pp. 1–5.

[5] A. Zhang and G. Shi, "Closed-form distortion formulas for continuous-time sigma-delta modulators," *IEEE Access*, vol. 7, pp. 75 642–75 647, 2019.

[6] S. Kiran, A. Shafik, E. Z. Tabasy, S. Cai, K. Lee, S. Hoyos, and S. Palermo, "Modeling of adc-based serial link receivers with embedded and digital equalization," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 9, no. 3, pp. 536–548, 2019.

[7] W. Wei, P. Ye, J. Song, H. Zeng, J. Gao, and Y. Zhao, "A behavioral dynamic nonlinear model for time-interleaved adc based on volterra series," *IEEE Access*, vol. 7, pp. 41 860–41 873, 2019.

[8] Y. Chiu, C. W. Tsang, B. Nikolic, and P. R. Gray, "Least mean square adaptive digital background calibration of pipelined analog-to-digital converters," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 1, pp. 38–46, 2004.

[9] C. Cao, "A 91.0-db sfdr single-coarse dual-fine pipelined-sar adc with split-based background calibration in 28-nm cmos," *IEEE Transactions on Circuits and Systems I: Regular Papers, Early Access*, vol. 99, p. 99, 2021.

[10] M. Guo, J. Mao, S. W. Sin, H. Wei, and R. P. Martins, "A 1.6-GS/s 12.2-mW Seven-/Eight-Way Split Time-Interleaved SAR ADC Achieving 54.2-dB SNDR with Digital Background Timing Mismatch Calibration," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 3, pp. 693–705, 2020.

[11] Jeffrey, C., Lagarias, James, A., Reeds, Margaret, H., Wright, and P. and, "Convergence properties of the nelder–mead simplex method in low dimensions," *SIAM Journal on Optimization*, 1998.

[12] L. Kull, S. Member, D. Luu, S. Member, C. Menolfi, M. Brändli, P. A. Francese, S. Member, T. Morf, S. Member, M. Kossel, S. Member, A. Cevrero, I. Ozkaya, S. Member, T. Toifl, and S. Member, "A 24-72-GS/s 8-b time-interleaved SAR ADC with 2.0-3.3-pJ/Conversion and $>$ 30 dB SNDR at Nyquist in 14-nm CMOS FinFET," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 12, pp. 3508–3516, 2018.