Power Benefits of Imprecise Computer Arithmetic

paragk@ucla.edu

Outline

- What is Imprecise Arithmetic
- Motivation
- Prevalent directions in research
- Applications
- Shortcomings
- References

Imprecise Arithmetic: Overview

- **In essence, not the exact output all of the time**
- **Certain input-vectors, will give an incorrect output**
- **Look to bound the error values and quantify their probability of occurrence**
- **Find applications that can absorb computed error values and probabilities**
  - Typically apply well to signal/media-processing
  - Analyze the benefits of inaccuracy: usually timing or power

Imprecise Arithmetic: Motivation

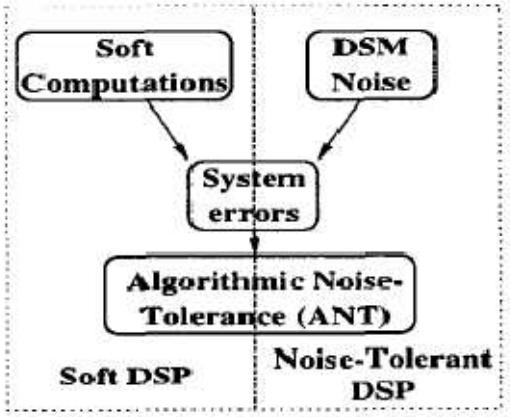
- **One of the biggest motivations used to be area, for example –**
  - Nearly half of the hardware in a tree or array multiplier is there just to get the last bit right [1]
- **Sometimes even performance/timing improvement, for example**
  - In n-bit addition, on average the carries propagate across $\log_2 n$ positions, but we still design for the worst case
- **Recently most though imprecision has been looked at from a power perspective**
  - In large part due to [5], which showed through thermo-dynamic arguments that a probabilistic switch's energy consumption can be as low as $kt*\ln_2*(1-p)$ joules per switching

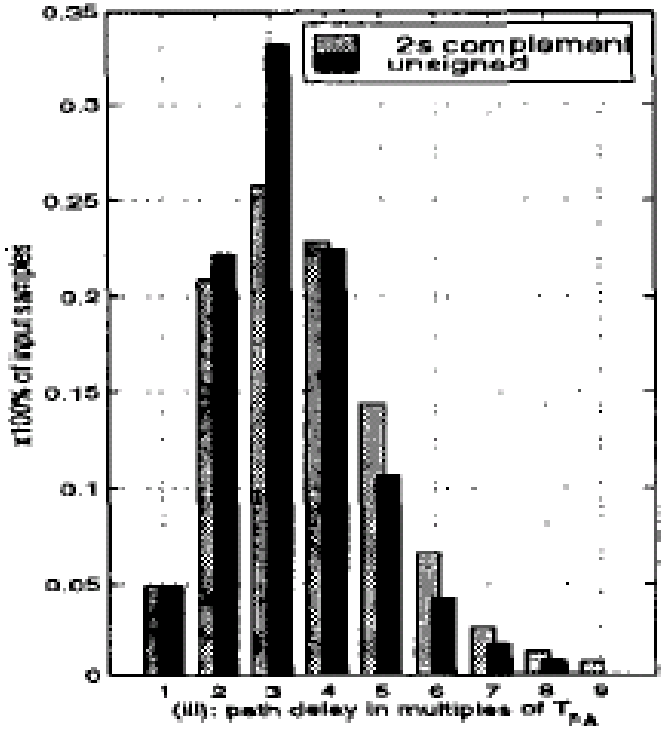
Energy Efficient Signal Processing via Algorithmic Noise Tolerance [2]

- **Basic Ideas –**
  - Scale supply voltage beyond the critical voltage required
  - Compensate for "noise" introduced by Algorithmic Noise-Tolerance (ANT)
- **Voltage scaling is an effective means of achieving reduction in energy since a reduction in supply voltage by a factor K, reduces the capacitive component of energy dissipation by a factor of K^2**
- **Typically a system is designed such that V_{dd} meets the worst case delay over all possible input patterns**
- **$V_{dd} = KV_{dd\text{-}crit}$, where $0 < K < 1$, leads to erroneous output when the critical path is excited**
- **Velocity saturation in modern devices favors this form of "Soft DSP"**

Results and ANT

- ANT system is to compensate for the degradation in the system output due to errors from soft-computations
- The fraction of inputs that excite the critical paths is small
- Energy savings of between 34-51% reported, with a degradation in performance by 0.5dB





Energy Efficient Computing Through Approximately Correct Arithmetic [3]

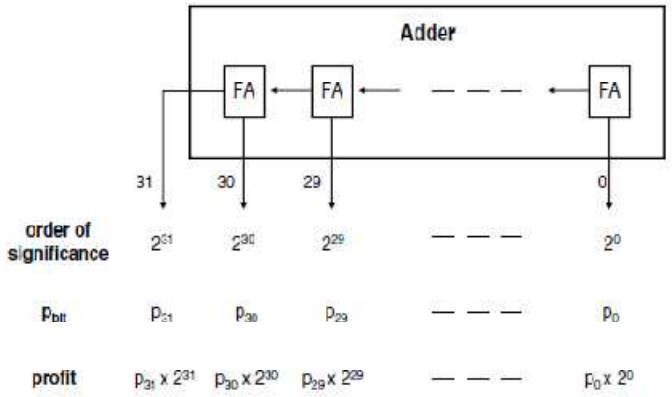
- The i^{th} bit is correct with the probability p_i, related to propagation delays
- Probabilistic arithmetic units derived from such probabilistic gates are shown to form a natural basis for primitives for DSP
- The resultant error at their outputs, manifests as degradation in the SNR
- Images obtained from such "probabilistic DSP" are nearly visually indistinguishable from those obtained via "deterministic approaches"
- Significant power-benefits for little SNR loss
- Effectively error has become a "design-parameter" that can be traded off for power savings (using the (1-p) factor from previous slide)

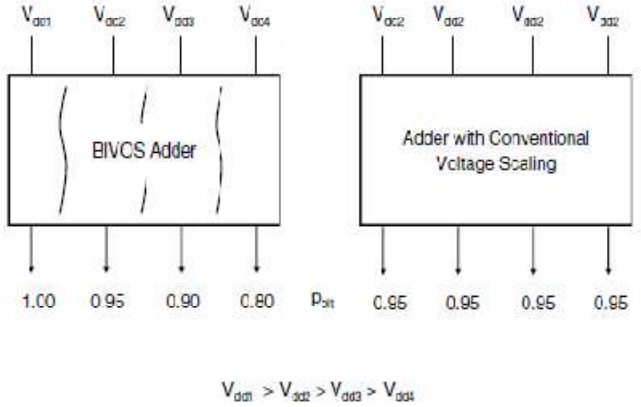
Guiding Principles

- ***There is a trade off between energy consumption and error induced by propagation delay, in circuits which implement arithmetic operations, that can be exploited to garner energy savings***
- ***Errors in bits of a higher value affect the quality of solution more than similar number of errors in bits of a lower value***
- ***Hence - will operate the full adders in the more significant position with a higher supply voltage, when compared to the full adders in the less significant positions. Thus decreasing the error rate or "probability" of error in the more significant positions when compared to the less significant positions.***

How this is different -

- Not all errors were created equal! Hence -each bit has an investment and profit associated with it
- Invest more (energy/V_{dd}) where profit (accuracy) is more
- Implies non-uniform bit error rates among the individual bit positions (of say an adder) with the MSB having lowest bit error rates and the LSB having the highest
- Exponentially better power savings compared to uniform scaling – no need for ANT either





Results

- **Images –**
  - (a) Baseline operation
  - (b) Non-Uniform Voltage Overscaling – 21.7% power reduction
  - (c) Uniform Voltage Overscaling – for same energy as (b)

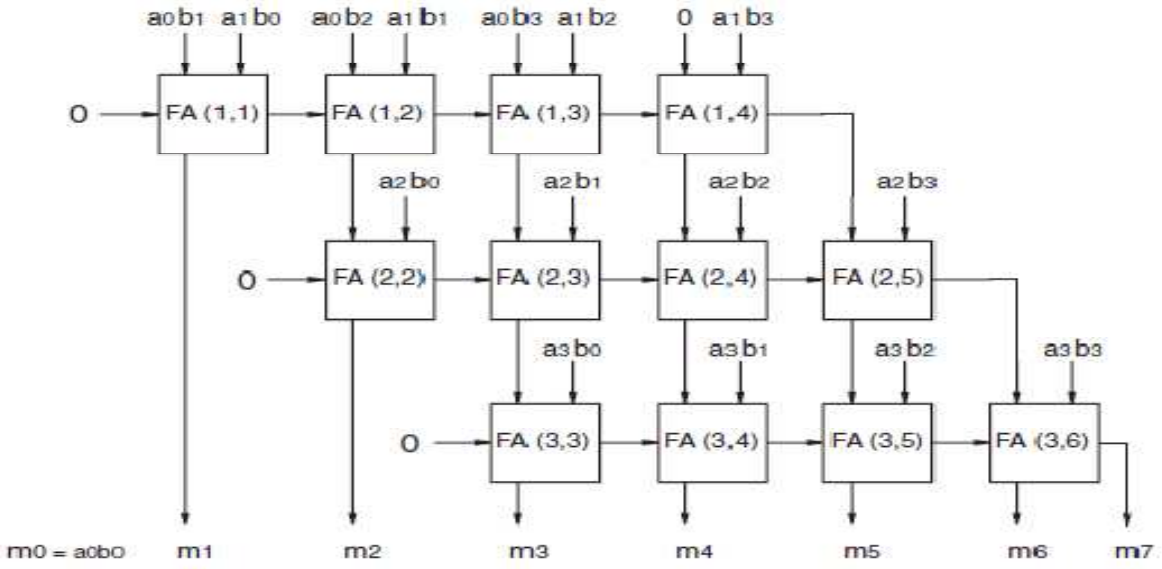
(a)

(b)

(c)

Energy Aware Probabilistic Multiplier [4]

- Basic principle is same – relax energy requirements by allowing possibly incorrect computation – extend the idea to multipliers
- Energy assignment for multipliers can be substantially different as compared to ripple-carry-adders
- Apply it to image enhancement
- Guided more by experimental results and Monte-Carlo Analysis
- First attempt in literature to consider energy assignment for a probabilistic multiplier



Multiplier Continued ..

- **Why is it different –**
  - Errors caused by adders propagate in two directions
  - Horizontal (via carry out) and vertical (via sum)
  - Leads to dependencies which in turn make it hard to derive an analytical expression
- **Based on analysis of simulation results they propose the following approach –**
  - First divide a given energy budget into sub-budgets, each is assigned to a column of adders.
  - Each of these sub-budgets are decided by optimization with Monte-Carlo simulation
  - Then a columns sub-budget is equally divided for each adder

Results

- **Image Sharpening –**
  - (a) Original blurred image
  - (b) Enhanced used deterministic multiplier
  - (c) Probabilistic multiplier with proposed assignment
  - (d) Uniform energy assignment with same energy as (c)

(a) (b) (c) (d)

Shortcomings

- **Wire delay – since error depends on delay, possibly wire delay should have been factored into the equation**
- **Gate sizing – up/downsizing of even a single gate affects the error probabilities and hence the effectiveness of the circuit**
- **Feasibility – the overhead of up-to four different voltage domains within the compact layout of an arithmetic unit is not accounted for**
- **Automation – *"We envision our circuits to be used in the context of application specific integrated circuits (ASICS)"* [3]**
  - But any standard optimization from gate sizing, to buffer insertion would affect the error probability
  - Might imply going back to the drawing board and re-computing energy assignments for every minor change in netlist

References

- [1] Gaining Speed and Cost Advantage from Imprecise Computer Arithmetic – Presentation – Prof. Behrooz Parhami
- [2] Rajamohana Hegde , Naresh R. Shanbhag, Energy-efficient signal processing via algorithmic noise-tolerance, Proceedings of the 1999 international symposium on Low power electronics and design, p.30-35, August 16-17, 1999, San Diego, California, United States
- [3] Lakshmi N.B. Chakrapani , Kirthi Krishna Muntimadugu , Avinash Lingamneni , Jason George , Krishna V. Palem, Highly energy and performance efficient embedded computing through approximately correct arithmetic: a mathematical foundation and preliminary experimental validation, Proceedings of the 2008 international conference on Compilers, architectures and synthesis for embedded systems, October 19-24, 2008, Atlanta, GA, USA
- [4] Energy-aware probabilistic multiplier: design and analysis, MSK Lau et al, Proceedings of the 2009 international conference on Compilers, architecture, and synthesis for embedded systems
- [5] Krishna V. Palem, Energy Aware Computing through Probabilistic Switching: A Study of Limits, IEEE Transactions on Computers, v.54 n.9, p.1123-1137, September 2005