Machine-Learning-Based Constrained Optimization of a Test Coupon Launch Using Inverse Modeling

Andrew Page, Xu Chen
Department of Electrical and Computer Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61820, USA
andrew.page@ieee.org, xuchen1@illinois.edu

***Abstract*—This paper demonstrates the forward modeling and inverse design of a test coupon launch structure used in the board measurement practice known as the delta-L method. An inverse model is trained to synthesize a launch design to exhibit a desired electrical performance and to be physically realizable. A forward model is constructed and used to evaluate the electrical performance of the designs synthesized by the inverse model during training. The training of this inverse model is treated as a convex optimization with constraints on the synthesized designs. These constraints inspire a novel implementation of constraint loss by a pair of everywhere-differentiable barrier functions. The finished inverse model is applied to a swift multi-criteria design optimization and the forward model is used to perform uncertainty analysis about the synthesized design. Considerations for further applications and improvement of the procedure are discussed.**

***Index Terms*—neural network, forward/inverse model, delta-L method, convex optimization, barrier function**

I. Introduction

A test coupon is a simple interconnect built on the margins of a printed circuit board (PCB) to characterize the capability of the board while avoiding intensive full-board measurement. Knowledge of the embedded channel behavior allows material and cross-section characterization to be unaffected by any other structures including pads and vias. A measurement of a test coupon interconnect will include such launch structures, obscuring the desired embedded channel characteristics.

Techniques exist to calibrate out the effects of launch structures, an example being automatic fixture removal. Such a technique requires proprietary software and the results can be sensitive to measurement error. The delta-L method is a simple alternative that requires two measurements and no proprietary software [1]. These measurements have been automated for large batches and an example launch has been optimized in the literature [2], [3]. Delta-L provides less detail than a full calibration but is often sufficient for estimating loss.

The delta-L method relies on low discontinuity presented by the via-and-pad launch [3]. A dogbone via pair, a commonly used structure, will be optimized for minimal discontinuity at four frequencies chosen to mimic the existing literature [3]: 10, 20, 30, and 40GHz. This will be done by way of a tandem forward- and inverse-model method. A forward model is a machine learning model trained to predict the performance of a design, mimicking and replacing electromagnetic simulation. An inverse model is one that synthesizes a design to meet a prescribed electrical performance. Training an inverse model is significantly more challenging as the synthesized designs must obey constraints imposed by geometric limitations and signal integrity (SI) rules.

979-8-3503-1798-5/23/$31.00 ©2023 IEEE

Section II will discuss the training of the forward model along with the developments necessary for the inverse model to synthesize within design constraints. Section III will demonstrate the use of these models to optimize a dogbone via pair over a 7-dimensional design space and perform uncertainty analysis about an optimized design. Concluding remarks and potential improvements are offered in section IV.

II. Training the Forward and Inverse Models

A. Forward model — predicting design performance

A launch structure can be electrically characterized with full-wave simulation, done here with Ansys HFSS. The upcoming inverse model training will require many such evaluations, rendering HFSS prohibitively slow. A forward model, implemented as a two-layer fully-connected neural network, will serve as a replacement. This model will map seven design parameters $\vec{\xi} = (r_{via}, r_{pad}, r_{ap}, p_{via}, p_{gnd}, l_{stub}, \theta)^T$ into the board-side differential-mode reflection coefficient magnitude

$$\mathrm{FM}(\vec{\xi}_i) = |\bar{S}_{dd11}(\vec{\xi}_i, f_j)| =: |\bar{S}|_i \in \mathbb{R}^4, \tag{1}$$

where $f_j = \{10, 20, 30, 40\}$GHz. A design includes the via, pad and anti-pad radii, ground and signal pitches, stub length and launch angle. The model is trained with HFSS simulations of vias sampled over the design space using latin hypercube sampling to minimize prediction loss $\mathcal{L}$, a mean-square error between training data $\{|S|_i\}$ and network predictions $\{|\bar{S}|_i\}$:

$$\mathcal{L}(\{|S|_i, |\bar{S}|_i\}) = \left\langle \left(|S|_i - |\bar{S}|_i\right)^2 \right\rangle_i. \tag{2}$$

B. Inverse model — coupled training for constrained design

The inverse model must synthesize designs to exude specified electrical characteristics while abiding by constraints. It is trained with a newly generated parameter set $\{\vec{\xi}_i\}$ with performances $\{|S|_i\}$ predicted using the forward model. A design $\bar{\vec{\xi}}$ is synthesized by the inverse model to meet performance $|S|$:

$$\mathrm{IM}(|S|) =: \bar{\vec{\xi}} \implies \mathrm{FM}\left(\bar{\vec{\xi}}\right) =: |\bar{S}| \approx |S|. \tag{3}$$

The performance of the inverse model is quantified both by its ability to produce a proper design, $|\bar{S}| \approx |S|$, and the compliance of this design with all provided constraints:

$$\tilde{\mathcal{L}}\left(\left\{|S|, |\bar{S}|, \bar{\vec{\xi}}\right\}\right) = \tilde{\mathcal{L}}_1(\{|S|, |\bar{S}|\}) + \lambda\tilde{\mathcal{L}}_2\left(\left\{\bar{\vec{\xi}}\right\}\right). \tag{4}$$

TABLE I
PARAMETER RANGES FOR FORWARD MODEL DATA GENERATION

Parameter	r_{via}	r_{pad}	r_{ap}	p_{via}	p_{gnd}	l_{stub}	θ
Min. (mm, °)	0.2	0.3	0.4	1.0	3.0	0.1	30
Max. (mm, °)	0.4	0.5	0.6	1.2	3.2	0.3	60

Loss $\tilde{\mathcal{L}}_1$ is a mean-square error between electrical performances as in (2). Loss $\tilde{\mathcal{L}}_2$ quantifies how far a synthesized design is from obeying all constraints, which will require care to implement. Weight λ balances the losses during training.

Such constraints will be separated into two categories in this work: *interval* and *mutual*. The former is the constraint that each parameter of a synthesized design must lie within its allowed range, specified for this problem in Table I. This restriction may represent manufacturing limitations or SI rules. The latter represents geometric relationships between parameters that must be upheld to avoid faulty designs, an example being a required margin between the pad and anti-pad radii r_{pad} and r_{ap}. There are four such constraints present:

$$\begin{aligned} c_1 &:= (r_{via} + 0.05\text{mm}) - r_{pad} \leq 0, \\ c_2 &:= (r_{pad} + 0.05\text{mm}) - r_{ap} \leq 0, \\ c_3 &:= (2r_{pad} + 0.2\text{mm}) - p_{via} \leq 0, \\ c_4 &:= (p_{via} + 2r_{ap} + 2r_{pad}) - p_{gnd} \leq 0. \end{aligned} \tag{5}$$

A cursory approach to implement these is to punish a faulty design with high loss and reward otherwise with low loss by use of an indicator function. Such an approach fails in practice due to this loss not being differentiable [4]. This will be amended by approximating the indicator function smoothly, referred to here as a *barrier* function.

A barrier function should reward acceptable designs while exhibiting high gradient outside spec to encourage network optimization. It should be relatively flat within spec so as to not bias admissible designs. This work introduces the exponential and $[-1, 0]$-bucket barrier functions and their gradients:

$$\begin{aligned} \hat{I}_{-,\text{exp}}(u) &= e^{tu}/t \implies \nabla \hat{I}_{-,\text{exp}} = e^{tu} \nabla u. \\ \hat{I}_{[-1,0]}(u) &= \frac{e^{tu} + e^{-t(u+1)} - 2e^{-t/2}}{t} \\ &\implies \nabla \hat{I}_{[-1,0]} = \left(e^{tu} - e^{-t(u+1)} \right) \nabla u. \end{aligned} \tag{6}$$

These two functions, shown in Fig. 1, indicate whether argument u is within range. The exponential barrier rewards $u < 0$, intended for use with the mutual constraints (5). The bucket barrier rewards $u \in [-1, 0]$, to be used with interval constraints. Each barrier has a stiffness parameter t; higher stiffness accentuates the barrier properties, though this may introduce instabilities in training. The peculiar arrangement of (5) now becomes clear; each c_k can be fed directly into $\hat{I}_{-,\text{exp}}$. The interval constraints are normalized to regularize units before evaluation against the bucket barrier:

$$\bar{\xi}_k \in [a, b] \mapsto c_{k+4} = \frac{\bar{\xi}_k - b}{b - a} \in [-1, 0]. \tag{7}$$

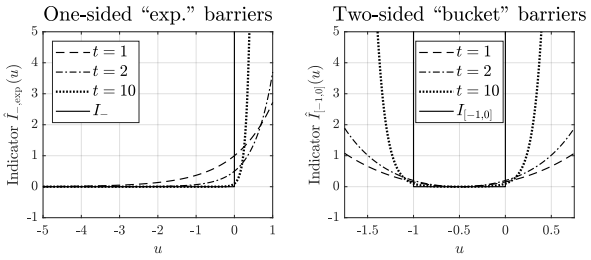

Fig. 1. Exponential and bucket barriers for $t = 1, 2, 10$, with indicators I.

Normalization is unnecessary for the mutual constraints in this problem as each is of similar order. The barriers are summed and averaged over a training batch to produce $\tilde{\mathcal{L}}_2$:

$$\tilde{\mathcal{L}}_2 \left(\left\{ \bar{\vec{\xi}}_i \right\} \right) = \left\langle \sum_{k=1}^{4} \hat{I}_{-,\text{exp}} \left(c_{k,i} \right) + \sum_{k=5}^{11} \hat{I}_{[-1,0]} \left(c_{k,i} \right) \right\rangle_i . \tag{8}$$

Here, $c_{k,i}$ is the kth condition of the ith design. The barriers depend implicitly on t, which should be chosen in combination with λ to ensure the final model exhibits both matching and constraint adherence. A rule of thumb is to make sure that the worst-case acceptable $\lambda \tilde{\mathcal{L}}_2$ is of the same order of magnitude as the desired $\tilde{\mathcal{L}}_1$ so that constraint adherence is trained for first, fine-tuning for prediction accuracy afterwards.

III. NUMERICAL RESULTS

A. Forward- and inverse-model training

The process outlined in Section II was carried out to form forward and inverse models representing a differential launch shown in Fig. 2 with a design space summarized by Table I and (5). The substrate is based on low-loss Doosan DS7409DV with $\epsilon_r = 3.25$ and $\tan\delta = 0.004$ [3]. The forward model training was done with 2,000 samples and tested against another 1,000, resulting in $\{0.045, 0.144, 0.065, 0.111\}$dB RMS loss at each frequency. Similar results were attained with less data, but this was chosen to maximize accuracy.

The inverse model was trained repeatedly to balance λ and t for this problem. Each model is used to synthesize 2,000 designs from a simulated dataset; a violation is counted if any constraint is disobeyed. $\tilde{\mathcal{L}}_1$ is measured between the true data and the forward model of the synthesized design. Fig. 3 summarizes this experiment, demonstrating the trade-off between the two goals; too much emphasis on $\tilde{\mathcal{L}}_2$ worsens $\tilde{\mathcal{L}}_1$, and

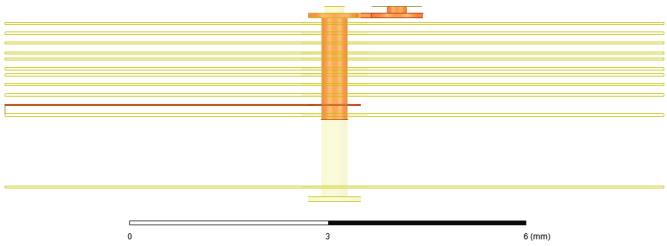

Fig. 2. Side profile of backdrilled symmetric differential via pair.

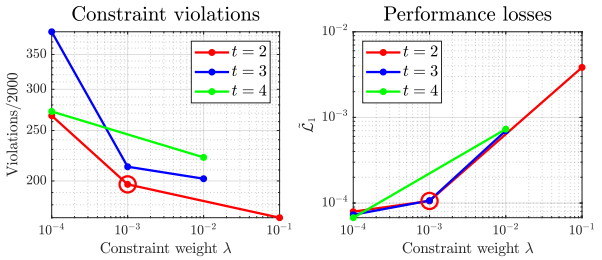

Fig. 3. Inverse models trained varying (λ, t), tested against 2,000 simulations.

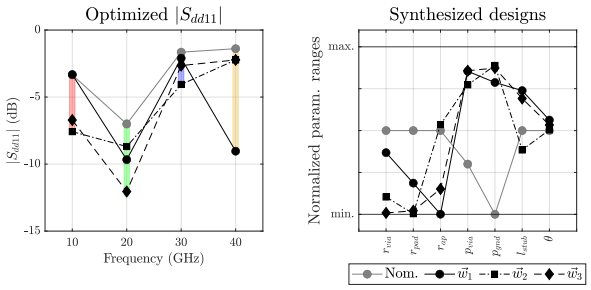

Fig. 4. Optimal performances and synthesized physical designs for each $\vec{w}_i$.

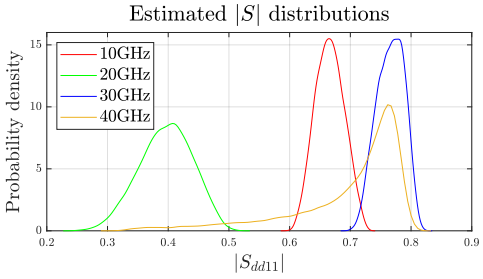

Fig. 5. S-parameter distributions about design optimized for equal frequency weighting, calculated with trained forward model.

vice versa. The best trained models for each combination are used to show their capability. The chosen model is indicated at $\lambda = 10^{-3}$ and $t = 2$. Runs with $t > 4$ tended to never converge; this may be due in part to the random network initialization providing designs initially far out of spec which contribute massive $\tilde{\mathcal{L}}_2$ with extreme gradients.

B. Design optimization

The inverse model will be applied to optimize the design by minimizing a weighted combination of its S-parameters:

$$\text{minimize } \vec{w} \cdot |\bar{S}|. \tag{9}$$

The weights $\vec{w} = (w_1, w_2, w_3, w_4)^T$ can be chosen to emphasize desired behaviors analogous to the spectra of transient signals incident on the via pair; $\vec{w}_1 = (1, 1, 1, 1)^T$ will weight each frequency equally while $\vec{w}_3 = \left(1, 1, \frac{1}{4}, \frac{1}{4}\right)^T$ emphasizes low frequency. Vector $\vec{w}_2 = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right)^T$ resembles the inverse-frequency envelope of a square wave spectrum.

This optimization starts with a known design and its electrical performance, then randomly sampling thousands of improved S-parameter vectors. The inverse model will synthesize the design of each, discarding any non-physical design. It is important to sample many combinations of $|\bar{S}|$, as not all combinations are achievable within the design space. The realizable design satisfying (9) for given $\vec{w}$ is chosen. Fig. 4 shows an optimal performance and accompanying design for each weight $\vec{w}_i$; the bounding bars show independent performance ranges at each frequency. From the design trends we gather that smaller via features improve performance, and that there is potentially an impedance mismatch due to undersized via pitches. Design-based criteria including fixing parameters may be added into such an optimization with ease.

C. Performance uncertainty analysis

A benefit of the surrogate model approach to design optimization is the potential to perform sensitivity or uncertainty analysis of the optimized design for little extra cost. This was done for the equal-frequency optimized design produced by the previous section by modeling the variables as uniformly random with a 5% tolerance. The resulting S-parameter spread is shown in Fig. 5. The spread is relatively low at 10 and 30GHz, while wide variations at 20 and 40GHz may indicate resonances near the optimal design.

IV. Conclusion

The presented tandem forward- and inverse-modeling scheme is capable of synthesizing test coupon launch designs that exude prescribed performance and obey constraints imposed by manufacturing or physical limitations. Careful choice of weight and stiffness parameters λ and t allows acceptable model training. The resulting inverse model can be used for swift multi-purpose design optimization. The forward model can perform uncertainty analysis around a design for little cost, a capability not possessed by an optimization based only on full-wave simulation. Further work may include implementing a wider design space for a multi-purpose inverse model.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. CNS 2137288 - Center for Advanced Electronics through Machine Learning (CAEML) and its industry members.

References

[1] J. Hsu, T. Su, K. Xiao, X. Ye, S. Huang, and Y. L. Li, "Delta-l methodology for efficient pcb trace loss characterization," in *2014 9th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT)*, 2014, pp. 113–116.

[2] J. Hsu, C. Lin, M. Ku, T. Su, and D. Chen, "Semi-automation delta-l for pcb loss electrical characterization," in *2018 13th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT)*, 2018, pp. 154–156.

[3] J. Lee, K. Kim, S. Kim, K. Kim, and K. Lee, "Optimization of pcb si coupon design that minimizes discontinuity through via-in-pad plated over (vippo) technique," *Journal of Microelectronics and Electronic Packaging*, vol. 17, pp. 128–137, 10 2020.

[4] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.