Determining worst-case eye height in low BER channels using Bayesian optimization

Majid Ahadi Dolatsara, Madhavan Swaminathan
School of Electrical and Computer Engineering, 3D Systems Packaging Research Center (PRC), Georgia Institute of Technology, Atlanta, GA, USA

Abstract— **Eye diagram simulation and bit error rate (BER) estimation is an essential task in signal integrity. A lengthy time domain simulation is required for non-LTI systems where statistical methods are generally inaccurate. However, with the BER reaching less than 10^{-12}, and with exponential increase in bandwidth, this task is expected to become more challenging and exorbitantly time consuming. In particular, the concern is with inter-symbol interference (ISI) effect, which can be caused by the state of several earlier bits. Therefore, this paper suggests an optimization method to find the bit patterns causing the lowest received high symbol, and the highest received low symbol, at the sampling time point. Difference of these values can be used to estimate the worst-case eye height. The proposed approach is based on a mapping method and Bayesian optimization, which provides a significant speedup compared to the traditional transient eye. This optimization technique is capable of solving both non-linear and non-convex problems. A numerical example is provided to show performance of the proposed approach.**

Keywords—eye diagram, eye height, high speed channels, Bayesian optimization, machine learning.

I. Introduction

Eye diagram simulation is an essential tool for signal integrity analysis; however, as the bit error rate (BER) of high-speed channels drops, deriving the eye diagram becomes more challenging. In many cases, a lengthy transient simulation is required. For instance, if the BER is 10^{-12}, it is expected to see only one failure in transmission of 10^{12} bits. Thus, development of approaches that are accurate and more efficient than the transient eye analysis is necessary.

Eye diagram represents the noise and jitter picked by a signal travelling through a channel. Jitter and noise are generally unwanted time and amplitude distortions, respectively. They are divided to random and deterministic types. The latter includes the data dependent jitter and noise, which is partially caused due to the intersymbol interference (ISI). ISI happens when the pulse response of the channel is longer than a unit interval (UI) and affects the next received symbol [1]. ISI is particularly challenging to model because the memory effect can span through many bits in the pulse sequence. This matter can be evaluated with an accurate eye analysis, which quantifies several characteristics of the signal. However, engineers sometimes only consider the eye height and eye width values for a quick evaluation. Therefore, in this paper we focus on efficient evaluation of the data dependent jitter and noise caused by ISI. The proposed approach quickly determines the eye height. Evaluation of eye width is left for future work.

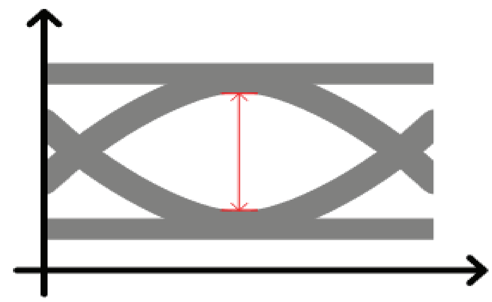


Fig. 1. Eye height is difference of the lowest received high logic and the highest received low logic at the sampling time point.

Various eye analysis methods have been suggested in the literature. Statistical eye analysis methods have been widely developed based on the channel's response and superposition [2], [3], [4], [5]. However, in general these methods are only accurate for LTI systems, or have limited applications. Nonetheless, nonlinearity can be present in high-speed channels due to non-linear IO component, compression in the receiver, or single ended signaling. Alternatively, development of surrogate models with Polynomial Chaos theory [6] and neural networks [7], [8] have been suggested to derive the full eye diagram. However, development of such models can introduce inaccuracy and still be computationally expensive. Hence, in this paper we develop an optimization approach that determines the bit pattern resulting in the worst-case eye opening. Since, only the data dependent jitter is considered, we only need to find the sequence of low and high logics resulting in the inner-most waveforms. Searching for the worst-case eye has been previously suggested in [9] and [10] using evolutionary algorithms, which can be inefficient. Moreover, [11] increases the efficiency in finding the worst-case eye by using random tree optimization. However, this method can still be computationally expensive due to curse of dimensionality which appears when many previous bits must be considered for ISI.

In the proposed approach, we initially map the search space to a one-dimensional space. The specific suggested mapping simplifies the optimization. Next, the Bayesian optimization (BO) is used to find the bit pattern resulting in the lowest received high symbol and the bit pattern resulting in the highest received low symbol, at the sampling time point. Difference of the two shows the eye height, which is illustrated in Fig. 1.

978-1-7281-3427-7/20/$31.00 ©2020 IEEE

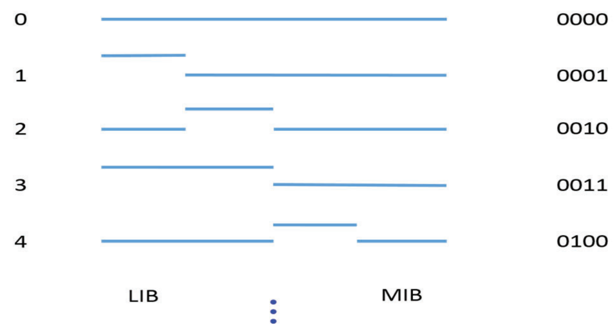


Fig. 2. Indexing of the bit patterns in the proposed approach for n=4.

II. Bayesian Optimization

Bayesian Optimization is used in the proposed approach for finding the worst-case eye height. Hence, in this section an overview of BO is provided. BO is an active learning algorithm that is proved to be effective for optimization of nonconvex and nonlinear functions [12], [13]. It is based on the Bayes theorem, in the sense that a prior distribution for a function, $f(x)$, is considered; then another sample is evaluated, and the posterior distribution is derived. This relationship is shown as:

$$P(f(x)|D_{1:t}) \propto P\big(D_{1:t}|f(x)\big)P\big(f(x)\big), \qquad (1)$$

where $D_{1:t} = \{x_{1:t}, f_{1:t}\}$ is the set of evaluated samples. A popular choice for calculating the posterior is the Gaussian Process (GP). In this method, initially a few random samples are evaluated, and the prior is set as a normal distribution, $f_t \sim \mathcal{N}(\mu, \mathbf{K})$, where μ is the mean and $\mathbf{K}$ is the covariance matrix. μ is initially set to zero, and $\mathbf{K}$ can be found using a kernel function such as the squared exponential function or its variations:

$$\begin{aligned} \mathbf{K}_{i,j} &= k\big(x_i, x_j\big), 1 \le i, j < t, \\ k\big(x_i, x_j\big) &= exp\left(\tfrac{1}{2}\left\|x_i - x_j\right\|^2\right) \end{aligned} \qquad (2)$$

Next, distribution of f_{t+1} is predicted:

$$\begin{gathered} P(f_{t+1}|D_{1:t}, x_{t+1}) = \mathcal{N}(\mu_{t+1}, \sigma_{t+1}^2) \\ \mu_{t+1} = \vec{K}^T \boldsymbol{K}^{-1} f_{1:t} \ , \ \sigma_{t+1}^2 = k(x_{t+1}, x_{t+1}) - \vec{K}^T \boldsymbol{K}^{-1} \vec{K} \end{gathered} \qquad (3)$$

where $\vec{K} = [k(x_1, x_{t+1}), k(x_2, x_{t+1}), \ldots, k(x_t, x_{t+1})]$. Finally, we need an acquisition function to find the best potential x_{t+1} sample to evaluate next. This function should provide a tradeoff between exploration and exploitation. A commonly used acquisition function is Expected Improvement (EI):

$$\begin{gathered} EI(x) = (\mu(x) - f(x^+) - \xi)\Phi(Z) + \sigma(x)\varphi(Z) \\ Z = (\mu(x) - f(x^+) - \xi)/\sigma(x) \end{gathered} \qquad (4)$$

with $\varphi(.)$ and $\Phi(.)$ being PDF and CDF of the standard normal distribution, respectively. Moreover, $f(x^+)$ is the maximum value that is observed so far. ξ is a hyper parameter that sets the tradeoff between exploration and exploitation. (4) is valid when $\sigma(x)$ is positive; however, $EI(x)$ is equal to zero if $\sigma(x)$ is zero. The next sample to be evaluated is the x that maximizes (4). The algorithm continues until converging to the global maximum or being stopped.

III. Proposed Bayesian Eye Height Approach

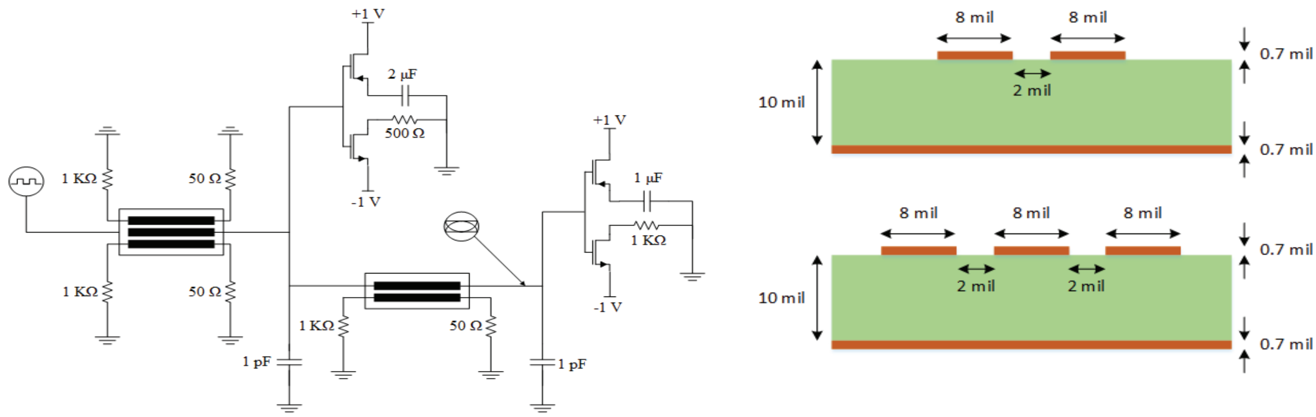
BO has been recently used for optimization problems in high-speed electronics, providing promising results [14], [15], [16]. These papers have used BO to find optimal values for physical parameters of a system. However, this paper differs from the previous studies since we have used BO to find optimal data patterns resulting in the worst-case eye height in eye analysis of a high speed channels.

For fast estimations it would be immensely useful if we had a way to identify the waveforms resulting in the worst-case eye height. This is calculated by finding the difference of the lowest received high symbol and the highest received low symbol, at the sampling time point. Since in this work we are only concerned with ISI and data dependent jitter and noise, one can use optimization techniques to find the sequences of symbols, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_n]$, that results in the waveforms passing through such points. However, this is a challenging optimization problem for the following reasons. First, it is high-dimensional because n can easily be higher than 20, which makes the optimization technique face the curse of dimensionality. Secondly, it is a discrete problem, and each variable only takes two possible values, which are zero and one. Thus, derivative based approaches are not applicable.

Therefore, we suggest reformulating the problem and mapping the search space based on our knowledge of high-speed channels and ISI. We know that ISI of each symbol decreases as it gets further from the current symbol shown by λ_n. Hence, the state of λ_1 has the least impact on noise and jitter, and the state of λ_n is the most impactful. By setting λ_1 as the least important bit (LIB) and λ_n as the most important bit (MIB), we obtain a unique index number, $I(\boldsymbol{\lambda})$, for each possible input. Note that $0 \le I(\boldsymbol{\lambda}) \le 2^n - 1$. This format is shown in Fig. 2. To find the highest low symbol, we set $\lambda_n = 0$. In addition, precursor ISI from the next bit can affect the eye height. We name the next bit λ_{n+1} and set it to 1 because, intuitively speaking, it can push the lowest high upward. Precursor ISI from further bits is often trivial, and it is not considered in this paper. Afterwards, we find $I(\boldsymbol{\lambda})|_{\lambda_n=0,\, \lambda_{n+1}=1}$ that maximizes the receiver voltage at the sampling time point. With the same logic, to find the lowest high symbol, we set $\lambda_n = 1$, and $\lambda_{n+1} = 0$. Then we find $I(\boldsymbol{\lambda})|_{\lambda_n=1,\, \lambda_{n+1}=0}$ that minimizes the receiver voltage at the sampling time point. Minimizing is done by maximizing negative of a function. Subsequently, the eye height is calculated as:

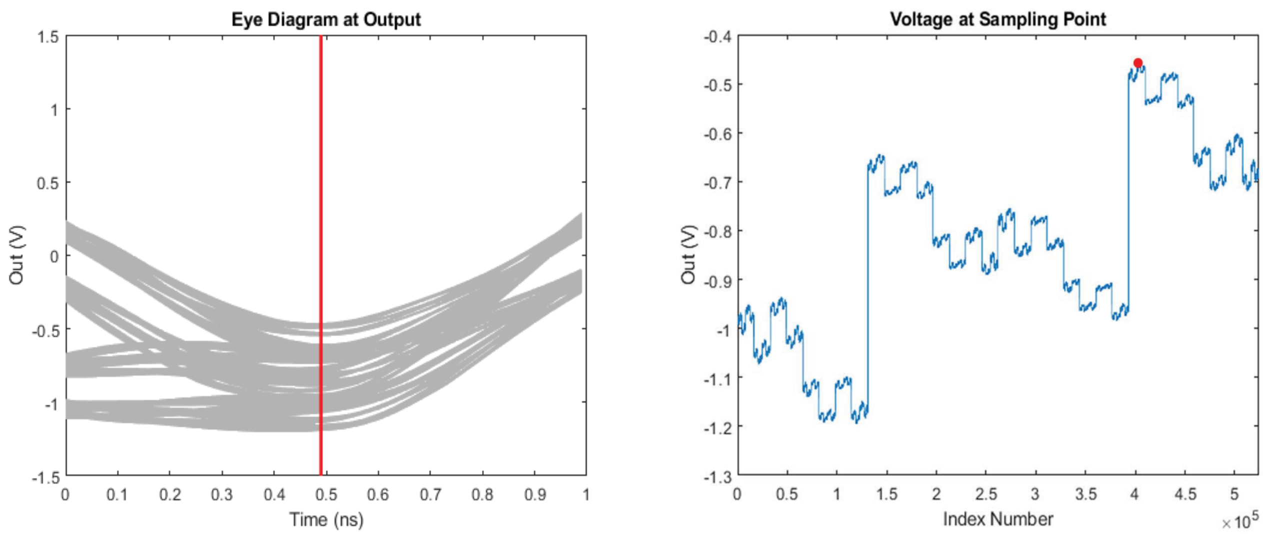
$$\begin{aligned} \mathrm{EH} = \min_I \ & V_r\big(t_s, I(\boldsymbol{\lambda})|_{\lambda_n=1,\, \lambda_{n+1}=0}\big) - \\ \max_I \ & V_r(t_s, I(\boldsymbol{\lambda})|_{\lambda_n=0,\, \lambda_{n+1}=1}) \end{aligned} \qquad (4)$$

where V_r is the receiver voltage shown on the eye diagram, and t_s is the sampling time point. We use BO for the optimization in (4). It is worth noting that the proposed mapping approach simplifies the optimization since it reduces the n-dimensional problem to a 1-dimensional problem, where the new 1-D



(a) (b)

Fig. 3. Example: A high-speed link with nonlinear terminationations and two sets of coupled microstrip lines. a) Circuit schematics. b) Cross section of the microstrip lines.



(a) (b)

Fig. 4. All the received waveforms ending in a 01 pattern. a) Waveforms ending in 01, overlayed as in the eye diagram. b) Received voltage values at the red line in (a) as a function of the index number.

objective function has a reasonable and optimizable shape. This function has such a shape because significant changes only happen when more important bits switch. Therefore, these changes do not happen close to each other. The objective function and the proposed approach are further illustrated with a numerical example in the following section.

IV. Numerical Example

To evaluate performance of the proposed approach for calculation of eye height, the system in Fig. 3 (a) is considered. This system is a high-speed link with single-ended signaling. It includes nonlinear terminations where the NMOS and PMOS transistors have Schicman-Hodges models. Moreover, two sets of microstrip lines are present in the link. Cross section of the lines is shown in Fig. 3 (b). The first set includes 3 lines with length of 4 inches, and the second set has 2 lines which are 2 inches long. Conductors of the microstrip lines are copper. Dielectric has permittivity of 4.5, and $\tan\delta = 0.02$. Furthermore, input of the channel is a random trapezoidal pulse with $V_{high} = 1$ V, $V_{Low} = -1$ V, bit rate = 1Gb/s, and rise/fall time = 100 ps. The input voltage source has an inner resistance equal to 50 Ω. Finally, the output is observed at the marked eye probe. The transient simulation is done in the circuit simulator of ANSYS Electronics Desktop 17.2, and the proposed approach is implemented in MATLAB R2018a where we have taken advantage of the Bayesian optimization tool.

We have considered the ISI effect of 19 previous symbols in this example (i.e. $n = 20$). For illustration, all the eye diagram waveforms ending in $\lambda_n = 0$, and $\lambda_{n+1} = 1$ are shown in Fig. 4 (a). Additionally, Fig. 4. (b) represents $V_r(t_s, I(\boldsymbol{\lambda})|_{\lambda_n=0,\lambda_{n+1}=1})$ at the sampling time point t_s, which is marked by a red line in Fig. 4 (a). The function shown in Fig. 4. (b) is one of the objective functions, and the goal is to find its global maximum, shown by a red dot. It is observed that, solving this one dimensional problem is easier than the original 19 dimensional problem. For Bayesian optimization, a variation of the expected

TABLE I. EYE HEIGHT RESULTS OF THE PROPOSED APPROACH AND THE TRANSIENT EYE.

	Optimum value (mV)	Eye height (mv)
Bayesian lowest high	459.55	921.73
Bayesian highest low	-462.18	
Transient eye lowest high	458.36	918.23
Transient eye highest low	-459.87	

improvement is used as the acquisition function; exploration ratio is set to 0.5, and the algorithm is stopped after 100 iterations. Two rounds of BO are performed to find the highest low and the lowest high received voltages at the sampling time point. Difference of these values shows the worst-case eye height. For comparison, we have performed a transient eye analysis with 1 million random bits. Results are compared in Table. I, which shows that the proposed approach provides good accuracy. Additionally, we have represented the transient eye, and the waveforms that pass through the lowest high and the highest low in Fig. 5, where the worst-case waveforms are shown in red. Note that the sampling point is at $t_s = 0.5\ ns$.

V. CONCLUSION

Eye analysis of high speed channels can be time consuming and computationally expensive. Therefore, in this paper a novel approach for fast calculation of the worst-case eye height without evaluation of the full eye diagram is proposed. This approach is focused on the data deterministic jitter and noise, and the ISI effect. Initially, sequential index numbers are assigned to the possible bit patterns that have nontrivial ISI. Then Bayesian optimization is used to find index numbers corresponding to the lowest high symbol and the highest low symbol, received at the sampling time point. Difference of these values shows the worst-case eye height. Finally, a numerical example is provided, which shows high accuracy of the proposed approach. Evaluation of the worst-case eye width, and integration of cross-talk is considered for future work.

ACKNOWLEDGMENT

This work has been supported by the DARPA CHIPS project under the award number N00014-17-1-2950.

REFERENCES

[1] M. P. Li, Jitter Noise and Signal Integrity at High-Speed, Upper Saddle River, NJ: Prentice Hall, 2007.

[2] B. K. Casper, M. Haycock and R. Mooney, "An accurate and efficient analysis method for multi-Gb/s chip-to-chip signaling schemes," in *IEEE Symp. Very Large Scale Integr. (VLSI) Circuits*, 2002.

[3] A. Sanders, M. Resoo and D. Ambrosia, "Channel Compliance Testing Using Novel Statistical Eye Methodology," in *DesignCon*, Santa Clara, CA, 2004.

[4] M. Tsuk, D. Dvorscak, C. S. Ong and J. White, "An electrical-level superposed-edge approach to statistical serial link simulation," in *IEEE/ACM Int. Conf. Comput.-Aided Design Dig. Tech. Papers*, 2009.

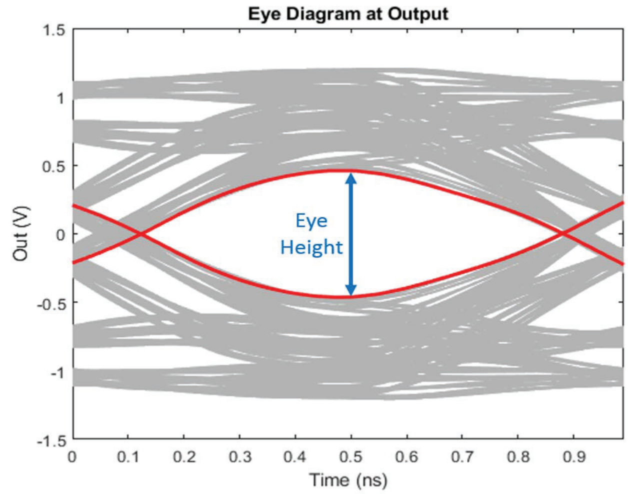


Fig. 5. Comparison of the transient eye and the calculated eye height.

[5] J. Ren and K. S. Oh, "Multiple edge responses for fast and accurate system simulations," *IEEE Trans. Adv. Packag.,* vol. 31, no. 4, pp. 741-748, Nov. 2008.

[6] M. Ahadi Dolatsara, J. Hejase, W. D. Becker and M. Swaminathan, "A Hybrid Methodology for Jitter and Eye Estimation in High-Speed Serial Channels Using Polynomial Chaos Surrogate Models," *IEEE ACCESS,* vol. 7, pp. 53629-53640, 2019.

[7] H. Yu, H. Chalamalasetty and M. Swaminathan, "Behavioral modeling of steady-state oscillators with buffers using neural networks," in *27th IEEE Conf. Elect. Perform. Electron. Packag. Syst.*, Oct. 2018.

[8] T. Nguyen and e. al., "Transient simulation for high-speed channels with recurrent neural network," in *27th IEEE Conf. Elect. Perform. Electron. Packag. Syst.*, 2018.

[9] S. M. Ulrich, A. M. Wirick, D. de Araujo, N. Pham and M. Cases, "The nittany genome project: A genetic algorithm approach to optimize a worst case bitstream for package simulation," in *IEEE Electrical Performance of Electrical Packaging*, 2003.

[10] N. Singh, B. Mutnury, N. Pham, M. Cases and C. Wesley, "Bit-pattern optimization for accurate analysis of complex high-speed interfaces," in *IEEE 58th Electronic Components and Technology Conference*, 2008.

[11] S. N. Ahmadyan, C. Gu, S. Natarajan, E. Chiprout and S. Vasudevan, "Fast eye diagram analysis for high-speed CMOS circuits," in *Design, Automation & Test in Europe Conference & Exhibition*, 2015.

[12] E. Brochu, V. M. Cora and N. De Freitas, "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," *arXiv preprint,* arXiv:1012.2599, 2010.

[13] P. I. Frazier, "A tutorial on bayesian optimization," *arXiv preprint,* arXiv:1807.02811, 2018.

[14] S. J. Park, B. Bae, J. Kim and M. Swaminathan, "Application of machine learning for optimization of 3-D integrated circuits and systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 25, no. 6, pp. 1856-1865, 2017.

[15] H. M. Torun, M. Swaminathan, A. K. Davis and M. L. F. Bellaredj, "A global bayesian optimization algorithm and its application to integrated system design," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 26, no. 4, p. 792.802, 2018.

[16] H. M. Torun and M. Swaminathan, "High-Dimensional Global Optimization Method for High-Frequency Electronic Design," *IEEE Transactions on Microwave Theory and Techniques,* vol. 67, no. 6, pp. 2128-2141, 2019.