Behavioral macromodeling of high-speed drivers via compressed tensor representations

Behavioral Macromodeling of High-Speed Drivers via Compressed Tensor Representations

C. Siviero†, S. Grivet-Talocia†, I. S. Stievano†, G. Signorini‡
† Dept. Electronics and Telecommunications, Politecnico di Torino, Torino, Italy
‡ Intel Corporation, Munich, Germany

***Abstract*—This paper addresses the behavioral modeling of digital drivers for Signal and Power Integrity co-simulations. State-of-the-art two-piece model representations are combined with a compact description of the device static characteristics. The latter are considered as multivariate mappings that are functions of the device electrical variables, and of additional parameters defining process corners and device settings. Overall model complexity is reduced through a compressed tensor representation obtained via a high-order singular value decomposition. Several application examples demonstrate the feasibility and the advantages of the proposed approach.**

***Index Terms*—Digital integrated circuits, I/O buffers, signal and power integrity, macromodeling, circuit simulation, singular value decomposition.**

I. Introduction

The robust design of high-performance electronic devices requires the assessment of system performance since early design phases. Such an assessment is mainly achieved via system-level numerical simulations, in order to accurately predict the transient evolution of the analog waveforms at signal and power pins, thus enabling verification of design constraints and, if necessary, design optimization. In this framework, the availability of efficient numerical models describing the external behavior of digital integrated circuits (ICs) plays a key role, thus becoming a primary target for designers.

Detailed models of real transceivers based on transistor-level descriptions cannot be effectively used in a system-level simulation environment. On one hand, these models are overly complex and hardly portable. Moreover, they are seldom available, since they disclose information on the internal structure and technology of devices that represent confidential proprietary information of IC suppliers. Most of the above limitations can be overcome by resorting to behavioral or surrogate models. Such models try to mimic the electrical behavior of devices, as can be observed from "external" ports, often providing an excellent compromise between complexity and accuracy. Different approaches are currently in use for generating behavioral models, that can be simulated in any electronic design automation tool [1], [2], [3].

Most state-of-the-art approaches share the common limitation of using custom solutions for modeling both the static and the dynamic characteristics of devices, with an unavoidable impact on accuracy, and expecially compactness and efficiency. This problem is even more relevant when the effects of different Process-Voltage-Temperature (PVT) operating conditions and device settings are considered. This paper proposes a framework based on high-order singular value decompositions, that is able to compress the large amount of information that is required to describe the multivariate driver characteristics, without any impact on model portability and implementation.

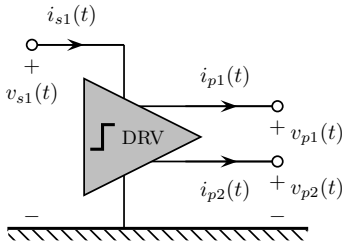

Fig. 1. Schematic illustration of a multi-pin driver

II. Problem statement and model structure

We consider a general transceiver topology, depicted in Fig. 1 for the particular case of a high-speed driver. The structure interacts with its environment through a set of output ports (henceforth labeled with subscript $_p$), and a set of supply ports (subscript $_s$). For instance, a single-ended driver characterized by a single power domain will have a single output port with variables (v_{p1}, i_{p1}) and a single supply port with variables (v_{s1}, i_{s1}). A differential driver will instead be characterized by a pair of output ports, with two port voltages collected as components of vector $\boldsymbol{v}_p$, and corresponding port currents in vector $\boldsymbol{i}_p$. A similar generalization holds for structures with multiple power domains.

In this work, we adopt the well-established two-piece model representation, which has been documented in [1], [2], [3] to be appropriate for the behavioral representation of high-speed drivers. Denoting with $i_\mu(t)$ any output port current, we consider the following representation

$$i_\mu(t) = w_\mu^H(t) f_\mu^H(\boldsymbol{v}_p, \boldsymbol{v}_s, \boldsymbol{\lambda}) + w_\mu^L(t) f_\mu^L(\boldsymbol{v}_p, \boldsymbol{v}_s, \boldsymbol{\lambda}), \quad (1)$$

where superscript $\nu = \{H, L\}$ labels submodels that characterize the device in fixed high (H) or low (L) logic state. Each of these submodels is further split as

$$f_\mu^\nu(\boldsymbol{v}_p, \boldsymbol{v}_s, \boldsymbol{\lambda}) = F_\mu^\nu(\boldsymbol{v}_p, \boldsymbol{v}_s, \boldsymbol{\lambda}) + g_\mu^\nu(\boldsymbol{v}_p, \boldsymbol{v}_s, \boldsymbol{\lambda}; \mathsf{D}), \quad (2)$$

where F_μ^ν denotes a static submodel describing the device under DC steady-state operation, and where g_μ^ν is a linear dynamic submodel, as emphasized by the presence of the time derivative operator $\mathsf{D} = d/dt$. The additional vector argument $\boldsymbol{\lambda}$ in (1) and (2) denotes any parameter upon which the device behavior may depend, such as temperature and process variables. The submodels g_μ^ν are linear state-space representations, estimated via standard system identification methods.

The model representation (1) can also be used for the description of the power supply currents (such as the current i_{s1} in the scheme of Fig. 1), possibly with an additional term accounting for the current absorbed by previous driver stages. The readers are referred to [2] and references therein for additional details.

III. Static characterization and modeling via tensor compression

Any of the static maps F_μ^ν in (2) can be described in abstract notation as a multivariate map

$$y = F(\boldsymbol{x}), \tag{3}$$

where the input vector $\boldsymbol{x} = (x_1, \ldots, x_N)^\mathsf{T}$ collects all port (output and supply) voltages, as well as additional parameters. An accurate behavioral macromodel should capture the variation of the output variables y with respect to all components x_n. To this end, standard characterization approaches perform a set of DC simulations, by computing the static response of the device by fixing each of the independent variables $x_n = X_{j_n}$ with $j_n = 1, \ldots, J_n$ and $n = 1, \ldots, N$. The result is a multidimensional tensor $\boldsymbol{\mathcal{Y}}$ with elements

$$Y_{j_1,\ldots,j_N} = F(X_{j_1}, \ldots, X_{j_N}) \tag{4}$$

with N being the *order*, and J_n being the dimension (number of components) along the n-th *mode* or *direction*. The complexity of the resulting dataset, which includes $|\boldsymbol{\mathcal{Y}}| = \prod_{n=1}^N J_n$ independent data points, must be reduced in order to obtain a tractable model.

Standard macromodeling approaches compute a parametric representation of $\boldsymbol{\mathcal{Y}}$, e.g., as a superposition of some multivariate basis functions, through some fitting process. Such an approach is however hardly scalable to orders N larger than 2 or 3 at most, due to a curse of dimensionality. The approach that we pursue in this paper is aimed at a dimensionality reduction of the tensor $\boldsymbol{\mathcal{Y}}$, by seeking an approximate representation that compresses the dataset before proceeding to any subsequent parametric identification and/or approximation. The process is described for the case $N = 2$ in Sec. III-A and extended to the general case in Sec. III-B.

A. The two-dimensional case

This scenario applies, e.g., to the case of a single-ended driver with a single power supply. In such case, the two independent variables are the output voltage $x_1 = v_p$ and the supply voltage $x_2 = v_s$. A double DC sweep leads to a tensor dataset with order $N = 2$, which is nothing else than a matrix $\mathbf{Y} \in \mathbb{R}^{J_1 \times J_2}$, with elements Y_{j_1,j_2}. A well-known result in linear algebra states that the optimal approximation of $\mathbf{Y}$ with a matrix $\bar{\mathbf{Y}}$ having fixed rank ρ is provided by the truncated Singular Value Decomposition (SVD)

$$\mathbf{Y} \approx \bar{\mathbf{Y}} = \mathbf{U}_1 \boldsymbol{\Sigma} \mathbf{U}_2^\mathsf{T}, \tag{5}$$

where $\boldsymbol{\Sigma} = \mathrm{diag}\{\sigma_1, \ldots, \sigma_\rho\}$ collects the largest ρ singular values. The columns $\boldsymbol{u}_{n,k_n}$ with $k_n = 1, \ldots, \rho$ of the two (orthogonal) matrices $\mathbf{U}_n \in \mathbb{R}^{J_n \times \rho}$ for $n = 1, 2$ collect the corresponding singular vectors. The above approximation minimizes the induced 2-norm of the residual $\|\bar{\mathbf{Y}} - \mathbf{Y}\|_2$ and is therefore optimal [4].

B. The general case

Here, we discuss how to generalize the SVD-based approximation (5) to a generic higher order $N > 2$. We start by rewriting (5) in the more abstract form

$$\mathbf{Y} \approx \bar{\mathbf{Y}} = \boldsymbol{\Sigma} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2, \tag{6}$$

where the operator $\times_n$ performs matrix multiplication along the n-th direction (here, $n = 1$ for rows and $n = 2$ for columns). The generalization of (6) to approximate a given tensor $\boldsymbol{\mathcal{Y}}$ with order $N > 2$ is straightforward [5], [6], as

$$\boldsymbol{\mathcal{Y}} \approx \bar{\boldsymbol{\mathcal{Y}}} = \boldsymbol{\mathcal{S}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \cdots \times_N \mathbf{U}_N \tag{7}$$

where the orthogonal matrices $\mathbf{U}_n \in \mathbb{R}^{J_n \times \rho_n}$ for $n = 1, \ldots, N$ multiply the *core* tensor $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{\rho_1 \times \rho_2 \cdots \times \rho_N}$ along the n-th direction. Note that, differently from (6), the core tensor $\boldsymbol{\mathcal{S}}$ is in general full and can be characterized by a different size ρ_n along each direction. The columns $\boldsymbol{u}_{n,k_n}$ of each matrix $\mathbf{U}_n$ can be interpreted as an orthogonal basis of the subspace that approximates the collection all vectors obtained by freezing all indices of $\boldsymbol{\mathcal{Y}}$ except along the n-th direction (also called n-th mode *fibers*). The componentwise expansion of (7) reads

$$\bar{Y}_{j_1,j_2,\ldots,j_N} = \sum_{k_1=1}^{\rho_1} \sum_{k_2=1}^{\rho_2} \cdots \sum_{k_N=1}^{\rho_N} S_{k_1,k_2,\ldots,k_N} (\mathbf{U}_1)_{j_1,k_1} (\mathbf{U}_2)_{j_2,k_2} \ldots (\mathbf{U}_N)_{j_N,k_N} \tag{8}$$

This expression shows that the original tensor $\boldsymbol{\mathcal{Y}}$ is represented by a much smaller tensor $\boldsymbol{\mathcal{S}}$ with $|\boldsymbol{\mathcal{S}}| = \prod_{n=1}^N \rho_n$, plus a collection of N basis sets, each having ρ_n vector elements. An effective data compression is achieved if $\rho_n \ll J_n$ for each direction n. The quality of the approximation can be measured by the Frobenius norm, defined as

$$\|\bar{\boldsymbol{\mathcal{Y}}} - \boldsymbol{\mathcal{Y}}\|_F^2 = \sum_{j_1,\ldots,j_n} |\bar{Y}_{j_1,j_2,\ldots,j_N} - Y_{j_1,j_2,\ldots,j_N}|^2 \tag{9}$$

The computation of (7) is here performed according to an Alternating Least Squares (ALS) algorithm [5], which refines an initial estimate of the matrices $\mathbf{U}_n$ by iterative reprojection of the original tensor along the subspaces available from previous iterations.

C. Static model construction

Once the approximation (7) is available, the components of $\mathbf{U}_n$ are combined with the corresponding input parameter values to construct a collection of one-dimensional datasets $\Omega_{n,k_n} = \{[X_{j_n}, (\mathbf{U}_n)_{j_n,k_n}],\ j_n = 1,\dots,J_n\}$, with one dataset for each $n = 1,\dots,N$ and $k_n = 1,\dots,\rho_n$. A corresponding parametric submodel $\varphi_{n,k_n}(x_n)$ is obtained through a piecewise linear interpolation process applied to Ω_{n,k_n}, and the approximation to the original map (3) is constructed as

$$y \approx \sum_{k_1=1}^{\rho_1} \cdots \sum_{k_N=1}^{\rho_N} S_{k_1,\dots,k_N} \varphi_{1,k_1}(x_1) \dots \varphi_{N,k_N}(x_N). \quad (10)$$

This result is an approximation based on one-dimensional submodels, assembled through a multidimensional tensor product, with coefficients available from the core tensor $\boldsymbol{\mathcal{S}}$. Due to the limited number of such coefficients, an equivalent circuit implementation of (10) becomes viable through behavioral voltage-controlled current sources.

IV. Results

In this section, the proposed modeling approach is applied to the output and supply ports of two CMOS devices, a differential driver $\mathcal{D}_A$ ($V_{\rm DD} = 3.3\,$V) and a single-ended LPDDR3 driver $\mathcal{D}_B$ ($V_{\rm DD} = 1.2\,$V). The responses of the corresponding transistor-level descriptions obtained by HSPICE are assumed as the reference curves hereafter.

As a first validation test, Figure 2 shows the static characteristic of the current i_{1H} at the non-inverting output terminal of the differential driver $\mathcal{D}_A$, forced to operate in the H state. The three-dimensional surface of Fig. 2, that also includes the information on the approximation error, clearly highlights the accuracy of the proposed approach, which provides a very accurate representation with a relatively small number of components $(8, 8, 6)$ along each direction (v_{p1}, v_{p2} and v_{s1}), respectively. The same accuracy level was obtained for all the other current static characteristics.

Following the guidelines reported in [2], a SPICE macromodel belonging to the class (1) has been synthesized and validated in a realistic transient simulation. The selected validation setup includes the switching driver, terminated at the output ports by a mismatched coupled transmission line structure. The supply port is connected to an ideal $V_{\rm DD}$ battery through a realistic power distribution network described by a lumped equivalent. Figure 3 compares the macromodel to the reference, respectively for the differential voltage at the output ports $v_d = v_{p1} - v_{p2}$ and for the supply voltage v_{s1}. The accuracy of the macromodel predictions is excellent.

Figure 4 collects a similar comparison for the single-ended driver $\mathcal{D}_B$, thus confirming the accuracy of the proposed solution for different device types. The proposed models offer a remarkable simulation speed up (for the latter case, $20\times$ faster than transistor-level models). We conclude that the proposed approch provides a simple yet effective solution for drastic complexity reduction in Signal and Power Integrity co-simulations.

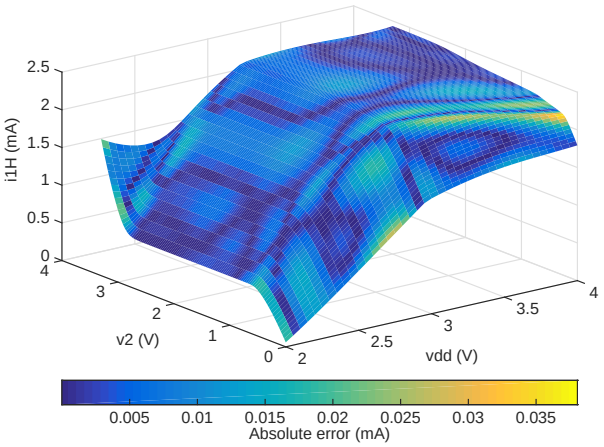

Fig. 2. Static characteristic of the current $i_1 = i_{p1}$ of the H submodel for the differential driver $\mathcal{D}_A$, plotted versus $v_2 = v_{p2}$ and $v_{\rm dd} = v_{s1}$, with $v_1 = v_{p1} = 1.65$ V. The colormap depicts the pointwise absolute error of the compressed model with ranks $(8, 8, 6)$.

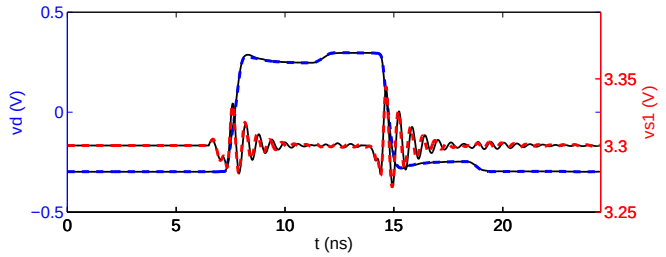

Fig. 3. Output differential voltage $v_{\rm d}(t) = v_1(t) - v_2(t)$ and corresponding power supply fluctuation $v_{s1}(t)$ for driver $\mathcal{D}_A$. Solid line: reference; dashed line: model.

References

[1] IBIS (I/O Buffer Information Specification) – Ver. 5.1, http://www.eda.org/ibis/, Aug. 24, 2011.

[2] I. S. Stievano, I. A. Maio, and F. G. Canavero, "Mπlog, macromodeling via parametric identification of logic gates", *IEEE Trans. on Adv. Packag.*, Vol. 27, No. 1, pp. 15–23, Feb. 2004.

[3] B. Mutnury, M. Swaminathan, and J. P. Libous, "Macromodeling of nonlinear digital I/O drivers," *IEEE Trans. on Adv. Packag.*, Vol. 29, No. 1, pp. 102–113, Feb. 2006.

[4] G. W. Stewart and J. Sun, *Matrix perturbation theory*, Academic Press, Boston, 1966.

[5] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications", *SIAM Review*, Vol. 51, No. 3, 2009, pp. 455–500.

[6] L. R. Tucker, "Some mathematical notes on three-mode factor analysis", *Psychometrika*, Vol. 31, 1966, pp. 279–311.

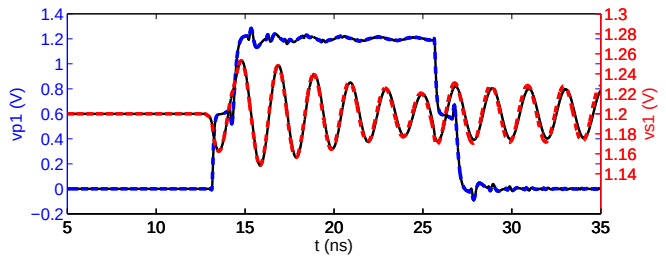

Fig. 4. Output port voltage response $v_{p1}(t)$ and corresponding power supply fluctuation $v_{s1}(t)$ for driver $\mathcal{D}_B$. Solid line: reference; dashed line: model.