A Scalable and Instantaneously Wideband RF Correlator Based on Margin Computing

Kareem Rashed, *Graduate Student Member, IEEE*, Aswin Undavalli, *Graduate Student Member, IEEE*, Shantanu Chakrabartty, *Senior Member, IEEE*, Aravind Nagulu, *Member, IEEE*, and Arun Natarajan, *Senior Member, IEEE*

***Abstract*— Correlation is a fundamental operation in radar/communication signal processing. Enabling efficient reconfigurable correlation at high frequencies with wide bandwidths and high dynamic range is a fundamental challenge. This article presents a direct-RF wideband analog correlator that utilizes a novel margin computation paradigm replacing traditional multiply-and-accumulate (MAC) with analog addition and thresholding to enable energy-efficient analog computation. A high-efficiency charge-domain realization is proposed and implemented in 65-nm CMOS occupying 0.97 mm^2. The analog correlator IC supports 5-GS/s inputs, a large correlation length of 1024, and 8-bit computing accuracy with a high energy efficiency of 152 TOPS/W. Practical applications of typical high-speed correlation are also demonstrated through system-level measurements such as radar signal detection, and code-domain processing.**

***Index Terms*— Analog computing, approximate computing, code-domain radars, correlation, inner-product, multiplier-free, RF sensing, spread spectrum communication.**

I. Introduction

CORRELATION is a fundamental operation at the heart of signal-processing techniques in radar and communications (Fig. 1) [1], [2], [3], [4], [5], [6], [7], [8]. For instance, in a spread spectrum radar, the distance to a target can be measured by correlating the received signal with the transmitted template. Matched filters in communication systems can also be viewed as a correlation with a template and are core blocks in direct-sequence code division multiple access (DS-CDMA) systems, where signals modulated with a code can be selectively received by correlating the received signal with the same code. Similarly, in global positioning system (GPS), signal acquisition is carried out by correlating received CDMA signals to recover precise timing and positioning [9], [10].

Received 6 May 2024; revised 31 July 2024 and 3 September 2024; accepted 8 September 2024. Date of publication 3 October 2024; date of current version 24 October 2024. This article was approved by Associate Editor Rabia Tugce Yazicigil. This work was supported in part by the National Science Foundation under Award 2128535 and in part by the Defense Advanced Research Projects Agency (DARPA) Massive Cross Correlation (DARPA MAX) Program under Award FA8650-23-2-7309. *(Corresponding author: Kareem Rashed.)*

Kareem Rashed and Arun Natarajan are with the School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR 97330 USA (e-mail: rashedk@oregonstate.edu).

Aswin Undavalli, Shantanu Chakrabartty, and Aravind Nagulu are with the Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130 USA (e-mail: nagulu@wustl.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSSC.2024.3465244.

Digital Object Identifier 10.1109/JSSC.2024.3465244

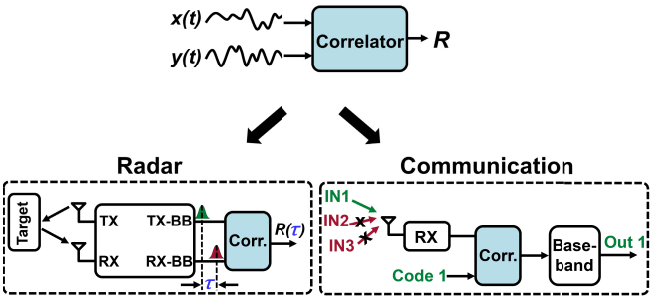

Fig. 1. Correlators are fundamental building blocks in radar and communication signal processing, where correlation can be broadly generalized as computing the cross correlation between two signals, $x(t)$ and $y(t)$.

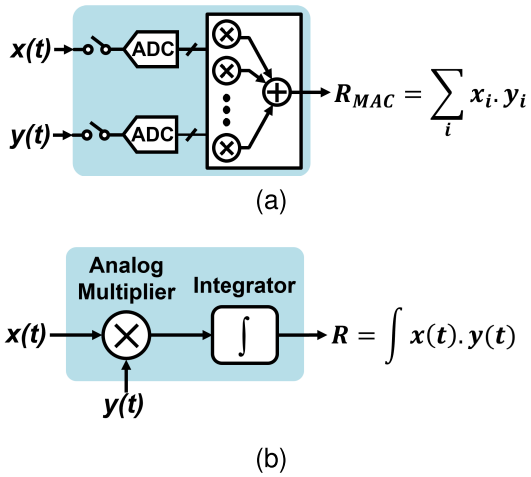

Fig. 2. Correlation can be performed in (a) digital domain through successive MAC operation (requiring ADCs) and (b) analog domain employing a mixer or a multiplier followed by an integrator, potentially requiring high power/large area for long correlator lengths.

Correlation with known templates, auto-correlation, and cross correlation can be generalized as cross correlation between two signals, which is typically performed, as shown in Fig. 2(a), in the digital domain using successive multiply-and-accumulate (MAC) operations [7], [8], [11], [12], [13]. Digital implementation offers high dynamic range, scalability, and power efficiency that improve with process scaling, albeit with higher implementation costs. However, the analog-to-digital converter (ADC) power consumption reduces overall system energy efficiency, especially for large bandwidths [14]. Additionally, digital MAC power consumption increases with frequency and compute length.

0018-9200 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

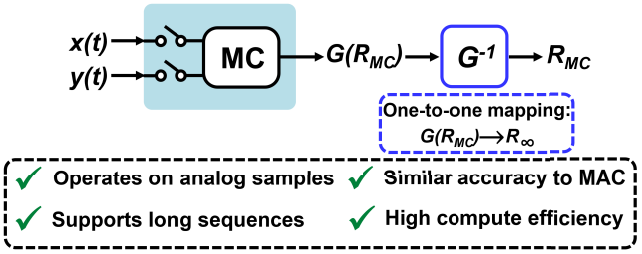

Fig. 3. Correlation can be estimated using multiplier-free approximation with performance similar to MAC correlation.

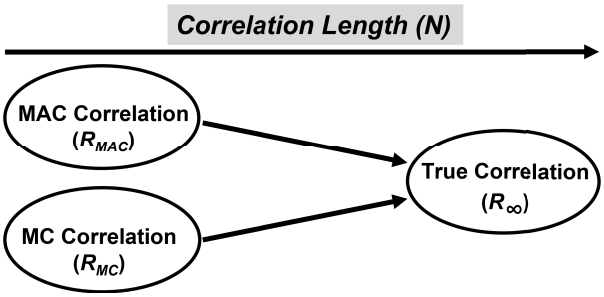

Fig. 4. Correlators (including digital MAC schemes) converge to the true cross correlation between two random sequences over long sequence lengths; alternate multiplier-free correlators can also exhibit such convergence.

Analog correlators, typically implemented, as shown in Fig. 2(b), using a mixer or multiplier followed by an integrator, can operate directly on RF inputs eliminating high-speed ADCs [1], [2], [3], [15], [16]. However, multiplier-based analog correlators suffer from high power/area, have short correlation lengths, and do not scale well to advanced process nodes [17], [18], [19]. In addition, analog correlators that employ charge-coupled devices (CCD) [5] or surface-acoustic wave (SAW) filters [20] are difficult to integrate in standard CMOS technologies.

Importantly, the correlation between two signals can be viewed as a computation resembling pattern matching, where the redundancy within the ensemble provides resilience to approximation errors. This has motivated the use of approximation techniques for efficient low-power computation. For example, compute in-memory (CiM) digital-domain approximation for the MAC operation based on $L1$-distance is employed in [21], achieving ~100 TeraOps/s/W (TOPS/W). Similarly, an approximate arithmetic hardware based on digital OR/AND operation is adopted in [22]. Although these approximate CiM approaches achieve high compute efficiency, they still suffer from the ADC/digital-to-analog converter (DAC) overhead at the input/output interface. In addition, they operate with lower speeds for multibit inputs and require new materials/memory IP [23], [24], [25], [26]. It is noteworthy that practically viable analog-domain approximation in correlators must 1) accommodate long sequences (>1000); 2) exhibit improved accuracy with longer sequences; and 3) achieve high compute energy efficiency comparable to digital MAC while eliminating input ADC.

In this article, we present a direct-RF wideband margin-computing (MC)-based analog correlator (Fig. 3) in standard 65-nm CMOS that achieves 1) correlation of instantaneously wideband RF inputs (dc-2.5 GHz); 2) large correlation length of 1024 in the analog domain (>100× better than prior work); 3) 8-bit computing accuracy (a.k.a. hardware-dynamic range of 50.3 dB); and 4) high compute efficiency of 152 TOPS/W traditionally provided only by digital-intensive compute schemes. We also demonstrate system-level measurements/applications such as radar signal detection and code-domain processing. We have expanded on the work in [27], including a complete description of system architecture, extended analysis, and implementation details of key blocks.

Section II presents the MC-based correlation approach that builds on margin-propagation theory in [28], [29], [30], [31], and [32]. Section III describes a low-power implementation of the MC-correlator in the charge domain. In Section IV, circuit-level implementations of key building blocks are presented. Section V discusses the non-idealities associated with the MC-correlator implementation and their effect on the computation precision. Measured performance is discussed in Section VI. Finally, this work is summarized in Section VII.

II. Correlation Using Margin-Computing Paradigm

Generally, cross correlation (R) between two time-domain signals $x(t)$ and $y(t)$ is defined over a finite period of time T as

$$R = \frac{1}{T}\int_0^T x(t)y(t)\,dt \tag{1}$$

with the true cross correlation between the two signals, R_∞, defined over infinite time ($R \to R_\infty$ when $T \to \infty$).

As shown in Fig. 3, analog correlators can be generalized as a function operating in the analog sample domain to generate an output that is a monotonic function of the cross correlation, with a subsequent inverse mapping function to recover the correlation. The true cross correlation R_∞ between two random input sequences, X and Y, can only be determined over long sequence lengths. For a given sequence pair, even MAC correlators only converge to R_∞ as sequence length is increased, with estimation errors for finite lengths depending upon sequence periodicity and probability distributions (i.e., for finite sequence length $R_{\text{MAC}} \neq R_\infty$). Accordingly, as shown in Fig. 4, a non-MAC-based approximation could potentially show similar performance to an ideal MAC approximation for finite sequence lengths while converging to the true correlation following a different convergence path than MAC. In this work, the correlation between input sequences X and Y is estimated using a computation paradigm we call "margin compute" that uses only analog addition and thresholding.

A. Margin Computing

In the following, we elaborate on the proposed MC approach, which relies on the margin-propagation theory outlined in [28], [29], [30], [31], and [32]. In machine learning, a margin refers to the distance of a data point from a decision boundary, where the decision boundary can be considered as a thresholding level. Such a margin can be defined for an input

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

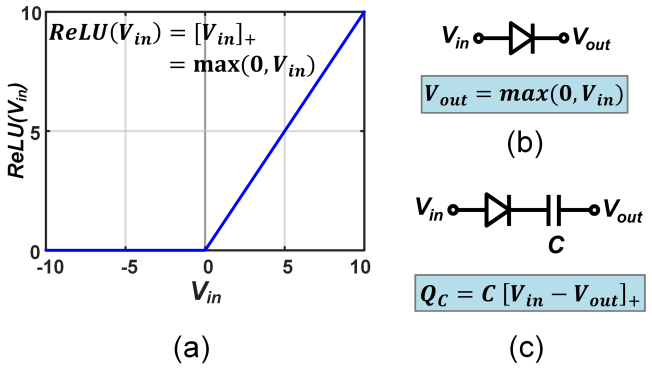

Fig. 5. (a) Definition of the ReLU at the core of the proposed correlator. (b) Since ReLU can be easily realized using diodes on-chip. (c) Diode-capacitor circuit can realize a ReLU-based margin computation.

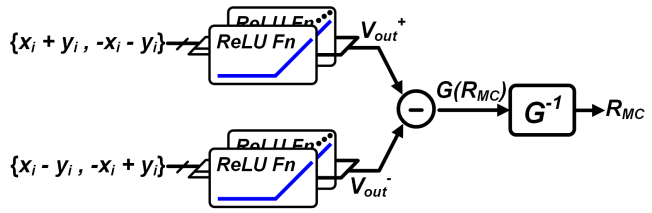

Fig. 6. Conceptual diagram of MC approach in which the correlation between sequence X and sequence Y is computed using ReLU function that operates on additive/subtractive combinations of X and Y inputs.

sequence. For example, assuming a normalization level, V_{norm}, the margin for an input, $V_{\text{in},k}$, can be defined as

$$\begin{aligned}\text{Margin}_k &= \left[V_{\text{in},k} - V_{\text{norm}}\right]_+ \\ &= \max\left(0,\, V_{\text{in},k} - V_{\text{norm}}\right) = \text{ReLU}\left(V_{\text{in},k}\right) \end{aligned} \tag{2}$$

with the ReLU function being defined for a threshold V_{norm}. In the following, we show how a computation that constrains the total margin across a set of inputs can be used to find correlation.

B. Proposed MC-Based Correlator

The MC-correlator uses a thresholding-based non-linear function that operates on four-quadrant operands generated from the inputs, such as $\pm X \pm Y$ [33]. While various possibilities exist for the non-linear function, the thresholding-based rectified linear unit (ReLU) function, shown in Fig. 5(a), is well-suited for low-power analog-friendly implementations. As shown in Fig. 5(b), ReLU function can be easily realized using diodes on-chip. As noted in Fig. 5, the output voltage of an ideal diode is the maximum of zero and V_{in}, which matches the ReLU definition.

This can be implemented using a diode-capacitor circuit, as shown in Fig. 5(c), with the circuit output being the normalization level.

The correlation between two input sequences can be estimated using the ReLU-based margin computation across four-quadrant inputs as shown in Fig. 6. In this ReLU-based correlation scheme, four sets of operands are applied in parallel as inputs for the margin calculation. Accordingly, the normalization levels across the parallel margin computations,

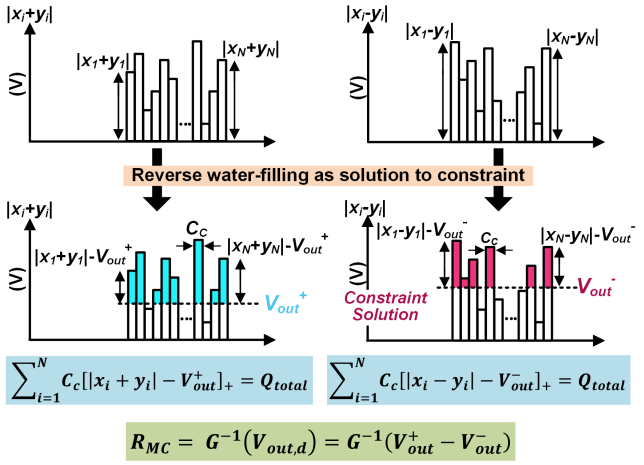

Fig. 7. Proposed MC-correlation uses ReLU function that operates on four sets of inputs $(X+Y, -X-Y)$ and $(X-Y, -X+Y)$ to estimate correlation. Reverse water-filling solves for the output voltage for each pair of sets, V_{out}^{+} and V_{out}^{-}, where correlation is proportional to the differential output $V_{\text{out}}^{+} - V_{\text{out}}^{-}$.

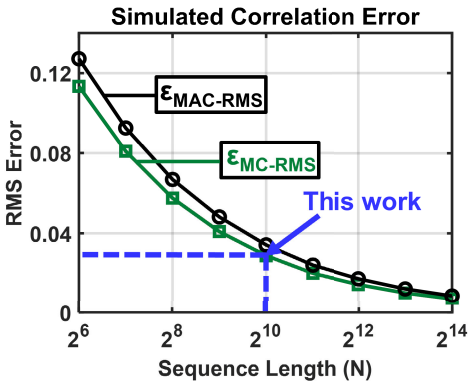

Fig. 8. Simulations demonstrating the proposed MC-correlation approach estimate correlation with improving accuracy as sequence length increases, with performance comparable to MAC correlators.

shown in Fig. 6 as V_{out}^{+} and V_{out}^{-}, estimates the inner (dot) product between the input sequences and hence the correlation. The estimated correlation based on MC, R_{MC}, can be determined based on an inverse mapping function that operates on the differential output $V_{\text{out},d}$ and maps it to R_{MC}.

Analytically, the MC-correlation single-ended outputs can be expressed in terms of the input sequences elements x_i and y_i as

$$\begin{aligned} V_{\text{out}}^{+} \quad \text{s.t.} \quad \sum_{i=1}^{N} \text{ReLU}\left(|x_i + y_i| - V_{\text{out}}^{+}\right) = \gamma \\ V_{\text{out}}^{-} \quad \text{s.t.} \quad \sum_{i=1}^{N} \text{ReLU}\left(|x_i - y_i| - V_{\text{out}}^{-}\right) = \gamma \end{aligned} \tag{3}$$

where γ is a hyperparameter. Consequently, R_{MC} can be expressed as

$$R_{\text{MC}} = G^{-1}\left(V_{\text{out}}^{+} - V_{\text{out}}^{-}\right) \tag{4}$$

where G^{-1} is the monotonic one-to-one mapping function.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

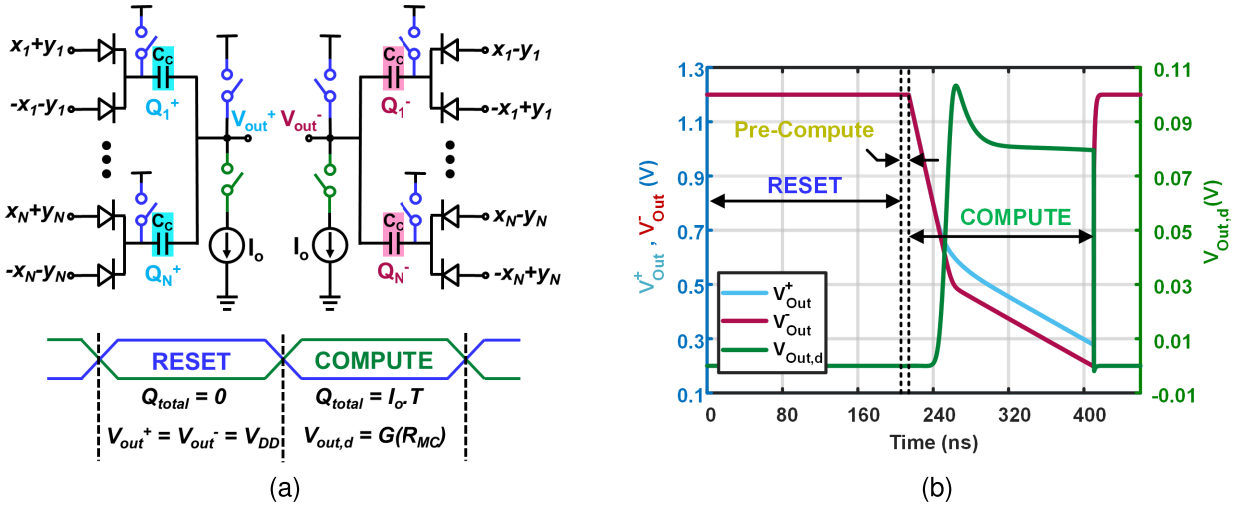

Fig. 9. (a) Energy-efficient diode-capacitor circuit to estimate correlation by performing the MC-computation outlined in Fig. 7 in the charge domain. (b) Simulation example of the charge-thresholding correlator for random inputs with $R_\infty = 1$.

III. Charge-Domain Thresholding for Margin-Compute-Based Correlation

As shown in Fig. 7, the ReLU-based MC is analogous to a reverse water-filling problem of finding an output voltage for a given set of inputs when the total charge, Q_{total}, across all compute capacitors, C_c, is constrained. This can be expressed as

$$V_{\text{out}}^{+} \quad \text{s.t.} \quad \sum_{i=1}^{N} C_c\left[|x_i + y_i| - V_{\text{out}}^{+}\right]_{+} = Q_{\text{total}}$$
$$V_{\text{out}}^{-} \quad \text{s.t.} \quad \sum_{i=1}^{N} C_c\left[|x_i - y_i| - V_{\text{out}}^{-}\right]_{+} = Q_{\text{total}} \qquad (5)$$

where the hyperparameter γ in (3) is now equal to Q_{total}/C_c. In this reverse water-filling analogy, the height of each vertical bar represents the value of the input argument, and the width of the bars represents the compute capacitor value. V_{out}^{+} and V_{out}^{-} are the output voltages for the four-quadrant inputs such that the total amount of charge (or the sum over the margins), represented by the colored-shaded area in Fig. 7, is Q_{total} in (5).

It can be noted that, for larger Q_{total}, $V_{\text{out}}^{\pm}$ is pushed lower. Since correlation can be viewed as a computation resembling pattern matching, the redundancy within the ensemble provides sufficiency for estimating the correlation using a subset of the input sequence. This can be noted in Fig. 7, where some bars are below the output levels $V_{\text{out}}^{\pm}$ indicating that no charge is stored in such caps and such operands are not part of the margins constraining and hence not part of the output computing. This is enabled through the use of ReLU providing high-efficiency operation.

From a hardware point of view, we can think of $V_{\text{out}}^{\pm}$ as the solution that ensures that the total charge is Q_{total} for a given set of inputs and a specific compute capacitor value, and we need a circuit to find $V_{\text{out}}^{\pm}$ by implementing the reverse water-filling constraint. Notably, Q_{total} has an optimal range for given circuit parameters such as capacitor value, the voltage range of the input operand, and the correlation length. For example, a large Q_{total} will reduce the energy efficiency of the compute scheme, as $2Q_{\text{total}}$ is used in every computation. Alternatively, for small values of Q_{total}, $V_{\text{out}}^{\pm}$, based on the reverse water-filling shown in Fig. 7, will be dominated by the maximum margin and will not represent the correlation between the inputs. In the ReLU-based MC scheme, the correlation is calculated by operating on four sets of operands, as shown in Fig. 6 and (5). The differential output voltage $(V_{\text{out}}^{+} - V_{\text{out}}^{-}) = G(R_{\text{MC}})$ estimates the correlation between the input sequence pair (Fig. 7).

To have a better understanding of the reverse water-filling operation, let us consider the case where inputs are highly correlated. In this case, $|X + Y|$ will be much larger than $|X - Y|$, and hence the heights of the bars to the left in Fig. 7 will be larger than the height of the bars to the right. Now, for the same amount of the constrained charge Q_{total}, V_{out}^{+} will be much larger than V_{out}^{-} indicating a high positive correlation. As mentioned earlier, correlation as pattern-matching computation has sufficient redundancy across samples to tolerate approximation errors. The proposed MC-correlation errors are compared with the MAC-approximation errors for random input sequences in Fig. 8. The correlation predicted by margin computation follows a similar error distribution to MAC schemes (i.e., $\varepsilon^2_{\text{MC-rms}} = |R_{\text{MC}} - R_\infty|^2 \approx \varepsilon^2_{\text{MAC-rms}} = |R_{\text{MAC}} - R_\infty|^2$ with both R_{MC} and R_{MAC} converging to R_∞ as sequence length, $N \to \infty$). Interestingly, for short sequence lengths, the higher-order terms in the MC approximation result in lower correlation errors when compared to the MAC correlation which only has the inner-product term.

Prior current-domain ReLU-based compute circuits require bias currents that lead to high power for long correlation lengths [28]. In this work, we present a charge-domain energy-efficient implementation shown conceptually in Fig. 9(a). The thresholding operation, described by (5), can be realized using charge-coupled diode-capacitor circuits, and a current source controls the hyperparameter $\gamma = Q_{\text{total}}/C_c$. For four-quadrant operands $(\pm x_i \pm y_i)$ applied to the diodes, the current source at the output node enforces a total amount of charge Q_{total} on the capacitors implementing the reverse water-filling constraint that solves for V_{out}^{+} and V_{out}^{-}. Notably, the operation

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

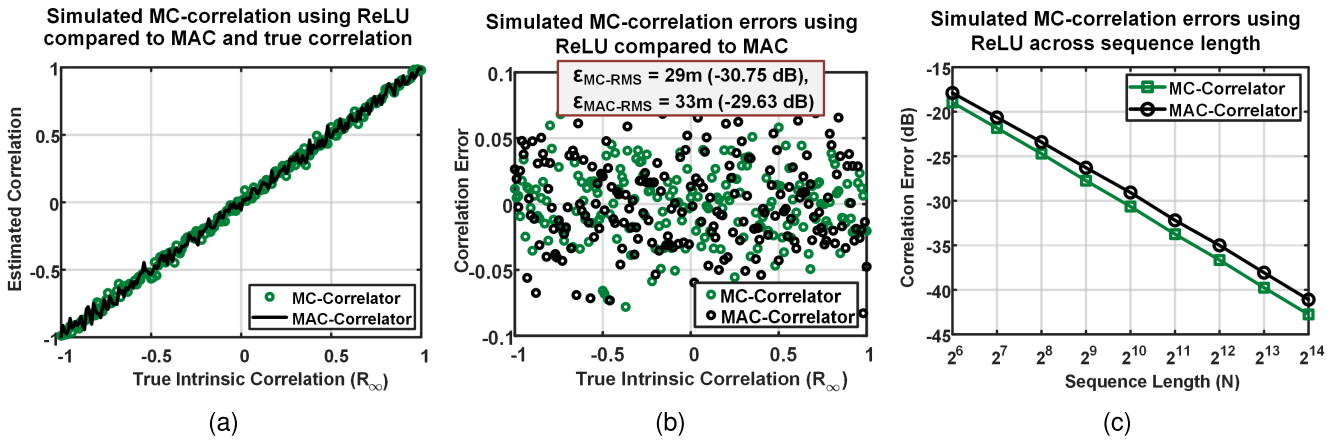

Fig. 10. Simulations of the proposed charge-thresholding correlator. (a) Similar to MAC, MC-correlation follows R_∞. (b) Error in correlation estimation using the proposed MC approach is comparable to MAC. (c) Precision of the MC-correlation improves with the increase of the correlation length.

of the reverse water-filling in Fig. 7 and its charge-domain implementation in Fig. 9(a) as well as the describing equation in (5) show that the output pair $V_{\text{out}}^{\pm}$ will be unique for a given input sequence and a given hyperparameter Q_{total}/C_c. This is confirmed through measurements as will be shown in Section VI.

The operation of the proposed MC-compute scheme has two phases: reset and compute. At reset, all compute capacitors' terminals are connected to the supply ensuring $V_{\text{out}}^{+} = V_{\text{out}}^{-} = \text{VDD}$ and storing zero total charge on the compute capacitors. During the compute phase, the reset switches are released, and the tail current sources impose the reverse water-filling constraint by storing a total charge of $Q_{\text{total}} = I_o t$ on the compute capacitors. Diodes with inputs greater than $V_{\text{out}}^{\pm}$ turn on and its outputs become $V_i^{\pm} = |x_i \pm y_i|$. Compute capacitors connected to these diodes contribute to the Q_{total} constraining storing a charge of $C_c[|x_i \pm y_i| - V_{\text{out}}^{\pm}]$. In contrast, diodes with inputs smaller than $V_{\text{out}}^{\pm}$ remain off preventing a charging path to their compute capacitors. Since compute capacitors are initially discharged to zero charge, the voltages across them remain zero with the diodes outputs $V_i^{\pm} = V_{\text{out}}^{\pm}$. By the end of the compute phase, the diode outputs are $V_i^{\pm} = \max(|x_i \pm y_i|, V_{\text{out}}^{\pm})$ and the thresholding condition, expressed in (5), is satisfied. Thus, the differential output voltage $V_{\text{out},d} = V_{\text{out}}^{+} - V_{\text{out}}^{-}$ estimates the correlation. In this charge-domain compute scheme, a total compute energy of $2Q_{\text{total}}\text{VDD}$ is drawn per N-length correlation resulting in a high energy efficiency of ~150 TOPS/W as we discuss in detail in Section VI.

Fig. 9(b) shows a simulation example of the output voltages for inputs with $R_\infty = 1$. The waveforms of the outputs follow the two-phase operation described earlier with $V_{\text{out},d} = 0$ in the reset phase, since single-ended outputs are connected to the supply, and $V_{\text{out},d}$ represents the correlation between the inputs by the end of the compute phase. The simulated performance of the proposed charge-domain correlation scheme is shown in Fig. 10. For random inputs, the simulated output voltage estimates correlation with error performance similar to MAC. In addition, the precision of the MC-correlation improves with the increase of the correlation length N, that is, the correlation converges to R_∞ with increasing N.

IV. CMOS Implementation of Proposed Correlator

In this section, we discuss the implementation of the correlator architecture proposed in Section III. Fig. 11 shows the implemented low-power 5 GS/s, 1024-sample analog correlator using the charge-thresholding core shown in Fig. 9(a). The RF correlator can be divided into three core functional blocks: 1) a sampler that sequentially stores input samples on 1024 capacitors; 2) an operand generator that generates the four-quadrant inputs required by the MC-compute engine; and 3) the correlation compute engine that estimates the correlation between the input signals. The operational timing is shown in Fig. 12(a), where input sampling and compute engine reset occur in the first phase. Following this, the operands required by the correlation compute engine are generated in a pre-compute phase. Finally, the correlation is estimated in the compute phase. At 5 GS/s, the 1024 samples require ~200 ns, the operand generation settles in <2 ns, and the correlator output is sampled after a compute time of ~100 ns. In addition to the core functional blocks, a multiphase non-overlapping clock generator drives the sampler, and a high input impedance amplifier with 1-GHz bandwidth is used to drive the analog output off-chip. Detailed implementation of the building blocks is discussed below.

A. Two-Layer 5-GS/s Sampler Operating on RF Input

A two-layer sampling scheme is utilized to support 5-GS/s operation while reducing parasitic capacitance at the RF input, as shown in Fig. 11. This approach also simplifies the clock generation for the sampling operation. Fig. 12(b) shows the timing diagram of the clocks, CLK_P and CLK_Q, driving the switches in the two layers. Layer 1 unit blocks consist of 32 capacitors and switches driven using 32 non-overlapping phase clocks, CLK_Q, with a pulsewidth of $32 \times T_s$ and a period of $32 \times 32 \times T_s$, where T_s is the sampling period. There are 32 Layer 1 blocks, resulting in 1024 capacitors for sample storage. Each of the Layer 1 unit blocks comes in series with a Layer 2 switch. The Layer 2 switches are controlled by CLK_P, which are a set of 32-phase non-overlapping clocks, with a period of $32 \times T_s$ and a pulsewidth of T_s. The proposed Layer 1 and Layer 2 ordering approach ensures that the faster

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

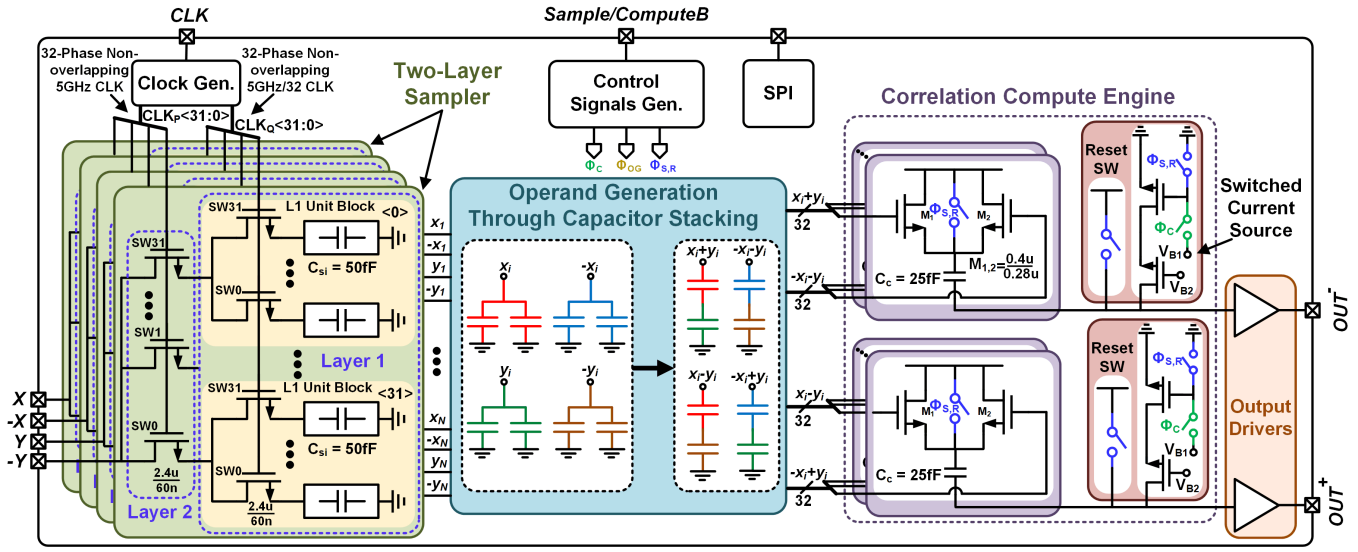

Fig. 11. Schematic of MC-based correlator implemented in 65-nm CMOS. The IC includes a 5-GS/s sampler that samples the two input sequences, switched-capacitor circuits that generate the $\pm x \pm y$ operands and an MC block that computes the correlation.

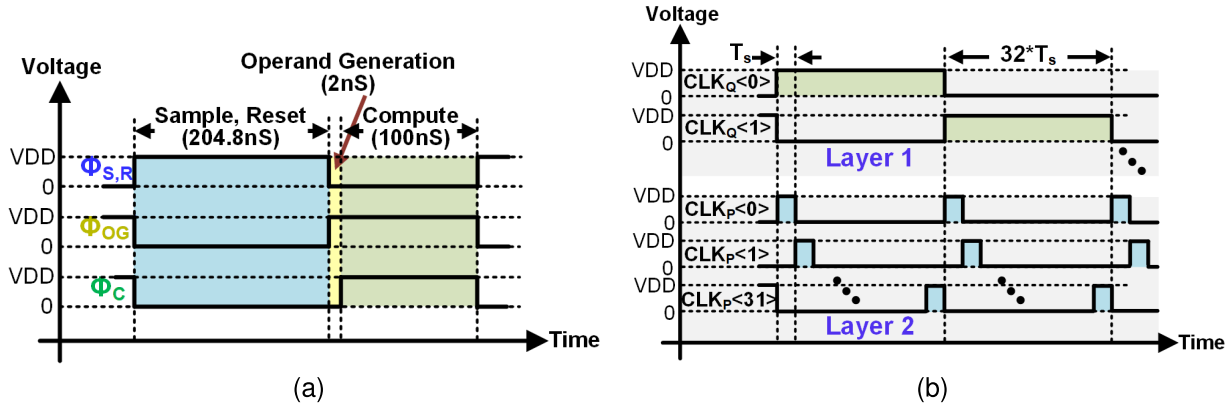

Fig. 12. (a) Operational timing of the MC-based correlator in Fig. 11. The correlator operates in three different phases: sample/reset, operands generation, and correlation compute. (b) Clocking operation of the two-layer sampler.

CLK_P multiphase clocks drive fewer switches than the slower CLK_Q multiphase clocks, reducing overall power consumption in the clock generation and sampler.

The sampling-switch design is constrained by the targeted correlator speed and length (5GS/s, 1024 samples), driver power consumption, and errors introduced by sampler non-idealities. NMOS devices with $(W/L) = (2.4\ \mu\text{m})/(0.06\ \mu\text{m})$ are used as the sampling switches to avoid extra parasitic capacitance associated with transmission-gate switches on high-frequency input nodes at the cost of limiting the maximum input swing.

Systematic sampling errors due to clock feedthrough do not impact correlation computation since the errors appear as dc shift for all the sampled values. On the other hand, capacitor leakage during the OFF-state and charge-injection errors during sampling are input-dependent and can, therefore, lead to errors in correlator output. The size of the sampling switch ensures that charge injection has a negligible impact on correlator accuracy in this implementation. Significant sampling capacitor leakage with an exponential decay in the OFF-state implies higher errors in samples that are stored earlier during the sampling when compared to later samples. Such sample-dependant gain error results in incorrect correlation computation. Notably, the operand generation scheme described in Section IV-B results in the effective storing capacitance being $4\times$ smaller than the actual sampling capacitor C_{si}, which increases the leakage in the compute phase. However, leakage during the compute phase impacts all operands equally and results in a fixed gain error. In this work, a relatively large sampling capacitor of 50 fF is adopted to minimize leakage errors. For a time scale of 1 μs, the employed C_{si} sees less than 1% error in the stored value due to leakage. With a conservatively sized sampling capacitor, the sampler puts the upper limit on the correlator speed.

B. Generation of Operands for Compute Phase

The operands required by the compute engine are generated through capacitor-stacking operation. As shown in Fig. 13(a), each input gets sampled on two capacitors. This allows the generation of four-quadrant operands $(\pm x \pm y)$ by stacking the capacitors appropriately. The conceptual switched-capacitor circuit shown in Fig. 13(b) is used to implement capacitor reconfiguration. Fig. 13(c) describes the operand generation across two phases. In the sampling phase, the sampling

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

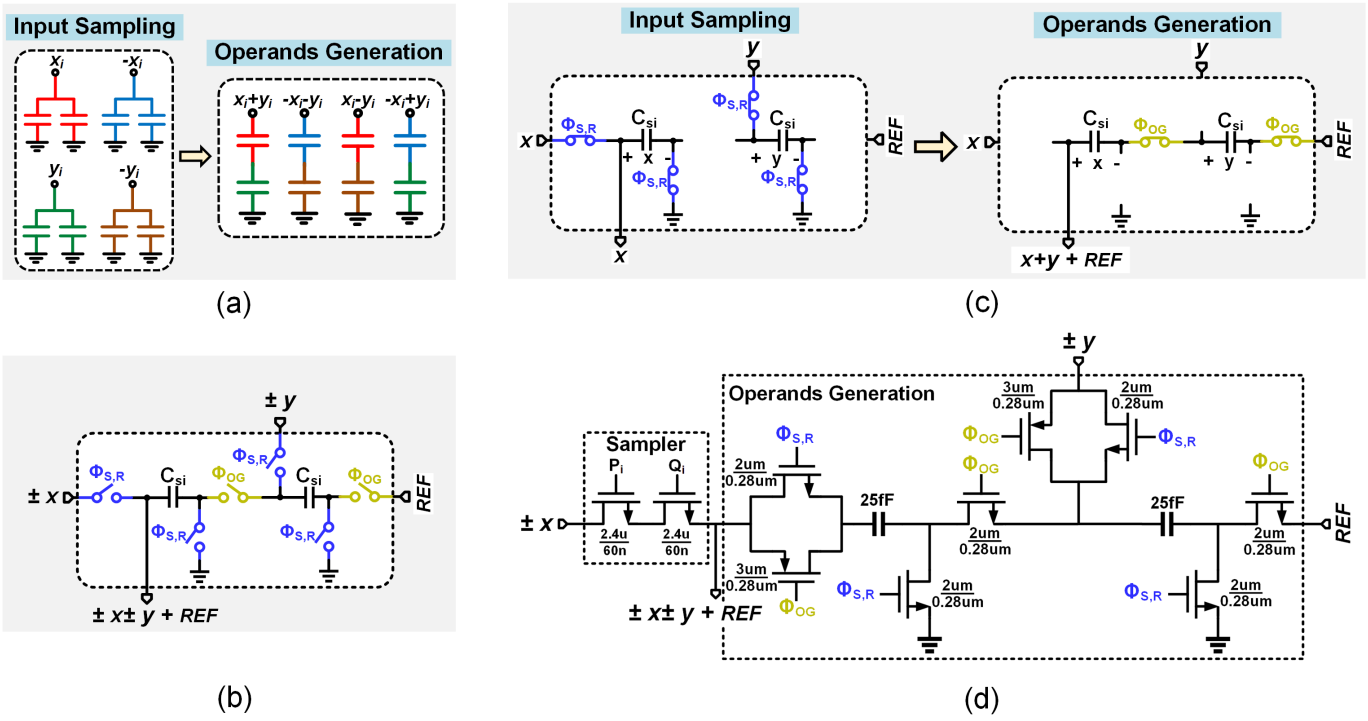

Fig. 13. (a) Conceptual diagram of operands generation operation. (b) Switched-capacitor circuit implementing capacitor reconfiguration required for the operands generation. (c) Two-phase operation of the operands generator. (d) Schematic of the operands generation circuit with the sampling switches.

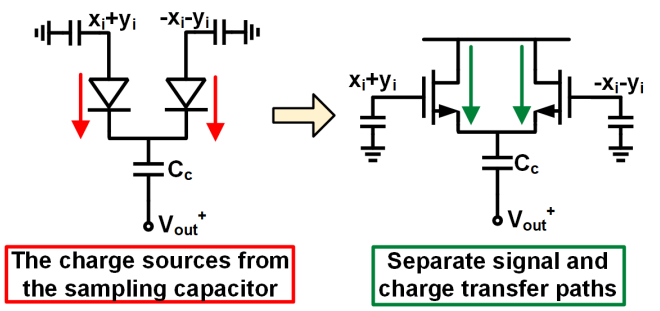

Fig. 14. Correlation compute engine replaces the diode-capacitor configuration with a common-drain transistor and a capacitor to separate the signal and compute charge transfer paths.

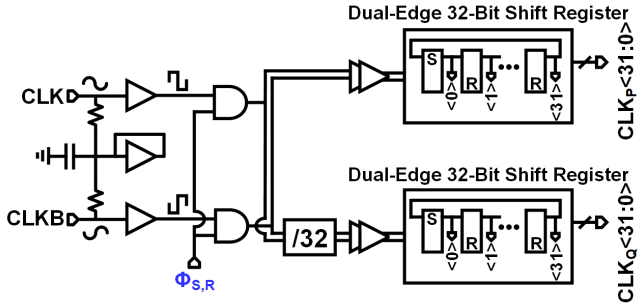

Fig. 15. Schematic of the multiphase clock generator.

capacitors, C_{si}, store the input samples. In the operand generation phase, the capacitors are stacked such that the stored values add together to generate the required operand. The detailed schematic of the operands generation circuit with the sampling switches is shown in 13(d). Transmission-gate switches with NMOS $(W/L) = (2\ \mu\text{m})/(0.28\ \mu\text{m})$ and PMOS $(W/L) = (3\ \mu\text{m})/(0.28\ \mu\text{m})$ are used in the operands generation circuit, with a time constant of 22 ps during sampling/operand generation. Notably, the operand generation is relatively insensitive to parasitic capacitance, as it results in a relatively small gain error that scales all inputs and does not impact the correlator accuracy. Additionally, the implemented operand generator supports a dc shift of all the generated operands through the "REF" input. This can be used to guarantee a positive voltage of the correlator single-ended outputs with enough headroom for the compute-engine current sources.

C. Margin-Computing-Based Correlation Engine

In the conceptual charge-domain compute scheme explained earlier, the compute charge Q_{total} flowing through the diodes is sourced from the sampling capacitors, thus changing the operand values during the compute operation, and resulting in computing errors. In the CMOS implementation, the diode-capacitor configuration is replaced with a common-drain transistor and capacitor to separate the signal and Q_{total} transfer paths (Fig. 14). The common-drain configuration puts a lower limit on the input signal swing to maintain the thresholding operation. The compute devices are sized with $(W/L) = (0.4\ \mu\text{m})/(0.28\ \mu\text{m})$. The compute engine uses a cascode current source to impose the ReLU constraint. A 50-μA current value is adopted balancing the compute speed and power trade-offs.

Compute capacitors, as well as sampling capacitors, are implemented using metal–oxide–metal (MOM) capacitors that are placed above active devices taking no additional area.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

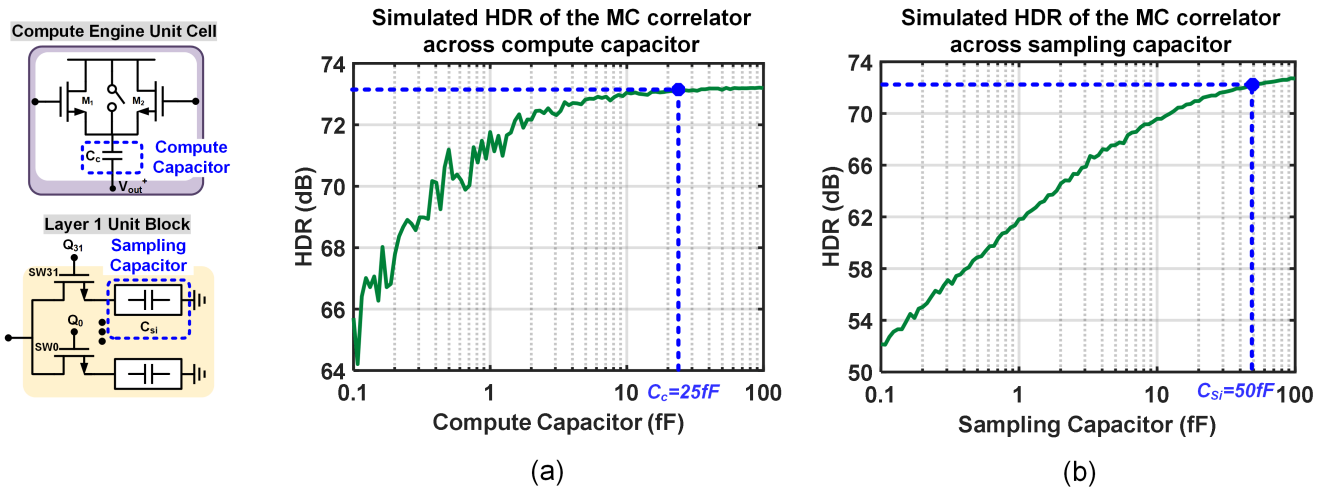

Fig. 16. Correlation error (a.k.a. HDR) simulations across (a) compute capacitor and (b) sampling capacitor.

Furthermore, the low voltage and temperature dependence of MOM capacitors and good matching enable accurate compute operation. The compute speed is determined by the charging rate of the compute capacitor C_c. However, the minimum capacitance size is limited by mismatch and process variations. While errors due to mismatch and noise are discussed in-depth in Section V, this implementation uses a conservative 25-fF compute capacitor to minimize the impact of unanticipated parasitic capacitances at the compute-engine outputs.

D. Multiphase Clock Generation for Sampling

The clock generator, shown in Fig. 15, uses an input buffer matched to 50 Ω. The clock generator uses a divide-by-32 divider based on a divide-by-two counter to generate CLK_Q. A dual-edge 32-bit shift register is used to generate the 32 non-overlapping phases for CLK_Q and CLK_P. The clock generation is disabled during the compute phase to reduce overall power consumption.

V. Impact of Noise and Mismatch on Proposed Correlator

This section details how the implementation non-idealities affects the correlation computation. Error in the correlation computation can be classified into: 1) errors due to the finite sequence length, $\varepsilon_{\text{len-rms}} \approx 1/\sqrt{N}$, with $\varepsilon_{\text{len-rms}} \to 0$ for larger sequence lengths; and 2) errors due to the hardware implementation, $\varepsilon_{\text{HW-rms}}$, that includes the margin-based computation approximation, noise, and mismatch. Therefore, the total error in correlation estimation can be expressed as, $(\varepsilon^2_{\text{len-rms}} + \varepsilon^2_{\text{HW-rms}})^{1/2}$. Notably, errors due to finite sequence lengths do not occur for periodic inputs where exact correlation can be computed across a period. Therefore, correlation errors with periodic inputs are dominated by hardware implementation errors. Accordingly, the hardware-dynamic range (HDR) can be expressed as

$$\text{HDR} = 20\log\frac{1}{\varepsilon_{\text{HW-rms}}}. \tag{6}$$

Correlator HDR can be determined using periodic sinusoidal inputs that are phase-shifted relative to each other. Since the true correlation R_∞ can be calculated in this case as $\cos(\theta)$, and assuming θ is the phase shift between the two inputs, $\varepsilon_{\text{HW-rms}}$ can be calculated as

$$\varepsilon_{\text{HW-rms}} = \sqrt{\overline{\left(G^{-1}\left(V_{\text{out},d}\right) - \cos(\theta)\right)^2}} \tag{7}$$

where $V_{\text{out},d}$ is the output of the margin-compute block and G^{-1} is the mapping function. G^{-1} depends on the adopted non-linearity (ReLU in the proposed MC-correlation) and maps the margin-compute output $V_{\text{out},d}$ to R_{MC} [33]. In this work, G^{-1} is approximated as a 5th-order polynomial function. G^{-1} is defined through an initial training phase where the correlation between input sequences with different degree of correlation (different true correlation R_∞) is estimated using the margin-compute correlator. The correlator output is then fit to the true correlation, used initially to create the correlated input sequences, using a polynomial fitting, and the 5th-order polynomial coefficients are obtained.

Through the following, the root-mean-square (rms) error $\varepsilon_{\text{HW-rms}}$ and hence the HDR will be used to characterize the impact of hardware non-idealities on the correlator accuracy. By applying sinusoidal inputs with phase shift $\theta \in (0, \pi)$, $\varepsilon_{\text{HW-rms}}$ and HDR can be found using (7) and (6), respectively.

A. Compute Capacitor Mismatch

As noted in Section IV-C, the compute capacitor is implemented using native MOM capacitors. MOM capacitor mismatch, which arises from lithography, etching process, and oxide film thickness variation, has been reported in the literature. Reference [34] reports the standard deviation of the unit capacitance relative variation, $\sigma(\frac{\Delta C}{C})$, of 2 fF MOM capacitor of different structures in 180-nm CMOS to be less than 0.13%. In [35], a 15-fF sandwich structure in 130-nm CMOS shows a standard variation of 0.09%. We used these variations to model the impact of compute capacitor mismatch. We estimated mismatch numbers for different capacitor values following Pelgrom's inverse-area mismatch

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

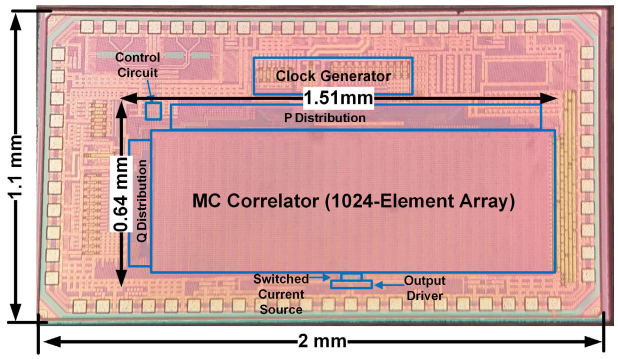

Fig. 17. Die photograph of charge-thresholding RF correlator implemented in 65-nm CMOS. The correlator occupies 0.97 mm^2.

model [36]. With the presence of a mismatch in the compute capacitor, (5) can be expressed as

$$\sum_{i=1}^{N}\left(C_c+\Delta C_{c,i}^{+}\right)\left[|x_i+y_i|-V_{\text{out}}^{+}\right]_{+}=Q_{\text{total}}$$
$$\sum_{i=1}^{N}\left(C_c+\Delta C_{c,i}^{-}\right)\left[|x_i-y_i|-V_{\text{out}}^{-}\right]_{+}=Q_{\text{total}} \tag{8}$$

where $\Delta C_{c,i}^{+}$ and $\Delta C_{c,i}^{-}$ are drawn from Gaussian distribution with zero-mean and a standard deviation that represents the capacitor mismatch. Fig. 16(a) shows the HDR plotted across different compute capacitor values. Mismatch of $\sigma = 0.2\%$ for 2 fF capacitor is assumed. This translates to a mismatch coefficient, the proportionality coefficients in Pelgrom's model, of 2.5 m. For 25-fF capacitance, the HDR is simulated to be ~73 dB. Notably, even with a 10× higher mismatch, that is, 25-m mismatch coefficients, HDR can be as high as 67.3 dB, supporting the use of smaller compute capacitance in future implementations.

B. Compute Device Mismatch

Compute device process mismatch can be modeled as an error term added to the operands. Accordingly, (5), will be expressed as

$$\sum_{i=1}^{N} C_c\left[|x_i+y_i|+\Delta_i^{+}-V_{\text{out}}^{+}\right]_{+}=Q_{\text{total}}$$
$$\sum_{i=1}^{N} C_c\left[|x_i-y_i|+\Delta_i^{-}-V_{\text{out}}^{-}\right]_{+}=Q_{\text{total}} \tag{9}$$

where Δ_i^{+} and Δ_i^{-} are drawn from a Gaussian distribution with zero-mean and a variance equal to the input-referred device V_T mismatch. Monte-Carlo simulations on the correlation compute engine show an HDR of 67 dB.

C. Sampling Noise

The thermal noise of the sampling circuit, determined by the KT/C_{si} limit, where K is Boltzmann's constant, T is the absolute temperature, and C_{si} is the sampling capacitor, can

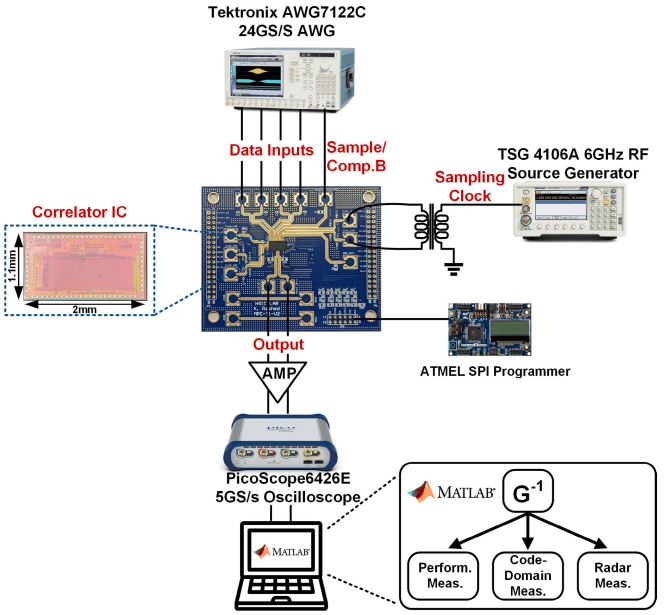

Fig. 18. Measurement setup for MC-correlator performance characterization and application-level demonstrations.

be modeled as part of the compute equations

$$\sum_{i=1}^{N} C_c\left[|x_i+y_i|+\sqrt{2}\Delta_i^{+}-V_{\text{out}}^{+}\right]_{+}=Q_{\text{total}}$$
$$\sum_{i=1}^{N} C_c\left[|x_i-y_i|+\sqrt{2}\Delta_i^{-}-V_{\text{out}}^{-}\right]_{+}=Q_{\text{total}} \tag{10}$$

where Δ_i^{+} and Δ_i^{-} are drawn from a Gaussian distribution of zero-mean and variance of KT/C_{si}. Fig. 16(b) plots the HDR across different sampling capacitor values. With a 50-fF sampling capacitor, an HDR of ~73 dB is expected.

D. Compute Device Noise

Compute device noise can be modeled; the same way its mismatch has been modeled, by adding the input-referred noise to the input operands (9). Transient noise simulations on the compute engine show an HDR of 70.4 dB.

E. Compute Current Source Noise

Noise in the compute-engine's current source will introduce an error in correlation computation by changing the compute charge Q_{total} from one computation to another. This can be modeled by an error term added to Q_{total} in (5) as

$$\sum_{i=1}^{N} C_c\left[|x_i+y_i|-V_{\text{out}}^{+}\right]_{+}=Q_{\text{total}}+\Delta^{+}$$
$$\sum_{i=1}^{N} C_c\left[|x_i-y_i|-V_{\text{out}}^{-}\right]_{+}=Q_{\text{total}}+\Delta^{-}. \tag{11}$$

For the employed 50 μA current, transient noise simulations on the compute engine show an HDR of 61.3 dB limiting the proposed correlator computation precision.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

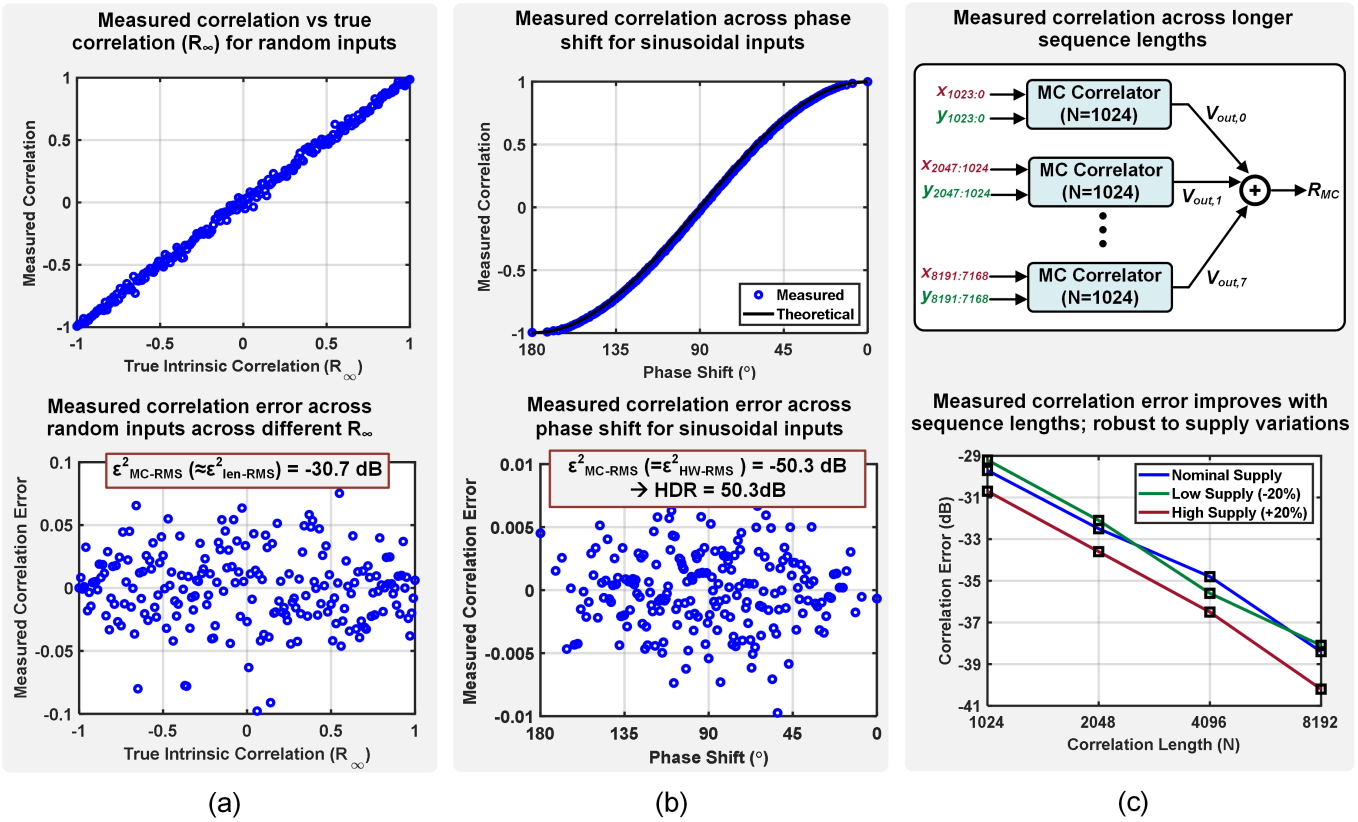

Fig. 19. (a) Measured correlation for random 5-GS/s inputs shows performance matching theory/simulations. (b) With sinusoidal inputs, measured MC-correlator demonstrates 8-bit ENOB HDR. (c) Proposed MC-correlator scales to higher correlation lengths and is robust to supply variations.

VI. Measurement Results

The proposed correlator is implemented in 65-nm CMOS process. Fig. 17 shows the die photograph, with the correlator occupying 0.97 mm^2. Each unit compute-cell of the correlator array includes the sampling and operand generation circuits with area minimized through the use of MOM capacitors. The die is packaged in a 60-lead 5 × 9 mm quad-flat no-lead (QFN) package that is soldered to the test PCB. Fig. 18 shows the measurement setup.

The correlator is characterized using periodic and random input sequences. To create random sequences with known correlations, two uncorrelated sequences A and B are mixed using the following method to create two correlated sequences X and Y with the required degree of correlation R_∞

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ R_\infty & \sqrt{1-R_\infty} \end{bmatrix} \times \begin{bmatrix} A \\ B \end{bmatrix}. \tag{12}$$

The X and Y sequences are digitally filtered to limit bandwidths and provided to an arbitrary-waveform generator (AWG) that drives the chip. The AWG also provides the control signal for sampling/compute phases. The correlator output is captured using an oscilloscope, and offline analysis is performed using MATLAB.

An RF source is used to generate periodic sine waves for HDR measurements, due to the limited resolution and the high noise of the AWG. The relative phase shift between the two inputs is created using a power divider and a phase shifter.

A. Compute Precision-HDR

Fig. 19 shows measured MC-correlator performance. As shown in Fig. 19(a), measurements using random 5-GS/s input sequences with known correlations show the measured correlation tracking R_∞. The correlation errors are dominated in this case by $\varepsilon_{\text{len-rms}}$ of −30.7 dB. Measured performance follows simulations in Fig. 10. The HDR is characterized using periodic 10-MHz sinewave inputs of 0.3 V amplitude sampled at 5.12 GS/s such that an integer number of input periods fits in 1024 samples. As shown in Fig. 19(b), measurements across relative phase shifts lead to a measured HDR of 50.3 dB (limited by the oscilloscope input-referred noise). The measured HDR can be translated to an 8.06 ENOB, demonstrating the proposed correlator's precision.

The scalability of the proposed approach is demonstrated by measuring longer inputs, shown in Fig. 19(c). Longer input sequences (up to 8 × 1024 = 8192) are partitioned into 1024-sample subsequences with correlator output for subsequences added together to compute sequence correlation. The measured accuracy with increasing sequence length follows the expected 3 dB lower error for a 2× increase in sequence length when errors are limited by $\varepsilon_{\text{len-rms}}$. These measurements demonstrate the feasibility of the MC-correlation approach for longer correlations through larger arrays tiling multiple ICs and/or through partitioning the sequences into subsequences.

As noted in Section II and Fig. 6, the MC-correlator differential output has a monotonic one-to-one mapping to the true

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

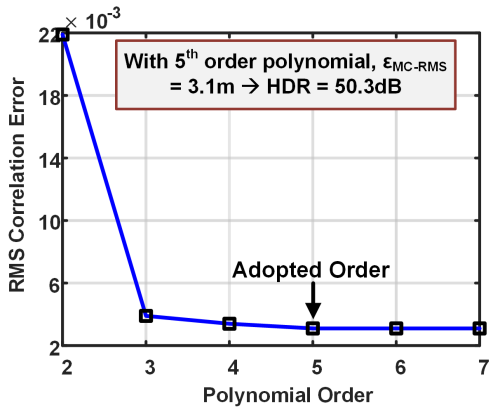

Fig. 20. Measured correlation error for sinusoidal inputs across mapping polynomial order—a 5th-order polynomial is used in all measurements.

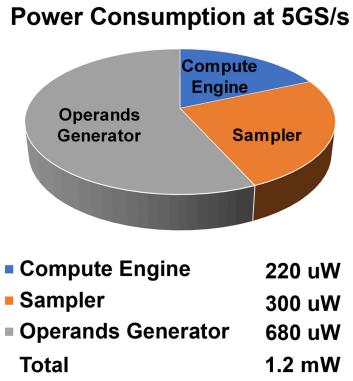

Fig. 21. Power consumption of core functional blocks in the MC-correlator at 5 GS/s.

correlation. This work uses a polynomial function to do this mapping. Fig. 20 shows the error in correlation computation for sinusoidal inputs versus the polynomial order. A 5th-order polynomial, of different coefficients with sinusoidal inputs and random inputs, is adopted in this work and is used in all measurements in Fig. 19.

B. Power Consumption and Energy Efficiency Calculation

Fig. 21 shows the power consumption breakdown across core functional blocks in the MC-correlator. The correlation compute engine consumes only 0.22 mW, operand generator consumes 0.68 mW, and the sampler drivers consume 0.3 mW. For a correlation length of 1024 and a sampling speed of 5 GS/s, the total correlation time (sampling time + compute time) is 409.6 ns, leading to a total correlation energy of 491.52 pJ. In a fully digital approach, computing the correlation between two sequences of length N and a B-bit precision requires a total of (N) B-bit multiplications and $(N-1)$ B-bit additions. Equivalently, the effective number of operations of the MC-correlation computation can be expressed as

$$\text{Effective Number of Operation} = (N)\text{ENOB}^2 + (N-1)\text{ENOB} \tag{13}$$

with a measured precision of 8.06 ENOBs, the compute efficiency of the correlator can be calculated in terms of TOPs/s/W or TOPS/W as

$$\text{Compute Efficiency} = \frac{\text{Effective Number of Operations}}{\text{Power Consumption} \times \text{Compute Time}}. \tag{14}$$

For a measured 1.2-mW power consumption and a total of 409.6-ns correlation time, the total energy efficiency is calculated to be ~152 TOPS/W. The multiphase clock generation consumes 4.2 mW, and clock input buffers that convert the off-chip sinewave to square wave consume 22.2 mW. It is worth mentioning that the clock generator in the presented IC is not optimized, since the main focus of this work is the correlator itself. In simulation, we can achieve significantly lower power consumption when the clock generation devices are resized for this application and frequency. In addition, in a complete system-on-chip (SoC), the clock generation can be combined with the rest of the transmit/receive chains adding a minimum overhead on the whole system. For complete characterization, the energy efficiency of the whole correlator IC is calculated, considering the clock generation power consumption to be 6.5 TOPS/W.

C. Performance Across Process, Supply, and Temperature Variations

The MC approach is relatively insensitive to supply variations in the compute engine. As shown in Fig. 19(c), measurements across ±20% supply variations demonstrate robustness to voltage variations. Similarly, Fig. 22(a) shows correlation measurements of 5-GS/s random inputs across temperatures. The correlation error increases <1 dB across a 60°C change in temperature, demonstrating good performance across temperature. In order to understand the impact of different sequences and ICs, the measurement in Fig. 19(a) is repeated for ten different data sets across three different chips, and the correlation error is reported in the box chart shown in Fig. 22(b) with less than 0.1 dB change in the mean value. Measured correlation of random inputs across sampling speed is shown in Fig. 22(c). The correlator IC shows almost no degradation in correlation error between 100 MS/s and 5 GS/s.

D. Applicability to Radar and Code-Domain Communications Systems

The feasibility of using the MC-correlator for typical correlation applications is shown through system-level measurements (Fig. 23). A noisy 1.25-GHz bandwidth radar pulse (SNR = 0 dB) is sampled at 5 GS/s and correlated against pulse templates. The correlation across time-shifted templates is shown in Fig. 23(a), demonstrating performance comparable to MAC approaches. A communication system application is shown in Fig. 23(b), where an input PN code at 2.5 GHz chip rate with SNR = 0 dB is correlated with delayed versions of the PN code. The correlation shows the expected impulse response for zero lag. In Fig. 23(c), a DS-CDMA signal with 2.5 GHz chip rate is correlated with the modulation PN code showing the matching between the demodulated signal and the original message. These measurements demonstrate that

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

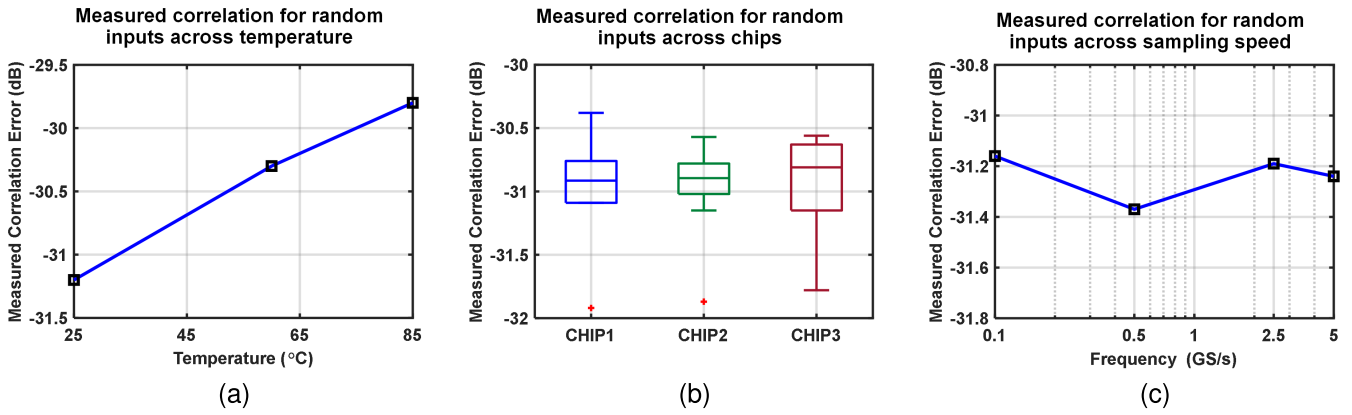

Fig. 22. (a) Measured correlation error for random 5-GS/s inputs across temperature. (b) Correlation error is measured for ten different random 5-GS/s sequences for three different chips. (c) Correlator IC shows almost no degradation in measured correlation for random inputs between 100 MS/s and 5 GS/s.

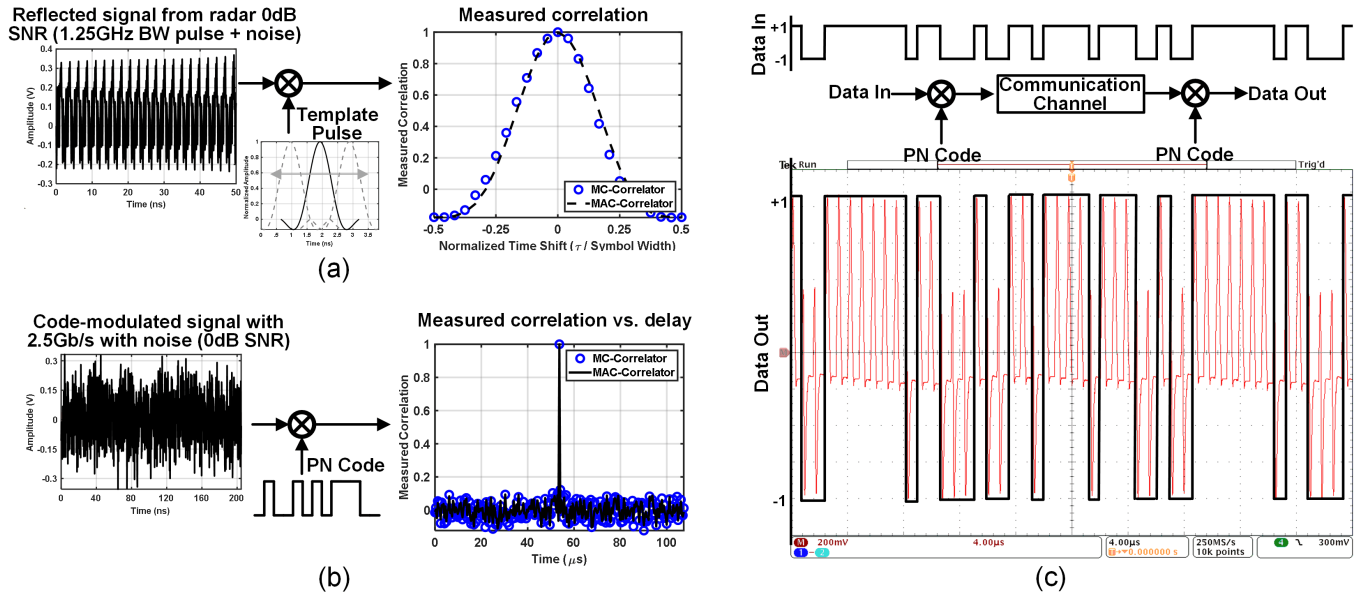

Fig. 23. Application demonstrations using MC-compute IC. (a) Radar measurement for input 1.25-GHz bandwidth pulse correlated with pulse template. (b) Measurement for 2.5-Gb/s PN code correlated with time-shifted versions. (c) Measurements for 2.5-Gb/s DS-CDMA signal correlated with the modulating PN code.

the MC-correlator can operate on samples directly in the analog domain, enabling low-latency and low-power signal processing.

E. Performance Summary and Comparison With Prior Work

Table I summarizes the measured performance and compares this approach with digital-intensive compute correlators and analog/RF correlators. When compared to state-of-the-art vector multipliers implemented in advanced technology nodes, the proposed compute scheme achieves higher 8-bit energy efficiency of 150 TOPS/W with input precision of 8.06 ENOBs. This work also demonstrates an improvement compared to state-of-the-art analog/RF correlators in terms of energy efficiency and area efficiency figures of merit. In addition, it achieves a large correlation length of 1024, which is $>100\times$ higher than prior work. Overall, the high energy efficiency of the proposed correlation scheme arises from the following:

1) Analog-domain operation. Being an all-analog correlator enables a high multibit operation (limited by noise in this work). Compared to digital implementation where power consumption scales with the number of bits, the all-analog implementation offers significantly better energy efficiency.
2) Charge-domain operation. The charge-domain implementation of the proposed correlation computation scheme where only $2Q_{\text{total}}\text{VDD}$ is drawn for 1024-length correlation computation contributes to its exceptional energy efficiency.
3) Multiplier-free operation. The proposed compute scheme avoids energy-intensive multiplications, instead utilizing low-power thresholding operations, making it more power-efficient.
4) Margin-compute approximation. By adopting a thresholding-based ReLU compute scheme, only a subset of the input sequence is used for the correlation estimation. Compared to MAC correlation where the whole input set is involved in the computation, the proposed approximate computing scheme enables high-efficiency operation.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PERFORMANCE SUMMARY OF THE MC-CORRELATOR AND COMPARISON TO STATE-OF-THE-ART VECTOR MULTIPLIERS AND ANALOG CORRELATORS

	Vector Multipliers					Analog Correlators			
	Y. He ISSCC23 [21]	D. Wang ISSCC22 [22]	L. Fick ISSCC22 [37]	H. Wang JSSC23 [38]	H. Jia JSSC22 [39]	A.R. Javed BCTM16 [18]	Q. Wu TMTT23 [19]	V. Mangal JSSC20 [17]	This Work
Tech.	28nm CMOS	28nm CMOS	40nm CMOS	22nm CMOS	16nm CMOS	130nm NPN	1um GaAS HBT	65nm CMOS	65nm CMOS
Computing Method	Digital ABS+ADD	Digital OR/AND+ADD	Current Multiply+ADD	Charge Multiply+ADD	Charge Multiply+ADD	Analog Multiply and Integrate	Pumped T-line	Analog Time Domain	Charge ReLU-MC
Area (mm^2)	1.3	1	189.8	1.5	25	0.15 (estimated)	78.61	0.09	0.97
Input & Weight Precision	2b-8b & 2b-8b	1b-4b & 1b	2b-8b & 2b-8b	8b & 8b	1b-8b & 1b-8b	Analog‡, 1 bit	Analog‡, LO	Analog‡, 1 bit	Analog, Analog 8.06 & 8.06
8b Throughput (TOPS)	0.117	0.127	21.3	1	3	-	-	-	0.18
8b Energy Efficiency (TOPS/W)	101.9	15.5	5.2	32.2	30	-	-	-	150* (6.5**)
8b Area Efficiency (TOPS/mm^2)	4.2	2.6	0.116	4	0.679	-	-	-	0.186
Input Sampling Rate	-	-	-	-	-	33 GS/s	0.6-1.1GS/s	100 S/s	5 GS/s
Code Length	-	-	-	-	-	1	8	11	1024
Correlator FoM (pJ/Corr/Len)	-	-	-	-	-	3.7	293	400	0.24
Compute Area/Corr. Len (mm^2)	-	-	-	-	-	0.15	39.304	0.0082	0.00095

‡ Dynamic range not reported. * Includes power consumption of correlator only. ** Includes power consumption of correlator and clock generation.

VII. CONCLUSION

In this work, we proposed a direct-RF wideband correlator. Utilizing a novel margin-based computation, the proposed correlator achieves high energy efficiency, traditionally provided only by digital-intensive compute schemes and high-precision operation. The margin-compute-based correlation is tested to be scalable by measuring longer sequence lengths. In addition, measurements across supply and temperature variations demonstrate the robustness of the MC-correlation scheme to environment non-idealities. The MC-correlator shows good performance across different ICs as well. The demonstrated system-level measurements show how the MC-correlator can be applied across signal-processing applications, with correlation occurring in the analog domain at RF.

Notably, the proposed correlation scheme can be scaled to advanced CMOS technologies with higher energy efficiency and area efficiency. Overall, the power and area-efficiency of the proposed correlator depend upon the area and power of sampler/clock generation, operand generation, and compute blocks. In particular, the sampler and clock-generation circuits benefit greatly from technology scaling due to lower switching dynamic power and lower ON-resistance for a given switch size. In the discussion in Section V-A, we showed that a smaller capacitor value can be employed given the reported mismatch in 65-nm CMOS. As reported in [40], the MOM capacitor mismatch in the 32-nm SOI process is 0.8% for a capacitor value of 1.2 fF, and simulations show this still results in an HDR higher than 65 dB for a compute capacitor value <5 fF. Smaller compute and operand generation devices will also reduce the physical layout size, leading to smaller routing parasitic capacitance at the output nodes. In summary, a smaller compute capacitor and lower parasitics enable the use of a smaller Q_{total} in the compute circuits, leading to improved energy efficiency as discussed in Section III. Similarly, smaller compute capacitors, compute devices, sampling switches, and clock circuits will lead to smaller correlator area in advanced technology nodes.

Future work includes further improvements in area and power efficiency through the development of new compute core circuits that enable analog computing by leveraging the underlying physics of on-chip devices in advanced CMOS technologies. Such approaches can also be extended to in-memory computing by relying on the physics of memory storage cells to perform computations.

ACKNOWLEDGMENT

The authors would like to thank Dr. James Wilson and his team from DARPA for their valuable feedback.

REFERENCES

[1] J. Li, M. Parlak, H. Mukai, M. Matsuo, and J. F. Buckwalter, "A reconfigurable 50-Mb/s-1 Gb/s pulse compression radar signal processor with offset calibration in 90-nm CMOS," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 1, pp. 266–278, Jan. 2015.

[2] I. Gresham et al., "Ultra-wideband radar sensors for short-range vehicular applications," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 9, pp. 2105–2122, Sep. 2004.

[3] B. Schleicher, I. Nasr, A. Trasser, and H. Schumacher, "IR-UWB radar demonstrator for ultra-fine movement detection and vital-sign monitoring," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 5, pp. 2076–2085, May 2013.

[4] T. Yamasaki, T. Nakayama, and T. Shibata, "A low-power and compact CDMA matched filter based on switched-current technology," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 926–932, Apr. 2005.

[5] E. Nishimori, C. Kimura, A. Nakagawa, and K. Tsubouchi, "CCD matched filter in spread spectrum communication," in *Proc. 9th IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, vol. 1, May 1998, pp. 396–400.

[6] T. Yamasaki and T. Shibata, "A low-power floating-gate-MOS-based CDMA matched filter featuring coupling capacitor disconnection," *IEEE J. Solid-State Circuits*, vol. 42, no. 2, pp. 422–430, Feb. 2007.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

[7] J.-S. Wu, M.-L. Liou, H.-P. Ma, and T.-D. Chiueh, "A 2.6-V, 44-MHz all-digital QPSK direct-sequence spread-spectrum transceiver IC [wireless LANs]," *IEEE J. Solid-State Circuits*, vol. 32, no. 10, pp. 1499–1510, May 1997.

[8] M.-L. Lieu and T.-D. Chiueh, "A low-power digital matched filter for direct-sequence spread-spectrum signal acquisition," *IEEE J. Solid-State Circuits*, vol. 36, no. 6, pp. 933–943, Jun. 2001.

[9] S. Skrzyniarz et al., "A 36.8 2b-TOPS/W self-calibrating GPS accelerator implemented using analog calculation in 65 nm LP CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Jan. 2016, pp. 420–422.

[10] J. Li, W. He, B. Zhang, L. Qi, G. He, and M. Seok, "CCSA: A 394 TOPS/W mixed-signal GPS accelerator with charge-based correlation computing for signal acquisition," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2023, pp. 430–432.

[11] K.-C. Liu, W.-C. Lin, and C.-K. Wang, "A pipelined digital differential matched filter FPGA implementation and VLSI design," in *Proc. Custom Integr. Circuits Conf.*, Aug. 1996, pp. 75–78.

[12] Y.-D. Chih et al., "An 89TOPS/W and 16.3TOPS/mm^2 all-digital SRAM-based full-precision compute-in memory macro in 22 nm for machine-learning edge applications," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 64, Feb. 2021, pp. 252–254.

[13] H. Fujiwara et al., "A 5-nm 254-TOPS/W 221-TOPS/mm^2 fully-digital computing-in-memory macro supporting wide-range dynamic-voltage-frequency scaling and simultaneous MAC and write operations," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 65, Feb. 2022, pp. 1–3.

[14] H. Hashemi and H. Krishnaswamy, "Challenges and opportunities in ultra-wideband antenna-array transceivers for imaging," in *Proc. IEEE Int. Conf. Ultra-Wideband*, vol. 42, Sep. 2009, pp. 586–591.

[15] D. Shen, F. Lin, and W. Gan Yeoh, "An analog correlator with dynamic bias control for pulse based UWB receiver in 0.18um CMOS technology," in *Proc. IEEE Radio Freq. Integr. Circuits (RFIC) Symp.*, Dec. 2006, pp. 277–280.

[16] H. Xie, X. Wang, A. Wang, B. Zhao, L. Yang, and Y. Zhou, "A broadband CMOS multiplier-based correlator for IR-UWB transceiver SoC," in *Proc. IEEE Radio Freq. Integr. Circuits (RFIC) Symp.*, vol. 4, Jun. 2007, pp. 493–496.

[17] V. Mangal and P. R. Kinget, "Clockless, continuous-time analog correlator using time-encoded signal processing demonstrating asynchronous CDMA for wake-up receivers," *IEEE J. Solid-State Circuits*, vol. 55, no. 8, pp. 2069–2081, Aug. 2020.

[18] A. R. Javed, J. C. Scheytt, and U. V. D. Ahe, "Linear ultra-broadband NPN-only analog correlator at 33 Gbps in 130 nm SiGe BiCMOS technology," in *Proc. IEEE Bipolar/BiCMOS Circuits Technol. Meeting (BCTM)*, Sep. 2016, pp. 78–81.

[19] Q. Wu, X. Zou, and Y. E. Wang, "Monolithically integrated directly pumped time-varying transmission lines," *IEEE Trans. Microw. Theory Techn.*, vol. 71, no. 4, pp. 1509–1517, Apr. 2023.

[20] H. Nakase, T. Kasai, Y. Nakamura, K. Masu, and K. Tsubouchi, "One chip demodulator using RF front-end SAW correlator for 2.4 GHz asynchronous spread spectrum modem," in *Proc. 5th IEEE Int. Symp. Pers., Indoor Mobile Radio Commun., Wireless Netw.-Catching Mobile Future.*, vol. 1, Aug. 1994, pp. 374–378.

[21] Y. He et al., "A 28 nm 38-to-102-TOPS/W 8b multiply-less approximate digital SRAM compute-in-memory macro for neural-network inference," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2023, pp. 130–132.

[22] D. Wang, C. T. Lin, G. K. Chen, P. Knag, R. K. Krishnamurthy, and M. Seok, "DIMC: 2219TOPS/W 2569F2/b digital in-memory computing macro in 28 nm based on approximate arithmetic hardware," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2022, pp. 266–268.

[23] Q. Liu et al., "33.2 A fully integrated analog ReRAM based 78.4 TOPS/W compute-in-memory chip with fully parallel MAC computing," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 500–502.

[24] C.-X. Xue et al., "A 22 nm 2Mb ReRAM compute-in-memory macro with 121–28TOPS/W for Multibit MAC computing for Tiny AI edge devices," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 244–246.

[25] R. Khaddam-Aljameh et al., "HERMES-core—A 1.59-TOPS/mm^2 PCM on 14-nm CMOS in-memory compute core using 300-ps/LSB linearized CCO-based ADCs," *IEEE J. Solid-State Circuits*, vol. 57, no. 4, pp. 1027–1038, Apr. 2022.

[26] P. Narayanan et al., "Fully on-chip MAC at 14 nm enabled by accurate row-wise programming of PCM-based weights and parallel vector-transport in duration-format," *IEEE Trans. Electron Devices*, vol. 68, no. 12, pp. 6629–6636, Dec. 2021.

[27] K. Rashed, A. Undavalli, S. Chakrabartty, A. Nagulu, and A. Natarajan, "A scalable and instantaneously wideband 5GS/s RF correlator based on charge thresholding achieving 8-bit ENOB and 152 TOPS/W compute efficiency," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2024, pp. 222–224.

[28] M. Gu and S. Chakrabartty, "Bias-scalable inner-product approximation circuit using analog margin propagation," in *Proc. IEEE 56th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2013, pp. 525–528.

[29] M. Gu and S. Chakrabartty, "Synthesis of bias-scalable CMOS analog computational circuits using margin propagation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 2, pp. 243–254, Feb. 2012.

[30] C. Kong and S. Chakrabartty, "Analog iterative LDPC decoder based on margin propagation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 54, no. 12, pp. 1140–1144, Dec. 2007.

[31] M. Gu and S. Chakrabartty, "A 100 pJ/bit, (32,8) CMOS analog low-density parity-check decoder based on margin propagation," *IEEE J. Solid-State Circuits*, vol. 46, no. 6, pp. 1433–1442, Jun. 2011.

[32] A. R. Nair, P. K. Nath, S. Chakrabartty, and C. S. Thakur, "Multiplierless MP-kernel machine for energy-efficient edge devices," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 30, no. 11, pp. 1601–1614, Nov. 2022.

[33] Z. Xiao and S. Chakrabartty, "A framework for analyzing cross-correlators using Price's theorem and piecewise-linear decomposition," 2023, *arXiv:2304.09242*.

[34] H. Omran, H. Alahmadi, and K. N. Salama, "Matching properties of femtofarad and sub-femtofarad MOM capacitors," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 6, pp. 763–772, Jun. 2016.

[35] A. Verma and B. Razavi, "Frequency-based measurement of mismatches between small capacitors," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2006, pp. 481–484.

[36] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, Oct. 1989.

[37] L. Fick, S. Skrzyniarz, M. Parikh, M. B. Henry, and D. Fick, "Analog matrix processor for edge AI real-time video analytics," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, Jun. 2022, pp. 260–262.

[38] H. Wang, R. Liu, R. Dorrance, D. Dasalukunte, D. Lake, and B. Carlton, "A charge domain SRAM compute-in-memory macro with C-2C ladder-based 8-bit MAC unit in 22-nm FinFET process for edge inference," *IEEE J. Solid-State Circuits*, vol. 58, no. 4, pp. 1037–1050, Apr. 2023.

[39] H. Jia et al., "Scalable and programmable neural network inference accelerator based on in-memory computing," *IEEE J. Solid-State Circuits*, vol. 57, no. 1, pp. 198–211, Jan. 2022.

[40] V. Tripathi and B. Murmann, "Mismatch characterization of small metal fringe capacitors," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 8, pp. 2236–2242, Aug. 2014.

Kareem Rashed (Graduate Student Member, IEEE) received the B.Sc. (Hons.) and M.Sc. degrees in electronics and communications engineering from Cairo University, Cairo, Egypt, in 2016 and 2020, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with Oregon State University, Corvallis, OR, USA.

From 2016 to 2021, he was a Design Engineer at Analog Devices (ADI), Cairo, working on a wide range of analog/RF/millimeter-wave circuits for wireless communications. His current research interests include exploring new analog/RF integrated circuits and systems for communications, sensing, and approximate computing.

Mr. Rashed was a recipient of the 2024 Analog Devices Outstanding Student Designer Award.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.

Aswin Undavalli (Graduate Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from NIT Delhi, New Delhi, India, in 2022. He is currently pursuing the Ph.D. degree in electrical and systems engineering with Washington University in St. Louis, St. Louis, MO, USA.

His research interests include RF integrated circuits, radar signal processing, and analog computing.

Shantanu Chakrabartty (Senior Member, IEEE) received the B.Tech. degree from Indian Institute of Technology, New Delhi, India, in 1996, and the M.S. and Ph.D. degrees in electrical engineering from Johns Hopkins University, Baltimore, MD, USA, in 2002 and 2004, respectively.

From 1996 to 1999, he worked at Qualcomm Inc., San Diego, CA, USA. From 2004 to 2015, he worked as an Associate Professor at the Department of Electrical and Computer Engineering, Michigan State University (MSU), East Lansing, MI, USA. He is currently working as a Clifford W. Murphy Professor and the Vice-Dean for Research with the McKelvey School of Engineering, Washington University in St. Louis, St. Louis, MO, USA.

Dr. Chakrabartty was a Catalyst Foundation Fellow, from 1999 to 2004. He is currently a fellow of American Institute of Medical and Biological Engineering (AIMBE). He was a recipient of the National Science Foundation's CAREER Award, the University Teacher-Scholar Award from MSU, and the 2012 Technology of the Year Award from MSU Technologies. He previously served as an Associate Editor for IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.

Aravind Nagulu (Member, IEEE) received the B.Tech. and M.Tech. degrees in electrical engineering from Indian Institute of Technology Madras, Chennai, India, in 2016, and the Ph.D. degree in electrical engineering from Columbia University, New York, NY, USA, in 2022.

In 2022, he joined the Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO, USA, where he is currently working as an Assistant Professor. His research interests include RF and millimeter-wave circuits, wireless systems, analog computing, cryogenic CMOS, and superconducting circuit design for large-scale quantum systems.

Dr. Nagulu has been a member of the Technical Program Committee of IEEE Custom Integrated Circuits Conference (CICC), since 2024. He was a recipient of the IEEE RFIC Symposium Best Student Paper Award (First Place), in 2018, the IEEE Solid-State Circuits Society Predoctoral Achievement Award, the ISSCC Analog Devices Outstanding Student Designer Award, the IEEE MTTS Graduate Fellowship in 2019, the IEEE Microwave Magazine Best Paper Award, Columbia University Electrical Engineering Collaborative Research Award in 2021, and the Eli Jury Award from Columbia University, in 2022.

Arun Natarajan (Senior Member, IEEE) received the B.Tech. degree in electrical engineering from IIT Madras, Chennai, India, in 2001, and the M.S. and Ph.D. degrees in electrical engineering from California Institute of Technology, Pasadena, CA, USA, in 2003 and 2007, respectively.

From 2007 to 2012, he was a Research Staff Member at the IBM T. J. Watson Research Center, Yorktown Heights, NY, USA. He worked on millimeter-wave (mm-wave) phased arrays for multiGb/s data links/airborne radar and self-healing circuits for increased yield in submicrometer process technologies. In 2012, he joined Oregon State University, Corvallis, OR, USA, where he is currently working as a Professor with the School of Electrical Engineering and Computer Science. He was part of the leadership team of Mixcomm Inc., Chatham, NJ, USA, a start-up focused on mm-wave beamformers for 5G and Satcom, which Sivers Semiconductors acquired, in 2022. His research interests include RF, mm-wave integrated circuits/systems for high-speed wireless communication/imaging, and low-power circuits for pervasive sensing and communication.

Dr. Natarajan was a recipient of the National Talent Search Scholarship from the Government of India, from 1995 to 2000, the Caltech Atwood Fellowship, in 2001, the IBM Research Fellowship, in 2005, the 2011 Pat Goldberg Memorial Award for the Best Paper in Computer Science, Electrical Engineering, and Mathematics published by IBM Research, the CDADIC Best Faculty Project Award in 2014 and 2016, the NSF CAREER Award, in 2016, the Oregon State University Engelbrecht Young Faculty Award, in 2016, and the DARPA Young Faculty Award, in 2017. He has served on the Technical Program Committee of the IEEE International Solid-State Circuits Conference (ISSCC), IEEE Radio frequency Integrated Circuits Conference (RFIC), and IEEE Compound Semiconductors IC Symposium (CSICS). He served as an Associate Editor for IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES, and IEEE JOURNAL OF SOLID-STATE CIRCUITS.

Authorized licensed use limited to: OREGON STATE UNIV. Downloaded on December 11,2024 at 07:01:28 UTC from IEEE Xplore. Restrictions apply.