Eye Height/Width Prediction From *S*-Parameters Using Learning-Based Models

Nikita Ambasana, *Student Member, IEEE*, Gowri Anand, Bhyrav Mutnury, *Senior Member, IEEE*, and Dipanjan Gope, *Senior Member, IEEE*

***Abstract*— Design space exploration and sensitivity analysis for electrical performance of high-speed serial links is a critical and challenging task for a robust, cost-efficient, and signal-integrity-compliant channel design. The generation of time-domain (TD) metrics like eye height and eye width at higher bit error rates requires longer bit sequences in TD circuit simulation, which is compute time intensive. Intelligent techniques to identify smaller design sets that cover the design space optimally may provide incorrect sensitivity analysis. This paper explores learning-based modeling techniques that rapidly map relevant frequency-domain metrics like differential insertion loss and total cross talk, in the presence of equalization, to TD metrics like eye height and eye width, thus facilitating a full-factorial design space sweep. Numerical results performed with multilayer-perceptron-based artificial neural network as well as least-squares support vector machine (LS-SVM) on Serial ATA 3.0 and Peripheral Component Interconnect Express Gen3 channels generate an average error of less than 2%.**

***Index Terms*— Artificial neural network (ANN), eye height, insertion loss (IL), multilayer perceptron (MLP), Peripheral Component Interconnect Express, return loss (RL), Serial ATA, signal integrity.**

I. Introduction

HISTORICALLY, eye diagrams and their characteristics, eye height (EH) and eye width (EW), have served as a metric for the quality of high-speed channels. With increasing data rates, the interconnect lengths on package boards start becoming electrically long, and hence, accurate transmission line models [1] and *S*-parameters [2] for SPICE time-domain (TD) simulations, using convolution of channel pulse response and input bit streams, were developed. However, with more stringent bit error rate requirements, this process became compute power and time intensive. Based on methods to estimate probability of error for binary symmetric channels in the presence of intersymbol interference (ISI) and cochannel interference [3], algorithms like LinkLab [4] and StatEye [5] utilizing probability density function estimates of jitter and ISI were formulated to obtain eye contours and bath tub curves [6]. However, these approaches did not capture the nonlinear nature of the transmitter and receiver accurately.

Hybrid approaches were developed to capture the goodness of statistical estimates and parallel TD simulations of select bit sequences [7]–[9]. However, given all the elements to speed up channel analysis, for a selected channel topology, there are several variables impacting the signal quality. These variables include controllable variables like trace length, impedance termination, and uncontrollable variables such as process and manufacturing tolerances.

Design rules are needed to ensure that signal quality is maintained. In order to come up with design rules, several TD simulations are needed to cover the design space. State-of-the-art existing methods involve statistical or optimization techniques using design of experiments (DoE), involving a systematic approach to select the smallest set of designs that optimally captures the design space using different ideologies such as orthogonality, replication, randomization, and blocking [10]. Response surface method [11] can be used to fit the DoE-based designs to predict EH and EW and perform Monte Carlo analysis. Parametric mapping using the multilayer perceptron (MLP)-based artificial neural network (ANN) can also predict EH and EW from channel topology variables [12]. A DoE-based training mechanism was used to train the network in order to capture the dependence of EH and EW on channel parameters. From [13], it is also evident that for a design space with large dimensions and highly nonlinear EH/EW behavior, a DoE-based approach can sometimes provide an inaccurate sensitivity analysis compared with a full-factorial sweep across all possible design and process variations. In spite of all the advanced techniques for accurate eye diagram generation discussed above, a full-factorial sweep TD simulation is prohibitive due to its excessive time requirements. An end-to-end channel simulation in frequency domain (FD) involves cascading available macromodels of individual channel components and is almost 100× faster than the TD simulation. Hence, a full-factorial design space sweep may be possible in FD.

In this paper, building on the work presented in [14], a methodology for creating learning-based models for mapping *S*-parameters directly onto eye diagram metrics, EH/EW, is discussed. The goal of this paper is not to reconstruct the TD waveform at the receiver or to construct a complete eye diagram, but predicting only EH and EW metrics from *S*-parameter data in order to facilitate a full-factorial design space exploration, which is currently computationally prohibitive. The input parameters to the mapping process are

Manuscript received September 8, 2015; revised January 21, 2016 and February 26, 2016; accepted March 15, 2016. Recommended for publication by Associate Editor S. Roy upon evaluation of reviewers' comments.

N. Ambasana and D. Gope are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail: ambasananikita@gmail.com; dipanjan.gope@gmail.com).

G. Anand is with Dell Enterprise Server Group, Bengaluru 560071, India (e-mail: gowri_anand@dell.com).

B. Mutnury is with the Enterprise Server Simulation Group, Dell Inc., Round Rock, TX 78682, USA (e-mail: bhyrav_mutnury@dell.com).

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/TCPMT.2016.2552081

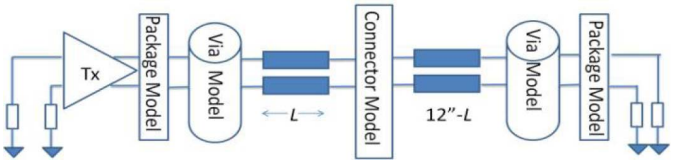


Fig. 1. PCIe Gen2 topology with one differential pair and variable interconnect length L.

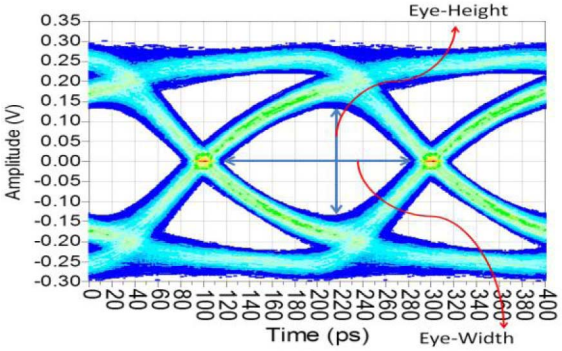


Fig. 2. Eye diagram profile for the PCIe Gen2 topology of Fig. 1 for $L = 4$ in.

the chosen S-parameters that affect the EH and EW under consideration. For generating an effective map, learning-based modeling techniques are explored. Based on experimental results, least-squares support vector machine (LS-SVM) and MLP-based ANN are chosen for this application with appropriate model parameter selection. A DoE-based training set is used to train the ANN and a full-factorial design space sweep is enabled by translating FD parameters to TD metrics EH and EW through learning-based models. With increasing design space dimensions in the DoE process, the size of the training set and correspondingly training time may be adversely affected. A reduced training set (RTS) method is proposed to alleviate this problem and generate a bound in the number of vectors used for training.

This paper is organized as follows. In Section II, FD-to-EH/EW mapping is proposed as a viable alternative for full-factorial design sweep and the constituent steps are outlined. In Section III, a brief overview of filtering-based feature selection (FS) techniques for identifying frequency steps for mapping is presented. In Section IV, learning-based modeling techniques used in the proposed approach are explained. In Section V, the training set generation methods, DoE and RTS, are described. In Section VI, the numerical results are presented to demonstrate the efficacy of the proposed approach for Serial ATA (SATA) 3.0 and Peripheral Component Interconnect Express (PCIe) Gen3 interfaces. Section VII gives a critical analysis of the proposed algorithm, and Section VIII concludes this paper.

II. FD-to-EH/EW Mapping

DoE for the design space exploration of high-speed Serializer/Deserializer (SerDes) (HSS) channels may provide incorrect sensitivity analysis as demonstrated in the following example. A PCIe Gen2 topology with a single pair of differential ports is shown in Fig. 1. The design space consists of the length of the interconnect L.

TD SPICE simulations are performed with S-parameter models of the channel for L ranging from 2 to 5 in at

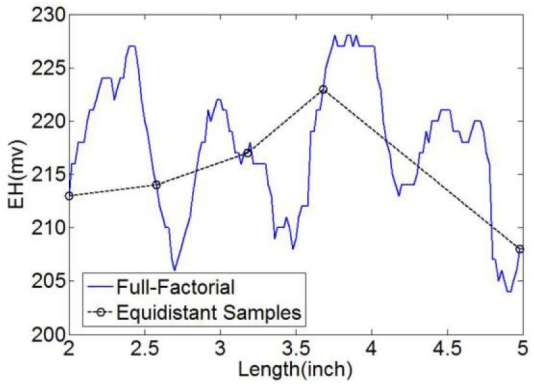


Fig. 3. Sensitivity of EH as a function of L for the PCIe Gen2 topology of Fig. 1. The equidistant samples, representative of a typical DoE, do not capture the real profile as evident from the full-factorial finely sampled simulations.

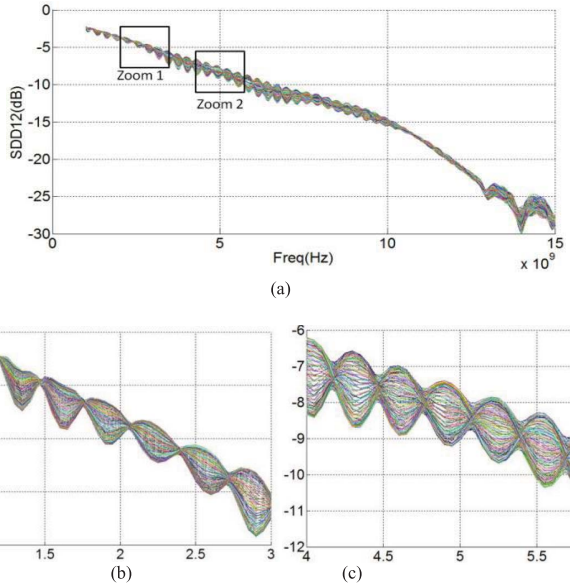


Fig. 4. (a) IL profiles for the PCIe Gen2 channel of Fig. 1. (b) Zoom 1. (c) Zoom 2.

20-mil variations. Fig. 2 shows the eye diagram at $L = 4$ in for a 10b8b nonreturn to zero code simulated over 8000 b.

Fig. 3 represents the plot of EH versus L. The bold line shows the actual variation of EH with L measured over 151 different values of L, whereas the dotted line and circles indicate the EH at five equally spaced points between 2 and 5 in representative of a DoE-based sampling. It is evident that the DoE-based smaller set gives inaccurate information of the sensitivity of EH with L over the range under consideration.

From this example, it is seen that a finer sampling may be required for an appropriate estimation of EH and its sensitivity. However, such a large number of TD simulations are prohibitive due to excessive compute time requirements.

Since the FD parameters of the channel are readily available from cascading the macromodels of the individual channel components, it is tempting to resort to a map from FD to EH/EW in order to alleviate the time bottleneck. An exercise was performed to identify such a relationship for PCIe Gen2 channel shown in Fig. 1 without any cross talk. In Figs. 4 and 5, the differential insertion loss (IL) and

TABLE I
RL AND IL PROFILES FOR DIFFERENT DESIGN CORNERS OF PCIe Gen2 CHANNEL OF FIG. 1

		IL (dB)				RL (dB)			
Length (mils)	Eye-Height (mv)	1GHz	2GHz	4GHz	6GHz	1GHz	2GHz	4GHz	6GHz
4900	204	-2.41347	-3.96574	-8.12408	-10.0467	-12.4397	-10.1199	-6.29275	-12.9277
2680	207	-2.66821	-3.79643	-6.20099	-10.4171	-9.8047	-10.7203	-36.1231	-6.63771
2000	213	-2.70503	-3.51285	-7.49776	-10.6304	-9.53155	-23.7923	-6.77449	-5.97518
3760	228	-2.2299	-3.58206	-6.83026	-10.3176	-15.4531	-20.8018	-10.7232	-7.83625

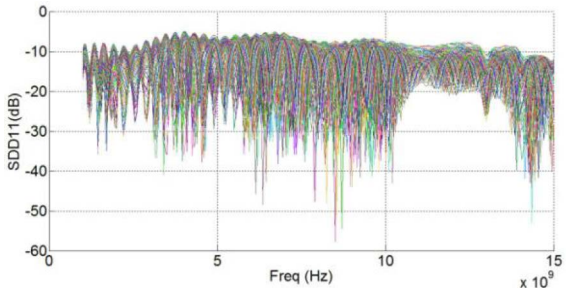


Fig. 5. RL profiles for the PCIe Gen2 channel of Fig. 1.

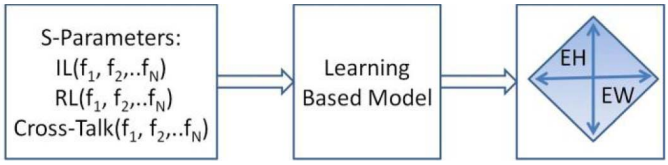


Fig. 6. Top-level block diagram of information flow in the proposed mapping methodology.

differential return loss (RL) profiles for several values of L for the channel in Fig. 1 are shown. In Table I, the first column lists the EH for four designs. The second and third main columns list the IL and RL at the indicated frequencies for the same designs. It can be seen from column 3 and rows 3 and 6 that a negligible variation of 0.38 dB in differential IL at 2 GHz in FD appear as a 24-mV difference in EH. In addition, comparing rows 5 and 6 for column 7, it is observed that better RL values lead to worse EH. Searching for a pattern in the IL and/or RL profiles at the second and third harmonics results in no meaningful conclusions either. Therefore, it can be concluded that the FD-to-EH/EW mapping is not straightforward.

This leads to an exploration of learning-based modeling techniques to map FD *S*-parameter data to EH/EW. A top-level block diagram indicating the objective of the mapping process is shown in Fig. 6, where f_1, f_2, and f_n represent the frequency points where the *S*-parameter profiles are sampled.

The process of creating such a model to predict EH/EW from *S*-parameters can be broken down into three steps for a given interface.

Step 1: Identifying the *S*-parameters and frequency step size for optimal mapping.

Step 2: Applying learning-based modeling and identifying optimal modeling parameters.

Step 3: Identifying a suitable training set.

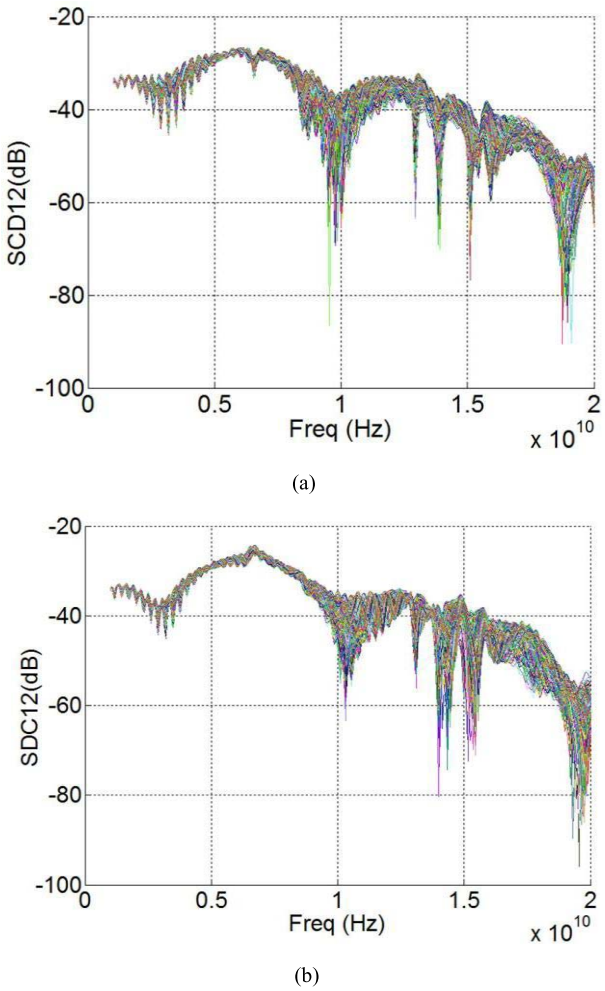


Fig. 7. *S*-parameter profiles for the PCIe Gen2 channel of Fig. 1. (a) Differential–common. (b) Common–differential.

III. PARAMETER AND FREQUENCY SELECTION

For the channel mentioned above, the differential IL and RL profiles for several values of L are shown in Figs. 4 and 5(a). However, there are other *S*-parameters like common–differential and differential–common. Fig. 7(a) and (b) shows the profiles for SDC12 and SCD12, respectively, for the channel under consideration. In the frequency range of interest, from 1 to 9 GHz, the profiles of SCD12 and SDC12 are considerably lower in magnitude as well as contain less information compared with the differential IL, and hence, for this case, only differential *S*-parameters for the through

channel are considered as input for the mapping process. In addition, the channels considered here are highly linear in the phase response over the frequency range selected, and hence, there is no impact of the phase of S-parameters on the EH and EW.

Such a study needs to be done to satisfy Step 1 and identify relevant S-parameters. In addition, a critical challenge in this process is the selection of appropriate frequency points that would be used to train and use the network. In [15], a discrete frequency sweep using a fixed start frequency, end frequency, and frequency step size was employed. However, this is a nonoptimal solution and the choices of frequencies are heavily dependent on the channel type.

In [16], a frequency selection process for FD-to-EH/EW mapping is described using a filter-based FS algorithm that identifies the frequencies containing maximum information for each of the EH/EW values observed in the DoE set and removes frequencies that contain redundant data. This process of intelligent input selection facilitates accurate modeling while reducing the input set size. In the numerical results presented in this paper, uniform frequency sampling based on channel characteristics has been used.

IV. Learning-Based Modeling

Learning-based modeling is a way to establish a parametric relation between two different models of the same physical system, wherein one model is accurate but requires expensive functional evaluations and the other is cheap to evaluate but less accurate, called the coarse model [17], [18]. In the context of FD to EH/EW, it can be explained as follows.

Let X be the set of designs in the design space of a given HSS channel. Let $D \subset X$ be the set of DoE of size d for that channel and $x_i \in X$ be the ith design. Let $t_i \in T \subseteq R^1$ be the TD metric (EH/EW) and $f_i \in F \subseteq R^m, m \geqslant 1$, be the FD metric or vector of S-parameters, for the ith design x_i. Let $\{t_1, t_2, \ldots, t_d\}$ be the set of d EH (EW) simulation results and $\{f_1, f_2, \ldots, f_d\}$ be the set of d FD simulation results. If $P : F \rightarrow T$ is the modeling function that maps FD to EH/EW, then the learning-based modeling problem can be summarized as obtaining P, such that

$$t = P(f) \tag{1}$$

$$\text{s.t. } \|t_i - P(f_i)\| \leq \epsilon \quad \forall x_i \in D \tag{2}$$

where ϵ is a nonzero positive real number. This process of obtaining the modeling function P for learning-based modeling techniques is iterative in nature. Many surrogate modeling techniques exist, but none of them has proven to be globally optimal across a range of problems [19]–[25]. A brief overview of four of the techniques studied for this paper is presented in Table II. Two techniques, namely, ANNs using MLP topology and LS-SVM, are found to be more suited to handle a high-dimensional mapping problem such as the FD-to-EH/EW mapping discussed here. These two techniques are described in detail in this section.

A. Least-Squares Support Vector Machines

The support vector machine (SVM) algorithm creates a classification function that is tuned to maximize the margin between training examples and classification boundary, in the process, identifying support vectors that are crucial to the classifier and discarding outliers. If the sample space is highly nonlinear for classification, the algorithm uses kernel functions to map this space into a higher dimensional feature space wherein an optimal hyperplane separating the classes or passing through all the training data is created [19], [20], as shown in Fig. 8.

In the work presented in this paper, LS-SVM is used for mapping, where the kernel is the L2-Norm.

TABLE II
Overview of Learning-Based Modeling Techniques

Modeling Technique	Description	Observation
Polynomial	Rational polynomial approximation, simple to evaluate [24]	Fails at higher input dimensions
Kriging	Kolmogorov-Weiner based predictive interpolation, requires accurately spaced training data [25]	Loss of accuracy at higher input dimensions
ANN	Described below	Fast and accurate at higher input dimensions
LS-SVM	Described below	Fast and accurate at higher input dimensions

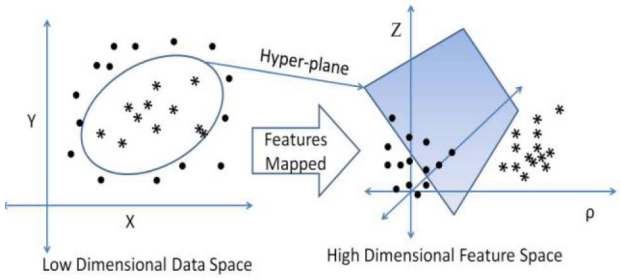


Fig. 8. SVM feature mapping technique.

B. ANNs and MLP

Neural networks are machines that are made to function in a manner similar to a brain. Their two main functions are to acquire knowledge from the environment through a learning process and store this knowledge in weights. The prediction function for a single-hidden-layer MLP ANN with N inputs, M hidden nodes, and one output is

$$y = f(\boldsymbol{x}) = \sum_{j=1}^{M} k_j * G\left(\sum_{i=1}^{N} w_{ij} x_i + b_j\right) + d \tag{3}$$

$$G(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{4}$$

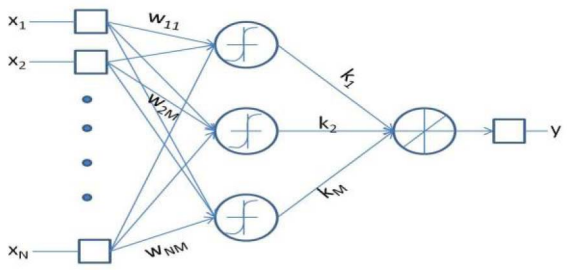


Fig. 9. Generic architecture of an ANN.

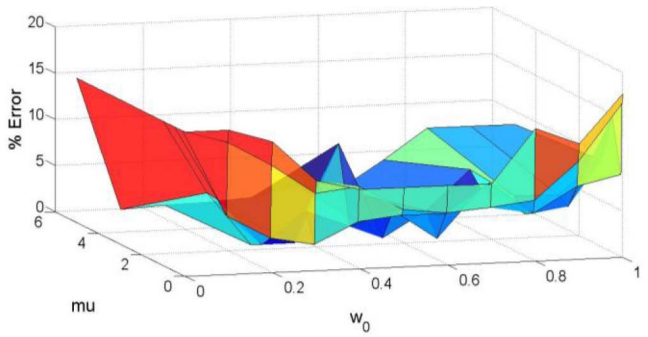


Fig. 10. Error surface over training parameters μ and $\boldsymbol{w_0}$. It can be observed that the error metric is highly nonlinear and steeply varying function in the training space.

where $\boldsymbol{x}$ is the input vector, w_{ij} is the weight connecting the ith input node to the jth hidden perceptron, and k_j is the weight connecting the jth hidden perceptron to the output node. An MLP is a two or more layer neural network with neuron units as perceptrons that have nonlinear activation functions in the hidden layer/s [21], as shown in Fig. 9. The intrinsic equation of an MLP makes it suitable for mapping the input M-D vector onto continuous output functions. In this work, the Levenberg–Marquardt (LM) algorithm [22] is used for training and gradient descent approaches is used on mean-squared error for weight optimization in back-propagation learning for a single-layer feedforward network with one hidden layer having a tan sigmoid activation function. The output layer is a linear mapping function.

C. Model Parameter Selection

Designing ANN for a particular task requires several parameter settings to be chosen by the user. Some of these settings are the number of hidden layers, the interconnections between neurons, the number of neurons in each layer, the activation function in the hidden and output layers, the training algorithm, and performance metric and constants controlling the training process. The goodness of a model and the speed with which it converges depends largely on the model parameters that impact the training function. The performance of the LM algorithm is influenced by the damping factor μ. At every iteration of the algorithm, the value of μ is modified by a positive increment of $\Delta\mu_+$ or a decrement of $\Delta\mu_-$. The learning iterations stop when the mean-squared error approaches a target goal ϵ or the gradient of the error function exceeds a threshold δ_f. The network is initialized with a set of weights w_0.

In Fig. 10, it is demonstrated using two of the training parameters (μ and w_0) that the error metric is a highly nonlinear and steeply varying function in the training parameter space. A method of nested binning in the parameters, μ, $\Delta\mu_+$, $\Delta\mu_-$, δ_f, and w_0, mentioned above is used to arrive at a set of training parameter values for a particular problem space. However, more elegant global optimization techniques like the genetic algorithm [23] can be used to determine the network parameters for optimal performance.

V. Training Set Generation Methods

One of the two primary tasks of learning-based models is to acquire knowledge from the system through a learning process. This task can be broken down into two phases: 1) training and 2) validation. While training the model, it is necessary to cover the design space optimally as well as keep the set as small as possible. The validation set is required to ensure that the network is not overtrained. Two different methods of generating a training set are described here.

A. Design of Experiments

DoE is a method to generate a reduced set of designs mathematically calculated to cover the entire design space in the fewest possible set of vectors. A detailed explanation of methods of generating a DoE set can be found in [10]. In this paper, the central composite design method of generating DoE is used, resulting in a training set size of $(2^{d-1} + 2d + 1)$ designs, where d is the dimension of the design space. Designs selected based on the DoE algorithm are simulated in FD and TD to get the EH/EW. These data are used for training the model, which is then used to predict EH/EW values for a full-factorial set of FD simulations.

B. Reduced Training Set

A considerable improvement in speedup of the design space exploration/sensitivity analysis process can be obtained if a training set is generated so as to not have the handicap of scaling with the design space dimensions d. In the case of FD-to-EH/EW mapping, such a set can be devised if a full-factorial design simulation in FD is available.

The concept of RTS is to ensure that the model is trained across the entire input span irrespective of the range of inputs. Hence, it is sufficient to train the model on designs that form the outliers and median at each sampled frequency point. As an extension of the example presented in Section II, differential IL for a particular topology is obtained by full-fledged all corner FD simulations for a given set of design variables. For each of the sample points lying on the vertical lines indicated in Fig. 11, three values are identified, the minimum, the maximum, and the nearest to median.

The designs corresponding to each of the identified differential IL values at discrete preselected frequencies are then combined and used as the RTS. TD simulations are conducted for the selected designs in the RTS. These TD and FD data are then used for training the ANN.

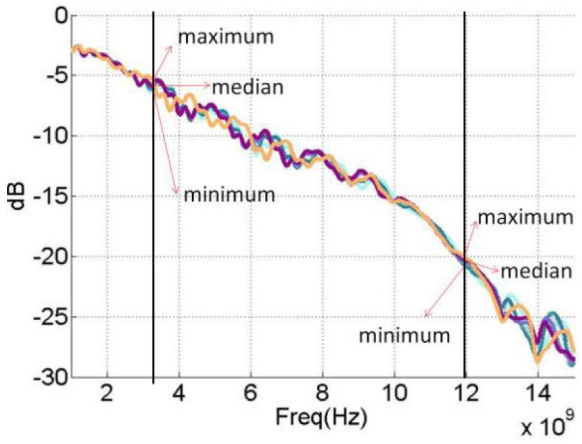


Fig. 11. Outlier and median-based design selection for differential IL.

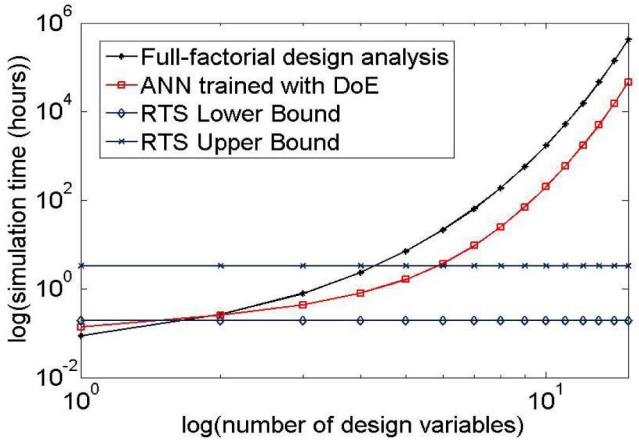


Fig. 12. RTS over DoE speedup comparison curve.

The benefit of this method of identifying a training set for a given topology can be seen from Fig. 12. The size of a DoE set for a given topology (S_d) scales exponentially with the number of design variables (5), whereas the size of the training set obtained by the RTS technique (S_r) is always bounded by the limits indicated in (6). If d is the design space size (number of design variables), M is the number of FD metrics considered for mapping, and K is the number of samples of FD metrics fed into the space map as input, then

$$S_d = 2^{d-1} + 2d + 1 \tag{5}$$

$$3M \leq S_r \leq 3MK. \tag{6}$$

VI. Numerical Results

In this section, numerical experiments are performed using the proposed learning-based mapping methodology to accurately predict EH/EW values from S-parameters.

The experimental process is summarized in Fig. 13. The diagram also highlights the methodology for testing the learning-based models generated.

Numerical experiments are performed on two HSS interfaces, SATA 3.0 and PCIe Gen3, to quantify the performance of ANN and SVM to map S-parameters to EH/EW. The

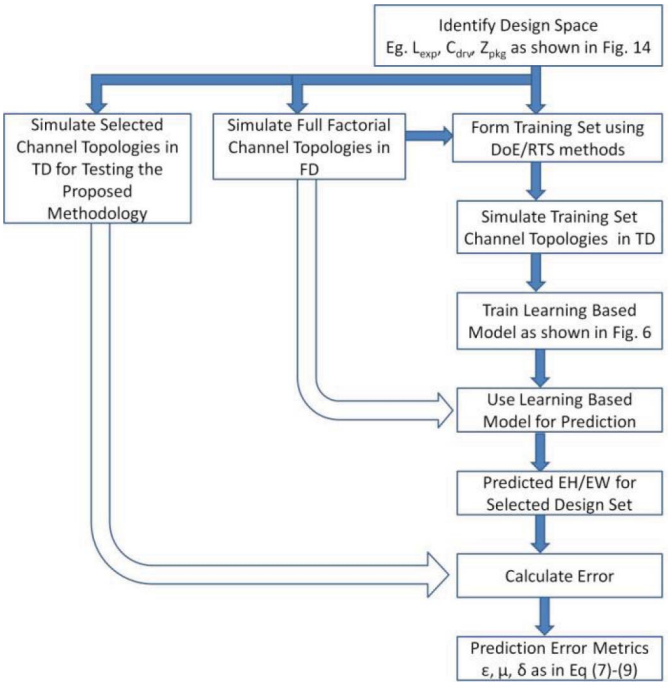


Fig. 13. Learning-based model generation and testing procedure used for the testing prediction accuracy for SATA 3.0 and PCIe Gen3 channels, presented in Section VI.

metrics used to measure mapping accuracy are, L2-Norm error (ε), average error (μ), and maximum error (δ), as defined below

$$\varepsilon(\vec{a}, \vec{b}) = \sqrt{\frac{\sum_{i=1}^{N} |a_i - b_i|^2}{\sum_{i=1}^{N} |a_i|^2}} \tag{7}$$

$$\mu(\vec{a}, \vec{b}) = \frac{\sum_{i=1}^{N} |a_i - b_i|}{N} \tag{8}$$

$$\delta(\vec{a}, \vec{b}) = \underset{i \in [1,N]}{\text{MAX}} |a_i - b_i| \tag{9}$$

where a_i are the predicted EH/EW values for the ith channel topology using the proposed mapping methodology and b_i are the values obtained after simulating the channel in TD using a SPICE simulator. N is the size of the test set.

A. SATA 3.0

The topology shown in Fig. 14 is used in three serial-link differential pairs shown in Fig. 17. The design space is formed of seven variables: the driver port capacitance (C_{drv}), expander board impedances ($Z_{\text{exp}\,1}$ and $Z_{\text{exp}\,2}$), length of expander board (L_{exp}), backplane impedance (Z_{bp}), length of the backplane (L_{bp}), and receiver port capacitance (C_{rcv}). Figs. 15 and 16 show the IL and total far-end cross talk (FEXT) profiles for all the test cases over 0–10 GHz. The S-parameters chosen for input to the mapping process are differential IL of the victim and total FEXT at the victim receiver port, sampled at 500 MHz from 1–9 GHz. The models are formed by training over a set generated by DoE and RTS techniques discussed in Section V. Fig. 18 shows the simulated eye diagram for the variable values C_{rcv} and C_{drv} set at 0.2 pF, L_{exp} to 2 in, L_{bp} to 9 in, and $Z_{\text{exp}\,1}$, $Z_{\text{exp}\,2}$, and Z_{bp} to 115 Ω.

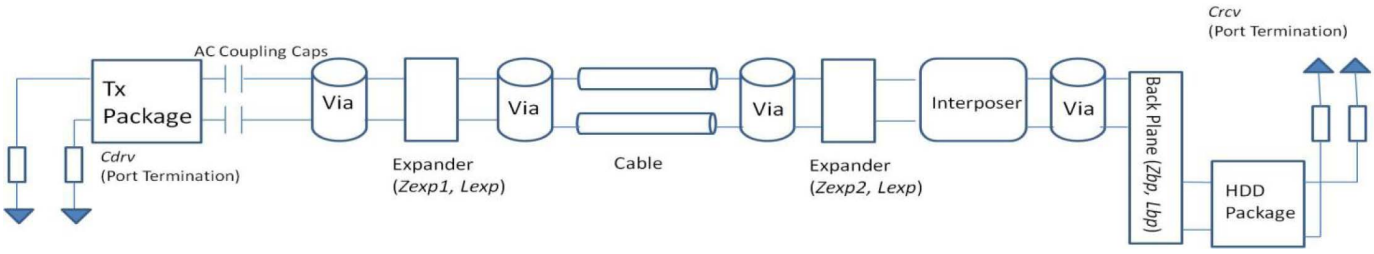


Fig. 14. SATA 3.0 topology from the controller to the hard disk drive for a single differential pair.

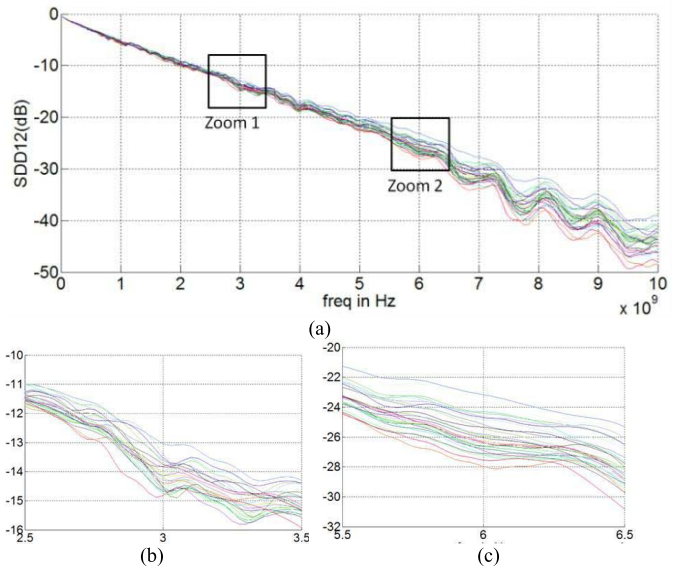


Fig. 15. (a) IL profiles for different combinations of design variables for SATA 3.0 of Fig. 14. (b) Zoom 1. (c) Zoom 2.

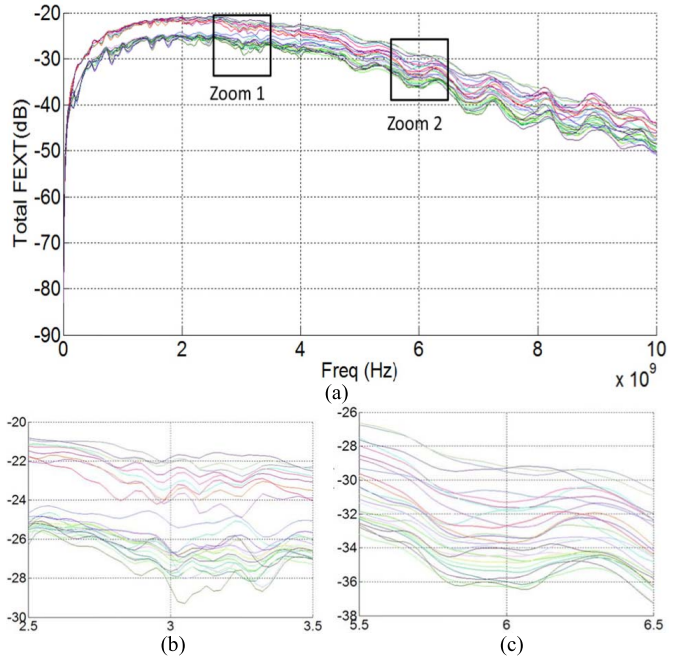


Fig. 16. (a) Total FEXT for different combinations of design variables for SATA 3.0 of Fig. 14. (b) Zoom 1. (c) Zoom 2.

The number of input nodes to the ANN is 34 and 17 points in frequency each for the differential IL and total FEXT. The size of DoE is 79 designs, size of RTS is 49, and the size of the test set is 129 randomly selected sets. The ANN used is of

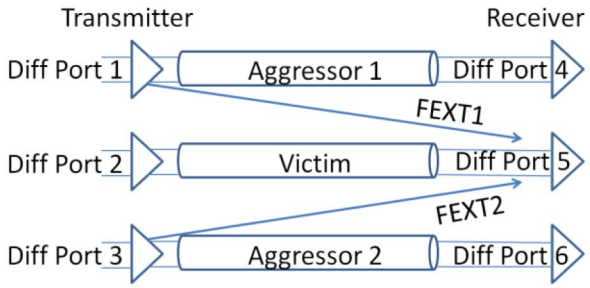


Fig. 17. Aggressor victim topology used for FD-to-EH/EW mapping in all numerical examples.

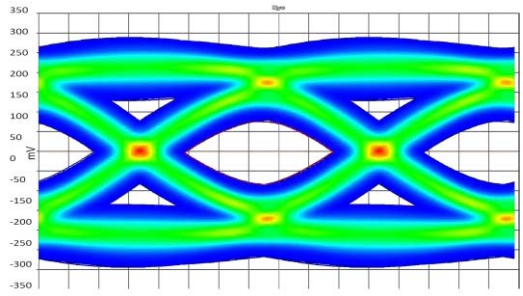


Fig. 18. Eye diagram for the SATA 3.0 channel for the design C_{rcv} and C_{drv} set to 0.2 pF, L_{exp} to 2 in, L_{bp} to 9 in, and $Z_{\text{exp}\,1}$, $Z_{\text{exp}\,2}$, and Z_{bp} to 115 Ω.

MLP type with one hidden layer formed of ten neurons having a sigmoid activation function and one output neuron with a linear activation function. The SVM used in all cases hence is from the Surrogate Modeling (SuMo) toolbox [26], [27] for the same number of inputs and outputs. Fig. 19(a) shows the correlation between the simulated and predicted EHs using ANN and SVM trained with DoE and RTS set of designs, and Fig. 19(b) shows the correlation between the simulated and predicted EWs using ANN and SVM trained with DoE and RTS set of designs. In Table III, the error metric values of the prediction accuracy, as defined by (7)–(9), are tabulated.

The values used to calculate the time of analysis are as follows.

1) TD simulation time td = 7 min.
2) FD simulation time tf = 45 s.
3) Number of variable $n = 7$.
4) Number of FD metrics $M = 2$ corresponding to magnitude in decibels of IL, total FEXT.
5) Number of frequency points sampled $K = 17$ corresponding 1 to 9 GHz at 500 MHz discretization.
6) Number of parallel processing cores $P = 4$.

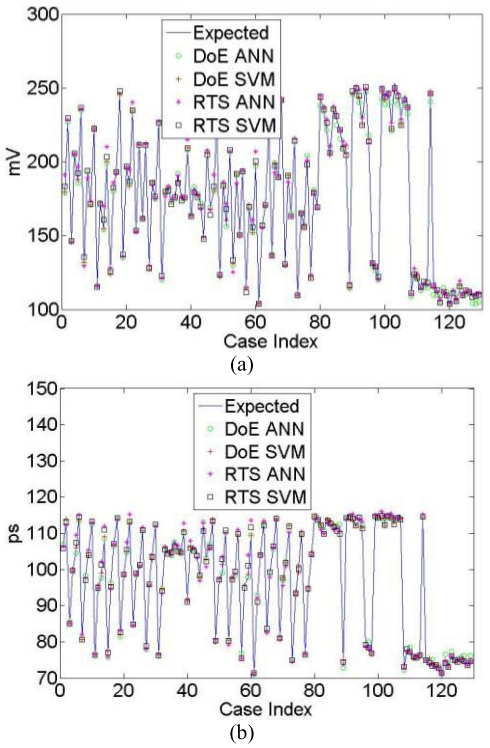


Fig. 19. Correlation of predicted data obtained from the proposed mapping methodology and simulated data obtained from TD simulations on selected designs for the SATA 3.0 channel of Fig. 14. (a) EH. (b) EW.

TABLE III
SATA 3.0 PREDICTION ERROR METRIC VALUES [ε, μ, AND δ AS PER (7)–(9)]

EH	DoE (79 Sets)			RTS (49 Sets)		
Modeling Technique	ε (%)	μ (mV)	δ (mV)	ε (%)	μ (mV)	δ (mV)
ANN	1.6	1.4	11.5	2.2	2.3	16.0
SVM	1.0	1.3	5.3	1.4	1.7	11.9
EW	DoE (79 Sets)			RTS (52 Sets)		
Modeling Technique	ε (%)	μ (ps)	δ (ps)	ε (%)	μ (ps)	δ (ps)
ANN	0.6	0.4	3.0	0.5	0.3	1.2
SVM	0.7	0.5	2.5	0.5	0.3	1.4

TABLE IV
SATA 3.0: SPEEDUP COMPARISON FOR THE EH/EW PREDICTION METHODOLOGY WITH DOE AND RTS TRAINING METHODS

Test Case	Full Factorial (min)	DoE (min)	RTS (min)
	$\frac{3^n t_d}{P}$	$\frac{3^n t_f + (2^{n-1} + 2n + 1)(t_f + t_d)}{P}$	$[3M - 3MK]$
7 Variables	3827	563	[12-198]

The time taken to optimize ANN parameters is negligible and not considered in the time analysis. Table IV gives a quantification of speedup for the SATA 3.0 test case obtained across the two different training methods mentioned above.

TABLE V
PCIe Gen3 PREDICTION ERROR METRIC VALUES [ε, μ, AND δ AS PER (7)–(9)]: CONSTANT EQUALIZATION

EH	DoE Training (45 Sets)			RTS Training (90 Sets)		
Modeling Technique	ε (%)	μ (mV)	δ (mV)	ε (%)	μ (mV)	δ (mV)
ANN	3.3	1.6	7.2	1.8	0.8	5.2
SVM	3.9	1.9	9.3	1.8	0.7	6.3
EW	DoE Training (45 Sets)			RTS Training (52 Sets)		
Modeling Technique	ε (%)	μ (ps)	δ (ps)	ε (%)	μ (ps)	δ (ps)
ANN	2.1	1.1	3.3	0.9	0.4	2.0
SVM	2.0	1.0	4.8	0.7	0.3	2.1

B. PCIe Gen3

A PCIe Gen3 channel with the topology shown in Fig. 20 and IL and total FEXT profiles shown in Figs. 21 and 22 are considered for the second set of experiments. The design space consists of six variables, *Zpckg1* and *Zpckg2* being the package impedances, *Z1* and *Z4* being the termination impedances, and *Z2* and *Z3* being the connector impedances, as marked in the topology with L set to 8.9 in. PCIe topology involves receiver equalization, which is not needed for SATA 3.0. For the experiments here, a three-tap feedforward equalizer (FFE) with a dc bias is used at the transmitter and a single-tap decision feedback equalizer (DFE) is used at the receiver end. For the first set of simulations, the equalization was kept constant across the different designs in both the training and test sets, whereas in the second experiment, the equalizer coefficient was allowed to adapt to the channel response and the impact of increasing the training set size on the accuracy was observed. The difference in the eye diagram due to adaptive equalization can be seen in Fig. 23. Fig. 23(a)–(c) shows the eye diagrams for a PCIe Gen3 channel without equalization, for constant equalization, and for the same channel with adaptive equalization, respectively. Equalization is a TD phenomenon for HSS interfaces and its effect is not simulated in the FD. However, the physics of the equalization is captured in the EH/EW of the designs provided for the model building, and hence, the FD-to-EH/EW mapping occurs with acceptable accuracy as can be seen in the following. A discussion on the nonlinearity introduced by the multiple-tap DFE and the prediction accuracy of the learning-based techniques, in presence of this nonlinearity, is presented in Section VI-B3.

1) Constant Equalization: The DoE training set consists of 45 designs and RTS training set consists of 90 designs. The test set consists of 300 randomly selected designs. The S-parameter inputs are differential IL and total FEXT at 500 MHz sampling from 500 MHz to 12 GHz (46 inputs). The ANN used is of MLP type with one hidden layer formed of 15 neurons having a sigmoid activation function and one output neuron with a linear activation function. Fig. 24(a) shows the correlation between the simulated and predicted EHs, for some cases, using ANN and SVM trained with

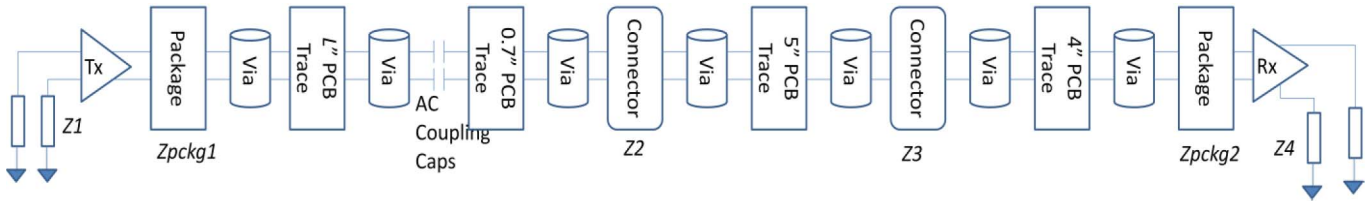


Fig. 20. PCIe Gen3 topology for a single differential pair.

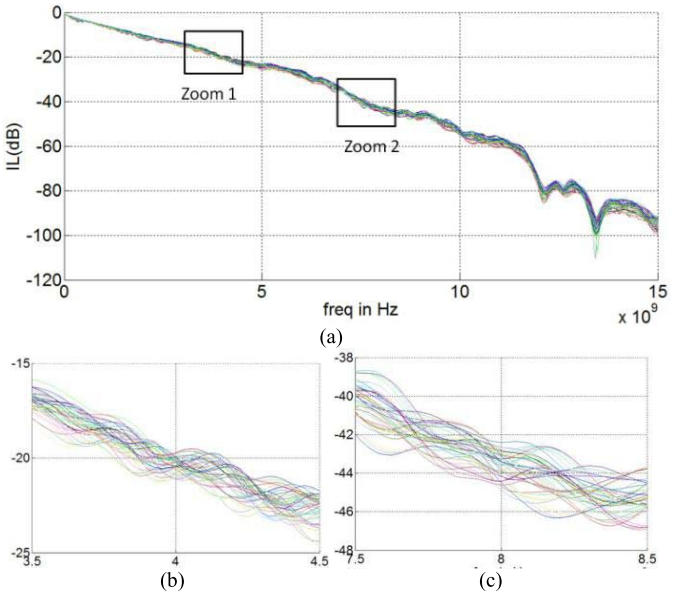


Fig. 21. (a) IL profiles for the PCIe Gen3 channel of Fig. 20. (b) Zoom 1. (c) Zoom 2.

DoE and RTS set of designs, and Fig. 24(b) shows the correlation between the simulated and predicted EWs using ANN and SVM trained with DoE and RTS set of designs. In Table V, the error metric values for the prediction accuracy as defined by (7)–(9) are tabulated for 300 test cases.

2) Adaptive Equalization: The test set consists of $d = 6$ design variables leading to $3^d = 729$ designs. The inputs are differential IL and total FEXT at 500 MHz sampling from 1 to 12 GHz, three-tap FFE coefficients, dc bias, and the DFE coefficient (51 input nodes). The ANN used is same as that for constant equalization. Fig. 25(a) and (b) demonstrates the EH and EW correlations, respectively, for the first 146 cases between the simulated and predicted values using ANN and SVM trained with DoE and RTS set of designs. In Table VI, the error metric values for the prediction accuracy as defined by (7)–(9) are tabulated.

3) Discussion on DFE Nonlinearity: A typical HSS channel has a linear FD response from the transmitter to the receiver. However, certain channels like PCIe require receiver equalization, which could have a nonlinear behavior. This raises a fundamental question as to how well the learning-based models would operate in the presence of such nonlinearity. In the PCIe Gen3 case study presented here, a single-tap DFE was used at the receiver end. Fig. 26 shows the values of EH to a sweep of the DFE tap value across a wide range, for a single channel with fixed impedances.

TABLE VI

PCIe Gen3 PREDICTION ERROR METRIC VALUES [ε, μ, AND δ AS PER (7)–(9)]: ADAPTIVE EQUALIZATION

EH	DoE Training (45 Sets)			RTS Training (118 Sets)		
Modeling Technique	ε (%)	μ (mV)	δ (mV)	ε (%)	μ (mV)	δ (mV)
ANN	2.6	1.4	5.4	1.7	0.8	5.0
SVM	3.7	1.9	5.8	1.7	2.4	8.6
EW	DoE Training (45 Sets)			RTS Training (52 Sets)		
Modeling Technique	ε (%)	μ (ps)	δ (ps)	ε (%)	μ (ps)	δ (ps)
ANN	2.1	1.1	4.7	1.1	0.5	2.8
SVM	2.1	1.1	4.9	0.9	0.4	2.3

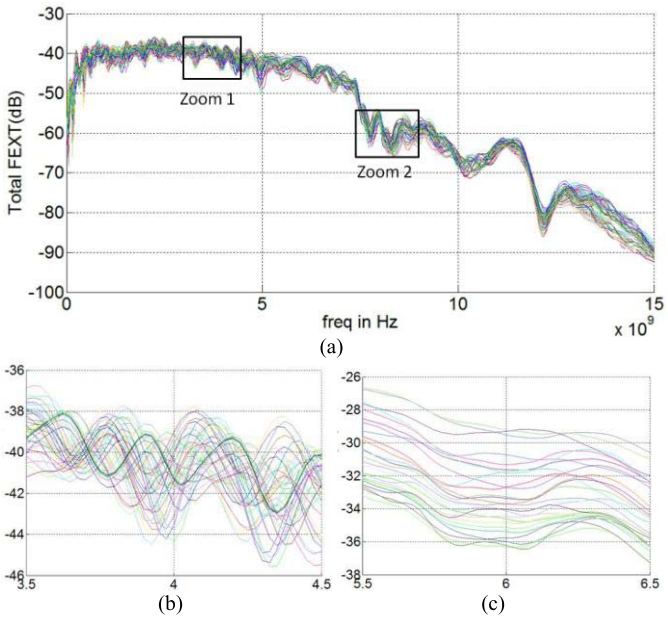
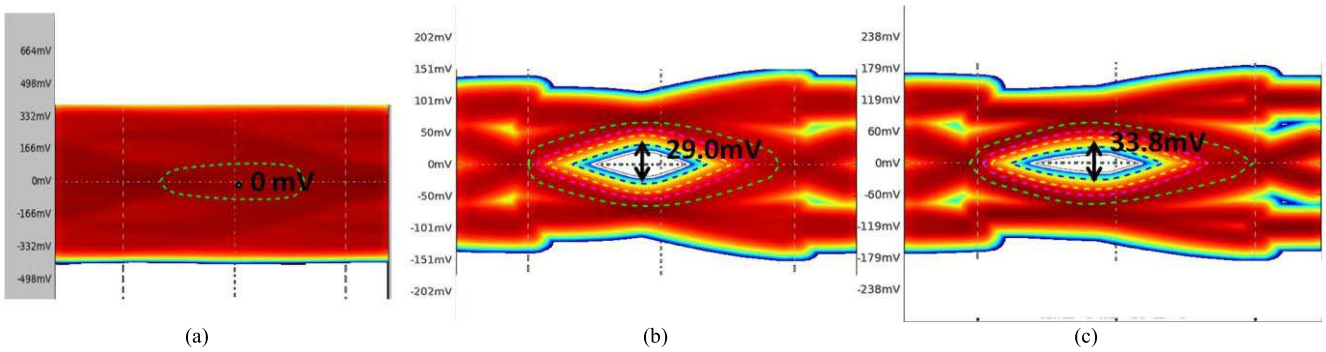


Fig. 22. (a) Total FEXT Profiles for the PCIe Gen3 channel of Fig. 20. (b) Zoom 1. (c) Zoom 2.

The EH is a nonlinear function of DFE tap value, but the prediction accuracy for cases with forced constant equalization is within 3%, as indicated in Table V. This example shows that the learning-based models work well for receiver equalization. To further strengthen the claim, the one-tap DFE was replaced with a three-tap DFE. Different tap combinations were forced for the said channel to study the behavior of EH with three-tap



(a) (b) (c)

Fig. 23. Eye diagram for the PCIe Gen3 channel of Fig. 20 (a) without equalization, (b) with constant equalization, and (c) with adaptive equalization.

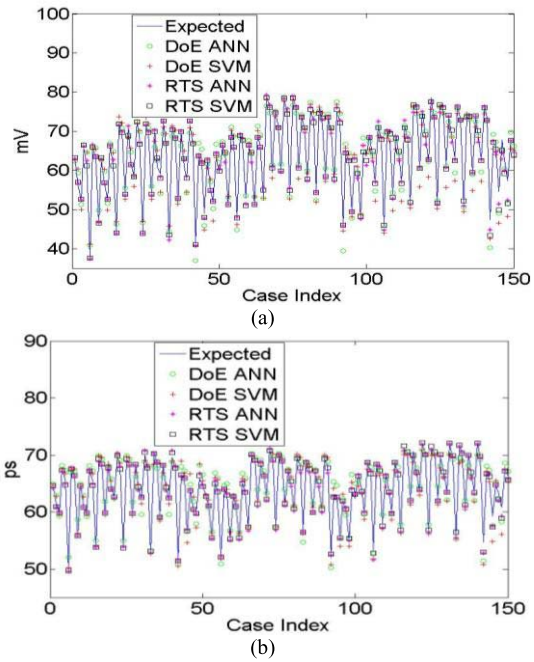


Fig. 24. Correlation of predicted data obtained from the proposed mapping methodology and simulated data obtained from TD simulations on selected designs for the PCIe Gen3 channel of Fig. 20 with constant equalization. (a) EH. (b) EW.

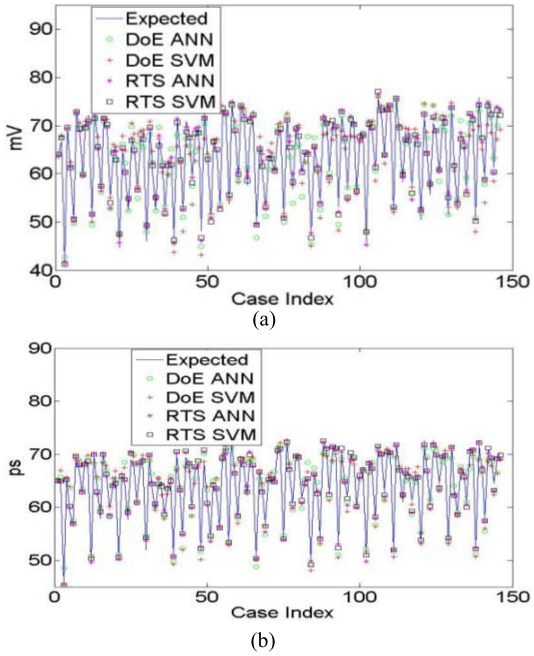


Fig. 25. Correlation of predicted data obtained from the proposed mapping methodology and simulated data obtained from TD simulations on selected designs for the PCIe Gen3 channel of Fig. 20 with adaptive equalization. (a) EH. (b) EW.

DFE coefficients. It would be expected that higher cumulative coefficient values should lead to better EH if the DFE response was linear; however, several degrees of nonlinearity were encountered in this paper. Fig. 27 shows the variation in EH for increasing cumulative tap values and different combinations of tap value arrangements.

When the DFE tap values are allowed to adapt to the received bit sequences, they are a function of the channel response and hence would lead to a highly linear behavior. However, by forcing the tap values, the DFE was made to operate in a weak nonlinear region and strong nonlinear region. The prediction accuracies for all the three possibilities are presented in Table VII.

VII. Limitations of the Methodology

The methodology of FD-to-EH/EW mapping is dependent largely on the training data used to generate and validate the learning-based models. Three common cases where the methodology could result in inaccurate prediction have been listed here.

The first case is when the number of training sets used to generate the learning-based model is insufficient. Fig. 28 shows the drop in prediction accuracy with a decreasing number of training sets for the PCIe Gen3 channel presented in Section VI. This paper outlines the two methods used to generate training sets and the impact they have on model generation time and accuracy in Section V.

The second case where prediction accuracy drops is when the *S*-parameter data are undersampled or the number of frequency points selected for FD-to-EH/EW mapping is insufficient. Fig. 29 shows the drop in accuracy with coarser sampling in FD. The numbers on the plot indicate the number of frequency points for the corresponding discretization value. The data given are for the PCIe Gen3 channel presented in Section VI.

The third case is when the TD simulation settings are different for generating the model than those used for testing

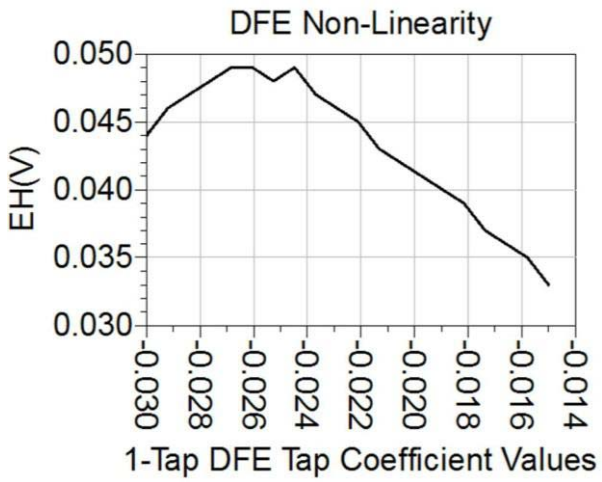


Fig. 26. Nonlinearity in EH with respect to the DFE tap coefficient value for the one-tap DFE.

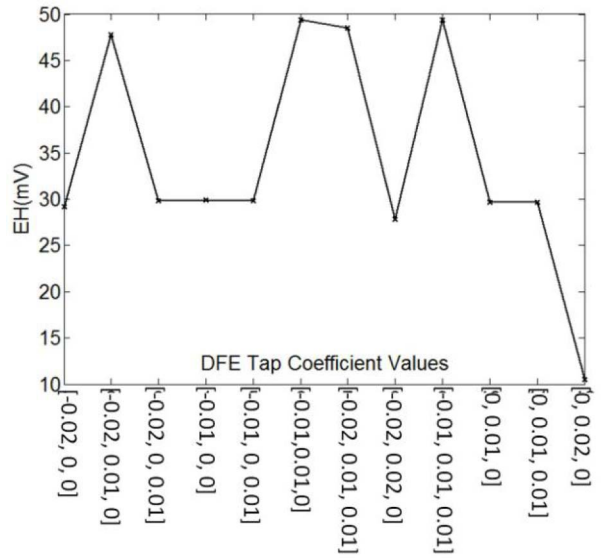


Fig. 27. Nonlinearity in EH with respect to sets of DFE tap coefficient values for the three-tap DFE.

TABLE VII
PREDICTION ACCURACY WITH DIFFERENT DEGREES OF DFE NONLINEARITY

Case : Three Tap DFE	Prediction Error (%)
Adaptive Equalization	2.2
Constant Equalization (Weak Non-Linearity)	1.8
Constant Equalization (Strong Non-Linearity)	2.3

the model. Simulation settings may correspond to rise time, pulse width, and voltage swing of the input bit stream. In addition, changes in channel topology like length of interconnects, package model impedances, connector/via models, which have not been accounted for in the model generation phase, may lead to inaccurate prediction.

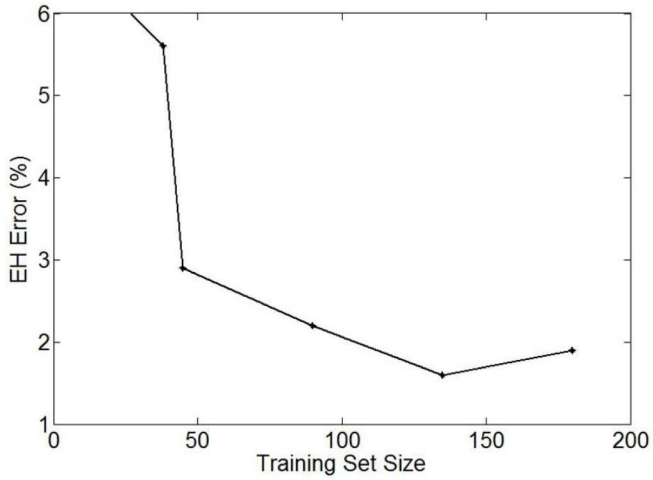


Fig. 28. Breakdown scenario: increasing error in prediction using the proposed mapping methodology for a smaller training data set size.

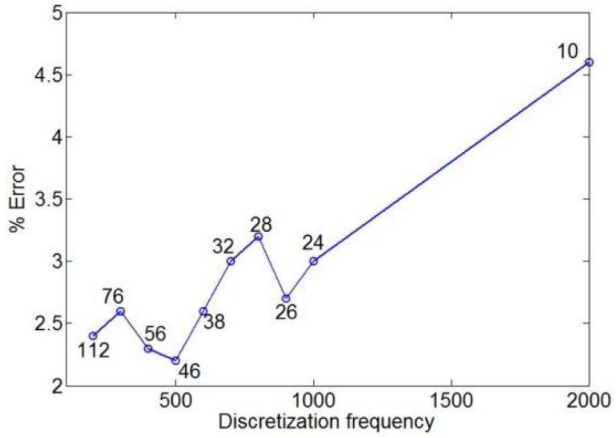


Fig. 29. Breakdown scenario: increasing error in prediction using the proposed mapping methodology for coarser sampling in frequency points.

VIII. CONCLUSION

A DoE-based design space analysis may be insufficient to capture sensitivity in a complex nonlinear design space for HSS channels. A full-factorial TD analysis is prohibitive due to excessive time requirements. Hence, modeling TD metrics, EH/EW, from FD *S*-parameters can be used for accurate EH/EW prediction. Learning-based modeling algorithms like MLP-based ANN and LS-SVM for space mapping from FD to EH/EW are demonstrated to work well for two different interfaces, SATA 3.0 and PCIe Gen3. A reduced training set algorithm is also proposed to bind the time for training the space-mapping network. The numerical results are presented to demonstrate the accuracy and speedup of the proposed methodology.

REFERENCES

[1] J. I. Alonso, J. Borja, and F. Perez, "A universal model for lossy and dispersive transmission lines for time domain CAD of circuits," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Boston, MA, USA, Jul. 1991, pp. 991–994.

[2] P. A. Perry and T. J. Brazil, "Forcing causality on S-parameter data using the Hilbert transform," *IEEE Microw. Guided Wave Lett.*, vol. 8, no. 11, pp. 378–380, Nov. 1998.

[3] C. W. Helstrom, "Calculating error probabilities for intersymbol and cochannel interference," *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 430–435, May 1986.
[4] B. K. Casper, M. Haycock, and R. Mooney, "An accurate and efficient analysis method for multi-Gb/s chip-to-chip signaling schemes," in *Symp. VLSI Circuits Dig. Tech. Papers*, Honolulu, HI, USA, Jun. 2002, pp. 54–57.
[5] V. Stojanovic and M. Horowitz, "Modeling and analysis of high-speed links," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2003, pp. 589–594.
[6] A. Sanders, M. Resso, and J. D'Ambrosia, "Channel compliance testing utilizing novel statistical eye methodology," in *Proc. DesignCon*, 2004, pp. 1–25.
[7] K. Xiao, B. Lee, and X. Ye, "A flexible and efficient bit error rate simulation method for high-speed differential link analysis using time-domain interpolation and superposition," in *Proc. IEEE Int. Symp. Electromagn. Compat.*, Detroit, MI, USA, Aug. 2008, pp. 1–6.
[8] D. Oh, J. Ren, and S. Chang, "Hybrid statistical link simulation technique," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 1, no. 5, pp. 772–783, May 2011.
[9] Z. Chen, W. D. Becker, and G. Katopis, "A new approach to deriving packaging system statistical eye diagram based on parallel non-linear transient simulations using multiple short signal bit patterns," in *Proc. IEEE-ECTC*, San Diego, CA, USA, May/Jun. 2012, pp. 160–167.
[10] K. Rekab and M. Shaikh, *Statistical Design of Experiments With Engineering Applications*. Boca Raton, FL, USA: CRC Press, 2005.
[11] B. Mutnury, F. Paglia, M. Wang, G. K. Singh, and A. C. Scogna, "Modeling and characterization of high speed interfaces in blade and rack servers using response surface model," in *Proc. IEEE-ECTC*, May/Jun. 2011, pp. 183–190.
[12] W. T. Beyene, "Application of artificial neural networks to statistical analysis and nonlinear modeling of high-speed interconnect systems," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 26, no. 1, pp. 166–177, Jan. 2007.
[13] R. Mellitz, V. Ragavassamy, and M. Brownell, "Fast, accurate, 'simulation-lite' approach for multi-GHz board design," in *Proc. DesignCon*, 2011. [Online]. Available: http://www.designcon.com/santaclara/downloads/DC11-Paper-Index.pdf
[14] N. Ambasana, D. Gope, B. Mutnury, and G. Anand, "Eye-height/width prediction from S-parameters using bounded size training set for ANN," in *Proc. IEEE-EDAPS*, Bengaluru, India, Dec. 2014, pp. 17–20.
[15] N. Ambasana, B. Mutnury, D. Gope, and G. Anand, "Application of artificial neural networks for eye-height/width prediction from S-parameters," in *Proc. IEEE-EPEPS*, Portland, OR, USA, Oct. 2014, pp. 99–102.
[16] N. Ambasana, D. Gope, B. Mutnury, and G. Anand, "Automated frequency selection for machine-learning based EH/EW prediction from S-Parameters," presented at the 24th Int. Conf. Elect. Perform. Electron. Packag. Syst. (EPEPS), San Jose, CA, USA, Oct. 2015, pp. 53–56.
[17] J. W. Bandler, R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 42, no. 12, pp. 2536–2544, Dec. 1994.
[18] M. H. Bakr, J. W. Bandler, K. Madsen, and J. Søndergaard, "An introduction to the space mapping technique," *Optim. Eng.*, vol. 2, no. 4, pp. 369–384, 2001.
[19] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory*, New York, NY, USA, 1992, pp. 144–152.
[20] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002.
[21] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1999.
[22] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, Nov. 1994.
[23] M. Bashiri and A. F. Geranmayeh, "Tuning the parameters of an artificial neural network using central composite design and genetic algorithm," *Sci. Iranica*, vol. 18, no. 6, pp. 1600–1608, Dec. 2011.
[24] W. Hendrickx and T. Dhaene, "Sequential design and rational metamodelling," in *Proc. Winter Simulation Conf.*, Dec. 2005, pp. 290–298.
[25] J. P. C. Kleijnen and W. C. M. van Beers, "Application-driven sequential designs for simulation experiments: Kriging metamodelling," *J. Oper. Res. Soc.*, pp. 876–883, Aug. 2004.
[26] D. Gorissen, "Grid-enabled adaptive surrogate modeling for computer aided engineering," Ph.D. dissertation, Dept. Comput. Sci., Univ. Antwerp, Antwerp, Belgium, 2010.
[27] *SUMO (SUrrogate MOdeling) Lab*. [Online]. Available: http://www.sumowiki.intec.ugent.be/General_guidelines

Nikita Ambasana (S'13) received the B.E. degree in electronics and communication engineering from the Dharmsinh Desai Institute of Technology, Dharamsinh Desai University, Nadiad, India, in 2009, and the M.Tech. degree in information and communication technology from the Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, India, in 2012. She is currently pursuing the Ph.D. degree with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India.

Her current research interests include methods of using neural networks for faster design space analysis for signal integrity of high speed channels and inverse imaging algorithms for channel design.

Gowri Anand received the M.Sc. degree in electrical engineering from Arizona State University, Tempe, AZ, USA, in 2002.

She is currently a Senior Electrical Engineer with the Enterprise PG Group, Dell India, Bangalore, India, where she has been serving as the Storage and Server Signal Integrity Leader for SATA technologies for the past four years. She has over six years of experience in electrical modeling and simulation of package designs. She is focused on analyzing SATA channels in rack and blade server systems with Dell India. She has authored four IEEE conference papers and holds two patents with Dell India.

Bhyrav Mutnury (SM'11) received the M.Sc. and D.Phil. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2002 and 2005, respectively.

He was involved in numerous research projects with the Georgia Institute of Technology, the Missouri Institute of Science and Technology, Rolla, MO, USA, and Pennsylvania State University, State College, PA, USA. From 2000 to 2005, he was a Research Assistant with the Department of Electrical and Computer Engineering, Georgia Institute of Technology. His work was focused on modeling and analyzing high speed power distribution networks and macro-modeling of active circuits to cut down simulation time by 10–100X. He was with the IBM System x Server Group, Austin, TX, USA, from 2005 to 2010, where he created an electrical design, modeling, and analysis algorithms that can reduce the modeling, simulation, and analysis of electrical interfaces anywhere from 100–1000x. He is currently a Distinguished Engineer and Global Team Leader with the Enterprise Signal Integrity Group, Dell USA, Austin, where he is responsible for storage, network, rack, and blade server designs. He is driving the next-generation high speed interfaces and modeling methodologies with Dell Inc. His innovative design of experiments-based methodologies and integrated tool suite to model high speed SerDes for electrical and physical designs have provided huge improvement in productivity. He is the Electrical Interface Expert on Gigabit Ethernet, X Attachment Unit Interface, Fibre Cable, Universal Serial Bus, PCIe, Quick Path Interconnect, Double Data Rate, and Fully Buffered DIMM. He has over 15 years of progressive experience in system design with strong focus on electrical modeling, analysis, and optimization of complex high speed servers. He has authored or co-authored over 55 refereed publications in various IEEE and non-IEEE conferences. These publications covered various disciplines, including package and interconnect modeling, analysis, and optimization; active circuit and transistor level circuit macromodeling; high speed serial and multidrop interface design and modeling; and services science and management engineering. He has filed more than 105 invention disclosures in the fields of electrical cable design, package and printed circuit board design, and electrical design space exploration.

Dr. Mutnury was a Secretary for the Austin Component Packaging and Manufacturing Technology Chapter from 2007 to 2012.

Dipanjan Gope (M'05–SM'12) received the B.Tech. degree in electronics and electrical communication engineering from IIT Kharagpur, Kharagpur, India, in 2000, and the M.S. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, WA, USA, in 2003 and 2005, respectively.

He was a Senior CAD Engineer with Intel, Hillsboro, CO, USA, from 2005 to 2007. He was a Founding Member with Nimbic, Inc., Stafford, TX, USA, a start-up pioneering in cloud-based 3-D full-wave electromagnetic simulation and acquired by Mentor Graphics, Inc., Wilsonville, OR, USA, in 2014, where he served as the Vice President of Research and Development from 2007 to 2011 and from 2013 to 2014. As part of the Ph.D. dissertation, he co-developed the PILOT technology for the fast circuit-electromagnetic simulation, now licensed by the University of Washington. He has been an Assistant Professor with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India, since 2011. He has authored over 50 journal and conference publications. His current research interests include computational electromagnetics with applications in signal integrity, power integrity, electromagnetic interference for high speed chip-package-systems, antenna analysis and design, fast solver algorithms, parallel programming for many-core, and cloud computing.

Dr. Gope received the Divisional Award from Intel for contributions toward providing signal integrity solution to Intel architecture platforms. He has served on the Technical Program Committee of many international conferences, including the IEEE Design Automation Conference, the IEEE International Conference on Computer-Aided Design, the IEEE Design, Automation and Test in Europe Conference and Exhibition, and the IEEE Asia and South Pacific Design Automation Conference. He was the General Co-Chair of the 2014 IEEE Electrical Design of Advanced Packaging and Systems.