2019

Principal component analysis of the magnetic transition in the three-dimensional Fermi-Hubbard model

Ehsan Khatami, *San Jose State University*

bepress™

Principal component analysis of the magnetic transition in the three-dimensional Fermi-Hubbard model

Ehsan Khatami

Department of Physics and Astronomy, San Jose State University, San Jose, CA 95192, USA

Abstract. Machine learning techniques have been widely used in the study of strongly correlated systems in recent years. Here, we review some applications to classical and quantum many-body systems and present results from an unsupervised machine learning technique, the principal component analysis, employed to identify the finite-temperature phase transition of the three-dimensional Fermi-Hubbard model to the antiferromagnetically ordered state. We find that this linear method can capture the phase transition as well as other more complicated and nonlinear counterparts.

1. Background and Introduction

We humans and other animals owe our often effortless recognition of objects, faces, voices and other cues to the highly complex neural network in our brain. While we cannot exactly model the tens of billions of neurons, each with a few thousand connections to other neurons, in our brains using the current computer technology, biological neural networks have certainly been the model and inspiration for the development of deep artificial neural networks (ANN), and more broadly, machine learning techniques. Their practical applications, mostly in the technology industry, such as Google's deepmind [1], Apple iPhone's voice and face recognition, Netflix's narrowing down of movie preferences for its customers, are often designed to perform specific tasks.

Machine learning paradigms in science have also been around for decades, for example, in the studies of connections between neural networks and spin glass models and how they can be used as memories [2, 3, 4, 5, 6, 7, 8]. However, it was the development of efficient algorithms for training deep networks that led to the rise of popularity of machine learning and its physical applications in recent years. For example, the seminal paper by Carrasquilla and Melko [9], which focused on learning algorithms for distinguishing phases of matter, or by Carleo and Troyer [10], which focused on using ANNs as superior variational ansatz for obtaining the ground state wavefunction of quantum many-body systems, both in 2016 paved the way for hundreds of other studies in the field of statistical mechanics and quantum many-body physics that followed.

In Ref. [9], the authors replace images of animals, objects, or handwritten digits, often used to showcase the power of novel machine learning algorithms in the industry, by Ising "images"; configurations of binary spins from Monte Carlo (MC) simulations of the Ising model on two-dimensional (2D) geometries. In one case, a simple fully-connected feed-forward ANN was trained, using these images at different temperatures, to identify the location of the critical

temperature of the 2D Ising model. They not only showed that their network could locate the phase transition using images it had not seen during the training, but that the trained network could also be used to identify the same transition on other geometries (e.g., the triangular lattice) with over 99% accuracy. They showed that the network output in this case is essentially the magnetization.

In studies that followed immediately [11, 12], the question explored was whether similar techniques can be applied to quantum mechanical models, more specifically, to fermionic lattice models. For quantum lattice models, such as the Fermi-Hubbard model [13, 14], and their quantum Monte Carlo (QMC) treatments [15], the question is, what type of "images" can be used for training? Unlike for the Ising model, where MC simulations sample directly over configurations of the constituents (spin-1/2), the configurations sampled over in QMC simulations of fermionic models are auxiliary fields and involve an "imaginary time" dimension. In Ref. [11], it was shown that ANNs cannot be trained to identify the quantum phase transition between a semi-metal and an antiferromagnetic Mott insulator in the honeycomb lattice Hubbard model using these fields directly, at least not when geometrical information was not preserved. Instead, using calculated Green's functions as inputs for ANNs, they showed that the networks can learn to accurately predict the location of the transition even if training has been performed using data far from the transition point. In Ref. [12], it was shown that raw auxiliary field configurations can in fact be used by ANNs to learn to identify the finite-temperature transition of the three-dimensional (3D) Fermi-Hubbard model from a paramagnetic to an antiferromagnetic phase using ANN architectures that respect the geometry of the model and involve so-called convolutional layers for extracting local features. Moreover, it was shown that the knowledge learned during training can be transferred to locate the same phase transition in a model with a different strength for fermionic interactions or at incommensurate fillings.

The question of machine-assisted phase discovery was also explored early on using the so-called *unsupervised* machine learning algorithms for both classical [16, 9, 17, 18, 19] and quantum systems [20, 21, 22]. For this class of methods, labels for the data (e.g., whether or not they belong to the ordered phase) need not be provided by a "supervisor" during the training process. Instead, unsupervised algorithms are designed to learn about similarities and dissimilarities among the data on their own during the training and categorize data based on that information. The categorization is meant to reduce the huge dimensionality of the configuration space (e.g., 2^N for the Ising model, where N is the number of spins) to a few "latent" dimensions where the similarity of configurations manifests itself in the clustering of corresponding data points. Such visualizations can be clearly seen in Refs. [16, 9, 18] and other studies. One can then classify a new configuration that has not been involved in the dimension reduction routine by projecting it onto the latent space. A summary of these unsupervised techniques is provided in Ref. [21].

Here, the question is again, which techniques can see through quantum fluctuations and effectively classify auxiliary configurations obtained during QMC simulations? And, how does the classification depend on the nature of the phases under investigation? Features observed in the reduced dimensional space can provide insights into properties and symmetries of the system as a tuning parameter like the temperature is varied, and sudden changes can hint at phase transitions. Several nonlinear algorithms have already been shown to exhibit nontrivial features in dimensionally-reduced QMC data for the 2D and 3D Hubbard models [21]. However, a linear technique called the principal component analysis (PCA) [23] has not shown a consistent performance. For example, no discernible features could be found for the quantum many-body localization transition of the random-field Heisenberg model [20], whereas transitions to certain magnetic or charge density wave phases could be captured for the Fermi-Hubbard model [22].

Here, we consider the case of the half-filled 3D Fermi-Hubbard model and its finite-temperature transition to the long-range antiferromagnetic phase and examine to what extent this transition can be captured using the PCA. As a benchmark, we first reproduce results from

the PCA for the 2D Ising model.

2. Models
2.1. 2D Ising Model
The energy of the ferromagnetic Ising model on the square lattice is written as

$$E = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \tag{1}$$

where $\sigma_i = \pm 1$, $\langle ij \rangle$ denotes that i and j are nearest neighbors, and J is the strength of the coupling, which we set to 1 as the unit of energy for this model. The critical temperature for this model is $T_c = \frac{2}{\ln(1+\sqrt{2})} = 2.27$.

We use a Markov chain MC algorithm with the Metropolis algorithm to obtain 10,000 configurations, 100 configurations per each of the 100 temperature values on a uniform grid between 1.5 and 4.0 for a 20×20 system. Starting with a random configuration, in this algorithm, a site i is picked at random and a move $\sigma_i \to -\sigma_i$ is proposed. The move is accepted with the probability $\min[1, e^{-\beta \Delta E}]$, where β is the inverse temperature, and ΔE is the total energy after the spin slip minus that before the flip. We employ a simulated annealing algorithm in which the temperature is reduced adiabatically using an exponential protocol during the simulations. This will allow the system to settle into one of the two degenerate ferromagnetic ground states that break the time reversal symmetry (all up spins or all down spins).

2.2. 3D Fermi-Hubbard Model
The Fermi-Hubbard Hamiltonians is expressed as

$$H = -t \sum_{\langle ij \rangle \sigma} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \tag{2}$$

where $c_{i\sigma}$ ($c_{i\sigma}^\dagger$) annihilates (creates) a fermion with spin σ on site i, $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$ is the number operator, U is the onsite Coulomb interaction, and t is the hopping amplitude, which we take to be unity. We consider the model with an average of one fermion per lattice site (half-filled limit) on a cubic lattice where it has been shown that the critical temperature for the transition to the antiferromagnetic phase is maximal around $T_N = 0.35$ for $U \sim 9$ [24, 25].

We use the determinant QMC algorithm [15, 26, 27] on a $N = 4 \times 4 \times 4$ cluster and 32 temperatures ranging from 0.10 to 0.65. As mentioned above, the auxiliary spin configurations here live in a $D + 1$ dimensional space, where D is the physical dimensionality (here 3). The Metropolis algorithm, which is performed separately for each temperature, proceeds with the same Markov chain as for the Ising model. However, the acceptance probability for a spin flip has to be computed using "fermion determinants" [15, 26]. In Fig. 1, we show a $D + 1$ dimensional spin configuration for $D = 1$. The vertical axis β is discretized with a spacing $\Delta \tau$. In order to keep the number of auxiliary spin variables ("features" in our PCA analysis below) fixed, we fix the number of imaginary time slices to 200. As a result, $\Delta \tau$ varies between 0.050 and 0.008 in our temperature range, rendering systematic errors negligible. Details of the method and other parameters we have used can be found in Ref. [12].

3. Principal Component Analysis
PCA is essentially a transformation of the coordinate system in a high-dimensional space to a set of principal axes determined through diagonalizing the covariance matrix of the data. The eigenvectors corresponding to the largest eigenvalues of the covariance matrix indicate the principal directions along which the data have the largest variance. After the PCA is performed we often project the data onto the two principal directions with the largest eigenvalues.
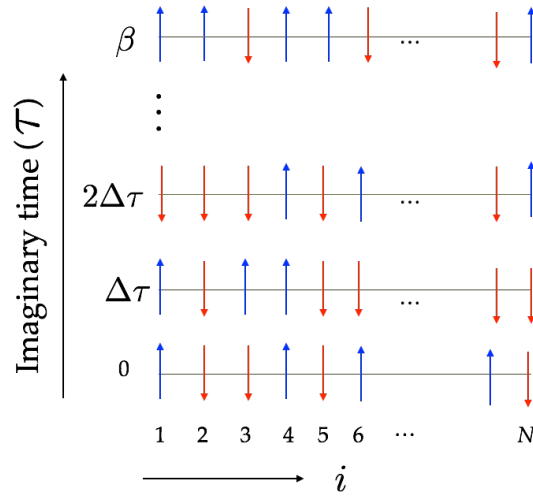
Figure 1. A random configuration of auxiliary spins during the QMC simulation of the one-dimensional Fermi-Hubbard model with a system of N sites at an inverse temperature of β. The horizontal axis represents the spatial dimension and the vertical axis represents the imaginary time dimension that extends from 0 to β. The latter is discretized with the step $\Delta\tau$.

3.1. The Algorithm
The PCA algorithm is as follows:

(i) Center the data. Each data point is a collection of "features", or components in the original space. For example, an Ising configuration can be thought of as a data point in a N dimensional space, where N is the system size and components are σ_i. In this step, the mean value of every feature among all the data is subtracted.

(ii) Create the matrix of data X. Each data point forms a row, columns represent features,

(iii) Compute the covariance matrix of data: $X^T X$.

(iv) Compute eigenvalues λ_n and normalized eigenvectors v_n of the covariance matrix. Eigenvalues represent the variance of each principal axis.

(v) Project data onto the eigenvectors: $P_n = X \cdot v_n$. These are the principal components.

4. Results
4.1. 2D Ising Model
In Fig. 2, we show results after applying the above PCA algorithm to the MC Ising configurations. The size of the covariance matrix, and therefore, the number of principal components is $N = 400$. The left panel in this figure shows the eigenvalues of the covariance matrix, where the largest eigenvalue is an order of magnitude larger than the rest. This implies that the data are mostly spread out along the first principal direction. We note that we perform the PCA on data from all the temperatures at once as opposed to a separate PCA calculation for every temperature.

　　Projecting the configurations onto the principal directions corresponding to the largest two eigenvalues of the covariance matrix of data, we can visualize any patterns developing as the temperature dips below T_c. In the right panel of Fig. 2, we plot the data from all temperatures in the space of the first two principal components. The color of symbols denotes the temperature indicated on the color bar. We observe a well-defined disk of data points corresponding to high temperatures above T_c. As the temperature decreases to near the critical value, the data spread out in both the P_1 and P_2 directions. At lower temperatures below T_c, we find that an overwhelming number of data points gather in a cluster at large P_1. We infer that these data
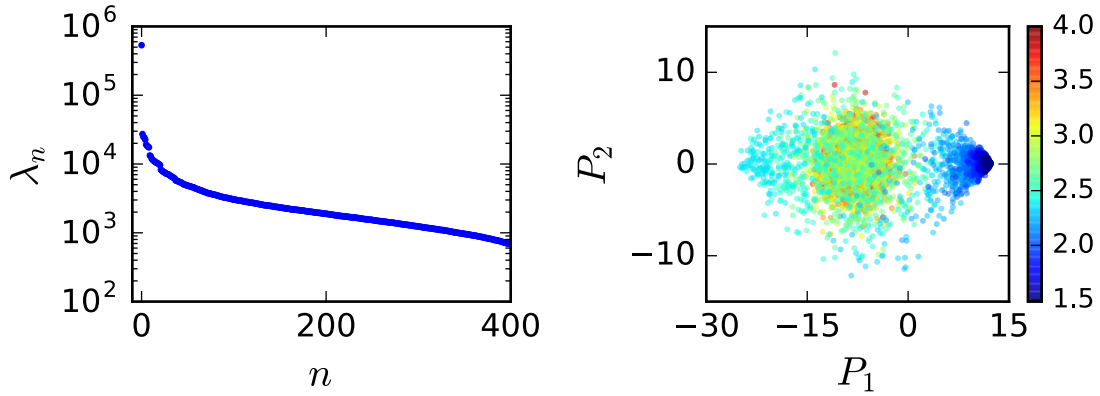
Figure 2. PCA applied to the 2D Ising model. Left: Eigenvalues P_i of the covariance matrix of spin configurations on a 20×20 square lattice. We find that the second eigenvalue is more than an order of magnitude smaller than the first eigenvalue. Right: The projection of data to the space of first two principal components. Colors denote the temperature. The high-temperature disk of data points suddenly changes to a cluster for low-temperature data points at larger P_1 as one crosses the critical temperature.

points correspond to the ordered phase. We point out that the reason for observing a single low-temperature cluster as opposed to two, e.g., as in Refs. [16, 9, 18], is that we have used simulated annealing which causes the system to gradually settle in one of the two degenerate ferromagnetic ground states of the model, as opposed to the possibility of systems at close temperatures falling into phases with magnetizations of opposite sign.

4.2. 3D Fermi-Hubbard Model

We now turn to the quantum model and apply the same technique to data obtained from QMC for the auxiliary spin configurations of the 3D Fermi-Hubbard model. Figure 3 summarizes our findings. In this case, the dimension of the covariance matrix (the number of features) is $N \times 200 = 12,800$. In the left panel of Fig. 3, we show the first 3,000 eigenvalues of the covariance matrix. The largest is more than an order of magnitude larger than the rest of the eigenvalues, again suggesting that there is mostly one special direction in the high-dimensional configuration space where the data have a relatively large variance. Finding the first two principal components allows us to project the data in a two-dimensional space and visualize them.

The results are shown in the right panel of Fig. 3. There, the evolution of clustering of data points as a function of temperature is less dramatic than in the Ising model across the critical temperature. As the temperature decreases the data simply spread out along P_1 and they do so in a symmetric way; the number of points with $P_1 > 0$ does not seem to be significantly different form those with $P_1 < 0$. The smoother evolution is partly a reflection of the fact that the determinant QMC used in this work is formulated using the fermion spins represented in the z basis, whereas the model has spin rotational symmetry and does not have a preferred direction; the antiferromagnetic ordering below T_N can happen in the xy plane for some instances, which will manifest itself essentially as randomness in the configuration of (auxiliary) spins represented in the z basis. We point out that although the auxiliary spins do not exactly represent particle spins, it can be shown that their correlation functions are proportional to the correlation functions of the actual fermion spins [28, 29].

We believe that another contributor to the lack of distinct low-temperature clusters in the Hubbard data shown in Fig. 3 is the existence of quantum fluctuations in the system, i.e., preventing a perfectly anti-aligned spin configuration in any spin direction from being the exact
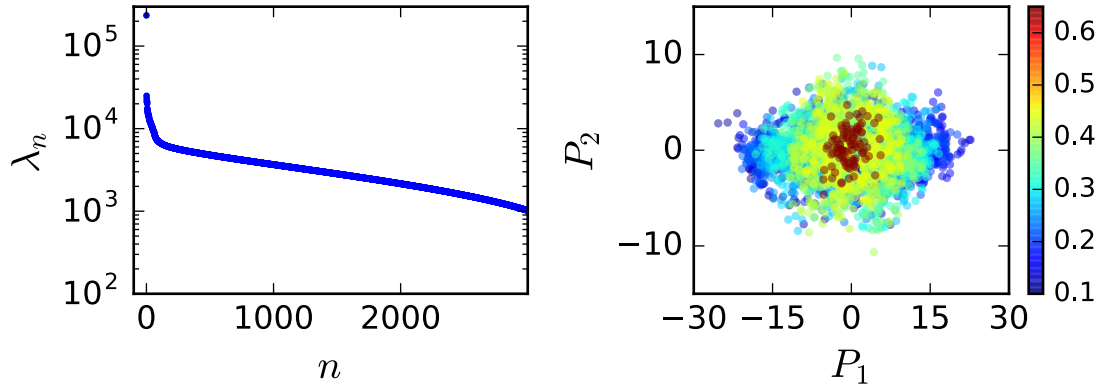
Figure 3. Same as Fig. 2 but for the half-filled 3D Hubbard model with $U = 9$. We show only the first 3,000 eigenvalues in the left panel. Similarly to the Ising case, the largest eigenvalue is at least an order of magnitude larger than the rest. Here, the evolution of clustering of data points as a function of temperature is less dramatic than in the Ising model across the critical temperature. As the temperature decreases the data points simply spread out along P_1.

ground state of the model, unlike for the classical model. Nevertheless, it has been shown that indicators that quantify the transformation of the distribution of data in the reduced-dimensional space as the temperature is varied can represent physically meaningful properties such as the magnetic structure factor [21]. The distribution in the right panel of Fig. 3 obtained through a linear machine learning algorithm seems to be as useful in gaining insight into the phase transition of the model as those obtained through more sophisticated and nonlinear methods applied to the same data [21].

5. Summary
Considering two classical and quantum many-body systems, both exhibiting a continuous finite-temperature phase transition to magnetically ordered phases, we study how a linear machine learning technique, the principal component analysis, can provide clues for the location and nature of the transitions using configurations generated during MC or QMC simulations of the model and without the calculation of any conventional properties. We find that, similar to what was found in past studies with the 2D Ising model, the PCA can signal the transition for this model with a high degree of accuracy. Around T_c, we observe a sudden shift in the distribution of dimensionally reduced Ising configurations in the space of the first two principal components from a central disk at high temperatures to a different cluster at $T \lesssim T_c$. This is consistent with the use of simulated annealing in our MC simulations. We find that the picture is different for the data corresponding to the 3D Fermi-Hubbard model. Due to the specific basis chosen for the QMC simulations and the existence of quantum fluctuations, in that case, the dimensionally-reduced data simply spread out from a central region as the temperature is decreased. However, despite that, quantifying the spread can likely results in reliable indicators for the phase transition.

6. Acknowledgments

7. References
[1] https://deepmind.com/.

[2] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.

[3] Daniel J. Amit, Hanoch Gutfreund, and H. Sompolinsky. Spin-glass models of neural networks. *Phys. Rev. A*, 32:1007–1018, Aug 1985.

[4] G. Toulouse, S. Dehaene, and J. P. Changeux. Spin glass model of learning by selection. *Proceedings of the National Academy of Sciences*, 83(6):1695–1698, 1986.

[5] N. Parga and M. A. Virasoro. The ultrametric organization of memories in a neural network. *J. Phys. France*, 47(11):1857–1864, 1986.

[6] G. Parisi. Asymmetric neural networks and the process of learning. *Journal of Physics A: Mathematical and General*, 19(11):L675, 1986.

[7] M. Mézard, J. P. Nadal, and G. Toulouse. Solvable models of working memories. *J. Phys. France*, 47(9):1457–1462, 1986.

[8] R. Scalettar and A. Zee. Perception of left and right by a feed forward net. *Biological Cybernetics*, 58(3):193–201, Feb 1988.

[9] Juan Carrasquilla and Roger G. Melko. Machine learning phases of matter. *Nat Phys*, 13(5):431–434, May 2017. Letter.

[10] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[11] Peter Broecker, Juan Carrasquilla, Roger G. Melko, and Simon Trebst. Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific Reports*, 7(1):8823, 2017.

[12] Kelvin Ch'ng, Juan Carrasquilla, Roger G. Melko, and Ehsan Khatami. Machine learning phases of strongly correlated fermions. *Phys. Rev. X*, 7:031038, Aug 2017.

[13] N. F. Mott. The basis of the electron theory of metals, with special reference to the transition metals. *Proceedings of the Physical Society. Section A*, 62(7):416, 1949.

[14] J. Hubbard. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 276(1365):238–257, 1963.

[15] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar. Monte carlo calculations of coupled boson-fermion systems. i. *Phys. Rev. D*, 24(8):2278–2286, Oct 1981.

[16] Lei Wang. Discovering phase transitions with unsupervised learning. *Phys. Rev. B*, 94:195105, Nov 2016.

[17] Sebastian J. Wetzel. Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. *Phys. Rev. E*, 96:022140, Aug 2017.

[18] Wenjian Hu, Rajiv R. P. Singh, and Richard T. Scalettar. Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination. *Phys. Rev. E*, 95:062122, Jun 2017.

[19] Ce Wang and Hui Zhai. Machine learning of frustrated classical spin models. i. principal component analysis. *Phys. Rev. B*, 96:144432, Oct 2017.

[20] Evert P. L. van Nieuwenburg, Ye-Hua Liu, and Sebastian D. Huber. Learning phase transitions by confusion. *Nat Phys*, 13(5):435–439, May 2017. Letter.

[21] Kelvin Ch'ng, Nick Vazquez, and Ehsan Khatami. Unsupervised machine learning account of magnetic transitions in the Hubbard model. *Phys. Rev. E*, 97:013306, Jan 2018.

[22] Natanael C. Costa, Wenjian Hu, Z. J. Bai, Richard T. Scalettar, and Rajiv R. P. Singh. Principal component analysis for fermionic critical points. *Phys. Rev. B*, 96:195138, Nov 2017.

[23] I. Jolliffe. *Principal component analysis*. John Wiley and Sons, Ltd, 2002.

[24] R. Staudt, M. Dzierzawa, and A. Muramatsu. Phase diagram of the three-dimensional Hubbard model at half filling. *The European Physical Journal B - Condensed Matter and Complex Systems*, 17(3):411–415, 2000.

[25] Ehsan Khatami. Three-dimensional Hubbard model in the thermodynamic limit. *Phys. Rev. B*, 94:125114, Sep 2016.

[26] S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar. Numerical study of the two-dimensional Hubbard model. *Phys. Rev. B*, 40:506–516, Jul 1989.

[27] *Codes available at http://quest.ucdavis.edu/ and https://code.google.com/p/quest-qmc.*

[28] J. E. Hirsch. Discrete Hubbard-Stratonovich transformation for fermion lattice models. *Phys. Rev. B*, 28:4059–4061, Oct 1983.

[29] J. E. Hirsch. Connection between world-line and determinantal functional-integral formulations of the Hubbard model. *Phys. Rev. B*, 34:3216–3220, Sep 1986.