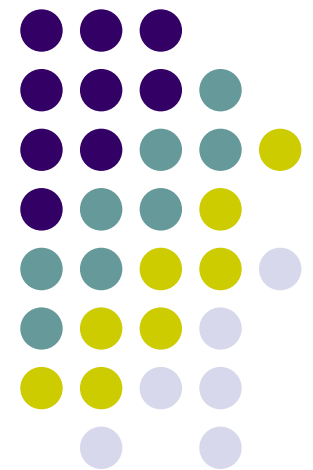


# CS159 Introduction to Computational Complexity

---

## The VLSI Model II





# The VLSI Model

- **Architectural Model:**

- Chips realize FSMs
- Wires are rectilinear.
- Wires have bounded width and separation  $\lambda$ .
- Gates can be binary or non-binary.
- Gates/memory cells occupy area proportional to  $\lambda^2$ .
- I/O pads also occupy area proportional to  $\lambda^2$ .
- Wires on at most  $v \geq 1$  levels. Gates on one level.
- Time for signal to travel wire of length  $l$  is constant.
- (Time for *diffusion model* is proportional to  $l^2$ !)



# The VLSI Model

- **Performance Measures:**

- Chip area  $A$  - manufacturing cost increases with  $A$
- Number of steps  $T$  - reflects computational cost



# The VLSI Model

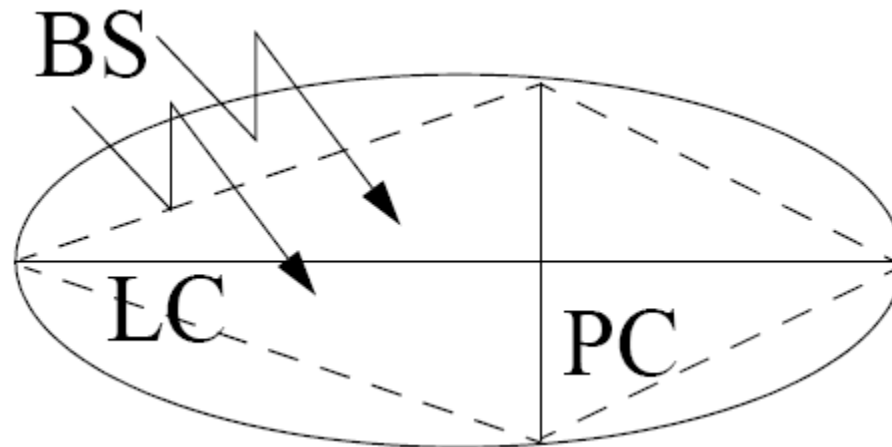
- **Functional Model:** Finite functions computed  $f : B^n \rightarrow B^m$  where  $B$  is any set, usually  $\{0,1\}$ .
- **Algorithmic Model:** Limitations on I/O
  - Inputs provided at times and places on the chip that are data-independent.
  - Each input is supplied once (*semilective alg.*) or provided multiple times (*multilective alg.*).
  - Outputs are produced once at times and places on the chip that are data-independent.

# Basic Area-Time Tradeoffs

## Ideas



- Convex chip boundary, feature size =  $\lambda$
- LC = longest chord, PC = chord  $\perp$  to LC divides inputs into two equal size sets. One side, BS, has at least half of outputs.



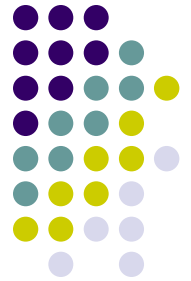
# Basic Area-Time Tradeoffs

## Ideas



- Find amount of information,  $I$ , that must move to outputs in BS from inputs on other side.
- At most  $|PC|/2\lambda$  wires cross PC.
- Number of steps  $T \geq I / (|PC|/2\lambda)$ .
- Area  $A \geq |PC||LC|/2 \geq |PC|^2/2$ . Thus,

$$AT^2 \geq 2\lambda^2 I^2$$

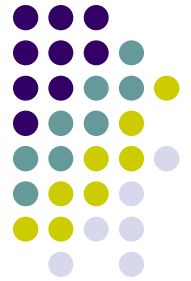


# Flow Properties of Functions

**Definition**  $f : A^n \rightarrow A^m$  is  $(\alpha, n, m, p)$ -**independent** for  $\alpha \geq 1$  and  $p \leq m$  if it has a  $w(u, v)$ -flow satisfying  $w(u, v) > (v/\alpha) - 1$  for  $n - u + v \leq p$ .

**Note:** If  $|X_1| = u$  and  $|Y_1| = v$ ,  $|X_0| + |Y_1| = n - u + v$ .

# Lower Bounds on Planar Circuit Size



If  $f$  on  $n$  inputs and  $m$  outputs is  $(\alpha, n, m, p)$ -ind. for  $p \leq m$  and  $\alpha \geq 1$ , there are  $> |A|^{(v/\alpha)-1}$  points in the image of its domain when  $\leq p - v$  inputs are fixed.

<i>Name</i>	<i>In's (n)</i>	<i>Out's (m)</i>	<i>Ind. Property</i>
Wrapped conv	$2n$	$n$	$(2, 2n, n, n/2)$
Cyclic shift	$n + \log n$	$n$	$(2, n + \log n, n, n/2)$
Int. multiply	$2n$	$2n$	$(2, 2n, n, n/2)$
$n$ -point FFT	$n$	$n$	$(2, 2n, n, n/2)$



# Area-Time Lower Bounds

**Theorem** Area  $A$  and time  $T$  required to compute wrapped convolution and cyclic shift of  $n$  inputs, integer multiplication of  $n$ -bit integers, or the FFT on  $n$  inputs on semiregular VLSI chip satisfy following:

$$AT^2, A^2T = \Omega(n^2)$$

$AT^2$  lower bound can be achieved up to a constant multiplicative factor for each of these functions for  $\Omega(\log n) \leq T \leq \sqrt{n}$ .



# Area-Time Lower Bounds

**Proof** Lower bound follows from circuit lower bounds.

From previous lecture, for any fully normal algorithm  $AT^2 = O(n^2)$  for  $\Omega(\log n) \leq T \leq \sqrt{n}$  on an embedded CCC network. Since cyclic shift and FFT are shown to be fully normal, we have matching upper and lower bounds for them.

From Prob. 12.13 wrapped convolution can be realized with matching bounds on  $AT^2$  over the same range of values for  $T$ . Same applies to integer multiplication. (Prob. 12.16.)



# Comments

The planar circuit size for cyclic shifting is  $\Omega(n^2)$  but the function can be realized with  $O(n \log n)$  gates.

Since  $AT^2 = \Omega(n^2)$ , wires occupy more area than gates when  $T = O(\sqrt{n/\log n})$ .

Matrix multiplication lower bound requires more refined analysis.

# Area-Time Bounds for Matrix Multiplication

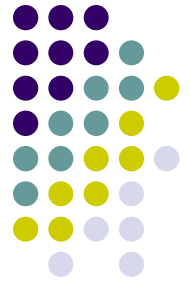


**Theorem** The area  $A$  and time  $T$  required to multiply two  $n \times n$  matrices with a semirecursive algorithm satisfies the following lower bound:

$$AT^2, A^2T = \Omega(n^4)$$

The  $AT^2$  lower bound can be met to within a constant multiplicative factor by chip seen in Lecture 1.

# Area-Time Bounds for Matrix Multiplication



**Proof** Apply Theorem on p. 6 to matrix multiplication by replacing the number of input variables  $n$  by  $2n^2$  and the number of output variables  $m$  by  $n^2$ . The  $w(u, v)$ -flow function has value

$$w(u, v) = (v - (2n^2 - u)^2 / 4n^2) / 2$$
$$w(u, v) \geq (n^2 / 2) (1 / (3P) - (3 / (2P))^2)$$

The r.h.s. is maximized when  $P = 14$  and the coefficient of  $n^2$  has value greater than  $1/163$ .