

Letters

Generation of Stochastic Interconnect Responses via Gaussian Process Latent Variable Models

Simon De Ridder , Dirk Deschrijver, Paolo Manfredi , Tom Dhaene , and Dries Vande Ginste 

Abstract—We introduce a novel generative model for stochastic device responses using limited available data. This model is oblivious to any varying design parameters or their distribution and only requires a small set of “training” responses. Using this model, new responses are efficiently generated whose distribution closely matches that of the real data, e.g., for use in Monte-Carlo-like analyses. The modeling methodology consists of a vector fitting step, where device responses are represented by a rational model, followed by the optimization of a Gaussian process latent variable model. Passivity is guaranteed by *a posteriori* discarding of nonpassive responses. The novel model is shown to considerably outperform a previous generative model, as evidenced by comparing accuracies of distribution estimation for the case of differential-to-common mode conversion in two coupled microstrip lines.

Index Terms—Gaussian process latent variable model (GP-LVM), generative models, high-speed connectors and links, statistical link analysis, stochastic modeling.

I. INTRODUCTION

MANUFACTURING variability in modern electronics has recently sparked a broad interest in stochastic modeling techniques capable of capturing this uncertainty. A key part of most of such analyses consists of obtaining a large number of statistical samples or instances, which is time-consuming and costly. Some techniques that are more efficient than a Monte Carlo approach have been developed to allow collection of such instances in large number, such as generalized polynomial chaos (gPC)-based methods [1]–[6]. A limitation of gPC is the requirement to know the varying parameters and their distributions or a number of samples at specific points in the design space. When these parameters or distributions are difficult to determine, or when they are too numerous, these techniques experience difficulties, to the point of intractability.

In [7], a model was proposed that uses only a limited number of device responses to generate additional ones, regardless of the number or distribution of the varying design parameters. In this letter, a considerably improved version of this approach is put forward, which is more flexible in capturing nonlinear dependencies and correlations, by

Manuscript received March 2, 2018; revised April 9, 2018; accepted April 21, 2018. Date of publication May 15, 2018; date of current version January 20, 2019. This work was supported in part by the Research Foundation Flanders and in part by the Interuniversity Attraction Poles Program BESTCOM initiated by the Belgian Science Policy Office. (*Corresponding author: Simon De Ridder.*)

S. De Ridder, D. Deschrijver, T. Dhaene, and D. Vande Ginste are with the Internet Technology and Data Science Lab, Department of Information Technology, Ghent University/imec, 9000 Ghent, Belgium (e-mail: simon.deridder@ugent.be; dirk.deschrijver@ugent.be; tom.dhaene@ugent.be; dries.vandeginste@ugent.be).

P. Manfredi is with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy (e-mail: paolo.manfredi@polito.it).

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

Digital Object Identifier 10.1109/TEMC.2018.2830104

making use of a Gaussian process latent variable model (GP-LVM) [8]. First, a small number of S-parameters (or any other frequency-domain responses), dubbed the “training set,” are transformed into a pole-residue form by means of the vector fitting (VF) algorithm [9], [10]. Next, in contrast to [7], the distribution of the fitted residue matrices is modeled using a GP-LVM, which is itself a generative model. Thus, in the last step, new sets of residue matrices can be generated, which, in turn, are transformed to new response instances whose distribution closely matches that of the true underlying distribution, of which the training set is a limited subset. Stability and reciprocity are guaranteed due to the VF characterization. Other physical constraints such as passivity and causality are enforced by rejection. The proposed technique is applied to the case of a pair of coupled microstrip lines, prone to manufacturing variability. In particular, we investigate the effect of stochastic geometric imbalance on the differential-to-common mode conversion, which causes electromagnetic interference issues [11]. The novel generative modeling strategy is carefully compared to the technique presented in [7], clearly demonstrating its improved performance. The remainder of this letter is organized as follows. In Section II, the problem statement and the modeling approach are outlined. Validation using the representative application example of a pair of coupled microstrip lines is provided in Section III. Finally, conclusions are drawn in Section IV.

II. GENERATIVE MODELING FRAMEWORK

A. Goal

Due to manufacturing tolerances, measurement errors, or variations in material properties, the response of an electronic device is subject to stochastic variability. The goal of the newly proposed model is to generate new instances of device responses (typically S-parameters), based on a limited number of measured or simulated instances (the training set), that accurately represent the distribution of this training set. Moreover, the generated instances should obey the same physical constraints (such as passivity) among the different elements of the responses for a multiport structure.

The training set consists of a small number of distinct random instances of the device response. They can be obtained via Monte Carlo simulation, by varying the physical parameters, or by measuring a limited number of different fabricated instances of a given device. The more complex the distribution of instances is, the larger the size of the training set should be to obtain the same accuracy. A detailed discussion on this matter is outside the scope of this letter, but was treated in [7].

B. Vector Fitting

As a first step of the modeling process, the device responses in the training set (sampled at a number of frequencies) are fitted using the

well-known VF algorithm [9], [10]. For a training set of size K , the N_p -port S-parameter matrix $\bar{S}_k(s)$ of the k th instance is represented as a finite sum of partial fractions dependent on the complex frequency variable s

$$\bar{S}_k(s) \approx \sum_{i=1}^N \frac{\bar{R}_{k,i}}{s-a_i} + \bar{D}_k + \bar{E}_k s, \quad k = 1, \dots, K. \quad (1)$$

Here, a_i and $\bar{R}_{k,i}$ are called poles and residue matrices, respectively. These are real or pairwise complex conjugate valued. If the device is stable, all poles must lie in the left half of the complex plane. \bar{D}_k and \bar{E}_k are optional terms that define the asymptotic behavior for frequencies $s \rightarrow 0$ and $s \rightarrow \infty$, respectively. All K S-parameter instances are first fitted using a common pole set $\{a_i\}_{i=1}^N$. Second, each instance k is fitted separately to allow both poles and residues to vary. The problem is now reduced to building a generative model for these poles and residue matrices.

C. Gaussian Process Latent Variable Model

A Gaussian process (GP) is an infinite collection of random variables, any finite subset of which is jointly Gaussian distributed [12], [13]. This means that for a set of K M -dimensional inputs $\bar{x}^1, \bar{x}^2, \dots, \bar{x}^K$, the GP can be represented by a set of K P -dimensional random variables $\bar{y}^1, \bar{y}^2, \dots, \bar{y}^K$, which, for each of the P dimensions, are jointly Gaussian distributed

$$\bar{y}_p \sim \mathcal{N}(\bar{0}, \bar{\Sigma}), \quad p = 1, \dots, P. \quad (2)$$

The dimension of each \bar{y}^k is given by $P = (N_p^2 + 1) \times N$, being all elements of the residue matrices and poles for that instance. The model is thus completed by specifying a covariance matrix function or kernel

$$\sigma_{ij} = \mathcal{K}(\bar{x}^i, \bar{x}^j). \quad (3)$$

A popular choice for \mathcal{K} , also adopted here, is the automatic relevance detection kernel. This kernel features length scales l_m for each of the M input dimensions as hyperparameters. A long length scale indicates that a certain dimension is not very relevant to the change in the function, hence its name. It takes the following form:

$$\mathcal{K}(\bar{x}^i, \bar{x}^j) = \sigma_y^2 \exp\left(-\frac{1}{2} \sum_{m=1}^M \frac{(x_m^i - x_m^j)^2}{l_m^2}\right). \quad (4)$$

Usually, GPs are employed to interpolate or predict new function values, given data in the form of observed (\bar{x}, \bar{y}) pairs. The GP-LVM [8], however, leverages the GP infrastructure to infer the distribution of the input variables from a given number of observed output points $\{\bar{y}^k\}_{k=1}^K$. The hyperparameters σ_y and $\{l_m\}_{m=1}^M$ are optimized using a variational lower bound to the marginal likelihood. A GP-LVM can thus be used to reduce the dimensionality of a distribution, while modeling it in a nonparametric and nonlinear way. For the model proposed here, the \bar{y}^k are constructed by vectorizing all poles and residue matrices for each training instance. The inferred \bar{x}^k are not readily interpretable, but represent a training instance in a lower number of dimensions ($M \ll P$), while their prior distribution is an independent standard Gaussian.

D. Generative Statistical Model

New instances can now easily be generated as follows. A sample \bar{x}^* is drawn from the GP-LVM's input prior distribution, a standard Gaussian distribution. The GP-LVM then provides an output \bar{y}^* for this sample, which is reshaped into a set of complex poles and residue matrices. This constitutes a VF model for a new instance.

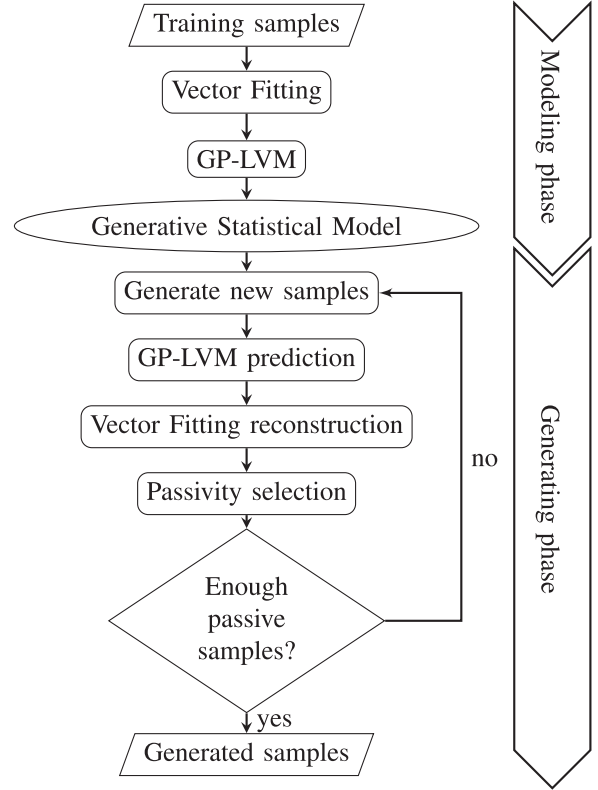


Fig. 1. Flowchart of the proposed modeling and analysis framework.

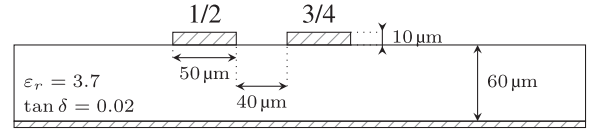


Fig. 2. Cross section of the microstrip lines with indication of mean, nominal values. The lines have a length of 10 cm. Above each conductor pair, the corresponding port numbers are shown (near-end/far-end).

In order to ensure passivity of the generated instances without creating a bias, those instances that violate passivity are discarded, and new instances are generated in their stead.

Owing to the sparse inference used (and required) by the GP-LVM, the asymptotic complexity of inference for the GP-LVM model is $\mathcal{O}(PN_i^2) = \mathcal{O}((N_p^2 + 1)NN_i^2)$, with N_i being the number of inducing variables, typically a small fraction of P . That of the VF step is $\mathcal{O}(KN_p^2(N_s + N)N^2)$, with N_s being the number of frequency points. Generating new instances is done in $\mathcal{O}(PK^3N^*)$, where N^* is the number of required instances.

III. APPLICATION EXAMPLE AND NUMERICAL RESULTS

In this section, the proposed methodology is demonstrated on a representative example of two coupled microstrip lines.

A. Validation

The differential-to-common mode S-parameters [14] of two coupled microstrip lines, depicted in Fig. 2, are studied. One thousand S-parameter instances were simulated by varying five parameters: the relative permittivity, the line separation, the substrate thickness, and both of the linewidths. These parameters were varied according to

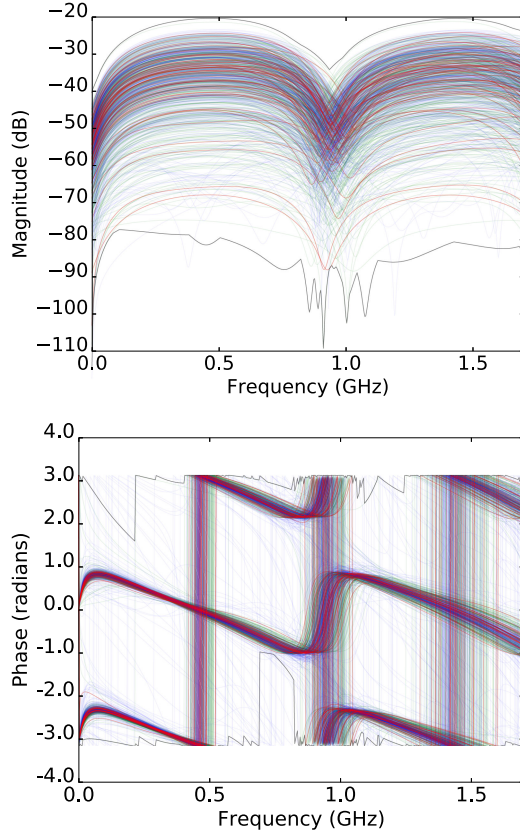


Fig. 3. Magnitude and phase of the mode conversion parameter S_{C1D1} for the coupled microstrip lines. The 50 training instances are shown in red on top, the 1000 generated instances are plotted in blue, and the 950 validation instances are shown in the background in green. The black lines represent the minimum and maximum of all simulated samples.

independent Gaussian distributions, each with a standard deviation that is 10% of their mean nominal value that is indicated in Fig. 2. Only 50 of these instances were used to train the proposed model. The model was then used to generate another set of 1000 instances, which was, in turn, compared to the 950 unused simulated instances (validation set).

Figs. 3 and 4 show the validation, generated, and training instances for some of the S-parameters. A good correspondence between the validation instances and the generated instances is apparent. A more detailed assessment is given in Fig. 5, where the cumulative distribution function (CDF) of S_{C1D1} at a frequency of 1.5 GHz for each set (training, validation, and generated) is displayed, showing good agreement.

On a Dell Optiplex 760 PC with an Intel Core 2 Quad CPU (at 3 GHz) and 7.7-GiB RAM, building the model took 254 s, while generating new samples took 22 s without rejection of nonpassive samples, and 27 s with rejection.

B. Comparison to the Kernel-Density-Estimation-Based Generative Model

The model introduced in [7] uses a principal component analysis (PCA) and a kernel density estimation (KDE) to model nonlinear correlations in a nonparametric way. Because PCA reduces dimensionality by assuming only linear correlation, and because each kernel in the KDE shares the same Gaussian shape, the flexibility of this model is limited. The GP-LVM models the high-dimensional data directly and only assumes a covariance measure between latent input variables,

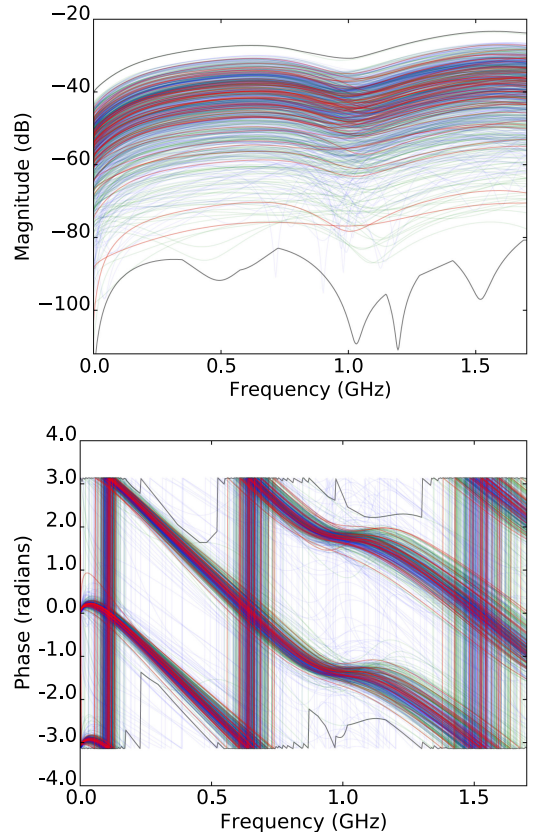


Fig. 4. Magnitude and phase of the mode conversion parameter S_{C2D1} for the coupled microstrip lines. Colors are as in Fig. 3.

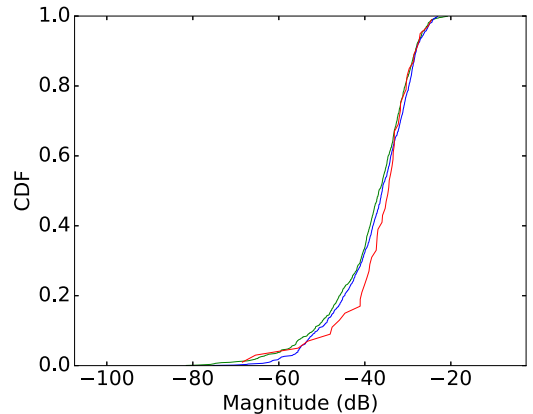


Fig. 5. CDF of the validation (green), generated (blue), and training (red) sets for the magnitude of the S_{C1D1} -parameter of the coupled microstrip example at 1.5 GHz.

which can be optimized. This makes the GP-LVM more flexible to deal with nonlinear data, as is shown by the following accuracy comparison.

To compare both models, we introduce a goodness-of-fit measure for two empirical distributions, called the two-sample Cramér–von Mises (CM) test statistic [15]. This is the integral over the squared difference of two empirical CDFs $F(x)$ and $G(x)$ of sample sizes m and n , respectively

$$T_{CM}(F, G) = \frac{mn}{m+n} \int_{-\infty}^{+\infty} [F(x) - G(x)]^2 dH(x) \quad (5)$$

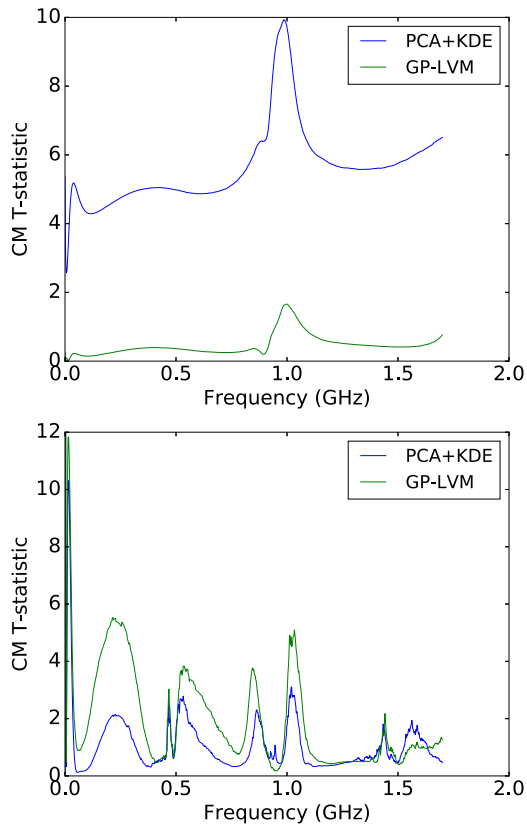


Fig. 6. CM statistic for the PCA + KDE model outlined in [7] and for the GP-LVM model proposed here for the magnitude (top) and phase (bottom) of S_{C1D1} .

where $H(x)$ is the empirical CDF of both samples together. With this measure, we can compare the distribution of the $n = 1000$ generated instances to that of the $m = 950$ withheld validation instances at a single frequency point. A comparison of the CM statistic across frequencies for different generative models can indicate which model produces a more accurate approximation of the true distribution.

Fig. 6 shows the CM statistic for both the model outlined in [7] (blue) and the GP-LVM model introduced here (green). It is apparent that the GP-LVM variant outperforms the model in [7] for the magnitude in the entire frequency range, while performing in a comparable manner for the phase.

IV. CONCLUSION

This letter introduces a novel generative model capable of generating a large set of device responses from a small training set of simulated or measured responses. It starts by representing the responses as a rational model using the VF algorithm. The poles and residues of these rational expansions are modeled by means of a GP-LVM. New instances are generated by the GP-LVM and transformed back into device responses by the VF expansion. Rejecting nonpassive instances ensures unbiased and physically consistent responses.

The modeling power of the generative model is demonstrated by applying it for a four-port coupled microstrip line, focusing on mode conversion induced by geometric unbalance. The distributions of simulated and generated S-parameters are found to be in very good agreement. A quantitative comparison shows that the GP-LVM-based model significantly outperforms the KDE-based model previously proposed in [7] to model the magnitude of S-parameters.

REFERENCES

- [1] P. Manfredi, D. Vande Ginste, D. De Zutter, and F. G. Canavero, "Uncertainty assessment of lossy and dispersive lines in SPICE-type environments," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 3, no. 7, pp. 1252–1258, Jul. 2013.
- [2] Z. Zhang, T. A. El-Moselhy, I. M. Elfadel, and L. Daniel, "Stochastic testing method for transistor-level uncertainty quantification based on generalized polynomial chaos," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 32, no. 10, pp. 1533–1545, Oct. 2013.
- [3] M. R. Rufuie, E. Gad, M. Nakhla, and R. Achar, "Generalized Hermite polynomial chaos for variability analysis of macromodels embedded in nonlinear circuits," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 4, no. 4, pp. 673–684, Apr. 2014.
- [4] P. Manfredi, D. Vande Ginste, D. De Zutter, and F. G. Canavero, "Stochastic modeling of nonlinear circuits via SPICE-compatible spectral equivalents," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 7, pp. 2057–2065, Jul. 2014.
- [5] P. Manfredi, D. Vande Ginste, D. De Zutter, and F. G. Canavero, "Generalized decoupled polynomial chaos for nonlinear circuits with many random parameters," *IEEE Microw. Wireless Compon. Lett.*, vol. 25, no. 8, pp. 505–507, Aug. 2015.
- [6] M. Ahadi and S. Roy, "Sparse linear regression (SPLINER) approach for efficient multidimensional uncertainty quantification of high-speed circuits," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 35, no. 10, pp. 1640–1652, Oct. 2016.
- [7] S. De Ridder *et al.*, "A generative modeling framework for statistical link analysis based on sparse data," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 8, no. 1, pp. 21–31, Jan. 2018.
- [8] M. Titsias and N. D. Lawrence, "Bayesian Gaussian process latent variable model," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 844–851. [Online]. Available: <http://www.jmlr.org/proceedings/papers/v9/titsias10a/titsias10a.pdf>
- [9] 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.
- [10] D. Deschrijver, M. Mrozowski, T. Dhaene, and D. De Zutter, "Macro-modeling 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.
- [11] F. Grassi, Y. Yang, X. Wu, G. Spadacini, and S. A. Pignari, "On mode conversion in geometrically unbalanced differential lines and its analogy with crosstalk," *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 2, pp. 283–291, Apr. 2015.
- [12] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006. [Online]. Available: <http://www.gaussianprocess.org/gpml>
- [13] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [14] C. Gazda, D. V. Ginste, H. Rogier, R. B. Wu, and D. D. Zutter, "A wideband common-mode suppression filter for bend discontinuities in differential signaling using tightly coupled microstrips," *IEEE Trans. Adv. Packag.*, vol. 33, no. 4, pp. 969–978, Nov. 2010.
- [15] T. W. Anderson, "On the distribution of the two-sample Cramer-von Mises criterion," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1148–1159, Sep. 1962.