Abstract

In this paper, we have modelled the flip-flop clock to output delay dependency on the data arrival time and introduced this phenomenon in timing analysis. Traditionally, finding the minimum clock period of a flip-flop based sequential design was based on the assumption that the setup-time and clock to output delay of a flip-flop are constant and hence each stage of the pipeline can be analyzed independently. However, it is well known that the delay of a flip-flop depends on the data arrival time at its input and hence there exists an interdependence among different pipeline stages. The problem of finding the minimum clock period of such a coupled system is a non-trivial problem. In this paper, we formulate the problem of finding the minimum clock period of a flip-flop based sequential circuit accounting for these dependencies. We show that the problem is a non-linear convex optimization problem. We propose three different solution approaches and compare their results on ISCAS '89 sequential benchmark circuits. Modeling these data arrival time dependencies we have seen a consistent decrement of approximately 50-60ps compared to the traditional approach using constant setup-time and flip-flop delays. We also show how the analysis can be extended to account for hold time constraints for short paths in the circuit.

1. Introduction

Static timing analyzers are widely used to verify the behavior of large digital circuit designs in various stages of design. They have also become the core engine used inside circuit optimization tools, such as the transistor and gate sizing tools and logic synthesis. Ever since the frequency of operation of circuits has gained marketing focus, accurate and efficient computation of a circuit’s clock period is one of the most important tasks in design.

Flip-flop delay is an increasing concern for high performance circuit designers. Cycle times have been shrinking dramatically driven by both faster gate delays and by more aggressive designs using fewer gates in a single pipeline stage. Thus, flip-flop delay accounts for a significant portion of the clock period and therefore, accurate modelling of this delay is critical. Also, less pessimistic STA algorithms based on better models help reduce the overall clock period of the circuit, by aiding better design and verification techniques.

Traditionally, the task of determining the minimum clock period of a flip-flop based sequential circuit has consisted of finding the longest combinational path between any two flip-flops of the circuit and adding setup-time, flip-flop delay and clock skew to it. (Note that the effect of clock jitter on clock period has been neglected in the entire paper without the loss of generality.) The above formulation is based on the assumption that the setup time of a flip-flop is fixed and the clock to output delay of a flip-flop is constant. This results in the major simplification that each pipeline stage can be considered independently from every other pipeline stage of the circuit. However, it is well known that the delay of a flip-flop is in fact a function of the difference between the clock and the data arrival time, as shown in Figure 1. We define $T_{DFF}$ as the difference between the clock and data arrival time and refer to $T_{fl}$ as the clock to output delay of the flip-flop. As shown in Figure 1, $T_{fl}$ increases with decreasing $T_{DFF}$ until the flip-flop goes metastable and $T_{fl}$ approaches infinity. Hence, both $T_{fl}$ and $T_{DFF}$ of a flip-flop must be considered as variables with mutual dependencies rather than being treated as constants. We show in this paper that the simplification of using a fixed setup time and constant flip-flop delay may yield a pessimistic minimum clock period, and it is therefore necessary to consider these interdependencies between pipeline stages. The problem of finding the clock period of circuit considering this dependency has similarity with the problem of finding the clock period of a latch based sequential circuit in [1][10][11] except for the fact that this problem is non-linear while latter in can be solved linearly.

In this paper, we model the flip-flop $T_{fl}$ dependencies on data arrival times and propose a new timing analysis algorithm that accounts for these dependencies, which to our knowledge is the first effort to address this problem formally. We describe the problem of determining the minimum clock period as non-linear optimization problem and prove that it is convex. We propose three solution methods, a sequential quadratic programming approach, simple rectilinear manhattan decent approach, a satisfiability approach, and have compared their performance in terms of their runtime and number of iterations on ISCAS’89 sequential benchmark circuits. We have also discussed their ease of integration with the existing static timing analyzers. We compared the exact solution obtained from the proposed formulation with the traditional approach using different $T_{setup}$ times and observed a reduction in the clock period of approximately 50-60ps, which is equivalent to approximately one fan out of four delay in this technology.

The treatment of $T_{fl}$ as a variable also has implications on the short path constraints of a given circuit. As the data arrival time gets closer to the clock the $T_{fl}$ increases thus reducing the number of hold time violations. We therefore show how our analysis can be extended to determine the lowest frequency that fulfills all the hold-time constraints. The long path analysis therefore finds the maximum frequency for which the long path constraints are met whereas the short path analysis finds the minimum frequency for which all short path constraints are satisfied. For any functionally correct design, its is therfore necessary for $f_{short} < f_{long}$. This results in a window of valid operating frequency as shown in Figure 2.

The remainder of the paper is organized as follows. Section 2 discusses the traditional method of finding the clock period and
the model of the flip-flop delay. Section 3 discusses the non-linear problem formulation and discusses its properties. Section 4 discusses three different approaches used to solve this non-linear problem. Section 5 describes the short path problem and its analysis, Section 6 presents results and Section 7 concludes this paper.

2. Background and Delay Model

In this section, we describe the traditional model used to compute the clock period of a circuit. We then present a new model for flip-flop clock to output delay Vs. data arrival time. We end the section with an example that demonstrates that the clock period computed based on traditional model is pessimistic. We introduce the following nomenclature.

Nomenclature

- \( T_c \): Clock period of the circuit.
- \( T_{\text{setup}} \): Setup time for the flip-flop.
- \( T_f \): Flip-flop clock to output delay.
- \( T_{\text{load}} \): Hold time of flip-flop.
- \( T_{\text{skew}} \): Characterized skew of the clock network.
- \( n \): Number of flip-flops in the circuit.
- \( T_{f,i,j} \): Effective delay between \( i \)th and \( j \)th flip-flops.
- \( T_{\text{Logic},i,j} \): Maximum combinational delay between \( i \)th and \( j \)th flip-flops.
- \( T_{\text{Logic},i,j,\min} \): Minimum combinational delay between the \( i \)th and \( j \)th flip-flop.
- \( T_{\text{Diff},i} \): Difference between clock and data arrival at \( i \)th flip-flop.
- \( T_{\text{Diff},ij} \): Difference between the clock and data arrival for a path between \( i \)th and \( j \)th flip-flops.

Traditionally, computing clock period of a sequential circuit is performed as described below.

\[
T_{c,ij} = T_{\text{setup}} + T_{f,i,j} + T_{\text{Logic},i,j,max} + T_{\text{skew}} \quad \text{(EQ 1)}
\]

\[
T_{\text{setup}} = \max\{T_{c,ij}\} \quad \forall (i,j \leq n) \quad \text{(EQ 2)}
\]

\( T_{\text{setup}} \) is defined as the difference between the clock and the arrival time where \( T_f \) increases by 5% from its nominal value as suggested in [2] and is shown as Figure 1.

The fundamental assumption in the above timing model is that the setup time and the flip-flop delay are constant, with the flip-flop going metastable and the delay approaching infinity as soon as the difference between the clock and arrival time is less than \( T_{\text{setup}} \). However, as shown in Figure 1, the transition between proper functionality and metastability is not as discrete as this model describes. The delay of a flip-flop increases gradually as the difference between the arrival time and clock decreases, creating a dependency of \( T_f \) on the data arrival time.

In order to include this effect of variable flip-flop delay in timing analysis, it is important to accurately model the delay of a flip-flop as a function of its data arrival time. Based on the model used to resolve the metastability of a latch we propose the model below in EQ 3 for the flip-flop delay.

\[
T_f = A \exp(B \cdot T_{\text{Diff}}) + C \quad \text{(EQ 3)}
\]

where A, B, C are real non-zero constants.

This model can be easily fitted based on the nonlinear least square method using the trust region dogleg algorithm, with weighted robust regression[12][3].

In Figure 3, we show a comparison between the proposed model in EQ3 and the HSPICE delay for a D-flip-flop from an industrial 0.18 micron library with a loading of 50 ff at 75°C. We fitted the proposed model for four different flip-flop cells from a commercial library. Table 1 summarizes the quality of the fit for different loads, with input slope of 1.8V/.1ns and temperature of 70°C. In Table 1, RMSE is the Root Mean Square Error whereas R-square is a statistic metric used to denote how successful the fit is in explaining the variation of the curve, about its mean. The ideal value for R-square is 1. From Table 1 it is evident that the model mentioned in EQ3 is a good match for describing the flip-flop delay dependence on its data arrival time.

We now show with an example how the traditional approach computes a pessimistic clock period. Consider a hypothetical cyclical sequential circuit as shown in Figure 4. For the sake of simplicity, we assume that both flip-flops in this circuit are of the same kind with the same loading and identical delay dependency curves. Furthermore, we assume that the circuit delay parameters are as follows: \( T_f = 150 \)ps, \( T_{\text{Logic}} = 700 \)ps, \( T_{\text{setup}} = 108 \)ps, \( T_{\text{skew}} = 0 \)ps.

Based on these parameters and applying EQ 1 we obtain:
The objective in the above formulation is to minimize $T_c$ through the optimal allocation of all $T_{\text{Diff},ij}$ in the circuit. From the above formulation it is clear that $T_c$ is a convex function in terms of the optimization variables $T_{\text{Diff},ij}$. Though the proof of this convexity is straightforward, it has significant implication on the optimization problem and hence we include the proof in this paper.

**Theorem 1.** $T_c$ is a convex function with $T_{\text{Diff},ij}$ as variables.

**Proof:**

We first observe that $T_{\text{Diff},ij}$ in (EQ5) is a convex function of variables $T_{\text{Diff},ij}$ as the minimum function retains convexity if its arguments are convex. Also, $T_{\text{ff},i}$ (EQ6) is convex as $T_{\text{Diff},ij}$ is convex and exponential functions are convex. From this follows that $T_{c,ij}$ (EQ4) is convex as the sum of convex functions is a convex function and consequently we find that $T_c$ (EQ7) is a convex function as $T_{c,ij}$ is convex and the maximum function also retains convexity if its arguments are convex. □

We now examine the properties of this optimization function both when the delay of the logic stages are balanced and when they are unbalanced.

**Balanced Logic.**

We consider the circuit as shown in Figure 4. Since both the pipeline stages are identical, we can replace the circuit by a single pipeline stage as shown in the Figure 6, as far as the task of finding clock period is concerned. Note that the circuit in Figure 6 is simply the unrolled version of the circuit in Figure 4. The clock period for the circuit in Figure 4 is given by:

$$T_c = T_{\text{Logic},11} + T_{\text{Dff},11} + A \cdot \exp(B \cdot T_{\text{Dff},11})$$  \hspace{1cm} (EQ 9)$$

To determine the minimum $T_c$ we must satisfy $\frac{dT_c}{dT_{\text{Dff},1}} = 0$. Applying this criteria in EQ9, we obtain the condition $\frac{dT_{\text{ff},i}}{dT_{\text{Dff},i}} = -1$. This means that the flip-flop operates at the point where the slope of the curve in Figure 5 reaches -1, as shown by point D. This is intuitively clear since we can only trade-off $T_{\text{Dff}}$ for $T_{\text{ff}}$ until the increase in $T_{\text{ff}}$ exceeds the decrease in $T_{\text{Dff}}$ leading to the optimal operating point of $\frac{dT_{\text{ff},i}}{dT_{\text{Dff},i}} = -1$.

Operating beyond this point would add more $T_{\text{ff}}$ than decreasing $T_{\text{ff}}$ and hence would increasing $T_c$.

The optimization space for $T_c$ as a function of $T_{\text{Dff},1}$ and $T_{\text{Dff},2}$ for the circuit in Figure 7 with balanced loads is shown in Figure 7. Clearly, $T_c$ is convex and its symmetric nature is due to the balanced stage delays $T_{\text{Logic},1}$ and $T_{\text{Logic},2}$. The optimization space can be viewed as the maximum of two intersecting surfaces defined by $T_{\text{c},12}$ and $T_{\text{c},21}$. From this, follows the useful observa-
We therefore repeatedly decrease the variable $T_{\text{Diff}}$ to which $T_c$ is most sensitive, by a small step. In the circuit, the most sensitive $T_{\text{Diff}}$ can be identified as the delay of the critical stage and the subsequent stage are updated and a new critical path is selected. We compare the new $T_c$ with the previous $T_c$ in each iteration and continue iterating until $T_c$ increases by a greater amount that the step size. The steps of the algorithm are as follows.

1. Calculate the critical path.
2. Decrease corresponding $T_{\text{Diff}}$ by a small step and update the timing of the critical stage and the subsequent stage.
3. If the increase in new $T_c$ is less than the step size, goto step 1 else, previous $T_c$ was optimal.

3. Satisfiability Approach (SA)

This method performs a bounded binary search on the optimization space by testing whether the circuit can satisfy a given clock period. We start with an initial upper and lower bound on $T_c$ and terminate the optimization process when the bounds converge. This approach is described as follows:

- $T_{\text{upperbound}} = \max{T_{\text{Logic,ij}}} + K$ (where $K$ is a constant that ensures the upper boundness)
- $T_{\text{lowerbound}} = \max{T_{\text{Logic,ij}}}$
- $T_{\text{current}} = \text{mean} \{T_{\text{upperbound}}, T_{\text{lowerbound}}\}$
- Pick an initial pipeline stage between two flip-flops and calculate $T_{\text{Diff}}$ as the difference of $T_{\text{current}}$ and $(T_f + T_{\text{Logic,ij}})$ for the corresponding pipeline.
- Using this $T_{\text{Diff}}$, calculate $T_f$ based on EQ3 for the next flip-flop and propagate this effect through the circuit.
- If at any stage $T_{\text{Diff}}$ becomes negative, then the circuit is not able to function at $T_{\text{current}}$. Change $T_{\text{lowerbound}} = T_{\text{current}}$ and goto step 3.
- Else if the propagation of $T_{\text{Diff}}$ in the circuit converges, change $T_{\text{current}} = T_{\text{upperbound}}$ and goto step 3.
- Continue iteration until $T_{\text{upperbound}}$ and $T_{\text{lowerbound}}$ converge.

It is important to note that in this method we compute $T_{\text{Diff}}$ based on $T_{\text{current}}$ such that for all stages $T_{\text{Logic,ij}} = T_{\text{current}}$. By computing $T_{\text{Diff}}$ in this manner, we ensure that every path in the circuit is critical. Each time we change a $T_{\text{Diff}}$, based on a change in $T_{\text{current}}$, we impact the flip-flop delay of the next stage propagate this change through the whole circuit. If there are cyclic paths in the circuit, iterations may be required before the circuit stabilizes. The circuit is said to satisfy the clock period $T_{\text{current}}$ if all the $T_{\text{Diff}}$ remain unchanged from the previous iteration with none of them are negative. If a $T_{\text{Diff}}$ becomes negative during the iterations, the circuit cannot meet $T_{\text{current}}$.

We now show that the optimization method reaches the global minimum. Figure 9 shows the entire optimization space $T_c$. By calculating $T_{\text{Diff}}$ based on $T_{\text{current}}$, as mentioned earlier, we ensure that every path in the circuit is critical. This corresponds to solutions in the optimization space that lie on the black line in Figure 9. Henceforth, we shall refer this black line as the critical line. It is clear that the global minimum is on this critical line. Also, in the above proposed method we effectively test whether there exist some intersection point between the $T_{\text{current}}$ surface and critical line. In each iteration we change the bounds and hence $T_{\text{current}}$ in such a way that we descend down the critical line as indicated by the arrow in Figure 9. Hence with each iteration our bounds converge to the minimum point. This method can be argued as a variant of steepest decent algorithm. Figure 9 shows the critical
5. Short Path Problem Formulation and its Analysis

In this section, we first mention the traditional method of identifying short paths and then discuss the implication of variable flip-flop delays on this problem. We then formulate the problem of finding the maximum clock period such that all the short path constraints are met.

In a traditional analysis, a hold time violation is reported when the sum of the flip-flop clock to output delay and the minimum combinational logic delay between the pipeline registers is less than the sum of characterized clock skew and hold time of a flip-flop:

\[ T_{ff, i} + T_{Logic, ij, \min} < T_{skew} + T_{hold, i} \quad \forall (i, j \leq n) \quad (EQ\ 10) \]

It is important to mention that the hold time of a flip-flop is a constant with respect to data arrival time. As suggested earlier in section 2, the clock to output delay of a flip-flop is not a constant but a variable, and a function of data arrival. As the data arrival time gets closer to the clock, the delay of the flip-flop increases as shown in Figure 3. This increase in \( T_{ff} \) helps reduce the short path violations as evident from EQ10. Thus, as we decrease the clock period, the flip-flop delay increases and more flip-flops in a design will meet short path constraint.

We illustrate this effect using the circuit shown in Figure 4 with following parameters, \( T_{skew} = 35ps, \ T_{hold} = 225ps, \ T_{Logic, i} = 50ps \)

Traditional analysis assumes that Dff1 always operates at point C in Figure 5. With this operation point a short path violation is reported for the above case. But it is quite possible for DFF1 to operate at point A, where the short path violation no longer exists. In this case, the short path is analyzed pessimistically by

\[ T_{c, ij} = T_{Diff, ij} + T_{ff, i} + T_{Logic, ij, \min} \quad \forall (i, j \leq n) \quad (EQ\ 11) \]

\[ T_{Diff, j} = \max\{T_{Diff, ij}\} \quad \forall (i \leq n) \quad (EQ\ 12) \]

\[ T_{ff, j} = A \cdot \exp\{B \cdot T_{Diff, j}\} + C \quad (EQ\ 13) \]

Note that the above formulation finds the minimum clock period such that there does exist at least one short path violation. Hence \( T_c - \delta \) is the maximum clock period such that all hold time constraints are met.

The solution of this problem as suggested in EQ16 is the lowest frequency for which all the short path constraints are satisfied whereas the solution of former problem in EQ8 is the highest frequency for which all the long path constraints are satisfied. In Figures 2 these two frequency constraints \( f_{short} < f_{long} \) are shown. For any properly functional design \( f_{short} < f_{long} \) to allow a region of correct operation. By padding the short path delays, we essential lower the lowest short path frequency and the amount of padding determines the size of the feasible region of operation.

6. Results

The discussed methods discussed in Section 4 were implemented and tested on ISCAS ’89 sequential benchmark circuits. We used the netlist as given by cbl.ncsu.edu and optimized them using SYNOPSIS Design Analyzer[9] keeping the sequential circuit elements intact. Table 2 shows the runtime and the number iterations for the three approaches discussed in Section 4. Table 2 presents the following information in column order: Circuit name, number of flip-flops in that circuit, runtime for Sequential Quadratic Programming (SQP), runtime for Rectilinear Manhattan Decent (RMD), runtime for Satisfiability Approach (SA), number of iterations for Rectilinear Manhattan Decent and number of iterations in Satisfiability Approach. The three solutions reached equivalent solutions, given the numerical accuracy of the analysis.

In Table 3, we compare the exact solution obtained by modeling flip-flop delay dependencies on the data arrival time with the
traditional method when using various fixed $T_{\text{setup}}$ times on the delay curve (Figure 1). Table 3, column 2 shows the result obtained by modelling the flip-flop delay dependency using the non-linear problem formulation and column 3, 4, 5 shows the results obtained with traditional timing analysis using the 5% delay point mentioned in Section 2, the delay change curve point where $d/dx = -1$ mentioned in Section 4 and shown as D in Figure 5, and a constant $T_{\text{setup}}$ delay point as shown in Figure 1. The results show significant improvement by modelling the flip-flop delay dependency over the traditional approach.

In addition to this we did compare our predicted highest long path frequency with HSPICE simulation. This simulation was done on some very smaller circuits that allowed for full SPICE simulation. The flip-flop delays were characterized using the same slopes at the input as seen in the HSPICE. The results are tabulated in Table 4. We see that the predicted clock period matches that with HSPICE results with a maximum deviation of 11ps and average deviation of 5ps.

### 7. Conclusion

In conclusion, we have presented a new timing model for computing the clock period of flip-flop based sequential circuits. The timing analysis based on this models is more accurate as compared to the traditional models as it accounts for flip-flop delay dependencies on the data arrival time. We showed that the problem of finding the maximum clock period of a flip-flop based sequential circuit considering these dependencies is a nonlinear but convex optimization problem. We proposed three approaches to solve this problem. Modelling flip-flop delay dependencies on data arrival time, we obtained a decrease in clock period of 50-60ps, which in current circuits with shallower pipelines can be a significant part of the clock period. We have also formulated the problem of finding the lowest frequency such that all the short path constraints are met. Together these two formulation form an

### Table 3. Clock Period calculated by Modelling flip-flop delay dependency and Traditional approach Using various different $T_{\text{setup}}$ time (ps).

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th>New Approach</th>
<th>5% $T_{\text{setup}}$ Time</th>
<th>d/dx = -1 $T_{\text{setup}}$ Time</th>
<th>Constant $T_{\text{tr}}$ $T_{\text{setup}}$ Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>s953</td>
<td>2774.2</td>
<td>2828</td>
<td>2825</td>
<td>2854</td>
</tr>
<tr>
<td>s9234_1</td>
<td>4304.2</td>
<td>4578</td>
<td>4555</td>
<td>4584</td>
</tr>
<tr>
<td>s9234</td>
<td>4484.2</td>
<td>4538</td>
<td>4535</td>
<td>4564</td>
</tr>
<tr>
<td>s838_1</td>
<td>6531.2</td>
<td>6588</td>
<td>6585</td>
<td>6614</td>
</tr>
<tr>
<td>s713</td>
<td>5404.2</td>
<td>5358</td>
<td>5355</td>
<td>5384</td>
</tr>
<tr>
<td>s641</td>
<td>5404.2</td>
<td>5358</td>
<td>5355</td>
<td>5384</td>
</tr>
<tr>
<td>s526n</td>
<td>2654.2</td>
<td>2708</td>
<td>2705</td>
<td>2734</td>
</tr>
<tr>
<td>s526</td>
<td>2634.2</td>
<td>2688</td>
<td>2685</td>
<td>2714</td>
</tr>
<tr>
<td>s400</td>
<td>2254.2</td>
<td>2308</td>
<td>2305</td>
<td>2334</td>
</tr>
<tr>
<td>s420_1</td>
<td>3574.22</td>
<td>3628</td>
<td>3625</td>
<td>3654</td>
</tr>
<tr>
<td>s382</td>
<td>2594.2</td>
<td>2648</td>
<td>2645</td>
<td>2674</td>
</tr>
<tr>
<td>s349</td>
<td>3264.2</td>
<td>3318</td>
<td>3315</td>
<td>3344</td>
</tr>
<tr>
<td>s344</td>
<td>3264.2</td>
<td>3318</td>
<td>3315</td>
<td>3344</td>
</tr>
<tr>
<td>s27</td>
<td>1192.2</td>
<td>1198</td>
<td>1195</td>
<td>1224</td>
</tr>
<tr>
<td>s13207_1</td>
<td>5034.2</td>
<td>5088</td>
<td>5085</td>
<td>5114</td>
</tr>
<tr>
<td>s13207</td>
<td>4994.2</td>
<td>5048</td>
<td>5045</td>
<td>5074</td>
</tr>
<tr>
<td>s15850_1</td>
<td>6344.22</td>
<td>6448</td>
<td>6445</td>
<td>6474</td>
</tr>
<tr>
<td>s1238</td>
<td>3344.22</td>
<td>3398</td>
<td>3395</td>
<td>3424</td>
</tr>
</tbody>
</table>

### Table 4. Comparison of HSPICE and Predicted results for long path analysis.

<table>
<thead>
<tr>
<th>Circuit</th>
<th>Load Type</th>
<th>Hspice Result</th>
<th>Predicted Model</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuit A</td>
<td>Balanced</td>
<td>22/0ps</td>
<td>22/0ps</td>
<td>0ps</td>
</tr>
<tr>
<td>Circuit B</td>
<td>Unbalanced</td>
<td>2310ps</td>
<td>2313ps</td>
<td>3ps</td>
</tr>
<tr>
<td>Circuit C</td>
<td>Balanced</td>
<td>692ps</td>
<td>681ps</td>
<td>-11ps</td>
</tr>
<tr>
<td>Circuit D</td>
<td>Unbalanced</td>
<td>2290ps</td>
<td>2291ps</td>
<td>1ps</td>
</tr>
</tbody>
</table>

basis of a static timing analyzer that models flip-delay dependencies.

### References