Fixed-Phase Retiming for Low Power Design

Kumar N. Lalgudi Marios C. Papaefthymiou

Department of Electrical Engineering
Yale University
New Haven, CT 06520

Abstract— In this paper we introduce fixed-phase retiming, an optimization technique for reducing the power dissipation of digital circuits without sacrificing their performance. In fixed-phase retiming, we first transform any given edge-triggered circuit into a two-phase level-clocked circuit by replacing each flip-flop by two level-sensitive latches. Subsequently, while keeping the latches clocked on one of the phases fixed, we relocate the remaining latches onto interconnections with high glitching activity and capacitive load. We formulate fixed-phase retiming as a boolean monotonic linear program and give an $O(V^6 \log V)$-time algorithm for solving it, where V is the number of combinational blocks in the circuit.

1 Introduction

The average power dissipation of a circuit may be significantly reduced by changes in its architecture. This paper describes *fixed-phase retiming*, an optimization technique that relocates the storage elements of digital CMOS circuits in order to reduce their power dissipation while maintaining their performance. In fixed-phase retiming, a given edge-triggered circuit is first transformed into a two-phase level-clocked circuit by replacing each edge-triggered flip-flop by two back-to-back level-clocked latches. Subsequently, latches clocked on one phase are relocated, while latches clocked on the other phase are kept fixed (hence the name fixed-phase). One objective of this transformation is to place latches on interconnections with high glitching activity, thereby shielding the glitches from large capacitive loads. Since in standard cell design the capacitance of a latch is typically smaller than the input capacitance of a combinational gate, this transformation reduces power dissipation during the opaque phase of the latch. Another objective of fixed-phase retiming is to reduce the number of latches in the circuit, thus reducing the power dissipated for storing data.

Fixed-phase retiming has several advantages over conventional edge-triggered retiming as described in [3]. First, since the latches clocked on one phase are kept fixed, the values of the state variables of the synchronous circuit can still be obtained at the same interconnections. Therefore, the testability characteristics of the original edge-triggered circuit remain virtually unchanged. Second, since only one latch per weighted edge is allowed to move, the final circuit is only marginally different from the original circuit, and changes in layout are not significant. Third, fixed-phase retiming can always be applied to further optimize edge-triggered circuits that have already been optimized using edge-triggered retiming. Finally, fixed-phase retiming can reduce power dissipation without sacrificing performance. In fact, performance may improve.

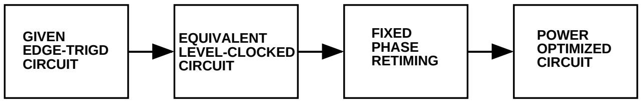

Figure 1: The fixed-phase retiming methodology for low power design.

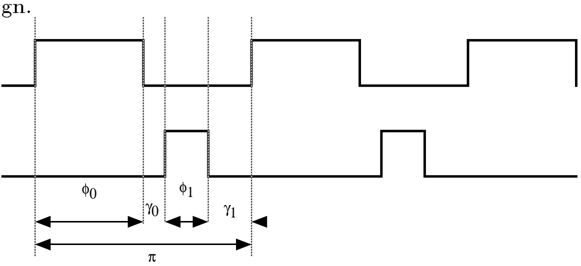

Figure 2: A two-phase clocking scheme $\pi = \langle \phi_0, \gamma_0, \phi_1, \gamma_1 \rangle$.

Fixed-phase retiming is best illustrated by the example in Figure 3. Figure 3(a) shows a section of an edge triggered circuit. The numbers on the edges indicate the potential reduction in power dissipation when an edge-triggered flip-flop is present on that edge, assuming that the rest of the circuit remains unchanged. Negative values of power reduction indicate an *increase* in power dissipation when a flip-flop is placed on an interconnection. This reduction in power dissipation can be achieved if the edge has a high *glitching-capacitance* product [3]. After replacing each edge-triggered flip-flop by two back-to-back level-clocked latches, the resulting circuit is fixed-phase retimed to obtain the circuit in Figure 3(b). Assuming a non-overlapping two-phase clocking scheme $\pi = \langle \phi_0 = 4, \gamma_0 = 1, \phi_1 = 4, \gamma_1 = 1 \rangle$ such as the one shown in Figure 2, power dissipation can be reduced by 11.8 power units. Specifically, the glitching on edges $B \xrightarrow{12} D$, $E \xrightarrow{13} F$ and $E \xrightarrow{-2} H$ is "masked" for 60% of the clock cycle which decreases power dissipation by $0.6 \times (12 + 13 - 2) = 13.8$ units of power. At the same time, the glitching on edges $G \xrightarrow{10} J$ and $H \xrightarrow{-5} K$ is "exposed" for 40% of the clock cycle which increases power dissipation by $0.4 \times (10 - 5) = 2$ power units. In

ISLPED 1996 Monterey CA USA
0-7803-3571-8/96/$5.00 © 1996

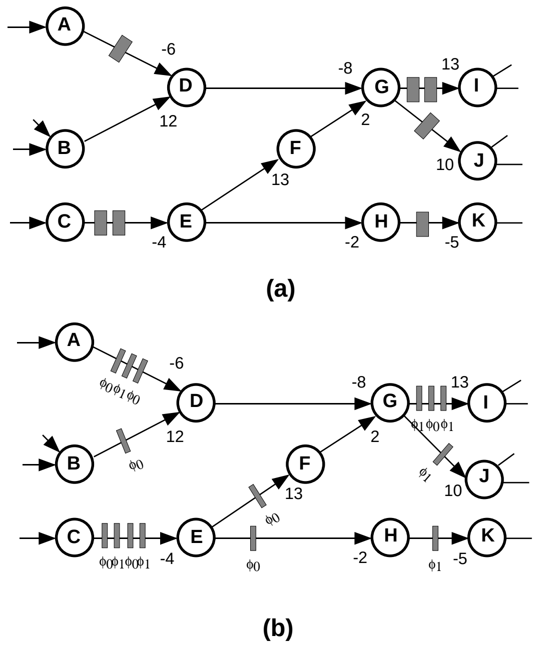

Figure 3: Illustration of fixed-phase retiming. (a) Initial edge-triggered circuit. (b) Fixed-phase retimed circuit that dissipates 11.8 units less power under a $\langle 4, 1, 4, 1\rangle$ two-phase non-overlapping clocking scheme.

order to simplify the computation of changes in power dissipation for this example, we assumed that glitching is uniformly distributed over the entire clock period and that the relocation of latches does not change glitching significantly.

In this paper, we show that fixed-phase retiming can be expressed as an efficiently solvable boolean monotonic quadratic program. Specifically, we describe an efficient algorithm for computing a fixed-phase retiming that minimizes the power dissipation of any given circuit while maintaining its performance. The algorithm runs in $O(V^6 \log V)$ steps, where V denotes the number of combinational blocks in the circuit.

The remainder of this paper has five sections. In Section 2 we describe our graph representation model and give an overview of retiming edge-triggered circuits for low power. In Section 3, we analyze the effects of fixed-phase retiming on power dissipation and derive a mathematical expression for the reduction in power dissipation. In Section 4 we express the fixed-phase retiming problem as a boolean monotonic quadratic program. In Section 5, we describe a linearization of the quadratic program and an $O(V^6 \log V)$-time algorithm for solving it.

2 Preliminaries

In this section we describe the graph representation of a circuit and discuss previous work on the application of retiming for reducing power dissipation in edge-triggered circuits. We also state our assumptions about power dissipation in level-clocked circuits.

2.1 Graph representation

Given an edge-triggered circuit, we obtain an equivalent level-clocked circuit by replacing each edge-triggered flip-flop by two level-clocked latches clocked on alternate phases of a nonoverlapping two-phase clocking scheme $\pi = \langle \phi_0, \gamma_0, \phi_1, \gamma_1\rangle$ Figure 2. We model a two-phase level-clocked circuit as a directed multigraph $G = \langle V, E, d, w, \chi, E_g, C\rangle$. The vertices V in the graph correspond to the combinational elements in the circuit. The directed edges E model the interconnections between the combinational blocks. For a combinational element v, the propagation delay is given by $d(v)$ and its input phase by $\chi(v)$. If the input phase of a vertex v is $\chi(v)$, then $\phi_{\chi(v)}$ clocks the last latch on any path that ends at v. Each edge $u \xrightarrow{e} v \in E$ connects an output of some combinational block u to the input of another block v and is associated with a weight $w(e)$ that gives the latch count on the wire. Each edge $u \xrightarrow{e} v$ is also associated with a pair $(E_g(e), C(e))$, where $E_g(e)$ denotes the average glitching frequency at the output of node u, and $C(e)$ denotes the capacitive load presented by node v to the output of node u. The product $E_g(e) \cdot C(e)$ is a measure of the power dissipation due to glitching on the edge e.

2.2 Retiming edge-triggered circuits

When an edge-triggered flip-flop is placed on a zero-weight edge $u \xrightarrow{i} v$, power dissipation is reduced, since the glitching at the output of u is shielded from the rest of the circuit by the flip-flop. Assuming that the rest of the network remains unchanged, the reduction in power dissipation is given by [3]

$$p_m^{ET}(i) = E_g(i) \cdot C(i) + E_g(i) \cdot \sum_{u \xrightarrow{j} v}^{fanout_i} (s_{j,i} \cdot C(j)) - E_g(i) \cdot C_{ff} \ . \tag{1}$$

The sum in Equation (1) has three terms. The first term $E_g(i) \cdot C(i)$ gives the reduction in power dissipation at the input of v due to masking.

Since the glitching on edge i also propagates through its transitive fanout $fanout_i$, the masking effect of the flip flop also affects power dissipation on each edge $u \xrightarrow{j} v$ in the combinational fanout $fanout_i$. The term $E_g(i) \cdot \sum_{u \xrightarrow{j} v}^{fanout_i} (s_{j,i} \cdot C(j))$ denotes the reduction in power dissipation on the edges of the transitive fanout of i, where $C(j)$ denotes the capacitive load presented by node v to the output of u. The probability that a transition on edge i propagates to edge j is denoted by $s_{j,i}$ and is given by [3]

$$s_{j,i} = Prob(j \updownarrow | i \updownarrow) \ , \tag{2}$$

where $Prob(i \updownarrow | i \updownarrow)$ denotes the probability of a transition at edge j given that there is a transition at edge i.

The third term $E_g(i) \cdot C_{ff}$ denotes an *increase* in power dissipation due to glitching at the flip-flop inputs, where C_{ff} denotes the input capacitive load of the flip-flop.

2.3 Retiming level-clocked circuits

Retiming level-clocked circuits has a similar effect on power dissipation as edge-triggered retiming. In contrast to flip-flops which shield glitching for the entire clock period, level-clocked latches shield glitching for the part of the clock period they are opaque.

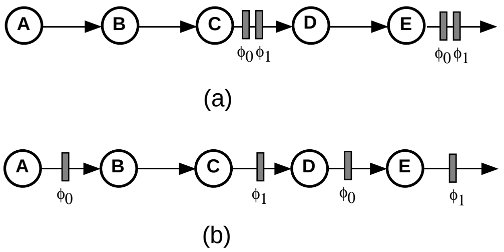

Figure 4: Masking, exposing and remasking. (a) A two-phase level-clocked circuit with back-to-back latches clocked by a two-phase non-overlapping scheme. (b) Fixed-phase retimed circuit. Moving the ϕ_0 latch from $C \rightarrow D$ to $A \rightarrow B$ masks the glitching of A's output from B, but exposes D and E to the glitching of C's output. Placing a ϕ_0 latch on $D \rightarrow E$ remasks the glitching of C from E.

Our analysis of fixed-phase retiming relies on certain simplifying assumptions. In level-clocked circuits, signals that flow through a latch during its transparent phase can initiate computations in the next combinational stage, a phenomenon termed as *cycle stealing.* As a result, data can ripple through several stages of storage elements before their propagation is complete. Our treatment of fixed-phase retiming does not take into account the effects of cycle stealing for the following reasons. First, our approach seeks to minimize the glitching component of power dissipation. Due to the inertial delay of the combinational blocks, we do not expect glitching to propagate through many combinational stages, and thus it is not an issue. Second, cycle stealing is a theoretical potential of level-clocked circuits and it is not clear how many practical circuits employ it extensively. Moreover, since our methodology applies to circuits which were originally edge triggered, we do not expect cycle stealing to be significant after fixed-phase retiming. Another simplifying assumption that we make in our analysis is that glitching is evenly distributed over each clock cycle.

3 Dissipation and fixed-phase retiming

Fixed-phase retiming affects power dissipation by shielding or exposing capacitive nodes to glitches, and by changing the number of latches in the circuit. In this section, we present a mathematical analysis and derive an expression that captures the changes in power dissipation due to fixed-phase retiming.

Figure 4 illustrates the effects of fixed-phase retiming to the glitching-related power dissipation of a circuit. First, when a latch is placed on an edge $u \xrightarrow{e} v$ with $w(e) = 0$, the glitching at the output of u is *masked* from v's input while the latch is opaque. In Figure 4(b), for example, when a latch is introduced on $A \rightarrow B$, the glitching at A's output is not visible to B when ϕ_0 is low. Second, when a latch is removed from an edge $u \xrightarrow{e} v$ with $w(e) = 2$, the glitching at the output of u, which was previously masked from v, is now *exposed* to v while the latch is transparent. For the edge $C \rightarrow D$ in Figure 4(b), the glitching of C is visible to D while ϕ_1 is high. Third, exposed glitching can get *remasked.* For example, the glitching at C's output is remasked from E when a latch is placed on the edge $D \rightarrow E$.

In the following four subsections, we give a detailed analysis of the changes in a circuit's power dissipation due to fixed-phase retiming. We first focus on each individual edge e, based on whether the corresponding wire in the original circuit has 0, 2, or more than 2 latches. We then consider the power dissipation effects of changes in the latch count of the circuit. Finally, we derive an expression for the overall change in power dissipation by adding up the contributions of the individual circuit components.

3.1 Edges with zero latches in original circuit

Consider an edge $u \xrightarrow{i} v \in E_c = \{e \in E : w(e) = 0\}$. Before retiming, power dissipation due to glitching occurs only on the combinational block v and its combinational transitive fanout $fanout_i$. This dissipation is given by the expression

$$p_{bef}(i) = E_g(i) \cdot \{C(i) + \sum_{j}^{fanout_i} (s_{j,i} \cdot C(j))\} \ . \qquad (3)$$

Power dissipation at latch inputs is zero since $w(i) = 0$.

After retiming, dissipation at the input of the combinational block depends on whether the retiming process introduces a latch on i or not. When a ϕ_0-latch is introduced on the edge i, i.e. $r(v) - r(u) = 1$, power is dissipated at the input of v during the transparent phase of the latch's operation. When there is no latch on the edge i, i.e. $r(v) - r(u) = 0$, power dissipation is the same as before retiming. Thus, the power dissipation associated with i due to dissipation in combinational blocks is given by

$$\begin{aligned} p_{aft}(i) \quad = \quad & \{\frac{\phi_0}{\pi} \cdot (r(v) - r(u)) + [1 - (r(v) - r(u))]\} \\ & \times E_g(i) \cdot \{C(i) + \sum_{j}^{fanout_i} (s_{j,i} \cdot C(j))\} \ . \qquad (4) \end{aligned}$$

The first term in Equation (4) denotes the power dissipation when $r(v)-r(u) = 1$, and the second term denotes the dissipation when $r(v) - r(u) = 0$.

For the kind of level-clocked latch implementations we consider, the capacitive load presented by a latch is the same, regardless of whether it is open or closed. As a result, the contribution to power dissipation by a latch on the edge i is given by

$$p^L_{aft}(i) = E_g(i) \cdot (r(v) - r(u)) \cdot C_L \ . \qquad (5)$$

For certain other implementations of a level-clocked latch such as a pass transistor inverter combination, the input capacitance C_L may depend on whether the latch is open or closed, in which case Equation (5) would have to be adjusted accordingly.

From Equations (3), (4), and (5), it follows that the reduction in power dissipation, due to the masking effect of fixed-phase retiming, associated with an edge $u \xrightarrow{i} v \in E_c$ is given by

$$\Delta p_{E_c}(i) = p_{bef}(i) - (p_{aft}(i) + p^L_{aft}(i)) \ . \qquad (6)$$

3.2 Edges with two latches in original circuit

Consider an edge $u \xrightarrow{i} v \in E_l = \{e \in E : w(e) = 2\}$. Before retiming, the power dissipation associated with the combinational block v of this edge is zero, since the output of the latches is a clean transition without glitches. The glitching at the input of the latches is seen by the ϕ_0-latch for the entire clock period and by the ϕ_1-latch for the duration when the ϕ_0-latch is transparent. As a result, the contribution to dissipation by the inputs of the latches is given by the expression

$$p^L_{bef}(i) = \frac{\pi + \phi_0}{\pi} \cdot E_g(i) \cdot C_L \ . \tag{7}$$

After fixed-phase retiming, the power dissipation associated with the combinational block of the edge i is given by the expression

$$\begin{aligned} p_{aft}(i) \quad &= \quad \frac{\phi_1}{\pi} \cdot E_g(i) \cdot r(u) \times \{C(i) \cdot (1 - r(v)) \\ &+ \sum_{u_j \xrightarrow{j} v_j}^{fanout_i} s_{j,i} \cdot C(j) \cdot (1 - r(v_j))\} \ . \end{aligned} \tag{8}$$

When $r(u) = 0$, i.e. the ϕ_0-latch remains on i, the combinational power dissipation remains zero. When $r(u) = 1$, Equation (8) is the sum of two terms. The first term is non-zero when $r(v) = 0$, i.e. when no ϕ_0-latch is introduced on the edge i, and the glitching $E_g(i)$ is visible to node v for the duration the ϕ_1 latch is open. When $r(v) = 1$, however, ϕ_0-latch is introduced which remasks glitching and thus no power is dissipated in the combinational fanout of i. The second term represents the effect of glitching $E_g(i)$ on each edge $u_j \xrightarrow{j} v_j$ in the combinational transitive fanout of i. When $r(v_j) = 0$, the glitching effect propagates, but it gets remasked when $r(v_j) = 1$.

The dissipation at the latch input after fixed-phase retiming is given by the expression

$$\begin{aligned} p^L_{aft}(i) \quad &= \quad \frac{\pi + \phi_0}{\pi} \cdot E_g(i) \cdot C_L \cdot (1 - r(u)) \\ &+ E_g(i) \cdot r(u) \{C_L + \frac{\phi_1}{\pi} \cdot C_L \cdot r(v)\} \ . \end{aligned} \tag{9}$$

Equation (9) has three terms. When $r(u) = 0$, power dissipation remains unchanged and is given by the first term. When $r(u) = 1$, the first term reduces to zero, and the second term denotes the dissipation in the ϕ_1-latch. When $r(v) = 1$, the third term denotes the power dissipation in the incoming ϕ_0-latch on i. The term denoting power dissipation in the incoming latches on edges $u_j \xrightarrow{j} v_j$ in the combinational transitive fanout of i is not included in Equation (9), since it is taken into account in the power dissipation term associated with edges $u_j \xrightarrow{j} v_j$, where $w(j) = 0$.

It follows from Equations (7), (8), and (9) that the change in power dissipation for the edge i is given by the expression

$$\Delta p_{E_l}(i) = p^L_{bef}(i) - (p_{aft}(i) + p^L_{aft}(i)) \ . \tag{10}$$

3.3 Edges with more than two latches

For edges that have more than two latches in the original circuit, the power dissipation does not change with fixed-phase retiming.

3.4 Latch power dissipation

We now consider the change in power dissipation due to a change in the number of latches in the circuit. Since the number of ϕ_1 latches remains unchanged, this component of power dissipation is entirely due to the transition activity E_{ϕ_0} on the ϕ_0 clock line and is given by the expression

$$\Delta p_{clk}(\phi_0) = E_{\phi_0} \cdot C_{L\phi_0} \cdot \sum_{u \xrightarrow{e} v \in E} r(v) - r(u) \ . \tag{11}$$

The switching activity on the ϕ_0 clock line is given by E_{ϕ_0}, and the capacitive load presented by the latch to the ϕ_0 clock line is given by $C_{L\phi_0}$.

It follows from Equations (6), (10), and (11) that the net reduction in power dissipation by means of fixed-phase retiming is given by

$$\mathcal{PR} = \sum_{e \in E_c} \Delta p_{E_c}(e) + \sum_{e \in E_l} \Delta p_{E_l}(e) + \Delta p_{clk}(\phi_0) \ . \tag{12}$$

Thus, the power optimization problem by fixed-phase retiming is equivalent to maximizing the objective function $\mathcal{PR}$ while maintaining the performance of the circuit.

4 Quadratic programming formulation

In this section we define the fixed-phase retiming problem for reducing power dissipation while maintaining performance. We show that this problem can be reduced to a boolean quadratic programming problem with monotone inequalities and positive quadratic coefficients. We exploit this property in the next section to design a polynomial-time algorithm for fixed-phase retiming.

The following lemma gives necessary and sufficient conditions for a retimed circuit G_r to be properly timed by a given clocking scheme π.

Lemma 1 (Lemma 36, [5]) *Let $G = \langle V, E, d, w, \chi \rangle$ be a two-phase, level-clocked circuit, let $\pi = \langle \phi_0, \gamma_0, \phi_1, \gamma_1 \rangle$ be a clocking scheme, and let $r : V \longrightarrow \mathbf{Z}$ be a retiming function. Moreover, let p be the shortest (least-weight) path from u to v in the graph $G' = \langle V, E, w' \rangle$ with edge-weight function $w'(e) = \pi w(e)/2 - d(j)$ for each edge $i \xrightarrow{e} j$ in E. Then, the retimed circuit G_r is properly timed by π if and only if for every edge $u \xrightarrow{e} v \in E$, we have*

$$r(u) - r(v) \leq w(e) \ , \tag{13}$$

and for every vertex pair $u, v \in V$, we have

$$\begin{aligned} d(p) \quad \leq \quad & \pi \left(\frac{1 + w(p)}{2} \right) + \phi_{\chi(u)} \\ & + \pi \left\lfloor \frac{r(v)}{2} \right\rfloor + (r(v) \bmod 2)(\gamma_{\chi(u)} + \phi_{1-\chi(u)}) \\ & - \pi \left\lfloor \frac{r(u)}{2} \right\rfloor - (r(u) \bmod 2)(\gamma_{\chi(u)} + \phi_{\chi(u)}) \ , \end{aligned} \tag{14}$$

if $\chi(u) \neq \chi(v)$*, and*

$$\begin{aligned} d(p) \quad \leq \quad & \pi\left(\frac{2+w(p)}{2}\right) - \gamma_{1-\chi(u)} \qquad (15) \\ & + \pi\left\lfloor\frac{r(v)}{2}\right\rfloor + (r(v) \bmod 2)(\gamma_{1-\chi(u)} + \phi_{\chi(u)}) \\ & - \pi\left\lfloor\frac{r(u)}{2}\right\rfloor - (r(u) \bmod 2)(\gamma_{\chi(u)} + \phi_{\chi(u)}) \ , \end{aligned}$$

if $\chi(u) = \chi(v)$*.*

Thus, the fixed-phase retiming problem for power optimization under performance constraints can be defined as follows.

Definition 2 (Problem FPR – Fixed-Phase Retiming) *Let* $G = \langle V, E, p, w, \chi, E_g, C\rangle$ *be a synchronous circuit, and let* $\pi = \langle \phi_0, \gamma_0, \phi_1, \gamma_1\rangle$ *be a two-phase clocking scheme. Moreover, let* p *be the shortest (least-weight) path from* u *to* v *in the graph* $G' = \langle V, E, w'\rangle$ *with edge-weight function* $w'(e) = \pi w(e)/2 - d(j)$ *for each edge* $i \xrightarrow{e} j$ *in* E*. The fixed-phase retiming problem for power minimization is to compute a retiming* $r : V \longrightarrow \{0,1\}$ *such that we maximize the objective* $\mathcal{PR}$ *from Expression (12) subject to the constraints that for every edge* $u \xrightarrow{e} v \in E$*, we have*

$$r(u) \leq r(v) + w(e) \ , \qquad (16)$$

and for every vertex pair $u, v \in V$*, we have*

$$\begin{aligned} d(p) \quad \leq \quad & \pi(\frac{2+w(p)}{2}) - \gamma_0 \qquad (17) \\ & + r(v)(\gamma_0 + \phi_1) - r(u)(\gamma_1 + \phi_1) \ . \end{aligned}$$

The objective $\mathcal{PR}$ gives the reduction in power dissipation under fixed-phase retiming. Inequalities (16) and (17) guarantee that the performance of the circuit is maintained. These inequalities follow from the inequalities in Lemma 1 for $r \in \{0,1\}$ and $\chi(u) = \chi(v) = 1$, since the original circuit is edge-triggered.

Definition 3 (Problem BMQP – Boolean Monotonic Quadratic Programming) *Let* $d_i \in \mathbb{R}$ *for* $i = 1, \ldots, n$*. Moreover, let* $d_{ij} \in \mathbb{R}$ *for* $i, j = 1, \ldots, n$*, and let* $P = \{(i,j) : d_{ij} > 0, 1 \leq i, j \leq n\}$*. The boolean monotonic quadratic programming problem is to compute* $x_i \in \{0,1\}$ *for* $i = 1, \ldots, n$*, such that the objective*

$$\sum_i^n d_i x_i + \sum_{(i,j)\in P} d_{ij}\, x_i \cdot x_j \qquad (18)$$

is maximized subject to the constraints

$$a_k x_i - b_k x_j \leq c_k \ , \qquad k = 1, \cdots, m \ , \qquad (19)$$

where, $a_k, b_k, c_k \in \mathbb{R}$ *and* $a_k, b_k \geq 0$ *for* $k = 1, \ldots .m$*.*

Problem FPR can be brought to the form of Problem BMQP, that is, a boolean quadratic program with monotone inequalities and a maximization objective with positive quadratic coefficients. The terms $r(u) \cdot r(v)$ and $r(u) \cdot r(v_i)$ in the objective $\mathcal{PR}$ are quadratic. Moreover, the unknowns r take integer values from the set $\{0,1\}$ by the definition of fixed-phase retiming. Boolean quadratic programs are intractable, in general. Two additional properties of Problem FPR enable us to solve it in polynomial time, as we describe in Section 5. First, all constraints in Definition 2 are monotone inequalities with at most two variables per inequality. Second, all quadratic terms in the objective $\mathcal{PR}$ have positive coefficients, assuming that the latch capacitance is smaller than the input capacitance of any combinational block, that is, $C_L \leq C(i)$ for all $i \in E$. A comprehensive comparison of the cells in the CMOS3 Cell Library [1] confirms that this assumption is reasonable. With the exception of a high impedance inverter whose input capacitance is comparable to that of a level-clocked latch, all other combinational blocks have higher input capacitances than a level-clocked latch. We conclude this section with the following lemma.

Lemma 4 *Problem* FPR *can be reduced to Problem* BMQP*.* □

5 Polynomial-time algorithm

In this section we give a polynomial-time algorithm for fixed-phase retiming. As we showed in the previous section, Problem FPR can be reduced to a boolean monotonic quadratic program. Although such programs are intractable in general, we can solve Problem FPR efficiently by reducing it even further to a boolean monotonic program with a *linear* objective. The key to the linearization of the quadratic objective is that the coefficients of its quadratic terms are positive. Using a technique proposed by Hochbaum and Naor from [2] we can obtain an $O(V^6 \log V)$-time algorithm for the problem.

Problem BMQP can be transformed into a boolean monotonic linear program by introducing a new boolean variable y_{ij} for each quadratic term $x_i \cdot x_j$, and by constraining y_{ij} to take the value $x_i \cdot x_j$.

Definition 5 (Problem BMLP – Boolean Monotonic Linear Programming) *Let* $d_i \in \mathbb{R}$ *for* $i = 1, \ldots, n$*. Moreover, let* $d_{ij} \in \mathbb{R}$ *for* $i, j = 1, \ldots, n$*, and let* $P = \{(i,j) : d_{ij} > 0, 1 \leq i, j \leq n\}$*. The boolean monotonic linear programming problem is to compute* $x_i, y_{ij} \in \{0,1\}$ *for* $i, j = 1, \ldots, n$*, such that the objective*

$$\sum_i^n d_i x_i + \sum_{(i,j)\in P} d_{ij}\, y_{ij} \qquad (20)$$

is maximized subject to the constraints

$$\begin{aligned} y_{ij} - x_i \quad & \leq \quad 0 \qquad (i,j) \in P \ , \qquad (21) \\ y_{ij} - x_j \quad & \leq \quad 0 \qquad (i,j) \in P \ , \qquad (22) \\ a_k x_i - b_k x_j \quad & \leq \quad c_k \qquad k = 1, \cdots, m \ , \qquad (23) \end{aligned}$$

where $a_k, b_k, c_k \in \mathbb{R}$ *and* $a_k, b_k \geq 0$ *for* $k = 1, \ldots, m$*.*

```
SolveFPR$(G, \pi)$
Compute objective $\mathcal{PR}$ and constraints $\mathcal{C}$ for Problem FPR
Linearize $\mathcal{PR}$ to obtain a BMLP by replacing every
  term of the form $r(u) \cdot r(v)$ by $r_{uv}$ such that
  $\mathcal{C} \leftarrow \mathcal{C} \cup \{r(u) \geq r_{uv}\} \cup \{r(v) \geq r_{uv}\}$
Compute $r$ that maximizes $\mathcal{PR}$ subject to $\mathcal{C}$ using the
  Hochbaum-Naor technique.
return $r$.
```

Figure 5: Algorithm SolveFPR for solving Problem FPR.

It is known that unconstrained boolean quadratic programs with objectives in the form of Expression (18) and their corresponding linearized programs with objectives in the form of Expression (20) have the same optimum value [4]. We prove a similar result for boolean monotonic quadratic programs.

Lemma 6 *The optimum values of the objectives for Problem* BMQP *and Problem* BMLP *are equal.*

Proof. Given a solution x_i, $i = 1, \ldots, n$, for Problem BMQP, we can construct a feasible vector for Problem BMLP with equal objective value. Such a vector can be obtained by using the same x_i's and by setting $y_{i,j} = x_i \cdot x_j$ for every $(i,j) \in \{(i,j) : d_{ij} > 0, 1 \leq i,j \leq n\}$. Conversely, given integers x_i, y_{ij}, $i,j = 1, \ldots, n$, that solve Problem BMLP, the x_i's are also feasible for Problem BMQP and achieve the same value for the objective. □

The following lemma is used to obtain a polynomial-time algorithm for Problem FPR.

Lemma 7 (Theorem 3.7, [2]) *The integer optimal solution of a monotone linear system of inequalities with respect to an arbitrary linear objective can be computed in $O(m(\sum_{i=1}^{n} |V_i|)^2 \log(\sum_{i=1}^{n} |V_i|))$ steps, where m is the number of inequalities, n is the number of variables, and V_i is the set of integers between the largest and smallest integer feasible values of variable x_i.*

Theorem 8 *Problem* BMLP *can be solved in $O(mn^2 \log n)$ steps.*

Proof. Problem BMLP is a monotone linear system of inequalities with boolean variables and a linear objective. We thus have $|V_i| = 2$ for every variable x_i, and the running time follows immediately from Lemma 7. □

We now present an efficient algorithm for solving Problem FPR which is based on the results of Hochbaum and Naor in [2]. The algorithm, given in Figure 5, first computes the objective and the set of constraints to obtain a boolean quadratic program. This boolean quadratic program is then linearized to obtain an instance of BMLP which is then solved using the algorithm due to Hochbaum-Naor [2]. We conclude this section with the following theorem.

Theorem 9 *Algorithm* SolveFPR *solves Problem* FPR *in $O(V^6 \log V)$ steps.*

Proof. From Lemma 4, we know that Problem FPR can be reduced to Problem BMQP with $O(V)$ variables and $O(V^2)$ constraints. Since $\mathcal{PR}$ can have $O(V^2)$ quadratic terms, from Lemmas 4 and 6 we infer that Problem FPR can be reduced to Problem BMLP with $O(V^2)$ variables and $O(V^2)$ constraints. The running time of the algorithm follows from Theorem 8. □

6 Conclusion

In this paper we have investigated the fixed-phase retiming transformation for reducing power dissipation in CMOS digital circuits. We have shown that the ensuing optimization problem can be formulated as a boolean monotonic linear program which can be solved in $O(V^6 \log V)$ steps. We are currently evaluating the effectiveness of fixed-phase retiming. Preliminary experiments with a 4-bit carry-lookahead adder indicate power savings of about 15%. Our experiments reveal that in addition to shielding highly capacitive nodes from glitching, fixed-phase retiming has the potential of reducing power dissipation by equalizing arrival times of signals at gate inputs.

References

[1] D. V. Heinbuch. *CMOS3 Cell Library.* Addison Wesley, 1988.

[2] D. S. Hochbaum and J. Naor. Simple and fast algorithms for linear and integer programs with two variables per inequality. *SIAM J. Computing*, 23(6):1179 – 1192, December 1994.

[3] J. Monteiro, S. Devadas, and A. Ghosh. Retiming sequential circuits for low power. In *Digest of Technical Papers of the 1993 IEEE International Conference on CAD*, pages 398–402, November 1993.

[4] P. Hansen P. L. Hammer and B. Simeone. Roof duality, complementation and persistancy in quadratic 0-1 optimization. *Mathematical Programming*, 28:121–155, 1984.

[5] M. C. Papaefthymiou. *A Timing Analysis and Optimization System for Level-Clocked Circuitry.* PhD thesis, Massachusetts Institute of Technology, September 1993. Available as MIT/LCS/TR-605.