A Novel Analog Mixed-Signal Multiplier and Corresponding Dot Product Calculation Circuit

TLB
Technologie-Lizenz-Büro
der Baden-Württembergischen
Hochschulen GmbH

An innovative Analog Mixed-Signal (AMS) Multiply-Accumulate Circuit (MAC) with a nonlinear transfer function offers energy-efficient calculations for Artificial Neural Networks (ANNs), overcoming the limitations of digital logic gates in AI applications and beyond.

- Extremely low power consumption (more durable and more environment friendly)
- Disruptive but still can be implemented on standard commercial CMOS technologies.
- Analog coprocessor IP can be placed together with other analog or digital blocks on a single CMOS ASIC

Fields of application

Circuits, electronics, and edge AI and IOT

Kontakt
Dr. Frank Schlotter
TLB GmbH
Ettlinger Straße 25
76137 Karlsruhe | Germany
Telefon +49 721-79004-0
schlotter@tlb.de | www.tlb.de

Entwicklungsstand

TRL 2 – Basic Technology Research

Patentsituation

DE 102020113088 A1 pending
EP21728011.4 pending
US 20230185530 A1 pending
Priority date: 14/ May/2020

Referenznummer

19/095TLB

Service

Technologie-Lizenz-Büro GmbH has been entrusted with exploiting this technology and assisting companies in obtaining licenses.

Background

Today's electronic system design relies on CMOS integrated circuits and binary logic for Digital Signal Processing (DSP), known for their robust noise margin between logic high and low levels.

However, the surge in demand for processing power, driven by Artificial intelligence (AI) and Deep Learning (DL) methods integrated into various devices and applications, is outpacing Moore's Law. This threatens the energy efficiency of traditional digital signal processing, especially at the internet edge.

By using only two MOSFETs for multiplication, this innovative analog solution offers at least a tenfold energy saving compared to a digital CMOS field multiplier using static CMOS gates, is compatible with existing CMOS technologies, and presents a novel AMS circuit design for neural network layers. It can also be seamlessly integrated with analog, Radio Frequency (RF), digital, and memory components on a single CMOS "System-on-Chip" approach.

Problem

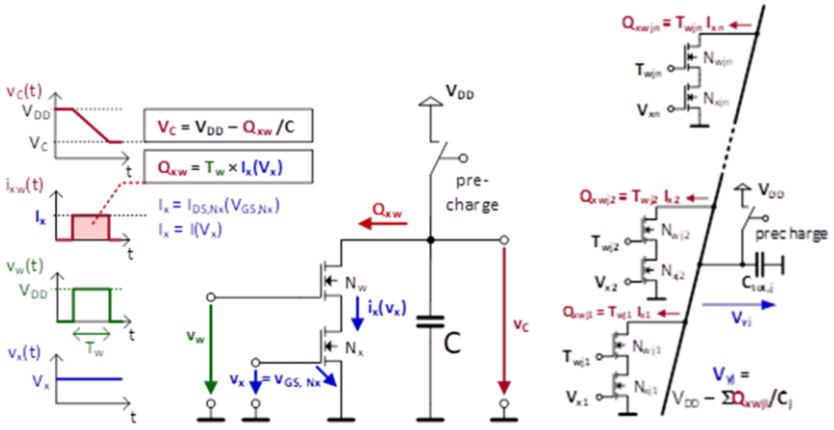
The predominant method to implement digital operations is to use static logic circuits on integrated circuits (ICs) that are based on CMOS technology with MOSFETs. There are two main advantages of static CMOS logic: First, the power consumption is mainly dynamic and thus proportional to the logical activity of the circuit. Second, the power consumption has strongly decreased with technology scaling. CMOS technology has strongly profited from the semiconductor technology's smallest feature size downscaling. The progressive downscaling driven by a technology roadmap following "Moore's Law" lasted over a period of nearly 60 years. Thus, technology-driven energy efficiency gains balanced the ever-increasing processing power demand until today. The outcomes are acceptable battery lifetimes of mobile devices and an acceptable ratio of electronic signal processing energy dissipation to the total human energy consumption. However, the traditional method of MOSFET feature size downscaling (i.e., "Dennard Scaling") has already passed, resulting in processor clock frequency stagnation at 2 to 3 GHz from about 2005, and the end of the MOSFET device scaling era is in sight. On the other side, the growth of signal processing power demand currently even accelerates due to the integration of processing-intensive artificial intelligence (AI) and deep learning (DL) methods to a rising number of devices and applications like speech recognition, automatic translation services, automatic face recognition, autonomous driving etc. Especially the deep neural network (DNN) compute power requirements increase much faster than the scaling gains from the underlying CMOS technology. Particularly, inference (i.e. the application of a trained network to new data) on mobile or autonomous devices with limited power budget needs to be revised. This can be seen as a challenge for all across the battery supply chain, from mining critical minerals to recycling. In summary, the foreseeable end of Moore’s Law and the following stagnation of the energy efficiency of the

dominating digital signal processing is putting a threat to the widespread deployment of upcoming signal processing intensive AI methods, especially at the internet edge. Disruptive circuit technology or circuit topology innovations are urgently required to keep the growth speed of AI deployment in a sustainable manner.

Solution

Researchers at the Institute of Electrical and Optical Communications Engineering of the University of Stuttgart have developed a novel, disruptive, patent-pending innovation to solve these problems by challenging the (almost) exclusive use of digital two-level binary signals. According to this invention, circuits employ an analog-mixed-signal processing approach where selected electrical nodes carry significantly more information in the analog domain, physically limited only by device noise and leakage. The invention allows executing complex arithmetic operations on single capacitive circuit nodes, involving as few devices as possible and drastically reducing the number of required circuit nodes.

The core innovation here is an Analog Mixed-Signal (AMS) Multiply-Accumulate Circuit (MAC) with a nonlinear transfer function. It comprises two series-connected Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) serving as analog multipliers and a capacitor for charge accumulation. The first multiplicand is represented as an analog voltage applied to the first MOSFET gate, while the second multiplicand is expressed by the temporal pulse width of a voltage pulse connected to the second MOSFET gate, that acts as a switch for the current supplied by the first MOSFET. A charge packet, proportional to the product of the multiplicands, i.e., the first MOSFET current and the pulse duration, accumulates on the capacitor that is connected in parallel to the two stacked MOSFET branches (see Figure left). By connecting two more MOSFET stacks in parallel to the capacitor, an efficient analog dot product calculation with minimal devices is enabled (see Figure right). This innovative AMS circuit configuration can efficiently evaluate neuron layers in Artificial Neural Networks (ANNs).



Figures: Left: The basic principle of the two-MOSFET analog multiplier with a capacitor for charge accumulation Right: Parallel connection of several two-MOSFET multiplier branches for analog dot product calculation [Source: German patent application DE 102020113088 A1]

Literature and links

R. Nägele, J. Finkbeiner, M. Grözing and M. Berroth, "Design of an Energy Efficient Analog Two-Quadrant Multiplier Cell Operating in Weak Inversion," 2022 20th IEEE Interregional NEWCAS Conference (NEWCAS), Quebec City, QC, Canada, 2022, pp. 5-9, doi: 10.1109/NEWCAS52662.2022.9842251

J. Finkbeiner, R. Nägele, M. Grözing and M. Berroth, "Design of an Energy Efficient Voltage-to-Time Converter with Rectified Linear Unit Characteristics for Artificial Neural Networks," 2022 20th IEEE Interregional NEWCAS Conference (NEWCAS), Quebec City, QC, Canada, 2022, pp. 327-331, doi: 10.1109/NEWCAS52662.2022.9842074

R. Nägele, J. Finkbeiner, V. Stadtlander, M. Grözing and M. Berroth, "Analog Multiply-Accumulate Cell With Multi-Bit Resolution for All-Analog AI Inference Accelerators," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 70, no. 9, pp. 3509-3521, Sept. 2023, doi: 10.1109/TCSI.2023.3268728