Statistical Analysis of the Efficiency of an Integrated Voltage Regulator by means of a Machine Learning Model Coupled with Kriging Regression

Statistical Analysis of the Efficiency of an Integrated Voltage Regulator by means of a Machine Learning Model Coupled with Kriging Regression

R. Trinchero*, M. Larbi†, M. Swaminathan† and F. G. Canavero*
* EMC Group, Department of Electronics and Telecommunications, Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
E-mail: riccardo.trinchero@polito.it
† School of Electrical and Computer Engineering, Georgia Institute of Technology,
Atlanta, GA 30332, USA

***Abstract*—This paper presents a preliminary version of a probabilistic model for the uncertainty quantification of complex electronic systems resulting from the combination of the least-squares support vector machine (LS-SVM) and the Gaussian process (GP) regression. The proposed model, trained with a limited set of training pairs provided by a set of full-wave expensive simulations, is adopted for the prediction of the efficiency of an integrated voltage regulator (IVR) with 8 uniformly distributed random parameters. The accuracy and the feasibility of the proposed model have been investigated by comparing the model predictions and its confidence intervals with the results of a Monte Carlo (MC) full-wave simulation of the device.**

***Index Terms*—IVR, Machine Learning, least-squares support vector machine, Gaussian process, error bounds.**

I. Introduction

In the last decades, several techniques have been presented for the efficient analysis of the responses of complex electronics systems affected by uncertain stochastic parameters [1]–[5]. The system uncertainties are usually related to manufacturing processes, tolerances and uncertain device characteristics, which unavoidably affect the system responses.

Polynomial chaos (PC) expansion [1], [2] and recently support vector machine (SVM) [3] and least-squares support vector machine (LS-SVM) regressions [4], [5] can be considered as reference techniques in the field of uncertainty quantification, since they allow building accurate and fast-to-evaluate surrogate model of the system output as a function of its stochastic parameters. The surrogate model is built from a limited set of training samples generated by the full computational model, thus providing an effective and viable alternative to the traditional Monte Carlo (MC) simulation [6].

However, all the above mentioned techniques provide a *deterministic* model. In the sense that the surrogate model can be seen as a function providing for any configuration of the input parameters of the system a prediction of the system output without any information on the model uncertainty (i.e., the confidence intervals (CI)). In fact, the degree of confidence of the model prediction is usually known only for the training samples, but it is completely unknown when the model is evaluated for a generic point in the parameters space.

The Gaussian process (GP) regression [7]–[10], also known as Kriging model, can be seen as a viable solution to the above problem. Specifically, the GP regression can be used to enrich a generic deterministic model by adding the CI to its predictions, for any configuration of its input parameters, thus providing the user with a *probabilistic* model. It is important to remark that the proposed procedure works under the assumption that the error of the deterministic model on the training sample is different from zero. This means that the deterministic model can be estimated from a regression, but not from an interpolation of the training data.

In this work, we are presenting a preliminary probabilistic model based on the combination of two techniques: the LS-SVM and the GP regressions. The proposed model is applied in order to predict the efficiency of an integrated voltage regulator (IVR) as a function of 8 uniformly distributed uncertain parameters. The accuracy of the proposed probabilistic model is investigated by comparing the model prediction and its CI with the resulting provided by a MC simulation of the above electronic system in Ansys HFSS.

II. GP Regression

Let us considering the problem of approximating a given set of L train pairs $\{(\mathbf{x}_i, y_i)\}_{i=1}^{L}$ provided by a full computational model i.e., $y_i = \mathcal{M}(\mathbf{x}_i)$ as a function of its stochastic parameter $\mathbf{x}_i \in \mathcal{P}$ with $\mathcal{P} \subset \mathbb{R}^d$, via the following GP regression [7]–[10]:

$$\mathcal{M}(\mathbf{x}) \approx \mu(\mathbf{x}) + \varepsilon(\mathbf{x}) \tag{1}$$

where $\mu(\mathbf{x})$ is a deterministic function representing the mean, also called *trend*, of the GP and $\varepsilon(\mathbf{x}) \sim \mathcal{GP}(0, k(\mathbf{x}, \mathbf{x}'))$ is a Gaussian process GP with zero mean and covariance $k(\mathbf{x}, \mathbf{x}')$ approximating the deviation of the deterministic model from the actual function $\mathcal{M}$, for any configuration of the parameters $\mathbf{x}$. A GP can be consider as an extension of the concept of multivariate Gaussian distributions to infinite dimensionality which returns for any arbitrary value of $\mathbf{x}_*$, a prediction of the mean value $\mu_{\mathbf{x}_*}$ and the variance $\sigma^2_{\mathbf{x}_*}$ of a normal distribution.

The covariance function $k(\mathbf{x}, \mathbf{x}')$ of the GP specifies the correlation between pairs of random variables under the as-

sumption that $\mathcal{M}$ is *smooth*, this means that points with similar predictor values are aspected to have close response values. The covariance function can be defined by several kernels, however, hereafter in this paper we will considered the ARD Matern 5/2 kernel function which writes:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \sqrt{5}r + \frac{5}{3}r^2\right) \exp(-\sqrt{5}r), \qquad (2)$$

with,

$$r = \sqrt{\sum_{m=1}^{d} \frac{(x_m - x'_m)^2}{\sigma_m^2}} \qquad (3)$$

where σ_f and σ_m for $m = 1, \ldots, d$ are the hyper-parameters of the kernel estimated via an optimization procedure e.g., by minimizing the negative log marginal likelihood of the GP [7].

Different from the standard GP regression, hereafter in this work, we are assuming that the trend $\mu(\mathbf{x})$ of the GP in (1) is a known function previously estimated via a generic regression technique. In the particular case in which the GP regression uses a *fixed* mean function, the hyper-parameters of the covariance function in (2) can be easily estimated by considering as training samples, the distance between the mean function and the actual responses of the full computational model as follows:

$$\tilde{y}_i = \mathcal{M}(\mathbf{x}_i) - \mu(\mathbf{x}_i), \;\; \text{for } i = 1, \ldots, L. \qquad (4)$$

It is important to remark that the proposed approach does not work with the interpolation methods for which the error on the training sample is exactly zero.

Once the optimal hyper-parameters of the covariance function in (2) have been estimated, as an example via the GP regression tool of MATLAB, the prediction of the probabilistic model at a new point $\mathbf{x}_* \in \mathcal{P}$, knowing the training pairs $\{(\mathbf{x}_i, \tilde{y}_i)\}_{i=1}^{L}$, can be approximated by a Gaussian random variable $M(\mathbf{x}_*) \sim \mathcal{N}(\mu_{\mathbf{x}_*}, \sigma^2_{\mathbf{x}_*})$ with mean $\mu_{\mathbf{x}_*}$ and variance $\sigma^2_{\mathbf{x}_*}$:

$$\mu_{\mathbf{x}_*} = \mu(\mathbf{x}_*) + \mathbf{k}_* \mathbf{K}^{-1} \tilde{\mathbf{y}} \qquad (5a)$$

$$\sigma^2_{\mathbf{x}_*} = k_{**} - \mathbf{k}_* \mathbf{K}^{-1} \mathbf{k}_*^T \qquad (5b)$$

where $\mu_{\mathbf{x}_*}$ and $\sigma^2_{\mathbf{x}_*}$ are the estimation of the mean and the variance of the normal distribution associated to the GP in (1), $\tilde{\mathbf{y}} = [\tilde{y}_1, \ldots, \tilde{y}_L]^T$, $\mathbf{K} \in \mathbb{R}^{L \times L}$ is the correlation matrix given by $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{k}_* = [k(\mathbf{x}_*, \mathbf{x}_1), \ldots, k(\mathbf{x}_*, \mathbf{x}_L)] \in \mathbb{R}^{1 \times L}$ and $k_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$.

The prediction mean $\mu_{\mathbf{x}_*}$ is used instead of deterministic model $\mu(\mathbf{x})$, whereas the variance $\sigma^2_{\mathbf{x}_*}$ gives a local error indicator about its precision. The resulting probabilistic interpretation of the model realization allows estimating the confidence bounds on the model prediction for a given confidence interval of $100(1-\alpha)\%$ as follows:

$$\left(\mu_{\mathbf{x}_*} - z_{1-\frac{\alpha}{2}} \sqrt{\sigma^2_{\mathbf{x}_*}}\right) \leq \mathcal{M}(\mathbf{x}_*) \leq \left(\mu_{\mathbf{x}_*} + z_{1-\frac{\alpha}{2}} \sqrt{\sigma^2_{\mathbf{x}_*}}\right) \qquad (6)$$

where z denotes the inverse of the Gaussian cumulative distribution function evaluated at $1 - \frac{\alpha}{2}$.

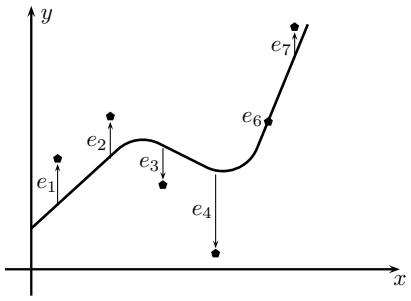

Fig. 1. Graphical interpretation of the LS-SVM regression.

III. Deterministic model via LS-SVM regression

The LS-SVM regression allows building accurate and compact surrogate models of the response of a generic high-dimensional nonlinear function $\mathcal{M}$ starting from a limited number of training samples [4]. Therefore, it can be seen as a good candidate to approximate the trend $\mu(\mathbf{x})$ of the GP in (1). Given a set of training samples $\{(\mathbf{x}_i, y_i)\}_{i=1}^{L}$, provided by a full computational model $\mathcal{M}$ as a function of its parameter $\mathbf{x}_i \in \mathcal{P}$ with $\mathcal{P} \subset \mathbb{R}^d$, a generic nonlinear LS-SVM regression in the dual space can be written as:

$$M_{LSSVM}(\mathbf{x}) = \sum_{i=1}^{L} \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b \qquad (7)$$

where $\alpha_i \in \mathbb{R}$ are scalar coefficients, $K(\cdot, \cdot) : \mathbb{R}^d \to \mathbb{R}$ is the *kernel function* and $b \in \mathbb{R}$ is the bias term.

The goal of the LS-SVM regression is to minimize the error e_i computed between the model prediction and the training samples in the L^2 norm (see Fig. 1 for additional detail). The above constraint, along with the one on the model flatness, allows estimating the parameters α_i and b via the solution of the following linear system:

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \boldsymbol{\Omega} + \mathbf{I}/\gamma \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \qquad (8)$$

where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_L]^T$, $\mathbf{y} = [y_1, \ldots, y_L]^T$, $\mathbf{1}^T = [1, \ldots, 1] \in \mathbb{R}^{1 \times L}$, $\boldsymbol{I} \in \mathbb{R}^{L \times L}$ is the identity matrix and $\boldsymbol{\Omega} \in \mathbb{R}^{L \times L}$ is the kernel matrix for which the element $\Omega_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ for any $i, j = 1, \ldots, L$.

The most common kernels used in both the SVM and the LS-SVM regression are listed below:

- linear: $K(\mathbf{x}_i, \mathbf{x}) = \mathbf{x}_i^T \mathbf{x}$;
- polynomial of order q: $K(\mathbf{x}_i, \mathbf{x}) = (1 + \mathbf{x}_i^T \mathbf{x})^q$;
- Gaussian radial basis function (RBF): $K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma^2\right)$.

The LS-SVM regression is already implemented in MATLAB within LS-SVMLab Toolbox version 1.8 [11].

IV. Application Example

The combination of the GP and the LS-SVM regression presented in Sec. II and III has been used to develop the probabilistic model in (1) with the aim of estimating the efficiency of the IVR converter shown in Fig. 2, as a function of

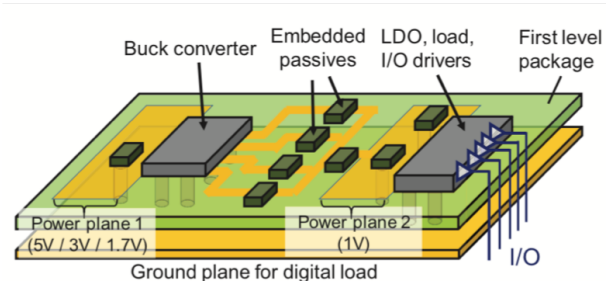

Fig. 2. Illustration of the two-chip SiP IVR architecture considered in Sec. IV.

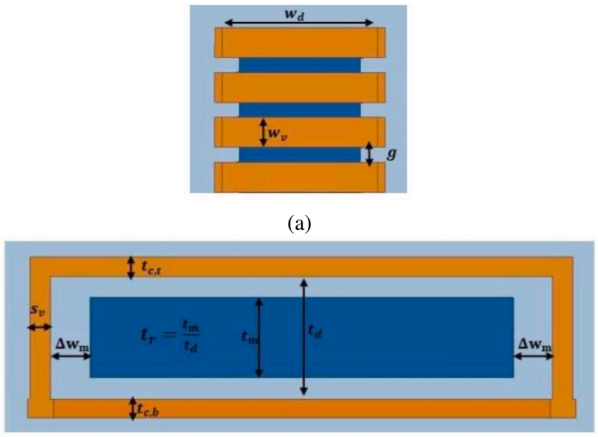

Fig. 3. Top view (panel (a)) and side view (panel (b)) of the geometrical parameters of a solenoidal inductor with magnetic core [9].

8 uncertain geometrical parameters of its embedded inductor. The considered architecture consists of a system-in-package (SiP) solution including two chips (buck converter and low-dropout (LDO)/load) with an integrated inductor on an organic package [12]. The inductor is a solenoid with Nickel-Zinc (NiZn) ferrite magnetic core as shown in Fig. 3(a) and (b).

The effect of the tolerances of the inductor geometrical parameters on the IVR efficiency has been characterized via a MC simulation based on the full-wave solver of Ansys HFSS. Specifically, 8 geometrical parameters have been considered as uniform random variables, as shown in Table I, whereas the number of windings and the magnetic core thickness ratio of the inductor have been fixed to $N_w = 6$ and $t_m = 0.9$ [2].

The probabilistic model in (1) has been trained from $L = 200$ samples drawn from a latin hypercube sampling scheme and resulting from a set of full-wave simulations. The deterministic model $\mathcal{M}_{LSSVM}$ in (7) has been estimated from the available training samples via the LS-SVM regression with RBF kernel. Then, the obatained model has been used to train a GP regression with fixed mean function (i.e., $\mu(\mathbf{x}) = \mathcal{M}_{LSSVM}(\mathbf{x})$) by using as covariance the ARD Matern 5/2 kernel function in (2). The hyper-parameters of the covariance function have been estimated based on the training pairs $\{(\mathbf{x}_i, \tilde{y}_i)\}$ defined in (4).

First of all, the accuracy of the deterministic model has been

TABLE I
UNCERTAIN GEOMETRICAL PARAMETERS OF THE SOLENOIDAL INDUCTOR IN FIG- 3.

Uniform random variables		Unit	$\mathcal{U}[Min; Max]$
Gap between windings	g	mil	$\mathcal{U}[4; 6]$
Size of via	s_v	μm	$\mathcal{U}[80; 120]$
Copper Trace Width	w_v	mil	$\mathcal{U}[9; 11]$
Copper Thickness Bottom	$t_{c,b}$	μm	$\mathcal{U}[64; 96]$
Copper Thickness Top	$t_{c,t}$	μm	$\mathcal{U}[64; 96]$
Dielectric Thickness	t_d	μm	$\mathcal{U}[180; 220]$
Dielectric Width	w_d	mil	$\mathcal{U}[59.4; 60.6]$
Magnetic Core Width offset	Δw_m	mil	$\mathcal{U}[9; 11]$

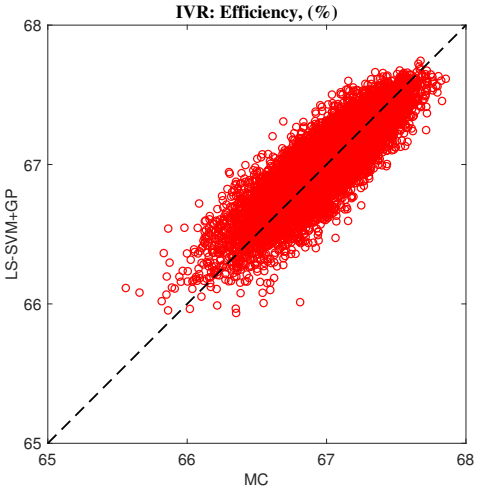

Fig. 4. Scatter plot (10000 samples) comparing the IVR efficiency calculated via a MC simulation in Ansys with the prediction of the probabilistic model in (1).

investigated via the scatter plot in Fig. 4. The plot compares the IVR efficiency estimated by the mean of the GP via (5a), with the result calculated through a MC simulation with 10000 samples. The plot highlights the good accuracy of the model as the samples are very close to the dashed line, which represents the perfect agreement between the model and the reference samples.

Then, the full probabilistic model is validated in Fig. 5 by comparing the efficiency of the IVR obtained from the Ansys simulations for 15 realizations (black crosses) randomly selected among the results of the MC simulation, with the mean values (red dots) and the 95% CI (horizontal blue bars) predicted by the proposed probabilistic models which combine the GP and the LS-SVM regression with RBF kernel. The results clearly highlight the capability of the model of accurately predicting the actual efficiency of the IVR, since most of the simulation results lay between the 95% CI given by the proposed model.

As a further validation, in Fig. 6 (top panel) the probability

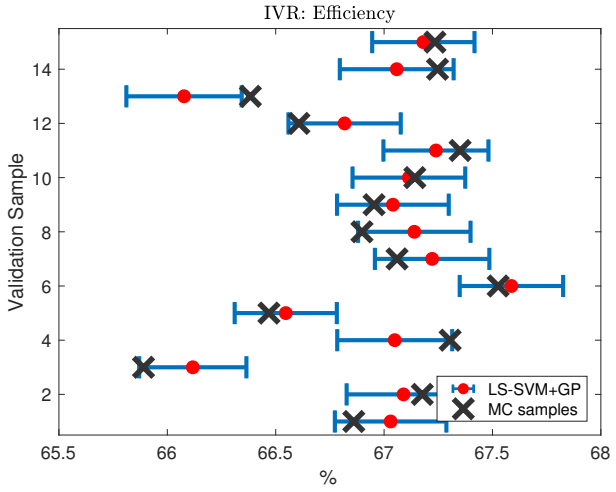

Fig. 5. Comparison between the results of 15 Ansys simulations (black crosses) randomly chosen among the 10000 MC samples with the corresponding main value (red dots) and 95% CI (horizontal blue bars) estimated via the proposed probabilistic model.

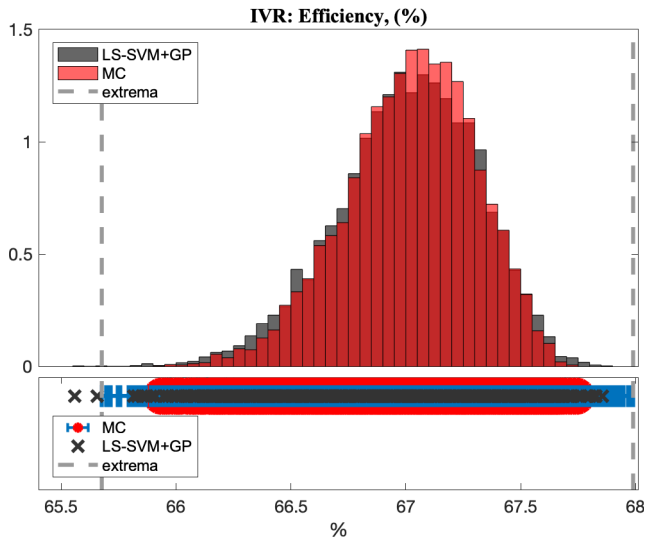

Fig. 6. Top panel: PDFs of the IVR efficiency obtained from the mean values of the probabilist model (red bins) compared with the histogram of 10000 MC samples (black bins). Bottom panel: comparison among the bound predicted by the MC simulation (black crosses), the mean values of the GP (red dots) and the mean values with their CI (gray dashed lines).

density functions (PDFs) provided by the mean values of the GP via (5a) are compared with the histogram of 10000 MC samples. We see that the main variability of the IVR efficiency is well captured by the model. Also, Fig. 6 (bottom panel) compares the bounds provided by a MC simulation with 10000 samples (black crosses) with the one predicted by both the mean value (red dots) and the CI (dashed gray lines) of the probabilistic model. The results clearly highlight the capability of the proposed probabilistic model to provide more reliable bounds of the system responses, thus proving a better estimation of the minimum and maximum IVR efficiency for the considered realizations. In terms of computational cost, the computational time required to generate the $L = 200$ training pairs is 3 h 27 min. Once the probabilistic model has been built, it allows predicting the IVR efficiency for 10000 samples in less than 1 s, while 10000 MC simulations required about 7 days.

V. Conclusions

This paper presented a probabilistic model for the prediction of the efficiency of an IVR as a function of 8 uncertain parameters related to the geometry of its embedded inductor. The proposed probabilistic model has been built from a limited set of training samples provided by a set of deterministic simulation in Ansys HFSS, based on a latin hypercube sampling scheme. The probabilistic model combines the LS-SVM and the GP regression. It allows estimating the IVR efficiency for any configuration of the inductor geometrical parameters, also providing an estimation of the model uncertainty, such as the CI of the model prediction. The accuracy of the proposed probabilistic model has been proven both from the deterministic and the statistical perspective by comparing its prediction with the results of a MC simulations.

VI. Acknowledgment

The Authors are grateful to Mr. Hakki M. Torun, Ph.D student at School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA , for providing the IVR simulation data.

References

[1] P. Manfredi, *at al.*, "Stochastic transmission line analysis via polynomial chaos methods: an overview," *IEEE Electromagn. Compat. Mag.*, vol. 6, no. 3, pp. 77–84, Third Quarter 2017.

[2] M. Larbi, *et al.*, "Uncertainty quantification of SiP based integrated voltage regulator," in *Proc. IEEE 22nd Workshop on Signal and Power Integrity (SPI)*, Brest, 2018, pp. 1-4.

[3] R. Trinchero, *at al.*, "Machine Learning for the Performance Assessment of High-Speed Links", *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 6, pp. 1627–1634, Dec. 2018.

[4] R. Trinchero, *at al.*, "Machine Learning and Uncertainty Quantification for Surrogate Models of Integrated Devices With a Large Number of Parameters," in *IEEE Access*, vol. 7, pp. 4056-4066, 2019.

[5] J. A. K. Suykens, *at al.*, "Least Squares Support Vector Machines", World Scientific Pub Co Inc, 2002.

[6] R. Spence and R. S. Soin, *Tolerance Design of Electronic Circuits.* London: Imperial College Press, 1997.

[7] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press. Cambridge, Massachusetts, 2006.

[8] F. Perez-Cruz, *at al.*, "Gaussian Processes for Nonlinear Signal Processing: An Overview of Recent Advances," *IEEE Signal Processing Magazine*, vol. 30, no. 4, pp. 40–50, July 2013.

[9] H. M. Torun, M. Swaminathan, A. Kavungal Davis and M. L. F. Bellaredj, "A Global Bayesian Optimization Algorithm and Its Application to Integrated System Design," in *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 4, pp. 792–802, April 2018.

[10] R. Schobi, B. Sudret and J. Wiart, "Polynomial-chaos-based Kriging," *Int. J. Uncertainty Quantification*, vol. 5, pp. 171–193, 2015.

[11] LS-SVMlab, version 1.8; Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven: Leuven, Belgium, 2011. Available online: http://www.esat.kuleuven.be/sista/lssvmlab/.

[12] S. Mueller, et al., "Design of high efficiency integrated voltage regulators with embedded magnetic core inductors," in *Proc. IEEE 66th Electronic Components and Technology Conference*, May 2016, pp. 566–573.