Distance Computation Based on Coupled Spin-Torque Oscillators: Application to Image Processing

Minsuk Koo,[1,*] M.R. Pufall,[2] Yong Shim,[3] A.B. Kos,[2] Gyorgy Csaba,[4] Wolfgang Porod,[5] W.H. Rippard,[2] and Kaushik Roy[6,†]

[1] *Department of Computer Science and Engineering, Incheon National University, Incheon, South Korea*
[2] *Quantum Electromagnetics Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA*
[3] *School of Electrical and Electronics Engineering, Chung-Ang University, Seoul, South Korea*
[4] *Faculty of Technology and Bionics, Pazmany Peter Catholic University, Budapest, Hungary*
[5] *Center for Nano Science and Technology, University of Notre Dame, Notre Dame, Indiana 46556, USA*
[6] *Department of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA*

(Received 4 March 2020; revised 28 May 2020; accepted 22 June 2020; published 1 September 2020)

Recent research on nano-oscillators has shown the possibility of using a coupled-oscillator network as a core-computing primitive for non-Boolean computation. The spin-torque oscillator (STO) is an attractive candidate because it is CMOS compatible, highly integrable, scalable, and frequency and phase tunable. Based on these promising features, we propose an alternative coupled-oscillator-based architecture for hybrid spintronic and CMOS hardware that computes a multidimensional norm. The hybrid system, composed of an array of four injection-locked STOs and a CMOS detector, is experimentally demonstrated. The measured performance is then used as the input to simulations that demonstrate the hybrid system as both a distance metric and a convolution computational primitive for image-processing applications. Energy and scaling analysis shows that the STO-based coupled-oscillatory system has a higher efficiency than the CMOS-based system with an order of magnitude faster computation speed in distance computation for high-dimensional input vectors.

DOI: 10.1103/PhysRevApplied.14.034001

I. INTRODUCTION

Distance computation between multidimensional vectors is used in numerous applications, particularly for data and workload intensive problems such as combinatorial optimization, recognition, and classification. In order to process the massive amount of data effectively in real time, it is desirable to realize an energy-efficient hardware for the distance computation that utilizes parallelism. The computation of the Euclidean distance ($L2$ norm) requires expensive operations in hardware for squaring compared to that of the Manhattan distance ($L1$ norm) [1,2]. Since it is inefficient to implement such expensive operators in digital CMOS circuits, analog computation-based approaches that use device physics to directly compute complex functions such as squaring [3–7]. Such analog computations obtain better energy efficacy at the cost of tolerable errors, and are beneficial for applications where an approximate result is sufficient instead of an exact result. The paradigm of "let physics do the computing" has motivated researchers to look at "alternative computing models" that explore the use of non-CMOS devices as functional units for better energy efficiency and speed. One of those alternative models is based on the coupled-oscillator network in which the oscillator array is used to compute (say) "similarity" between two multidimensional vectors. The similarity can be defined in terms of the distance between the two vectors. [8–15].

Coupled networks of nonlinear oscillators are widely found in nature, with examples ranging from pendulum clocks on a wall [16], to flashing fireflies [17], animal flocking [18], to the coupled oscillations in the human heart and brain [19,20]. Inspired by such systems, researchers have proposed coupled-oscillator systems to solve image- or pattern-recognition problems [9,12,21–26] in a "preferred way of nature." Indeed, recent work has shown that a convolutional neural network can be implemented in an energy-efficient manner with a coupled array of CMOS ring oscillators [27,28]. Further research into unconventional oscillators with different characteristics (e.g., smaller footprints, higher efficiency and frequency, different nonlinearities and coupling strategies) could open additional network possibilities.

*koo@inu.ac.kr
†kaushik@purdue.edu

With the development of emerging oscillators such as spin-torque oscillators (STOs) and vanadium dioxide (VO_2), recent research has demonstrated fabricated oscillators [29–32]. However, they use external capacitors or bonding wires for coupling, making it difficult to couple large number of oscillators or to build high-density networks. Nanoscale oscillators such as STOs are attractive for implementing the large number of coupled-oscillator network for computation because they provide potential scalability of the functional units to smaller dimensions, along with faster computation time and less energy consumption compared to standard digital or analog CMOS implementations [3–5,7,14,29,33,34]. In this work, we focus our attention on a STO-based coupled-oscillatory system for approximate Euclidean distance (ED) computation. Such a hybrid nano-oscillatory system comes with multiple challenges that must be overcome to be practically implemented and adopted. First, the system has to be CMOS compatible (in terms of the fabrication processes, operating currents, and voltages) and scalable. Second, the system should perform computations in parallel for better energy efficiency, taking advantage of the properties of nano-oscillators. STOs are back-end process compatible with CMOS, and their oscillating amplitudes are also CMOS compatible. Moreover, STOs offer tunable frequency and phase, and generate microwave rf oscillating signals that enable fast computation. These features motivated us to explore STOs with the injection-locking scheme for the coupling as the primitive for a distance-computing architecture. We experimentally demonstrate a system composed of giant magnetoresistance- (GMR) based STO devices [35–37] and a CMOS detector as the core-computing primitive for distance computation. In addition, we theoretically show that phase dynamics of the system inherently introduces nonlinearity that makes the system appropriate for measuring the $L2$ distance between two multidimensional vectors. The hybrid system is used to measure the distance between two input vectors whose output follows the $L2^2$ norm. Our experimental results on four coupled systems along with CMOS peripherals for distance measurement is used to parameterize larger-scale simulations of coupled-oscillatory system. The hardware for $L2^2$ distance calculation is also used to compute approximate convolution, which in turn has been used for an edge-detection task.

II. COUPLED SPIN-TORQUE OSCILLATOR ARRAY

Here we describe the characteristics of STO devices and show how an array of STOs with these characteristics can be used as part of a system to compute the $L2$ distance between two multidimensional input vectors.

STOs are compact, current-controlled rf oscillators that utilize spin-transfer torques to drive magnetization precession. The STOs used here consist of a 70-nm electrical nanocontact made to a giant magnetoresistive) spin valve, which is a magnetic multilayer consisting of a fixed magnetic reference layer and a free magnetic layer separated by a nonmagnetic spacer layer. A dc current flowing through the nanocontact produces an antidamping torque on the free layer, causing the free layer to precess about the net effective field. Increasing the current changes the net effective field (primarily by changing the demagnetizing field via the cone of precession), thereby changing the precession frequency [38]. The resultant oscillating resistance via the GMR effect leads to an alternating voltage across the device, producing microwave signals in the GHz range. Furthermore, such oscillators are nonlinear, causing them to phase lock to impressed rf signals, rf fields, and to other STOs [35,39].

An array of STO devices is shown in Fig. 1(a). The GMR multilayer stack consists of a CoFe reference layer and CoFe/Ni multilayered free layer (see Appendix A for fabrication details). This stack, and the nanocontact geometry are chosen for the simplicity of their magnetic structure that minimizes device-to-device variations due to patterning. The left panel is a SEM showing the microstrip to deliver rf fields for injection locking, and the coplanar waveguide (CPW)-to-microstrip device lines to supply dc current. As shown in the cross-section schematic in the right panel of Fig. 1(a), the STOs are patterned 3 μm apart, which ensures that each STO receives an injection signal of similar amplitude and phase. This also ensures minimal interdevice interaction, which is not needed in this implementation. Also shown in the schematic is the dc magnetic field applied at approximately 5° to the surface normal. The combination of this external field and the internal (demagnetizing and anisotropy) fields of the free magnetic layer set the frequency of precession and the orientation of the precession cone (see Appendix A for sample details).

A typical individual device spectral response to an injected rf magnetic field is shown in the top panel of Fig. 1(b), which shows the device frequency pulling to the injected signal and phase locking. The degree of locking—that is, the fraction of the time the device is phase stable relative to the injected signal—increases as the STO frequency approaches the injection frequency. By measuring the amount of power outside of the injection signal and comparing it to the free running power, we can estimate the degree of locking of the device. For the devices in the measured array, we restrict our operation to regions where the estimated degree of locking is greater than 95%.

The injection signal is delivered to the STOs via a 1-μm-wide microstrip patterned over the devices [see Fig. 1(a)], producing an ac magnetic field at the devices of approximately 0.4 mT for the data presented here. Both the amplitude and the phase of an injection-locked oscillator

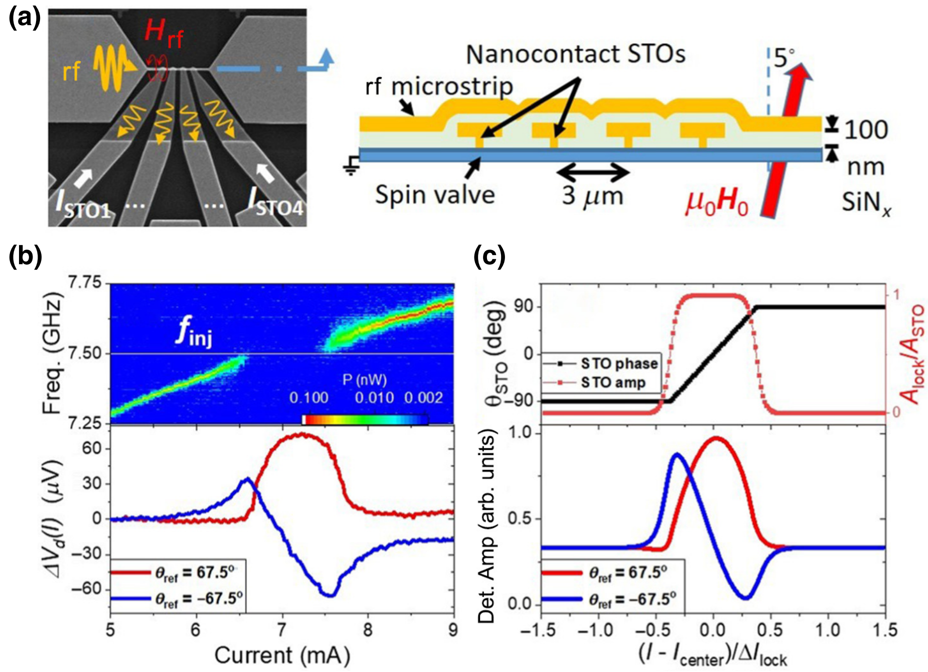

FIG. 1. (a) Left panel: scanning electron micrograph of device array, bias lines and signal paths, and rf microstrip. Direction of $H_{\rm rf}$ injection signal shown on SEM. Right panel: schematic cross section along rf microstrip showing spin-valve mesa, STO nanocontacts (approxiamtely 70-nm diameter), bias and rf lines, and applied field direction $\mu_0 H_0$. (b) Measurements of injection locking of an STO device. Top panel: spectral output from injection-locked STO device. Bottom panel: voltage from microwave diode across locking range for two different values of phase of the microwave reference signal, measured relative to the phase minimizing the sum of the parasitic and reference signal. (c) Phasor model of homodyne detection of sum of STO and reference signals. Top panel: assumed STO amplitude and phase response across locking range. Bottom panel: phasor model of expected signal at $f = f_{\rm inj}$, for different values of reference phase. For these curves an additional overall phase offset of $\pi/4$ is needed to mimic the measured response. The origin of such an offset is unknown, but could be due to a phase difference between the parasitic voltage and the injection field at the devices.

change across the locking range. For STOs, the phase has been shown to shift from $-\pi/2$ to $+\pi/2$ across the locking range [39]. In Ref. [39], this phase was detected via time-domain measurements of the STO output waveform. Here we show detection of relative phase via a homodyne scheme, in which a reference signal is added to the device signal, and the combined signal measured by a microwave diode detector. The diode detector is sensitive to microwave power, and results in a signal that is a function of the product of the STO and reference signal amplitudes, and their relative phase.

The bottom panel of Fig. 1(b) shows the diode detector voltage across the locking range of the STO for two different values of the reference phase. The phase angle of the reference is measured relative to that phase, which minimizes the sum of the parasitic and reference signals. The detected voltage has the same form as the power detected at $f = f_{\rm inj}$ via spectrum analyzer measurements (not shown), and is due to the large amplitude of the reference signal relative to the STO signal. The derivation of this relation is in Sec. III. Note that the shape of the diode detector output can change dramatically with the phase of the reference. A simple phasor model describing the STO amplitude and phase across the locking range is shown in Fig. 1(c). The bottom panel shows the expected response for those phase angles, based on a STO model whose amplitude and phase at $f_{\rm inj}$ vary as shown in the top panel. For the results presented here, we choose a reference phase and amplitude such that the detection curve has a maximum near the center of the locking range. These response curves are obtained using a zero-bias microwave Schottky diode. In Sec. IV, we show that a custom CMOS-based detector produces similar output curves.

Based on these observed features, STOs can be used to encode input information as the frequency in the free-running mode or as the phase in the injection-locking mode. The latter is the case in this paper and STOs are used as current-to-phase converters. The relative phase of STOs to the injected rf magnetic field represents input information mapped to the bias current of STOs. The introduction of an additional reference signal plays two additional roles in our degree of match system, first by enabling the detection of the STO phase locking without a spectrally resolved detection system, and second by improving the approximation of the $L2^2$ distance computation, as explained in the next section.

III. COUPLED OSCILLATOR-BASED DISTANCE-COMPUTATION SYSTEM

In this section, we first describe the functional configuration of the oscillator-based computing system, and explain how our system computes the distance between two input vectors by exploiting the device characteristics described above. We analyze the system with the derived equations to show the relationship between the input phase information and the corresponding output signal amplitude. Finally, the impact of noise from the device is considered for the operation of the system.

In our system, the coupled STOs are injection-locked in frequency by an ac magnetic field. Locked STOs emit the same frequency as the injection signal but can have different phases depending on their input currents. We choose the input currents to be proportional to the difference between the two input vectors, such that the coupled spin-torque oscillator network maps input information into phases of the oscillatory signals produced by the GMR STOs. As depicted in Fig. 2(a), each injection-locked oscillator is biased with a current corresponding to an elementwise difference between two vectors of an input. The output signals of the STOs are merged through a summing element [denoted as $\sum$ in Fig. 2(a)] before presenting to the detector unit. Different techniques such as resistive coupling [40] and capacitive coupling [12] have been used to sum incoming signals from the oscillators. In our implementation (described in the next section) we use Wilkinson coherent power combiners to avoid additional phase and amplitude offsets. The output signal from the summing unit exhibits different amplitude as a function of the relative phases to the reference signal. The detector unit measures this amplitude, and returns a digitized code that is proportional to the $L2^2$ between input vectors as described below.

We analyze how the oscillator network converts input information into the amplitude of the combined signal and how the nonlinear distance metric can be obtained. When injection locked, STOs emit a constant frequency as a function of current, but change phase monotonically across the locking range as explained in the previous

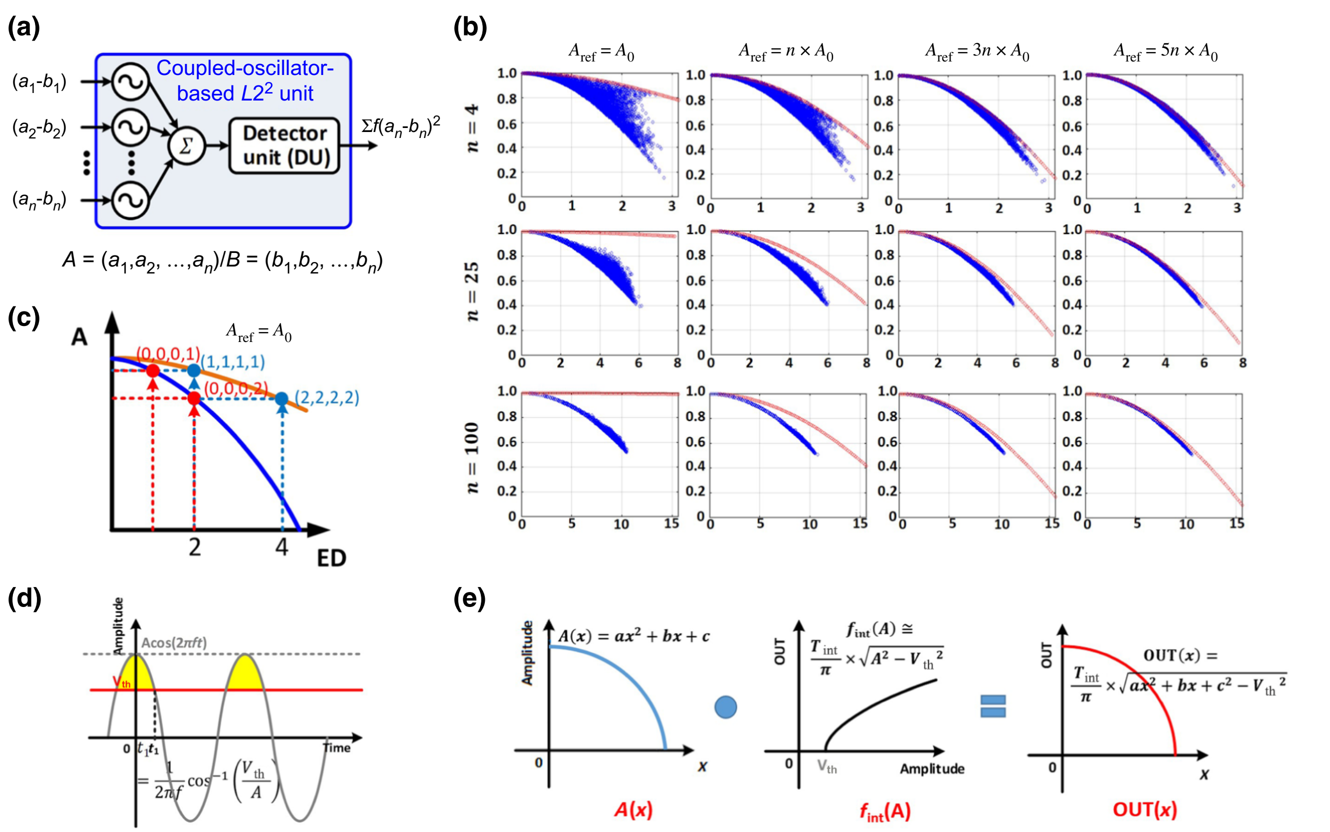

FIG. 2. (a) Block diagram of a coupled oscillator-based $L2^2$ unit proposed as a distance-computing primitive. (b) Model showing response for oscillators with the same A when the unit is programmed with different Euclidean distances when $A_{ref} = A_0$. (c) Normalized detector output amplitude, A [see Eq. (3)], versus $L2$ for $n = 4$, 25, and 100 with different $A_{ref} = A_0$, nA_0, $3nA_0$, and $5nA_0$. The red curve in each plot is the detector output in the limit that all STOs have the same phase. The x axis in all plots is for a theoretical $L2$ in arbitrary units. (d) Principle of detection via integration of a thresholded sinusoid. (e) Plots of and expressions describing the overall system operation.

section [39,41,42]. Here, the combined signal of the oscillator array can be expressed as $\sum_{i=1}^{n} A_i \cos(\omega t + \delta_{\rm ref} + \delta_i)$, where ω is the reference signal frequency, A_i is the amplitude of STO i at ω, n is the number of STOs (i.e., the dimension of the input vector), $\delta_{\rm ref}$ is the phase of the reference signal, and δ_i is the phase relative to that of the injection signal. Using the harmonic addition theorem [43], this signal can be rewritten as $A\cos(\omega t + \delta)$, where the amplitude and phase are defined as (see Appendix C for details):

$$A^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} A_i A_j \cos\left(\delta_i - \delta_j\right), \tag{1}$$

$$\delta = \tan^{-1}\frac{\sum_{i=1}^{n} A_i \sin\delta_i}{\sum_{i=1}^{n} A_i \cos\delta_i}. \tag{2}$$

In addition to the n oscillator signals, we have the additional reference signal that has the frequency of ω and provides the reference phase, $\delta_{\rm ref}$, to the oscillators. Therefore, we have the amplitude of the combined signal as

$$\begin{aligned} A^2 = {} & \sum_{i=1}^{n}\sum_{j=1}^{n} A_i A_j \cos(\delta_i - \delta_j) \\ & + 2A_{\rm ref}\sum_{j=1}^{n} A_j \cos(\delta_j) + A_{\rm ref}^2, \end{aligned} \tag{3}$$

where $A_{\rm ref}$ is the amplitude of the reference signal. We can expand Eq. (3) with the approximation of $\cos(\theta) \approx 1 - \theta^2/2$, where the approximation error is less than 5% when $|\theta|$ is less than $\pi/3$. The amplitude of the combined signal can be expressed, assuming all STOs have the same amplitude A_0, as (see Appendix C for details)

$$\begin{aligned} A^2 = {} & (A_0 n + A_{\rm ref})^2 - [(n-1)A_0^2 + A_{\rm ref}A_0]\sum_{i=1}^{n}\delta_i^2 \\ & + A_0^2\left[\left(\sum_{i=1}^{n}\delta_i\right)^2 - \sum_{i=1}^{n}\delta_i^2\right]. \end{aligned} \tag{4}$$

The first term of Eq. (4) is constant with respect to δ, and the second term is proportional to the $L2^2$ of an n-dimensional vector $[\delta_1, \delta_2, \ldots, \delta_n]$. The third term is referred to as an error term since we can obtain the approximated Euclidean distance if the error term is small enough. The maximum of the error can be derived by using Chebyshev's sum inequality, $(\sum_{i=1}^{n}\delta_i)^2 \le n(\sum_{i=1}^{n}\delta_i^2)$. The error term can be expressed as

$$A_0^2\left[\left(\sum_{i=1}^{n}\delta_i\right)^2 - \sum_{i=1}^{n}\delta_i^2\right] \le A_0^2(n-1)\left(\sum_{i=1}^{n}\delta_i^2\right), \tag{5}$$

where the equality holds if all δ_is are identical. Therefore, this error term produces a range of outputs depending on the distribution of the oscillator phases, and is maximized when the phases of the n oscillators are the same.

Here, the role of the reference signal becomes important to reduce the contribution of the error term to A. For example, let us assume that our system has four STOs and one reference signal, where the amplitude of the reference signal is the same as that of STOs ($A_0 = A_{\rm ref}$). Under this condition, the reference signal works as another single STO whose phase is $\delta_{\rm ref}$ or "0" relative phase. Figure 2(c) shows the relative phase information for four STOs, which can be represented as [0, 0, 0, 1] and [1, 1, 1, 1], where the $L2$ distance of the two cases are 1 and 2, respectively. The two cases can be considered as five STOs with phase information [0, 0, 0, 0, 1] and [0, 1, 1, 1, 1]. Here, the "1" is the phase difference in arbitrary units relative to the phase of the reference signal. From Eq. (3), we see that both of these cases result in the same amplitude of the combined signals, although their Euclidean distance is different. Furthermore, vectors with equivalent Euclidean distance ([0,0,0,2] and [1,1,1,1]) produce different outputs. Both of these results are a consequence of the error term represented by the third term in Eqs. (4) and (5). In Fig. 2(c), the value of A will be on the orange line, i.e., the upper bound, when the phase is evenly distributed among all four STOs, whereas the value of A will be on the blue line, i.e., the lower bound, when only one STO has a large phase value, despite the two cases having the same Euclidean distance. These two extreme cases give the maximum and minimum error represented by the third term in Eq. (4). A strong reference signal resolves these problems by increasing the contribution of the second term in Eq. (4), thereby reducing the relative contribution of the error term.

Figure 2(b) shows the normalized amplitude of the combined signal from n oscillators ($n = 4$, 25, and 100) compared to the expected $L2$ output. The simulation considers an additional reference signal of phase equal in the phase to the injection signal but with increasing amplitude. These simulations test how the system responds to the range of input vectors that can occur for a given $L2$, as a function of the number of oscillators and reference signal size. The output obtained when the n oscillators, when locked, have n random phases is shown by blue points, whereas the red points show the output when the n oscillators when locked all have equal phase. For a given $L2$, the amplitudes of all blue points are less than the amplitude of the red curve, showing the aforementioned claim that A has maximum error when the phases of n oscillators are the same. When the reference signal has the same amplitude as the STOs, the system shows a broad, noisy, quadratic dependence on $L2$. With the aid of a stronger reference signal, the output shows a stronger quadratic dependence on $L2$ by reducing the relative contribution of the error term as described above. We also observe that the quadratic dependence on $L2$ becomes less noisy as n increases for the same strength of the reference signal. This is because the variance of the error term for the case of n independent random variables δ_i linearly increases with increasing n, whereas that of the output amplitude (A) increases as $\sqrt{n}$. In other words, a clearer quadratic dependence on $L2$ can be obtained for a fixed strength of the reference signal as the number of dimension increases if the programmed input phases are randomly distributed in the locking range.

The STO devices have small output amplitudes (approximately -59 dBm, see Sec. IV). Thus, the amplitude variations in the combined signal from the oscillator array, occurring due to the phase differences between STO signals, are even smaller. In the previous section, a monolithic microwave Schottky diode is used to demonstrate this method of locking detection, whereas a CMOS detector integrated into other control circuitry is desirable to show system scalability. The key challenge for the CMOS detector circuit is to differentiate the small amplitude differences in the incoming signal. Rather than measuring the exact amplitude based on the complex analog circuitry [44,45], we propose a simple but adjustable integrator that is able to represent the relative amplitude difference of the inputs for a wide range of amplitudes. The integrator only tracks the region of our interest, which is the small amplitude change in the incoming sinusoidal signal around the peak value. Thresholding the signal around its peaks enables the integrator to easily detect the change in signal amplitude.

The integrator outputs a voltage proportional to the definite integral of a thresholded sinusoidal signal, which represents the enclosed yellow area between the threshold level $V_{\rm th}$ and the oscillatory signal as shown in Fig. 2(d). In a given integration time $T_{\rm int}$, the area is a function ($f_{\rm int}$) of the amplitude of the sinusoidal signal A and $V_{\rm th}$, written as

$$f_{\rm int}(A) \cong \frac{T_{\rm int}}{\pi} \times \sqrt{A^2 - V_{\rm th}{}^2}, \tag{6}$$

where $T_{\rm int}$ is long enough compared to the period of the sinusoidal signal $2\pi/\omega$. Thus, for a $\approxeq 7$ GHz signal from the STOs, an integration time of a few nanoseconds is sufficient. As shown in Fig. 2(e), the overall system can be expressed as Eq. (7) by replacing A in Eq. (6) with an arbitrary second-order polynomial, $A = ax^2 + bx + c$, where x is set to the $L2$ norm:

$$\mathrm{OUT}(x) = \frac{T_{\rm int}}{\pi} \times \sqrt{\left(ax^2 + bx + c\right)^2 - V_{\rm th}{}^2}. \tag{7}$$

This representation of the integrated value, $\mathrm{OUT}(x) = \mathrm{OUT}(L2)$ is a quadratic function of $L2$, thereby illustrating that the proposed system can perform approximate distance computation ($L2^2$ norm).

For real devices, the phase and amplitude of the oscillating signals from STOs deviate from the expected values due to process variations and noise. Based on the measured results in Refs. [39,46], the effect of phase (10°, 20°) and amplitude variations (10%, 20%) versus $L2$ are estimated in Figs. 3(a) and 3(b) for different amplitudes of the reference signal $A_{\rm ref} = nA_0$, and $3nA_0$, respectively. Note that the quadratic dependence on $L2$ is still maintained even with process variations and device noise.

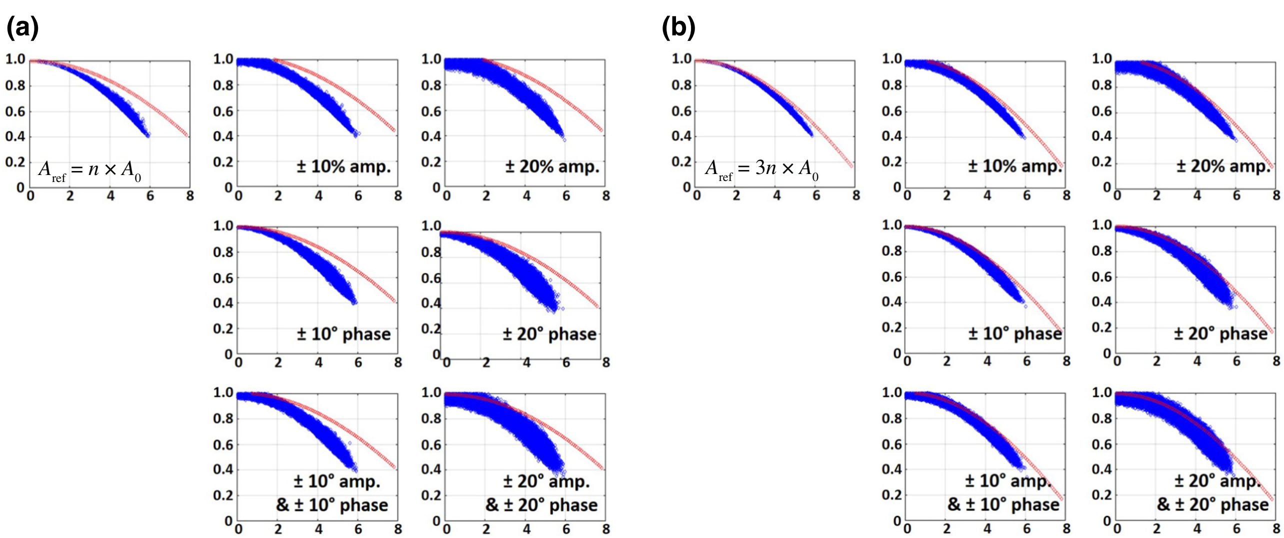

FIG. 3. Effect of phase (10°, 20°) and amplitude variations (10%, 20%) on the detector output versus $L2$ for the cases (a) $A_{\rm ref} = nA_0$, and (b) $A_{\rm ref} = 3nA_0$, with the number of oscillators $n = 25$, using the same simulation setup as in Fig. 2(b).

IV. IMPLEMENTATION AND RESULTS

In this section, the implementations of the coupled STO network and CMOS detector are discussed. Figures 4(a) and 4(b) show a block diagram of the entire hybrid system and its hardware implementation, respectively. See Appendix A for details on the hardware implementation. The entire system is divided into two subgroups: (1) GMR STO module with external components and (2) CMOS detector ($L2^2$ unit) along with bias current sources. A conventional microwave diode detector is also employed as a comparison to the CMOS detector. For the chosen computation architecture of Fig. 2(a), first the STOs must couple to the injection signal, which works as a phase reference. Subsequently, the individual oscillator output signals must be combined with equal phase and amplitude. The STOs are arranged in a bank of four oscillators excited by a common microstrip field line. Each STO is contacted by a coplanar-to-microstrip waveguide enabling independent current biasing and high bandwidth signal detection. The injected signal couples parasitically to the STO output lines (approximately 35-dB isolation), resulting in a signal with amplitude (-31 dBm) much larger than the STO (approximately -59 dBm). This parasitic is at a fixed phase relative to the locked STO signal, and thus coherently mixes with the STO signal at the power detector, providing a phase reference for phase-sensitive detection as described in the previous section. To further tune the phase and amplitude of the reference signal, an additional reference with adjustable phase and amplitude is also added to the summed STO [see Fig. 4(b)]. The reference signal enables the detection of this locking curve directly without spectrally resolving the STO output. The resulting homodyne signal at the detector during locking is proportional to $A_{\mathrm{STO}} \times A_{\mathrm{ref}}$ [see Eq. (1)], and is much larger than the STO signal itself. For the measurements in this research, A_{ref} is approximately $660A_{\mathrm{STO}}$, and locking is easily detectable as a change in the detector output voltage.

Once the signal is presented to the CMOS module, additional amplification stages are used inside the CMOS detector [Fig. 4(c)]. The first two stages are a low-noise amplifier (LNA) and a differential amplifier to make the signal compatible with CMOS circuits. Inverter-based amplifiers are used to amplify only the portion of interest—the peak of the incoming signal—based on the threshold of the inverter. Consequently, the signal at the input to the integrator behaves like the fourth waveform shown in Fig. 4(c) (INV2 amp.) where the signal normally stays at the supply voltage (VDD) level, and only goes below when the INV1 amp signal exceeds the threshold level of the second inverter (V_{th2}). The amplitude of the incoming signal V_{amp} determines the depth of the dip in voltage

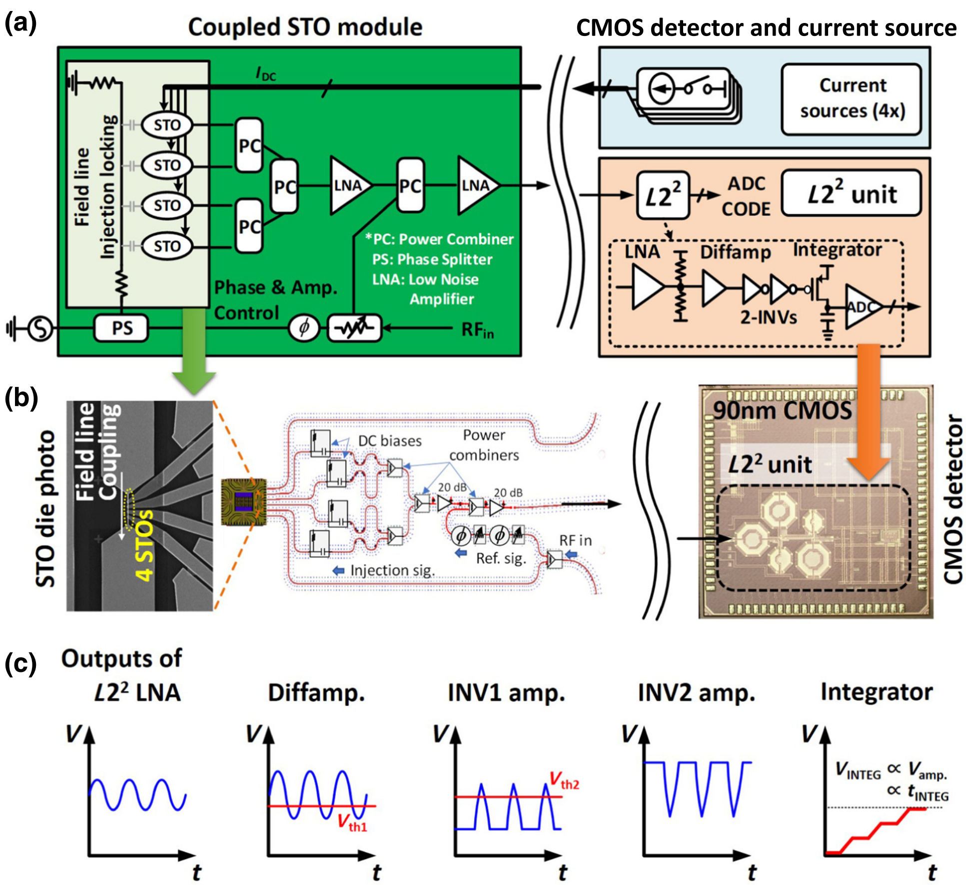

FIG. 4. (a) Block diagram for entire system, and (b) its implementation. (c) The expected signals at different nodes in the CMOS detector module.

from the VDD level, and thus the amount of current is stored onto the capacitor. Accordingly, the voltage across the capacitor rises during the integration time. Finally, the integrator output voltage is converted into a digital code at the analog-to-digital converter (ADC) stage for further image processing. The CMOS detector is fabricated in the 90-nm CMOS process.

We first measure the spectral response of each individual STO device in an array of four STOs to characterize the coupled-oscillator network. These responses are shown in Fig. 5(a) with an applied field $\mu_0 H_0 = 0.379$ T to place the oscillation frequencies near 7 GHz. The variations in the frequency and amplitude versus current (i.e., the spectral response) from device-to-device are one of the most challenging aspects of creating arrays of oscillators with STOs. The devices in the array have sufficient overlap near 7 GHz to allow injection locking of all devices simultaneously. The locking response for each device in the array is shown in Fig. 5(b). The difference in shapes of the locking curves for a given reference signal amplitude and phase occurs due to a combination of extrinsic device differences (injection-locking phase at a given STO, effective microwave path lengths) and intrinsic differences (STO locking dynamics). From the locking response, a reference current I_{0j} for each device is chosen, and a random set of test points mapped as current deviations from I_{0j} are applied. The resulting distribution is shown in Fig. 5(c) (see Appendix D for details of this analysis). As the test current moves away from the reference point, the detected signal shown in Fig. 5(c) decreases monotonically, albeit with significant noise. The noise on the detected voltage at a given test current is a combination of electrical noise (primarily due to the large gain needed to detect the small GMR signals) and the asymmetry in the response curve

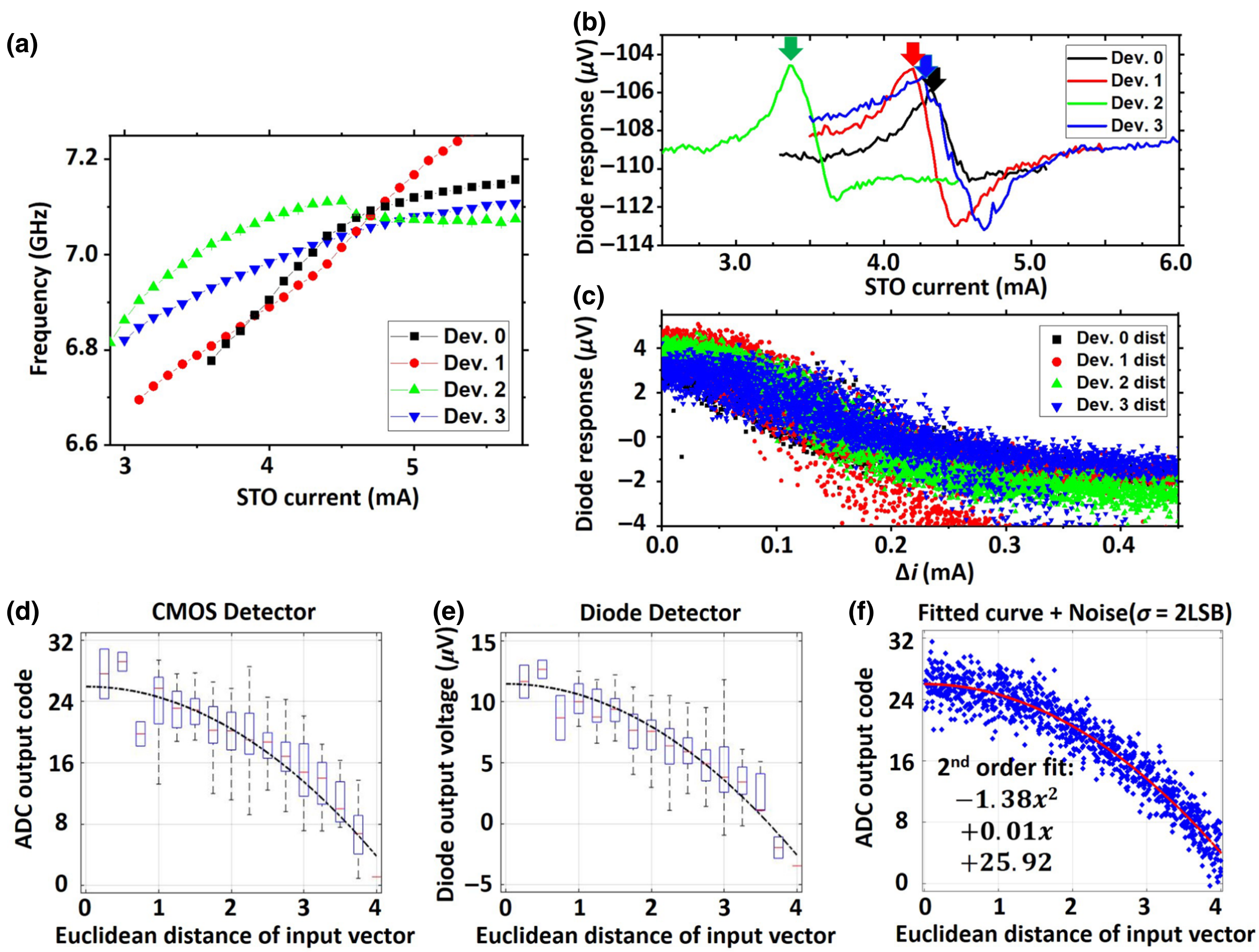

FIG. 5. (a) Fitted frequency responses of all devices versus current, showing an overlap at the 7-GHz operating point. (b) Diode output voltage curves for all devices thorugh locking range. (c) Distribution of diode response for each device for 5000 test ΔI points. Box plots of the response of the four-device array for test vectors within $\Delta I < 230\ \mu A$, (d) with CMOS detector and (e) with diode detector plotted versus Euclidean distance (defined as mA/10^{-4} mA). (f) The fitted curve from (e) with normal distributed noise [σ of the 2-least-significant-bit (LSB) ADC code].

for $I_j = I_{0j} \pm \Delta I$. Phase noise of the injection-locked STO can also add to noise. However, our measurements of the close-in power spectral density across the locking range do not show significant fluctuations out of the locking band.

Finally, the response curve of the four devices, which are simultaneously programmed based on Euclidean distance of their currents ($[I_1, \ldots, I_4]$) from the reference point I_{0j}, is shown in Figs. 5(d) and 5(e) using the CMOS detector and the diode detector, respectively. Approximately 400 randomly generated input vectors are applied to the STOs and the corresponding output voltage from the system is measured. On each box, the central mark indicates the median, the ends of the vertical boxes indicate the 25th and 75th percentiles, respectively, and the lines indicate the min and max values. The diode response remains monotonic with $|\Delta I|$ and it is similar to the curve calculated from the individual response curves in Fig. 5(c). This is typical for noninteracting devices with amplitudes much less than the reference signal. The response of the CMOS detector approximately follows that of the microwave diode detector, suggesting that the two detectors introduce similar nonlinearities into the signal detection channel.

The output voltage of the CMOS detector is sampled with a 5-bit resolution ADC implemented in the detector. Note that the output analog voltage can be postprocessed with ADC with different resolution. For the case of $A_{\text{ref}} = 5(nA_0)$ and $n = 4$ in Fig. 2(b), the standard deviations (σ) from the fitted curve are 0.32 and 1.98 LSB (5-bit ADC) for the case of no variation and 20° phase and 20% amplitude variation, respectively. The σ of the measured average CMOS detector response from the fitted curve is 2.67 LSB, which is larger than the σ obtained for 20° phase and 20% amplitude variation. Based on the power spectral density across the locking range, we believe that the measured noise comes mainly from the gain-amplification stages, which can be reduced or removed if the system is built on a single chip and the device signal amplitudes improve through increased magnetoresistance. Figure 5(f) shows the second-order fitting curve from Fig. 5(d) with normal distributed noise (σ for 2 LSB of the 5-bit ADC code). We utilize the fitted curve from the measured results with normal distributed noise for approximate distance computation targeting image-processing applications as described in the following section.

V. APPLICATIONS USING AN STO-BASED $L2^2$ NORM

To check the feasibility of using the proposed $L2^2$ unit for distance computation and convolution, we parameterize the simulations using the experimentally obtained response functions [Fig. 5(f)] to perform a facial recognition task and an edge-detection task using Gabor filtering [49]. For facial recognition, 40 images [47] are compared to each other in a pixelwise fashion, and $L2^2$ calculated as shown in Fig. 6(a). These 92×112 images are converted to 5-bit grayscale. As described in the previous section, each STO in a 2×2 array is biased with an amount of current corresponding the pixelwise difference between the reference image and the template image, and by sliding this 2×2 array cross the image, the approximate $L2^2$ [degree of match (DOM)] between two images is calculated. The distance is mapped onto a gray scale ranging from white (similar) to dark (dissimilar) and plotted as a 40×40 matrix comparing 40 images in the left panel of Fig. 6(b) for both the ideal and STO-based $L2^2$ calculation. The diagonal of the matrices are the match of the image to itself, and off-diagonal elements are the DOM of one face to another. It is possible to observe the similarity of response of the oscillator-based $L2^2$ to the ideal result, with the STO DOM simulation capturing these ideal and partial matches in a similar manner to the ideal Euclidean distance.

The right panel of Fig. 6(b) gives a different means of comparing the STO $L2^2$ to the ideal $L2^2$ by plotting the STO $L2^2$ versus the ideal $L2^2$. This STO $L2^2$ simulation includes 1σ Gaussian white random noise (GWRN) on the coefficients (a, b, c) in Eq. (7) [σ uncertainty derived from the fit to Fig. 5(d)]. This plot shows that the STO $L2^2$ and ideal $L2^2$ are strongly linearly correlated, showing that the STO $L2^2$ system is a good representation of the ideal $L2^2$. Furthermore, the plot also includes $L2^2$ distances between the original face images to those with added GWRN noise of different bit depths. By comparing the normalized breadth of uncertainty introduced by these noisy images to that introduced by the 1-σ uncertainty on the STO parameter values, we can get a measure of the impact of STO noise. In the inset to Fig. 6(b), we can see that the uncertainty added by the noise in the STO transfer curve (i.e., the vertical spread in the data) is roughly equivalent to the uncertainty of match added by the addition of approximately 3 bits of Gaussian white noise to the images in an ideal $L2^2$ calculation.

The energy consumption for distance computing has been estimated to project the efficiency of the proposed system. Specifically, we compare our system with CMOS-based analog distance calculation circuits (DCCs) as a separate block. Note that the power consumption of our hybrid system is dominated by the LNA [see Fig. 6(c)] needed to amplify the STO outputs to CMOS levels, a consequence of the GMR STOs used in the system. If higher magnetoresistance STOs are used, both amplifier and STO power consumption (through lower operating currents) could be reduced. However, even without considering such devices, our current hybrid system with GMR STOs becomes more power efficient as the number of dimensions increase (because the LNA is shared by all STOs in our method) and shows comparable or better

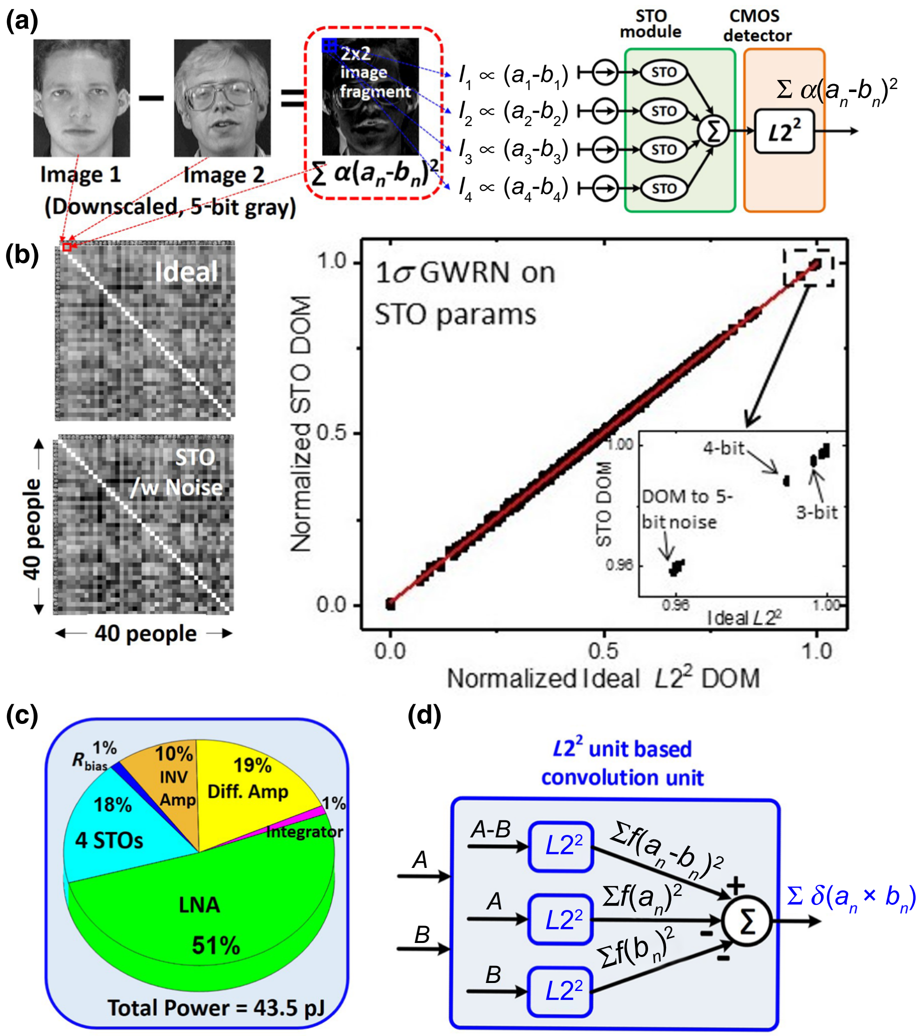

FIG. 6. (a) $L2^2$ distance computation with the AT&T face database [47]. (b) Left panel: matrices of ideal calculated face-to-face $L2^2$ (top) and STO approximate $L2^2$ (bottom) simulated using fitting curve from the measured CMOS detector [Fig. 5(d)], including (σ of the 2-LSB ADC code) noise. Right panel: face-to-face STO $L2^2$ versus ideal $L2^2$, each normalized from 0 (worst match) to 1 (best match). Also included are L22 distances to images with added GWRN of different bit depths. STO $L2^2$ calculated using 1σ GWRN on parameters in fitted CMOS detector curve. Inset: magnified view of high-DOM region. (c) The power breakdown of hybrid system for the four-dimensional case is shown. (d) Block diagrams for a convolution computing primitive based on Refs.[40,48].

energy efficiency to CMOS analog DCCs with at least an order of magnitude faster computation speed (Table I).

When n increases, the increase in total power consumption is only due to the power consumed by the STOs. In our system, each STO consumes 490 μW, can be improved via the use of higher magnetoresistive materials, tunneling magnetoresistive (TMR) structures, or spin-orbit excitation, and thus has the potential to be significantly more efficient than the CMOS oscillators in Ref. [27]. The power consumed by STOs is only 18% of the total power in our system and is less than 1% when we assume the per-STO power consumption as 1 μW [29,51]. Therefore, power per dimension of our system decreases and converges to the power consumption of a single STO as n increases. As mentioned in Sec. III, we need a stronger reference signal as n increases to obtain a more accurate distance measure. Furthermore, as n increases, it is desirable to have higher resolution ADC to resolve the smaller relative

TABLE I. Comparison with CMOS distance-computation circuits.

Ref.	No. inputs	Power [mW]	f_s [MHz]	Energy/dim.[pJ]
[4]	4 × 5	14.95	1	747.50
[7]	16 × 16	0.7	0.33	8.29
[50]	2 × 1	0.733	20	18.33
[5]	3 × 1	0.085	10	2.83
This work	4 × 1	10.87	250	10.87
	16 × 1	16.75	250	4.19
This work + projection [29]	4 × 1	8.92	357	6.25
	16 × 1	8.93	357	1.56

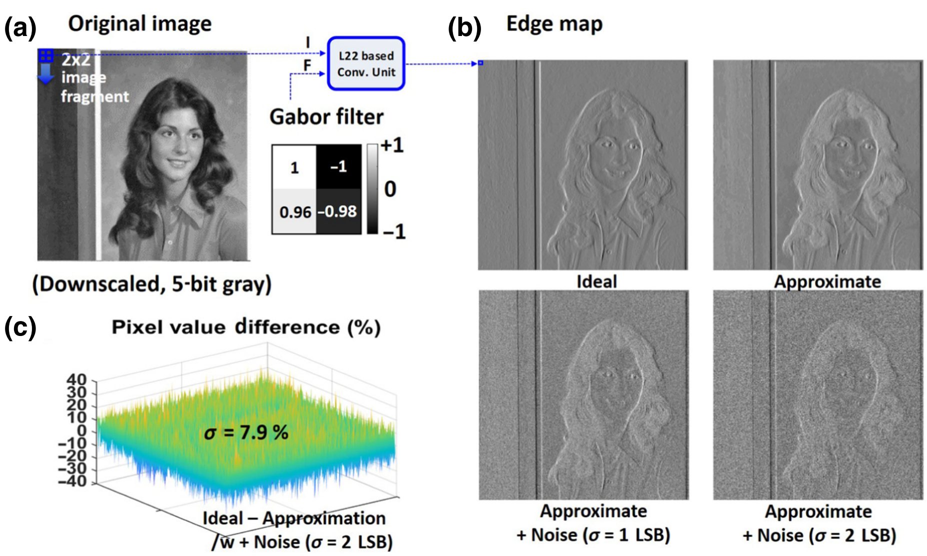

FIG. 7. (a) Edge detection on image (4.1.04 from SIPI database [52]) with the proposed computing primitive. (b) The edge map results from simulation (ideal), experiment with the fitting curve (approximate), and experiment with the fitting curve in addition to σ of 1- and 2-LSB noise. (c) The difference of the edge map between the ideal case and the approximation in addition to σ of 2-LSB noise.

changes of each STO. Depending on the noise in the system, this will set a limit on how large the STO arrays can scale.

In addition to computing the $L2$ between $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$, the $L2^2$ distance units [outlined in blue in Fig. 2(a)] can also be used to estimate convolution of two input vectors, A and B. Let us consider three $L2^2$ units having inputs of $(A - B)$, A, and B as shown in Fig. 6(d). The output of $L2^2(A - B) - L2^2(A) - L2^2(B)$ can be represented as $\alpha \sum_{i=1}^{n}(a_n \times b_n)$, which is proportional to the dot product of A and B (or convolution) [40,48].

The computing block for convolution is used in edge detection of images to determine the efficacy of the proposed coupled oscillatory network. Edge detection using a 2×2 kernel {using the Gabor filter [52] [Fig. 7(a)]} is performed through the following process: the image is converted to 5-bit grayscale and convolved with Gabor filter kernel of size 2×2. The Gabor filter kernels have been generated based on the model in Ref. [48]. For an image fragment I and the filter kernel F, the pixelwise differences $I - F$, $I - 0$, and $F - 0$ are calculated, and the corresponding bias currents are applied to three $L2^2$ units, shown in Fig. 6(d). The system outputs a level proportional to convolution, thus generating a single pixel of the output edge map. The entire edge map is obtained by sequentially sliding the image fragment window across the image. Note that 5-bit quantization is applied for a pixel intensity of the image and the Gabor filter kernel. Edge-detection results from the ideal convolution and approximate convolution based on our system are shown in Fig. 7(b) for different levels of system noise. Despite the additional noise, the images clearly show the edges present in the image, particularly when the noise is confined to the LSB. The difference between the ideal edge map and approximate one with normal distributed noise ($\sigma = 2$ LSB) is plotted in Fig. 7(c), where the variance of the difference is 7.9%.

VI. SUMMARY

We experimentally demonstrate a core distance-computing primitive based on an STO-based coupled-oscillator array. Starting from the theoretical background of obtaining an $L2^2$ norm from a coupled-oscillator array, we show that the combination of injection locking of the oscillators and their interference with a reference signal can be efficiently used to realize the distance-computation unit. The performance of the system as an $L2^2$ unit is examined by applying randomly generated test input vectors as bias current to the STOs and generating corresponding output digital codes from the CMOS detector. The characteristic curve from the experiment approximates an $L2^2$ norm which, in turn, is used to determine the feasibility of the STO-based coupled system for image-processing applications. The approximate distance and convolution output based on our system shows reasonable accuracy as compared to the ideal results. Energy and scaling analysis shows that GMR-based STOs for distance computation have higher efficiency than CMOS-based DCCs for high-dimensional input vectors. Modest improvements in STO critical currents and magnetoresistance (through the use of magnetic tunnel junctions) can make oscillator-based systems even more attractive.

ACKNOWLEDGMENTS

The authors thank David Payton (HRL Laboratories) for helpful discussions. The work is supported in part by the DARPA UPSIDE program, the Center for Brain Inspired Computing (C-BRIC), a MARCO and DARPA sponsored JUMP center, by the Semiconductor Research

Corporation, the National Science Foundation, Intel Corporation and by the Vannevar Bush Fellowship.

APPENDIX A: STO FABRICATION

The nanocontact STOs used in this work have a pseudospin valve magnetic heterostructure of the form (thicknesses in nm): Ta2/Cu(N)12/Ta3/Cu3/CoFe5/Cu4/[Co Fe0.33/Ni0.37]x4/CoFe0.33/Cu2/Ta4 grown by sputter deposition. The Cu(N) layer is sputter deposited in an Ar:N gas, which produces a smoother underlayer for better subsequent growth of the magnetic layers. The multilayer CoFe/Ni free layer has an effective in-plane magnetization $\mu_0 M_{\text{eff}} = 0.15$ T due to the surface anisotropy of the Ni interfaces [53,54].

The nanocontacts are formed via a self-aligned process. Electron-beam lithography is used to define an approximately 70-nm diameter resist pillar on the film stack using a negative tone resist, and a 50-nm SiN_x conformal layer then deposited. Mechanical polishing shears off the SiN_x and resist asperity, and after an O_2 plasma ash, results in a 70-nm nanocontact through the SiN_x. A Ta3/Au100 lead forming the center conductor of a coplanar waveguide contacts this nanocontact via to bias and read out the STO. 50 nm of SiN_x is then deposited over this structure, upon which the microwave microstrip is patterned to deliver the rf injection magnetic field.

APPENDIX B: RF CONDITIONING BOARD

The STO device array chip is mounted in an rf delivery and signal combining board that uses grounded CPW lines to deliver the injection signal, read out the STO output voltage dynamics, and combine the STO signals. The layout schematic of the board is shown in Fig. 8, and is designed for the measurement of STO arrays with an rf operation frequency in the 5–10 GHz range. The STO coplanar waveguide traces are contacted using spring-finger contacts (effective bandwidths > 10 GHz), enabling efficient changing of chips. The traces on the board are grounded, stitched, coplanar waveguides (bandwidths > 20 GHz), which connect microwave ICs for power combining, phase shifting, amplitude trimming, and amplification. Power combiners are Wilkinson coherent combiners. The reference signal line has microwave IC phase shifters (5.625° steps) and attenuators (0.5-dB steps) that enable phase-stable compensation of the parasitic signal on the STO lines. The STOs are each contacted by tapered coplanar waveguide to quasimicrostrip lines (effective bandwidth of > 10 GHz), and connected to pads at the edge of the chip. The different CPW path length to each STO are compensated on the signal conditioning board when combining the STO signals. These connect to bias tees with a cut-off frequency of approximately 10 MHz allow separation of microwave signals and bias currents. The dc lines from the bias-tee ICs to the off-board current supplies are not optimized for speed in this implementation, however, and limit the vector-loading operation to less than approximately 10 kHz. This speed restriction can be mitigated in an integrated implementation.

The Wilkinson power combiners and splitters used have an insertion loss of 4 dB, a typical imbalance in phase of 2.3° and amplitude of 0.1 dB. The two amplifiers used each have a gain of 20 dB, and an 8-GHz upper operating frequency. To adjust the amplitude and phase of the coupled reference signal, the injection signal is first split on the signal-conditioning board by a coherent power divider, adjusted in amplitude and phase by a digital attenuator (0.5-dB steps) and phase shifter (5.625° steps), and then coherently combined with the summed STO signals, producing the "net" reference signal referred to in the text. The board is constructed of nonmagnetic elements, to allow usage in the gap of an electromagnet.

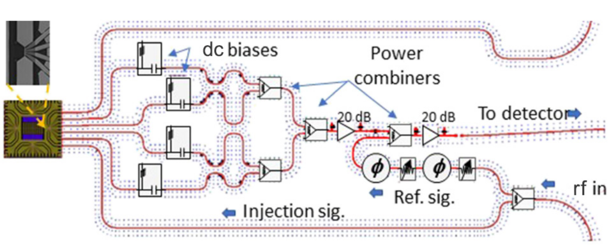

FIG. 8. Layout of microwave signal-conditioning board.

APPENDIX C: ADDITION OF SINUSOIDS AT THE SAME FREQUENCY BUT DIFFERENT PHASE

Note that this analysis with simple approximation is only for helping understanding of the system mathematically. The simulations shown in Figs. 2 and 3 used the equations without any approximation.

Addition of sinusoids at the same frequency but different phase can be expanded

$$\begin{aligned}\varphi &= \sum_{i=1}^{n} A_i \cos(\omega t - \delta_i)\\ &= \sum_{i=1}^{n} A_i[\cos(\omega t)\cos(\delta_i) + \sin(\omega t)\sin(\delta_i)]\\ &= \cos(\omega t) \times \sum_{i=1}^{n} A_i \cos(\delta_i) + \sin(\omega t) \times \sum_{i=1}^{n} A_i \sin(\delta_i).\end{aligned} \tag{C1}$$

Defining $A\cos\delta = \sum_{i=1}^{n} A_i \cos(\delta_i)$ and $A\sin\delta = \sum_{i=1}^{n} A_i \sin(\delta_i)$, Eq. (C1) becomes

$$\begin{aligned}\varphi &= A\cos(\omega t) \times \cos(\delta) + A\sin(\omega t) \times \sin(\delta)\\ &= A\cos(\omega t - \delta),\end{aligned} \tag{C2}$$

where

$$A^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} A_i A_j \cos(\delta_i - \delta_j),$$
$$\delta = \tan^{-1}\frac{\sum_{i=1}^{n} A_i \cos(\delta_i)}{\sum_{i=1}^{n} A_i \cos(\delta_i)}. \tag{C3}$$

Note that the amplitude A is a function of δ_i, which are the variables from which we calculate distance. If we define the last element of the sum as the reference signal $A_n = A_{\mathrm{ref}}$, $\delta_n = \delta_{\mathrm{ref}}$, then the amplitude A^2 can be written as

$$A^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} A_i A_j \cos\left(\delta_i - \delta_j\right) + 2A_{\mathrm{ref}}\sum_{j=1}^{n} A_j \cos\left(\delta_j\right) + A_{\mathrm{ref}}^2. \tag{C4}$$

We expand Eq. (C4) with the approximation of $\cos\theta \approx 1 - \theta^2/2$ as below.

$$A^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} A_i A_j \cos\left(\delta_i - \delta_j\right) + 2A_{\mathrm{ref}}\sum_{j=1}^{n} A_j \cos\left(\delta_j\right) + A_{\mathrm{ref}}^2 \cong \sum_{i=1}^{n}\sum_{j=1}^{n} A_i A_j \left(1 - \frac{(\delta_i - \delta_j)^2}{2}\right) + 2A_{\mathrm{ref}}\sum_{j=1}^{n} A_j \left(1 - \frac{\delta_j^2}{2}\right) + A_{\mathrm{ref}}^2$$
$$= (\alpha + A_{\mathrm{ref}})^2 - (\alpha + A_{\mathrm{ref}})\sum_{j=1}^{n} A_i\delta_i^2 + \left(\sum_{j=1}^{n} A_i\delta_i\right)^2, \tag{C5}$$

where $\alpha = \sum_{i=1}^{n} A_i$. The first term in this expression is an offset, and the second term is the desired response, proportional to the Euclidean distance squared of an n-dimensional vector $[\delta_1, \delta_2, \ldots, \delta_n]$. If the amplitudes are similar and the phase offsets are normally distributed around zero, the last term should average to zero for sufficiently large n. To see this, if we further assume that all oscillators have the same amplitude A_0, then $\alpha = nA_0$ and

$$A^2 = (A_0 n + A_{\mathrm{ref}})^2 - (nA_0^2 + A_{\mathrm{ref}}A_0)\sum_{j=1}^{n}\delta_j^2 + (A_0\sum_{j=1}^{n}\delta_j)^2$$
$$= (A_0 n + A_{\mathrm{ref}})^2 - [(n-1)A_0^2 + A_{\mathrm{ref}}A_0]\sum_{j=1}^{n}\delta_j^2 + A_0^2\left[(\sum_{j=1}^{n}\delta_j)^2 - \sum_{j=1}^{n}\delta_j^2\right]. \tag{C6}$$

The third term is an error term, the maximum of which can be derived using Chebyshev's sum inequality, as described in the text. Thus, we can obtain the approximated Euclidean distance if the error term is small enough to be tolerable.

APPENDIX D: DISTRIBUTION ANALYSIS

The distributions shown in Fig. 5 are formed with 5000 trials, a measurement that took 20 min due to the laboratory equipment used. Due to the phase-sensitive nature of the measurement, the zero point of the measurement drifted over this time, likely due to small temperature changes in the board changing effective path lengths. The amplitude of the drift is on the order of the device signal over the course of the measurement, and so can mask the device response. Consequently, a low-frequency smoothing procedure is applied to the data to remove this slow variation in the data. This is possible because in the measurement, a series of test bias currents are applied in random order to the STO to determine the distribution of the detector response. A time trace of the output of a single device is shown in Fig. 9(a) for both the diode and CMOS detectors. Since these current offsets are applied at random and are drawn from a Gaussian distribution around zero, the envelope of these traces should be uniform with time. Hence, the slow variations observed are not device response, but variations likely due to small changes in the relative phase condition of the signal and reference signals. These variations are slow, and are significant only due to the slow current sources used in this test (roughly 2-Hz data rate).

To remove these slow variations, a smoothing function is applied to the data [Savitsky-Golay second-order smoothing with 200 points, shown as the red curves in Fig. 9(a)] and subtracted. The diode detector is a standard microwave diode detector, sampled with a digital multimeter with approximately 6-Hz bandwidth. The CMOS detector, on the other hand, is controlled via an field-programmable gate array (FPGA) board, which is queried once per current bias. Each query returns a value that is the average of 5000 samples of the detector itself (the FPGA code is also modified to measure the RMS error on these samples is measured separately.) The banding evident in the CMOS detector response in Fig. 9(a) is due to the bit depth of the analog-to-digital converter on the detector. Note that the two detector traces are taken at the same time, so that the slow time variations of the two signals are similar but not identical. There is a slow quasiperiodic

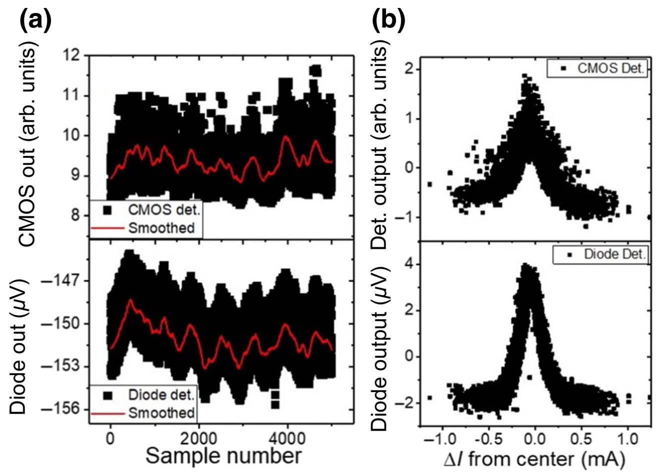

FIG. 9. Response of CMOS and microwave diode detectors to different current biases across phase-locking range of a single STO device. (a) Response as a function of time, for which random offsets are applied to the device. Smoothed curve shows slow drift of the signal. Samples taken at approximately 2-Hz rate (b) Resultant response versus current after slow drift signal shown in (a) subtracted from detector signals.

oscillation (possibly due to temperature variations) visible in both, while an even slower variation is also evident in the diode detection. It is unknown what the source of this additional variation is, but could be due to the larger bandwidth of the diode detector. The response functions with the slow variations removed are shown as a function of bias current difference in Fig. 9(b), showing a proper peaked response of the device signal suitable for the distance computation. The use of faster measurement times (as would occur in an integrated system), or the use of a self-referencing algorithm that dynamically accounts for phase drift, can remove the need for this additional step of removing the slow drift in the detector signal.

[1] S.-R. Kuang, J.-P. Wang, and C.-Y. Guo, Modified booth multipliers with a regular partial product array, IEEE Trans. Circuits Syst. II: Express Briefs **56**, 404 (2009).

[2] E. Antelo, P. Montuschi, and A. Nannarelli, Improved 64-bit radix-16 booth multiplier based on partial product array height reduction, IEEE Trans. Circuits Syst. I: Regular Papers **64**, 409 (2016).

[3] A. Gopalan and A. H. Titus, A new wide range euclidean distance circuit for neural network hardware implementations, IEEE Trans. Neural Netw. **14**, 1176 (2003).

[4] B.-D. Liu, C.-Y. Chen, and J.-Y. Tsao, A modular current-mode classifier circuit for template matching application, IEEE Trans. Circuits Syst. II: Analog Digital Signal Process. **47**, 145 (2000).

[5] T. Talaśka, M. Kolasa, R. Długosz, and W. Pedrycz, Analog programmable distance calculation circuit for winner takes all neural network realized in the cmos technology, IEEE Trans. Neural Netw. Learn. Syst. **27**, 661 (2015).

[6] G. T. Tuttle, S. Fallahi, and A. A. Abidi, in *1993 IEEE International Solid-State Circuits Conference Digest of Technical Papers* (IEEE, San Francisco, 1993), p. 38.

[7] G. Cauwenberghs and V. Pedroni, A low-power cmos analog vector quantizer, IEEE J. Solid-State Circuits **32**, 1278 (1997).

[8] A. Raychowdhury, A. Parihar, G. H. Smith, V. Narayanan, G. Csaba, M. Jerry, W. Porod, and S. Datta, Computing with networks of oscillatory dynamical systems, Proc. IEEE **107**, 73 (2019).

[9] D. E. Nikonov, G. Csaba, W. Porod, T. Shibata, D. Voils, D. Hammerstrom, I. A. Young, and G. I. Bourianoff, Coupled-oscillator associative memory array operation for pattern recognition, IEEE J. Exploratory Solid-State Comput. Devices Circuits **1**, 85 (2015).

[10] S. P. Levitan, Y. Fang, D. H. Dash, T. Shibata, D. E. Nikonov, and G. I. Bourianoff, in *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications* (IEEE, Turin, 2012), p. 1.

[11] V. Narayanan, S. Datta, G. Cauwenberghs, D. Chiarulli, S. Levitan, and P. Wong, in *2014 Design, Automation Test in Europe Conference Exhibition (DATE)* (IEEE, Dresden, 2014), p. 1.

[12] D. Fan, S. Maji, K. Yogendra, M. Sharad, and K. Roy, Injection-locked spin Hall-induced coupled-oscillators for energy efficient associative computing, IEEE Trans. Nanotechnol. **14**, 1083 (2015).

[13] J. A. Carpenter, Y. Fang, C. N. Gnegy, D. M. Chiarulli, and S. P. Levitan, in *2014 14th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA)* (IEEE, Notre Dame, 2014), p. 1.

[14] T. Shibata, R. Zhang, S. P. Levitan, D. E. Nikonov, and G. I. Bourianoff, in *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications* (IEEE, Turin, 2012), p. 1.

[15] G. Csaba, M. Pufall, D. E. Nikonov, G. I. Bourianoff, A. Horvath, T. Roska, and W. Porod, in *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications* (IEEE, Turin, 2012), p. 1.

[16] C. Huygens, *Oeuvres Complètes de Christiaan Huygens* (M. Nijhoff, The Hague, 1899).

[17] B. Ermentrout, An adaptive model for synchrony in the firefly pteroptyx malaccae, J. Math. Biol. **29**, 571 (1991).

[18] S.-Y. Ha, E. Jeong, and M.-J. Kang, Emergent behaviour of a generalized Viscek-type flocking model, Nonlinearity **23**, 3139 (2010).

[19] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine* (MIT Press, Cambridge, 1965), Vol. 25.

[20] G. Buzsáki and A. Draguhn, Neuronal oscillations in cortical networks, Science **304**, 1926 (2004).

[21] P. Baldi and R. Meir, Computing with arrays of coupled oscillators: An application to preattentive texture discrimination, Neural Comput. **2**, 458 (1990).

[22] F. C. Hoppensteadt and E. M. Izhikevich, Synchronization of mems resonators and mechanical neurocomputing, IEEE Trans. Circuits Syst. I: Fundamental Theory Appl. **48**, 133 (2001).

[23] F. C. Hoppensteadt and E. M. Izhikevich, Pattern recognition via synchronization in phase-locked loop neural networks, IEEE Trans. Neural Netw. **11**, 734 (2000).

[24] E. Vassilieva, G. Pinto, J. A. De Barros, and P. Suppes, Learning pattern recognition through quasi-synchronization of phase oscillators, IEEE Trans. Neural Netw. **22**, 84 (2010).

[25] D. Vodenicarevic, N. Locatelli, F. A. Araujo, J. Grollier, and D. Querlioz, A nanotechnology-ready computing scheme based on a weakly coupled oscillator network, Sci. Rep. **7**, 44772 (2017).

[26] H. Arai and H. Imamura, Neural-Network Computation Using Spin-Wave-Coupled Spin-Torque Oscillators, Phys. Rev. Appl. **10**, 024040 (2018).

[27] D. E. Nikonov, P. Kurahashi, J. S. Ayers, H. J. Lee, Y. Fan, and I. A. Young, A coupled cmos oscillator array for 8ns and 55pj inference in convolutional neural net works, arXiv:1910.11803 (2019).

[28] D. E. Nikonov, P. Kurahashi, J. S. Ayers, H. J. Lee, Y. Fan, and I. A. Young, Convolution inference via synchronization of a coupled cmos oscillator array, arXiv:1910.11802 (2019).

[29] M. Romera, P. Talatchian, S. Tsunegi, F. A. Araujo, V. Cros, P. Bortolotti, J. Trastoy, K. Yakushiji, A. Fukushima, and H. Kubota, *et al.*, Vowel recognition with four coupled spin-torque nano-oscillators, Nature **563**, 230 (2018).

[30] N. Shukla, A. Parihar, M. Cotter, M. Barth, X. Li, N. Chandramoorthy, H. Paik, D. G. Schlom, V. Narayanan, A. Raychowdhury, and S. Datta, in *2014 IEEE International Electron Devices Meeting* (IEEE, San Francisco, 2014), p. 28.7.1.

[31] N. Shukla, W.-Y. Tsai, M. Jerry, M. Barth, V. Narayanan, and S. Datta, in *2016 IEEE Symposium on VLSI Technology* (IEEE, Honolulu, 2016), p. 1.

[32] S. Dutta, A. Parihar, A. Khanna, J. Gomez, W. Chakraborty, M. Jerry, B. Grisafe, A. Raychowdhury, and S. Datta, Programmable coupled oscillators for synchronized locomotion, Nat. Commun. **10**, 3299 (2019).

[33] N. Gala, S. Krithivasan, W.-Y. Tsai, X. Li, V. Narayanan, and V. Kamakoti, An accuracy tunable non-Boolean coprocessor using coupled nano-oscillators, ACM J. Emerging Technol. Comput. Syst. (JETC) **14**, 1 (2018).

[34] P. Talatchian, M. Romera, F. A. Araujo, P. Bortolotti, V. Cros, D. Vodenicarevic, N. Locatelli, D. Querlioz, and J. Grollier, Designing large arrays of interacting spin-torque nano-oscillators for microwave information processing, arXiv:1908.09908 (2019).

[35] S. Kaka, M. R. Pufall, W. H. Rippard, T. J. Silva, S. E. Russek, and J. A. Katine, Mutual phase-locking of microwave spin torque nano-oscillators, Nature **437**, 389 (2005).

[36] D. Houssameddine, U. Ebels, B. Delaët, B. Rodmacq, I. Firastrau, F. Ponthenier, M. Brunet, C. Thirion, J.-P. Michel, L. Prejbeanu-Buda, M.-C. Cyrille, O. Redon, and B. Dieny, Spin-torque oscillator using a perpendicular polarizer and a planar free layer, Nat. Mater. **6**, 447 (2007).

[37] P. M. Braganca, B. A. Gurney, B. A. Wilson, J. A. Katine, S. Maat, and J. R. Childress, Nanoscale magnetic field detection using a spin torque oscillator, Nanotechnology **21**, 235202 (2010).

[38] W. H. Rippard, M. R. Pufall, S. Kaka, S. E. Russek, and T. J. Silva, Direct-Current Induced Dynamics in $co_{90}fe_{10}/ni_{80}fe_{20}$ Point Contacts, Phys. Rev. Lett. **92**, 027201 (2004).

[39] W. H. Rippard, M. R. Pufall, S. Kaka, T. J. Silva, S. E. Russek, and J. A. Katine, Injection Locking and Phase Control of Spin Transfer Nano-Oscillators, Phys. Rev. Lett. **95**, 067203 (2005).

[40] D. M. Chiarulli, B. Jennings, Y. Fang, A. Seel, and S. P. Levitan, in *2015 IEEE Computer Society Annual Symposium on VLSI* (IEEE, Montpellier, 2015), p. 125.

[41] Y. Zhou, J. Persson, and J. Åkerman, Intrinsic phase shift between a spin torque oscillator and an alternating current, J. Appl. Phys. **101**, 09A510 (2007).

[42] Y. Zhou, J. Persson, S. Bonetti, and J. Åkerman, Tunable intrinsic phase of a spin torque oscillator, Appl. Phys. Lett. **92**, 092505 (2008).

[43] N. Oo and W.-S. Gan, On harmonic addition theorem, Int. J. Comput. Commun. Eng. **1**, 200 (2012).

[44] G. D. Geronimo, P. O'Connor, and A. Kandasamy, Analog cmos peak detect and hold circuits. Part 1. Analysis of the classical configuration, Nucl. Instrum. Methods Phys. Res. Sec. A **484**, 533 (2002).

[45] G. D. Geronimo, P. O'Connor, and A. Kandasamy, Analog cmos peak detect and hold circuits. Part 2. The two-phase offset-free and derandomizing configuration, Nucl. Instrum. Methods Phys. Res. Sec. A **484**, 544 (2002).

[46] W. Rippard, M. Pufall, and A. Kos, Time required to injection-lock spin torque nanoscale oscillators, Appl. Phys. Lett. **103**, 182403 (2013).

[47] F. S. Samaria and A. C. Harter, in *Proceedings of the Second IEEE Workshop on Applications of Computer Vision, 1994* (IEEE, Sarasota, 1994), p. 138.

[48] D. E. Nikonov, I. A. Young, and G. I. Bourianoff, Convolutional networks for image processing by coupled oscillator arrays, arXiv:1409.4469 (2014).

[49] D. Gabor, Theory of communication. Part 1: The analysis of information, J. Inst. Electr. Eng.-Part III: Radio Commun. Eng. **93**, 429 (1946).

[50] D. Fernandez, L. Martínez-Alvarado, and J. Madrenas, A translinear, log-domain fpaa on standard cmos technology, IEEE J. Solid-State Circuits **47**, 490 (2011).

[51] M. Gajek, J. Nowak, J. Sun, P. Trouilloud, E. O'Sullivan, D. Abraham, M. Gaidis, G. Hu, S. Brown, and Y. Zhu, *et al.* Spin torque switching of 20 nm magnetic tunnel junctions with perpendicular anisotropy, Appl. Phys. Lett. **100**, 132408 (2012).

[52] U. Signal, Image processing institute, USC–SIPI image database, Aviliable on http://sipi.usc.edu/database.

[53] W. H. Rippard, A. M. Deac, M. R. Pufall, J. M. Shaw, M. W. Keller, S. E. Russek, G. E. Bauer, and C. Serpico, Spin-transfer dynamics in spin valves with out-of-plane magnetized coni free layers, Phys. Rev. B **81**, 014426 (2010).

[54] J. M. Shaw, H. T. Nembach, and T. J. Silva, Damping phenomena in co90fe10/ni multilayers and alloys, Appl. Phys. Lett. **99**, 012503 (2011).