nano-tera.ch
swiss scientific initiative in health / security / environment systems

IcySoc RTD 2013

FNSNF

Inexact and Approximate Circuits for Error Tolerant Applications

Jérémy Schlachter, Vincent Camus, Christian Enz

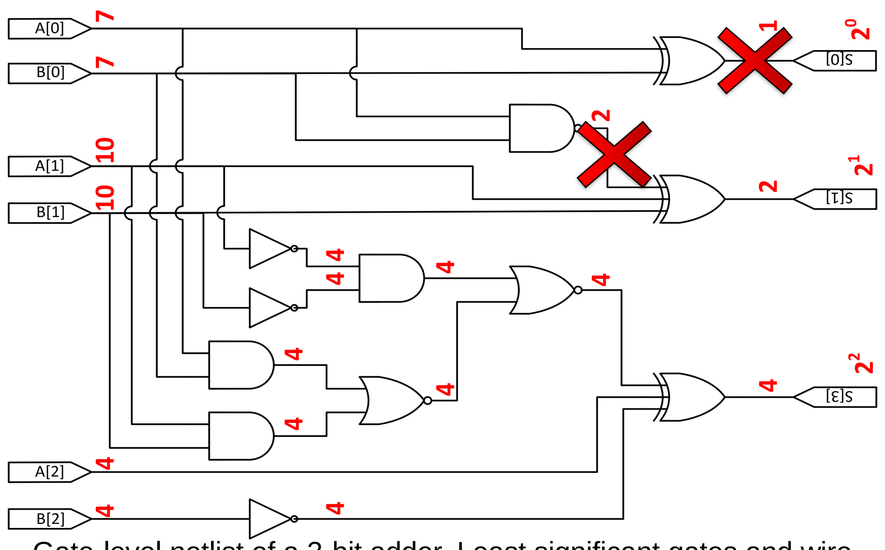
Ecole polytechnique fédérale de Lausanne (EPFL) - Integrated Circuits Laboratory (ICLAB)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Abstract

Inexact or approximate computing is a new paradigm wherein a small accuracy loss can be traded against significant efficiency improvements. While cell phones and other battery powered devices can benefit of this inaccuracy to reduce their energy consumption, high performance computing benefit of a higher number of operation for a given energy budget, as well as a significant speed up of the program execution.

Gate-level Circuit Pruning



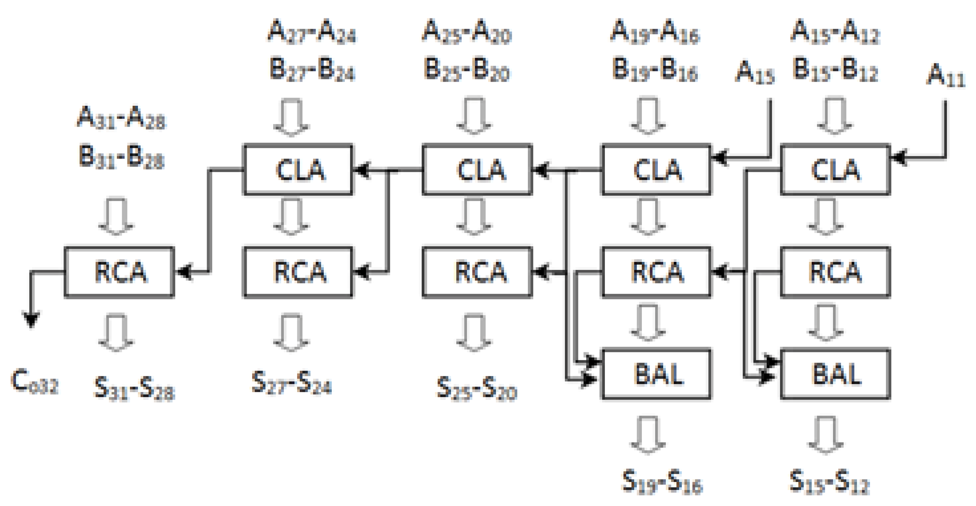
Gate-level netlist of a 3-bit adder. Least significant gates and wire are pruned.

Pruning[1] consists in removing circuit's parts such as full adder cells (coarse grain), or gates (fine grain), that are not often activated, and which do not have a significant impact on the final output. The error is proportional to the number of removed elements.

This technique is compatible with any combinational circuit and can show up to one order magnitude savings in power and area.

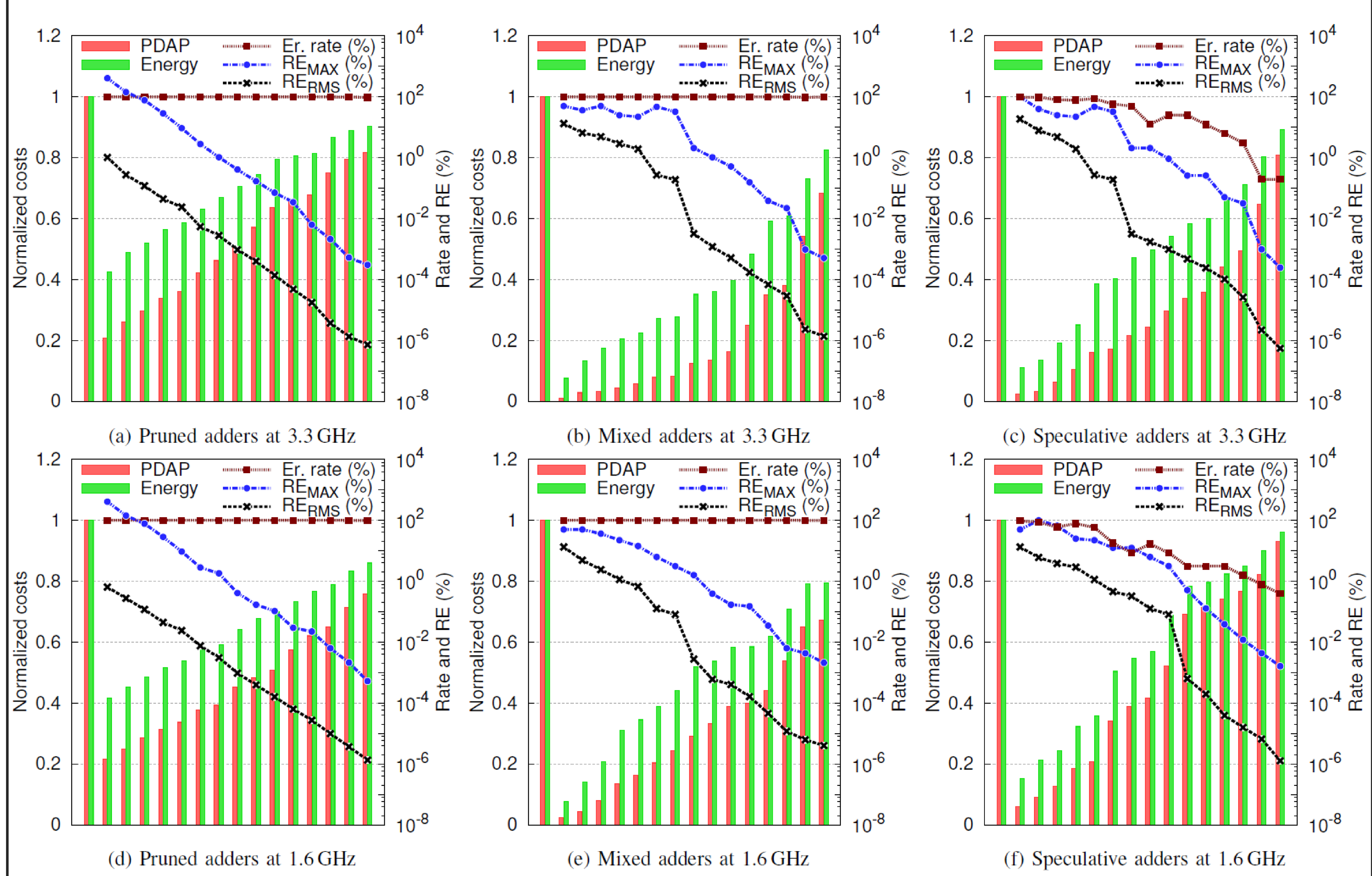
Inexact Speculative Adder

The main idea is to slice the circuit structure and to speculate internal signals from a reduced number of inputs. An additional balancing technique can be used to reduce the potential errors and tune them to application specification. This technique can multiply circuit speed and strongly relax its timing and mapping constraints, allowing huge energy and area savings, up to 73 % Energy-Delay-Area (EDAP) reduction for 0.001 % relative error RMS and 88 % reduction for 1 % relative error RMS.



Speculative adder ETBA [2]

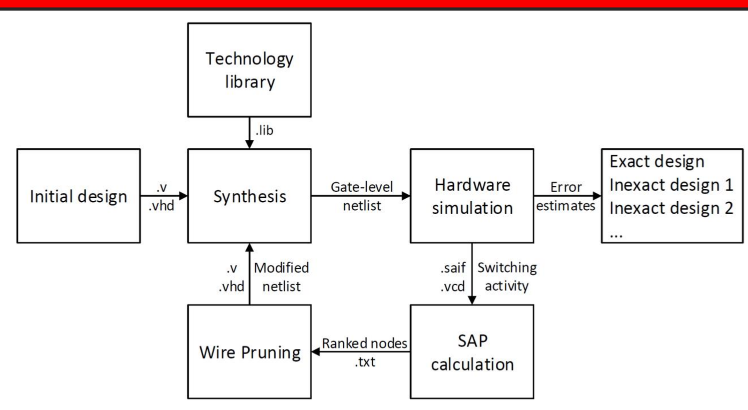
Circuit costs of pruned and speculative adders



Error characteristics and normalized cost of 32-bit pruned, mixed and speculative adders synthesized at 3.3 GHz and 1.6 GHz.

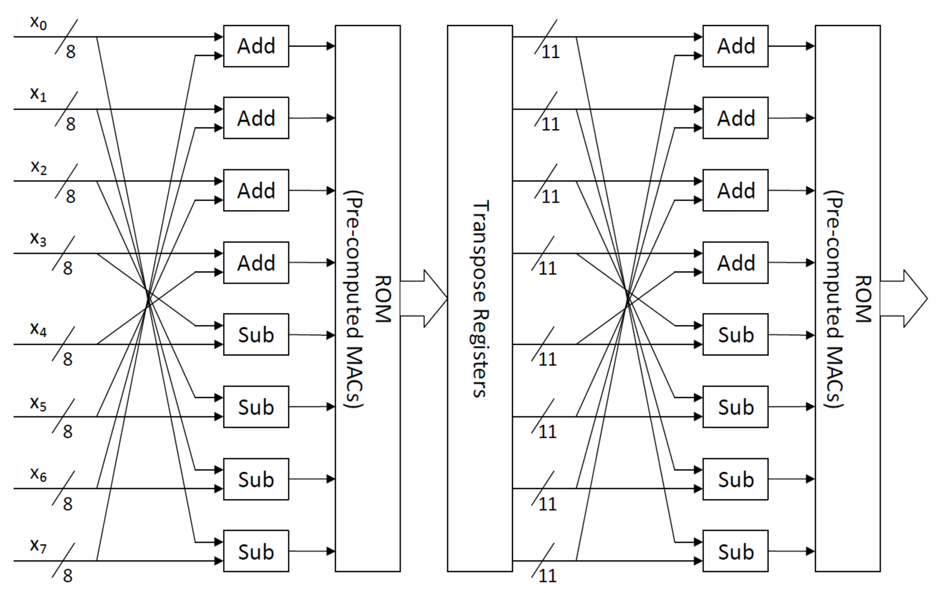
Automatic CAD Flow

Automatic approximate circuits generator, with tunable accuracy for specific applications. This tool is fully integrated in standard digital design flow and compatible with any synthesizable design.



Automatic gate level pruning CAD tool

JPEG encoding with pruned DCT



2D DCT architecture

Discrete Cosine Transform is one of the key building block for many video encoding standard such as JPEG / MPEG and is a perfect candidate for approximations.

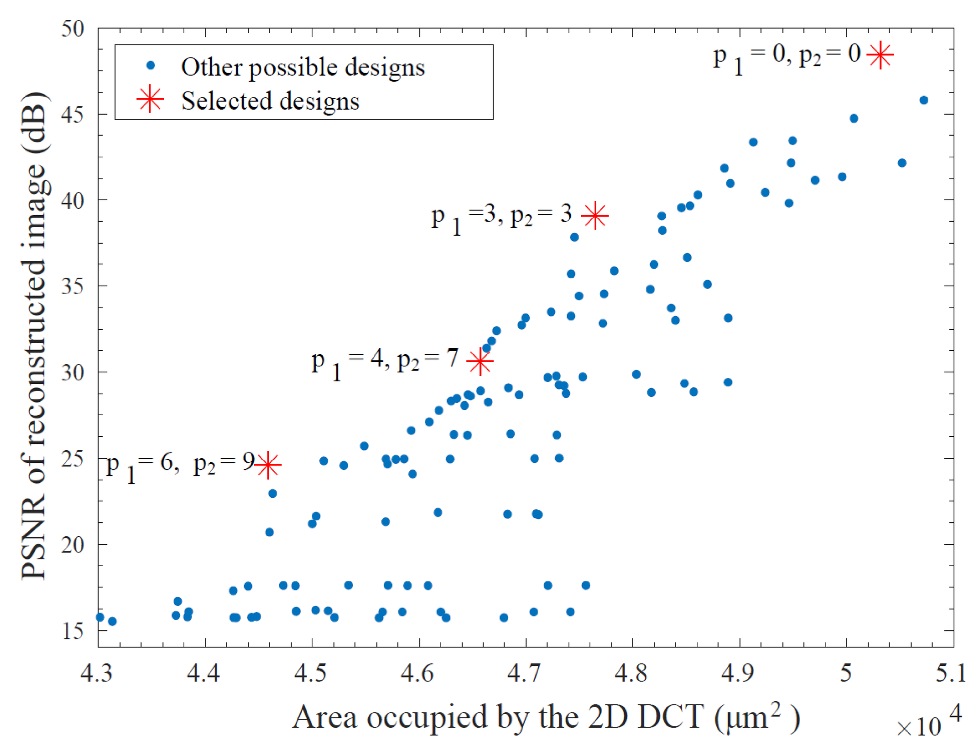


Image quality versus DCT area

Applying Gate-Level Pruning to each of the arithmetic blocks inside the DCT leads to savings of up to 12% area and 10% power on the entire DCT block, even though arithmetic blocks occupy less than 4% of the total area. p_1 denotes the number of pruned nodes in the first stage of the DCT while p_2 is the number of pruned nodes in the second stage.

(a) $p_1 = 0$ $p_2 = 0$ PSNR = 48.4 dB (b) $p_1 = 3$ $p_2 = 3$ PSNR = 39.1 dB (c) $p_1 = 4$ $p_2 = 7$ PSNR = 30.6 dB (d) $p_1 = 6$ $p_2 = 9$ PSNR = 24.6 dB

Pictures of Lena compressed with the exact DCT (a) and the 3 approximate variations (b,c,d)

[1] J. Schlachter *et al.*, Automatic generation of inexact digital circuits by gate-level pruning, *ISCAS*, 2015.
[2] M. Weber *et al.*, Balancing Adder for error tolerant applications, *ISCAS*, 2013.
[3] N. Zhu *et al.*, An enhanced low-power high-speed Adder For Error-Tolerant application, *ISIC*, 2009.

http://iclab.epfl.ch/inexact