

多媒體系統晶片設計技術之研究(III)

系統晶片的功率估測與低功率技術之研究

The Study of Power Estimation and Low Power Techniques for System-on-Chip

計劃編號: NSC89-2218-E-009-076

執行期間: 89年8月1日至90年7月31日

主持人: 沈文仁 國立交通大學電子資訊學院院長

國立交通大學電子工程學系教授

計劃參與人員: 黃恆亮、許智揚、許文俊、陳志榮、陳宏慶、鄭吉成

一、中文摘要

Monte Carlo 方法是一個普遍被用來估計一個族群分佈平均值的方法。而這個方法也被引用到數位電路的功率估測上[1]。在 *Monte Carlo* 方法裏面有一個重要的假設是它的取樣空間必需是 *Normal* 分佈, 但是對一個數位電路而言, 這個假設並不一定會成立。例如一個電路如果在不同的操作模式之下通常會有不同的功率消耗狀況。本計劃提出了一個強化的 *Monte Carlo* 方法 (*Advance Monte Carlo, AMC*), 在這個方法裏面可以對取樣空間所取出來的樣本做分析, 如果取出來的樣本顯示取樣空間和 *Monte Carlo* 方法的說取樣空間是 *Normal* 相違背的話, *AMC* 可以自動加大取樣時的樣本大小, 以提高取樣空間接近 *Normal* 分佈的程度。這樣 *Monte Carlo* 方法才能適用而做出正確的估計。另外如何對取樣空間分層以減少取樣時所需的樣本數目我們也做了一些研究, 並提出了一個全新的取樣空間動態分層方法 (*Adaptive Stratification*)。這個動態分層方法可以利用取出的樣本, 再加上該線路的高階模型模擬結果來做多變數回歸分析 (*Multiple Regression*)。以這個分析結果再對取樣空間的樣本可能值來做分析, 並重新分層, 由實驗結果我們可以看出, 我們所提出的 *Advanced Monte Carlo* 搭配 *Adaptive Stratification* 所構成的 *AMCAS* 確實能有效地減少在估計邏輯線路時所需要的模擬時間。
關鍵詞: 分層, 動態分層, 蒙地卡羅, 強化的蒙地卡羅

英文摘要

Monte Carlo approach of power estimation is based on the assumption that the samples are *Normally* distributed. However, in many cases for the power distribution of a circuit to be non-*Normal*. In this report, an *Advanced Monte Carlo (AMC)* method is proposed with a measure of the normality of the samples. The proposed method can increase the sample size when needed to improve the normality. Besides, a novel input sequence stratification scheme for power estimation is also proposed that needs only one power model for the simulator. The stratification technique utilizes a multiple regression method to compute the coefficient matrix of the input transitions, which is used in the stratification function. This new stratification technique can adaptively update the coefficient matrix and keep the population of input variables a better stratification. The experimental results demonstrate that the proposed *Advanced Monte Carlo with Adaptive Stratification (AMCAS)* effectively reduce the simulation time while keeping the user-specified confidence level and error level, and is thereby an efficient estimation technique for estimating the power consumption of IP level modules.

Key words: Stratification, Adaptive Stratification, Monte Carlo, Advanced Monte Carlo

二、計劃緣由與目的

近幾年來隨著積體電路技術的大飛躍, 晶片愈做愈小而且密度和操作頻率愈高。另外, 由於可攜式計算器材和無線通訊的蓬勃發展, 例如手提電腦, 手提攝錄, 個人通訊手機, 個人數位助理、等等

裝備需要快速和複雜的運算功能，而且也需要體積小、重量輕的電池，這使得功率消耗的問題漸漸的受重視而成為設計者優先考慮的因素。事實上，消耗大量功率的晶片會降低晶片的壽命、可靠度和效率，甚至需要額外的排熱裝置來降低溫度，所以不適用於可攜式的產品。因此，如何估測功率消耗並進而降低晶片的功率消耗已成為近幾年來最熱門的研究課題之一。但是要設計一個低功率電路，一個快而且準確的功率消耗估測工具是非常需要的，因為它可幫助設計者很快驗證他的設計是否比別的方法省功率。

由於製程技術的進步，使得將複雜的系統整合在單一晶片的設計方法成為未來的一種趨勢。晶片設計者不必自己設計所須要的所有模組，而可以用虛擬元件(Virtual components)來代替，然後再向廠商購買符合虛擬元件功能且可重複使用的 Intellectual properties (IPs)，就可以很快的設計出自己所要的晶片了。這種系統晶片除了面臨以上所說的電晶體密度及操作頻率高而消耗大量功率的問題之外，很重要的一點是它是由不同廠商所提供的 IP 所組成，而不是單由自己發展的模組所組成，這造成了功率估測及低功率設計的困難度。一個晶片可能由不同的廠商所提供不同類型的 IP 所組成。對於系統晶片設計商(IP Consumer)，雖然他可以從模組提供者(IP Provider)那邊取得一些功率消耗的值，或是更理想的有功率消耗模型，但是這個設計是關係到系統晶片設計商的一代產品的成功與否，甚至可以影響到系統晶片設計商的生死存亡。從 IP Provider 那兒拿到的功率模型是否準確？或者是功率消耗值是否是在合理的環境下所量出來的？可信度有多高？這些問題的答案都不是系統晶片設計商願意冒險去猜測的。所以在拿到一個硬 IP (hard IP) 之後，系統晶片設計商通常會想盡辦法自己去做功率估測。另外，由於高階功率估測軟體(如 *Synopsys* 的 *PrimePower*) 或是功率最佳化軟體(如

Synopsys 的 *PowerCompiler*)的訂價都非常高，所以目前在許多小型的 IC 設計公司裏可花錢來請有經驗的 IC 設計者。還有另一種情況是，IC 設計者使用了其他公司所提供者是 cell library，但是由於這些 IP 或 cell 的 power model 並不是那麼準或是根本不准，上述三種情況都是 IC 設計者會想要跑 transistor-level 的 power simulation 的情況。但是目前大部份的 IC 設計者只知道輸入的訊號可以測出某個 fault (也就是 test 或是輸入什麼樣的訊號可以量出線路的延遲 (critical-path delay)，但是要估測功率消耗者大概只能隨便挑幾個輸入訊號來模擬，而且目前邏輯線路的大小，也很難讓做一個線路的真正在使用時會遇到的所有情況都做模擬。所以，這也就是本計劃為什麼要在 IP 設計的環境中找出一個可以替換者挑出用來做功率估測的輸入訊號做原因。

本計劃的目的就是，利用一個 transistor-level net-list 以及它的 HDL net-list 有效率的功率估測用輸入訊號選擇系統

三、研究方法及成果

首先，讓我們先來看看傳統的 *Monte Carlo* 是如何估計一個取樣空間(population)的均值(mean)的。

A. 傳統的 *Monte Carlo* 方法

先定義各層的取樣空間。假設我們有一串估測平均功率的輸入訊號(input sequence)度為 $N+1$ 。也就是這一串 input sequence 有 $N+1$ 個輸入模式(input pattern)，每二個 input pattern 組成一個輸入向量(input vector)就是這一串 input sequence 可以拆成 N 個 input vector。第 i 個 input vector 以 V_i 來代表，第 i 個 input vector 所會產生的功率消耗以 P_i 來表示。在真正將 input vector 送進 simulator 去做功率模擬(power simulation)

這些 $Power(V_i)$ 都是未知的。所謂的取樣就是從 N 個 input vector 裏面抽出一個或一些 input vector 送進 simulator 去做 power simulation 以得到被選出的 input vectors 對應的 $Power(V_i)$ 。所以我們可以把所有的 input vector 的 $Power(V_i)$ 看成是如下圖的最上層的 Level 0 population:

圖一：各層取樣空間的定義

經由亂數對 Level 0 的 population 取出 l 個 input vector 的 $Power(V_i)$ 之後，這 l 個 $Power(V_i)$ 的平均值就是 Level 1 population 裏面的一個圈圈，又叫做一個 sample。在圖一的例子裏面，因為 Level 1 的每個小圈圈裏面包含 4 個 level 0 的 population 的點，也就代表在這個例子裏面，Level 1 的一個 sample 是由 4 個 Level 0 的 $Power(V_i)$ 的平均值所組成。這個 l 值又被叫做 sample size。在取出 n 個 sample 之後，這 n 個 sample 的平均值就形成一個 Level 2 population 的一員，又叫做一個 sample average。

由圖一的各層 population 可以看出，Level 2 population 裏的一個圈，是由很多個 ($l*n$ 個) Level 0 的圈圈的平均所形成。根據中央極限定理 (Central Limit Theorem) 所陳述，不管 Level 0 population 是什麼樣的分布，當 $l*n$ 趨近無限大的時候，Level 2 population 都會趨近常態分布。傳統的 Monte Carlo 方法就是假設 Level 2 的 population 是常態分布 (Normal distribution)，在這個假設之下來分析 sample average 和 Level 0 population 的平均值，也就是線路的平均功率消耗之間的差會有多少。若這個差比使用者所要

求的值還大的話，那麼就必需多取一些 vectors 一直到達到使用者的要求為止。斷式如下所示：

$$\frac{t_{(r+1)/2, n-1} S}{\bar{x} \sqrt{n}} \leq V$$

其中各個變數的定義為：

n : number of samples of level 1 population

r : User defined confidence level

$t_{(r+1)/2, n-1}$: $(r+1)/2$ percentile of a t -distribution of $n-1$ degree of freedom.

S^2 : Variance of the n samples of population.

\bar{x} : Mean of the n samples of level 1 population

V : User defined error level.

由於傳統的 Monte Carlo 方法對於它的假設: Level 2 population 為 Normal distribution 沒有根據實際上 level 2 population 接近的程度加以測量，所以在常常會發生無效的估計 sample average 和 level 0 population 平均值之間的差距。

B. Advanced Monte Carlo

針對這個缺點，我們提出了一個加進 Monte Carlo 方法 (Advanced Monte Carlo) 的流程圖如下圖二所示。Advanced Monte Carlo 和傳統的 Monte Carlo 之間的差別就在於，當傳統的 Monte Carlo 已經覺得取出的 input vector 夠多，Advanced Monte Carlo 會開始針對 Level 2 population 的 Normal 程度做分析，並針對 Level 2 population 的 Normal 程度，這與 Monte Carlo 的 Normal 程度的分析，簡稱 Normality Test。我們是採用第三階及第四階標準化 (Standardized 3rd and 4th moment)，這在統計學上又稱為不對稱度 (3rd standardized moment: Skewness) 以及陡峭度 (4th standardized moment: Kurtosis) [4]。

在遇到如下圖的 *Bi-modal* distribution 很容易因誤判 Level 2 population 的 *Norm* 而過早停止。

圖二：Advanced Monte Carlo 的流程圖

對一個 *Normal* distribution 而言，它的 *Skewness* 應該等於 0，也就是說它是左右對稱的，而它的 *Kurtosis* 應該等於 3 所以我們就利用 Level 2 的 *Skewness* 和 *Kurtosis* 來判斷 Level 2 是否真的很接近 *Normal* distribution。計算 Level 2 的 *Skewness* 和 *Kurtosis* 的方法，我們利用下面式子來做說明：

$$std_mom3(1) = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3 \quad (2)$$

$$std_mom4(1) = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4$$

上式中的 $std_mom3(1)$ 和 $std_mom4(1)$ 代表的是 Level 1 的 standardized 3rd moment 和 4th moment，而 n 還是 sample 的數目， x_i 是各個 sample 的值， \bar{x} 值是 sample mean 而 s^2 是 sample variance。利用(2)式所得到的 Level 1 population 的 standardized moments，代入下面這個式子就可以求出 Level 2 population 的 standardized moments：

$$std_mom3(n) = std_mom3(1) / \sqrt{n} \quad (3)$$

$$std_mom4(n) = 3 + (std_mom4(1) - 3) / n$$

上式中的 $std_mom3(n)$ 和 $std_mom4(n)$ 就是當 sample 數目等於 n 的時候 Level 2 population 的 standardized moments。

由我們的實驗中發現，傳統的 *Monte Carlo* 方法

圖三：Bi-modal 分佈及其參數

所以我們利用實驗的方法，對各種 *Bi-modal* 分佈做實驗，以求得 level 2 的 standardized moments 要多靠近 distribution 才能讓 level 2 是 *Normal* dis 的假設成立。最後我們得到一個經驗值

$$|std_mom3(n)| \leq 0.04,$$

$$|std_mom4(n) - 3| \leq 0.002$$

利用上面所得到的經驗公式，我們分別所提出的方法和傳統的 *Monte Carlo* *ISCAS85* 的測試線路進行功率估測，手到的結果如下面表列：

表一：MC VS. AMC with Confidence level $\Gamma=$

Circuit	MC			AMC	
	l	$viol_r$	$nVecs$	$final\ l$	$viol_r$
C432	30	0.0085	4597	53.67	0.00:
C499	30	0.0140	1625	42.69	0.00:
C880	30	0.0161	1703	45.13	0.00:
C1355	30	0.0143	1753	42.49	0.00:
C1908	30	0.0121	2539	45.6	0.00:
C3540	30	0.0167	510	36.09	0.00:
C6288	30	0.0096	3934	44.09	0.00:
Average	30	0.0130	2380	44.25	0.00:

表二：MC VS. AMC with Confidence level $\Gamma=$

Circuit	MC			AMC	
	l	$viol_r$	$nVecs$	$final\ l$	$viol_r$
C432	30	0.0554	2613	49.33	0.03:
C499	30	0.0788	903	38.44	0.03:
C880	30	0.0775	947	42.32	0.03:
C1355	30	0.0761	973	40.72	0.03:
C1908	30	0.0658	1433	43.33	0.03:

C3540	30	0.0597	292	35.49	0.0331	374
C6288	30	0.0666	2219	43.62	0.0379	2554
Average	30	0.0685	1340	41.89	0.0344	1584

表三 : *MC* VS. *AMC* with Confidence level $r=0.90$

Circuit	<i>MC</i>			<i>AMC</i>		
	l	$viol_r$	$nVecs$	$final_l$	$viol_r$	$nVecs$
C432	30	0.1141	1810	47.92	0.0756	2142
C499	30	0.1434	617	36.53	0.0815	784
C880	30	0.1393	642	40.44	0.0766	840
C1355	30	0.1395	662	37.96	0.0745	852
C1908	30	0.1317	984	41.83	0.0819	1214
C3540	30	0.1011	214	35.75	0.0682	278
C6288	30	0.1182	537	43.29	0.0802	1826
Average	30	0.1267	780	40.53	0.0769	1133

表一~三分別是在 confidence level $\alpha=0.99$ 、 0.95 以及 0.90 而 error level $\epsilon=0.05$ 的時候，我們對傳統的 *Monte Carlo* 方法 (*MC*) 和我們所提出的 *Advanced Monte Carlo (AMC)* 所做的比較。在上面的表中， l 是 sample size，由於 *MC* 並沒有動態提高 sample size 的機制，所以它的 sample size 都設成統計學家咸認足夠的 30。另一方面，由於 *AMC* 則會在偵測到 Level 2 的 Normal 程度不夠的時候，把 sample size 提高，所以我們把它的初始 sample size 設為 6，而表中所列的 $final_l$ 為最後 *AMC* 結束的時候所使用的 sample size。下一欄 $viol_r$ 則為最後 *MC* 或是 *AMC* 對 power 的估計值和真正的 power 值之間的差異超過 ϵ 的比例。理論上，當使用者要求對估計結果有 99% 的把握的時候，這個 $viol_r$ 就不應該超過 $1-r$ 。我們可以看到不管在那一個表裏面，傳統的 *MC* 方法在 $viol_r$ 這一欄常常都會超過 $1-r$ 。例如表三裏面， $r=0.99$ ，所以 $viol_r$ 不應該超過 0.01。但是傳統的 *MC* 的 $viol_r$ 全部都超過，反觀我們所提出的 *AMC* 則在每一個表的 $viol_r$ 全部都能保持在 $1-r$ 以下，由此可以看出 *AMC* 確實成功的利用測量 Level 2 population 的方法來達到使用者要求的 confidence level 的目的。

C. Adaptive Stratification

除了 *AMC* 之外，我們還提出了一個新的取樣空

間的動態分層方法。我們知道一個邏輯工作時，常常會出現很多種不同的功例如一個 CPU 在做 $A+1$ ， $A+B$ ，或者是時候就會消耗不同的功率。如果我們在估測的時候，能在取樣空間裏面將已耗不同功率的 input vector 加上 (stratification)，就可以降低在取樣時的 variance，進而達到減少 sample number。也就是可以提高 input sequence 的 (compaction ratio)[2]，上述關係也可由 (1) 當 $t_{(r+1)/2, n-1}$ ， \bar{x} 以及 l 固定的時候，代入 (1) 就可以得到一個較小的 n 。在之前的論文提出的 stratification 的方法中，是利用 g 的功率估測來做為指出各個 input vector 值，但是因為對一個硬 IP 來說，要有 g 的功率模型可以拿來做 gate-level 功率估測無法如願。有時是因為沒有 gate-level 模型，或是沒有 gate-level 的功率估測軟件可能連 gate-level 的線路都沒有。所以我們用了一個只要有該線路的最高層的 HDL 即可的 stratification 方法。我們所憑藉的線路輸入訊號的轉換 (transition) 較多，則該線路能消耗功率就較高，而如果輸出 transition 較多的話，就表示有較多的 transition 從輸入端被傳到輸出端，也會導致較多的消耗。利用這個基礎，我們使用 *Regression* 的方法來計算每一個輸入訊號有 transition 的時候，可能會引發的 transition 相對值有多少。在這裏我們用相對值與 $Power(V_j)$ 和別的值比起來可能是大或小到分層的目的。

令對應於 input vector V_j 在線路輸出的 output vector 為 W_o 。令線路輸入的個數為 n_{in} ，輸出的個數為 n_{out} 。若有一函數定義

$$A(b' b'^{+1}, V_i) = [a_1, a_2, \dots, a_{n_{in}}] \quad (5)$$

$$a_j = \begin{cases} 1, & \text{when } v'_{i,j} = b' \text{ and } v'^{+1}_{i,j} = b'^{+1} \\ 0, & \text{otherwise} \end{cases}$$

也就是在 V_i 裏面若 input j 是從 b' 變成 b'^{+1} 的話, 那麼 a_j 就等於 1, 否則就等於 0。令

$$X_{in}(V_i) = [A(00, V_i) A(01, V_i) A(10, V_i)] \quad (6)$$

$$X_{out}(W_i) = [A(00, W_i) A(01, W_i) A(10, W_i)]$$

設已取出的 input vectors 有 m 個 ($m=n^*$), 且將這些被 sample 出來的 input vectors 重新編號為 $V_{s,i}$, 而它們對應的 output vectors 則編號為 $W_{s,i}$ 那麼, 令

$$X = \begin{bmatrix} X_{in}(V_{s-1}) & X_{out}(W_{s-1}) \\ X_{in}(V_{s-m}) & X_{out}(W_{s-m}) \end{bmatrix} \quad (7)$$

$$P = \begin{bmatrix} Power(V_{s-1}) \\ Power(V_{s-m}) \end{bmatrix}$$

利用 *Multiple Regression*[4], 我們可以求出滿足下式的 C

$$XC = \hat{P} \text{ s. t. } \text{MINIMUM}[RMS(P - \hat{P})] \quad (8)$$

由上式所解出的 C , 我們就可以利用下式來對每一個 input vector 都算出一個估計的 power 值。

$$Power_{est}(V_i, W_i) = [X_{in}(V_i) X_{out}(W_i)] C \quad (9)$$

利用這個估計值, 我們就可以指出那些 input vector 是可能會造成較大的功率消耗, 而那些是會有較小的功率消耗, 並加以分類以降低 level 1 population 的 sample variance。我們將這個方法稱為動態分層法 (*Adaptive Stratification*), 因為這個方法可以隨著每次有新的 input vector V_i 被 sample, 得到新的 $Power(V_i)$ 之後, 馬上再算出最新的 C 值, 並用來分層, 以隨時保持 sample variance 在最小的情況。我們將 *Adaptive Stratification* 和前半段所述的 *AMC* 實現在同一個程式裏面並稱它為 *AMCAS* (*Advanced Monte Carlo with Adaptive Stratification*)。並將它和其他的分層方法一起做比較, 結果如表四。

四、結論與討論

在表四裏面, 第一個方法是只有 *AMC* *Adaptive Stratification*, 由於它沒有 sample variance 的動作, 所以它的平均的 input vector 數目是下列幾個方法裏面第二個方法是用 *Hamming Distance* 來作依據, 這個方法在使用上是最簡單的, 於它只考慮輸入訊號是否有 transition, 了有些輸入訊號是 level-triggered 的, t 功率消耗對它們的邏輯值比對它們 transition 更敏感, 因此, 這個方法在有可能效果不錯, 但是卻不穩定。如它 C1908 和 C6288 的時候效果不錯, 但是的線路則沒有那麼有效。第三個方法是提出的 *AMCAS*, 我們的方法在各個線路有效率的對 input vector 做挑選, 除了 C C6288 之外, 其他的線路都比 *HDM* 改訂最後一欄是將所有的 input vector 都分然後再用這些 $Power(V_i)$ 來做分層。雖然法在實際上是可能使用的, 但是這比用 gate-level 的 power simulation 低 sample variance s^2 的方法, 也就這個方法的壓縮比, 可以拿來當做壓縮比參考值。

在本計劃裏面, 我們不但提出了一個可傳統的 Monte Carlo 方法裏的取樣 Normal 程度的方法, 另外我們還提出。需線路的 HDL 模型就可以動態對取樣層以降低 sample variance。由我們的證明我們提出的方法確實能夠在使用 confidence level 以及 error level 有效率地對 input sequence 做壓 designer 能夠利用本計劃所提出的方法地為他們的線路挑出一些具代表性的, 壓縮過後的 input vectors 來做功率估

五、參考文獻

- [1] R. Burch, F. N. Najm, P. Yang, and T. N. Trick, "A Monte Carlo Approach for Power Estimation," IEEE Trans. on VLSI, March 1993, pp63-71.
- [2] C. S. Ding, Q. Wu, C. T. Hsieh, M. Pedram, "Stratified Random Sampling for Power Estimation," Trans. on CAD, June 1998, pp465-478.
- [3] H. Stark, J. W. Woods, "Probability, Random Process, and Estimation Theory for Engineers," *Prentice-Hall* 1986
- [4] Lyman Ott, "An Introduction to Statistical Methods and Data Analysis," PWS-Kent Pub. Co., 1988.

- [5] Heng-Liang Huang, Wen-Zen Shen and Jing- "Advanced Monte Carlo with Adaptive Stratification Power Estimation," submitted to IEEE Trans. 2001.

表四：各式分層法的比較 ($\Gamma=0.95$)

Circuit	NO STRAT			HDM			AMCAS			REAL	
	final l	nVecs	viol r	final l	nVecs	viol r	final l	nVecs	viol r	final l	nVecs
C432	49.33	2995	0.0315	216.68	1287	0.0255	71.96	835	0.0172	39.69	35
C499	38.44	1096	0.0354	271.69	848	0.0002	64.28	495	0.0126	36.83	12
C880	42.32	1173	0.0332	367.00	1410	0.0078	79.23	667	0.0129	35.75	15
C1355	40.72	1193	0.0338	272.68	851	0.0005	72.24	519	0.0099	37.62	12
C1908	43.33	1705	0.0363	210.47	715	0.0042	100.24	723	0.0039	36.28	15
C3540	35.49	374	0.0331	305.93	1275	0.0112	38.22	308	0.0257	38.52	12
C6288	43.62	2554	0.0379	206.13	685	0.0028	124.83	905	0.0027	35.06	15
Avg	41.89	1584	0.0344	264.36	1010	0.0074	78.71	636	0.0121	37.10	17