β-Variational Autoencoder as an Entanglement Classifier

Nahum Sá* and Itzhak Roditi†
Centro Brasileiro de Pesquisas Físicas,
Rua Dr. Xavier Sigaud 150,
22290-180 Rio de Janeiro, Brazil and
(Dated: October 21, 2020)

We focus on using a β-Variational Autoencoder (β-VAE) to discriminate if a quantum state is entangled or separable based on measurements. We split the measurements into two sets, the set of local measurements and correlated measurements. Using the latent space, which is a low dimensional representation of the data, we show that restricting ourselves to the set of local measurement data it is not possible to distinguish between entangled and separable states. Meanwhile, when considering both correlated and local measurements an accuracy of over 80% is attained and show up in the structure of the latent space of the β-VAE.

I. INTRODUCTION

Entanglement is one of the most outstanding properties in quantum physics. It is associated with correlations that are non classical and may occur between otherwise unconnected quantum systems. Its usefulness emerges in such areas as quantum information, quantum computation, quantum cryptography and quantum metrology. It is also crucial to the phenomena of quantum teleportation [1]. Due to the importance of entanglement, which appears in so many instances, it is no wonder that there is a huge interest in finding methods that can detect, classify and analyze it [2–5].

On the other hand, deep learning (DL) techniques are becoming one of the most important assets in the physicists' toolbox, for instance helping us understand patterns that have little or no bias from the previously established theoretical framework. In general, DL techniques have been used in relation to computer science, and some successful examples are the technologies associated with pattern recognition, especially computer vision [6, 7], as well as some applications that received praise from the media, like Alpha GO [8]. Anyhow, lately this kind of approach found its way in physics, going beyond computer science. Some recent applications emerge in condensed matter physics [9], quantum many-body physics [10, 11] and molecular modeling [12]. DL techniques are even being applied as a way to unveil how physical concepts emerge [13, 14].

In this paper, our interest goes to quantum physics, particularly quantum information. One important problem in this domain is to distinguish between entangled and separable states, this is known to be an NP-hard problem [15], therefore there is no known classical algorithm that could solve this problem efficiently.

Hence, we analyse a way to encode the high dimensional labeled data coming from measurements of 2 qubit states, that corresponds to density matrices with 15 parameters, into a lower-dimensional representation which we call latent space. We deal with this problem through a Neural Network architecture called a β-Variational Autoencoder [16] which is used as a tool to distinguish between entangled and separable states.

In section II we explain how we simulated and labelled the data. In section III we explain how to use the β-Variational Autoencoder architecture for our problem and specify the loss function used. We discuss and show the results of the model in section IV and we finish with our conclusions and future works on section V.

II. DATA

There are several methods to distinguish between entangled and separable states, summarized in Horodecki et al. [4]. Here we chose to study the case of bipartite entanglement of 2 qubit states, applying the Positive Partial Trace (PPT) criterion (sometimes called Perez-Horodecki criterion) that was proposed first by Perez [17] and extended by Horodecki et al. [18] where it was shown that it provides a sufficient and necessary condition for a two-qubit system to be entangled.

In Quantum Theory the most general way to describe a quantum state is using a density matrix which is represented by a linear operator having the following properties: (1) Unity trace ; (2) Positive.

The PPT Criterion consists of using the partial transpose transformation on a density matrix, if after the transformation the state is completely positive, then it is separable (SEP). Otherwise, if the system consists of two-qubits, it is entangled (ENT). Considering $\rho \in C^2 \otimes C^2$ as a density matrix of two-qubits and T as the transpose map, the PPT can be synthesized by the following expression:

$$(I \otimes T)\rho \geq 0 \Rightarrow \rho \in \mathrm{SEP} (I \otimes T)\rho < 0 \Rightarrow \rho \in \mathrm{ENT} \quad (1)$$

Since we are going to use supervised learning we need labeled data. To generate the data we used Qutip [19, 20]

* nahumsa@cbpf.br
† roditi@cbpf.br; Also at Institute for Theoretical Physics, ETH Zurich 8093, Switzerland.

to simulate random density matrices and use the PPT criterion to label the data, we chose the label "1" for entangled states and the label "0" for separable states. We then measure on the Pauli matrices basis $\{\sigma_i \otimes \sigma_j\}$, where $i, j \in 0, x, y, z$ and $\sigma_0 = I$. All measurements are labeled using the following convention:

$$M_{ij} = \mathrm{Tr}\big[\rho(\sigma_i \otimes \sigma_j)\big] \tag{2}$$

For the two-qubit case we need 15 measurements, we can exclude the measurement M_{00} that always equals to 1 because of the definition of a density matrix, in order to have a complete tomography of the state. One can split the tomographic-complete measurements into two disjoint sets: correlated measurements, M_{ij} such that $i, j \neq 0$, having 9 measurements and local measurements, M_{ij} such that $i = 0$ or $j = 0$, having 6 measurements.

Using this convention we have three types of training and test data, each with 5000 samples and 3000 samples respectively. For convenience, we call these three types: the tomographic-complete dataset, correlated measurements dataset and local measurements dataset.

III. MODEL

The Variational Autoencoder (VAE) was proposed by Diederik and Welling [21] and is most commonly used for generative modeling. It has been extended by Higgins et al. [16] for an architecture which is called β-Variational Autoencoder (β-VAE) which creates disentangled representations on the latent space (usually a lower-dimensional representation of the data). Both models can be represented by the same graph shown in figure 1.

The main principle beneath the VAE, or β-VAE, is the use of a probabilistic latent space, which is a lower-dimensional representation of the data, that we assume follows some prior distribution. The most common choice is the Gaussian distribution $\mathcal{N}(0, 1)$ which will be used in this work.

For our problem, the idea is to encode the high dimensional input of measurements into a two-dimensional latent space that acts as a classifier of entanglement using a β-VAE.

We trained from end-to-end using backpropagation with a two-component loss function that consists of a categorization loss and a latent loss. Our choice of categorization loss $\mathcal{L}_{\mathrm{cat}}(\mathbf{y}, \hat{\mathbf{y}})$ is the categorical cross-entropy, where y_i is the true label and $\hat{y}_i$ is the predicted label. For the latent loss $\mathcal{L}_{\mathrm{KL}}(\mu, \sigma)$ we chose the Kullback-Leibler (KL) Divergence. The total loss is given by the sum of these two losses:

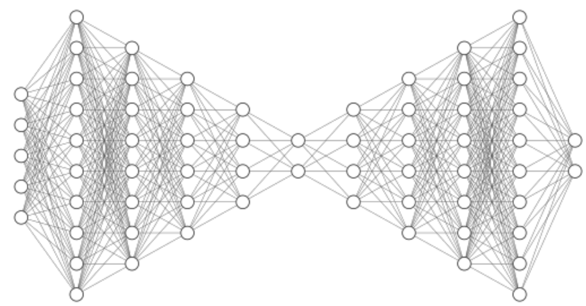

FIG. 1. The architecture of a VAE or β-VAE. In this paper, only the input depends on the dataset, all others are independent of the dataset. The encoding structure is of size 512, 256, 128, 64, 32 the latent space is of size 2, the decoder structure is of size 32, 64, 128, 256, 512 and the output of size 2.

$$\begin{aligned} \mathcal{L}_{\mathrm{total}} = r_{cat} & \overbrace{\left[\sum_i y_i \log(\hat{y}_i)\right]}^{\mathcal{L}_{\mathrm{cat}}(\mathbf{y},\hat{\mathbf{y}})} + \\ & + \beta \overbrace{\left[\frac{1}{2}\sum_i(\sigma_i + \mu_i^2 - 1 + \log(\sigma_i)\right]}^{\mathcal{L}_{\mathrm{KL}}(\mu,\sigma)} \end{aligned} \tag{3}$$

Where r_{cat} and β are weighting coefficients which are hyperparameters of our model and should be optimized for our task.

For the training, we adopt the Adam algorithm [22] as the optimizer with the 'reduce learning rate on plateau' method callback on Keras framework [23]. In all layers, except the last layer, we used LeakyReLU Activation in conjunction with a Dropout layer [24] to avoid overfitting. For the last layer, our choice was the softmax activation in order to capture the probability of the state being separable or entangled.

We trained the model for 100 epochs, starting with learning rate 0.005, using $r_{cat} = 500$ and $\beta = 1$ for each data set. In order to find those hyperparameters we discuss the methods for hyperparameter tuning on sec IV.

IV. RESULTS AND DISCUSSION

We trained and evaluated our model for the three datasets, as specified in sections II and III. In our model, we will encode the information of the 15-dimensional input into a 2 dimensional latent space as represented in figure 3, in which is possible to see that there are correlations between the dataset used and the latent space. The loss, for the tomographically-complete set, regarding training is showed in figure 2. For the other datasets, the loss behaves in a similar manner varying only the accuracy.

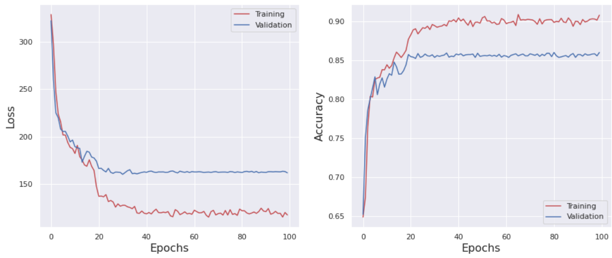

FIG. 2. The plot of the Loss function (left) and accuracy of the model(right) during training and evaluation on unseen data. The behavior of the loss function is the same for all datasets but with different accuracy, as specified in the text.

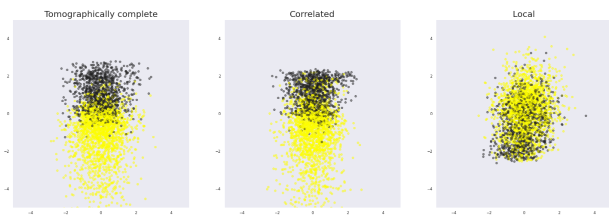

FIG. 3. The plot of the latent space for each test dataset (examples that weren't previously seen by the model). Yellow dots represents entangled states and black dots represents separable states. From left to right tomographic-complete measurements, correlated measurements and local measurements. We see that there is clustering on both tomographic-complete measurements and correlated measurements, but we find no clustering on local measurements, showing that the importance of each measurement for entanglement detection is different, therefore we could use only the correlated measurements for detecting entanglement.

For tomographic-complete measurements, we see a clear distinction between entangled and separable states, therefore we can use the latent space as an entanglement classifier. For instance, if we chose all points with $y > 0$ to be separable we find an accuracy of 84%, comparing to the accuracy of the whole model, which is 88%. Therefore, to use only the latent space is effective as a discriminator of entanglement.

The same can be done for correlated measurements. Indeed, choosing $y > 0$ to be separable states we find an accuracy of 80% and for the full model, we find an accuracy of 83%. On the other hand, for local measurements choosing $y > 0$ on the latent space gives approximately the same accuracy as the whole model, namely, 61%.

It is interesting to note, as stated before, that the latent space representation depends on the type of measurement being made. For the correlated measurements (M_{ij} such that $i, j \neq 0$) we see that the latent space still clusters into two different classes just as the case where we use tomographically-complete measurements. On the other hand, for local measurements(M_{ij} such that $i = 0$ or $j = 0$) it is not possible to distinguish between separable and entangled states using the latent space.

To evaluate the hyperparameters of the model we varied the β factor of the loss equation (Eq. 3) as shown in figure 4. As can be seen, when $\frac{\beta}{r_{cat}} > 0.3$ the accuracy of the model goes down considerably.

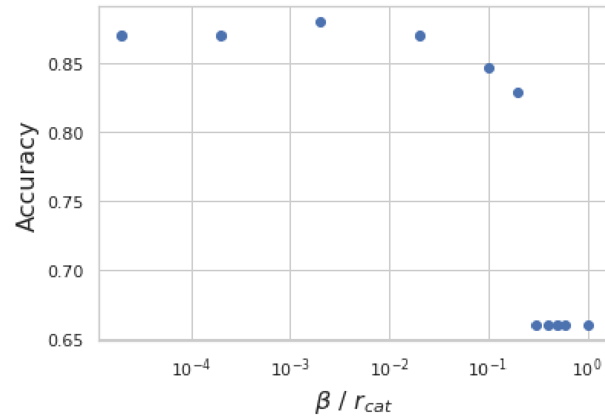

FIG. 4. Plot showing the β/r_{cat}-dependency of the accuracy for tomographic-complete measurements. There are two regimes that are defined by the ratio between β and r_{cat}, this happens because the β factor enforces the latent space distribution for being equal to a Gaussian distribution $\mathcal{N}(0, 1)$ if the β is considerably smaller than r_{cat} the latent space distribution doesn't need to be Gaussian.

Our assumption is that the β factor multiplying the KL Divergence forces the latent space distribution to the same prior Gaussian $\mathcal{N}(0, 1)$ distribution. Analysing the shape of our result distribution, mainly the entangled states, we see that it is not Gaussian at all, therefore enforcing the KL Divergence condition will diminish the accuracy of the model.

There are a handful of techniques to optimize the training on Deep Learning, one example is to use an unsupervised learning method before the supervised learning fine-tuning. In order to see if there is a significant difference in using this technique, we used Restricted Boltzmann Machines for pretraining [25]. In order to use this method, it is needed to convert all variables be binary, so encode binary representations of the measurements, labeling 1 when the measurement is positive and 0 when the measurement is negative. However no significant improvement on both the latent space and loss have been found using this technique.

V. CONCLUSION

In this paper, we propose a novel way to use the latent space of a β-Variational Autoencoder to encode the information concerning the entanglement of the quantum state using a set of tomographically-complete measurements. We divide this set into two disjoint sets, one for correlated measurements and the other for local measurements in order to analyse if there is any difference between these two types of measurements.

Applying our model on a tomographically-complete set of measurements of two-qubit system, we can distinguish between entangled and separable states with high precision both on the prediction of the model (88%) and using the latent space as an entanglement classifier (84%). In addition, for correlated measurements, of type $\sigma_{x,y,z} \otimes \sigma_{x,y,z}$, we also can distinguish between entangled and separable states, but with less precision for the whole model (83%) and using the latent space (80%) compared to the set of tomographically-complete measurements.

On the other hand, applying for local measurements the model is not able to learn any representation of entanglement in the latent space, showing that local measurements of type $\sigma_{x,y,z} \otimes I$ or $I \otimes \sigma_{x,y,z}$ is not able to characterize if the state is entangled or separable.

This result is supported by quantum theory because entangled states are expected to show non-locality given by correlated measurements, on the other hand, separable states are expected to be characterized by local measurements. Our model provides a novel way to identify local and correlated measurements.

In the future, we intend to analyse if this type of model can find accurate description of bipartite or multipartite entanglement for more than two-qubits.

ACKNOWLEDGMENTS

We would like to thank Luiz Bispo, and Leonardo Cirto for useful comments on the manuscript. I. R. thanks the Quantum Information Group of ETH-Zurich for their warm hospitality. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

The source code is available on GitHub.

[1] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).

[2] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Entanglement in many-body systems, Reviews of modern physics **80**, 517 (2008).

[3] Y.-C. Ma and M.-H. Yung, Transforming bell's inequalities into state classifiers with machine learning, npj Quantum Information **4**, 1 (2018).

[4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Reviews of modern physics **81**, 865 (2009).

[5] M. B. Plenio and S. S. Virmani, An introduction to entanglement theory, in *Quantum Information and Coherence* (Springer, 2014) pp. 173–209.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in neural information processing systems* (2012) pp. 1097–1105.

[7] P. Sermanet, K. Kavukcuoglu, S. Chintala, and Y. LeCun, Pedestrian detection with unsupervised multi-stage feature learning, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2013) pp. 3626–3633.

[8] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, *et al.*, Mastering the game of go without human knowledge, Nature **550**, 354 (2017).

[9] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, Nature Physics **13**, 431 (2017).

[10] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[11] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, Nature Physics **14**, 447 (2018).

[12] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. Von Lilienfeld, Fast and accurate modeling of molecular atomization energies with machine learning, Physical Review Letters **108**, 058301 (2012).

[13] R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, Discovering physical concepts with neural networks, Physical Review Letters **124**, 010508 (2020).

[14] H. P. Nautrup, T. Metger, R. Iten, S. Jerbi, L. M. Trenkwalder, H. Wilming, H. J. Briegel, and R. Renner, Operationally meaningful representations of physical systems in neural networks, arXiv preprint arXiv:2001.00593 (2020).

[15] L. Gurvits, Classical deterministic complexity of Edmonds' problem and quantum entanglement, in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing* (2003) pp. 10–19.

[16] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner, beta-vae: Learning basic visual concepts with a constrained variational framework., Iclr **2**, 6 (2017).

[17] A. Peres, Separability criterion for density matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[18] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Physics Letters A **223**, 1 (1996).

[19] J. Johansson, P. Nation, and F. Nori, Qutip: An open-source python framework for the dynamics of open quantum systems, Computer Physics Communications **183**, 1760 (2012).

[20] J. Johansson, P. Nation, and F. Nori, Qutip 2: A python framework for the dynamics of open quantum systems, Computer Physics Communications **184**, 1234 (2013).

[21] D. P. Kingma and M. Welling, Auto-encoding variational bayes, arXiv preprint arXiv:1312.6114 (2013).

[22] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[23] F. Chollet *et al.*, Keras, `https://github.com/fchollet/keras` (2015).

[24] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, The journal of machine learning research **15**, 1929 (2014).

[25] G. E. Hinton and R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, science **313**, 504 (2006).