# Concurrent Sizing, Vdd and V<sub>th</sub> Assignment for Low-Power Design*

Concurrent Sizing, Vdd and $V_{th}$ Assignment for Low-Power Design*

Ashish Srivastava     Dennis Sylvester     David Blaauw

University of Michigan, EECS Department, Ann Arbor, MI 48109, USA {ansrivas, dennis, blaauw}@eecs.umich.edu

**Abstract**

We present a sensitivity based algorithm for total power including dynamic and subthreshold leakage power minimization using simultaneous sizing, Vdd and Vth assignment. The proposed algorithm is implemented and tested on a set of combinational benchmark circuits. A comparison with traditional CVS based algorithms demonstrates the advantage of the algorithm including an average power reduction of 37% at primary input activities of 0.1. We also investigate the impact of various low Vdd values on total power savings.

## I. Introduction

Early implementations of dual-Vdd designs have showed very promising results with power savings on the order of 40-50% [1]. However, the authors of [2] claim that the power reduction achievable by dual-Vdd can be expected to decrease with reducing power supplies. More recently, [3] shows that using a second threshold voltage in conjunction with a second Vdd can be used to maintain the achievable power reduction with scaling process generations. It was also demonstrated that using more than two power supplies or threshold voltages provides minimal reduction as compared to that provided by two Vdd or Vth [2,3].

Using multiple power supplies in a design imposes the topological constraint that gates operating at a lower supply voltage cannot fan-out to gates operating at a higher supply voltage without the use of dedicated level converters. Two approaches that obey this constraint have been proposed in the literature. Clustered Voltage Scaling (CVS) [4] allows only one transition from high Vdd to low Vdd gates along a path, and level converts low Vdd signals to high Vdd at the flip-flops. Extended CVS (ECVS) allows for level conversion on paths in between flip-flops and thus can improve the achievable power reduction. Also, there has been a large amount of recent work in power optimization using dual Vth and sizing e.g. [5]. But existing work fails to consider the optimization of total power dissipation and are restricted to either dynamic or leakage power optimization. Hence we observe that there is a pressing need to integrate all three of these low-power design variables concurrently in an efficient algorithm. This paper is the first to perform simultaneous gate-level sizing, Vdd, and Vth assignment in a dual-Vdd/Vth environment to minimize total power consumption (defined as the sum of static and dynamic power). Since our algorithm enables simultaneous optimization of total power using Vdd and Vth allocation and sizing we refer to the complete algorithm as VVS.

## II. Algorithm Description

We propose a two stage sensitivity-based approach to minimize total power using dual Vdd, sizing and dual Vth. All the gates in the design are initially assumed to be operating at the higher supply and lower threshold voltage. Throughout the flow of the VVS algorithm a front is maintained located at the interface between the low and high Vdd gates. Similar to CVS we do not allow level conversion within the logic itself and hence, we must strictly observe this topological constraint. The timing constraints on the design remain fixed throughout the flow of the algorithm.

In the first stage of the VVS algorithm Vdd assignment and sizing are combined to minimize total power while we move the front from the primary outputs to the primary inputs. The second stage uses the optimal point found in the first stage as the starting point for the optimization and then relies on both Vdd and Vth assignment along with sizing to further reduce total power while the front is moved back to the primary outputs.

VVS is initialized by creating a list of primary outputs of the design that represents the front of the design. A predictive metric is then used to order gates in this list. This metric could be based on simple parameters such as the fanout capacitance or the slack of the gate for example. The gate with the maximum value for the predictive metric is selected as the candidate gate, which is then assigned to low Vdd if the timing constraints are not violated. Gates are identified that can be included in the backward front as a result of the assignment of the previous gate to low Vdd

At the end of CVS, none of the gates on the front can be assigned to low Vdd without violating the timing constraints. Gate sizing is then employed to compensate for the delay added during the assignment of a gate to low Vdd. A sensitivity measure to upsizing for all of the gates in the circuit is calculated which is used to identify gates to be up-sized. Let $\Delta D$ represent the change in delay and $\Delta P$ the change in power dissipation due to upsizing. The sensitivity of each gate to up-sizing is defined as

$$Sensitivity = \frac{1}{\Delta P}\sum_{arcs}\frac{\Delta D}{Slack_{arc} - S_{min} + K} \quad (1)$$

where Smin is the worst slack seen in the circuit and K is a small positive quantity. The form of the sensitivity measure gives a higher value to gates lying on the critical paths of the circuit. The arcs represent the falling and rising arcs associated with each of the inputs of the gate. The gate with the maximum sensitivity is then selected and sized up. This process is repeated until all slacks in the circuit become positive. The up-sizing required can result in an increase in total power and such moves in certain cases can be accepted if they allow us to move out of local minima.

At all points during the first stage the best-seen solution is saved and this solution is restored at the end of the first stage. The end of the first stage is signaled when the list containing the gates on the backward front becomes empty or else none of the gates in the list can be assigned to low Vdd without violating timing (even with the maximum allowed amount of upsizing).

We now define the front to consist of all gates that are operating at low Vdd and have all of their fanins operating at high Vdd. Importantly, assigning a gate on this front to operate at high Vdd will not lead to a violation of the topological constraint. We now calculate 1) a sensitivity measure for gates on the front with respect to high Vdd operation and 2) a sensitivity measure for all gates in the circuit with respect to upsizing.

Table 1: Power savings at various phases of the algorithm for activity factor of 0.1

| | | | | % Savings compared to initial design | | | | | | | | |
| | Initial Power (uW) | | | CVS only | | | CVS+Sizing | | | VVS | | |
| Circuit | Leakage | Switching | Total | Leakage | Switching | Total | Leakage | Switching | Total | Leakage | Switching | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c432 | 19.7 | 92.4 | 112.1 | 1.48 | 1.16 | 1.21 | -10.37 | 7.91 | 4.70 | 24.55 | 7.54 | 10.53 |
| c880 | 41.8 | 211.1 | 252.9 | 33.79 | 28.85 | 29.67 | 28.47 | 42.11 | 39.86 | 66.05 | 41.68 | 45.70 |
| c1908 | 54.2 | 182.2 | 236.4 | 12.42 | 10.03 | 10.57 | 12.42 | 10.03 | 10.57 | 74.91 | 7.48 | 22.93 |
| c2670 | 93.8 | 466.2 | 559.9 | 31.58 | 25.17 | 26.25 | 40.59 | 45.68 | 44.82 | 58.08 | 45.48 | 47.59 |
| c3540 | 106.8 | 423.3 | 530.0 | 5.38 | 3.59 | 3.95 | 10.25 | 31.79 | 27.45 | 36.39 | 31.06 | 32.13 |
| c5315 | 175.6 | 740.7 | 916.3 | 26.60 | 21.97 | 22.86 | 41.15 | 58.01 | 54.78 | 51.59 | 57.80 | 56.61 |
| c6288 | 325.7 | 400.8 | 726.5 | 1.90 | 1.81 | 1.85 | 1.90 | 1.81 | 1.85 | 81.37 | -8.02 | 32.06 |
| c7552 | 201.4 | 598.9 | 800.3 | 47.47 | 34.06 | 37.43 | 46.91 | 47.69 | 47.49 | 47.71 | 47.66 | 47.67 |
| Average | | | | 20.08 | 15.83 | 16.72 | 21.41 | 30.63 | 28.94 | 55.08 | 28.83 | 36.90 |

Both these sensitivities are calculated as the ratio of the change in delay to the change in power dissipation as a result of the corresponding operation. The gate with the maximum sensitivity is then either assigned to high Vdd or up-sized based on the operation to which the maximum sensitivity corresponds.

Once a gate is up-sized or reset to high Vdd operation, timing slack has been created in the circuit. To exploit this slack and reduce total power, the next step begins by computing the sensitivity of all gates in the circuit with respect to operation at high Vth. This sensitivity is calculated as the ratio of the change in power to change in delay in order to identify gates that provide the maximum decrease in power for the minimum increase in delay. Based on this sensitivity measure gates are assigned to high Vth as long as the timing constraints of the design are met. This set of moves (assignment to high Vdd or upsizing a gate followed by the associated high Vth assignments) is then accepted if the total power is found to decrease otherwise the set of moves is reversed.

This two-stage VVS algorithm allows us to make intelligent choices to trade-off dynamic power for leakage power in order to obtain a reduction in the total power dissipation. It effectively directs the algorithm to automatically provide either more leakage or dynamic power reduction based on the initial design point.

### III. Results

The algorithm described in Section II was implemented in C and tested on ISCAS85 benchmark circuits that vary in size from 169 to 2500 gates [6]. The circuits were synthesized using an industrial 0.13μm library with a nominal Vdd of 1.2V and a nominal Vth of ±0.23V (these are fixed throughout) which represent the high Vdd and high Vth respectively. The standard cells in the library are also characterized at various design points including low Vdd = {0.6, 0.7, 0.8} V and low Vth ={0.14, 0.12, 0.1,0.08} V. We also created duplicate low Vdd libraries in which gate delays are computed with inputs toggling at high Vdd rather than low Vdd. All energies (static, short-circuit, and dynamic) and capacitance variations due to varying thresholds [5] are inherently considered using these SPICE-derived library files.

The synthesized design is first sized using a TILOS-like [7] sensitivity based sizing algorithm to obtain the power-delay curve for the design. The design is then resized from the initial synthesized point to a delay point that is backed off from the minimum achievable delay by 20%, which still maintains an aggressive delay since the initial design is synthesized using the fastest combination of Vdd and Vth. Subsequent phases of the algorithm maintain this timing and

no further relaxation in timing is used to obtain power improvement.

Table 1 shows the results obtained for the ISCAS benchmark circuits, with an activity factor of 0.1. The columns corresponding to the initial power list the actual power numbers. The remaining columns show the % reduction in leakage, switching and total power at the end of three distinct phases of the algorithm; 1) CVS only, 2) CVS and sizing only, and 3) VVS. The results clearly show the advantage offered by each of the steps of the algorithm. CVS coupled with sizing increases the average savings in switching power by a factor of 2 from approximately 15% to 30%. The leakage power also shows a significant reduction of ~20% which can be attributed to the roughly cubic dependence of leakage power on Vdd [8]. The last phase shows that a small amount of switching power (an average of 1.8% of the initial switching power) can be traded off to obtain substantial savings (~33%) in leakage power due to the exponential dependence of leakage current on Vth.

Fig. 2 shows the variation in power savings when using different values for the low Vdd and Vth. It is important to note that the same design at different low Vth's are operating at different frequencies and hence the power savings are relative to different initial design points. Thus a comparison of the power savings between various low Vth's is not justified. The figure clearly shows that a low Vdd of 0.6V provides an increase in power reduction of approximately 10% compared to a low Vdd of 0.8V. This is expected on the basis of the rules of thumb proposed in [2] which show that the optimal low Vdd is typically about half of the high Vdd in a dual-Vth environment.
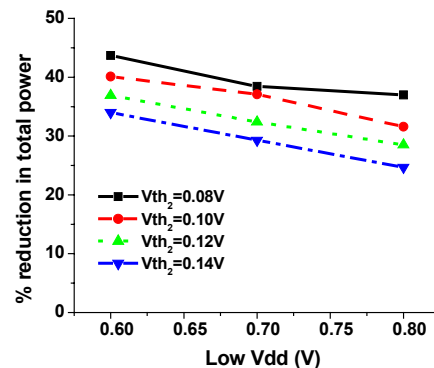


Fig. 2 Dependence of average power savings on low Vdd

### References
[1] K. Usami, *et al., IEEE JSSC*, pp. 1772-1780, March 1998.
[2] M. Hamada, *et al., Proc. CICC*, pp.89-92, 2001.
[3] A. Srivastava, *et al., Proc. ASPDAC,* pp.400-403, 2003.
[4] K. Usami, *et al., Proc. ISLPED*, pp.3-8, 1995.
[5] S. Sirichotiyakul, *et al., Proc. DAC*, pp.436-441, 1999.
[6] F. Brglez, *et al., Proc. ISCAS*, pp.695-698, 1985.
[7] J. Fishburn, *et al., Proc. ICCAD*, pp.326-328, 1985.
[8] R. Krishnamurthy, *et al., Proc. CICC,* pp.125-128, 2002.