Problems in VLSI design

• wire and transistor sizing
  – signal delay in RC circuits
  – transistor and wire sizing
  – Elmore delay minimization via GP
  – dominant time constant minimization via SDP

• placement problems
  – quadratic and $\ell^1$-placement
  – placement with timing constraints
Signal delay in RC circuit

\[
C \frac{dv}{dt} = -G(v(t) - 1), \quad v(0) = 0
\]

- capacitance matrix \( C = C^T \succ 0 \)
- conductance matrix \( G = G^T \succ 0 \)
• $v$: node voltages

• as $t \to \infty$, $v(t) \to 1$

• delay at node $k$:

$$D_k = \inf\{T \mid v_k(t) \geq 0.5 \text{ for } t \geq T\}$$

• critical delay: $D = \max_k D_k$
Transistor sizing

RC model of transistor

\[ R_{sd} \propto \frac{1}{w} \]

\[ C_g \propto w \]
\[ C_s \propto w \]
\[ C_d \propto w \]

nMOS transistor (width \( w \))
example
• to first approximation: linear RC circuit

• design variable: transistor width $w$

• drain, source, gate capacitance affine in width

• ‘on’ resistance inversely proportional to width
**Wire sizing**

interconnect wires in IC: distributed RC line
lumped RC model:

\[ C_i \propto w_i \ell_i \]
\[ R_i \propto \frac{\ell_i}{w_i} \]

- replace each segment with \( \pi \) model
- segment capacitance proportional to width

Problems in VLSI design
• segment resistance inversely proportional to width

• design variables: wire segment widths $w_i$
Optimization problems involving delay

\[ C(x) \frac{dv}{dt} = -G(x)(v(t) - 1), \quad v(0) = 0 \]

- design parameters \( x \): transistor & wire segment widths
- capacitances, conductances are affine in \( x \):

\[
C(x) = C_0 + x_1 C_1 + \cdots + x_m C_m
\]

\[
G(x) = G_0 + x_1 G_1 + \cdots + x_m G_m
\]
tradeoff between

• delay, complicated function of \( x \)

• area, affine in \( x \)

• dissipated power in transition \( v(t) = 0 \rightarrow 1 \)

\[
\frac{1^T C(x) 1}{2}
\]

affine in \( x \)
Elmore delay

• area above step response

\[ T_{elm}^k = \int_0^\infty (1 - v_k(t)) dt \]

• first moment of impulse response

\[ T_{elm}^k = \int_0^\infty tv_k(t)' dt \]
\( T_k \) is mean, \( D_k \) is median

- \( T_k^{elm} \geq 0.5D_k \)
- good approximation of \( D_k \) only when \( v_k \) is monotonically increasing
- interpret \( v_k' \) as probability density:

\( T_k \) is mean, \( D_k \) is median
Elmore delay for RC tree

- one input voltage source
- resistors form a tree with root at voltage source
- all capacitors are grounded
Elmore delay to node $k$:

$$
\sum_i C_i \left( \sum R's \text{ upstream from node } k \text{ and node } i \right)
$$

Example:

$$
T_{3}^{\text{elm}} = C_3(R_1 + R_2 + R_3) + C_2(R_1 + R_2) + C_1R_1 \\
+ C_4R_1 + C_5R_1 + C_6R_1
$$
Elmore delay optimization via GP

in transistor & wire sizing, \( R_i = \alpha_i/x_i \), \( C_j = a_j^T x + b_j \)
\((\alpha_i \geq 0, a_j, b_j \geq 0)\)

Elmore delay:

\[
T_{elm}^k = \sum_{ij} \gamma_{ij} R_j C_i = \sum_k \beta_k \prod_{i=1}^m x_i^{\alpha_{ik}}
\]

\((\gamma_{ij} = +1 \text{ or } 0, \ \beta_k \geq 0, \alpha_{ij} = +1, 0, -1)\)

... a posynomial function of \( x > 0 \)

hence can minimize area or power, subject to bound on Elmore delay using geometric programming

commercial software (1980s): e.g., TILOS
Limitations of Elmore delay optimization

• not a good approximation of 50% delay when step response is not monotonic
  (capacitive coupling between nodes, or non-diagonal $C$)

• no useful convexity properties when
  – there are loops of resistors
  – circuit has multiple sources
  – resistances depend on more than one variable
Dominant time constant

\[ C(x) \frac{dv}{dt} = -G(x)(v(t) - 1), \quad v(0) = 0 \]

- eigenvalues \( 0 > \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \) given by
  \[ \det(\lambda_i C(x) + G(x)) = 0 \]

- solutions have form
  \[ v_k(t) = 1 - \sum_i \alpha_{ik} e^{\lambda_i t} \]

- slowest ("dominant") time constant given by \( T^{\text{dom}} = -1/\lambda_1 \) (related to delay)
• can bound $D$, $T^{elm}$ in terms of $T^{dom}$

• in practice, $T^{dom}$ is good approximation of $D$
Dominant time constant constraint as linear matrix inequality

**upper bound** $T_{\text{dom}} \leq T_{\text{max}}$

$T_{\text{dom}} \leq T_{\text{max}} \iff T_{\text{max}}G(x) - C(x) \succeq 0$

- convex constraint in $x$ (linear matrix inequality)
• no restrictions on $G, C$

• $T^{\text{dom}}$ is quasiconvex function of $x$, i.e., sublevel sets

$$\{ x \mid T^{\text{dom}}(x) \leq T_{\text{max}} \}$$

are convex
Sizing via semidefinite programming

minimize area, power s.t. bound on $T^{\text{dom}}$, upper and lower bounds on sizes

\[
\begin{align*}
\text{minimize} & \quad f^T x \\
\text{subject to} & \quad T_{\text{max}} G(x) - C(x) \succeq 0 \\
& \quad x_i^\text{min} \leq x_i \leq x_i^\text{max}
\end{align*}
\]

- a convex optimization problem (SDP)
- no restrictions on topology
  (loops of resistors, non-grounded capacitors)
Wire sizing

minimize wire area subject to

• bound on delay (dominant time constant)

• bounds on segments widths

RC-model:
as SDP:

\[
\begin{align*}
\text{minimize} & \quad \sum_i \ell_i x_i \\
\text{subject to} & \quad T_{\max}G(x) - C(x) \succeq 0 \\
& \quad 0 \leq x_i \leq 1
\end{align*}
\]
area-delay tradeoff

- globally optimal tradeoff curve
- optimal wire profile tapers off
step responses (solution (a))
Wire sizing and topology

not solvable via Elmore delay minimization

min area s.t. max dominant time constant (via SDP):

minimize $\sum x_i$

subject to $T_{\text{max}}G(x) - C(x) \succeq 0$
• usually have more wires than are needed
• solutions usually have some $x_i = 0$

• different points on tradeoff curve have different topologies
solution (a)

\[ x_4 = 0.15 \quad x_6 = 0.11 \]

\[ v_1(t) \]

\[ v_2(t) \]

\[ v_3(t) \]
solution (b)

\[ x_3 = 0.02 \]

\[ x_4 = 0.03 \quad x_6 = 0.03 \]
solution (c)

\[ x_3 = 0.014 \]
Placement

- list of cells: cells \( i = 1, \ldots, N \) are placeable, cells \( i = N + 1, \ldots, N + M \) are fixed (e.g., I/O)
- input and output terminals on boundary of cells
- group of terminals connected together is called a net
• placement of cells determines length of interconnect wires, hence signal delay

• problem: determine positions \((x_k, y_k)\) for the placeable cells to satisfy delay constraints

• practical problem sizes can involve 100,000s of cells

• exact solution (including delay, area, overlap constraints) is very hard to compute

• heuristics (often based on convex optimization) are widely used in practice
Quadratic placement

assume for simplicity:

- cells are points \((i.e.,\) have zero area)
- nets connect two terminals \((i.e.,\) are simple wires)

**quadratic placement:**

\[
\text{minimize} \sum_{\text{nets } (i,j)} w_{ij} \left( (x_i - x_j)^2 + (y_i - y_j)^2 \right)
\]

weights \(w_{ij} \geq 0\)

unconstrained convex quadratic minimization
(called ‘quadratic programming’ in VLSI)
• solved using CG (and related methods) exploiting problem structure (e.g., sparsity)

• physical interpretation: wires are linear elastic springs

• widely used in industry

• constraints handled using heuristics (e.g., adjusting weights)
$\ell^1$-placement

minimize $\sum_{\text{nets } (i,j)} w_{ij} (|x_i - x_j| + |y_i - y_j|)$

- measures wire length using Manhattan distance (wire routing is horizontal/vertical)

- motivation: delay of wire $(i, j)$ is $RC$ with

\[ R = R_{\text{driver}} + R_{\text{wire}}, \quad C = C_{\text{wire}} + C_{\text{load}} \]

$R_{\text{driver}}, C_{\text{load}}$ are given, $R_{\text{wire}} \ll R_{\text{driver}},$

$C_{\text{wire}} \propto \text{wire length (Manhattan)}$
• called ‘linear objective’ in VLSI
Nonlinear spring models

\[
\text{minimize } \sum_{\text{nets } (i,j)} h(|x_i - x_j| + |y_i - y_j|)
\]

\(h\) convex, increasing on \(\mathbb{R}_+\)

example

- flat part avoids ‘clustering’ of cells
• quadratic part: for long wires $R_{\text{wire}} \propto \text{length}$

• solved via convex programming
Timing constraints

• cell $i$ has a processing delay $D_{i}^{\text{proc}}$

• propagation delay through wire $(i, j)$ is $\alpha l_{ij}$, where $l_{ij}$ is the length of the wire

• minimize max delay from any input to any output
problem is:

$$\text{minimize} \quad T$$

subject to

$$\sum_{\text{cells \ in \ path}} D^\text{proc}_i + \sum_{\text{wires \ in \ path}} \alpha \ell_{ij} \leq T$$

• one constraint for each path

• variables: $T$, positions of placeable cells (which determine $\ell_{ij}$)

• a very large number of inequalities
A more compact representation

• introduce new variable $T_{i}^{\text{out}}$ for each cell

• for all cells $j$, add one inequality for each cell $i$ in the fan-in of $j$

\[
T_{i}^{\text{out}} + \alpha l_{ij} + D_{j}^{\text{proc}} \leq T_{j}^{\text{out}} \tag{1}
\]

• for all output cells

\[
T_{i}^{\text{out}} \leq T \tag{2}
\]

• minimize $T$ subject to (??) and (??)

convex optimization problem:

• with $\ell^1$-norm, get LP

• with $\ell^2$-norm, get SOCP
extensions (still convex optimization):

- delay is convex, increasing fct of wire length
- max delay constraints on intermediate cells
- different delay constraints on cells
Non-convex constraints and generalizations

non-convex constraints

• cells are placed on grid of legal positions

• cells are rectangles that cannot overlap

• reserved regions on chip
generalizations

• multi-pin nets: share interconnect wires

• combine placement with wire and gate sizing