Perceptrons with Hebbian learning based on wave ensembles in plastic potentials

T. Espinosa-Ortega$^1$ and T. C. H. Liew$^1$

$^1$Division of Physics and Applied Physics, Nanyang Technological University 637371, Singapore

(Dated: September 1, 2014)

We present a general theoretical model to realize a bilayer perceptron for hardware neural networks with applications in pattern recognition. In the network, multiple interconnections are allowed, by using the Schrödinger wave function as input and outputs signals; moreover, microscopic plastic potentials allow to process information and “train” the system in micrometer’s scale. As particular cases, we present the calculations for two devices where the information is carried by light and the plastic potentials emerge due polariton-phonon and polariton-nuclear spin interactions. We show both designs are capable to perform digit recognition.

PACS numbers:

Neural networks exploit massive interconnectivity to become highly efficient at certain tasks, such as classification, and graphic [1] or acoustic [2] pattern recognition. For example, biological neurons may operate individually on millisecond timescales yet their simultaneous connection to several thousands of other neurons allows a parallelization of tasks far beyond the capabilities of CMOS logic. Naturally, this observation has motivated research into artificial neural networks including inorganic synapses [3, 4] and hardware implementations [5]. Electrically connected systems have been based on memristors [6] or spin torque devices [7] such as nanomagnetic spin switches [8], while optical systems can exploit vector matrix multiplication [9], fractional fourier transform [10] and might be implemented using photonic [11] or photorefractive [12] crystals.

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

$$\xi_{out}^i = f \left( \sum_j w_{ij} \xi_{in}^j + b_i \right)$$

(1)

where $f$ is some function depending on the sum of inputs from several input neurons, $\xi_{in}^i$, each multiplied with a different weight $w_{ij}$. The optional constant $b_j$ is known as the bias, which may be different for different neurons. The weights represent the “knowledge” in the system and they are typically adjusted in a “learning” process to achieve the desired network function.

Given Eq.1 an artificial neural network requires three ingredients: 1) A large number of interconnections between different neurons; 2) The possibility of different weights for each interconnection; 3) 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, a layer of electronic connections to control the weights is still needed.

We present a general microscopic model to design a perceptron with Hebbian learning, where the input and output signals are given by the Schrödinger wave function and the weights of the system can be encoded in the inputs. The two layers of the system are substitute by dividing the $k$ space into two; one half for the inputs and another for the outputs; using a plastic potential, the weights of the system can be adjust in the same layer where the inputs are sent, therefore the traditional system of 3 layers [input-weights-output], can be reduced to a single unique layer where the plastic potential acts on the wave functions. We discuss two particular cases where the described architecture might be implemented and show, that the proposed devices can perform pattern recognition in a micrometer scale, after a performing Hebbiang learning during microseconds.

Scattering in Reciprocal Space.— Let us consider the Schrödinger equation for particles with a wavefunction $\psi(x)$ moving in a spatially varying potential $V(x)$:

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

(2)

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

$$V(x) = \sum_{ij} V_{ij} \cos ((k_i - k_j) \cdot x)$$

(3)

Where the spatial dependent part of the wavefunction and driving terms can be decomposed into Fourier components in reciprocal space:

$$\psi(x) = \frac{1}{\sqrt{S}} \sum_i \psi_i e^{i k_i \cdot x}$$

(4)

$$F(x) = \frac{1}{\sqrt{S}} \sum_i F_i e^{i k_i \cdot x}$$

(5)
Let us restrict the driving to wavevectors of the same magnitude, such that all states in the system have the same energy. Substituting Eqs. 3, 4 into Eq. 2 and collecting terms oscillating at the same wavevector, we obtain:

\[ i\hbar \frac{\partial \psi_i}{\partial t} = (E_0 - i\Gamma_0) \psi_i + \sum_j V_{ij} \psi_j + F_i \]  

(6)

where \(E_0\) and \(\Gamma_0\) are the energy and decay rates at \(|k_i|\) (now the same for all particles). Let us divide reciprocal space into a half containing driven wavevectors, which will characterize a 35 dimension vector of inputs in the system, \(F_1, F_2\), etc. and half for a 10 dimension output vector. The amplitudes \(F_i \in (0, F_0)\) can take one of two values each, representing a binary input.

The 35 input set will map into a 5 \times 10 array where every \(|\psi_i|^2\) represents a pixel of a number digit. Correspondingly each element of the output vector identifies an input set [see Fig.4].

Assuming that \(V_{ij}\) can be treated as a perturbation, the steady state values in the input half of reciprocal space are:

\[ \psi_i = \frac{-F_i}{E_0 - i\Gamma} \]  

(7)

States in the output half of reciprocal space develop amplitudes:

\[ \psi_i = \frac{\sum_j V_{ij} \psi_j}{E_0 - i\Gamma} = \frac{\sum_j V_{ij} F_j}{(E_0 - i\Gamma)^2} \]  

(8)

In other words the output intensities, \(|\psi_i|^2\), are a function of \(\sum_j V_{ij} F_j\), where \(F_j\) are the inputs and \(V_{ij}\) are the weight connections, in analogy to 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 set is applied and the desired output state is simultaneously activated.

See for instance Fig.4, here the corresponding wave function of the input vectors \(k_i\) map into “number one”, simultaneously, the desire output \(\kappa_1\) is activated and the plastic potential \(V(x)\), is modified. Repeating the process over a “training set” of input vectors allows the system to learn how to distinguish different inputs. The result is that the neuron connection weights should adopt the form:

\[ V_{ij} \propto \sum_{\{v\}} F_i^{(v)} F_j^{(v)} \]  

(9)

where we sum over different vectors in the training set, which are labelled by the index \(v\). 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 \(F_i^{(v)}\).

Note that the driving field intensity \(F_i^{(v)}\) for each training step is given by:

\[ \left| F_i^{(v)} e^{i\kappa_i \cdot x} + F_j^{(v)} e^{i\kappa_j \cdot x} \right|^2 = 2F_i^{(v)} F_j^{(v)} \cos((\kappa_i - \kappa_j) \cdot x) + \text{const.} \]  

(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).

FIG. 1: We show the input set of signals \(k_i\) as blue rays, with \(|\psi_i|^2\) mapped into blue pixels and the output signals \(\kappa_1\), represented with a red rays, with \(|\psi_{\kappa_1}|^2\) mapped into red pixels. a) Training phase. The input set that corresponds to number one, is activated simultaneously with the desire output; the activated output \(\kappa_1\) will identify this number. b) Operation phase. The input set represents a digit with defects, there is no driven field over the output, nevertheless the maximum output corresponds to \(|\psi_{\kappa_1}|^2\) and the number can be recognized.

Semiconductor Microcavities. — Exciton-Polaritons are particles that appear in semiconductor microcavities due the strong coupling of quantum well excitons with cavity photons [14]; their injection into the system can be controlled via optical excitation and are good candidates for operating with plastic, variable, potentials. For example, the injection of a large number of polaritons causes a blueshift of the polariton potential, due to repulsive polariton-polariton interactions. This allows optical engineering of the potential, which is indeed shifted in proportion to the intensity of an applied optical field [13]. However, due to the short lifetime of polaritons such deformation of the potential is short-lived and the system would quickly forget anything learnt from a Hebbian training process. Other possible excitations in the system include acoustic phonons and nuclear spin polarizations. These are much longer lived than the typical picosecond scale of polariton dynamics.
Acoustic Phonons. — The interaction between polaritons and acoustic phonons is described by the Fröhlich Hamiltonian [12, 16]:

\[ \mathcal{H}_{p} = X \sum_{q, k} \left( G_q \hat{a}_k \hat{b}_q + G^*_q \hat{b}_k^\dagger \hat{a}_q + G^*_q \hat{b}_k \hat{a}_q + G_q \hat{b}_q^\dagger \hat{a}_k \right) \]  

(11)

where \( \hat{a}_k \) and \( \hat{b}_q \) are polariton and field operators in reciprocal space, respectively; \( X \) is the excitonic fraction (Hopfield coefficient), which is the same for all polaritons given that they all have the same magnitude of \( k \), [wave vector in-plane the quantum well]; \( G_q \) is the exciton-phonon scattering amplitude, \( q = (q_x, q_z) \), and

\[ G_{q,q_z} = i \sqrt{\frac{\hbar(q^2_1 + q^2_2)^{1/2}}{2 \rho V u}} \left[ a_e (1 + \frac{m_e}{2M} |q| a_B)^{-3/2} \right. \\
- a_h (1 + \frac{m_h}{2M} |q| a_B)^{-3/2} \left. I^\perp(q_z) \right] \]  

(12)

where \( V \) is the quantized volume, \( M \) is the exciton mass, \( \rho \) and \( u \) are the density and the longitudinal sound velocity, \( a_B \) is the exciton Bohr radius and \( a_c, a_h \) corresponds to the deformation potentials of electron and hole respectively. The function \( I^\perp(q_z) \) is equal to the superposition of the exciton envelope function with the phonon plane wave in the quantum well grow direction.

In order to account for the polarit-on-polariton scattering in the \( z \) direction, we take \( G_q \) as the integral of \( G_{q,q_z} \) over \( q_z \) in the interval \([-2\pi/L_z, 2\pi/L_z]\), with \( L_z \) equal to the width of the quantum well.

Using the Heisenberg equation, we obtain a mean-field evolution of the polariton wavefunction \( \psi \), of the form in Eq. 4 where now \( F_i \) represents the amplitude of the optical driven field that injects polaritons into the system, \( \Gamma \) is equal to the polariton decay rate and the plastic potential \( V(x) \), appears due the interaction of the excitonic part of the polaritons with acoustic phonons. In this case, the Fourier components of the potential are given by,

\[ V_{ij} = X |G_{ij}| \left( \chi_{ij} - \chi^*_j \right) \]  

(13)

where during the scattering process, the phonon wavevector \( q_{ij} \) is equal to the difference of polariton wave vectors \( k_i \) and \( k_j \), i.e \( G_{ij} = G_{|k_i - k_j|} \). The phonon wavefunction \( \chi = \langle \hat{b} \rangle \), have been decomposed into components with wavevectors given by differences of the polariton wavevector states:

\[ \chi(x) = \sum_{ij} \chi_{ij} e^{i(k_i - k_j) \cdot x} \]  

(14)

Note that the phonon component \( \chi_{ij} \) is strongly stimulated when the polaritons with states \( k_i \) and \( k_j \) are simultaneously activated. Phonons generated in this way inherit coherence from the polariton states and have an evolution given by:

\[ i\hbar \frac{\partial \chi_{ij}}{\partial t} = \left( \hbar u q_{ij} - i\Gamma \right) \chi_{ij} + X G^*_j \psi_i^\dagger \psi_j \]  

(15)

we introduced the phonon decay rate \( \Gamma \). We neglect other wavevector states in the system that are only weakly occupied given our excitation scheme, where only an specific set of inputs and outputs will be driven optically [Fig.4].

The exciton-phonon scattering amplitude \( G_q \), remains approximately constant for small \( |q| \), therefore we can safely substitute \( G_q \) for \( G_0 \approx 6.5meV \). Finally, for timescales larger than the polariton lifetime, the polariton field will reach the steady state Eq.(7), then the solution of Eq.(15) becomes,

\[ \chi_{ij} = \frac{X G_0 F_i F_j e^{-i\hbar u q_{ij} t} \psi_i^\dagger \psi_j - 1}{(\hbar u q_{ij} - i\Gamma \chi) (i\Gamma - E_0)^2} \]  

(16)

where, as before, \( F_i \) and \( F_j \) represent the amplitude of the optical driven field for the inputs and outputs.

If we choose to work with small wavevectors in reciprocal space, \( |k_i| = 0.5\mu m^{-1} \), the phonon lifetime is expected to be in the range of 100ns [17]. This represents the memory time of the system and is sufficient for operation afterwards [recalling that the network can be operated with picosecond repetition times due to the fast polariton dynamics].

As shown in Fig.1a, during training, the optical driving field stimulates simultaneously the input and the desire output in the optical microcavity; in this way, the driving field leaves a phonon-lifetime lasting mark through the polariton-phonon interactions Eqs.(13,16), that will allow the system to “memorize” the digit. In Fig.3a we show the contour plot of \( V_{ij} \), as a function of the pulse duration [see time dependence in Eq.(16)], and as a function of the polariton field. As we expected, the polariton-phonon interaction is amplified when the polariton field increases and the coherent stimulation of the wavevectors \( |k_i - k_j| \) last longer. These characteristics allow to control optically the weights of the system Eq.(1) and the training time.

During the “operation phase” [Fig.1b], the output of the system can be calculated straightforward using Eq.(5). After training the system during 100ps for each input set [each digit], we achieve full recognition of all digits as shown in Fig.8. This clearly shows how the perceptron corrects errors and is therefore different from simply copying the input.

Dynamic Nuclear Polarization. — If during the training phase, the driving field is circularly polarized and laser energy \( E_p \) is so as to excite free electrons, then one can consider the dynamic polarization of nuclei. The hyperfine interaction between a single electron and single
nuclear spin is given by the Hamiltonian \[ \mathcal{H}_{hf} = \nu_0 A |\Phi(R)|^2 \left( \hat{I}_x \hat{S}_x + \hat{I}_y \hat{S}_y + \hat{I}_z \hat{S}_z \right) \] (17)

where \( \nu_0 \) is the unit cell volume and \( A \) is the hyperfine coupling constant. \( \Phi(R) \) represents the electron envelope function, evaluated at the position of the nuclear spin. \( \hat{I} \) and \( \hat{S} \) represent the nuclear and electron spin operators, respectively. An electron with spin polarized in say the \( z \)-direction can undergo a spin flip, transferring its spin to a nucleus.

The average nuclear spin polarization in the system is determined by the rate equation:

\[
\frac{d\langle I_z(x) \rangle}{dt} = \Gamma_{hf}(x) \left( Q \langle S_z \rangle - \langle I_z(x) \rangle \right) - \frac{\langle I_z(x) \rangle}{\tau} \]

where \( \tau \) is the nuclear spin relaxation time, which takes values at least on the order of microseconds \[21\, \text{to} \, 22\] and even up to minutes \[23\]. The quantity \( Q = \frac{I}{I_{\text{sat}}} \), where \( I \) is the total nuclear spin and \( S_z = 1/2 \) is the total spin of an electron. The hyperfine scattering rate \( \Gamma_{hf} \), was calculated for a semiconductor microcavity in Ref. \[24\] and is proportional to the electron density. Compared to the wavelength of polaritons, electrons do not move significant distances during their lifetime such that we can consider the hyperfine scattering rate as being proportional to the optical field intensity. Solving Eq. (18), we found,

\[
\langle I_z(x) \rangle = \frac{\Gamma_{hf} Q S_z \tau \left( 1 - e^{-\left( \Gamma_{hf} + \tau^{-1} \right) t} \right)}{\Gamma_{hf} + \tau^{-1}} \] (19)

For small quantum wells \( [L_z \sim 10 \text{nm}] \) and electron density around \( 10^{12} \text{cm}^{-2} \) we have that \( \Gamma_{hf} \ll \tau^{-1} \), therefore we can approximate the above expression as \( \langle I_z(x) \rangle = \Gamma_{hf} Q S_z \tau \left( 1 - e^{-t/\tau} \right) \). Consequently, the polariton potential given by \( V(x) = XA(I_x(x)) \propto |F_0|^2 \) as desired.

Then the potential that the nuclear spin polarization induce in the polaritons, can be used in a similar way to the first example, where now,

\[
V_{ij} = \frac{8 \pi X A^3 Q S_z m_e \nu_0^2 I(I + 1)}{3 \hbar^3 V L_z E_p^2} F_i F_j \] (20)

with \( E_p = E_c(\lambda) \), the energy of the cavity mode at the pump wave vector. Notice that in this case first we need to induce the potential with an optical pump tuned to excite the electrons and later inject the polariton that carry the information with a laser tuned to the lower polariton band; nevertheless, this system has the advantage that the memory last longer since the dynamics of the nuclear spin polarization is slower [see Fig. 2b]. We perform the calculations for this system and show that it is capable of realize Hebbian learning with results similar than in Fig. 3. The parameters used are given in Ref. \[25\].

![FIG. 2](image_url)

**FIG. 2:** a) Phonon potential Eq. (13), as a function of the duration of the laser that generates the input \( k_i \) and \( k_j \), and the polariton density [see Eq. (16)]. b) Same as a) but for the Nuclear Spin potential.

![FIG. 3](image_url)

**FIG. 3:** We show pattern recognition of the numbers when defects are present using the plastic potential due polariton-phonon interaction, for nuclear spin polarization interactions, we obtained a similar result.

Unlike other hardware implementations of neural networks one does not need to store and control each weight, \( w_{ij} \), in an independent feedback loop. All the weights are encoded in the form of the polariton potential, which automatically adapts when exposed to the correct training conditions.

While the lifetimes of nuclear spins and phonons greatly exceed that of polaritons, they still pose a limitation to the operation time of the network before it forgets its memory. However, in both schemes it would be possible to continuously apply input and output vectors in the training set while making calculations using the opposite spin polarization. The two spin polarizations of polaritons are only weakly coupled allowing “revision” to be applied to the network while it gives a sustained useful output.

In conclusion, we have developed a microscopic theory to construct perceptrons using wave ensembles and plastic potentials that allow to operate the system with an unlimited number of interconnections in a single layer and in picoseconds-micrometers scale. We performed the calculations for two possible devices where the information is carried by the polariton wave function and the system can be controlled optically, we show how the
polariton-phonon and the polariton-nuclear spin polarization, interactions can be implemented to develop a memory in the system that allow to recognize patterns as in neural networks with Hebbian learning. Finally, due the characteristics of the system we believe this approach is a promising step forward to create microscopic neural networks.

[1] R. Rojas, Neural Networks, Springer-Verlag, New York, (1996).
[2] L. Appeltant, M.C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C.R. Mirasso and I. Fischer, Nature Commun. 2, 468 (2011).
[3] R. H. R. Hahnloser, et al. , Nature 405, 947 (2000).
[4] T. Ohno, T. Hasegawa, T. Tsuruoka, K. Terabe, J. K. Gimzewski and M. Aono, Nature Mater. 10, 591 (2011).
[5] Misra & I. Saha, Neurocomputing, 74, 762 (2012).
[6] N. Locatelli, V. Cros, & J. Grollier, Nature Mater., 13, 11 (2014).
[7] V. Q. Diep, B. Sutton, B. Behin-Aein, & S. Datta, arXiv:1401.2654 (2014).
[8] J. Ohta, K. Kojima, Y. Nitta, S. Tai, & K. Kyuma, Opt. Lett., 15, 1362 (1990).
[9] B. Barshan and B. Ayrulu, Neural Net. 15, 131 (2002).