Information Theoretic Modeling and Analysis for Global Interconnects With Process Variations

Stojan Z. Denic, *Member, IEEE*, Bane Vasic, *Senior Member, IEEE*,
Charalambos D. Charalambous, *Senior Member, IEEE*, Jifeng Chen, *Student Member, IEEE*, and
Janet M. Wang, *Senior Member, IEEE*

Abstract—As the CMOS semiconductor technology enters nanometer regime, interconnect processes must be compatible with device roadmaps and meet manufacturing targets at the specified wafer size. The resulting ubiquitous process variations cause errors in data delivering through interconnects. This paper proposes an Information Theory based design method to accommodate process variations. Different from the traditional delay based design metric, the current approach uses achievable rate to relate interconnect designs directly to communication applications. More specifically, the data communication over a typical interconnect, a bus, subject to process variations (“uncertain” bus), is defined as a communication problem under uncertainty. A data rate, called the achievable rate, is computed for such a bus, which represents the lower bound on the maximal data rate attainable over the bus. When a data rate applied over the bus is smaller than the achievable data rate, a reliable communication can be guaranteed regardless of process variations, i.e., a bit error rate arbitrarily close to zero is achievable. A single communication strategy to combat the process variations is proposed whose code rate is equal to the computed achievable rate. The simulations show that the variations in the interconnect resistivity could have the most harmful effect regarding the achievable rate reduction. Also, the simulations illustrate the importance of taking into account bus parasitic parameters correlations when measuring the influence of the process variations on the achievable rates.

Index Terms—Achievable rate, bit error rate process variations, global interconnect, information theory.

I. INTRODUCTION

PROPER design of on-chip and off-chip interconnects appears as one of the major issues in the development of new VLSI technologies [1]–[3]. Since interconnect networks are one of the key parts of new integrated circuit (IC) paradigms [such as network-on-chip (NOC) and multi-core processors], unreliable interconnects would affect the reliability of an IC as a whole.

Manuscript received November 11, 2007; revised April 28, 2008. First published December 11, 2009; current version published February 24, 2011. This work was supported in part by the NSF under Grant CCF-0634969 and ITR-0325979. The work of C. D. Charalambous was supported by the Research Promotion Foundation of Cyprus under the grant ARTEMIS PLHROx0505/24.

S. Z. Denic was with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721 USA. He is now with the Telecommunications Research Lab, Toshiba Research Europe Limited, BS1 4ND, Bristol, U.K. (e-mail: stojan.denic@toshiba-trel.com).

B. Vasic, J. Chen, and J. Wang are with the Electrical and Computer Engineering Department, University of Arizona, Tucson, AZ 85721 USA (e-mail: vasic@ece.arizona.edu; chenjf@ece.arizona.edu; wml@ece.arizona.edu).

C. D. Charalambous is with the Electrical and Computer Engineering Department, University of Cyprus, Nicosia 1678, Cyprus (e-mail: chadcha@ucy.ac.cy).

When a feature size of technology is below 65 nm, about 40% of on-chip global interconnects may have performance variations larger than 3σ [1], [4]. For nontraditional nanoscale electronics, such as molecular or self-assembly, uncertainty in the interconnect operation is present as well. This uncertainty originates from the inherent stochastic switching behavior of interconnects. Hence, interconnects will have more process variability and thereby more nonuniform behavior across a chip. In addition to nonuniform behavior, it is expected that future nanoscale interconnects will be operating at reduced noise margins due to the low power requirement. As a consequence, an error-free transfer of bits along the interconnects cannot be further considered as granted.

The uncertainty introduced by process variability represents a fundamentally new challenge for IC design. From the point of view of interconnects, the presence of uncertainty due to process variability means that interconnect parasitic parameters may deviate from their nominal values depending on, for example, the interconnect location on the chip. When an interconnect is understood as a communication channel, this means that interconnect communication channel characteristics can vary in an unpredictable fashion making the task of reliable data communication very difficult. In information theory, this problem is called a communication subject to uncertainty. In this paper, to address the problem of communication over uncertain interconnects, we propose an information theoretical technique known as universal encoding/decoding which is introduced shortly.

There are currently two main complementary directions used in coping with the uncertainties due to the process variations. One applies the VLSI design tools (such as wire spacing, layering, shielding, buffer insertion, staggering, metal fill, and accurate litho simulation [5]–[12]), while the other suggests the implementation of signal processing techniques such as error correction and constrained coding. We say that these two approaches are complementary since only one of them without the other cannot resolve the problem of reliable data transmission over uncertain interconnects. For example, although the VLSI design tools could be effective in reducing the effects of crosstalk, radiation and process variations, they cannot completely remove the uncertainty, and therefore they cannot solely provide the reliable transmission over the interconnects.

The techniques the VLSI designers use to deal with the uncertainty can be summarized as follows: 1) compensate for it later; 2) reduce the uncertainty itself; 3) reduce the uncertainty effect; and 4) guardband against uncertainty (for more details see [13]). Major limitations of the “compensate for it later” approach (which includes a post-process metal fill) are: correc-

tions are not completely doable, it destroys hierarchy since it is used at the end of design, and post-process behavior is itself uncertain. “Reduce the uncertainty itself” method implements tools such as shielding and buffer insertion to compensate for the uncertainty directly. Shielding reduces the crosstalk between interconnects, while a buffer compensates for the signal attenuation along the interconnects. The limitations of these methods are: they require a detailed designer knowledge and the amount of uncertainty reduction is not very clear. “Reduce the uncertainty effect” methods apply techniques that will cover as many different working conditions as possible. This approach is robust to many variations but the drawback is that design tools are not well developed. “Guardband against uncertainty” is the most commonly used approach. Its main weakness is that it can be too pessimistic from the performance point of view if the design parameters are chosen too conservatively.

Previously mentioned approaches and methodologies are usually either technology dependent or implementation dependent. Further, the improvements of the CMOS fabrication technology can hardly help reduce the variability because the lithograph accuracy is approaching its sub-wavelength limit. While there is no clear consensus on how far and how fast the current copper/aluminium interconnect technology will continue to scale and which of the emerging technologies eventually enters production, it is certain that future nanotechnologies will have significant manufacturing defect rate.

Besides the above mentioned solutions, the interconnect reliability can be achieved by using an elegant, technology and circuit implementation independent alternative: error-correction and constrained coding. As shown by [14], coding has two important advantages over repeater/buffer insertion. First, coding allows the data flow in both directions, whereas buffers in general limit buses to operate in one direction. Second, instead of increasing wire delays and power consumption like buffer insertion, coding reduces both of them during the data transmission.

Error-correction and constrained coding has been earlier proposed as solutions for several problems for on-chip and off-chip interconnects. More specifically, error-correction and constrained codes are used to minimize the crosstalk between different lines of the bus [15], correct noise corrupted bits, improve energy efficiency of the data transmission and deal with some combination of the previous problems [16]–[22].

An assumption made by previous work on error-correction and constrained coding for IC interconnects is that the communication channel is perfectly known to the encoder and decoder. However, as discussed above, one cannot expect the perfect knowledge of interconnect parameters leading to the uncertainty regarding its corresponding communication channel characteristics. This can result in poor performance of proposed interconnect communication schemes.

The main goal of this paper is to address the problem of conveying the data over uncertain interconnects for new VLSI technologies, focusing on the example of a specific interconnect, i.e., a bus. Since, this is in essence the problem of data communication over uncertain channel, our approach applies the tools of Communication and Information Theory and error control coding. We are particularly interested in computing the quantity, which is called the achievable data rate R_a measured in bits per

second. The achievable data rate R_a is a rigorously defined measure of transmitted information, and its importance stems from the fact that it provides a lower bound on the data rate for reliable transmission. If the data transmission rate R over the interconnect is smaller than R_a , the reliable transmission of data can be guaranteed [23]. This implies the use of error-correction codes at the transmitting end of the bus, and decoding at the receiving end to protect the data against the bus variability, crosstalk and noise. The stress is on the word “reliably”, which means that if $R < R_a$ then each data bit will be delivered to the other side of interconnect with small probability of error, i.e., small bit error rate (BER). In fact, a BER can be made arbitrarily small subject to certain conditions that will be explained shortly. For the problem considered in this paper, R_a gives a designer a sense of how many bits can reliably be conveyed over the bus, regardless of its variability.

The main contributions of this paper are as follows.

- 1) First, the bus will be modeled as an RC interconnect network and described as a compound multiple-input-multiple-output (MIMO) communication channel [24]. A compound model is standard in information theory. It is applied to address the issues of nonuniform communication channel behavior. Thus, it is an appropriate model to capture the variability of interconnects across the IC. Instead of representing the bus by only one fixed transfer function, the bus will be described by the set of all possible transfer functions determined by the variability of bus parasitic parameters (R and C) that depend on the manufacturing process. This set of all possible transfer functions represents one unique model called “Compound Bus” (CB). The exact definition of the communication over a compound channel or CB is the following: from the known set of all possible buses pick one randomly and transmit the information over that unknown bus. Thus, the set of buses is known but the particular bus over which the transmission is carried out is unknown.
- 2) A MIMO communication channel is used to capture the crosstalk between bus lines.
- 3) For this bus model, the achievable data rate R_a shall be computed by using the mutual information (MI) [23].
- 4) We propose a practical data transmission strategy, which accompanied with error-correction codes, enables reliable transmission over the CB. Moreover, the proposed strategy guarantees the existence of a single code (so-called universal code) which can be used over all buses from the CB [25]. This approach mitigates a bus design and lowers the cost of solution since it uses only one encoding/decoding procedure for all buses on the chip regardless of their variability.
- 5) In addition, we identify a number of problems related to data communication along the uncertain IC interconnects which are of interest for future research.

This paper is organized as follows. Section II provides a background description of bus codes, data code rate, and bus communication channel model. Section III introduces the variability modeling for the bus interconnect. Section IV shows how the achievable data rates can be computed by using the information theoretic concept of the MI. Section V proposes practical

Fig. 1. p Bus lines with encoder and decoder.

communication strategy for attaining the achievable data transmission rate. Section VI gives an example of a CB modeling and computing achievable data transmission rate. Section VII considers several practical error correction schemes and their suitability for the CB in terms of their complexity. Section VIII concludes this work with an elaboration with future research directions.

II. BACKGROUND

A. Notation

Fig. 18 lists all the notations and parameter definitions used in this paper. $\mathbf{Z}$ denotes the set of integers. $\mathbf{R}$ denotes the set of real numbers. $\mathbf{C}$ denotes the set of complex numbers. x^* denotes the complex conjugate of the complex number x . x^T denotes the transpose of the vector $x \in \mathbf{C}^p$. $\|x\|$ denotes the Euclidean norm of the vector x . I_p denotes the identity matrix of size p . A^H denotes the complex conjugate transpose of matrix A . $\bar{\sigma}(A)$ denotes the largest singular value of matrix A . $E[x]$ denotes the expected value of the random variable x . $\text{Var}[x]$ denotes the variance of the random variable x . $x(k) * h(k)$ denotes the operation of convolution between two signals $x(k)$ and $h(k)$. $X(z) \triangleq Z(x(t))$ denotes the z -transform of a discrete-time signal $x(k)$. $X(e^{j\theta})$ denotes the discrete Fourier transform of a discrete-time signal $x(k)$.

B. Basic Concepts

Consider a bus structure with p parallel interconnects (see Fig. 1) as a MIMO communication channel. A MIMO model is used to take into account crosstalk among p lines of the bus. This model is supported, for instance, by the work done in [22].

Ideally, each line of the bus should convey a pulse representing a bit of information without any distortion as shown in Fig. 2(a). Here, $x(t)$ is the input signal to one line of the bus, while $y(t)$ is its output. Corresponding discrete-time representations of the continuous-time signals $x(t)$ and $y(t)$ are given in Fig. 2(b). (Notice that while in the VLSI circuit design research area, data propagation, and data are regarded as continuous in time ($t \in \mathbf{R}$), in this case data information is represented by a discrete-time model ($k \in \mathbf{Z}$) which is consistent with the sampling of data bits at the bus end.)

While traveling along the bus, each pulse conveying a bit of information, experiences different types of distortions such as attenuation, noise, intersymbol interference, and crosstalk. Fig. 2(c) illustrates one type of distortion caused by parasitic parameters of the bus (such as capacitance). Here, one impulse at the bus input at $k = 1$ invokes a train of impulses at the bus output. Hence, instead of having the zero output at the moments

Fig. 2. Continuous-time versus discrete-time bus representation.

$k > 1$ due to the zero input $x_2(k)$ for $k > 1$, the output consists of the train of impulses. This kind of output of a system caused by an impulse is called the impulse response. Non-ideal impulse response introduces a distortion into transmitted signal called inter-symbol interference. For example, imagine that the next transmitted bit is 0 represented by $x_2(2) = 0$. Then, since $y_2(2)$ is different from 0 (because of the impulse response caused by $x_2(1) = 1$), the transmitted bit $x_2(2)$ can be erroneously recognized as 1. If, in addition, an additive noise is added to the output, the reliable transmission of information along the bus becomes even more complicated.

One way to protect data transmitted over the bus, is by employing an error-correction code. Roughly speaking, the role of an error-correction code is to add redundant bits on the top of data bits which enables a decoder at the other end of the bus (see Fig. 1) to extract useful information regardless of distortions. By introducing additional, redundant bits, high reliability can be achieved which effectively reduces the data rate over the bus. This is a price that has to be paid in order to reduce the number of erroneously transmitted data bits, i.e., the BER. Next, the process of encoding is described more formally.

In Fig. 1, an encoder maps the sequence of binary vectors $\{d(k)\}_{k \geq 0}$ (useful data) to $\{x(k)\}_{k \geq 0}$ which is sent over the bus. $d(k)$ is defined by $[d_1(k), \dots, d_p(k)]^T$, where $d_i(k) \in \{0, 1\}$, $1 \leq i \leq p$. The input to the bus structure is the set of data $x(k) \triangleq [x_1(k), \dots, x_p(k)]^T$, of p bits ($x_i(k) \in \{0, 1\}$, $1 \leq i \leq p$). Here, $x_i(k)$, $1 \leq i \leq p$ carries voltage values that may be interpreted as logic value 0 or 1. The output is a vector $y(k) \triangleq [y_1(k), \dots, y_p(k)]^T$. The mapping between sequences $\{d(k)\}_{k \geq 0}$ and $\{x(k)\}_{k \geq 0}$, is carried out in the following way. The encoder takes a block of K data vectors $d(k)$, $k \in \{0, 1, \dots, K-1\}$, $\mathbf{d} = [d(0) \dots d(K-1)]^T$ which is called a source word and maps it into another block of length N , $\mathbf{x} = [x(0) \dots x(N-1)]^T$ which is called a codeword. The introduced redundancy is $p(N - K)$ bits. Here, both $\mathbf{d}$ and $\mathbf{x}$ are represented as discrete-time data vectors.

To measure the encoder's performance, we use the code rate defined as

$$R \triangleq \frac{\log_2 2^{Kp}}{NT_s} = \frac{Kp}{NT_s}. \quad (1)$$

Code rate R has unit bits per second, and T_s is a sampling interval in seconds. A smaller value T_s indicates faster transmission speed for data. The physical meaning of code rate R is obvious. Since there are K vectors of size p in the source word, the number of all possible transmitted messages is 2^{Kp} . The number of bits transmitted is therefore $\log_2 2^{Kp}$. A code rate R is the ratio between the number of useful data bits sent and the overall time NT_s required for this information to be sent. If $\log_2 2^{Kp}$ is divided only by N , this represents the number of bits sent per one vector, or one sample as it is often called. The higher the code rate R is, the more efficient the encoder is. One of the goals of this paper is to find an upper bound for R when the bus is subject to variability. This quantity is called the achievable data rate R_a which is a lower bound for the Shannon's channel capacity [23]. If we know R_a then we will know what code to choose for a particular channel. If $R < R_a$, the probability that a sent data bit, after decoding, will be erroneous can be made arbitrarily small when the codeword length N becomes very large.

Another important performance measure is the power used for data transmission. Here

$$\frac{1}{N} \sum_{k=0}^{N-1} \|x(k)\|^2 \leq P \quad (2)$$

defines the average power used by the transmitted data $x(k)$. This power is limited by P , a user defined bound. In general, we prefer to limit and minimize the power during the data transmission.

As demonstrated in Fig. 1, a vector sequence $\{x(k)\}_{k \geq 0}$ is the input of the bus. A sequence of matrices $\{h(k)\}_{k \geq 0}$ in the time domain, $h(k) \in \mathbf{R}^{p \times p}$, represents the impulse responses of the bus. Entry $h_{ij}(k)$ in matrix $h(k)$ with $1 \leq i, j \leq k$ models how i th input $x_i(k)$ affects the j th output $y_j(k)$. Taking an additive noise $n(t)$ into consideration, the output of the bus can be expressed as

$$y(k) = \sum_{j=0}^k h(k-j)x(j) + n(k). \quad (3)$$

The first term $\sum_{j=0}^k h(k-j)x(j)$ is the convolution of $x(k)$ and $h(k)$, while the second term $n(k)$ is the additive white Gaussian noise in our case. The choice of Gaussian noise in this case is justified as follows. In a deep-submicrometer environment, a large number of sources contribute towards overall noise that affects the detection of bits at the end of the bus. Therefore, it is difficult to characterize all noise sources and their distributions. That is why, we use the central limit theorem to represent the distribution of the overall noise by a Gaussian distribution as it has been done in [14]. Even, if this assumption is not quite correct, from the information theoretical point of view, the Gaussian noise represents the worst case noise in a sense that such assumption produces the lower bound on achievable rate R_a .

If there is no interference between neighboring symbols $x(k)$ (i.e., $h(1) = \dots = h(k) = 0$), the bus is described by

$$y(k) = h(k)x(k) + n(k). \quad (4)$$

A corresponding presentation in the frequency domain can be obtained by taking z -transform of (3) giving

$$Y(z) = H(z)X(z) + N(z) \quad (5)$$

where $X(z) \triangleq Z(x(k))$, $Y(z) \triangleq Z(y(k))$, $N(z) \triangleq Z(n(k)) \in \mathbf{C}^p$ and $H(z) \triangleq Z(h(t)) \in \mathbf{C}^{p \times p}$. The z -transform is defined by

$$X(z) \triangleq \sum_{k=0}^{\infty} x(k)z^{-k}. \quad (6)$$

By substituting $z = e^{j\theta}$ into (5), we get discrete Fourier transform (DFT) representation

$$Y(e^{j\theta}) = H(e^{j\theta})X(e^{j\theta}) + N(e^{j\theta}) \quad (7)$$

where $\theta \in [-\pi, \pi]$ is the normalized frequency; $\theta \triangleq \omega T_s$, and ω is the continuous-time frequency. ω is related to the continuous-time Fourier transform defined by

$$X(j\omega) \triangleq \int_{-\infty}^{+\infty} x(t)e^{-j\omega t} dt \quad (8)$$

where $x(t)$ is the continuous-time counterpart of the discrete-time (sampled) signal $x(k)$. When $-\omega_N \leq \omega \leq \omega_N$ ($\omega_N \triangleq \pi/T_s$ is the Nyquist frequency), θ changes from $-\pi$ to π . Consequently, the i th output is given by

$$Y_i(e^{j\theta}) = H_{ii}(e^{j\theta})X_i(e^{j\theta}) + \sum_{k \neq i} H_{ik}(e^{j\theta})X_k(e^{j\theta}) + N_i(e^{j\theta}) \quad (9)$$

where $X_i(e^{j\theta})$, $Y_i(e^{j\theta})$, $N_i(e^{j\theta})$, and $H_{ij}(e^{j\theta})$ are entries in $X(e^{j\theta})$, $Y(e^{j\theta})$, $N(e^{j\theta})$, and $H(e^{j\theta})$, respectively. Here, (5) reveals three sources of uncertainty. One is the process variation induced uncertainty on transfer function $H(e^{j\theta})$, the other is the white noise and the third is given by the third term on the right-hand side of (9) modeling the crosstalk. In the following section, the uncertainty modeling will be considered in more detail.

III. VARIABILITY MODELING

As discussed in the earlier sections, we can model the bus by its transfer function $H(z)$. Without loss of generality, assume that the noise spectrum $N(z)$ remains the same across the bus. The process variations, however, may cause the transfer function $H(z)$ to vary from one location to the other along the bus and throughout the IC. In this section, first it is shown how the transfer function matrix $H(z)$ can be constructed in terms of the resistance-capacitance (RC) interconnect network to model the bus, and secondly, the model, which includes bus variability and utilizes previously constructed transfer function matrix $H(z)$, is introduced.

A. Nominal Bus Model

To model a bus interconnect, we use the RC network model found in [16], which is briefly described next. Introduce the matrix $T(s)$ given in Laplace domain by

$$T(s) \triangleq \begin{bmatrix} Y & Z & 0 & \cdots & 0 & 0 \\ Z & X & Z & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & X & Z \\ 0 & 0 & 0 & \cdots & Z & Y \end{bmatrix}. \quad (10)$$

Here

$$X = 1 + sR_w(2C_c + C_g) \quad (11)$$

$$Y = 1 + sR_w(C_g + C_c) \quad (12)$$

$$Z = -sR_wC_c \quad (13)$$

where R_w is the interconnect wire resistance, C_g is the interconnect wire-to-substrate capacitance, and C_c is the inter-wire capacitance called the coupling capacitance. These parameters are computed as follows:

$$R_w \triangleq \frac{\rho L}{w_w t_1} \quad (14)$$

$$C_g \triangleq \frac{\epsilon L w_w}{h} \quad (15)$$

$$C_c \triangleq \frac{2\epsilon L t_1}{d}. \quad (16)$$

ρ is the resistivity of copper, w_w is the wire width, d is the separation between two wires, t_1 is the wire height, h is the height of the wire above the substrate, ϵ is the permittivity of the oxide between wires, and L is the length of the interconnect. A short section of the bus (of length L) can be described by the MIMO transfer function matrix

$$H(s) \triangleq T^{-1}(s). \quad (17)$$

The transfer function matrix of longer interconnect is obtained by multiplying transfer function matrices $H(s)$ of shorter interconnects.

The variability of the bus characteristics as seen through the variability of the transfer function matrix $H(s)$ depends on the variability of parasitic parameters R_w , C_g , and C_c . According to [4], in a $0.1\text{-}\mu\text{m}$ technology, these parameters may vary as follows:

$$\rho = \rho \pm \Delta\rho = 2.2 \times 10^{-8} \pm 30\% \Omega\text{m} \quad (18)$$

$$d, w_w = (d, w_w) \pm \Delta(d, w_w) = 237 \times 10^{-9} \pm 20\% \text{ m} \quad (19)$$

$$t_1, h = (t_1, h) \pm \Delta(t_1, h) = 498 \times 10^{-9} \pm 15\% \text{ m}. \quad (20)$$

To illustrate and better understand how the process variability affects the parasitic parameters R_w , C_g and C_c , Fig. 3 is given which depicts two bus interconnects with parameters that define their geometry. The bus on the right is an ideal interconnect, while one on the left suffers from the erosion. The erosion is one type of process variabilities that occurs in the case of dense array thin interconnects, and it is characterized by the oxide and metal removal. The erosion results in effective decrease of the

Fig. 3. Example of process induced bus variability.

wire height t_1 leading to an increase of the wire resistance R_w and decrease of the coupling capacity C_c . This shows that the process variability can cause correlated variations of the parasitic parameter values. As noted in [4], the examples of parameters' correlations are as follows.

- 1) Given a fixed wire pitch, the variation of the separation between two wires d is negative of the variation of the wire width w_w .
- 2) The wire height t_1 and the height of the wire above the substrate h are negatively correlated with a correlation coefficient of -0.5 .

The effect of previous correlations are taken into account in the case-study given in Section VI.

B. Compound Bus Model

In this section, we introduce the model which has been widely used in control and information theory to describe the variability in the frequency domain [25]–[27]. The model relies on the H^∞ normed linear space and $\|\cdot\|_\infty$ norm (for precise definitions see the Appendix). The $\|\cdot\|_\infty$ norm is defined by using the singular value decomposition (SVD) of the matrix. In essence, any matrix $H \in \mathbb{C}^{p \times p}$ assumes a decomposition $H = U\Sigma V^H$, where $U \in \mathbb{C}^{p \times p}$ and $V \in \mathbb{C}^{p \times p}$ are unitary matrices, and $\Sigma \in \mathbb{C}^{p \times p}$ is a diagonal matrix with positive entries $\{\sigma_i\}_{i=1}^r$ representing the singular values of H . When H is a transfer function matrix of the communication channel, the entries of Σ can be understood as the attenuation coefficients of the channel, while U and V define the phase. When H is scalar, the analogy is even more clear since the singular value is equivalent to the modulus $|H|$ of the transfer function $H = |H|e^{j\phi_H}$. To understand why the SVD is a natural way to model uncertainty, notice that the Shannon's capacity of communication channels defined by (3) is directly related to the singular values of H [28].

Now, since the manufacturing process variability determines the variability of the bus parasitic parameters such as capacitance and resistance, it determines the variability of the transfer function matrix $H(z)$ and the set of all possible frequency responses $H(e^{j\theta})$ of the bus. Therefore, the model, which takes into account the manufacturing process variability, is the set of all possible transfer function matrices $H(e^{j\theta})$. One way to describe such a set (call it A_2) is to represent it by a ball in the frequency domain by using H^∞ space and $\|\cdot\|_\infty$ norm such as (21), shown at the bottom of the next page. The ball A_2 is centered at a known transfer function $H_{nom}(e^{j\theta})$, the so-called nominal transfer function. The size of the ball is determined by a perturbation $w(e^{j\theta})\Delta(e^{j\theta})$. The perturbation is defined by: 1) an unknown matrix $\Delta(e^{j\theta})$ that takes care of the uncertainty in the

Fig. 4. Nyquist plot of variability modeling.

phase because $\|\Delta\|_\infty \leq 1$ and 2) a known scalar transfer function $w(e^{j\theta})$ that defines A_2 radius. To understand why $w(e^{j\theta})$ defines the radius of the ball A_2 observe that a ball in the complex domain is given by

$$|z - z_0| \leq r, \quad z, z_0 \in \mathbf{C}. \quad (22)$$

Since we deal with matrices, to define the ball we need to use matrix maximum singular value $\bar{\sigma}(\cdot)$. Then the ball in the frequency domain is given by

$$\begin{aligned} \bar{\sigma}(H(e^{j\theta}) - H_{nom}(e^{j\theta})) &\leq |w(e^{j\theta})| \bar{\sigma}(\Delta(e^{j\theta})) \\ &\leq |w(e^{j\theta})| \|\Delta\|_\infty \\ &\leq |w(e^{j\theta})| \end{aligned} \quad (23)$$

since $\|\Delta\|_\infty = \sup_{\theta \in [0, 2\pi)} \bar{\sigma}(\Delta(e^{j\theta})) \leq 1$, and $\bar{\sigma}(AB) \leq \bar{\sigma}(A)\bar{\sigma}(B)$. At each frequency, the size of the set A_2 is determined by the radius $|w(e^{j\theta})|$.

To further clarify this point, we refer to Fig. 4 which shows the Nyquist plot of a scalar transfer function $H_{nom}(e^{j\theta})$. A Nyquist plot is a imaginary part versus real part of a transfer function. Each point on the graph corresponds to one frequency θ . The length of the vector connecting the origin and any point on the graph corresponds to a transfer function modulus (attenuation) at some frequency θ , while the angle between the vector and the positive real line corresponds to a phase. The perturbation superimposes the variability at each frequency by a ball centered at the apex of the vector with the radius determined by $|w(e^{j\theta})|$. Thus, any point within the ball is possible because $\Delta(e^{j\theta})$ has an unknown phase and $|\Delta(e^{j\theta})| \leq 1$. Based on Fig. 4, A_2 can be seen as a tube around $H_{nom}(e^{j\theta})$.

The definition of A_2 also implies

$$\bar{\sigma}(H_{nom}) - |w| \leq \bar{\sigma}(H) \leq \bar{\sigma}(H_{nom}) + |w| \quad (24)$$

which gives an upper and lower bound on the singular values (attenuation coefficients) of the transfer function matrices that belong to A_2 . $H_{nom}(e^{j\theta})$ and $w(e^{j\theta})$ are appropriately chosen

according to (17) to describe all possible matrices $H(e^{j\theta})$. Usually, $H_{nom}(e^{j\theta})$ is determined by the nominal values of the parasitic parameters of the bus, while $|w(e^{j\theta})|$ follows from the parasitic parameters variations.

IV. ACHIEVABLE DATA RATE OF THE BUS

In this section, the formula for computing the achievable data rate is provided for the bus model described by (3) and (21). It is assumed that the transmitted data sequence $\{x(t)\}_{t \geq 0}$ is random as it is in reality. First, the definition of power constraint for random signals is introduced. Second, the formula for the achievable data rate is given by the maximin optimization problem where the objective function is the mutual information (MI) rate. Third, the solution of the maximin optimization problem is presented and the implications of solution are discussed.

A. Power Constraint for Random Signals

Since we deal with the random signals, the power constraint equivalent to (2) should be introduced. It is given in terms of the power spectral density (PSD) matrix of $x(t)$ denoted by $W_{\mathbf{x}}(\theta)$. PSD $W_{\mathbf{x}}(\theta)$ is related to its correlation function $R(k)$ through the inverse Fourier transform in the following way:

$$R(k) \triangleq E[x(t+k)x^H(t)] \triangleq \frac{1}{2\pi} \int_0^{2\pi} W_{\mathbf{x}}(\theta) e^{j\theta k} d\theta \quad (25)$$

or for $k = 0$

$$R(0) \triangleq E[x(t)x^H(t)] \triangleq \frac{1}{2\pi} \int_0^{2\pi} W_{\mathbf{x}}(\theta) d\theta. \quad (26)$$

Then, the power constraint is given by the set of PSD matrices

$$\begin{aligned} A_1 \triangleq & \left\{ W_{\mathbf{x}}(\theta) : \text{Trace} E[x(t)x^H(t)] = \int_0^{2\pi} \text{Trace}(W_{\mathbf{x}}(\theta)) d\theta \leq P \right\}. \end{aligned} \quad (27)$$

It is clear that

$$\text{Trace} E[x(t)x^H(t)] = \sum_{i=1}^m E[x_i^2(t)] \quad (28)$$

represents the average power of the transmitted data over all p bus interconnects.

B. Achievable Rate

The achievable rate subject to variability in the transfer function matrix $H(e^{j\theta})$ described by (21) is going to be related to the MI rate. A nice property of the MI rate formula used here is that it is given in the frequency domain. This enables us to use a frequency domain variability description, which is of a great practical importance. On the other hand, it should be observed that

$$A_2 \triangleq \{H(e^{j\theta}) \in H^\infty : H(e^{j\theta}) = H_{nom}(e^{j\theta}) + w(e^{j\theta})\Delta(e^{j\theta}), H, H_{nom}, w, \Delta \in H^\infty, \|\Delta\|_\infty \leq 1\} \quad (21)$$

the formula for the MI rate presented here makes an assumption regarding the input signal $x(k)$, i.e., the input sequence $\{x(k)\}_{k \geq 0}$ is a discrete-time Gaussian stochastic process which is different from a standard positive IC input consisting of 0 and 1 s. Yet another problem is that 0 and 1 s are uniformly distributed. This is the reason we introduce an additional assumption and give further explanations. To ensure $x(k) \geq 0$, we assume that $x_i(k)$, $1 \leq i \leq r \leq p$ have mean values $a_i = E[x_i(k)] > 0$ such that $a_i = c\sigma_{x_i}$, $\sigma_{x_i}^2 = \text{Var}[x_i]$. In our case, we choose $c = 4$ ensuring 99.99% positive values for $x(k)$. Further, the application of the Gaussian input is justified by the fact that a Gaussian distribution gives a tight lower bound for the achievable rates subject to a discrete uniform input for low to moderate SNRs (see [29]). This means, that the Gaussian input provides a good estimation for the achievable rate for the uniform input in the case of low SNRs. A low SNR condition is true for new ICs since it is implied by a low power supply demand.

In [27], the achievable rate subject to variability in the transfer function matrix $H(e^{j\theta})$ described by (21) is given by

$$R_a \triangleq \max_{W_{\mathbf{x}} \in A_1} \min_{H \in A_2} J(W_{\mathbf{x}}, W_{\mathbf{n}}) \quad (29)$$

where $J(W_{\mathbf{x}}, W_{\mathbf{n}})$ defined by

$$\frac{1}{4\pi} \int_0^{2\pi} \log \det(I_p + H(e^{j\theta})W_{\mathbf{x}}(\theta)H^H(e^{j\theta})W_{\mathbf{n}}^{-1}(\theta))d\theta \quad (30)$$

is the MI rate between $\{x(k)\}_{k \geq 0}$ and $\{y(k)\}_{k \geq 0}$. $W_{\mathbf{n}}(\theta)$ is the PSD matrix of the noise $n(k)$. The solution of (29) is given by [27]

$$\begin{aligned} \max \quad & \frac{1}{4\pi} \sum_{i=1}^r \int_0^{2\pi} \times \log \left(1 + \frac{W_{x_i}[(\sigma_i - |w|)^+]^2}{W_{n_i} + \sum_{k \neq i} |w|^2 W_{x_k}} \right) d\theta \\ \text{s.t.} \quad & \sum_{i=1}^r \int_0^{2\pi} W_{x_i} d\theta = P_{\mathbf{x}}. \end{aligned} \quad (31)$$

Here, $W_{x_i}(\theta)$ and $W_{n_i}(\theta)$, $1 \leq i \leq r \leq p$, are diagonal entries of PSD matrices $W_{\mathbf{x}}(\theta)$ and $W_{\mathbf{n}}(\theta)$, respectively. $r \leq p$ is the number of nonzero singular values of $H_{\text{nom}}(e^{j\theta})$ or the rank of $H_{\text{nom}}(e^{j\theta})$. $\{\sigma_i(e^{j\theta})\}_{i=1}^r$ are the singular values of $H_{\text{nom}}(e^{j\theta})$. Without loss of generality, it is assumed that $W_{\mathbf{n}}(\theta)$ is diagonal. If this is not the case, then $W_{\mathbf{n}}(\theta)$ and $H(e^{j\theta})$ can jointly be treated as an equivalent channel assuming that the PSD of the noise is an identity matrix.

The formula for the achievable rate suggests that the bus can be decomposed into the r parallel channels along the singular values of the channel matrix $H_{\text{nom}}(e^{j\theta})$ as shown in Fig. 5. The transmission along each singular value is possible if $\sigma_i(e^{j\theta}) - |w(e^{j\theta})|$ is positive, which is denoted by a notation $(\cdot)^+$. $(\cdot)^+$ takes the value of an argument if the argument is positive; otherwise, it is zero. Hence, the size of the set A_2 (or the uncertainty set as it is often called), $|w(e^{j\theta})|$, determines how the achievable rate is affected by the uncertainty in the manufacturing process. The larger the uncertainty $|w(e^{j\theta})|$, the smaller the achievable rate R_a . If there is no uncertainty (the size of the uncertainty set $|w(e^{j\theta})|$ is zero), (31) reduces to the classical capacity of the

Fig. 5. Interpretation of R_a in terms of parallel channels.

MIMO communication channel [28]. If $\sigma_i(e^{j\theta}) < |w(e^{j\theta})|$, the i th channel disappears (see Fig. 5) because the i th channel coefficient becomes zero.

The formula for the achievable rate R_a given by (31) is interesting for other reasons as well. Notice that (31) is given in terms of the logarithm of the signal-to-noise ratio (SNR), i.e., $\log(1 + \text{SNR}_i(\theta))$, where

$$\text{SNR}_i(\theta) \triangleq \frac{W_{x_i}(\theta)(\sigma_i(e^{j\theta}) - |w(e^{j\theta})|)^2}{W_{n_i}(\theta) + \sum_{k \neq i} |w(e^{j\theta})|^2 W_{x_k}(\theta)}, \quad 1 \leq i \leq r. \quad (32)$$

The numerator is the power of the received useful signal at the i th interconnect, and the denominator is the equivalent noise consisting of the white noise W_{n_i} for the i th interconnect, and the crosstalk $\sum_{k \neq i} |w|^2 W_{x_k}$ from other interconnects as shown in Fig. 5. Therefore, it follows that the crosstalk noise is zero if there is no uncertainty. However, if there is uncertainty, the channel gain will be diminished ($\sigma_i - |w|$), and, in the same time, the crosstalk noise will appear. In Section V, it will be demonstrated how it is possible to eliminate the crosstalk noise when there is no variability introduced by the manufacturing process.

Remark 4.1: How should (31) be interpreted? If the data transmission rate R is smaller than R_a , then there exists at least one encoding/decoding procedure (so called universal encoding/decoding) providing the probability of the decoding error arbitrarily close to zero regardless of the bus variability when the codeword length N becomes very large. In other words, the strength of the results stems from the fact that we can apply only one code for any bus that comes from the set A_2 , and at the same time, we can guarantee the probability of the decoding error arbitrarily close to zero. This is the simplest and effective solution that one can have for this type of problems. The pros and cons of the approach employed here will be discussed in more detail in the next section.

Remark 4.2: Although the Gaussian input is the most optimal input for the Gaussian noise, we cannot claim that the signals on the IC have Gaussian distribution. Therefore, the future research should address the problem of computation of the achievable rates for the same channel model, but for different input distributions such as uniform.

Fig. 6. Transmission scheme for uncertain channel matrix.

Remark 4.3: The unit of achievable rate given by (31) is bits per sample. Here, “sample” refers to the value of the signal taken at discrete-time moments which are T_s seconds apart. To get the achievable rate in bits per second, the value computed by (31) should be multiplied by $1/T_s$ samples/s.

V. ATTAINING ACHIEVABLE RATE

When a MIMO channel is known, an efficient communication is achieved by decomposing the MIMO channel into the r scalar independent, parallel channels, where r is the rank of the transfer function matrix. “Independent” means that there is no crosstalk among them. The idea is very simple; it consists of finding singular value decomposition of the transfer function matrix $H(e^{j\theta}) = U(e^{j\theta})\Sigma(e^{j\theta})V^H(e^{j\theta})$ and using unitary matrices $U(e^{j\theta})$ and $V(e^{j\theta})$ in processing the transmitted signal. Since, for the bus with variability we do not know the channel exactly, we apply the singular value decomposition on the nominal channel $H_{\text{nom}}(e^{j\theta}) = U(e^{j\theta})\Sigma(e^{j\theta})V^H(e^{j\theta})$. However, in this case the SVD does not produce independent parallel channels because of the uncertainty caused by variability. There will be some residual crosstalk that will be interpreted as an additional noise. Besides, the uncertainty will affect the channel coefficients, i.e., singular values, as explained in Section IV.

The strategy that attains the achievable rate (31) is the one that finds the worst-case communication channel representing the bus, and designs an encoder/decoder pair for this channel. The worst-case bus occurs when the channel coefficients, in terms of the singular values, take the smallest values for each frequency (the attenuation is the largest) and when the equivalent noise $N_{eq,i}(e^{j\theta})$ is the largest [see Section IV and (34)]. The reasoning behind this is the following: if the encoder and decoder performs well over the worst-case channel, it will perform even better over “better” channels, having larger channel coefficients and smaller noise. The rest of the section gives detailed description of the optimal communication scheme.

The communication scheme that attains achievable rate (31) is given in Figs. 6–8. First, the transmitted signal $x(k)$ is processed by the filter having the transfer function matrix $V(e^{j\theta})$ at the transmitter end of the bus (see Fig. 7), and processed by the filter $U^H(e^{j\theta})$ at its receiving end (see Fig. 8). $V_{li}(e^{j\theta})$, $1 \leq l, i \leq p$, are the entries of the l th row of $V(e^{j\theta})$, representing the scalar filters with impulse responses $v_{li}(k)$. $U_{li}(e^{j\theta})$, $1 \leq l, i \leq p$, are the entries of the l th row of $U(e^{j\theta})$, representing the scalar filters with impulse responses $u_{li}(k)$. Each encoder output $x_i(k)$, $1 \leq i \leq p$, is convolved with $v_{li}(k)$ to contribute to its bus input. At each bus output, received signal $y'_l(k)$ is convolved with $u_{li}^*(-k)$ to produce the l th output of the bus $y_l(k)$.

The matrices $U(e^{j\theta})$ and $V(e^{j\theta})$ can also be understood as a part of equivalent communication channel defined by

Fig. 7. Signal processing at the transmitter's end of the bus.

Fig. 8. Signal processing at the receiver's end of the bus.

$U^H(e^{j\theta})H(e^{j\theta})V(e^{j\theta})$. Then, encoder/decoder pair sees the equivalent channel given by

$$\begin{aligned} H_{eq}(e^{j\theta}) &= U^H(e^{j\theta})H(e^{j\theta})V(e^{j\theta}) \\ &= \Sigma(e^{j\theta}) + U^H(e^{j\theta})\Delta(e^{j\theta})V(e^{j\theta})w(e^{j\theta}) \\ &= \Sigma(e^{j\theta}) + (\Delta_{s,\text{diag}}(e^{j\theta}) + \Delta_{s,\text{off-diag}}(e^{j\theta}))w(e^{j\theta}) \end{aligned}$$

where $\Delta_s(e^{j\theta}) \triangleq U^H(e^{j\theta})\Delta(e^{j\theta})V(e^{j\theta})$, and $\Delta_s(e^{j\theta}) \triangleq \Delta_{s,\text{diag}}(e^{j\theta}) + \Delta_{s,\text{off-diag}}(e^{j\theta})$. $\Delta_{s,\text{diag}}(e^{j\theta})$ contains diagonal, while $\Delta_{s,\text{off-diag}}(e^{j\theta})$ contains off-diagonal entries of $\Delta_s(e^{j\theta})$, respectively. We can group further the first two terms, which are diagonal matrices, that determine the i th bus output as the function of the i th input solely. The third term is the crosstalk coming from the inputs different from the i th one. The output of the i th channel is represented in the frequency domain by

$$\begin{aligned} Y_i(e^{j\theta}) &= (\sigma_i(e^{j\theta}) + \delta_{s,ii}(e^{j\theta})w(e^{j\theta}))X_i(e^{j\theta}) \\ &\quad + \sum_{k \neq i} \delta_{s,ik}(e^{j\theta})w(e^{j\theta})X_k(e^{j\theta}) + N_i(e^{j\theta}). \end{aligned} \quad (33)$$

Here, $\delta_{s,li}(e^{j\theta})$, $1 \leq l, i \leq p$, are the entries of $\Delta_s(e^{j\theta})$. The off-diagonal elements $\delta_{s,li}(e^{j\theta})$, $l \neq i$, are absorbed into an equivalent noise

$$N_{eq,i}(e^{j\theta}) = N_i(e^{j\theta}) + \sum_{k \neq i} \delta_{s,ik}(e^{j\theta})w(e^{j\theta})X_k(e^{j\theta}). \quad (34)$$

Because $\|\Delta_s(e^{j\theta})\|_\infty \leq 1$, it follows $|\delta_{s,li}(e^{j\theta})| \leq 1$. From previous discussion, we can determine the worst-case communication channel representing the bus. The worst-case bus occurs when the channel coefficient $(\sigma_i(e^{j\theta}) + \delta_{s,ii}(e^{j\theta})w(e^{j\theta}))$ is the smallest for each frequency (the attenuation is the largest) and when the equivalent noise $N_{eq,i}(e^{j\theta})$ is the largest. This

will happen if $\delta_{s,ii}(e^{j\theta}) = e^{j\pi-j\phi_w}$ and $|\delta_{s,li}(e^{j\theta})| = 1$, $l \neq i$, where $w(e^{j\theta}) \triangleq |w(e^{j\theta})|e^{j\phi_w}$. Then, the channel coefficients become $\sigma_i(e^{j\theta}) = |w(e^{j\theta})|$, and the PSD of the noise is given by $W_{n_i}(\theta) + \sum_{k \neq i} |w(e^{j\theta})|^2 W_{x_k}(\theta)$. Thus, one can use r scalar error-correction codes, each tuned to the corresponding worst-case channel $\sigma_i(\theta) = |w(e^{j\theta})|$, and the equivalent noise. Since the codes work well for the worst case bus, it will work even better for other transfer function matrices from the set A_2 .

Remark 5.1: There are a couple of practical questions that should be addressed when the suggested method is applied to deal with the crosstalk and uncertainty.

The first question is the choice of practical error-correcting codes that are going to be employed. This depends on system parameters such as delay. Since the encoding and decoding procedures introduce the delay, the designer would want to use short codes and low-complexity encoders and decoders in order to minimize overall system delay. However, we leave this question for future research.

The other question is related to the cost of applied filters in terms of power. From the point of view of information theory, in [28], it is shown that applied filters do not affect the received SNR. In other words, the performance of the overall communication system will not be tarnished by introducing filters U^H and $V(e^{j\theta})$.

VI. CASE-STUDY OF BUS ACHIEVABLE RATES

The following example illustrates the impact of different types of uncertainties on the achievable transmission rate. It is assumed that the bus consists of four parallel interconnects. Although, this is not usually the case, we choose $p = 4$ since for larger p , the achievable rate computation becomes more numerically involved, and the choice of p does not change the conclusions we want to deliver.

Next, it will be demonstrated how it is possible to construct the uncertainty set A_2 based on the introduced MIMO model and values for parasitic parameters. To compute the nominal matrix $H_{\text{nom}}(s)$, we take the values $\rho = 2.2 \times 10^{-8} \Omega\text{m}$, $d = 237 \times 10^{-9} \text{ m}$ and $t_1, h = 498 \times 10^{-9} \text{ m}$, and use (10)–(17).

We also compute the transfer function matrices for three other scenarios: 1) $d = 237 \times 10^{-9} - 20\% \text{ m}$, $w_w = 237 \times 10^{-9} + 20\% \text{ m}$; 2) $d = 237 \times 10^{-9} - 10\% \text{ m}$; $w_w = 237 \times 10^{-9} + 10\% \text{ m}$ and $\rho = 2.2 \times 10^{-8} + 15\% \Omega\text{m}$; and 3) $t_1 = 498 \times 10^{-9} - 10\% \text{ m}$.

Denote the bus transfer function matrices for previous three scenarios by $H_c(s)$, $H_{rc}(s)$, and $H_e(s)$, respectively. They are going to be used to construct three examples of uncertainty sets A_2 .

In the simulations, we use $L = 12 \mu\text{m}$ and the relative permittivity of $\epsilon_r = 3.9$ [16].

Fig. 19 summarizes all parameters of transfer functions (tf) used in the case-study.

In all three scenarios, we take into account correlation among bus parameters. Scenario 1 shows the case when the space between two interconnects of the bus is smaller than the nominal, implying that the interconnect width is larger than the nominal. This uncertainty leads to the increase of C_g and C_c and decrease in R_w .

Fig. 9. Bode plots of the nominal bus and the bus with larger than nominal C_g and C_c capacitances having length of 400 sections [(1,1) entries of transfer function matrices].

Scenario 2, in addition to previous assumptions, treats the case when the resistivity ρ is larger by 15%. The main reason for the resistivity uncertainty is that the metal used for interconnect is not a pure material, but a compound with the main base of copper or aluminum. It is further surrounded by another thin layer as the barrier layer. The overall resistivity is the average of all these materials. Further, as the wire size is scaled down getting close to a metal grain size, the deposition process itself becomes harder to control. That is another reason causing a random variation of the resistivity. This uncertainty leads to the increase of all three bus parasitic parameters.

Scenario 3 simulates the erosion illustrated by Fig. 3. This uncertainty leads to the increase of R_w and decrease of C_c .

Figs. 9, 11, and 13 show the Bode plots of (1,1) entries of the $H_c(s)$, $H_{rc}(s)$, and $H_e(s)$, respectively, each accompanied with (1,1) entry of the $H_{\text{nom}}(s)$, where $s \triangleq j\omega$. It can be seen that $H_{rc}(s)$ and $H_e(s)$ introduce larger attenuation comparing to $H_{\text{nom}}(s)$. The same behavior is observed for all other diagonal entries of $H_{rc}(s)$ and $H_e(s)$ that describe how the i th entry affects the i th output of the bus. Further, Figs. 12 and 14 illustrate a typical characteristic of off-diagonal entries of $H_{rc}(s)$ and $H_e(s)$. Namely, off-diagonal entries of $H_{rc}(s)$ and $H_e(s)$ experience larger crosstalk than $H_{\text{nom}}(s)$. On the other hand, Figs. 9 and 10 show that the transfer function matrix $H_c(s)$ is almost identical with the nominal $H_{\text{nom}}(s)$. This can be understood by noticing that the scenario 1 has opposite effects on the parasitic parameters; although C_g and C_c are increased, R_w is decreased.

The previous analysis suggests that it is very important to take into account the correlation of bus parameters when considering the effect of the uncertainty on the bus performance. Also, it can be seen that the resistivity can play a major role in defining the bus performance subject to the uncertainty.

Further, we are going to deal with the 60-section-long bus, and all transfer functions refer to that particular bus. To obtain the transfer function matrices that correspond to discrete-time domain, denoted by $H_{rc}(z)$, $H_e(z)$, and $H_{\text{nom}}(z)$ (switching

Fig. 10. Bode plots of the nominal bus and the bus with larger than nominal C_g and C_c capacitances having length of 400 sections [(1,2) entries of transfer function matrices].

Fig. 12. Bode plots of the nominal bus and the bus with larger than nominal resistivity ρ having length of 60 sections [(1,2) entries of transfer function matrices].

Fig. 11. Bode plots of the nominal bus and the bus with larger than nominal resistivity ρ having length of 60 sections [(1,1) entries of transfer function matrices].

Fig. 13. Bode plots of the nominal bus and the bus subject to erosion having length of 60 sections [(1,1) entries of transfer function matrices].

from Laplace transform to Z transform), Tustin discretization of the continuous-time transfer function matrices $H_{rc}(s)$, $H_e(s)$, and $H_{nom}(s)$, is carried out. The output of the system is sampled with the sampling interval $T_s = 10^{-12}$ s. This sampling interval has been chosen since from Fig. 11, it can be noticed that attenuation of 10 dB is achieved between $\omega = 10^{11}$ rad/s and $\omega = 10^{12}$ rad/s.

The frequency response matrix from $H(z)$ is generated by substituting z with $e^{j\omega T_s} = e^{j\theta}$. When $-\omega_N \leq \omega \leq \omega_N \triangleq \pi/T_s$, then $-\pi \leq \theta \leq \pi$, where ω_N is the Nyquist frequency.

Now, the uncertainty set is constructed in the following way. To ensure that $H_{rc}(z)$ belongs to the uncertainty set A_2 , observe that

$$\bar{\sigma}(H_{rc}(e^{j\theta}) - H_{nom}(e^{j\theta})) = \bar{\sigma}(\Delta(e^{j\theta})w(e^{j\theta})) \leq |w(e^{j\theta})|.$$

It is chosen that the size of the uncertainty set $|w(e^{j\theta})|$ is equal to $\bar{\sigma}(H_{rc}(e^{j\theta}) - H_{nom}(e^{j\theta}))$ for each frequency $-\pi \leq \theta \leq \pi$, thus to its lower bound, which represents the distance from the nominal channel transfer matrix $H_{nom}(e^{j\theta})$ to the worst-case matrix $H_{rc}(e^{j\theta})$. Fig. 15 shows the frequency dependence of the singular values of $H_{nom}(e^{j\theta})$ accompanied with the size of uncertainty set $|w(e^{j\theta})|$. In fact, Fig. 15 depicts the magnitudes of two transfer functions $w_1(e^{j\theta})$ and $w_2(e^{j\theta})$ that correspond to the uncertainty sets constructed around $H_{rc}(e^{j\theta})$ and $H_e(e^{j\theta})$, respectively. For all frequencies $|w_1(e^{j\theta})| \geq |w_2(e^{j\theta})|$ demonstrating the larger distance from $H_{nom}(e^{j\theta})$ to $H_{rc}(e^{j\theta})$ than from $H_{nom}(e^{j\theta})$ to $H_e(e^{j\theta})$. The loss of achievable data rate due to the uncertainty can be anticipated by observing that the achievable rate depends on the difference between the singular values of $H_{nom}(e^{j\theta})$ and the magnitude of the uncertainty set $w(e^{j\theta})$. The contribution of each frequency to the achievable rate is measured by this

Fig. 14. Bode plots of the nominal bus and the bus subject to erosion having length of 60 sections [(1,2) entries of transfer function matrices].

Fig. 15. Singular values of H_{nom} and the sizes of uncertainty sets versus frequency θ .

difference. If for some frequency $\sigma_i(\theta) - |w(e^{j\theta})| < 0$, the contribution of this frequency is zero.

Fig. 16 shows the impulse responses of $\sigma_i(\theta)$, $\sigma_i(\theta) - |w_1(e^{j\theta})|$ and $\sigma_i(\theta) - |w_1(e^{j\theta})|$ $1 \leq i \leq 4$, which correspond to the equivalent channel representation for $H_{\text{nom}}(e^{j\theta})$, $H_{rc}(e^{j\theta})$, and $H_e(e^{j\theta})$. A main observation is that the uncertainty reduces the amplitude of the impulse response in the vicinity of $k = 0$ and broadens the impulse response for higher values of k . This affects the reliability of the transmission in two ways: 1) due to the lower amplitude of the impulse response in the vicinity of $k = 0$, the decision about the sent bit at $k = 0$ is more sensitive to the additive noise and 2) broaden impulse response introduces larger inter-symbol interference for subsequent bits making the decision regarding these bits less reliable (see Fig. 2). Therefore, to improve the performance, a communication technique for suppressing inter-symbol interference has to be employed.

Fig. 16. Impulse responses of the nominal and perturbed buses.

Fig. 17. Achievable rates for nominal bus and two CBs.

In Fig. 17, the achievable rates are plotted for the nominal channel, and for the CBs when the worst-case channel matrices are determined by $H_{rc}(e^{j\theta})$ and $H_e(e^{j\theta})$, respectively. The noise PSD matrix $W_n(\theta)$ is taken to be diagonal with diagonal entries being equal and equal to 0.04 that make sense for the power supply of $V_{DD} = 2$ V. The plots reveal noticeable reduction in the achievable rates when the uncertainty is determined by the variation in the resistivity ρ . For instance, for $\text{SNR} = 17.85$ dB, the achievable rate for the nominal channel is $R_a = 1.045 \times 10^{12}$ bit/s, while for the $H_{rc}(e^{j\theta})$ based CB is $R_a = 0.85 \times 10^{12}$ bit/s which corresponds to 15% loss. For $\text{SNR} = 20$ dB, the achievable rate for the nominal channel is $R_a = 1.3 \times 10^{12}$ bit/s, while for the $H_{rc}(e^{j\theta})$ based CB is $R_a = 1.04 \times 10^{12}$ bit/s corresponding to 23% loss.

Remark 6.1: The example reveals the weakness of universal encoding approach. If only one universal code is to be used, the price has to be paid in terms of the reduced achievable rate. Another alternative approach for NOC advocated in [3] is the

symbol	meaning
$x(k)$	bus input signal
$y(k)$	bus output signal
$h(k)$	impulse response
$n(k)$	additive noise
P	power constraint
R_a	achievable code rate
R	code rate
T_s	sampling interval
H_{nom}	nominal transfer function
$w(e^{j\theta})$	size of uncertainty set
σ_i	singular value of matrix H_{nom}
ρ	metal resistivity
w_w	wire width
d	separation between bus lines
t_1	wire height
h	height of wire above substrate
ϵ	permittivity of oxide between wires
L	length of bus
R_w	wire resistance
C_g	wire-to-substrate capacitance
C_c	coupling capacitance

Fig. 18. Notations and parameter definitions.

tf	parameters
$H_{nom}(s)$	$d, w_w = 237\text{nm}, t_1, h = 498\text{nm}, \rho = 22\text{n}\Omega\text{m}$
$H_c(s)$	$d = 237\text{nm} - 20\%, w_w = 237\text{nm} + 20\%$
$H_{rc}(s)$	$d = 237\text{nm} - 10\%, w_w = 237\text{nm} + 10\%, \rho = 22\text{n}\Omega\text{m} + 15\%$
$H_e(s)$	$t_1 = 498\text{nm} - 10\%$

Fig. 19. Parameters of transfer functions.

use of feedback and automatic-repeat-request (ARQ) schemes. However, this requires some kind of adaptivity of the encoding procedure and complex decoding techniques since the channel is not precisely known. Consequently, everything boils down to what price one wishes to pay. Universal codes offer low complexity but reduced achievable rate, while ARQ could give higher achievable rate with increased complexity. At the end, one might explore combinations of both techniques in addition to manufacturing technology improvement to find an optimal solution.

Remark 6.2: As pointed out in Section IV, (31) includes the term describing the crosstalk noise. In Fig. 17, SNR is defined as

$$10 \log_{10} \frac{P}{\frac{1}{2\pi} \int_0^{2\pi} \text{Trace}(W_{\mathbf{n}}(\theta)) d\theta} \quad (35)$$

where $W_{\mathbf{n}}(\theta)$ is primarily concerned with the thermal noise and other noises which arise in electronic circuits excluding the crosstalk noise. Thus, overall SNR will actually be smaller.

VII. ERROR CORRECTION CODES FOR CB

Consider the case of the bus consisting of 4 parallel interconnects which has been discussed earlier. The choice of error correcting codes is constrained by a number of factors, foremost of which is the achievable data rate that upper bounds the rate of the codes for achieving error free transmission over the bus. For example, Fig. 17 suggests that between 12 and 14 dB, the achievable data rate is around 0.5×10^{12} bit/s. Here, we are going to assume that this rate is equally divided to four scalar error correcting codes, each operating on one line of the

bus. Actually, the encoders and decoders operate on the equivalent channel shown in Fig. 6 which includes filters V and U^H . Therefore, a code rate for each code should not be larger than $0.5/4 = 0.125$. One possible choice for the codes would be a (63,7) Euclidian Geometry codes [30], whose code rate is $7/63 = 0.111 < 0.125$.

The choice of codes is also limited by complexity, delay and power requirements, which often require the codes to have conflicting characteristics. For instance, codes with longer lengths usually achieve better performance but at the cost of higher latency and decoder complexity. In the following discussion, we attempt to point out several examples of codes and decoders with varied implementation complexities, which could be good candidates for the CB.

For this high-speed application, it is generally suitable to use codes of very small length to reduce latency and decoder complexity. There exist numerous decoders that range from very-low complexity to high complexity, whose performances are usually proportional to their complexity. In general, decoders can be classified as low-complexity hard-decoders and high-complexity soft-decoder. Among the simplest decoders are one-step and multistep majority logic decoders (MLD) [30], [31], which usually require only a few logic gates to implement. These decoders are ideal candidates for extremely high-speed applications that can be operated at 10–40 Gb/s, though their performance pales in comparison to more sophisticated decoding algorithms. Many well-known algebraic codes like BCH, Reed–Muller codes and codes from finite geometries can be decoded using such decoders [30]. The hardware circuitry for these codes consists of a set of multi-input XOR and multi-input majority-logic gates. For example, a one-step MLD for a (15,7) BCH code requires only four four-input XOR gates, a four-input majority logic gate, and a 15-bit shift-register. Multi-step MLD performs better than one-step MLD but are marginally more complex. A two-step MLD for a (15,5) BCH code requires only 36 four-input XOR gates, seven six-input majority-logic gates, and a 15-bit shift register. There are many low-rate very short codes that can be used with this class of decoders.

Soft-decoders that make use of soft-values perform better than hard-decoders by many orders of magnitude, but require far more complex circuitry. Best performance for any code is achieved by the optimal maximum *a posteriori* (MAP) decoder [30]. Normally, it is impractical to implement this algorithm for most codes. But it is possible to implement it for certain very short codes. In [32], MAP decoder for a (8,4) Hamming code is implemented on a $0.5 \mu\text{m}$ CMOS technology with a decoder circuit size of 0.82 mm^2 . Bit-rates at 2–10 Mb/s are significantly lower than what is capable with hard-decoders. The most important class of soft-decoders is the class of iterative decoders, which are suboptimal, but of relatively lower complexity than MAP decoders. Some of the best performances reported so far are due to such decoders applied to a class of codes called low-density parity-check (LDPC) codes. In [33], an iterative decoder for a (32,8) LDPC code is implemented on a $0.18 \mu\text{m}$ CMOS technology requiring a core area of 0.57 mm^2 with a throughput of 6 Mb/s. For 65-nm CMOS technology, the

LDPC code implementation has 690 mW as power dissipation, 1 Gb/s as speed and 0.25 mm² as the area based on values from [34].

Finally, the overall complexity of the scheme also depends on the encoder/decoder architecture over the bus lines. For instance, the data through the bus lines may be transmitted by independently encoding and decoding data in each bus line. Alternately, a single stream of input data may be encoded and transmitted by equally distributing over all four bus lines. Other variations over this architecture are also possible. Again, the type of architecture to be used is determined by the specific system requirements.

VIII. CONCLUSION

New VLSI manufacturing processes inevitably introduce IC interconnect variability. This paper considers several different key issues such as: 1) modeling interconnect variability; 2) measuring the variability effect on the interconnect performance; and 3) proposing possible design solutions based on the Information Theory concepts. We suggest frequency-domain variability modeling by using H^∞ space norm. For this model, the achievable data transmission rate is computed, and encoding/decoding strategy is proposed that guarantees the reliable transmission of information along the bus interconnects regardless of their variability. We identified several problems such as intersymbol interference and encoder/decoder complexity that should be studied in the future research.

APPENDIX

To introduce the H^∞ space and $\|\cdot\|_\infty$ norm, we need the notions of transfer function stability.

A transfer function $H(z)$ is called stable if and only if all its poles lie within the unit circle in the z complex plain. If $H(z)$ is a matrix transfer function, $H(z)$ is stable if and only if all its entries are stable [26].

The H^∞ space is the set of all stable transfer function matrices.

The H^∞ norm of the system given by the matrix $H(e^{j\theta})$ ($\|H\|_\infty$) is computed as

$$\|H\|_\infty \triangleq \sup_{\theta \in [0, 2\pi)} \bar{\sigma}[H(e^{j\theta})] \quad (36)$$

where $\bar{\sigma}[H(e^{j\theta})]$ denotes the maximum singular value of the matrix $H(e^{j\theta})$. In the case of the scalar $H(e^{j\theta})$, (36) reduces to

$$\|H\|_\infty \triangleq \sup_{\theta \in [0, 2\pi)} |H(e^{j\theta})| \quad (37)$$

i.e., to the peak of the magnitude of the Bode plot of $H(e^{j\theta})$.

REFERENCES

- [1] ITRS, "The International Technology Roadmap for Semiconductors," 2005. [Online]. Available: <http://www.itrs.net>
- [2] I. Bahar, J. Harlow, D. Hammerstrom, W. Joyner, C. Lau, D. Marculescu, A. Orailoglu, and M. Pedram, "Executive summary: Architectures for silicon nanoelectronics and beyond," presented at the NSF Nanoarchitecture Workshop, Portland, OR, 2005.

- [3] V. Raghunathan, M. B. Srivastavat, and R. K. Gupta, "A survey of techniques for energy efficient on-chip communication," in *Proc. DAC*, Anaheim, CA, Jun. 2003, pp. 900–905.
- [4] Y. Cao, P. Gupta, A. B. Kahng, D. Sylvester, and J. Yang, "Design sensitivities to variability: Extrapolations and assessments in nanometer VLSI," in *Proc. ASIC/SOC Conf.*, Sep. 2002, pp. 411–415.
- [5] J. Cong, L. He, C. K. Koh, and Z. Pan, "Interconnect sizing and spacing with consideration of coupling capacitance," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 20, no. 9, pp. 1164–1169, Sep. 2001.
- [6] J. Cong, C. K. Koh, and P. H. Madden, "Interconnect layout optimization under higher order RLC model for MCM designs," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 20, no. 12, pp. 1455–1463, Dec. 2001.
- [7] J. Cong, T. Kong, and Z. Pan, "Buffer block planning for interconnect planning and prediction," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 9, no. 6, pp. 929–937, Dec. 2001.
- [8] C. J. Alpert, J. Hu, S. S. Sapatnekar, and P. G. Villarrubia, "A practical methodology for early buffer and wire resource allocation," in *Proc. DAC*, Las Vegas, NE, Jun. 2001, pp. 189–194.
- [9] C. J. Alpert, C. Chu, G. Gandham, M. Hrkic, J. H. C. Kashyap, and S. T. Quay, "Simultaneous driver sizing and buffer insertion using a delay penalty estimation technique," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 23, no. 1, pp. 136–141, Jan. 2004.
- [10] H. Su, J. Hu, S. S. Sapatnekar, and S. R. Nassif, "A methodology for the simultaneous design of supply and signal networks," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 23, no. 12, pp. 1614–1624, Dec. 2004.
- [11] P. Yu and D. Z. Pan, "TIP-OPC: A new topological invariant paradigm for pixel based optical proximity correction," in *Proc. IEEE ACM Int. Conf. Comput.-Aided Des. (ICCAD)*, San Jose, CA, Nov. 2007, pp. 847–853.
- [12] P. Yu and D. Z. Pan, "Novel intensity based OPC algorithm with speedup in lithography simulation," in *Proc. IEEE ACM Int. Conf. Comput.-Aided Des. (ICCAD)*, San Jose, CA, Nov. 2007, pp. 854–859.
- [13] L. Scheffer, "On-chip interconnect variation," 2005. [Online]. Available: http://sliponline.org/SLIP06/Presentations/2_1_lou.pdf
- [14] S. R. Sridhara and N. R. Shanbhag, "Coding for system on chip networks: A unified framework," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 6, pp. 655–667, Jun. 2005.
- [15] M. Favalli and C. Metra, "Bus crosstalk fault-detection capabilities of error-detecting codes for on-line testing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 7, no. 3, pp. 392–396, Sep. 1999.
- [16] R. Singhal, G. S. Choi, and R. Mahapatra, "Information theoretic capacity of long on-chip interconnects in the presence of crosstalk," in *Proc. 7th Int. Symp. Quality Electron. Des. (ISQED)*, San Jose, CA, 2006, pp. 407–412.
- [17] K. N. Patel and I. L. Markov, "Error-correction and crosstalk avoidance in dsm buses," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 12, no. 10, pp. 1076–1080, Oct. 2004.
- [18] D. Bertozzi, L. Benini, and B. Ricco, "Energy-efficient and reliable low-swing signaling for on-chip buses based on redundant coding," in *Proc. IEEE ISCAS*, 2002, vol. 1, pp. 93–96.
- [19] Y. M. Chee, C. J. Colbourn, and A. C. H. Ling, "Optimal memoryless encoding for low power off-chip data buses," in *Proc. IEEE ICCAD*, Nov. 2006, pp. 369–374.
- [20] R. R. Rao, H. S. Deogun, D. Blaauw, and D. Sylvester, "Bus encoding for total power reduction using a leakage-aware buffer configuration," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 12, pp. 1376–1383, Dec. 2005.
- [21] R. Hedge and N. R. Shanbhag, "Toward achieving energy efficiency in presence of deep submicron noise," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 8, no. 4, pp. 379–390, Aug. 2000.
- [22] S. R. Sridhara, A. Ahmed, and N. R. Shanbhag, "Area and energy-efficient crosstalk avoidance codes for on-chip buses," in *Proc. ICCD*, San Jose, CA, 2004, pp. 12–17.
- [23] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623–656, 1948.
- [24] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inf. Theory*, vol. 44, no. 10, pp. 2148–2177, Oct. 1998.

- [25] W. L. Root and P. P. Varaiya, "Capacity of classes of Gaussian channels," *SIAM J. Appl. Math.*, vol. 16, no. 6, pp. 1350–1393, Nov. 1968.
- [26] J. Doyle, B. Francis, and A. Tannenbaum, *Feedback Control Theory*. New York: Macmillan, 1992.
- [27] S. Z. Denic, "Communication subject to normed channel uncertainty," Ph.D. dissertation, Sch. Inf. Technol. Eng., Univ. Ottawa, Ottawa, ON, Canada, 2006.
- [28] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm.*, vol. 6, no. 10, pp. 585–596, Nov. 1999.
- [29] S. Denic, I. Djordjevic, J. Anguita, B. Vasic, and M. Neifeld, "Information theoretic limits for free-space optical channels with and without memory," *IEEE J. Light Wave Technol.*
- [30] S. Lin and D. J. Costello, *Error Control Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1983, pp. 273–290.
- [31] I. S. Reed, "A class of multiple-error correcting codes and the decoding of scheme," *IRE Trans.*, vol. 4, pp. 38–49, Sep. 1954.
- [32] C. Winstead, J. Dai, S. Yu, C. Myers, R. R. Harrison, and C. Schlegel, "CMOS analog MAP decoder for (8,4) Hamming code," *IEEE J. Solid-State Circuits*, vol. 39, no. 1, pp. 122–131, Jan. 2004.
- [33] S. Hemati, A. H. Banihashemi, and C. Plett, "A 0.18- μ m CMOS analog min-sum iterative decoder for a (32,8) low-density parity-check (LDPC) code," *IEEE J. Solid-State Cir.*, vol. 41, no. 11, pp. 2531–2540, Nov. 2006.
- [34] A. J. Blanksby and C. J. Howland, "A 690 mW 1 Gb/s 1024-b, rate-1/2 low-density parity-check codedecoder," *IEEE J. Solid-State Circuits*, vol. 37, no. 3, pp. 404–412, Mar. 2002.

Stojan Z. Denic (S'04–M'06) received the B.Sc. and M.Sc. degrees in electrical engineering from the Faculty of Electronic Engineering, University of Nis, Nis, Serbia, in 1994 and 2001, respectively, and the Ph.D. degree in electrical engineering from the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada, in 2006.

He currently holds a Senior Research Engineer post with the Telecommunications Research Laboratory, Toshiba Research Europe Limited, Bristol, U.K. From 1994 to 1996, he was a Research Assistant with the Telecommunications Lab, the Faculty of Electronic Engineering, University of Nis, Nis, Serbia. In 1996, he joined Telekom Srbija, Serbia, as a Network Design Engineer. From 2000 to 2006, he was a graduate student with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada. In the meantime, from 2005 to 2006, he was visiting the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus. In 2006, he joined the Department of Electrical and Computer Engineering, University of Arizona, Tucson, where he was a Research Assistant Professor until 2008. His research interests include information theory, communication over uncertain channels, control over communication channels, optical communications, and coding for wireless and digital recording channels.

Bane Vasic (S'92–M'93–SM'02) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Nis, Nis, Serbia, in 1989, 1991, and 1994, respectively.

Currently, he is a Professor with the Department of Electrical and Computer Engineering and Mathematics, University of Arizona, Tucson. Prior to this appointment, he was at Bell Laboratories, Allentown, PA. He authored a number of journal and conference articles and book chapters, and edited two books. His patents are implemented in Bell Lab chips. His research interests include coding theory, communication theory, constrained systems, and digital communications and recording.

Dr. Vasic is a Member of the Editorial Board of the IEEE TRANSACTIONS ON MAGNETICS, and was a Chair or Technical Program Chair for several workshops and conferences including 2003 and 2007 IEEE Communication Theory Workshop (CTW); 2004 DIMACS Workgroup and Workshop on Theoretical Advances in Information Recording; 2004 LANL Workshop on Applications of Statistical Physics to Coding Theory; and 2006 Communication Theory Symposium within ICC.

Charalambos D. Charalambous (S'90–M'91–SM'05) received the B.S., M.E., and Ph.D. degrees from Old Dominion University, Norfolk, VA, in 1987, 1988, and 1992, respectively, all in electrical engineering.

In 2003, he joined the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus, where he is currently an Associate Professor. In 2005, he was elected Associate Dean of the School of Engineering for a three-year term. He was an Associate Professor with University of

Ottawa, School of Information Technology and Engineering from 1999 to 2003, and an Adjunct Professor with McGill University from 1999 to 2002. He has served on the faculty of McGill University, Department of Electrical and Computer Engineering, as a visiting faculty, from 1995 to 1999. From 1993 to 1995, he was a Post-Doctoral Fellow with Idaho State University, Engineering Department. His current research interests are in theory and applications of stochastic processes and systems subject to uncertainty, communication and control systems, large deviations, information theory, and connections between robustness, information theory, and statistical mechanics.

Dr. Charalambous is currently an Associate Editor of the IEEE COMMUNICATIONS LETTERS, *Systems and Control Letters*, and from 2002 to 2004 he served as an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He was a member of the Canadian Centers of Excellence through MIT ACS. (the mathematics of information technology and complex systems), from 1998 to 2001. In 2001, he received the Premier's Research Excellence Award of the Ontario Province of Canada.

Jifeng Chen (S'08) received the B.S. and M.S. degrees in the area of telecommunications engineering from the Xidian University, China, in 2002 and 2006, respectively, and the M.S. degree in computer engineering from the University of Arizona, Tucson, in 2009.

His current research interests mainly include the high-speed interconnect, transceiver and receiver design and modeling for nanoscale circuits under sub-micron technologies.

Janet M. Wang (M'98–SM'08) received the B.S. degree from Nanjing University of Science and Technology (former East China Institute of Technology), Nanjing, China, in 1991, the M.E. degree from the Chinese Academy of Sciences, Beijing, China, in 1994, and the M.S. and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley, in 1997 and 2000, respectively.

She was with Intel, Santa Clara, CA, and Cadence, Santa Clara, CA, from 2000 to 2002. Since 2003, she has been a Faculty Member in the Electrical Computing Engineering Department, University of Arizona, Tucson.

Dr. Wang is a member of the IEEE Circuits and Systems Society. She was a recipient of the 2005 National Science Foundation Career Award and 2006 Presidential Early Career Award for Scientists and Engineers.