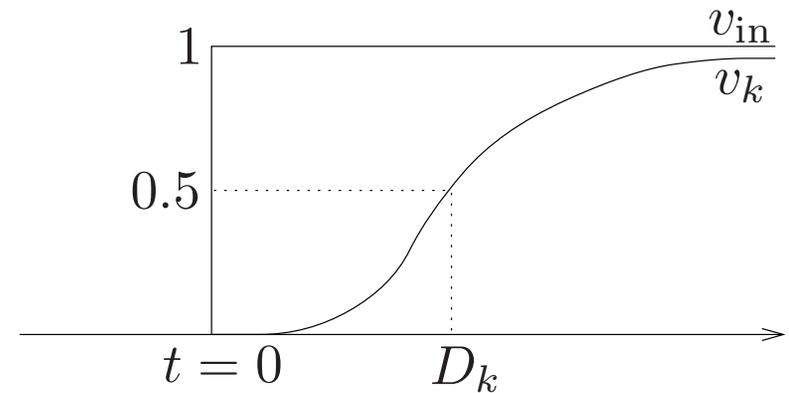
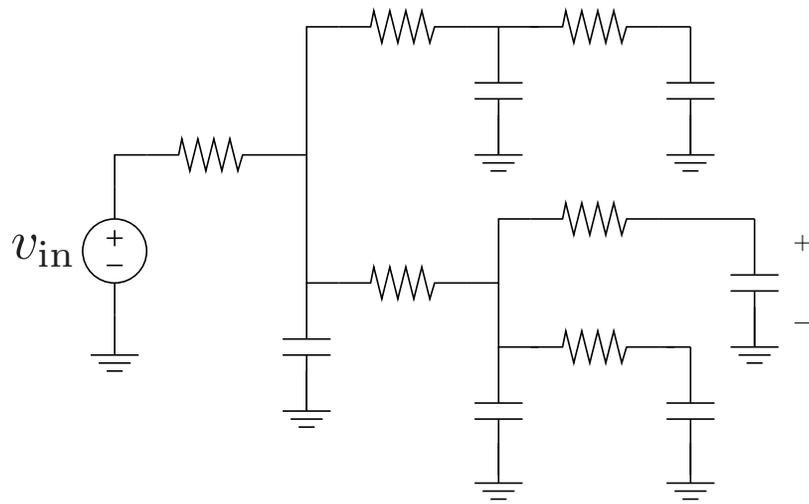


Problems in VLSI design

- 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 ℓ^1 -placement
 - placement with timing constraints

Signal delay in RC circuit



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

- capacitance matrix $C = C^T \succ 0$
- conductance matrix $G = G^T \succ 0$

- v : node voltages
- as $t \rightarrow \infty$, $v(t) \rightarrow \mathbf{1}$

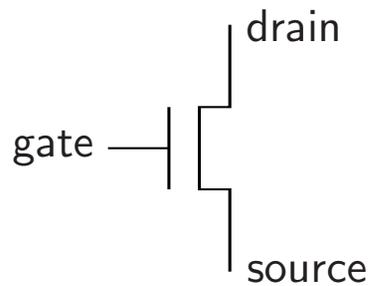
- delay at node k :

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

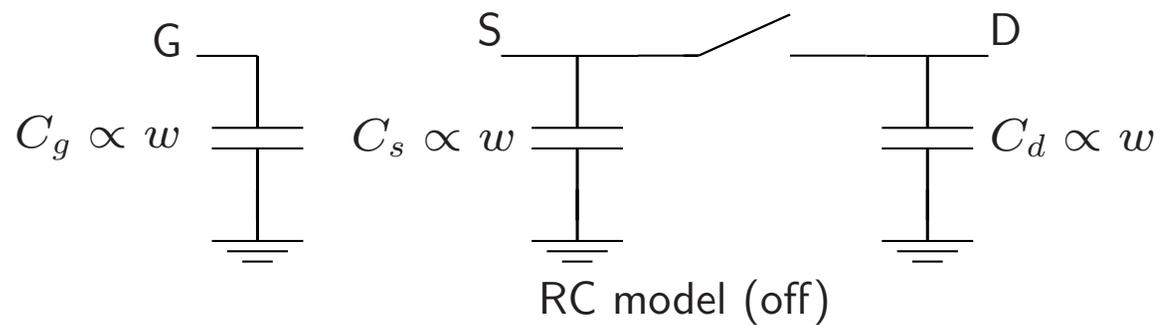
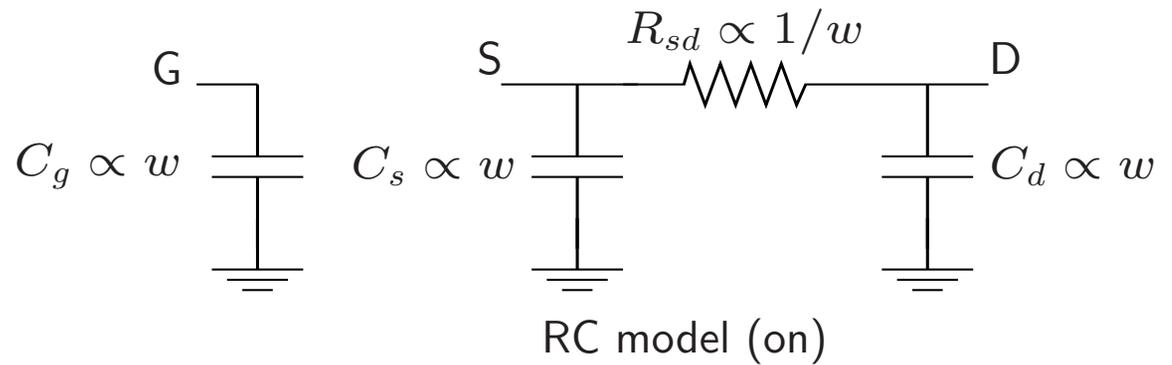
- critical delay: $D = \max_k D_k$

Transistor sizing

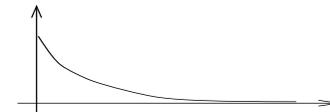
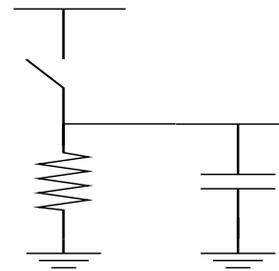
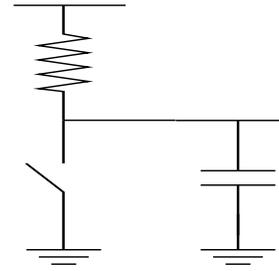
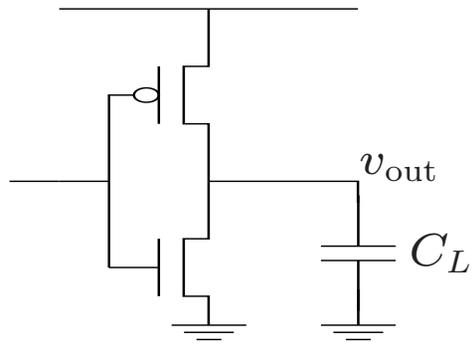
RC model of transistor



nMOS
transistor
(width w)



example

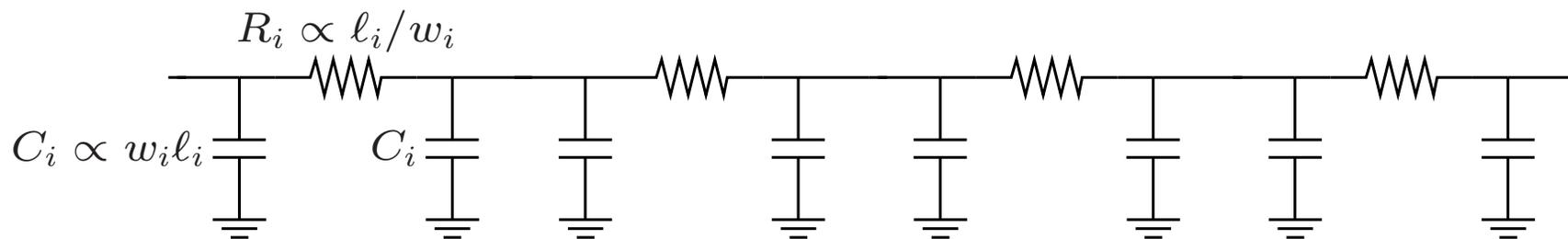
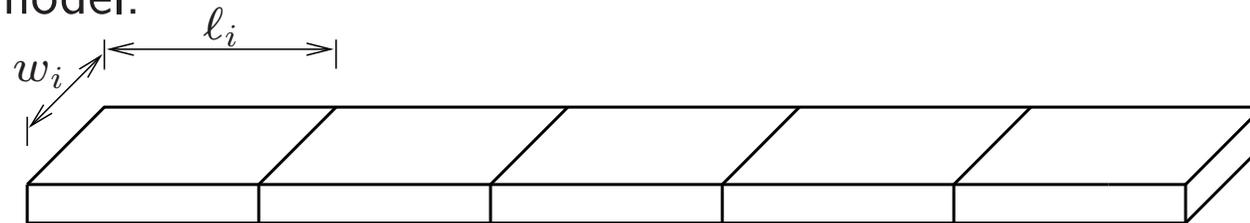


- to first approximation: linear RC circuit
- design variable: transistor width w
- drain, source, gate capacitance affine in width
- 'on' resistance inversely proportional to width

Wire sizing

interconnect wires in IC: distributed RC line

lumped RC model:



- replace each segment with π model
- segment capacitance proportional to width

- segment resistance inversely proportional to width
- design variables: wire segment widths w_i

Optimization problems involving delay

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

- 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 $v(t) = 0 \rightarrow \mathbf{1}$

$$\frac{\mathbf{1}^T C(x) \mathbf{1}}{2}$$

affine in x

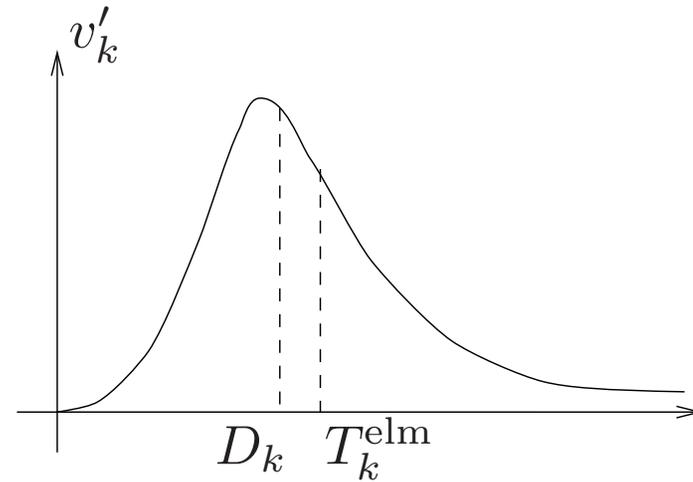
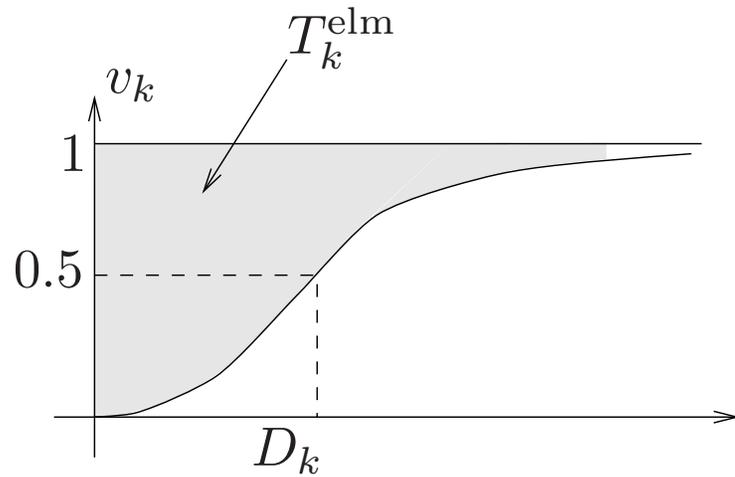
Elmore delay

- area above step response

$$T_k^{\text{elm}} = \int_0^{\infty} (1 - v_k(t)) dt$$

- first moment of impulse response

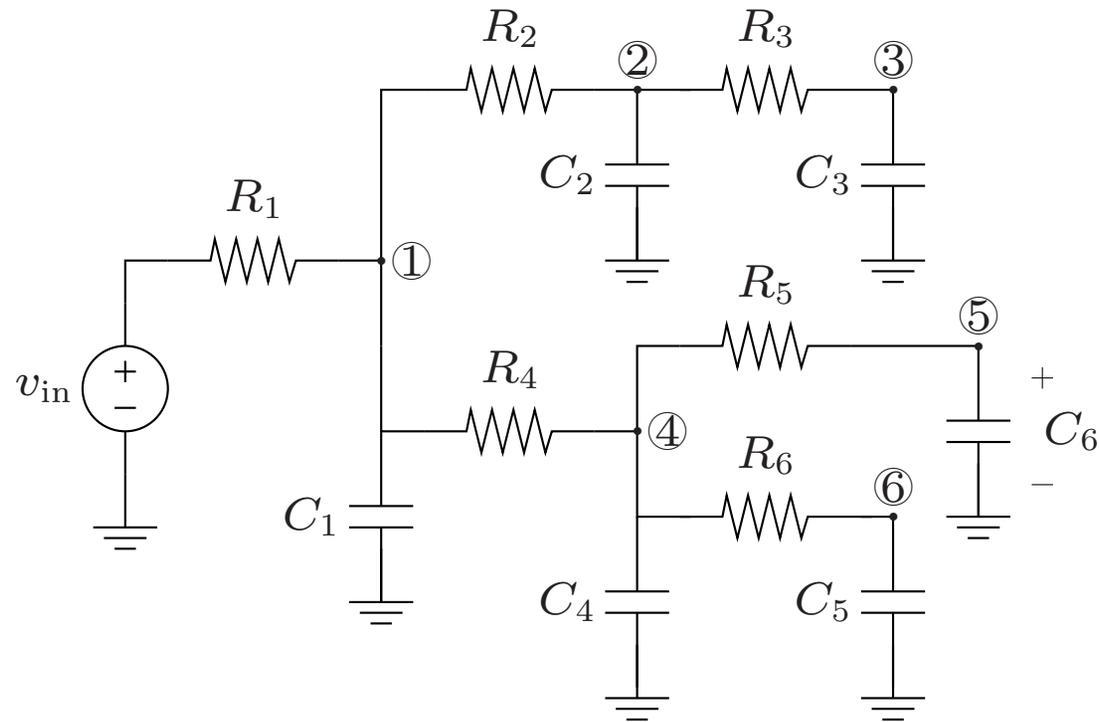
$$T_k^{\text{elm}} = \int_0^{\infty} t v_k(t)' dt$$



- $T_k^{\text{elm}} \geq 0.5D_k$
- 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

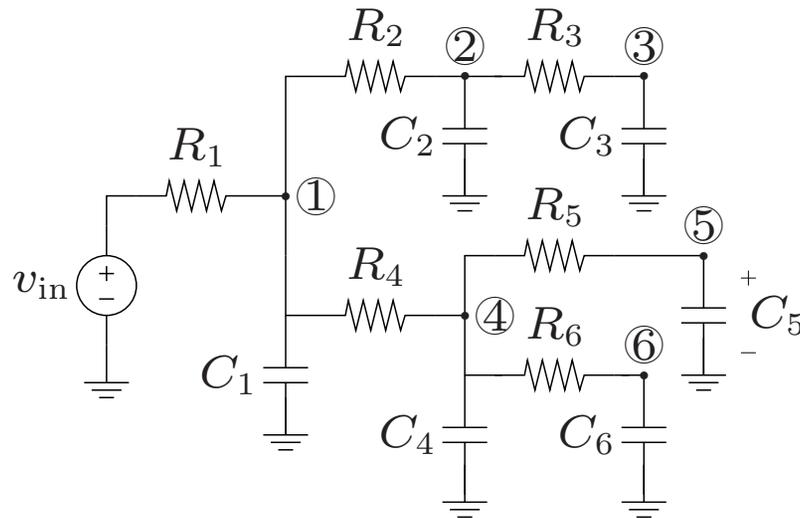
RC tree



- one input voltage source
- resistors form a tree with root at voltage source
- all capacitors are grounded

Elmore delay to node k :

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



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:

$$T_k^{\text{elm}} = \sum_{ij} \gamma_{ij} R_j C_i = \sum_{k=1} \beta_k \prod_{i=1}^m x_i^{\alpha_{ik}}$$

($\gamma_{ij} = +1$ or 0 , $\beta_k \geq 0$, $\alpha_{ij} = +1, 0, -1$)
... a *posynomial* function of $x \succ 0$

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

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

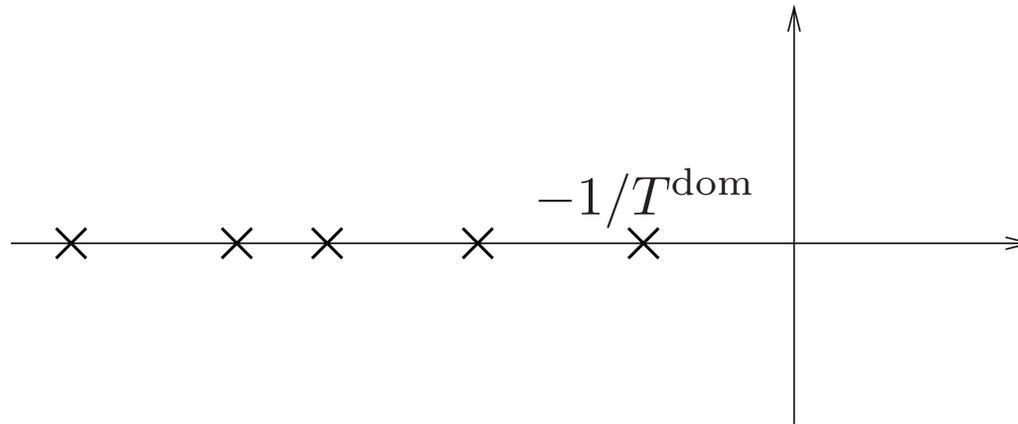
- eigenvalues $0 > \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ given by

$$\det(\lambda_i C(x) + G(x)) = 0$$

- solutions have form

$$v_k(t) = 1 - \sum_i \alpha_{ik} e^{\lambda_i t}$$

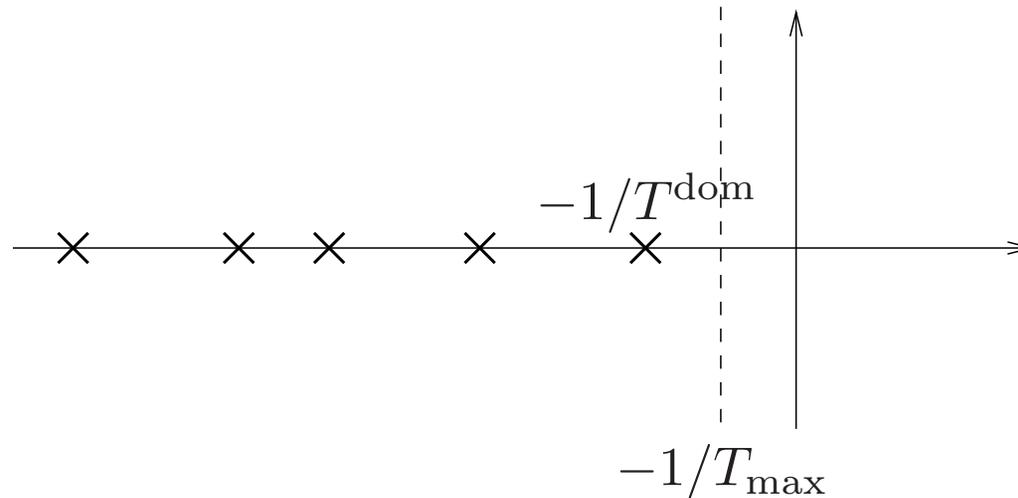
- slowest (“dominant”) time constant given by $T^{\text{dom}} = -1/\lambda_1$ (related to delay)



- 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^{dom} is quasiconvex function of x , *i.e.*, sublevel sets

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

are convex

Sizing via semidefinite programming

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

$$\begin{aligned} &\text{minimize} && f^T x \\ &\text{subject to} && T_{\max} G(x) - C(x) \succeq 0 \\ &&& x_i^{\min} \leq x_i \leq x_i^{\max} \end{aligned}$$

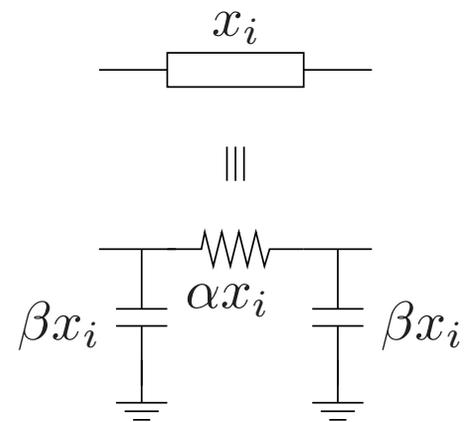
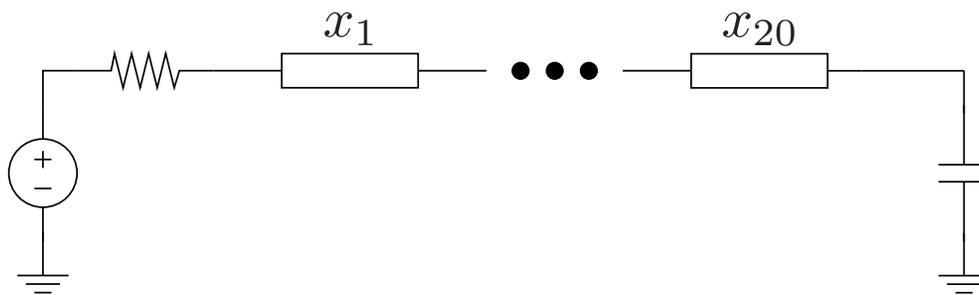
- a convex optimization problem (SDP)
- no restrictions on topology
(loops of resistors, non-grounded capacitors)

Wire sizing

minimize wire area subject to

- bound on delay (dominant time constant)
- bounds on segments widths

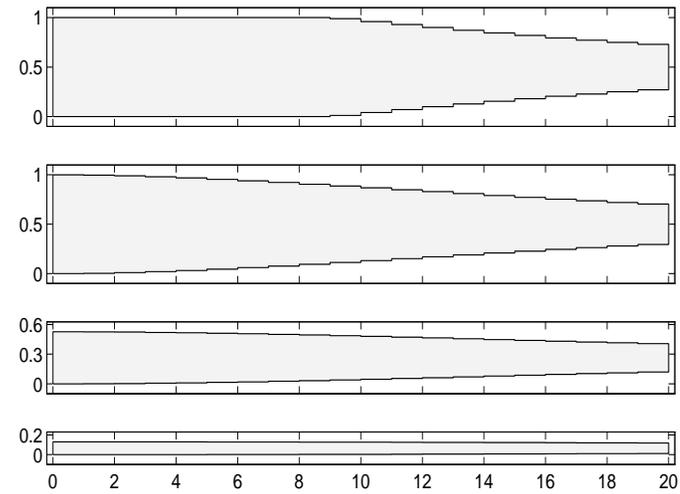
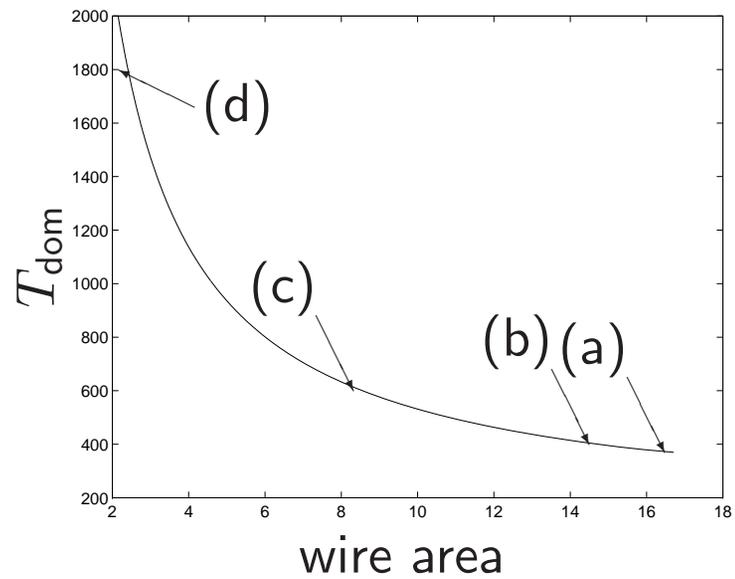
RC-model:



as SDP:

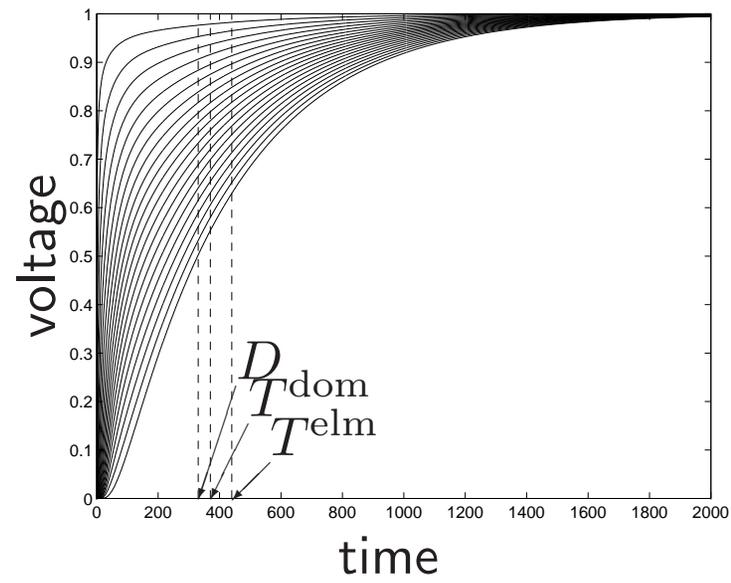
$$\begin{aligned} &\text{minimize} && \sum_i \ell_i x_i \\ &\text{subject to} && T_{\max} G(x) - C(x) \succeq 0 \\ &&& 0 \leq x_i \leq 1 \end{aligned}$$

area-delay tradeoff

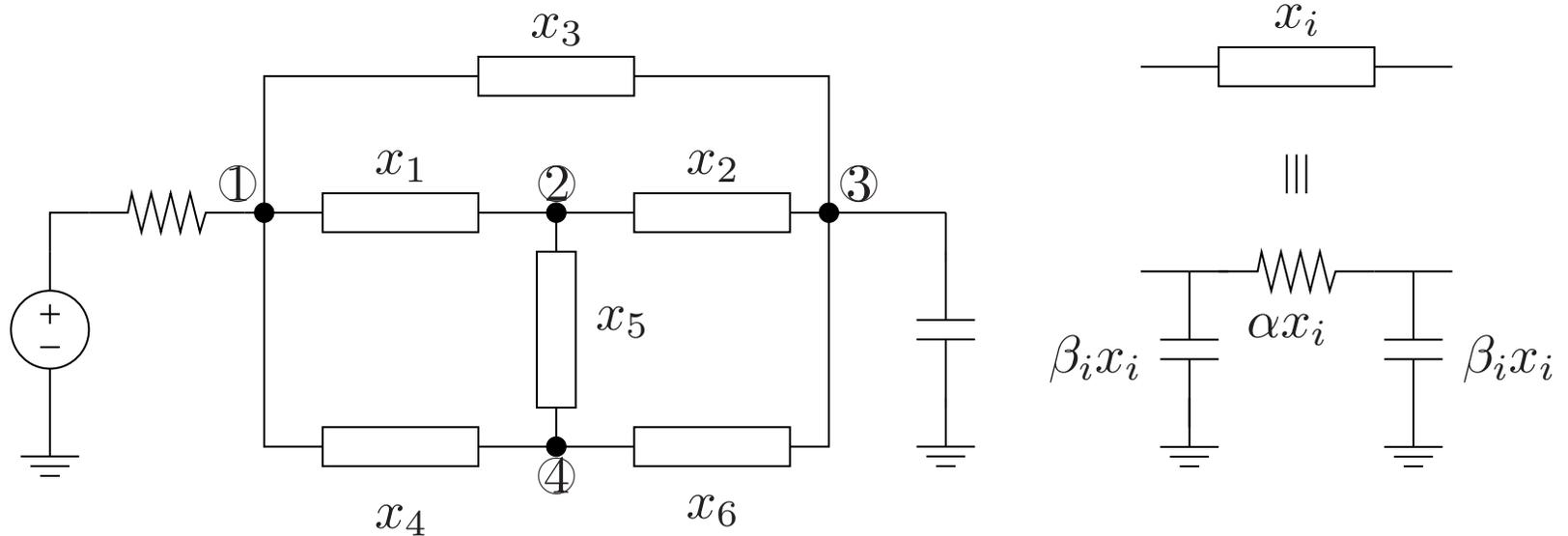


- globally optimal tradeoff curve
- optimal wire profile tapers off

step responses (solution (a))



Wire sizing and topology

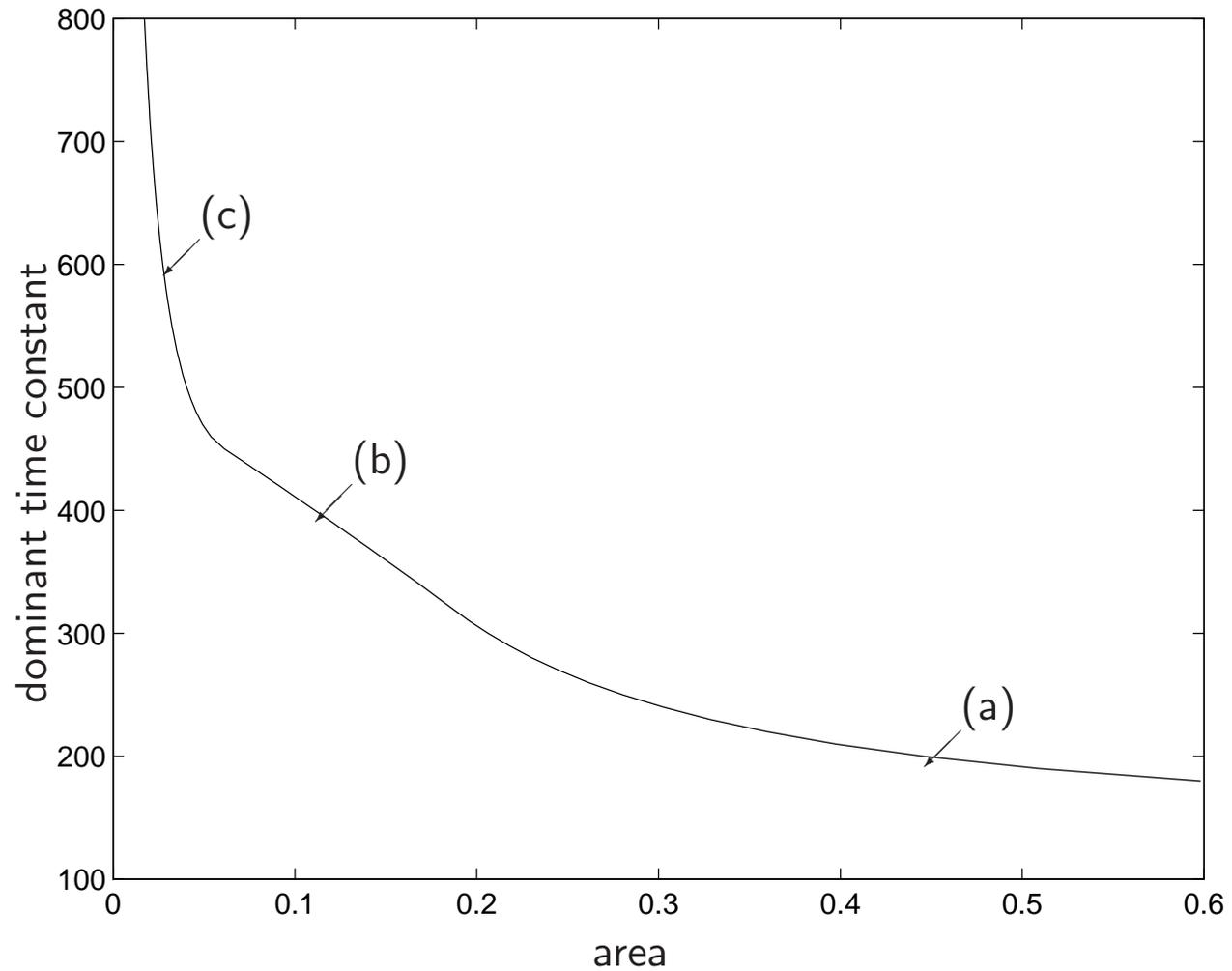


not solvable via Elmore delay minimization

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

$$\begin{aligned} & \text{minimize} && \sum x_i \\ & \text{subject to} && T_{\max} G(x) - C(x) \succeq 0 \end{aligned}$$

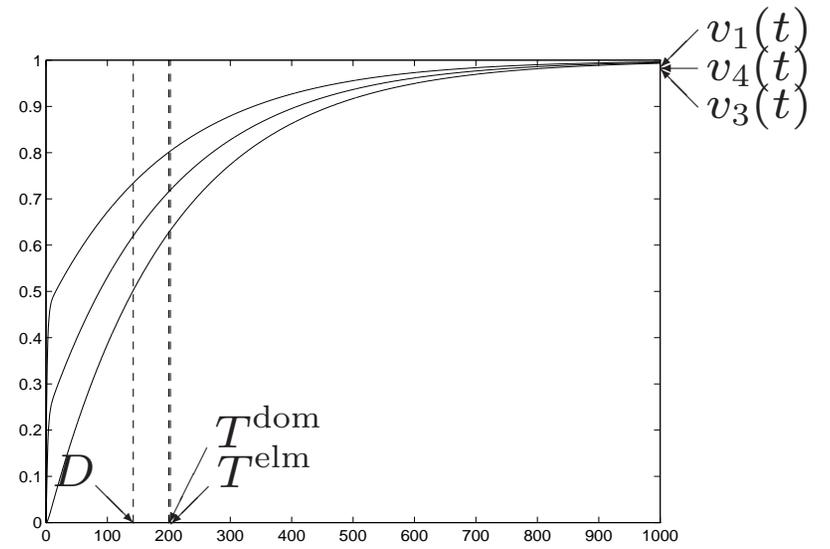
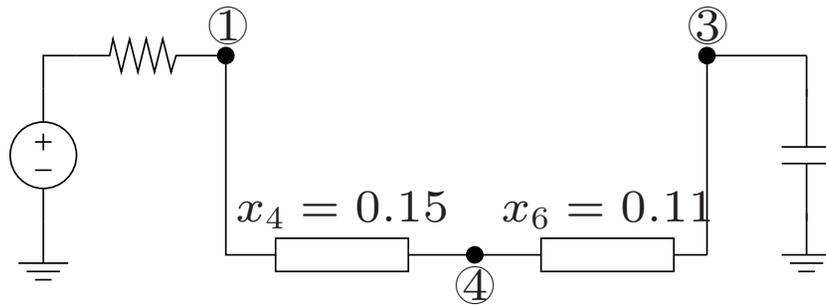
tradeoff curve



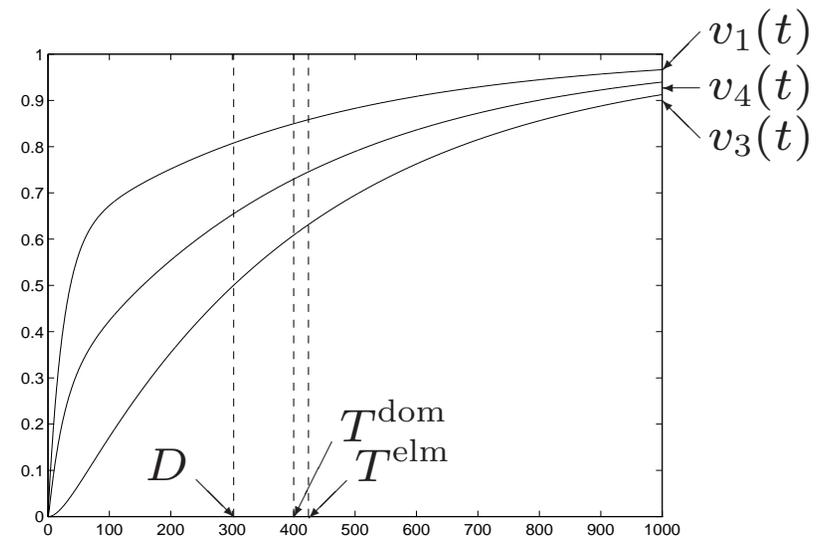
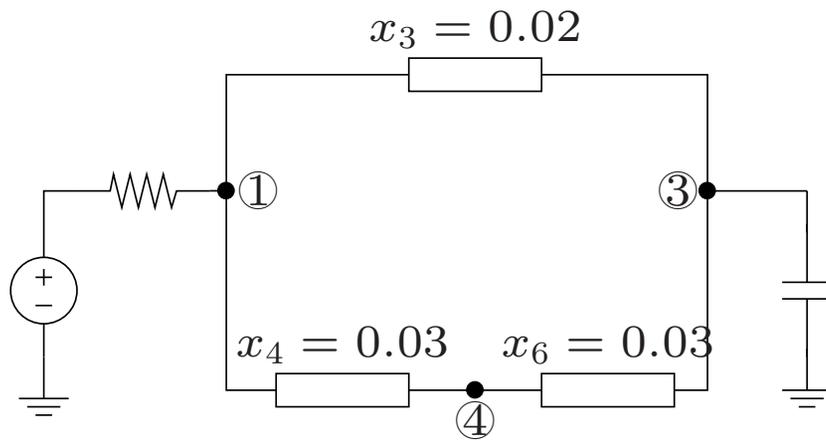
- usually have more wires than are needed

- solutions usually have some $x_i = 0$
- different points on tradeoff curve have different topologies

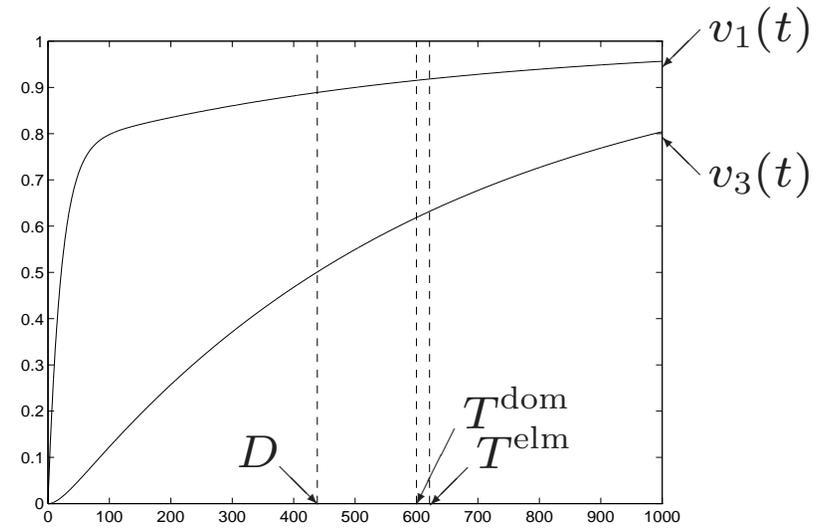
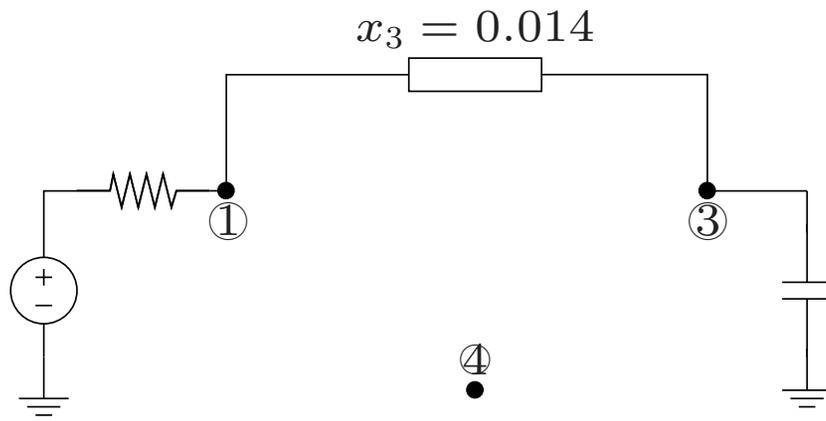
solution (a)



solution (b)

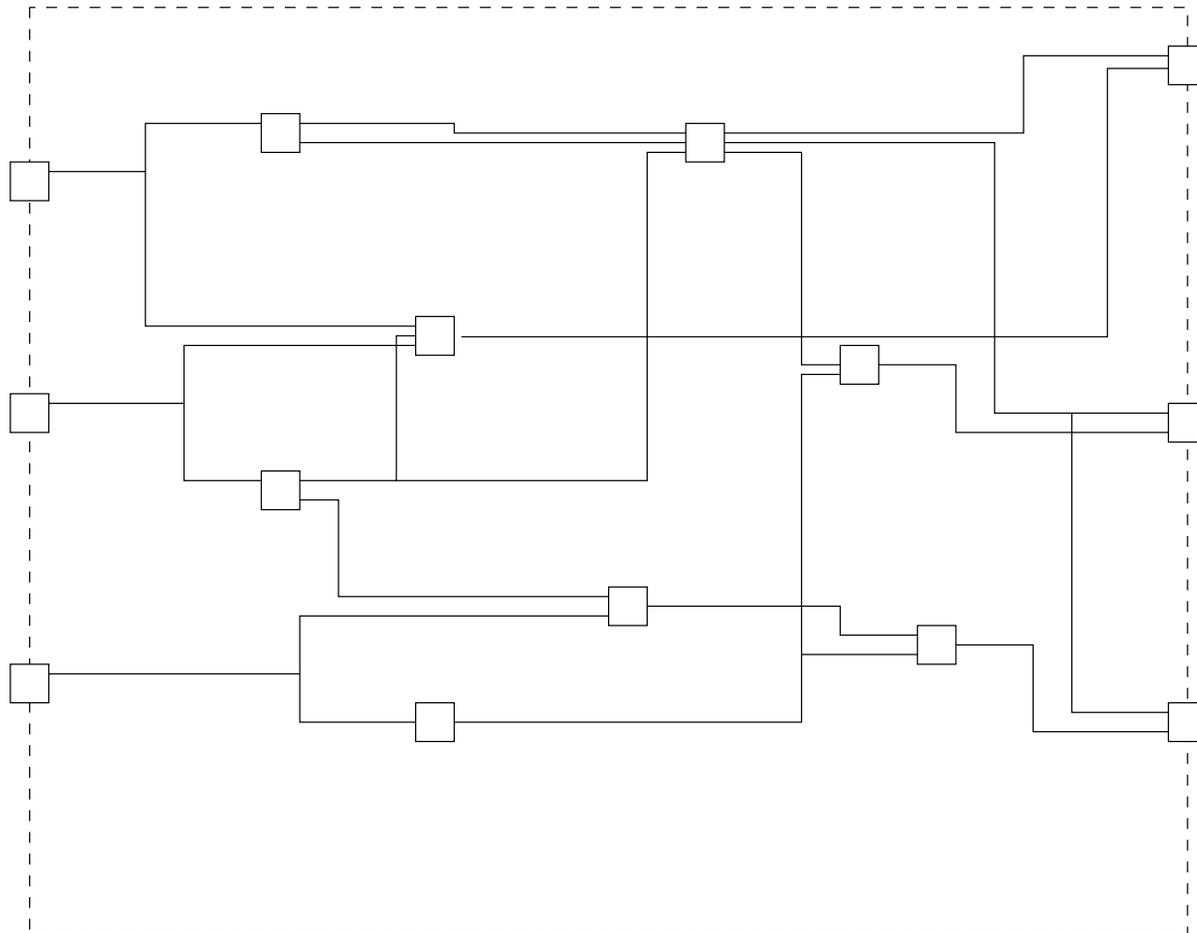


solution (c)



Placement

- list of cells: cells $i = 1, \dots, N$ are placeable, cells $i = N + 1, \dots, N + M$ 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:

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

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)

ℓ^1 -placement

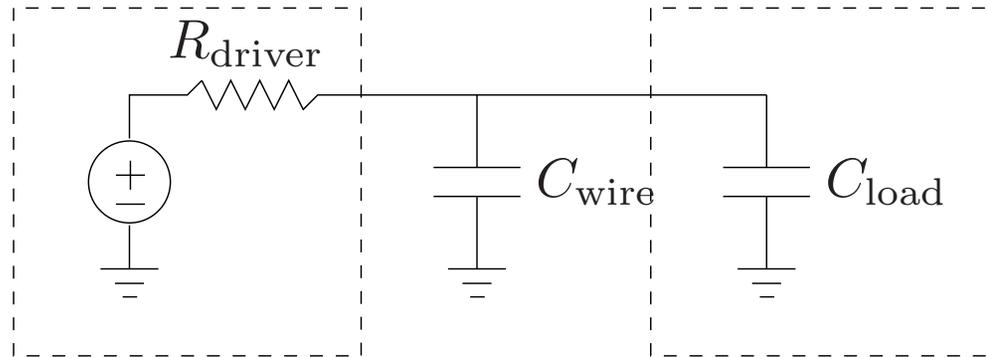
$$\text{minimize} \quad \sum_{\text{nets } (i,j)} w_{ij} (|x_i - x_j| + |y_i - y_j|)$$

- 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, $R_{\text{wire}} \ll R_{\text{driver}},$

$$C_{\text{wire}} \propto \text{wire length (Manhattan)}$$



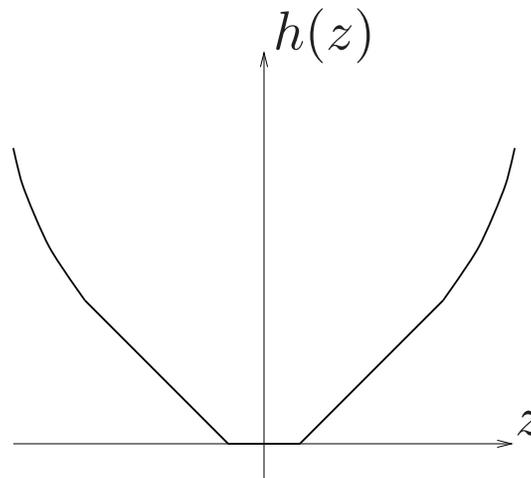
- called 'linear objective' in VLSI

Nonlinear spring models

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

h convex, increasing on \mathbf{R}_+

example

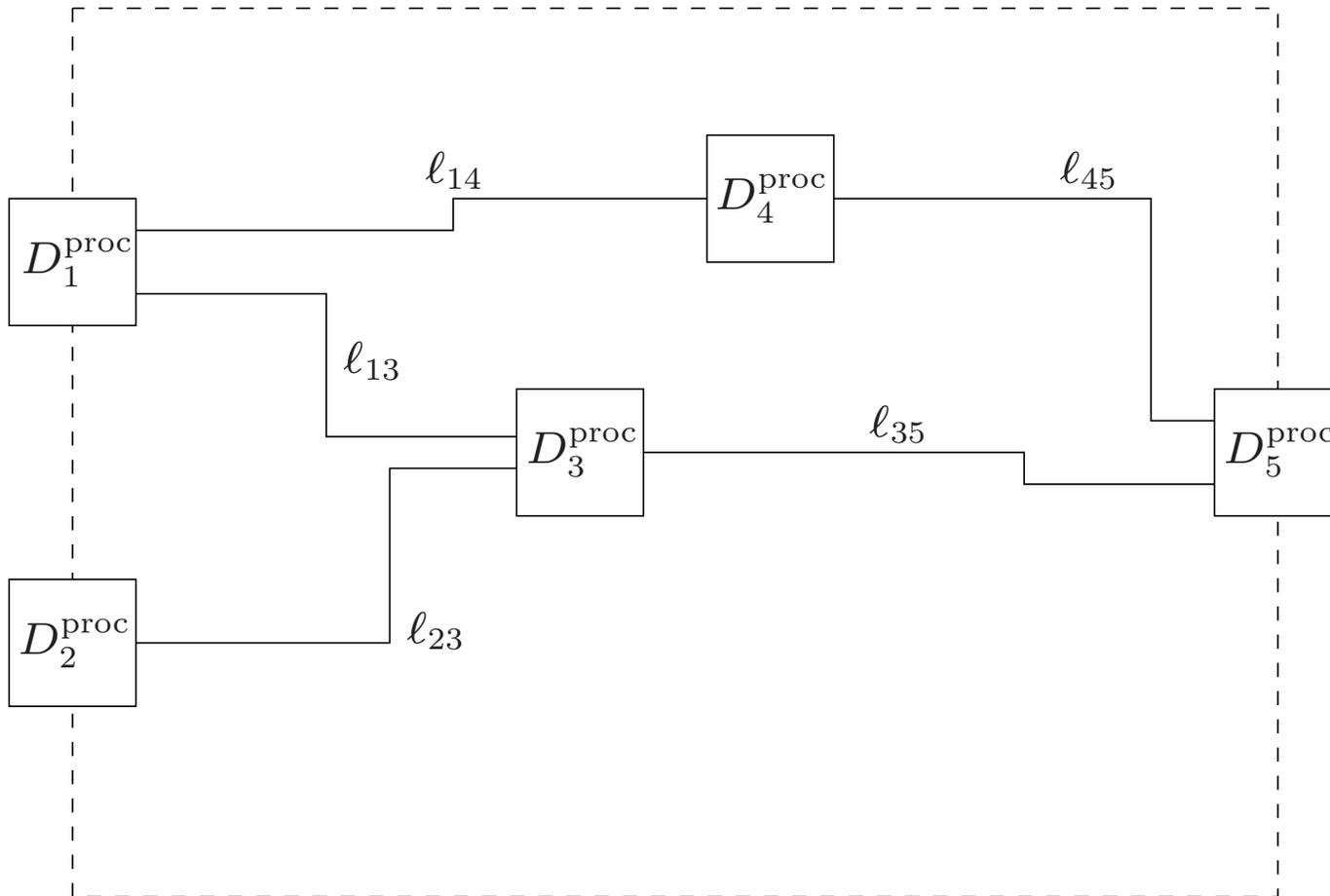


- flat part avoids 'clustering' of cells

- 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^{proc}
- propagation delay through wire (i, j) is αl_{ij} , where l_{ij} is the length of the wire
- minimize max delay from any input to any output



problem is:

$$\begin{array}{l} \text{minimize} \\ \text{subject to} \end{array} \quad T \sum_{\substack{\text{cells} \\ \text{in path}}} D_i^{\text{proc}} + \sum_{\substack{\text{wires} \\ \text{in path}}} \alpha l_{ij} \leq T$$

- one constraint for each path
- variables: T , positions of placeable cells (which determine l_{ij})
- a *very* large number of inequalities

A more compact representation

- introduce new variable T_i^{out} for each cell
- for all cells j , add one inequality for each cell i in the fan-in of j

$$T_i^{\text{out}} + \alpha l_{ij} + D_j^{\text{proc}} \leq T_j^{\text{out}} \quad (1)$$

- for all output cells

$$T_i^{\text{out}} \leq T \quad (2)$$

- minimize T subject to (??) and (??)

convex optimization problem:

- with ℓ^1 -norm, get LP
- with ℓ^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

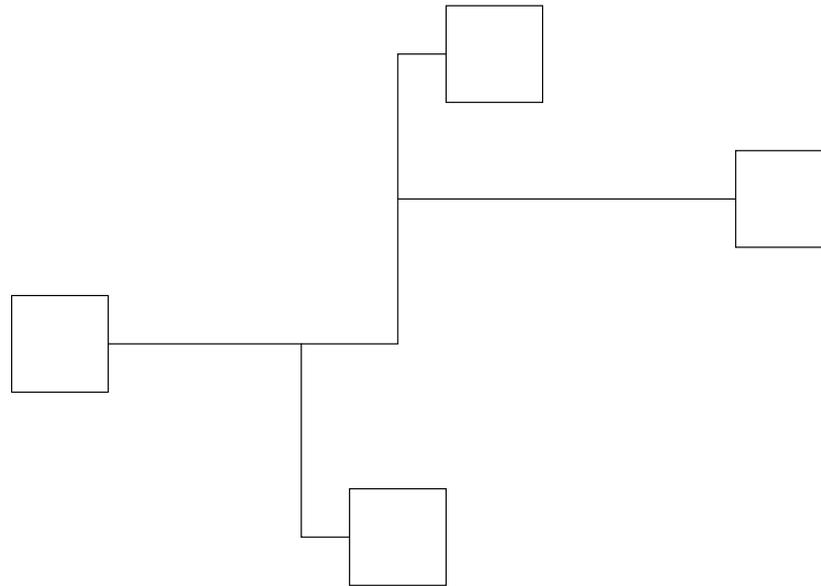
Non-convex constraints and generalizations

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