POLITECNICO DI TORINO
Repository ISTITUZIONALE

Comparative Analysis of Prior Knowledge-Based Machine Learning Metamodels for Modeling Hybrid Copper–Graphene On-Chip Interconnects

Original
Comparative Analysis of Prior Knowledge-Based Machine Learning Metamodels for Modeling Hybrid Copper–Graphene On-Chip Interconnects / Kushwaha, Suyash; Soleimani, Nastaran; Treviso, Felipe; Kumar, Rahul; Trinchero, Riccardo; Canavero, Flavio; Roy, Sourajeet; Sharma, Rohit. - In: IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY. - ISSN 0018-9375. - STAMPA. - 64:6(2022), pp. 2249-2260. [10.1109/TEMC.2022.3205869]

Availability:
This version is available at: 11583/2973545 since: 2022-12-05T10:35:11Z

Publisher:
IEEE

Published
DOI:10.1109/TEMC.2022.3205869

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

01 May 2026

IEEE EMC SOCIETY
TRANSACTIONS ON
Electromagnetic Compatibility

Comparative Analysis of Prior Knowledge Based Machine Learning Metamodels for Modeling Hybrid Copper-Graphene On-Chip Interconnects

Journal:	*Transactions on Electromagnetic Compatibility*
Manuscript ID	TEMC-167-2022
Manuscript Type:	Regular Paper
Date Submitted by the Author:	11-Apr-2022
Complete List of Authors:	Kushwaha, Suyash; Indian Institute of Technology Ropar, Electrical Engineering Soleimani, Nastaran; Politecnico di Torino Treviso, Felipe; Politecnico di Torino, Dip. Elettronica Kumar, Rahul; Indian Institute of Technology Ropar, Electrical Engineering; IIT Ropar, EE Trinchero, Riccardo; Politecnico di Torino, Electronics and Telecommunications Canavero, Flavio; Politecnico di Torino, Dip. Elettronica; Roy, Sourajeet; Indian Institute of Technology Roorkee, ECE; Indian Institute of Technology Roorkee Sharma, Rohit; Indian Institute of Technology Ropar, Electrical Engineering
Key Words:	High speed interconnects < Signal integrity and power integrity (SIPI)

SCHOLARONE™
Manuscripts

Authors' Response to Key Paper Submission Questions

Has this manuscript been submitted previously to any journal? **If your paper was submitted previously to T-EMC, enter the Manuscript ID of the previous submission.** **If your paper was submitted to another journal, enter the Journal Title.** **If none of these apply, enter N/A.**	No
Is this paper an extension of the author's conference paper?	Yes
If yes, please indicate if the conference paper is:	
Cited and discussed in the manuscript?	Yes
Has been extended with a significant amount of new material?	Yes
Included with the submission?	No
What research problem is addressed by the manuscript and why is it important to the EMC community?	1. In this paper, machine learning metamodels have been developed in order to predict the per-unit-length (p. u. l.) parameters of hybrid copper-graphene on-chip interconnects based on their structural geometry and layout. 2. The salient feature of all these machine learning metamodels is that they exploit the prior knowledge of the p. u. l. parameters of the interconnects obtained from cheap empirical models to reduce the number of expensive full-wave electromagnetic (EM) simulations required to extract the training data. 3. Thus, the proposed machine learning metamodels are referred to as prior knowledge-based machine learning (PKBML) metamodels.
What is the novelty of your work?	In this paper, two different prior knowledge-based machine learning (PKBML) approaches are developed and their performance relative to conventional ML approaches and full-wave EM simulations are explored for predicting the p. u. l. parameters of hybrid copper-graphene on-chip

	interconnects. In particular, the PKBML approaches is tested for ANN, SVM, and LS-SVM metamodels for thoroughness of analysis. From the detailed explorations performed, it is concluded that the PKI approach best accelerates the training of all ML metamodels provided no noise is in present in the data. However, in the presence of noise, the SD approach ends up providing better acceleration.

Comparative Analysis of Prior Knowledge Based Machine Learning Metamodels for Modeling Hybrid Copper-Graphene On-Chip Interconnects

Suyash Kushwaha, Nastaran Soleimani, *Student Member, IEEE*, Felipe Treviso, *Student Member, IEEE*, Rahul Kumar, Riccardo Trinchero, *Member, IEEE*, Flavio Canavero, *Life Fellow, IEEE*, Sourajeet Roy, *Senior Member, IEEE* and Rohit Sharma, *Senior Member, IEEE*

***Abstract* — In this paper, machine learning metamodels have been developed in order to predict the per-unit-length (p. u. l.) parameters of hybrid copper-graphene on-chip interconnects based on their structural geometry and layout. Machine learning metamodels within the context of this article include artificial neural networks (ANNs), support vector machines (SVMs), and least-square support vector machines (LS-SVMs). The salient feature of all these machine learning metamodels is that they exploit the prior knowledge of the p. u. l. parameters of the interconnects obtained from cheap empirical models to reduce the number of expensive full-wave electromagnetic (EM) simulations required to extract the training data. Thus, the proposed machine learning metamodels are referred to as prior knowledge-based machine learning (PKBML) metamodels. The PKBML metamodels offer the same accuracy as conventional ML metamodels trained exclusively by full-wave EM solver data, but at the expense of far smaller training time costs. In this paper, detailed comparative analysis of the proposed PKBML metamodels have been performed using multiple numerical examples.**

***Index Terms*— Artificial neural networks (ANN), copper-graphene interconnects, least-square support vector machine (LS-SVM), per-unit-length parameters, support vector machine (SVM), transient simulation.**

I. Introduction

When IC technology nodes scale below the 22 nm mark, the performance of on-chip copper interconnects start being affected by various scattering mechanisms such as grain boundary scattering, surface roughness scattering, and sidewall and top/bottom surface scattering [1]. These scattering mechanisms significantly increase the value of the per-unit-length resistance of the interconnects from their nominal bulk value. Increase in the resistance of interconnects results in more signal attenuation, power losses, and latency [1]-[4]. Moreover, when the copper ions from the interconnects get diffused into the dielectric layer, dielectric conductivity and leakage losses increases while discontinuities are created in the interconnects [5]. As a remedy for the diffusion of copper ions, a barrier layer of tantalum (Ta) or tantalum nitride (TaN) is generally placed around the copper trace [3], [6], [7]. The conductivity of this barrier layer is usually lower than copper, and hence, electrons flow is difficult through the barrier layers. In addition, by using these barriers layers around the copper trace, the effective cross-sectional area of the interconnect is decreased, which ultimately results in lowering the current carrying capacity of the interconnect.

Explorations into better barrier layer materials has led to the use of graphene nanoribbons as barrier layers around copper interconnects [8]-[12]. Due to the long mean free path of graphene compared to copper, the graphene barrier layers exhibit lower scattering resistances, and thereby, significantly lower the equivalent resistance of the trace plus the barrier layers [11], [12]. Moreover, graphene barrier layers being ultra-thin, they can stop diffusion of copper ions into the dielectric without significantly reducing the effective cross-sectional area of the interconnect. Finally, graphene barrier layers possess good thermal stability at very high temperatures. These advantages of using graphene barrier layers around a copper trace makes hybrid copper-graphene interconnects a good candidate for on-chip applications.

When designing hybrid on-chip copper-graphene interconnects for peak performance, design space exploration of the interconnects in SPICE is essential. Typically, SPICE simulation of interconnects is done in two-steps. At the very first step, the per-unit-length (p. u. l.) parameters of the interconnects are extracted from a quasi-TEM approximation of the interconnect structure, usually using full-wave electromagnetic (EM) solvers [13], [14]. In the second step, the extracted p. u. l. parameters are used in SPICE multiconductor transmission lines (MTL) representative models of the

Suyash Kushwaha and Rohit Sharma are with the Department of Electrical Engineering, Indian Institute of Technology Ropar, Rupnagar 140001, India. (e-mail: suyash.20eez0016@iitrpr.ac.in; rohit@iitrpr.ac.in).

Nastaran Soleimani, Felipe Treviso, Riccardo Trinchero and Flavio Canavero are with the Department of Electronics and Telecommunication, Politecnico de Turino, Corso Duca degli Abruzzi, 24 10129 Torino, Italy. (e-mail: nastaran.soleimani@polito.it; felipe.treviso@polito.it; riccardo.trinchero@polito.it; flavio.canavero@polito.it).

Rahul Kumar and Sourajeet Roy are with the Department of Electronics and Communication Engineering, Indian Institute of Technology Roorkee, Roorkee 247667, India. (e-mail: rahul.kumar@iitrpr.ac.in; sourajeet.roy@ece.iitr.ac.in).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org

interconnects for performing a complete transient analysis [13], [14]. In this two-step approach, the computational time cost of the first step is usually much larger than that of the second step. For the case of design space explorations, the two-step simulation process of above needs to be repeated thousands of times with different values of the geometrical, physical, and material parameters of the interconnects, thus leading to even larger simulation time costs.

To address the high simulation time costs of design space explorations of interconnects, machine learning (ML) regression based surrogate models or metamodels have been used [30], [34]. Interconnect modeling using artificial neural networks (ANNs) and support vector machine (SVM) regression have already been reported in several works [22]-[25]. The basic idea behind these ML metamodels is that once they are trained, they can accurately and analytically predict the p. u. l. parameters of interconnects for different value of the geometrical, physical, and material parameters of the structure without having to resort to full-wave EM simulations. Therefore, ML metamodels can perform parametric sweeps and design space explorations of interconnects at a fraction of the time cost required by the aforementioned conventional two-step process using SPICE.

Despite the obvious advantage of ML metamodels, the one challenge they face is that they are very data hungry. In effect, they need to be thoroughly trained on massive datasets for them to be reliably accurate. These massive training datasets are usually generated using many repeated full-wave EM simulations at extremely high time costs [22]. To reduce this large computational cost of training ML metamodels, a variety of methods such as the source difference method [25], prior knowledge input method [26], [27], and space-mapping [28], [29] have been reported in the literature. The basic requirement for all of these methods to work is the availability of computationally cheaper empirical models, hereafter referred to as low-fidelity models. These low-fidelity models sacrifice their predictive accuracy for numerical efficiency. Thus, they can give a fast but only an approximate prediction of the p. u. l. parameters of interconnects based on the design parameters of the structure. These approximate predictions can, however, be then refined using ML metamodels trained with very small number of full-wave EM solver data. Therefore, ML models can be trained more efficiently by strategically using less EM solver data and more empirical model data [25]-[30]. These ML metamodels are known as prior knowledge-based machine learning (PKBML) metamodels. Recently, PKBML metamodels using the source difference (SD) method have been applied to the problem of predicting the p. u. l. parameters of hybrid copper-graphene interconnects [34]. However, one possible challenge of the SD method is that it only supports a linear correlation between the results of the low-fidelity model and those of the full-wave EM simulations. As a result, the SD approach does not provide the best possible accuracy.

In this paper, three key contributions to the state-of-the-art in ML based modeling of the p. u. l. parameters of hybrid copper-graphene interconnects are presented.

(i) The above accuracy issue of the SD method is addressed by using a more general prior knowledge input (PKI) method for predicating the p. u. l. parameters. The PKI method has the capacity to include nonlinear correlation between the results of the low-fidelity model and those of the full-wave EM simulations. As a result, the PKI method can provide smaller modeling errors compared to the SD method for the same number of training points. The improved performance of the PKI method compared to the SD method is demonstrated in this paper using various numerical examples. This contribution of the paper is based on the authors preliminary work of [31].

(ii) In this paper, the comparative analysis between the SD and PKI methods is performed across different ML metamodels such as the ANN, SVM, and least square SVM (LS-SVM) metamodels. Such a detailed comparative analysis was not included in the earlier work of [34]. In fact, such a thorough comparative analysis of different PKBML metamodels for interconnects in general and hybrid copper-graphene interconnects in particular have not been reported before.

(iii) Finally, in this paper, the comparative analysis between the SD and PKI methods across different ML metamodels is conducted both in the presence and absence of simulation noise. Such an analysis reveals critical insights as to which combination of PKBML method and ML metamodeling technique is most suitable depending on the level of noise in the training data – something that has not been reported in [34].

The organization of this paper is as follows: Section II presents a brief overview of the different ML metamodels used in this paper (i.e., the ANN, SVM, and the LS-SVM metamodels). Section III describes the development of the proposed PKI method for the above metamodels and compares its performance with conventional ML metamodels as well as the SD method. Section IV includes all details of the low-fidelity model used for the SD and PKI methods specifically for hybrid copper-graphene interconnects. Section V covers the numerical examples and Section VI concludes the paper.

II. Review of Conventional Machine Learning Metamodels for Interconnects

Performing transient analysis of high-speed interconnects using circuit simulators such as SPICE is essential for signal integrity verification [13], [14]. SPICE models of interconnects are based on the concept of multiconductor transmission lines and require the prior knowledge of the p. u. l. parameters of the interconnect structure for their solution [35]. Unfortunately, the p. u. l. parameters of interconnects are extracted from full-wave EM simulations of the interconnect structure at high computational time costs [22], [23]. This problem is exacerbated during design space explorations where the SPICE simulation of the interconnects, and consequently, the p. u. l.

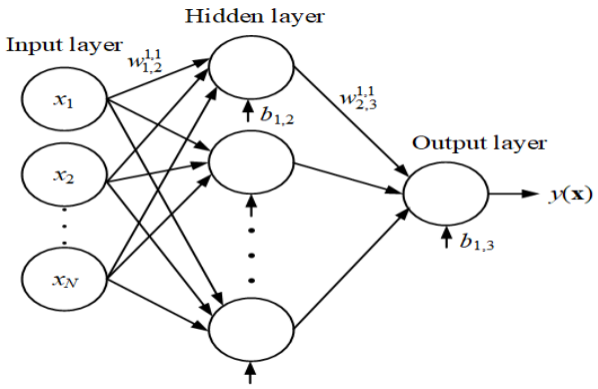


Fig. 1: Basic three-layer architecture of ANN

parameter extraction has to be repeated for thousands of different values of the design parameters of the interconnects. In such scenarios, ML based surrogate models or metamodels can be employed to emulate the p. u. l. parameters of interconnects as analytic functions of the design parameters [30], [31], [34]. These metamodels enable the analytic calculation of the p. u. l. parameters of interconnects for different values of the design parameters free from expensive full-wave EM simulations. For this reason, ML metamodels are seen as powerful tools for design space exploration of interconnects [15]-[21].

In order to explain how ML metamodels work, consider the problem of approximating the unknown output $y \in \Re$ of a general non-linear system as

$$y = M(\mathbf{x}) \tag{1}$$

where M is the input-output model of the system and $\mathbf{x} = [x_1, \mathrm{K}, x_N]^T \in \Re^N$ is a vector containing the normalized representation of the system parameters for $-1 \le x_i \le 1$ and i = 1, …, N. For example, in the case of interconnects, the variable y can be one of the entries of the p. u. l. resistance, inductance, or capacitance matrices of the structure and the vector **x** can represent the geometrical, physical, and material parameters of the structure. Now, ML metamodels approximate the input-output behavior of the nonlinear model M using a closed-form analytical map $\tilde{M}$. This analytic map is constructed from the training dataset $D = \{(\mathbf{x}_l, y_l)\}_{l=1}^{L}$ extracted using rigorous solvers (e.g., a full-wave EM solver for interconnects) such that

$$y_l \approx \tilde{M}(\mathbf{x}_l) \tag{2}$$

Several ML metamodels can be used to identify the analytic map $\tilde{M}$ of (2). This paper will consider three different metamodels: the ANN, the SVM, and the LS-SVM regressions which are reviewed next.

A. ANN Metamodels

A multi-layer perceptron (MLP) architecture of an ANN is illustrated in Fig. 1. The architecture shown in the figure is a feed-forward network in which all the neurons are arranged in multiple layers. The first layer on the left-hand side is called the input layer and the last layer on the right-hand side is called the output layer. All the layers in between are called hidden layers. The neurons making up each hidden layer are defined by non-linear activation functions, whereas the activation function of the output layer is generally linear. The input layer takes the values of **x** corresponding to a certain design point and propagates these values through the hidden layers. In the hidden layers, these values undergo nonlinear transformations till the output layer predicts the desired output (i.e., $y = \tilde{M}_{ANN}(\mathbf{x})$). Indeed, assuming a single hidden layer, the predicted output takes the form

$$y = \tilde{M}_{ANN}(\mathbf{x}) = \sigma_3\left(b_{1,3} + \sum_{i=1}^{N_h} w_{2,3}^{(i,1)} \sigma_2\left(b_{i,2} + \sum_{j=1}^{N} w_{1,2}^{(j,i)} x_j\right)\right) \tag{3}$$

In (3), σ_p refers to the nonlinear activation function used in the neurons of the p-th layer, $b_{p,q}$ is the bias value entering the p-th neuron of the q-th layer, and $w_{p,q}^{\alpha,\beta}$ is the synaptic weight linking the α-th neuron of the p-th layer to the β-th neuron of the q-th layer. The goal of the ANN metamodel of Fig. 1 is to tune the values of all weight and bias terms in (3) to minimize the error loss function [15]

$$f_{ANN}(\mathbf{w}, \mathbf{b}) = \arg\min_{\mathbf{w},\mathbf{b} \in \Re}\left(\frac{1}{L}\sum_{l=1}^{L}\left(y_l - \tilde{M}_{ANN}(\mathbf{x}_l)\right)^2\right) \tag{4}$$

where (**w**, **b**) refer to the real valued set of weights and bias terms in the ANN architecture (i.e., the weights and bias terms in (3)). This process is referred to as training of the ANN metamodel. Typically, the optimization of (4) can be performed using standard gradient-based back-propagation techniques [22], [33]. In this work, a thorough investigation of all the ANNs is done and it is found that the hyperbolic tangent activation function is best candidate function for predicting the p. u. l. values of hybrid copper-graphene interconnects. This activation function takes the form

$$\sigma_2(z) = \frac{e^{2z} - 1}{e^{2z} + 1} \tag{5}$$

in equation (3).

B. SVM and LS-SVM Regression Metamodels

The SVM and LS-SVM regressions admit the same primal space formulation given by [36]

$$\tilde{y} \approx M_{(LS)SVM}(\mathbf{x}) = \langle \mathbf{w}, \mathbf{\Phi}(\mathbf{x}) \rangle + b, \tag{6}$$

where $\mathbf{\Phi}(\mathbf{x}) = [\phi_1(\mathbf{x}), ..., \phi_P(\mathbf{x})]^T$, $\mathbf{w} = [w_1, ..., w_P]^T$, and b are vectors collecting the basis functions, regression coefficients, and bias term, respectively. The unknowns **w** and b are estimated by solving the following optimization problem:

$$\min_{w,b} \frac{1}{2}\|\mathbf{w}\|_{L_2}^2 + \gamma \sum_{l=1}^{L} \ell(y_l, \langle \mathbf{w}, \mathbf{\Phi}(\mathbf{x}_l) \rangle + b), \tag{7}$$

where $\ell(y_l, \langle \mathbf{w}, \mathbf{\Phi}(\mathbf{x}_l) \rangle + b)$ is a predefined error loss function computed over the training samples and γ is the regularizer term [36]. The latter is the regression hyperparameter which must be tuned during the training phase to reduce the model variance and thus prevent overfitting.

The difference between the SVM and LS-SVM depends on the loss function used. While for the LS-SVM, $\ell(\cdot)$ is the traditional squared loss divided by two, the SVM uses the so-called linear ε-insensitive loss. The latter adds a penalty equal to the excess model $\ell 1$-error with respect to ε. For this reason, the region [-ε, +ε] is called ε-insensitive zone. Interested readers are directed to the literature [18], [19], [36] regarding the differences between SVM and LS-SVM.

At this point, it is important to remark that the primal space formulation in (6) provides a parametric model, i.e., the number of coefficients that should be estimated is equal to the number of basis functions P. However, both SVM and LS-SVM metamodels also allow an equivalent dual space formulation expressed as

$$\tilde{y} \approx M_{(LS)SVM}(\mathbf{x}) = \sum_{l=1}^{L} \alpha_l k(\mathbf{x}, \mathbf{x}_l) + b \tag{8}$$

where the number of model coefficients $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_L]^T$ is equal to L (i.e., the number of samples) and $k(\mathbf{x}, \mathbf{x}_l) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ is the co-called kernel function. In this paper, a traditional radial basis function (RBF) kernel will be used to train both the SVM and the LS-SVM regressions. The RBF kernel takes the form

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(\frac{-\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2} \right) \tag{9}$$

where σ is the hyperparameter referred to as the width of the RBF.

For the dual-space formulation of (8), the coefficients α and the bias coefficient b can be obtained in a closed form by solving a linear system of equations. On the other hand, the SVM coefficients are computed numerically as the solution of a quadratic convex optimization problem [36]. Importantly, due to the ε-insensitive loss function, this optimization leads to a sparse solution (i.e., most coefficients α_i are equal to zero) while others are restricted by the condition $\alpha_i \in [0, \gamma]$.

III. Proposed Prior Knowledge Based Machine Learning (PKBML) Techniques

The main computational expense in constructing ML metamodels is in the training of these metamodels. This is because in order to achieve good accuracy when working with conventional ML metamodels such as the ones described in the previous section, the required amount of training data extracted from full-wave EM simulations for hybrid copper-graphene interconnects can be very large, thereby leading to massive training time costs. In this paper, this issue is resolved by utilizing prior knowledge-based ML (PKBML) frameworks. In such frameworks, the full-wave EM solvers used, which are highly accurate but computationally very slow, are referred to as the high-fidelity models. In addition to these high-fidelity models, approximate empirical models which are computationally cheap to solve are also developed and referred to as low-fidelity models [34], [22],[25]. Thereafter, intelligently combining the data extracted from the low and high-fidelity models enables a better accuracy versus training time tradeoff than what is possible for conventional ML metamodels trained on data from the high-fidelity model alone. The metamodels constructed from such PKBML frameworks are called PKBML metamodels. Specifically, this work presents two PKBML frameworks for hybrid copper-graphene interconnect modeling – one based on the source difference (SD) technique and another based on the prior knowledge input (PKI) technique.

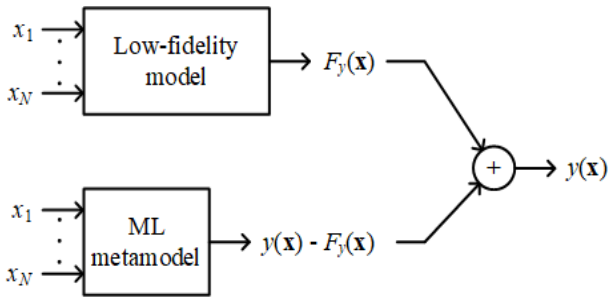


Fig. 2: Block diagram illustrating how a low-fidelity model and a ML metamodel can be utilized in a SD technique.

A. Source Difference (SD) Technique

The SD technique uses the predictions of the fast empirical (low-fidelity) model as prior knowledge to accelerate the training of the ML metamodel. In particular, the ML metamodel is trained to learn the difference between the predictions made by the high-fidelity model $y = M(\mathbf{x})$ and the low-fidelity model, say $F_y(\mathbf{x})$, as shown in Fig. 2. Thus, the training dataset used is $D_{SD} = \{(\mathbf{x}_l, E(\mathbf{x}_l))\}_{l=1}^{L}$ where:

$$E(\mathbf{x}_l) = y_l - F_y(\mathbf{x}_l) \tag{10}$$

where $y_l = M(\mathbf{x}_l)$ is the prediction obtained via the high-fidelity model and $F_y(\mathbf{x}_l)$ is the output calculated via the low-fidelity model for the same input parameters $\mathbf{x}_l$. Now, the variance of the error term of (10) is given as

$$Var(E(\mathbf{x})) = Var(y(\mathbf{x})) + Var(F_y(\mathbf{x})) - 2Cov(y(\mathbf{x}), F_y(\mathbf{x})) \tag{11}$$

It is noted from (11) that as the correlation between the predicted outputs of the low and high-fidelity models increases, the variance of the error quantity of (10) decreases. This, in turn, implies that a smaller number of training samples in the dataset D_{SD} will be sufficient to capture the variability of the error quantity of (10). So, for an appropriate low-fidelity model, the number of training samples required to train a ML metamodel

to emulate the error quantity of (10) will be much lower than required to emulate the true output $y = M(\mathbf{x})$. Once the error quantity of (10) is emulated by any metamodel, the true output can be recovered simply as the sum of the predicted outputs from the metamodel and the low-fidelity model as (see Fig. 2)

$$y(\mathbf{x}) = E(\mathbf{x}) + F_y(\mathbf{x}) \tag{12}$$

Therefore, in summary, when correlation between the high-fidelity model $y = M(\mathbf{x})$ and the low-fidelity model $F_y(\mathbf{x})$ is high, the SD approach enables a dramatic reduction of the number of training samples required during the training phase of ML metamodels [22], [25], [34]. In the work of [34], a SD approach was developed for the fast ANN based modeling of hybrid coper-graphene interconnects.

B. Prior Knowledge Input (PKI) Technique

Based on (12), it is observed that the SD technique assumes a linear correlation between the low and high-fidelity models. However, such a strong assumption might lead to loss of convergence of the SD technique when the correlation between the low and high-fidelity models is actually nonlinear. In such scenarios, the PKI framework as shown in Fig. 3 can be a promising technique to overcome the above limitation [31]. The underlying concept of the PKI technique is to represent the true output as a nonlinear function of the predictions of the low-fidelity model $F_y(\mathbf{x})$ added to the error quantity of (10) as

$$y(\mathbf{x}) = G\left(F_y(\mathbf{x})\right) + E(\mathbf{x}) \tag{13}$$

where $G(.)$ is an appropriate nonlinear function. The expression of (13) can be expressed even more compactly using a new nonlinear function $H(.)$ as

$$y(\mathbf{x}) = H\left(F_y(\mathbf{x}), \mathbf{x}\right) \tag{14}$$

From (14), it is concluded that the output of the high-fidelity model $y = M(\mathbf{x})$ is a nonlinear combination of the inputs $\mathbf{x}$ and the corresponding predicted output of the low-fidelity model $F_y(\mathbf{x})$. Therefore, a ML metamodel can be trained to emulate the nonlinear function of (14) using a new training dataset $D_{PKI} = \left\{\left(\tilde{\mathbf{x}}_l, y_l\right)\right\}_{l=1}^{L}$ where the new input space is augmented as

$$\tilde{\mathbf{x}}_l = \begin{bmatrix} \mathbf{x}_l \\ F_y(\mathbf{x}_l) \end{bmatrix} \tag{15}$$

According to (14), (15) and as shown in Fig. 3, in the PKI formulation it is the output of the low-fidelity model that serves as prior knowledge and becomes additional inputs to the ML metamodel. This prior knowledge guides the ML metamodel to learn the nonlinear function of (14) using much fewer training datapoints than what is conventionally required. Moreover, in the formulation of (14), it is clear that a generic nonlinear correlation exists between the outputs of the low and high-fidelity models. Therefore, the hard assumption of a linear correlation of the SD technique is remedied in the PKI framework.

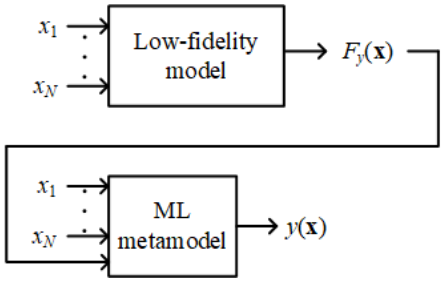


Fig. 3: Block diagram illustrating how a low-fidelity model and a ML metamodel can be utilized in a PKI technique.

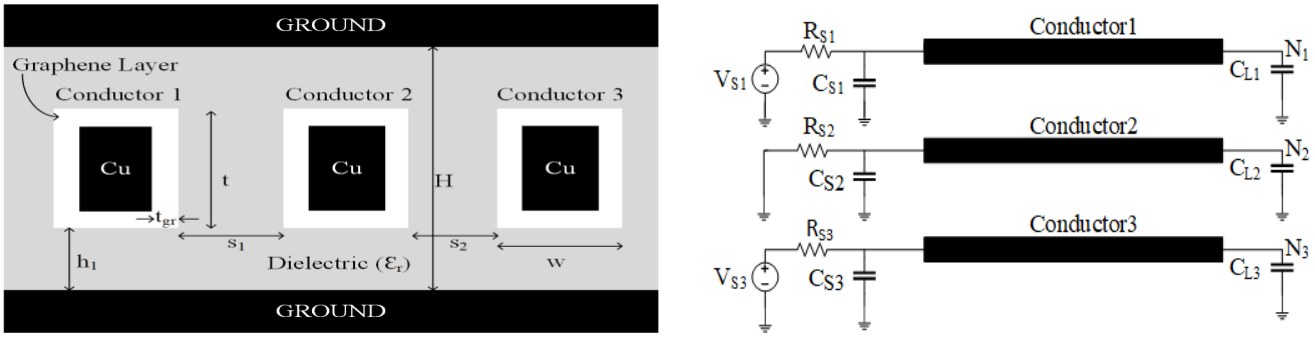
IV. Appropriate Low-Fidelity Model for Copper-Graphene Interconnects

In the work of [11], an empirical model was developed to study the effect of the input parameters **x** on the p. u. l. resistance (**R**), inductance (**L**), and capacitance (**C**) matrices of hybrid copper-graphene interconnects. The hallmark of this empirical model was that it was analytic in nature, and hence, numerically very efficient to solve. However, in the work of [34], it was demonstrated that this empirical model was not very accurate for signal integrity verification. So, in this paper, the empirical model is considered to be a suitable low-fidelity model that can be used in the PKI and SD techniques described in Section III.

In this low-fidelity model, the grain boundary scattering and surface scattering mechanisms are assumed to be the primary contributors of the p. u. l. resistance of copper interconnects. The effects of these scattering mechanisms are modeled by the Mayadas-Shatzkes (M-S) and the Fuchs-Sondheimer (F-S) models respectively. Thus, using these specific models, first the p. u. l. resistance of the copper trace of the hybrid interconnect is calculated as [11]

$$\begin{aligned} R_{Copper} &= \frac{\rho_0\left(F_{FS} + F_{MS}\right)}{wt} \\ F_{MS} &= \left[1 - 1.5\alpha + 3\alpha^2 - 3\alpha^3 \ln\left(1 + \frac{1}{\alpha}\right)\right]^{-1} \\ \alpha &= \frac{\lambda_0 R_f}{D_g\left(1 - R_f\right)}, \quad F_{FS} = 0.45\lambda_{Cu}\left(1 - p_{Cu}\right)\frac{w+t}{wt} \end{aligned} \tag{16}$$

In (16), the term ρ_0 is the bulk resistivity of the copper conductor, D_g is the grain size of copper, R_f is the reflection coefficient at the grain boundary, p_{Cu} is the specularity parameter, λ_0 is the bulk mean free path of copper, and w and t are the interconnect width and thickness respectively. Graphene possesses a very high mean free path because of limited space for the scattering of electrons. Next, for the graphene nanoribbons with specularity parameter (p) in the range of 0 to 1, the p. u. l. resistance of the barrier layers is calculated as [11]



(a) (b)

Fig. 4: Details of the hybrid copper-graphene interconnect network used in Example 1 & 2. (a) Cross-sectional view of the hybrid interconnect structure. (b) Schematic of the hybrid interconnect network showing the driver and load circuits modeled as linear RC circuits.

$$R_{gr} = \frac{h}{2N_l e^2 N_{ch}} \left[\sum_{i=1}^{m} \left(1 + \frac{L}{\lambda_{effi}} \right)^{-1} \right]^{-1} \tag{17}$$

Here, N_l is the number of the graphene layers, N_{ch} is the number of conducting channels, and λ_{effi} is the effective mean free path of the *i*-th sub-band given by Matthiessen's rule. Now, by combining (16) and (17), the effective p. u. l. resistance parameter of the entire copper-graphene conductor in a hybrid interconnect is given as [11]

$$\frac{1}{R_{Cu-gr}} = \frac{1}{R_{Copper}} + \sum_{k=t,b,l,r} \frac{1}{R_{gr}^k} \tag{18}$$

where the quantities $\left\{ R_{gr}^t, R_{gr}^b, R_{gr}^l, R_{gr}^r \right\}$ in (18) refer to the p. u. l equivalent resistance contributions from the top, bottom, left and right side of the graphene barrier layers respectively.

Next, in the low-fidelity model, the effective p. u. l. capacitance and inductance parameters of each conductor of a hybrid copper-graphene interconnect is calculated as [11]

$$\begin{aligned} C_{Cu-gr} &= \left[\sum_{k=t,b,l,r} (C_{rec}^k) + \frac{1}{C_e} \right]^{-1} \\ L_{Cu-gr} &= \left(\sum_{k=t,b,l,r} \frac{1}{L_{rec}^k} \right)^{-1} + L_e \end{aligned} \tag{19}$$

where C_e and L_e refer to the p. u. l. electrostatic capacitance and magnetic inductance components and $L_eC_e = \mu_0\varepsilon_0\varepsilon_r$. The quantities $C_{rec}^t, C_{rec}^b, C_{rec}^r, C_{rec}^l$ and $L_{rec}^t, L_{rec}^b, L_{rec}^r, L_{rec}^l$ in (19) refer to the p. u. l. equivalent capacitance and inductance contributions from the top, bottom, right, and left side of the graphene barrier layers for each conductor, respectively and is calculated using the recursive methodology detailed in [11]. Once the p. u. l. parameters of each conductor are calculated as described above, they can be used to construct the resistance, inductance, and capacitance matrices (i.e., the **RLC** matrices) of hybrid copper-graphene interconnects as outlined in [34].

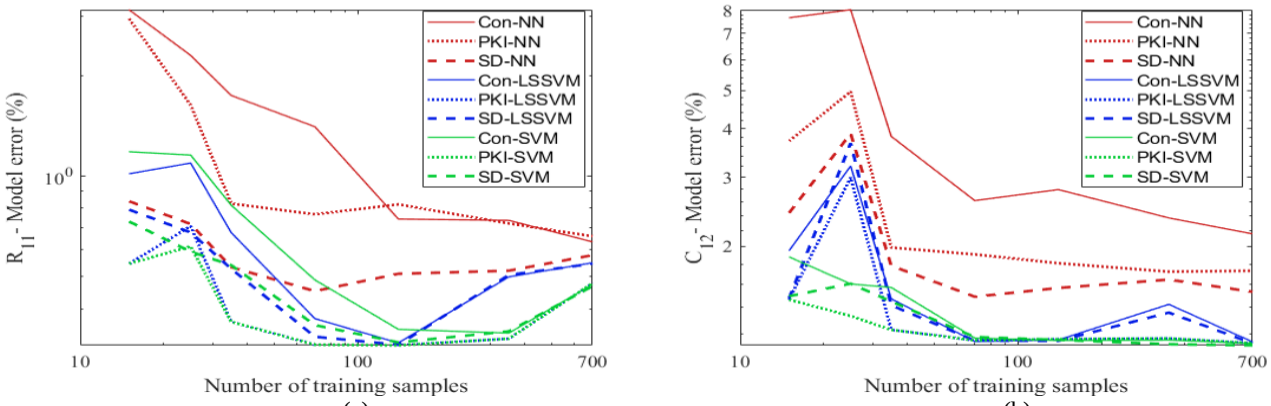
V. NUMERICAL EXAMPLES

In this section, two numerical examples are presented to demonstrate the advantages of the proposed PKBML metamodels of Section III over conventional ML metamodels and full-wave EM simulations for the signal integrity verification of hybrid copper-graphene interconnects. In both these examples, at the first step, the SD and PKI metamodels presented in Section III are applied to predict the p. u. l. parameters of hybrid copper-graphene interconnects. Then, the obtained p. u. l. parameters are embedded in SPICE multi-conductor transmission line (MTL) models to perform transient analysis and signal integrity verification of the interconnects. All the ANN metamodels used in this section have a single hidden layer and are trained using the Levenberg-Marquardt optimizer with back-propagation as available in the ML toolbox in MATLAB. On the other hand, the SVM and the LS-SVM metamodels used in this section are trained in MATLAB via built-in functions and the LS-SVMlab v1.8 toolbox [37]. The SVM metamodels are optimized using a Bayesian optimizer with 5-fold cross-validation (CV) error [18] for a maximum of 30 iterations while the LS-SVM metamodel uses the same Bayesian optimizer with a leave-one-out CV error [36] for a maximum of 25 optimization steps.

TABLE I
DESIGN PARAMETERS FOR THE NETWORK OF EXAMPLE 1

	Design Parameters	Nominal Values	% Variation
1	Width(w)	18nm	+/- 15 %
2	Thickness(t)	37.8nm	+/- 15 %
3	Spacing between line 1 and 2 (s_1)	13nm	+/- 15 %
4	Spacing between line 2 and 3 (s_2)	17nm	+/- 15 %
5	Height of Interconnect from GND layer (h_1)	37nm	+/- 15 %
6	Total height of dielectric (H)	113.8nm	+/- 10 %
7	Barrier layer thickness(t_{gr})	1nm	+/- 15 %
8	Dielectric constant (ϵ_r)	3.9	+/- 15 %

Example 1- Signal Integrity Analysis: In this example, a 3-conductor hybrid copper graphene interconnect structure is considered as shown in Fig. 4(a). The parametric variability for this structure is detailed in Table I. In the first part of this example, attention is laid on extracting the p. u. l. **RLC** parameter matrices of the interconnects of Fig. 4(a). For this purpose, multiple distinct methods are adopted – a commercial



(a) (b)

Fig. 5: Scaling of the testing NRMSE with increasing number of training points for the values of the R_{11} (panel (a)) and C_{12} (panel (b)) p. u. l. parameters using the different ML metamodel and their PKI and SD variants.

TABLE II
ACCURACY AND TIME COST COMPARISON BETWEEN DIFFERENT METAMODELS FOR TRAINING DATASET OF 35, 140, AND 700 SAMPLES

Method		L = 35 (Cost = 105 MINUTES)			L = 140 (Cost = 420 MINUTES)			L = 700 (Cost = 2100 MINUTES)		
		Ave. RMSE (%)	Max RMSE	Optimization time (s)	Ave. RMSE (%)	Max RMSE	Optimization Time (s)	Ave. RMSE (%)	Max. RMSE	Optimization Time (s)
SVM	CON	0.85	1.57	190	0.63	1.43	241	0.63	1.42	1804
	PKI	0.77	1.56	205	0.62	1.42	253	0.62	1.41	1790
	SD	0.82	1.61	220	0.63	1.41	285	0.62	1.40	2645
LS-SVM	CON	0.81	1.5	2.38	0.62	1.41	5.47	0.64	1.46	157.6
	PKI	0.77	1.56	2.38	0.62	1.42	5.96	0.62	1.41	154
	SD	0.81	1.61	2.07	0.62	1.40	5.41	0.64	1.46	145
ANN	CON	2.28	4.50	39	1.38	2.79	41	1.16	2.45	45
	PKI	1.24	2.42	31	1	1.84	32	0.92	1.80	35
	SD	1.17	2.71	24	0.90	1.69	32	0.89	1.68	45

full-wave EM solver which in this case is the ANSYS Q3D Extractor tool [32], the conventional ANN, SVM, and LS-SVM metamodels trained using data obtained from the ANSYS Q3D Extractor tool, and the proposed SD and PKI variants of the ANN, SVM, and LS-SVM metamodels trained using a combination of data obtained from the ANSYS Q3D Extractor tool and the low-fidelity model of Section IV [11]. All the ML metamodels of this example are trained using the same training dataset where the number of points in the dataset is progressively increased as L= {15, 25, 35, 70, 140, 350, 700}. All these training points are selected using a Latin hypercube sampling (LHS) scheme. For testing, all the ML metamodels use a single testing dataset comprising of 500 points uniformly distributed over the entire input parameter space.

Now, for the given training dataset, the decay of the testing errors for the R_{11} and C_{12} p. u. l. parameters achieved by different ML metamodels is shown in Fig. 5. The testing error considered in the above plot is the normalized root mean square (NRMS) error expressed as:

$$\varepsilon_{NRMS} = \frac{1}{N_{test}} \sqrt{\frac{\sum_{l=1}^{N_{test}} \left(y(\mathbf{x}_l) - \hat{y}(\mathbf{x}_l) \right)^2}{\sum_{l=1}^{N_{test}} y^2(\mathbf{x}_l)}} \text{ x } 100 \,, \tag{20}$$

where $y(\mathbf{x}_l)$ is the true result obtained from the full-wave EM solver and $\hat{y}(\mathbf{x}_l)$ is the result predicted by the ML metamodel. For this example, a NRMS testing error of less than 2% for all p. u. l. parameter implies that the metamodel is sufficiently accurate, due to the fact that for such low errors in the p. u. l. parameter values, there is virtually no errors in the SPICE simulation results [34]. From Fig. 5, the following three observations are made.

(i) Irrespective of the ML metamodel used, their SD and PKI formulations converge much faster than their conventional formulations. This is to be expected given that the SD and PKI approaches exploit the correlation between the results of the low-fidelity model and the full-wave EM simulations.

(ii) For most ML metamodels, the PKI approach shows a faster convergence than the SD approach (except for ANNs). This too is expected given that the PKI approach assumes a general nonlinear correlation between the results of the low-fidelity model and the full-wave EM simulations as opposed to the

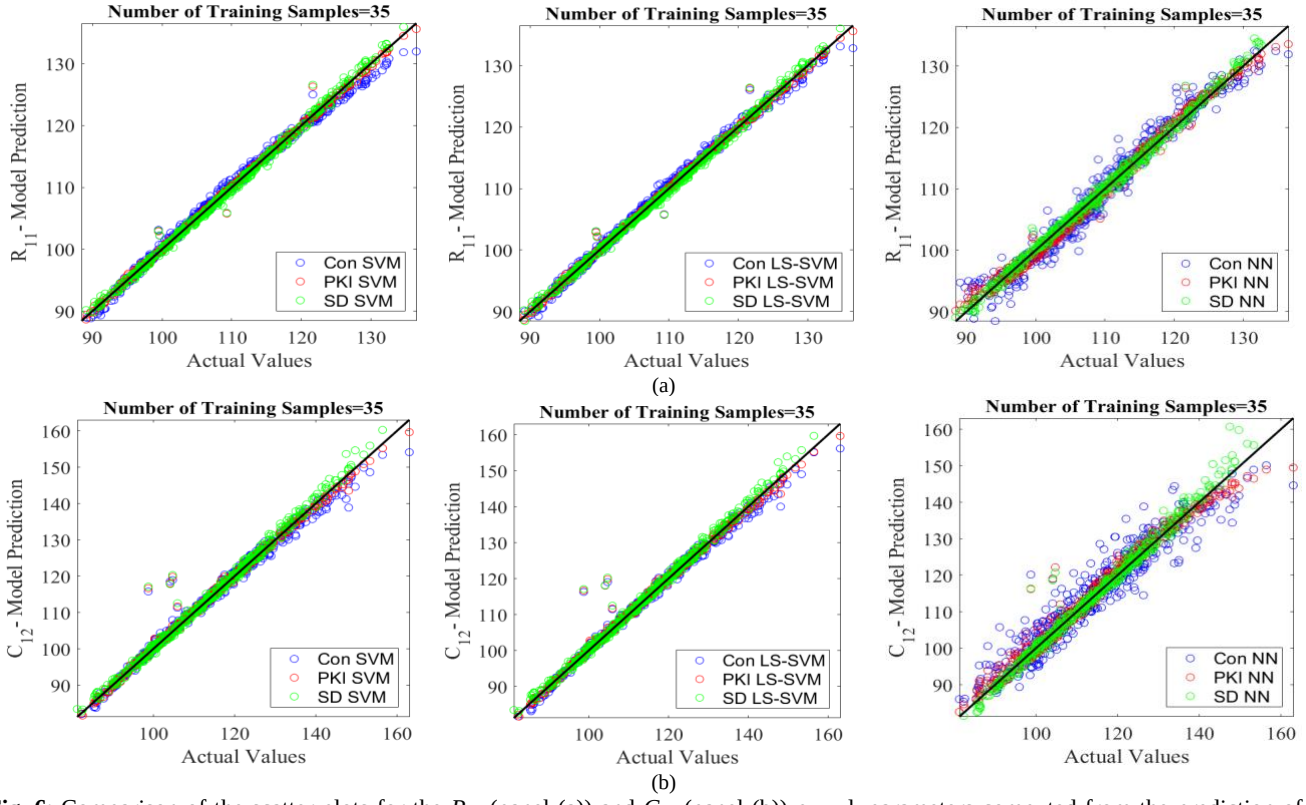


Fig. 6: Comparison of the scatter plots for the R_{11} (panel (a)) and C_{12} (panel (b)) p. u. l. parameters computed from the prediction of the proposed metamodels by using as reference the corresponding values computed via a full-wave EM simulation in ANSYS Q3D Extractor.

restrictive linear correlation assumed by the SD approach. This is a key conclusion that was not explored in [34].

(iii) The SVM and LS-SVM metamodels, whether of the conventional or the SD and PKI variety, are much more accurate than their ANN counterparts. This is another important conclusion that was not explored in the earlier work of [34].

As further proof of the above observations, Table II provides an exhaustive comparison among the considered metamodels in terms of their average and maximum NRMSE computed over all the entries of the p. u. l. matrices and their training time costs for 35, 140, and 700 samples. The training time costs are broken into two parts – the time taken to generate the training dataset (given at the top of the table) and the time taken to calculate the optimal weights and bias values that will minimize the loss function (referred to as the 'optimization time'). The results of Table II clearly show that for the same number of training samples, the proposed PKI and SD techniques provide a lower testing error compared to the conventional metamodels whether of the ANN, SVM, or LS-SVM type. This clearly underlines the universal improved convergence of the proposed PKI and SD techniques. In fact, the PKI and SD variants of ANN metamodels are sufficiently well trained using only 35 training samples as opposed to the conventional ANN metamodel that requires roughly 140 (i.e., 4 times larger) number of training samples.

Regarding the computational time costs, Table II shows that the largest time cost is incurred during the generation of the

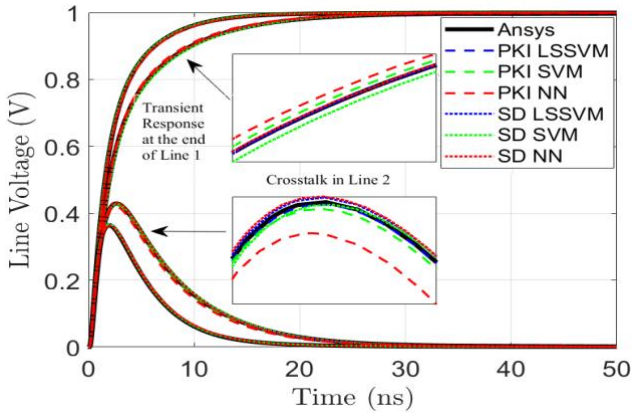


Fig. 7: Transient responses at the end of line 1 and crosstalk in the at the end of line 2 simulated in SPICE and based on the p. u. l. parameters predicted by the different metamodels for the two different design points.

training dataset than during the optimization. This is to be expected because it is only during the generation of the training dataset that the repeated expensive full-wave EM simulations have to be performed. For example, the computational time cost required by the ANSYS Q3D Extractor tool for generating 35 training samples is around 6300 seconds, which is more than 31 times the time required to optimize the slowest metamodel. On the other hand, the remarkable fact of all these techniques is that the time taken by the metamodels to predict their outputs is extremely small – indeed, once trained, all of the metamodels take less than one second for predicting the p. u. l. parameter at

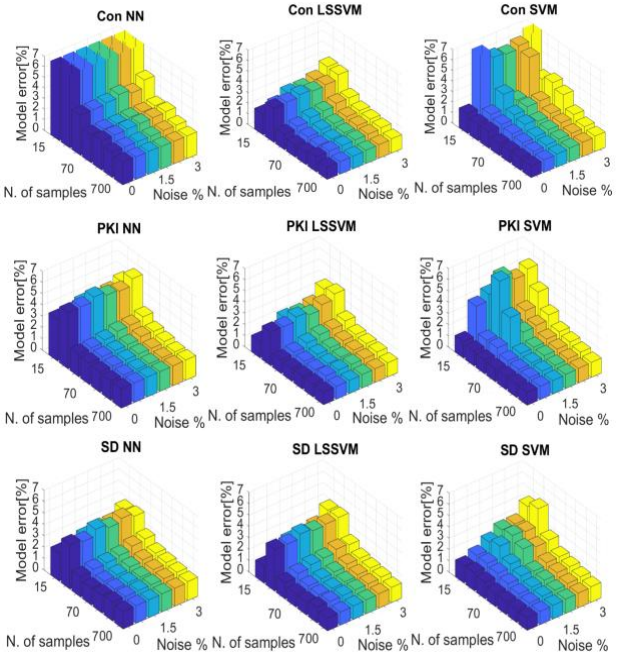


Fig. 8: Comparison among the NRMS testing errors for C_{12} provided by the different metamodels for different number of training samples (i.e., L) and noise level σ_n.

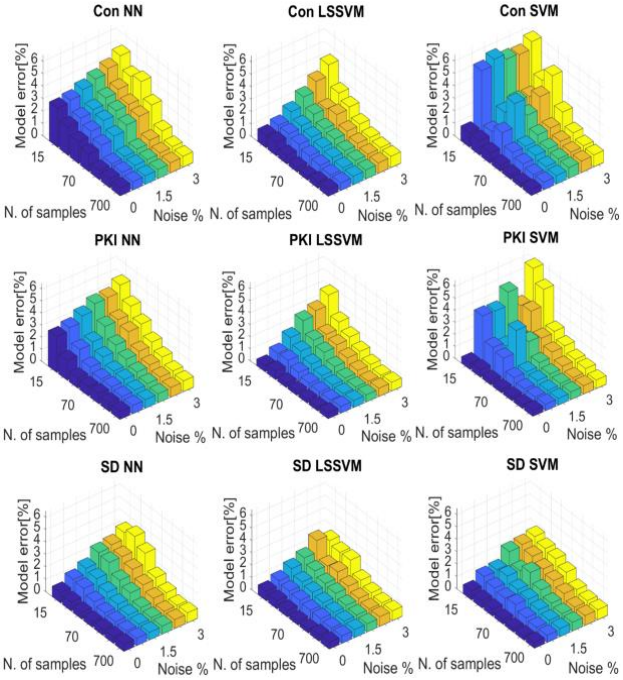


Fig. 9: Comparison among the NRMS testing errors for R_{11} provided by the different metamodels for different number of training samples (i.e., L) and noise level σ_n.

any design point. This makes these metamodels far better suited than full-wave EM simulators for tasks that require several evaluations, e.g., uncertainty quantification, design space exploration, sensitivity analysis, etc. The results of Table II are validated using the scatter plots of Fig. 6.

In the final part of this example, the values of the p. u. l. parameters of the interconnects of Fig. 4(a) predicted by the different ML metamodels when trained to below a 2% NRMS testing error threshold for ANN metamodels and 1.5% testing error threshold for SVM and LS-SVM metamodels are used for a SPICE transient analysis. For this purpose, lines 1 and 3 of the interconnects are excited by voltage sources with saturated ramp waveforms of rise time T_r = 0.1 ps and amplitude 1 V. Line 2 is the victim line. The outputs of interest for this example are the transient responses at the far-end of line 1 and 2. In Fig. 7, the transient responses of the interconnects obtained using the p. u. l. parameters predicted by the ML metamodels is illustrated for two arbitrary points in the normalized input parameter space [-0.9860 0.9750 -0.1731 -0.5867 -0.2937 -0.7857 0.3705 0.7532] and [0.9665 -0.9654 0.7426 0.1931 0.6529 0.8909 0.8300 -0.0865]. In addition, Fig. 7 also includes the transient responses obtained using the p. u. l. parameters extracted from full-wave EM simulations for the same two points. It is noted from Fig. 7 that the results obtained from ML metamodels exhibit very good agreement with the results obtained from the full-wave EM simulations. This demonstrates that the accelerated training of the ML metamodels provided by the proposed SD and PKI approaches does not lead to any loss

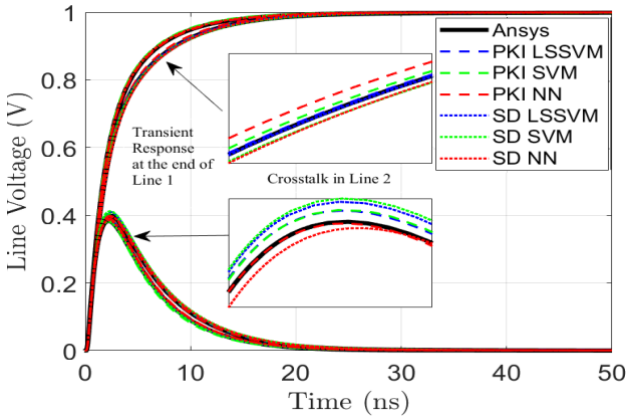


Fig. 10: Transient responses at the end of line 1 and crosstalk in the at the end of line 2 simulated in SPICE and based on the p. u. l. parameters predicted by the different metamodels for the two different design points for noise level σ_n = 3%

in model accuracy either when predicting the p. u. l. parameters of the interconnects or when performing the SPICE simulations.

Example 2 - Robustness Analysis: In this example, the robustness of the different PKBML metamodels of Section III to simulation noise is examined. For this purpose, the pristine training data of the previous example is now corrupted by adding random Gaussian noise as

$$y_{noisy}(\mathbf{x}_l) = y(\mathbf{x}_l)(1+\xi) \tag{21}$$

Where $\xi \sim N\left(0, \sigma_n^2\right)$ is a Gaussian random variable whose standard deviation lies in the range 1% - 3%, such that in the worst-case 99.7% of the noisy realizations falls within $\pm 9\%$

variation around the noiseless case. Next, the p. u. l. parameters of the interconnects of Fig. 4(a) are once again predicted using the different methods of Example 1. In Figures 8 and 9, the scaling of the NRMS testing error of (20) achieved by different ML metamodels with respect to the noise level (i.e., the value of σ) and the number of training points when predicating the C_{12} and R_{11} p. u. l. parameters is shown. From Figs. 8 and 9, the following three conclusions are made.

(i) As the noise level increases, for the same number of training samples the testing error of each ML metamodel increases. This increase in the testing error is noticed for both conventional and PKBML formulations.

(ii) Even in the presence of noise, the SD and PKI formulations are still able to provide a lower testing error relative to the conventional ML metamodels for the same number of training points. Thus, while the effectiveness of the SD and PKI formulations may be diminished by the presence of noise, their ability to accelerate the training of ML metamodels remains intact.

(iii) The SD formulation shows significantly lower error than even the PKI formulation, and hence, is the most robust among the proposed PKBML formulations to simulation noise.

In the final part of this example, the values of the p. u. l. parameters of the interconnects of Fig. 4(a) predicted by the different ML metamodels when trained to below a 2% NRMS testing error threshold in the presence of maximum noise (i.e., σ_n = 3%) are used for SPICE transient analysis of the hybrid interconnect network depicted in Fig. 4(b). The active and victim lines of the interconnects are exactly as used in Example 1 and excited by the same sources. In Fig. 10, the transient responses of the interconnects obtained using the p. u. l. parameters predicted by the ML metamodels is illustrated for two arbitrary points in the input normalized parameter space (different points than those used in Example 1) [0.9857 -0.3199 0.3784 0.7145 -0.8059 -0.9908 0.5916 0.7275] and [0.1431 0.9127 0.7862 -0.9728 0.8036 0.9752 0.3177 0.8134]. In addition, Fig. 10 also includes the transient responses obtained using the p. u. l. parameters extracted from full-wave EM simulations for the same two points. It is noted from Fig. 10 that the results obtained from ML metamodels exhibit very good agreement with the results obtained from the full-wave EM simulations even despite the fact that the training data is no longer pristine and is corrupted by noise.

VI. Conclusion

In this paper, two different prior knowledge-based machine learning (PKBML) approaches are developed and their performance relative to conventional ML approaches and full-wave EM simulations are explored for predicting the p. u. l. parameters of hybrid copper-graphene on-chip interconnects. In particular, the PKBML approaches is tested for ANN, SVM, and LS-SVM metamodels for thoroughness of analysis. From the detailed explorations performed, it is concluded that the PKI approach best accelerates the training of all ML metamodels provided no noise is in present in the data. However, in the presence of noise, the SD approach ends up providing better acceleration. Finally, the SVM and LS-SVM metamodels generally outperform the ANN metamodels both for the conventional and PKBML approaches.

References

1. J. A. Davis *et al.*, "Interconnect limits on gigascale integration," *Proc. IEEE*, vol. 89, no. 3, pp. 305-324, March 2001
2. W. Steinhögl, G. Schindler, G. Steinlesberger, M. Traving, and M. Engelhardt, ''Comprehensive study of the resistivity of copper wires with lateral dimensions of 100 nm and smaller,'' *J. Appl. Phys.*, vol. 97, no. 2, pp. 023706-1–023706-7, 2005
3. International Technology Roadmap for Semiconductors, 2015. [Online]. Available: http://www.itrs2.net/itrs-reports.html
4. A. Todri-Sanial, J. Dijon, and A. Maffucci (editors), *Carbon Nanotubes for Interconnects*. Switzerland: Springer International, 2017
5. P. S. Peercy, "The drive to miniaturization," *Nature*, vol. 406, pp. 1023-1026, August 2000.
6. W. Steinhögl, G. Schindler, G. Steinlesberger, M. Traving, and M. Engelhardt, "Size-dependent resistivity of metallic wires in the mesoscopic range" *Physical Review B*, vol. 66, no. 7, pp. 075414-1–075414-4, Aug. 2002
7. L. Y. Wang, D. H. Zhang, C. Y. Li, and P. D. Foo, "Comparative study of Ta, TaN, and Ta/TaN bi-layer barriers for Cu-ultra low-*k* porous polymer integration." *Thin Solid Films*, vol. 462-463, pp. 176-181, 2004
8. R. Mehta, S. Chugh, and Z. Chen, "Enhanced electrical and thermal conduction in graphene-encapsulated copper nanowires," *Nano Letters*., vol. 15, no. 3, pp. 2024–2030, Feb. 2015
9. P. Goli, H. Ning, X. Li, C. Y. Liu, K. S. Novoselov, and A. A. Balandin, "Thermal properties of graphene–copper–graphene heterogeneous films," *Nano Letters*, vol. 14, no. 3, pp. 1497–1503, 2014
10. C.-H. Yeh et al., "Scalable graphite/copper bishell composite for high performance interconnects," *ACS Nano*, vol. 8, no. 1, pp. 275–282, 2014.
11. Z. Cheng et al., "Analysis of Cu-Graphene Interconnects," *IEEE Access*, vol. 6, pp. 53499-53508, 2018
12. W. S. Zhao, D. W. Wang, G. Wang, and W. Y. Yin, "Electrical modeling of on-chip cu-graphene heterogeneous interconnects," *IEEE Electron Device Lett.*, vol. 36, no. 1, pp. 74–76, Jan. 2015
13. R. Achar and M. Nakhla, "Simulation of high-speed interconnects," *Proc. IEEE*, vol. 89, no. 5, pp. 693–728, May 2001
14. C. R. Paul, *Analysis of Multiconductor Transmission Line*. New York: Wiley-Interscience, 2008
15. M. Swaminathan, H. K. Torun, H. Yu, J. A. Hejase, and W. D. Becker, "Demystifying machine learning for signal and power integrity problems in packaging," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 10, no. 8, pp. 1276-1295, Aug. 2020
16. H. M. Torun and M. Swaminathan, "High-dimensional global optimization method for high-frequency electronic design," *IEEE Trans. Microwave Theory Tech.*, vol. 67, no. 6, pp. 2128–2142, Jun. 2019
17. R. Medico, D. Spina, D. Vande Ginste, D. Deschrijver, and T. Dhaene, "Machine-learning-based error detection and design optimization in signal integrity applications," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 9, no. 9, pp. 1712–1720, Sep. 2019
18. 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," *IEEE Access*, vol. 7, pp. 4056-4066, 2018
19. R. Trinchero, P. Manfredi, I. S. Stievano, and F. G. Canavero, "Machine learning for the performance assessment of high-speed links," *IEEE Trans. Electromagn. Compatibility*, vol. 60, no. 6, pp. 1627–1634, Dec. 2018.
20. H. M. Torun, A. C. Durgun, K. Aygun, and M. Swaminathan, "Causal and passive parameterization of s-parameters using neural networks," *IEEE Trans. Microwave Theory Tech.*, vol. 68, no. 10, pp. 4290–4304, Oct. 2020
21. H. Park *et al.*, "Deep reinforcement learning-based optimal decoupling capacitor design method for silicon interposer-based 2.5-D/3-D ICs," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 10, no. 3, pp. 467–478, Mar. 2020

22. Q.-J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*, Norwood, Massachusetts: Artech House 2000
23. A. Veluswami, M. S. Nakhla, and Q.-J. Zhang, "The application of neural networks to EM based simulation and optimization of interconnects in high-speed VLSI circuits," *IEEE Trans. Microwave Theory and Tech.*, vol. 45, no. 5, pp. 712-723, May 1997
24. L. Xia, J. Meng, R. Xu, B. Yan and Y. Guo, "Modeling of 3-D Vertical Interconnect Using Support Vector Machine Regression," in IEEE Microwave and Wireless Components Letters, vol. 16, no. 12, pp. 639-641, Dec. 2006, doi: 10.1109/LMWC.2006.885585.
25. P. M. Watson, K. C. Gupta, and R. L. Mahajan, "Development of knowledge based artificial neural network models for microwave components," in *Proc. IEEE Int. Microwave Symp.*, Baltimore, MD, June 1998, pp. 9-12.
26. P. M. Watson, K. C. Gupta, and R. L. Mahajan, "Applications of knowledge-based artificial neural network modeling to microwave components," *Int. J. RF Microwave Computer-Aided Eng.*, vol. 9, pp. 254–260, May 1999
27. J. Bandler *et. al.*, "Neuromodeling of microwave circuits exploring space mapping technology," in *Proc. IEEE Int. Microwave Symp.*, Anaheim, CA, 1999, pp. 149-152
28. J. W. Bandler, R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 42, no. 12, 1994, pp. 2536-2544.
29. M. H. Bakr, J. W. Bandler, M. A. Ismail, J. E. Rayas-Sánchez, and Q. J. Zhang, "Neral space mapping optimization for EM-based design," *IEEE Trans. Micrwave Theory Tech.*, vol. 48, pp. 2307–2315, Dec. 2000
30. R. Kumar *et al*., "Estimating per-unit-length resistance parameter in emerging opper-graphene hybrid interconnects via prior knowledge based accelerated neural networks," in *Proc. IEEE 29th Conf. Electrical Perf. Electronic Packag. and Systs.*, San Jose, CA, USA, 2020, pp. 1-3.
31. S. Kushwaha, A. Attar, R. Trinchero, F. Canavero, R. Sharma and S. Roy, "Fast Extraction of Per-Unit-Length Parameters of Hybrid Copper-Graphene Interconnects via Generalized Knowledge Based Machine Learning," 2021 IEEE 30th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS), 2021, pp. 1-3, doi: 10.1109/EPEPS51341.2021.9609233.
32. *ANSYS Q3D Extractor*, 2020, [online] Available: http://www.ansys.com/products/electronics/ansys-q3d-extractor.
33. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," in Parallel Distributed Processing vol. 1, D. E. Rumelhart and J. L. McClelland, Editors, Cambridge, MA: MIT Press 1986, pp. 3l8-362.
34. R. Kumar et al., "Knowledge-Based Neural Networks for Fast Design Space Exploration of Hybrid Copper-Graphene On-Chip Interconnect Networks," in IEEE Transactions on Electromagnetic Compatibility, doi: 10.1109/TEMC.2021.3091714.
35. C. R. Paul, Analysis of Multiconductor Transmission line, New York: Wiley,1994.
36. Johan A. K. Suykens et al. "Least Squares Support Vector Machines," World Scientific, 2002. ISBN 9812381511.
37. De Brabanter K. et al., "LS-SVMlab Toolbox User's Guide version 1.8", Internal Report 10-146, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 2010.

Dear Editor,

On behalf of the authors, I take this opportunity to submit our paper titled: ***Comparative Analysis of Prior Knowledge Based Machine Learning Metamodels for Modeling Hybrid Copper-Graphene On-Chip Interconnects*** in this esteemed journal.

In this paper, three key contributions to the state-of-the-art in ML based modeling of the p. u. l. parameters of hybrid copper-graphene interconnects are presented.

(i) The above accuracy issue of the SD method is addressed by using a more general prior knowledge input (PKI) method for predicating the p. u. l. parameters. The PKI method has the capacity to include nonlinear correlation between the results of the low-fidelity model and those of the full-wave EM simulations. As a result, the PKI method can provide smaller modeling errors compared to the SD method for the same number of training points. The improved performance of the PKI method compared to the SD method is demonstrated in this paper using various numerical examples.

(ii) In this paper, the comparative analysis between the SD and PKI methods is performed across different ML metamodels such as the ANN, SVM, and least square SVM (LS-SVM) metamodels. Such a detailed comparative analysis was not included in the earlier works. In fact, such a thorough comparative analysis of different PKBML metamodels for interconnects in general and hybrid copper-graphene interconnects in particular have not been reported before.

(iii) Finally, in this paper, the comparative analysis between the SD and PKI methods across different ML metamodels is conducted both in the presence and absence of simulation noise. Such an analysis reveals critical insights as to which combination of PKBML method and ML metamodeling technique is most suitable depending on the level of noise in the training data.

In a nutshell, the topics covered in this article are quite comprehensive and align with the wide scope of this journal.

Should you have any questions, please feel free to write to me.

Sincerely,

Rohit Sharma

(On behalf of the authors)