Unsupervised learning of thermal phase diagram of the XXZ model

Daniel del Canto de Guzmán.*

The XXZ model is a paradigmatic system in quantum many-body physics, extensively studied due to its rich phase diagram and relevance in condensed matter theory. Identifying phase transitions within this model is crucial for understanding quantum magnetism, yet traditional approaches often require prior knowledge of order parameters [2] or computationally expensive numerical methods that require exact diagonalization of a hamiltonian of a Hilbert space that grows exponentially with the number of spins. In recent years, unsupervised learning techniques have emerged as powerful tools for detecting phase transitions without requiring labeled data. For instance, applying PCA in the latent space of the autoencoders is shown to be useful to identify the order parameters of phase transitions [8]. In the reference [8], a variational autoencoder is applied to detect phase transitions of the two-dimensional Ising model and the three-dimensional XY model. Among unsupervised learning methods, autoencoders have shown great potential in extracting meaningful low-dimensional representations of complex quantum states. In this work, we employ an autoencoder trained on different samples of input data to explore the phase transitions of the XXZ model in an unsupervised manner. The code is implemented in Python in [3].

I. INTRODUCTION.

In the field of physics, especially in the study of complex systems, dimensionality reduction techniques are essential for uncovering fundamental structures and behaviors within high-dimensional datasets. Techniques such as Principal Component Analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE), and autoencoders have become pivotal in analyzing data from various physical phenomena, including phase transitions, quantum states, and material properties.

PCA, introduced by Karl Pearson in 1901 [6], serves as a powerful statistical tool for analyzing variance in multidimensional data. In the context of physics, PCA helps identify the principal directions of variation in experimental data, allowing researchers to focus on the most relevant features that influence the system's behavior. This is particularly important in the analysis of phase transitions, where identifying the key parameters that govern the transition can lead to a deeper understanding of the underlying physics.

Developed by Laurens van der Maaten and Geoffrey Hinton in 2008 [7], t-SNE is a nonlinear dimensionality reduction technique that excels in visualizing complex high-dimensional data. In physics, t-SNE provides a valuable method for examining the intricate relationships between different states of a system, revealing clusters that may correspond to distinct phases or behaviors. This capability is particularly beneficial in the study of quantum systems, where the landscape of states can be highly complex and non-linear.

Autoencoders, a type of artificial neural network introduced by Geoffrey Hinton and colleagues in the 1980s, are particularly useful in unsupervised learning applications within physics [5]. They enable the compression of high-dimensional data into lower-dimensional latent representations while preserving essential physical characteristics. This feature is crucial for analyzing large datasets generated from experiments or simulations, allowing researchers to explore the relationships and structures inherent in the data without prior knowledge of the relevant parameters.

The one-dimensional Heisenberg XXZ model for spin-1/2 particles is a cornerstone in the study of quantum magnetism and strongly correlated systems. This model describes a chain of interacting spins where each spin-1/2 particle interacts with the rest through anisotropic exchange interactions. Besides that, each particle is subject to a magnetic field. Its simplicity and exact solvability make it an ideal testbed for developing and benchmarking new theoretical methods, including the autoencoder to detect phase transitions.

That's why we will study, from an academic point of view, the phase transitions of this model depending on the anisotropy, the magnetic field and the range of interaction. The PCA will be used so as to analyze the principal components of our data, along which the observables may be found in a clusterized way. The t-SNE will be useful when the required transformation to obtain the principal components is non-linear. Finally, if the autoencoder is able to detect these principal components properly, it will reconstruct, from a subset of the whole data set, the whole phase diagram.

* Master's Student, Quantum Science and Technologies, University of Barcelona.; ddelcade7@alumnes.ub.edu

II. ANALYSIS OF PHASE TRANSITIONS.

A. Phase transitions in the XXZ model.

The hamiltonian of the XXZ model for spin-1/2 particles in 1D is given by:

$$H = -\sum_{i<j}^{L} J_{ij}\left(S_i^x S_j^x + S_i^y S_j^y + \Delta S_i^z S_j^z\right) - \sum_{i=1}^{L} h_z S_i^z \quad (1)$$

where Δ is the anisotropy parameter and S_i^x, S_i^y, S_i^z are the spin-1/2 operators at site i. The parameter Δ controls the nature of the spin interactions while h_z is the magnetic field. On the other hand, the interaction between particles is governed by $J_{ij} = \frac{1}{|i-j|^\alpha}$, where α is a positive integer that dictates the range of the interaction. For $\alpha = 0$, all the particles interact with each other in the same magnitude and, for $\alpha = 3$, the system shows a dipolar-type interaction. However for $\alpha \geq 6$ the interactions are effectively limited to nearest neighbours. In [4], it is shown the phase diagram for nearest neighbours interaction.

B. Autoencoder to detect phase transitions.

The aim is to reconstruct a whole dataset from a given subset in order to detect phase transition without needing to compute all the phase diagram. The autoencoder is a neural network [5] with a specific architecture. It is divided into an encoder and a decoder. The last layer of the encoder is the input layer of the decoder. The space in which all the encoded data lies is called the latent space and its dimensions should large enough to capture the main features of the input data. Finally, as its objective is to reconstruct the input, the input and output layers have the same number of neurons.

III. PHASE DIAGRAM INFERENCE METHODS.

To apply the unsupervised analysis, we will need to divide the work in some steps, that will be explained in this section.

A. Definition of the dataset.

At first, we need to define a dataset from which we can select samples to infere the behaviour of the system. For this, we develop `Python` with `netket` libraries, to create the Hilbert space and the hamiltonian of the problem for $\alpha = 0, 3, 6$, $\Delta \in [-1, 1]$ and $h_z \in [0, 2]$. We will work with a chain of $L = 4$ spins and obtain the Gibbs states [1] for different values of T from the set $T = 0, 0.05, 0.1, 1$. From the Gibbs states, we will compute the expected values for observables such as S_z, the total spin, ΔS_z, S_x, ΔS_x and we will obtain too the bipartite entanglement entropy. This will allow us to see the true phase diagrams.

$$\begin{aligned} H =& \hbar\omega_q(t)(\sigma_A^\dagger\sigma_A^- + \sigma_B^\dagger\sigma_B^-) + \hbar\omega_w a^\dagger a + \\ & \hbar g_{qw}(t)(\sigma_A^\dagger a + \sigma_A^- a^\dagger + \sigma_B^\dagger a + \sigma_B^- a^\dagger) \end{aligned} \quad (2)$$

$$(\kappa, \gamma) \quad (3)$$

$$\omega_q(t) = \omega_q(0) + \Delta\omega(1 - e^{-t/T}) \quad (4)$$

$$\Delta\omega = (\omega_w - \omega_q(0)) \quad (5)$$

$$g_{qw}(t) = g_{qw,f}(1 - e^{-t/T}) \quad (6)$$

$$(g_{qwf}, \Delta\omega, T) \quad (7)$$

$$\kappa, \gamma \in \{1.0, 0.1, 0.01, 0.001\}\text{MHz} \quad (8)$$

$$T \in (0, 20.0)\text{ns} \quad (9)$$

$$\Delta\omega \in (0, 2)(2\pi)\text{GHz} \quad (10)$$

$$g_{qwf} \in (0, 2)(2\pi)\text{GHz} \quad (11)$$

$$T = 19.65ns, \Delta\omega = 1.156(2\pi)GHz \quad (12)$$

$$Lat = 7.73ns \quad (13)$$

$$g = 2 \cdot 2\pi GHz \quad (14)$$

$$\kappa = 0.001 \cdot 2\pi MHz, \gamma = 0.001 \cdot 2\pi MHz \quad (15)$$

B. Data clusterization. PCA and t-SNE analysis.

We will analyze all the observables in order to find those directions in which their variance is bigger. This will give us a combination of the input independent parameters. We will use the PCA technique to find a linear transformation of the axes and also t-SNE for non-linear. When we can distinguish different values of an observable per different combination of the new axes, we will say that we have found clusterization for that observable.

C. Define autoencoder architecture and generate phase diagram.

In this step, we define the autoencoder architecture, based on the results that provide different architectures, training it with different samples of the dataset to explore which combination best reveals the true phase diagram. The input data of the autoencoder consists in the Cholesky vector representing the Gibbs state of the system, i.e., the density matrix. This means that the autoencoder will be trained to reconstruct these vectors, that have 272 components.

The chosen encoder will consist in three layers, where the second layer has 200 neurons, and the innermost layer has 150 neurons. The decoder will follow a symmetric architecture, though with different parameters. The last layer of both the encoder and decoder employs a LeakyReLU activation function with a slope of 0.3, whereas the remaining layers use ReLU. Additionally, the first and second layers of both the encoder and decoder are trained with a dropout rate of 20%.

The training will be done with Stochastic Gradient Descent algorithm, using the Adam optimizer and batches of 5 random samples. The whole traning will finish after 100 epochs and we will check how many samples are necessary to obtain a good reconstruction of the phase diagram. We will try with $M = [1, 2, 10, 20, 50, 100, 200]$ samples.

D. Latent space analysis.

We need enough complexity in the latent space of the autoencoder, should we want to capture all the features of the data. That's why we will analyze whether we find the same number of principal components as we had previously found with PCA and t-SNE analysis.

IV. RESULTS.

In this section, we will show the results for null temperature and non null temperature with the methods explained in the previous section.

A. Temperature T=0K

We start analyzing the system in $T = 0K$, where we find the system in the ground state. For $\alpha = 0$ we plot the results for the exact diagonalization in Fig. 1, where we can see continues transitions for the values in the energy or no transition for S_x, whereas we observe divergent gradients for S_z, ΔS_z, ΔS_x and the Von-Neumann entropy, giving three different phases [4].

In the case of dipole-type interaction, we obtain the Fig. 2, where we observe similar results.

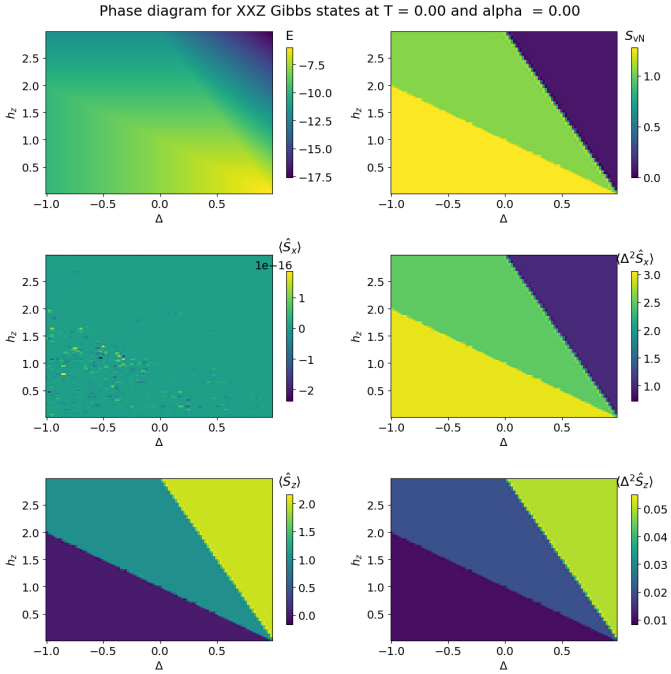

FIG. 1. Energy, entropy, S_z, ΔS_z, S_x and ΔS_x with respect to Δ, h_z for $\alpha = 0$.

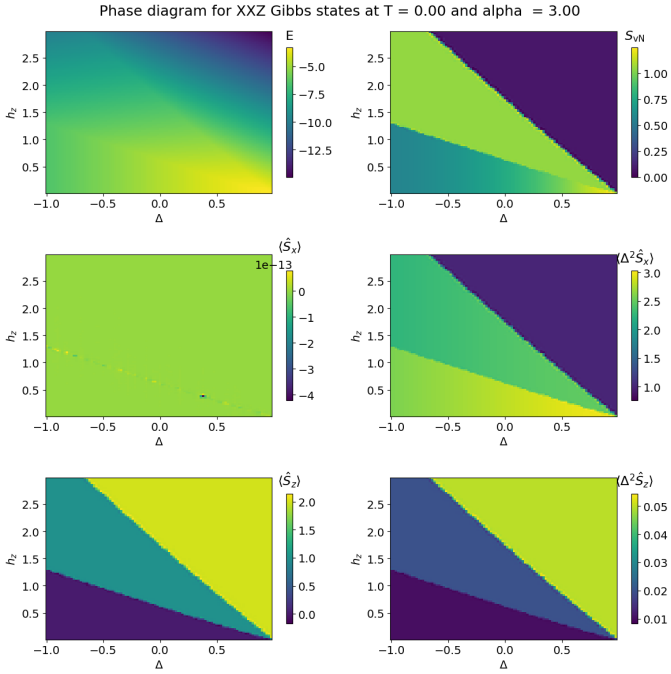

FIG. 2. Energy, entropy, S_z, ΔS_z, S_x and ΔS_x with respect to Δ, h_z for $\alpha = 3$.

Finally, for the nearest neighbours interactions, we obtain the Fig. 3, where we observe again three phases through the same observables.
Thus, we have already defined the whole dataset that we want to reconstruct.

Now, we will apply the PCA and t-SNE analysis. We

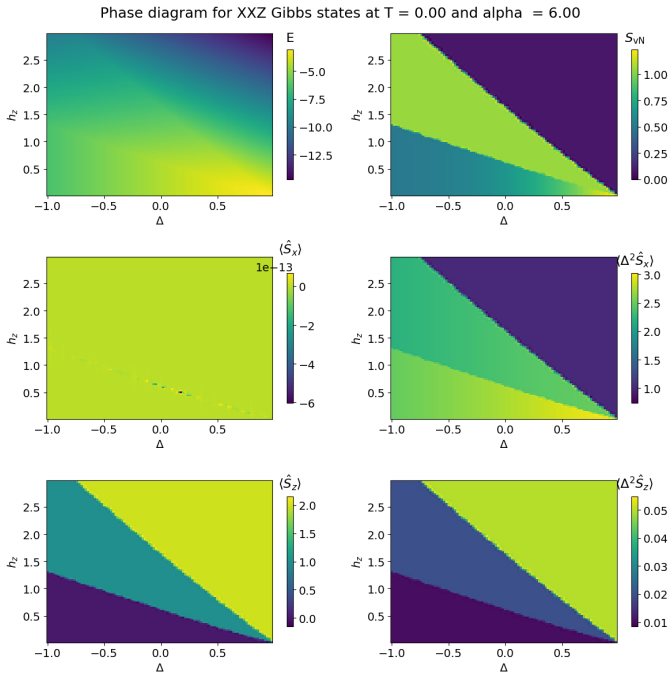

FIG. 3. Energy, entropy, S_z, ΔS_z, S_x and ΔS_x with respect to Δ, h_z for $\alpha = 6$.

find two principal components in our three datasets, and we show two observables with respect to them in the Figs 4, 5 and 6.

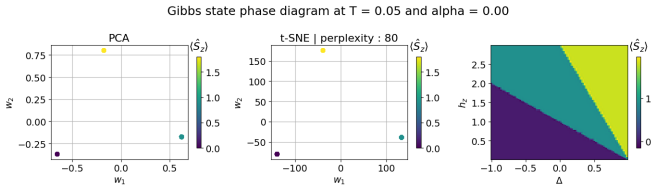

FIG. 4. Analysis of total spin S_z with respect to principal components found through PCA and t-SNE for $\alpha = 0$. The third graph shows S_z with respect to Δ and h_z to compare.

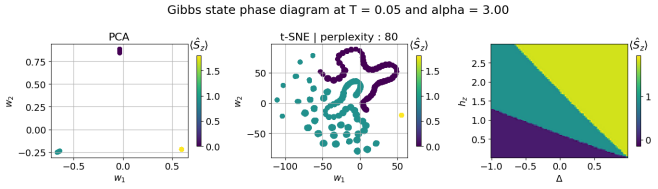

FIG. 5. Analysis of total spin S_z with respect to principal components found through PCA and t-SNE for $\alpha = 3$. The third graph shows S_z with respect to Δ and h_z to compare.

In this analysis, we can see clusterization the spin S_z. As it can be seen in the mentioned figures, different values of S_z, that correspond to different phases, are well separated through the principal components in both the PCA and t-SNE analysis. Similar results are obtained for all the values of α with respect to S_z, however, in the Fig 6, we show the entropy to see that there is no

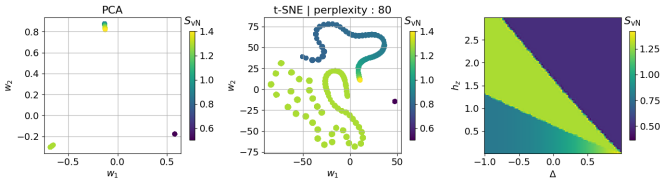

FIG. 6. Analysis of bipartite Von Neumann entropy with respect to principal components found through PCA and t-SNE for $\alpha = 6$. The third graph shows S_{vN} with respect to Δ and h_z to compare.

clusterization in that observable.

Now, after studying the results of our autoencoder with different number of random samples we will show the most relevant results. Let's observe the results with a training of 10 samples in the figures 8 and 7. If we

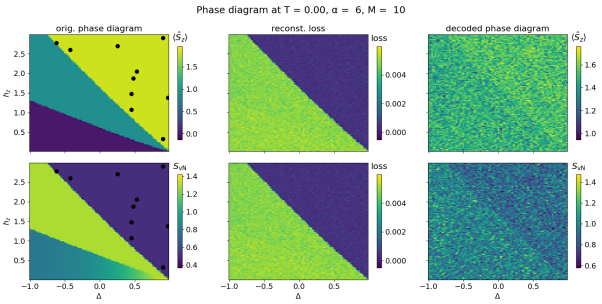

FIG. 7. Phase diagram in S_z and in the entropy with respect to Δ and h_z. It's also shown the reconstruction with the autoencoder. The reconstruction loss function is shown in the middle. They are taken 10 samples from just one phase, as it can be observed in the black points.

compare these figures, we can see the importance of choosing the points from all the phases of the diagram. The neural network needs samples from every phase to identify all the different patterns and reconstruct the whole diagram. That's why we can observe a better reconstruction in the first figure, though the absolute values of the reconstructed diagram can be improved. We can see also a comparison between the decoded density matrix and the exact one in the figures 10 and 9, proving that the density matrix is better reconstructed when samples from different phases are given.

Thus, if we obtain these results with just 10 samples, it would be interesnting to increase the number of samples to see how many of them are necessary to obtain a accurate reconstruction. Although with 20 samples we obtain very accurate results for the phase diagram in S_z, we do not achieve the same accuracy in the entropy diagram. This is because, as a result of the random choice, only two points are taken from the phase at the bottom of the diagram. However, if we increase the number of samples to 50, that part is also

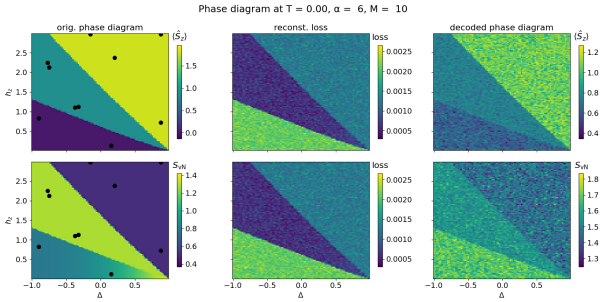

FIG. 8. Phase diagram in S_z, and in the entropy with respect to Δ and h_z. It's also shown the reconstruction loss function and the distance between the decoded density matrix and the exact one in terms of Hilbert-Schmidt distance and the trace distance. The samples are the black points, from the whole phase diagram.

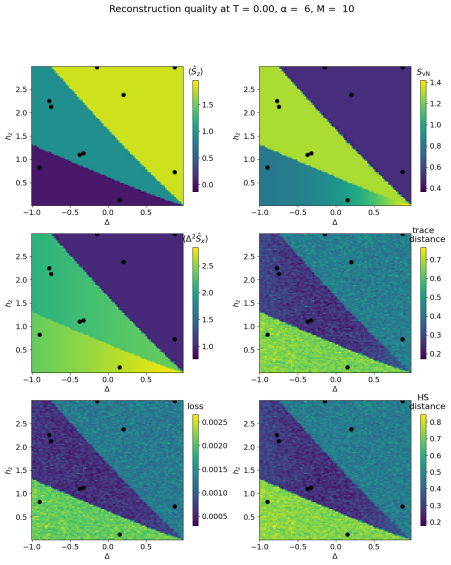

FIG. 9. Phase diagram in S_z, ΔS_z and in the entropy with respect to Δ and h_z. It's also shown the reconstruction loss function and the distance between the decoded density matrix and the exact one in terms of Hilbert-Schmidt distance and the trace distance. The samples are the black points, from the whole diagram.

covered and we obtain the results of the Fig 11. As it may be observed, the reconstruction loss function starts to be vanishingly small. Similar results are obtained for the density between density matrices and they can be verified in the code.

As expected, the quality of the reconstruction improves when we increase the number of samples to 200 [3]. However, such calculations are not necessary with this autoencoder. This highlights the importance of defining a well-designed unsupervised algorithm for phase detection.
As the last check, we will provide the analysis of the PCA latent space of the autoencoder in the Fig 12, where

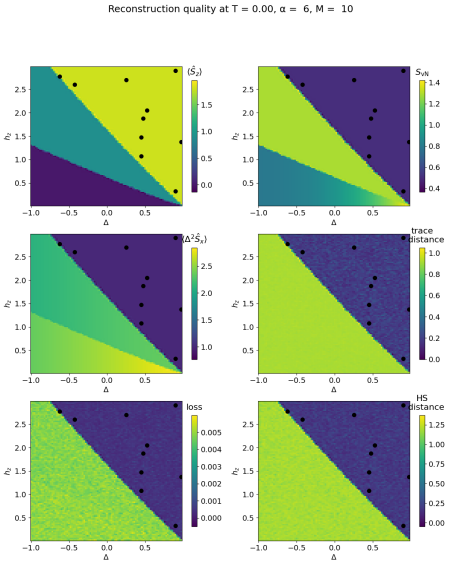

FIG. 10. Phase diagram in S_z, ΔS_z and in the entropy with respect to Δ and h_z. It's also shown the reconstruction loss function and the distance between the decoded density matrix and the exact one in terms of Hilbert-Schmidt distance and the trace distance. The samples are the black points, from just one phase of the diagram.

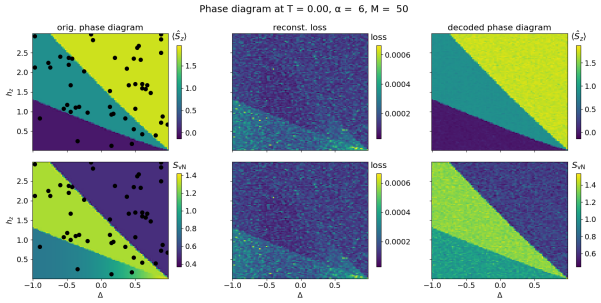

FIG. 11. Phase diagram in S_z and in the entropy with respect to Δ and h_z. It's also shown the reconstruction with the autoencoder. The reconstruction loss function is shown in the middle. The 50 samples are taken all the phase diagram, as it can be observed in the black points.

we show the total magnetization, S_z, and its variance, with respect to the two most relevant components of the latent space vectors, that accumulate the 95% of the variance - a detail that can be also verified in the code.

B. Temperature $T > 0K$

If we obtain the phase diagrams for non null temperatures, we observe that the phases start to be less defined. It can be observed in the Fig 13. In fact, we observe in the PCA analysis that there isn't clusterization for neither S_z nor the entropy, as it may be observed in the Fig 14. Since the phases are less distinguished, our PCA analysis doesn't give separation between points with

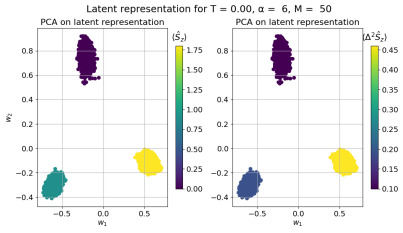

FIG. 12. Latent space analysis: The most relevant components have been found through PCA analysis, w_1 and w_2, that explain the 95% of the variance of the latent space. S_z and ΔS_z are shown with respect to them.

different values of S_z and the entropy.

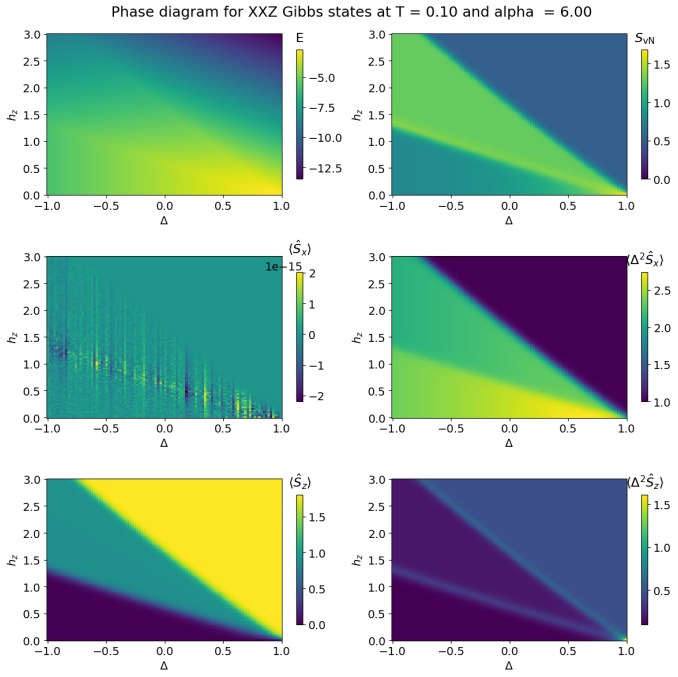

FIG. 13. Energy, entropy, S_z, ΔS_z, S_x and ΔS_x with respect to Δ, h_z for $\alpha = 6$ and $T > 0$.

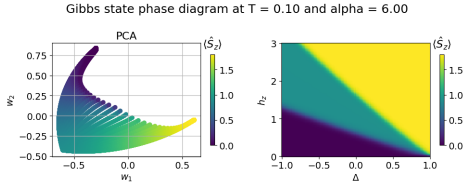

FIG. 14. Analysis of total spin S_z with respect to principal components found through PCA for $\alpha = 6$ and $T = 0.1K$. The second graph shows S_z with respect to Δ and h_z to compare.

If we continuously increase the temperature, we observe that the phase transition completely disappears, meaning that we no longer see a diverging gradient in either the entropy or S_z. It can be verified in the Fig. 15. On the other hand, we observe similar results in the PCA analysis to those obtained for $T = 0.1K$.

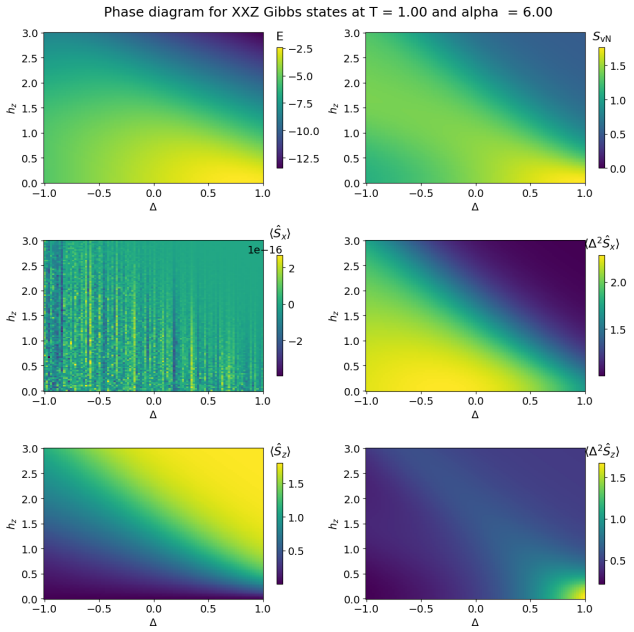

FIG. 15. Energy, entropy, S_z, ΔS_z, S_x and ΔS_x with respect to Δ, h_z for $\alpha = 6$ and $T > 0$.

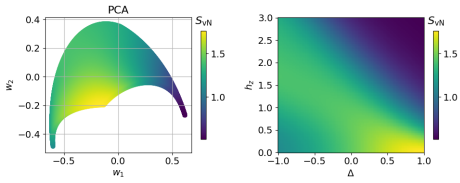

FIG. 16. Analysis of total spin S_z with respect to principal components found through PCA for $\alpha = 6$ and $T = 0.1K$. The second graph shows S_z with respect to Δ and h_z to compare.

Although here are no well-defined phases for $T = 1.0K$, the autoencoder shows a good performance when we take only 20 samples, as it may be observed in the Fig. 17. This result is very interesting, because, with just a small subset of the dataset, the autoencoder has been able to identify the continuous transitions of the diagram.

V. OUTLOOK.

Since the most challenging part of this work has been defining the architecture of the autoencoder and the training strategy, an interesting research direction would be to develop a general method for designing the architecture of an unsupervised algorithm for a generic spin

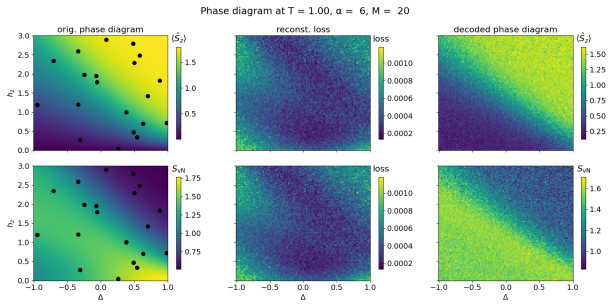

FIG. 17. Phase diagram in S_z and in the entropy with respect to Δ and h_z. It's also shown the reconstruction with the autoencoder. The reconstruction loss function is shown in the middle. The 20 samples are taken from all the dataset, as it can be observed in the black points.

chain of length L. Additionally, we have found that the optimal batch size is 5, though this value may vary depending on the specific problem.

On the other hand, since I do not yet have a background in Condensed Matter Physics, gaining deeper insights into this field would allow for a more thorough analysis.

[1] F Bagarello, C Trapani, and S Triolo. Gibbs states defined by biorthogonal sequences. *Journal of Physics A: Mathematical and Theoretical*, 49(40):405202, September 2016.

[2] P. M. Chaikin and T. C. Lubensky. *Contents*, page vii–xvi. Cambridge University Press, 1995.

[3] Daniel Del Canto de Guzmán. Code for unsupervised learning in google colab. `https://colab.research.google.com/drive/1G4O5ypsMbwIWrijfGuhf9juqQw0tnNJj?pli=1`, 2025.

[4] Fabio Franchini. *An Introduction to Integrable Techniques for One-Dimensional Quantum Systems*. Springer International Publishing, 2017.

[5] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[6] Karl Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[7] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008.

[8] Sebastian J. Wetzel. Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. *Physical Review E*, 96(2), August 2017.