

AUTOENCODER-AIDED ANALYSIS OF LOW-DIMENSIONAL HILBERT SPACES

G. Žlabys, M. Račiūnas, and E. Anisimovas

Institute of Theoretical Physics and Astronomy, Vilnius University, Saulėtekio 3, 10257 Vilnius, Lithuania

Emails: giedrius.zlabys@tfai.vu.lt, mantas.raciunas@tfai.vu.lt, egidijus.anisimovas@ff.vu.lt

Received 14 September 2021; accepted 23 September 2021

We study the applicability of feedforward autoencoders in determining the ground state of a quantum system from a noisy signal provided in a form of random superpositions sampled from a low-dimensional subspace of the system's Hilbert space. The proposed scheme relies on a minimum set of assumptions: the presence of a finite number of orthogonal states in the samples and a weak statistical dominance of the targeted ground state. The provided data is compressed into a two-dimensional feature space and subsequently analyzed to determine the optimal approximation to the true ground state. The scheme is applicable to single- and many-particle quantum systems as well as in the presence of magnetic frustration.

Keywords: feedforward autoencoder, low-dimensional Hilbert spaces, numerical ground-state estimation

PACS: 03.65.Aa, 02.70.Rr, 02.90.+p

1. Introduction

A broad range of numerical simulations in quantum physics are aimed at an approximate determination of the ground state of an interacting quantum many-body system [1]. In many cases of fundamental or practical importance, solutions can potentially enhance our understanding of quantum phases and phase transitions, and lead towards engineering and exploitation [2, 3]. The source of the complexity and ensuing numerical intractability of the problem is an enormous dimension of the associated Hilbert space (HS) [4, 5]. Even for discretized problems – i.e. problems defined on a lattice with a countable or even finite set of sites – the dimension of the HS scales exponentially with the number of particles and quickly exceeds the available computational resources.

The design of approximate approaches to the quantum many-body problem has been demonstrating very dynamic and fruitful activity, pro-

ducing numerous schemes and insights [1, 6–8]. The ideas that have been tried to tackle quantum complexity are very diverse and range from rather simple-minded – such as variational ansatzes [9] and truncation of the single particle basis to limit the combinatorics or imaginary-time evolution [10] – to quite intricate schemes that rely on representations based on the entanglement structure (tensor networks and related approaches [4, 11, 12]) and artificial neural network (ANN) quantum states [5, 13–16].

In our work, we take the view that an approximate computation of the ground state using one of many available schemes can be regarded as a random shot at the HS which can be expected to land in its low-dimensional subspace spanned by a few, a few dozen or even a few hundred (but certainly not an exponentially large number) low-energy excited states in addition to the sought ground state. In other words, the resulting approximate wave function is most likely a superposition of a limited

number of low-energy states. This is insured by a general observation that the ground-state energy is obtained with a certain accuracy, and the computational procedure is optimized to make the residual error as small as possible. Therefore, this error defines a window of energies, and only excited states characterized by excitation energies commensurate with this window can contribute to the superposition that defines the approximate wave function of the ground state.

Let us also stress that there may be two classes of the mentioned randomness. Besides the randomness (or rather arbitrariness) introduced by the approximation, e.g. truncation of the basis, choice of the ansatz or identification of the relevant sector of the HS to be explored, there is the random sampling built into the algorithm itself. Such sampling – guided by weighted rejection of energy-increasing moves in the configuration space – insures an efficient exploration of the relevant states and lies at the heart of numerous approaches collected under the umbrella term of Monte-Carlo techniques [7, 17].

We thus focus on the following generic problem: Assume that one has a certain computational black box generating random points from a (relatively) low-dimensional subspace representing the low-energy sector of the HS of a certain physical system. We then ask the question if the obtained results, i.e. a sequence of such random samples, can be further analyzed and the actual ground state can be filtered out with precision superior than that of the provided noisy data. The paramount assumption that we rely on is that the contribution of the ground state in the random superpositions is dominant. More precisely, the weights that describe contributions of the ground- and various excited states are assumed to be random numbers drawn from a distribution that ensures that the mean of the weight corresponding to the ground state is higher than the mean of the weight corresponding to any of the excited states. Analysis of the random data is performed using a feedforward autoencoder [18, 19] – a basic type of ANN [19] whose task is to copy multi-dimensional data from its input to its output through a bottleneck layer of just a few (in most of our work, just two) nodes. This forces the autoencoder to find efficient low-dimensional representations for the significant features of the data and to produce explicit distributions in the feature space

that aids the subsequent analysis and selection. Our results indicate that the feature space can indeed be used to efficiently sample the HS spanned by low-energy excitations and – even more ambitiously – to pinpoint the most successful representation of a true ground state.

2. Motivation

Even though the rest of the paper aims for a general treatment and is not tailored to suit any particular situation, for the sake of orientation we now provide a specific practical example drawn from a numerical simulation of a small optical lattice [20] of 5×5 sites pierced by an artificial flux [21] and populated by 4 bosonic cold atoms. The lattice is schematically sketched in the inset of Fig. 1. The scaled (dimensionless) Hamiltonian reads

$$\hat{H} = -\sum_{\langle ij \rangle} \hat{a}_i^\dagger \hat{a}_j e^{i\gamma_{ij}} + \frac{1}{2} \sum_i \hat{n}_i (\hat{n}_i - 1), \quad (1)$$

and consists of two parts representing, respectively, the hopping transitions on all directed pairs of nearest-neighbour links $\langle ij \rangle$ and on-site interactions between particles. Here, \hat{a}_i^\dagger are the bosonic annihilation (creation) operators on site i and $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$ are the on-site occupation numbers. The Peierls phases γ_{ij} are chosen to ensure a uniform flux through all plaquettes. Note that we take the hopping matrix elements and the interaction strength to be of equal magnitude, which is chosen as the energy unit in Eq. (1). Hence, the model is parameter free.

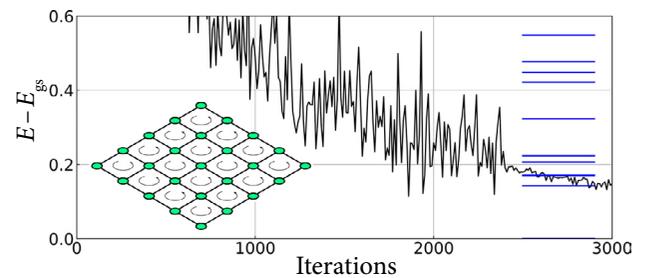


Fig. 1. Typical results of the ANN simulation of an interacting four-particle system on a square lattice of 5×5 sites (shown in the inset) in the quantum Hall regime. Blue (online) lines show the energies of the nine lowest excited states with respect to the ground state. A black line shows the evolution of the estimate of the ground-state energy versus the number of iterations.

This system supports topological band structures [22] and, in the presence of strong particle interactions, is expected to host fractional Chern insulating states [22] in a clean, experimentally accessible and tunable setting. It has been shown [23] that even such small systems provide access to charge fractionalization that characterizes fractional quantum Hall systems [24]. Previous research shows that ANN-inspired algorithms can be quite successful in determining the ground states of model systems [5, 13, 14], and following [13] we apply an ansatz that encodes quantum states as a feedforward ANN and uses the Metropolis sampling algorithm to train the network following the approach of Ref. [15]. However, a fractional quantum Hall system on a finite lattice supports numerous low-energy edge modes and the absence of an energy gap (more precisely, the presence of multiple low-energy excitations) poses a considerable challenge. In a typical situation of 4 particles on a 5×5 lattice with a flux of 0.175 flux quanta per plaquette, the ANN converges to energies close to the first or second excited states. In Fig. 1, we show the evolution of the ground-state energy estimate (black line) as a function of the number of iterations. The blue lines on the right edge depict the 9 lowest excited states of energies (known from the exact diagonalization). All energies are measured relative to the exact ground-state energy, i.e. differences $E - E_{\text{gs}}$ are plotted. In this rather typical case, the ground-state wave function encoded in the trained ANN after 3,000 iterations is a superposition of the true ground state (with the weight 0.765), the first excited state (weight 0.124) and several dozen higher-energy excited states. If better accuracy is needed, the following step is to filter out the ground state by following iterations as a random walk: with the contribution of the ground state systematically present, the excited-state components will behave in an erratic fashion.

3. Model

To make our treatment as general as possible and not tied to specifics of any particular system, we study quantum states $|\psi\rangle$ constructed as superpositions

$$|\psi\rangle = \sum_{j=1}^N \omega_j^{1/2} e^{i\theta_j} |j\rangle, \quad (2)$$

of normalized orthogonal basis vectors $|j\rangle$ that span an N -dimensional HS. The weights ω_j are real

and positive random numbers chosen from a distribution that we describe shortly and the phases θ_j are random and uniformly distributed in the range $[0, 2\pi]$.

The distribution of weights ω_j (satisfying the normalization condition $\sum_i \omega_i = 1$) is modelled by a random division of the unit interval $[0, 1]$, a well-known mathematical model [25]. To achieve this goal, a random number generator with the probability density function $f(x) \propto x^p$ with the domain $[0, 1]$ is sampled to obtain a set of $N-1$ point coordinates. These points then divide the unit interval into N subintervals of various lengths, and these lengths – numbered consecutively with $j = 1, \dots, N$ from the left to the right – act as the weights ω_j of the basis vectors $|j\rangle$. Note that any positive value of p ensures the dominance of the ‘ground’ state $|j = 1\rangle$ and monotonously decaying influences of ‘excited’ states with $j = 2, 3, \dots, N$. We stress that p and N are the only tunable parameters introduced in this model.

The sets of basis vectors $\{|j\rangle\}$ are specific to a particular problem. It turns out, however, that the procedure of analyzing the low-dimensional HS and filtering out an improved ground state is insensitive to the system-specific details, and is universally applicable. For illustrative purposes in this paper we consider relatively small systems where each basis vector is defined on the support of typically $M \approx 500$ sites or configurations, i.e.

$$|j\rangle = \sum_{\mu=1}^M c_{\mu}^{(j)} |\mu\rangle. \quad (3)$$

The sets of the complex-valued coefficients $c_{\mu}^{(j)}$ describe the chosen model and, to assert universality, were generated using four distinctly different ways:

1. The vectors $c^{(j)}$ were copied from the orthonormal vectors obtained from the reduced QR decomposition of a random complex matrix A of dimensions $M \times N$. The real and imaginary parts of each matrix element were generated by uniformly sampling a $[-1, 1]$ interval. Here, $M = 500$ is the number of coefficients $c_{\mu}^{(j)}$ and N is the number of basis vectors. The QR factorization provides mutually orthogonal column vectors which are interpreted as the basis vector set not related to any particular physical system. This is our choice for the data generation below.

2. Eigenstates of a two-dimensional square 23×23 ($M = 529$ sites) lattice system described by the scaled tight-binding Hamiltonian $\hat{H} = \sum_{(ij)} \hat{a}_j^\dagger \hat{a}_i$. In this case, the obtained eigenvectors correspond to a particular single-particle model and have rather regular shapes; e.g. the ground-state wave function is well approximated by a product of two sines, with the exception of a small number of sites close to the boundary.

3. Eigenstates of a frustrated square 23×23 ($M = 529$ sites) lattice. To introduce frustration, the hopping matrix elements of the previously discussed lattice are multiplied by Peierls phases chosen to describe a uniform magnetic flux of $1/8$ of the flux quantum through each lattice plaquette. In this case, the energy spectrum is fractal and low-energy wave functions have irregular shapes.

4. Eigenstates of a frustrated *many-body* system described by the Hamiltonian (1). We take a 3×3 square lattice pierced by a uniform flux as above and filled with 4 bosons (this gives $M = 495$ configurations).

The autoencoder used for data compression and analysis of this work is implemented in Python using the Keras library [26] and consists of nine densely connected layers. The input and output layers are at the edges, the code (bottleneck) layer in the middle, plus three encoding and three decoding ‘hidden’ layers that connect, respectively, the input and output to the code layer. The input and output layers have $2M$ nodes which is twice the number of coefficients (to represent the real and the imaginary parts). The encoding and decoding parts of the autoencoder decrease in node count with each layer while approaching the code layer. The code layer consists of just two nodes, and narrows the information about the input states to two real variables. As the typical number of state-vector coefficients is $M \approx 500$, the node-count structure of the hidden layers is taken to be 100-50-25-2-25-50-100. This choice is sufficient to represent the information about the samples. If the node count is increased, the training time increases as well because of a larger space of parameters to optimize. On the other hand, if the network is too small, it might lose the ability to encode the samples efficiently.

All layers use the hyperbolic-tangent activation function [19]. This is motivated by the fact that the initial data is a normalized wave vector hence

its coefficients can never exceed the active range of $\tanh(\cdot)$. A benefit of this choice is a compactly bounded parameter region of the code layer from which the relevant part of HS is sampled.

Before the random superpositions (2) are provided as samples to the input layer some initial data pre-processing is done. Firstly, the gauge (global phase) of $|\psi^{(j)}\rangle$ is fixed by setting one of its coefficients in the resolution (3) to be real and positive. This procedure improves the separability of the feature-space parameter distribution and does not reduce generality. Secondly, the ANN used in this work has real-valued weights, therefore the sample coefficients are split into the real and imaginary parts. These parts are concatenated into a new vector where all the real parts are followed by the imaginary parts. In this work, for each explored system we provide 1,000 samples and train the autoencoder for 3,000 updates of the network weights. These weights are optimized using the AdaMax algorithm [27] guided by the mean squared error between the input and the output layers.

4. Results

Let us now proceed to numerical simulations of the encoding of low-dimensional Hilbert subspaces. To set the foundation for the forthcoming study of more involved cases, we begin with the simplest two-state quantum-mechanical system. Here, one can firmly rely on geometrically intuitive visualization, the Bloch sphere.

4.1. Sectors of the Bloch sphere

In this subsection, we temporarily step aside from the generation of random weights based on the division of an interval. Instead, we now generate uniformly distributed random points within the sectors of the Bloch sphere delimited by sharp boundaries, as illustrated in the top line of the subplots of Fig. 2. In the first two cases, we restrict the range of the covered polar angles to, respectively, $\theta \in [0, \pi/4]$ and $\theta \in [0, \pi/2]$ thereby focusing on the areas close to the ‘North pole’, i.e. the ground state. In these two situations, the contribution of the ground state in the resulting superposition is dominant. In contrast, in the remaining two cases we sample symmetric distributions where both

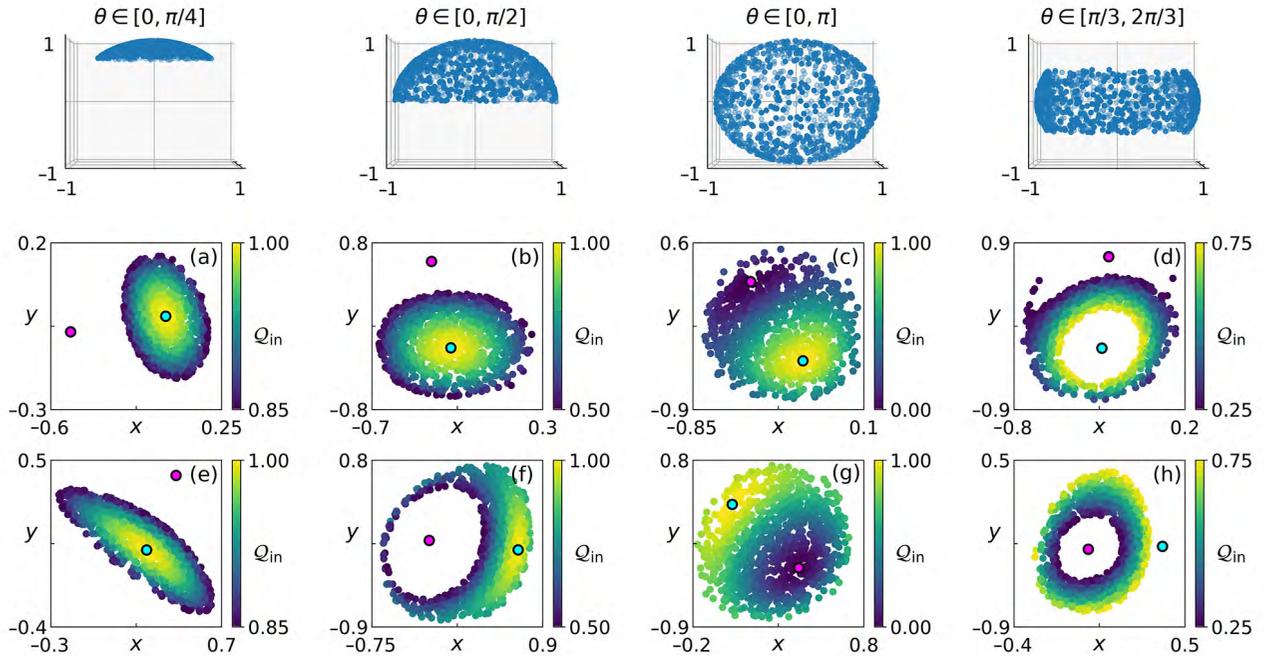


Fig. 2. Feature-space distributions of 1,000 random samples covering the indicated sectors of the Bloch sphere. Data is organized in columns, and in each case two samples corresponding to two choices of the random seed are shown. Colours encode the overlap Q_{in} of the sampled vectors with the true ground state. Note that the range of covered values is specific to each case, as indicated by the colour bars. Cyan (magenta) marker (dark and light in printed version) denotes the ground (excited) state corresponding to the North (South) pole of the Bloch sphere.

states have statistically equal opportunities to contribute. To illustrate the latter situation, we sample either the whole Bloch sphere, $\theta \in [0, \pi]$, or the ‘equatorial’ region, $\theta \in [\pi/3, 2\pi/3]$.

In the leftmost column of Fig. 2, panels (a) and (e), we show the results obtained from sampling the ‘polar’ region $\theta \in [0, \pi/4]$ with 1,000 uniformly distributed random points. The two-dimensional plots show the distribution of the encoded states in the feature plane parametrized by the values of its two nodes, x and y . The colours of the symbols indicate the overlap of each input sample with the ground state, $Q_{\text{in}} = |\langle \psi | \psi_{\text{gs}} \rangle|^2$. Due to the used geometrical restriction these overlaps range from unity to $\cos^2(\pi/8) \approx 0.85$. Here – as well as in the remaining columns – we show two example distributions obtained with two different random seeds. This is a reminder that the obtained feature distributions are statistical samples drawn from a certain ensemble. The encoding of the pure ground (excited) state is shown by a cyan (magenta) circular marker. The results indicate that the autoencoder is doing its job of relevant-feature extraction well: we deal here with

intrinsically two-dimensional data (the polar region of the Bloch sphere) which was provided in the form of vectors living in a space of superficial dimensionality of $2M = 1000$. As it is evident from the elliptical distributions, the autoencoder was able to identify the two intrinsic components and reduce the dimensionality of the data. The true ground state (cyan marker) is promptly positioned in the centre of the elliptical distribution, while the excited state is placed at a safe distance on the side. In Fig. 2(b, f), we see that a clear separation of the two Bloch hemispheres is also achieved, even though (see panel (f)) the manifold corresponding to one of the hemispheres may not always be compact. When the whole Bloch sphere is sampled, the autoencoder faces a more complicated task: the sampled manifold is no longer topologically equivalent to a disk but covers the entire sphere. The results shown in Fig. 2(c, g) again confirm the separability of the data: the two poles of the Bloch sphere are still systematically mapped to two well-separated positions, and the distribution of overlaps is resolved as a smooth gradient covering the whole range $Q_{\text{in}} \in [0, 1]$. With this much

of gained intuition, panels (d) and (h), which correspond to sampling of the equatorial region, are not surprising and clearly follow the same pattern. We note, however, that in the last two cases none of the two eigenstates is statistically dominant. Therefore – if the provided cyan and magenta markers were missing – one may be able to infer their positions from the distributions in the feature plane but, in view of the symmetry, there is no way to tell which is which. Proceeding to the analysis of more interesting cases, let us stress again that for practical reasons one is ultimately interested in the determination of the true ground state, and the only assumption that is built into the proposed method is the weak asymmetry in the data, i.e. some dominance of the ground state with respect to one, several or several dozen excited states.

4.2. General few-level systems

We now turn to the general application of the autoencoder-aided analysis of randomly sampled low-dimensional HS. This is a good place to recapitulate that in order to keep the exposition free from any superfluous assumptions, we have only two parameters in the proposed model. One is the number of contributing states N , and – as discussed in the motivational section – we focus on the regimes where the number of such states ranges from just a few to fifty. The other parameter is the power-law exponent p . In the limit $p = 0$, the contributions from all N participating eigenstates become uniform, and one is left with no information that could help identify the individual states. In contrast, for growing values of p the contribution of the ground state becomes progressively stronger.

We summarize the essential numerical results in Fig. 3 where the two columns correspond to the two choices of $p = 2$ and $p = 10$, and the three rows cover the cases of $N = 2$, $N = 5$ and $N = 50$ states. As above, the two-dimensional plots zoom into the relevant portion of the feature space, and the colours of the symbols represent the overlap Q_{in} of the given sample with the true ground state. Starting with two-state systems, we observe that Fig. 3(a) retains the overall similarity to those presented in the previous subsection despite the fact that instead of sampling the fixed regions of the Bloch sphere we now determine

the weights from the random-interval-division model. With $p = 2$, the weights of the two states can in principle be any from the interval $[0, 1]$, however, their means equal $3/4$ and $1/4$, respectively, for the ground and the excited state. We contrast this situation with that shown in panel (b): here the exponent of the probability distribution function $\sim x^p$ has a much higher value $p = 10$ and the mean weights of the ground and the excited states are, respectively, 0.917 and 0.083 . This results in a more compact clustering of samples around the true ground state.

The next crucial step is to verify that the scheme is still applicable beyond two-level systems. An important observation is that two-level systems do allow a two-dimensional representation of the HS in the form of the Bloch sphere. For richer few-level systems, a possible approach could involve

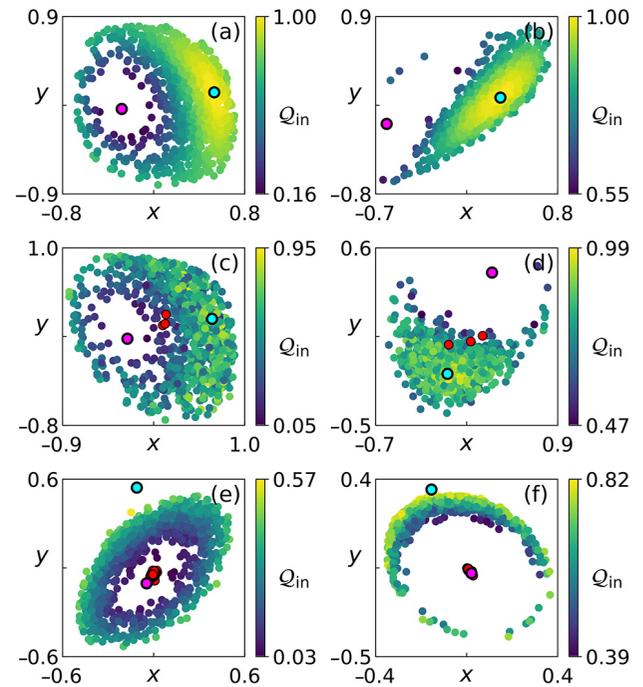


Fig. 3. Feature-space distributions of 1,000 random superpositions of $N = 2$ (top row, panels (a) and (b)), $N = 5$ (middle row, panels (c) and (d)) and $N = 50$ (bottom row, panels (e) and (f)) orthogonal states. The weights are generated from the random-interval-division model with, respectively, $p = 2$ and $p = 10$ in the left/right column. Colours encode the overlap Q_{in} of the sampled vectors with the true ground state. Cyan (magenta) marker (dark and light in printed version, respectively) denotes the ground (first excited) state and intermediate (red online) markers denote all remaining excited states.

a dimensional enlargement of the feature space, however, we did not find such an approach to be particularly promising. Instead, we keep the dimensionality of the feature space fixed, and proceed with the analysis of the resulting two-dimensional distributions. As it turns out, it is possible to treat systems with a vastly different number of states on equal footing thus defining a universal (N -independent) scheme. In addition to the previously discussed independence of the nature of the states in the superposition, this aspect of universality is another important asset.

We found that there are no dramatic differences between few-level systems with different numbers of states (3, 4, 5, ...), and in panels (c) and (d) of Fig. 3 we show the representative results obtained for $N = 5$. In panel (c) the power-law exponent is $p = 2$ and the weights of all five states are still comparable (the mean values of the weights range from 0.534 to 0.074) and all of them can appreciably contribute to the typical superpositions. The range of overlaps (see the colour bar) is thus broad and covers the range [0.05, 0.95]. Nevertheless the autoencoder is able to efficiently separate the data, and the resulting distribution in the feature space is broadly similar to those seen for two-state systems. The presence of three higher excited states is illustrated with three intermediate (red online) symbols indicating the positions of their encodings. We observe that the algorithm was able to clearly distinguish the ground and the first excited state (magenta marker in online version). The remaining excited states are quite well separated from the ground and the first excited state but lumped together – this is the consequence of their small weights that carry too little useful information. In panel (d), the simulation is repeated with $p = 10$. Here, the dominance of the ground state is much stronger (its mean weight is 0.832); this naturally results in an even clearer separation of the ground and the first excited state as well as in stronger clustering of data around the ground state.

Proceeding to larger HS such as $N = 50$ shown in Fig. 3(e, f), we see that the two-dimensional distributions in the feature space still retain familiar features and overall the autoencoder is able to successfully separate the data. However, the increasing number of contributing states leads to a more problematic identification of the positions of the best encoding of the true ground state. In

panel (e) we show the results pertaining to $p = 2$. Here, the dominance of the ground state is rather weak as the mean weight of the ground state is just 0.243. We see that the ground state is clearly separated from the excited states but the excited states are hard to distinguish. However, due to the lack of data with $Q_{\text{in}} > 0.57$, the true ground state is beyond the edge of the cluster. If the position of the true ground state was not indicated and the task was to determine it from the analysis, it would be quite hard to infer its position. In panel (f), the situation is more favourable; here $p = 10$ and the mean weight of the ground state is 0.670. However, the data is not well clustered and forms an extended crescent, therefore, it is again not easy to pinpoint an optimal representation of the ground state on the extended outer edge of the cluster. Nevertheless, in the following subsection we will see that the observed uncertainty of optimal representation does not pose a significant problem.

4.3. Quality of decoding

So far, we concentrated solely on the encoding phase, i.e. the mapping of the input data to the feature space. This provided answers about the algorithm's ability to separate and sort input data by distributing input superpositions of different quality Q_{in} to different regions of the feature plane. Let us now turn to decoding, i.e. the complementary mapping from the feature plane to the reconstructed output states $|\psi_{\text{out}}\rangle$. For the sake of qualitative discussion, we define two additional metrics $Q_{\text{out}} = |\langle \psi_{\text{out}} | \psi_{\text{gs}} \rangle|^2$ and $Q = |\langle \psi | \psi_{\text{out}} \rangle|^2$. They describe, respectively, the quality of the decoding (mapping from the feature space to the output wave function) and overall closeness of the input and the output states.

Figure 4 summarizes the findings. For two-state systems (panels (a) and (b)), the result clearly reflects the intrinsically two-dimensional nature of the data. The input states are encoded in the feature space with one-to-one correspondence, therefore, the overall quality Q shown in panel (a) is very close to the unity. For five- and fifty-level systems the input-to-output mapping is lossy. As shown in panel (c), in the case of $N = 5$, many different input states are encoded as the same pair of feature-space parameters (x, y) , and subsequently reconstructed into a unique output state. The quality of reconstruction

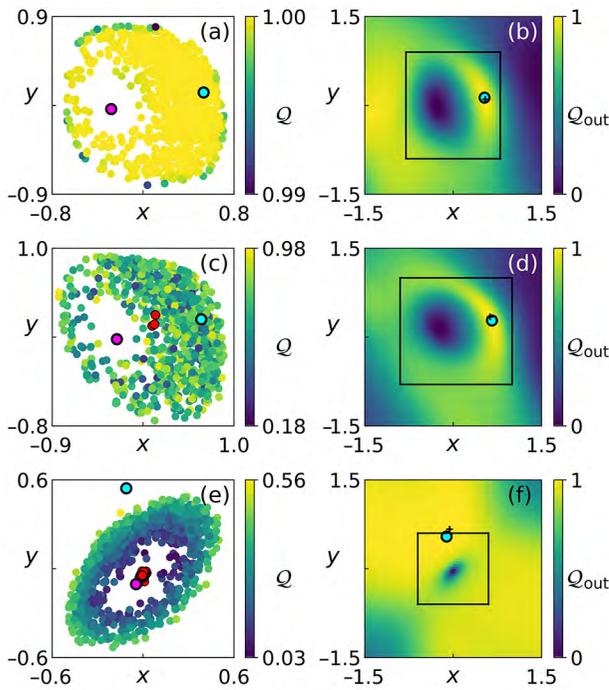


Fig. 4. The quality Q_{out} of mapping from the feature space to the output (or reconstructed) states (right column, panels (b), (d) and (f)) and the overall mapping Q from input to output. The rows, top to bottom, correspond to, respectively, $N = 2$, $N = 5$ and $N = 50$ contributing states and the colours have the usual meaning. The '+'-shaped markers indicate the position of the optimal representation, i.e. the position in the feature that is decoded to the output state that has the largest overlap with the true ground state.

(see panel (d)) covers the entire range of values from 0 to 1, that is, the whole range of output states can be generated by sampling the two-dimensional feature space. In panel (d), the blue region close to the origin $(x, y) = (0, 0)$ accommodates encodings characterized by the nearly-zero overlap between the reconstructed state and the true ground state. On the other hand, the (yellow online) crescent situated close to the geometric centre of input data points signifies the region where encodings are nearly perfectly, with $Q_{\text{out}} \approx 1$, reconstructed to the true ground state. Quite interestingly, in many-level systems such as $N = 50$ depicted in panels (e) and (f) the overall situation is similar, but the light (yellow online) region of nearly perfect reconstructions (see panel (f)) is much larger and covers a significant portion of the feature space. This is fortunate, since the uncertainty of pinpointing the optimal representation for the ground state turns out to matter very little. Even if the optimal

location in the feature space is located with a poor accuracy, its broad neighbourhood accommodates encodings that are nearly as good. This phenomenon can be understood by realizing that the task of separation of the useful signal (the component belonging to the ground state) from the noise (contribution from excited states) actually becomes easier with the growing number of excited states: their superpositions are drawn from spaces with larger dimensionality and consequently are more irregular and can be clearly identified as noise.

The potential improvements may be even more convincing when visualized directly at the wavefunction level. For the sake of a specific example, we focus on a single particle moving on a 23×23 square lattice. The true ground state of this system is described by a product of two sine functions, and the ground-state density distribution is shown in Fig. 5(a). We generate 1,000 superpositions of the $N = 50$ lowest-energy eigenstates whose weights are sampled from the random-interval-division model with $p = 2$. The density distribution of the best – that is, having the largest overlap with the true ground state – of the samples is shown in panel (b). Having trained the autoencoder with these 1,000 samples for the usual 3,000 epochs, we find the pair of feature space parameters which minimize the energy of the system. The density distribution computed from the resulting wave function is plotted in Fig. 5(c) and shows a radical improvement.

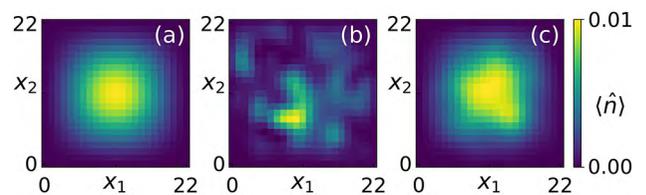


Fig. 5. (a) True ground-state density $\langle \hat{n} \rangle$ of a single particle on a square 23×23 lattice. (b) Density of the sample that has the best overlap with the ground state. (c) Density of the vector generated by the autoencoder which minimizes the energy of the system.

CONCLUSIONS

To summarize, we performed a study of the potential of autoencoder ANN in denoising information of a quantum system's ground state. Using the minimum set of assumptions, such as presence

of a certain number of orthogonal contributing states and a slight dominance of the special ‘ground’ state, we were able to construct a broadly applicable scheme to analyze and systematically sample the minimal (two-dimensional) feature space. The encouraging results motivate further studies based on availability of additional information, e.g. noisy information of the state energy that is typically available in numerical simulations.

Acknowledgements

This work was funded by the European Social Fund under Grant No. 09.3.3-LMT-K-712-01-0051. We thank Artūras Acus for enlightening discussions.

References

- [1] A. Avella and F. Mancini, *Strongly Correlated Systems: Numerical Methods*, Springer Series in Solid-State Sciences (Springer Berlin Heidelberg, 2015), <https://doi.org/10.1007/978-3-662-44133-6>
- [2] P.A. Lee, From high temperature superconductivity to quantum spin liquid: progress in strong correlation physics, *Rep. Prog. Phys.* **71**, 012501 (2007), <https://doi.org/10.1088/0034-4885/71/1/012501>
- [3] K. von Klitzing, Quantum Hall effect: Discovery and application, *Annu. Rev. Condens. Matter Phys.* **8**, 13 (2017), <https://doi.org/10.1146/annurev-conmatphys-031016-025148>
- [4] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys.* **349**, 117 (2014), <https://doi.org/10.1016/j.aop.2014.06.013>
- [5] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017), <https://doi.org/10.1126/science.aag2302>
- [6] R.O. Jones, Density functional theory: Its origins, rise to prominence, and future, *Rev. Mod. Phys.* **87**, 897 (2015), <https://doi.org/10.1103/RevModPhys.87.897>
- [7] W.M.C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, *Rev. Mod. Phys.* **73**, 33 (2001), <https://doi.org/10.1103/RevModPhys.73.33>
- [8] A. Weiße and H. Fehske, *Computational Many-Particle Physics*, Lecture Notes in Physics (Springer Berlin Heidelberg, 2008).
- [9] F. Becca and S. Sorella, *Correlated Models and Wave Functions* (Cambridge University Press, 2017) pp. 3–36.
- [10] L. Lehtovaara, J. Toivanen, and J. Eloranta, Solution of time-independent Schrödinger equation by the imaginary time propagation method, *J. Comput. Phys.* **221**, 148 (2007), <https://doi.org/10.1016/j.jcp.2006.06.006>
- [11] R. Orús, Tensor networks for complex quantum systems, *Nat. Rev. Phys.* **1**, 538 (2019), <https://doi.org/10.1038/s42254-019-0086-7>
- [12] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96 (2011), <https://doi.org/10.1016/j.aop.2010.09.012>
- [13] H. Saito, Solving the Bose–Hubbard model with machine learning, *J. Phys. Soc. Jpn.* **86**, 093001 (2017), <https://doi.org/10.7566/JPSJ.86.093001>
- [14] H. Saito and M. Kato, Solving the Bose–Hubbard model with machine learning, *J. Phys. Soc. Jpn.* **87**, 014001 (2017), <https://doi.org/10.7566/JPSJ.87.014001>
- [15] G. Carleo, K. Choo, D. Hofmann, J.E.T. Smith, T. Westerhout, F. Alet, E.J. Davis, S. Efthymiou, I. Glasser, S.-H. Lin, et al., NetKet: A machine learning toolkit for many-body quantum systems, *SoftwareX* **10**, 100311 (2019), <https://doi.org/10.1016/j.softx.2019.100311>
- [16] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019), <https://doi.org/10.1103/RevModPhys.91.045002>
- [17] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017), <https://doi.org/10.1017/9781316417041>
- [18] G.E. Hinton and R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science* **313**, 504 (2006), <https://doi.org/10.1126/science.1127647>
- [19] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016).

- [20] M. Lewenstein, A. Sanpera, and V. Ahufinger, *Ultracold Atoms in Optical Lattices: Simulating Quantum Many-body Systems* (Oxford University Press, 2012), <https://doi.org/10.1093/acprof:oso/9780199573127.001.0001>
- [21] A. Eckardt, *Colloquium: Atomic quantum gases in periodically driven optical lattices*, *Rev. Mod. Phys.* **89**, 011004 (2017), <https://doi.org/10.1103/RevModPhys.89.011004>, <https://arxiv.org/abs/1606.08041>
- [22] E. J. Bergholtz and Z. Liu, Topological flat band models and fractional Chern insulators, *Intl. J. Mod. Phys. B* **27**, 1330017 (2013), <https://doi.org/10.1142/S021797921330017X>, <https://arxiv.org/abs/1308.0343>
- [23] M. Račiūnas, F. N. Únal, E. Anisimovas, and A. Eckardt, Creating, probing, and manipulating fractionally charged excitations of fractional Chern insulators in optical lattices, *Phys. Rev. A* **98**, 063621 (2018), <https://doi.org/10.1103/PhysRevA.98.063621>
- [24] D. Tong, *Lectures on the Quantum Hall Effect* (2016), <https://arxiv.org/abs/1606.06687>
- [25] R. Perline, Zipf's law, the central limit theorem, and the random division of the unit interval, *Phys. Rev. E* **54**, 220 (1996), <https://doi.org/10.1103/PhysRevE.54.220>
- [26] Keras: *The Python Deep Learning Library*, <https://keras.io>
- [27] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization* (2014), <https://arxiv.org/abs/1412.6980>

NEDIDELIO MATMENŲ SKAIČIAUS HILBERTO ERDVĖS ANALIZĖ PASITĖLKIANČIAMS AUTOENKODERĖM

G. Žlabys, M. Račiūnas, E. Anisimovas

Vilniaus universiteto Teorinės fizikos ir astronomijos institutas, Vilnius, Lietuva

Santrauka

Straipsnyje aprašomas tiesioginio sklaidimo autoenkoderių pritaikomumo siekiant patikimai nustatyti kvantinės sistemos pagrindinę būseną iš triukšmingo signalo, kurį sudaro seka atsitiktinių superpozicijų, paimtų iš sistemos Hilberto erdvės nedidelio matmenų skaičiaus poerdvio, tyrimas. Siūloma schema remiasi tik minimaliomis prielaidomis: (i) laikoma, kad imties elementus sudaro baigtinio skaičiaus ortogonalinių būsenų superpozicijos ir (ii) yra silpnas statistinis ieškomos pagrindinės būsenos dominavimas. Gaunami atsitiktiniai

duomenys pasitelkiant autoenkoderį yra suspaudžiami į dvimatį požymių sluoksnį ir šis atvaizdavimas yra analizuojamas siekiant nustatyti optimalų tikrosios pagrindinės būsenos artinį. Siūlomas metodas yra universalus ir tinka tiek viendalelinėms, tiek daugiadalelinėms kvantinėms sistemoms tirti. Taip pat parodoma, kad jis yra pritaikomas ir daugiadalelinėms sistemoms stipriuose magnetiniuose laukuose, kai sistemos energijos spektras yra fraktalinio pobūdžio, o banginės funkcijos pasižymi nereguliaria struktūra.