Dual Voltage Design for Minimum Energy Using Gate Slack

Kyungseok Kim and Vishwani D. Agrawal
Department of ECE, Auburn University, Auburn, AL 36849, USA
kyungkim@auburn.edu, vagrawal@eng.auburn.edu

***Abstract*—This paper presents a new slack-time based algorithm for dual V_{dd} design to achieve maximum energy saving. Although a global optimum is sought computation time is kept low. The slack of a gate is defined as the difference between the critical path delay for the circuit and the delay of the longest path through that gate. A linear-time algorithm is used for computing slacks for all gates of the circuit. Positive non-zero slack gates are classified into two groups, one in which all gates can be unconditionally assigned low voltage and the other where only a selected subset can be assigned low voltage without violating the positive non-zero slack requirement. Multiple voltage boundaries are given special consideration. The overall complexity of this power optimization algorithm is linear in number of gates as compared to a previously published exponential-time exact algorithm using mixed integer linear program (MILP). We apply the new algorithm to optimize ISCAS'85 benchmark circuits and compare the results with those from MILP. We avoid the use of level converters at multiple voltage boundaries. Energy savings from the new slack-time based algorithm is very closed to those from MILP. For c880, the energy saving is 22% for subthreshold voltage operation and 50% for nominal operation in PTM CMOS 90nm. For c2670 nominal voltage design, time of dual voltage optimization is reduced 43X compared to the MILP method. This new algorithm is beneficial for large circuits with many large positive slack paths that would require enormous time for optimization by the MILP approach.**

I. Introduction

Scaling power supply voltage (V_{dd}) is widely used as a technique for reducing both dynamic power and static leakage power. This degrades the performance. To scale the power supply voltage of a circuit without degrading the performance, a dual V_{dd} technique exploits time slacks in the circuit by assigning a lower supply voltage (V_{DDL}) to gates on non-critical paths and a higher supply voltage (V_{DDH}=V_{dd}) to gates on critical paths [14], [2], [11], [15], [16]. For maximizing power saving, an optimal lower supply voltage is determined such that it can be assigned to as many positive slack gates as possible without exceeding the critical path delay. However, signal levels at multiple voltage boundaries require special attention.

The literature gives two main algorithms for dual V_{dd} design, namely, *clustered voltage scaling* (CVS) [15] and *extended clustered voltage scaling* (ECVS) [16]. To insure correct logic function of gates across multiple voltage boundaries without a level converter (LC), the CVS algorithm does not allow the V_{DDL} cells to feed directly into V_{DDH} cells and level converting is only implemented within special level conversion flip-flops (LCFF). This topological constraint limits the use of the full slack available to reduce power consumption. The ECVS algorithm eliminates this constraint using a level converter that is inserted when a V_{DDL} cell drives a V_{DDH} cell. More power saving is achievable than CVS algorithm. Both algorithms use the backward heuristic traversing the circuit from primary outputs to primary inputs in level order. Two heuristic algorithms have theoretical runtime complexity $O(n^2)$, where n is total number of gates in a circuit [4].

Most researches have focused on improving power saving implementing their own greedy algorithms [2], [3], [11]. These are still heuristic approaches and provide a suboptimal solution for dual V_{dd} assignment. Mixed integer linear programs (MILP) [6] are widely used to optimize a circuit for minimizing power or energy consumption using sizing, multiple V_{dd}, multiple threshold voltage (V_{th}) and combinations of those [5], [9], [10], [13]. MILP searches global optimal solution for an objective function, minimize power, considering the entire design space. Thus, it may take huge time to optimize large circuits used in modern VLSI systems. Time complexity of MILP optimization may not be acceptable in practice.

For dual V_{dd} design, we need to find the optimal V_{DDL} and its assignments to positive slack gates in a circuit for minimum power. If we can quickly find all positive slack gates that can be assigned to V_{DDL}, it reduces much optimization work of dual V_{dd} design and saves computation time. We propose a new slack-time based algorithm to save computation time and obtain a nearly global solution similar to that obtained by an MILP. The new technique is highly efficient and gives a quality of solution very close to the MILP.

The paper is organized as follows. Section II gives recent background on MILP models [10] for comparison purpose with a proposed algorithm. In Section III, we propose the new slack-time based algorithm for dual V_{dd} design. Section IV reports SPICE simulation results of our algorithm and MILP to validate results. Finally, a conclusion of this work is given in Section V.

978-1-4244-9066-0/11/$26.00 ©2011 IEEE

II. MILP FOR OPTIMAL V_{DDL} AND ASSIGNMENTS OF DUAL V_{dd}

There are two ways to find the optimal lower supply voltage V_{DDL} and its assignments for dual voltage design in the literature. First, the optimal V_{DDL} is simultaneously searched by applying V_{DDL} assignment algorithm to a circuit with different V_{DDL} values, then it selects a pair of the optimal V_{DDL} and its assignments for minimum power consumption [3], [15], [16]. Otherwise, theoretical path delay model to power supply voltage V_{dd} is developed to determine the optimal V_{DDL} for maximum power saving, then V_{DDL} assignments are executed to achieve lowest power consumption considering multiple voltage boundaries [7], [11]. Most of dual V_{dd} techniques are based on heuristic greedy algorithms and applied to nominal operating circuits for lowering power consumption.

For energy constrained applications, dual V_{dd} technique is applied to a subthreshold logic circuit for further reducing minimum energy operating point [10]. In this paper, MILP models similar to CVS are formulated to find the best optimal V_{DDL} and its assignments for dual V_{dd} design. This global optimum algorithm is applicable to a circuit operating in both subthreshold and nominal supply voltage, but is needed to multiple runs considering all available V_{DDL} to given V_{DDH}. Now, we extends this MILP models to select automatically the optimal V_{DDL} and its assignments by introducing new variables for one-time run. We briefly explain new variables and parameters here before presenting MILP models.

- $X_{i,v}$: supply voltage assignment integer variable that is 1 for gate i with power supply voltage v.
- V_v: supply voltage integer variable that is 1 for two selected V_{DDH} and V_{DDL} in available power supply voltage v.
- $td_{i,v}$: gate delay for gate i with supply voltage v.
- $V_{dd,v}$: power supply voltage value for v.
- G_{tot}: total number of gates in a circuit.

MILP models are reformulated from [10]:

$$\text{Minimize} \sum_i \sum_v E_{tot,i,v} \cdot X_{i,v}$$

$\forall i \in$ all gates and $\forall v \in$ power supply voltage domain V (1)

$$E_{tot,i,v} = \alpha_i \cdot C_{L,i,v} \cdot V_{dd,v}^2 + P_{leak,i,v} \cdot T_c \quad (2)$$

Subject to timing constraints:

$$T_i \geq T_j + t_{d,i,v} \cdot X_{i,v} \quad \forall j \in \text{all fanin gates of gate } i \quad (3)$$

$$T_i \leq T_c \quad \forall i \in \text{all primary output gates} \quad (4)$$

Subject to topological constraints:

$$\sum_{v \in V} V_{dd,v} \cdot X_{i,v} \leq \sum_{v \in V} V_{dd,v} \cdot X_{j,v} \quad \forall j \in \text{all fanin gates of gate } i \quad (5)$$

Subject to dual supply voltages selection:

$$\sum_{v \in V} V_v = 2 \quad (6)$$

$$V_{V_{DDH}} = 1 \quad (7)$$

$$\sum_{v \in V} X_{i,v} = 1 \quad \forall i \in \text{all gates} \quad (8)$$

$$\sum_i X_{i,v} \leq G_{tot} \cdot V_v \quad \forall i \in \text{all gates}, \ \forall v \in V \quad (9)$$

Main difference of MILP models from [10] is dual V_{dd} selection conditions. T_c is critical path delay and given by the performance requirement, thus V_{DDH} is selected from (7) in power supply domain V. Using a bin-packing technique [1] all gates must be assigned to one of power supply voltage in V from (8) and (9).

MILP always guarantees that a dual V_{dd} circuit with the optimal V_{DDL} and it assignments achieve minimum energy consumption at same performance. We use absolute optimal results of MILP as a reference to check accuracy of our slack-time based algorithm that is presented in next section.

III. NEW SLACK-TIME BASED ALGORITHM FOR DUAL V_{dd} DESIGN

In this section, we propose a new slack-time based algorithm that finds the optimal V_{DDL} and its assignments for dual V_{dd} design as much close as global optimization from MILP solution for minimum energy consumption.

First, Our algorithm generates slack time distribution of a given circuit. We have developed an expanded version of static timing analysis (STA) [8]. For the output of gate i, let $T_{PI}(i)$ be the longest time for an event to arrive from a PI and $T_{PO}(i)$ be the longest time for an event to reach a PO. The delay of the longest path through gate i is given by,

$$D_{p,i} = T_{PI}(i) + T_{PO}(i) \quad (10)$$

The critical path delay for the circuit is,

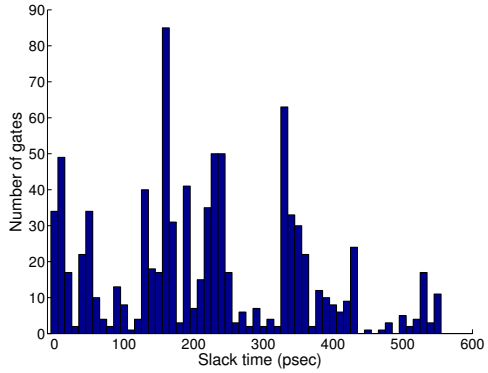
$$T_c = \text{Max}\{D_{p,j}\} \quad \forall \text{ gate } j \quad (11)$$

Slack time for gate i is found as follows:

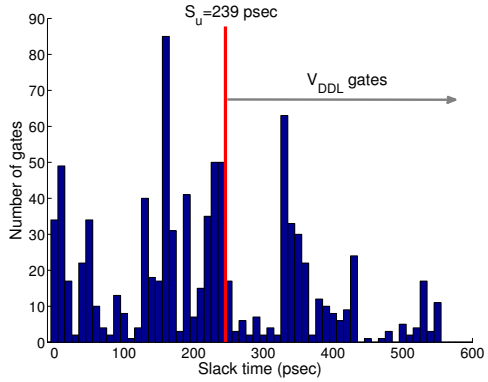
$$S_i = T_c - D_{p,i} \quad (12)$$

The time for calculating slack time for all gates of a circuit is $O(n)$, where n is total number of gates. Figure 1(a) shows the slack time distribution for c2670 in ISCAS'85 benchmark circuits in PTM 90nm CMOS technology [18].

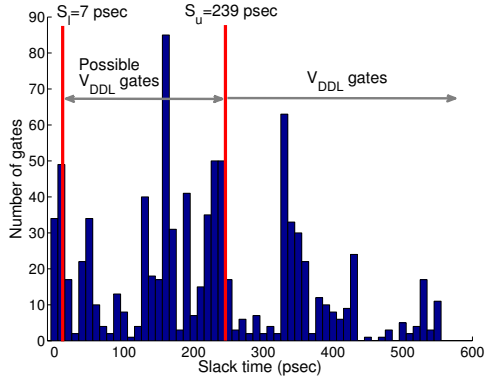
To quickly identify the possible V_{DDL} gates on non-critical paths, we introduce an upper slack time (S_u) that guarantees that any gate with slack time larger than S_u will be free from timing violation, i.e., *negative slack*, irrespective of the voltage assignment for other gates. The slack time of a V_{DDH} gate that is equal to S_u becomes zero after assigning V_{DDL} to all gates on the longest path through it. We find S_u using (12).



(a) Slack time distribution for a single nominal $V_{dd} = 1.2V$.



(b) Upper slack time S_u at V_{DDH}=1.2V and $V_{DDL} = 0.69V$.



(c) Lower slack time S_l at $V_{DDH} = 1.2V$ and $V_{DDL} = 0.69V$.

Fig. 1. Procedure of slack-time based algorithm for c2670 in ISCAS'85 benchmark circuits in PTM 90nm CMOS.

Let S'_i be the slack time of gate i after assigning V_{DDL} to all gates on the longest path through it. Now, $D'_{p,i}$ is the longest path delay through the gate i.

$$\begin{aligned} S'_i &= T_c - D'_{p,i} \\ &= T_c - \beta \cdot D_{p,i} \\ &= T_c - \beta \cdot (T_c - S_i) \end{aligned} \tag{13}$$

Where β is the ratio of $D_{p,i}$ to $D'_{p,i}$. It is approximated by

$$\beta = \frac{D'_{p,i}}{D_{p,i}} \approx \frac{T'_c}{T_c} \tag{14}$$

T'_c is the critical path delay when V_{DDL} is supplied to the entire circuit. It is determined by the static timing analysis in the same way as [8]. By substituting S_u for S_i in (13), S'_i become zero. Thus, S_u is obtained as:

$$S_u = \frac{\beta - 1}{\beta} \cdot T_c \tag{15}$$

In Figure 1(b), any gate that has a positive slack time larger than S_u, i.e., in the range covered by the right arrow, is safely assigned to V_{DDL} without timing violation. S_u serves as a slack threshold. Any gate with slack above this threshold is unconditionally assigned to V_{DDL} irrespective of voltages of other gates on paths passing through it.

The slack time of all gates on critical paths is zero. Hence, there is no room to assign V_{DDL} to those gates. But, if there is a gate with a positive slack time that is close to zero, it may be possible to assign V_{DDL} provided other gates on paths through it remain with V_{DDH}, such that no path delay exceeds T_c.

Let t_d be a gate delay in the circuit. After assigning V_{DDL}, t_d is increased. Suppose, it becomes t'_d. The amount $t'_d - t_d$ is the increase in path delay through the gate. This is also the reduction in the slack of other gates on paths through the V_{DDL} gate. Therefore, a gate that has the slack time larger than $t'_d - t_d$ can be assigned to V_{DDL}. Let us call this slack time the lower slack time (S_l). Because each logic gate has the different value of $t'_d - t_d$, the minimum value of $t'_d - t_d$ is used to define S_l.

$$\begin{aligned} S_l &= \text{Min}\left[(t'_d - t_d)_{\text{gates } j}\right] \\ &= \text{Min}\left[(\beta - 1) \cdot t_{d \text{ gates } j}\right] \quad \forall j \in \text{all gates} \\ &\text{assume } \frac{t'_{d,j}}{t_{d,j}} \approx \frac{D'_{p,j}}{D_{p,j}} = \beta \end{aligned} \tag{16}$$

For simplicity, we assume that path delay is proportional to the delay of a gate on it. Timing violation from this assumption is checked later when V_{DDL} gates are chosen, finally.

As shown in Figure 1(c), S_l can be used to search possible V_{DDL} gates between S_l and S_u, the range shown by a double arrow. The gates with positive slack time less than S_l are unconditionally assigned to V_{DDH} and are located near or on critical paths.

Until now, we have demonstrated how to select gates that can be assigned to V_{DDL} using simple two slack times, S_u and S_l. A gate with slack time larger than S_u is assigned to V_{DDL}, while a gate with slack time less than S_l is assigned to V_{DDH}. For a gate with slack time between S_l and S_u, we need to carefully select the power supply voltage. V_{DDL} assignment for these gates affects the assignment of other gates on paths if we have to hold the path delay within T_c. The order of V_{DDL} assignment to these gates affects the energy saving of the dual

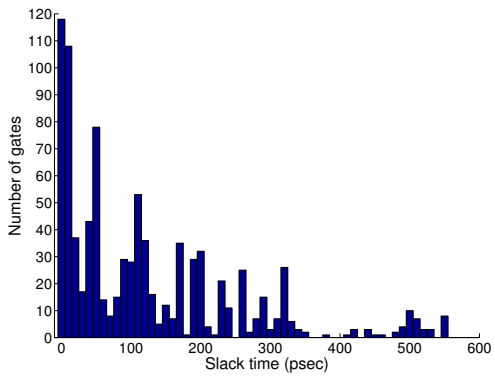


Fig. 2. Slack time distribution of an optimized c2670 with $V_{DDH} = 1.2V$ and $V_{DDL} = 0.69V$.

V_{dd} design, when we consider multiple voltage boundaries. Thus, we need to use a greedy approach depending on the type of dual V_{dd} design. If we allow V_{DDL} gates to drive V_{DDH} gates like ECVS, the selection order should minimize the use of level converters to maximize energy saving. Because CVS does not use level converters, there exists a topological constraints that a V_{DDL} gate can not drive a V_{DDH} gate. Therefore, the selection order is chosen to maximize V_{DDL} assignment to gates with this topological constraint.

In this paper, we use the slack time distribution to implement dual V_{dd} algorithm like CVS. The result of the algorithm is compared to MILP solution in terms of energy saving and run-time. To maximize V_{DDL} assignment with topological constraints, first, higher logic depth gates between S_l and S_u should be assigned to V_{DDL}. This priority reflects the fact that V_{DDL} gates do not feed into V_{DDH} gates directly. The timing violation should be checked when a gate between S_l and S_u is assigned to V_{DDL}. We find all V_{DDL} gates, which do not violate the critical path timing constraint T_c. Additionally, checking topological constraints for these V_{DDL} gates, we ascertain that all V_{DDL} gates satisfy both timing and topological constraints.

The final stage of the algorithm searches for the optimal V_{DDL} value to give maximum energy saving. We already know all V_{DDL} gates for each available V_{DDL} value from previous procedures. Thus, we simply calculate the energy saving from V_{DDL} gates, then select the optimal V_{DDL} to meet best energy saving. Figure 2 shows that the slack time distribution of an optimized c2670 circuit that has the optimal $V_{DDL} = 0.69$V from our algorithm. In next section, we show the results of optimization from the slack-time based algorithm for ISCAS'85 benchmark circuits, which operate in either subthreshold or nominal supply voltage.

IV. RESULTS

As example circuits, ISCAS'85 benchmark circuits are synthesized with four types of basic standard cells, namely, INV, NAND2, NAND3, and NOR2. Average activity of a synthesized circuit is found from logic simulation with randomly generated input vectors. We extract gate delay, capacitance and leakage power of basic standard cells through SPICE simulation by varying power supply voltage from 0.1V to 1.2V in 10mV steps. All SPICE simulations were run for room temperature (300K) using PTM 90nm CMOS process, where CMOS device threshold voltages are $V_{th,pmos} = 0.21$V and $V_{th,nmos} = 0.29$V at nominal $V_{dd} = 1.2$V.

For comparing the algorithm of Section III with MILP of Section II, we measure the energy consumption of benchmark circuits using SPICE simulation for a single V_{dd} as a reference. Random input vectors for each circuit in SPICE simulation are the same as those used in logic simulation to measure the average activity. To find the optimal V_{DDL} and its assignments for maximum energy saving, MILP algorithm is applied for a synthesized circuit. With MILP solution, the SPICE netlist of an optimized circuit is generated, where each gate has its voltage assignment either as the given V_{DDH} or an optimal V_{DDL}. SPICE simulation runs with this netlist to measure energy consumption of the optimized dual V_{dd} circuit. Same procedure is repeated for the design obtained by the slack-time based algorithm.

First, we apply both algorithms to benchmark circuits operating in subthreshold region. We assume that V_{DDH} at minimum energy is given with the corresponding speed for each benchmark circuit. Table I shows SPICE simulation results from two algorithms. The results of two algorithm exactly match each other, thus we present it once here. Using dual V_{dd} design, total energy saving for c880 (8-bit ALU) is 22.2% as the best case. In Table II, both algorithm are applied to optimize benchmark circuits operating with the nominal supply voltage. We set 1.2V as a nominal power supply voltage for PTM 90nm CMOS by referring to the industry standard 90nm CMOS technology. The results from from the two algorithms do not match for c880 and c6288, but energy savings are very close. Evidently, the result of the slack-time based algorithm is very close to the global optimization, even though it uses a greedy heuristic to select the best V_{DDL} gates from all V_{DDL} gates that pass timing constraint. For c880, energy savings from MILP and our algorithm are 51.0% and 50.8%, respectively. Compared to the energy savings in the subthreshold region, these energy savings are much larger. This is because lower supply voltage increases the gate delay exponentially in the subthreshold region, while the gate delay increase for the nominal voltage operation is polynomial according to the alpha-power law model [12], [17]. It means that positive slack of gates in a circuit is reduced quicker by assigning V_{DDL} in subthreshold region. Thus, we obtain an optimal V_{DDL} that is closer to V_{DDH} and there are fewer V_{DDL} gates as well. Figure 3 shows slack time distributions before and after optimization by our algorithm

TABLE I

ENERGY SAVING AND OPTIMAL V_{DDL} FROM MILP [10] OR SLACK-TIME BASED ALGORITHM FOR GIVEN V_{DDH} IN ISCAS'85 BENCHMARK CIRCUITS IN SUBTHRESHOLD REGION IN PTM 90NM CMOS. BOTH ALGORITHMS PRODUCED IDENTICAL RESULT.

Benchmark circuit	Total gates	Activity α	V_{DDH} (V)	V_{DDL} (V)	V_{DDL} gates (%)	E_{single} (fJ)	E_{dual} (fJ)	$E_{reduc.}$ (%)	Freq. (MHz)	MILP CPU time(s)*	Slack CPU time(s)*
c432	154	0.19	0.25	0.23	5.2	7.9	7.8	1.1	14.4	0.3	2.5
c499	493	0.21	0.22	0.18	9.7	20.2	19.8	2.0	11.9	0.3	19.2
c880	360	0.18	0.24	0.18	46.4	14.4	11.2	22.2	13.6	5.8	17.9
c1355	469	0.21	0.21	0.18	10.2	19.5	19.0	2.5	9.8	0.2	13.3
c1908	584	0.20	0.24	0.21	24.3	26.5	25.0	5.8	11.8	3.2	47.6
c2670	901	0.16	0.25	0.21	46.4	32.8	28.0	14.8	17.4	35.9	134.4
c3540	1270	0.33	0.23	0.14	7.0	88.0	84.6	3.8	7.2	3.2	256.5
c5315	2077	0.26	0.24	0.19	47.1	116.8	98.0	16.1	9.8	852.3	692.0
c6288	2407	0.28	0.29	0.18	2.7	165.4	162.0	2.1	9.4	2.6	1293.7
c7552	2823	0.20	0.25	0.21	42.3	131.7	117.1	11.1	13.6	1452.2	1408.3
Average					24.1			8.2			

*Intel Core 2 Duo 3.06GHz, 4GB RAM.

TABLE II

ENERGY SAVING AND OPTIMAL V_{DDL} FROM MILP [10] AND SLACK-TIME BASED ALGORITHM FOR ISCAS'85 BENCHMARK CIRCUIT OPERATING IN NOMINAL V_{dd} IN PTM 90NM CMOS.

	Single V_{dd}			Dual V_{dd}									
				MILP					Slack-time based algorithm				
Benchmark circuit	V_{DDH} (V)	E_{single} (fJ)	Freq. (GHz)	V_{DDL} (V)	V_{DDL} gate (%)	E_{dual} (fJ)	$E_{reduc.}$ (%)	CPU (s)*	V_{DDL} (V)	V_{DDL} gate(%)	E_{dual} (fJ)	$E_{reduc.}$ (%)	CPU (s)*
c432	1.20	160.1	1.7	0.75	5.2	153.9	3.9	0.6	0.75	5.2	153.9	3.9	15.8
c499	1.20	460.6	2.3	0.79	19.5	433.4	5.9	403.8	0.79	19.5	433.4	5.9	194.4
c880	1.20	277.6	2.0	0.59	56.9	136.1	51.0	455.0	0.60	57.5	136.6	50.8	62.1
c1355	1.20	453.0	2.3	0.69	13.6	433.6	4.3	340.2	0.69	13.6	433.6	4.3	132.0
c1908	1.20	496.5	1.5	0.67	26.9	402.4	19.0	2146.9	0.67	26.9	402.4	19.0	247.8
c2670	1.20	647.6	1.8	0.69	57.9	337.9	47.8	20848.9	0.69	57.9	337.9	47.8	480.7
c3540	1.20	1844.0	1.1	0.70	11.6	1667.0	9.6	601.0	0.70	11.6	1667.0	9.6	1243.5
c6288	1.20	3066.0	0.5	1.18	53.1	2976.0	2.9	10523.7	0.47	2.9	2985.0	2.6	6128.0
Average					30.6		18.0			24.4		18.0	

*Intel Core 2 Duo 3.06GHz, 4GB RAM.

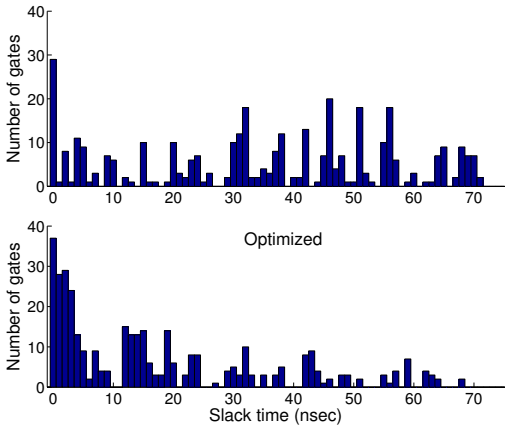
applied to c880 for both subthreshold and nominal voltage operations.

We measure the run-time of two algorithms based on CPU time in seconds. Our algorithm is written in the Perl script language. Thus, it has inherently slower execution than a program in the C language. The run time for MILP depends on the number of integer variables, the complexity of inequalities that specify the linear constraints, and the size of optimization space. From Table I, MILP is mostly faster than our algorithm except for c5315 and c7552. Both circuits have large slacks and have larger optimization spaces to be searched. Also, available V_{DDL} as an integer variable in MILP is limited by minimum operating voltage that guarantees correct logic function for the lower supply voltage. It is 0.1V below which the circuit function fails. This limitation reduces the size of the optimization space for MILP algorithm.

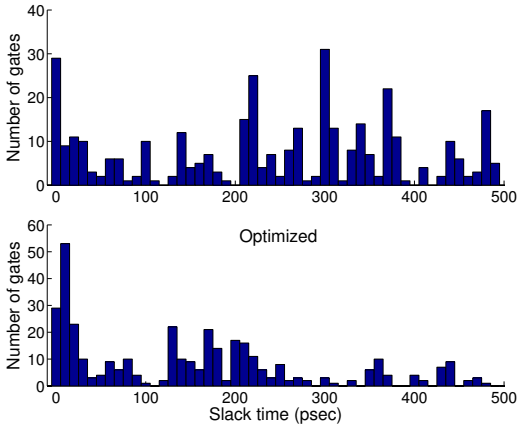
In Table II, the run time of our algorithm is ∼43X faster than MILP for c2670, because a larger range for V_{DDL} in nominal operation needs to be searched by MILP. For available V_{DDL} from power supply domain, our algorithm has linear time complexity $O(n)$ for finding the best energy saving by reducing time of searching for V_{DDL} gates using the thresholds S_l and S_u. While MILP searches recursively V_{DDL} gates from all gates inside the circuit to obtain the best energy saving. Thus, MILP displays an exponential time complexity for some benchmark circuits. Therefore, we can use our algorithm to optimize large circuits for dual V_{dd} design within reasonable time instead of using the exponential complexity MILP.

V. CONCLUSION

This paper presents a new slack-time based algorithm for dual V_{dd} design. Emphasis is on saving computation time and effort for maximizing energy saving in a given circuit. In a dual V_{dd} design, the given performance for a circuit determines the higher supply voltage V_{DDH}. The method of selecting a lower supply voltage V_{DDL} and the use of positive slack gates are the main ideas presented in this paper. The proposed algorithm classifies all positive slack gates into V_{DDH}, possible V_{DDL}, and V_{DDL} groups, respectively, based on the slack time of gates. After classification, the algorithm only investigates the "possible V_{DDL} gates" for available V_{DDL} considering multiple voltage boundaries in the energy optimization procedure. This reduces complexity of energy optimization process and the computation time remains tolerable for large circuits compared to the other available



(a) Subthreshold: $V_{DDH} = 0.24V$ and $V_{DDL} = 0.18V$.



(b) Nominal: $V_{DDH} = 1.2V$ and $V_{DDL} = 0.60V$.

Fig. 3. Slack time distribution before and after optimization of slack-time based algorithm for c880.

MILP methods. SPICE simulation for ISCAS'85 benchmark circuits shows energy savings up to 22.2% in subthreshold operation and 50.8% in nominal operation, which are the same as were obtained by the higher-complexity MILP method [10]. Computation time is reduced up to 43X compared to MILP. Our proposed algorithm has linear time complexity of $O(n)$ with n being the number of gates in the circuit. This novel slack-time based algorithm is useful because the MILP method is limited by its exponential run time cost.

ACKNOWLEDGMENT

This research has been supported in part by the Wireless Engineering Research and Education Center at Auburn University.

REFERENCES

[1] M. Anis, S. Areibi, M. Mahmoud, and M. Elmasry, "Dynamic and Leakage Power Reduction in MTCMOS Circuits Using an Automated Efficient Gate Clustering Technique," in *Proc. 39th Design Automation Conference*, 2002, pp. 480–485.

[2] C. Chen, A. Srivastava, and M. Sarrafzadeh, "On Gate Level Power Optimization Using Dual-Supply Voltages," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, vol. 9, no. 5, pp. 616–629, 2001.

[3] J. C. Chi, H. H. Lee, S. H. Tsai, and M. C. Chi, "Gate Level Multiple Supply Voltage Assignment Algorithm for Power Optimization Under Timing Constraint," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 6, pp. 637–648, 2007.

[4] D. Chinnery and K. Keutzer, *Closing the Power Gap Between ASIC & Custom: Tools and Techniques for Low Power Design*. Springer, 2007.

[5] D. G. Chinnery and K. Keutzer, "Linear Programming for Sizing, Vth and Vdd Assignment," in *Proc. International Symposium on Low power Electronics and Design*, 2005, pp. 149–154.

[6] R. Fourer, D. M. Gay, and B. W. Kernighan, *AMPL: A Mathematical Programming Language*. Brooks/Cole-Thomson Learning, 2003.

[7] M. Hamada, Y. Ootaguro, and T. Kuroda, "Utilizing Surplus Timing for Power Reduction," in *Proc. IEEE Custom Integrated Circuits Conf.*, 2001, pp. 89–92.

[8] R. B. Hitchcock, Sr., "Timing Verification and the Timing Analysis Program," in *Proc. 19th Design Automation Conference*, 1982, pp. 594–604.

[9] K. Kim and V. D. Agrawal, "Minimum Energy CMOS Design with Dual Subthreshold Supply and Multiple Logic-Level Gates," in *12th International Symposium on Quality Electronic Design*, Mar. 2011.

[10] K. Kim and V. D. Agrawal, "True Minimum Energy Design Using Dual Below-Threshold Supply Voltages," in *Proceedings of 24th International Conference on VLSI Design*, Jan. 2011.

[11] S. H. Kulkarni, A. N. Srivastava, and D. Sylvester, "A New Algorithm for Improved VDD Assignment in Low Power Dual VDD Systems," in *Proc. International Symp. Low Power Electronics and Design*, 2004, pp. 200–205.

[12] T. Sakurai and A. Newton, "Alpha-Power Law MOSFET Model and Its Applications to CMOS Inverter Delay and Other Formulas," *IEEE Journal of Solid-State Circuits*, vol. 25, no. 2, pp. 584–594, Apr. 1990.

[13] A. Srivastava, D. Sylvester, and D. Blaauw, "Power Minimization using Simultaneous Gate Sizing, Dual-Vdd and Dual-Vth Assignment," in *Proc. 41st Design Automation Conference*, 2004, pp. 783–787.

[14] V. Sundararajan and K. K. Parhi, "Synthesis of Low Power CMOS VLSI Circuits Using Dual Supply Voltages," in *Proc. 36th Design Automation Conference*, 1999, pp. 72–75.

[15] K. Usami and M. Horowitz, "Clustered Voltage Scaling Technique for Low-Power Design," in *Proc. International Symposium on Low Power Design*, 1995, pp. 3–8.

[16] K. Usami, M. Igarashi, F. Minami, T. Ishikawa, M. Kanzawa, M. Ichida, and K. Nogami, "Automated Low-Power Technique Exploiting Multiple Supply Voltages Applied to a Media Processor," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 3, pp. 463–472, 1998.

[17] A. Wang, B. H. Calhoun, and A. P. Chandrakasan, *Sub-Threshold Design for Ultra Low-Power Systems*. Springer, 2006.

[18] W. Zhao and Y. Cao, "New Generation of Predictive Technology Model for Sub-45 nm Early Design Exploration," *IEEE Trans. Electron Devices*, vol. 53, no. 11, pp. 2816–2823, Nov. 2006.