Synthesis of Quality Configurable Systems

Seogoo Lee, Behzad Boroujerdian

Andreas Gerstlauer, Lizy K. John

Electrical and Computer Engineering

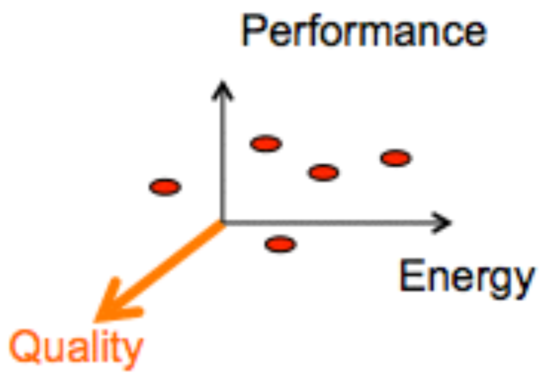
University of Texas at Austin

(This research is supported by NSF & Intel)

The University of Texas at Austin
Electrical and Computer Engineering
Cockrell School of Engineering

Approximate Computing (AC)

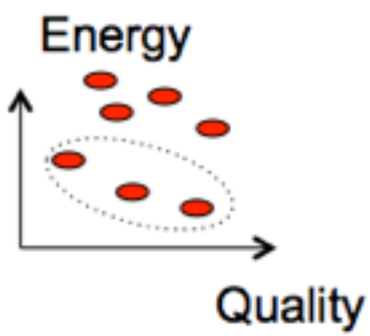
- **Welcoming quality degradation**
  - Gain in performance, power, …
- ***Quality* as a new design objective**



- **Optimize in the quality-energy-performance design space**
  - Need systematic treatment across the compute stack

Quality Adaptiveness

- **Quality States (Q-states)**
  - Pareto optimal in Q-*x* space

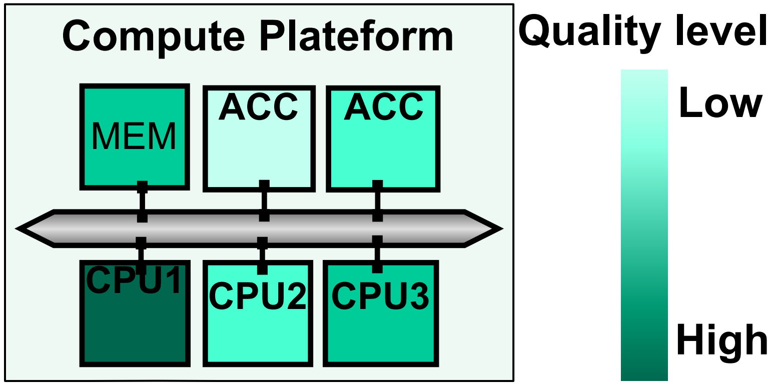


- **Dynamic nature of quality**
  - Varies based on: application
  - Varies based on: input, temperature, …
  - ➢ Quality adaptive systems are desirable

- **Ideal system**
  - Identifies Q-states
  - Navigates through Q-states

Quality Configurable Architecture

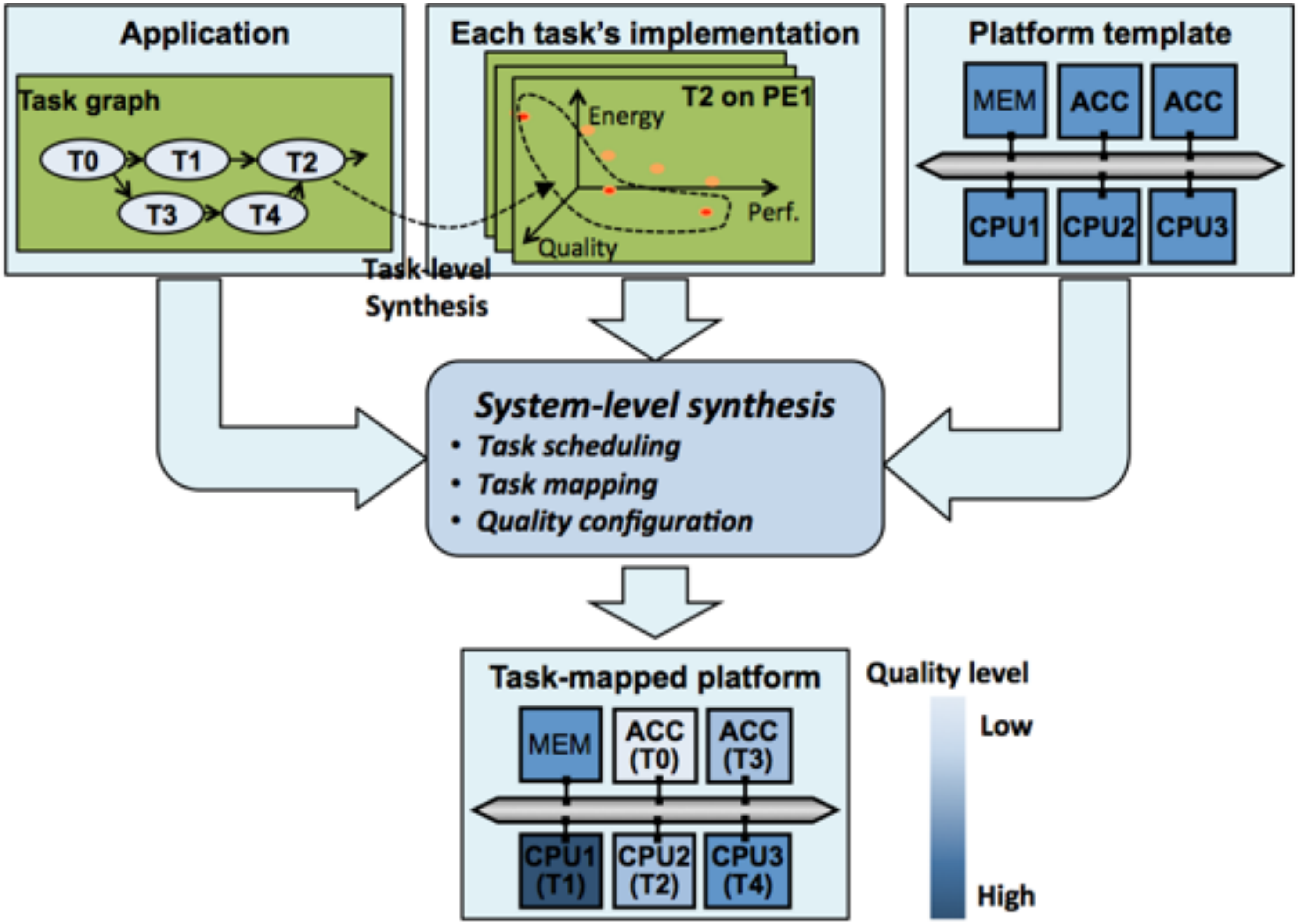
- **Heterogeneous systems**
  - Mapping and scheduling
  - ➢ Quality-configurable PEs



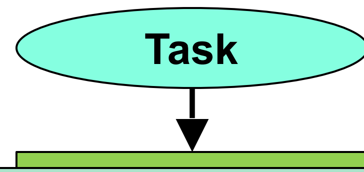
➢ **Software processors**
  - Quality exposed to ISA
  - One binary per Q-state

➢ **Hardware processors**
  - Control capable of transitioning among quality states

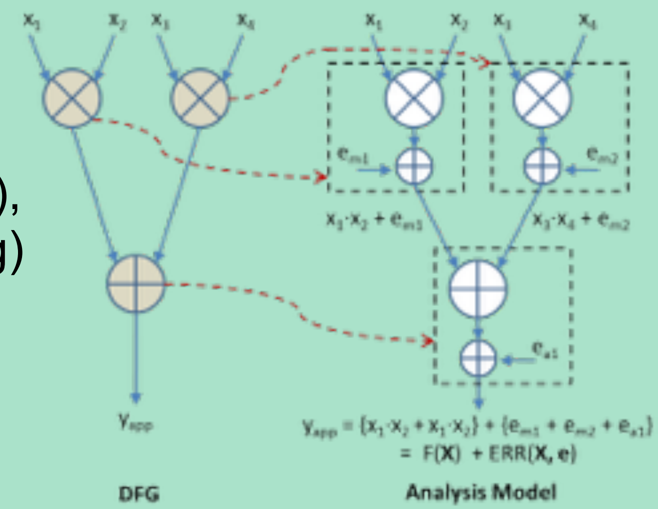
Quality Management



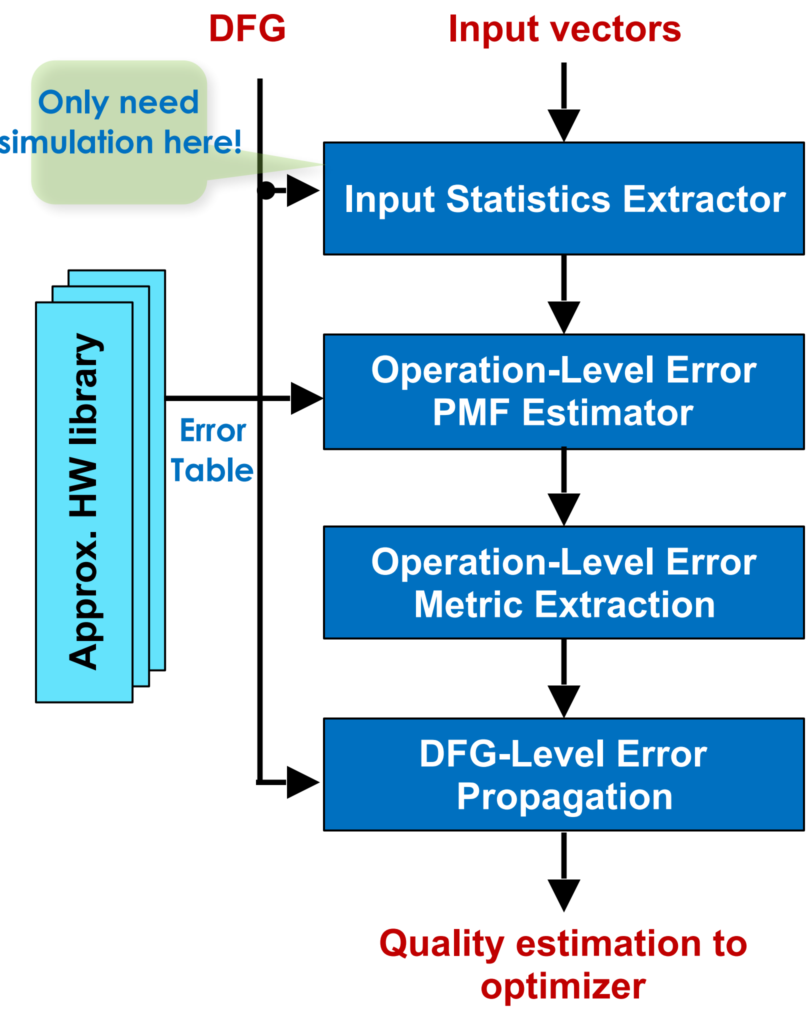
Approximating HW/SW Synthesis



- **Analytical quality/energy models [ISQED16]**
  - Estimate operation-level quality metrics (statistical variance, statistical bounds, min/max) & propagate through (C)DFG
  - Energy model considering area (leakage), switching activity & delay (voltage scaling)
- **(Meta-)heuristics for optimization**
  - Quality-energy design space

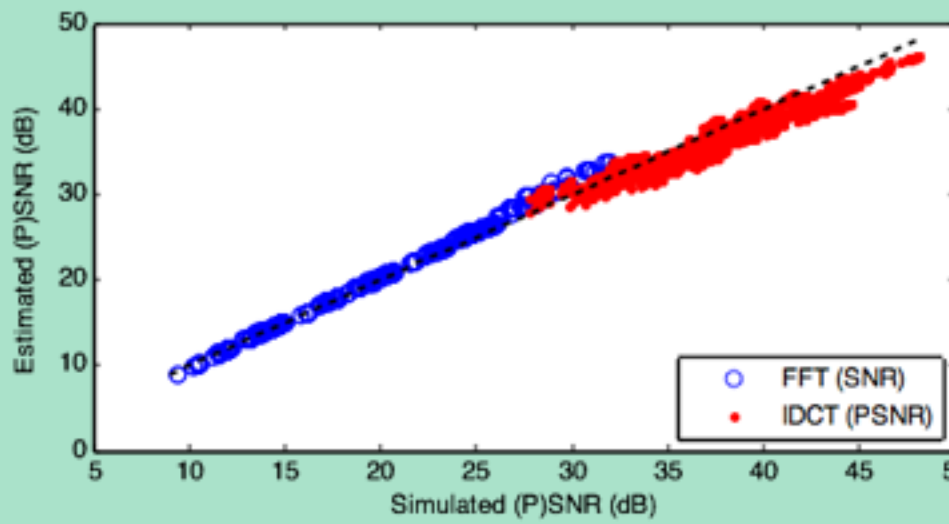


Quality Model



Our quality estimation is

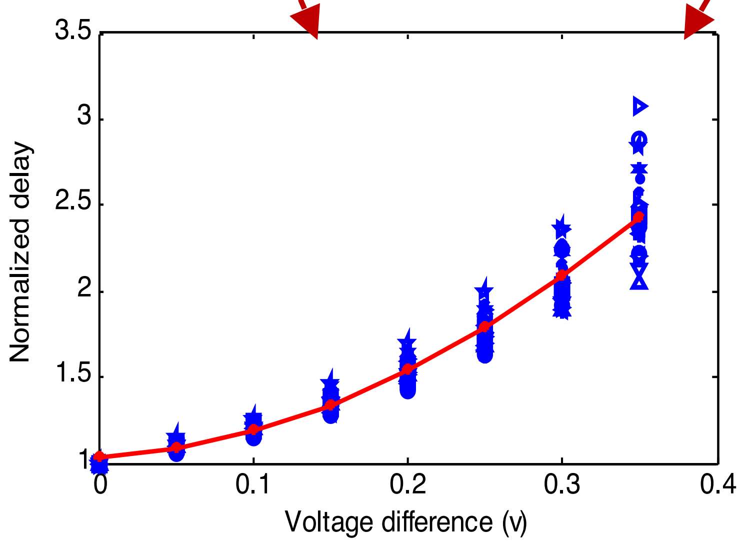
- GENERIC: work with various quality metrics and HW approximations
- FAST: semi- semi-analytical only with one-time error-free simulation
- ACCURATE: consider data/error dependency

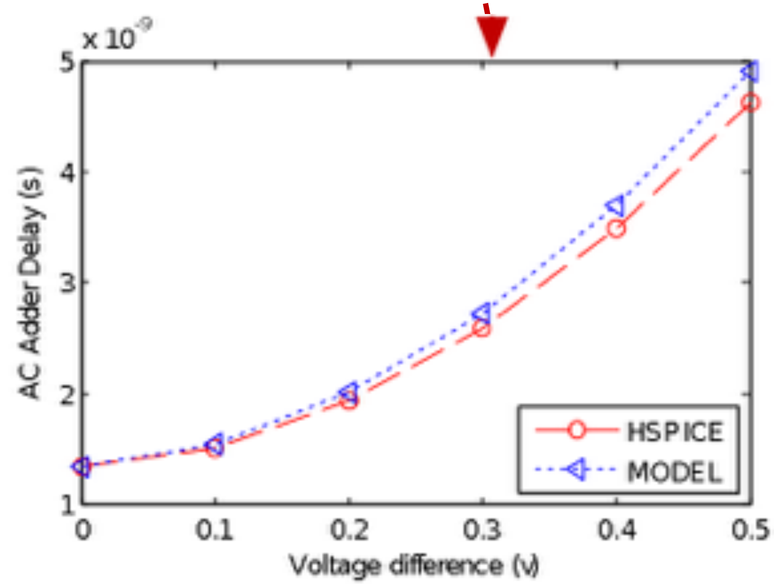


Energy Model

- **Approximate computing reduces**
  - Area/switching activity: directly convertible to energy
  - Critical path timing delay: useful with voltage reduction
    - Need voltage-delay model of standard cells

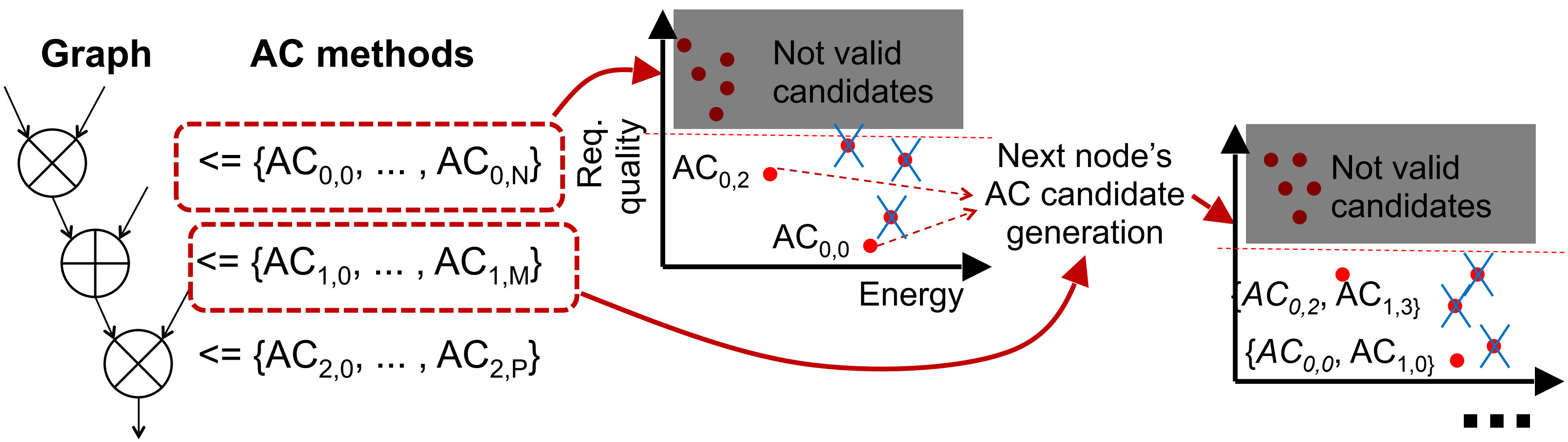
HSPICE simulation with voltage sweep → **Find a common model all standard cells** → **Calculate delay of critical paths for different voltages**





Optimization Heuristic

- **Branch & bound like heuristic**
  - Considers that a node's approximation changes its output
    - Start approximation from the first node
    - Forward traversal of a graph
  - Early elimination of dominated solution candidates



Experiments

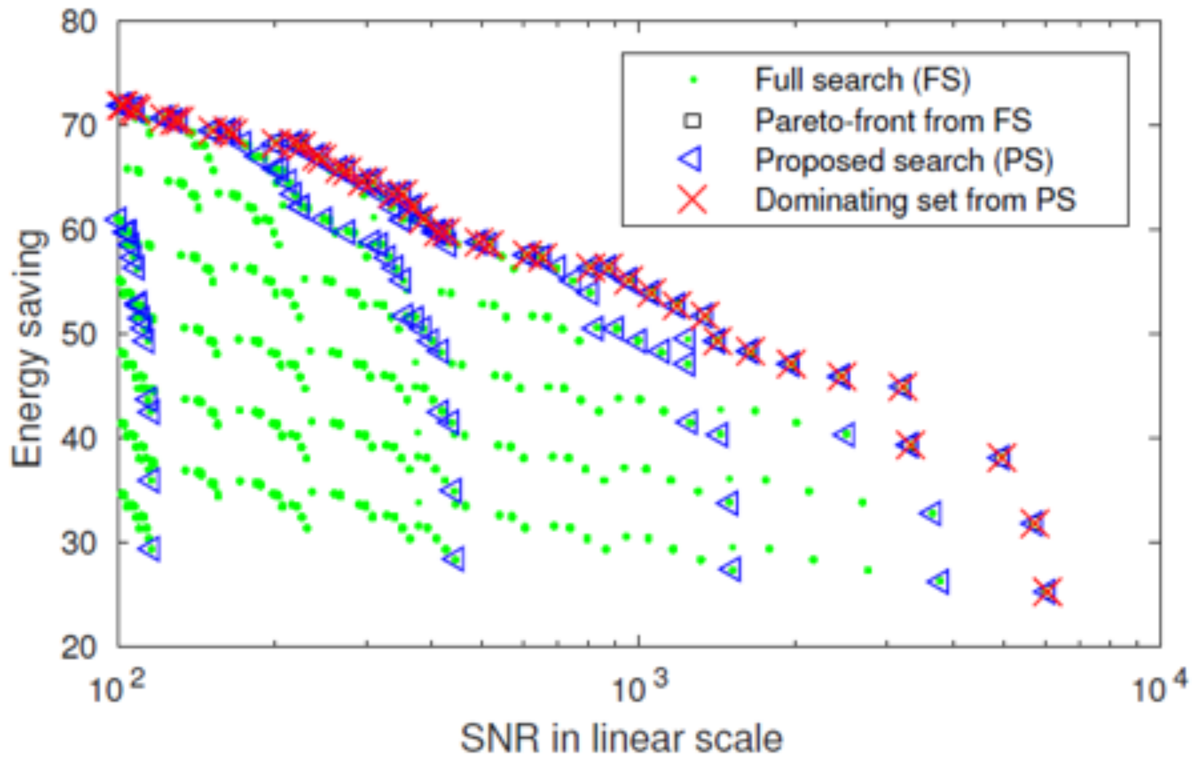
- **Setup**
  - Examples

Design	# of s_i	Dataset	Description
2MM	6	Gaussian	2×2 matrix multiplication
branch	7	Gaussian	Simple branch example
idct	3	Lena image	1D-IDCT at JPEG decoder
sad	2	SD-VDS [17]	Sum of absolute difference
had8x8	5	H.265	8×8 Hadamard matrix computation
gblur	6	SD-VDS	5-tap Gaussian blur

  - Delay characterization: Synopsys 32 nm (V_{ref} = 1.05V)
  - Quality: SNR with analytical method

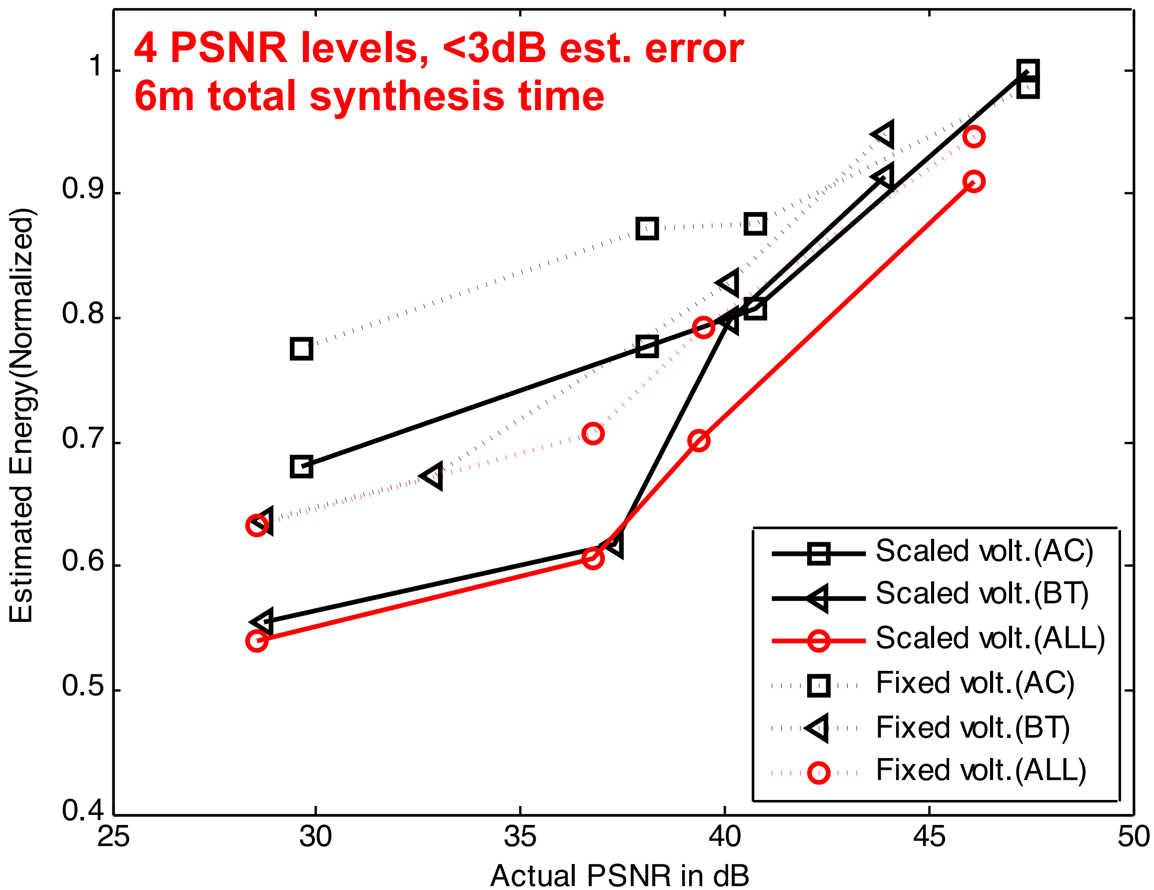
Efficiency of Exploration Heuristic

- **Optimality comparison**
  - For example, Gaussian blur function
    - With target SNR = 100.0 (linear)
  - Obtain a set of solutions that are close to Pareto frontiers



IDCT Optimization Results

- **2D-IDCT, 4 SNR target levels**
  - Bit-truncated (BT) & approximate (AC) adders/multipliers



45dB

46% energy savings w/ Vdd scaling

30dB

Summary & Conclusions

- **Quality as a new design objective**
  - Quality-energy-performance design spaces
- **Quality-adaptive systems**
  - Identifying Q-states
  - Navigating through them
- **Quality management**
  - Compilers for quality-programmable processors
  - Synthesis of quality-configurable accelerators
  - System-level task mapping/scheduling
  - Run-time quality governors (OS)

Thank You!