

Haar-based Interconnect Coding for Energy Effective Medium/Long Range Data Transport

N. Cucu Laurenciu, S.D. Cotofana
Computer Engineering Laboratory, Delft University of Technology, The Netherlands.
{N.CucuLaurenciu, S.D.Cotofana}@tudelft.nl

Abstract—In this paper we introduce and evaluate Haar based codec assisted medium and long range data transport structures, e.g., bus segments, Network on Chip interconnects, able to deal with technology scaling related phenomena (e.g., increased susceptibility to proximity coupling noise and transmission delay variability), targeting energy savings at the expense of a reasonably small overhead, i.e., 1 extra wire, a 2-gate encoder, and a 2-gate decoder, for each and every pair of uncoded wires. For practical evaluation we employed a 45nm commercial CMOS technology and different random, uncorrelated workload profiles. For 5mm and 10mm long 8-bit buses (without repeaters), we obtain energy savings of 55% and 34%, and a transmission frequency increase of 35% and 41%, respectively, at the expense of less than 1% area overhead with respect to the reference system (i.e., 8-wire synchronous uncoded bus), which prove energy and delay effectiveness. We further augment our proposal with a Single Error Correction and Double Error Detection (SECDED) scheme particularly adapted to its structure, in order to cope with very deep sub-micron noise (e.g., supply voltage variations, electromagnetic interference) induced transmission errors. When compared to the reference system (not SECDED protected), for 10mm long buses, our Haar tailored SECDED approach consumes 27% less energy at the expense of 2% area overhead.

Index Terms—Bus coding, Haar, dynamic power.

I. INTRODUCTION

As technology aggressively down-scales, wires are getting narrower and taller, the wire pitch smaller, the interconnect parasitics worsen, and the inter-wire capacitance increases (and thus the susceptibility to neighboring wires interference) negatively impacting the transmission latency and power consumption (the interconnects dominating figures [1]), as well as the transmitted signals integrity. In view of these considerations, in concert with logic optimization, a multi-criteria (e.g., energy, delay, physical area, reliability) design-time interconnect-centric avenue becomes a critical desideratum for high-performance and/or low power SoCs.

En route to address the interconnect issues, several solutions have been proposed, e.g., wire shielding and spacing, low-swing signaling, charge recycling, buffer/repeater insertion, and coding techniques [2]. The coding-based strand of research constitutes a technology and implementation independent compelling alternative, appealing from the power-delay-reliability multi-objective optimization standpoint. Several coding techniques have been previously investigated [3], [4], [5], [6], [2], most of them being focused on a single bus desideratum criterion (i.e., low power or low latency or reliability). Low power and low delay codes have as salient target switching activity reduction by taking advantage in the coding process of the temporal/spatial signature of the to be transmitted data. Most of these techniques are either tailored for data buses [7], [8], [9] or for address buses [10], [11], [12], [13], [14], [15], and as a result whilst effective for specific buses data profile correlations and peculiarities, they are less suitable for a general transmission context with random, uncorrelated data. To combat and decrease interconnects

susceptibility to the Very Deep Sub-Micron (VDSM) noise (e.g., supply voltage variations, electromagnetic interference) induced errors that can arise during the data transport, error control bus coding was employed [2]. However, as the error detection and correction codec is generally very computationally involved, its combination with low power or latency bus coding techniques has not been envisaged.

In light of the above perspective, in this paper we introduce and evaluate in a commercial 45nm technology codec assisted energy effective reliable data transport structures, e.g., bus segments, Network on Chip (NoC) interconnects, able to deal with technology scaling related phenomena, e.g., crosstalk and transmission delay variability, at the expense of a reasonably small area overhead. To this effect, we propose a low complexity 2 : 3 single stage Haar Transform based codec, which enables energy savings while also alleviating data transport time related aspects (i.e., we diminish the data transmission latency and obtain a lower variability data arrival profile, which is a key issue for interconnect robustness (reliability) in the context of VDSM fabrication technologies high process parameter variability). Moreover, we augment the Haar-based codec with a Single Error Correction Double Error Detection (SECDED) capability adapted to its peculiarities, such that we can combat errors arising during the data transport process. To assess the practical implications of our proposal, an 8-bit wide interconnect segment is equipped with the proposed codec infrastructure and evaluated for varying interconnect length and width. Our simulations indicate that when compared to the reference uncoded interconnect our proposal enables 30%, 55%, and 34% energy savings for an interconnect length of 1mm, 5mm, and 10mm, respectively. Moreover, given that the considered data encoding schemes diminish the crosstalk occurrence, codec augmented interconnects longer than 1mm can be operated at a higher frequency than the uncoded ones. In particular, a clock frequency increase of about 35% and 41% is enabled for an interconnect length of 5mm and 10mm, respectively. The energy and data transmission delay reductions are obtained at an area increase of less than 1% with respect to the reference uncoded design. The direct Error Correcting Code (ECC) augmentation of a 10mm long 8-wire reference system results in a $1.33\times$ energy increase, while for the same bus length the ECC enhanced Haar system requires 2% area overhead, consumes with 27% less energy and operates at a slightly higher frequency than the reference uncoded counterpart.

The remaining of the paper is organized as follows: The Haar coding scheme algorithmic aspects and codec architecture are discussed in Section II. Section III deals with the evaluation of the proposed codec practical implications. Section IV briefly reviews recent related work on bus coding and compares our proposal with existing state-of-the-art. Section V concludes the paper.

II. HAAR CODEC MODUS OPERANDI

For the nanometer technology, as the inter wire capacitances dominate the total bus capacitance, crosstalk between adjacent wires becomes a prominent concern. The increased capacitive coupling effects include glitches and/or increase of the transmission delay along the bus (when adjacent wires are switching in opposite directions, the transition on one wire might be slowed down), and in turn to the overall power consumption increase. One way to diminish these effects is to encode the transmitted data, such that coupling transitions ("1" \rightarrow "0" and viceversa) between adjacent wires are as scarce as possible. To this end, subsequently, we: (i) introduce a codec suitable for medium and longer range interconnects, whose implementation is presented Section II-A, that simultaneously targets energy, area, and delay merits, and (ii) augment the codec with error detection and correction circuitry particularly tailored to its structure, as described in Section II-B. At the crux of our codec, lies a 2 : 3 stage 1 Haar Transform [16] with a lightweight implementation, which takes advantage of bits compression benefits to reduce both the wires own transition count (less switching in time along each wire), and coupling transition count (less occurrences of adjacent wires switching concomitantly in opposite directions). For brevity, we assume a byte-wise synchronous data transmission but the discussion is general and can be easily extended for other interconnect widths.

A. Haar Codec

The Haar encoder receives as input per each clock cycle, a data byte subsequently denoted by $\{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, and generates as output a 12-bit wide encoded word to be sent over the bus wires.

To this end, the input data byte is divided into 4 2-bit groups as follows: $\{x_0, x_1\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_6, x_7\}$. For each such bit pair, the encoder performs 1-bit pair-wise addition and subtraction. Normally, the sum of two bits in 2's complement notation requires two bits for representation (as the exact sum value is required). However, because we compute both the sum and the difference of the same bits only the sum MSB (which corresponds to the carry out signal of a 1-bit full adder) needs to be evaluated. This is because the sum LSB is the parity bit, which is identical to the difference LSB (since the sum of 2 bits has the same parity as their difference). As concerns the difference of two input bits, it is likewise performed in 2's complement and requires a 2-bit representation. Specifically, for the input bits $\{x_0, x_1\}$, for instance, the encoder computes the following three bits:

$$S^{(0)} = x_0 \wedge x_1, C_1^{(0)} = x_1, C_0^{(0)} = x_0 \vee x_1,$$

where \wedge denotes a logical AND operation, and \vee denotes a logical OR operation.

The Haar decoder receives as input 12 bits of encoded data $\{S^{(0)}, C_1^{(0)}, C_0^{(0)}, S^{(1)}, C_1^{(1)}, C_0^{(1)}, \dots, S^{(3)}, C_1^{(3)}, C_0^{(3)}\}$ and outputs the data byte $\{\hat{x}_0, \hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4, \hat{x}_5, \hat{x}_6, \hat{x}_7\}$. Conceptually speaking, if we know the sum and difference of two numbers, it is straightforward to compute the two numbers in cause. Exemplifying for the bit pair $\{x_0, x_1\}$, the following equations govern:

$$\hat{x}_0 = S^{(0)} \oplus C_1^{(0)} \oplus C_0^{(0)}, \hat{x}_1 = C_1^{(0)},$$

where \oplus denotes a logical XOR operation, and for an error-

free transmission $\hat{x}_0 = x_0$ and $\hat{x}_1 = x_1$.

The following architectural related observations are in order:

– *Codec Implementation Complexity, Delay, and Energy.*

As concerns the hardware implementation, the Haar encoder and decoder exhibit very low complexity, consisting of one logic level for the encoder (one OR/AND gate) and two logic levels for the decoder (two XOR gates for the decoder, respectively). As a result, the Haar encoder/decoder delay is very small, i.e., the delay of a single logic gate for the encoder and two logic gates the decoder. The codec simplicity has also positive implications on its energy consumption.

– *Codec Scalability to Wider Interconnects.*

The low hardware complexity enables its utilisation for wider interconnects, as it scales linearly with respect to the number of wires (e.g., for 4 wires, the encoder requires 2 parallel OR gates and 2 parallel AND gates).

B. SECDED

For ECC data protection, two avenues can be followed, i.e., either protect the original bits to be transmitted over the bus, or the actual bits that are sent on the bus. Subsequently, we shall present the former approach algorithmic details, as it enables us to take advantage of the particular Haar codec structure for energy reduction.

Let m_0 denote the encoded sequence of transmitted bits $m_0 = (S^{(0)}, C_1^{(0)}, C_0^{(0)})$, and ϵ the transmission error pattern. Table I summarizes all possible 1-bit error scenarios affecting m_0 . The first two columns in the table represent the original data bits x_0 and x_1 , the third and the fourth column denote the Haar encoded message at the interconnect transmitting end, and the message at the receiving end (Haar encoded message + noise), respectively; while the last two columns correspond to the estimated data bits after Haar decoding. As for single bit errors, all possible values that ϵ can take are: (0 1 0), (1 0 0), and (0 0 1), 3 situations should be analyzed at the interconnect receiving end for each possible x_0 and x_1 bit combination. One

TABLE I
ONE-BIT ERROR SCENARIOS FOR HAAR SYSTEM.

x_0	x_1	(m_0)	$(m_0 \oplus \epsilon)$	\hat{x}_0	\hat{x}_1
0	0	(0 0 0)	(1 0 0)	1 (F)	1 (F)
			(0 1 0)		1 (F)
			(0 0 1)		0
0	1	(0 1 1)	(1 1 1)	1 (F)	0 (F)
			(0 0 1)		0 (F)
			(0 1 0)		1
1	0	(0 0 1)	(1 0 1)	0 (F)	1 (F)
			(0 1 1)		1 (F)
			(0 0 0)		0
1	1	(1 0 0)	(0 0 0)	0 (F)	0 (F)
			(1 1 0)		0 (F)
			(1 0 1)		1

may note in Table I, that in all one-error scenarios, the decoded value \hat{x}_0 is always erroneous. This is expected, as all three encoded bits $(S^{(0)}, C_1^{(0)}, C_0^{(0)})$ are involved in the computation of the decoded bit \hat{x}_0 . It follows that any single error affecting the encoded bit sequence $(S^{(0)}, C_1^{(0)}, C_0^{(0)})$, will always result in an erroneous \hat{x}_0 value. Thus it is mandatory to protect x_0 in order to be able to correct an erroneously decoded value \hat{x}_0 . On the other hand, when \hat{x}_0 is erroneous, we also need to discriminate the correct value of \hat{x}_1 , in which case $C_0^{(0)}$ has to be protected. Generalizing from the two input bits $\{x_0, x_1\}$ to the entire input byte $\{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, we

TABLE II
HAAR SYSTEM SINGLE ERROR CORRECTION.

Case 1	Case 2
$\hat{x}_0 = S^{(0)} \oplus C_1^{(0)} \oplus C_0^{(0)}$	$\hat{x}_0 = E_1 \oplus E_2 \oplus E_4$
$\hat{x}_2 = S^{(1)} \oplus C_1^{(1)} \oplus C_0^{(1)}$	$\hat{x}_2 = E_1 \oplus E_2 \oplus E_3$
$\hat{x}_4 = S^{(2)} \oplus C_1^{(2)} \oplus C_0^{(2)}$	$\hat{x}_4 = E_2 \oplus E_3 \oplus E_4$
$\hat{x}_6 = S^{(3)} \oplus C_1^{(3)} \oplus C_0^{(3)}$	$\hat{x}_6 = E_1 \oplus E_3 \oplus E_4$

propose to append the following 5 error control coding bits to the 12-bit Haar encoded sequence:

$$\begin{aligned} E_1 &= x_0 \oplus x_2 \oplus x_6 & E_3 &= x_2 \oplus x_4 \oplus x_6 \\ E_2 &= x_0 \oplus x_2 \oplus x_4 & E_4 &= x_0 \oplus x_4 \oplus x_6 \\ E_5 &= C_1^{(0)} \oplus C_1^{(1)} \oplus C_1^{(2)} \oplus C_1^{(3)} \end{aligned}$$

The $\{E_1, E_2, E_3, E_4\}$ bits correspond to a (7, 4) Hamming code and are used for the correction of the input bits $\{x_0, x_2, x_4, x_6\}$, while bit E_5 is simply a parity bit used for the correction of the input bits $\{x_1, x_3, x_5, x_7\}$.

Single Error Detection and Correction

Any single bit-flip error affecting any bit of the 17-bit sequence $\{S^{(0)}, C_1^{(0)}, C_0^{(0)}, \dots, S^{(3)}, C_1^{(3)}, C_0^{(3)}, E_1, \dots, E_5\}$, can be corrected as follows: We compute in parallel each of the bits $\{\hat{x}_0, \hat{x}_2, \hat{x}_4, \hat{x}_6\}$ in two manners, as summarized in Table II. The bits $\{w_1, w_2, w_3, w_4\}$ computed as $w_1 = \hat{x}_0$ (case 1) $\oplus \hat{x}_0$ (case 2), $w_2 = \hat{x}_2$ (case 1) $\oplus \hat{x}_2$ (case 2), $w_3 = \hat{x}_4$ (case 1) $\oplus \hat{x}_4$ (case 2), and $w_4 = \hat{x}_6$ (case 1) $\oplus \hat{x}_6$ (case 2) are utilized to discriminate the correct set of values between the case 1 and case 2 estimates, as presented hereafter.

- $\hat{x}_0, \hat{x}_2, \hat{x}_4, \hat{x}_6$ bits correction

In the error free scenario, \hat{x}_0 (case 1) coincides with the value of \hat{x}_0 (case 2) and thus $\{w_1, w_2, w_3, w_4\} = \{0, 0, 0, 0\}$.

If one error occurs in the sequence $\{S^{(0)}, C_1^{(0)}, C_0^{(0)}, \dots, S^{(3)}, C_1^{(3)}, C_0^{(3)}\}$, then one value of \hat{x} (case 1) is computed wrong, and one of the $\{w_1, w_2, w_3, w_4\}$ bits is equal to "1". In this situation, the case 1 decoded bits $\{\hat{x}_0, \hat{x}_2, \hat{x}_4, \hat{x}_6\}$ are the correct ones.

If one error occurs in the sequence $\{E_1, E_2, E_3, E_4\}$, then three values of \hat{x} (case 2) are computed wrong, and three of the bits $\{w_1, w_2, w_3, w_4\}$ are equal to "1". In this situation, the case 2 decode bits $\hat{x}_0, \hat{x}_2, \hat{x}_4, \hat{x}_6$ are the correct ones.

Thus, to summarize the discrimination bits are used as:

- If $w_1 + w_2 + w_3 + w_4 = 3$ then choose case 2 $\{\hat{x}_0, \hat{x}_2, \hat{x}_4, \hat{x}_6\}$ decoded bits;
- Otherwise, choose case 1 $\{\hat{x}_0, \hat{x}_2, \hat{x}_4, \hat{x}_6\}$ decoded bits.

- $\hat{x}_1, \hat{x}_3, \hat{x}_5, \hat{x}_7$ bits correction

Suppose one error occurred and the correct value of \hat{x}_0 was obtained using the above methodology. However, there are two Haar encoded 3-bit sequences for the correct value of $\hat{x}_0 = 0$, per se. Specifically, $(S^{(0)}, C_1^{(0)}, C_0^{(0)})$ is then either (0, 0, 0) or (0, 1, 1), which means we do not know exactly whether the value of \hat{x}_1 is "0" or "1". However, since bit E_5 is correct and equal to $C_1^{(0)} \oplus C_1^{(1)} \oplus C_1^{(2)} \oplus C_1^{(3)}$, it allows for the determination of the correct value of \hat{x}_1 . In this case $\hat{x}_1 = E_5 \oplus C_1^{(1)} \oplus C_1^{(2)} \oplus C_1^{(3)}$. Note that the situation when one error affects E_5 is of no relevance, as all the 12 S and C bits required to restore the correct input byte $\{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ are correct.

Double Error Detection

A first case is when any two bits from the sequence $\{S^{(0)}, C_1^{(0)}, C_0^{(0)}, \dots, S^{(3)}, C_1^{(3)}, C_0^{(3)}, E_1, \dots, E_4\}$, are affected, which results in a value of $w_1 + w_2 + w_3 + w_4$ which is either equal to 2, or equal to 4.

A second case is when one bit-flip error is in the previous 16-bits sequence, while the other error affects bit E_5 . In such a case, the error in the sequence can be detected with the previous single detection flow (otherwise stated, $w_1 + w_2 + w_3 + w_4$ has to be either equal to 1 to equal to 3), while bit E_5 can be duplicated and transmitted twice over the wires, at the extra cost of an additional redundant wire.

We note that even if full SECDEC protection requires up to 18 wires, if area is a foremost design optimization goal, a plausible alternative is to make use of both clock edges and transmit the 18 bits over a 9-wire bus (9 bits on the rising edge, and the other 9 bits on the falling edge).

III. SIMULATION RESULTS

To gain insight on the practical implications of the proposed Haar codec enhanced data transport approach we evaluate by means of SPICE simulations the following systems:

- "Ref" - An 8-wire reference system, for which uncoded, raw data are transmitted over the wires. The 8-wire system serves as comparison reference from timing, energy, and area standpoints, as no coding scheme for optimizing the energy and reliability characteristics of the data transmission is applied in this case.
- "H12" (Haar) - A 12-wire system, which makes use of the scheme proposed in Section II.
- "H8" - A 12-wire system which transmits encoded data by means of the scheme proposed in Section II, the design being done with smaller distance between wires such that the bus width is preserved, i.e., $W(H8) = W(Ref)$.
- "BI" (Bus Invert) - A 9-wire system implementing the coding scheme in [3], a simple and efficient precursor of many coding methods and typical comparison reference.
- "Ref + Hamm" - A 12-wire system which corresponds to the "Ref" system protected with a single error detection and correction (15, 11) Hamming code (8 bits information, 12 bits code length).
- "H12 + Hamm" - An 8-wire system which corresponds to the "H12" system protected with a single error detection and correction (7, 4) Hamming code (4 bits information, 7 bits code length). The 3 Hamming bits protect the even bits $\{x_0, x_2, x_4, x_6\}$. One more parity bit protects the odd bits $\{x_1, x_3, x_5, x_7\}$. As in total $12 + 3 + 1 = 16$ bits are required to be sent over the bus, an 8-wire bus can be employed (the data being sent over the bus on both clock edges).
- "H12 + ECC" - A 9-wire system afferent to the "H12" system protected using the SECDED scheme proposed in Section II. In this case a total of $12 + 5 = 17$ bits are required, which can be sent over a 9-wire bus with a double data rate (on both edges of the clock).

For each system, the SPICE simulation setup consists of encoders & input buffers, interconnect, and output buffers & decoders. Figure 1 depicts the simulation setup for the proposed interconnect codec augmented systems. The setup for the 8-wire, reference system is similar, with the exception of the encoder and decoder blocks which are excluded. As concerns the interconnect, for given specifications (e.g.,

wire length, number of parallel conductors), and technology parameters (e.g., related to the dielectric and metal layer stack conductivity, dielectric permittivity, wire pitch, aspect ratio etc.), a SPICE RLGC compatible model was obtained using the Synopsys Raphael electromagnetic field solver. The simulations were performed in SPICE, employing a commercial 45nm technology, at nominal operating conditions, for different bus lengths (1mm to 10mm to cover medium and long range interconnects) and bus widths (8 to 512, constructed from multiple 8-wire bus subsystems). As data to be transmitted over the wires, 10000 randomly generated bytes are provided as system input, one byte per clock cycle.

A. Energy & Area

The consumed energy is measured for the entire system (encoder/input buffers + interconnect + decoder/output buffers), and over the entire transmission duration, i.e., $10000 \times T_{clk}$. To provide a fair comparison, T_{clk} is tailored for each analyzed system (as a function of the wire length, and of the encoder/decoder maximum operation frequency), such that the data at each system output can be correctly sampled. The energy is measured in SPICE using the supply current integrated over the entire transmission duration, thus we capture both the static and the dynamic energy components.

1) Energy Oriented Haar Systems

Figure 2 graphically illustrates the energy reduction achieved by the 8-bit "H12", "H8", and "BI" schemes for different interconnect lengths. The energy percentages are reported relative to the energy figures obtained for the Ref 8-wire system. A similar trend is observed for all systems: As the interconnect length increases, the energy gain also increases, which is as expected, since longer interconnects are more energy demanding than shorter ones, and thus can benefit more from a switching activity reduction on its wires. For the "H12" and "H8" systems, an energy gain apex of 56% and 30% is manifested at 6 – 8mm and 6 – 7mm, respectively. The lowest energy reduction potential is observed for shorter interconnects (30% and 23%, respectively for 1mm wires), where the codec energy is more significant, and thus the Haar scheme becomes less effective. When moving to longer, e.g., 10mm interconnects lower energy saving of 34% and 25% are achieved for the "H12" and "H8" systems, respectively, which can be attributed to the longer wires and driving buffers consumed energy counterpoising the codec enabled energy benefits. The "BI" design is clearly less effective than the Haar based designs and even results in energy increase for shorter than 5 mm wires. As concerns the area footprint, a less than 1% hardware overhead corresponds to the Haar systems, which is not unexpected, as the codec requires very simple logic. Since the chip area footprint is determined by the logic and local interconnects, and not by the global interconnects, which are implemented in the upper metal layers, and their afferent vias, the 8 to 12 bus width increase does not add any additional overhead to the area footprint. Thus the "H12" system has \approx the same area as the 8-wire "Ref", while from the energy point of view is more effective with 27% on average than the "Ref" system. To assess the potential sensitivity of the codec augmented interconnect performance w.r.t. the bus width, we evaluated the energy reduction - as depicted in Figure 3 - for the transmission of a 512-bit wide input data vector over 8, 16, 32, 64, 128, 256, and 512-bit wide 5mm long buses (the Haar system bus being composed of multiple shielded 12-wire

wide bus segments). We observe an energy gain inflexion point is obtained for the 64-wire bus (55%), w.r.t. which 6% and 2% energy increase is obtained for the 512-wire bus and 8-wire bus, respectively. We attribute this to: (i) the switching activity profile changes (e.g., as the bus size increases from S to $2S$ one potential bus switching is eliminated, thus leading to smaller energy dissipation), and (ii) the transmission time reduction (e.g., as the bus size increases from S to $2S$ the transmission time is reduced by half). On the other hand, as the bus size increases the parasitics also become more complex, and interconnect and area requirements are doubled, with negative implications on the timing and energy figures.

2) Energy and Reliability Oriented Haar Systems

Figure 4 depicts the percentage energy gain for 12-wire "Ref + Hamm", 8-wire "H12 + Hamm", and 9-wire "H12 + ECC" systems, relative to the energy figures of the 8-wire "Ref" system, for 1mm to 10mm bus lengths. We observe that the direct augmentation of the reference with SECDED capabilities (the "Ref + Hamm" system) comes with a great energy consumption increase (more than $2 \times$ "Ref") with an aggravation trend as the interconnect length increases, which can be explained by a worse parasitics profile for the 12-wire bus and an activity profile higher in coupling transitions. Conversely the two Haar-based systems become more energy effective with the interconnect length increase and even consume less energy than the "Ref" unprotected baseline for longer than 7mm interconnects. For interconnect length < 7 mm, the two Haar schemes enhanced with error protection consume more energy than the "Ref" system. Specifically, on average the 8-wire "H12 + Hamm" and 9-wire "H12 + ECC" consumes with 15% and 28% more energy than the "Ref" system, respectively. However, for interconnect length above 7mm both "H12 + Hamm" and "H12 + ECC" are more energy effective than the "Ref" system consuming 11% and 18% less energy on average, respectively. When compared to the 8-wire "H12 + Hamm" system, the 9-wire "H12 + ECC" systems enables an activity profile with fewer transitions on the extra error detection and correction lines, which is reflected in higher overall energy gain when the interconnect energy dissipation is the dominant contributor (≥ 7 mm). Note that the energy dissipated by the two ECC augmented Haar-based systems can be further diminished, if the power supply voltage is reduced (as the afferent timing faults single errors can be corrected by the SECDED logic). Area wise, the ECC augmented systems require 1%, $\approx 2\%$, and $\approx 2\%$ area overhead with respect to the "Ref" area, for "Ref+Hamm", "H12 + Hamm", and "H12 + ECC", respectively.

B. Delay

1) Energy Oriented Haar Systems

Simulation results reveal that for the "Ref" 8-wire system, the bit arrival time for each wire exhibits a smaller spread when compared to the "H12" system. However, the "Ref" maximum arrival time is larger than the one provided by the Haar system counterpart, except for smaller length wires (1mm). This has positive implications on the transmission clock period, as indicated by Figure 5, which can be decreased with 35%, and 41% for 5mm, and 10mm, respectively. For 1mm, the clock period is negatively impacted, as it is increased with 31% vs. "Ref", which can be attributed to the effects of bus switching activity diminution that are more prominent for medium and longer wires than for shorter wires (the total delay

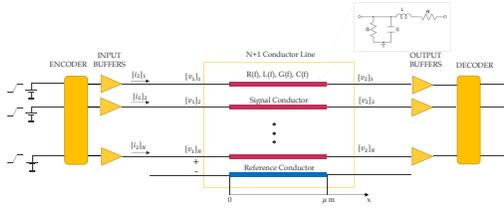


Fig. 1. SPICE Simulation Setup.

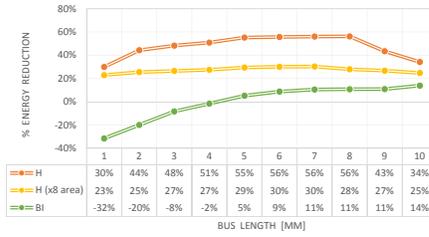


Fig. 2. Energy profile vs. interconnect length.

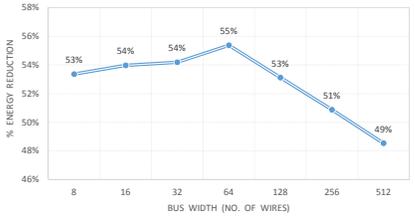


Fig. 3. Energy profile for "H12" system for bus length of 5mm vs. bus width.

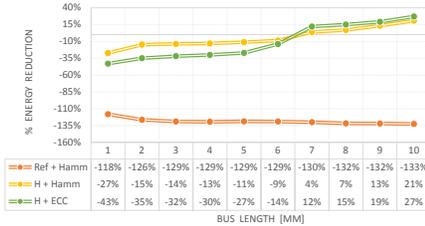


Fig. 4. Energy profile vs. interconnect length for the ECC protected systems.

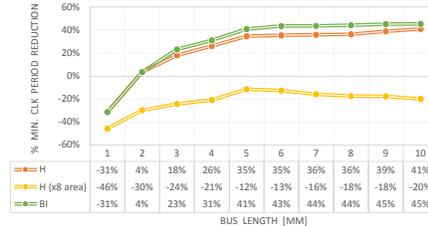


Fig. 5. Minimum clock period vs. interconnect length.

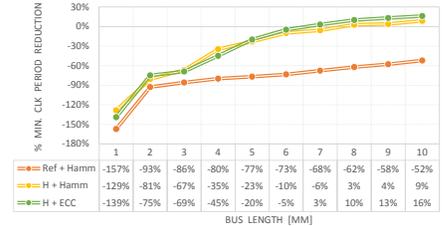


Fig. 6. Minimum clock period vs. interconnect length for the ECC protected systems.

- encoder+decoder+bus - of the coding-based system having to counterbalance the 8-wire bus delay, in order to obtain clock period benefits). The "H8" system is negatively impacted for all considered lengths, requiring a clock period increase vs. "Ref" of 46%, 12%, and 20% for 1mm, 5mm, and 10mm, respectively. We attribute this to the bus design parameters (e.g., decreased spacing), since the "H8" bus width is smaller than the one of the H12 bus, occupying the same metal layer area as the "Ref" 8-wire bus. We note that the delay figures are afferent to the "H12" system without repeaters, reflecting the maximum wire length for which the signal integrity is preserved and which enables the maximum energy savings. If we buffer the "H12" system, the delay can be further improved, but at the expense of consuming extra energy, and increasing the area footprint. Additionally, the system without repeaters benefits the stringent time-to-market constraints, as it enables a faster timing-closure progress.

2) *Energy and Reliability Oriented Haar Systems*
 Figure 6 depicts the percentage reduction of the minimum clock period for the ECC protected systems w.r.t. the "Ref" clock period. The "Ref + Hamm" system clock period is decreased for all wire lengths, while for the "H12 + Hamm" and "H12 + ECC" systems, the frequency is decreased only until 7mm and 6mm, respectively. At 10mm the "H12 + Hamm" system can properly operate at a clock period 9% smaller than "Ref" clock period, while the "H12 + ECC" system enables a clock period decrease of 16% w.r.t. "Ref". We note that even if for the two Haar systems, the propagation delay across the wires is reduced, the added delay of the ECC/Hamming encoder and decoder makes the overall delay reduction gain possible only for longer wires. We note that the reported delay figures correspond to the minimum bus delay for which the signal integrity is preserved. However, the delay can be further reduced under the safe operation value, since an ECC scheme is in place and can correct potential errors.

IV. STATE-OF-THE-ART COMPARISON

Subsequently, we give a brief account of the most recent prevalent articles documenting bus coding, and a comparison against

the state-of-the-art performance figures for the "H12" and the "H12 + ECC" systems. We note that a direct comparison with state-of-the-art is not always straightforward, e.g., when the implementations are done in different technologies, to which effect we apply the Dennard scaling [17], in the case of an analytical evaluation, or one which doesn't account for the codec performance penalties.

• Power Reduction Codes

Generally speaking the power-reduction oriented bus coding research corpus can be broadly divided into two main categories: methods which reduce the per wire self-switching activity and methods which reduce both self and coupling (between neighboring wires) transitions. The former category disregards the inter-wire coupling parasitics, which for VDSM technologies results in delay and power penalties. The latter category is in better alignment with state of the art interconnect physical phenomena and as a result enables power savings. However, delay and reliability aspects are not considered. With a few exceptions (e.g., Bus Invert coding [3], a simple and efficient precursor of several methods), the preponderant existing low power bus coding techniques are only effective either for data buses (e.g., [7], [8], [9]) or for address buses (e.g., irredundant permutation-based codes - Gray code [10], data dependent reordering codes [11]; redundant codes - Odd/Even Bus Invert [12] which extends [3] for coupling activity reduction, T0 code [13], Beach code [14], Limited Weight code [15])). Furthermore, most of these techniques mainly exploit the spatial and/or temporal correlations of the transferred data, which render them less effective for random data transmission. Following the same philosophy, most recent work includes Conditionally Coded Blocks (CCB) code, Sign Extension (SEM) code, XOR/XNOR code, and Qaudro code [2]. State-of-affairs CCB and SEM [2] yield 58% and 60%, respectively power reduction in 130nm for a 16-bit transmission. The Quadro coding [2] achieves up to 47% reduction for byte-wise transmission, the technology node not being specified. Comparatively, we obtain in 45nm 56% energy reduction for

6mm to 8mm long 8-bit buses, and 54% energy reduction for 5mm 16-bit buses (2 shielded 8-bit bus segments).

- *Delay Reduction Codes*

Crosstalk Avoidance Codes (CACs) were proposed to reduce crosstalk induced delay by forbidding certain transitions (e.g., opposite direction switching) on adjacent wires [4] or bit patterns (e.g., '010' and '101') [5], which would cause the highest delay (affluent to the worst adjacent capacitive coupling cases). However, as integral part of an IC, interconnects are exposed to various environmental aggression factors (e.g., supply voltage fluctuation, electromagnetic interference), which pose in the absence of an error resiliency mechanism, signal integrity and reliability problems. [5] reports 21% power savings in a 90nm process with an encoder area of 369 2-input gates for 12-bit transmission, while [4] estimate a delay of 300ps in 65nm for a pipelined implementation, at the expense of 17% less area overhead than [5]. In our case the operating frequency can be increased by 35% for 5mm, with a lightweight hardware implementation (a codec gate count of 16 and total logic depth of 3 vs. a state-of-art gate count in the order of hundreds), thus surpassing state-of-the-art.

- *Reliability Improvement Codes*

To combat errors that may occur during bus transmission and thereby diminish the necessity and impact on performance of timing and voltage margins over-designing traditional approach, error control codes were explored, most notably Hamming and cyclic linear block codes [6]. However, they usually incur significant performance penalties caused by the codecs high complexity, which is not the case for the proposed SECEDED Haar based systems.

- *Joint Codes*

Combining in a serial manner low power codes with crosstalk avoidance codes proved to be inefficient, as the crosstalk avoidance properties per se, will be altered/canceled by a subsequent low power coding. However, crosstalk avoidance codes benefit also, as an aside the power consumption (even if to a lesser extent when compared to the savings achieved by low power codes). Conversely, error control codes can be combined with either low power codes or crosstalk avoidance codes, towards joint power/delay-reliability merits, with the chief caveat of significant coding overhead, as the joint code is merely a concatenation of two individual, independent codes. In this regard, very few joint coding techniques have been investigated, e.g., [18] and [19], which attain single and double, respectively error correcting capacity, by combining a Hamming code with crosstalk avoidance codes; single error correcting [20] and [21] which introduce the Duplicate-Add-Parity code, Modified-Dual-Rail code and Boundary-Shit code, respectively. State-of-affair ECC based methods [20] report $\approx 40\%$ energy reduction when compared to standard Hamming code for a byte-wise transmission, using 10mm wires in 130nm, while our savings w.r.t. standard Hamming are $1.2\times$ bigger in 45nm. [21] achieves the ECC capability for 9 extra wires and 45 FO4 codec gate delay, while [18] reports a codec delay of 27 FO4 (unpipelined) and an energy dissipation 15.1% and 18.3% lower over [20] and [21], in 130nm. To conclude, our overall simulation results indicate significant energy savings, while increasing the operating frequency, and having also SECEDED capabilities, thus outperforming state-of-the-art counterparts.

V. CONCLUSIONS

In this paper we proposed an energy effective bus coding scheme, that facilitates a higher operating frequency when

compared to the uncoded counterpart, as well as with state-of-affairs. We further augmented the Haar codec with a tailored SECEDED scheme with energy, area, and frequency merits outperforming a direct ECC augmentation of the uncoded system. We analyzed two systems: the Haar-based system which targets energy efficiency, and the SECEDED augmented Haar-based system which targets a reliable and energy effective data transport. Simulation results in 45nm of the Haar system for an 8-wire interconnect and various workload profiles, indicate energy savings of 55% and 34% and operation frequency increase of 35% and 41% for 5mm and 10mm, respectively, at the expense of an area increase of less than 1%, with respect to the reference uncoded system. The SECEDED enhanced Haar system consumes for the 10mm case $1.2\times$ less energy than the reference uncoded system, and requires 2% area overhead when compared to the "Ref" area.

REFERENCES

- [1] Borkar, S., "Role of Interconnects in the Future of Computing." in *Journal of Lightwave Technology*, 2013, pp. 3927–3933.
- [2] Saini, S., *Low Power Interconnect Design*. Springer, 2015.
- [3] Stan, M.R. and Burleson, W.P., "Bus-Invert Coding for Low-Power I/O." in *IEEE Transactions on VLSI Systems.*, 1995, pp. 49–58.
- [4] Duan, C. and Chengyu, Z. and Khatri, S.P., "Forbidden Transition Free Crosstalk Avoidance CODEC Design." in *Design Automation Conference.*, 2008, pp. 986–991.
- [5] Duan, C. and Calle, V.H.C. and Khatri, S.P., "Efficient On-Chip Crosstalk Avoidance CODEC Design." in *IEEE Transactions on Very Large Scale Integration Systems.*, 2009, pp. 551–560.
- [6] Bertozzi, L. Benini, and G. De Micheli, "Error Control Schemes for On-Chip Communication Links: The Energy Reliability Tradeoff." in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems.*, 2005, pp. 818–831.
- [7] Sathish, A. and Latha, M.M. and Kishore, K.L., "A Technique to Reduce Transition Energy for Data-Bus in DSM Technology." in *IJCSI International Journal of Computer Science.*, 2011, pp. 402–406.
- [8] Natesan, J. and Radhakrishnan, D., "Shift Invert coding (SINV) for Low Power VLSI." in *IEEE Conference on Digital System Design.*, 2004, pp. 190–194.
- [9] Sathish, J. and Rao, T.S., "Bus Regrouping Method to Optimize Power in DSM Technology." in *IEEE-Int. Conference on Signal processing, Communications and Networking.*, 2008, pp. 432–436.
- [10] Su, C.L. and Tsui, C.Y. and Despaigne, A. M., "Saving Power in the Control Path of Embedded Processors." in *IEEE Design and Test of Computers.*, 1994, pp. 24–30.
- [11] Murgai, R. and Fujita, M., "On Reducing Transition Through Data Modifications." in *DATE.*, 1999, pp. 82–88.
- [12] Lin, R.B. and Fujita, M., "Inter-Wire Coupling Reduction Analysis of Bus-Invert Coding." in *IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 55.*, 2008, pp. 1911–1920.
- [13] Benini, L. and De Micheli, G. and et. al., "Asymptotic Zero-Transition Activity Encoding for Address Buses in Low-Power Microprocessor-based Systems." in *Great Lakes VLSI Symposium.*, 1997, pp. 77–82.
- [14] Benini, L. and De Micheli, G. and et.al., "System-level Power Optimization of Special Purpose Applications: The Beach Solution." in *Int. Symp. on Low Power Electronics and Design.*, 1997, pp. 42–49.
- [15] Stan, M.R. and Burleson, W.P., "Coding a Terminated Bus for Low Power." in *Great Lakes Symp. on VLSI.*, 1995, pp. 70–73.
- [16] Benedetto, J.J. and Frezier, M.W., *Wavelets: Mathematics and Applications*. CRC Press, 1994.
- [17] Dennard, R., "Design of Ion-Implanted MOSFETs with Very Small Physical Dimensions." in *IEEE Jrn. of Solid State Circuits*, 1974, pp. 256–268.
- [18] Ganguly, A. and Pande, P.P. and et.al., "Addressing Signal Integrity in Networks on Chip Interconnects through Crosstalk-Aware Double Error Correction Coding." in *IEEE Computer Society Annual Symposium on VLSI.*, 2007, pp. 317–324.
- [19] R. Srinivasa and N. Shanbhag, "Coding for Reliable On-Chip Buses: Fundamental Limits and Practical Codes." in *International Conference on VLSI Design.*, 2005, pp. 417–422.
- [20] Rossi, D. and Cavallotti, S. and Metra, C., "New ECC for Crosstalk Impact Minimization." in *IEEE Int. Symp. on Defect and Fault Tolerance in VLSI Systems.*, 2003, pp. 257–264.
- [21] Patel, K.N. and Markov, I.L., "Error-Correction and Crosstalk Avoidance in DSM Busses." in *IEEE Tr. on VLSI Systems.*, 2004, pp. 1076–1080.