<table>
<thead>
<tr>
<th>Title</th>
<th>Perceptrons with Hebbian learning based on wave ensembles in spatially patterned potentials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Espinosa-Ortega, Tania; Liew, Timothy Chi Hin</td>
</tr>
<tr>
<td>Date</td>
<td>2015</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10220/25402">http://hdl.handle.net/10220/25402</a></td>
</tr>
<tr>
<td>Rights</td>
<td>© 2015 American Physical Society. This paper was published in Physical Review Letters and is made available as an electronic reprint (preprint) with permission of American Physical Society. The paper can be found at the following official DOI: [<a href="http://dx.doi.org/10.1103/PhysRevLett.114.118101">http://dx.doi.org/10.1103/PhysRevLett.114.118101</a>]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.</td>
</tr>
</tbody>
</table>
Perceptrons with Hebbian Learning Based on Wave Ensembles in Spatially Patterned Potentials

T. Espinosa-Ortega and T. C. H. Liew
Division of Physics and Applied Physics, Nanyang Technological University, Singapore 637371, Singapore
(Received 31 July 2014; published 18 March 2015)

A general scheme to realize a perceptron for hardware neural networks is presented, where multiple interconnections are achieved by a superposition of Schrödinger waves. Spatially patterned potentials process information by coupling different points of reciprocal space. The necessary potential shape is obtained from the Hebbian learning rule, either through exact calculation or construction from a superposition of known optical inputs. This allows implementation in a wide range of compact optical systems, including (1) any nonlinear optical system, (2) optical systems patterned by optical lithography, and (3) exciton-polariton systems with phonon or nuclear spin interactions.

DOI: 10.1103/PhysRevLett.114.118101 PACS numbers: 42.30.Sy, 07.05.Mh, 71.36.+c, 84.35.+i

Neural networks exploit massive interconnectivity to become highly efficient at certain tasks, such as classification, and pattern recognition [1,2]. While biological neurons may operate individually on millisecond time scales, their simultaneous connection to several thousands of other neurons allows a parallelization of tasks far beyond the capabilities of complementary metal-oxide-semiconductor (CMOS) logic. Naturally, this observation has motivated research into artificial neural networks, including hardware implementations [3]. Electrically connected systems have been based on inorganic synapses [4], spiking silicon circuits [5,6], feedback on Bose-Einstein condensates [7], memristors [8], or spin torque devices [9] such as nanomagnetic spin switches [10]. Optical systems can exploit vector matrix multiplication [11] or fractional Fourier transforms [12].

In models of neuron behaviour or in artificial neural networks, an individual neuron gives an output given by

$$g_{ij}^{out} = f\left(\sum_j w_{ij} g_{in}^{in}\right)$$

(1)

where $f$ is some function depending on the sum of inputs from several input neurons, $g_{in}^{in}$, each multiplied with a different weight $w_{ij}$. The weights represent the “knowledge” in the system and are adjusted in a “learning” process or set externally for the desired function.

Given Eq. (1), an artificial neural network requires three ingredients: (a) A large number of interconnections between different neurons, (b) The possibility of different weights for each interconnection, (c) Flexibility in the choice of weights to allow learning. A two-layer network with one set of inputs and one layer of output neurons is already a powerful unit, known as a perceptron, capable of the linear classification of data and pattern recognition. However, a small device with 35 inputs and 10 outputs already requires 350 independent weight connections.

In electrical devices the engineering of so many connections is a serious task. Optical designs can benefit from the overlapping of several different light rays to build these connections; however, the controlled weighting of connections often requires separate electronic connections or bulky spatial light modulators.

Here a simple scheme of a perceptron is demonstrated, based on the superposition of an ensemble of waves. A model based on the Schrödinger equation illustrates the wide applicability of the scheme, which is compatible with a range of micron-sized solid-state implementations where a spatially varying potential is available. Examples include (1) nonlinear optical systems, (2) systems microstructured by optical lithography, and (3) exciton-polariton systems with (i) acoustic phonon interactions or (ii) nuclear spin interactions. In general, the weights in the system can be calculated using the Hebbian learning rule or constructed via a superposition of optical waves. In examples 1 and 2, the weights are then written with a fixed potential. In example 3, the change (plasticity) of the effective potential, under a given set of inputs and outputs, allows a direct demonstration of Hebbian learning in which the system adapts automatically to perform the desired network function. The advantage of the proposed general scheme is that it is not necessary to construct or control each weight individually; it is only necessary to provide the appropriate stimulus as a combination of known input and output waves. Both permanent weightings for repetitive tasks as well as reconfigurable networks are possible. For illustration, the task of pattern recognition is considered. Operation on ultrafast (picosecond) time scales is expected.

General scheme.—Let us consider the 2D Schrödinger equation for particles with a wave function $\psi(x)$ moving in a spatially varying potential $V(x)$:

$$i\hbar \frac{\partial \psi}{\partial t} = \left(\hat{E} - i\Gamma + V(x)\right)\psi + F(x)$$

(2)
where $\dot{E}$ represents the kinetic energy and we allowed for driving ($F$) and decay terms ($\Gamma$), corresponding to a nonequilibrium system. The kinetic energy is measured from the energy associated with the driving field energy. Let us imagine a potential of the form

$$V(x) = \sum_{ij} V_{ij} \cos \left( (|k| - |j|) \cdot x \right)$$  \hspace{1cm} (3)

where the spatially dependent part of the wave function and driving terms can be decomposed into Fourier components in reciprocal space:

$$\psi(x) = \frac{1}{\sqrt{S}} \sum_i \sqrt[2]{\psi_i e^{i|k| \cdot x}}$$ \hspace{1cm} (4)

$$F(x) = \frac{1}{\sqrt{S}} \sum_i F_i e^{i|k| \cdot x}$$ \hspace{1cm} (5)

where $S$ is the area of the system and we restrict the driving field to wave vectors of the same magnitude, such that all states in the system have the same energy (which is conserved). Substituting Eqs. (3)–(5) into Eq. (2) and collecting terms oscillating at the same wave vector, we obtain

$$i\hbar \frac{\partial \psi_j}{\partial t} = (E_0 - i\Gamma) \psi_j + \sum_k V_{jk} \psi_k + F_j$$ \hspace{1cm} (6)

where $E_0$ and $\Gamma$ are the energy and decay rates at $|k|_i$ (now the same for all particles). We divide reciprocal space into a half containing driven wave vectors, which will characterize a vector of inputs in the system, $F_1, F_2, \text{etc.}$, and a half for the output. During training the amplitudes $F_j \in (0, F_0)$ can take on one of two values each, representing a binary input [see illustration in Fig. 1(a)].

Assuming that $V_{ij}$ is small, the steady state values in the input half of reciprocal space are

$$\psi_j = \frac{F_j}{E_0 - i\Gamma}$$ \hspace{1cm} (7)

States in the output half of reciprocal space are reached via scattering with $V_{ij}$, giving the amplitudes

$$\psi_j = \frac{\sum_i V_{ij} \psi_i}{E_0 - i\Gamma} = \frac{\sum_i V_{ij} F_j}{(E_0 - i\Gamma)^2}.$$ \hspace{1cm} (8)

In other words, the output intensities, $|\psi_j|^2$, are a function of $\sum_j V_{ij} F_j$, where $F_j$ are the inputs and $V_{ij}$ are the weight connections, as required for a neural network.

**Hebbian learning.**—For a network to be useful, the weights $V_{ij}$ must be chosen to give the desired network function. Under Hebbian learning, neuron connection weights are increased when input and output neurons fire simultaneously. This allows training of the network where an input vector is applied and the desired output state is simultaneously activated. Repeating the process over a “training set” of input vectors allows the system to learn how to distinguish different inputs, its ability being encoded in neuron connection weights of the form

$$V_{ij} \propto \sum_r F_i^{(r)} F_j^{(r)}.$$ \hspace{1cm} (9)

where we sum over different vectors in the training set, which are labeled by the index $r$. Here the index $i$ represents vectors in the input half of reciprocal space. To activate states in the output half of reciprocal space, we apply the driving field represented by $F_j^{(r)}$.

Note that the driving field intensity for each training step contains terms given by

$$|F_i^{(r)} e^{i|k|_i \cdot x} + F_j^{(r)} e^{i|k|_j \cdot x}|^2 = 2 F_i^{(r)} F_j^{(r)} \cos[(|k|_i - |k|_j) \cdot x] + \text{const.}$$ \hspace{1cm} (10)

Consequently, we seek a mechanism of varying the potential $V(x)$ in proportion to the driving field intensity during the training phase. This gives a field with the same form as in Eq. (3) with weights $V_{ij}$ following the Hebbian learning rule [Eq. (9)]. The contribution of the constant in Eq. (10) would also give a constant shift of the potential; however, this can be fully compensated by varying the energy of all the particles in the system (determined by the driving field frequency).

As an illustration of Hebbian learning, we demonstrate the task of pattern recognition, where the potential is constructed from a superposition of known inputs and outputs from a training set of numerical characters ranging from 0–9. Figure 2 shows that input characters (top row) are scattered to particular output wave vectors. In this way the system recognizes the characters (bottom row), even if the inputs contain multiple errors.
A key ingredient of the scheme is that the required potential has a spatial pattern proportional to the intensity of a known optical field. In the rest of the Letter we consider different possible physical realizations of the potential. All cases follow the general recipe outlined above and allow direct reproduction of Fig. 2.

**Nonlinear optical systems.**—A wide variety of nonlinear optical systems are described by the nonlinear Schrödinger equation, identical to Eq. (2) with an additional term $a|\psi|^2\psi$. If we imagine that the system can be excited by two different frequencies, $\omega_1$ and $\omega_2$, giving field components $\psi_1$ and $\psi_2$, then we can use the frequency $\omega_1$ for creating a potential of the form $a|\psi_1|^2$ which is experienced by $\psi_2$. Each frequency component should again be decomposed in reciprocal space and the potential can be written either by direct calculation of the necessary optical field or by cycling over the input and output superpositions in the training set. The advantage of an optically induced potential [13] is that it can be changed at will, allowing rapid reconfiguration of the network function. Ideally, a system has multiple resonant modes to allow the efficient injection at different frequencies. Alternatively, one could make use of excitation of two orthogonal polarizations, provided they all have the same energy; $q = (q_1, q_2)$. The explicit form of the exciton-phonon scattering amplitude $G_q$ is written in Ref. [21].

The phonon scattering rate can be calculated using the Fermi golden rule,

$$W_{ij,q} = \frac{2\pi}{\hbar} \left| \langle k_j | n_{q_i,q_j} + 1 | H_p | n_{q_i,q_j} | k_i \rangle \right|^2 \delta E_{k_j,q_i} \delta E_{k_i}$$

$$= \frac{2\pi}{\hbar} |XG_{ij,q}^s| \mu_{ij} \bar{n}_{ij}^{\text{pol}} (n_{ij}^{\text{pol}} + 1) (n_{ij}^{\text{pol}} + 1),$$

where $q_i$ is the phonon density of states and $n_{ij}^{\text{pol}}, n_{ij}^{\text{pol}}$ stand for the polariton and phonon densities respectively. We integrate the expression $2\pi |XG_{ij,q}^s|^2 \bar{n}_{ij}^{\text{pol}}$ over $d_\zeta$ (from $-2\pi/L_z$ to $2\pi/L_z$, with $L_z$, quantum well width) to obtain the in-plane component of the scattering rate $w_{ij} \approx w(q \to 0) \equiv w$. The dynamics of the phonon density is then given by

$$\hbar \frac{\partial n_{ij}^{\text{ph}}}{\partial t} = w_{ij} n_{ij}^{\text{pol}} + 1) (n_{ij}^{\text{pol}} + 1) - \Gamma \bar{n}_{ij}^{\text{pol}}.$$ 

In analogy to Eq. (7), the polariton densities for the driven signals follow $n_{ij}^{\text{pol}} = |F|^2 / (E_0^2 + \Gamma^2)$. This allows straightforward calculation of the phonon densities obtained during the training phase [Fig. 3(b)]. The optical driving field stimulates the excitation of phonons with specific wave vectors. In GaAs based systems the phonon lifetime has...
been measured in the range of 100 ns [22] for low frequencies, which can be arranged by choosing small wave vectors, $|k_i| = 0.5 \mu m^{-1}$, in reciprocal space. Consequently, the phonons leave a long-lasting mark through the polariton-phonon interactions, which will allow the system to “memorize” the digit.

During the operation phase, the output signals obey

$$\hbar \frac{\partial n_{j}^{\text{pol}}}{\partial t} = \sum_i \omega n_{j}^{\text{pol}}(n_{ij}^{\text{ph}} + 1)(n_{j}^{\text{pol}} + 1) - \Gamma n_{j}^{\text{pol}}. \tag{14}$$

For a weak polariton density during operation and $\Gamma \gg \Gamma_x$, the polariton output density is given by

$$n_{j}^{\text{pol}} = \frac{\sum_i \omega n_{j}^{\text{pol}}(n_{ij}^{\text{ph}} + 1)}{\sum_i \omega n_{j}^{\text{pol}}(n_{ij}^{\text{ph}} + 1) - \Gamma (e^{\sum_i \omega n_{i}^{\text{pol}}(n_{ij}^{\text{ph}} + 1) - \Gamma}) - 1). \tag{15}$$

The analogy with Eq. (8) is most visible when $\Gamma$ is large. We verified that the obtained outputs allow the reproduction of Fig. 2 for both GaAs and GaN based systems. Testing shorter phonon lifetimes revealed that $\Gamma_x$ is not a critical parameter for obtaining the correct outputs. $\Gamma_x$ is however important for determining the memory time of the system, which is on the one hand longer than the polariton lifetime to allow multiple repetitions and short enough to allow rapid reconfiguration of the device.

Within our theory we have neglected polariton-polariton scattering, which can in principle cause the redistribution of polaritons between modes on the circle in reciprocal space [23]. During training, this would have little effect as the polaritons between modes on the circle in reciprocal space scattering, which can in principle cause the redistribution of polariton density of the driving field is circularly polarized and the laser density of dominant fields would still be those directly driven by the polaritons [23]. During training, this would have little effect as the polaritons between modes on the circle in reciprocal space scattering, which can in principle cause the redistribution of polariton density of dominant fields would still be those directly driven by the polaritons. The potential that the nuclear spin polarization induces on the polaritons is given by

$$\langle I_z(x) \rangle = \frac{\Gamma_{\text{hf}} Q S_z (1 - e^{-t/\tau})}{\Gamma_{\text{hf}} + \tau^{-1}}. \tag{18}$$

For small quantum wells ($L_z \sim 10 \text{ nm}$) and electron density around $10^{12} \text{ cm}^{-2}$ we have that $\Gamma_{\text{hf}} \ll \tau^{-1}$, therefore, we can approximate the above expression as

$$\langle I_z(x) \rangle \approx \frac{\Gamma_{\text{hf}} Q S_z (1 - e^{-t/\tau})}{\Gamma_{\text{hf}} + \tau^{-1}}. \tag{19}$$

The potential that the nuclear spin polarization induces on the polaritons is given by

$$V_{ij} = \frac{8 \pi \chi A^2 Q S n_e \nu_0^2 (I + 1)}{3 \hbar^2 V L c E_p} F_i F_j \tag{19}$$

with $E_p = E_c(|k|)$, the energy of the cavity mode at the pump wave vector. The nuclear spin induced potential is shown in Fig. 3(a), for different exposure times and electron densities.

FIG. 3 (color online). (a) Nuclear spin induced potential, given by Eq. (19). (b) Phonon-induced weight $\omega_{n_{ij}}$ for a GaAs based system [34]. The plots show the variation with the duration of application of input $k_i$ and output $k_j$, as well as the variation with electron density (a) and polariton density (b).
For the considered densities, the nuclear spin induced potentials correspond to nuclear spin polarizations of around 1%. While nuclear spin polarizations up to 3.3% and 10% were observed in quantum wells [35] and quantum dots [36], respectively, and higher electron densities were considered in the literature [37], the small nuclear spin induced potential is in fact sufficient for our task. If one compares to the strength of disorder, which is around 10 μeV in state-of-the-art samples, then the nuclear spin induced potential may seem insignificant. However, for the determination of the scattered polariton intensity one must examine the Fourier components of the potentials. In reciprocal space, the nuclear spin induced potentials are localized at specific points while a random disorder potential is spread widely. In fact, if one compares the Fourier amplitude $V_{ij}$ to the typical Fourier amplitude $V_{\text{disorder}}$ of a Gaussian correlated disorder potential, one finds $V_{ij}/V_{\text{disorder}} = \sqrt{\sigma^2 \left(k_i - k_j\right)^2}^{1/2}/(\sqrt{\sigma V_{\text{rms}}})$, where $\sigma$ is the system area, $\sigma$ the disorder correlation length, and $V_{\text{rms}}$ is the root mean squared amplitude of disorder. After all training steps are completed $V_{ij} \sim 2 \mu$eV. Typically, microcavities have a larger disorder than in Ref. [38], characterized by a height of 0.1 meV or root mean squared variation $V_{\text{rms}} \approx 10 \mu$eV and correlation length $\sigma \approx 1 \mu$m. Taking these values, $k_i - k_j = 1 \mu$m$^{-1}$, and a wide excitation area $S = 200 \mu$m$^2$, one finds that the ratio of scattered polariton intensity in reciprocal space by the nuclear spin induced potential and disorder potential is $n_j/n_{\text{disorder}} = (V_{ij}/V_{\text{disorder}})^2 \sim 1600$. Consequently, the nuclear spin induced potential has a much stronger effect on polariton scattering than disorder, and would be clearly visible experimentally on picosecond time scales [25].

Substituting Eq. (19) into Eq. (8), we verified that Fig. 2 was reproducible for typical parameters [34].

Conclusion.—We developed a general scheme to construct perceptrons using wave ensembles, appropriate to a wide range of optical systems. Unlike other hardware implementations of neural networks, our general scheme avoids the need to store and control the weight of each network connection, $w_{ij}$, in an independent feedback loop. A large number of weights are encoded in the form of the spatially structured potential, which is constructed from exposure to a known set of training conditions.

We considered three classes of example systems: (1) nonlinear optical systems, (2) systems patterned by optical lithography, and (3) exciton-polariton systems. In nonlinear optical systems, the optically induced potential can be rapidly reconfigured, while optical lithography allows the engineering of a permanent network for a repetitive task. The exciton-polariton systems allow network weights to be automatically chosen by phonons or nuclear spins inside the system, allowing the system to learn in response to a given set of stimuli. Because of the wide availability of suitable systems, we believe this approach will lead to ultrafast and compact hardware implementations of neural networks as well as a new platform for machine learning.

We thank I. A. Shelykh for discussion and acknowledge support from the NTU Grant No. M4080853.110.706022. T. L. was funded by the Lee Kuan Yew Endowment Fund.

[34] \[ m_s = 0.069 m_0, \quad m_b = 0.18 m_0, \quad \Gamma = 0.16 \text{meV} \quad \text{and} \quad \Gamma_s = 3.3 \text{meV}, \quad E_0 = 0, \quad \rho = 5.3 \text{g/cm}^3, \quad u = 5.37 \times 10^3 \text{m/s}, \quad a_j = 5.8 \text{ nm,} \quad a_e = -7 \text{eV,} \quad a_h = 2.7 \text{ eV} \quad \nu_0 = (5.65 \text{ Å})^{1/2}, \quad A = 90 \text{ µeV}, \quad Q = 5, \quad S_z = 1, \quad \tau = 10 \text{ ms,} \quad w = 15 \text{ neV}. \]