

ENERGY MACRO-MODEL FOR ON-CHIP INTERCONNECTION BUSES

R. Mendoza, M. Pons, F. Moll & J. Figueras
Technical Report DOCT-07-0001
June 29, 2006

1. GOAL

Given an on-chip bus driver-interconnection-receiver design of N parallel lines (Fig. 1), the objective is to develop its energy consumption macro-model. With this model we will be able to evaluate the energy metrics for the bus under a certain traffic and information coding.

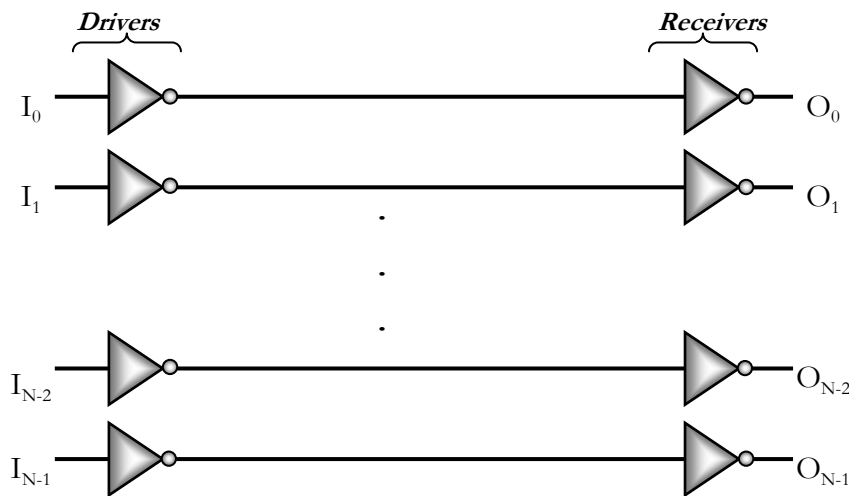


Figure 1. Bus to be modelled.

2. INTRODUCTION

In Nanometric technologies, with every new technology generation, device dimension, threshold voltage and supply voltage are scaled to make possible higher complexity and improved performance. This trend has resulted in a significant increase of on-chip communications needs implying the growth of the number of global buses, drivers, repeaters and receivers required to communicate different distant modules in a chip. Moreover the lateral capacitance component of interconnections is continually growing due to the reduction of wire pitch and the increase of aspect ratio [ITRS05].

Thus, several encoding bus techniques had been developed to control these effects. There are three families of them [Sridhara05]: Low Power (LPC), Crosstalk avoidance (CAC) and Error Control (ECC). LPC minimises the self-switching activity in each line and switching activity between lines, CAC reduces the delay and ECC focuses on error detecting and correcting.

In our work, we will evaluate the simple ECC parity code, which adds one bit to every code word to make the number of bits with value 1 odd (odd-parity code) or even (even-parity code). With this coding scheme, if the counted number of bits with value 1

in the receiver is not the expected (odd or even) a single error can be detected. We will also evaluate the most common ECC Hamming code, which uses the concept of overlapping parity maintaining a Hamming distance greater than 2 between all the code words. The Hamming code following the rule $2^c \geq d + c + 1$, where c is the number of redundant check bits and d is the number of data bits, corrects single errors or it can be used to detect double errors. [TUCS05]

A number of strategies utilizing bus encoding have been proposed to reduce the energy consumption in buses. This energy is composed by three factors: energy for capacitive switching and energy for short-circuit current and leakage energy during the complete duration of the bus transmission (clock period P).

Several approaches seek to minimize the capacitive consumption in buses [Gray53] [Stan95] [Benini97] [Benini97x] [Victor01] [Ghoneima04] [Lind05] [Ravindra05] [Lyuh06].

In [Gray53] [Stan95] [Benini97] [Benini97x] the ‘Gray’, ‘Bus Inverter’ (BI), ‘Zero Transition’ (T0) and ‘Beach’ codes were proposed respectively. [Victor01] presents a theoretical analysis of ‘self-shielding’ or ‘crosstalk-immune’ codes. [Ghoneima04] proposes a low energy bus scheme which introduces ‘Dynamic Delay’ between oppositely switching adjacent lines (DDL). [Lind05] proposes modifications to BI coding scheme called ‘Deep-Sub-Micron Bus Inverter’ coding (DSMBI) and ‘Minimal Redundancy’ code (MiniRed). [Ravindra05] and [Lyuh06] present a bus encoding scheme that takes into account coupling and self transitions of bus lines. However all of these code schemes neglect the impact of leakage energy.

More recently, some works include the leakage energy in global buses [Rao04] [Deogun04] [Rao05] [Deogun05].

In [Rao04] the authors compare traditional bus schemes with different designs based on Multi-Threshold CMOS (MTCMOS) technique. [Deogun04] and [Rao05] propose a self-shield encoding algorithm combined with Staggered Threshold Voltage (STV, dual- v_{th} design) for crosstalk aware and low energy. In [Deogun05] the authors present a Dynamically Pulsed MTCMOS (DPM) bus system to attend the problems of runtime and standby leakage.

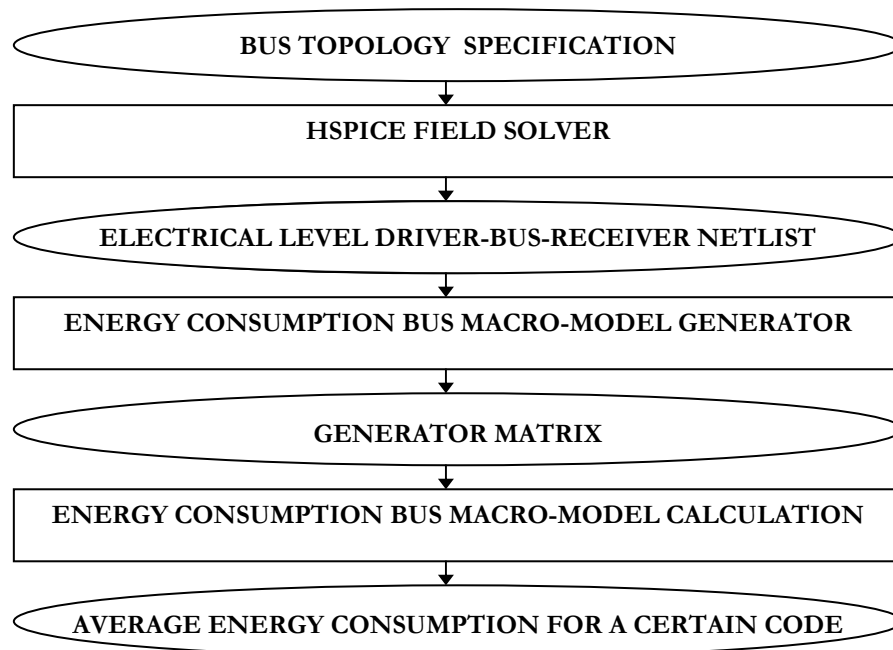


Figure 2. Flow diagram for the extraction of the bus energy consumption.

In our study, what we will present is a comprehensive bus energy consumption macro-model considering all of the three components of energy. The estimation of the energy consumption of a bus for nanometric technologies will be developed in a hierarchical way. First of all, the bus topology specification is needed. Then, using the HSPICE tool we will translate this specification to the electrical level obtaining the bus model netlist. With this model we will extract a matrix containing the energy consumption per line depending on the transition concerned. This matrix is the macro-model generator because it gives all the information to calculate the bus energy consumption.

3. BUS TOPOLOGY SPECIFICATION

Using the HSPICE W Element we will be able to specify the target bus model through the use of Built-in Field Solver tool [HSPICE05]. It provides the ‘RLC’ values per unit length from the material and geometrical two dimensional definition of the bus. With these values we will calculate the capacitances between lines and the substrate for the specified bus length. While other authors like [Sortiriadis02] only consider the coupling effects to one adjacent line, our macro-model takes into account the coupling effects to the two adjacent lines of the line under study.

In the description of the lines of the bus through W-element we will be able to specify the following aspects showed in Figure 3 depending on the technology node chosen:

- Number of lines of the bus N
- Line cross-section:
 - Width of the lines W
 - Thickness of the lines T
 - Aspect ratio W/T
- Length of the lines L
- Line separation S
- Type of conductor of the lines:
 - Aluminium Al

- Copper Cu
- Chip material configuration (layer stack):
 - Silicon Oxide SiO_2
 - Silicon Si (N or P)
- Height H from the lines to the P-Si or N-Si substrate

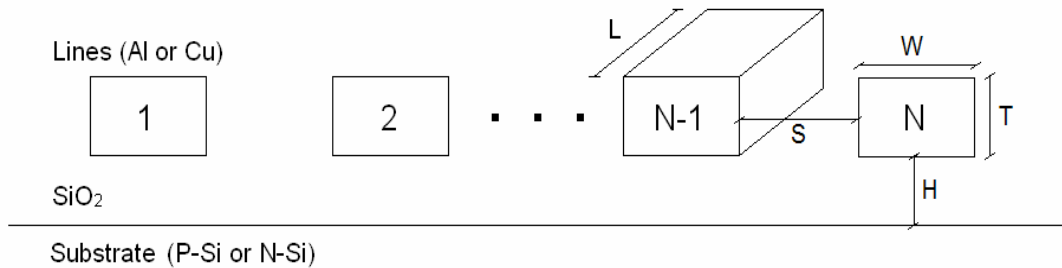


Figure 3. Geometrical parameters of the bus lines.

On the other hand, there are several possible routing configurations for the bus. In order to unify the study, we will assume some other characteristics:

- Straight parallel bus lines
- A part of the total line length L is over GND substrate and the rest is over V_{DD} substrate (Figure 3). This is defined by the parameter $\eta \in [0;1]$:
 - $\eta \cdot L$ is over GND
 - $(1 - \eta) \cdot L$ is over VDD
- Only the critical parasitic capacitances are considered
 - Capacitance to ground per unit length C_{GND}'
 - Capacitance to V_{DD} per unit length C_{VDD}'
 - Coupling capacitances to one (C_{i+1}') and two lines (C_{i+2}') away (coupling capacitances of lines at distances greater than two are assumed negligible)

The capacitances will be calculated multiplying in each case the length of the line concerned by the capacitance per unit length.

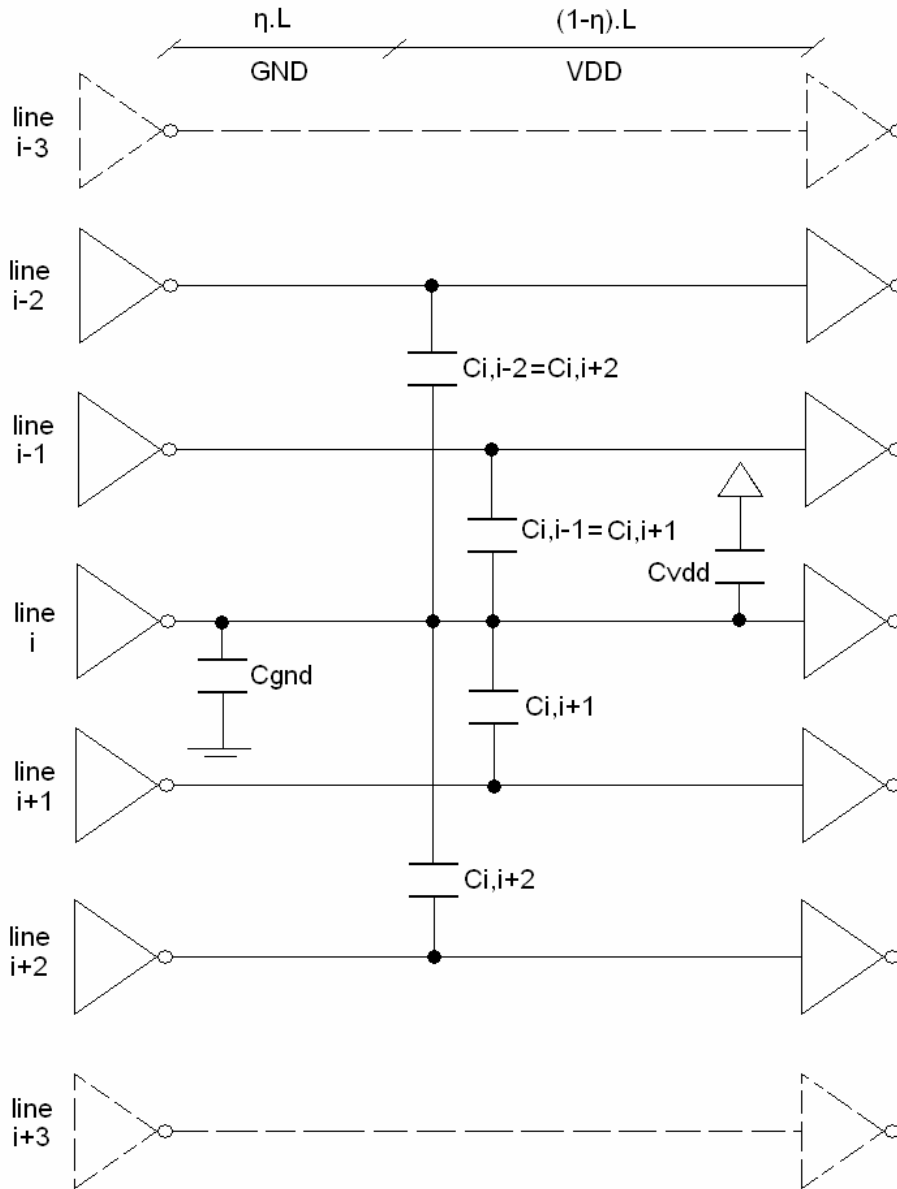


Figure 4. Capacitance model

4. ELECTRICAL LEVEL DRIVER-BUS-RECEIVER NETLIST

In addition to the bus capacitances, to define the electrical level HSPICE netlist we have to include the driver and receiver descriptions. In fact, the energy consumption of the bus also depends on the chosen configuration of these drivers and receivers.

In the design of drivers and receivers we can chose the next parameters:

- Technology node (nanometric)
- Width of the PMOS and NMOS transistors W_p and W_n
- Length of the PMOS and NMOS transistors L_p and L_n

As we did before for the bus topology, we have to make some assumptions:

- Drivers and receivers are identical and assumed inverters
- The length of the transistors is the minimum length according to the technology node

- Drivers' width is chosen to guarantee that the line's delay is smaller than $P/5$ where P is the clock period

5. ENERGY CONSUMPTION BUS MACRO-MODEL GENERATOR

The energy consumption per line generator matrix (Table 1) of our N lines bus energy macro-model will be constructed from the simulation results obtained with the HSPICE driver-bus-receiver netlist for a bus of only five lines. This is possible because we only consider the coupling effects of two adjacent lines. Thus, the energy consumption in the line i , the middle line of the five simulated, is a function of the bit transitions in lines $i-2$, $i-1$, i , $i+1$, $i+2$, as showed the Equation 1:

$$E_i = f(i-2, i-1, i, i+1, i+2) \quad (1)$$

In the case of the corner lines that have no adjacent lines in one direction, or only have one of them, we will be able to use the same generator matrix. This is due to the fact that we will accept that the corner line energy consumption is equivalent to the middle line energy consumption obtained when we have adjacent lines with the same transition than the middle line. We will call these lines 'Phantom Lines' (PL) and the way they work is shown in Figure 5. We are assuming that the energy consumption is not modified in these conditions because no capacitance has to be charged or discharged between the middle line and the PL.

Therefore we will only have to obtain simulations to build a matrix containing the energy consumption of the middle line of a five line bus for each possible transition from input to output.

Table 1. Generator Matrix containing the energy consumption for the middle line of five lines.

		Final state of the 5 lines				
		00000	00001	...	11110	11111
Initial state of the 5 lines	00000	$E_i(00000>00000)$	$E_i(00000>00001)$		$E_i(00000>11110)$	$E_i(00000>11111)$
	00001	$E_i(00001>00000)$	$E_i(00001>00001)$		$E_i(00001>11110)$	$E_i(00001>11111)$
	...					
	11110	$E_i(11110>00000)$	$E_i(11110>00001)$		$E_i(11110>11110)$	$E_i(11110>11111)$
	11111	$E_i(11111>00000)$	$E_i(11111>00001)$		$E_i(11111>11110)$	$E_i(11111>11111)$

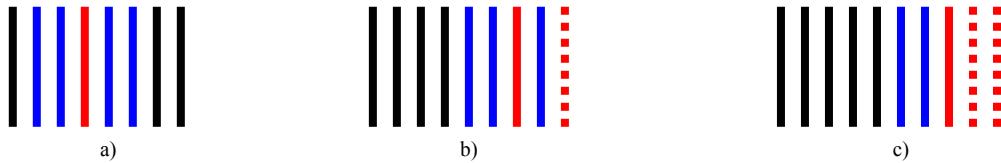


Figure 5. Phantom lines: (a) no phantom lines, (b) one phantom line and (c) two phantom lines. Red lines are middle lines under study. Blue lines are adjacent lines. Dotted red lines are phantom lines added with the same transition than the middle line.

6. ENERGY CONSUMPTION BUS MACRO-MODEL CALCULATION

Using the matrix generator shown in Table 1, we will be able to calculate the total consumption E_t of the N lines bus by adding the consumption for each line E_i depending on the coding transition:

$$E_t = \text{sum}(E_i) \quad (2)$$

Now, we can obtain the average energy consumption for a coding scheme by adding each of the energy consumptions from one word to another and dividing the total sum by the number of possible transitions:

$$E_{avg} = \text{sum}(E_i) / \# \text{transitions} \quad (3)$$

With this information, we can compare the different coding schemes in the bus under study. For making the calculations, we can use a program that implements the following pseudo-code:

Inputs:

GeneratorMatrix [2⁵][2⁵];

CodingWords;

NumWords, N;

Outputs:

TotalEnergy;

AverageCodingEnergy;

Main:

i,j,k;

TotalEnergy=0;

Word previousword[N], nextword[N];

For(i=0; i<NumWords; i++)

{

 previousword= Read word i from CodingWords;

 For (j=0; j<NumWords; j++)

 {

 nextword= Read word j from CodingWords;

 for(k=0; k<N; k++)

 {

 TotalEnergy= +LineEnergy(k, previousword, nextword);

 }

 }

}

AverageCodingEnergy=TotalEnergy/(NumWords²);

LineEnergy(k, previousword, nextword):

m,n;

If(k<2)or(k>N-3)

{

 If (k==0)

 {

 m=previousword[k]x2^4+ previousword[k]x2^3 + previousword[k]x2^2 +
 + previousword[k+1]x2 + previousword[k+2];

 n=nextword[k]x2^4+ nextword[k]x2^3 + nextword[k]x2^2 +
 + nextword[k+1]x2 + nextword[k+2];

 }

 If (k==1)

 {

 m=previousword[k]x2^4+ previousword[k-1]x2^3
 + previousword[k]x2^2 + previousword[k+1]x2 + previousword[k+2];

 n=nextword[k]x2^4+ nextword[k-1]x2^3 + nextword[k]x2^2 +
 + nextword[k+1]x2 + nextword[k+2];

 }

 If (k==N-1)

 {

 m=previousword[k-2]x2^4+previousword[k-1]x2^3
 + previousword[k]x2^2 +previousword[k]x2 + previousword[k];

 n=nextword[k-2]x2^4+ nextword[k-1]x2^3 + nextword[k]x2^2 +
 + nextword[k]x2 + nextword[k];

 }

 If (k==N-2)

 {

 m=previousword[k-2]x2^4+ previousword[k-1]x2^3 +
 previousword[k]x2^2 + previousword[k+1]x2 + previousword[k];

 n=nextword[k-2]x2^4+ nextword[k-1]x2^3 + nextword[k]x2^2 +
 + nextword[k+1]x2 + nextword[k];

 }

}

Else

{

 m=previousword[k-2]x2^4+ previousword[k-1]x2^3 + previousword[k]x2^2 +
 + previousword[k+1]x2 + previousword[k+2];

 n=nextword[k-2]x2^4+ nextword[k-1]x2^3 + nextword[k]x2^2 +
 + nextword[k+1]x2 + nextword[k+2];

}

Return GeneratorMatrix[m][n];

7. BUS TOPOLOGY EXAMPLE

To evaluate the proposed macro-model, we have chosen the following parameters to specify the driver-bus-receiver topology:

- $N=32$
- $W=0.16\mu\text{m}$
- $T=0.1\mu\text{m}$
- $W/T=1.6$
- $L=100\mu\text{m}$
- $S=0.14\mu\text{m}$
- Type of conductor of the lines = Cu
- Chip material configuration (layer stack):

```

Half Space, AIR
----- Z = 4.100000e-07m
  siO2  H = 3.000000e-07m
----- Z = 1.100000e-07m
  si    H = 1.000000e-07m
----- Z = 1.000000e-08m
////// Bottom Ground Plane //////////////////////////////////
----- Z = 0m

```

- $H=0.21\ \mu\text{m}$
- $\eta=0.5$
- Technology node=90nm
- $W_p=270\text{nm}$ and $W_n=90\text{nm}$
- $L_p=90\text{nm}$ and $L_n=90\text{nm}$
- $P=1\text{ns}$

8. MACRO-MODEL

Part of the generator matrix obtained for the example defined in the previous section is shown in Table 2. In fact, this matrix has 32×32 entries. Some values are negative because capacitors charged for the previous state give their charge to make the transition.

Table 3 presents for some transitions the error between the energy consumption calculation and simulation for a 32 lines bus. We can see that for this cases, this error varies from almost 0% when no transition is done to 5.5% in an arbitrary case. To be able to evaluate in more detail the mean error introduced by the macro-model it will be necessary to simulate and check all the possible transitions. That is to say $(2^{32}) \times (2^{32})$ cases, and this is precisely what we want to avoid by the means of the macro-model.

Table 2. Generator Matrix obtained containing the energy consumption (in Joules) for the middle line of five lines.

		Final state of the 5 lines				
		00000	00001	...	11110	11111
Initial state of the 5 lines	00000	2.3406E-18	1.4828E-16		3.6503E-15	3.6443E-15
	00001	-1.4460E-16	2.3406E-18		3.6590E-15	3.6501E-15
	...					
	11110	5.1670E-15	5.3275E-15		4.7739E-18	5.3732E-18
	11111	5.0195E-15	5.1668E-15		3.9703E-18	4.7739E-18

Table 3. Comparison of the energy consumption obtained using the proposed macro-model and the HSPICE simulator for five transitions on a 32 lines bus.

	Initial state (hex)	Final state (hex)	Macro-model Energy Consumption (J)	Simulation Energy Consumption (J)	Error (%)
Case 1	00.00.00.00	FF.FF.FF.FF	1.1662E-13	1.17200E-13	0.497
Case 2	FF.FF.FF.FF	FF.FF.FF.FF	1.5276E-16	1.52770E-16	0.003
Case 3	00.00.00.00	AA.AA.AA.AA	1.9414E-13	1.96170E-13	1.035
Case 4	00.00.00.00	FF.00.AA.FF	1.1732E-13	1.20170E-13	2.373
Case 5	4D.A8.F1.4B	56.29.AA.F3	1.4582E-13	1.54190E-13	5.429

9. CODING EVALUATION

Implementing the pseudo-code presented in section 6 we have evaluated the average energy consumption for a transmission of 8 bits (256 code words, from 0 to 255), for one error detecting coding scheme (128 code words, 7 bits of data plus 1 bit for parity) and for one error correcting coding scheme (16 code words, 4 bits of data plus 4 bits for Hamming redundancy). The results are shown in Table 4. One can expect that for detecting or correcting errors the average consumption increases, but we can see that the difference between the three schemes is minimal. The price that is paid for having fault tolerance is the reduction of the number of code words that can be transmitted on a given line of N bits. Therefore, the efficiency for parity and Hamming codes is smaller than for the uncoded transmission.

Table 4. Average energy consumption for different coding schemes on a 8 lines bus.

	Average Energy Consumption (Joules)	Number of Code Words
Uncoded (8 bits)	3.3412e-014	256
Even-Parity (7+1 bits)	3.3421e-014	128
Hamming (4+4 bits)	3.0389e-014	16

10. CONCLUSIONS

We have presented a comprehensive energy macro-model for on-chip interconnection N lines buses. It takes into account the second level of coupling capacitances between lines and also transistor leakage currents to calculate the consumption. It allows the designer to quickly evaluate his coding schemes without having to simulate all the possible N bits transitions but only the transitions for 5 bits using the concept of Phantom Lines, thus reducing the design time. The generator matrix is the only data needed by the program that we have developed to evaluate the average energy consumption for different fault tolerant coding schemes. The error introduced by this model is around 5% in an arbitrary transition. Future works will try to understand in more detail and to reduce the sources of error (PL, HSPICE tool inaccuracies and number of coupling levels considered).

11. REFERENCES

- [ITRS05] International Roadmap for Semiconductor 2005. <http://public.itrs.net/>
- [Sridhara05] S. Sridhara and Naresh Shanbhag, "Coding for system-on-chip networks: A unified framework". IEEE Transactions on VLSI, vol. 13, no. 6, pp.665-667, June 2005.
- [TUCS05] Lehtonen, Teijo and Plosila, Juha and Isoaho, Jouni, "On Fault Tolerance Techniques towards Nanoscale Circuits and Systems", TUCS Technical Report Number 708 August 2005, pp.1 – 32.
- [Gray53] F. Gray, "Pulse code communication", U. S. Patent 2 632 058, March 17 1953.

- [Stan95] M. R. Stan, and W. P. Burleson, W.P.: “*Bus-invert coding for low-power I/O*”, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Volume 3, Issue 1, March 1995 pp.49 – 58.
- [Benini97] L. Benini, G. De Micheli, E. Macii, D. Sciuto, and C. Silvano, “*Asymptotic zero-transition activity encoding for address buses in low-power microprocessor-based systems*”, Proceedings. Seventh Great Lakes Symposium on VLSI, 1997. 13-15 March 1997, pp.77 – 82.
- [Benini97x] L. Benini, G. De Micheli, E. Macii, M. Poncino, and S. Quer, “*System-level power optimization of special purpose applications the beach solution*”, Proceedings., 1997 International Symposium on Low Power Electronics and Design, 1997. 18-20 Aug 1997, pp.24 – 29.
- [Victor01] B. Victor and K. Keutzer, “*Bus encoding to prevent crosstalk delay*”, IEEE/ACM International Conference on Computer Aided Design, 2001. ICCAD 2001, pp. 57-63
- [Ghoneima04] M. Ghoneima and Y. L. Ismail, “*Utilizing the effect of relative delay on energy dissipation in low-power on-chip buses*”, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Volume 12, Issue 12, Dec 2004 pp.1348 – 1359.
- [Lind05] T. Lindkvist and J. Löfvenberg, “*Minimal Redundancy, Low Power Bus Coding*”, IEEE NORCHIP 2005, Oulu, Finland, 2005
- [Ravindra05] J. V. R. Ravindra, K. S. Sainarayanan and M.B. Srinivas, “*An efficient power reduction technique for low power data I/O for military applications*”, The 24th Digital Avionics Systems Conference, 2005. DASC 2005. Volume 2, 30 Oct.-3 Nov. 2005 pp 7.E.3-1 to 7.E.3-8 .
- [Lyu06] C.-G. Lyuh and T. Kim, “*Low-power bus encoding with crosstalk delay elimination*”, IEE Proceedings-Computers and Digital Techniques, Volume: 153, Issue: 2 pp.93- 100.
- [Rao04] R. Rao; K. Agarwal, D. Sylvester, R. Brown, K. Nowka and S.Nassif, “*Approaches to Run-Time and Standby Mode Leakage Reduction in Global Buses*”, Proceedings of the 2004 International Symposium on Low Power Electronics and Design, 2004. ISLPED '04. pp. 188 – 193.
- [Deogun04] H.S. Deogun, R. Rao, D. Sylvester and D. Blaauw, “*Leakage-and crosstalk-aware bus encoding for total power reduction*”, Proceedings 41st Design Automation Conference, 2004, pp. 779 – 782.
- [Rao05] R. Rao, H.S. Deogun, D. Blaauw and D. Sylvester, “*Leakage Bus encoding for total power reduction using a leakage-aware buffer configuration*”, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Volume 13, Issue 12, Dec. 2005 pp.1376 – 1383.
- [Deogun05] H.S. Deogun, R. Rao, D. Sylvester, R. Brown and K. Nowka, “*Dynamically pulsed MTCMOS with bus encoding for total power and crosstalk minimization*”, Sixth International Symposium on Quality of Electronic Design, 2005. ISQED 2005. 21-23 March 2005, pp.88 – 93.
- [HSPICE05] [HSPICE® Signal Integrity User Guide](#), Version X-2005.09, September 2005. Synopsys®.
- [Sotiriadis2002] P.P.Sotiriadis, and A.P.Chandrakasan, “*A bus energy model for deep submicron technology*”, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Volume 10, Issue 3, June 2002 pp. 341 – 35.