Machine Learning Based Design Space Exploration and Applications to Signal Integrity Analysis of 112Gb SerDes Systems

Alex Manukovsky[#1], Yuriy Shlepnev[*2], Zurab Khasidashvili[#3]

[#]*Intel Inc.,*
IDC, Haifa, Israel
[1]*alex.manukovsky@intel.com* , [3]*zurab.khasidashvili@intel.com*

[*]*Simberian Inc.,*
2629 Townsgate Rd., Suite #235, Westlake Village, CA 91361, USA
[2]*shlepnev@simberian.com*

***Abstract*— This paper describes a systematic approach for system design space exploration through the application of machine learning (ML) methods for advanced system analysis. A demonstration of applying this method for signal integrity analysis, and a case study of 112Gb SerDes systems analysis based on Channel Operating Margin (COM) simulation methodology, are provided.**

I. INTRODUCTION

Design space exploration plays an extremely important role for SerDes system design, since design teams usually face a complex landing zone for SerDes products. This is due to a large number of customer use cases to be addressed, multiple constraints in the solution space for multi-protocol support, and a high degree of system variations to be covered to ensure proper operation in various configurations. This challenge holds true for both Si products and hard IP SerDes design cases.

It is often the case for the SerDes systems that a single device needs to support multiple applications and will target several markets all at once – NIC (network interface card), servers, switches, etc. Hence the device is required to properly operate within different channel topologies unique to each market segment, taking into account the variety of potential system design implementations of the end customer. Moreover, a given SerDes system will usually support multiple usage modes, short/medium/long reach communication over a variety of media types such as direct point to point channels on PCB or package, communication over backplanes and various cabling solutions of a variety of lengths: copper cables and direct attach or fiber channels. As a result, the same device is required to comply with multiple standard specifications according to its landing zone.

Even when considering a single case of the target system, the variation between channels of the same design due to implementation constrains might have a significant impact on the system performance. The same holds true for the manufacturing variation of the multiple channel components. The manufacturing process impact on package and PCB transmission lines performance can result from variation of: material properties, copper roughness, transmission lines geometry, via stubs, etc. Additional performance variation results from connectors and cables, manufacturing tolerances and Si IOs PVT impact.

To cope with this large solution space, multiple equalization (EQ) mechanisms and their capabilities and various configurations need to be explored, such as the number and range of FFE taps at the Tx, the number of DFE sliding taps for the Rx, CTLE characteristics, etc., in order to maintain an adequate system performance over the entire solution space. Moreover, the EQ mechanisms must be optimized with respect to their ability to support the target solution space, factoring all the costs, the performance, and the design aspects of the final product.

It becomes obvious that a systematic approach is required to address this challenge. In this paper we demonstrate a practical application of Machine Learning (ML) based methods for advanced design space exploration.

First, the solution space is mapped along with its multiple constraints, and multiple channel models, corresponding to the cases of interest required to cover this space, are generated with an EM simulator. Then, an investigation of the system level performance is conducted, covering channel topologies for various market segments, variation within the same segment, variation within the same design, and manufacturing tolerances. In this work IEEE 802.3 STD Channel Operating Margin (COM) methodology is used, which enables an evaluation of overall system performance as well as channel quality when used with a specification defined reference transmitter and receiver with configurable equalization capabilities. This method allows the channel designers to gain insight into their expected product quality without the need for proprietary simulators or detailed information regarding their device. Finally, we perform a design/system exploration as follows: given a response variable (an output of the design/system), we find the parameters (features, in ML terminology) having the greatest effect on the response. Moreover, we look for combinations (conjunctions) of ranges of numeric features and values of nominal features having the greatest effect on the response variable. To explore the design, some of the main questions we answer using the ML

2377-5726/21/$31.00 ©2021 IEEE
DOI 10.1109/ECTC32696.2021.00201

techniques above include the following: (a) If the response variable does not satisfy the spec, by having values outside the designated ranges, what are the parameter combinations accounting for that? (b) What are the feature ranges for which, for most samples/tests, the response is well inside the required range? First, the root-cause of the failures in the design (failure to comply with a spec or a standard) is identified, and then an insight on how to optimize the design is provided. This method is implemented on a 112Gb system case study.

This work tackles four main challenges of practical design space exploration of Ethernet (ETH) systems:

- *Generating a large quantity of link models to cover the solution space*
- *Evaluating the performance of a large quantity of links and system configuration*
- *Methodically analyzing the large volume of results*
- *Enabling an automated ML based decision support procedure to cope with system complexity and decisions based on big data*

While the design space definition remains an expert choice, the above four challenges could be solved in an automated process, as this paper will demonstrate, while providing practical examples of 112Gb C2C link.

II. Range Analysis for Decision Support Applications

The ML approach that we use for design exploration is Feature Range Analysis, or Range Analysis (RA), for short [1,2]. Range Analysis is an algorithm resembling Rule Learning (RL) [3], Rule Induction (RI) [4], and Subgroup Discovery (SD) [5]. From the algorithmic perspective, the main distinguishing feature of RA is that it heavily employs Feature Selection [6] in two basic building blocks of the algorithm: the ranking and basis procedures. The third basic building block in RA, the procedure called quality, resembles the technique used in RL, RA and SD, where the selection of rules or subgroups is done solely based on a quality function (or based on multiple quality functions). These three procedures will be explained below.

The purpose of Range Analysis is to identify combinations of ranges of numeric (or continuous) features and levels of nominal (or categorical) features that explain positive samples – samples whose characteristics and the behaviour we want to explore in the data. Binary (or dichotomous) features are a special case of nominal features with two levels, 0 and 1. For binary responses O_bin, it is conventional to encode the value of positive samples as 1 and value of negative samples as 0. For numeric responses O_num, there is no definition of positive and negative samples, but one might be interested in finding ranges where the values in the response are in its 'high range' or in its 'low range'. A high range or a low range in the response values is not defined in general via a specific threshold value. When there is a threshold k for the response values that can distinguish between high values and the rest (or between low values and the rest), it is often convenient to model numeric responses as binary by applying transformation O_bin = O_num > k (or O_bin = O_num < k, respectively). In this work we do not consider nominal responses with more than two levels as this slightly more general case can easily be reduced to (multiple instances of) the binary case. For simplicity, we will assume that there is only one response variable O in each analysis.

The RA algorithm generates range features that are most relevant for the response, where 'most relevant' might mean (a) having a strong correlation or high mutual information with the response, based on one or more correlation measures; (b) explaining part of the variability in the response not explained by the strongest correlating features; or (c) maximizing a quality function. Important examples of quality functions include the ones listed below, where Pos and N denote the counts of positive and all samples in the entire dataset, respectively, p0 = Pos/N, R denotes a range, and n(R) denotes the count of all samples within R:

- *True Positive Rate (also known as sensitivity, recall, or hit rate): TPR(R) = TP(R)/Pos, where TP(R) denotes the count of true positive samples, that is, positive samples within the range R.*
- *Predictive Positive Value (also known as precision): PPV(R) = TP(R)/n(R)*
- *The lift: Lift(R) = PPV(R)/p0*
- *Weighted Relative Accuracy [7]: WRAcc(R) = (n(R)/N)*(PPV–p0)*

For numeric responses, the counterpart of PPV(R) is the mean value of the response on samples within R, and the counterpart of p0 is the mean value of the response on all samples, thus Lift(R) and WRAcc(R) also make sense for numeric responses [8]. While positive and negative samples only make sense for binary responses, the concepts like True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) can be generalized to numeric responses as well. That allows to generalize the quality functions that are based on these concepts for numeric responses [1,2]. Each range feature is a binary feature where the value 1 on a sample is interpreted as "the sample is within the range", and the value 0 is interpreted as "the sample is outside the range".

The RA algorithm works as follows:

1. The RA algorithm first ranks features highly correlated to the response; this can be done using an Ensemble Feature Selection procedure, we refer to it as ranking procedure. In addition, RA uses the Maximal Relevance Minimal Redundancy (MRMR) procedure to select a subset of features which both strongly correlate to the response and provide a good coverage of the entire variability in the response, we refer to this procedure as basis.

2. For the nominal features selected in the first stage, or optionally, for all nominal features, from each level a binary range feature is generated thru a one-hot encoding. In a similar way, a fresh binary feature is generated for each selected numeric feature and each constructed range. These features are called single-range features. Note that an important part of all the above-mentioned algorithms (RL, RI, SD, RA) is to define candidate ranges of numeric features. RL,

RI and SD generate non-overlapping ranges while in RA the candidate ranges can be overlapping. This helps to significantly improve the accuracy of RA compared to RL, RI, SD. The RA algorithm then applies ranking and basis procedures to select the most relevant single ranges; in addition, RA selects single range features that maximize one or more quality functions. We refer to the quality-function based selection of ranges as quality.

3. For each pair of selected single range features associated with different original features, RA generates range-pair features which have value 1 on each sample where both the component single-range features have value 1 and have value 0 on the remaining samples. RA then applies again ranking, basis and quality procedures to select the most relevant range pairs.

4. Similarly, from the selected single ranges and selected range pairs, the RA algorithm builds range triplets, and applies ranking, basis and quality procedures to select most relevant ones.

In the implementation of Range Analysis in Intel's ML tool EVA [1, 2, 9, 10, 11], for practical considerations the dimensionality of the range features is limited to three (single ranges, range pairs, and range triplets). The ML experiments reported in this paper are performed with EVA.

Both the features and the responses can be systematically explored providing an answer to the following:

- What are the important features that impact the system behaviour the most?
- What are the ranges of each of the important features in which we are most likely to achieve desired system response?
- What are the combinations of feature ranges that enable us to achieve desired system response?

Simply put, we could systematically identify:

- What are the system characteristics required for achieving good performance?
- What are the system characteristics required for achieving excellent performance?
- What are the system characteristics responsible for bad performance?

This analysis can be performed for complex systems with a large amount of system variables and complex output behaviour, while bad, good or excellent can be determined by the specification, an expert opinion, or relative system performance.

Applying Range Analysis to design space exploration of system performance, factoring in a variety of operating conditions, controlled and uncontrolled factors, and multiple system configurations, allows a methodical, automated analysis of the solution space. This analysis provides a feasible way to handle the complexity of Ethernet systems, yielding the desirable insight on system behaviour comprehendible for engineers, and can be used as a decision support tool for design choices in the hands of the system architect, Si designer, Si Engineer, and more.

III. COM AS A QUALITY METRIC FOR 112 GB LINKS

The goal is to evaluate the performance of an Ethernet system over a large solution space for multiple channels, system configurations, various choices of equalization mechanisms with various capabilities, a variety of package choices, process voltage and temperature (PVT) Si characteristics, etc. The IEEE COM tool is selected for this task as the industry standards for the Ethernet protocol specifications. It allows simulating various system configurations of interest with a sufficient computational speed, and is thus applicable for dealing with the required high volume of simulations for such an analysis.

Channel Operating Margin [12-14] is a signal to noise ratio defined as

$$COM = 20 \cdot \log\left(\frac{A_{Signal}}{A_{Noise}}\right) \quad (1)$$

Where A_{Signal} is the peak signal and A_{Noise} is the peak Bit Error Rate (BER) noise defined through the peak signal minus the peak BER eye opening. Signal in this context includes all losses and dispersion in the link from chip to chip and the effect of equalization. Noise in this context includes all possible signal degradation effects with some assumptions. It includes return loss, reflections and couplings (crosstalk) as well as equalization by Tx and Rx. COM metric is computed in the time domain as the voltage ratio of signal available in a reference signalling architecture (Tx and Rx) to noise at the reference receiver's sampler – basically, it characterizes the complete link from chip to chip. The noise is calculated for the specified Detector Error Ratio (DER). DER is a generalization of BER for NRZ and SER for PAM4. Equalized single bit or symbol response (SBR or SSR) and major signal degradation factors are used to calculate the vertical slice of the eye diagram centered at the sampling point where DER is minimal. For IEEE802.3bj, bm and ck (C2C) the reference architecture is a Rx CTLE , along with Tx FFE and Rx DFE whose number taps vary, optional reference packages, and filters [12].

COM is a simulation of a reference transmitter and receiver system with a baseline equalization capability. It serves as a common reference for chip design and board design. The COM parameters represent the expected capability of a realizable PHY design. Channels that meet COM requirements are expected to work with compliant PHYs with the specified BER or better. Thus, COM is a practical metric to make decisions on materials selection, package construction, PCB construction, and SerDes design. COM can be used to budget between loss, reflections, coupling, and noise, supporting a wide range of platform configurations. Though, it may not be obvious how different features affect the COM and how to identify the ranges of features within which the design will work with high confidence. We use the machine learning algorithm to help with those decisions in this paper.

In order to demonstrate a realistic application of Range Analysis based decision support tool for Ethernet systems design space exploration, we chose the 112Gb Chip to Chip link example. The chosen system has just enough complexity

to require a decision support tool to gain a comprehensive insight on the one hand, but on the other hand could be understood by an experienced SI engineer in order to validate the findings of the demonstrated method. The link under investigation is illustrated in Fig. 1.

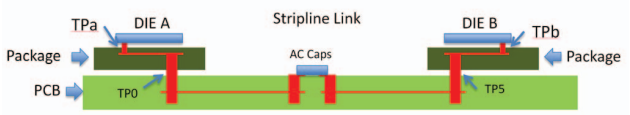


Fig. 1. Link under investigation: TP0 to TP5 are locations of ports for analysis with the reference package; TPa to TPb are locations of ports for analysis with the custom package model.

To compute S-parameters of the link Simbeor MLKit was developed for signal integrity analysis automation for generating the complete link as well as for the PCB section. We use a reference script from IEEE 802.3ck task force [12] with most of the parameters fixed to the reference values, unless defined otherwise. The input to the COM algorithm is a collection of 4 port s-parameters (s4p files). h link consists of the channel, represented by an s4p file (TROUGH), and all the relevant crosstalk s4p aggressor files (FEXT and NEXT). The channel is modelled from pin to pin or from the BGA pad to BGA pad - this includes both BGA escapes and DC blocking caps (TP0 to TP5 as shown in Fig. 1). The end-to-end channel COM is computed from a reference transmitter pad to the sampler input in the receiver. All computed S-parameters should also be suitable for the time domain conversion for 112 Gb PAM4 link, considering bandwidth and sampling requirements (covered in the next chapter).

The output of the COM script is the COM value defined by (1) that can be ranked as follows:

- *Compliant (Good) channel characteristics (COM > 3dB)*
- *Non Compliant (Bad) channel characteristics (COM < 3dB)*
- *Excellent channel characteristics (COM > 4dB)*

IV. De-compositional Analysis of 112 Gb Links

To cover the design space, multiple link models are generated. To do so, we use a hybrid de-compositional electromagnetic analysis, or the 1D+3D technique [15]. De-composition of a simple link is illustrated in Fig. 2. The link is partitioned into discontinuities and transmission line segments and uses package models defined in COM script.

1D models are built as solutions of Telegrapher's equations. 3D models are built with solution of full-wave Maxwell's equations. Both models are used to compute S-parameters of a complete link. Modal or per unit length parameters for the Telegrapher's equations (Z, Y) are computed with static or quasi-static field solver (2D problems for Laplace's equations) or an electromagnetic fields solver (3D problems for Maxwell's equations). Not only straight single line segments, but also lines with coupling, multimodal waveguides, periodic structures (BGA breakout routing) can be accurately modeled with this approach.

1D+3D Hybrid de-compositional analysis with transmission line models for traces (1D) and 3D models for discontinuities or transitions is the best technique for the serial interconnects under the localization condition, both for the design exploration and post-layout analysis. This approach usually works for PCB and packaging problems with relatively long traces, but may fail if trace segments are too short – complete 3D analysis of adjacent discontinuities is required in this case. A differential transmission line segment can be used as a very simplistic model of a link with a possible coupling to other differential links for preliminary investigations.

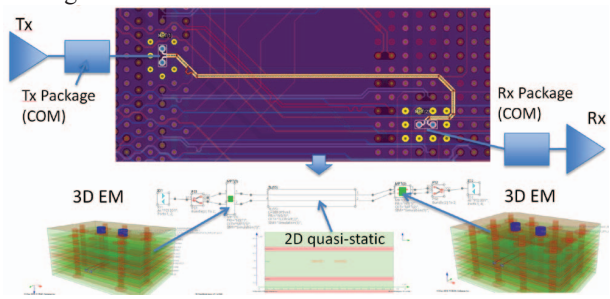


Fig. 2. Decomposition of a simple link into 1D transmission line segment models with parameter extracted with 2D quasi-static field solver and optional 3D model built with electromagnetic solver.

The accuracy of the 1D+3D approach depends on some conditions. First, proper localization of every single transition in the link is required. It is relatively difficult to do on PCB for the bandwidth of 112 Gb signals. Possible breakout of localization at very high frequencies will be neglected in this investigation. Proper de-embedding of 3D discontinuities is required to avoid artificial reflections on the boundaries between 1D and 3D models. The accuracy also depends on the availability of broadband dielectric and conductor roughness models. Such models can be identified with GMS-parameters or SPP Light techniques. Broadband dielectric models can be constructed with data available from manufacturer. However, conductor roughness models require the identification of realistic conductor roughness parameters. These parameters vary corresponding to Cu foil type and treatment procedures used for PCB manufacturing process for a specific stackup case. However, several groups of similar roughness characteristics can be identified as common for the ultra-high speed market segment. We will use statistical conductor roughness parameters previously identified in [16]. Additional necessary conditions for the accuracy of the 1D+3D approach are discussed in [17] and are not relevant to this investigation.

V. 112 Gb Links Modeling: Channel Features and Signal Degradation Factors

To assess the validity of the results obtained by our method, an understanding of the considered link model behavior is required, yielding an insight into the parameters affecting the link performance and the ranges in which these parameters vary.

There are three major groups of signal degradation factors to model for PCB and packaging interconnects: thermal losses and dispersion, reflections and couplings.

Thermal losses include dielectric polarization loss and dispersion and also conductor + conductor surface roughness loss and dispersion. We can call it thermal loss, because the useful energy of the signal is dissipated in dielectric and conductor as heat. Causal Wideband Model (Djordjevic-Sarkar) defined with two parameters (Dk and LT) at one frequency point is used to model the dielectric. The range of losses considered here is from ultra-low loss dielectric with loss tangent LT=0.001 to a medium-loss dielectric with LT=0.01. For conductor roughness modeling we use causal Huray-Bracken model with two parameters, SR and RF, identified previously in [16]. The simplest model for HVLP copper were described with SR=0.14um and RF=8.7 or with SR=0.075um, RF=24.5 – both models provide sufficient accuracy up to 50 GHz as was demonstrated in [16]. The range of the losses effect is illustrated in Fig. 3. The focus of this investigation is on how the thermal losses affect the link performance.

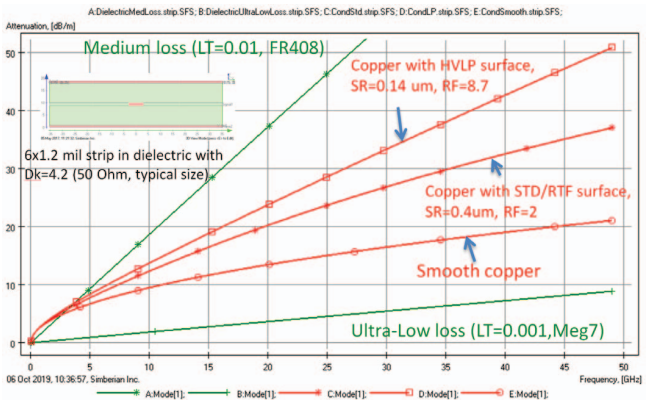
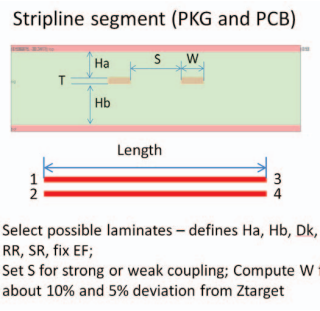


Fig. 3. Attenuation in dB/m for the range of dielectric and conductor surface treatment used in this investigation.

Reflections are the second group of the signal degradation factors included in this investigation. Trace impedance mismatch and single discontinuities (package bumps and balls, via transitions, AC caps ...) cause reflections and resonances due to multiple reflections. As a result, some energy of the signal will be reflected back to the transmitter (return loss) and some energy will propagate to the receiver with multiple reflections on the way and cause additional signal degradation due to the dispersion of the insertion loss and phase delay (usually called ISI).

Both thermal losses and reflections due to the impedance mismatch are defined by the material properties and by the geometry of the transmission line segments that define the link. Striplines are usually used for the high-speed links. Features affecting practically all electrical properties of a single stripline segment are illustrated in Fig. 4.

Considering variabilities in PCB or package manufacturing, we will use target impedance Ztarget to define the geometry of the cross-section with approximately 5% and 10% deviation from the target value.



Feature	Description
Ha	Height of dielectric layer above strips
Hb	Height of dielectric layer below strips
T	Strip thickness
W	Strip width in the middle of strip layer
S	Spacing between strips (middle)
EF	Strip etch factor – EF=0.5(Wtop-Wbot)/T
RR	Conductor resistivity relative to annealed copper
SR	Surface roughness in um for Huray-Bracken roughness model
RF	Surface roughness factor for Huray-Bracken roughness model
Dk	Dielectric constant at 1 GHz for Wideband Debye dielectric model
LT	Loss tangent at 1 GHz for Wideband Debye dielectric model
Length	Differential strip line length

Fig. 4. Features affecting loss, dispersion and reflections for a differential stripline segment and model construction.

There are two major discontinuities in the package – transition from bumps to stripline and transition from the stripline to BGA balls. The last model may include the transition from package balls to PCB stripline (BGA-PCB vias). The other possible discontinuities on PCB are PCB vias and AC coupling capacitors.

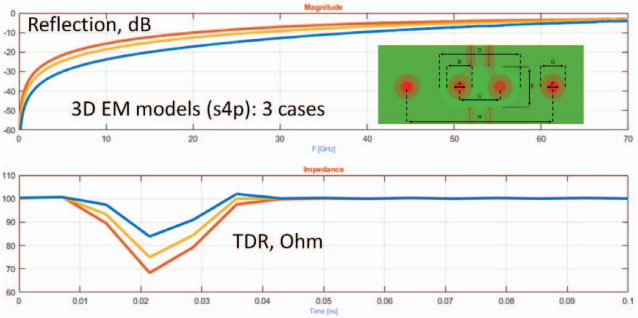


Fig. 5. Models for three possible via-hole transitions from strip to microstrip traces.

For this investigation we created S-parameter models for each discontinuity in the link based on realistic implementations. Next, the structures were optimized with respect to DFM design for manufacturing constraints of common PCB manufacturing process, as it is assumed that a significant optimization effort will be considered for realistic 112Gb link. The simplest model of a possible link will be just stripline segment with the reference package transmission line and capacitive discontinuities for the bumps and balls. Reflections from planar transitions such as bends, transitions from one cross-section to another are neglected in this investigation.

Example of 3D discontinuity models are shown in Fig. 5 with various geometry optimization levels of the same structure – not optimized, optimized and highly optimized.

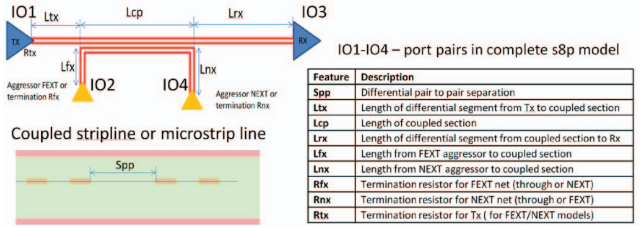
Couplings are the third group of the signal degradation factors. It includes a very broad range of physical effects listed in Table I, which can be further separated into leaks

(useful signal energy loss) and interference (unwanted energy added to the signal).

Table I. Coupling types and modeling or possible mitigation.

Coupling type	Model
Crosstalk – leaks and interference in parallel traces	INCLUDE
Via localization breakout – leaks and interference and through parallel planes and between via	Localize vias up to 40-50 GHz and do not simulate via coupling
Couplings through slots and cutouts in reference planes	Prohibit in layout
Modal transformations in diff. pairs (aka skew) – bends, asymmetry in routing, FWE	Mitigate with proper length compensation – do not simulate
Multipath propagation, radiation, EMI, EMC,...	Suppress with localization – do not simulate

For the crosstalk coupling investigation effect, a simplified model shown in Fig. 6 with the features defined in the table can be used. A cross-section of the differential stripline segments in such a link is defined as shown in Fig. 4. This is in addition to the separation parameter Spp and lengths of the coupled and un-coupled segments. For simplicity, lengths of the un-coupled segments can be set to zero. Coupled segments of the more realistic C2C link are defined with Spp and lengths. The simplest link (Fig. 4), link with coupling (Fig. 6) and C2C cases are included into Simbeor MLKit.



Feature	Description
Spp	Differential pair to pair separation
Ltx	Length of differential segment from Tx to coupled section
Lcp	Length of coupled section
Lrx	Length of differential segment from coupled section to Rx
Lfx	Length from FEXT aggressor to coupled section
Lnx	Length from NEXT aggressor to coupled section
Rfx	Termination resistor for FEXT net (through or NEXT)
Rnx	Termination resistor for NEXT net (through or FEXT)
Rtx	Termination resistor for Tx (for FEXT/NEXT models)

Fig. 6. A simplified model for the investigation of a link with crosstalk.

The output of the decompositional model for the simplified and realistic C2C structures is the complete s8p model illustrated in Fig. 6. It is used to derive s4p models for the victim link (THROUGH – IO1 to IO3), far-end crosstalk aggressor to victim S-parameters (FEXT – IO2 to IO3) and near end crosstalk aggressor to victim S-parameters (NEXT – IO4 to IO3). All IOs here are just pairs of ports terminated by specified termination resistor. Note that the THROUGH model will include the losses from leaks to the terminated aggressor link (near and far end leaks). The port numeration used here is [1 2 3 4] as illustrated in Fig. 4. Model building for transmission lines, some discontinuities and complete link analysis is automated in Simbeor MLKit that was developed for this purpose.

We should note that the key enabling technology for design space exploration is the capability to quickly and automatically construct realistic link models.

VI. Simple 112 Gb Link Case Study Results

To demonstrate the viability of the proposed analysis and to gain an initial confidence in the result of the proposed method, a case of a simple link is examined. The simplest link consist of a segment of differential stripline on PCB with all features defined in Fig. 4 and the reference package model as defined in IEEE 802.03ck specifications. The package model for the Tx and Rx will have capacitance Cd=120fF, Cp=70fF, package Zc=92.5Ω, package_tl_τ=6.14ps/mm, package_tl_γ0_a1_a2=[0 0.0009909 0.0002772]. Packages with 1mm, 5mm, 12mm and 31mm transmission line segments are investigated (assuming Tx and Rx package length are equal for each case). The number of the features in the model shown in Fig. 4 is further reduced to just six with the values defined in Table II.

Table II. DOE table for the simple stripline link investigation.

	Feature						
1	Tx_PCB_TL_S (S/H)	1	2	3			
2	Tx_PCB_TL_L (Length) [in]	0.5	1	3	6	9	12
3a	PCB_Dk (Dk)	2.8	3.2	3.5	3.8		
3b	PCB_LT (LT)	0.001	0.002	0.004	0.009		
4	PCB_TL_H (Ha,Hb) [mil]	3	5	8	10		
5	PCB_Imp [Ω]	85	90	95	100	105	
6	PKG_Len, [mm]	1	5	12	31		

The dielectric material is defined simultaneously with two parameters, PCB_Dk and PCB_LT. That choice corresponds to a usual practical selection between high-end materials with extremely low losses and a relatively low dielectric constant, and medium-loss materials with a higher dielectric constant. The PCB_TL_H feature corresponds to Ha=Hb in Fig. 4. Half ounce copper (T=0.6mil) with the conductor surface roughness defined with SR=0.075um and RF=24.5 is used [16]. Tx_PCB_TL_S is the separation between differential strips defined as the multiple of PCB_TL_H. Each analysis starts with the synthesis of strip width (W) for a given impedance PCB_Imp defined for 5 cases. The feature Tx_PCB_TL_L is the link length in inches. The total number of cases covered by Table II. is 5760 and the range of total PCB link losses is illustrated in Fig. 7. We further separated all cases into two groups – with very short package (1mm and 5mm case) and with the reference package line length (12mm and 31mm).

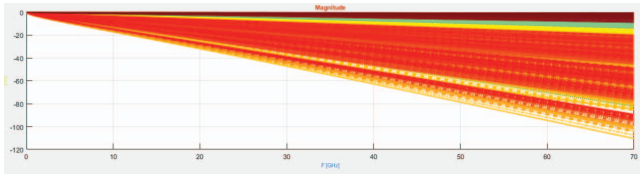


Fig. 7. Range of PCB channel losses in investigated links.

The range of the PCB channel losses in all links is illustrated in Fig. 7. It includes practically lossless extremely short 0.5in links with high-end dielectric on one end (top brown line) and 12in link with medium dielectric losses on the other end (bottom red line). This demonstrates the variety of channels considered for this analysis.

First, we evaluate the performance of the described system with package length of 12 and 31 mm representing a medium to long package length in the ETH market over the entire population of representative PCB channels defined by the DOE. The question is what are the system characteristics required for achieving excellent performance, COM > 4 dB.

The results of the Range Analysis with EVA are shown in Tables III, IV, V and VI. In these and other similar tables, the column "selection" specifies one or more methods that selected that original feature or range. The selection value "correlation" corresponds to the procedure called ranking in the Range Analysis algorithm as described in Section 2. Selection value "coverage" corresponds to the procedure basis of Range Analysis algorithm, and selection value "target" corresponds to selection based on the procedure called quality in the Range Analysis algorithm. These names – correlation, coverage, and target – were chosen to reflect the intuition behind the respective procedures, for the users who are not experts in Machine Learning and are not familiar with the details of the Range Analysis algorithm.

The first step would be to establish what are the important features having the greatest impact on the system behavior the most, and to rank them according to importance. The future ranking based on correlation or by coverage is presented in Table III. It is in good agreement with feature importance ranking based on range analysis of a single feature maximizing a quality function of Max lift presented in Table IV.

Table III. Important single range features based on having a strong correlation or high mutual information with the response (selection method :correlation) or by explaining part of the variability in the response not explained by the strongest correlating features (selection method: coverage).

feature	score	selection
Pkg_len_RX	1.0000	correlation-coverage
Tx_PCB_L	0.3949	correlation-coverage
PCB_lmp	0.2309	correlation-coverage
PCB_Dk	0.0859	correlation-coverage
PCB_H	0.0529	correlation-coverage
Tx_PCB_DS	0.0079	coverage
Tx_PCB_S	0.0011	coverage

Table IV. Important single range features for system with package length of 12 and 31 mm.

	feature	max score	max lift
0	Pkg_len_RX	1.0000	2.001159
1	Tx_PCB_L	0.3490	1.653923
2	PCB_lmp	0.1999	1.292459
3	PCB_H	0.0284	1.072431
4	Tx_PCB_S	0.0011	1.001120

Surprisingly, the most important single range feature is not the PCB channel length, as commonly thought, but the package length. The package length is also the most important feature in the pairs of features (Table V) and in the triplets of the features (Table VI). PCB link length and impedance and thickness of dielectric importance are also rated very high – we will examine such non trivial findings in depth later on. It can be noticed that based on single feature range analysis considering the package length feature alone, a range of package lengths can be defined in which the probability of having a "excellent" performance (COM > 4 dB) is over double than that in the overall population, since the max lift score for this feature is 2.001159. Furthermore, as we investigate the range defined by two features simultaneously (Table V), the max lift score for the pair of features identified as the most important in our system, package length and PCB channel length, is higher compared to the single feature defined range. This trend further continues for the triple feature defined range, demonstrated in Table VI. It can be seen that the triplet of characteristics identified to have the strongest impact on the performance consists of: (1) package length, (2) PCB channel length and (3) PCB channel impedance. Having certain values of this triplet of features will increase the chance for "excellent" performance (COM > 4dB) by more than 3.8 times.

Table V. Important range pair features for system with package length of 12 and 31 mm.

	feature1	feature2	max score	max lift
0	Pkg_len_RX	Tx_PCB_L	1.0000	3.307847
1	Pkg_len_RX	PCB_lmp	0.8773	2.586416
2	Pkg_len_RX	PCB_H	0.4797	2.136974
3	Pkg_len_RX	PCB_Dk	0.6936	2.131448
4	Tx_PCB_L	PCB_lmp	0.3344	1.453289

Table VI. Important range triplet features for system with package length of 12 and 31 mm.

	feature1	feature2	feature3	max score	max lift
0	PCB_lmp	TK_PCB_L	Pkg_len_RX	0.8263	3.835184
1	Pkg_len_RX	TK_PCB_L	PCB_lmp	0.9639	3.835184
2	PCB_Dk	TK_PCB_L	Pkg_len_RX	0.8480	3.216147
3	Pkg_len_RX	Tx_PCB_L	PCB_Dk	0.7398	2.951053
4	TK_PCB_L	PCB_lmp	Pkg_len_RX	1.0000	2.907210
5	PCB_Dk	PCB_lmp	Pkg_len_RX	0.7385	2.805366
6	PCB_H	Tx_PCB_L	Pkg_len_RX	0.8333	2.766596
7	Pkg_len_RX	PCB_lmp	PCB_Dk	0.6248	2.262003

Table VII. Important single range features for system with package length of 1 and 5 mm.

	feature	max score	max lift
0	Tx_PCB_L	1.0000	1.653923
1	PCB_lmp	0.5292	1.292459
2	PCB_Dk	0.1917	1.126741
3	PCB_H	0.0660	1.072431
4	Tx_PCB_S	0.0024	1.001120

Table VIII. Important range triplet features for system with package length of 1 and 5 mm.

	feature1	feature2	feature3	max score	max lift
0	Tx_PCB_L	PCB_lmp	PCB_Dk	0 4146	1.917592
1	Tx_PCB_L	PCB_lmp	PCB_H	0 3714	1.917592
2	Tx_PCB_L	PCB_lmp	PCB_Dk	1.0000	1.917592
3	Tx_PCB_L	PCB_lmp	Tx_PCB_S	0.1956	1.797743
4	Tx_PCB_L	PCB_Dk	PCB_H	0.5625	1.718731
5	PCB_lmp	Tx_PCB_L	PCB_Dk	0 9301	1.668179
6	PCB_lmp	Tx_PCB_L	PCB_H	0.8936	1.648301
7	Tx_PCB_L	PCB_lmp	PCB_H	0 9479	1.543401
8	PCB_lmp	PCB_Dk	PCB_H	0.2440	1.223649

Next, the same type of analysis is performed on the 1 and 5 mm case packages and the results of range analysis for single and triplet features are displayed in Tables VII, VIII. Surprisingly, the package length is no longer an important feature (this will be examined later). Moreover, the max lift scores are relatively low, and at first glance it seems that our

analysis has failed and is unsuccessful in identifying the important system characteristics. However, a more in depth examination (as will be presented later on in this section) reveals than in this case, most of the configurations have COM > 4dB and qualify as having an "excellent" performance. As a result, most of the characteristics will satisfy the performance requirement, and no "unique" properties are required, meaning having a preferred characteristic will only slightly improve the odds of having "excellent" performance.

In this case, to get an insight on system performance, a different question needs to be examined: what are the system characteristics responsible for bad performance, COM < 3dB. Such cases are expected to have some distinguished characteristics that separate them from the general population. The results of the Range Analysis with EVA considering what types of systems should be avoided are shown in Tables IX, X.

Table IX. Systems to avoid (COM<3dB): Important single-range features for system with package length of 1 and 5 mm.

	feature	max score	max lift
0	Tx_PCB_L	1.0000	2.626174
1	Pkg_len_RX	0.3192	1.471736
2	PCB_lmp	0.1406	1.371255
3	PCB_Dk	0.0564	1.094238
4	PCB_H	0.0476	1.079895
5	Tx_PCB_S	0.0088	1.010703

Table X. Systems to avoid (COM<3dB): Important range triplet features for system with package length of 1 and 5 mm.

	feature1	feature2	feature3	max score	max lift
0	Tx_PCB_L	PCB_Dk	Pkg_len_RX	1.0000	4.329620
1	Tx_PCB_L	PCB_H	Pkg_len_RX	0.8922	4.190758
2	Tx_PCB_L	PCB_lmp	Pkg_len_RX	0.9155	3.875588
3	Tx_PCB_L	Tx_PCB_S	Pkg_len_RX	0.7841	3.863166
4	Tx_PCB_L	PCB_lmp	PCB_H	0.7859	3.467138
5	Tx_PCB_L	PCB_lmp	PCB_Dk	0.8488	3.141867
6	Tx_PCB_L	PCB_H	PCB_Dk	0.7878	3.115134
7	Tx_PCB_L	PCB_Dk	PCB_H	0.6252	2.780420
8	Tx_PCB_L	PCB_lmp	PCB_H	0.5220	2.133623

It can be noticed, that in this case, the ranking based on a single feature and a triplet of features are very similar to the ranking of excellent systems with 12 and 31 mm package cases. This means that the same characteristics are important for the system operation.

Next, the findings of the proposed analysis are examined in detail and their validity is evaluated. All COM results are plotted in Fig. 8 and Fig. 9 as a function of the total link loss at the Nyquist frequency (including loss in package). Fig. 8 compares COMs for two very short packages with length 1mm and 5mm and Fig. 9 compares COMs for two reference cases of packages with length 12mm and 31mm. Graphs on the left show all data for each package case in the same color. Graphs on the right show additional information about PCB link length coded with colors – from blue for the shortest 0.5in link to red for the longest 12in link. First, we can observe that the shortest 1mm package provides the best performance – almost no failure cases. The 1mm case has only a few failures when total link losses exceed 40dB, which is expected. The 5mm package fails for very lossy links and very short links. A further increase of the package length to 12 mm makes things much worse – almost all cases fail. However, the longer 31mm package improves the situation. Note that 5mm is just a little smaller than the wavelength at the Nyquist frequency – we can expect resonances between the discontinuities when the package size becomes a multiple of half of the wavelength in the package. The presence of two strong discontinuities in the Rx package – bumps and balls explains the signal degradation for the packages with relatively small lengths that exceed half of the wavelength in the package. Transmission lines in the package are relatively lossy and a further increase of the package length helps to dump the resonances, as is clearly visible on the Single Symbol Response (SSR) shown in Fig. 10.

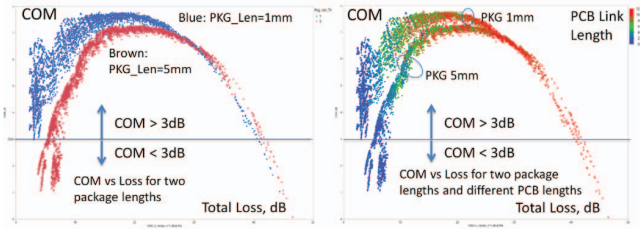


Fig. 8. COM vs link total loss at Nyquist frequency for two cases with very short packages (both graphs). Right graph shows PCB link length in color from blue (0.5in) to red (12in).

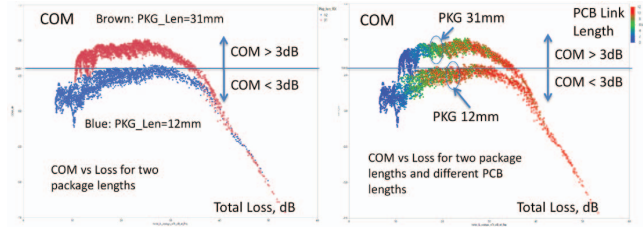


Fig. 9. COM vs link total loss at Nyquist frequency for two reference package lengths (both graphs). Right graph shows PCB link length in color from blue (0.5in) to red (12in).

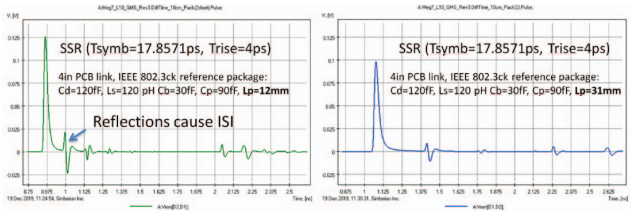


Fig. 10. Impulse response for 12mm package (left) and 31mm package (right).

The impact of the dielectric material selection (dielectric constant and loss tangent) is illustrated in Fig. 11. The picture shows 8 graphs – 4 graphs at the top for the shorter 12mm package and 4 graphs at the bottom for the longer 31mm package. Each graph is for a different PCB dielectric choice. The selection of better dielectrics does not help at all and

makes things worse for the shorter package lengths. Also the dielectric selection does not matter even for the longer package as long as the total link losses are below 35dB. We can also observe that dielectrics with more losses help to mitigate failures for very short PCB links.

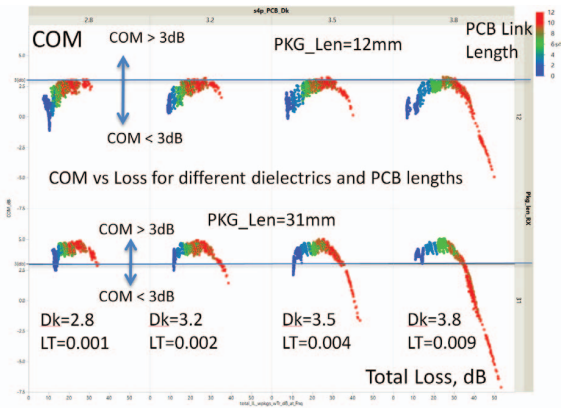


Fig. 11. COM vs link total loss for two reference package lengths (12mm and 31mm), different dielectrics (Dk&LT) and PCB link length (from blue 0.5in to red 12in).

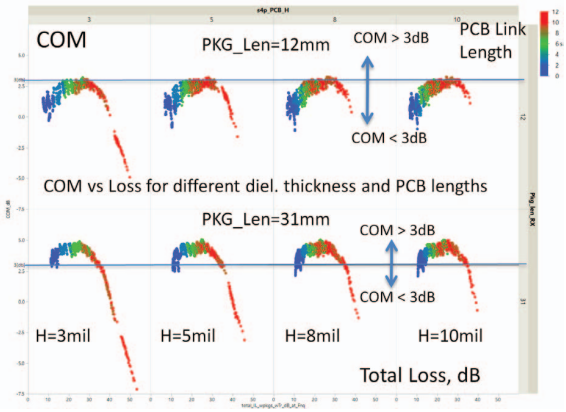


Fig. 12. COM vs link total loss for two reference package lengths (12mm and 31mm), different dielectric thickness (H) and PCB link length (from blue 0.5in to red 12in).

Yet another factor affecting the total losses is the dielectric thickness – it defines the trace width for the target impedance. Traces will have more losses with thinner dielectrics. The dependence of COM on the total loss at the Nyquist frequency is shown in Fig. 12. The picture shows 8 graphs – 4 graphs at the top for the shorter 12mm package and 4 graphs at the bottom for the longer 31mm package. Each graph is for a different thickness of the dielectric. We can observe that the thinner dielectrics help to reduce failure for the shortest PCB links, while the thickest dielectric helps to transmit the signal over longer PCB traces. This is an expected behavior.

In addition to the electrical and mechanical parameters of the dielectric, our numerical experiment included typical PCB production impedance variations. Graphs for COM vs the total losses at the Nyquist frequency are shown in Fig. 13 for two reference package length. The left graph shows additional information about the impedance variations coded with colors from blue for 85Ω to red form 105Ω. Graph on the right show ERL metric in colors. We can conclude that the lower impedance values were beneficial for the link performance.

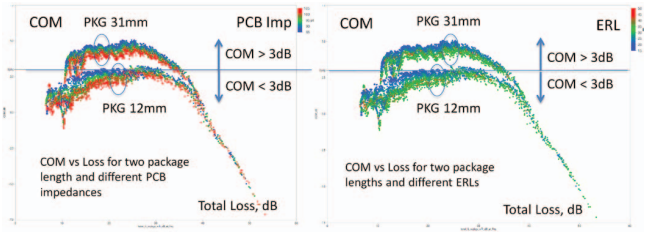


Fig. 13. COM vs link total loss at Nyquist frequency for two reference package lengths (12mm and 31mm, both graphs). Left graph shows PCB trace impedance in color from blue (85Ω) to red (105Ω). Right graph shows ERL coded in colors.

The design space can now be systematically explored using the proposed method. Note that the actual feature ranges identified by EVA are not shown for this simplified case. We illustrate the same process with the ranges for more realistic links in the next section.

VII. Realistic 112 Gb link with PCB crosstalk and discontinuities

More realistic link contains two identical reference package models (Rx and Tx) and PCB link with viahole transition from BGA to strip line, viahole transition from stripline to microstrip and AC coupling capacitors, model of AC coupling capacitors, viahole transition from microstrip back to strip line and viahole transition from strip line to BGA as illustrated in Fig 14. Example of viahole model is shown in Fig. 5. Only the best-case discontinuities (blue lines in Fig. 5) are used in this case study (typical case after optimization).

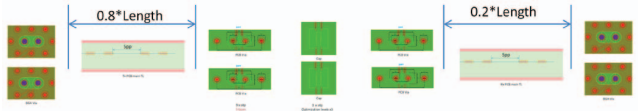


Fig. 14. Schematic view of the realistic link with coupling and discontinuities.

All features and values are shown in Table XI. The total number of links is 25920.

Table XI. DOE table for link with crosstalk and discontinuities.

	Feature						
1	Tx_PCB_TL_S (S/H)	1	2	3			
2	Tx_PCB_TL_L (Length) [in]	0.5	1	3	6	9	12
3a	PCB_Dk	2.8	3.2	3.5	3.8		
3b	PCB_LT	0.001	0.002	0.004	0.009		
5	PCB_TL_H (Ha, Hb) [mil]	3	5	8	10		
6	PCB_Imp [Ohm]	85	90	95	100	105	
7	Tx_PCB_TL_DS (Spp/H)	3	5	10			
8	SR [um]	0.075	0				
9	PKG_Len [mm]	5	12	31			

The AC caps are located at 0.8 of the total link length. As in the case of the simple link, the reference model was used to simulate the package with lengths 5, 12 and 31 mm (Pkg_len_TX feature). Tx and Rx packages are assumed

identical. To include the effect of crosstalk, an additional feature is required, Tx_PCB_TL_DS, defining the separation between differential pairs (pair to pair separation Spp in Fig. 6 divided by dielectric thickness H). To investigate the effect of conductor roughness, we added a feature for conductor roughness SR with just 2 values – zero for smooth conductor (the best-case scenario) and 0.075 um as in the simplest case (RF=24.5). All other features are defined exactly as in the case of the simple link.

The realistic link was investigated with 1 DFE tap as in the case of the simple link. EVA ranking of the features is shown in Table XII. Features for the most relevant single ranges are listed in Table XIII. We can see that the package length remains the most important single range feature as in the simple link case. PCB link length (Tx_PCB_L), impedance (PCB_Imp) and dielectric properties (PCB_Dk/LT) are ranked higher then conductor roughness (PCB_SR), and pair to pair separation defining the crosstalk (Tx_PCB_DS).

As an example, the single ranges identified by EVA for the impedance are shown in Table XIV. The table shows multiple outcomes with different lifts and maximal lift in line 44. It states that the impedance equal to 90 Ohm covers 2691 cases (Positive In) with COM > 3dB. 8249 cases are out of this range (Positive Out). It has a relatively small number of cases with COM < 3dB (Negative In), comparing to all cases (Negative Out + Negative In). The Range Lift of about 1.23 also indicates the importance of the impedance (see the definition of Lift in Section 2). The effect of the impedance is further illustrated in Fig. 15 – we can see more cases with COM > 3dB when the impedance range is extended from 85 to 95 Ohm (line 46 in Table XIV).

Table XII. All feature ranking results for the realistic link.

	feature	score	selection
56	Pkg_len_TX	1.0000	correlation-coverage
57	Tx_PCB_L	0.5426	correlation-coverage
58	PCB_Imp	0.2433	correlation-coverage
59	PCB_Dk	0.0544	correlation-coverage
60	PCB_LT	0.0544	correlation-coverage
61	PCB_H	0.0400	coverage
62	PCB_SR	0.0210	coverage
63	Tx_PCB_DS	0.0128	coverage
64	Tx_PCB_S	0.0064	coverage

Table XIII. Important single range features for the realistic link.

	feature	max score	max lift
0	Pkg_len_TX	1.0000	1.713138
1	Tx_PCB_L	0.8063	1.482123
2	PCB_Imp	0.3451	1.229225
3	PCB_H	0.0559	1.040834
4	PCB_Dk	0.0630	1.027400
5	PCB_SR	0.0302	1.022407
6	Tx_PCB_DS	0.0093	1.009892

Table XIV. Single ranges identified for the impedance feature.

	PCB_Imp	score	selection	Positive OUT	Negative OUT	Positive IN	Negative IN	Range Lift
44	90:90	0.1672	target	8249	12449	2691	2487	1.229225
45	-Inf:90	0.2865	target	5571	9957	5369	4979	1.227204
46	-Inf:95	0.3451	coverage -target	3210	7143	7730	7793	1.177833
47	-Inf:100	0.2770	target	1401	3777	9539	11159	1.090070

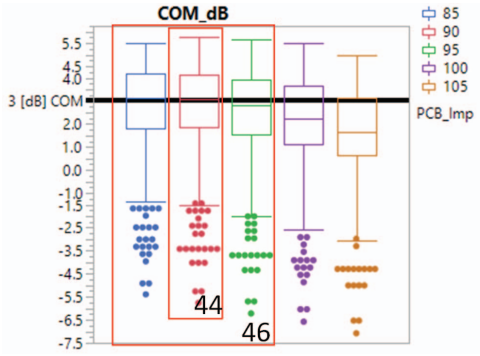


Fig. 15. Box plots for COM distributions for 5 PCB impedances (color-coded) with outlined ranges from row 44 and 46 of Table XIV.

Pairs of features for the most relevant ranges are listed in Table XV. Now we can see that the package length and the PCB impedance have the maximal lift that is also similar to the simple case. Four important ranges for pair of package length and PCB link impedance detected by EVA that substantially increase COM are shown in Table XVI and further illustrated on the box plot in Fig. 16.

Table XV. Important range pair features for the realistic link.

	feature1	feature2	max score	max lift
0	PCB_Imp	Pkg_len_TX	0.9959	2.213154
1	Pkg_len_TX	Tx_PCB_L	0.9782	2.182943
2	PCB_Dk	Pkg_len_TX	0.7254	1.882133
3	PCB_Dk	Tx_PCB_L	0.5314	1.505301
4	PCB_Imp	Tx_PCB_L	0.6828	1.390276
5	PCB_H	Tx_PCB_L	0.4950	1.323446

Table XVI. Possible ranges for PCB link impedance and package length.

	PCB_Imp	Pkg_len_TX	score	selection	Positive OUT	Negative OUT	Positive IN	Negative IN	Pair Lift
7	90:90	31:Inf	0.4957	coverage-target	9325	14825	1615	111	2.213154
8	-Inf:90	31:Inf	0.8086	target	7716	14711	3224	225	2.210964
18	-Inf:95	31:Inf	0.9959	correlation-coverage-target	6337	14365	4603	571	2.104236
23	-Inf:100	31:Inf	0.9741	correlation-target	5394	13583	5546	1353	1.901400

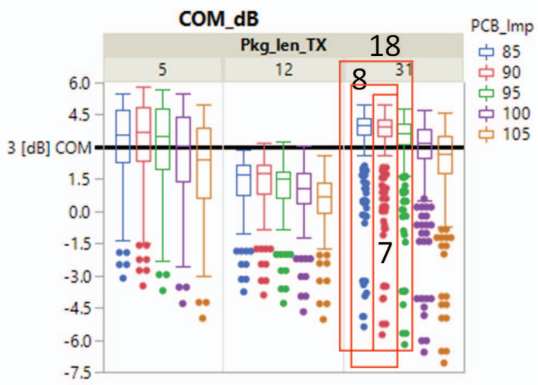


Fig. 16. Box plots for COM distributions for 3 package lengths (columns) and 5 PCB link impedances (color-coded) with outlined ranges from rows 7, 8 and 18 of Table XVI.

Restriction of the impedance to 90 Ohm and package length to 31 mm produces the maximal lift or maximises the likelihood of susses for this link. Proper selection of ranges for pair of package length and PCB link length as shown in Table XVII and illustrated in Fig. 17 can also increase likelihood of success. In brief, COM is better with longer packages, lower impedance and medium PCB link length in this case.

Table XVII. Possible ranges for package length and PCB link length.

	Pkg_len_TX	Tx_PCB_L	score	selection	Positive OUT	Negative OUT	Positive IN	Negative IN	Pair Lift
10	31:Inf	9:9	0.4275	target	9611	14825	1329	111	2.182943
11	31:Inf	6:9	0.6725	target	8301	14695	2639	241	2.167338
14	31:Inf	6:6	0.4140	target	9630	14806	1310	130	2.151734
19	31:Inf	6:Inf	0.8290	coverage	7140	14416	3800	520	2.080557
20	31:Inf	3:Inf	0.9782	correlation-coverage-target	6024	14092	4916	844	2.018688

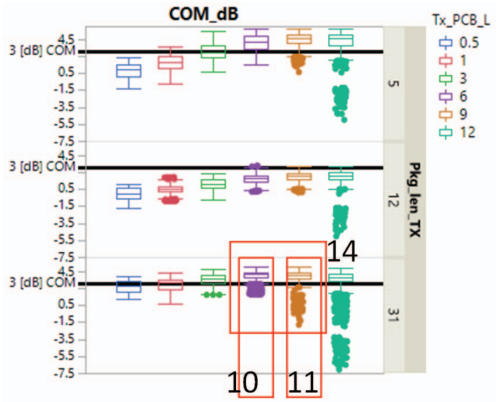


Fig. 17. Box plots for COM distributions for 3 package lengths (rows) and 6 PCB link lengths (color-coded) with outlined ranges from rows 10, 11 and 14 of Table XVII.

Finally, EVA-selected range triplets for the realistic link with coupling are shown in Table XVIII. Again, PCB link impedance and length and package length are the most important range triplets maximizing the lift. Four possible ranges for the first triplet are shown in Table XIX and further illustrated on the box plot in Fig. 18. Longer package links produce better COM. We can also conclude that the shorter package links will work better with more losses in the PCB.

The worst case scenario is the package link length that produces resonances that case failure almost always as illustrated in Fig. 19 (brown dots are almost all below the 3 dB threshold). However, usually we do not have control over the package link length.

Table XVIII. Important range triplet features.

	feature1	feature2	feature3	max score	max lift
0	PCB_Imp	Tx_PCB_L	Pkg_len_TX	1.0000	2.365265
1	PCB_Dk	Tx_PCB_L	Pkg_len_TX	0.7511	2.365265
2	PCB_Dk	PCB_Imp	Pkg_len_TX	0.8383	2.136133
3	PCB_Dk	Tx_PCB_L	PCB_Imp	0.5385	1.414642
4	PCB_H	Tx_PCB_L	PCB_Imp	0.5225	1.404740
5	PCB_Dk	Tx_PCB_L	PCB_H	0.4255	1.360514

Table XIX. Four possible ranges identified for the first triplet of features (PCB link impedance and length and package length).

	PCB_Imp	Pkg_len_TX	Tx_PCB_L	score	selection	Positive OUT	Negative OUT	Positive IN	Negative IN	Triplet Lift
9	-Inf:90	31:Inf	-Inf:3	0.5050	coverage	9335	14820	1605	116	2.365265
0	-Inf:95	31:Inf	6:6	0.3789	target	10076	14936	864	0	2.365265
1	-Inf:90	31:Inf	6:6	0.2971	target	10364	14936	576	0	2.250287
17	-Inf:95	31:Inf	-Inf:9	0.9038	correlation-target	7064	14502	3876	434	2.205840

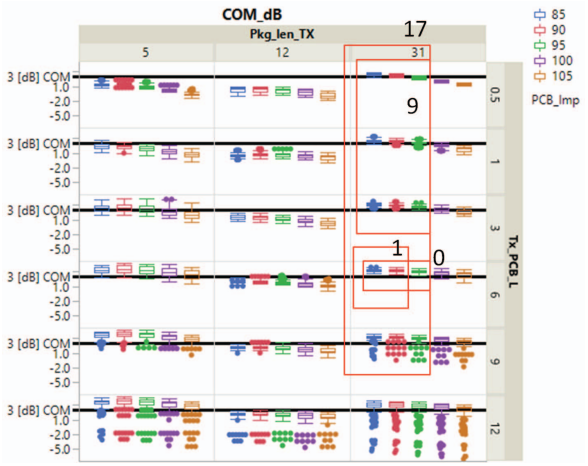


Fig. 18. Box plots for COM distributions for 3 package lengths (columns), 6 PCB link lengths (rows) and 5 PCB impedances (color - coded) with outlined ranges from rows 0, 1 and 9 of Table 7.8.

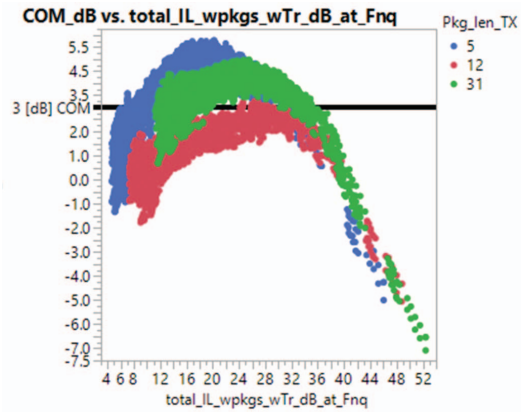


Fig. 19. COM vs link total loss at Nyquist frequency for 3 package lengths (color-coded) for all cases.

Expertise-based analysis becomes complex when tens or hundreds of features are involved. In such cases identifying important range combinations becomes practically impossible even for the best experts in the domain, and with a manual analysis there is little confidence that no important range combinations have been missed. The machine learning in general and the Range Analysis in particular are available for PCB or package designers who are unfamiliar with the signal integrity at all or are dealing with new technologies that have not yet been deeply studied. It is a formal process, where an initial set of features (length, thickens, dielectric,…) is identified with some knowledge about the features contributing to signal degradation. The rest of the process is a completely automated design exploration with the Machine Learning algorithms. The conclusion on relevant ranges of the features becomes very formal in this case – it does not require expertise or tedious manual simulations.

VIII. CONCLUSION

Real life design challenges require an enormous amount of parameter value combinations to populate the design space. Simple sweeps of design parameters may be not suitable. Particularly, a systematic approach is required to address SerDes design solution space coverage for multiple equalization mechanisms and various channel configurations affecting the system performance.

We demonstrate a practical application of ML based methods to identify the parameters and combinations thereof having the greatest effect on the design/system output, account for failure to meet the specs/standards, and provide an insight on how to optimize the design to meet a selected performance metric. This method is implemented on a 112Gb system case study.

This method allows a methodical, automated analysis of the solution space, yielding the desirable insight on system behaviour comprehendible for engineers, and can be used as a decision support tool for design choices in the hands of the system architect, Si designer, Si Engineer, and more.

REFERENCES

[1] Z. Khasidashvili, A. J. Norman. Range Analysis and Applications to Root Causing. In: 6th IEEE International Conference on Data Science and Advanced Analytics, DSAA 2019.

[2] Z. Khasidashvili, A. J. Norman. Feature Range Analysis. Journal of Data Science and Analytics, 2021.

[3] J. Furnkranz, D. Gamberger, and N. Lavra˜c. Foundations of Rule Learning. Cognitive Technologies. Springer, 2012.

[4] P. Clark, T. Niblett. The CN2 induction algorithm, Machine Learning, vol. 3, 1989.

[5] S. Wrobel. An algorithm for multi-relational discovery of subgroups. PKDD 1997.

[6] I. Guyon, A. Elisseeff. An Introduction to Variable and Feature Selection. Journal of Machine Learning Research 3, 2003.

[7] N. Lavra˜c, B. Kavsek, P. Flach, L. Todorovski. Subgroup Discovery with CN2-SD. Journal of Machine Learning Research 5, 2004.

[8] M. Atzmueller, F. Lemmerich. Fast Subgroup Discovery for Continuous Target Concepts. Foundations of Intelligent Systems, LNCS 5722, 2009.

[9] C.W. Koay, A.J. Norman, Z. Khasidashvili. Analog Circuit Process Monitoring, IEEE Intl. Workshop on Defects, Adaptive Test, Yield and Data Analysis, 2017.

[10] A. Manukovsky, Y. Juniman, Z. Khasidashvili. A Novel Method of Precision Channel Modeling for High Speed Serial 56Gb Interfaces. DesignCon 2018.

[11] A. Manukovsky, Z. Khasidashvili, A.J. Norman, Y. Juniman, R. Bloch. Machine Learning Applications for Simulation and Modeling of 56 and 112 Gb SerDes Systems. DesignCon 2019.

[12] IEEE P802.3ck Task Force - Tools and Channels http://www.ieee802.org/3/ck/public/tools/index.html

[13] M. Brown, M. Dudek, A. Healey, E. Kochuparambil, L. Ben-Artsi, R. Mellitz, C. Moore, A. Ran, P. Zivny, "The state of IEEE 802.3bj 100 Gb/s Backplane Ethernet", DesignCon 2014, January 2014, Santa Clara

[14] Measuring Channel Operating Margin – Anritsu app note https://dl.cdn-anritsu.com/en-us/test-measurement/files/Technical-Notes/White-Paper/11410-00989A.pdf

[15] Y. Shlepnev, Decompositional Electromagnetic Analysis of Digital Interconnects, IEEE International Symposium on Electromagnetic Compatibility (EMC13), Denver, CO, 2013, p.563-568.

[16] A. Manukovsky, Y. Shlepnev, Measurement-assisted extraction of PCB interconnect model parameters with fabrication variations, 2019 IEEE 28st Conference on Electrical Performance of Electronic Packaging and Systems, Oct. 6-9, 2019, (EPEPS 2019), October 6-9, 2019, Montreal, Canada.

[17] M. Marin, Y. Shlepnev, Systematic approach to PCB interconnects analysis to measurement validation, 2018 IEEE Symposium on Electromagnetic Compatibility, Signal and Power Integrity, July 30-August 3, 2018, Long Beach Convention Center, Long Beach,.