Modeling of IC Buffers from Channel Responses Via Machine Learning Kernel Regression

Modeling of IC Buffers from Channel Responses via Machine Learning Kernel Regression

Riccardo Trinchero †*, Tommaso Bradde †, Mihai Telescu ‡, Igor S. Stievano †

† Dept. Electronics and Telecommunications, Politecnico di Torino, Italy
‡ Univ Brest, Lab-STICC, CNRS, UMR 6285, F-29200 Brest, France
* corresponding: {riccardo.trinchero@polito.it}

***Abstract*—This paper investigates the potential of a fully behavioral approach for the generation of accurate models of digital IC buffers based on conventional kernel regressions. The proposed approach does not assume a specific model structure like the classical two-piece model representation which has been massively used in literature, offering a promising and viable alternative to facilitate the modeling of nonlinear electrical devices. The collected results represent a first proof-of-concept, aimed at demonstrating the strengths of the proposed alternative modeling approach.**

***Index Terms*—digital integrated circuits, buffer modeling, signal integrity, high-speed interconnects, kernel regression.**

I. Introduction

Since the early 1990s, there has been an growing interest of the signal and power integrity community in generating accurate and efficient simulation models of digital IC buffers for the quality and reliability assessment of high-speed digital channels.

The above trend was stimulated by the birth and by the subsequent improvements of the input output buffer information specification (IBIS) which was strongly supported by the electronic design automation (EDA) and silicon vendors [1]. IBIS suggests a physical-consistent modeling approach based on basic building blocks inspired by the constitutive elements composing the various IC buffers, with all the required features and recent technological advancements (including drivers and receivers, single-ended and differential technology, power-supply effects and pre/de-emphasis). IBIS models for single-ended and differential drivers assume a two-piece structure in which the output port current of the buffer is defined by the weighted combination of two submodels describing the device behavior in either the high or the low logic state. This assumption facilitates model estimation, allowing researchers concentrating on model advancements and additional features.

Based on the above picture, IC buffer modeling inspired by the IBIS idea has always been the mainstream solution which is well supported by both the EDA tools and companies, representing the natural playground for research (e.g., see [2] and references therein). A wide variety of research papers can be found in the literature, with the common background of assuming a two-piece structure and with the main differences on the type and structure of submodels, ranging from physical-consistent modeling approaches to more behavioral solutions (mostly Artificial Neural Networks (ANNs) or Recurrent Neural Networks (RNNs)) [3]–[11].

In addition, some attempts in addressing the problem of IC modeling via a fully black-box behavioral approach based on ANNs have been followed [12], [13]. Even if a good accuracy was observed, some inherent critical aspects still have to be solved in order to have a robust modeling framework. ANNs generally require a possibly large number of neurons, a clever algorithm for training and for forcing a physical consistent behavior of models in a typical simulation environment, and a careful design of the training sequences, being this latter aspect also shared by the approaches in which a two-piece structure is assumed. In addition, a possible alternative blind modeling solution should be able to be flexible and general, to accommodate the inherent multiport electrical characteristic of IC buffers.

Among the above requirements, a key role is played by the simplification of the modeling procedure, including the training which should be done on-the-fly, i.e., based on the observation of device responses in a typical application scenario, without specific requirements and setup for custom excitation. Some preliminary efforts were tried along the years [14]–[16], however failing to match an ultimate solution to the development of a general, simple, accurate and scalable approach to device modeling.

Undoubtedly kernel-based machine learning regression can be seen as a promising candidate for the above modeling problem [17], [18]. The linear model structure adopted by kernel regressions has the key advantage of radically simplifying the training phase when a "relatively small" set of training data is used, compared to the more flexible ANN structures [17], [19], [20].

In this paper, the kernel ridge regression (KRR) is used for the identification of the dynamic nonlinear characteristic of a single-ended IC buffer, offering a first preliminary study on the feasibility and strengths of this alternate approach. Emphasis is given on model generation through a simple behavioral procedure based on the observation of the IC transient responses during normal operation of devices.

II. Statement of the problem

In this work, we propose the generation of a mathematical model allowing to mimic the electrical behavior of the output

port current of a digital IC driver like the one shown in the scheme of Fig. 1. Specifically, we seek for an approximated relation describing the multiport constitutive characteristic of the driver for which the implicit form writes:

$$g\left(\boldsymbol{u}(t), y(t), t, \frac{d}{dt}, \ldots, \frac{d^p}{dt^p}\right) = 0, \tag{1}$$

where $g(\cdot)$ is a generic nonlinear dynamic map, $\boldsymbol{u}(t) = [v_i(t), v_o(t)]$ and $y(t) = i_o(t)$.

A fully behavioral approach is followed, avoiding the need of a dedicated test setup and cumbersome device control for the observation of the transient voltage and current responses required for model generation (i.e., for the estimation of the parameters defining g).

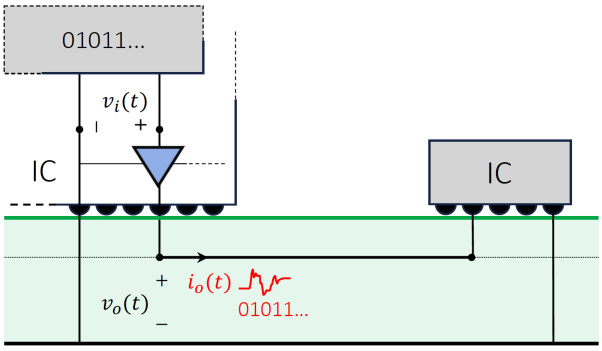

Fig. 1. Typical interconnect structure with the main blocks and the relevant input and output electrical variables of an IC driver.

III. System Identification via Kernel Ridge Regression

Let us consider the problem of building a model of the explicit dynamic characteristic of a generic non-linear multiport circuit element with input $\boldsymbol{u}(t)$ and output $\hat{y}(t)$ of the form [17], [18]:

$$\hat{y}_k = f(\hat{y}_{k-1}, \ldots, \hat{y}_{k-p}, \boldsymbol{u}_k, \ldots, \boldsymbol{u}_{k-p}), \tag{2}$$

where f is a generic nonlinear dynamic mapping, $\hat{y}$ and $\boldsymbol{u} = [u^{(1)}, \ldots, u^{(n)}]^T$ are the estimated output and corresponding input vector collecting the voltages and/or currents at the device ports with discrete time index k (e.g., $\hat{y}_k = \hat{y}(kT)$, being T the assumed sampling period) at different time instants. The value p denotes the model order.

The above model is usually referred to in the literature as NOE (nonlinear output error) model and consists of a recursive equation in the estimated output. Indeed, the model prediction at a given time step depends on the predictions of the model at previous p time steps. Data-driven modeling techniques and machine learning approaches can be used to learn a NOE model, such as the one in (2), starting from a set of time samples of the input and output signals [17], [18].

Specifically, given a training set $\mathcal{D}_{NOE} = \{(\boldsymbol{x}_l, y_l)\}_{l=1+p}^{L}$ where the vector $\boldsymbol{x}_l = [\boldsymbol{u}_l, \ldots, \boldsymbol{u}_{l-p}]$ collects the current and past values of the discrete time input $\boldsymbol{u}_l$ and y_l is the corresponding discrete time output signals the dynamic map in (2) can be learnt via recurrent regression techniques, such as the RNNs [8], [17], [18].

However, the NOE model is not fully compatible with the feedforward structure used by conventional kernel regressions. Indeed, even if a recurrent formulation of kernel regressions is in principle feasible, the training of such model would required to solve a non-convex problem, thus limiting the advantages of such approaches with respect to RNNs (additional details in this regards can be found in [17]–[20]).

A viable solution that facilitates the application of traditional kernel-based regressions in the context of system identification is given by the nonlinear autoregressive with exogenous input (NARX) model, which writes [17], [18]:

$$\hat{y}_k = f(y_{k-1}, \ldots, y_{k-p}, \boldsymbol{u}_k, \ldots, \boldsymbol{u}_{k-p}), \tag{3}$$

where the entries $y_{k-1}, \ldots, y_{k-p}$ and $\boldsymbol{u}_k, \ldots, \boldsymbol{u}_{k-p}$ denote the true output and the inputs at different time instants, and $\hat{y}_k$ is the estimated output at time k.

Different from the recursive model in (2), the NARX model does not have any recursion in the variable y_k, since it uses as input the true output values $y_{k-1}, \ldots, y_{k-p}$ available in the training set. Therefore, the NARX model in (3) is a static model which can be suitably learnt via conventional feedforward kernel regressions [17] by using as training data $\mathcal{D}_{NARX} = \{(\tilde{\boldsymbol{x}}_l, y_l)\}_{l=1+p}^{L}$, where now $\tilde{\boldsymbol{x}}_l = [\boldsymbol{x}_l, y_{l-1}, \ldots, y_{l-p}]^T$.

The dynamic map $f(\cdot)$ can be learnt from the NARX model in (3) via a conventional kernel ridge regression (KRR). The resulting model can be evaluated as a recurrent model as:

$$\hat{y}_k = \sum_{l=1+p}^{L} \alpha_l K(\tilde{\boldsymbol{x}}_l, [\boldsymbol{x}_k, \hat{y}_{k-1}, \ldots, \hat{y}_{k-p}]^T), \tag{4}$$

where $\{\alpha_l\}_{l=1+p}^{L}$ are the coefficients to be estimated during the learning phase and $K(\cdot,\cdot)$ is a conventional scalar kernel function.

The coefficients $\boldsymbol{\alpha} = [\alpha_{1+p}, \ldots, \alpha_L]^T$ can be suitably estimated from the $\mathcal{D}_{NARX}$ training set via the solution of the following linear system [19], [21]:

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \mathbf{I}_L)^{-1} \boldsymbol{y}, \tag{5}$$

where $\mathbf{K}$ is a Gramian matrix such that $[K]_{ij} = K(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{x}}_j)$ defined by evaluating the kernel function on each configuration pairs belonging to the training input set and λ is a Tikhonov regularizer. In this work a Gaussian RBF kernel is used and its hyperparameter together with λ are estimated via a 3-fold cross-validation. The RBF kernel is used instead of other state-of-the-art kernel functions such as the polynomial kernel, since it has shown superior performance in several realistic complex regression problems (see as an example the results

in [22]). The effectiveness of alternative kernel functions will be considered in future research.

It is important to remark that even if the training of the NARX model is carried out by considering the static map in (3), the resulting model obtained from the KRR in (4) is evaluated as a recurrent model.

IV. Results

The results in this section are aimed at verifying the feasibility of the proposed approach, also stressing its overall benefits in terms of model accuracy. The considered test case is a plain CMOS driver composed of four cascaded stages which has demonstrated to be representative for the typical rich nonlinear dynamical behavior of this class of circuits.

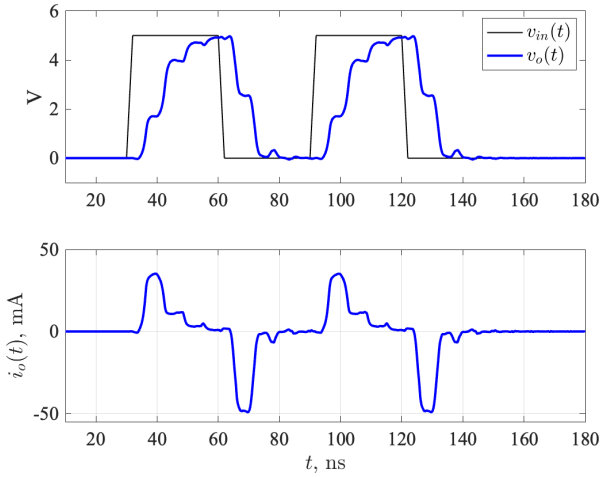

Fig. 2. Test waveforms: transient responses of the example device driven by a "010100" bit stream and loaded by an interconnect with $Z_0 = 50\Omega$ characteristic impedance and $T_d = 3$ ns delay.

Figure 2 shows the output port voltage $v_o(t)$ and current $i_o(t)$ responses of the example driver for a typical distributed load defined by a mismatched interconnect. The device is driven by a "010100" bitstream forced by a trapezoidal input signal $v_{in}(t)$. The sampled output port current response is assumed as the reference sequence for model testing (i.e., validation).

Figure 3 shows, for the same logic activity "010100" of the driver, the output port current and voltage responses observed for three different loads. In the figure, the paired voltage and current responses are labeled as set #1, #2 and #3. Two lumped loads and one distributed interconnect with characteristics different form the one considered for model testing are considered.

A first experiment is carried out by generating a KRR-based model built via the procedure outlined in Sec. III and by considering one set of training waveforms only (i.e. set #1 in this case). Figure 4 compares the reference output port current waveform to either the static mapping or the recurrent response obtained by the estimated KRR model, highlighting that a qualitative signature of the current response is captured only.

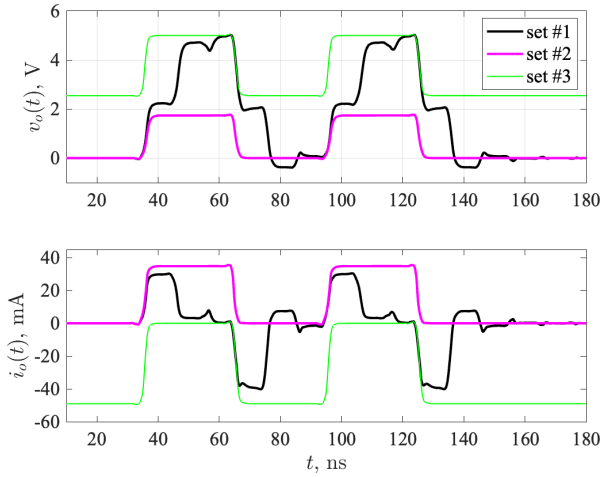

Fig. 3. Training waveforms: transient responses of the example device loaded by three different distributed or lumped loads different from the one considered for validation. Set #1: transmission line with $Z_0 = 75\,\Omega$ and $T_d = 5$ ns; set #2: shunt connection between a $50\,\Omega$ resistor and a 10 pF capacitor; set #3: $50\,\Omega$ in series with the power supply battery

One single dynamical load appeared not sufficient to allow the generalization of the KRR regression to obtain accurate results. The positive aspect is the nearly overlapped responses of the static mapping and the recurrent KRR prediction.

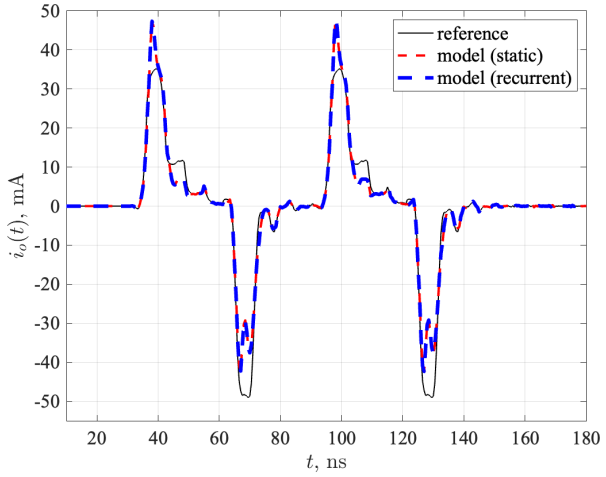

Fig. 4. Model validation (*single training sequence*): comparison between the reference output current of the test set shown in Fig. 2 and the model prediction obtained by both the static and the recurrent versions of the proposed machine learning regression. The training responses labeled as set#1 in Fig. 3 are used for the model generation.

A second experiment is instead carried out by generating a model with a procedure which uses all the three sets of device responses of Fig. 3, yielding a major improvement on model accuracy, as can be appreciated in the comparison of Fig. 5. This additional test confirms the potential of the proposed method in matching a very good accuracy of models. KRR regressions allows to implement a simple modeling procedure.

Also, the time required by model estimation is limited and it does not introduce critical practical aspects. The training time on a MacBook Pro (M1, 2022) is on the order of dozens of seconds (104 s for the latter test).

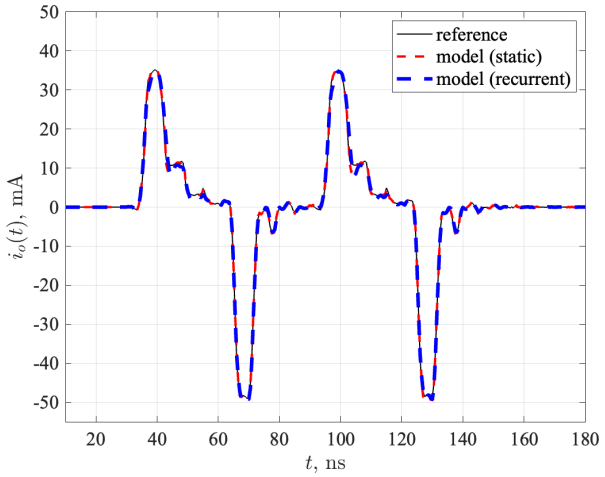

Fig. 5. Model validation (*multiple training sequences*): comparison between the reference output current of the test set shown in Fig. 2 and the model prediction obtained by both the static and the recurrent versions of the proposed machine learning regression. All the three training responses in Fig. 3 are used for the model generation.

V. Conclusions & Future Work

This paper investigates the effectiveness of an alternative approach for the modeling of IC buffers based on the KRR. Specifically, the KRR is used to learn the static map provided by a NARX model approximating the actual nonlinear dynamic characteristic of a digital IC driver. The regression model is trained by considering the observations of voltages and currents at the device ports for different load conditions defined with the aim of exploring the possible operating conditions of the device. The resulting model can be then evaluated as a recurrent model. The overall model accuracy has been assessed by considering a new load configuration (i.e., different from the load used during the training phase) by comparing the model predictions with the corresponding results obtained by SPICE. The results highlighted the excellent accuracy of the proposed model. On the other hand, additional research work has to be carried out to assess the effectiveness of the presented modeling framework on different structures (e.g., differential IC drivers and/or the inclusion of the power supply effects), as well as the possibility of integrating the resulting model in a SPICE-like solver.

References

[1] *I/O Buffer Information Specification, Ver. 7.2. Accessed: Dec. 12, 2023 [online]. Available: https://ibis.org/*

[2] G. Signorini, C. Siviero, M. Telescu, I.S. Stievano "Present and future of I/O-buffer behavioral macromodels," IEEE Electromagnetic Compatibility Magazine, vol. 5, no. 3, pp. 79–85, 2016.

[3] I.S. Stievano, I.A. Maio, F.G. Canavero, C. Siviero, "Reliable eye-diagram analysis of data links via device macromodels," IEEE Transactions on Advanced Packaging, vol. 29, no. 1, pp. 31–38, 2006.

[4] I.S. Stievano, I.A. Maio and F.G. Canavero, "Mπlog, macromodeling via parametric identification of logic gates," IEEE Transactions on Advanced Packaging, vol. 27, no. 1, pp. 15–23, Feb. 2004.

[5] T. Zhu, M.B. Steer and P.D. Franzon, "Accurate and Scalable IO Buffer Macromodel Based on Surrogate Modeling," IEEE Trans. Compon. Packag. Manuf. Technol., vol. 1, no. 8, pp. 1240–1249, Aug. 2011.

[6] B. Mutnury, M. Swaminathan and J.P. Libous, "Macromodeling of nonlinear digital I/O drivers," IEEE Transactions on Advanced Packaging, vol. 29, no. 1, pp. 102–113, Feb. 2006.

[7] H. Yu and M. Swaminathan, "A Bit-Time-Dependent Model of I/O Drivers for Overclocking Analysis," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 28, no. 7, pp. 1630–1637, 2020.

[8] H. Yu, T. Michalka, M. Larbi and M. Swaminathan, "Behavioral Modeling of Tunable I/O Drivers With Preemphasis Including Power Supply Noise," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 28, no. 1, pp. 233–242, Jan. 2020.

[9] W. Dghais, M. Souilem and M. Alam, "Mixed-Signal Overclocked I/O Buffers Model Abstraction for Signal Integrity Assessment," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 27, no. 3, pp. 691–699, March 2019.

[10] M. Souilem, N. Zgolli, T.R. Cunha, W. Dghais and H. Belgacem, "Signal and Power Integrity IO Buffer Modeling Under Separate Power and Ground Supply Voltage Variation of the Input and Output Stages," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 31, no. 6, pp. 874–886, June 2023.

[11] M. Souilem, J.N. Tripathi, W. Dghais and H. Belgacem, "An IBIS-like Modelling for Power/Ground Noise Induced Jitter under Simultaneous Switching Outputs (SSO)," *2019 IEEE 23rd Workshop on Signal and Power Integrity (SPI)*, Chambéry, France, 2019, pp. 1–4.

[12] Yi Cao, Runtao Ding and Qi-Jun Zhang, "State-space dynamic neural network technique for high-speed IC applications: modeling and stability analysis," in IEEE Trans. Microw. Theory Tech., vol. 54, no. 6, pp. 2398–2409, June 2006.

[13] Y. Cao and Q. -J. Zhang, "A New Training Approach for Robust Recurrent Neural-Network Modeling of Nonlinear Circuits,", IEEE Trans. Microw. Theory Tech., vol. 57, no. 6, pp. 1539–1553, June 2009.

[14] C. Siviero, R. Trinchero, S. Grivet-Talocia, G. Signorini, M. Telescu, "Constructive Signal Approximations for Fast Transient Simulation of Coupled Channels," IEEE Trans. Compon. Packag. Manuf. Technol., vol. 9, no. 10, pp. 2087–2096, 2019.

[15] C. Diouf, M. Telescu, I.S. Stievano, N. Tanguy, F.G. Canavero, "Simplified topology for integrated circuit buffer behavioural models," IET Circuits, Devices & Systems, vol. 11, no. 2. pp. 183–187, Mar. 2017.

[16] I.S. Stievano, I.A. Maio and F.G. Canavero, "On-the-fly Estimation of IC Output Port Macromodels," *2006 IEEE Workshop on Signal Propagation on Interconnects*, Berlin, Germany, 2006, pp. 109–112.

[17] J.A.K. Suykens, et al., *Least Squares Support Vector Machines*, World Scientific Pub Co Inc, 2002.

[18] J. A. K. Suykens and J. Vandewalle, "Recurrent least squares support vector machines,"in IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, vol. 47, no. 7, pp. 1109–1114, July 2000.

[19] N. Soleimani, R. Trinchero and F. G. Canavero, "Bridging the Gap Between Artificial Neural Networks and Kernel Regressions for Vector-Valued Problems in Microwave Applications," IEEE Trans. Microw. Theory Tech., vol. 71, no. 6, pp. 2319-2332, 2023.

[20] A. Rudi, L. Carratino, and L. Rosasco, "Falkon: An optimal large scale kernel method," Advances in neural information processing systems, vol. 30, 2017.

[21] J. Shawe-Taylor and N. Cristianini, *Kernel methods: an overview.* Kernel Methods for Pattern Analysis, Cambridge: Cambridge University Press, 2004, pp. 25–46.

[22] R. Trinchero, M. Larbi, H. M. Torun, F. G. Canavero and M. Swaminathan, "Machine Learning and Uncertainty Quantification for Surrogate Models of Integrated Devices With a Large Number of Parameters," in IEEE Access, vol. 7, pp. 4056-4066, 2019.