Accurate Waveform Modeling using SVD with Applications to Timing Analysis

Anand Ramalingam, Ashish Kumar Singh
Sani R. Nassif, Michael Orshansky, David Z. Pan
The University of Texas at Austin, IBM Research - Austin
http://www.cerc.utexas.edu/utda
Motivation

- On-chip waveforms do not look like “ramps”
  - Devices are operating in more complex regimes and do not at all look like current sources
  - Loads, both interconnect as well as gate inputs, are resistive and non-linear

- We persist in trying to fit an outdated waveform model onto far more complicated behaviors
  - Applications like Statistical Static Timing Analysis (SSTA) require accurate modeling
  - Model inaccuracy must be « expected variability to reliably estimate performance variability
Ramp Based Timing Model

Expressed in terms of a Ramp approximation of input and output waveforms.

- Arrival time, Slope and Load Capacitance

Timing Model: $[t_o, s_o] = F([t_i, s_i], C_L)$
Source of Error in Timing Models

1. Inability of the waveform function (ramp) to fit the real waveform.
2. Estimation of a complex load by a single capacitance.
3. Lack of complete modeling support (coupling noise, multiple input switching etc...).

We are focusing in this paper on the first two sources of error:

- Part 1: $\pi$-model for interconnect
- Part 2: Accurate waveform modeling
Part 1: Better Load Modeling

Benefits:
- A $\pi$ model of the load is clearly a better representation than a single capacitor.
- We did not do a complete study to exactly quantify the improvement achieved.

Costs:
- Modeling a gate’s behavior as a function of a $\pi$ load means we have 2 more variables to vary.
- When using traditional (e.g. full factorial) experiment designs to create the timing models, adding 2 variables can be quite costly.
Existing Work on $\pi$-Models

O’Brien and Savarino developed an algorithm for reducing an RC tree to a driving-point $\pi$–model.
Solving the Dimensionality Problem

- A naïve implementation of a gate model builder may use a full factorial design, resulting in an exponential number of simulation vs. modeling variables.

- We use Latin-Hypercube Sampling, a well established statistical sampling technique instead.
  - Number of simulation $\sim$ linear in number of modeling variables.
Part 2: Accurate Waveform Modeling

History:

- **Heuristic models**
  - Equivalent waveform model [Hashimoto, ICCAD ’03]
  - Weibull distribution [Amin, ICCAD ’03]

- **Change of basis models**
  - Model current not voltage, CSM [Amin, DAC ’06]

- **Data based models**
  - Basis decomposition [Jain, ICCAD ’05]
  - PCA based approach [Nassif, TAU ’04]

Our approach: extend the PCA approach

- Use more appropriate SVD instead of PCA
- Generate waveform model based on complete library
- Demonstrate application to interconnect as well
Data based model

- Divide $[0...V_{dd}]$ into $n$ intervals.
- Measure $t_0$ through $t_{n-1}$ for waveforms of interest.
  - $t_i =$ time at which waveform crosses $V_{dd} \times (i/(n-1))$
Intuition for our work

- The time/voltage \((t_i, v_i)\) pairs that define a waveform are **not independent** of each other.

- To verify this we analyzed waveforms obtained from various cells in the library under varying input and interconnect load conditions.
  - We expect the crossing times \((t_i)\) of these waveforms to be inter-related.

```
<table>
<thead>
<tr>
<th>#</th>
<th>t₀</th>
<th>t₁</th>
<th>t₂</th>
<th>t₃</th>
<th>...</th>
<th>tₙ₋₁</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Exploratory Data Analysis Scatter plot of \((t_i, t_j)\)
Slide 11

Scatter plot of \((t_i, t_j)\) pairs

Observation: Very high of correlation, + correlation depends on how close points are
Comments on Crossing Time Stats

- The crossing times ($t_i$) are obviously not independent of each other
  - Strong correlation across all the crossing times
  - Then $t_i$ can be expressed as a function of a smaller number of independent variables
- How to find a smaller subset of independent variables?
  - Previous work used PCA, which works best when the distribution of the $t_i$ is Gaussian (not the case in general)
- We use an alternative dimension reduction technique, Singular value decomposition (SVD)
  - Designed for the more general case where the $t_i$ are simply linearly related.
Singular Value Decomposition (SVD)

- SVD of a matrix $\mathbf{T}$ is given by
  \[ \mathbf{T} = \mathbf{U} \Sigma \mathbf{V}' \]
- $\mathbf{U}$ – orthonormal basis for columns in $\mathbf{T}$
- $\mathbf{V}$ – orthonormal basis for the rows in $\mathbf{T}$
  - Thus the basis for waveforms
  - Note that the basis are obtained from the data
  - Thus, data speaks for itself (no assumptions needed)
- $\Sigma$ – diagonal matrix contains singular values
  - Singular values are ordered in a non-decreasing order
  - Singular values $\sigma_i$ “weighs” the basis columns $\mathbf{V}_i$
  - First few basis are sufficient to capture the data (i.e. the waveform) accurately
Singular values of a Matrix

A plot of singular values of $\mathbf{T}$
- Only the first few dominate

Values

$\sigma_i$
Interpreting the basis vectors ($\mathbf{V}$)

Each of the columns of $\mathbf{V}$ represents a weighted sum of the times $t_i$

$\mathbf{V}_{.1}$ represents an “average” of all the time points, i.e. $t_{50\%}$

$\mathbf{V}_{.2}$ represents linear weighting of all the time points -- slope --
Example

- If the \( n \) time points of a waveform are represented as pairs \((voltage_j, t_j)\)

<table>
<thead>
<tr>
<th>Voltage</th>
<th>0</th>
<th>1/13</th>
<th>...</th>
<th>12/13</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>t [ps]</td>
<td>599</td>
<td>632</td>
<td>...</td>
<td>788</td>
<td>802</td>
</tr>
</tbody>
</table>

- Consider \( \mathbf{V}_{.2} \), which we interpreted as \( \sim \) slope in the previous slide

<table>
<thead>
<tr>
<th>( \mathbf{V}_{.2} )</th>
<th>( \mathbf{v}_{1,2} )</th>
<th>( \mathbf{v}_{2,2} )</th>
<th>...</th>
<th>( \mathbf{v}_{13,2} )</th>
<th>( \mathbf{v}_{14,2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>-0.52</td>
<td>-0.41</td>
<td>...</td>
<td>0.31</td>
<td>0.39</td>
</tr>
</tbody>
</table>

- The value in the new basis is given by dot product \( \langle t, \mathbf{V}_{.2} \rangle = -202.31 \), which approximates the slope.
  - We call this dot product a moment \( (m_2) \)
Summary of SVD Analysis

Library of cells

Sample Inputs

Simulate (Spice)

Post-Process

Ex: Ramp Model

Get \([t_0, s_0]\) to represent waveform

Waveform Analysis

T matrix

SVD: \( T = U\Sigma V' \)

New representation: \( TV \)

TV matrix

Only first few \( m_i \) required to accurately represent a waveform

Get \([t_0, s_0]\) to represent waveform
What Does This All Mean?

- It is possible to sample, simulate and analyze the set of waveforms that a *library of cells* would produce.
  - From this we can determine precisely how many independent variables are required in order to represent waveforms with a specified accuracy.
  - When we analyze the entire library we might need more independent variables than for a given cell.

- Once the independent variables are selected, we also get a transformation that allows us to go from the independent variables to the waveform.
  - So the complete waveform can be readily re-generated from the values of those variables.
In normal (ramp-based) STA, we propagate waveforms through gates and through wires.

Gate Timing Model:

\[ [t_o, s_o] = F_G([t_i, s_i], C_L) \]

Implemented using standard timing models

Wire Timing Model:

\[ [t_w, s_w] = F_W([t_o, s_o], C_L) \]

Implemented using tools such as RICE
We need to propagate waveforms through gates and through wires.

**Gate Timing Model:**

\[ m_o = F_G([m_i], C_L) \]

**Wire Timing Model:**

\[ m_w = F_W([m_o], C_L) \]

Implemented using enhanced timing models.

We will explore this next.
SVD + STA (Interconnects)

- We will still use RICE to propagate the waveform.
- We will use the SVD transforms to convert back and forth between real waveform representation (voltage vs. time) and moments!
Waveform recovery from moments

Waveform recovered at the far end of the interconnect of an inverter

Voltage [V] vs. Time [ps]

SPICE
4Moments approx
Where Is This Model Needed

- Any time that an interface between the analog and digital world is required.
  - Input/Output from wire loads.
- Any time that knowledge of the waveform details is desired
  - SSTA, where model inaccuracy must be « expected variability to reliably estimate performance variability
- Another Example: Estimating $I_{DD}$ in power grid simulation
  - A ramp waveform estimate not useful for since it makes the current look like a step
Features of our model

- Precisely quantify the error we commit in modeling waveforms
- A model which can gracefully expand to model additional effects
  - Resistive wires, process variations, ...
- A model which is a natural extension of the existing models
  - Allows us to use existing models where possible
Conclusions

- Advent of SSTA is causing a re-examination of how cell delay models are generated.
  - Additional dependencies are required.
  - More accuracy is needed.

- Empirical enhancements are costly in development time.

- A data-driven approach which re-uses existing data to drive improvement has the best chance of success.