High-Level Current Macro-Model For Power-Grid Analysis *

Srinivas Bodapati
ECE Dept. and Coordinated Science Lab.
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801, USA
bodapati@uiuc.edu

Farid N. Najm
ECE Department
University of Toronto
Toronto, Ontario, Canada M5S 3G4
f.najm@utoronto.ca

ABSTRACT

We present a frequency domain current macro-modeling technique for capturing the dependence of the block current waveform on its input vectors. The macro-model is based on estimating the *Discrete Cosine Transform* (DCT) of the current waveform as a function of input vector pair and then taking the inverse transform to estimate the time domain current waveform. The input vector pairs are partitioned according to Hamming distance and a current macro-model is built for each Hamming distance using regression. Regression is done on a set of current waveforms generated for each circuit, using HSPICE. The average relative error in peak current estimation using the current macro-model is less than 20%.

Categories and Subject Descriptors

B.7 [**Hardware**]: Integrated Circuits—*CAD*; B.7.2 [**Integrated Circuits**]: Design Aids—*Modeling*

General Terms

Algorithms

Keywords

Power grid, Current macro-model, DCT

1. INTRODUCTION

Integrated circuits are drawing increasingly large currents from the power grid. Such large currents in the power grid aggravate physical reliability of IC's due to electromigration. They also cause significant voltage drop leading to circuit slow-down or soft errors. As a result, power grid analysis and design is now an important concern during chip design, not only for power dissipation and reliability reasons, but for performance reasons as well. Therefore, it is important to do early design planning of the power grid, in order to reduce the chance of having to redesign large parts of it [4]. To enable this, in an environment where previously designed blocks are being reused (hard IP blocks) we propose a bottom-up current waveform macro-model for logic blocks. We work at a level of abstraction that may be called structural RTL: the circuit is described as an interconnection of flip-flops and Boolean (combinational) logic blocks. In this context, we develop a cycle-based model for the current waveform of each combinational block that captures the dependence of the current waveform per-cycle as a function of the *vector pair* applied at the block inputs. Previous work has targeted either the average power [11, 9, 13, 6, 2] or energy-per-cycle [12, 7].

Current waveform macro-modeling is difficult because of the large variations that are possible in current waveform shapes, and due to the very large number of possible vector pairs. To overcome this problem, we have developed a macro-modeling technique that is based on Discrete Cosine Transform (DCT). In this approach, we estimate the DCT of the current waveform, instead of the the time domain current waveform and then use the inverse DCT to get the time-domain current waveform. Unlike time-domain current waveforms, the DCT of current waveform is more regular. Specifically, large variations in waveform shapes in the time domain translate to variations only in the *parameters* of the DCT but not its overall shape. Therefore, in our modeling approach, we use a *template* function for the DCT of the per-cycle current waveform. The *template* parameters (discussed in section 4) depend on the input vector pair. We have found that one can use low-order polynomial models to capture this dependence. We use regression to generate these polynomials, based on a number of randomly generated vector pairs for which the circuit is simulated in HSPICE.

The form of this template function is not exactly *derived*, but *inferred* by examining the forms of the frequency transforms of two types of current waveform shapes that one commonly sees in practice. In the next section, we will describe the frequency transforms that will be used as the basis for inferring a DCT model template.

2. DISCRETE COSINE TRANSFORM

The current waveform obtained from circuit simulation (HSPICE) is a discrete-time signal, which can be considered

*This research was supported by the Semiconductor Research Corporation (SRC), under SRC 97-DJ-484 (with funds from Texas Instruments Inc.) and SRC 99-TJ-682 (with funds from IBM Corp.).

DAC 2002, June 10-14, 2002, New Orleans, Louisiana, USA.

as being obtained from the periodic sampling of a continuous-time current waveform, i.e, $i[n] = i_c(nT),\ 0 \le n \le N-1$, where N is the length of the current sequence, $i_c(\cdot)$ is the continuous-time current waveform and T is the *sampling period*, whose reciprocal is the *sampling frequency* [10]. In case of circuit simulation, T is equal to the (fixed) time step specified in the transient analysis (we have used a time step of 0.01 ns). The 1-dimensional Discrete Cosine Transform [8] (DCT) of a sequence $\{i\,[n]\,, 0 \le n \le N-1\}$ is defined as:

$$I[k] = \alpha\,(k) \sum_{n=0}^{N-1} i\,[n] \cos\left[\frac{\pi\,(2n+1)\,k}{2N}\right],\ 0 \le k \le N-1 \quad (1)$$

where $\alpha(0) = \sqrt{1/N}$, $\alpha(k) = \sqrt{2/N}$, for $1 \le k \le N-1$, and the inverse transform is given by:

$$i[n] = \sum_{k=0}^{N-1} \alpha\,(k)\, I[k] \cos\left[\frac{\pi\,(2n+1)\,k}{2N}\right],\ 0 \le n \le N-1 \quad (2)$$

In order to gain some insight into the form that the DCT current macro-model should take, we will look at the analytical forms of the Fourier transform (FT) corresponding to simplified representations of typical current waveforms. For example, we will consider a piecewise-linear triangular current waveform and construct its FT as an analytical closed-form expression. The *form* of this FT will suggest what forms we should use in our DCT model template. The reason this makes sense is that there is a relationship between the two transforms, as follows: the DCT is closely related to the DFT (Discrete Fourier Transform), which is a sampled form of the DTFT (Discrete Time Fourier Transform), which is itself related to the FT. We will now explain these relationships.

The N-sample DCT is related to a $2N$-sample DFT [8], as follows. The DCT of an input sequence of N-samples can be obtained by extending the input to a $2N$-sequence sample with even symmetry, taking a $2N$-point DFT, and saving N terms of it. The even extension of $i[n]$ is defined as:

$$i'[n] = \begin{cases} i[n] & n = 0, 1, \ldots, N-1 \\ i[2N-1-n] & n = N, N+1, \ldots, 2N-1 \end{cases} \quad (3)$$

and the $2N$-point DFT of $i'[n]$, is given by:

$$\begin{aligned} I'[k] &= \frac{1}{\sqrt{2N}} \sum_{n=0}^{2N-1} i'[n] e^{-j(2k\pi/2N)} \\ &= \sqrt{\frac{2}{N}} e^{(jk\pi/2N)} \sum_{n=0}^{N-1} i[n] \cos\left[\frac{(2n+1)k\pi}{2N}\right] \end{aligned} \quad (4)$$

which shows that the $2N$-point DFT and the DCT are closely related. The relationship between the DFT and the DTFT [10], is as follows. The DFT of a sequence, $\{i'_n\}_{n=0}^{2N-1}$, is a set of evenly spaced samples of the DTFT over the frequency range 0 to 2π, multiplied by a constant factor to make the DFT an orthonormal transform. Thus, $I'[k] = \frac{1}{\sqrt{2N}} I_d\,(\omega)\,|_{\omega=\frac{2\pi k}{2N}}$ where $k = 0, 1, \ldots, 2N-1$ and $I_d(\omega)$ is the DTFT of the sequence $\{i'_n\}$, defined by: $I_d(\omega) = \sum_{n=0}^{2N-1} i'[n] e^{-j\omega n}$, where ω is the frequency in radians. Finally, if we consider that the sequence $\{i'_n\}$ is obtained by sampling from an even-extended continuous time current waveform, then the relationship between $I_d(\omega)$ and the Fourier transform (FT) of the continuous waveform, denoted by $I_c(\Omega)$, is given by [10]: $I_d(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} I_c\left(\frac{\omega+2\pi n}{T}\right)$ where Ω has been replaced by $\left(\frac{\omega+2\pi n}{T}\right)$. When $n = 0$, which corresponds to a DTFT between $-\pi \le \omega \le \pi$, we get $\Omega = \frac{\omega}{T}$ [10] . Thus, in summary, the DCT of a discrete time sequence is related to the Fourier transform of an even-extended version of the continuous-time function from which the given discrete-time samples were taken. In the next two sections, we discuss different possible approximations to the actual current waveform, present their corresponding continuous transforms, analyze them and then develop the DCT current macro-model templates based on them.

3. ANALYSIS OF CURRENT WAVEFORMS

A CMOS combinational logic circuit draws current over one cycle in response to a vector pair at its inputs, and these waveform shapes typically fall into three categories:

Approximately Triangular: In this case, the time domain current waveform has a triangular geometry, and the DCT of the current waveform is similar to the DCT of a regular triangular wave.

Approximately Trapezoidal: In this case, the time domain current waveform has a trapezoidal geometry, and the DCT of the current waveform is similar to the DCT of a regular trapezoid.

Multiple peaks: In this case the time domain current waveform has multiple peaks which are separated in time. Currently we use partitioning in time to convert such a waveform to a sequence of waveforms that are of one of the above forms.

In the following, we will explore the FT of the triangular and the trapezoidal current waveforms, which will be used to infer reasonable forms for our DCT model template.

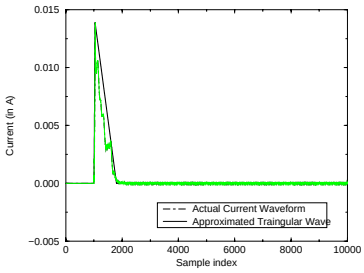

Figure 1(a). Typ. triangular current waveform and its approx.

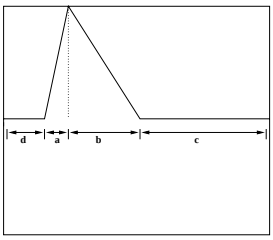

Figure 1(b). Explanation of $f_{\text{tri}}(t)$.

3.1 Triangular Current Waveform

A typical triangular current waveform and its triangular waveform approximation are shown in Fig. 1(a). Before taking the DCT, we increase the waveform time duration and assume zero values for the waveform over the time extension. This is referred to as "zero padding," and its main advantage is that it increases the resolution of the DCT (the corresponding DTFT is sampled more closely), which gives a smoother DCT curve. Zero padding in the time domain is equivalent to using a higher sampling frequency in the frequency domain of the DTFT to get the corresponding DFT/DCT [10]. Thus, even though the original current waveform ends at 2000 samples (which corresponds to 20 ns for our sampling period of 0.01 ns), we have shown 10,000 sample points.

We construct an even extension of this triangular waveform and derive the corresponding continuous Fourier transform. The triangular waveform shown in Fig. 1(a) can be

analytically expressed as:

$$f_{\rm tri}(t) = \begin{cases} 0, & \text{for } 0 \le t \le d \\ A(t-d)/a, & \text{for } d \le t \le a+d \\ A(a+b+d-t)/b, & \text{for } a+d \le t \le a+b+d \\ 0 & \text{for } a+b+d \le t \le Q \end{cases} \quad (5)$$

where a, b, c, d are the dimensions shown in Fig. 1(b), and $Q = a+b+c+d$. The Fourier transform of the even extension of $f_{\rm tri}(t)$, denoted by $F_{\rm tri}(\Omega)$ is given by:

$$\begin{aligned} F_{\rm tri}(\Omega) &= Ae^{-j\Omega(Q)}\left[\frac{2\cos\Omega(b+c) - 2\cos\Omega(a+b+c)}{a\Omega^2}\right] \\ &+ Ae^{-j\Omega(Q)}\left[\frac{2\cos\Omega(b+c) - 2\cos\Omega(c)}{b\Omega^2}\right] \quad (6) \end{aligned}$$

where Ω is the frequency in the analog (continuous) domain and is related to ω, the digital frequency by $\omega = \Omega T$. In our case, we have $T = .01$ ns and $N = 10,000$ before the even extension. After extension, we have $2N$ samples therefore $\omega = 2\pi k/2N = \pi k/10,000$ and $\Omega = \pi k/100$ rad/ns. It is important to note that for $\Omega = \pi k/100$, $F_{\rm tri}(\Omega)$ is real for all integral values of k because $Q = a+b+c+d = 100$ ns.

In Fig. 2(a), we show the first 100 samples (for clarity) of the 10,00 point DCT of a current waveform, DCT of its triangular waveform approximation and the plot of the continuous transform. Fig. 2(a) is typical of all triangular current waveform shapes that we have seen. The DCT plot of the current waveform appears to be a decaying sinusoid. Therefore, instead of constructing a model for each point on the DCT (every frequency component), we use a generic function (a template) to model the entire DCT. In order to infer the form of this template, we compare the plot of continuous transform with the DCT plot. Except for a scale factor, it is obvious that $F_{\rm tri}(\Omega)$ has the right shape and thus may help us define the form of the current macro-model.

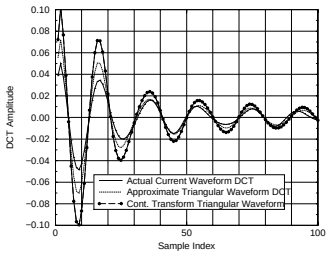

Figure 2(a). DCT of a typ. triangular current waveform.

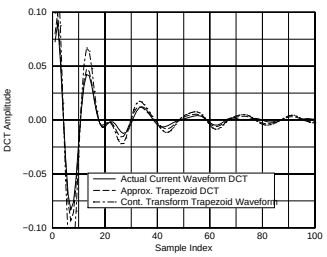

Figure 2(b). DCT of a typ. trapezoidal current waveform.

The most important aspect to observe is that the amplitude of $F_{\rm tri}(\Omega)$ decays as the square of the frequency, a fact which we will make use of in developing our DCT model template. Since the DCT looks like a decaying sinusoid, we can use a simple sinusoid which decays as a square of the frequency. But unlike a decaying sinusoid which has a constant time period, the plots in Fig. 2(a) show that in case of a current waveform and its approximations (both discrete and continuous transform plots) the time period is varying (we can use the difference between consecutive/alternate zero crossings to measure the time period). We use this fact in our model as well. It actually helps simplify our current macro-model, in the sense that we do not use multiple cosine terms as in $F_{\rm tri}(\Omega)$, instead we use a simplified expression with a variable time period, as discussed in section 4.

3.2 Trapezoidal Current Waveform

In some cases, the current waveform has *trapezoidal* shape, as shown in Fig. 3(a), which also shows a piecewise linear trapezoidal approximation to the current waveform. The equation of the trapezoidal wave is given by:

$$f_{\rm tra}(t) = \begin{cases} 0, & \text{for } 0 \le t \le a \\ A(t-a)/b, & \text{for } a \le t \le a+b \\ A, & \text{for } a+b \le t \le a+b+c \\ A(U-t)/d, & \text{for } a+b+c \le t \le U \\ 0 & \text{for } U \le t \le W \end{cases} \quad (7)$$

where the dimensions are illustrated in Fig. 3(b), $U = a + b + c + d$ and $W = a + b + c + d + e$. The Fourier transform of the even extension of $f_{\rm tra}(t)$, denoted by $F_{\rm tra}(\Omega)$ is given by:

$$\begin{aligned} F_{\rm tra}(\Omega) &= Ae^{-j\Omega(W)}\left[\frac{2\cos\Omega(c+d+e) - 2\cos\Omega(W-a)}{b\Omega^2}\right] \\ &+ Ae^{-j\Omega(W)}\left[\frac{2\cos\Omega(e+d) - 2\cos\Omega(e)}{d\Omega^2}\right] \quad (8) \end{aligned}$$

As before, it turns out that $F_{\rm tra}$ is real for the values of Ω under consideration. In Fig 2(b), we show the DCT of the current waveform, the DCT of the trapezoidal waveform approximation, and the plot of the continuous transform. Again, we can observe that, except for a scale factor, the continuous transform plot compares well with the actual current waveform DCT. Thus, we can use the continuous transform to infer the shape of the DCT model template. The most obvious inference we can draw from the continuous transform, is, as before, the quadratic decay in the amplitude of the DCT, in terms of frequency.

If we compare the plots in Fig. 2(a) and Fig. 2(b), we can see that the DCT plot of a trapezoid shows some deviation from the perfectly decaying sinusoid of triangular waves. The same observation holds true for the respective continuous transform as well, therefore we cannot use a decaying sinusoid with a varying time period to simplify our model as in the case of triangular waveforms (see section 4). Instead, we will use a slightly different analytical expression for the trapezoidal current waveform, which is based on $F_{\rm tra}(\Omega)$ to some extent, but also based on observations related to the plots. Specifically, for trapezoidal waveforms, we will use two cosine terms (the continuous transform has 4 cosine terms), and again use a varying time period. We use the difference between successive maxima (or minima) as as a measure of the time period. The DCT plots show that this difference varies for successive maxima (or minima). Details are given in section 4.

3.3 Waveform Partitioning

For multiple peak current waveforms we use a simple partitioning algorithm in the time domain, in which we detect the index of the first peak, and then traverse to a point where the current value is less 10% of the maximum value. We repeat this process and partition the time domain waveform into sub-intervals so that each sub-interval contains a single peak, and then use the same macro-modeling technique discussed in section 4. This technique is used only in cases where we have peaks which are very far apart. Partitioning also keeps the number of current macro-model templates to only two. Since in the current macro-model characterization phase we need to automatically detect which template to use depending on the error criterion (the error criterion

Table 1: Dummy Coding for transitions at a single node.

y	x_1	x_2	x_3	y
$0 \rightarrow 1$	0	0	1	1
$1 \rightarrow 0$	0	1	0	2
$1 \rightarrow 1$	1	0	0	3
$0 \rightarrow 0$	0	0	0	4

would detect which of the two templates is a better fit), it is good to minimize the number of possible templates.

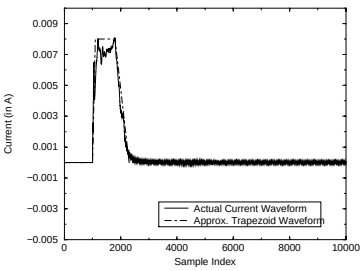

Figure 3(a). Typ. trapezoidal current waveform and its approx.

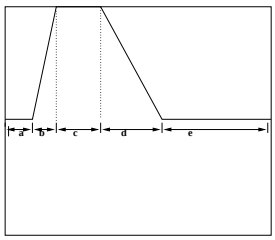

Figure 3(b). Explanation of $f_{\rm tra}(t)$.

4. MACRO-MODEL CONSTRUCTION

In order to formulate a macro-model, one needs to develop a mapping between the input Boolean vector pair and certain variables that can be the input variables to the macro-model. These variables are used as *predictor variables* in regression. For a given input vector, a single node at the primary input can undergo one of the following four transitions $\{0 \rightarrow 1, 1 \rightarrow 0, 1 \rightarrow 1, 0 \rightarrow 0\}$. We treat the set of four possible transitions at each primary input as a categorical variable [5]. In general, a categorical variable with k levels (in our case 4) is transformed into $k-1$ variables (in our case 3) each with 2 levels. This process of creating variables from categorical variables is called **dummy coding**. In our case, we use three variables x_1, x_2 and x_3 to code the categorical variables. The dummy coding for the 4 levels of our categorical variable is shown in the second, third and fourth columns of Table 1 (and also shown in [12]). These three columns are labeled as x_1, x_2, x_3 in Table 1. Basically, each input transition is mapped to a vector of $[x_1\ x_2\ x_3]$, but this has a clear disadvantage in that it increases the number of variables as well as the number of regression coefficients. But we have found that, depending on the Hamming distance of the vector pair, we can introduce a less expensive solution, as follows. For large Hamming distances (more than 60% of the inputs are switching), very few inputs undergo $0 \rightarrow 0$ and $1 \rightarrow 1$ transition, therefore we can assume that the categorical variable has just two levels, so that a single two-valued variable is required to represent a primary input. The mapping for large Hamming distance is shown in the fifth column of Table 1, labeled y.

4.1 Triangular Template

The triangular template is a simplified as well as a modified version of $f_{\rm tri}(\Omega)$. The simplification is achieved because we explicitly incorporate the effect of the changing time period (which happens because of multiple cosine terms in $f_{\rm tri}(\Omega)$). We estimate the first few time periods (in our case, five) directly and use them in a single cosine term, whose other parameters do not change. We can do this only because of the energy compaction achieved with orthogonal transforms like DCT. The template equation for the DCT of the triangular current waveform is given by: $I_{\rm tri}(k) = D(k)A\cos\left(\frac{2\pi(k-1)}{T_i}\right),\ k = 1, 2, \ldots$, where k is the sample index, $D(k)$ is a decay factor, A is a scale factor, which we refer to as amplitude, and $T_i,\ i \in \{1, 2, 3, 4, 5\}$ is the variable time period. The DCT value corresponding to $I_{\rm tri}(k)|_{k=0}$ is called the DC Value. These terms are the *parameters* of the template, which we relate to the variables associated with the input vector pair. The triangular template is similar to what was presented in [1] (the trapezoidal template is new), where it was shown that the time period T_i does not vary significantly for a given Hamming distance of the input vector pair, prompting a partitioning of the model according to Hamming distance. Thus, T_i is estimated as in [1]. As for $D(k)$, we use the following template for it, which is motivated by the inverse square dependence on frequency seen above: $D(k) = \frac{f(x)k}{g(x)k^2+h(x)}$ where x is either y (a vector of length one) or the vector $[x_1\ x_2\ x_3]$ (refer to Table 1). This expression is similar to that defined in [1], except that we have modified $f(x), g(x)$ and $h(x)$, to account for possible interaction between the input variables, through cross-product terms, so that $f(x), g(x)$ and $h(x)$ are now of the form (where $\mathcal{H}$ stands for Hamming distance):

$$\Pi_{i=1}^{p}(\alpha_i y_i + \beta_i),\ \text{for large } \mathcal{H}. \quad (9)$$

$$\Pi_{i=1}^{p}(\alpha_{i1}x_{1,i} + \alpha_{i2}x_{2,i} + \alpha_{i3}x_{3,i} + \beta_i),\ \text{for small } \mathcal{H}. \quad (10)$$

where p is the number of primary inputs. The coefficients α and β are estimated using regression as in [1]. The Amplitude A, and the DC Value $I_{\rm tri}(0)$ are also modeled along the lines of $f(x), g(x)$ and $h(x)$. Their coefficients are also estimated using regression. The triangular template can provide a good model in cases of both large and small Hamming distances, but in some cases for small Hamming distance, the trapezoidal template is better. However, estimating the coefficients of the parameters of the triangular template is simpler.

4.2 Trapezoidal Template

A trapezoidal template equation model seems to best fit the current waveforms for low Hamming distances for some circuits, when the fraction of primary inputs switching is less than 60% or so. Our trapezoidal template is given by:

$$\begin{aligned} I_{\rm tra}(k) &= \frac{A}{D'(k)}\left[B(k)\cos\left(\frac{2\pi(k-1)}{T_i}\right)\right] \\ &+ \frac{A}{D'(k)}\left[C(k)\cos(2\pi(k-1)\omega)\right] \end{aligned} \quad (11)$$

where $k = 1, 2, \ldots$ is the sample index, $T_i,\ i \in \{1, 2, 3, 4, 5\}$, is the variable time period, A is the amplitude and $B(k) = p(x)k,\ C(k) = q(x)k,\ D'(k) = r(x)k^2 + s(x),\ \omega = t(x),\ A = w(x)$ where x is the vector $[x_1\ x_2\ x_3]$. The functions $p(x),\ q(x),\ r(x),\ s(x),\ t(x)$ and $w(x)$, all have the same functional form (with different coefficients), given by: $\Pi_{i=1}^{p}(\alpha_{i1}x_1 + \beta_{i2}x_2 + \gamma_{i3}x_3 + \delta_i)$, where the coefficients are estimated using non-linear regression. Since the number of parameters in the template is large compared to $I_{\rm tri}$, we have a large number of coefficients and estimating the coefficients of the parameters of the trapezoid is computationally more expensive than estimating the coefficients of a triangular

wave. The polynomial function used to model the DC value $I_{\text{tra}}(k)|_{k=0}$ is similar to that used for A. The coefficients of the function corresponding to $I_{\text{tra}}(k)|_{k=0}$ are estimated separately, using simple linear regression. The form of the above template is motivated by the functional form of the the continuous Fourier transform expression $F_{\text{tra}}(\Omega)$ for a trapezoid waveform, which was given above (8) and which involves four cosine terms, and decays as the square of the frequency. Since four cosine terms may be too many (computationally expensive model), and a single cosine term cannot capture the deviation from a simple decaying sinusoid, we have used only two cosine terms in our template, decaying with the square of frequency. In order to maintain a good fit with just two cosine terms, we have found that using separate frequency terms in the numerator, as we have done with $B(k)$ and $C(k)$, is advantageous.

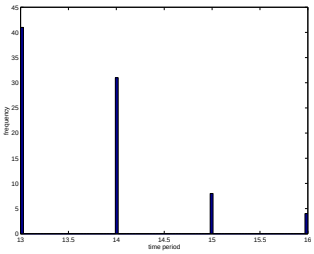

Figure 4. Dist. of the first time period for alu2.

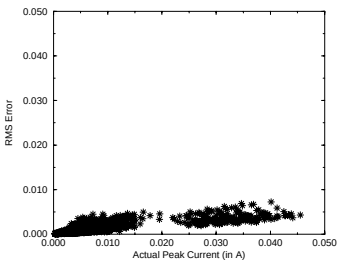

Figure 5. RMS error vs. peak current.

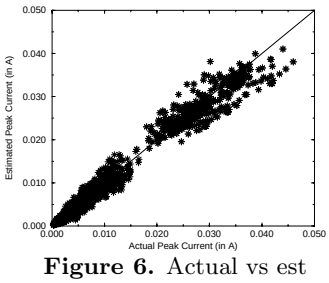

Figure 6. Actual vs est peak current.

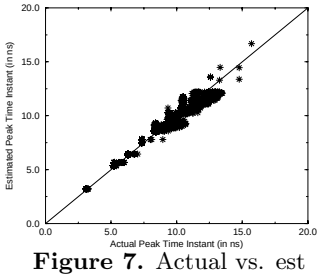

Figure 7. Actual vs. est time inst. when peak occurs.

The philosophy behind using two cosine terms is that one cosine term would produce a basic decaying sinusoid (as in the triangular case) and the other would account for the deviations from the decaying sinusoid that are seen at low Hamming distance. Therefore, the variable time period T_i is used with only one of the two cosine terms, the one which is aimed at generating the decaying sinusoid. The frequency ω of the other cosine term is estimated as a function of input vector pair using regression. The variable time period terms T_i are approximated as the difference between consecutive maxima (or minima) of the DCT. As in the triangular case and [1], the time period does not vary much for a given Hamming distance, as shown in Fig. 4. Thus, we use the same method as in [1] to get an estimate of the first five time periods. Basically, from the distribution of time periods for a given Hamming distance, we choose the time period value which occurs most often (in statistical terms, we choose the *mode* of the time period distribution as the time period for a given Hamming distance). The first five time periods are enough to get an estimate of the dominant terms of the DCT.

5. EXPERIMENTAL RESULTS

A set of randomly generated vectors were used to simulate various benchmark circuits in HSPICE, and the corresponding current macro-models were built for these circuits. The resulting macro-models were tested for accuracy on a different set of randomly generated vector pairs for each of the benchmark circuits listed in Table 2, which also shows the number of inputs and gates for each circuit. In Table 2, e_{pavg} (5th column) denotes the **average relative error** in peak current estimation, and e_{tavg} (6th column) denotes the **average relative error** in the estimation of the time instant at which the peak current occurs. The 4th column in Table 2, denoted by RMS_{avg} lists the **average Root Mean Square Error** [8] between the estimated and the actual current waveforms. The average error in all the three cases has been computed over the number of input vector pairs over which the macro-models have been tested. The last column in Table 2 shows the total number of simulations (number of vector pairs) needed to build the macro-model for all possible hamming distances, for each circuit. The number of simulations can be reduced by combining models for various hamming distances.

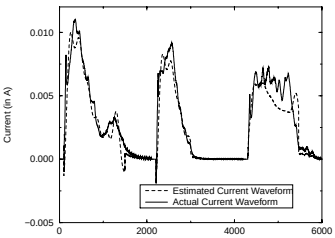

Figure 8. For alu2, ham dist =4.

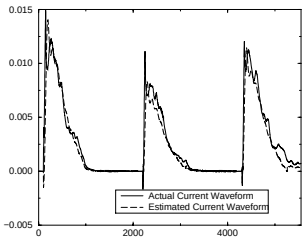

Figure 9. For alu2, ham dist =8.

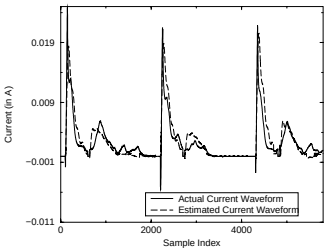

Figure 10. For c432, ham dist =32.

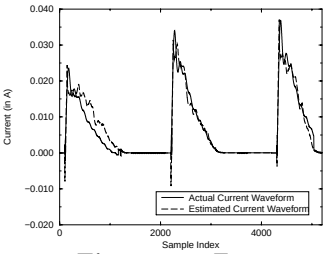

Figure 11. For c880, ham dist = 55.

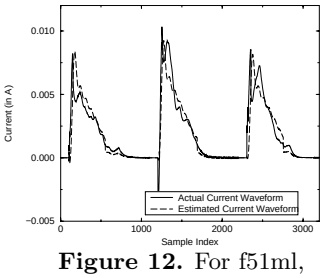

Figure 12. For f51ml, ham dist = 7.

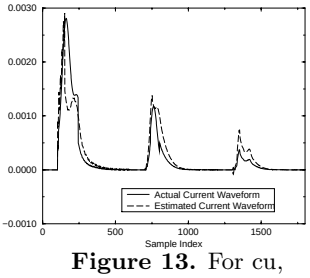

Figure 13. For cu, ham dist = 2.

In order to give some intuition as to the goodness of these average error numbers, and some measure of spread, we compared the RMS error for each vector pair with the actual peak current as shown in Fig 5. The plot shows that the

RMS error is more or less bounded across the various current peak values for the benchmark circuits, and increases very slowly with current peak. Figure 6 shows the correlation plot between actual and estimated peak current and similarly in Fig. 7 we show the correlation plot between actual and estimated *time instant* at which peak occurs. Both these correlation plots are mostly linear which implies that both the peak current and the time instant at which the peak occurs can be estimated fairly accurately using our macro-model.

Finally, in Figs. 8–17, we present a comparison of the actual and estimated current waveform for the benchmark circuits listed in Table 2, for some example Hamming distances, where we superimposed the the estimated current waveform on top of the actual current waveform. These plots show good accuracy in most of the cases, especially for large Hamming distances. This occurs because when a small fraction of inputs are switching, the current waveform significantly depends on what inputs are switching, which leads to large variations in the current amplitude. However, the current magnitude itself is much lower in case of low Hamming distance, so that the absolute error is actually small. Having said this, we must point out that the model in fact works very well, capturing the required current waveforms in a high-level black box macro-model. Finally, we should mention that even though our macro-model has been built using HSPICE, one can build current macro-models for much larger circuits using more efficient simulators, like PowerMill, or gate-level simulation techniques such as [3].

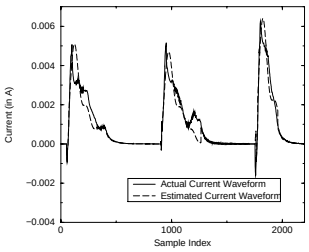

Figure 14. For x2, ham dist = 8.

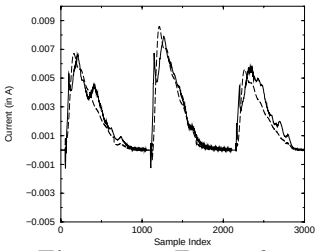

Figure 15. For random8, ham dist = 4.

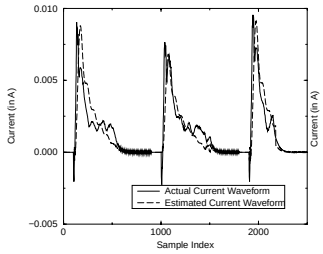

Figure 16. For parity, ham dist = 14.

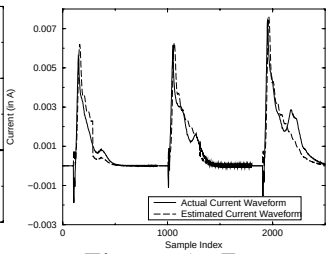

Figure 17. For mux, ham dist = 19.

6. CONCLUSION

In order to enable early block-level analysis of the power grid, when using hard IP blocks, we have proposed a cycle-based current waveform modeling technique that involves predicting the DCT of the current waveform from the input vector pairs, and using the inverse DCT to get back the time-domain waveform. The parameters of the DCT template function are expressed in terms of input vector pairs, using regression. While we continue to improve the estimation and validate this technique on more circuits, the data so-far show that this type of estimation is indeed possible, enabling early block level current estimation for power grid design.

Table 2: Benchmark circuits used in Figs. 5–17.

Circuit	#I	#Cells	RMS_{avg}	$e_{\text{pavg}}\%$	$e_{\text{tavg}}\%$	#sim.
c880	60	383	.0032	11.98	5.92	7000
alu2	10	368	.0023	17.70	7.51	700
c432	36	217	.0030	12.43	6.45	4500
cu	14	48	.00049	18.91	1.99	770
f51ml	8	105	0.0011	19.21	2.12	320
mux	21	91	.0004	17.86	3.00	2000
random8	8	158	0.0010	14.14	6.02	320
parity	16	68	.00077	15.37	3.98	640
x2	10	50	.00054	14.35	2.43	500

7. REFERENCES

[1] S. Bodapati and F. N. Najm. Frequency-domain supply current macro-model. In *ISLPED*, pages 295–298. IEEE, 2001.

[2] A. Bogliolo and L. Benini. Node sampling: A robust RTL power modeling approach. In *IEEE ICCAD*, pages 461–467. ACM/IEEE, 1998.

[3] A. Bogliolo, L. Benini, G. D. Micheli, and B. Ricco. Gate-level current waveform simulation of CMOS integrated circuits. In *ISLPED*, pages 109–112. IEEE, 1996.

[4] A. Dharchoudhury, R. Panda, D. Blaauw, R. Vaidyanathan, B. Tutuianu, and D. Bearden. Design and analysis of power distribution networks in PowerPC microprocessor. In *ACM/IEEE DAC*, pages 738 – 743. IEEE, 1998.

[5] R. F. Gunst and R. L. Mason. *Regression Analysis And Its Application, A Data-oriented approach*. Marcel Dekker, Inc, 1980.

[6] S. Gupta and F. N. Najm. Power macromodeling for high level power estimation. In *34th ACM/IEEE DAC*, pages 365–370. IEEE, 1997.

[7] S. Gupta and F. N. Najm. Energy-per-cycle estimation at RTL. In *ISLPED*, pages 121–126. IEEE, 1999.

[8] A. K. Jain. *The Fundamentals Of Digital Image Processing*. Prentice-Hall, Inc, 1989.

[9] P. E. Landman and J. M. Rabaey. Architectural power analysis: The dual bit type method. *IEEE Trans. on VLSI Systems*, 3:173–187, June 1995.

[10] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, Inc, 1990.

[11] S. R. Powell and P. M. Chau. Estimating power dissipation of VLSI signal processing chips: The PFA technique. *VLSI Signal Processing IV*, pages 250–259, March 1990.

[12] Q. Qiu, Q. Wu, C.-S. Ding, and M. Pedram. Cycle-accurate macro-models for RT-level power analysis. *IEEE Tans. on VLSI Systems*, 6:420–528, December 1998.

[13] A. Raghunathan, S. Dey, and N. K. Jha. Register-transfer level estimation techniques for switching activity and power consumption. In *IEEE ICCAD*, pages 158 – 165. IEEE, 1996.