Transmitter and Receiver Equalizers Optimization Methodologies for High-Speed Links in Industrial Computer Platforms Post-Silicon Validation

Francisco E. Rangel-Patiño¹,², José E. Rayas-Sánchez¹, and Nagib Hakim²

¹Department of Electronics, Systems, and Informatics, ITESO – The Jesuit University of Guadalajara
Tlaquepaque, Jalisco, 45604 Mexico
²Intel Corp.
e-mail: francisco.rangel@intel.com

Abstract — As microprocessor design scales to nanometric technology, traditional post-silicon validation techniques are inappropriate to get a full system functional coverage. Physical complexity and extreme technology process variations introduce design challenges to guarantee performance over process, voltage, and temperature conditions. In addition, there is an increasingly higher number of mixed-signal circuits within microprocessors. Many of them correspond to high-speed input/output (HSIO) links. Improvements in signaling methods, circuits, and process technology have allowed HSIO data rates to scale beyond 10 Gb/s, where undesired effects can create multiple signal integrity problems. With all of these elements, post-silicon validation of HSIO links is tough and time-consuming. One of the major challenges in electrical validation of HSIO links lies in the physical layer (PHY) tuning process, where equalization techniques are used to cancel these undesired effects. Typical current industrial practices for PHY tuning require massive lab measurements, since they are based on exhaustive enumeration methods. In this paper, direct and surrogate-based optimization methods, including space mapping, are proposed based on suitable objective functions to efficiently tune the transmitter and receiver equalizers. The proposed methodologies are evaluated by lab measurements on realistic industrial post-silicon validation platforms, confirming dramatic speed up in PHY tuning and substantial performance improvement.

Index Terms — ANN, Broyden, channel, crosstalk, CTLE, DoE, equalization, Ethernet, eye diagram, FIR, HSIO, ISI, jitter, Kriging, metamodels, optimization, PCIe, post-silicon validation, receiver, SATA, SFP, signal integrity, space mapping, surrogates, system margining, transmitter, tuning, USB.

I. INTRODUCTION

Technology scaling and advanced silicon packaging techniques are allowing high density integration. However, as process technologies scale down to nanometric dimensions, traditional IC design methods are challenged by the problem of increased silicon process variation. The combined effects of increased product complexity, performance requirements, and time-to-market (TTM) commitments have added tremendous pressure on post-silicon validation [1].

A significant portion of the circuits to be validated in modern microprocessors corresponds to high-speed input/output (HSIO) links. Undesired effects such as jitter, inter-symbol interference (ISI), crosstalk and others, can create multiple signal integrity problems in HSIO circuits, making maximum bus speeds difficult to achieve in practice. This problem is aggravated by the fact that channel speeds keep increasing from one generation bus technology to the next one. This is of particular concern for HSIO interfaces, such as Peripheral Component Interconnect Express (PCIe), Serial Advanced Technology Attachment (SATA), Universal Serial Bus (USB), and Ethernet interfaces.

Training algorithms and physical layer (PHY) tuning are two important components in modern HSIO links. PHY tuning knobs are usually embedded in the I/O links, and can be digitally tuned to appropriate values. Considering the large die-to-die process variations, as well as the typical fluctuations in operating conditions, board impedance, channel loss, and different add-in cards/DIMMs, the performance of HSIO links can exhibit large variation. PHY tuning provides a way to reconfigure I/O links to cancel various fluctuations. However, it is usually unknown in pre-silicon which configuration gives the overall best performance, becoming necessary to search for “optimal” PHY tuning knob configurations. Most current industrial practices to perform PHY tuning consist of exhaustive enumeration methods, turning them into the most time-consuming processes in post-silicon validation [2].

This paper presents several optimization techniques based on novel objective functions to optimize the transmitter (Tx) and receiver (Rx) equalizers in a server post-silicon validation platform. This paper essentially summarizes [2], [11] and [25].

The rest of the paper is organized as follows. Section II presents an overview on PHY tuning. Section III describes a holistic optimization approach that merges system margining and jitter tolerance measurements to optimize the Rx equalizer. An efficient optimization methodology is proposed in Section IV to find out the optimal coefficients for a reconfigurable finite impulse response (FIR) filter used on a Tx Ethernet interface. In Section V, a new optimization methodology is proposed to find optimal coefficients for the Tx and Rx in a PCIe equalization (EQ) process. In Section VI, a neural modeling approach is described to efficiently simulate the silicon equalizer Rx. In Section VII, the Broyden-based input space mapping algorithm is exploited to optimize the PHY tuning Rx equalizer. Finally, in Section VIII are discussed the overall results of the proposed techniques.
II. POST-SILICON PHY TUNING

As mentioned before, modern process technologies introduce large silicon process variation. Different techniques exist to maximize yield based on statistical design for analog circuits, and these techniques usually fall into two categories: design-time optimization and post-silicon tuning [3].

Design-time optimization techniques explore the design space at system- and device-level to maximize the yield. However, accurate simulation models for the complete system are computationally very expensive. On the other hand, post-silicon tuning has been widely adopted to confront the silicon process variation. Tunable elements are proposed to adjust the analog circuit performance after chip fabrication [4], [5], allowing to reconfigure I/O links to cancel the effects of system channels’ variability [6]. PHY tuning settings include: parameters of an equalizer at the Tx, Rx, or both; the clock and data recovery circuit settings; variable gain amplifiers; baud-spaced FFE in the Tx, and the bias voltages or currents values, among others [7]. A typical system may have hundreds of combinations of EQ parameter values. Finding the optimal PHY settings that guarantee the bit error rate (BER) required by an industrial specification is called PHY tuning.

III. HSIO RECEIVER EQUALIZATION BY SURROGATE BASED OPTIMIZATION

To perform PHY tuning at the Rx, either Rx eye diagram margins [8] are measured and optimized, or jitter tolerance (JTOL) tests [9] are executed until measurements comply with the link specifications. Next, a trade-off analysis is done to arrive at a single set of EQ values that satisfy both test scenarios. In [2] we present a holistic approach to concurrently optimize Rx system margins and JTOL.

A. System Test Setup

The proposed holistic methodology was tested in a post-silicon industrial environment, using an Intel server platform (see Fig. 1), comprised mainly of a host central processing unit (CPU) and a platform controller hub (PCH). Within the PCH, our methodology was tested on a USB3 Gen 1 HSIO link [10].

A new test setup was designed to combine both types of measurements, as shown in Fig. 2. We stress the Rx with a BER tester, sending a USB3 compliant pattern including all jitter impairments as per specification. The channel configuration is set as “far-end” using test fixtures and a 3-meter cable. The host computer is capable of accessing Rx knobs and sending commands to the BER tester in order to increase the jitter amplitude and frequencies. Then, we measure the system margins based on a process called system margin validation (SMV) [11], which is a methodology to assess how much margin is in the design with respect to silicon processes, voltage, and temperature, by using an on-die test circuitry. We sweep the jitter amplitude at the specification frequencies to obtain JTOL results. The pass/fail criterion is given by the specification limits, known as JTOL mask.

B. Objective Function Formulation

Let \( R_{m} \in \mathbb{R}^2 \) denote the electrical system margins response, consisting of the width \( e_w \in \mathbb{R} \) and height \( e_h \in \mathbb{R} \) of the functional eye diagram,

\[
R_{m} = R_{m}(x, \psi, \delta) = \left[ e_w(x, \psi, \delta), e_h(x, \psi, \delta) \right]^T
\]  

(1)

Both the eye width and height are function of the EQ knobs settings \( x \), the operating conditions \( \psi \), and the devices \( \delta \). We aim at finding the optimal knobs \( x^* \) to maximize the functional eye diagram area. However, depending on \( \psi \) and \( \delta \), the eye diagram can be decentralized with respect to the eye-width (asymmetry \( e_{w_{as}} \)), eye-height (asymmetry \( e_{h_{as}} \)) or both. Hence, the objective function must consider the asymmetries. The area of the eye diagram and the asymmetries must be scaled by weighting factors \( w_1, w_2, w_3 \) such they become comparable. Hence, an objective function is defined as

\[
u(x) = -w_1 e_w(x, \psi, \delta) e_h(x, \psi, \delta) + w_2 e_{w_{as}}(x, \psi, \delta) + w_3 e_{h_{as}}(x, \psi, \delta)
\]  

(2)

and the optimization problem for system margining is

\[
x^* = \arg \min_x u(x)
\]  

(3)
The holistic approach is realized by adding a JTOL penalty function to (3), such that we find EQ knobs settings that optimize the functional eye diagram and simultaneously satisfies the JTOL specified mask. The JTOL system response, \( R_j \in \mathbb{R} \), consists of measurements of the sinusoidal jitter amplitude,

\[
R_j = R_j(x, \psi, \delta) = S_{JA}(x, \psi, \delta)
\]

where \( S_{JA} \) is the sinusoidal jitter amplitude. The new optimization problem is then defined as

\[
x^* = \text{arg min}_x u(x) \quad \text{subject to} \quad g(x) \leq 0
\]

where \( g(x) = S_{JA} - S_{LA} \); \( S_{JA} \) is the JTOL spec mask.

We can define an objective function that covers both the electrical margaining system and the JTOL system responses,

\[
u(x) = -w_1 \left[ \epsilon_w(x, \psi, \delta) \right] + w_2 \left[ \epsilon_{na}(x, \psi, \delta) \right] + w_5 \left[ \epsilon_{na}(x, \psi, \delta) \right] + n_0^5 \left\| G(x) \right\|_2^2
\]

where \( n_0^5 \in \mathbb{R} \) is a penalty coefficient and \( G(x) \) is the JTOL penalty vector function defined as,

\[
G(x) = \text{max} \left\{ 0, g(x) \right\}
\]

IV. ETHERNET TRANSmitter EQUALIZATION BY DIRECT OPTIMIZATION

Transceiver modules, such as some Ethernet protocols like the 10-Gigabit Small Form Factor Pluggable (XFP/SFP) and Enhanced SFP (SFP+), are regulated by specifications that ensure consistency between suppliers with requirements for eye mask measurements. These eye mask definitions specify Tx output performance in terms of voltage amplitude and time [14]. Per Ethernet IEEE standard [15], the equalization for SFP+ Tx may be accomplished with a feedforward equalizer (FFE) 3-tap FIR filter. The filter response can be adjusted by controlling the tap number and coefficients values.

Several FIR filter coefficients optimization techniques have been reported [16, 17, 18]; however, all of them are applied at design simulation level. SFP+ Tx FIR filter is not self-adaptive, and then PHY tuning is required during post-silicon validation, being the current practice based on exhaustive enumeration methods.

In [19], we propose a simple yet efficient optimization technique for a reconfigurable FIR filter used in a SFP+ Tx, by defining an effective objective function and by using direct numerical optimization in a post-silicon validation platform.

D. Results

The Rx knobs settings obtained through the optimization process were verified by measuring both the Rx inner eye height/width and jitter tolerance of the PCH. The optimized knobs setting showed an improvement of 175% on eye diagram area as compared to the initial knobs setting, and a 34% improvement as compared with the traditional (tradeoff) approach, as shown in Fig. 3. Similarly, the jitter tolerance results showed a substantial improvement with margins well above the specification limit template, as seen in Fig. 4. The efficiency of this approach was also demonstrated by a significant time reduction on post-silicon validation. While the traditional process requires days for a complete optimization, the method proposed here can be completed in a few hours. The technique can easily be applied to other interfaces such as SATA and PCIe, as demonstrated in [2].
B. System Measurements

The definition of eye height is derived from computing the difference between the inner 3σ points on the inside of the histograms of the one and zero levels, as shown in Fig. 6, where σ is the standard deviation of the histograms. The eye width is essentially the effective distance between the inner two 3σ points on the time histograms. To compute jitter, time variances of the rising and falling edges of an eye diagram at the crossing point are captured (see Fig. 6). The time histogram is analyzed to determine the amount of jitter. The peak-to-peak jitter is defined as the full width of the histogram, meaning all data points present.

C. Objective Function Formulation and Optimization

Let $R_E \in \mathbb{R}^3$ denote the signal integrity system response, which consists of the eye amplitude histogram mean high $h_{uid}$, the histogram mean low $h_{lid}$, and the total jitter $J_T$ on the eye diagram.

$$R_E = [h_{uid}(x, \psi), h_{lid}(x, \psi), J_T(x, \psi)]$$  \tag{8}

$R_E$ is a function of the PHY tuning settings $x \in \mathbb{R}^N$ (FIR tap coefficients) and the operating conditions $\psi$. The eye height $e_h \in \mathbb{R}$ is obtained from $e_h(x, \psi) = h_{uid}(x, \psi) + 3\sigma_h + h_{lid}(x, \psi) - 3\sigma_l$ \tag{9}

where $\sigma_h$ and $\sigma_l$ are the standard deviation of the histogram mean high and the histogram mean low, respectively.

Since we want to maximize the eye diagram, our initial objective function consists simply of $-e_h$, however, as the eye width is a function of the total jitter $J_T$, we must consider $J_T$ in the objective function formulation. $e_h$ and $J_T$ must be scaled by weighting factors $w_1, w_2, \in \mathbb{R}$ such that they become comparable. Therefore, the objective function is defined as

$$u(x) = -w_1[e_h(x, \psi)] + w_2[J_T(x, \psi)]$$  \tag{10}

The optimization problem for the signal integrity system is

$$x^* = \arg \min_x u(x)$$  \tag{11}

We now modify the optimization problem such that the optimal set of coefficients maximizes the eye diagram without exceeding the mask limits. The new optimization problem can be defined through a constrained formulation,

Fig. 5. Test setup for SFP+ Tx optimization. From [19].

$$x^* = \arg \min_x u(x) \text{ subject to } l_1(x) \leq 0, l_2(x) \leq 0$$  \tag{12}

with

$$l_1(x) = (h_{uid} + 3\sigma_h) - V_{H}^{ub}$$  \tag{13}

$$l_2(x) = V_{L}^{lb} - (h_{lid} - 3\sigma_l)$$  \tag{14}

where $V_{H}^{ub}$ and $V_{L}^{lb}$ are the eye mask specification limits: voltage high upper bound, and voltage low lower bound, respectively. A more convenient unconstrained formulation can be defined by adding a penalty term, as

$$U(x) = -w_1[e_h(x, \psi)] + w_2[J_T(x, \psi)] + \rho U(L(x))$$  \tag{15}

where $\rho \in \mathbb{R}$ is a penalty term and $L(x)$ is the eye mask limit penalty function defined as

$$L(x) = \max \{0, l_1(x), l_2(x)\}$$  \tag{16}

Our final objective function to optimize eye diagram and meet eye mask specification is

$$x^* = \arg \min_x U(x)$$  \tag{17}

We find the optimal set of FIR coefficients values $x^*$ by solving (17) with (15) using the Nelder-Mead method.

D. Results

When the FIR input signal becomes a pseudo-random bit sequence (PRBS) of length $(2^{23} - 1)$ with a 10.3125 Gbps data rate, the resultant eye diagram is shown in Fig. 7. The eye diagram is significantly distorted, with an eye height and eye width of 270 ticks and 189 ticks, respectively. Fig. 7 also shows the zero crossing points on the horizontal axis are not compressed enough, leading to high jitter measurements.

After applying the proposed optimization process, we get optimal Tx EQ coefficients in just 35 iterations. The optimized coefficients substantially improve $e_h$ and $e_w$, as shown in Fig. 8, being now 864 ticks and 257 ticks, respectively, which corresponds to a 252% improvement as compared to that one with the initial coefficients. The efficiency of this approach was also confirmed by a dramatic time reduction in post-silicon validation, from 4 days in the traditional process based on exhaustive search, to just 2 hours in the proposed method.

V. PCIe Transceiver EQ by Direct Optimization

PCIe is one of the most complex HSIO interfaces [20]. PCIe is a packet based high-speed point-to-point interconnection
technology that evolves with new computer industrial demands [21], and it is the primary interface for a CPU to connect with I/O devices.

The PCIe specification defines an adaptive mechanism for EQ to determine the optimum value of the Tx and Rx EQ coefficients within a fixed time limit. Testing every coefficients combination using an exhaustive enumeration method to find the best one is very time consuming. To speed up this selection, the current practice is to find a subset of coefficient combinations during post-silicon validation, and then program it into the system BIOS. The current industrial method to find the best subset of coefficients consists of using maps of EQ, which are intuitive visual indicators that help experienced post-silicon validation engineers to find the optimal coefficient combination by inspection.

In [22], we propose a simple yet efficient optimization methodology to find the optimal subset of coefficients for the Tx and Rx in a PCIe equalization process, here summarized.

A. Tx and Rx Equalizers

Most Tx serializer-deserializer implementations comprise a FFE 3-tap FIR filter. \( C_m, C_0, \) and \( C_p \) represent the three filter taps coefficients. The EQ topology at the Rx may be a combination of a continuous-time linear equalizer (CTLE) that works independently of the clock recovery circuit, and a decision feedback equalizer (DFE). The CTLE is a simple one-tap coefficient (\( C_t \)) continuous-time circuit with high-frequency gain boosting, whose transfer function can compensate the channel response [23].

B. Transmitter Equalization Coefficient Matrix

The values of the Tx coefficients are subjected to the following protocol constraints:

\[
|C_m| + |C_0| + |C_p| = 1 \quad \text{subject to} \quad C_0 > 0, \ C_m \leq 0, \ C_p \leq 0 \quad (18)
\]

These constraints are implemented by determining only \( C_m \) and \( C_p \), being \( C_0 \) implied by (18). Additionally, the coefficients range and tolerance are constrained by some requirements, as follows.

The coefficients must support all eleven values for the presets, and their respective tolerances, as defined by the Tx preset ratios table in the PCIe specification [21].

In order to keep the output-transmitted power constant with respect to coefficients, a full swing (FS) indicates the maximum differential voltage that can be generated by the Tx,

\[
FS = |C_m| + |C_0| + |C_p| \quad (19)
\]

The flat level voltage should always be greater than the minimum differential voltage that can be generated by the Tx, indicated as the low frequency (LF) parameter,

\[
C_0 - |C_m| - |C_p| \geq LF \quad (20)
\]

When the above constraints are applied, the resulting coefficients space may be mapped onto a triangular matrix, as shown in Fig. 9, where several EQ maps, one per \( C_i \) value, are superimposed. \( C_m \) and \( C_p \) coefficients are mapped onto the y- and x-axis, respectively. Each matrix cell corresponds to a valid combination of \( C_m \) and \( C_p \) coefficients, and \( u(x^*) \) correspond to a combination of \( C_m, C_p \) and \( C_t \) that results in an eye diagram qualified as optimum.

Three EQ maps are generated for each of \( C_i \) value, and each lane and device pairing may require one or more EQ maps. Current industrial methods, used by experienced validation engineers, consists of visually analyzing each EQ map to select the coefficients \( C_m \) and \( C_p \) for the FIR filter in the Tx, and \( C_t \) for the CTLE in the Rx, that correspond to an eye qualified as optimum. However, this has to be done by ensuring at the same time that the responses around the best \( C_m-C_p \) matrix cell are at least 80\% the value of that matrix cell (see Fig. 9). Due to the large number of EQ maps, finding the optimal subset of coefficients is usually a very challenging task.

C. Objective Function Formulation and Optimization

We aim at finding the optimal set of coefficients to maximize the functional eye diagram based on the margin response. Here we follow our work in [2] to define the corresponding initial objective function.

As described in previous section, we need to ensure the optimal system margin response is within a suitable area in the coefficients search space of the EQ map. In order to satisfy this requirement, the four margin responses around \( u(x^*) \) must be at least 80\% of the value of \( u(x^*) \), as shown in Fig. 9, where \( u_{i,j} \) are the objective function values for the \( i \)-th \( C_m \) and \( j \)-th \( C_p \) values, being \( C_m \) and \( C_p \) the vectors of Tx FIR pre-cursor and post-cursor values, respectively, and \( C_t \) is the vector of Rx CTLE coefficient values. This avoids selecting an optimal solution with a too high sensitivity.
We now modify the optimization problem such that the optimal set of coefficients maximizes the system margins response without exceeding the limit of $0.8u(x^*)$ in the vicinity. The optimization problem can be defined through a constrained formulation,

$$x^* = \arg \min_x u(x)$$

subject to  
\begin{align*}
l_1(x) &\leq 0, \quad l_2(x) \leq 0, \quad l_3(x) \leq 0, \quad l_4(x) \leq 0 \quad \text{(21)}
\end{align*}

with

\begin{align*}
l_1(x) &= u(C_{m1}, C_{p1}, \psi, \delta) - 0.8u(C_{m1}, C_{p1}, \psi, \delta) \quad \text{(22)} \\
l_2(x) &= u(C_{m1}, C_{p1}, \psi, \delta) - 0.8u(C_{m1}, C_{p1}, \psi, \delta) \quad \text{(23)} \\
l_3(x) &= u(C_{m1}, C_{p1}, \psi, \delta) - 0.8u(C_{m1}, C_{p1}, \psi, \delta) \quad \text{(24)} \\
l_4(x) &= u(C_{m1}, C_{p1}, \psi, \delta) - 0.8u(C_{m1}, C_{p1}, \psi, \delta) \quad \text{(25)}
\end{align*}

where $C_{m1}$ and $C_{p1}$ are the set of coefficients that maximize the margins response for each of the $C_i$ values.

A more convenient unconstrained formulation can be defined by adding a penalty term, as

$$U(x) = -w_1 [e_w(x, \psi, \delta) + e_h(x, \psi, \delta)] + w_2 [e_w(x, \psi, \delta)]$$

$$+ w_3 [e_u(x, \psi, \delta) + \gamma_0 |L(x)|^2]$$

(26)

where $\gamma_0 \in \Re$ is the penalty coefficient and $L(x)$ is a corner limits penalty function, defined as

$$L(x) = \max \{0, l_1(x), l_2(x), l_3(x), l_4(x)\}$$

(27)

Then, we aim at finding the optimal set of coefficients values $x^*$ by solving (17) with (26) as objective.

The combination of pattern search and the Nelder-Mead is a good approach to deal with our objective function (26) that contains many local minima. We start the optimization with pattern search, which serves for exploring the design space until finding a potential region where the global minimum is located. Then, the solution found by pattern search is used as seed for the Nelder-Mead method, which further minimizes the objective function for a more precise solution.

D. System Test Setup

The system under test is an Intel post-silicon validation platform. The PCIe link is exercised at the packet level with a protocol add-in test card which emulates the external device, as shown in Fig. 10. Measurements are based on the SMV...
perceptron (3LP) is in principle sufficient for universal approximation [26], we use a 3LP to implement our neuromodel, with \( n \) inputs (equal to the number of Rx knobs), \( h \) hidden neurons, and \( m \) outputs (number of system responses of interest). The required complexity of the ANN, determined by \( h \), depends on the generalization performance for a given set of training and testing data [27]. Following [28], we gradually increase \( h \) during training for regularization.

\[ \begin{align*}
R_{l}(x, \psi, \delta, w) &= R_{l}(x, \psi, \delta) \\
\text{The ANN main input-output relationship is denoted as} \quad R_s &= f(x) \\
\text{We aim to develop a fast and accurate ANN model for } f \text{ by training the ANN with a set of measured learning data. The learning data are pairs of } (x_l, t_l), \text{ with } L = 1, 2, \ldots, l, \text{ where } t_l \text{ contains the desired outputs or targets (obtained from measurements) for the ANN model at the } x_l \text{ inputs, with } l \text{ as the total number of learning samples. During training, we keep fixed the system at voltage/temperature (VT) nominal conditions and without changing the external devices. Under these conditions, } \psi \text{ and } \delta \text{ remain constant. Therefore, the ANN model during training is treated as} \\
R_{Ll} &= R_s(x_L, w) \\
\text{The ANN performance during training is evaluated by computing the difference between ANN outputs and the targets for all the learning samples,} \\
E_l(w) &= R_{Ll}(x_L, w) - t_l \\
\text{where } E_l \text{ is the learning error matrix. Following [28], the problem of training the ANN is formulated as} \\
w &= \arg \min_w \| E_L(w) \|_F \\
\text{To control the generalization performance while solving (32), we use } T \text{ testing base points } (x_T) \text{ not used during training. The scalar learning and testing errors are given by} \\
E_l &= \| R_{Ll}(x_L, w) - R_{lt} \|_F \\
E_T &= \| R_{Tl}(x_T, w) - R_{Tt} \|_F \\
\text{where } R_{lt} \text{ and } R_{Tt} \text{ are the output matrices of the fine and ANN model, respectively, at the } T \text{ testing base points, and } R_{Ll} \text{ is the fine model response at the } L \text{ learning base points.} \\
\text{The 3LP is trained by using the Bayesian regularization [29] method available in MATLAB Neural Network Toolbox. The algorithm for training the ANN is shown in [25]. We first define the learning ratio to split the pairs of inputs and targets into the learning and testing datasets. Then, we use a decoupling network process with initial set of inputs and outputs to compute initial weighting factors } w_0 \text{ and corresponding initial error } \epsilon_0^{old}. \text{ We start training the 3LP with just one hidden neuron } (h = 1), \text{ and calculate the corresponding learning and testing errors. We keep increasing the complexity of the ANN } (h) \text{ until the generalization performance starts to deteriorate [25].} \\
\text{C. Experimental System Configuration and DoE Approaches} \\
\text{The system under test is a server post-silicon validation platform, comprised of a CPU and a PCH. Within the PCH, our methodology was tested on two different HSIO links: USB3 Super-speed Gen 1 and SATA Gen 3. The measurement system is based on the SMV process.} \\
\text{We employ three different DoE techniques to explore the desired solution space with a reduced number of test cases. For each test case, we use seven input variables that represent Rx knobs } (n = 7), \text{ which are settings used in three main Rx circuitry blocks (CTLE, VGA, and CDR), and then we retrieve the eye measurements from the system under test. The employed DoE techniques are: 1) Box Behnken (BB), which is type of second order response surface methodology (RSM) [30], using 62 experiments; 2) orthogonal arrays (OA) [31], using an } L_3(3^8) \text{ array in order to capture non-linear effects in the objective function by only running 27 experiments; and 3) Sobol [32] low-discrepancy sequence to sample the solution space. Given the quasi-Monte Carlo sampling approach of} \]
neural model effectively simulates the actual physical measurements with a total relative error of 1.7% for the $e_w$ response and 2.5% for the $e_h$ response. In other words, the ANN metamodel is able to predict margins with up to 95% of accuracy when using equalization settings not used during the ANN training. The technique can easily be applied to other interfaces, such as USB3 Super-speed Gen 1, as shown in [25].

VII. HSIO RECEIVER EQUALIZATION BY SPACE MAPPING OPTIMIZATION

In [33], we reported how the Broyden-based input space mapping (SM) algorithm, better known as aggressive SM (ASM) [34], [35], is used for the first time in HSIO PHY tuning optimization. Our SM approach takes advantage of a coarse surrogate model developed following [36]. In our case, the fine model is a measurement-based post-silicon validation industrial platform, while the coarse model is based on a Kriging surrogate technique. Our approach is illustrated by optimizing the PHY tuning Rx equalizer settings for a SATA Gen 3 channel topology.

A. Broyden-based Input Space Mapping

SM optimization methods belong to the general class of surrogate-based optimization algorithms [37]. They are specialized on the efficient optimization of computationally expensive models. The most widely used SM approach to efficient design optimization is the ASM or Broyden-based input space mapping algorithm [35]. ASM efficiently finds an approximation of the optimal design of a computationally expensive model (fine model) by exploiting a fast but inaccurate surrogate representation (coarse model) [35]. ASM aims at finding a solution that makes the fine-model response close enough to the desired response.

B. Fine Model

Our fine model is an Intel server post-silicon validation platform, as shown in Fig. 14. Within the PCH, our methodology is applied to a HSIO link SATA Gen3 [38]. The measurement system is based on the SMV process. We follow our work in [2] to define the corresponding objective function.

We use five input variables that represent the SATA Rx PHY tuning coefficients, which are settings used in three main Rx circuitry blocks (CTLE, VGA, and CDR). $e_w \in \mathbb{R}$ and $e_h \in \mathbb{R}$ are obtained from measured parameters,

$$e_w(x, \psi, \delta) = e_{w1}(x, \psi, \delta) + e_{w2}(x, \psi, \delta)$$

$$e_h(x, \psi, \delta) = e_{h1}(x, \psi, \delta) + e_{h2}(x, \psi, \delta)$$

where $e_{w1} \in \mathbb{R}$ and $e_{w2} \in \mathbb{R}$ are the eye width-right and eye width-left measured parameters, respectively, and $e_{h1} \in \mathbb{R}$ and $e_{h2} \in \mathbb{R}$ are the eye height-high and eye height-low parameters, respectively.

C. Coarse Model

Here, we follow our work in [36] to develop a coarse
sage model for a HSIO link SATA Gen3. By using the PHY tuning setting coefficients as inputs $x$ and the corresponding eye height and width as outputs $R_c$, we select a Kriging surrogate modeling technique [12] with a Sobol [32] DoE approach with only 50 samples.

D. Objective Function

We want to find the optimal set of PHY tuning settings $x$ that maximize the functional eye diagram area. Therefore, our objective function is given by

$$u(x) = -[e_u(x, \psi, \delta)]_c e_b(x, \psi, \delta)$$

During SM optimization, both $\psi$ and $\delta$ are kept fixed.

E. ASM Optimization Results

After applying the Broyden-based input SM algorithm [35], we arrive to a space-mapped solution, $x^{SM}$, in just 6 iterations (or fine model evaluations). The set of Rx EQ coefficients contained in $x^{SM}$ makes the measured SATA Rx inner eye height and width of the PCH as open as that one predicted by the optimized coarse surrogate model. The SM solution ($x^{SM}$) found makes an improvement of 85% on the fine model eye diagram area as compared to that one with the initial settings ($x^{(0)}$), and a 33% improvement as compared to that one with the optimal coarse model solution ($x^{c*}$), as shown in Fig. 15.

The efficiency of this approach is also demonstrated by a very significant time reduction in post-silicon validation and PHY tuning Rx equalization. While the traditional industrial process requires days for a complete empirical optimization, the method proposed here can be completed in a few hours. The technique can easily be applied to other interfaces like USB and PCI express.

VIII. DISCUSSION AND CONCLUSION

In this paper, we proposed direct and surrogate-based optimization methods, including space mapping, based on suitable objective functions to efficiently tune the Tx and Rx equalizers coefficients. The experimental results, based on real industrial validation platforms, demonstrated the efficiency of the proposed methods, showing a substantial improvement as compared with the current industrial practice, and accelerating the typical required time for equalizers tuning.

A holistic optimization approach that merges system margining and jitter tolerance measurements for PHY tuning was demonstrated. The experimental results demonstrated the efficiency of the method to deliver optimal margins while ensuring jitter tolerance compliance, showing a substantial improvement for both system margins and jitter tolerance as compared with the current industrial practice, and dramatically accelerating the typical time required for PHY tuning.

It was also proposed a direct optimization approach based on a suitable objective function formulation to efficiently tune the Tx FIR filter for the Ethernet SFP+ interface. The optimized coefficients were evaluated by measuring the real eye diagram of the physical system, showing a great mitigation of the ISI effects, and accelerating the typical required time for Tx coefficients tuning.

A direct optimization approach for PCIe link equalization was also proposed based on a suitable objective function formulation to efficiently tune the Tx FIR filter and Rx CTLE EQ coefficients to mitigate ISI and other undesired channel effects, and successfully comply with the PCIe specification. The optimized EQ coefficients were evaluated by measuring the real eye diagram of the physical system, demonstrating a great mitigation of the ISI and channel effects, and significantly enhancing current PCIe Tx/Rx tuning industrial practices in post-silicon validation.

A metamodeling technique based on artificial neural networks was also presented to efficiently simulate the effects of the Rx EQ circuitry in industrial HSIO links. Through the proposed neural modeling procedure, an efficient surrogate model is found that approximates the system with a reduced set of testing and training data.
Finally, it was also described how the Bryden-based input space mapping algorithm can be used to efficiently optimize the PHY tuning Rx equalizer settings by exploiting a low-cost low-prediction Kriging model, and a measurement-based post-
silicon validation platform as the fine model. The experimental results,
based on a real industrial validation platform, demonstrated the efficiency of the method, showing a
substantial performance improvement and a dramatic acceleration of the typical required time for PHY tuning.

Ultimately, the present paper is based on the doctoral dissertation [39].

REFERENCES
[1] J. Keshava, N. Hakim, and C. Prudvi, “Post-silicon validation challenges: How EDA and academia can help,” in Proc. Design Automation Conf. (DAC), Anaheim, CA, June 2010, pp. 3–7.

[2] F. E. Rangel-Patiño, A. Viveros-Wacher, J. E. Rayas-Sánchez, I. Durón-Rosales, E. A. Vega-Ochoa, N. Hakim and E. López-Miralrio, “A holistic formulation for system margining and jitter tolerance optimization in industrial post-silicon validation,” IEEE Trans. Emerging Topics Computing, vol. 6, 2018. (published online: 29 Sep. 2017).

[3] W. Yao, Y. Shi, E. He, S. Mueck, and P. Ferguson, “Design and modeling of a 16-bit 1.5Msps successive approximation adc with non-binary capacitor array,” in Proc. Int. Great Lakes Symp. on VLSI, Washington, D. C, April 2003, pp. 161 - 164

[4] F. E. Rangel-Patiño, A. Viveros-Wacher, J. E. Rayas-Sánchez, E. A. Vega-Ochoa, I. Durón-Rosales, and N. Hakim, “A holistic methodology for system margining and jitter tolerance optimization in post-silicon validation,” in IEEE MTTS Latin America Microw. Conf. (LAMC-2016), Puerto Vallarta, Mexico, Dec. 2016, pp. 1-4.

[5] C. Gu, “Challenges in post-silicon validation of high-speed I/O links,” in Proc. Int. Conf. on Computer-Aided Design (ICCAD), IEEE/ACM, San Jose, CA, Nov. 2012, pp. 547-550.

[6] A. Viveros-Wacher et al. “SMV methodology enhancements for high speed I/O links of SoCs”, in IEEE VLSI Test Symposium (VTS), Napa, CA, Apr. 2014, pp. 1-5.

[7] Y. Fan, Z. Zilic, “Accelerating jitter tolerance qualification for high speed serial interfaces”, in IEEE Int. Symp. on Quality Electronic Design, San Jose, CA, Mar. 2009, pp. 360-365.

[8] IEEE Std 1596.1, 2016. Universal Serial Bus Revision 3.1 Specification [Online]. Available: http://www.usb.org/developers/doc

[9] F. E. Rangel-Patiño, J. L. Chávez-Hurtado, A. Viveros-Wacher, J. E. Rayas-Sánchez, and N. Hakim, “System margining surrogate-based optimization in post-silicon validation,” IEEE Trans. Microw. Theory Techn., vol. 65, no. 9, pp. 3109-3115, Sep. 2017.

[10] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, “Design and analysis of computer experiments,” Statistical Science, vol. 4, no. 4, pp. 409-423, Nov. 1989.

[11] S. N. Lophaven, H. B. Nielsen, J. Søndergaard, “Aspects of the Matlab toolbox DACE,” Internal Report IMM-TR-2002-13, Technical University of Denmark, Lyngby, 2002. (published online: 29 Sep. 2017).

[12] [ANRITSU] (2010). Understanding Eye Pattern Measurements, application note no. 11401-0053 (revision A) [Online]. Available: https://www.anritsu.com

[13] IEEE Standard for Ethernet. IEEE Standard 802.3-2015.

[14] K. S. Reddy, S. K. Sahoo, “An approach for FIR filter coefficient optimization using differential evolution,” in Int. J. Electron. Commun. (AEU), 2015, vol. 65, no. 1, pp. 101–108.

[15] M. Kunn, K. Möller, P. Zipf, “Reconfigurable FIR Filter Using Distributed Arithmetic on FPGAs,” in IEEE Int. Symp. Circuits and Systems (ISCAS), Kuala Lumpur, Malaysia, May 2013.

[16] S. Cheng, Y. C. Lai, and R. B. Wu, “Optimization of FIR filter to improve eye diagram for general transmission line systems,” in Design, Automation & Test in Europe Conf. & Exhibition, Dresden, Germany, Apr. 2010, pp. 1321-1324.

[17] I. Duron-Rosales, F. Rangel-Patino, J. E. Rayas-Sánchez, J. L. Chávez-Hurtado, and N. Hakim, “Reconfigurable FIR filter coefficient optimization in post-silicon validation platform for optical interconnects,” in Int. Caribbean Conf. Devices, Circuits, and Systems (ICCDCS-2017), Cozumel, Mexico, Jun. 2017, pp. 85-88.

[18] A. Wilen, J. Shade, and R. Thornburg, Introduction to PCI Express: A Hardware, and Softw. Developer’s Guide. Hillsboro, OR: Intel Press, 2003.

[19] PCI SIG Org. (2017), PCI Express® Base Specification Revision 4.0 Version 1.0 [Online]. Available: https://pe3.org/compspecifications.

[20] F. E. Rangel-Patiño, J. E. Rayas-Sánchez, E. A. Vega-Ochoa, and N. Hakim, “Direct optimization of a PCI Express link equalization in industrial post-silicon validation,” in IEEE Latin American Test Symp. (LATS 2018), Sao Paulo, Brazil, Mar. 2018, pp. 1-6.

[21] D. R. Stauffer, T. J. Mechler, M. Sorn, K. Dramstad, C. R. Ogilvie, A. Mohammad, and J. Rokkoh, High Speed SERDES Devices and Applications, New York, NY: Springer, 2008.

[22] Scipy.org, PCI SIG Org. (2017), Open-source Softw. for Math., Science, Eng. [Online]. Available: https://docs.scipy.org/doc/scipy/reference.

[23] F. E. Rangel-Patiño, J. E. Rayas-Sánchez, A. Viveros-Wacher, J. L. Chávez-Hurtado, E. A. Vega-Ochoa, and N. Hakim, “Post-silicon receiver equalization modeling and FPGAs at artificial neural networks,” IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, vol. 37, 2018. (published online: 8 May 2018).

[24] H. Hurtado, A. Viveros, J. E. Rayas, S. N. Lophaven, H. B. Nielsen, J. Søndergaard, “Aspects of the Matlab Optimization Toolbox,” IEEE Mind & Machine, vol. 11, no. 2, pp. 277-280.

[25] I. M. Sobol, “Monte Carlo methods for solving certain boundary value problems,” in Numerical Analysis of Computer Experiments,” in J. L. Chávez-Hurtado, E. A. Vega-Ochoa, and N. Hakim, “Electromagnetic optimization exploiting aggressive space mapping algorithm over two decades of development,” in Proc. Int. Conf. on Integrated Circuits, vol. 1, pp. 1-4.

[26] F. M. F. Rangel, O. Ochoa, and N. Hakim, “High-speed digital signal processing,” in IEEE Int. Symp. on Circuits and Systems, vol. 3, pp. 647-650, May 2009.

[27] [R. E. Wacker, J. E. Rayas-Sanchez, M. Kumm, K. Moller, P. Zipf, “Reconfigurable FIR Filter Using Distributed Arithmetic on FPGAs,” in IEEE Int. Symp. Circuits and Systems (ISCAS), Kuala Lumpur, Malaysia, May 2013.

[28] R. E. Wacker, J. E. Rayas-Sanchez, M. Kumm, K. Moller, P. Zipf, “Reconfigurable FIR Filter Using Distributed Arithmetic on FPGAs,” in IEEE Int. Symp. Circuits and Systems (ISCAS), Kuala Lumpur, Malaysia, May 2013.

[29] R. E. Wacker, J. E. Rayas-Sanchez, M. Kumm, K. Moller, P. Zipf, “Reconfigurable FIR Filter Using Distributed Arithmetic on FPGAs,” in IEEE Int. Symp. Circuits and Systems (ISCAS), Kuala Lumpur, Malaysia, May 2013.

[30] R. E. Wacker, J. E. Rayas-Sanchez, M. Kumm, K. Moller, P. Zipf, “Reconfigurable FIR Filter Using Distributed Arithmetic on FPGAs,” in IEEE Int. Symp. Circuits and Systems (ISCAS), Kuala Lumpur, Malaysia, May 2013.