Analysis of Bus-Invert coding in the presence of correlations

Alberto García-Ortiz, Daniel Gregorek, Christof Osewold
Institute of Electrodynamics and Microelectronics (ITEM), University of Bremen, Germany
{agarcia,gregorek,osewold}@item.uni-bremen.de

Abstract
The selection of the right low-power coding technique during the design of the interconnect architecture has a high potential. However, it requires the analysis and evaluation to be performed at high-levels of abstraction. Closed formulas to quantify the power reduction achieved by each low-power code are therefore highly desirable. Using a simplified Dual Bit Type model for the underlying signals, the current work provides closed formulas for calculating the transition activity of correlated DSP signals encoded with the Bus-Invert technique. It considers the effect of the temporal and spatial correlation present in typical DSP signals. The accuracy of the proposed formulas is validated with extensive simulations at the bit level.

Original published in IEEE

García-Ortiz, A.; Gregorek, D.; Osewold, C.; Analysis of Bus-Invert coding in the presence of correlations; In Electronics, Communications and Photonics Conference (SIECPC) , 2011 Saudi International; Digital Object Identifier: 10.1109/SIECPC.2011.5876944; PDF
Analysis of Bus-Invert Coding in the Presence of Correlations

Alberto García-Ortiz, Daniel Gregorek, Christof Osewold
Institute of Electrodynamics and Microelectronics (ITEM)
University of Bremen, Germany
{agarcia,gregorek,osewold}@item.uni-bremen.de

Abstract—The selection of the right low-power coding technique during the design of the interconnect architecture has a high potential. However, it requires the analysis and evaluation to be performed at high-levels of abstraction. Closed formulas to quantify the power reduction achieved by each low-power code are therefore highly desirable. Using a simplified Dual Bit Type model for the underlying signals, the current work provides closed formulas for calculating the transition activity of correlated DSP signals encoded with the Bus-Invert technique. It considers the effect of the temporal and spatial correlation present in typical DSP signals. The accuracy of the proposed formulas is validated with extensive simulations at the bit level.

I. INTRODUCTION

Low-power coding for on-chip and off-chip interconnects is emerging [1], [2], [3], [4] as an important technique for overcoming the increasing power budgets required for the communication architecture of modern chips. As technology shrinks, the impact of the interconnects in the overall system power and system performance is increasing. Moreover, the random variability of the buffers and wires that define the interconnects cannot be neglected anymore. Thus, it is necessary to consider simultaneously power, timing, and reliability metrics while optimizing the coding strategy [2]. An important consequence of the aforementioned scenario is the need for high-level estimation techniques which allow a fast analysis of different coding alternatives.

An effective way of reducing the power consumption in global interconnects is the Bus-Invert (BI) method, first proposed by Fletcher [5] and afterwards re-discovered by Stan et.al. [3]. Initially, the technique was designed to address the switching activity only, however, several extensions as Odd/Even Bus-Invert have been proposed for reducing the coupling activity as well (see [6], [4] and references there). Further on, a wide range of approaches have been developed to improve the performance of the basic Bus-Invert scheme. Several approaches referred as Partial Bus-Invert advocate for coding a part of the bus only (see [7], [4] and reference there), while other techniques combine Bus-Invert with other schemata as non-redundant memoryless coding [4] or even more complex approaches as Berger codes [1]. Without a suitable mathematical analysis of the different coding techniques (and their alternatives) it is not possible to exploit them early in the design flow.

In recent years, several works analyzing the characteristics of Bus-Invert have been reported in the literature [8], [6], [7]. Using concepts from the Viterbi algorithm, Rokhani et.al. prove that the Bus-Invert has an optimal low-power strategy [8]. Moreover, in [6], [7] Lin et.al. provide closed formulas for the transition activity (switching bits) and coupling activity of Bus-Invert. However, these two works are restricted to the simple case of uncorrelated signals. Typical DSP signals contain temporal and spatial correlations, and cannot be analyzed with the currently existing approaches. This work provides a framework to overcome that limitation.

The rest of the work is organized as follows. Sec. II describes the assumptions on the employed signal model. It also describes the proposed closed formulas for estimating the transition activity of BI encoded signals. Afterwards, Sec. III analyzes the accuracy of the proposed equations first for idealized correlated DSP signals and then for some examples of Gaussian signals. Finally, the conclusions and further work are described in IV. In order to make the exposition more fluent, the mathematical demonstrations are summarized in the appendix.

II. ANALYSIS

A. Signal model

Correlated DSP signals have a very well defined structure for the transition activity [9] called Dual Bit Type (DBT). On the one hand, the LSBs have a transition activity of 0.5 and they tend to be uncorrelated. On the other hand, the MSBs have a strong correlation and a fixed activity depending on the (word-level) temporal correlation of the signal.

As a simple theoretical model for a B-bits correlated DSP signal, we consider that it is composed by two ideal regions (see Fig. 1). The uncorrelated one, composed by L uncorrelated bits, and the correlated one composed by M = B − L equal bits. Each of those correlated bits have a probability of 0.5, and a transition activity given by the parameter t_{MSB}.

It can be noticed that the previous model has three main parameters, B, L, and t_{MSB}. Using these parameters, the total transition activity of the un-encoded signal, T_t, is given by:

\[ T_t = \frac{L}{2} + (B - L)t_{MSB} \] (1)
Fig. 1. Simplified DBT signal model consisting of $L$ uncorrelated LSBs and $B-L$ identical MSBs transition activity equal to $t_{\text{MSB}}$.

B. Statistical analysis of BI

Let us consider a stationary $B$-bits random signal $X[n]$, which is afterwards encoded using the classical Bus-Invert technique [3]. The resulting signal has two parts: a redundant bit, called $r[n]$, and the remaining non-redundant $B$ bits called $Y[n]$ in the sequel. The basic operation of BI is to send either $X[n]$ or $X[n]$ in order to reduce the switching activity. If the hamming distance between $Y[n-1]$ and $X[n]$ is smaller than $\frac{B}{2}$, the coder sets $r[n] = 0, Y[n] = X[n]$. In the opposite case, it sends $r[n] = 1, Y[n] = \overline{X[n]}$. An example is shown in Tab. I.

### Example of the operation of Bus-Invert

<table>
<thead>
<tr>
<th>Signal $X[n]$</th>
<th>$Y[n]$</th>
<th>$r[n]$</th>
<th>Hamming distance $X[n], X[n-1]$</th>
<th>$Y[n], Y[n-1]$</th>
<th>$r[n], r[n-1]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$11011$</td>
<td>$11011$</td>
<td>$0$</td>
<td>$4$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$00101$</td>
<td>$10100$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$0$</td>
</tr>
<tr>
<td>$11001$</td>
<td>$11001$</td>
<td>$0$</td>
<td>$2$</td>
<td>$1$</td>
<td>$0$</td>
</tr>
<tr>
<td>$11010$</td>
<td>$11010$</td>
<td>$0$</td>
<td>$2$</td>
<td>$1$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

Let us define the random variable $\theta_U$ as the number of bits that toggle in the LSBs. Clearly, it has a binomial distribution with number of trials equal to $L$ and probability equal to $1/2$.

$$P_R(\theta_U = h_u) = \binom{L}{h_u} 2^{-L}$$

In order to simplify the notation, we use the cumulative distribution function (CDF) of $\theta_U$ in the sequel. It is given by:

$$\text{bincdf}_{L}(x) = \sum_{h_u=0}^{x} \binom{L}{h_u} 2^{-L}$$

Moreover, we define the random variable $\theta_X$ that describes the hamming distance between two consecutive samples of $X[n]$ (i.e., the number of bits that toggle). Its probability density function (PDF) is given by:

$$P_R(\theta_X = h_x) = (1 - t_{\text{MSB}}) \ P_R(\theta_U = h_x) + t_{\text{MSB}} \ P_R(\theta_U = h_x - M)$$

The hamming distance between $X[n]$ and $X[n-1]$, and the hamming distance between $Y[n]$ and $Y[n-1]$, are related by the equation $h_Y = \min(h_X, B - h_X)$. Moreover, there is a transition in $r[n]$ whenever $h_x > \frac{B}{2}$. Thus, the total transition activity of $r[n]$ and $Y[n]$ are given by:

$$T_r = \sum_{h_x=0}^{L} P_R(\theta_X = h_x) \min(h_X, B - h_X)$$

$$T_y = \frac{L}{2} \beta + (B-L) \alpha_2$$

with the parameters:

$$\alpha_1 = 1 - \text{bincdf}_L(\frac{B}{2})$$

$$\alpha_2 = 1 - \text{bincdf}_L(\frac{B - 1}{2})$$

$$\beta = 1 - \binom{L-1}{\frac{B-1}{2}} 2^{-L+1}$$

These formulas provide compact and close expressions for the transition activity of correlated signals under the given signal model.

### III. Experimental results

This section discusses the experimental validation of the proposed equations. First, we analyze the accuracy of the approach with signals that can precisely be characterized by the simple DBT model described in Sec. II. The goal is to ensure that the equations are exact indeed. Afterwards, the general case of Gaussian signals is discussed.

In order to validate Eq. (7) and Eq. (8), appropriate random signals must be generated. A $B$-bits signal consisting of $L$ uncorrelated bits and $M$ correlated bits can be constructed using $L+1$ independent binary random signals (see Fig. 1).

Each of the first $L$ bits is generated as a temporally uncorrelated uniform random bit with $p = 1/2$. Additionally an uncorrelated uniform random variable is constructed with $p = t_{\text{MSB}}$. An XOR-decorrelator is used to transform that signal into another one with $p = 1/2$ and switching activity equal to $t_{\text{MSB}}$. This signal defines the $M$ MSBs of $X[n]$.

In these experiments the bus-width $B$ is set to 12 to allow a wide range of possible values of $L$. Moreover, the length of the data is set to $2^{14}$ samples to allow enough accuracy even in the presence of strong correlations. We use MATLAB to generate the sequences of samples.

In the first experiment we fix the $t_{\text{MSB}}$ and change $L$ from 1 to $B-1$. Using bit-level simulations we record the transition activity of the un-encoded signal $X[n]$, and of the Bus-Invert encoded signal $\{Y[n], r[n]\}$. We compare those experimental results with the closed formulas provided by Eq. (7) and Eq. (8). We observe a perfect match between theory and experiments. As an example, Fig. 2 depicts the
results for a highly correlated signal with $t_{MSB} = 0.1$ and a poorly correlated signal with $t_{MSB} = 0.4$. The figures also show the transition activity using [7] which does not consider the correlation of the signal. Clearly, that approach leads to large errors. For example, for $B = 12$, $L = 6$, and $t_{MSB} = 0.4$ the measured activity equals approx. 3.4, while the un-correlated approach from [7] would estimate approx. 5. It represents an overestimation of 43%. On the other hand, our Eq. (7) and Eq. (8) provide the right results.

In the next experiment, $L$ is fixed and $t_{MSB}$ varies. Again we observe a perfect match between theory and experiments. As an example, the results for $L = 6$ and $L = 10$ are depicted in Fig. 3. We can observe the different slopes of the curves regarding the $t_{MSB}$, as previously discussed.

Although the detailed analysis of the coding characteristics of Gaussian signals is outside the scope of the current work, we have experimentally analyzed the suitability of our formulation in this case. The goal is to prove the robustness of the developed framework.

In these experiments we generate samples of a Gaussian signal characterized by a given temporal correlation factor $\rho$ and a given standard deviation $\sigma$. The data is generated using an ARMA model with selected coefficients to obtain the desired $\sigma$ and $\rho$. The input of the model is uncorrelated Gaussian noise, generated by the randn function of MATLAB.

Fig. 2. Transition activity of un-encoded and BI encoded signal. Comparison between experimental results and proposed equations for different number of uncorrelated bits ($L$).

Fig. 3. Transition activity of un-encoded and BI encoded signal. Comparison between experimental results and proposed equations for different values of the transition activity in the MSB ($t_{MSB}$).
Again the length of the data is set to $2^{14}$ samples to allow enough accuracy even in the presence of strong correlations. In order to facilitate the comparison, we keep $B = 12$. Further on, we perform bit-level accurate simulations to obtain the transition activity of each bit of a Bus-Invert encoded Gaussian signal.

In order to use our framework, the parameters $t_{MSB}$ and $L$ are needed. The transition activity in the MSB, $t_{MSB}$, can exactly be obtained from the temporal correlation as $t_{MSB} = \frac{\text{atan}(\rho)}{\pi}$. However, the number of LSBs, $L$ can only be estimated. In [9] the so called break-points $BP_0$ and $BP_1$ are described. They define the region where the bits are changing from the un-correlated behavior to the correlated one. As a rough estimation of $L$, we are using the first term of $BP_0$, i.e., $L = \log_2(\sigma)$ here.

The activity in the redundant bit is estimated using Eq. (7). As suggested by Eq. (8) we use the factor $\frac{\beta}{2}$ for estimating the transition activity in the LSBs, while for the activity in the MSBs the factor $\alpha_2$ is employed. The location of the breakpoints is adjusted to match the expected value. The results of the simulation for $\rho = 0.8$ and $\rho = -0.3$ are depicted in Fig. 4.

We observe that there is a very good match between the model and the experimental results. As the standard deviation $\sigma$ decreases, the transition activity in the LSBs increases, while the transition activity in the MSBs decreases. The trend is properly modeled by the proposed framework. Moreover, even if the signal has a negative correlation (e.g. $\rho = -0.3$) which corresponds to a $t_{MSB}$ larger than 0.5, the model properly predicts that the activity after coding is smaller than 0.5 for the non-redundant bits. It can only be larger than 0.5 for the redundant bit, as predicted by Eq. (7).

IV. Conclusion

The current work provides a solid framework for the mathematical analysis of Bus-Invert under the presence of temporal and spatial correlations. Using a simplified DBT model, it provides simple closed formulas to determine the switching activity of signals encoded with Bus-Invert. Previously proposed results can be obtained as a particular case of Eq. (7) and Eq. (8) when $M = 0$.

Experimental evidence has been provided to show that the proposed formulation is exact for signals following the here defined simplified DBT model consisting of two regions. Moreover, several experiments show that the current framework could be used to estimate the transition activity in the more general case of Gaussian signals.

Further work is related with the analysis of coupling activity, and the detailed analysis of Gaussian signals.

REFERENCES


Fig. 4. Transition activity for each bit of Gaussian signal encoded with BI. Comparison between experimental results and proposed equations for $\rho = 0.8$ $\sigma = [328, 164, 82, 41]$ (upper figure) and $\rho = -0.3$ $\sigma = [205, 102, 51, 26]$ (lower figure) respectively.
If we now plug this result in Eq. (6), we can observe that the second term vanishes. The cost \( \min(h_x, B - h_x) \) is symmetric with respect to \( B/2 \), while \( \left( \frac{L}{2} - h_x \right) \) is antisymmetric. Thus:

\[
T_y = \sum_{h_x=0}^{L} \left( \frac{L}{h_x} \right) 2^{-L} \min(h_x, B - h_x)
\]

Now, since \( \min(h_x, B - h_x) = h_x + \min(0, B - 2h_x) \),

\[
T_y = \sum_{h_x=0}^{L} \left( \frac{L}{h_x} \right) 2^{-L} h_x + \sum_{h_x=[B/2]}^{L} \left( \frac{L}{h_x} \right) 2^{-L} [B - 2h_x]
\]

For the first term,

\[
T_{y1} = \sum_{h_x=0}^{L} \left( \frac{L}{h_x} \right) 2^{-L} h_x = \frac{L}{2}
\]

while for the second one, \( T_{y2} \), we can use the transformation \( x = L - h_x \) to obtain

\[
T_{y2} = \sum_{h_x=[B/2]}^{L} \left( \frac{L}{h_x} \right) 2^{-L} [B - 2h_x]
\]

\[
= \sum_{x=0}^{L-[B/2]} \left( \frac{L}{h_x} \right) 2^{-L} \left[ M - 2 \left( \frac{L}{2} - x \right) \right]
\]

Taking into account the following well known identity:

\[
\sum_{k=0}^{x} \left( \frac{L}{k} \right) \left( \frac{L}{2} - k \right) = \frac{L}{2} \left( \frac{L-1}{x} \right)
\]

we conclude that

\[
T_{y2} = M \text{bincdf}_L \left( L - \left[ \frac{B}{2} \right] \right) - L \left( \frac{L-1}{\left[ B/2 \right] - 1} \right) 2^{-L}
\]

Since \( \left[ B/2 \right] - 1 = \left[ (B-1)/2 \right] \) we can combine Eq. (12) and Eq. (13) to obtain:

\[
T_y = \frac{L}{2} + M \left[ 1 - \text{bincdf}_L \left( \left[ \frac{B+1}{2} \right] \right) \right] - L \left( \frac{L-1}{\left[ B/2 \right] - 1} \right) 2^{-L}
\]

which can be arranged in the form of Eq. (8).