

Maximum Entropy-Based Quantum Boltzmann Machines

Roshawn Terrell
Vilnius University
roshawn.terrell@ff.stud.vu.lt

Eleanor Watson
Atlantic Council GeoTech Center
nwatson@atlanticcouncil.org

Timofey Golubev
Michigan State University
golubevt@msu.edu

Abstract

We construct a Restricted Boltzmann Machine using the principle of maximum entropy and Shannon entropy as the cost function in the optimization problem. We demonstrate that the problem of optimization of entropy in RBM can be described as the Inverse Ising problem and that the optimal values of coefficients for the RBM are identical to the parameters in the Hamiltonian of the Ising model. We also show that real physical variables, such as magnetization and susceptibility, obtained using our RBM are in good correspondence to results from analytical or numerical methods. These results suggest that RBM neural networks using the principle of maximum entropy can be applied to modeling physical systems which can be described by discrete states, including fundamental quantum physics such as topological systems, and biological systems such as the correlated spiking of neurons.

1. Introduction

While neurons share the same characteristics and makeup as other cells of the body, their unique electrochemical aspect allows them to transmit signals through the body. Neuronal communication, therefore, is an electrochemical process that involves different transmitters responsible for different functions. Until recently, the nature of this transmission of signals remained elusive, with numerous theories, such as Hebbian theory, postulating that signal transmission is the result of increased synaptic efficacy that results from the persistent stimulation of the postsynaptic cells by the presynaptic cells. We attempt to resolve this question by suggesting the existence of a powerful connection between entropy maximisation and numerous other fields that include neuroscience, cognitive science, and machine learning, among others. With this understanding, we develop a maximum entropy quantum thermodynamical model for both biological and artificial neural networks.

Small correlations among very many pairs of neurons could add up to a strong effect on the overall network. There are an infinite number of models that are consistent with a given set of pairwise correlations. We cannot experimentally measure all states, but we can measure averages and correlations. The conventional approach to describe the dynamics of a network we are studying is to make some assumptions about the network and then calculate the consequences of these assumptions. However, our assumptions may be wrong. In such complex systems as neural networks, it is best to try to remove all our assumptions and find models of the system which have as little structure as possible while still being able to reproduce some experimental measurements. Such a minimally structured model is by definition the model with the maximum entropy [9, 16].

These Maximum Entropy models have been used to describe how local interactions can generate collective dynamics in neural activity [19].

Ising models from physics are the maximum-entropy probability distributions that can exactly reproduce measured pairwise correlations between spins. There has been a long history of physicists exploring analogies between statistical mechanics of Ising models and the dynamics of neural networks [10-13]. It has been shown that this analogy can be turned into a precise mapping and connected to experimental data where the pairwise interactions Ising model provides a very accurate description of the patterns of neural firing (spiking and silence) in retinal ganglion cells despite the fact that it excludes all higher order interactions between multiple cells [9, 15, 16]. In this work, we develop a maximum-entropy model for neural networks using an Ising-type Hamiltonian in the quantum thermodynamics formalism.

In recent years, one can notice a large number of investigations on the relations between neural-network (NN) algorithm descriptions and quantum thermodynamics (QTD) formalism. The key idea in such a correspondence is that we can put in compliance the description of entropy in the meaning of Shannon (so called information entropy):

$$S_A = - \sum_i p_i \ln(p_i) \quad (1)$$

where, p_i has the meaning of the probability, and the Gibb's entropy:

$$S = -k_B \sum_i p_i \ln(p_i) \quad (2)$$

here, we have the discrete set of microstates with the energies E_i and the probability p_i that the system is in each microstate i . The next very important relation was

shown in [1], stating that the Gibbs Entropy is equal to the classical "heat engine" entropy characterized as:

$$dS = \frac{dQ}{dT} \quad (3)$$

These well-known relations give the possibility to build model relations between different types of NN and quantum systems. As an example, we can provide the work of Deng et. al [2] where the authors use artificial neural networks to study some interesting phenomenon in quantum physics - the topological phases of matter. They described one-dimensional (1D) symmetry-protected topological cluster state and the 2D and 3D toric code states with intrinsic topological orders using short-range neural networks in an exact and efficient fashion. We can relatively easily see the correspondence between NN and quantum physics models in the 1D case. The next figure (Fig. 1) shows this relation. For the 1D cluster (Fig 1 (a)) we have the following Hamiltonian:

$$H = - \sum_{k=1}^N \sigma_{k-1}^z \sigma_k^x \sigma_{k+1}^z \quad (4)$$

where, σ^z , σ^x – are Pauli matrices and N defines the size of the system.

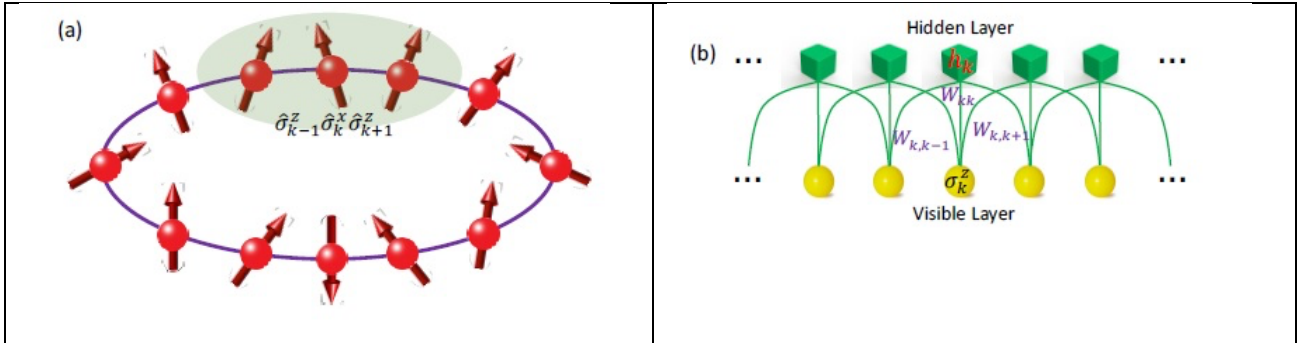


Fig 1. (Illustrations taken from [1]). 1D quantum cluster of $\frac{1}{2}$ spins (a) and corresponding neural network (b).

The Hamiltonian (Eqn. 4) has boundary conditions and two symmetric transformations. Due to such nature, the ground state of the Hamiltonian $|G\rangle$ obeys:

$$\sigma_{k-1}^z \sigma_k^x \sigma_{k+1}^z |G\rangle = |G\rangle \quad (5)$$

In the context of quantum information and computation, this state is called a *cluster state* or more generally a *graph state*. And for this cluster state the authors show how to construct an exact artificial neural network representation (Fig. 1(b)) with the following artificial-neural-network quantum state (ANNQS):

$$\Phi_M(\Xi; \Pi) = \sum_{\{h_k\}} \exp\left(\frac{i\pi}{4} \sum_k h_k (1 + 2\sigma_{k-1}^z + 3\sigma_k^x + \sigma_{k+1}^z)\right) \quad (6)$$

Here, $\{h_k\} = \{-1, 1\}^M$ denotes the possible configurations of M hidden auxiliary spins and the weights $\Pi = (\mathbf{a}_k, \mathbf{b}_{k'}, W'_{kk})$ are parameters needed to train to best represent the many-body quantum state.

Here, we reviewed one example of a correspondence between the NN and quantum physics models in detail but this work is not unique and we can see a number of similar papers where authors starting from NN and matching these with a physical model or vice-versa. For example we can mention the next work in this subject [3], where Ramacher conceived the NN as a system which can be derived from a partial differential equation of Hamilton-Jacobi type and described the neurons and weights as state variables of the dynamic system. Among others we can also admit the Luis M. Silva paper [4] where authors proposed neural network classification using as cost function Shannon's entropy of the error using the principle which minimizes error entropy in order to maximize the mutual information between the output and the desired target of a neural network. Or, we can also cite the paper [5] where they also used an entropy formalism but under the minimum description length principle.

We can use the language of physics to describe some NN or (and this is more importantly in our case) we can build the NN to study some physical model/phenomena (as it was done in the [2]). The main relation between these two cases is the entropy.

The physical meaning of the Gibbs entropy derives from the thermodynamic potential function of the *Grand canonical ensemble*:

$$\Omega = k_B T \ln \left(\sum_{\text{all microstates}} \exp \left(-\frac{E_{\text{states}}}{k_B T} \right) \right) \quad (7)$$

From the Grand canonical ensemble, we can write the entropy as the derivative:

$$S = -\frac{\partial \Omega}{\partial T} \quad (8)$$

2. Maximal Entropy Bose-Hubbard RBM Model for Neurons

2.1 Model Formalism

In accordance with previous works [8], a large system of interacting particles can be described by

$$H_{\text{tot}} = H_{\text{sys}}(x, \lambda(t)) + H_{\text{bath}}(y) + h_{\text{int}}(x, y) \quad (9)$$

where x accounts for all the coordinate degrees of freedom for the N particles in the system, y does the same for the bath, and the Hamiltonian functions H_{sys} , H_{bath} , and h_{int} define conservative interactions among the various position coordinates of system and bath. The function $\lambda(t)$ plays the role of a time-varying external field that acts exclusively on the system and can do work on the coordinates x . h_{int} is assumed to be small and may be ignored. This Hamiltonian has been used to model neural networks in past works [10, 11].

To model the dynamics of firing neurons we propose the next model Hamiltonian:

$$H = -J \sum_{\langle i,j \rangle} \delta_{ij} b_i^+ b_j + \sum_i \frac{U}{2} (n_i - 1) n_i - \sum_i \mu n_i \quad (10)$$

where b_i^+ is an operator of appearance of the neuron (in model we can say that this is boson) on the state i , B_i – operator of disappearance of the neuron on the state i , $J\delta_{ij}$ – are tunneling coefficients (in our case we can say that this is the nearest-neighbor hopping), U – is the on-site interaction strength between two bosons, and

the last term corresponds to the chemical potential (this is related to the contact with the “bath” of heat).

Written in this way, the Hamiltonian is called the Bose-Hubbard model. Bose, because of Bose-Einstein statistics to describe the particles; Hubbard, because of the Hubbard-like term ($b_i^\dagger b_j$) related to the strong correlation between particles. This model is widely used to describe the phase transition to the state of Bose-Einstein condensate in systems of ultracold bosons in optical lattice which occurs at some values of $\frac{J}{U}$. In our case, the term with J describes the interactions due to neighboring firing neurons. The important portion in this term is the matrix formed by J and the Kronecker delta function δ_{ij} which has the meaning of the network adjacency matrix A such that $A_{ij}=1$ when we have a link between i and j sites and $A_{ij}=0$ otherwise.

Similar models have been used to describe the behavior of the system in depend of network topology [17] and for networks in machine learning [18]. But in both mentioned papers (as well as in [10]), the authors used quite simple Mean-Field Approximation (MFA) while we will use the formalism of Hubbard operators which will allow us take into account the interaction between bosons in more reliable way than the MFA did. the X-operator transform the system from the state “m” to state “n” and can be written as:

$$X_i^{nm} = |n, i \rangle \langle m, i| \quad (11)$$

Then, the operators of creation and annihilation of bosons in the X-operators approach are:

$$b_i = \sum_n \sqrt{n+1} X_i^{n,n+1}, \quad b_i^\dagger = \sum_n \sqrt{n+1} X_i^{n+1,n} \quad (12)$$

And the Hamiltonian (10) can reformulated as:

$$H = \sum_{i,n} \alpha_n X_i^{nn} - J \sum_{ij} \delta_{ij} \left(\sum_n \sqrt{n+1} X_i^{n+1,n} \right) \left(\sum_n \sqrt{n+1} X_j^{n,n+1} \right) \quad (13)$$

Where,

$$\alpha_n = \frac{U}{2} n(n-1) - \mu n, \quad (14)$$

are single-site energies.

One can write the equation (13) in the matrix form as:

$$H = \sum_i H_i^{\text{latt}}$$

where:

$$H_i^{\text{latt}} = \begin{pmatrix} \alpha_0 & -J & 0 & 0 & \dots \\ -J & \alpha_1 & -\sqrt{2}J & 0 & \dots \\ \dots & -\sqrt{2}J & \alpha_2 & -\sqrt{3}J & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & -\sqrt{n}J & \alpha_n \end{pmatrix} \quad (15)$$

This is a symmetric tridiagonal matrix and can be diagonalized using standard numerical methods. We will get the diagonal matrix:

$$H'_{\text{latt}} = \begin{pmatrix} \beta_0 & 0 & 0 & 0 \\ 0 & \beta_1 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \beta_n \end{pmatrix} \quad (16)$$

where the new functions β_k are functions of α_k and J . Using these β_k (we have to notice that they have a meaning of energy levels) we can use the known relation between entropy and grand canonical potential GCP (7), to define the GCP as:

$$\Omega = k_b T \ln \left(\sum_k \exp \left(-\frac{\beta_k}{k_b T} \right) \right) \quad (17)$$

In general, we can take the derivative as:

$$\begin{aligned} \frac{\partial \Omega}{\partial T} &= k_b \ln \left(\sum_k \exp \left(-\frac{\beta_k}{k_b T} \right) \right) \\ &+ k_b T \frac{1}{\left(\sum_k \exp \left(-\frac{\beta_k}{k_b T} \right) \right)} \left(\sum_k \exp \left(-\frac{\beta_k}{k_b T} \right) \left(\frac{\beta_k}{k_b T^2} \right) \right) \end{aligned} \quad (18)$$

Relation (18) can be solved numerically using earlier obtained very important functional relation

$$\beta_k = F(\alpha_{k-1}, \alpha_k, \alpha_{k+1}, J) \quad (19)$$

In (19) we have the dependence only from α with indexes $k-1, k, k+1$ because in general, the resulting matrix (15) is tridiagonal. Let's limit our discussion on known physics of Bose-Einstein systems, the so-called hard-core bosons approximation. For the case of the Bose-Hubbard model this means that we have only possible states with number of bosons on-site $n_B \leq 1$. Thus, for matrix (15) we will get using (14):

$$H_i^{latt} = \begin{pmatrix} \alpha_0 & -J \\ -J & \alpha_1 \end{pmatrix} = \begin{pmatrix} 0 & -J \\ -J & -\mu \end{pmatrix} \quad (20)$$

We can find by solving the quadratic equation for eigenvalues of H_i^{latt} :

$$\beta_0 = -\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + J^2}, \quad \beta_1 = -\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + J^2} \quad (21)$$

Then, the GCP can be written as:

$$\Omega = k_b T \ln \left(\exp \left(-\frac{\beta_0}{k_b T} \right) + \exp \left(-\frac{\beta_1}{k_b T} \right) \right) \quad (22)$$

And entropy thus is:

$$S = -\frac{\partial\Omega}{\partial T} = -\left[k_b \ln \left(\exp\left(-\frac{\beta_0}{k_b T}\right) + \exp\left(-\frac{\beta_1}{k_b T}\right) \right) + \frac{1}{T} \frac{\left(\exp\left(-\frac{\beta_0}{k_b T}\right) \beta_0 + \exp\left(-\frac{\beta_1}{k_b T}\right) \beta_1 \right)}{\exp\left(-\frac{\beta_0}{k_b T}\right) + \exp\left(-\frac{\beta_1}{k_b T}\right)} \right]$$

We have derived the entropy function from model parameters as:

$$S = \Phi(T, \mu, J) \quad (23)$$

For maximal entropy we can take the derivatives of this function with respect to the model parameters to be equal zero (for example, $\frac{\partial\Phi}{\partial T} = 0$ if we want find critical temperature) to find the behavior of the system with respect to criticality (for Bose-Hubbard model this related to Mott Insulator (MI)-Superfluid (SF) phase transition).

The following figure presents the dependence of entropy calculated for the case of hard-core boson approximation and using relation (23) as function of tunneling coefficient J . All variables are in units of energy (for this we use $k_B T \rightarrow \theta$) and $\theta = 1$.

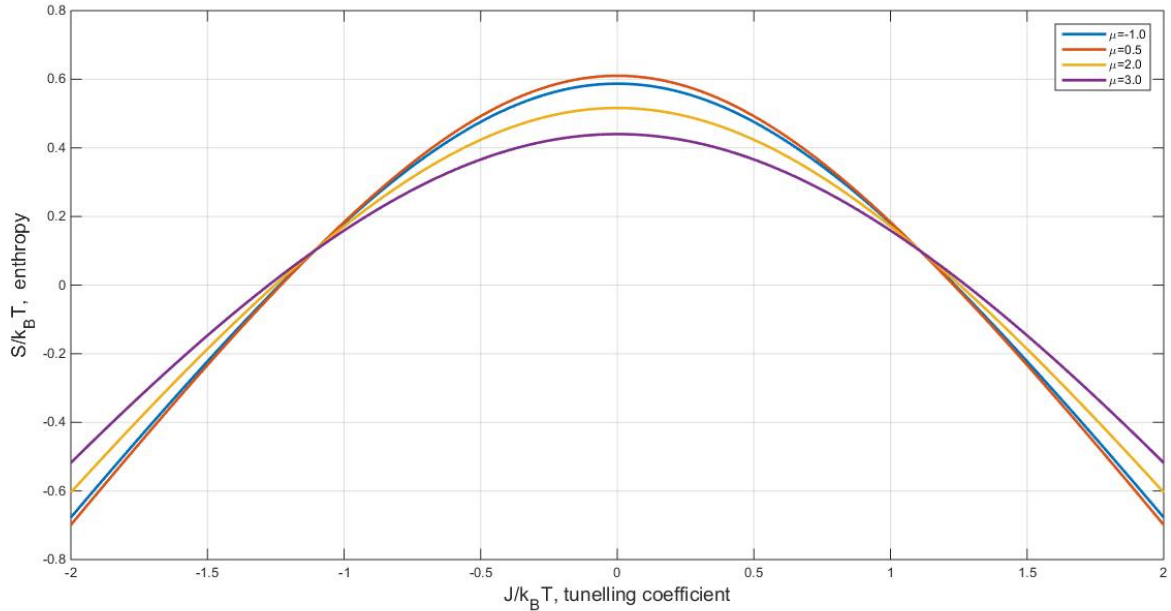


Fig.1. Entropy of hard-core bosons case of Bose-Hubbard model as function of tunneling

coefficient J, for different values of chemical potential

As one can see from Fig.1, the entropy has a maximum at the value $J=0$ for any value of chemical potential, but this result is for the hard-core version of Bose-Hubbard model. In the case when the number of bosons on-site can be more than 0 or 1, we will have one more important parameter U (related to the on-site interaction of bosons) and the behavior of entropy function will be more complicated.

2.2 Constructing the RBM Neural Network

Our neural network approach is based on the paper of Bausch and Leditzky [24] and will serve as a starting point for developing a NN which corresponds to the Bose-Hubbard model described above. We first formulate an optimization problem for Absolutely Maximally Entangled (AME) for dimension d and number of qubits n . For example, for an AME(n,d) we need to define state $|\psi_{n,d}\rangle$.

This state is the decomposition with respect to a known basis $\{|i\rangle\}_{i=0}^{d-1}$

$$|\psi_{n,d}\rangle = \frac{1}{C} \sum \psi(i^n) |i^n\rangle \quad (24)$$

where the function $\psi(i^n)$ is computed by the neural network and to encode these strings i^n we can use different approaches (binary, scaled, one-hot). These basis strings $\{|i\rangle\}_{i=0}^{d-1}$ are what is named in the program as the spin-list - i.e. some space of variables which we encoded to be used in the neural network. For example for binary encoding and AME(3,2) we will get the following spin list:

0 0 0 0 1 1 1 1

$$\begin{array}{cccccccc}
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1
\end{array}$$

Here we have d^n columns of binary coding of basis strings $\{|0\rangle\}$, $\{|1\rangle\}$ and $\{|2\rangle\}$. It is important to note that that basis functions ($|i\rangle$) as functions ($|h\rangle$) are the visible and hidden units from Restricted Boltzmann Machine (RBM):

$$H_{RBM} = \sum a_l i_l + \sum b_k h_k + \sum_{k<l} W_{kl} i_k h_l \quad (27)$$

The values of a , b and W coefficient are the NN weights that define the NN “architecture”. The RBM Hamiltonian has the same form as our Bose-Hubbard Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} \delta_{ij} b_i^+ b_j + \sum_i \frac{U}{2} (n_i - 1) n_i - \sum_i \mu n_i \quad (28)$$

therefore allowing the extension of the NN developed by Bausch and Leditzky to our model. The wave functions for the RBM can be written in terms of AME states as

$$|\psi_n\rangle = \sum_{i^n \in [0,1]^n} \psi_n(i^n) |i^n\rangle \quad (29)$$

where

$$|i^n\rangle = |i_1\rangle \otimes |i_2\rangle \otimes |i_3\rangle \otimes \dots$$

is the spin list. Then

$$|\psi_{RBM}\rangle = |\psi_n\rangle = \sum_{i^n \in [0,1]^n} \sum_{h^n \in \{0,1\}} \frac{\exp(-H(i^n, h^n))}{Z} |i^n\rangle \quad (30)$$

where Z is the partition function and equal to the sum of all possible combinations

$$Z = \sum_{i,h} \exp(-H(i^n, h^n)) \quad (31)$$

Using Eqn 30, we can formulate the optimization problem as follows. For a subset S of n states, $\rho_S = \text{tr}_S(|\psi_n\rangle)$, the constraint on $|\psi_n\rangle$ to be being absolutely maximally entangled is related to the linear entropy

$$S_L = \frac{2^n}{2^n - 1} (1 - \text{tr}(\rho_S^2)) \quad (32)$$

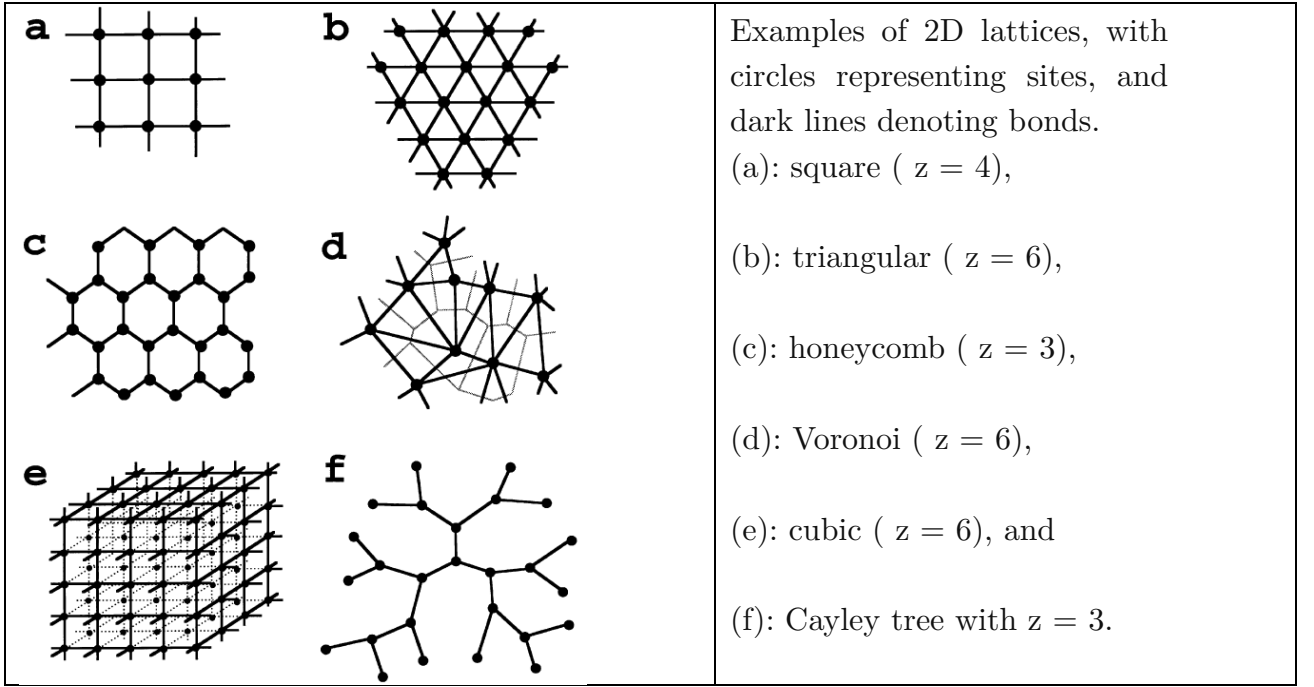
Then the average linear entropy is

$$Q_m(\psi_n) = \binom{n}{n/2}^{-1} \sum_S S_L(\rho_S) \quad (33)$$

which is the main equation for optimization. We use the artificial bee colony, pattern search, and gradient search optimization approaches to optimize $Q_m(\psi_{n,d})$ and generate the NN parameters and state functions ψ_n which obey the maximum entropy condition.

2.3 Analytical solution for Bose-Hubbard Model

Suppose we have some lattice with defined geometry (in 3D cases this can be cubic, hexagonal etc., for 2D cases it can be quadratic, triangle, honeycomb...).



The sites of this lattice are occupied by bosons where n_i is the occupation number of site i . The Bose-Hubbard Hamiltonian is written as:

$$H = - \sum_{\langle i,j \rangle} t_{ij} b_i^\dagger b_j + \sum_i \frac{U}{2} (n_i - 1) n_i - \sum_i \mu n_i \quad (1)$$

where b_i^\dagger is an operator of appearance of the boson i , b_i is operator of disappearance of the boson, $n_i = b_i^\dagger b_i$ is the occupation number of site i , t_{ij} is the nearest-neighbor hopping amplitude, U is the on-site interaction strength between two bosons and the last term corresponds to the chemical potential. For large enough interactions U , the ground state of the system will be a Mott insulating phase while it remains a superfluid for smaller interactions. This was shown in [25].

Using the formalism of Hubbard operators, which will allow us take into account the interaction between bosons, the X -operator transforms the system from the state “ m ” to state “ n ” and can be written as:

$$X_i^{nm} = |n, i\rangle \langle m, i| \quad (2)$$

Then, the operators of creation and annihilation of bosons in the X-operators approach are:

$$b_i = \sum_n \sqrt{n+1} X_i^{n,n+1}, \quad b_i^+ = \sum_n \sqrt{n+1} X_i^{n+1,n} \quad (3)$$

The Hamiltonian (1) can be re-written as

$$H = \sum_{i,n} \alpha_n X_i^{nn} - \sum_{ij} t_{ij} b_i^+ b_j \quad (4)$$

where,

$$\varepsilon_n = \frac{U}{2} n(n-1) - \mu n, \quad (5)$$

are single-site energies.

Now we have to introduce the order parameter of the system. This is the average number of the creation and annihilation operators

$$\varphi = \langle b_i \rangle = \langle b_i^+ \rangle \quad (6)$$

For the Mott insulator phase this parameter is equal to zero, while as soon as $\varphi \neq 0$, this means that the system is in a Bose-Einstein condensate state. Using the mean-field approximation for the hopping term as:

$$b_i^+ b_j = b_i^+ \langle b_i \rangle + \langle b_i^+ \rangle b_i - \langle b_i^+ \rangle \langle b_i \rangle = (b_i^+ + b_i) \varphi - \varphi^2 \quad (7)$$

(4) can be rewritten as:

$$H = \sum_{i,n} \varepsilon_n X_i^{nn} - \varphi t_0 \sum_i (b_i^+ + b_i) - N t_0 \varphi^2 \quad (8)$$

Here we introduce the mean-field average for hopping:

$$t_0 = \sum_{ij} t_{ij}$$

Thus, the BH model Hamiltonian can be written (using (3)) as the sum of matrix (matrix with respect to Hubbard operators) and constant parts

$$H = \sum_{i,pr} H_{pr} X_i^{pr} - N t_0 \varphi^2 \quad (9)$$

where:

$$\hat{H}_{pr} = \begin{pmatrix} \varepsilon_0 & -t_0\varphi & 0 & 0 & \dots \\ -t_0\varphi & \varepsilon_1 & -\sqrt{2}t_0\varphi & 0 & \dots \\ \dots & -\sqrt{2}t_0\varphi & \varepsilon_2 & -\sqrt{3}t_0\varphi & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & -\sqrt{n}t_0\varphi & \varepsilon_n \end{pmatrix} \quad (10)$$

Hard-core bosons case

For example, if we restrict ourselves to the hard-core bosons approximation, where only states with 0 or 1 possible numbers of bosons exist, we will get the following simple 2x2 matrix for \hat{H}_{pr}

$$\hat{H}_{pr} = \begin{pmatrix} \varepsilon_0 & -t_0\varphi \\ -t_0\varphi & \varepsilon_1 \end{pmatrix} = \begin{pmatrix} 0 & -t_0\varphi \\ -t_0\varphi & -\mu \end{pmatrix}$$

where we applied the single-site energies defined by (5). In the case when the interactions are considered, but we restrict the maximum boson occupation number to $n_B \leq 2$ we will have

$$\hat{H}_{pr} = \begin{pmatrix} 0 & -t_0\varphi & 0 \\ -t_0\varphi & -\mu & -\sqrt{2}t_0\varphi \\ 0 & -\sqrt{2}t_0\varphi & U - 2\mu \end{pmatrix}$$

In both cases, we have to diagonalize this matrix by using some canonical matrix relations and the lower eigenvalue of diagonalized matrix will be the ground state energy. For the hard-core bosons model, this can be done analytically:

$$\hat{H}'_{pr} = \begin{pmatrix} 0 & -t_0\varphi \\ -t_0\varphi & -\mu \end{pmatrix} \rightarrow \begin{pmatrix} -\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2\varphi^2} & 0 \\ 0 & -\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2\varphi^2} \end{pmatrix} \quad (11)$$

where we used the following canonical transformation:

$$U^T H U = H', \quad U = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad \sin(2\alpha) = \frac{t_0\varphi}{\sqrt{\frac{\mu^2}{4} + t_0^2\varphi^2}}$$

In the relation (11) we have still one unknown: order parameter φ . To find it we will use the grand-canonical potential and its derivative with respect to the order parameter. Defining, $\beta = 1/T$ the partition function of model using the Hamiltonian (9) will be:

$$Z = Sp e^{-\beta H} = Sp \exp \left(-\beta \left(\sum_{i,pr} H_{pr} X_i^{pr} - N t_0 \varphi^2 \right) \right) = \exp(\beta N t_0 \varphi^2) Z_0^N \quad (12)$$

Where:

$$Z_0 = e^{-\beta \left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right)} + e^{-\beta \left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right)} \quad (13)$$

Thus, the grand canonical potential is:

$$\Omega = - \left(\frac{1}{\beta} \right) \ln Z = -N t_0 \varphi^2 - \frac{N}{\beta} \ln \left(e^{-\beta \left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right)} + e^{-\beta \left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right)} \right) \quad (14)$$

And, then :

$$\left(\frac{\partial \Omega}{\partial \varphi} \right) = -2 t_0 \varphi N - \frac{N}{\beta} \frac{1}{Z_0} \frac{\partial Z_0}{\partial \varphi} \quad (15)$$

So, our equation for order parameter we have as:

$$-2 t_0 \varphi - \frac{1}{Z_0} \frac{\partial Z_0}{\partial \varphi} \frac{1}{\beta} = 0 \quad (16)$$

And, now:

$$-2 t_0 \varphi + \sum_{p'} \langle X^{p'p'} \rangle \frac{\partial \varepsilon_{p'}}{\partial \varphi} = 0, \quad \langle X^{p'p'} \rangle = \frac{1}{Z_0} e^{-\beta \varepsilon_{p'}} \quad (17)$$

Using:

$$\frac{\partial \varepsilon_{0'}}{\partial \varphi} = -t_0 \sin(2\alpha), \quad \frac{\partial \varepsilon_{1'}}{\partial \varphi} = t_0 \sin(2\alpha) \quad (18)$$

we can write the equation:

$$\varphi = \frac{t_0 \varphi}{2} \left(\frac{\langle X^{1'1'} \rangle - \langle X^{0'0'} \rangle}{\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}} \right) \quad (19)$$

So, finally, for the case of $\varphi \neq 0$:

$$\frac{2}{t_0} = \frac{\langle X^{1'1'} \rangle - \langle X^{0'0'} \rangle}{\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}} \quad (20)$$

The equation (20) is the final equation with which we can numerically define the value of order parameter and use this value to find the ground state energy.

Ground state energy for the case of hard-core bosons approximation

We can find the quite simple solution of (20) for the case of $T = 0$. For this condition we can state that for $\langle X^{1'1'} \rangle = 1$, $\langle X^{0'0'} \rangle = 0$, the equation is:

$$\frac{2}{t_0} = \frac{1}{\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}}, \quad \text{and solution } \varphi = \frac{1}{2} \sqrt{1 - \mu^2/t_0^2} \quad (21)$$

Then for (11) we will have:

$$\hat{H}'_{pr} = \begin{pmatrix} -\frac{\mu}{2} - \frac{t_0}{2} & 0 \\ 0 & -\frac{\mu}{2} + \frac{t_0}{2} \end{pmatrix}$$

Therefore, for the case of hard-core bosons we get the following spectrum of the model for the state with $\varphi \neq 0$ (SF phase) and for the state $\varphi = 0$ (Normal phase)

$$E_1^{SF} = -\frac{\mu}{2} - \frac{t_0}{2}, \quad E_2^{SF} = -\frac{\mu}{2} + \frac{t_0}{2}, \quad E_1^{NO} = 0, \quad E_2^{NO} = -\mu \quad (22)$$

Also, the condition (21) gives the solution to when the SF phase exists

$$\varphi = \frac{1}{2} \sqrt{1 - \mu^2/t_0^2} \neq 0, \quad 1 - \frac{\mu^2}{t_0^2} > 0, \quad \mu \in (-t_0, t_0) \quad (23)$$

For numerical evaluation, we can state the unit of energy as t_0 i.e. $t_0 = 1$ and build the energy spectrum as a function of chemical potential.

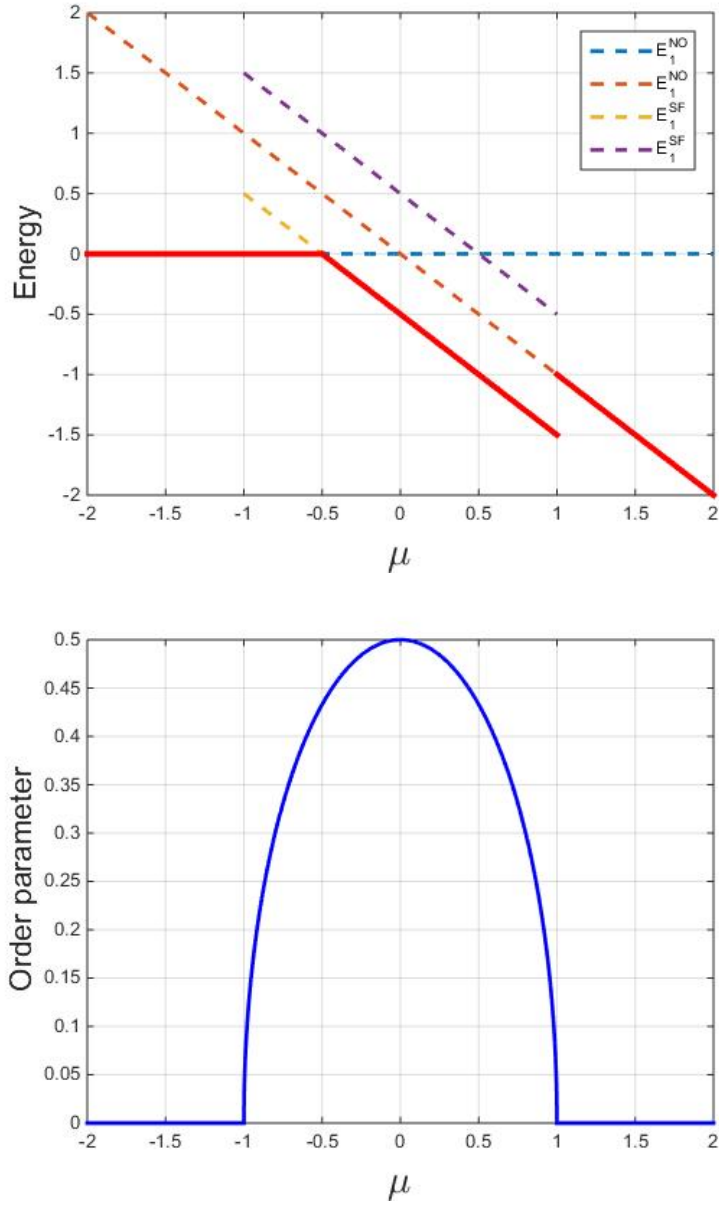


Figure. Energy spectrum for both Normal (NO) and SF phases as function of chemical potential (Upper), Corresponding value of order parameter (Lower).

Note that energy levels for SF phase exist only in the range $\mu \in (-t_0, t_0)$. The ground state energy as function of chemical potential is marked with a filled line in the figure. We see that for small values of $\mu < -t_0/2$, the ground state energy is equal to zero

which means that the ground state is the Normal state $E_1^{NO} = 0$ (state with 0 bosons on-site). For values of $\mu > t_0$, the ground state energy is $E_2^{NO} = -\mu$ (state with one boson on-site). For the intermediate values of chemical potential the ground state will be the state with energy $E_1^{SF} = -\frac{\mu}{2} - \frac{t_0}{2}$ which appears as a result of level splitting of state E_1^{NO} induced by the appearance of the BE- condensate in the system.

Described above is the simple case of hard-core bosons and $T = 0$ which is only case that can be described analytically. For the case of $T \neq 0$ the equation (20) can be solved numerically.

We can return to the case $n_B \leq 2$ and include the interaction between bosons on site U , but this can be done only numerically, for example in [26] the maximal number of bosons on-site was $N_{max} = 5$ and they used the method of exact diagonalization to obtain the values for ground state energy for different values of model parameters.

Entropy for the case of hard-core bosons approximation.

The known relation for Gibb's entropy for the grand canonical ensemble $\Omega = \Omega(T, \mu, t_0, \varphi)$ is

$$S = -\frac{\partial \Omega}{\partial T} \quad (24)$$

where Ω can be written as (14). Thus, for entropy we can write:

$$S = -\frac{\partial(Nt_0\varphi^2)}{\partial T} + N \partial \left[\ln \left(e^{-\frac{1}{T} \left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right)} + e^{-\frac{1}{T} \left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right)} \right) \right] / \partial T$$

And

$$\frac{S}{N} = \frac{1}{T^2} \left(\frac{\left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right) e^{-\left(\frac{1}{T}\right)\left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)} + \left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right) e^{-\left(\frac{1}{T}\right)\left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)}}{e^{-\left(\frac{1}{T}\right)\left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)} + e^{-\left(\frac{1}{T}\right)\left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)}} \right)$$

So,

$$\frac{S}{N} = \frac{1}{T^2} \left(\frac{\left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right) e^{\left(\frac{1}{T}\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)} + \left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \right) e^{\left(-\frac{1}{T}\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)}}{e^{\left(\frac{1}{T}\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)} + e^{\left(-\frac{1}{T}\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)}} \right) \quad (25)$$

And (25) is the final relation for entropy as functions $S = S(T, \mu, t_0, \varphi)$.

So, now we have to solve equation (20) numerically, which can be written in form:

$$\frac{2}{t_0} \frac{\exp\left(-\frac{1}{T}\left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)\right) - \exp\left(-\frac{1}{T}\left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)\right)}{\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2} \left(\exp\left(-\frac{1}{T}\left(-\frac{\mu}{2} + \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)\right) + \exp\left(-\frac{1}{T}\left(-\frac{\mu}{2} - \sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}\right)\right) \right)}$$

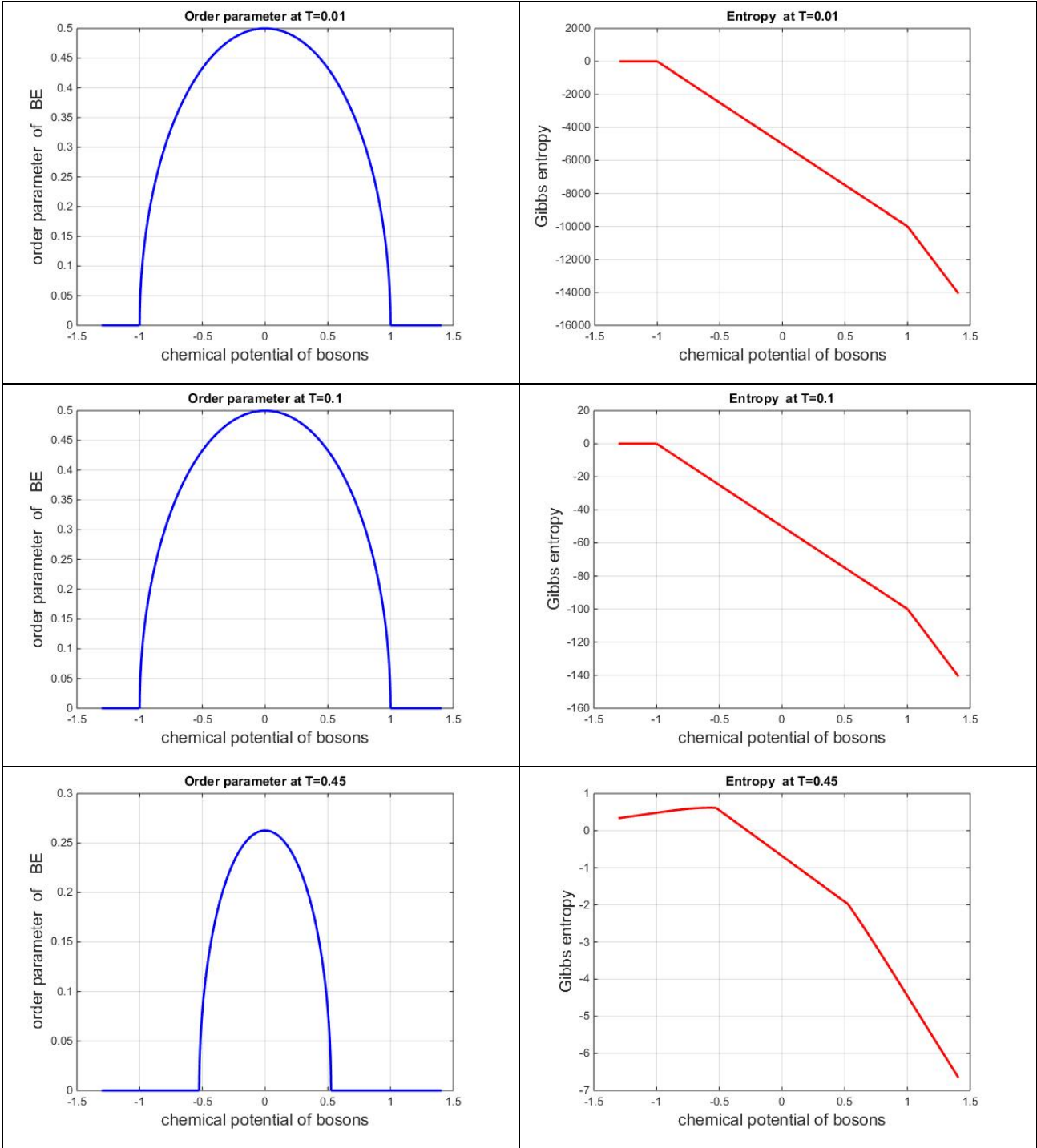
Or, after some simplifications:

$$\sqrt{\frac{\mu^2}{t_0^2} + 4\varphi^2} = \frac{\left(e^{\frac{2}{T}\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}} - 1 \right)}{\left(e^{\frac{2}{T}\sqrt{\frac{\mu^2}{4} + t_0^2 \varphi^2}} + 1 \right)} \quad (26)$$

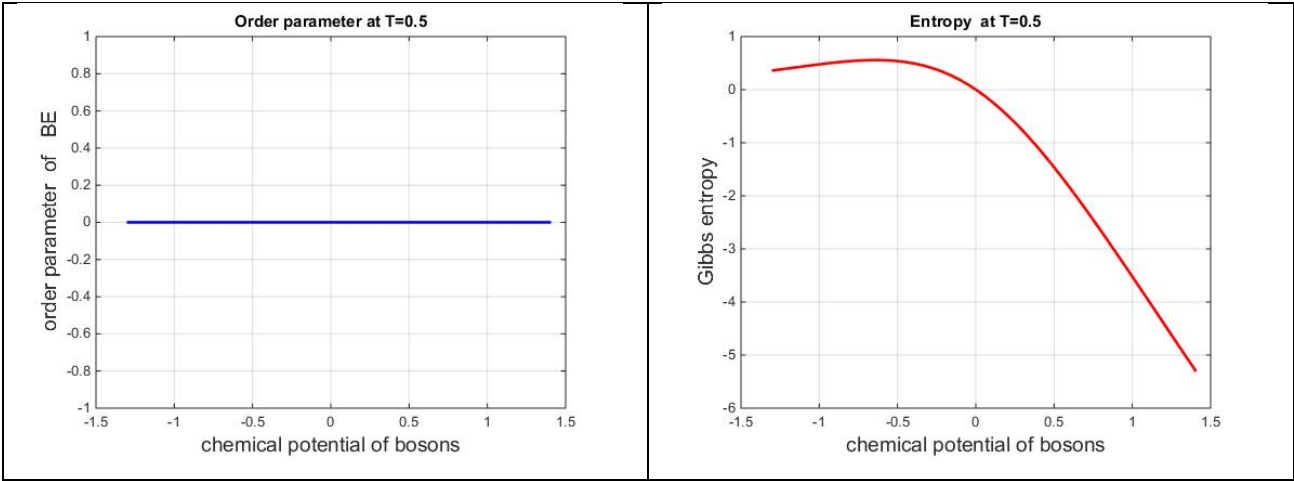
to find the value of order parameter for the given values of T, μ, t_0 . Knowing this value of φ we can find the entropy per site of the Bose-Hubbard model.

Results.

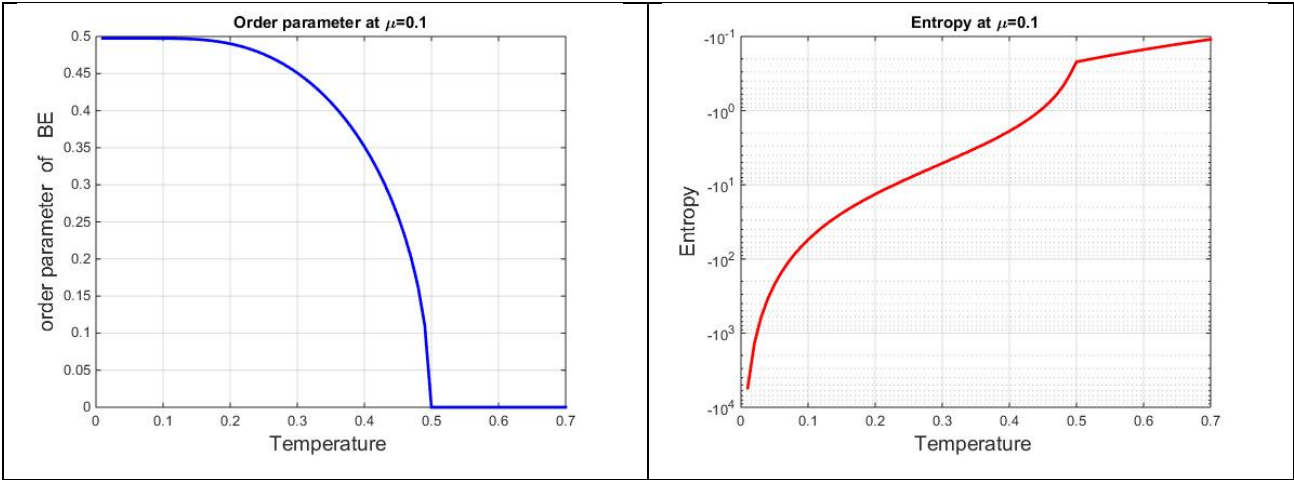
In this section, we present the results of the solution of the equation for the order parameter (26) and related dependencies of Gibb's entropy using (25). We took the value of boson hopping as $t_0 = 1$ so all other variables are normalized relative to this value. The first part is the dependence of order parameter and entropy from chemical potential of bosons for different temperatures:

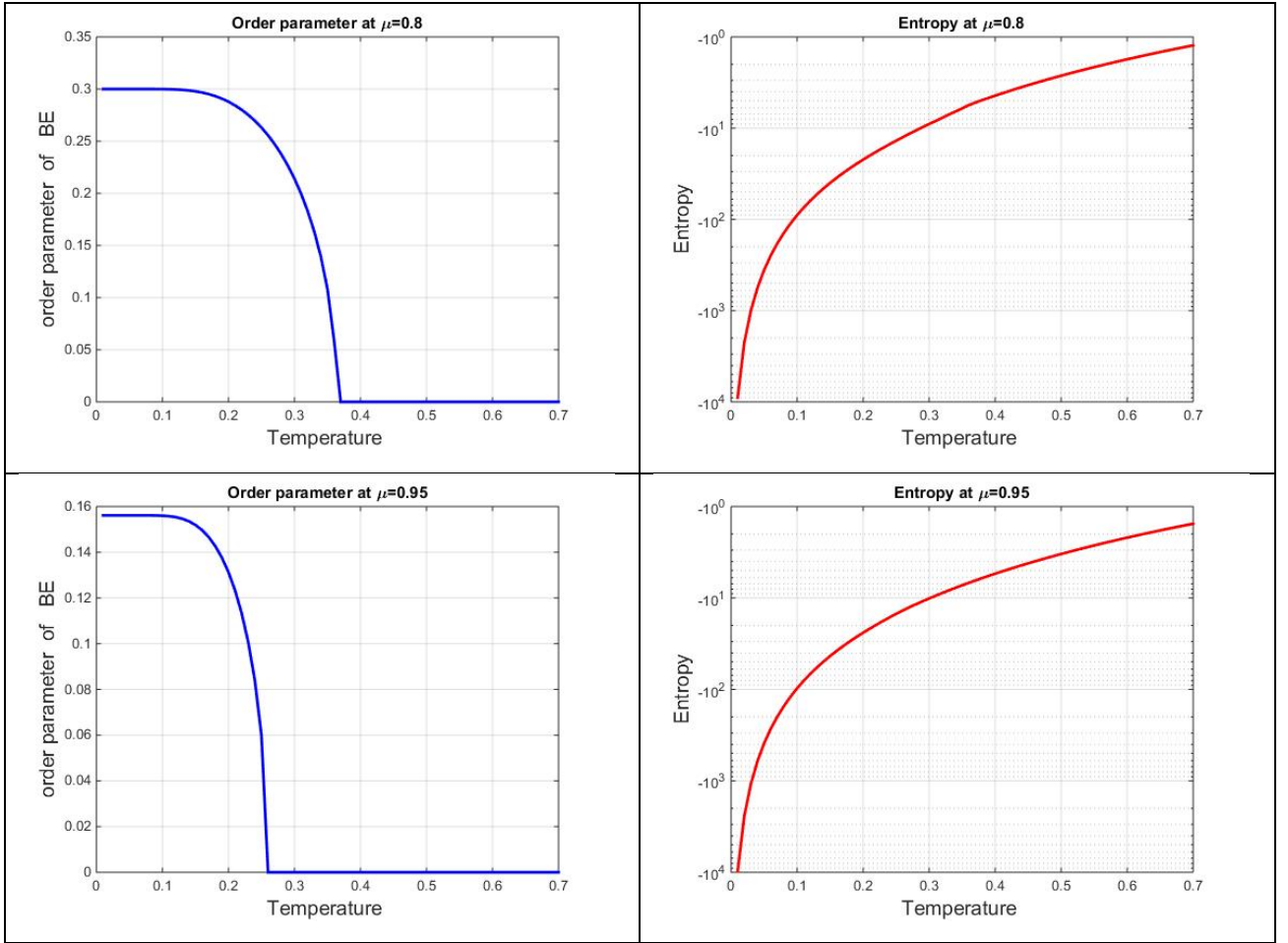


Here one can see that for non-zero values of order parameter (this corresponds to the existence of Bose-Einstein condensate), the entropy line has a discontinuity in slope which is related to the phase transition from normal to superfluid phase in the system. With increasing temperature, we notice a reduction in the region of existence of the BE condensate, and with further increasing of temperature the SF phase disappears.



The next part of the results is related to dependence of the order parameter and entropy on temperature for a given chemical potential of bosons.





Here we also can see that for non-zero values of order parameter, the entropy line has a discontinuity in slope (this is more obvious for small values of chemical potential) and with increasing of temperature, we notice a reduction of the region of existence of the BE condensate.

3. Ising Hamiltonian Neural Network

A network of neurons can be described by the Ising model, borrowed from condensed matter physics, which is written in terms of the following Hamiltonian:

$$H = - \sum_{i,j} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i \quad (1)$$

where σ_i is spin with possible orientations (up, $\sigma_i = 1$ and down $\sigma_i = -1$), h_i is external magnetic field, and J_{ij} is pairwise coupling between spins. Figure 1 shows a depiction of this form of the Ising model with different values of magnetic field on every site, as well as possible pair correlations between all sites. It can be also described as 1D spin chain with periodic boundary and fully pairwise interaction. In terms of biological neurons, the spin can be re-interpreted as the “on” or “off” state of neurons where the “on” state is when the neuron is emitting an action potential (electrochemical pulse).

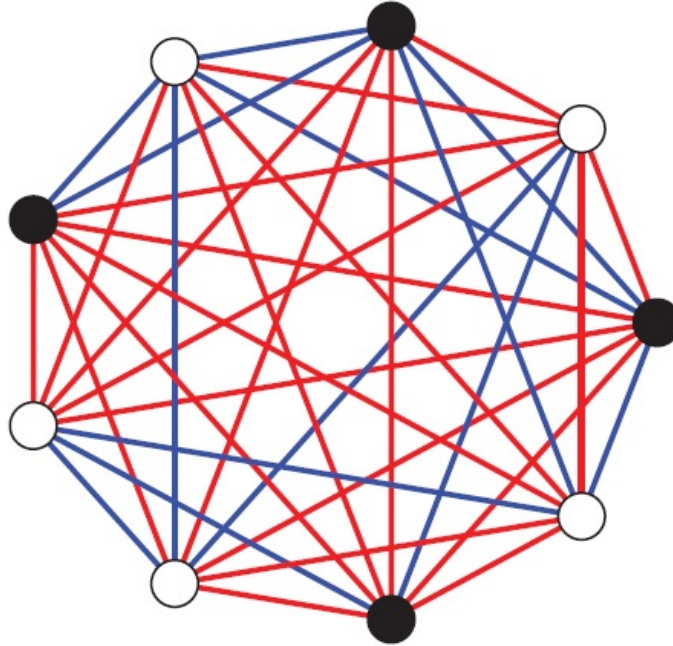


Fig. 1: Fully connected pairwise Ising model. Each spin can have value equal to -1 (black) or 1 (white), and can be connected to every other spin with positive (red) or negative (blue) value. (figure taken from [1]).

In the maximum entropy method we describe the Shannon entropy which is written as follows:

$$S[p] = \sum p(s) \log(p(s)) \quad (2)$$

where, s are possible configurations of system and $p(s)$ are probabilities of these configurations. It is important to note that these configurations could be on-off patterns of firing in neurons or the orientation of spins in a material. More detail description of maximum entropy usage for solving inverse Ising model using the co-niii-3 package is described in the referenced paper [27].

Here we want to describe using the results obtained from the inverse Ising problem to find the thermodynamical variables of the Ising model. In thermodynamics, when we have the values of h_i, J_{ij} for the model (1) we can build the energy term for every possible combination of spins.

$$E(\boldsymbol{\sigma}) = - \sum_{i,j} J_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i \quad , \quad \boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_N) \quad (3)$$

Then, the partition function is a sum over all possible configurations (for N sites we have 2^N configurations)

$$Z_N = \sum_{\boldsymbol{\sigma}} \exp\left(-\frac{E(\boldsymbol{\sigma})}{T}\right) \quad (4)$$

(hereafter we use the ‘energy’ form of temperature which is $T \equiv k_B T$).

The free energy is then then:

$$F_N = -T \log(Z_N) \quad (5)$$

Now, if we want to calculate the average of some observable variable, we have to apply statistical averaging using:

$$\langle A \rangle = \frac{\sum_{\boldsymbol{\sigma}} A \exp\left(-\frac{E(\boldsymbol{\sigma})}{T}\right)}{Z_N} \quad (6)$$

For example, for magnetization per site this gives:

$$\langle m \rangle = \frac{\sum_{\sigma} \frac{(\sum_i \sigma_i)}{N} \exp\left(-\frac{E(\sigma)}{T}\right)}{Z_N} \quad (7)$$

We can also find the so-called Gibbs Entropy (thermodynamical characteristic of the system) as a derivative from Free Energy:

$$S = -\frac{\partial F}{\partial T} = \log(Z_N) + \frac{T}{Z_N} \frac{\partial Z_N}{\partial T} = \log(Z_N) - \frac{1}{Z_N T} \sum_{\sigma} E(\sigma) \exp\left(-\frac{E(\sigma)}{T}\right) \quad (8)$$

Results

We calculated the values for average magnetization and Gibbs Entropy as functions of temperature T using the values of magnetic field (H) and pairwise interaction (J) obtained from the inverse Ising problem. We used N=4,6,8 for the possible numbers of sites in the Inverse Ising Model. This choice can be explained the fact that even numbers are more suitable from symmetry considerations and N=2 is too small for good reliability of an inverse solution. The value of temperature was taken in units of maximal value of inter-site interaction, which is the standard procedure for Ising-like models. The values of H and J when the inverse Ising problem with the coniii-3 package are shown in the table below.

N	Mi n H	Av- er- age H	Ma x H	Mi n J	Av- er- age J	Ma x J
4	0.0	0.1	0.2	-	0.0	0.1
	542	267	126	0.0	449	679
	27	44	95	844	78	57
				96		

6	-	0.1	0.2	-	0.0	0.1
	0.0	066	523	0.2	134	653
	115	30	36	544	32	46
	54			83		
8	-	0.0	0.2	-	0.0	0.2
	0.0	856	185	0.2	133	323
	769	27	58	434	98	34
	62			08		

Figures 2 and 3 show the average magnetization and Gibbs Entropy calculated using Equations 7 and 8, respectively.

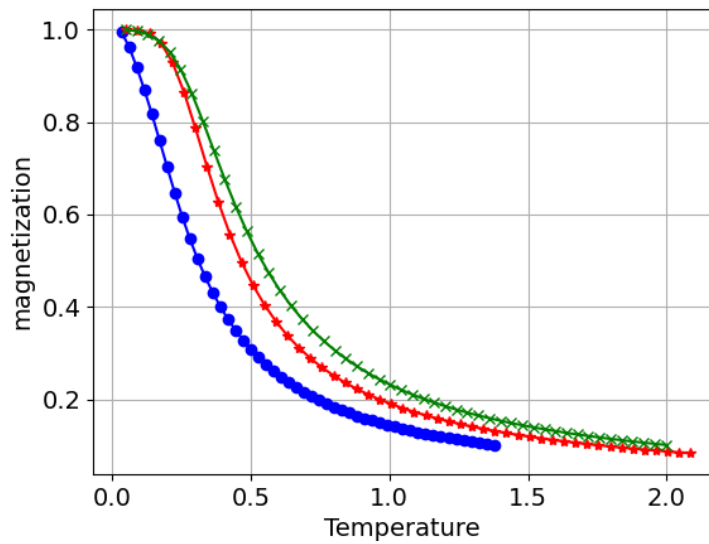


Fig. 2: Average magnetization per site as function of temperature ($N=4$ – blue, $N=6$ – red, $N=8$ – green).

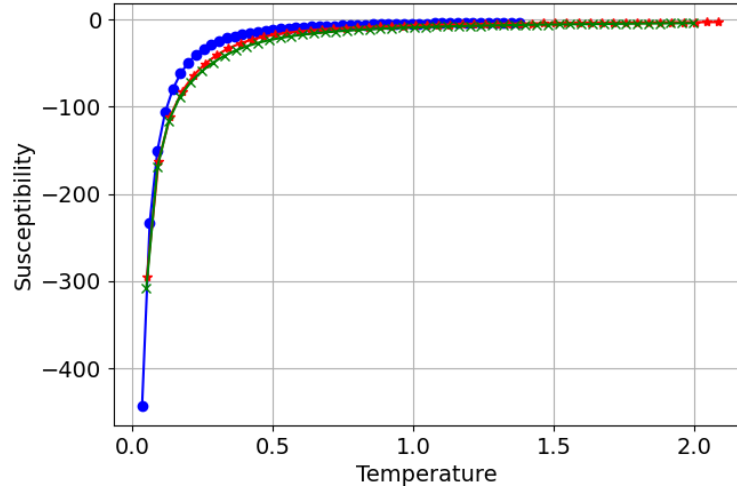


Fig. 3. Gibbs entropy as function of temperature ($N = 4$ – blue, $N=6$ – red, $N=8$ – green).

An additional useful quantity which can characterize the system is the Free Energy (in our case for the Ising model defined as in (5)). In the commonly known and one of the pilot works in Thermodynamics, [28] Gibbs stated the principle of minimum energy:

1) In an isolated system it is a necessary and sufficient condition that for any possible variations in the state of the system, which keeps the energy invariant, the variations of entropy shall either vanish or be negative. **Maximum Entropy Hypothesis.**

2) In an isolated system it is a necessary and sufficient condition that for any possible variations in the state of the system, which do not alter its entropy, the variations of its energy shall either vanish or be positive. **Minimum Energy Hypothesis.**

Therefore, it is known that in the closed system, as in the Ising model or neuron network, the state with the maximum entropy will correspond to the state of the

minimal energy. To illustrate this, we can calculate the Free Energy of Ising Model as a function of temperature (Figure 4).

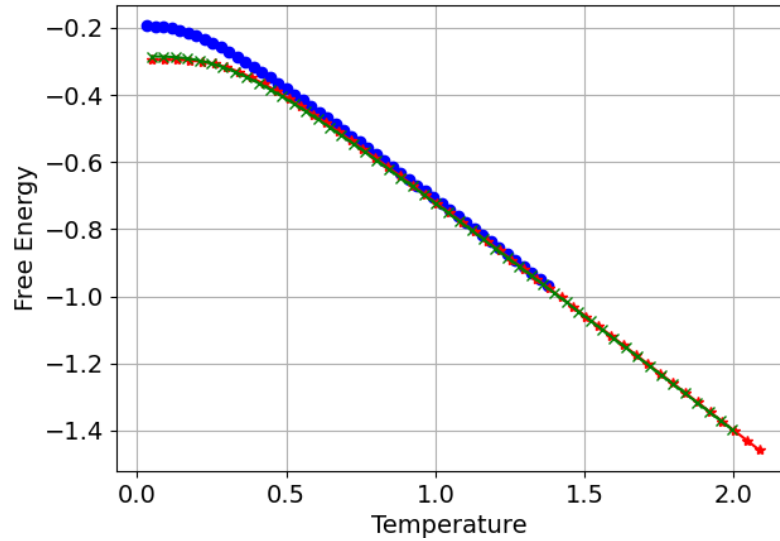


Fig. 4. Free energy as function of temperature (N =4 – blue, N=6 – red, N=8 – green).

When we compare Figures 4 and 3, one can notice that the maximization of entropy corresponds to minimization of the function of energy. Moreover, when the entropy becomes constant, we see a tendency to approach the minimum of energy at this given temperature.

In order to see if our results are correct, we can compare the obtained graphs for magnetization and Entropy to some literature results (Figure 5). As one can see, our results are in good correspondence with results obtained for more complicated systems.

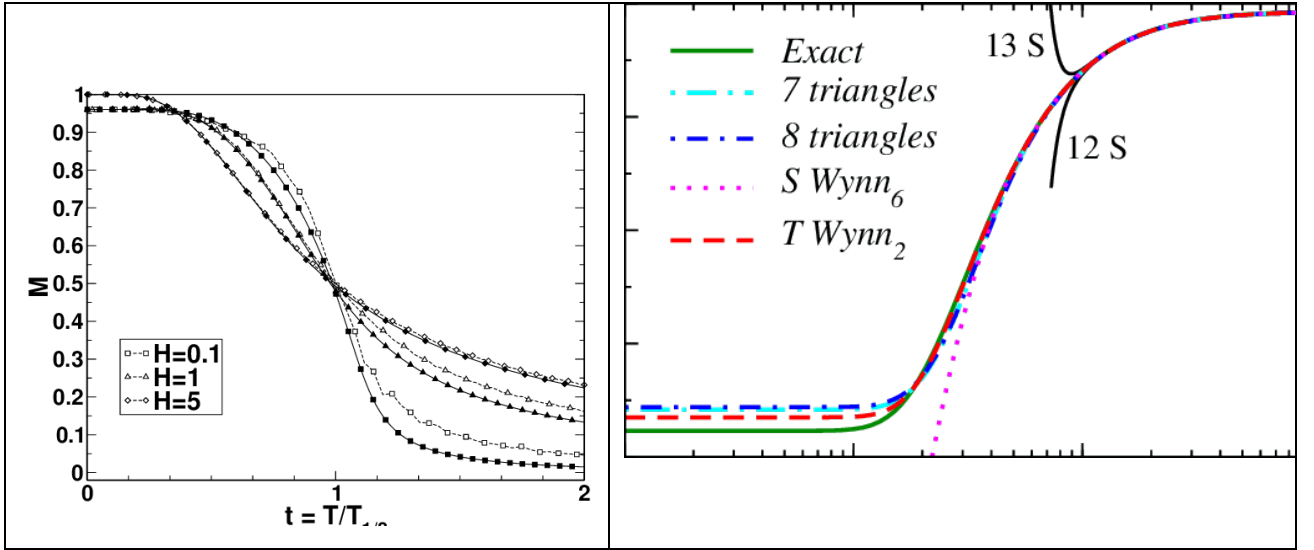


Fig. 5: Average magnetization per site (left) taken from [29] and Entropy (right) taken from [30]

4. Discussion

Identity of the Inversed Ising problem and the Restricted Boltzmann Machine

Here we show that the RBM can be simplified to an Ising model. The general description of the Restricted Boltzmann Machine with visible v_i , hidden h_i units, bias offsets and matrix of weights (α_i, b_i, W_{ij}) has the energy which can be written as:

$$E(v, h) = \sum_{i=1}^n \alpha_i v_i + \sum_{j=1}^m b_j h_j + \sum_{i,j} W_{i,j} v_i h_j \quad (9)$$

As one can see, this relation is very similar to the description of the Ising model if we assume that the visible and hidden units are equivalent (let's mark them $v_i = h_i = s_i$). Then, taking the opposite sign to the weights, we can rewrite (9) as

$$E(s) = - \sum_{i=1}^N \alpha_i s_i - \sum_{i,j} W_{i,j} s_i s_j \quad (10)$$

Now it is easy to see that this is the exact relation for energy of the Ising Model in the form (1) when we denote spins (+1 or -1) as visible/hidden units.

Restricted Boltzmann Machine uses the principle of maximum Entropy (search of minimum of cost function which is minus Entropy) to find the optimal set of coefficients a_i, b_i, W_{ij} in (9). Inversed Ising problem uses the principle of maximum Entropy (written as (2)) to find the optimal set of magnetic field h_i and interaction J_{ij} from (1), and these models are identical when we consider possible values of hidden/visible units in RBM as values of spins in the Inversed Ising problem, which should be equal to +1 or -1.

Conclusions

Taking in the general method of maximum entropy as the optimization function of neural networks, we can state that we can construct the Restricted Boltzmann Machine which will correspond to a real physical system. The optimization problem provided using known methods for RBM will give us the optimal parameters for its different possible physical representations. We have validated the approach by comparison of results for fundamental thermodynamical quantities from the Ising Hamiltonian description of a neural network (a simplified RBM) to the literature. This approach can be used for investigations into fundamental quantum physics (e.g. for different topological states [31]), as well as for biological systems. For example, Ising models have been used to describe the correlated spiking activity of populations of neurons in the retina [32], or for description of amino acid interactions in proteins subject to constraints pertaining to the mean numbers of various types of equilibrium contacts for a given sequence or a set of sequences [33].

References:

- [1] E. T. Jaynes, Gibbs vs Boltzmann Entropies, *American Journal of Physics* **33**, 391 (1965);
- [2] D-L Deng, X. Li, and S. Das Sarma, Machine learning topological states, *Phys. Rev. B* **96**, 195145, (2017)
- [3] U. Ramacher, Hamiltonian dynamics of neural networks, *Neural Networks*, **6**, 547, (1993)
- [4] L.M. Silva, J. M. de Sa and L.A. Alexandre, Neural Network Classification using Shannon's Entropy, *ESANN 2005, 13th European Symposium on Artificial Neural Networks, Bruges, Belgium, April 27-29, 2005, Proceedings*
- [5] S. Wiedemann, A. Marban, K-R Muller and W. Samek, Entropy-Constrained Training of Deep Neural Networks, <https://arxiv.org/abs/1812.07520>
- [6] J. L. England. Statistical physics of self-replication, *The Journal of chemical physics* **139** (12), 09B623_1 (2013)
- [7] J. L. England. Dissipative adaptation in driven self-assembly, *Nature nanotechnology* **10** (11), 919 (2015)
- [8] N. Perunov, R.A Marsland, J. L England Statistical physics of adaptation, *Physical Review X*. **6** (2), 021036 (2016)
- [9] El Schneidman, M. Berry, R. Segev, W. Bialek. *Nature* **440**, 1007-1012 (2006).
- [10] Kelvin Ch'ng, Nick Vazquez, and Ehsan Khatami, Unsupervised machine learning account of magnetic transitions in the Hubbard model, *Phys. Rev. E* **97**, 013306 (2018)
- [11] Hiroki Saito Solving the Bose–Hubbard Model with Machine Learning, *J. Phys. Soc. Jpn.* **86**, 093001 (2017)
- [12] A. D. Wissner-Gross, and C. F. Freer. Causal entropic forces. *American Physical Society*, **110** (168702), (2013).
- [13] JJ Hopfield, *Proc Nat'l Acad Sci (USA)* **79**, 2554 (1982).

- [14] DJ Amit, Modeling Brain Function: The World of Attractor Neural Networks (Cambridge University Press, Cambridge, 1989).
- [15] Tkačik G, Schneidman E, Berry MJ II & Bialek W (2006) Ising models for networks of real neurons. *arXiv.org*: q-bio/0611072.
- [16] G. Tkacik, O. Marre, D. Amodei, E. Schneidman, W. Bialek, M. Berry. Searching for Collective Behavior in a Large Network of Sensory Neurons. *PLoS Comput. Biol.* 1 (2014).
- [17] A. Halu, L. Ferretti, A. Vezzani and G. Bianconi, Phase diagram of the Bose-Hubbard Model on Complex Networks, *EPL (Europhysics Letters)*, **99**(1) , 18001, (2012)
- [18] P. Huembeli, A. Dauphin, and P. Wittek, Identifying quantum phase transitions with adversarial neural networks, *Phys. Rev. B*, **97**, 134109, (2018)
- [19] T. Nghiem, B. Telenczuk, O. Marre, Aa. Destexhe, U. Ferrari. Maximum entropy models reveal the correlation structure in cortical neural activity during wakefulness and sleep. *Physical Review E*. 98, 012402, (2018).
- [20] A. D. Wissner-Gross, and C. F. Freer. Causal entropic forces. *American Physical Society*, 110 (168702), 2013. DOI: 10.1103/PhysRevLett.110.168702.
- [21] L. Zyga. The thermodynamics of learning. February, 2017.
- [22] J, Humphries, L. Xiong, J. Liu, A. Prindle, F. Yuan, H.A. Arjes, L. Tsimring, and G. M. Suel. Species-independent attraction to biofilms through electrical signaling. *Cell*, 168 (1-2), 2017, pp. 200-209. <https://doi.org/10.1016/j.cell.2016.12.014>
- [23] A. Bubnof. A brain built from atomic switches can learn. September, 2017. Retrieved from <https://www.quantamagazine.org/a-brainbuilt-from-atomic-switches-can-learn-20170920/>

- [24] J. Bausch, F. Leditzky. Quantum Codes from Neural Networks. October 25, 2019.
- [25] Fisher M P A, Weichman P B, Grinstein G and Fisher D S 1989 Phys. Rev. B 40(1) 546.
- [26] McBrien K, Carleo G, Khatami E. Ground state phase diagram of the one-dimensional bose-Hubbard model from restricted Boltzmann machines. [arXiv:1903.03076](https://arxiv.org/abs/1903.03076).
- [27] Lee, E. D., & Daniles, Bryan (2019). Convenient interface to inverse Ising (ConIII): A Python 3 package for solving Ising-type maximum entropy models. *Journal of Open Research Software*, 7(1).
- [28] Gibbs, J.W., Criteria of Equilibrium and Stability'' in p.56 The Scientific Papers of J. Willard. 1875. Dover Publ.
- [29] Morawetz, Klaus & Olbrich, Carsten & Gemming, Sibylle & Schreiber, Michael. (2006). Reduction of surface coverage of finite systems due to geometrical steps. *European Physical Journal B*. 62. 10.1140/epjb/e2008-00168-y.
- [30] Rigol, Marcos & Bryant, Tyler & Singh, Rajiv. (2007). Numerical Linked-Cluster Algorithms. I. Spin systems on square, triangular, and kagome lattices. *Physical review. E, Statistical, nonlinear, and soft matter physics*. 75. 061118. 10.1103/PhysRevE.75.061118.
- [31] Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma (2017) Machine learning topological states, *Phys. Rev. B* 96, 195145
- [32] MJ Berry II, R Segev & W Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population. E Schneidman, *Nature* **440**, 1007-1012 (2006); q-bio.NC/0512013.
- [33] F Seno, A Trovato, JR Banavar & A Maritan. Maximum entropy approach for deducing amino acid interactions in proteins, *Phys Rev Lett* **100**, 078102 (2008).