On the statistics of a set of quantum mechanical harmonic oscillators. A machine learning approach.

Carlos Pegueros Denis*
Physics Department, Centro de Investigación y de Estudios Avanzados (CINVESTAV). Mexico City, Mexico.
(Dated: August 1, 2020)

In this work, a Restricted Boltzmann Machine (RBM) is trained to learn and reproduce the statistics of a set of quantum mechanical harmonic oscillators at a fixed frequency and in thermal equilibrium with a heat reservoir at fixed temperature. Thermodynamic quantities are computed as ensemble averages from the generated samples and compared to the theoretical values as a means of validation. Possible future work is discussed.

I. INTRODUCTION

Both classical and quantum statistical physics have provided a successful framework to study physical problems where the large number of particles and its immense amount of degrees of freedom makes it impossible (and, in most cases, not useful) to find an exact solution.

Even though a set of quantum mechanical harmonic oscillators demands the use of quantum statistics to give a correct description of its thermodynamics, it is a fortunate fact that some important statements can still be derived with the aid of classical statistics via the canonical distribution and that is the approach to be followed in the upcoming sections, where a RBM is trained as a generative model to mimic this distribution and reproduce the statistics of such systems. [1]

A. Set of Quantum Mechanical Harmonic Oscillators [1]

The quantum mechanical harmonic oscillator is a well-studied beast whose energy eigenvalues, for a frequency ω, are given by:

$$\epsilon_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{1}$$

Which can be used to calculate the probability ρ_n of finding a single-particle system in a state given by the quantum number n by replacing the Hamiltonian $H(q,p)$ in the classical canonical distribution with equation 1:

$$\rho_n = \frac{\exp\{-\beta\epsilon_n\}}{Z(T,V,1)} \tag{2}$$

This equation can be extended to a N-particle system (for distinguishable particles, which is **not** the case for true quantum statistics, but will serve for the purpose of this work) calculating the partition function as:

$$Z(T,V,N) = [Z(T,V,1)]^N = \left[2\sinh\frac{\beta\hbar\omega}{2}\right]^N \tag{3}$$

Which can be used to derive thermodynamic properties as ensemble averages such as the internal energy $U = \langle E\rangle$:

$$\langle E\rangle = N\hbar\omega\left[\frac{1}{\exp\{\beta\hbar\omega\}-1} + \frac{1}{2}\right] \tag{4}$$

This expression can be conveniently re-written to give a more meaningful result in terms of the mean energy of one oscillator $\langle\epsilon_n\rangle$ and the mean quantum number $\langle n\rangle$ (which is the mean level of excitation of an oscillator at temperature T):

$$\langle E\rangle = N\langle\epsilon_n\rangle \tag{5}$$

where:

$$\langle\epsilon_n\rangle = \hbar\omega\left(\langle n\rangle + \frac{1}{2}\right) \tag{6}$$

$$\langle n\rangle = \frac{1}{\exp\{\beta\hbar\omega\}-1} \tag{7}$$

These results are what is expected from the neural network to be able to reproduce.

II. DEFINING THE BOLTZMANN MACHINE

As a toy model, a set of 100 oscillators was used as the input. In order to use a vector of quantum numbers as the visible layer, its values were normalized to 1 (considering a maximum quantum number n_{max} discussed in section II A) and a Gaussian-Bernoulli RBM was used.

* Personal email: peguerosdc@gmail.com

[1] Code available at: `https://github.com/peguerosdc/ml4phy-quantum-oscillators`

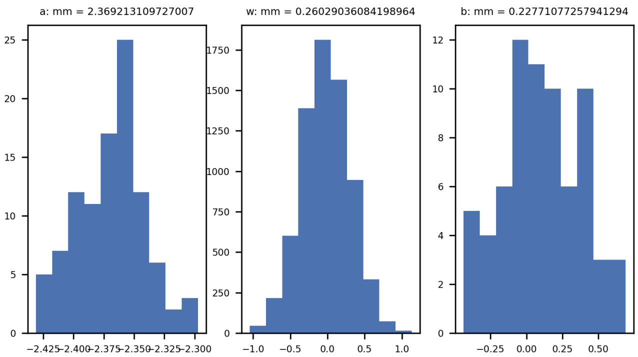

Figure 1. Weights and biases of the trained neural network

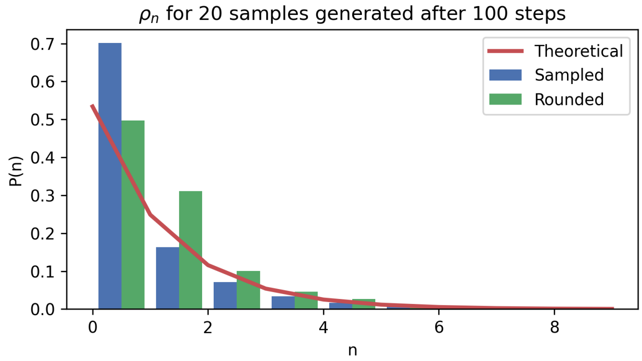

Figure 2. Theoretical and sampled probability distributions for a quantum number n.

A. Data Preparation

A training set of 100,000 samples was generated using a Monte Carlo approach with the aid of equation 2 and a uniform probability distribution to decide the quantum number of each element in the set.

The temperature and frequency were arbitrarily set to $T = 10 \times 10^{-6}[K]$ and $\omega = 1 \times 10^{6}[rad/s]$ to make the lowest quantum numbers the most probable so the Monte Carlo generator could be capped at a defined n_{max}, which in this case was set to 10 by looking at the theoretical probability distribution (see figure 2) and considering that this value could enclose enough quantum numbers to capture the physics of the system.

An example of 8 sets of 8 oscillators each is showed for the sake of clarity:

```
array([[0., 5., 0., 0., 3., 1., 0., 0.],
       [1., 0., 2., 1., 0., 0., 0., 0.],
       [1., 0., 1., 0., 0., 0., 0., 1.],
       [2., 1., 5., 0., 0., 0., 0., 0.],
       [2., 0., 0., 1., 0., 1., 0., 0.],
       [0., 1., 0., 0., 0., 0., 1., 0.],
       [0., 0., 2., 0., 0., 0., 1., 1.],
       [0., 1., 0., 0., 0., 2., 0., 2.]])
```

B. The Training Process

The hyper-parameters were chosen according to the guidelines described in [2] [3] and one combination that yielded good results (which are the ones presented in this paper) consists of 70 hidden units, a batchsize of 10 sets and a learning rate $\eta = 0.005$. No momentum, weight-decay nor sparse targets were used.

After 300,000 training steps using contrastive divergence, the final state of the neural network defined by the weights w, visible biases a and hidden biases b can be seen in the histograms shown in figure 1 along with its mean absolute magnitude mm.

III. RESULTS

Two compute the statistics of the resulting ensembles, two approaches were considered.

The first one consists of calculating averages from the direct output of the neural network. As this output is represented by a Gaussian visible layer as discussed in section II, after denormalizing, the quantum numbers would take real values between 0 and 10, which do not represent real physical situations, so a second approach is presented where these values are rounded to their closest integers before calculating averages.

The reconstructed probability distributions taken from 20 sets generated by the RBM after 100 steps is shown in figure 2. Comparing the plot with the histograms, it can be seen that with the first approach, there is an over population of oscillators in the "real base state", which corresponds to a quantum number $n \in [0, 1)$. When using the second approach, this is not a problem anymore as that excess of oscillators is forced to move to $n = 1$, which shows to be a good compensation as this histogram fits better with the theoretical distribution.

The expected values computed with both approaches are showed in table I along with the theoretical values calculated with equations 5, 6, 7. As opposite to the probability distribution, the sampled averages show to provide more accurate values than the rounded averages, which are supposed to be closer to what a real physical system would be.

	Theoretical	Sampled	Rounded
$\langle E \rangle$	1.4471307636e-26	1.441212123e-26	1.3788526515e-26
$\langle \epsilon_n \rangle$	1.4471307636e-28	1.4412121231e-28	1.3788526515e-28
$\langle n \rangle$	0.872244867016	0.866632503412	0.8075

Table I. Theoretical and computed average values (in the SI) for the two approaches.

IV. CONCLUSION

The Restricted Boltzmann Machine is proven to be a good generative model for a set of quantum mechanical harmonic oscillators, which serves as an instructive way of understanding how this kind of neural networks work and of showing its connection to physics in a simple application apart from the Ising Model, which is the base of how RBMs are defined and which has already been implemented and studied [4].

As future work, this RBM could be improved to cover a wider range of temperatures and frequencies to push the ability of neural networks to abstract and generalize the physics learned in this work even further.

A more ambitious work, would be to choose a completely different system (not necessarily in thermal equilibrium with a heat bath) with more than one phase to see if a RBM is capable of capturing its physics along with its phase transitions.

Another illustrative and different use of neural networks apart from the RBM, would be to train them to calculate thermodynamic observables from a given state (which should be a straight-forward averaging linear task) and some quantities derived from them (such as heat capacity).

[1] W. Greiner, L. Neise, H. Stocker, H. Stöcker, and D. Rischke, *Thermodynamics and Statistical Mechanics*, Classical theoretical physics (Springer-Verlag, 1995).

[2] Geoffrey Hinton, *A Practical Guide to Training Restricted Boltzmann Machines (Version 1)*, Tech. Rep. 2010-003 (University of Toronto, 2010).

[3] J. Yosinski and H. Lipson, "Visually debugging restricted boltzmann machine training with a 3d example," (2012).

[4] Giacomo Torlai and Roger G. Melko, "Learning thermodynamics with boltzmann machines," Phys. Rev. B **94**, 165134 (2016).