

# Nonboolean Pattern Recognition Using Chains of Coupled CMOS Oscillators as Discriminant Circuits

VAHNOOD POURAHMAD<sup>1,2</sup>, SASIKANTH MANIPATRUNI<sup>3</sup> (Member, IEEE),  
DMITRI NIKONOV<sup>3</sup> (Member, IEEE), IAN YOUNG<sup>3</sup> (Fellow, IEEE),  
and EHSAN AFSHARI<sup>1,2</sup> (Senior Member, IEEE)

<sup>1</sup>Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA

<sup>2</sup>Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI, USA

<sup>3</sup>Intel Corporation, Hillsboro, OR, USA

CORRESPONDING AUTHOR: V. Pourahmad (vp87@cornell.edu)

This work was supported by the MSR Intel program 2346.001 via Semiconductor Research Corporation. This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author. This consists of a PDF file containing additional information not included in the paper. This material is 2.14 MB in size.

**ABSTRACT** Networks of weakly coupled oscillators are special cases of Hopfield networks and as such may be used in building physical systems capable of associative memory recall and pattern recognition. However, the existing architectures are not suitable for hardware implementation mainly due to the complexity of required couplings between the oscillators. In this paper, we propose an alternative way of using coupled oscillators in what we call “discriminant circuits” in analogy to the concept of “discriminant functions” in the field of artificial intelligence and machine learning. The main advantage of our system is in the simplicity of its architecture, which relies only on local couplings between adjacent oscillators. Using this architecture, we design a network of coupled CMOS oscillators as the core of a physical non-Boolean pattern recognition engine. The simplicity of the proposed circuit makes it readily implementable on any standard CMOS technology.

**INDEX TERMS** Associative memory, CMOS integrated circuits, coupled oscillators, Hopfield neural networks, ring oscillators.

## I. INTRODUCTION

RECENTLY, there has been a rise of interest in using coupled oscillators as non-Boolean processors [1]–[6]. Non-Boolean processing refers to a set of computational tasks that are usually associated with the realm of artificial intelligence and machine learning, such as pattern recognition and associative memory recall. What these computational problems have in common is the inefficiency, if not impossibility, of developing algorithms based solely on combinational (Boolean) logic for their solutions.

Networks of coupled oscillators have already been studied as special examples of Hopfield [7]–[9] networks that are famously capable of image reconstruction and associative memory recall [10], [11], [13], [14]. However, it is difficult to see how these theoretical considerations can be used to realize a physical associative memory engine. The main reason for this is that in general, a network of  $n$  oscillators requires  $n^2$  programmable connections between oscillators [11], [12]. This can be easily seen from the well-known

Hebbian learning rule [11], [13], which determines the coupling properties of the network and is formulated as

$$C_{ij} = \sum_{k=1}^m \xi_i^{(k)} \overline{\xi_j^{(k)}} \quad (1)$$

where  $C_{ij}$  is a complex number that determines the properties (strength and inherent phase shift) of the coupling from the  $j$ th to the  $i$ th oscillator,  $\xi_i^{(k)}$  denotes the  $k$ th pattern to be stored in the network, and  $m$  is the total number of patterns to be stored. A short derivation of Hebbian learning rule from the Kuramoto [15]–[17] model of weakly coupled oscillators is presented in the Supplementary Material.

Hoppensteadt *et al.* [11] proposed an elegant way of bypassing the connectivity problem in the networks of coupled oscillators. In their approach, there are no direct connections between the oscillators and instead they are all coupled through a shared medium. The shared medium, however, needs to be excited by a carefully constructed signal so that

the final mathematical equations governing the dynamics of oscillators are identical to that of the fully connected network with coupling coefficients given by (1). This approach does not fully solve the connectivity problem and instead trades it with another, perhaps equally hard, problem of generating the required global signal that excites the shared medium [11]. Also the network design becomes complicated for larger  $n$ , making it difficult to realize large networks of coupled oscillators [11]. The possibility that oscillators will be connected to the shared medium by wires of different lengths that will also have different signal delays is a reason for concern. In fact, a physical implementation of this idea as reported in [14] was limited to a network of eight oscillators and also required an off-chip signal generator for exciting the shared medium.

In the Supplementary Material, we also show how the problem of **spurious pattern storage** [8], [9] naturally arises as a consequence of the Hebbian learning rule for weakly coupled oscillators. The number of stored spurious patterns quickly increases with the number of desired stored patterns [8], imposing additional limitations on the storage capacity of such networks.

Alternatively, an effective and widely adopted approach for solving a variety of pattern recognition problems is to use “discriminant functions” or more generally “artificial neural networks” [18]–[20].

The basic idea of discriminant function approach is to start with a function with some free parameters and then tune these parameters so that the function’s output mainly depends on the particular “class” to which its input belongs. As an example, suppose we want to distinguish between input patterns belonging to two classes,  $\Omega_1$ ,  $\Omega_2$ , using a discriminant function approach. We show the discriminant function by  $f(\mathbf{x}, \boldsymbol{\eta})$  where  $\boldsymbol{\eta}$  is a vector of the free parameters of the function and  $\mathbf{x}$  is the function’s input. To make  $f(\mathbf{x}, \boldsymbol{\eta})$  a good discriminant function, one possibility is to look for a vector  $\boldsymbol{\eta}$  such that

$$\begin{aligned} f(\mathbf{x}, \boldsymbol{\eta}) &> 0 & \text{if } \mathbf{x} \in \Omega_1 \\ f(\mathbf{x}, \boldsymbol{\eta}) &< 0 & \text{if } \mathbf{x} \in \Omega_2. \end{aligned}$$

The process of finding vector  $\boldsymbol{\eta}$  in above-mentioned is called the “training phase.” After training is complete, we can reliably use  $f(\mathbf{x}, \boldsymbol{\eta})$  as a discriminant function to classify input patterns to either class  $\Omega_1$  or  $\Omega_2$  by only looking at the sign of  $f(\mathbf{x}, \boldsymbol{\eta})$ .

Building upon the idea of discriminant functions, we propose to use the term “discriminant circuit” to describe any physical circuit with some free control parameters that can be used in pattern recognition applications. The control parameters of a discriminant circuit should similarly be tuned such that the circuit’s output depends mainly on some collective property of its input signals. The circuit’s output signal can then be used to distinguish different input patterns from each other.

While complex software algorithms, enjoying access to several layers of arbitrary discriminant functions, provide

the greatest degree of flexibility, discriminant circuits would offer much faster processing speeds. Convergence to the final steady state in the networks of coupled oscillators typically takes a few hundred oscillation cycles. Therefore, faster processing speeds are achieved if oscillators with higher free running frequencies are used in forming the network. It is also worth mentioning that while it is true that convergence to the steady state may take longer in longer chains of coupled oscillators, the average convergence time does not seem to get impractically long. In fact, our numerical simulations showed that in changing the length of an oscillator chain from 25 to 50 to 100, the median of the (normalized) convergence time increased from 93.75 to 112.3 and 128.8 (normalized time units), respectively. That is to say, the median convergence time seems to rise sublinearly with the chain size.

The rest of this paper is organized as follows. In Section II, we study the nonlinear dynamics of two coupled oscillators and show how arbitrary phase shifts can be enforced between them by controlling the ratio of two different coupling channel strengths. Accurate circuit simulation results using the device models of ST Bi-COMS 130-nm technology show remarkable agreement between the mathematical model and the simulated behavior of the circuit. In Section III, we present a simple chain of coupled oscillator with tunable coupling parameters as a discriminant circuit and show how it can be used to realize the core of a non-Boolean pattern recognition engine with CMOS oscillators. The simplicity of the coupling structure makes the proposed circuit design readily implementable in any standard CMOS technology. Finally, Section IV discusses future research directions and concludes this paper.

## II. NONLINEAR DYNAMICS OF TWO COUPLED OSCILLATORS

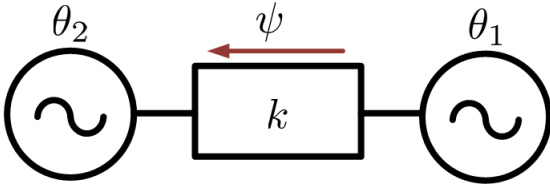
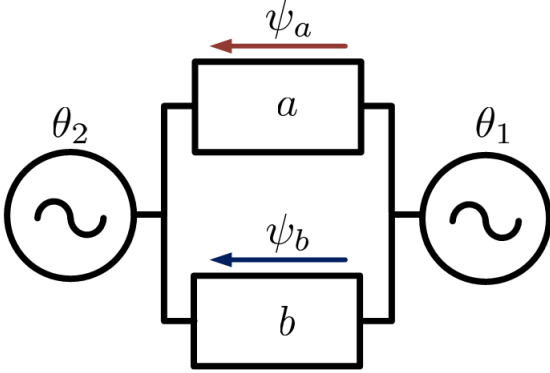
In this section, we study the nonlinear dynamics of two weakly coupled oscillators using the “phase-reduced,” or perhaps better known “Kuramoto,” model [15]–[17]. In particular, we analyze the case where two symmetric coupling channels exist between the oscillators and show how by controlling the ratio of the two coupling mode strengths, arbitrary phase shifts are enforced between the oscillators.

The celebrated Kuramoto’s approximation to the dynamics of two weakly coupled oscillators with a “symmetric” coupling gives [17]

$$\dot{\theta}_1 = k \sin(\theta_2 - \theta_1 - \psi) \quad (2)$$

$$\dot{\theta}_2 = k \sin(\theta_1 - \theta_2 + \psi) \quad (3)$$

where, as shown in Fig. 1,  $\theta_1$  and  $\theta_2$  are the phase variables associated with the two oscillators, and  $k$  and  $\psi$  show the strength and inherent phase shift of the coupler, respectively. As a consequence of symmetric coupling assumption [13], [17], the inherent phase shift term  $\psi$  appears in the first equation with a minus sign and in the second equation with a plus sign, consistent with the reference direction of coupling shown by the arrow in Fig. 1. With reference


**FIGURE 1.** Kuramoto model of two weakly coupled oscillators.

**FIGURE 2.** Bimodal, symmetric coupling of two oscillators.

to Fig. 2, let us now assume that two symmetric coupling channels with the strengths of  $a/2$  and  $b/2$  and inherent phase shifts of  $\psi_a$  and  $\psi_b$  exist between two oscillators. The Kuramoto model of this system is the sum of expressions for each coupling channel<sup>1</sup>

$$\begin{aligned}\dot{\theta}_1 &= \frac{a}{2} \sin(\theta_2 - \theta_1 - \psi_a) + \frac{b}{2} \sin(\theta_2 - \theta_1 - \psi_b) \\ \dot{\theta}_2 &= \frac{a}{2} \sin(\theta_1 - \theta_2 + \psi_a) + \frac{b}{2} \sin(\theta_1 - \theta_2 + \psi_b).\end{aligned}$$

Therefore, the dynamics of the relative phase shift between the oscillators,  $\varphi = \theta_2 - \theta_1$ , is governed by

$$\dot{\varphi} = a \sin(-\varphi + \psi_a) + b \sin(-\varphi + \psi_b). \quad (4)$$

We now proceed to find the steady-state solution of (4). First, notice that a Lyapunov function [21] exists for this system, which guarantees convergence to steady state

$$\begin{aligned}E(\varphi) &= -a \cos(-\varphi + \psi_a) - b \cos(-\varphi + \psi_b) \\ \Rightarrow \dot{E} &= -(\dot{\varphi})^2 \leq 0.\end{aligned}$$

A stable steady-state solution of (4) necessarily corresponds to a local minima of  $E(\varphi)$

$$\begin{aligned}\frac{dE}{d\varphi} &= 0 \\ \Rightarrow \frac{d}{d\varphi} \{a \cos(-\varphi + \psi_a) + b \cos(-\varphi + \psi_b)\} &= 0 \\ \frac{d^2E}{d\varphi^2} &> 0 \\ \Rightarrow \frac{d^2}{d\varphi^2} \{a \cos(-\varphi + \psi_a) + b \cos(-\varphi + \psi_b)\} &< 0.\end{aligned}$$

<sup>1</sup>This linear behavior is a consequence of approximating phase disturbances around the oscillator's limit cycle with the gradient of its phase function, see [16], [17] for more details.

In order to get a final closed formula for  $\varphi$ , it is better to rewrite  $E(\varphi)$  using Euler's formula

$$\begin{aligned}E(\varphi) &= -\frac{a}{2} [e^{i(-\varphi + \psi_a)} + e^{i(\varphi - \psi_a)}] \\ &\quad -\frac{b}{2} [e^{i(-\varphi + \psi_b)} + e^{i(\varphi - \psi_b)}].\end{aligned}$$

Simple algebraic manipulation reveals that steady-state condition,  $(dE/d\varphi) = 0$ , gives

$$e^{i\varphi}(ae^{-i\psi_a} + be^{-i\psi_b}) = e^{-i\varphi}(ae^{i\psi_a} + be^{i\psi_b}).$$

Since the left- and right-hand sides of the above-mentioned equation are complex conjugates of each other, they should be purely real to be equal, that is

$$e^{-i\varphi}(ae^{i\psi_a} + be^{i\psi_b}) \in \mathbb{R}.$$

There are only two possibilities for  $\varphi$  to satisfy this requirement

$$\varphi = \begin{cases} \text{Arg}(ae^{i\psi_a} + be^{i\psi_b}) \pm 2n\pi \\ \text{Arg}(ae^{i\psi_a} + be^{i\psi_b}) \pm (2n + 1)\pi. \end{cases}$$

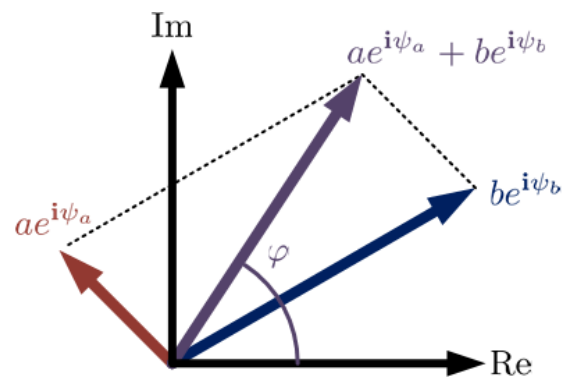
On the other hand, the stability condition,  $(d^2E/d\varphi^2) > 0$ , implies

$$\frac{d^2E}{d\varphi^2} > 0 \Rightarrow e^{i\varphi}(ae^{-i\psi_a} + be^{-i\psi_b}) > 0.$$

Therefore, the stable steady-state phase shift between the oscillators is

$$\varphi = \text{Arg}(ae^{i\psi_a} + be^{i\psi_b}). \quad (5)$$

The above-mentioned result has an interesting intuitive interpretation: if one considers the effect of each coupling channel in Fig. 2 separately, it is easy to see that the top coupling pushes the relative phase shift between the two oscillators to  $\psi_a$ , whereas the bottom coupler pulls it toward  $\psi_b$ . In the presence of both coupling channels, the final phase shift can be found from the weighted vector superposition shown in Fig. 3.


**FIGURE 3.** Weighted superposition of two coupling channels.

It is not difficult to realize different coupling channels between physical oscillators. In a three-stage CMOS ring

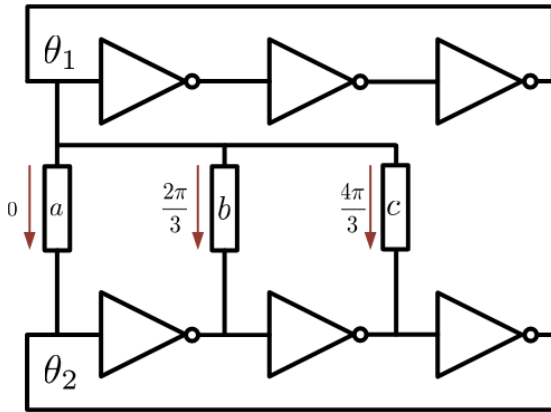


FIGURE 4. Different coupling channels of two CMOS ring oscillators.

oscillator shown in Fig. 4, adjacent inverter gate voltages naturally have a phase shift of  $2\pi/3$  with respect to each other and therefore depending on which points of two such oscillators are connected via a resistor, different coupling modes can be realized. For example, if only the middle coupling in Fig. 4 is active, then  $\theta_2 + 2\pi/3$  should eventually become equal to  $\theta_1$  to minimize the energy loss in the resistive connection between the two oscillators. This condition is then consistent with having  $\psi = 2\pi/3$  in (2). Similarly, one can argue that the leftmost and rightmost couplings shown in Fig. 4 have inherent phase shifts of 0 and  $4\pi/3$ , respectively.

The circuit level schematic of the coupled ring structure is shown in Fig. 5. The three inverter gates in oscillators (1) and (2) are labeled as **1A**, **1B**, and **1C** and **2A**, **2B**, **2C**, respectively. The input terminals of the two couplers in the middle are likewise labeled to imply connectivity to the respective oscillator terminals without drawing explicit wire connections. The couplers are CMOS transmission gates, which are essentially resistors that can be digitally switched ON or OFF by their gate control voltages,  $S_1$  and  $S_2$ . Since conductance of transistors are to the first-order proportional to their aspect ratios,  $(W/L)$ , a discrete set of coupling strengths can be realized for each coupling channel by laying out a bank of differently sized transmission gates on chip. The strength of each coupling channel can then be digitally controlled by activating the transmission gate with the desired size from this available set.

The circuit in Fig. 5 has been simulated using accurate device models of ST BiCMOS 130-nm technology in Cadence Virtuoso. As can be seen from Fig. 6, the simulated steady-state phase shifts of oscillators agree very well with what (5) predicts.

### III. PATTERN RECOGNITION WITH COUPLED OSCILLATOR CHAINS

In this section, we generalize the case of two coupled oscillators to a chain of coupled oscillators and show how this structure can be used as a “discriminant circuit” in the sense we explained in Section I.

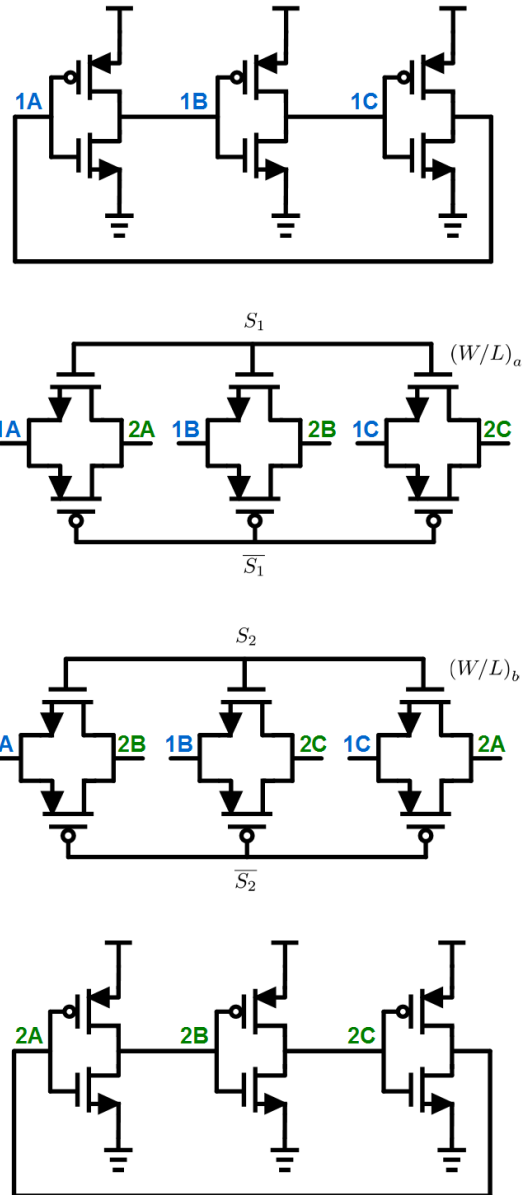
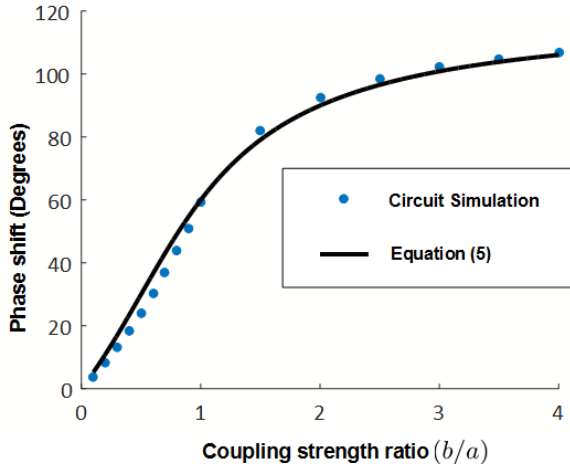


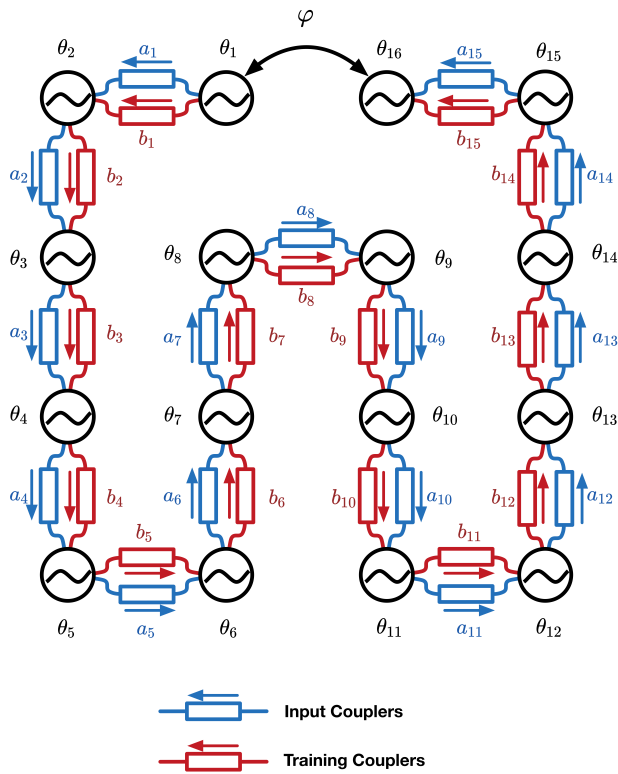
FIGURE 5. Circuit level schematic of CMOS-coupled oscillators.

Fig. 7 shows a folded chain of 16 ring oscillators where two types of tunable couplers, “input” and “training” couplers, are available in each link. The system response is defined as the phase shift between the oscillators at the two ends of the chain, which is shown by  $\varphi$ . In a simply connected chain of  $N$  oscillators, there are exactly  $l = N - 1$  links and therefore  $l$  different instances of “input” and “training” couplers each. For a practical implementation, it is best to fold the chain into a rectangular configuration similar to what is shown in Fig. 7 to make the actual layout as compact as possible. In the Supplementary Material, we outline the general way of constructing such folded chains for rectangular networks of oscillators with even number of rows and columns.

The “input couplers” have zero inherent phase shift and their strengths,  $a_k$ , are determined by a given 2-D input



**FIGURE 6.** Agreement between theoretical and simulation results.



**FIGURE 7.** Chain of oscillators with tunable couplings.

pattern as explained in the Supplementary Material. The “training couplers” on the other hand have an inherent phase shift of  $2\pi/3$  and their relative strengths,  $b_k$ , will be determined in the training phase. In other words, vectors  $\mathbf{a} = [a_1 \dots a_k]^T$  and  $\mathbf{b} = [b_1 \dots b_l]^T$  are analogous to vectors  $\mathbf{x}$  and  $\boldsymbol{\eta}$  in the discriminant function  $f(\boldsymbol{\eta}, \mathbf{x})$  formulation introduced in Section I.

Intuitively, we expect that since there are no closed loops in the folded chain structure, the phase shifts in each link can

be independently set to

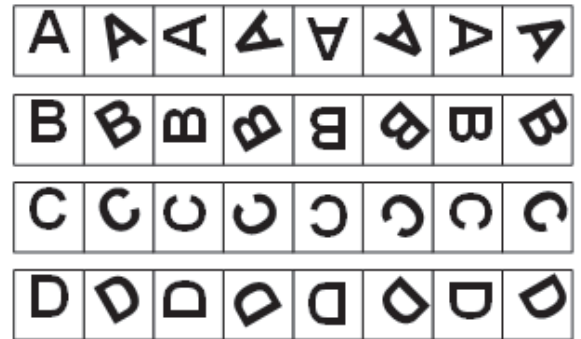
$$\varphi_k = \text{Arg}(a_k + b_k e^{i2\pi/3})$$

so as to minimize the total Lyapunov energy of the chain. Consequently, a simple formula for the circuit output,  $\varphi$ , as a function of the circuit coupling parameters exists

$$\varphi = \sum_{k=1}^l \text{Arg}(a_k + b_k e^{i2\pi/3}). \quad (6)$$

If we had closed loops in the network, a constrained minimization using Lagrange multipliers would be necessary to ensure that sum of phase shifts around each loop would add up to some multiple of  $2\pi$ , and (6) would not necessarily be valid. A more formal derivation of (6) can be found in the Supplementary Material.

It is perhaps best to explain the circuit’s operation with the aid of an example: suppose we want to train a coupled oscillator chain network to distinguish some alphabet letters irrespective of their orientation. To be more specific, imagine that we want to classify the 32,  $20 \times 20$  pixel images of four alphabet letters shown in Fig. 8.



**FIGURE 8.** Alphabet letter images used for network training.

As explained before, the input coupler strengths,  $a_k$ , are dictated by the input image. That is, we assign each  $a_k$  to one pixel in the input image and set its (normalized) value to a minimum of say, 0.1, if the corresponding pixel is a logical zero or a maximum of 1 if the corresponding pixel is a logical one. Although we are assuming that the input pattern is a binary image for simplicity in this example, this is not essential to the circuit’s operation and, in general,  $a_k$  parameters can assume gray-scale values.

Our objective in the “training phase” is to find a good choice of training coupler strengths,  $b_k$ , that would result in a discriminatory response of the circuit. For example, a good discriminatory circuit response happens when  $\varphi$  is concentrated around  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ ,  $270^\circ$  for the given samples of **A**, **B**, **C**, **D**, respectively.

A practical training algorithm should also be limited to a finite choice of  $b_k$  parameters from some available set of coupling strengths. We may thus define “training resolution” as the total number of bits allowed in specifying  $b_k$  parameters.

The system response to each input can be visually represented by placing a dot at an angle of  $\varphi$  on a unity circle where  $\varphi$  is calculated by (6). This visual representation is used in Fig. 9 to show constant improvement in the discriminatory behavior of the circuit with finer resolutions allowed in the training.

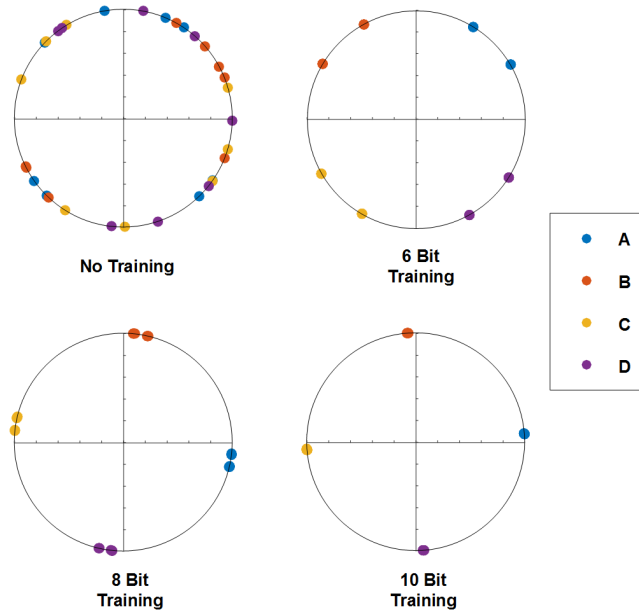


FIGURE 9. Effect of training on separation between different alphabet letter images.

The above-mentioned example is just meant to illustrate the operation of coupled oscillator chain as a “discriminant circuit” and should not be considered a complete pattern recognition system in itself. In particular, since there is no “preprocessing” or “feature extraction” [18], [19] steps involved, the system is prone to the infamous **curse of dimensionality** problem [18], [19]. In other words, the network parameters may currently be overfitted to give perfect outputs for the given samples in the training set but that makes the system more vulnerable to noise and random distortions and degrades its ability to perform useful “generalizations” [18], [19].

While it is certainly more convenient to have only one discriminant circuit that classifies several different objects, it is perhaps better to have a network of discriminant circuits where each circuit is specially trained in recognizing just one particular object. A collaborative network of such specialized, discriminant circuits would then become a more powerful pattern recognition engine.

This strategy not only increases the total number of samples of each object that the system would be able to recognize after training, but also has the added benefit of simplifying the final circuit block that is necessary to interpret different values of  $\varphi$  for either positive or negative classifications. This is so because a simple “**phase frequency detector**” (PFD) found in standard “**phase-locked loops**” [22]–[24] designs

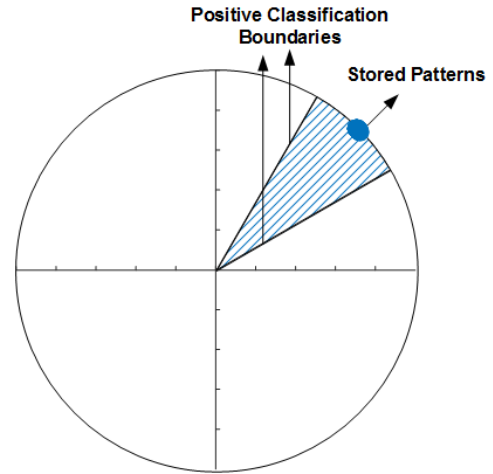


FIGURE 10. Behavior of a successfully trained network.

can be integrated with the circuit to determine whether  $\varphi$  is within some designated range or not. The output of the PFD block determines whether the circuit “**recognizes**” or “**rejects**” a pattern given to it for classification. Of course, as shown in Fig. 10, a successful training should be able to find a set of  $b_k$  parameters that results in  $\varphi$  to be within the designated range for all the samples used in training. It should be noted, however, that different discriminant circuits do not necessarily need to correspond to different physical instances on a chip. A single discriminant circuit can be sequentially loaded with the right training parameters developed for different objects to determine their presence in a given image.

The general way of using chains of coupled oscillators for pattern classification can thus be summarized as follows.

Let us show the object<sup>2</sup> we want the circuit to recognize by  $\xi$ . We also assume that several examples of  $\xi$ , which we show by  $\xi^{(1)} \dots \xi^{(m)}$ , are given that can be used in “training.” We pick a target value for the circuit’s output signal,  $\varphi = \varphi_0$ , load the network’s  $a_k$  parameters with each of the patterns  $\xi^{(1)} \dots \xi^{(m)}$ , and try to find a single set of  $b_k$  parameters that would result in the circuit producing an output in the vicinity of  $\varphi_0$  for every  $\xi^{(1)} \dots \xi^{(m)}$ . We store the above-mentioned set of  $b_k$  parameters in some library as the “training template” for  $\xi$  and call it  $b^\xi$  for future reference.

Later, when an unidentified image,  $\chi$ , is presented to the system by loading the corresponding values of  $a_k$  parameters, we use the set  $b^\xi$  from above-mentioned and see if  $\varphi$  happens to fall into the vicinity of  $\varphi_0$ . That is, we now have a testing method, which we may call the “**recognize/reject**” test for pattern  $\xi$ . If the result of the “recognize/reject” test is positive, then  $\chi$  is identified as another instance of  $\xi$ . Otherwise,  $\chi$  is considered to be an object different than  $\xi$  and consequently the next training template in the library is used to seek a match.

<sup>2</sup>By “object,” we mean some abstract concept, like “letter A” or “Cat,” that carries some meaning. By “example/sample,” we mean actual images associated with such concepts, like the image of a letter “A” or a picture of a cat

There still remains a couple of issues with our structure that need to be addressed at this point. These are what we may call the problems of “false” and “ambiguous” classifications and are analogous to the problem of “spurious storage” in Hopfield networks.

False classification happens when a pattern not similar to  $\xi^{(1)} \dots \xi^{(m)}$  is labeled as another example of  $\xi$ . This is possible to happen, because  $b^\xi$  was developed during the training phase to make sure that when input image happens to be any one of  $\xi^{(1)} \dots \xi^{(m)}$ , then  $\varphi \approx \varphi_0$  without necessarily preventing other, perhaps unrelated, images to result in producing the same values of  $\varphi$ .

Ambiguous classification is a consequence of false classification problem and happens when an input image,  $\chi$ , can be classified as an instance of more than one object, say both  $\xi$  and  $\lambda$ , because with both  $b^\xi$  and  $b^\lambda$ , the circuit output,  $\varphi$ , falls into the vicinity of  $\varphi_0$ .

Fortunately, there is an easy fix to the problem of false classification in our system. As shown in Fig. 11, the solution is to find multiple, independent training templates,  $b_{(1)}^\xi, b_{(2)}^\xi, \dots$ , for different target values of  $\varphi_0$ , say  $\varphi_0^{(1)}, \varphi_0^{(2)}, \dots$  and require passing several “recognize/reject” tests using these templates before an object is positively classified as  $\xi$ .

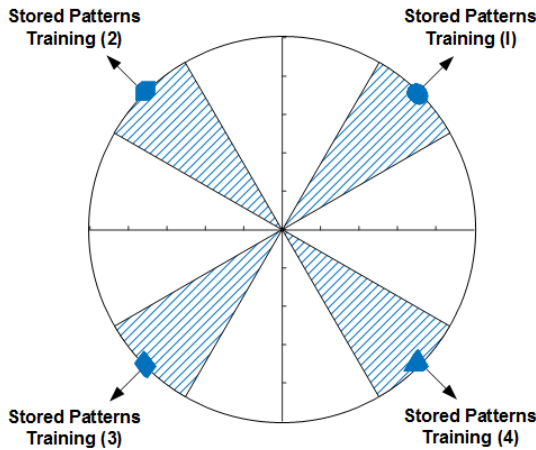


FIGURE 11. Developing several training templates using different target angles.

Numerical simulations confirm (see Fig. 12) that chances of false classification decrease rapidly (the fitted curve is an exponential) as more training templates are developed and consequently more “recognize/reject” tests are used for any positive classification.

As a final note, we investigate the storage capacity of a discriminant circuit, which can be defined as the maximum number of samples of a single object for which the network can be successfully trained (see Fig. 10). In a realistic application, different sample images used in training the network for identification of a particular object may be highly correlated with each other, which should make it easier to find a single training template that gives the desired circuit output for all of them. In quantifying the storage capacity of our network,

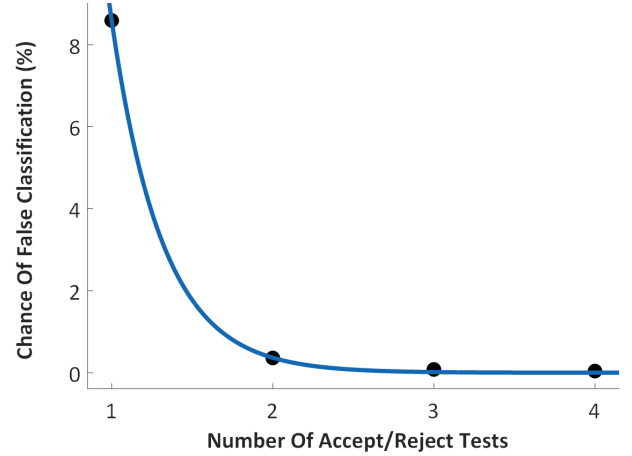


FIGURE 12. Chance of false classification versus number of recognize/reject tests.

however, we assume a more pessimistic situation where sample images used for training are totally uncorrelated to see if the network can still be successfully trained. We leave the details of the training algorithm that is written as an MATLAB code for the interested reader in the Supplementary Material. Also a complete version of the MATLAB code used for quantifying the network’s storage capacity is attached to this paper. The result of these numerical experiments is shown in Fig. 13 and can be summarized as follows.

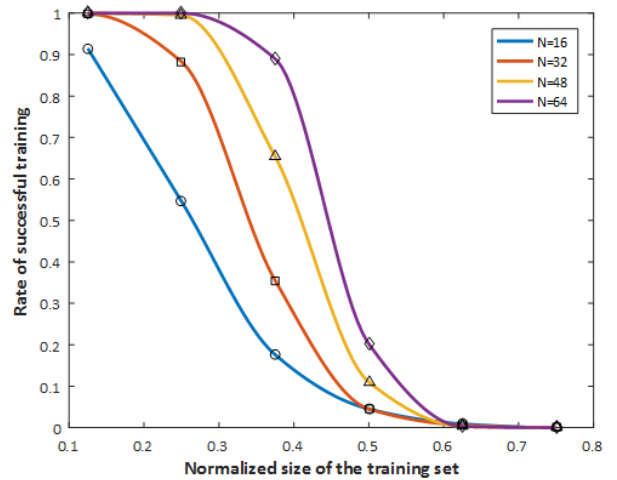


FIGURE 13. Rate of successful training versus relative size of training set.

For each fixed size of the network,  $N$ , an increasing number of uncorrelated patterns,  $m$ , are allowed for training. The  $x$ -axis shows the ratio  $x = m/N$ . The  $y$ -axis is the “success” rate in training the network for different values of  $N$  and  $m$ . The target value of  $\varphi_0$  is arbitrarily set to  $\pi/4$  in these experiments and an experiment is called successful if for all the  $\xi^{(1)} \dots \xi^{(m)}$  patterns, we have  $|\varphi - \varphi_0| \leq \pi/8$ . A total of 1000 numerical experiments were carried on for each pair of  $m$  and  $N$ .

Interestingly, the MATLAB simulation results indicate that the reliable storage capacity of network, which can be defined as the maximum value of  $x$  for which the training success rate is close to 1, increases with the network size and is more than 0.25 for  $N = 64$ .

#### IV. CONCLUSION

In conclusion, in this paper, we introduced the concept of “discriminant circuit,” proposed a novel way of controlling relative phase shifts between CMOS ring oscillators with arbitrary resolution, and showed how a simply connected chain of coupled oscillators can be employed as an ultrafast pattern classifier engine. The fast processing is due to the fact that the settling time for the coupled oscillator chain, i.e., for the link phase shifts to reach their final values, is typically on the order of a few hundreds oscillation periods. An important characteristic of this paper is the simplicity of the circuit architecture, which relies only on local couplings between adjacent oscillators and is readily implementable on any standard CMOS technology. As for future research directions, a prototype version of the circuit should be fabricated and tested. The circuit should then be used in conjunction with other standard pattern recognition techniques, like preprocessing and feature extraction, to be applied to more realistic pattern recognition examples. Also, while the focus in this paper was on coupled oscillators, a discriminant circuit can be built with other physical components that would result in even simpler and more compact designs. An example worth considering would be a tunable  $RC$  network with  $R$  and  $C$  values acting as “input” and “training” parameters and the circuit’s cutoff frequency as the discriminant circuit’s output.

#### REFERENCES

- [1] P. Maffezzoni, B. Bahr, Z. Zhang, and L. Daniel, “Oscillator array models for associative memory and pattern recognition,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 6, pp. 1591–1598, Jun. 2015.
- [2] D. E. Nikonov, “Coupled-oscillator associative memory array operation for pattern recognition,” *IEEE J. Exploratory Solid-State Comput. Devices Circuits*, vol. 1, pp. 85–93, Nov. 2015.
- [3] T. Shibata, R. Zhang, S. P. Levitan, D. E. Nikonov, and G. I. Bourianoff, “CMOS supporting circuitries for nano-oscillator-based associative memories,” in *Proc. 13th Int. Workshop Cellular Nanoscale Netw. Appl. (CNNA)*, Aug. 2012, pp. 1–5.
- [4] S. P. Levitan, Y. Fang, D. H. Dash, T. Shibata, D. E. Nikonov, and G. I. Bourianoff, “Non-Boolean associative architectures based on nano-oscillators,” in *Proc. 13th Int. Workshop Cellular Nanoscale Netw. Appl. (CNNA)*, Aug. 2012, pp. 1–5.
- [5] M. J. Cotter, Y. Fang, S. P. Levitan, D. M. Chiarulli, and V. Narayanan, “Computational architectures based on coupled oscillators,” in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI (ISVLSI)*, Jul. 2014, pp. 130–135.
- [6] Y. Fang, C. N. Gnegy, T. Shibata, D. Dash, D. M. Chiarulli, and S. P. Levitan, “Non-Boolean associative processing: Circuits, system architecture, and algorithms,” *IEEE J. Exploratory Solid-State Comput. Devices Circuits*, vol. 1, pp. 94–102, Dec. 2015.
- [7] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *Proc. Nat. Acad. Sci. USA*, vol. 79, no. 8, pp. 2554–2558, 1982.
- [8] R. Rojas, *Neural Networks: A Systematic Introduction*. Berlin, Germany: Springer-Verlag, 1996.
- [9] J. Hertz, R. G. Palmer, and A. Krogh, *Introduction to the Theory of Neural Computation*. Redwood City, CA, USA: Addison-Wesley, 1991.
- [10] T. Aoyagi, “Network of neural oscillators for retrieving phase information,” *Phys. Rev. Lett.*, vol. 74, no. 20, p. 4075, 1995.
- [11] F. C. Hoppensteadt and E. M. Izhikevich, “Oscillatory neurocomputers with dynamic connectivity,” *Phys. Rev. Lett.*, vol. 82, no. 14, p. 2983, 1999.
- [12] P. Maffezzoni, B. Bahr, Z. Zhang, and L. Daniel, “Analysis and design of boolean associative memories made of resonant oscillator arrays,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 11, pp. 1964–1973, Nov. 2016.
- [13] F. C. Hoppensteadt and E. M. Izhikevich, “Pattern recognition via synchronization in phase-locked loop neural networks,” *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 734–738, May 2000.
- [14] R. W. Hölzel and K. Krischer, “Pattern recognition with simple oscillating circuits,” *New J. Phys.*, vol. 13, p. 073031, Jul. 2011.
- [15] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*. Berlin, Germany: Springer, 1984.
- [16] G. B. Ermentrout and D. H. Terman, *Mathematical Foundations of Neuroscience*, vol. 35. New York NY, USA: Springer, 2010.
- [17] E. M. Izhikevich, *Dynamical Systems in Neuroscience*. Cambridge, MA, USA: MIT Press, 2007.
- [18] C. M. Bishop, *Neural Networks for Pattern Recognition*. London, U.K.: Oxford Univ. Press, 1995.
- [19] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
- [20] T. Kohonen, *Self-Organization and Associative Memory*, vol. 8. New York NY, USA: Springer, 2012.
- [21] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications in Physics, Biology, Chemistry, and Engineering*. Reading, MA, USA: Addison-Wesley, 2014.
- [22] B. Razavi, *RF Microelectronics*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2012.
- [23] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [24] I. A. Young, J. K. Greason, and K. L. Wong, “A PLL clock generator with 5 to 110 MHz of lock range for microprocessors,” *IEEE J. Solid-State Circuits*, vol. 27, no. 11, pp. 1599–1607, Nov. 1992.

**VAHNOOD POURAHMAD** received the B.S degree in civil engineering from Sharif University of Technology, Tehran, Iran, in 2008 and M.S. degrees in mechanical engineering and electrical engineering from Cornell University, Ithaca, NY, USA, in 2012 and 2016, respectively. He is currently working toward his Ph.D. degree in electrical and computer engineering from Cornell University.

His current research interest is in the study of electrical coupled oscillators and their various practical applications in engineering, specifically in designing non-Boolean/associative processors and phased array systems.

**SASIKANTH MANIPATRUNI** (M’07) received the B.S. degree in electrical engineering from IIT Delhi, New Delhi, India, and the Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA.

He is currently a Staff Scientist with the Exploratory Integrated Circuits Group, Components Research, Intel, Santa Clara, CA, USA, where he is involved in beyond-CMOS devices and circuits. He was a co-inventor of several building blocks for silicon photonic devices, such as modulators, switches, and WDM interconnects. He has authored 50 journal/conference articles and has over 2000 citations in international journals. He has over 35 issued and pending patents in the fields of integrated photonics, spin devices, and magnetic resonance imaging.

**DMITRI NIKONOV** (M’99–SM’06) received the M.S. degree in aeromechanical engineering from the Moscow Institute of Physics and Technology, Moscow, Russia, in 1992, and the Ph.D. degree in physics from Texas A&M University at College Station, College Station, TX, USA, in 1996.

He joined Intel, Santa Clara, CA, USA, in 1998. He is currently a Principal Engineer with the Components Research Group, Hillsboro, OR, USA, doing simulation and benchmarking of beyond-CMOS logic devices, and managing research programs with universities on nanotechnology.



**IAN YOUNG** (M'78–SM'96–F'99) received the B.E.E. and M.Eng.Sc. degrees from the University of Melbourne, Melbourne, VIC, Australia, and the Ph.D. degree in electrical engineering from the University of California at Berkeley, Berkeley, CA, USA.

He is currently a Senior Fellow and the Director of the Exploratory Integrated Circuits with the Technology and Manufacturing Group, Intel Corporation, Hillsboro, OR, USA. He leads a research group exploring the future options for the integrated circuit in the beyond-CMOS era.

Dr. Young is currently the Editor-in-Chief of the IEEE Journal of Exploratory Solid-State Computational Devices and Circuits.

**EHSAN AFSHARI** (S'98–M'07–SM'11) received the B.Sc. degree in electronics engineering from the Sharif University of Technology, Tehran, Iran, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 2003 and 2006, respectively.

In 2006, he joined the Faculty of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA, as an Assistant Professor and was promoted to an Associate Professor in 2012. In 2016, he joined the Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI, USA, as an Associate Professor. His current research interests are mm-wave and terahertz electronics and low-noise integrated circuits for applications in communication systems, sensing, and biomedical devices.