Composite Delay-Informed Rational Kernel for Modeling of Electrically Long Interconnects

Tekalegn Mezemr Yihune, Thijs Ullrick, *Member, IEEE*, Tom Dhaene, *Senior Member, IEEE*, Dries Peumans, *Member, IEEE*, and Domenico Spina, *Senior Member, IEEE*

***Abstract*—This work presents a novel composite Gaussian process (GP) model for representing the complex-valued scattering parameters (S-parameters) of electrically long microwave interconnects as a function of frequency. The proposed model integrates a rational Szegö kernel, which captures the holomorphic structure and Hermitian symmetry of S-parameters, with a delay-informed periodic kernel designed to capture oscillatory behavior induced by wave propagation, thus resulting in improved predictive accuracy.**

***Index Terms*—Gaussian processes (GPs), microwave interconnects, propagation delays, scattering parameters (S-parameters).**

I. Introduction

ACCURATE modeling of scattering parameters (S-parameters) is essential for designing and optimizing high-frequency circuits, especially in scenarios involving electrically long interconnects. As the bandwidth increases and wave propagation effects become dominant, interconnects' S-parameters exhibit a dynamic and complex behavior.

Classical modeling techniques, most notably vector fitting (VF) and its delay-aware extension, delayed VF (DVF), have long been the standard for rational approximation of such frequency responses [1]. Their primary advantage lies in the ability to generate compact, interpretable models suitable for both frequency- and time-domain simulations [2]. However, these methods are inherently deterministic and lack the ability to quantify predictive uncertainty, which limits their application in robust design workflows and sensitivity analysis.

In recent years, data-driven modeling techniques have gained traction for this task. Artificial neural networks have shown promise in capturing nonlinear dependencies across frequency and design parameters [3], [4]. Yet, their reliance on extensive training data poses challenges in high-frequency applications, where data acquisition can be costly (i.e., full-wave simulations or measurements). Support vector machines mitigate overfitting through strong regularization, but often use generic kernels ignoring the underlying physics of electromagnetic propagation [5].

Gaussian process (GP) models have emerged as a compelling alternative, offering high data efficiency and the ability to encode prior physical knowledge through kernel design. Yet, standard stationary kernels are inadequate for modeling the nonsmooth, oscillatory behavior induced by propagation delays. Recent efforts have introduced physics-informed kernels to address this, such as those incorporating rational structure or delay-induced periodicity. However, these approaches typically operate in isolation, focusing either on preserving holomorphy [6] or capturing delay effects [7].

This work introduces a hybrid kernel that extends the rational Szegö formulation [6] by integrating a delay-informed kernel through a structured Hadamard product. It retains essential analytic structure through the Szegö component, while enabling the model to capture delay-induced oscillatory behavior observed in electrically long interconnects. The proposed kernel preserves Hermitian symmetry and enforces physically consistent cross-correlations between real and imaginary components, reflecting the structure of passive linear systems. A suitable application example validates the accuracy of the proposed modeling method.

II. Methodology

GPs are nonparametric, probabilistic models that place a multivariate normal prior over functions. Unlike fixed-parameter models, GPs adapt their complexity with data volume, enabling efficient learning and accurate predictions even from limited data. GPs are fully specified by mean $m(x)$ and kernel $k(x, x')$ functions, which encode prior beliefs and correlations, respectively. The kernel's design is critical for capturing behaviors like delay-induced oscillations in electrically long interconnects. Predictions at new inputs are made by conditioning the prior on observed data, yielding the posterior via standard GP regression (GPR). This posterior provides both predictions and uncertainty estimates, supporting robust design and sensitivity analysis [8]. Note that, the formulation adopted in the following sections refers to the modeling of a single, complex-valued element of the S-parameters with respect to frequency, for simplicity.

A. Rational Szegö Kernel

The rational Szegö kernel was introduced in [9] and adapted for microwave systems in [6]. It is designed for complex-valued functions that are holomorphic in the frequency

Received 18 July 2025; revised 13 September 2025 and 22 October 2025; accepted 31 October 2025. This work was supported in part by the Strategic Research Program of the Vrije Universiteit Brussel (VUB) under Grant SRP-78 and in part by the Research Council of the VUB. *(Corresponding author: Tekalegn Mezemr Yihune.)*

Tekalegn Mezemr Yihune and Domenico Spina are with the Department ELEC, Vrije Universiteit Brussel, 1050 Brussel, Belgium (e-mail: tekalegn.mezemr.yihune@vub.be).

Thijs Ullrick is with the Department of Information Technology, IMEC, Ghent University, 9052 Ghent, Belgium, and also with the Department ELEC, Vrije Universiteit Brussel, 1050 Brussel, Belgium.

Tom Dhaene is with the Department of Information Technology, IMEC, Ghent University, 9052 Ghent, Belgium.

Dries Peumans is with the Department ELEC, Vrije Universiteit Brussel,1050 Brussel, Belgium, and also with IMEC, 3001 Leuven, Belgium.

Digital Object Identifier 10.1109/LMWT.2025.3629238

domain, such as S-parameters. It enforces key physical properties, including analyticity, Hermitian symmetry, and causality, making it particularly well-suited to modeling passive linear systems. This kernel is expressed as

$$K_{\text{sz}}(s_0, s_1) = \begin{bmatrix} \Re\left(\frac{k+c}{2}\right) & \Im\left(\frac{-k+c}{2}\right) \\ \Im\left(\frac{k+c}{2}\right) & \Re\left(\frac{k-c}{2}\right) \end{bmatrix} \quad (1)$$

with the Szegö covariance and pseudo-covariance functions

$$k(s_0, s_1) = \frac{\sigma^2}{2\alpha + s_0 + s_1^*}, \quad c(s_0, s_1) = \frac{\sigma^2}{2\alpha + s_0 + s_1} \quad (2)$$

where $s = j2\pi f$ is the Laplace variable, the symbols $\Re(\cdot)$ and $\Im(\cdot)$ indicate the real and imaginary part, respectively, while α and σ are hyperparameters to be learned during training. While powerful, the rational Szegö kernel is not able to capture accurately the complex patterns caused by propagation delays, as shown in Section III.

B. Delayed GP Kernel

In order to model the dynamic behavior of electrically long interconnects, [7] proposed a novel physics-inspired kernel

$$k_{\text{del}}(f, f') = k_{\text{mean}}(f, f') + k_{\text{env}}(f, f') \prod_{m=1}^{M} k_{\text{per},m}(f, f') \quad (3)$$

where f and f' indicate frequency values. Note that k_{mean} and k_{env} are Matérn 5/2 kernels [7], while the periodic kernel $k_{\text{per},m}$ is defined as

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

where τ_m for $m = 1, \ldots, M$ indicate the propagation delays, while ℓ_m for $m = 1, \ldots, M$ are lengthscale hyperparameters to be learned during training, controlling the length scale of oscillations. The fundamental idea is that k_{mean} captures the baseline behavior of the S-parameters, k_{env} models the S-parameters' envelope, while $k_{\text{per},m}$ describes the oscillatory behavior induced by distributed propagation delays. Note that the kernel in (3) is real-valued, hence two separate kernels must be adopted to model the real and imaginary part of each element of the S-parameters.

C. Composite Rational-Delayed Kernel

To accurately model the S-parameters of electrically long interconnects, it is necessary to account for both the causal behavior and oscillatory effects induced by propagation delays. To this end, we propose a composite kernel that combines the rational Szegö kernel K_{sz} in (1), which enforces analytic structure and Hermitian symmetry, with a delay-aware real-valued kernel k_{tau}, via an elementwise (Hadamard) product

$$K_{\text{composite}}(s_0, s_1) = K_{\text{sz}}(s_0, s_1) \circ \left[k_{\text{tau}}(f_0, f_1) \cdot \mathbf{1}_{2\times 2}\right] \quad (5)$$

where $\mathbf{1}_{2\times 2}$ denotes a 2×2 matrix of ones, f_0 and f_1 are the frequency values corresponding to s_0 and s_1, respectively, and

$$k_{\text{tau}}(f_0, f_1) = k_{\text{env}}(f_0, f_1) \prod_{m=1}^{M} k_{\text{per},m}(f_0, f_1). \quad (6)$$

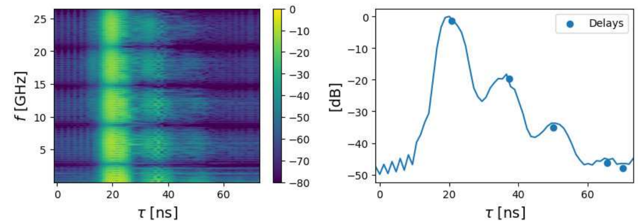

Fig. 1. Example of delay estimation via the Gabor transform: spectrogram (left) and energy spectrum (right) of the element S_{12} of the S-parameters of the DUT in Section III.

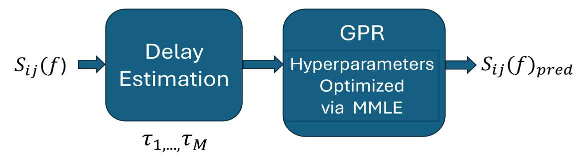

Fig. 2. Modeling framework: the delay estimation of each element S_{ij} of the S-parameters is followed by GPR, optimizing hyperparameters via MMLE.

TABLE I
PREDICTION ACCURACY OF THE GP MODELS

S-param	Delay	MaxAE (dB)			MAE (dB)			MAE (dB)
	(ns)	K_{sz}	K_{del}	$K_{composite}$	K_{sz}	K_{del}	$K_{composite}$	5-fold
S_{11}	4	-9.01	-17.32	**-29.58**	-19.43	-35.27	**-57.44**	-50.79 ± 0.86
S_{12}	3	-8.99	-19.92	**-25.38**	-26.53	-46.95	**-54.54**	-52.33 ± 0.60
S_{13}	16	-10.42	-20.57	**-23.34**	-27.62	-45.94	**-53.37**	-51.26 ± 0.37
S_{22}	2	-7.89	-19.11	**-30.90**	-26.27	-37.54	**-57.82**	-50.53 ± 0.71
S_{23}	16	-12.42	-22.69	**-27.91**	-36.62	-47.03	**-55.89**	-51.91 ± 0.46
S_{33}	6	-6.72	-10.21	**-21.69**	-28.62	-47.41	**-58.05**	-50.95 ± 0.65

The kernels $k_{\text{env}}(f_0, f_1)$ and $k_{\text{per},m}(f_0, f_1)$ are described in Section II-B: the first is a Matérn 5/2 kernel and the second is shown in (4). In (5), the rational kernel K_{sz} effectively models the mean behavior of the S-parameters, while the Hadamard product serves to modulate it by incorporating delay-induced oscillations and spectral envelope effects [7].

The proposed kernel tailors the GP prior to both the global analytic structure and the localized frequency behavior, a strategy consistent with broader practices in structured kernel design [10]. This reflects the physical intuition that oscillatory delay effects are superimposed on a globally analytic transfer function.

The resulting kernel remains positive semidefinite (PSD) due to the Schur product theorem [11], and thus defines a valid covariance function for multi-output GP regression. Specifically, $K_{\text{composite}}$ in (5) depends on three PSD kernels: the Szegö kernel, the Matérn-5/2 kernel, and the periodic kernel described in (4). Note that the matrix $\mathbf{1}_{2\times 2}$ in (5) is also PSD, with eigenvalues $\{0, 2\}$, so the block factor $k_{\text{tau}}(f_0, f_1) \cdot \mathbf{1}_{2\times 2}$ is PSD.

The proposed kernel retains the matrix structure of K_{sz} in (1), thereby preserving Hermitian symmetry and enforcing cross-correlations between the real and imaginary parts, which are critical for modeling causal systems. Note that, the hyperparameters of k_{env} and $k_{\text{per},m}$ in (5) and (6) are jointly trained for the complex-valued S-parameters, while they are trained separately and independently for the real and imaginary part of the S-parameters in [7]. However, the

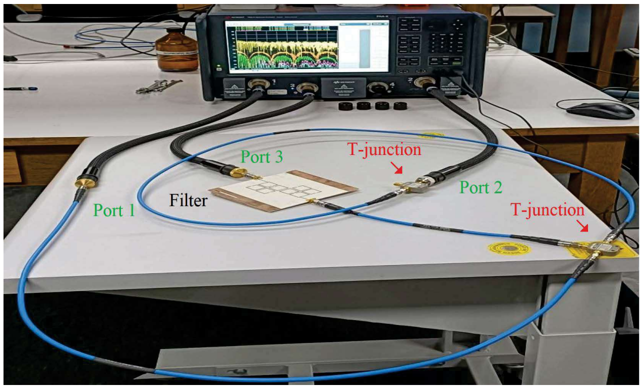

Fig. 3. Illustration of the measurement setup.

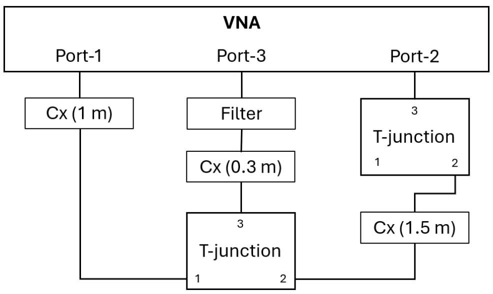

Fig. 4. Block diagram of the measurement setup showing how VNA ports, coaxial cables (Cx), filter, and T-junctions are connected.

passivity of the computed S-parameters is not guaranteed during training, given that each element of the S-parameter curve is modeled separately. Passivity can be enforced as a postprocessing step on the predicted S-parameter values by using suitable numerical techniques [3].

The value of the propagation delays τ_m is extracted via the Gabor transform of the S-parameters. First, a Gaussian window is used to filter the frequency response, localizing energy jointly in frequency and time. The squared magnitude of the transform produces a spectrogram, where local maxima along the time axis correspond to dominant delays, as shown in Fig. 1(a). To determine the number of delays M, the spectrogram is averaged over frequency, and each delay's energy contribution is calculated, as shown in Fig. 1(b). Only those delays whose relative energy exceeds a predefined threshold are retained [12]. All other hyperparameters in (5) and (6) are optimized via maximum marginal likelihood (MMLE) during model training. Since the composite kernel comprises more components compared to standard kernels, resulting in a larger set of hyperparameters, the optimization may be sensitive to initialization. Therefore, the optimizer was initialized with hyperparameter values obtained from earlier fits of the same model on the same dataset, leading to stable convergence. The proposed modeling framework is shown in Fig. 2.

III. Application Example

The proposed methodology can be applied to both simulated and measured S-parameter data. In this work, we consider the challenging setup shown in Fig. 3, where the S-parameters are measured via a vector network analyzer (VNA). The system under study is shown in the block diagram of the DUT in Fig. 4. It includes three interconnected Teledyne Reynolds model 269-0221 microwave-grade coaxial cables, a bandpass filter centered around 1 GHz with a 100 MHz bandwidth, and two T-junctions realized with HP 11667B resistive power splitters rated for up to 26.5 GHz, connected as follows. A 1-m cable connects Port 1 to a T-junction. From this junction, a 0.3-m branch leads to the bandpass filter connected to Port 3. The other end of this T-junction is connected to a 1.5-m cable, terminated at the second T-junction. One port of the latter T-junction is connected to a VNA (Port 2), while the remaining port is left open to induce a controlled reflection.

S-parameter measurements are performed using a four-port PNA-X VNA over the frequency range [10 MHz–26.5 GHz]. To evaluate the model's generalization ability, 4000 measured samples of S-parameters are considered, divided evenly into training (used to estimate the model's hyperparameters) and test sets (used to evaluate its prediction accuracy). For 2000 frequency samples per S-parameter curve, training the proposed kernel was completed within 1 min on a laptop with Intel i7 CPU, 64 GB RAM, and prediction over the full test set required less than 10 s. Here, we compare the performance of the proposed GPR method with respect to the techniques illustrated in Sections II-A and II-B. During training, the hyperparameters for each GP model are determined by minimizing MMLE, and the models' accuracy is assessed on the test sets using mean absolute error (MAE) and maximum absolute error (MaxAE). These metrics for each kernel, along with the number of propagation delays τ_m adopted, are summarized in Table I. Note that, the kernels in (3) and in (5) adopt the same number and values of propagation delays for each element of the S-parameters, estimated via the Gabor transform [12]. Additionally, the network considered is reciprocal, hence the corresponding scattering matrix is symmetric: the modeling accuracy in Table I is reported only for the upper triangular elements of the scattering matrix. The proposed complex-valued kernel incorporating propagation delay structure, indicated in Table I as $K_{\text{composite}}$, consistently achieves lower MAE and MaxAE than both the Szegö kernel K_{sz}, which do not consider delay information, and the kernel K_{del} proposed in [7], which includes delays but models the real and imaginary parts separately. The last column in Table I shows the MAEs' mean and standard deviation for $K_{\text{composite}}$ obtained via fivefold cross-validation over the 2000 training data (400 samples per fold): the proposed GP is robust to different training sets. Since delays are nonlearnable, they have not been estimated anew during cross-validation.

The improved predictive accuracy is further illustrated in Fig. 5. The left subplot shows the magnitude of S_{12}, which exhibits high-frequency oscillations due to long electrical paths. The Szegö kernel models only the smooth envelope, but fails to capture the oscillatory behavior. The novel $K_{\text{composite}}$ kernel estimates precisely notch depths and spectral variations. The right subplot presents the magnitude of S_{23}, which is challenging to model given that its magnitude is smaller than −40 dB for most of the spectrum. Here, the predictions of the real-valued kernel in (3) and the novel proposed kernel (5) are compared. Note that the latter improves precision across the band by accurately tracking rapid amplitude varia-

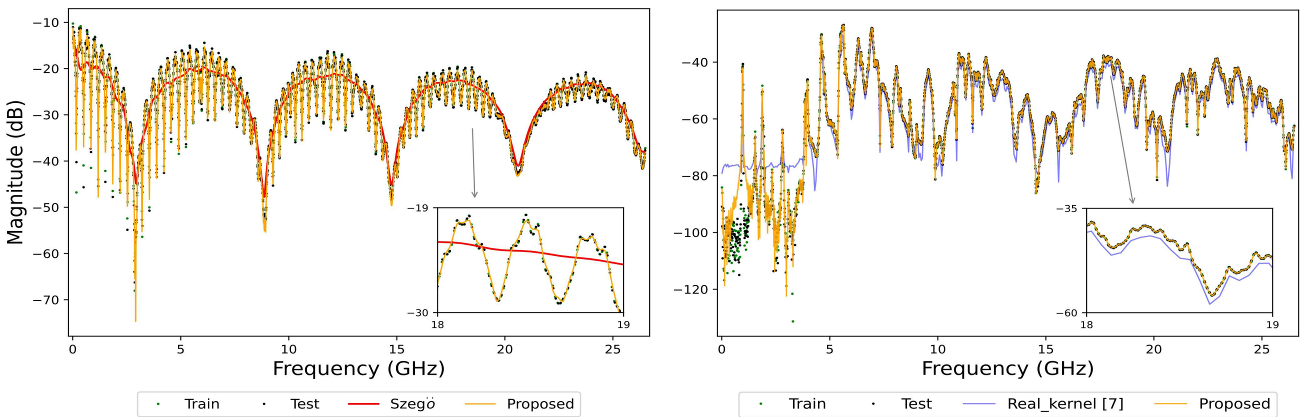

Fig. 5. Estimation of elements S_{12} (left) and S_{23} (right) via the novel proposed kernel (orange), the rational Szegö kernel (red), and the real-valued kernels in [7] (blue) with respect to measured training (green dots) and test samples (black dots). Insets show a detailed comparison for $f \in [18, 19]$ GHz, highlighting kernel accuracy differences.

tions and deep notches even when the magnitude of S_{23} is below –60 dB, due to the joint encoding of delay effects and rational structure. In the very low-frequency regime, however, the magnitude drops below –90 dB, where even the proposed kernel cannot fully reproduce the response. This is expected, since such extreme attenuation is close to the measurement noise floor and is difficult to capture with high accuracy.

IV. Conclusion

We introduced a novel composite GP model based on a delay-aware extension of the rational Szegö kernel, superimposing propagation delay effects on a global analytic structure for improved modeling of electrically long interconnect. This novel approach enforces physically consistent cross-correlations between real and imaginary parts of S-parameters and offers improved modeling accuracy compared to existing physics-informed kernels, as validated by suitable experimental results.

References

[1] S. Grivet-Talocia and B. Gustavsen, *Passive Macromodeling: Theory and Applications*. Hoboken, NJ, USA: Wiley, Nov. 2015.

[2] Y. Ye, T. Ullrick, W. Bogaerts, T. Dhaene, and D. Spina, "SPICE-compatible equivalent circuit models for accurate time-domain simulations of passive photonic integrated circuits," *J. Lightw. Technol.*, vol. 40, no. 24, pp. 7856–7868, Dec. 15, 2022.

[3] H. M. Torun, A. C. Durgun, K. Aygun, 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.

[4] 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 (SPI)*, Siegen, Germany, May 2022, pp. 1–4.

[5] 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.

[6] T. Ullrick, D. Deschrijver, W. Bogaerts, and T. Dhaene, "Modeling microwave S-parameters using frequency-scaled rational Gaussian process kernels," in *Proc. IEEE 33rd Conf. Electr. Perform. Electron. Packag. Syst. (EPEPS)*, Oct. 2024, pp. 1–3.

[7] F. Garbuglia, T. Reuschel, C. Schuster, D. Deschrijver, T. Dhaene, and D. Spina, "Modeling electrically long interconnects using physics-informed delayed Gaussian processes," *IEEE Trans. Electromagn. Compat.*, vol. 65, no. 6, pp. 1715–1723, Dec. 2023.

[8] J. Li and H. Wang, "Gaussian processes regression for uncertainty quantification: An introductory tutorial," 2025, *arXiv:2502.03090*.

[9] J. Bect, N. Georg, U. Römer, and S. Schöps, "Rational kernel-based interpolation for complex-valued frequency response functions," *SIAM J. Sci. Comput.*, vol. 46, no. 6, pp. A3727–A3755, Dec. 2024.

[10] P. Liu, *Bayesian Optimization: Theory and Practice Using Python*. Cham, Switzerland: Springer, Mar. 2023.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, Oct. 2012.

[12] 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.