Holistic Die-to-Die Interface Design Methodology for 2.5-D Multichip-Module Systems

Muhammad Waqas Chaudhary, Andy Heinig, and Bhaskar Choubey, *Senior Member, IEEE*

***Abstract*— More than Moore technologies can be supported by system-level diversification enabled by chiplet-based integrated systems within multichip modules (MCMs) and silicon interposer-based 2.5-D systems. The division of large system-on-chip dies into smaller chiplets with different technology nodes specific to the chiplet application requirement enables the performance enhancement at the system level while achieving lower power consumption. However, these chiplets need to communicate with each other. Routing resources in MCM and 2.5-D systems are limited due to system size and thickness restrictions. This work presents an energy/bit optimization approach for multichip systems with the possibility of co-optimization with the routing resources defined by the signaling pitch. Holistic design methodologies are shown which can be further extended by the designer to define the application-specific constraints. A detailed analysis of energy per bit relationship to the voltage swing requirement for different topologies is presented along with a specific CML signaling-oriented design flow for 2.5-D chip-to-chip interfaces as an example of topology-specific optimization possibilities within this methodology.**

***Index Terms*— Chip-to-chip communication, design methodology, energy per bit, multichip module (MCM), optimization, routing pitch.**

I. Introduction

THE typical modern processors have billions of transistors per chip, and their size has increased over the years [1]. At the same time, memory sizes have also increased. The size of a typical low-power double-data-rate (LPDDR3) memory package for 16-Gb or 2-GB size is $15 \times 15\,\text{mm}^2$. For 16-GB memory access, in a PCB-based solution, the design would require a minimum of eight packages of RAM around the CPU. While one can find solutions around this problem by increasing the number of ranks in memory access, the ideal parallel access capability for read/write of all memories at the same time would require 32×64 data channels, leading to a total of 2048 interconnects for just the data lines. The minimum size for such a PCB-based system would be $(8 \times 225.250)\ 2050\,\text{mm}^2$. This, therefore, calls for miniaturization toward smaller systems and placing many chips (chiplet) in a single package to reduce the number of interconnects on the eventual system board where the multichip system (MCS) is placed.

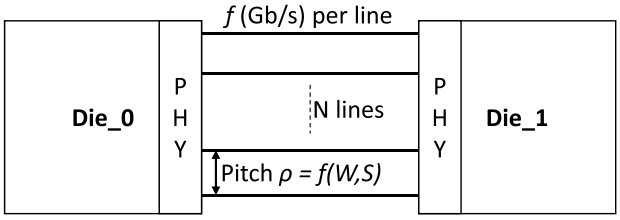


Fig. 1. Die-to-die interface in an MCS.

Loke *et al.* [2] demonstrated in beyond CMOS scaling that analog and mixed-signal blocks do not necessarily benefit greatly from the classical Moore scaling. Rather, special circuit design blocks, such as high-voltage IO cells and analog blocks, are more difficult to design with the same performance in sub-20-nm technology nodes than longer channel length nodes. Therefore, chiplet-based approaches with different technology nodes placed together within a package or 2.5-D silicon interposer-based system provide a solution to large PCB systems and beyond CMOS scaling-related problems in analog–mixed-signal design. A typical chip-to-chip interface is shown in Fig. 1. There are two main aspects to a chip-to-chip communication interface design—PHY or front-end for transmitter/receiver blocks and the channel/interconnect.

Optimal designs of these require an understanding of the entire structure including the transmitter, channel, and receiver. Balamurugan *et al.* [3] modeled and statistically analyzed high-speed input–output (IO) links with various noise sources in the entire link, obtaining a methodology for estimating the jitter and eye diagram at the receiver. Further statistical modeling of high-speed transceiver link has been presented by Stojanovic and Horowitz [4] and Sanders *et al.* Oh *et al.* [5] extended the statistical model by accurate jitter estimation methodology considering not just bounded Gaussian noise sources but also those with arbitrary spectrum.

Menin *et al.* [25] estimated the evolution of the least mean square (LMS) adaptive loop for equalization optimization. Compared to conventional statistical eye estimation for a provided channel, transmitter, and receiver, a probabilistic estimation methodology was shown to reduce the computational time overhead. Beyene *et al.* [6] described the optimal data rate estimation methodology for a provided channel based on

Manuscript received June 15, 2021; revised September 4, 2021; accepted September 11, 2021. Date of publication October 1, 2021; date of current version December 16, 2021. Recommended for publication by Associate Editor W. T. Beyene upon evaluation of reviewers' comments. *(Corresponding author: Muhammad Waqas Chaudhary.)*

Muhammad Waqas Chaudhary and Andy Heinig are with the Division Engineering of Adaptive Systems, Fraunhofer Institute for Integrated Circuits, Fraunhofer IIS/EAS, 01187 Dresden, Germany (e-mail: muhammad.chaudhary@eas.iis.fraunhofer.de; andy.heinig@eas.iis.fraunhofer.de).

Bhaskar Choubey is with the Department of Analog Circuits and Image Sensors, Siegen University, 57076 Siegen, Germany (e-mail: bhaskar.choubey@uni-siegen.de).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCPMT.2021.3117118.

Digital Object Identifier 10.1109/TCPMT.2021.3117118

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

the signaling power, extending previous optimal data rate estimation of Hatamkhani and Yang [7]. The frequency-dependent loss in the signal swing due to channel was used to predict the signaling power and calculate the optimum data rate with respect to power for a given channel. Kiran *et al.* [8] further extended the statistical eye estimation methodology to the ADC-based serial transceivers. They added the ADC-related noise parameters to the statistical approach. A hybrid scheme combining the statistical methodology with transient simulation was used to estimate the parameters related to the nonlinearity (INL), quantization noise, and time-interleaving mismatches in ADC-based links. Jitter modeling and estimation methodology for source-synchronous links has also been studied by Balamurugan and Shanbhag [9], [10]. They modeled the effect of the channel on the transmitter jitter, presenting techniques for jitter mitigation including slower forwarded clock and jitter equalization. Beyene *et al.* [11] presented a simulation and measurement correlation for system model parameters for 16-Gb/s memory interface, adjusting them to predict the link performance for similar memory interfaces.

In other works, Karim *et al.* [12] analyzed and simulated the different 2-D, 2.5-D, and 3-D interconnects for finding the optimum among them in terms of power. Jangam *et al.* showed the advantages of a dielet-based approach SuperCHIPS in comparison to the PCB- and SOC-based approaches in terms of latency, bandwidth, and power efficiency [13] by using different design libraries for lower power (LPE) in smaller cores and general purpose (GP) in larger cores for a multicore architecture. Zhang *et al.* [14] demonstrated the benchmarking of die-to-die interconnects in terms of latency and energy efficiency. Another work by Jangam and Iyer [15] proposed a comprehensive signaling figure-of-merit (si-FoM) for communication wireline links showing that some links can achieve very high energy efficiency but only work for very short lengths. However, these works have been optimizing the black-box IO circuits with extracted interconnect models to find the optimum interconnect only in terms of energy efficiency or delay.

In addition to these optimization studies, there have been several front-end physical interface (PHY) design implementations in the literature. Poulton *et al.* [16] showed a 25-Gb/s ground-referenced single ending transceiver for die-to-die communication over package interconnect length of maximum 10mm. They used length matching of data and clock lines on the package to simplify the clock and data recovery on the receiver side. Dehlaghi and Carusone [17] demonstrated silicon interposer-based die-to-die link for maximum 20-Gb/s data rate. The signaling topology uses passive termination on the transmitter for equalization. Dickson *et al.* [20]–[22] showed source-synchronous parallel links for chip-to-chip communication. For equalization at the receiver, decision feedback equalization (DFE) was implemented with infinite impulse response (IIR) along with standard digital DFE. Tajalli *et al.* [21] recently demonstrated a 1-pJ/bit 20-Gb/s/wire die-to-die link over package for up to 6-dB loss channels. The transceiver uses 5 bits over a six-wire coded non-return-to-zero (CNRZ) scheme to minimize crosstalk, simultaneous switching noise (SSN), and electromagnetic interference (EMI). These integrated circuit design studies utilize a known channel and are, hence, optimizing the PHY for best energy performance. Most system studies optimize the energy per bit for a provided signaling topology for a known channel [7], [22], [23].

However, codesign of the channel with analytical models of transmitters and receivers has received limited attention, if any. Furthermore, there are no codesign approaches specifically for IO and channel with respect to multichip module (MCM) and 2.5-D interconnect. Some limited studies have specified the energy cost calculation for a known topology in a specified channel [24]. However, they do not consider how a specific signaling driver topology could be selected for the desired aggregate bandwidth, die edge size, and given substrate properties. This article fulfills this gap by presenting detailed voltage swing-based energy per bit cost analysis for channel and noise constraints. One can perform detailed analysis with a known transmitter and interconnect models; this does not help the IO and channel designers at the start of the design, as no transmitter model is present at this stage. Hence, this article combines the analytical driver models for different topologies with interconnect width and spacing to determine the optimum quickly at the design initiation stage, itself. It then presents a holistic design methodology for energy/bit*pitch objective-based chip-to-chip signaling topology and channel width, and spacing and substrate selection. This is followed by a CML IO-based 2.5-D interface holistic approach description and topology-specific optimization analysis. Section II provides a theoretical framework of the voltage swing requirements. Section III presents a methodology for the energy performance optimization of these MCSs. Section IV provides an analysis of different signaling topologies used in the MCS interface. Section V details the methodology for energy and pitch combined optimization. Section VI summarizes this article and provides an outlook into future work.

II. Background on Voltage Swing

In digital communication wireline systems, the digital bits are encoded in the form of electrical symbols by the transmitter using a digital-to-analog converter (DAC) [4]. Additional noise is added to the symbols due to inherent noise in the transmitter. These transmitted symbols travel through the wireline channel and are distorted by the channel impulse response. The distorted symbols are further affected by the noise of the receiver input front end. An ideal sampler or slicer detects the symbols just like an analog-to-digital converter (ADC) [3]. A common coding scheme for wireline communication is non-return-to-zero (NRZ), which encodes each digital bit in the form of a symbol $\pm V_s$, where V_s is the symbol voltage. The system model for an NRZ coded wireline chip-to-chip communication is shown in Fig. 2 [25].

There are several sources of noise in the communication link from the transmitter to the receiver sampler. Noise in the transmitter is caused by transistors, resistors, and the crosstalk due to on-chip interconnect and power supply variations [4]. Crosstalk and power supply noise are often deterministic due to their dependence on the design imperfections, such as limited spacing between interconnects and high impedance of

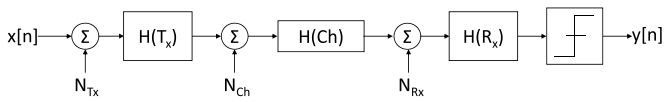


Fig. 2. Chip-to-chip communication link model.

the power supply network. Transistors and resistors in high-frequency operation, nevertheless, suffer from purely random thermal white noise [4].

Channel noise is mainly caused by intersymbol interference (ISI), which can be limited by using equalization at the receiver or transmitter end. This equalization or high-frequency peaking leads to an equivalent loss in the voltage swing. The receiver front end generally has a preamplifier with equalizer and sampler to convert the analog symbol to digital bit [8]. The noise at the receiver is due to the mismatch offset and the circuit noise in the sampler and preamplifier. The offset can be corrected by calibration techniques, with typical values of few millivolts, including the sampler sensitivity [11].

Considering a minimum required voltage margin at the receiver sampler, the minimum transmitted symbol voltage V_s is derived from the total deterministic and random noise in the $T_x \rightarrow Channel \rightarrow R_x$ link. The deterministic noise is added to the signal; however, the effect of random noise is dependent upon the minimum required bit error rate (BER) at the receiver sampler. Random noise probability is typically assumed to be a Gaussian function with zero mean μ and σ_n standard deviation. For NRZ-signaling, the BER P_{err} of link determines the minimum voltage swing based on the total random noise with standard deviation σ_n according to the Q-function as [26]

$$\text{BER}_{\text{NRZ}} = P_{\text{err}} = Q\left(\frac{V_s}{\sigma_n}\right) \tag{1}$$

where the Q-function is given as

$$Q(\tau) = \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \frac{-\mu^2}{2} d\mu. \tag{2}$$

From this, one can derive that for a typical BER of 10^{-12}, the voltage swing V_s/σ_n should be 7. In other words, the voltage peak-to-peak swing should be at least 14 times the rms noise. The BER can be visualized as the shaded area under the convolved noise and original signal PDFs shown in Fig. 3.

The peak-to-peak voltage V_{spp} can then be expressed as

$$V_{\text{spp}} = 14\sigma_n + k_c V_{\text{spp}} + n_{\text{Rx}} + k_{\text{eq}} V_{\text{spp}} + n_{\text{PS}} + V_{\text{Margin}} \tag{3}$$

where k_c is the crosstalk coefficient, which is the deterministic noise. It should be noted that the crosstalk noise is dependent on the transmitted voltage swing and cannot be removed by simply increasing the voltage swing. k_{eq} is the equalization or high-frequency peaking-dependent loss of voltage swing, which is dependent on the frequency of operation, the channel conductive, and dielectric losses. It is also a deterministic noise and can be estimated directly from the channel frequency response models, which are generally available in the form of RLGC tabular data or scattering parameters. The receiver R_x sensitivity and offset are included here by the term n_{Rx}, which can be in the order of few millivolts after offset correction techniques. The power supply network-dependent noise is denoted by n_{PS}. Generally, around 50 mV is a good voltage

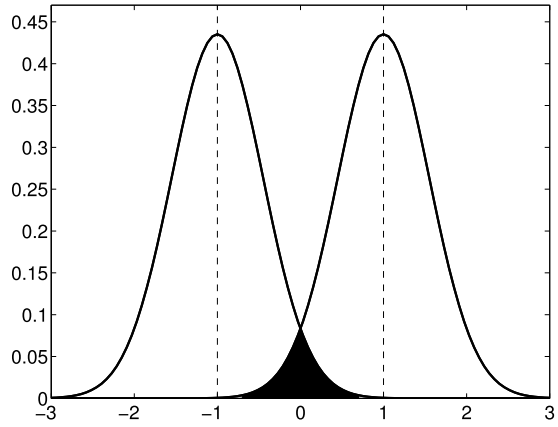


Fig. 3. Original signal convolved with noise PDF.

margin shown here as v_{Margin}. By including the V_s-dependent noise sources, the voltage swing from (3) can be rewritten as

$$V_{\text{spp}} = \frac{14\sigma_n + n_{\text{Rx}} + n_{\text{PS}} + v_{\text{Margin}}}{1 - k_c - k_{\text{eq}}}. \tag{4}$$

This shows that, in order to minimize the voltage swing, the coefficients k_c and k_{eq} have to be minimized [5].

As an example, let us consider a typical 400-mV swing peak-to-peak V_{spp} transmitted signal over a channel with 12-dB loss, received by a front end with offset and sensitivity of 10 mV after offset correction. Assume the k_c coefficient to be 0.1, σ_n rms value to be 1 mV, and power supply noise as 10 mV. The noise coefficient k_{eq} due to dB equalization Eq can be expressed as

$$k_{\text{eq}} = 1 - 10^{-Eq/20}. \tag{5}$$

For 12-dB equalization, k_{eq} is equal to 0.75, which means that 75% of the voltage swing is lost in the equalization and the available swing from the 400-mV signal is only 100 mV. The remaining margin v_{Margin} after adding all other random and deterministic noise sources is only 26 mV, which is less than the generally required margin of more than 50 mV. Thus, the voltage swing must be selected in correlation with the channel, required frequency of operation, and crosstalk behavior. Therefore, design methodologies and optimization approaches are required to minimize the voltage swing and select an optimum channel design. The circuits' noise parameters, such as receiver margin, random transmitter noise, and power supply noise, are kept constant in this design methodology to emphasize the correlation between driver topology and channel. The channel loss characterized by k_{eq} will be indirectly signified by the effective bandwidth (BW_{tot}) in Section V and presents a demonstration for energy efficiency with pitch combined optimization for current mode logic (CML) signaling on interposer interconnect.

III. Methodology for Energy Optimization

From (4), it can be concluded that the voltage swing is dependent upon the crosstalk and equalization characteristics of the channel. Crosstalk factor k_c is directly correlated with the spacing between the signal lines. Considering that the signal–signal–signal routing approach, as shown in Fig. 4,

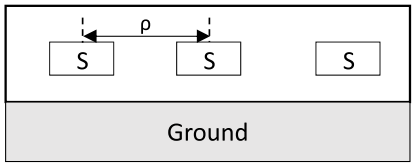


Fig. 4. Signal–signal–signal topology.

is used with a ground plane beneath, the edge length of die (E) is defined as

$$E = N\rho + W \tag{6}$$

where ρ is the signaling pitch and N is the number of signal lines. ρ is defined as [27]

$$\rho = W + S.$$

Then, the spacing (S) can be written as

$$S = E - (N+1)W.$$

Generally, the spacing between signal lines is defined as the multiple of the width (W) of line and is in the range of 1–3 W for 2.5-D interconnects [27]. Considering the multiple coefficient $a_s \in 1-3$, the width (W) can be written as function of N as

$$W = \frac{E}{N+1+a_s}. \tag{7}$$

The coefficient k_c as a function of spacing S can be estimated from the capacitance relationship between the signal lines, the width-to-dielectric thickness ratio, and the impedance. The impedance determines the required width-to-dielectric thickness [28]. If H is the height from signal to plane, the width must be about two times the height H for 50-Ω impedance. Assuming approximately 2.5% typical near-end-crosstalk coefficient k_c, it has been shown that the spacing should be two times the width [28]. Hence, for 2.5% typical k_c, the spacing is $S = 2W$. Therefore, the maximum number of lines N with 2.5% k_c, and given dielectric thickness H and 50-Ω impedance, can be described as

$$N_{\max} = \frac{E}{2H} - 3. \tag{8}$$

The relationship to aggregate bandwidth F and per line bandwidth f can be expressed as

$$N \le N_{\max} = \frac{F}{f_{\min}} = \frac{E}{2H} - 3 \tag{9}$$

where $f_{\min}$ denotes the minimum frequency/line achievable with maximum number of lines $N_{\max}$ for given aggregate F. It should be noted that $N_{\max}$ is limited by the integration technology for interconnect, along with the minimum pad pitch on the dies. $f_{\min}$ will also be indirectly limited due to $N_{\max}$ limitation. Furthermore, the maximum frequency could be limited by the silicon technology available for the dies.

The equalization factor k_{eq} is now the only factor in the denominator of (4), which can change the requirement of transmitted voltage swing V_{spp}. For a given ϵ_r and length of line L, the loss can be controlled by suitably selecting the dielectric constant ϵ_r value. By increasing the ϵ_r and tangent-loss tanδ, the losses increase, and k_{eq} factor increases; however, the cost of package or substrate is lower. When the dielectric constant is lower, the costs for package substrate are higher, but k_{eq} factor is lower, resulting in lower voltage swing requirement and lower energy costs.

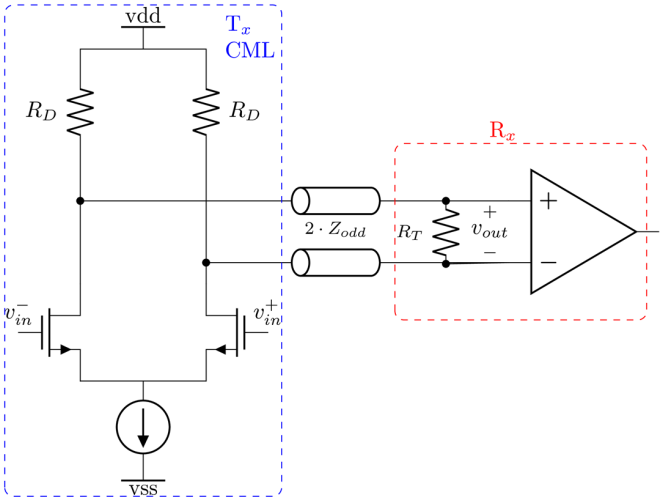


Fig. 5. CML topology for differential signaling.

The energy costs for a given voltage swing are dependent upon the signaling topology. These include $T_i \in [\text{SSTL} - \text{LCM}, \text{SSTL} - \text{HCM}, \text{CML}]$. The energy per bit pJ/bit estimation should then be performed for these topologies in relationship to the voltage swing requirement. Therefore, the methodology for minimum energy/bit for given aggregate bandwidth is described as follows.

1) For provided dielectric thickness H and dielectric constant ϵ_r, determine the maximum number of lines possible $N_{\max}$.
2) Determine k_{eq} for required W for given impedance Z_o and frequency per unit line $f_{\min}$.
3) Determine energy cost ϕ (pJ/bit) for each topology $T \in T_i$.
4) Select the topology with minimum ϕ cost at required $f = F/N_{\max}$.

IV. Transmitter Front-End Driver

For signal transmission over the wire, the most popular are CML differential signaling [29] and source series terminated logic (SSTL) [30]. Both of these topologies assume transmitter termination along with a far-end receiver termination for good signal integrity at high data rates. These topologies shall be studied below to make the ground for co-optimization of channel and driver discussion.

A. Current Mode Logic Differential Signaling

The energy per bit ϕ can be estimated as [27]

$$\phi = P \cdot t_{\text{UI}} = \frac{P}{f} \tag{10}$$

where t_{UI} is the data bit unit interval, P is the power consumption, and f is the signal data rate. The CML topology for differential signaling is shown in Fig. 5, where the tail current

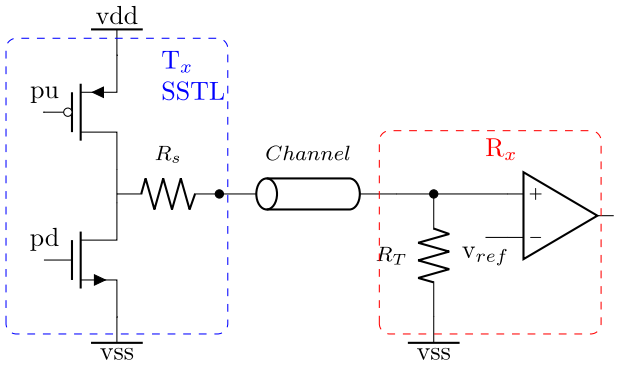


Fig. 6. SSTL-GND termination signaling.

source is constantly sinking current I from the power source vdd. Based on the differential inputs vin^+, vin^- to the driver, the current I is switched from one side of differential pair to the other side. This generates the differential peak-to-peak voltage swing $V_{\text{spp}} = \text{IZ}_o$. Therefore, for constant impedance, the voltage swing is directly proportional to the current I. The energy per bit for CML can, therefore, be expressed as

$$\phi_{\text{CML}} = \frac{V_{\text{DD}} \cdot I}{f} = \frac{V_{\text{DD}} \cdot 2V_s}{f \cdot R_T} \tag{11}$$

where V_s is the single-ended swing $V_s = V_{\text{spp}}/2$. It should be noted that the maximum number of lines available for the differential signaling is half of the single-ended signaling. Hence, f_{min} for CML is also doubled for a given F bandwidth

$$f_{\text{mindiff}} = \frac{F}{(N_{\text{max}}/2)} = \frac{2F}{N_{\text{max}}} = 2f_{\text{min}}. \tag{12}$$

B. Source Series Terminated Logic-LCM Using Ground Termination

The first voltage mode source series termination signaling topology to be studied here uses GND termination at the receiver end, as shown in Fig. 6 [30]. Here, the power source vdd sources current during high bit transmission resulting in voltage v_{oh} at the receiver. On the other hand, there is no current sourced during the low bit transmission v_{ol} [30]. Hence,

$$v_{\text{oh}} = V_{\text{DD}} - \frac{V_{\text{DD}}}{2R_T} \cdot R_T = 0.5V_{\text{DD}} \tag{13}$$

$$v_{\text{ol}} = 0. \tag{14}$$

These high and low voltages result in a single-ended swing V_s of

$$V_s = v_{\text{oh}} - v_{\text{ol}} = 0.5V_{\text{DD}}. \tag{15}$$

The energy per bit at a given frequency f is dependent upon the current sourced during high and low bits' transmission. The current during v_{oh} transmission is

$$I_h = \frac{V_{\text{DD}}}{2R_T}. \tag{16}$$

Hence, the *rms* current from the supply V_{DD} is

$$I_{\text{rms}} = \frac{I_h}{\sqrt{2}} = \frac{V_{\text{DD}}}{2\sqrt{2}R_T} = \frac{V_s}{\sqrt{2}R_T}. \tag{17}$$

The energy per bit can, therefore, be written as

$$\phi_{\text{sslcm}} = \frac{P}{f} = \frac{V_{\text{DD}} I_{\text{rms}}}{f} = \frac{V_{\text{DD}}^2}{2\sqrt{2}fR_T} = \frac{\sqrt{2}V_s^2}{fR_T}. \tag{18}$$

Since there is no current consumption during the v_{ol} transmission, the energy per bit ϕ consumption can be reduced if there are more number of low bits compared to high bits. Assuming that some type of data bus inversion (DBI) technique is applied, which leads to an increase in the number of low bits to 90%, with high bits being only 10%, then the *rms* current and energy per bit for this topology become

$$\begin{aligned} I_{\text{rmssslcm10\%}} &= \sqrt{\frac{1}{T}\left[0.1T\left(\frac{V_{\text{DD}}}{2R_T}\right)^2 + 0\right]} \\ &= \sqrt{0.1\left(\frac{V_{\text{DD}}}{2R_T}\right)^2} = \frac{V_{\text{DD}}}{2\sqrt{10}R_T} = \frac{V_s}{\sqrt{10}R_T} \end{aligned} \tag{19}$$

$$\begin{aligned} \phi_{\text{sslcm10\%}} &= \frac{P}{f} = \frac{V_{\text{DD}} I_{\text{rms}}}{f} = \frac{V_{\text{DD}}^2}{2\sqrt{10}fR_T} \\ &= \frac{2V_s^2}{\sqrt{10}fR_T} = \frac{0.63V_s^2}{fR_T}. \end{aligned} \tag{20}$$

C. Source Series Terminated Logic-HCM Vtt Termination

Another common variant of source terminated logic used in double-data-rate memories is the termination to a source-sink supply generally termed V_{tt}, as shown in Fig. 7 [31]. Here, the current is sourced from the power supply from V_{DD} during the v_{oh} transmission, and current is sourced from V_{tt} during v_{ol} transmission. The high and low voltages can be calculated as

$$v_{\text{oh}} = V_{\text{DD}} - \frac{V_{\text{DD}} - V_{\text{tt}}}{2R_T} \cdot R_T = 0.5V_{\text{DD}} + 0.5V_{\text{tt}} \tag{21}$$

$$v_{\text{ol}} = V_{\text{tt}} - \frac{V_{\text{tt}}}{2R_T} \cdot R_T = 0.5V_{\text{tt}}. \tag{22}$$

Therefore, the single-ended swing V_s is

$$V_s = v_{\text{oh}} - v_{\text{ol}} = 0.5V_{\text{DD}} + 0.5V_{\text{tt}} - 0.5V_{\text{tt}} = 0.5V_{\text{DD}}. \tag{23}$$

The current sourced from V_{DD} during v_{oh} transmission can be expressed as

$$I_{\text{hVdd}} = \frac{V_{\text{DD}} - V_{\text{tt}}}{2R_T}. \tag{24}$$

On the other hand, the current during v_{ol} transmission sourced from supply V_{tt} is

$$I_{\text{lVtt}} = \frac{V_{\text{tt}}}{2R_T}. \tag{25}$$

The current profile of this topology during high- and low-voltage symbols transmission is shown in Fig. 8.

Due to two supplies being used here, the *rms* current is calculated for both V_{DD} and V_{tt}. The rms currents for 50%

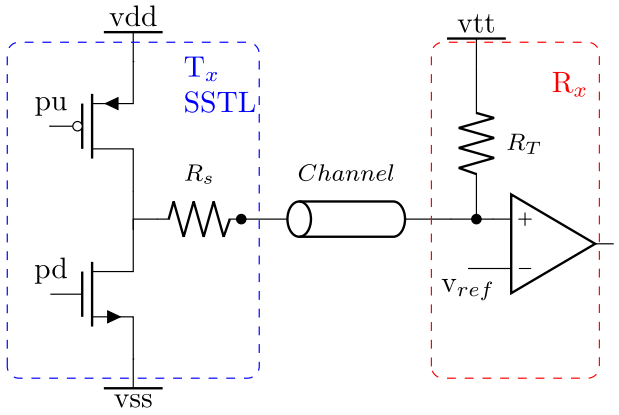


Fig. 7. SSTL-Vtt termination signaling.

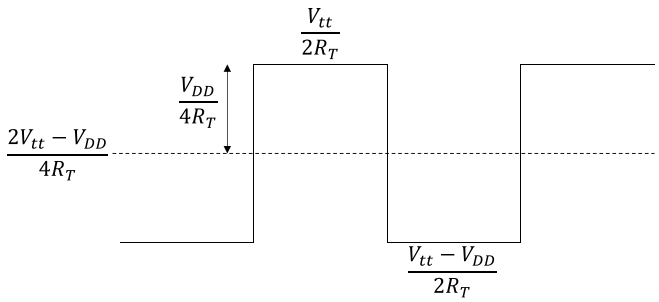


Fig. 8. SSTL-Vtt termination signaling current profile.

high and low symbols are

$$I_{\text{rmsVdd}} = \frac{I_{\text{hVdd}}}{\sqrt{2}} = \frac{\frac{V_{\text{DD}}-V_{\text{tt}}}{2R_T}}{\sqrt{2}} \tag{26}$$

$$I_{\text{rmsVtt}} = \sqrt{\left(\frac{2V_{\text{tt}}-V_{\text{DD}}}{4R_T}\right)^2 + \left(\frac{V_{\text{DD}}}{4R_T}\right)^2} = \frac{\sqrt{V_{\text{DD}}^2 + 2V_{\text{tt}}^2 - 2V_{\text{tt}}V_{\text{DD}}}}{2\sqrt{2}R_T}. \tag{27}$$

The total power consumption P is the sum of power in both supplies. The power and energy per bit for this topology can, therefore, be expressed as

$$P = P_{V_{\text{tt}}} + P_{V_{\text{DD}}} = V_{\text{tt}} I_{\text{rmsV}_{\text{tt}}} + V_{\text{DD}} I_{\text{rmsV}_{\text{DD}}} \tag{28}$$

$$E_{\text{bssthcm}} = P \cdot t_{\text{UI}} = \frac{P}{f} = V_{\text{DD}} \frac{V_{\text{DD}} - V_{\text{tt}}}{2\sqrt{2} f R_T} + V_{\text{tt}} \frac{\sqrt{V_{\text{DD}}^2 + 2V_{\text{tt}}^2 - 2V_{\text{tt}}V_{\text{DD}}}}{2\sqrt{2} f R_T}. \tag{29}$$

Typically, V_{tt} is assumed to be half of V_{DD} [31], and hence, energy per bit can be shown as

$$\phi_{\text{ssthcm}} = \frac{1.2V_s^2}{f R_T} = \frac{V_{\text{DD}}^2}{3.3 f R_T}. \tag{30}$$

D. 300-mV Swing Example

For a typical single-ended voltage swing requirements of 300 mV and V_{DD} of 0.6 V, Fig. 9 shows the comparison of these topologies for various data rates and achieved energy/bit. It should be noted here that this graph does not consider

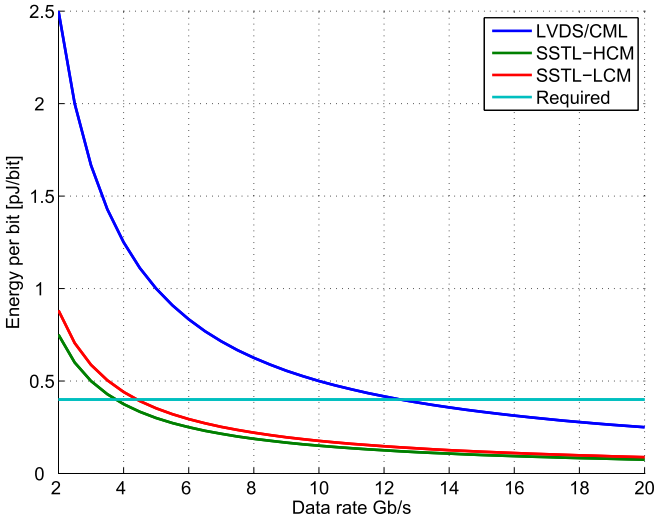


Fig. 9. Energy/bit ϕ comparison for $V_s = 0.3$V (no equalization considered).

equalization, assuming low enough losses limiting the use of equalization in drivers due to short interconnect length within the MCM or package. If the required energy/bit is 0.4 pJ/bit, one may notice that, even with such low power supply V_{DD}, CML can only support up to 13 Gb/s while using double the number of lines. It is also worth noting that the losses at double the f_{min} for CML are still low enough to provide a 300-mV swing.

V. Methodology for 2.5-D Interface Energy–Pitch Optimization

In an MCS, the routing resources are also critical. The whole system costs should be reduced especially the energy/bit and the routing pitch. The question of energy–routing pitch combined optimization is interesting as it attempts to find middle optima between the two performance metrics that are correlated. Especially, for silicon interposer-based 2.5-D chip-to-chip interfaces, the routing pitch dependent upon the width of the line is critical due to higher losses. In order to save the routing area, the energy/bit and pitch should be minimized together with a combined metric ψ(pJ/bit*mm) to find an optimal for the two performance objectives.

A. Holistic Multichip ψ Minimization Methodology

A holistic design flow for 2.5-D silicon interposer-based multichip interface is presented here. This flow expands the previous work by Hatamkhani and Yang [9], Hatamkhani *et al.* [24], and Hatamkhani and Yang [25] and attempts to find the optimum topology for minimum energy/bit and routing pitch combined. It derives the optimum width W, spacing S, and signaling topology T for a given aggregate bandwidth F. The design flow used in this work is shown in Fig. 10.

For any integration technology and the substrate available, the width W is restricted by W_{min} and the spacing restricted by the S_{min}, which leads to the minimum signaling pitch ρ_{min} in the given interface integration platform. For the case of single-ended signaling with routing in the format GSGSG, where W is the signal width, and W_{GND} is the ground

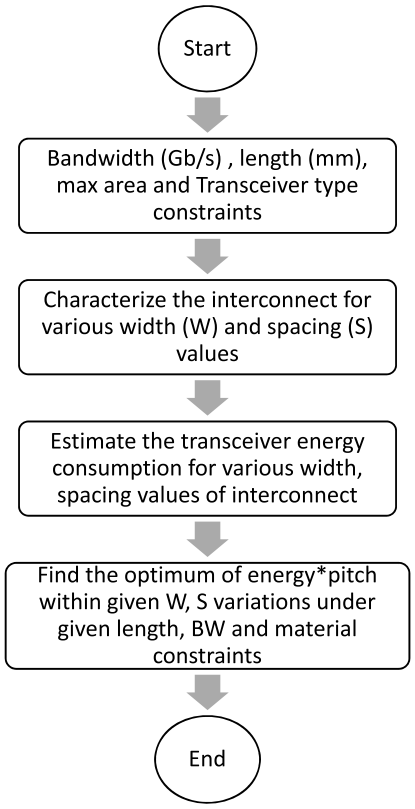


Fig. 10. Energy/bit*pitch ψ optimization methodology [27].

linewidth, then pitch ρ can be expressed as

$$\rho = W + W_{\text{GND}} + 2S. \tag{31}$$

For an example case, when the ground linewidth is set to minimum possible in the integration technology, the minimum signaling pitch $\rho_{\min}$ can be written as

$$\rho_{\min} = W + W_{\min} + 2S. \tag{32}$$

Therefore, for this signal ended signaling case with minimum width of ground and transceiver topology $T_i \in T$, the energy–pitch efficiency metric ψ is

$$\psi(T_i, \rho) = \frac{\phi}{f_b}(W + W_{\min} + 2S). \tag{33}$$

It is worth noting that one can also use an alternative metric of (energy/bit)/(bandwidth/mm) [15]. This is the same as cost metric ψ but normalized to f_b, where f_b is bit-rate per interconnect in the link.

The design flow shown in Algorithm 1 is run in its basic form through all the possible combinations of transceiver topology in the given set T, possible interconnect width W, and spacing W possibilities to find the optimum cost ψ solution consisting of T_{opt}, W_{opt}, and S_{opt} for a given aggregate bandwidth f_b. This algorithm is based on the assumption that the transmitter and receiver equalization settings or tap values are calculated based on the pulse response. The tap values or equalization settings are used to calculate the energy per bit and the cost metric ψ, which is computed in an iterative manner until the minimum is obtained, which corresponds to the optimal solution. Herein, the frequency is kept constant. Changing this would lead to a change in the amount of equalization and, hence, different energy efficiencies for the same line pitch. This would require reiteration of the algorithm. A faster approach would be to characterize the equalization and energy cost for different line pitch over a range of frequencies during the internal loop of the algorithm, itself.

Algorithm 1 Holistic MCS Communication Interface Design Flow

Result: Optimum solution $T_{\text{opt}}, W_{\text{opt}}, S_{\text{opt}}$
define width range: $W = \{W_{\min}, \ldots, W_{\max}\}$
define spacing range: $S = \{S_{\min}, \ldots, S_{\max}\}$
define Transceiver types: $T_i \in T$
define data bit rate: f_b
define interconnect average length: L
initialize ψ_{old}
while $T_i \in T$ **do**
  for $W \leq W_{max}$ **do**
    for $S \leq S_{max}$ **do**
      Find S-parameters for given W, S
      Find pulse response for given f_b
      Find required number of Taps for Tx
      Find required number of DFE Taps for Rx
      calculate power consumption in Tx, Rx:
      $\phi_{\text{Tx}} = \left[\phi_{\text{Drv}} + \phi_{\text{Eq}} + \phi_{\text{Ser}} + \phi_{\text{Ckbuf}}\right]$
      $\phi_{\text{Rx}} = \left[\phi_{\text{buf}} + \phi_{\text{Eq}} + \phi_{\text{DeSer}} + \phi_{\text{Ckbuf}}\right]$
      calculate signalling pitch:
      $\rho = W + W_{\min} + 2S$
      calculate interface energy-area cost:
      $\psi = \frac{\phi}{f_b}(W + W_{\min} + 2S)$
      if $\psi < \psi_{\text{old}}$ **then**
        $T_{\text{opt}} = T_i, W_{\text{opt}} = W, S_{\text{opt}} = S$
      end
      update $\psi_{\text{old}} = \psi$
    end
  end
end

It also should be noted that the length is a variable in the algorithm. Changing this changes the dependence of the total cost on the signaling topology and the interconnect parameters. As an example, consider a very short interconnect less than 1-mm length. Here, the width-dependent line capacitance is more significant than the equalization cost. Hence, minimum width and spacing could result in minimum energy/bit*pitch cost. However, for longer lines (5–10 mm), the losses in very thin lines are higher, requiring more profound equalization. Hence, the best energy/bit*cost could be even for the largest width interconnect or an intermediate value. Hence, the average length is kept a variable in the algorithm to signify this effect. Finally, for heterogeneous integration, the chiplets must be optimized in terms of silicon technology. The algorithm then needs to be extended to calculate the energy costs for all available technologies for different width values and correspondingly different equalization requirements.

In order to appreciate the basic operation of the above algorithm, consider an example of silicon interposer with a low resistivity of 100 $\Omega \cdot$cm and a dielectric constant of 11.9, as shown in Fig. 11 [27]. Here, two metal layers are present

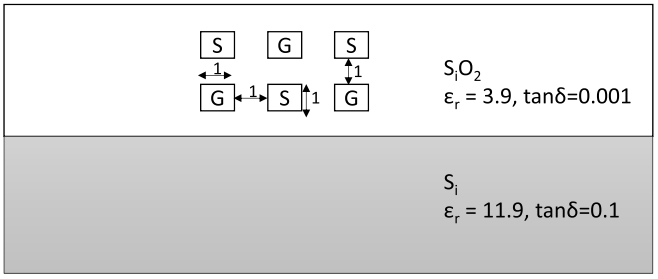


Fig. 11. Stackup for silicon interposer-based MCS.

around the insulator $S_i O_2$. Consider the length L to be 10 mm. For simplicity, consider the data rate per unit line to be 10 Gb/s only and the range of width variation to be 1–2 μm, while the spacing is kept constant to 1 μm, as shown in Fig. 11. The insertion loss variation by width variation is shown in Fig. 12. At the Nyquist frequency of 5 GHz for 10-Gb/s signaling, the insertion loss dependent upon frequency for the 2-μm-wide line is 7 dB higher than for 1-μm line. In addition, there is a 4-dB higher dc loss in the 1-μm-wide line, which means a reduced voltage swing at the Rx input.

For evaluating a channel regarding its insertion loss, the pulse response method is used in general [3]. It consists of sending a pulse with a width equal to one unit interval of 100 ps at 10 Gb/s on one end of the channel. Both ends of the channel are terminated properly with typical 50-Ω impedance to avoid any reflection, which is not analyzed in this case. A simulation is performed in the Hyperlynx tool for pulse response analysis of 1- and 2-μm-wide lines. The pulse has a bit rate of 10 Gb/s and a rise time of 1 ps. The simulation results are shown in Fig. 13, where the x-axis is normalized to one unit interval in order to observe the intersymbol interference with previous or later data bits.

The signal rises completely within 1-UI for both 1- and 2-μm-wide lines. This means that there is no precursor ISI with previous bits. Nevertheless, both lines cause the signal to extend further into later bits. The response of 2- and 1-μm line drops to zero after about 3-UI and 2-UI, respectively. For cancellation of these postcursor ISI, either a continuous-time linear equalizer (CTLE) or a decision feedback equalizer is used in receiver design [24]. For complete cancellation of ISI, a high-impedance peaking in CTLE or a large number of DFE taps will be required. If a DFE is used, two decision-feedback equalizer (DFE) taps will be required for a 2-μm line at the receiver end to cancel the second and third UI ISIs, while only one DFE tap shall be required for the 1-μm line to cancel the second-UI ISI.

Palaniappan and Palermo [24] demonstrated a method to estimate the power consumption of receiver circuit based upon the equalization value in dB for CTLE and the number of taps for DFE. It was shown that CTLE is used only up to 12-dB equalization, while, for higher values, a DFE is preferred. As in our case, though the total loss is much higher, the frequency-dependent insertion loss (total-dc) S21(dB) is less than 12 dB for both lines. Therefore, only CTLE equalization is used in this case. Based on this estimation methodology and 0.1-mW/Gb/s power per 6-dB CTLE equalization, extra

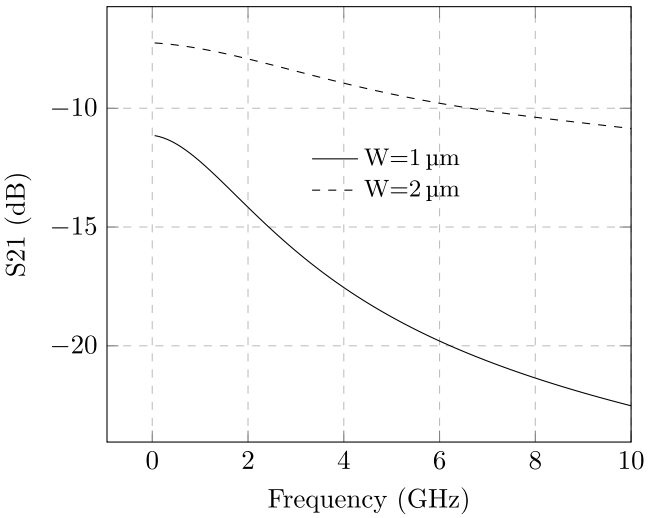


Fig. 12. S-parameters extracted using HSPICE 2-D field solver.

1-mW power ϕ_{Rx} is consumed by the receiver circuit interface with 1-μm-wide interconnect.

Ignoring other blocks in transmitter and receiver for basic understanding of this algorithm, the energy–pitch metric ψ is calculated based on just the front-end driver, receiver, and equalization blocks. Even as the power consumption for 1-μm interface increases, the energy–pitch metric ψ is still less by 0.1pJ/bit $\cdot$ μm for 1-μm line interface compared to the 2-μm interface. This means that, for combined energy–area or energy–pitch performance of a multichip interface, the 1-μm-wide line interface shall still be the better choice.

This methodology is rigorous though computationally complex. One can reduce its time complexity through a number of approaches. One can keep a check on the number of T_x and R_x equalization taps calculated within each loop. As the width (W) increases, there will be a point at which the equalization will not be required and power will not decrease any further. The width increment loop can be stopped now with minimum spacing as this would be the minimum energy/bit*pitch cost. Alternatively, one can remove the spacing (S) loop in the algorithm, especially if crosstalk and differential impedance are not a problem. The third approach could be to remove the step of equalization taps calculation. Rather the channel loss for each possible width value of interconnect can be calculated at the start, and only the dB equalization requirement would be determined. Hence, no pulse response analysis would be required. A final approach to reduce the complexity could be to use the si-FoM shown by Jangam and Iyer [15]. Combining factors, such as silicon cost, number of IO rows, and length of the link with energy/bit*pitch cost function, could result in a faster choice of technology and topology.

B. CML IO

In Section V-A, on holistic design flow, the width variation is the principal guiding factor. However, the spacing between interconnects is not considered. In another example being shown here, the spacing shall also play an important role along with the width of the interconnect and be critical to the opti-

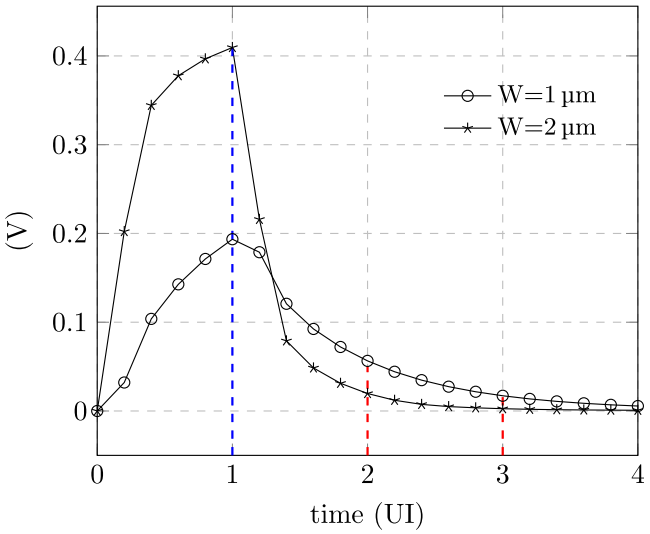


Fig. 13. Response for 10-Gb/s input pulse with 1-ps rise time.

mization of energy–pitch ψ metric. CML signaling topology is used for this example and a high resistivity of 10 000-$\Omega \cdot$ cm silicon substrate with low tan δ of 0.001 [32]. This example is based on the following consideration for energy–pitch minimization of chip-to-chip communication interface:

The increment in spacing between differential pairs in CML signaling should lead to higher impedance and lower power consumption; however, it will also increase the signaling pitch. The vice versa is true for the decrement in spacing. The methodology presented here combines the ψ energy–pitch analysis to optimize the width, spacing, and power consumption.

This section uses the holistic energy–pitch minimization methodology to optimize the design of CML transmitter front end and interposer for optimum energy and area efficiency [29]. Generally, the impedance of the CML drivers is designed as 50 Ω for PCB-based systems. However, when these circuits are used for data transmission between chips in 2.5-D integrated technology, then they can use the high-impedance design to lower the required current (I) for a given required voltage swing. The main drawback of CML drivers is their data rate independent power consumption. Regardless of the frequency of data transmission, the static I_{bias} current flows through the driver. Therefore, power consumption in CML drivers can be defined as $V_{\text{DD}} \cdot I_{\text{bias}}$.

Consider a stack shown in Fig. 14, which contains two metal layers of copper in S_iO_2 dielectric over a silicon substrate. A coplanar architecture is considered in which a differential pair is surrounded by ground lines for shielding purposes and has ground lines under it on the lower metal layer, all separated with a constant spacing (S).

The goal of the codesign is to investigate the performance of this coplanar architecture with different width and spacing values and then to find the W and S values for which minimum energy*pitch (Enpitch) cost is achieved at the maximum possible effective −3-dB bandwidth (BW_{eff}). Once the 2-D field solver has extracted the RLGC model for all possible values of W and S, then the first thing to do is to find out the odd

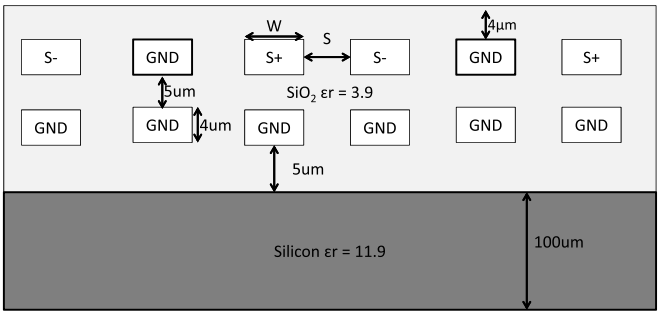


Fig. 14. Interposer stackup used for simulation.

mode impedance Z_{odd} for each W and S value combination. Z_{odd} can be calculated using the following equation:

$$Z_{\text{odd}} = \sqrt{\frac{L_o - L_m}{C_o + 2C_m}} \tag{34}$$

where L_o, L_m, C_o, and C_m represent the self and mutual inductances and capacitances, respectively.

This equation shows that, if the mutual coupling capacitance between the interconnects of a differential pair for CML signaling is increased due to reduced spacing between the pair lines, the odd mode impedance shall decrease. The impedance is inversely related to the power consumption for a given voltage swing and shall lead to poor energy performance.

Based upon the basic transmission line theory, the insertion loss of the differential pair is dependent upon both the conductor R and dielectric G conductance [29]

$$\alpha = \alpha_C + \alpha_D \tag{35}$$

where α_C represents the conductor resistance loss and α_D represents the dielectric conductance loss. For simplicity of the example, two assumptions are being made. One is that the dielectric loss is extremely small, as the dielectric conductance factor G of the RLGC model is not significant due to the extremely high resistivity of 10 000-$\Omega \cdot$ cm silicon substrate. Therefore, the attenuation over the line will only be due to the conductive losses due to interconnect resistance R. The second assumption is the low loss assumption of differential pair such that the inductive behavior of the transmission line is much larger than the resistive behavior. Similarly, the capacitive coupling is much larger than the dielectric conductance. This can be shown as the general transmission line propagation constant γ as [33]

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{36}$$

which can be expressed as

$$\gamma = j\omega\sqrt{LC}\sqrt{\left(1 + \frac{R}{j\omega L}\right)\left(1 + \frac{G}{j\omega C}\right)}.$$

Using the second assumption of low loss transmission line, $R \ll j\omega L$, and $G \ll j\omega C$, the Taylor series expansion leads to

$$\gamma = \frac{1}{2}\left(R\sqrt{\frac{C}{L}} + G\sqrt{\frac{L}{C}}\right) + j\omega\sqrt{\text{LC}} = \alpha + j\beta.$$

Hence, the attenuation α in Nepers/meter is

$$\alpha = \frac{1}{2}\left(\frac{R}{Z_{\text{odd}}} + \text{GZ}_{\text{odd}}\right).$$

Using the first assumption that the dielectric conductance loss is extremely small, the attenuation in Nepers/meter becomes

$$\alpha = \frac{R}{2Z_{\text{odd}}}.$$

Since 1 Neper = 8.686 dB, then attenuation factor α in decibels for 10-mm line differential pair can be calculated by the following equation:

$$\alpha_{\text{dB}} = 8.686\left[\frac{(R_o + R_s\sqrt{f})/100}{2Z_{\text{odd}}}\right] \tag{37}$$

where R_o and R_s are the dc resistance and the skin effect resistance factor values, respectively.

By plotting and finding the −3-dB bandwidth BW_{ch} for each W and S configuration, 10%–90% rise time tr_{ch} of the link interconnect [34] can be calculated by

$$\text{tr}_{\text{ch}} = \frac{0.35}{\text{BW}_{\text{ch}}}. \tag{38}$$

For perfect matching of driver along with channel and receiver impedance, an approximate effective or final rise time tr_{tot} needs to be calculated. With a single-pole very short length RC channel assumption for the interposer, a first-order estimation of the rise time tr_{tot} at the receiver end can be performed as [35]

$$\text{tr}_{\text{tot}} = \sqrt{(\text{tr}_{\text{Tx}})^2 + (\text{tr}_{\text{channel}})^2 + (\text{tr}_{\text{Rx}})^2}.$$

Since the output impedance of T_x and input impedance of R_x is equal to channel impedance Z_{odd} for best matching, their 10%–90% rise time is equal to $2.2RC$ or $2.2Z_{\text{odd}}C_{\text{pad}}$. Hence, the final rise time to input amplifier at receiver end is

$$\text{tr}_{\text{tot}} = \sqrt{(2.2C_{\text{pad}}Z_{\text{odd}})^2 + \left(\frac{0.35}{\text{BW}_{\text{ch}}}\right)^2 + (2.2C_{\text{pad}}Z_{\text{odd}})^2}. \tag{39}$$

This can be further simplified to

$$\text{tr}_{\text{tot}} = \sqrt{9.68(C_{\text{pad}}Z_{\text{odd}})^2 + \left(\frac{0.35}{\text{BW}_{\text{ch}}}\right)^2}. \tag{40}$$

Then, the total or final bandwidth BW_{tot} can be calculated using the inverse of (38) and can be written as

$$\text{BW}_{\text{tot}} = \frac{0.35}{\sqrt{9.68(C_{\text{pad}}Z_{\text{odd}})^2 + \left(\frac{0.35}{\text{BW}_{\text{ch}}}\right)^2}}. \tag{41}$$

The next step is to find the power and signal routing pitch cost for each configuration. For the CML driver, as shown in Fig. 5, the power consumed is only static, which can be calculated as the product of supply voltage V_{DD} and I_{bias}. For a given voltage swing V_{SW} requirement, the current required is $V_{\text{SW}}/Z_{\text{odd}}$, which is equal to half of the CML driver bias current, i.e., $I_{\text{bias}}/2$. The signal pitch for such coplanar configuration is

$$\rho = 3 \times (S + W).$$

The power consumption ϕ is product of current I_{bias} and voltage supply V_{DD}

$$P = I_{\text{bias}} \cdot V_{\text{DD}} = \frac{2V_{\text{DD}}V_{\text{SW}}}{Z_{\text{odd}}}.$$

Therefore, the final metric for our codesign methodology is energy*pitch/bit given by (42). It is worth noting that this is the same equation as (11); however, with detailed forms of power and pitch calculations for CML differential pair signaling on interposer

$$\psi = \frac{\phi}{\text{BW}_{\text{tot}}} * \rho = \left(\frac{2V_{\text{DD}}V_{\text{SW}} * 3(S + W)}{Z_{\text{odd}}\text{BW}_{\text{tot}}}\right). \tag{42}$$

The calculated attenuation values using (37) are plotted in Fig. 15. Odd mode differential impedance Z_{diff} is plotted in Fig. 16. It shows that, with an increasing spacing of metal lines, the inductance increases, which results in increased

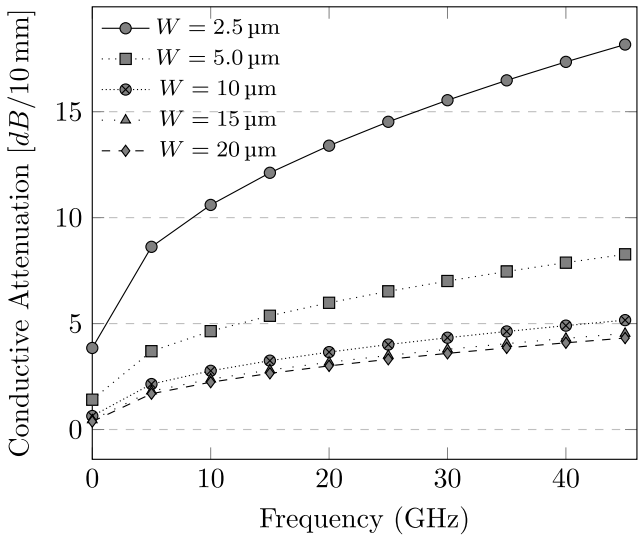


Fig. 15. Attenuation versus width for S = W.

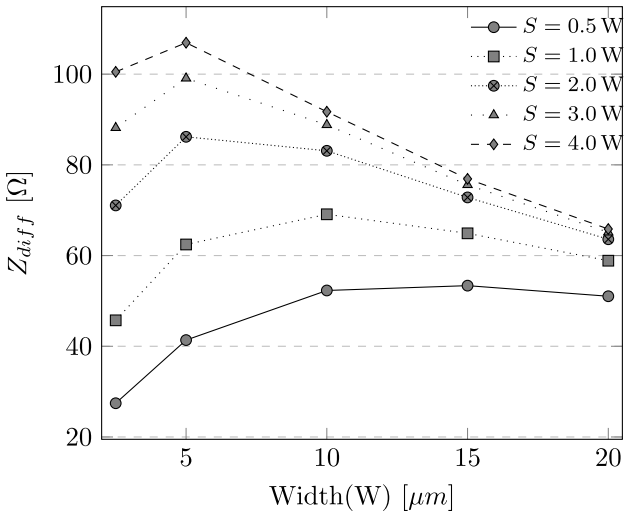


Fig. 16. Z_{diff} versus width (W) and spacing (S).

TABLE I
SUMMARY OF CONSTRAINTS AND RESULTS

Length	Width	Spacing	V_{sw}	ϵ_r	$tan\delta$	BW (GHz)	Z_{diff} (Ω)	Energy/bit * pitch
10 mm	2.5-20 μm	1-80 μm	300 mV	3.9	.001	0.24-13.3	27-107	83-368 $pJ/bit * \mu m$

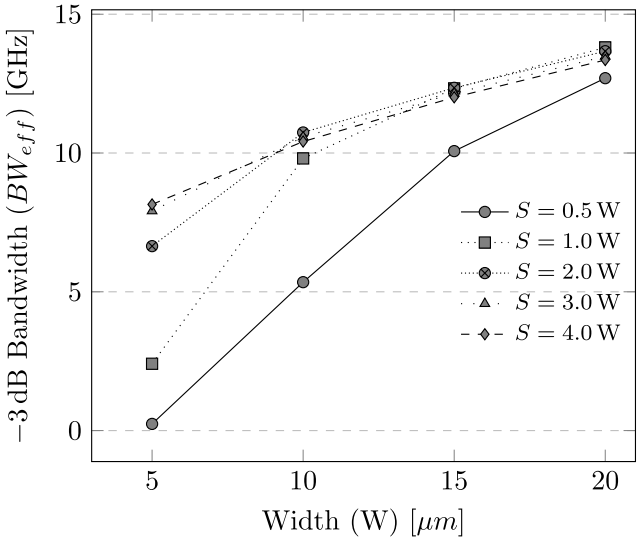


Fig. 17. −3-dB bandwidth variation with width (W) and spacing (S).

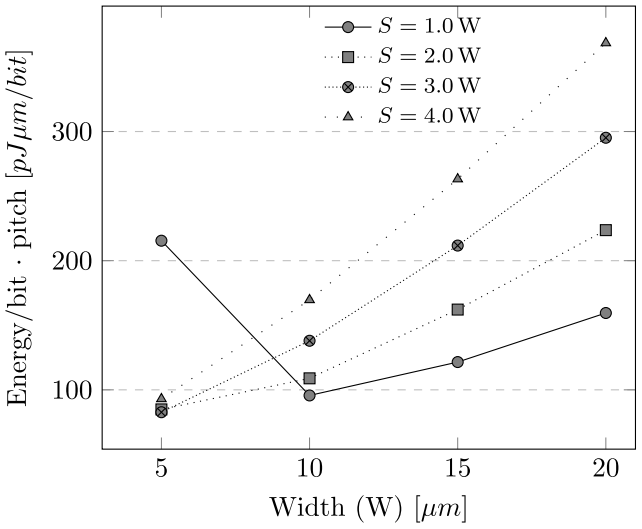


Fig. 18. Energy*pitch variation with width (W) and spacing (S).

impedance. With an increment in width, the capacitance increases, which reduces the impedance. As can be seen in the plot, Z_{diff} reaches a peak at 5-μm width but decreases with further increments in width. One conclusion from this plot is that, at 5-μm width, differential impedance is maximum for the given stackup, which could lead to the lowest current requirement in the CML driver design.

In order to calculate the bandwidth of driver and receiver, pad capacitance is selected as 0.2 pF to meet the JEDEC ESD requirements [36], [37]. The −3-dB effective bandwidth for the whole path from T_x to R_x is plotted in Fig. 17, which shows that the bandwidth increases with increasing width and spacing. However, this will increase the area cost of the design. Therefore, a combined energy/bit*pitch metric is needed for optimum configuration selection, as plotted in Fig. 18. Power supply value V_{DD} is 1.8 V, and the required V_{SW} is 300 mV. It can be seen from the plot that the cost metric reaches a minimum for 10-μm width with 10-μm spacing and supports the −3-dB bandwidth of 10 GHz. Utilizing the energy–area cost analysis, knowledge of the optimum spacing and width of lines simplified the signaling interface design. Table I summarizes the design constraints and results for this design optimization procedure. For example, a chip to be designed should have an edge length of 3 mm for maximum bandwidth and minimum power. Then, using the derived optimum 10-μm width and spacing for 50 CML differential pairs, running at 10 Gb/s results in optimum energy and area cost bandwidth of 500 Gb/s.

VI. CONCLUSION

Chiplet-based approach for future highly integrated systems at package or silicon interposer level is essential to More than Moore realization. Energy per bit optimization based on the channel constraints is generally performed but for highly miniaturized MCSs, where the routing resources are essential to the system-level optimization, and a codesign approach is required. This work presents holistic energy/bit*pitch objective-based system-level signaling topology, channel width, and spacing estimation. The CML 2.5-D signaling is used as an example for the holistic approach, which takes advantage of a detailed analysis presented for different signaling topologies with respect to equalization and noise-derived voltage swing requirements.

REFERENCES

[1] S. Arora, D. Bouvier, and C. Weaver, "AMD next generation 7 nm Ryzen 4000 APU, 'renoir,"' in *Proc. IEEE Hot Chips Symp. (HCS)*, Aug. 2020, pp. 1–30.

[2] A. L. S. Loke *et al.*, "Analog/mixed-signal design challenges in 7-nm CMOS and beyond," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Apr. 2018, pp. 1–8.

[3] G. Balamurugan, B. Casper, J. E. Jaussi, M. Mansuri, F. O'Mahony, and J. Kennedy, "Modeling and analysis of high-speed I/O links," *IEEE Trans. Adv. Packag.*, vol. 32, no. 2, pp. 237–247, May 2009.

[4] V. Stojanovic and M. Horowitz, "Modeling and analysis of high-speed links," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2003, pp. 589–594.

[5] K. S. Oh *et al.*, "Accurate system voltage and timing margin simulation in high-speed I/O system designs," *IEEE Trans. Adv. Packag.*, vol. 31, no. 4, pp. 722–730, Nov. 2008.

[6] W. T. Beyene *et al.*, "Signaling analysis and optimal channel data rates for high-performance FPGA interfaces," in *Proc. IEEE 27th Conf. Electr. Perform. Electron. Packag. Syst. (EPEPS)*, Oct. 2018, pp. 17–19.

[7] H. Hatamkhani and C.-K. Ken Yang, "A study of the optimal data rate for minimum power of I/Os," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 11, pp. 1230–1234, Nov. 2006.

[8] S. Kiran *et al.*, "Modeling of ADC-based serial link receivers with embedded and digital equalization," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 9, no. 3, pp. 536–548, Jul. 2019.

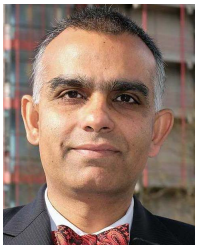
[9] G. Balamurugan and N. Shanbhag, "Modeling and mitigation of jitter in high-speed source-synchronous interchip communication systems," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2003, pp. 1681–1687.
[10] G. Balamurugan and N. Shanbhag, "Modeling and mitigation of jitter in multiGbps source-synchronous I/O links," in *Proc. 21st Int. Conf. Comput. Design*, 2003, pp. 254–260.
[11] W. T. Beyene *et al.*, "Advanced modeling and accurate characterization of a 16 Gb/s memory interface," *IEEE Trans. Adv. Packag.*, vol. 32, no. 2, pp. 306–327, May 2009.
[12] M. A. Karim, P. D. Franzon, and A. Kumar, "Power comparison of 2D, 3D and 2.5D interconnect solutions and power optimization of interposer interconnects," in *Proc. IEEE 63rd Electron. Compon. Technol. Conf.*, May 2013, pp. 860–866.
[13] S. Jangam, S. Pal, A. Bajwa, S. Pamarti, P. Gupta, and S. S. Iyer, "Latency, bandwidth and power benefits of the SuperCHIPS integration scheme," in *Proc. IEEE 67th Electron. Compon. Technol. Conf. (ECTC)*, May 2017, pp. 86–94.
[14] Y. Zhang, X. Zhang, and M. S. Bakir, "Benchmarking digital die-to-die channels in 2.5-D and 3-D heterogeneous integration platforms," *IEEE Trans. Electron Devices*, vol. 65, no. 12, pp. 5460–5467, Dec. 2018.
[15] S. Jangam and S. S. Iyer, "A signaling figure of merit (s-FoM) for advanced packaging," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 10, no. 10, pp. 1758–1761, Oct. 2020.
[16] J. W. Poulton *et al.*, "A 1.17-pJ/b, 25-Gb/s/pin ground-referenced single-ended serial link for off- and on-package communication using a process- and temperature-adaptive voltage regulator," *IEEE J. Solid-State Circuits*, vol. 54, no. 1, pp. 43–54, Jan. 2019.
[17] B. Dehlaghi and A. Chan Carusone, "A 0.3 pJ/bit 20 Gb/s/wire parallel interface for die-to-die communication," *IEEE J. Solid-State Circuits*, vol. 51, no. 11, pp. 2690–2701, Nov. 2016.
[18] T. O. Dickson *et al.*, "A 1.8 pJ/bit $16 \times 16Gb/s$ source-synchronous parallel interface in 32 nm SOI CMOS with receiver redundancy for link recalibration," *IEEE J. Solid-State Circuits*, vol. 51, no. 8, pp. 1744–1755, Aug. 2016.
[19] T. O. Dickson *et al.*, "A 1.4 pJ/bit, power-scalable 16×12 Gb/s source-synchronous I/O With DFE receiver in 32 nm SOI CMOS technology," *IEEE J. Solid-State Circuits*, vol. 50, no. 8, pp. 1917–1931, Aug. 2015.
[20] T. O. Dickson *et al.*, "An 8× 10-Gb/s source-synchronous I/O system based on high-density silicon carrier interconnects," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 884–896, Apr. 2012.
[21] A. Tajalli *et al.*, "A 1.02-pJ/b 20.83-Gb/s/wire USR transceiver using CNRZ-5 in 16-nm FinFET," *IEEE J. Solid-State Circuits*, vol. 55, no. 4, pp. 1108–1123, Apr. 2020.
[22] H. Hatamkhani, F. Lambrecht, V. Stojanovic, and C.-K. K. Yang, "Power-centric design of high-speed I/Os," in *Proc. 43rd ACM/IEEE Design Automat. Conf.*, E. M. Sentovich, Ed. Piscataway, NJ, USA: IEEE, Jul. 2006, p. 867.
[23] H. Hatamkhani and C.-K. K. Yang, "Power analysis for high-speed I/O transmitters," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2003, pp. 142–145.
[24] A. Palaniappan and S. Palermo, "A design methodology for power efficiency optimization of high-speed equalized-electrical I/O architectures," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 21, no. 8, pp. 1421–1431, Aug. 2013.
[25] D. Menin, A. De Pra, A. Bandiziol, W. Grollitsch, R. Nonis, and P. Palestri, "A simple simulation approach for the estimation of convergence and performance of fully adaptive equalization in high-speed serial interfaces," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 9, no. 10, pp. 2079–2086, Oct. 2019.
[26] J. Craig, "A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," in *Proc. MILCOM*, vol. 2, Nov. 1991, pp. 571–575.
[27] M. W. Chaudhary, A. Heinig, and B. Choubey, "Energy-area aware channel design for multi-chip interfaces," in *Proc. IEEE 29th Conf. Electr. Perform. Electron. Packag. Syst. (EPEPS)*, Oct. 2020, pp. 1–3.
[28] F. D. Mbairi, W. P. Siebert, and H. Hesselbom, "High-frequency transmission lines crosstalk reduction using spacing rules," *IEEE Trans. Compon., Packag., Technol.*, vol. 31, no. 3, pp. 601–610, Sep. 2008.
[29] M. W. Chaudhary and A. Heinig, "Co-design of CML IO and interposer channel for low area and power signaling," in *Proc. IEEE 19th Int. Symp. Design Diag. Electron. Circuits Syst. (DDECS)*, Apr. 2016, pp. 1–6.
[30] M. W. Chaudhary, A. Heinig, and B. Choubey, "13-Gb/s transmitter for bunch of wires chip-to-chip interface standard," in *Proc. IEEE 63rd Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2020, pp. 333–336.
[31] Y.-C. Bae *et al.*, "A 1.2 V 30nm 1.6Gb/s/pin 4Gb LPDDR3 SDRAM with input skew calibration and enhanced control scheme," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2012, pp. 44–46.
[32] R.-Y. Yang, C.-Y. Hung, Y.-K. Su, M.-H. Weng, and H.-W. Wu, "Loss characteristics of silicon substrate with different resistivities," *Microw. Opt. Technol. Lett.*, vol. 48, no. 9, pp. 1773–1776, 2006.
[33] M. B. Steer, *Microwave and RF Design: A Systems Approach*. Raleigh, NC, USA: SciTech Pub, 2010.
[34] E. Bogatin, *Signal Integrity Simplified* (Prentice-Hall Modern Semiconductor Design Series). Upper Saddle River, NJ, USA: Prentice-Hall, 2004.
[35] W. C. Elmore, "The transient response of damped linear networks with particular regard to wideband amplifiers," *J. Appl. Phys.*, vol. 19, no. 1, pp. 55–63, 1948.
[36] M. Huo *et al.*, "A case study of problems in JEDEC HBM ESD test standard," *IEEE Trans. Device Mater. Rel.*, vol. 9, no. 3, pp. 361–366, Sep. 2009.
[37] G. S. Alliance. (Jan. 2015). *Electrostatic Discharge (ESD) in 3D-IC Packages Version 1.0*. Accessed: Jun. 2, 2021. [Online]. Available: https://www.3dincites.com/wp-content/uploads/GSA-ESDA-3D-IC_ESD_Whitepaper_1.pdf

Muhammad Waqas Chaudhary was born in Islamabad, Pakistan, in 1986. He received the B.S. degree in electronics engineering from the GIK Institute, Topi, Pakistan, in 2008, the M.S. degree in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2013, and the Ph.D. degree in electrical engineering from the University of Siegen, Siegen, Germany, in 2021.

Since 2014, he has been a Researcher with Fraunhofer Institute IIS/EAS, Dresden, Germany. He has presented his work at several conferences. His research interests include system-level design optimization, methodology development, and circuit design for low-power high-speed parallel wireline communication circuits for extremely short distances.

Andy Heinig was born in Dresden, Germany, in 1975. He received the Diploma degree in information technology from the Technical University of Cottbus, Cottbus, Germany, in 2006.

He then joined the Fraunhofer Institute IIS/EAS, Dresden, Germany. Since 2013, he has been leading the Working Group of Advanced Packaging, Fraunhofer IIS/EAS. In this position, he started several Saxony, German, and European research projects in the field of design and design automation for advanced packing. He is also working in different standardization groups, e.g., Si2. His research is in the field of assembly design kits and assembly design flows for advanced packaging where he has presented several papers at different conferences around the world.



Bhaskar Choubey (Senior Member, IEEE) received the D.Phil. degree from the University of Oxford, Oxford, U.K., in 2007, and the B.Tech. degree from the Regional Engineering College, Warangal, India, in 2002.

He was an Associate Professor with the University of Oxford and a Lecturer with the University of Glasgow, Glasgow, U.K. He is currently the Chair of the Department of Analog Circuits and Image Sensors, Siegen University, Siegen Germany. His research interests are in analog and mixed-signal designs, CMOS image sensors, microelectromechanical sensors, hardware intelligent circuits, and nonlinear dynamics.

Mr. Choubey is also a fellow of the Institution of Engineering and Technology. He received the IEEE Sensors Council GOLD Early Career Achievement Award and the Myril B. Reed Best Paper Award from the IEEE Midwest Symposium of Circuits and Systems. He has served as the 2019 Chair of the IEEE EPPC Working Group on Information and Communication Technology (ICT) in Europe. He was an Associate editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II. He is also an Associate Editor of the IEEE SENSORS JOURNAL.