POWER ANALYSIS AND OPTIMIZATION

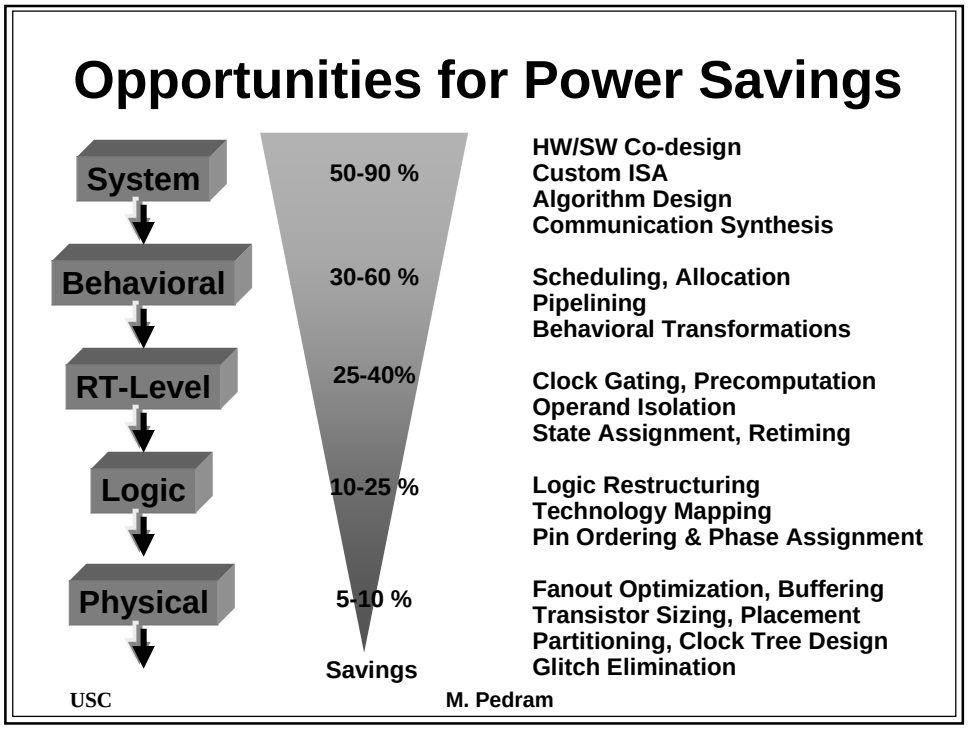
Massoud Pedram
University of Southern California
Department of EE-Systems
Los Angeles CA 90089

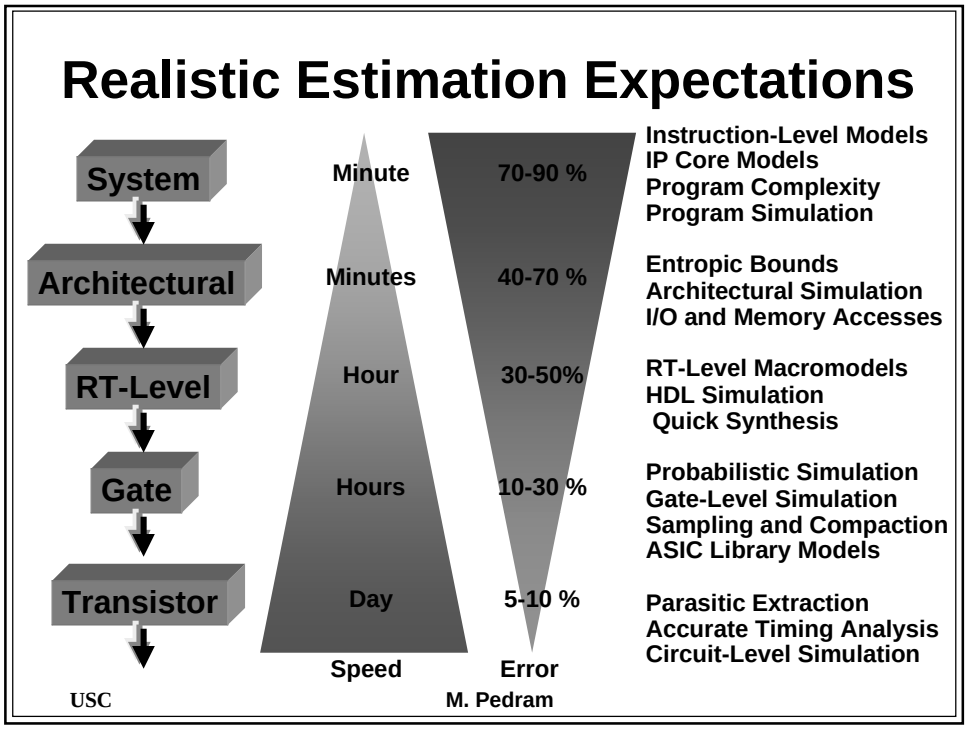
USC M. Pedram

Outline

- Motivation and Objectives
- Power Estimation Methodology
- Example Analysis/Estimation Tools
- Power Optimization Flow
- Example Minimization Techniques
- Summary

USC M. Pedram


Opportunities for Power Savings
System
Behavioral
RT-Level
Logic
Physical
50-90 %
30-60 %
25-40%
10-25 %
5-10 %
Savings
HW/SW Co-design
Custom ISA
Algorithm Design
Communication Synthesis
Scheduling, Allocation
Pipelining
Behavioral Transformations
Clock Gating, Precomputation
Operand Isolation
State Assignment, Retiming
Logic Restructuring
Technology Mapping
Pin Ordering & Phase Assignment
Fanout Optimization, Buffering
Transistor Sizing, Placement
Partitioning, Clock Tree Design
Glitch Elimination
USC
M. Pedram

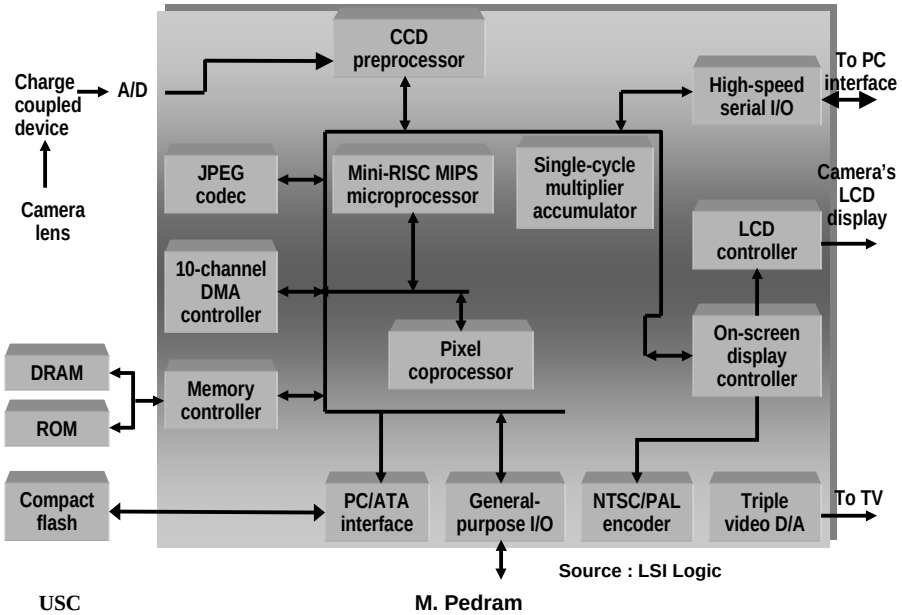

Realistic Estimation Expectations
System
Architectural
RT-Level
Gate
Transistor
Minute
Minutes
Hour
Hours
Day
Speed
70-90 %
40-70 %
30-50%
10-30 %
5-10 %
Error
Instruction-Level Models
IP Core Models
Program Complexity
Program Simulation
Entropic Bounds
Architectural Simulation
I/O and Memory Accesses
RT-Level Macromodels
HDL Simulation
Quick Synthesis
Probabilistic Simulation
Gate-Level Simulation
Sampling and Compaction
ASIC Library Models
Parasitic Extraction
Accurate Timing Analysis
Circuit-Level Simulation
USC
M. Pedram

Example Applications

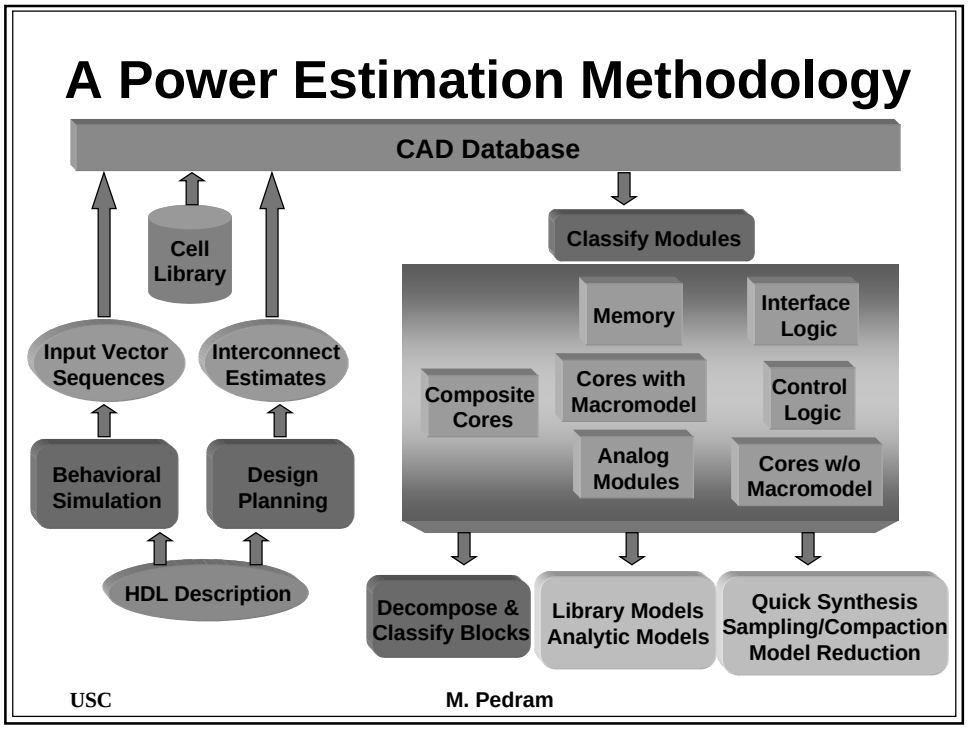
- **Portable Electronics (PC, PDA, Wireless)**
- **Ultra-Low-Power Circuits (Pacemaker)**
- **Space Missions (Miniaturized Satellites)**
- **IC Cost (Packaging and Cooling)**
- **Reliability (Electromigration, Latch-up)**
- **Signal Integrity (Switching Noise, DC Voltage Drop)**
- **Thermal Design**

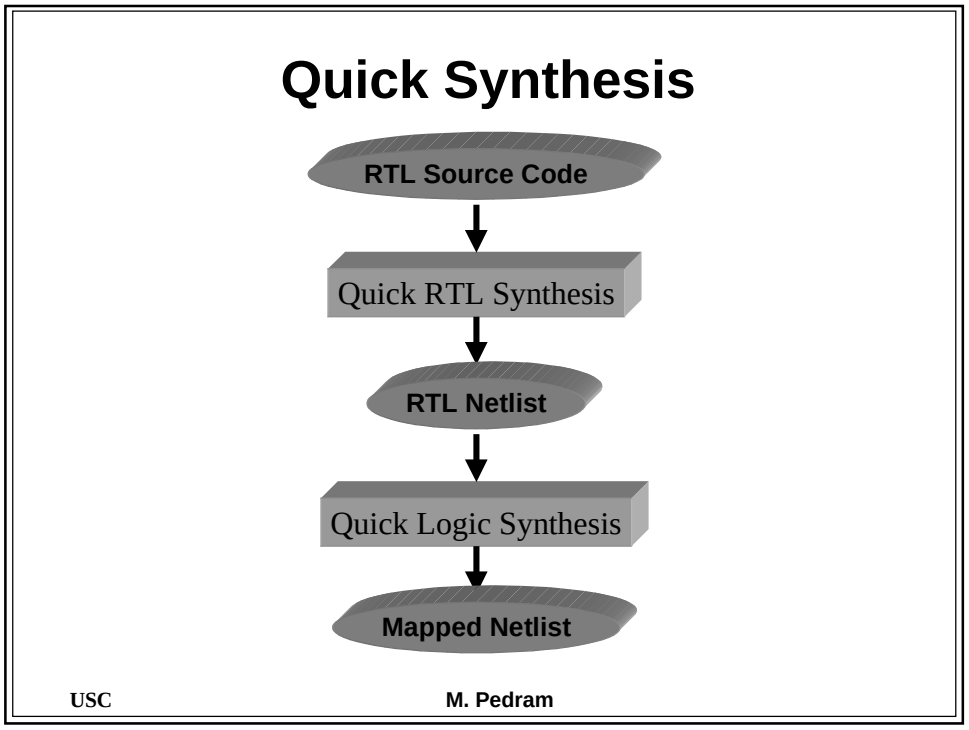
USC M. Pedram

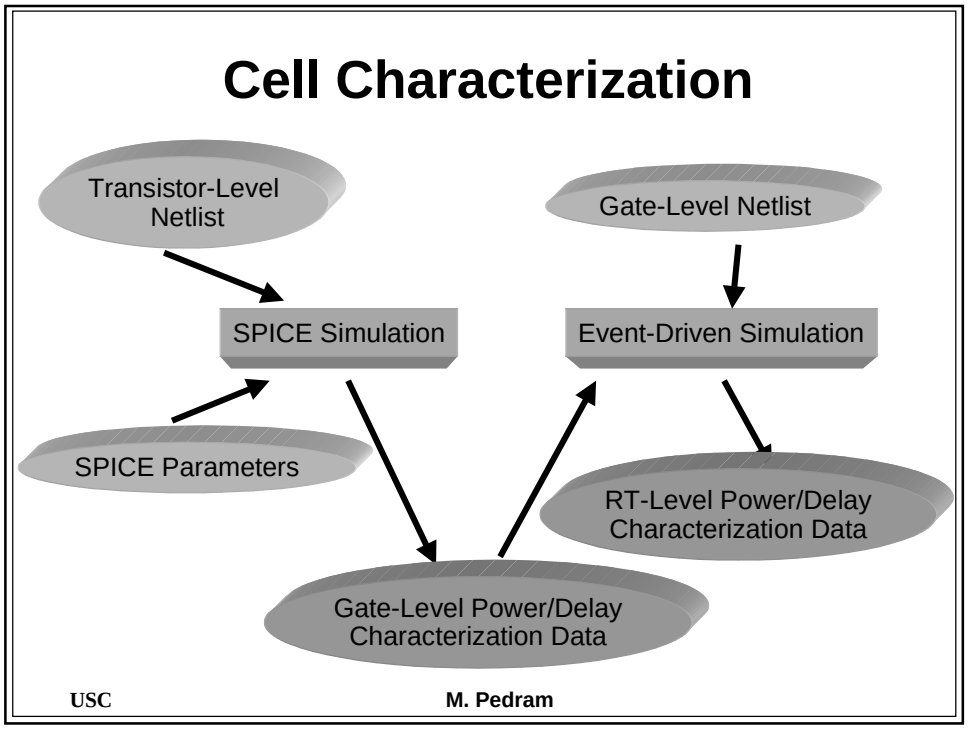
Digital Camera Circuit

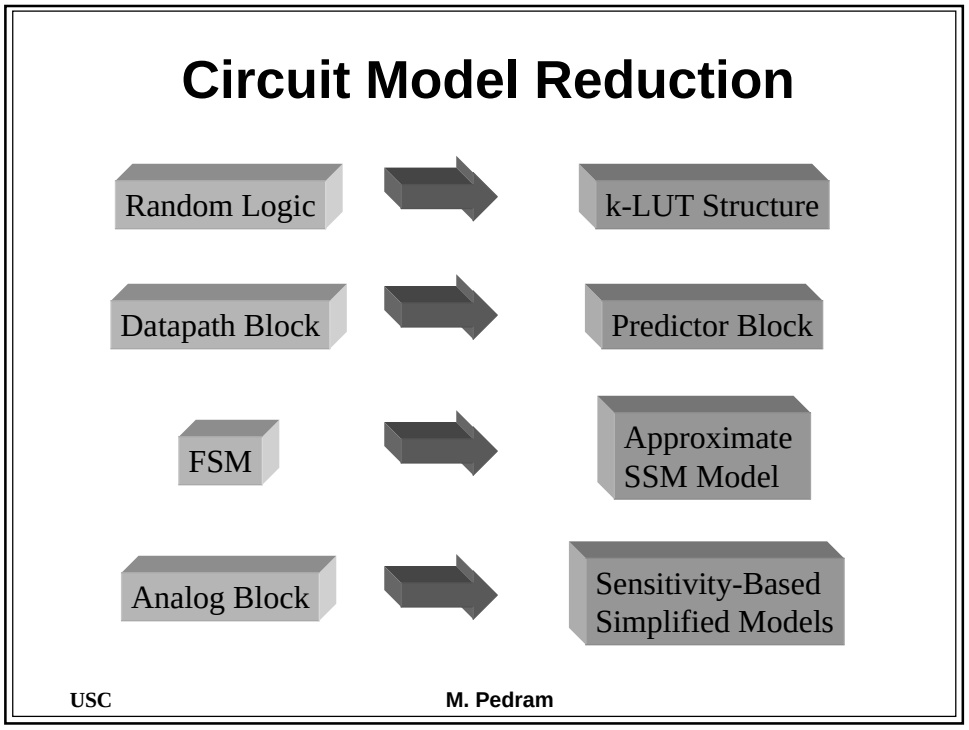


USC M. Pedram

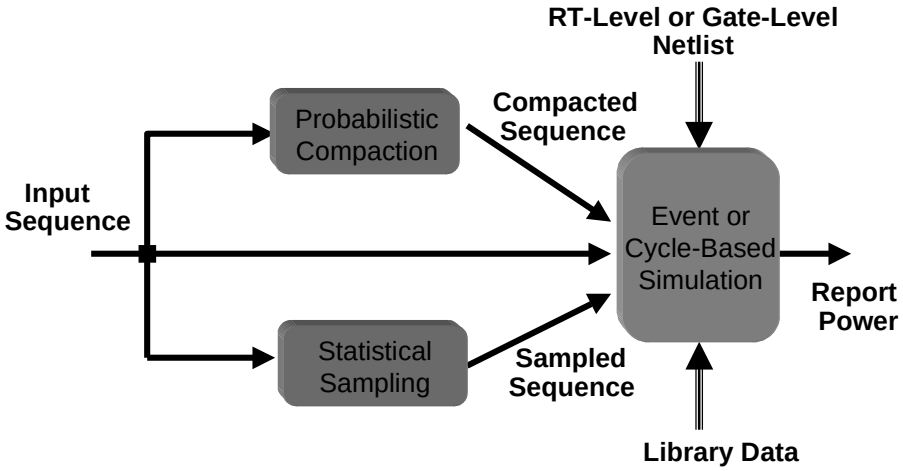

A Power Estimation Methodology
CAD Database
Cell Library
Classify Modules
Memory
Interface Logic
Input Vector Sequences
Interconnect Estimates
Composite Cores
Cores with Macromodel
Control Logic
Analog Modules
Cores w/o Macromodel
Behavioral Simulation
Design Planning
HDL Description
Decompose & Classify Blocks
Library Models Analytic Models
Quick Synthesis Sampling/Compaction Model Reduction
USC
M. Pedram


Quick Synthesis
RTL Source Code
Quick RTL Synthesis
RTL Netlist
Quick Logic Synthesis
Mapped Netlist
USC
M. Pedram


Cell Characterization
Transistor-Level Netlist
Gate-Level Netlist
SPICE Simulation
Event-Driven Simulation
SPICE Parameters
RT-Level Power/Delay Characterization Data
Gate-Level Power/Delay Characterization Data
USC
M. Pedram

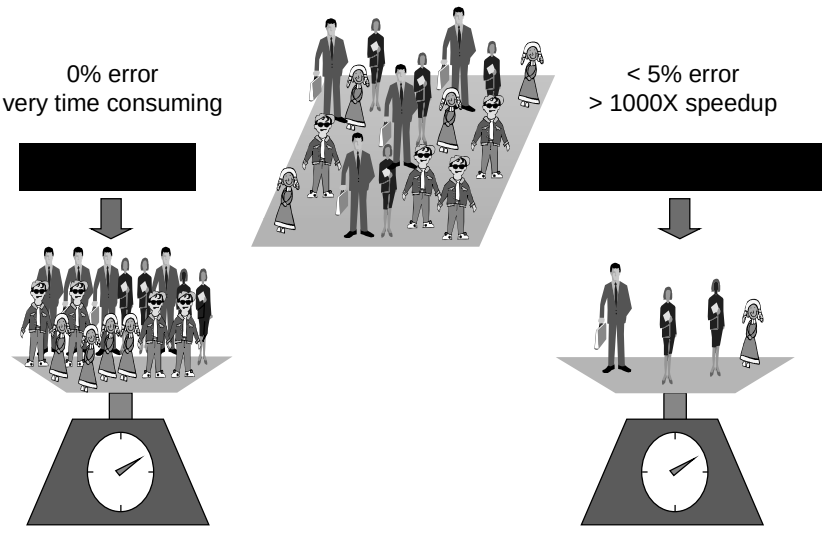

Circuit Model Reduction
Random Logic
k-LUT Structure
Datapath Block
Predictor Block
FSM
Approximate SSM Model
Analog Block
Sensitivity-Based Simplified Models
USC
M. Pedram

Dynamic (Simulative) Techniques

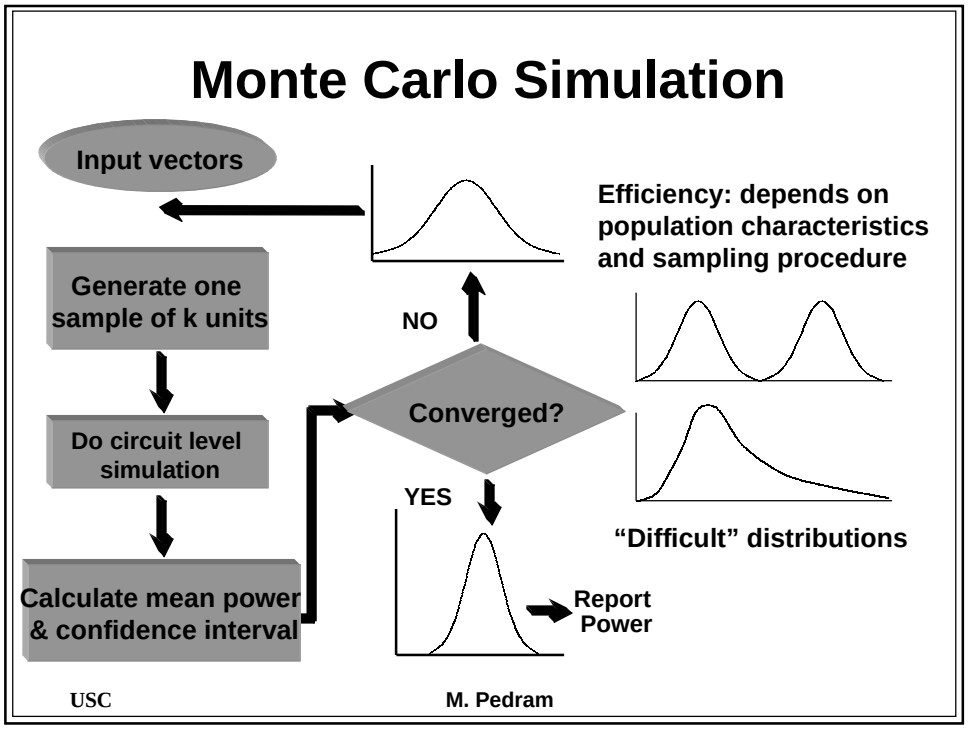


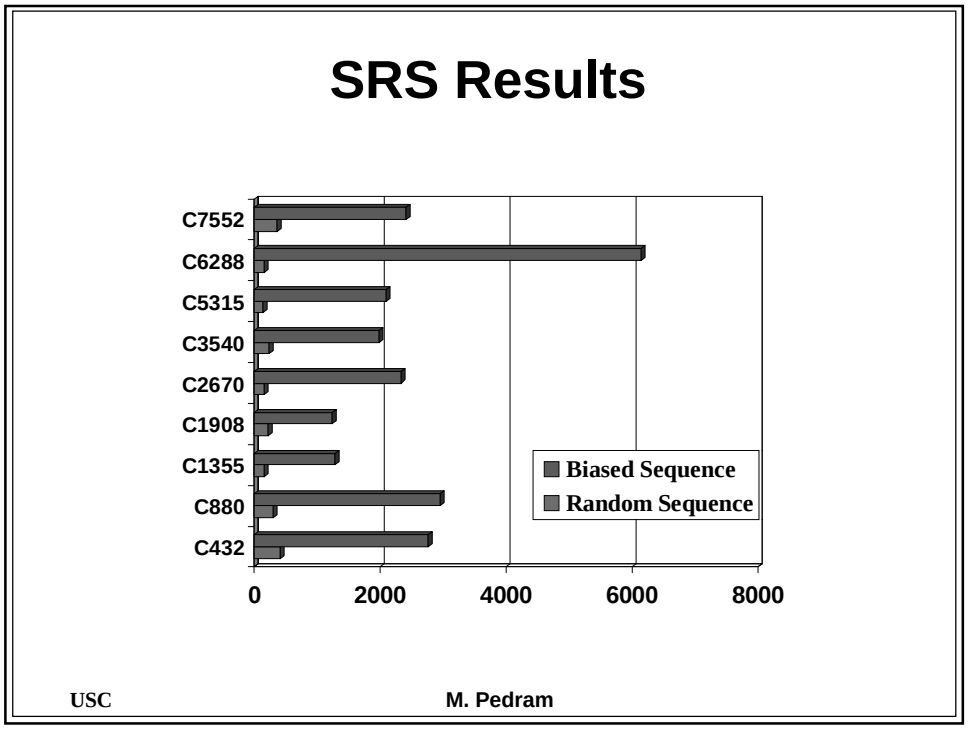
USC M. Pedram

Simple Random Sampling (SRS)



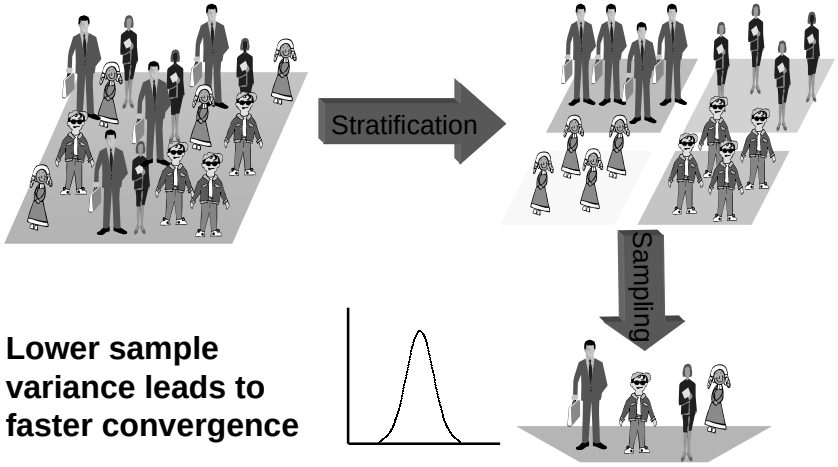
USC M. Pedram


Monte Carlo Simulation
Input vectors
Generate one sample of k units
Do circuit level simulation
Calculate mean power & confidence interval
Converged?
NO
YES
Report Power
Efficiency: depends on population characteristics and sampling procedure
"Difficult" distributions
USC
M. Pedram


SRS Results
C7552
C6288
C5315
C3540
C2670
C1908
C1355
C880
C432
0
2000
4000
6000
8000
Biased Sequence
Random Sequence
USC
M. Pedram

Stratified Random Sampling (StRS)

Estimate the average weight, assuming that gender and age of individuals are readily available

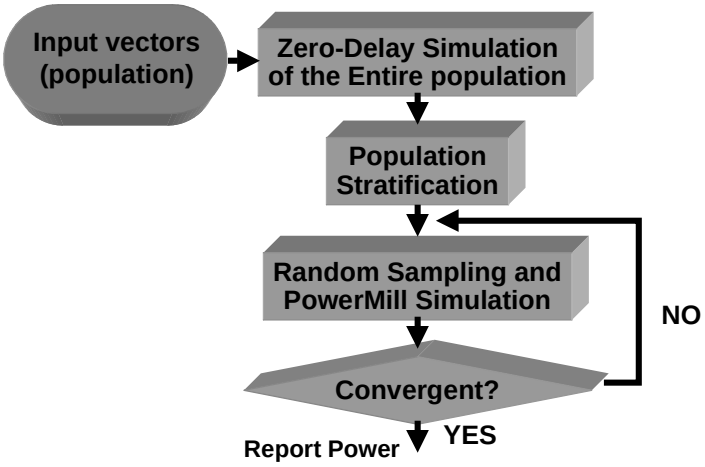


Lower sample variance leads to faster convergence

USC **M. Pedram**

Application to Power Estimation

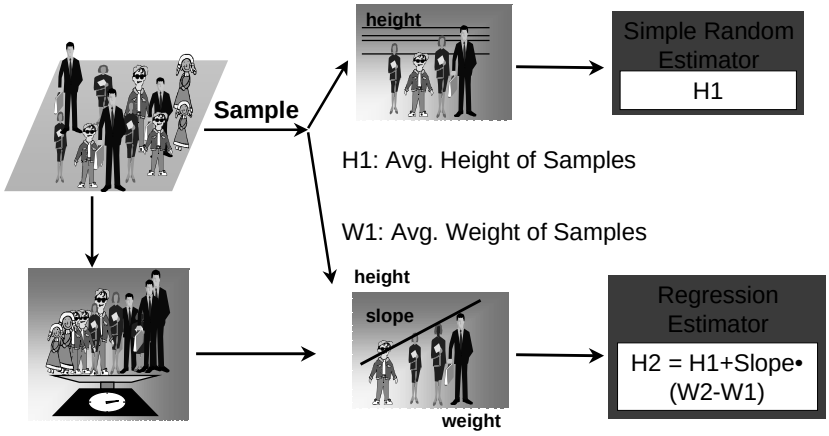
Age & Gender → Zero-delay power estimate
Weight → Powermill power estimate



USC **M. Pedram**

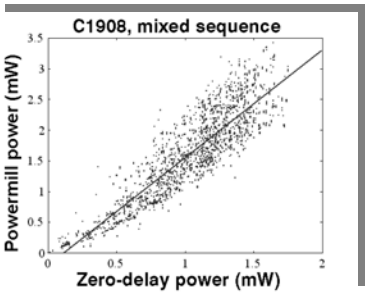
Regression Estimation

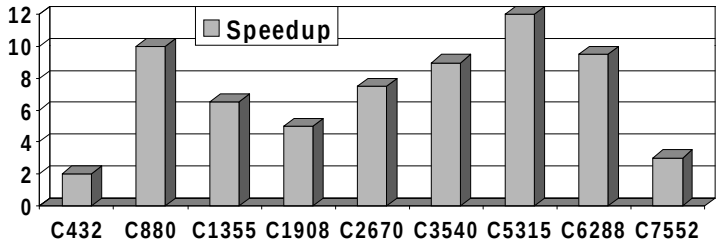
Estimate the average height, assuming that weight of individuals is readily available



USC M. Pedram

StRS Results

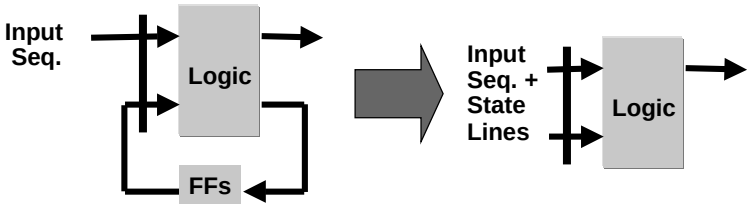




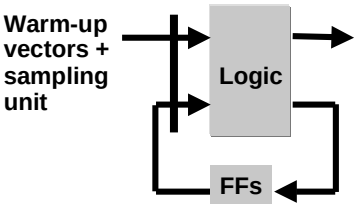
USC M. Pedram

Sampling on FSMs

Method 1: Functional simulation + sampling on comb. circuit



Method 2: Do machine warm-up for every sampling unit

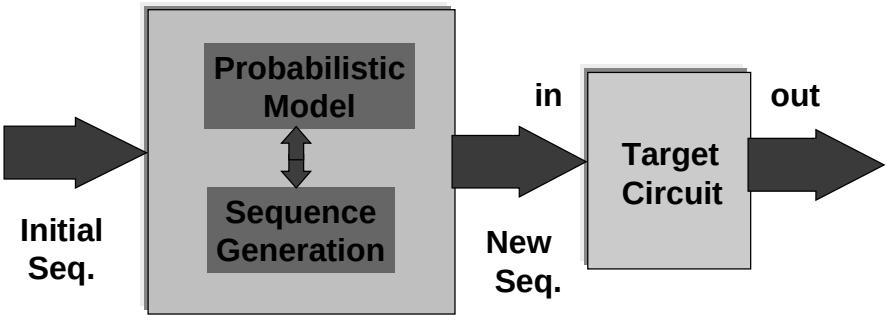


Warm-up sequence length can be quite large

USC M. Pedram

Probabilistic Compaction (PC)

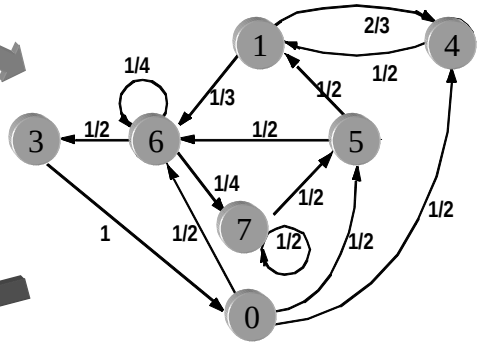
Sequence compaction : Generate a new, but shorter, sequence which exhibits nearly the same spatio-temporal correlations as the initial sequence



USC M. Pedram

PC by Example

Initial sequence
01101010011111010
11100000110111000
10000110101110101
Avg. bit activity
23/16



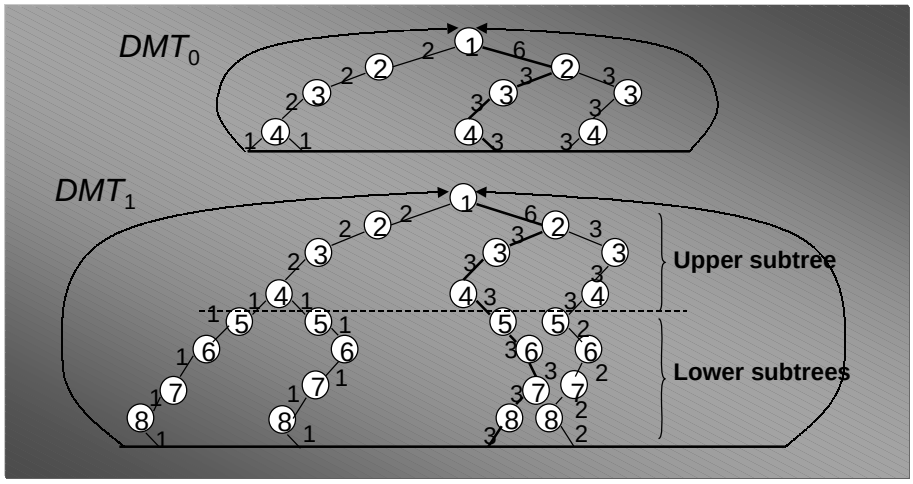
Compacted sequence
001010
011000
010001
Avg. bit activity
8/5

Stochastic State Machine Model

USC M. Pedram

Dynamic Markov Trees

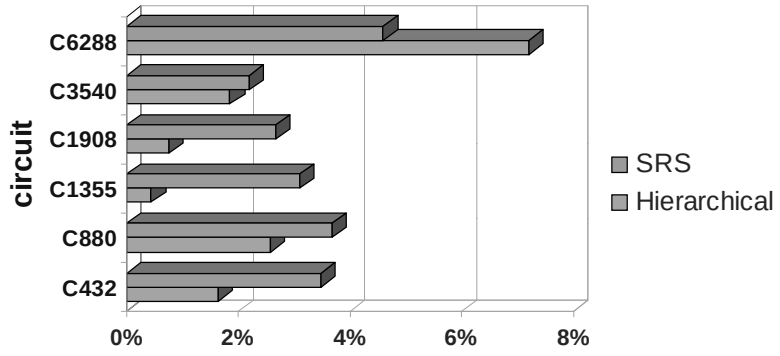
S_1 = (0000, 0001, 1001, 1100, 1001, 1100, 1001, 1100)



USC M. Pedram

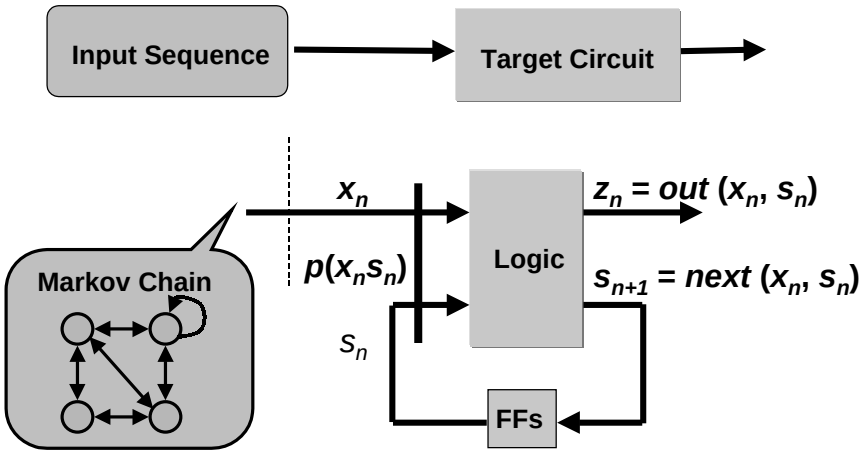
Comparison with SRS

L = 4,000
Compaction Ratio 10



USC M. Pedram

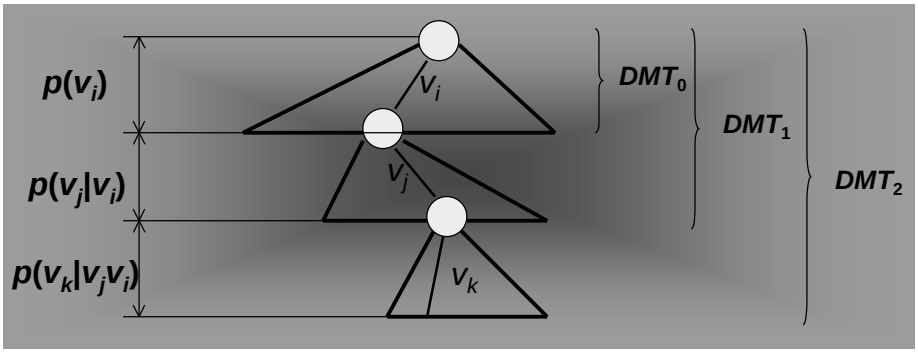
Compaction for FSMs



USC M. Pedram

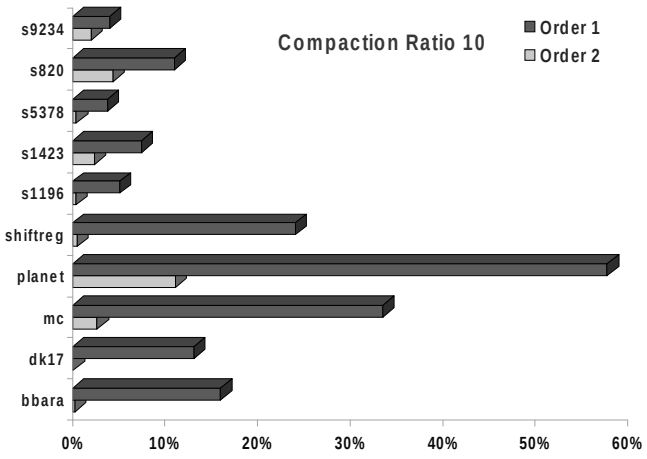
Higher Order DMTs

A lag-k Markov chain which correctly models the input sequence, also models the joint *k*-step conditional probabilities of the primary inputs and state lines



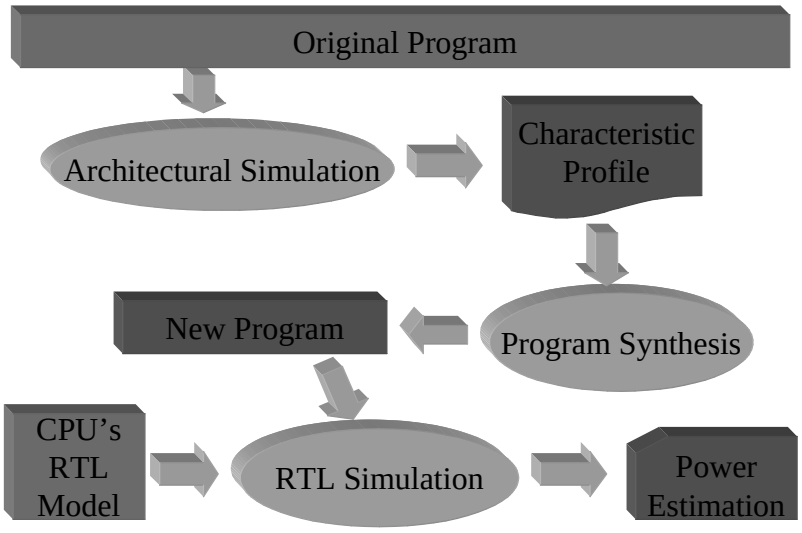
USC M. Pedram

High Order DMT Results



USC M. Pedram

Instruction Level Compaction

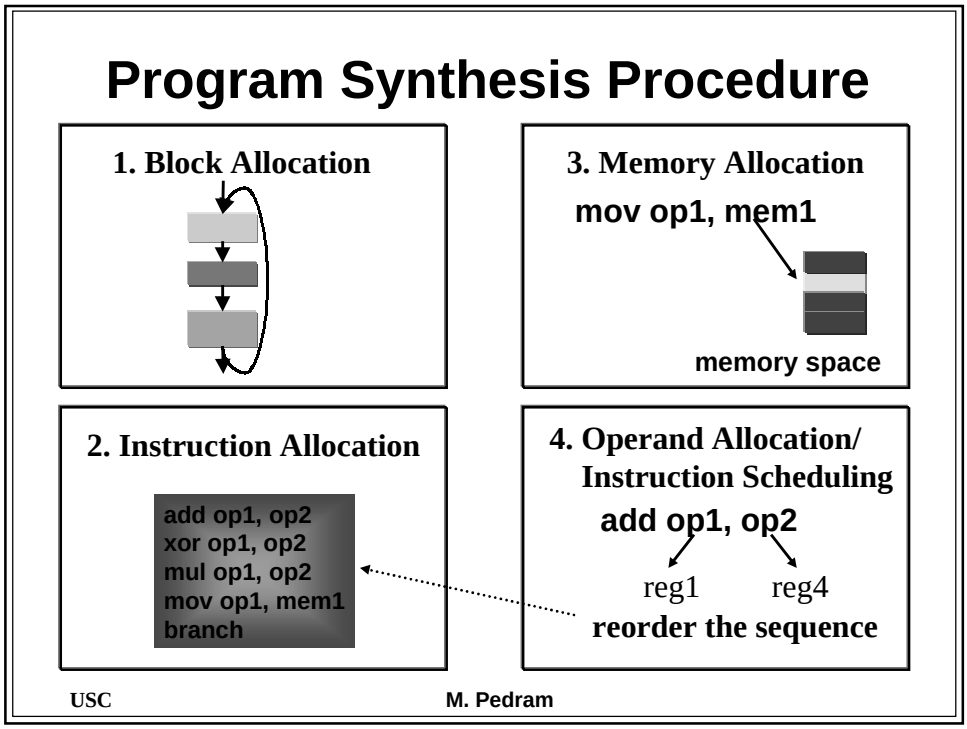


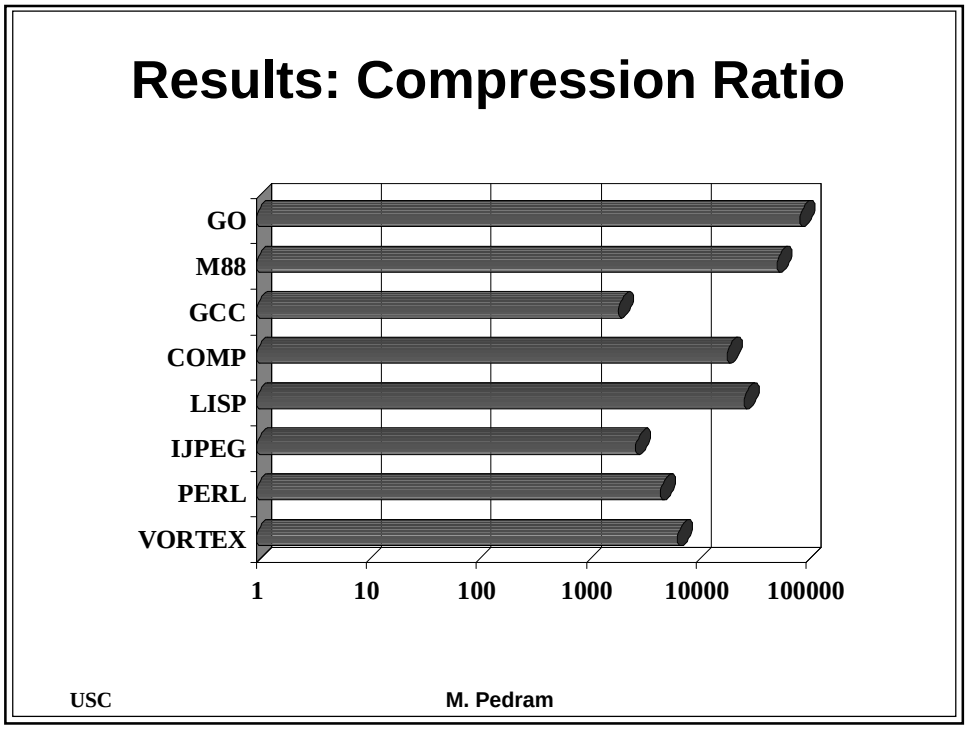
USC M. Pedram

Characteristic Profile

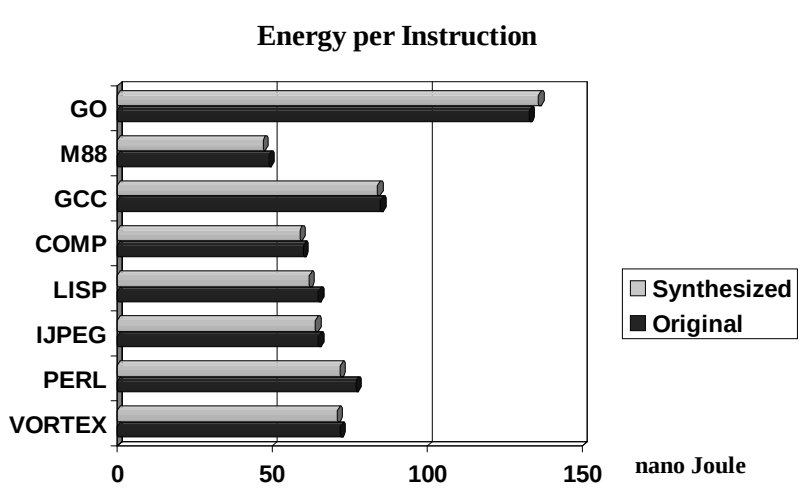
- **Instruction mix**
- **Average instruction size *(if applicable)***
- **Clocks per instruction**
- **Branch prediction miss rate**
- **Instruction cache miss rate**
- **Data cache read/write miss rate**
- **Pipeline stall rate**
- **Speculative execution and register renaming are not considered**
- **Target micro-processor: A super-scalar pipelined CPU with branch prediction (Intel's Pentium chip)**

USC M. Pedram


Program Synthesis Procedure
1. Block Allocation
3. Memory Allocation
mov op1, mem1
memory space
2. Instruction Allocation
add op1, op2
xor op1, op2
mul op1, op2
mov op1, mem1
branch
4. Operand Allocation/
Instruction Scheduling
add op1, op2
reg1
reg4
reorder the sequence
USC
M. Pedram

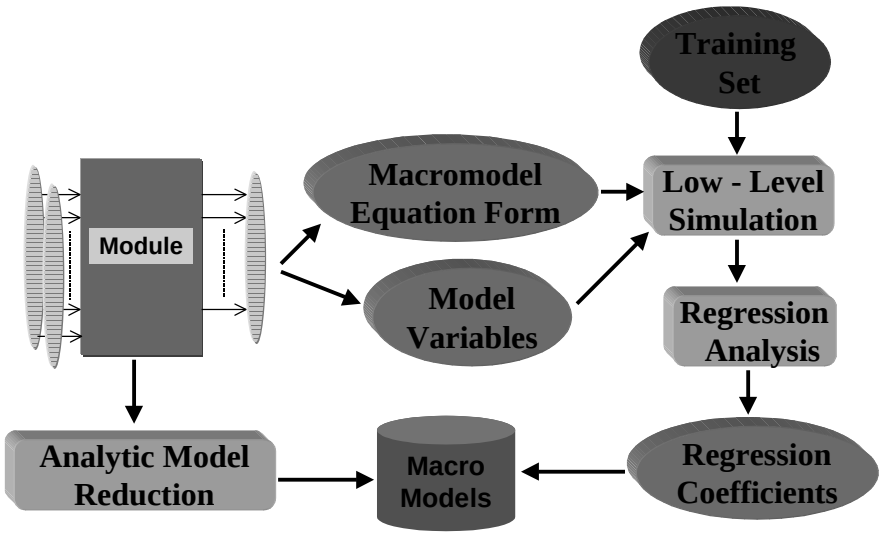

Results: Compression Ratio
GO
M88
GCC
COMP
LISP
IJPEG
PERL
VORTEX
1
10
100
1000
10000
100000
USC
M. Pedram

Results: Accuracy



USC M. Pedram

Power Macromodeling

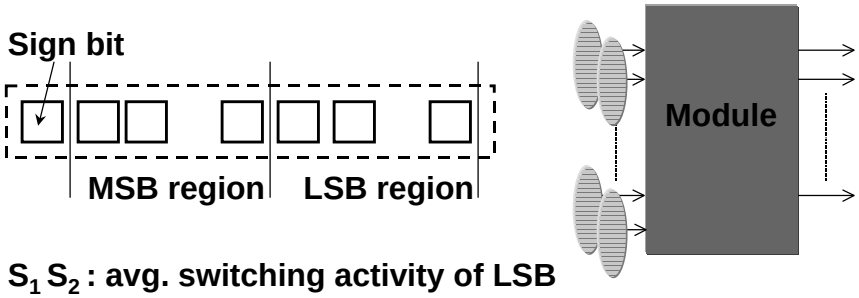


USC M. Pedram

Dual Bit Type Model

Consider a data path block:

$$Pwr = C_0 + C_1 \cdot S_1 + C_2 \cdot S_2 + C_3 \cdot S_3 + C_4 \cdot S_4$$



S_1 S_2 : avg. switching activity of LSB (MSB) region of operand 1
S_3 S_4 : avg. switching activity of LSB (MSB) region of operand 2

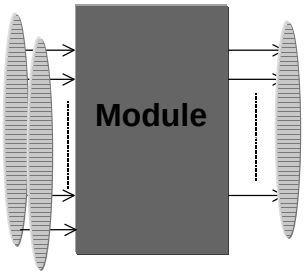
USC M. Pedram

Input/Output Data Model

Consider a data path block:

$$Pwr = C_0 + C_1 \cdot S_1 + C_2 \cdot S_2 + C_3 \cdot S_3$$

S_1 S_2 : avg. switching activity of operands 1 and 2
S_3 : avg. switching activity of output

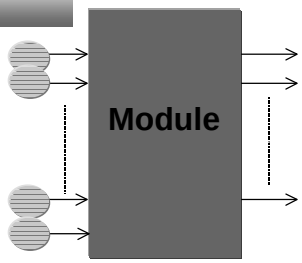


USC M. Pedram

Bitwise Data Model

Consider a random logic block:

$$Pwr = C_0 + \sum_{inputs} C_i \cdot S_i$$

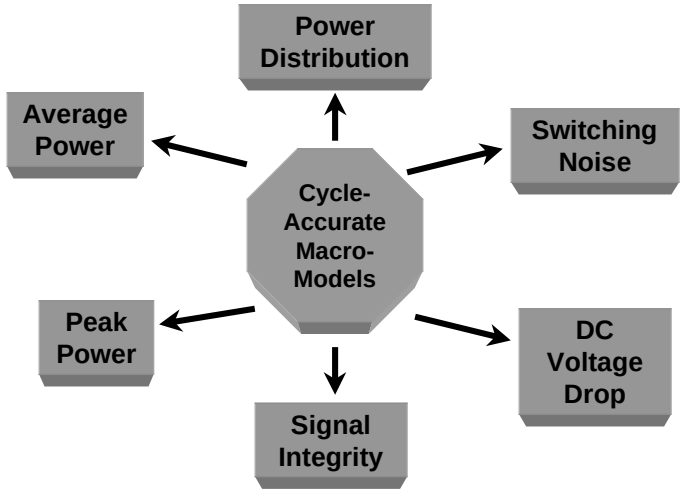


S_i : avg. switching activity of input signal i

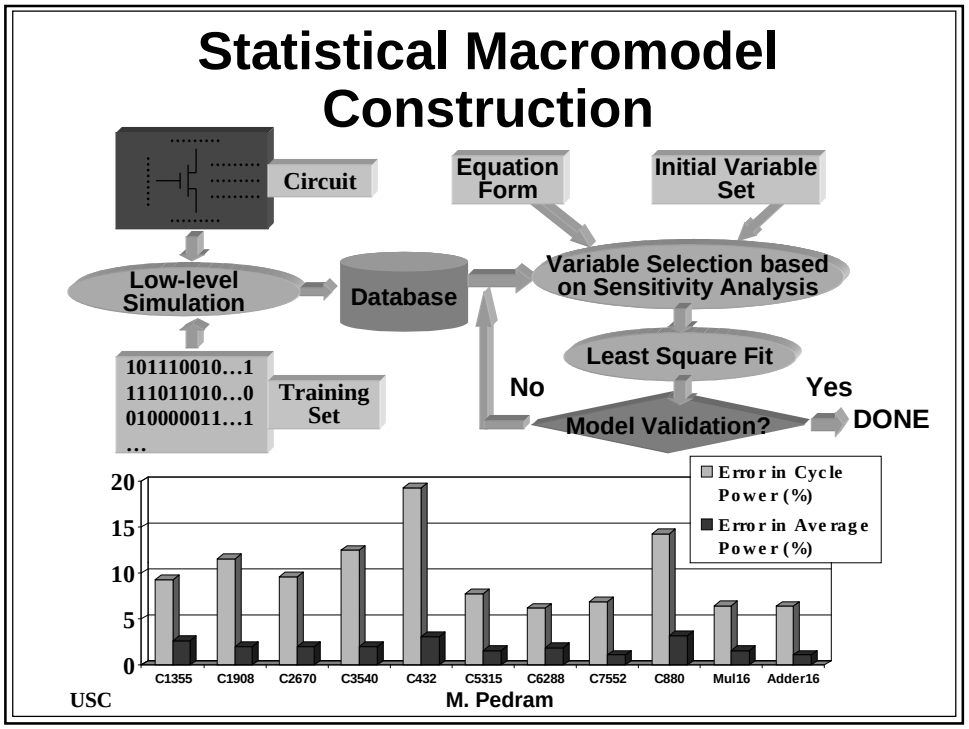
More parameters lead to a higher degree of accuracy, but increase the computational overhead

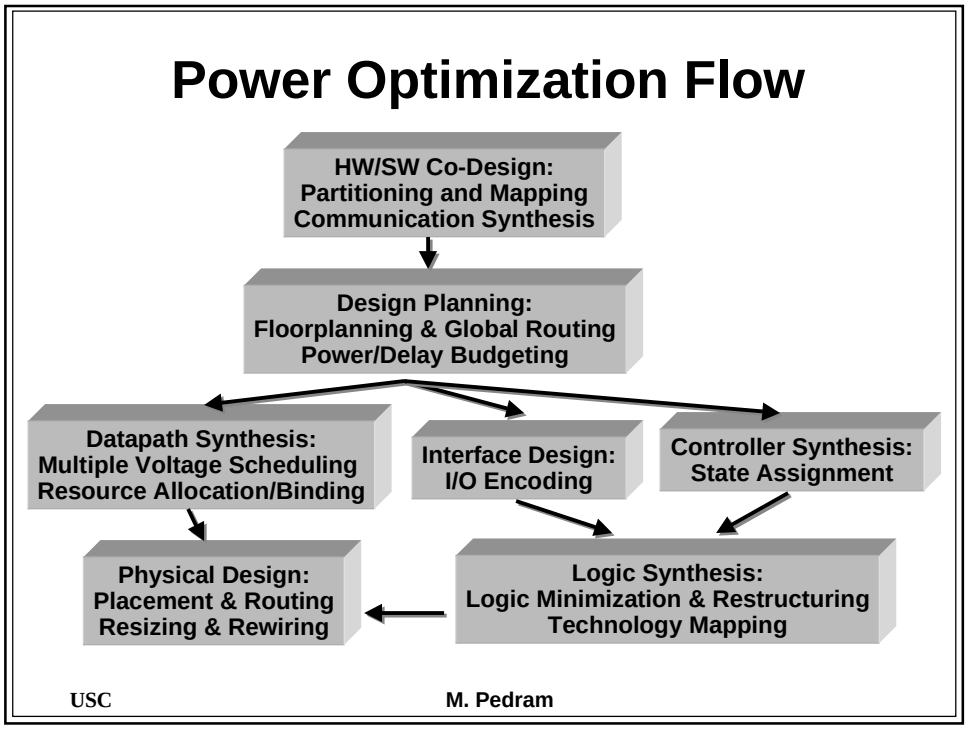
USC M. Pedram

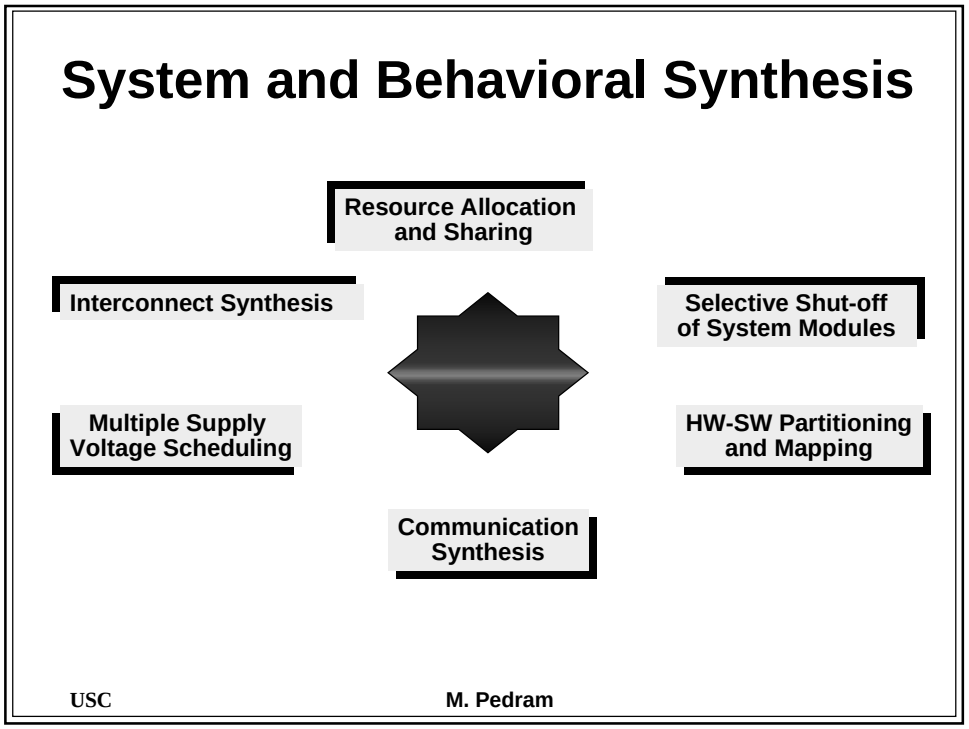
Cycle-Accurate Macro-Models

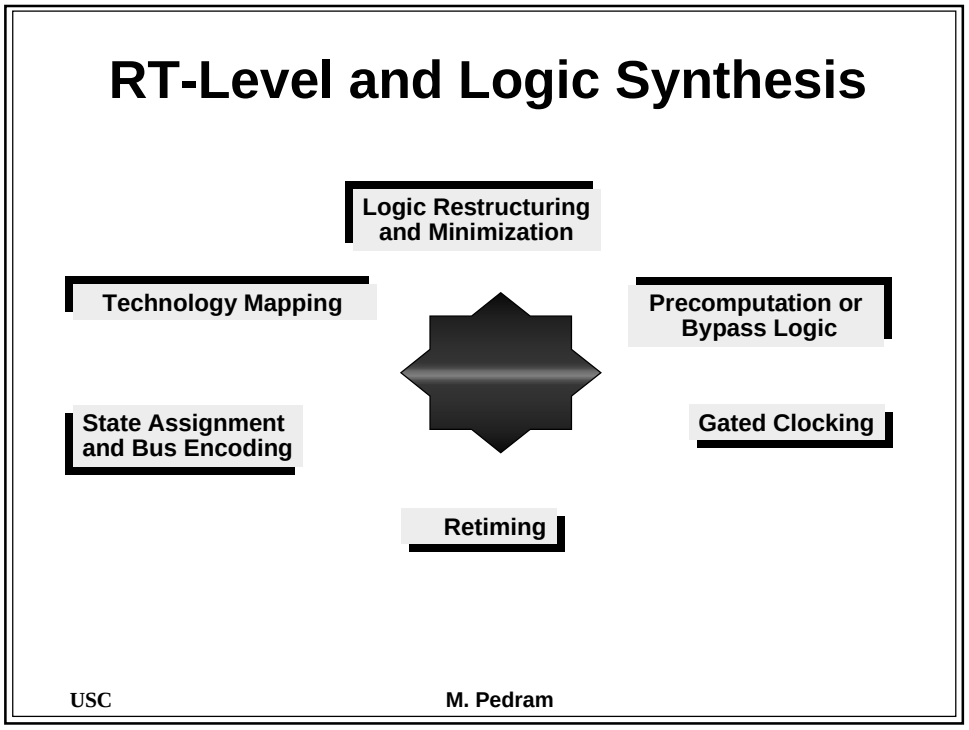


USC M. Pedram

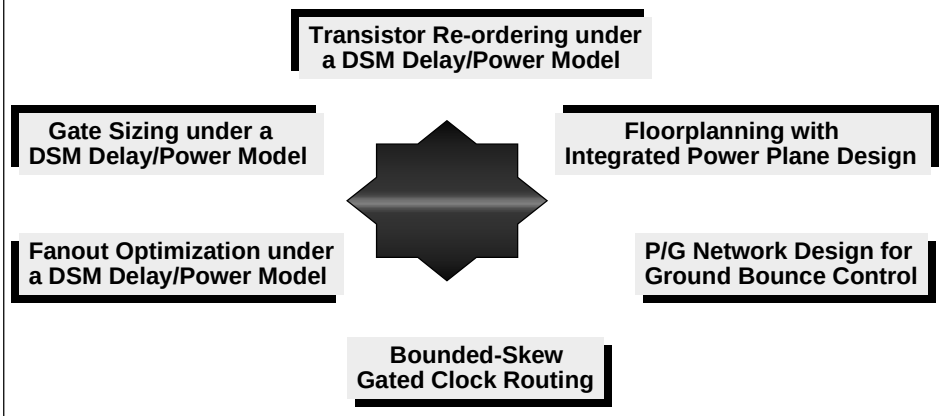

Statistical Macromodel Construction
Circuit
Equation Form
Initial Variable Set
Low-level Simulation
Database
Variable Selection based on Sensitivity Analysis
Least Square Fit
101110010…1
111011010…0
010000011…1
…
Training Set
No
Yes
Model Validation?
DONE
Error in Cycle Power (%)
Error in Average Power (%)
20
15
10
5
0
C1355 C1908 C2670 C3540 C432 C5315 C6288 C7552 C880 Mul16 Adder16
USC
M. Pedram


Power Optimization Flow
HW/SW Co-Design:
Partitioning and Mapping
Communication Synthesis
Design Planning:
Floorplanning & Global Routing
Power/Delay Budgeting
Datapath Synthesis:
Multiple Voltage Scheduling
Resource Allocation/Binding
Interface Design:
I/O Encoding
Controller Synthesis:
State Assignment
Physical Design:
Placement & Routing
Resizing & Rewiring
Logic Synthesis:
Logic Minimization & Restructuring
Technology Mapping
USC
M. Pedram


System and Behavioral Synthesis
Resource Allocation and Sharing
Interconnect Synthesis
Selective Shut-off of System Modules
Multiple Supply Voltage Scheduling
HW-SW Partitioning and Mapping
Communication Synthesis
USC
M. Pedram


RT-Level and Logic Synthesis
Logic Restructuring and Minimization
Technology Mapping
Precomputation or Bypass Logic
State Assignment and Bus Encoding
Gated Clocking
Retiming
USC
M. Pedram

Physical Design



USC M. Pedram

Summary

- CAD flows and tools can reduce power dissipation in VLSI circuits and systems by a factor of 5 - 8 X over the next three years
- Process and voltage scaling can provide another factor of 8 - 12 X
- Commercial tools for gate-level and circuit-level power estimation and optimization exist
- High level power analysis and estimation tools are a key enabling technology
- Early Design planning and system-level design and power optimization tools are needed

USC M. Pedram