Literature Review – Approximate Parallel Prefix Adders

Ananya G

MTech Student, BMSCE, Bangalore.

May 5, 2024

Sujatha K

Associate Professor, BMSCE, Bangalore.

May 5, 2024

Abstract

This paper offers a review of the literature on the theoretical and practical working of Approximate parallel prefix adders (APPAs). Approximate parallel prefix adders are innovative designs for high-speed arithmetic circuits used in digital signal processing and other computational applications. Unlike traditional adders that prioritize exactness, APPAs leverage approximations to achieve substantial performance gains in terms of speed and energy efficiency. This abstract explores the fundamental concepts behind APPAs, highlighting their design principles, operational advantages, and potential applications in modern computing systems. By sacrificing precision to optimize throughput and resource utilization, APPAs represent a promising frontier in the field of parallel prefix computation.

Introduction

1. **Parallel Prefix Adders**
   Basic Arithmetic Circuit Design

In the realm of digital circuit design, parallel prefix adders—also known as carry-save adders or carry-select adders—emerge as critical components for efficiently performing simultaneous addition of multiple binary numbers. This technology finds indispensable application in fields requiring rapid computation of large numbers, including digital signal processing, microprocessor arithmetic, and scientific computations. What distinguishes parallel prefix adders from conventional ripple-carry adders is their ability to compute carry bits in parallel, thereby significantly reducing critical path delays and enhancing overall performance.

A fundamental principle underlying parallel prefix adders is the concept of breaking down addition operations into smaller subproblems and resolving them concurrently—a strategy that harnesses the inherent parallelism inherent to digital circuits. This approach dramatically cuts down the number of stages required for carry propagation, resulting in accelerated addition operations crucial for high-speed applications.

Types of Parallel Prefix Adders

Diverse variants of parallel prefix adders exist, each characterized by its unique design trade-offs and advantages:

- **Ripple-Carry Prefix Adder (RCPA) :** This basic form of parallel prefix adder employs a ripple-carry mechanism for carry propagation. While straightforward in design, it suffers from longer critical path delays, limiting its performance in applications demanding high-speed computations.

- **Brent-Kung Adder [2]:** Proposed by Brent and Kung, this adder adopts a tree structure where addition is executed in parallel at each level of the tree. Despite its increased depth relative to other designs, it achieves a logarithmic speed improvement compared to ripple-carry adders while using fewer logic gates.
- **Kogge-Stone Adder [3]:** Characterized by a broader tree structure, the Kogge-Stone adder offers scalability for larger bit-width additions, promising faster operation and enhanced efficiency.
- **Ladner-Fischer Adder [4]:** Developed by Ladner and Fischer, this adder divides computations into stages where each stage computes both sum and carry bits. It strikes a balance between hardware complexity and speed, making it particularly efficient for parallel addition operations.
- **Sklansky Adder :** Noted for its efficient use of hardware resources, the Sklansky adder minimizes fanout and node count through a recursive structure. By processing input bits in pairs at each stage, it achieves reduced propagation delays and enhanced overall performance.

2. **Approximate Parallel Prefix Adders [1]**
   Innovating Arithmetic Circuit Design
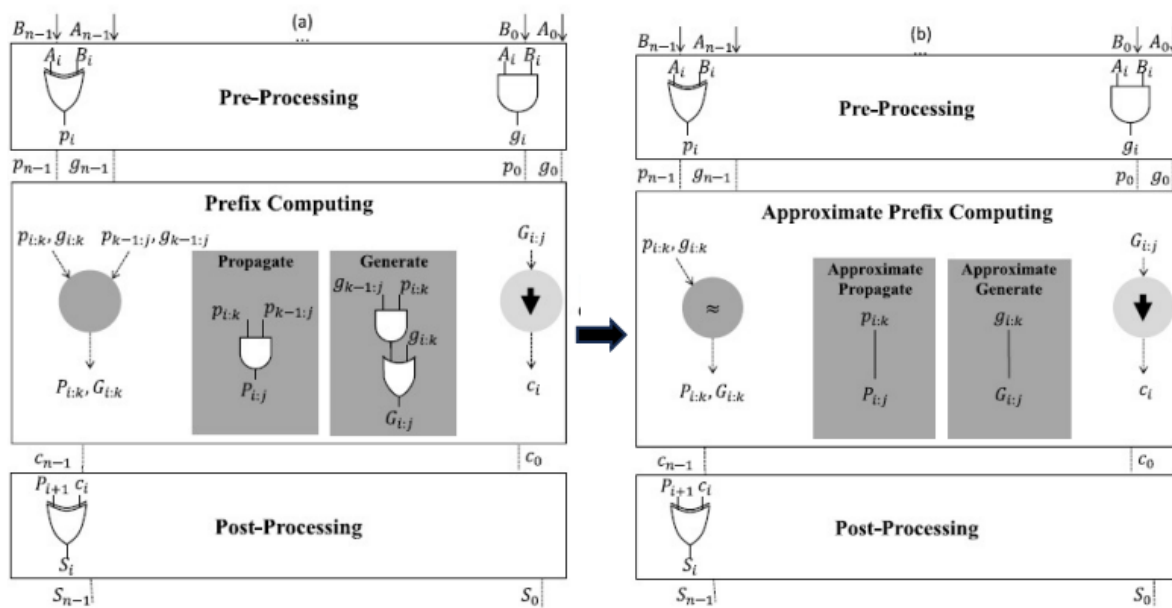
Approximate Parallel Prefix Adder Procedure

In the landscape of arithmetic circuit design, approximate parallel prefix adders introduce a fresh perspective that seeks to strike a balance between computational accuracy, efficiency, and resource utilization. Traditional parallel prefix

[2] R. Brent and H. Kung, "A Regular Layout for Parallel Adders," IEEE Trans. Computers, Vol. C-31, 1982, pp. 260-264.

[3] P. Kogge and H. Stone, "A Parallel Algorithm for the Efficient Solution of a General Class of Recurrence Equations," IEEE Trans. Computers, Vol. C-22, No. 8, 1973, pp. 786-793.

[4] R. Ladner and M. Fischer, "Parallel Prefix Computation," JACM, Vol. 27-4, 1980, pp. 831- 838.

[1] M. M. A. d. Rosa, G. Paim, P. Ü. L. d. Costa, E. A. C. d. Costa, R. I. Soares and S. Bampi, "AxPPA: Approximate Parallel Prefix Adders," in IEEE Transactions on Very Large-Scale Integration (VLSI) Systems, vol. 31, no. 1, pp. 17-28, Jan. 2023, doi: 10.1109/TVLSI.2022.3218021.

adders are renowned for their rapid and efficient addition operations but often demand increased hardware complexity and power consumption to achieve high precision.

In contrast, approximate parallel prefix adders diverge from the traditional norm by prioritizing speed and resource efficiency over absolute precision. These innovative adders leverage approximation techniques to trade a certain degree of accuracy for tangible gains in speed, area efficiency, or reduced power consumption. By relaxing the demand for perfect accuracy, approximate parallel prefix adders open exciting avenues for optimizing arithmetic circuits across diverse applications such as digital signal processing, image processing, and machine learning.

The design ethos behind approximate parallel prefix adders encompasses the integration of approximation algorithms, error estimation mechanisms, and optimization techniques. These components collaborate harmoniously to strike a delicate balance between the desired level of accuracy and the constraints imposed by the target application or hardware platform. Depending on specific requirements, designers can explore an array of approximation strategies— including truncated arithmetic, quantization, or selective precision—to tailor the adder's performance characteristics to the application's unique needs.

A key advantage offered by approximate parallel prefix adders lies in their ability to harness the inherent error resilience observed in many computational tasks. By capitalizing on the tolerance of certain applications to minor errors or inaccuracies, these adders can deliver substantial improvements in performance and resource utilization without jeopardizing overall system functionality. Furthermore, the inherent flexibility embedded in approximate arithmetic empowers designers to navigate nuanced trade-offs between accuracy, speed, and energy efficiency, enabling them to fine-tune the adder's behavior for specific use cases or deployment scenarios. This adaptive approach underscores the dynamic evolution of arithmetic circuit design, ushering in a new era where optimization transcends the confines of conventional precision to embrace the complexities of real-world computational tasks.

Literature Study

In contemporary research endeavors, there has been a discernible trend towards enhancing the computational efficiency of digital signal processing (DSP) blocks by integrating approximate adder These adders, pivotal in various computational kernels of error-tolerant applications like machine learning and signal, image, and video processing, serve as fundamental building blocks for mathematical operations such as subtraction, comparison, multiplication, squaring, and division Within this domain, the parallel prefix adders (PPAs) have garnered considerable attention due to their notable speed and efficiency. PPAs, organized as a parallel prefix graph comprising carry operator nodes known as prefix operators (POs), are specifically optimized for parallelizing carry generation (G) and propagation (P), rendering them among the fastest adders available. In addressing the challenges of integrating approximate adders into DSP blocks, a novel framework is introduced for analytically estimating output quality by conceptualizing the error of approximate adders as an additive noise that perturbs DSP block outputs. This framework employs a signal processing theoretical modeling approach, which characterizes the power of approximation noise through the integral of error spectral density over the bandwidth. Subsequently, the output qualities of DSP blocks employing approximate adders, including finite impulse response (FIR) filters, discrete cosine transforms, and fast Fourier transforms, are estimated using this framework. To validate its accuracy, the proposed framework's predictions are compared with simulation results using the signal-to-noise ratio (SNR) metric, with an observed inaccuracy of less than 2.5dB on average.

Expanding on the application spectrum, the pursuit of low-power requirements in portable multimedia devices has led to the exploration of error resiliency techniques, where voltage over scaling has traditionally been a predominant approach. However, an alternative avenue explored in recent research involves logic complexity reduction at the transistor level, wherein various imprecise or approximate full adder cells are proposed with reduced complexity. These innovations not only lead to reduced switched capacitance but also enable significantly shorter critical paths, facilitating voltage scaling and resulting in substantial power savings, up to 69%, when compared. to accurate adder implementations. Additionally, a novel approximate adder design, integrating error-reduced carry prediction and constant truncation schemes, demonstrates improved computation accuracy and hardware efficiency. Experimental results underscore the efficacy of this approach, with significant reductions observed in power, energy, and area-delay product metrics.

Further extending the discourse, a study focuses on enhancing fixed-point FIR adaptive filters using approximate distributed arithmetic (DA) circuits. Here, the integration of the radix-8 Booth algorithm is instrumental in reducing partial product numbers in the DA architecture, while partial products are generated approximately through data truncation with error compensation. Synthesis results affirm significant reductions in delay, area, and power consumption compared to accurate designs, indicating promising applications in system identification and saccadic systems. Moreover, a proposed synthesis of approximate adders aims to enhance the area and energy efficiency of FIR filters implemented in CMOS. This approach, validated through extensive experimentation, showcases notable energy per sample savings and hardware area reduction without compromising filter frequency response or signal-to-noise ratio.

Lastly, speculative adders, particularly speculative carry select addition (SCSA), have emerged as a compelling avenue for reducing critical path delays and achieving high performance with low area overhead. Analytical modeling of SCSA error rates facilitates design exploration, demonstrating superior performance compared to traditional adders in terms of speed and area reduction. Collectively, these advancements underscore the potential of approximate adders and related techniques in optimizing computational efficiency and power consumption across various DSP applications, including multimedia processing and error-resilient systems.

In recent years, there has been a growing interest in approximate computing as a means to trade off computational accuracy for improved speed and power efficiency. Various studies have explored the implementation of approximate circuits in digital signal processing (DSP) blocks to enhance performance while minimizing energy consumption.

One notable approach involves utilizing approximate adders and multipliers in the design of efficient digital filters, such as finite impulse response (FIR) filters and discrete cosine transforms (DCTs). These designs often leverage techniques like distributed arithmetic (DA) circuits and radix-8 Booth algorithms to reduce hardware costs and achieve significant reductions in delay, area, and power consumption. Moreover, approximate computing has found application in error-tolerant systems, particularly in domains like multimedia and signal processing. For instance, researchers have proposed approximate hardware architectures for tasks such as electrocardiogram (ECG) signal processing, demonstrating substantial energy efficiency gains while maintaining acceptable performance levels. Similarly, investigations into voltage over-scaling (VOS) have highlighted its potential for optimizing energy efficiency in hardware accelerators, particularly in closed-loop algorithms [16]. However, understanding the propagation of timing errors induced by VOS throughout the system remains a critical challenge.

Approximate arithmetic units, such as logarithmic multipliers and partial product generators, have been extensively studied to further enhance energy efficiency and performance in digital designs. By employing techniques like iterative approximation and novel compressor structures, researchers have achieved significant reductions in power consumption while minimizing accuracy losses. These advancements have led to the development of optimized designs for specific applications, such as squares and fast Fourier transform (FFT) accelerators.

In the context of video encoding, approximate computing has been leveraged to improve the efficiency of critical operations like sum of absolute differences (SAD) calculation. By integrating adder compressors into SAD hardware architectures, researchers have achieved notable reductions in power dissipation while maintaining real-time processing capabilities for high-definition video streams.

Furthermore, in the domain of neural network (NN) inference, approximate computing principles have been integrated into accelerator designs to mitigate energy consumption without significant loss of accuracy. Techniques such as weight-oriented approximation mapping enable run-time adjustment of accuracy levels, resulting in substantial energy savings for NN-based applications. Overall, the exploration of approximate computing techniques across various domains underscores their potential to address the energy-efficiency challenges faced by modern digital systems. From digital filters to neural network accelerators, approximate computing offers a versatile framework for balancing computational accuracy with performance and power efficiency, paving the way for future advancements in hardware design and optimization.

Conclusions and Discussion per Capstone:

In conclusion, the integration of approximate parallel prefix adders (AxPPAs) represents a significant advancement in enhancing computational efficiency and energy savings within error-tolerant applications like machine learning, signal processing, and multimedia tasks. Our base paper introduces AxPPAs by leveraging approximations in prefix operators (POs) of parallel prefix adders (PPAs), resulting in architectures such as AxPPA-BK, AxPPA-KS, AxPPA-LF, and AxPPA-SK. Through extensive evaluation, including stand-alone testing and integration into signal processing kernels like sum of squared differences (SSDs) video accelerators and finite impulse response (FIR) filters, our proposed AxPPA-LF demonstrates superior performance in energy-quality and area-quality trade-offs compared to existing energy-efficient approximate adders (AxAs). Specifically, the AxPPA-LF achieves exceptional savings in circuit area and power while maintaining high-quality standards across a range of approximation levels. This work highlights the potential of AxPPAs to significantly optimize digital system performance, demonstrating compelling benefits in terms of energy efficiency and computational quality in both generic and application-specific scenarios.

References

B. Silveira et al., "Power-Efficient Sum of Absolute Differences Hardware Architecture Using Adder Compressors for Integer Motion Estimation Design," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 64, no. 12, pp. 3126-3137, Dec. 2017, doi: 10.1109/TCSI.2017.2728802

D. Esposito, A. G. M. Strollo, E. Napoli, D. De Caro and N. Petra, "Approximate Multipliers Based on New Approximate Compressors," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 65, no. 12, pp. 4169-4182, Dec. 2018, doi: 10.1109/TCSI.2018.2839266

G. Paim, H. Amrouch, E. A. C. d. Costa, S. Bampi and J. Henkel, "Bridging the Gap Between Voltage Over-Scaling and Joint Hardware Accelerator-Algorithm Closed-Loop," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 1, pp. 398-410, Jan. 2022, doi: 10.1109/TCSVT.2021.3059229.

H. Jiang, L. Liu, P. P. Jonker, D. G. Elliott, F. Lombardi and J. Han, "A High-Performance and Energy-Efficient FIR Adaptive Filter Using Approximate Distributed Arithmetic Circuits," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 1, pp. 313-326, Jan. 2019, doi: 10.1109/TCSI.2018.2856513

H. B. Seidel, M. M. A. da Rosa, G. Paim, E. A. C. da Costa, S. J. M. Almeida and S. Bampi, "Approximate Pruned and Truncated Haar Discrete Wavelet Transform VLSI Hardware for Energy-Efficient ECG Signal Processing," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 68, no. 5, pp. 1814-1826, May 2021, doi: 10.1109/TCSI.2021.3057584.

J. Lee, H. Seo, H. Seok, and Y. Kim, "A novel approximate adder design using error reduced carry prediction and constant truncation," IEEE Access, vol. 9, pp. 119939–119953, 2021.

K. Du, P. Varman, and K. Mohanram, "High performance reliable variable latency carry select addition," in Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE), Mar. 2012, pp. 1257–1262.

K. C. Pathak, A. D. Darji and J. N. Sarvaiya, "Low power Dadda multiplier using approximate almost full adder and Majority logic based adder compressors," 2022 IEEE Region 10 Symposium (TENSYMP), Mumbai, India, 2022, pp. 1-6, doi: 10.1109/TENSYMP54529.2022.9864428.

L. B. Soares, S. Bampi, and E. Costa, "Approximate adder synthesis for area- and energy efficient FIR filters in CMOS VLSI," in Proc. IEEE 13th Int. New Circuits Syst. Conf. (NEWCAS), Jun. 2015, pp.

L. Shetkar and G. Hegde, "Design of Radix-8 Based Approximate Booth Squarer for Error Tolerant Computing," 2023 International Conference on Next Generation Electronics (NEleX), Vellore, India, 2023, pp. 1-6, doi: 10.1109/NEleX59773.2023.10421496.

M. M. A. d. Rosa, G. Paim, P. Ü. L. d. Costa, E. A. C. d. Costa, R. I. Soares and S. Bampi, "AxPPA: Approximate Parallel Prefix Adders," in IEEE Transactions on Very Large-ScaleIntegration (VLSI) Systems, vol. 31, no. 1, pp. 17-28, Jan. 2023, doi: 10.1109/TVLSI.2022.3218021.

M. Pashaeifar, M. Kamal, A. Afzali-Kusha and M. Pedram, "A Theoretical Framework for Quality Estimation and Optimization of DSP Applications Using Low-Power Approximate Adders," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 1, pp. 327-340, Jan. 2019, doi: 10.1109/TCSI.2018.2856757

M. Pashaeifar, M. Kamal, A. Afzali-Kusha and M. Pedram, "A Theoretical Framework for Quality Estimation and Optimization of DSP Applications Using Low-Power Approximate Adders," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 1, pp. 327-340, Jan. 2019, doi: 10.1109/TCSI.2018.2856757.

P. T. L. Pereira et al., "Energy-Quality Scalable Design Space Exploration of Approximate FFT Hardware Architectures," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 69, no. 11, pp. 4524-4534, Nov. 2022, doi: 10.1109/TCSI.2022.3191180

P. Kogge and H. Stone, "A Parallel Algorithm for the Efficient Solution of a General Class of Recurrence Equations," IEEE Trans. Computers, Vol. C-22, No. 8, 1973, pp. 786-793.

R. Ladner and M. Fischer, "Parallel Prefix Computation," JACM, Vol. 27-4, 1980, pp. 831- 838.

R. Brent and H. Kung, "A Regular Layout for Parallel Adders," IEEE Trans. Computers, Vol. C-31, 1982, pp. 260-264

Sanzhar Yergaliyev, Muhammad Tahir Akhtar, "A Systematic Review on Distributed Arithmetic-Based Hardware Implementation of Adaptive Digital Filters", IEEE Access, vol.11, pp.85165-85183, 2023.

V. Gupta, D. Mohapatra, A. Raghunathan and K. Roy, "Low-Power Digital Signal Processing Using Approximate Adders," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 32, no. 1, pp. 124-137, Jan. 2013, doi: 10.1109/TCAD.2012.2217962.

W. Liu, J. Xu, D. Wang, C. Wang, P. Montuschi and F. Lombardi, "Design and Evaluation of Approximate Logarithmic Multipliers for Low Power Error-Tolerant Applications," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 65, no. 9, pp. 2856-2868, Sept. 2018, doi: 10.1109/TCSI.2018.2792902

Z. G. Tasoulas, G. Zervakis, I. Anagnostopoulos, H. Amrouch and J. Henkel, "Weight Oriented Approximation for Energy-Efficient Neural Network Inference Accelerators," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 67, no. 12, pp. 4670-4683, Dec. 2020, doi: 10.1109/TCSI.2020.3019460