Modeling Electrically Long Interconnects Using Physics-Informed Delayed Gaussian Processes

Federico Garbuglia ⓘ, Torsten Reuschel ⓘ, *Member, IEEE*, Christian Schuster ⓘ, *Senior Member, IEEE*, Dirk Deschrijver ⓘ, *Senior Member, IEEE*, Tom Dhaene ⓘ, *Senior Member, IEEE*, and Domenico Spina ⓘ, *Senior Member, IEEE*

Abstract—**This work presents a machine learning technique to model wide-band scattering parameters (S-parameters) of interconnects in the frequency domain using a new Gaussian processes (GP) model. Standard GPs with a general-purpose kernel typically assume high smoothness and therefore are not suitable to model S-parameters that are highly dynamic and oscillating due to propagation delays. The new delayed Gaussian process (τGP) model employs a physics-informed kernel consisting of periodic components, whose fundamental frequencies are interpreted as tunable propagation delays. Then, the model hyperparameters are tuned using a combination of maximum marginal likelihood estimation (MMLE) and delay estimation using Gabor transform. The delay estimation allows one to automatically identify the optimal fundamental frequencies for the kernel, thus increasing the numerical stability of the hyperparameters tuning process. The resulting delayed Gaussian process model accurately predicts the S-parameter values at desired frequency points in the training interval. Two application examples demonstrate the increased accuracy of the new technique, compared to standard Gaussian processes, vector fitting (VF), and delayed vector fitting (DVF) rational models.**

Index Terms—**Delay estimation, Gabor transform, Gaussian processes (GP), interconnects, kernels, machine learning (ML), S-parameters.**

I. INTRODUCTION

MACROMODELING plays a crucial role in the characterization of high-frequency interconnects and finds application in various domains such as the optimization of

Federico Garbuglia, Dirk Deschrijver, and Tom Dhaene are with the Department of Information Technology, Ghent University-imec, 9052 Ghent, Belgium (e-mail: federico.garbuglia@ugent.be; dirk.deschrijver@ugent.be; tom.dhaene@ugent.be).

Torsten Reuschel is with the Department of Physics, University of New Brunswick (UNB), Fredericton, NB E3B 5A3, Canada (e-mail: torsten.reuschel@unb.ca).

Christian Schuster is with the Institut für Theoretische Elekrotechnik, Hamburg University of Technology (TUHH), 21079 Hamburg, Germany (e-mail: schuster@tuhh.de).

Domenico Spina is with the Department of Information Technology, Ghent University-imec, 9052 Ghent, Belgium, and also with Department ELEC, Vrije Universiteit Brussel (VUB), Pleinlaan 2 1050, Brussel (e-mail: domenico.spina@vub.be).

design, modeling of lossy transmission lines, and efficiency enhancement of general-purpose electromagnetic full-wave simulations [1]. Vector fitting (VF) is a popular macromodeling technique that is frequently utilized to construct a rational function approximation of interconnect transfer functions over the frequency variable [1], [2], [3]. One of the major advantages of VF is that it can model transfer functions with a relatively small number of rational coefficients (poles and zeros). In addition, the VF rational model can be converted into a state-space representation and integrated into SPICE-like solvers for simulation. Furthermore, compliance with respect to physical properties such as causality and passivity can be enforced on the VF rational model. A delayed vector fitting (DVF) version is also available for devices that present significant propagation delays or crosstalk, such as electrically long interconnects [4], [5]. However, when the behavior of the device under test deviates from the rational approximation, techniques based on VF require a high number of rational coefficients, leading to possible numerical instabilities. Moreover, extending VF to parameterized models over several design parameters is challenging. In fact, the rational model coefficients are typically nonsmooth and discontinuous, since they are extracted by VF individually for each instance of design parameters [6].

In recent years, several machine learning techniques, such as artificial neural networks (ANN) [7], [8], [9] or support vector machines (SVM) [10] have shown improved accuracy in interconnects macromodeling, especially for a large number of ports and design parameters. However, while ANNs are able to model high-dimensional and nonlinear functions, they require a high amount of training data and they tend to overfit. On the other hand, the excellent regularization capability of SVMs reduces the overfitting risk and improves data efficiency. As an alternative, Bayesian machine learning models such as Gaussian processes (GP) [11] have been explored. While GPs limit the overfitting risk and guarantee data efficiency similar to SVMs, their probabilistic modeling enables a more efficient hyperparameters optimization and physically informed interpretability thanks to the posterior variance estimation [16].

GPs have already demonstrated their efficacy in optimizing performance metrics of microwave devices using Bayesian learning [12], [14]. Thus, GPs are a promising technique for modeling transfer functions. However, modeling interconnects S-parameters, over wide frequency bands using GPs is problematic. While the data samples may exhibit highly nonsmooth and

TABLE I
PRIMARY PROPERTIES AND CAPABILITIES OF S-PARAMETER MODELING TECHNIQUES

Base model	VF	DVF	SVM	NN	VF + NN	GP	τ**GP**
Reference	[2]	[4]	[10]	[8], [9]	[7]	[12], [19]	**[this work]**
Stochastic modeling	no	no	no	no	no	yes	yes
Trainable model parameters	yes	yes	yes	yes	yes	no	no
Frequency modeling	yes	yes	no	yes	yes	**no**	**yes**
Causality/Passivity enforcement	yes	yes	no	yes	yes	**no**	no
Delayed structures	no	yes	no	no	no	**no**	**yes**
Overfitting risk	medium	medium	low	high	high	**low**	**low**

Note that only a summary of standard base models is reported here; the models' capabilities may be extended and improved in specific applications.

The bold values highlight the improvements of the presented technique over the standard Gaussian process modeling.

oscillating behavior, the GPs typically enforce a smooth mean and covariance, also known as *kernel* function [11].

The purpose of this article is to introduce a novel approach, named the delayed Gaussian process (τGP), to overcome the challenges of GP modeling. First, a physics-informed kernel function with periodic components is designed. Periodic kernels have been already used in various research areas, such as time-series [13] analysis and signal processing, to model functions with visible oscillatory behaviors or periodic domains [14]. Here, periodic kernels are applied to electrically long interconnects. Second, inspired by delay-based rational models [4], the fundamental frequency of each periodic kernel component is associated with a propagation delay in the transfer function of the device under test (DUT). The propagation delays for the kernel are automatically estimated from the available frequency data samples using the Gabor transform [4], [15]. The delay estimation is combined with the maximum marginal likelihood estimation (MMLE) [16] into the scheme in Fig. 1 in order to obtain the complete set of optimal hyperparameters for the τGP model.

As a result, the new τGP allows one to model frequency responses that are highly affected by propagation delays and crosstalk. To better recognize the utility of the aforementioned modeling techniques, their primary properties and capabilities are summarized in Table I. In particular, to the best of our knowledge, the τGP is the first technique that enables the representation of delayed structures over frequency using GPs. The rest of the article is organized as follows. Section II provides a full description of the proposed modeling technique, focusing on the new model kernel and hyperparameter tuning process. Next, in Section III, the new τGP is applied to two interconnects and compared to standard GPs, VF, and DVF rational models. Finally, Section IV concludes the article.

II. METHODOLOGY

A. GP Modeling: State of the Art

This article presents an application of GP for modeling interconnect S-parameter data across the frequency variable. GP models do not rely on a fixed number of trainable model parameters, therefore they are nonparametric models and their computational complexity grows polynomially with the amount of training data samples. Therefore, GPs can be trained on a relatively low amount of data points and are particularly suitable for representing low-dimensional functions or stochastic processes. The GP is a probabilistic, real-valued model, in which each data point is modeled as a realization of a Gaussian probability density with mean and variance, while the correlation between pairs of points is determined by the kernel function. Notably, the kernel can include prior knowledge about the function or stochastic process being modeled, such as smoothness, discontinuities, stationarity, or noise distribution in the data. A properly designed kernel allows the GP to achieve high accuracy when modeling dynamic functions. Once the kernel is defined, Gaussian process regression (GPR), also known as Kriging [11], can be executed to predict the value of new data samples. For the interested reader, a detailed description of the properties of GPR can be found in [11].

One main limitation of standard GPs is that they assume a relatively high smoothness of the quantity to be modeled with respect to the input domain variables. However, this is often not the case when modeling high-speed interconnects, whose frequency response may be dynamic and oscillating. As an alternative, the use of deep Gaussian process (DGP) has been shown to be beneficial for the optimization of frequency responses [17], [18]. Although, DGPs come at the cost of high computational complexity and instability when dynamic functions are modeled. Instead, this work proposes a new delayed Gaussian process (τGP) for interconnect modeling.

B. Delayed Gaussian Process Kernel

The novel τGP is built on a physics-informed kernel, denoted as k_{del}, which guarantees increased representation power without sacrificing data efficiency. Being real-valued, it is necessary to apply the τGP separately to the real and imaginary part of S-parameter $S_{a,b}(f)$, defined as a function of the frequency f between any pair of the electrical ports a, b. Note that, without loss of generality, the proposed τGP can be applied to other frequency-dependent figures of merit, including magnitude and phase of $S(f)$. The new kernel function obeys the following analytical expression:

$$k_{\text{del}} = k_{\text{mean}} + k_{\text{envelope}} \cdot \prod_{m=1}^{M} k_{\tau,m} \qquad (1)$$

where M is the number of delays. The three terms of the new kernel are defined as follows:

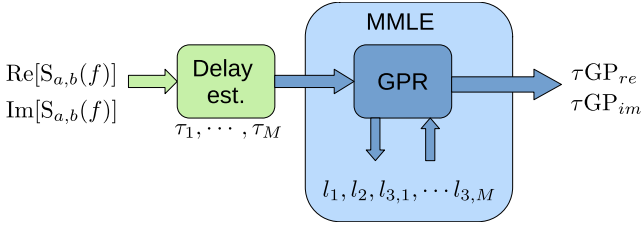$$k_{\text{mean}}(f, f') = k_{\text{Matern}}(f, f', l = l_1); \qquad (2)$$

Fig. 1. Training procedure for the proposed τGPs. First, a delay estimation algorithm (Section II-D) is applied to the real and imaginary parts of each S-parameter to produce the delays hyperparameters τ_m. Next, the lengthscale hyperparameters of the models are optimized using maximum marginal likelihood estimation (Section II-C).

$$k_{\text{envelope}}(f, f') = k_{\text{Matern}}(f, f', l = l_2); \quad (3)$$

$$k_{\tau,m}(f, f') = \exp\left(-2\frac{\sin^2(\pi\tau_m(f - f'))}{l_{3,m}^2}\right) \quad (4)$$

where the terms k_{mean} and k_{envelope} are standard Matern 5/2 kernel functions with lengthscale l, while $k_{\tau,m}$ is a periodic kernel with fundamental frequency τ_m [13]. The analytical expression for the Matern 5/2 kernel is

$$k(f, f') = \frac{2^{1-5/2}}{\Gamma(5/2)}\left(\sqrt{5}d\right)^{5/2} K_{\nu=5/2}\left(\sqrt{5}d\right) \quad (5)$$

where $d = |f - f'|^2/l^2$ and $K_{\nu=5/2}$ is the modified Bessel function of the second kind. The quantities l_1, l_2, and $l_{3,m}$ are trainable lengthscale hyperparameters: Intuitively, they represent the minimum distance over which the correlation between two frequency samples is considered negligible. It is worth noting that the kernel terms are stationary: the correlation between two frequency samples f and f' depends solely on their Euclidean distance $|f - f'|$. Additionally, k_{del} is stationary as it is a combination of stationary functions.

Each term of the new kernel encodes prior knowledge of the function to be modeled. In particular, for the S-parameter curves of an interconnect, it is reasonable to make the following assumptions:

Assumption 1: The propagation delays and crosstalk manifest as oscillating S-parameter values across frequency. This assumption [1] has been verified for devices constituted by a cascade of lumped elements or transmission line segments [4], [20].

Assumption 2: The mean values of the oscillations are continuous and highly correlated between adjacent frequency samples, and weakly correlated otherwise. This correlation is modeled by the decaying exponential in k_{mean}.

Assumption 3: The frequency samples separated by integer multiples of the oscillations' semiperiod are highly correlated. This dependency is expressed by the $\sin^2(\pi\tau_m(\cdot))$ operator in (1). This corresponds to the interpretation of τ_m as the main *propagation delays* in the examined device.

Assumption 4: The oscillations may have a nonconstant amplitude. Thus, the term k_{envelope} models the envelope of the oscillations. The amplitude values are highly correlated between adjacent frequency samples.

Moreover, it is possible to assume oscillations of several fundamental frequencies by combining multiple periodic terms $k_{\tau,m}$ into a single product, as indicated by (1). Since the previous considerations are statistical, the resulting k_{del} guarantees sufficient flexibility to model highly diverse S-parameter functions over frequency. However, the new kernel needs to be adjusted to the available data samples by tuning the vector of hyperparameters $\boldsymbol{\theta} = [l_1, l_2, l_{3,1}, \ldots, l_{3,M}, \tau_1 \ldots, \tau_M]$.

C. Hyperparameter Tuning Using MMLE

A common approach for tuning kernel hyperparameters is via the MMLE [16]. The marginal likelihood represents the probability of observing the available data samples when a particular prior distribution is assumed. In the S-parameter use case, the marginal likelihood can be expressed as $p(\boldsymbol{s}|F, \boldsymbol{\theta})$, where F is the set of observed frequencies and $\boldsymbol{s}$ is the vector of real or imaginary parts of the recorded S-parameter values. Conveniently, the GPR provides an analytical expression for the marginal likelihood, as a function of the chosen kernel

$$\log p(\boldsymbol{s}|F, \boldsymbol{\theta}) = -\frac{1}{2}\boldsymbol{s}^T K_s^{-1}\boldsymbol{s} - \frac{1}{2}\log|K_s| - \frac{n_f}{2}\log 2\pi \quad (6)$$

where K_y is the matrix of kernel function evaluations for each pair of training samples and n_f is the number of training samples. Then, the optimal hyperparameter values can be searched by maximizing the marginal likelihood

$$\boldsymbol{\theta}_{opt} \sim \arg\max_{\theta} p(\boldsymbol{s}|F, \boldsymbol{\theta}). \quad (7)$$

Since the $p(\boldsymbol{s}|F, \boldsymbol{\theta})$ is analytical, the maximization can be rapidly executed via gradient ascent. Furthermore, it is possible to integrate additional assumptions in the maximization problem in order to improve the regularization of the model and avoid overfitting. For instance, any hyperparameter θ_j can be constrained in a specific interval of variability. Alternatively, hyperparameter values can be drawn from an analytical prior, such as the Gaussian or the gamma probability density: $\theta_r \sim p(\theta_r)$. For example, by setting lengthscale priors that favor small values, the model will better capture low-periodicity patterns in the data.

The MMLE has become the common practice to tune lengthscale hyperparameters since it produces highly accurate GP models of smooth functions [16]. However, this approach may be suboptimal when all the hyperparameters need to be tuned together (hereby denoted as "full" MMLE), including the fundamental frequencies τ_m of periodic kernels. Since small errors in τ_m may cause the underfitting of oscillations, the MMLE solution may be narrow in the hyperparameter space and hard to converge on. Furthermore, the number of periodic components M has to be selected in advance. On one hand, increasing M allows to model more dynamic curves and may increase the fitting accuracy. On the other hand, if the number of periodic components is excessive, the GP may exhibit overfitting or numerical instabilities. Therefore, an upfront delay estimation approach is used in combination with the MMLE to better estimate the fundamental frequencies τ_m for the τGP. The two approaches are combined by following the scheme in Fig. 1 for the training of the τGPs. First, data samples of both real and

imaginary parts of an S-parameter $S_{a,b}$ are provided to the delay estimation algorithm. Next, the GPR is reiterated by updating the lengthscales according to the MMLE procedure. Lastly, the GPR is executed separately for the real and imaginary of the S-parameters, thus obtaining a $\tau GP_{re|a,b}$ and $\tau GP_{im|a,b}$ model.

D. Delay Estimation via Gabor Transform

The use of an upfront delay estimation method is motivated by the characterization of interconnects using transmission lines. In particular, using the extensively studied method of characteristics [21], [22] it is possible to factorize the transfer function $H_l(f)$ of a transmission line as follows:

$$H_l(f) = Q_l(f) \exp\left(-j 2\pi f \tau\right) \qquad (8)$$

where Q_l is a rational function representing the lossy and dispersive contributions and j is the imaginary unit, while the *propagation delay* $\tau = L\sqrt{\mathcal{LC}}$ is determined by the geometrical length L, the inductance $\mathcal{L}$ per unit length and capacitance $\mathcal{C}$ per unit length of the line. In [4], the same factorization is extended to a cascade of lumped elements or transmission line segments

$$H(f) \sim \sum_{m=1}^{M} Q_m(f) \exp(-j 2\pi f \tau_m) \qquad (9)$$

$$Q_m(f) = \sum_{c=1}^{C} \frac{r_{cm}}{j 2\pi f - p_c} \qquad (10)$$

where c are the indices of rational coefficients, τ_m are propagation delays, and r_{cm} and p_c denote residues and poles, respectively. Note that $Q_m(f)$ is the rational approximation used in the standard VF, while the DVF employs $H(f)$ for electrically long interconnects [4], [5]. Next, computing real and imaginary of $H(f)$ leads to

$$Re[H] = \sum_{m=0}^{M} \left(Re[Q] \cos\left(2\pi f \tau_m\right) - Im[Q] \sin\left(2\pi f \tau_m\right)\right) \qquad (11)$$

$$Im[H] = \sum_{m=0}^{M} \left(Re[Q] \sin\left(2\pi f \tau_m\right) + Im[Q] \cos\left(2\pi f \tau_m\right)\right) \qquad (12)$$

which are linear combinations of periodic functions of fundamental frequencies equal to τ_m. Therefore, for an interconnect modeled as a cascade of transmission lines, the real and imaginary parts of $H(f)$ are composed of periodic components determined by the propagation delays. This knowledge is incorporated into the τGP model of S-parameters via periodic kernel components, as stated in Section II-B. Consequently, the optimal τ_m hyperparameters for the τGP can be computed via a delay estimation algorithm. For this purpose, the Gabor transform is selected among the different techniques available in the literature due to its ease of implementation and accuracy in estimating propagation delays. Note that, the Gabor transform has been used in the literature in combination with the DVF modeling technique [4], [5]. A brief elucidation of the chosen delay estimation algorithm is hereby shown for the S_{11} of three coupled
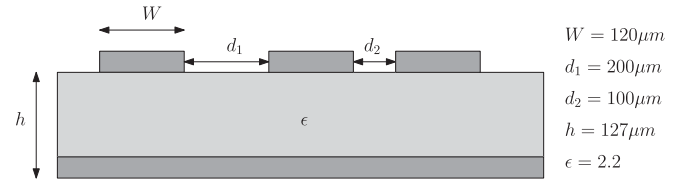

Fig. 2.   Application I. Cross section of the three coupled microstrips.
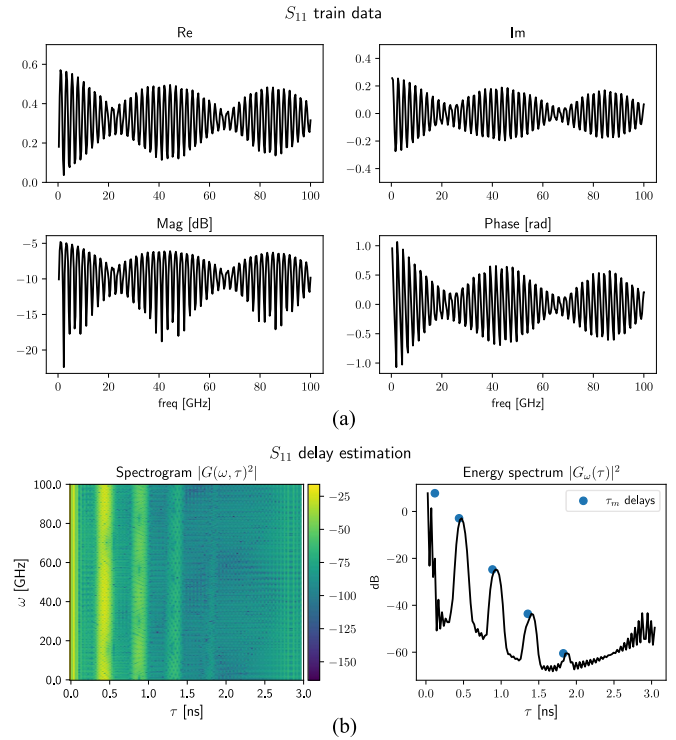

Fig. 3.   Application I: (a) Training samples for the S_{11} parameter. (b) Spectrogram (left) and energy spectrum over time (right) of the S_{11} parameter, computed using Gabor transform for the delay estimation.

microstrips (Application I). First, the Gabor transform [15] is computed on the S-parameters samples over the frequency [Fig. 3(a)]. The result is a time-domain spectrogram $G(\omega, \tau)$ (Fig. 3(b)), where ω, τ are the frequency and time variable, respectively. Second, the spectrogram $|G|^2$ is averaged across the frequency variable, resulting in the energy spectrum $|G_\omega(\tau)|^2$ [Fig. 3(b)]. Finally, the optimal τ_m hyperparameters are computed as the main peaks in the energy spectrum. The peaks are detected using a simple comparison of neighboring values. Note that for a more robust estimation, the τ_m delays are approximated with the central position of each energy peak, rather than with the position of the peak maximum. A detailed description of the delay estimation via Gabor transform is available in [4]. The resulting τ_m represents the principal time-domain components in which the transfer function can be decomposed. Using this algorithm, the number of delays M is given by the number of peaks that exceed a chosen threshold on the spectrogram energy $|G_\omega(\tau)|^2$. Note that the delay estimation via the Gabor transform may necessitate a manual tuning to the considered frequency
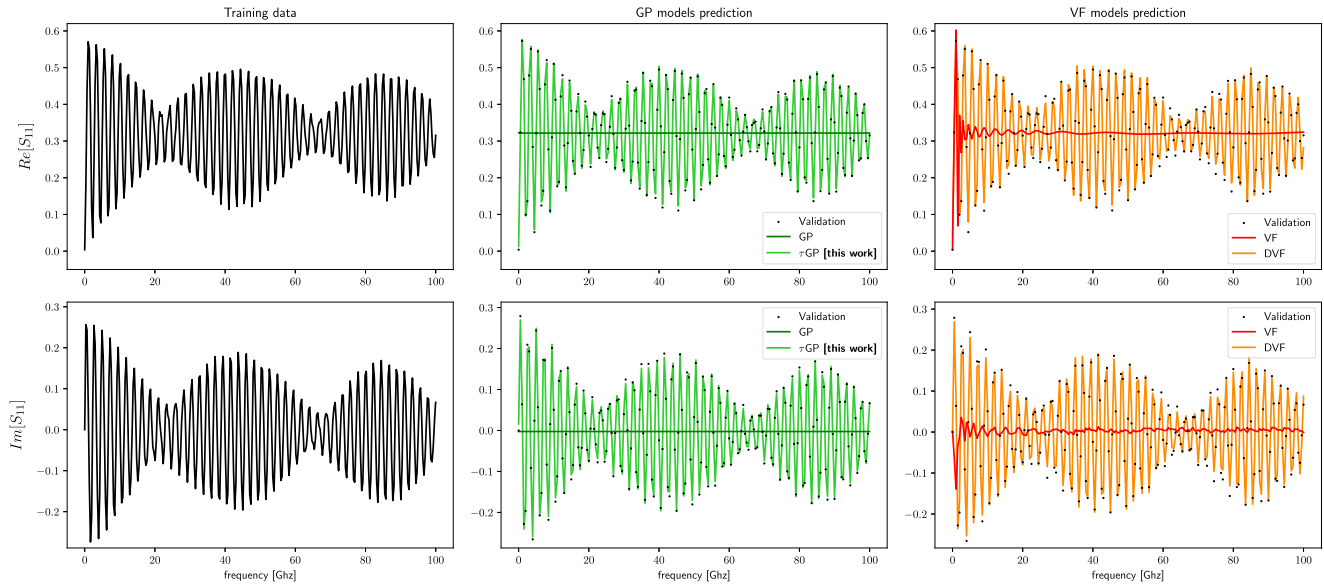
Fig. 4. Application I: S_{11} training samples (left) and predicted samples using the GP-based models (middle) and VF-based rational models (right); black dots represent the validation samples.
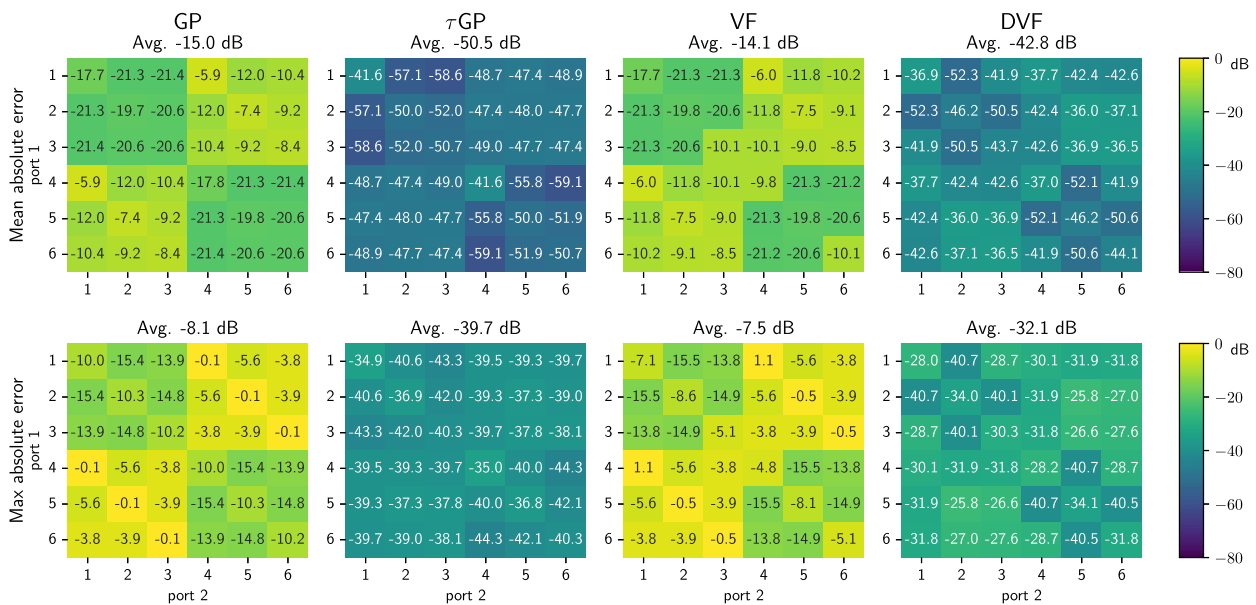

Fig. 5. Application I: Mean absolute error (top) and maximum absolute error (bottom) of S-parameter models on the validation set, across the frequency range, for each pair of ports.

band via several settings, such as the length of the filtering window for the spectrogram and the minimum peak width. The proposed delay estimation via Gabor transform, combined with the MMLE allows us to efficiently find the complete set of optimal hyperparameters.

III. APPLICATION EXAMPLES

In the next Sections, two simulated examples of interconnects are considered, with different ports and frequency samples. Each S-parameter is modeled separately for one instance of these devices. Therefore, the models are trained on only one realization of the frequency response. Four techniques are employed: VF, DVF, GP, and τGP. The other SVM and NN techniques mentioned in the Introduction are not suitable in this setting: Typically, multiple instances of frequency responses have to be provided to prevent them from overfitting. The VF and DVF models used in the examples are built on C = 50 and C = 10 pole/residue pairs (10), respectively. Note that, the total number of pole/residue pairs for DVF depends on the number of delays and is equal to $M \times C$, as indicated in (9).

Following the scheme in Fig. 1, two τGP models are separately built for both real and imaginary parts of each S-parameter.
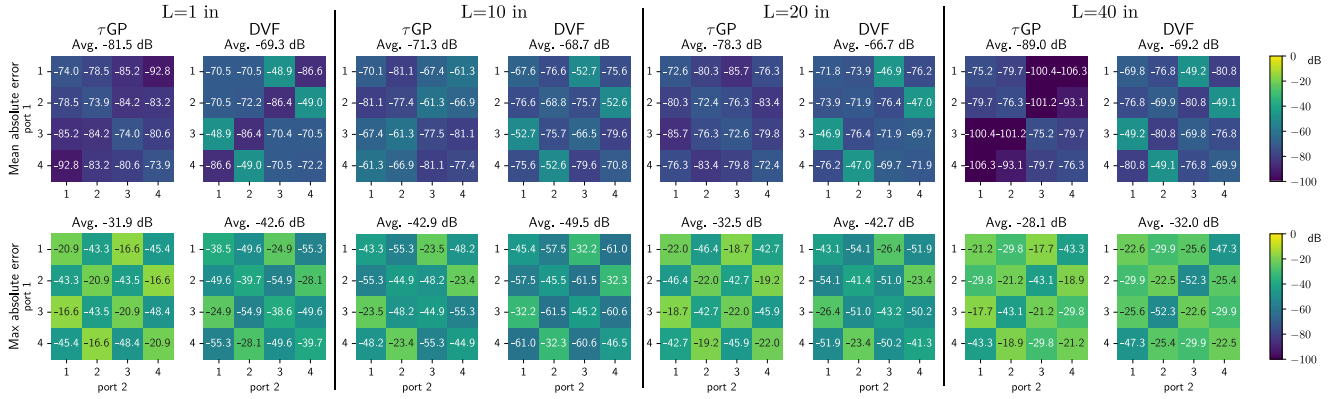
Fig. 6. Application II. Mean absolute error (top) and maximum absolute error (bottom) of single-ended S-parameter models, over the frequency, for each ports pair and for different routing lengths L.
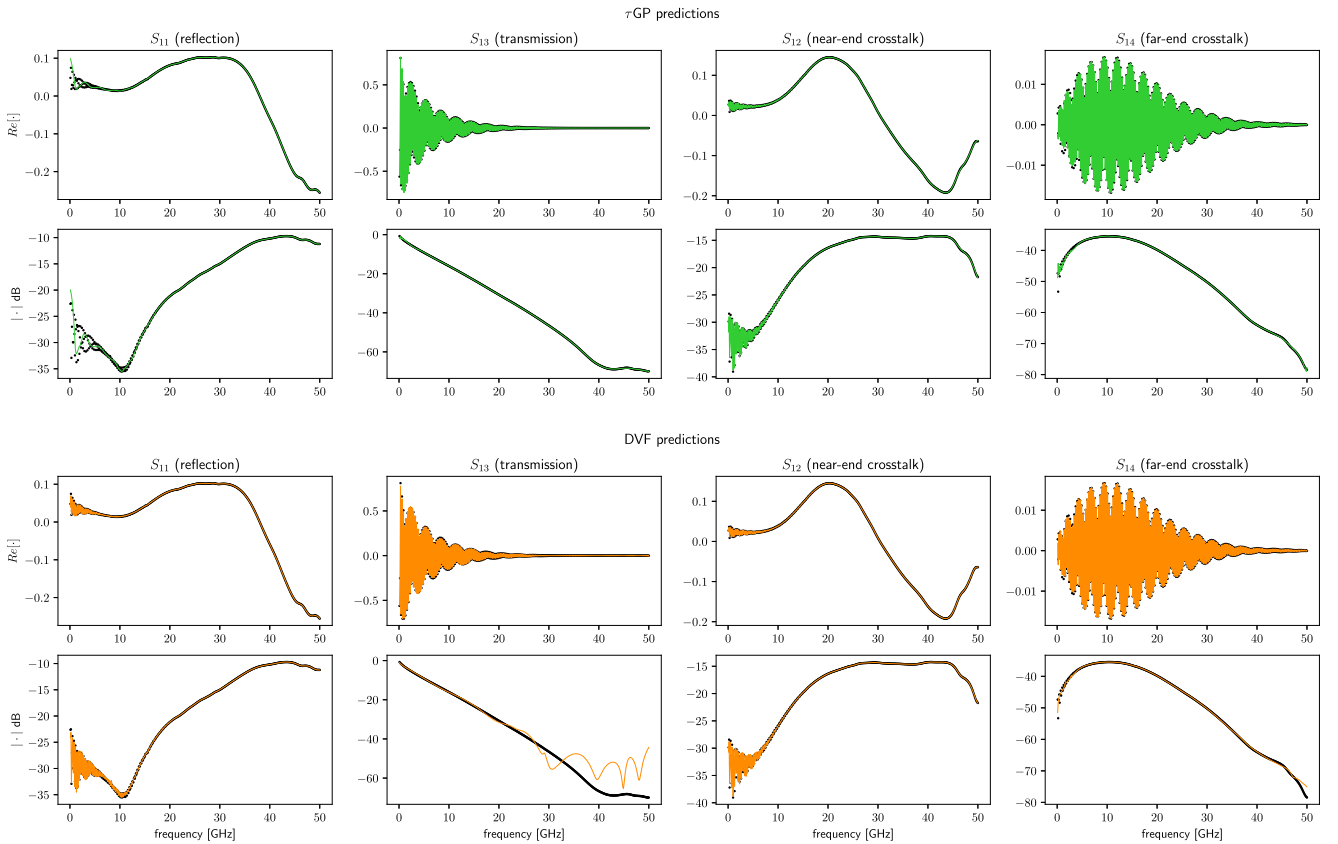

Fig. 7. Application II. Single-ended S-parameters of interest for $L = 1$ in, between the first pair of the interconnect: Comparison of validation samples (black dots) with predicted values using the τGP (top) and the DVF (bottom).

Therefore, the complex-valued S-parameters predicted by the τGPs are: $S_{a,b} = \tau GP_{re}|_{a,b} + j \cdot \tau GP_{im}|_{a,b}$.

In order to build the τGP and DVF models, the propagation delays could be computed from the geometrical and electrical characteristics of the interconnects: $\tau = L\sqrt{\mathcal{LC}}$. However, estimating the electrical characteristics is challenging due to the high coupling among the interconnect traces. Therefore, the upfront delay estimation via the Gabor transform is preferred to directly extract the delays from the S-parameter samples.

Note that, in both applications, the delay estimation settings are tuned separately for the DVF and for the proposed τGP models, in order to maximize the models' accuracy. Nonetheless, the

delay estimation settings are fixed among all the S-parameters of the device. In addition, to reach the optimal performance, in the DVF the estimated delays are adjusted by executing a local numerical optimization until the lowest possible fitting error is obtained on the training samples.

A. Application I: Coupled Microstrip Lines

In this example, three coupled transmission lines are considered. This 6-port device is formed by microstrips traces that are 5 cm long, with cross-section dimensions reported in Fig. 2. In this device, the crosstalk among pairs of transmission lines

TABLE II
APPLICATION I, FULL MMLE VS UPFRONT DELAY ESTIMATION: MEAN
ABSOLUTE ERROR (DECIBEL) OVER THE FREQUENCY

τ**GP** with:	**MMLE**			**Delay est.**
	M=2	M=3	M=4	
Avg. $S_{n,n}$	−23.8	−14.62	−16.04	**−46.44**
Avg. $S_{n,n+1}$	−20.12	−11.17	−12.9	**−50.92**
Avg. S (all)	−22.31	−13.33	−9.47	**−49.1**

TABLE III
SUMMARIZED COMPARISON BETWEEN τGP AND DVF

	τGP	DVF
Base model	Bayesian (GP)	rational fitting (VF)
Delays est. method	Gabor transform	Gabor transform
Delay implementation	kernel hyperparam.	exponential term
Complex function fitting	no	yes
Uncertainty estimation	yes	no
Comp. complexity	$\mathcal{O}(n_f^3)$	$\mathcal{O}(n_f)$
(w.r.t frequency samples)		
Application I		
No. training samples (n_f)	306	306
Avg. modeling error	$\sim$−50 dB	$\sim$−43 dB
Model training time	$\sim$10 s	$\sim$5 s
(6×6 S-matrix)		
Application II		
No. training samples (n_f)	500	500
Avg. modeling error	$\sim$−80 dB	$\sim$−68 dB
High-frequency accuracy	higher	lower
Model training time	$\sim$20 s	$\sim$5 s
(4×4 S-matrix)		

varies due to the different spacing (d_1, d_2). Using a 50 Ω reference impedance, the device's single-ended S-parameters are simulated with an EM solver. For this DUT, 306 linearly-spaced samples are simulated for training, and 200 linearly-spaced samples for validation, over the same frequency range from 1 to 100 GHz.

As an illustration, Fig. 3(a) shows the training samples for S_{11} parameter which manifest strong oscillatory behaviors.

Following the scheme in Fig. 1, a τGP model is built for both real and imaginary parts of each S-parameter. Therefore the predicted, complex-valued, S-parameters are: $S_{a,b} = \tau GP_{re}|_{a,b} + j \cdot \tau GP_{im}|_{a,b}$. Fig. 3(b) reports an example of the spectrogram and the energy spectrum obtained using the Gabor transform on the S_{11} parameter. As described in Section II-D, the peaks of the energy spectrum represent the estimated propagation delays. Subsequently, τGPs are built on the training data using the estimated delays τ_m and the lengthscale hyperparameters optimized via MMLE. For this application, the GP model uses the standard, general-purpose squared exponential kernel [13] for both real and imaginary parts. Fig. 4 illustrates the prediction of each model for the real part of the S_{11} parameter. Results show that the standard GP and the standard VF do not allow a correct representation of the oscillations across the frequency variable, while the proposed τGP and DVF closely match the validation data. In addition, Fig. 5 summarizes the mean and the maximum error of the models' prediction, averaged for different sets of S-parameters: The DVF and τGP consistently produce low prediction errors. On the contrary, the other models do not appear suitable for this modeling task. Additionally, it is worth testing the performance of the τGP when a full MMLE is employed to find both optimal delays and lengthscale hyperparameters. For this test, the number of delays M has to be fixed in advance. The results of this test are summarized in Table II. It evinces from the Table that the delay estimation improves the hyperparameter tuning process. In fact, unlike the Gabor estimation, the full MMLE does not produce an accurate fitting for all the elements of S-matrix and the maximum error strongly depends on the assumed number of delays M. Therefore, although the new kernel k_{del} allows sufficient modeling power, the delay estimation is necessary to consistently achieve high accuracy.

B. Application II: PCB-Based Data Link

This example consists of a single-board, PCB-based data link interconnect [20]. The S-parameters are extracted using a physics-based model for different routing lengths $L \in [1, 40]$ in. The interconnect, which occupies the outer nine layers of a board with around 40 layers and an effective dielectric constant of $\varepsilon_r = 3.7$ and a loss tangent of $\tan \delta = 0.01$, comprises 22 differential signal pairs. These signals are transmitted via pin fields with a pitch of 60 mil and are susceptible to crosstalk due to this configuration. The single-ended S-parameters for the first four interconnect ports are extracted at 1001 frequency points, ranging from 0.05 to 50 GHz. The data points are equally split into train and validation sets across the frequency variable. Next, the different modeling strategies are executed. In particular, the mean and envelope lengthscale hyperparameters of the τGP are drawn from a Gamma prior distribution: $l_1, l_2 \sim \Gamma(5, 5)$. This improves the conditioning of the MMLE procedure for this application, as suggested in Section II-C.

Similarly to Application I, the GP with standard kernels and VF do not allow for a proper fit of the training samples. On the contrary, Fig. 6 shows that both the new τGP and the DVF produce highly accurate results for all S-matrix elements and for all L values. In particular, τGP models seem to produce a lower mean error across frequency than DVF, but a higher maximum error. This is confirmed by Fig. 7 that illustrates different interconnect S-parameters: Reflection, transmission, near-end crosstalk, and far-end crosstalk for the first differential signal pair.

Furthermore, in this application, it is particularly interesting to test the prediction of insertion loss (IL) and return loss (RL) figures of merit. In fact, the high accuracy obtained by τGP and DVF for the real and imaginary part of single-ended S-parameters enables the extraction of magnitude, phase, mixed-mode S-parameters, or other frequency-dependent figures of merit. Fig. 8 reports the mixed-mode IL and RL calculated from the models' prediction. It appears that at high frequency τGP generally allows for better retrieval of the IL than DVF, while the model's accuracy is similar for the RL. Finally, a comparison between the proposed τGP and the DVF is summarized in Table III, including an approximate computation time for the model building.
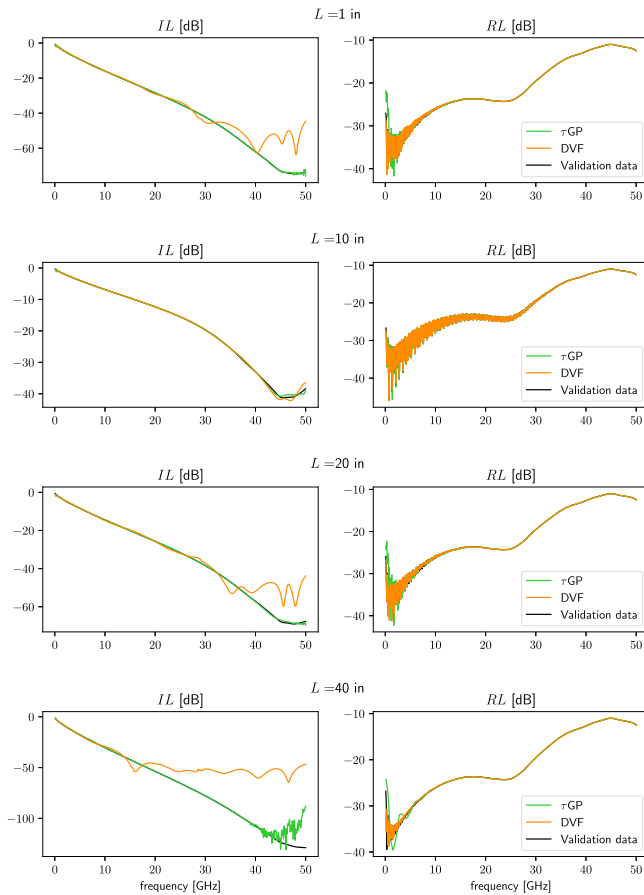
Fig. 8. Application II. Comparison of mixed-mode insertion loss (IL) and mixed-mode return loss (RL), that are derived from τGP and DVF, built on the original single-ended S-parameters, for different routing lengths L.

IV. Conclusion

The presented delayed Gaussian process model (τGP) is able to model S-parameters of electrically long interconnects over frequency. Unlike Gaussian processes with general-purpose kernels, the τGP is able to capture the S-parameters oscillations caused by propagation delays along the interconnect. Experimental results show that the use of τGP is to be preferred to the pre-existing delayed vector fitting (DVF) rational modeling technique when the maximum modeling accuracy is desired. The τGP appears to be especially more accurate than DVF at high frequencies, where the S-parameters likely deviate from rational approximation. The proposed technique introduces the possibility of modeling frequency-dependent figures of merit using stochastic machine-learning models. Thus, it constitutes an initial step towards a deeper integration of physical information in Bayesian models like Gaussian processes. Further efforts are needed to extend the new technique to the modeling of S-parameter values over additional design variables and to enforce physical properties, such as causality and passivity, on the predictions of the model.

References

[1] S. Grivet-Talocia and B. Gustavsen, "Black-box macromodeling and its EMC applications," *IEEE Electromagn. Compat. Mag.*, vol. 5, no. 3, pp. 71–78, Third Quarter 2016.

[2] B. Gustavsen and A. Semlyen, "Rational approximation of frequency domain responses by vector fitting," *IEEE Trans. Power Del.*, vol. 14, no. 3, pp. 1052–1061, Jul. 1999.

[3] D. Deschrijver, M. Mrozowski, T. Dhaene, and D. De Zutter, "Macromodeling of multiport systems using a fast implementation of the vector fitting method," *IEEE Microw. Wireless Compon. Lett.*, vol. 18, no. 6, pp. 383–385, Jun. 2008.

[4] A. Chinea, P. Triverio, and S. Grivet-Talocia, "Delay-based macromodeling of long interconnects from frequency-domain terminal responses," *IEEE Trans. Adv. Packag.*, vol. 33, no. 1, pp. 246–256, Feb. 2010.

[5] M. Sgueglia, A. Sorrentino, M. de Magistris, D. Spina, D. Deschrijver, and T. Dhaene, "A novel parametric macromodeling technique for electromagnetic structures with propagation delays," in *Proc. IEEE 21st Workshop Signal Power Integrity*, Lake Maggiore, Italy, 2017, pp. 1–4.

[6] Y. Zhuo, F. Feng, J. Zhang, and Q. -J. Zhang, "Parametric modeling incorporating joint polynomial-transfer function with neural networks for microwave filters," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 11, pp. 4652–4665, Nov. 2022.

[7] J. Jin, F. Feng, W. Zhang, J. Zhang, Z. Zhao, and Q. -J. Zhang, "Recent advances in deep neural network technique for high-dimensional microwave modeling," in *Proc. IEEE MTT-S Int. Conf. Numer. Electromagn. Multiphys. Model. Optim.*, Hangzhou, China, 2020, pp. 1–3.

[8] H. M. Torun, A. C. Durgun, K. Aygün, and M. Swaminathan, "Causal and passive parameterization of S-parameters using neural networks," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 10, pp. 4290–4304, Oct. 2020.

[9] M. Schierholz, I. Erdin, J. Balachandran, C. Yang, and C. Schuster, "Parametric S-parameters for PCB based power delivery network design using machine learning," in *Proc. IEEE 26th Workshop Signal Power Integrity*, Siegen, Germany, 2022, pp. 1–4.

[10] 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, 2019.

[11] C. E. Rasmussen and C. K. I. Williams, "Regression," in *Gaussian Processes for Machine Learning*, ch. 2, Cambridge, MA, USA: MIT Press, 2005, pp. 7–32.

[12] S. Koziel, S. Ogurtsov, I. Couckuyt, and T. Dhaene, "Variable-fidelity electromagnetic simulations and co-kriging for accurate modeling of antennas," *IEEE Trans. Antennas Propag.*, vol. 61, no. 3, pp. 1301–1308, Mar. 2013.

[13] A. Wilson and R. Adams, "Gaussian process kernels for pattern discovery and extrapolation," in *Proc. 30th Int. Conf. Mach. Learn.*, PMLR, Atlanta, GA, USA, 2013, pp. 1067–1075.

[14] F. Garbuglia, T. Claeys, I. Couckuyt, D. Deschrijver, and T. Dhaene, "Bayesian active learning for radiation pattern sampling over cylindrical surfaces," *IEEE Trans. Electromagn. Compat.*, vol. 64, no. 5, pp. 1391–1398, Oct. 2022.

[15] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, PA, USA: Soc. Ind. Appl. Math., 1992.

[16] C. E. Rasmussen and C. K. I. Williams, "Model selection and adaptation of hyperparameters," in *Gaussian Processes for Machine Learning*, ch. 5, Cambridge, MA, USA: MIT Press, 2005, pp. 105–129.

[17] A. Damianou and N. D. Lawrence, "Deep Gaussian processes," in *Proc. 16th Int. Conf. Artif. Intell. Statist.*, 2013, pp. 207–215.

[18] F. Garbuglia, D. Spina, D. Deschrijver, I. Couckuyt, and T. Dhaene, "Bayesian optimization for microwave devices using deep GP spectral surrogate models," *IEEE Trans. Microw. Theory Techn.*, vol. 71, no. 6, pp. 2311–2318, Jun. 2023.

[19] H. M. Torun and M. Swaminathan, "High-dimensional global optimization method for high-frequency electronic design," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 6, pp. 2128–2142, Jun. 2019.

[20] T. Reuschel et al., "Efficient prediction of equalization effort and channel performance for PCB-Based data links," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 7, no. 11, pp. 1842–1851, Nov. 2017.

[21] S. Grivet-Talocia, H.-M. Huang, A. E. Ruehli, F. Canavero, and I. M. Elfadel, "Transient analysis of lossy transmission lines: An efficient approach based on the method of characteristics," *IEEE Trans. Adv. Packag.*, vol. 27, no. 1, pp. 45–56, Feb. 2004.

[22] F.-Y. Chang, "The generalized method of characteristics for waveform relaxation analysis of lossy coupled transmission lines," *IEEE Trans. Microw. Theory Techn.*, vol. 37, no. 12, pp. 2028–2038, Dec. 1989.

Federico Garbuglia received the M.Sc. degree in electronic engineering from the Polytechnic University of Marche, Ancona, Italy, in 2019. He is currently working toward the Ph.D. degree in electrical engineering with the Surrogate Modelling Lab (SUMO Lab), IDLab, Ghent University, Ghent, Belgium.

His current research interests include machine learning and active learning for electronics design automation and optimization.

Torsten Reuschel (Member, IEEE) received the B.Sc. degree in general engineering sciences, the M.Sc. degree (Hons.) in electrical engineering, and the Ph.D. degree in electrical engineering from the Hamburg University of Technology (TUHH), Hamburg, Germany, and with the University of Waterloo, ON, Canada, in 2010, 2013, and 2018, respectively.

As of 2021, he is a Project Lead for high-frequency systems engineering with the Canadian High Arctic Ionospheric Network, University of New Brunswick (UNB), Fredericton, NB, Canada. Prior to that, he was with the Automotive Electronics division of the Robert Bosch GmbH, Gerlingen, Germany, advancing modeling and simulation methods for electromagnetic compatibility of advanced electronic control units and high-speed digital automotive communication. His current research interests include design of measurement instrumentation as well as combined analysis of signal integrity and electromagnetic wave propagation in high-speed digital systems.

Dr. Reuschel was the recipient of IEEE TRANSACTIONS ON COMPONENTS, PACKAGING, AND MANUFACTURING TECHNOLOGY Best Paper Award in 2016 and DesignCon Best Paper Awards in 2017 and 2018.

Christian Schuster (Senior Member, IEEE) received the Diploma degree in physics from the University of Konstanz, Konstanz, Germany, in 1996, and the Ph.D. degree in electrical engineering from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 2000.

From 2006, he is Full Professor and Head of the Institute of Electromagnetic Theory, Hamburg University of Technology (TUHH), Hamburg, Germany. Prior to that, he was with the IBM T. J. Watson Research Center, Yorktown Heights, NY, USA, where he was involved in high-speed optoelectronic package and backplane interconnect modeling and signal integrity design for new server generations. He is also an Adjunct Associate Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology, GA, USA. His current research interests include signal and power integrity of digital systems, multiport measurement and calibration techniques, and development of electromagnetic simulation methods for communication electronics.

Dr. Schuster was the recipient of IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY Best Paper Awards in 2001 and 2015, TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY Best Paper Awards in 2012 and 2016, IEC DesignCon Paper Awards in 2005, 2006, 2010, 2017, and 2018, three IBM Research Division Awards from 2003 to 2005, IBM Faculty Awards in 2009 and 2010, and the Sustained Service to the EMC Society Award in 2019. He is a Member of the German Physical Society (DPG) and several technical program committees of international conferences on signal and power integrity, and electromagnetic compatibility. He was Distinguished Lecturer for IEEE EMC Society from 2012–2013, as the Chair of the German IEEE EMC Chapter from 2016–2019, and as a Member of the Board of Directors of the EMC Society in 2015 and from 2020–2022. He is currently an Associate Editor for IEEE TRANSACTIONS ON ELECTROMAGNETIC COMPATIBILITY.

Dirk Deschrijver (Senior Member, IEEE) received the first Ph.D. degree in computer science from the Department of Mathematics and Computer Science, University of Antwerp, Antwerp, Belgium, in 2007, and a second Ph.D. degree in engineering from Ghent University, Ghent, Belgium, in 2012.

From 2008 to 2014, he was an FWO Postdoctoral Research Fellow with the IBCN Research Group, ID-Lab, Ghent University, where since 2016, he has been an Associate Professor, working on data analytics, machine learning, and surrogate modeling.

Tom Dhaene (Senior Member, IEEE) received the M.S. and the Ph.D. degrees in electrical engineering from Ghent University, Ghent, Belgium, in 1989 and 1993, respectively.

Since 2000, he has been Professor with the Computer Modeling and Simulation Research Group, University of Antwerp, Belgium. Since 2007, he is a Full Professor with IDLab, Ghent University, researching on distributed scientific computing, machine learning, bioinformatics, signal integrity, electromagnetic compatibility, optimal design, surrogate modeling, and circuit and EM modeling.

Domenico Spina (Senior Member, IEEE) received the M.S. degree (cum laude) in electronics engineering from the University of L'Aquila, L'Aquila, Italy, in 2010, and the joint Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, and the University of L'Aquila in 2014.

Since 2015, he has been a Member of the Internet and Data Science Lab (IDLab) of the Department of Information Technology (INTEC), Ghent University - imec, first as Postdoctoral Researcher and currently as R&D Technical Lead. His research interests include modeling and simulation, uncertainty quantification, and machine learning for analog electronic and photonic systems.