Dual-Parameterized Quantum Circuit GAN Model in High Energy Physics

Su Yeon Chang[1,2,*], *Steven* Herbert[3,4,**], *Sofia* Vallecorsa[1,***], *Elías F.* Combarro[5], and *Ross* Duncan[3,6,7]

[1]OpenLab, CERN, Geneva, Switzerland
[2]Department of Physics, EPFL, Lausanne, Switzerland
[3]Cambridge Quantum Computing Ltd, Cambridge, UK
[4]Department of Computer Science and Technology, University of Cambridge, Cambridge, UK
[5]Department of Computer Science, University of Oviedo, Oviedo, Spain
[6]Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK
[7]Department of Physics, University College London, London, UK

Abstract. Generative models, and Generative Adversarial Networks (GAN) in particular, are being studied as possible alternatives to Monte Carlo simulations. It has been proposed that, in certain circumstances, simulation using GANs can be sped-up by using quantum GANs (qGANs).

We present a new design of qGAN, the dual-Parameterized Quantum Circuit (PQC) GAN, which consists of a classical discriminator and two quantum generators which take the form of PQCs. The first PQC learns a probability distribution over N-pixel images, while the second generates normalized pixel intensities of an individual image for each PQC input.

With a view to HEP applications, we evaluated the dual-PQC architecture on the task of imitating calorimeter outputs, translated into pixelated images. The results demonstrate that the model can reproduce a fixed number of images with a reduced size as well as their probability distribution and we anticipate it should allow us to scale up to real calorimeter outputs.

1 Introduction

The next High Luminosity Large Hadron Collider (HL-LHC) phase will collect an overwhelming amount of data, with complex physics and small statistical error. To analyse this data, high precision methods which use only limited resources are needed. Traditional Monte Carlo based simulation, such as *Geant4* [1, 2] and the *GeantV* prototype [3] for *full simulation* of particle transport, is however very time-consuming, therefore new approaches using deep neural networks have been studied for *fast simulations*.

Generative Adversarial Networks (GAN) [4] are a strong candidate for such fast simulations. Based on two neural networks, generator and discriminator, trained alternatively, GANs have been widely explored thanks to their ability to generate images with complex structures at much high speed. In HEP, the variations of GAN, such as CaloGAN [5] and 3DGAN [6],

---

*e-mail: su.yeon.chang@cern.ch
**e-mail: steven.herbert@cambridgequantum.com
***e-mail: sofia.vallecorsa@cern.ch

have achieved similar performance as full Monte Carlo based simulation, but with reduced time taken.

At the same time, quantum computing has emerged as another important pillar in modern research attracting the attention of many researchers due to its potential to execute certain tasks with an exponentially reduced amount of resources both in time and space compared to classical processors [7]. It has already shown promising results in various fields, such as optimization [8, 9] and cryptography [10].

Advances in both deep learning and quantum computing suggest to merge them to benefit their advantages at once, leading to a new field of study, so-called *Quantum Machine Learning* (QML). Quantum Generative Adversarial Networks, which are the quantum version of GANs, are one of its examples. Several quantum GAN models have been investigated in the last few years, but the scientific community still confronts a need to further explore in order to apply the model to more realistic use-cases.

In this paper, we propose for the first time a *Dual Parameterized Quantum Circuit GAN* model (dual-PQC GAN model) as one of the improvements to overcome the remaining limitations of quantum GANs. This model uses two parametrized quantum circuits, which share the role of a single quantum generator: the first PQC learns the distribution over image samples, while the second PQC determines the amplitude distribution over pixels on a single image. Thanks to this separation, it is possible to exploit the continuous nature of probability distributions over output states in quantum circuits to represent continuous variables.

This paper is organized as follows. Section 2 summarizes the application of Generative Adversarial Networks in HEP. We then present a short overview of a quantum version of GAN in Section 3. In Section 4 and Section 5, the first prototype of a dual-PQC GAN model is proposed, with the results of its simulation. This paper concludes with Section 6, which summarizes and gives an outlook of future works.

2 Applications of GANs in HEP

GANs, designed by I. Goodfellow et al. in 2014 [4], are deep generative models which aim to reproduce new data from a given original training set. They are characterized by two deep neural networks, **Generator** *G* and **Discriminator** *D*, which are trained alternatively. During the training, *G* progressively generates data similar to the real one, while *D* increases the probability of assigning the correct labels to both real and fake data. Numerous improvements have been made since the initial proposal and, in particular, GANs have achieved remarkable success in image processing and generation, via variations such as Deep Convolutional GAN (DCGAN) [11], Auxiliary Classifier GAN (ACGAN) [12], Progressive GAN [13], etc.

The evolution of GANs has attracted strong interest in the high energy physics domain. In HEP, the detectors can be described as 3D cameras, recording pictures of particle collisions. Calorimeters, in particular, measure the energies deposited by a shower of the particles that traverse them. They generally consist of alternate arrays of active sensor material and passive dense layers to ensure that the incoming (primary) particle will deposit most of its energy inside their volume. These energy depositions can be compared to the monochromatic pixel intensities of a 3D image. Because of their high granularity, the detailed simulation of a calorimeter is particularly time-consuming. As a result, GANs come into the limelight to allow fast simulation of particle showers with high fidelity.

One possible application of GAN in HEP is 3DGAN [6, 14], which is a 3D extension of GAN, using 3D (de-)convolutional layers to capture the whole 3D energy profile. It simultaneously performs two additional tasks of estimating the incoming particle energy and measuring the total deposited energy to enhance the stability and convergence of networks. Details on 3DGAN architecture and its performance validation are available in [6].

3 Quantum Generative Adversarial Networks

The possibility of combining machine learning and quantum computing also led to the generalization of GAN to quantum systems by S. Lloyd and C. Weedbrook [15]. The main mechanism of the model, the adversarial training, is reproduced, but different scenarios are possible: the input data can be either quantum data or classical data embedded in quantum states and the discriminator/the generator can also be either classical or quantum.

Since its initial proposal, several quantum GAN (qGAN) variations have been suggested to generate either classical data [16–18] or quantum data [19–22]. Zoufal et al [16] proposed a hybrid qGAN model composed of a quantum generator and a classical discriminator to train on *classical* data. During the training, the generator learns an arbitrary probability distribution over discrete variables, which is encoded in the amplitudes of the final quantum state. This model was applied of quantum finance, and demonstrated using the real quantum hardware, *IBM Q Boeblingen*. Anand et al [23] also present similar results, but simulated on the real Rigetti quantum hardware, *Aspen-4-2Q-A*. Unlike the aforementioned models treating classical data, Situ et al [18] propose a qGAN model which aims to approximate an unknown pure quantum state with a quantum generator and a quantum discriminator. One problem with quantum machine learning models is the apparent difficulty of training PQCs, captured by the vanishing gradient and barren plateau problems. Fortunately, there have been also studies on the methods to improve the performance of qGAN, for instance, Quantum Multiplicative Matrix Weight (QMMW), which helps to avoid mode collapse or vanishing gradient problem in qGAN [24].

The results from the previous research are impressive, showing the potential of qGAN for the near-term quantum hardware. However, additional investigations are needed in order to fully understand the quantum advantages of qGAN and generate not only simple probability distributions but also more complex image samples.

4 Dual-PQC GAN model

Our preliminary experiments involved training a qGAN, as conceptualised in [16], for the calorimeter problem – however this immediately revealed a problem: as the image itself is encoded in the amplitudes of the computational basis states, this meant that only the average of all the training samples could be learned, and the GAN did not, therefore, sample typical images. To put this in more precise terms, in order to achieve the exponential compression of representing a $N = 2^n$ pixel image using n qubits, it follows that the qGAN prepares a quantum state of the form:

$$|\psi\rangle = \sum_{j=0}^{2^n-1} \sqrt{I_j}\,|j\rangle \tag{1}$$
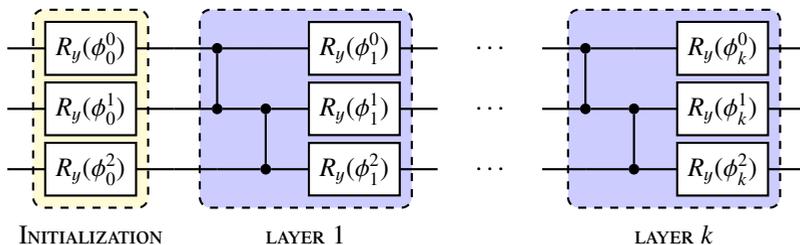
Figure 1: *Quantum variational form using Pauli-Y rotations and CZ gates with depth k.*

where I_j is the intensity of the jth pixel. So we can see that, in a sense, the qGAN encodes a *single* image as a probability distribution, and so there is no room left to also encode a probability distribution, representing the full dataset, to sample images from. This problem does not arise in classical GANs, where a single sample from a $O(N)$ sized neural network generates a single image in one go.

In this section we describe our solution, a new type of quantum GAN, the *dual Parameterized Quantum Circuit (PQC) GAN model* (dual-PQC GAN model). It aims to reproduce a set of image samples from real training data while preserving the exponential compression achieved by amplitude encoding. The work in [25] follows from similar motivation.

The dual-PQC GAN is a hybrid qGAN architecture which has one classical discriminator and two parameterized quantum circuits, PQC1 and PQC2, sharing the role of the generator. PQC1, with n_1 qubits, learns the probability distribution over image samples and PQC2, with n_2 qubits, learns the amplitude distribution over pixels of each image. The classical discriminator takes the training set and the images generated by PQC2, and it classifies them into real and fake. The predicted labels are used to tune alternatively ϕ_1, ϕ_2 and θ, the parameters for PQC1, PQC2, and the discriminator, respectively.

In this study, both PQCs consist of alternating layers of single-qubit Pauli-rotations and a set of two-qubit entanglement gates, as shown in Fig. 1 It is widely used in quantum machine learning thanks to its strong expressive power, offering an effective way of reconstructing an expected behaviour [26, 27]. We use RY rotation gates and CZ entanglement gates, but other choices are possible.
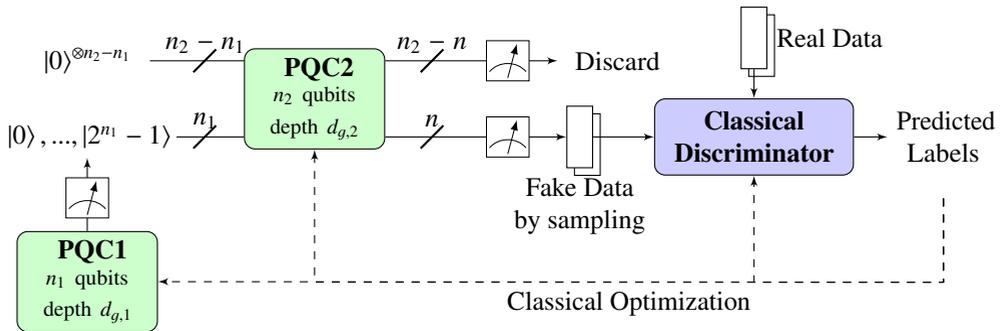
Figure 2: *Schematic Diagram of dual-PQC GAN to reproduce images of 2^n pixels.*

Consider a training set $X \subset \mathbb{R}^{2^n}$ of $N = 2^n$ pixel images. To begin, the output state of PQC1 is measured producing n_1 bits. Then, via a set of Pauli-X gates, this bit string is used to initialise PQC2 with the corresponding computational basis state in the 2^{n_1} dimensional Hilbert space, $|i\rangle \in \{|0\rangle, ..., |2^{n_1} - 1\rangle\}$. By repeatedly measuring n output qubits of PQC2, we finally construct the probability distribution over the computational basis in the 2^n dimensional Hilbert space. This distribution is *interpreted* as an image of size 2^n for each input state.

Note that since PQC2 performs a unitary operation, and since its inputs are always computational basis states, the output quantum states are necessarily orthogonal. This puts an unwanted restriction on the possible images. Relying on the Stinespring dilation theorem, we can remove this restriction by choosing $n_2 > n_1, n$ and using some ancilla qubits which are discarded at the end. We will return to this point at the end of the section.

Let p_g^i be the probability that state $|i\rangle$ is measured by PQC1; let $\mathcal{I}_i$ denote the normalized image produced by PQC2 when given the input state $|i\rangle$; and let I_{ij} be the amplitude of jth

pixel of $\mathcal{I}_i$ with $i \in \{0, ..., 2^{n_1} - 1\}$ and $j \in \{0, ..., 2^n - 1\}$. Then the output states generated by PQC1, G_{1,ϕ_1}, and PQC2, G_{2,ϕ_2}, are explicitly given as :

$$G_{1,\phi_1} |\psi_{initial}\rangle = |g_{1,\phi_1}\rangle = \sum_{j=0}^{2^{n_1}-1} \sqrt{p_g^j} |j\rangle , \tag{2}$$

$$G_{2,\phi_2} |\mathbf{0}\rangle |i\rangle = |g_{2,\phi_2}\rangle = \sum_{j=0}^{2^n-1} \sqrt{I_{ij}} |\Psi_j\rangle |j\rangle \tag{3}$$

where $|\psi_{initial}\rangle$ is the input state of PQC1 fixed during the whole training and $|\Psi_j\rangle$ are some $n - n_2$ qubit states that we discard.

During the training, PQC1 learns the distribution $p_g(i)$ over $\mathcal{I}_i$, so that it approaches to the real distribution, p_{real} over X. On the other hand, PQC2 learns the amplitude over 2^n pixels for 2^{n_1} images, to make $\mathcal{I}_i$ as close as possible to real images. At the end of the training, the true/fake probabilty predicted by the discriminator for $\mathcal{I}_i$, $D(\mathcal{I}_i)$ should converge to $1/2$.

For the following simulations, we use a modified min-max loss, given as:

$$L_G = -\frac{1}{m} \sum_{i=1}^{m} \log D(G(\mathbf{z}_i)) = -\sum_{i=0}^{2^{n_1}-1} p_g^i \log(D(\mathcal{I}_i)) \tag{4}$$

$$L_D = \frac{1}{m} \sum_{i=1}^{m} \left[\log D(\mathbf{x}_i) + \log(1 - D(G(\mathbf{z}_i))) \right] = \frac{1}{m} \sum_{i=1}^{m} \log D(\mathbf{x}_i) + \sum_{i=0}^{2^{n_1}-1} p_g^i \log(1 - D(\mathcal{I}_i)), \tag{5}$$

where m is the batch size, x_i the real data and z_i random input. The first equality gives the definition of the loss in the classical GAN and the second equality the practical formula used in dual-PQC GAN simulations.

Based on the calculated loss, the parameters in the quantum generator are tuned by computing the analytic quantum gradient descent [28], while the discriminator is optimized in the exactly same way as in classical GAN. Further details on the method used for analytic quantum gradient are explained in Ref. [29, 30].

Ultimately, the dual-PQC GAN model can generate 2^{n_1} images of size 2^n. Increasing the number of qubits used in PQC1 and PQC2 allows to increase both the number and size of produced images. This model shows an advantage in terms of computational resources by using only $O(\log(N))$ qubits, compared to the classical neural networks with $O(N)$ neurons to reproduce an image of size N. Specifically, the potential advantages are threefold: firstly, there is an exponential reduction in space (memory) requirement; secondly, the resultant exponential reduction in number of tunable parameters (i.e. the number of tunable parameters is proportional to the number of gates in the PQC or neurons in the classical neural network) suggests that training could be performed more efficiently; finally, it is potentially advantageous to have the images encoded in quantum states if further processing is to be performed (for example if that image processing can itself be performed more efficiently using quantum computing). It should, however, be noted that if all one wants to do is to generate an image, then the number of samples required from PQC2 is exponential in the number of qubits.

How many ancillas are needed?

Since we want to be able to reproduce any collection of images, we have to use some ancillary qubits. Stinespring's theorem gives an upper bound on the number of ancillas but it is not tight. We will now show that $n_2 = 2n$ will suffice when we assume for simplicity that $n_1 = n$.

Let U denote the unitary matrix corresponding to PQC2. When U is applied to the input state $|0\rangle^{\otimes n} \otimes |i\rangle$, where $|i\rangle$ is an n-qubit basis state, the resulting state is always one of the columns of U; in particular it always one of the first 2^n columns. We'll construct an example of U which realises 2^n arbitrary images. Let

$$|\mathcal{I}_i\rangle = (\sqrt{I_{i,0}}e^{i\phi_{i,0}}, \ldots, \sqrt{I_{i,0}}e^{i\phi_{i,0}})^T$$

be a state whose amplitudes encode the pixels of image i. Consider the state $|U(i)\rangle = |i\rangle \otimes |\mathcal{I}_i\rangle$, and observe that if its first n qubits are measured in the computational basis and discarded, then the remaining quantum state is precisely $|\mathcal{I}_i\rangle$. Since $\langle U(i)|U(j)\rangle = \delta_{ij}$, ie, they are orthonormal, we can construct the required *unitary* matrix, U, by setting the first 2^n columns to be $|U(i)\rangle$ for $i \in \{0, \ldots 2^n - 1\}$ and choosing the remainder arbitrarily. Since these columns correspond to states which will not be selected by any input to PQC2, they don't matter. It should be noted that it is unlikely that the trained dual-PQC GAN would actually converge on unitaries of such a form, and thus this construction is given merely to demonstrate that $2n$ qubits suffice; further improvements are surely possible.

5 Training dual-PQC GAN

This section tests dual-PQC GAN described in Section 4 using *qiskit* statevector simulator and shows its potential to generate a set of image samples from a training set with a certain degree of fidelity. Note that no explicit measurement of qubits has been performed, as we are only using the theoretical statevector simulator. We emphasize that, in order to work with a manageable number of qubits, this study simplifies the original 3DGAN problem by reducing it to a 1D problem: reproducing the energy pattern along the calorimeter depth. In other words, the training dataset is composed by 1D energy profiles along the calorimeter z dimension, averaged over $N = 4$ pixels. It should be noted that such drastic reductions in problem size are common-place in quantum machine learning, in order to obtain useful proof-of-principle results.

In order to evaluate the performance of the model, the original data set of 20,000 sample images is classified into 2^n classes of via K-means clustering [31] as shown on Fig. 3a for the case $n = 2$. The average image for each class, displayed on Fig. 3b, gives an insight on the

(a)   (b)
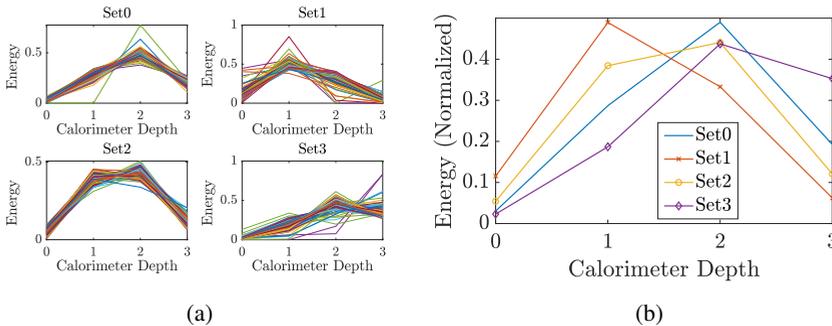
Figure 3: *Classification of 20,000 normalized real image samples (only 100 samples displayed) into* 4 *classes via K-means clustering [31] (a) and their average (b). The mean image of Set i is assigned to a computational basis state $|i\rangle$.*

shape of images which should be produced by PQC2. Note that this clustering is purely for evaluation of results - raw data are used for the training.

Although, theoretically, it is sufficient to take $n_1 = n = 2$, several preliminary simulations have shown that $n_1 = 2n = 4$ gives better stability in the results. Therefore, for the following simulations, PQC1 takes $n_1 = 4$ but still builds a probability distribution over $2^2 = 4$ images, by only measuring $n = 2$ qubits among four. PQC1 is initialized with an equiprobable superposition over the computational basis, $\{|0\rangle, ..., |2^{n_1} - 1\rangle\}$ with $n_1 = 4$. Furthermore, the initial parameters for both PQC1 and PQC2 are sampled from a uniform distribution over $[-\delta, \delta]$, with $\delta = 10^{-1}$. δ close to zero was chosen as we found this to speed up convergence. The discriminator is implemented in PyTorch, using an input layer with 4 nodes, two hidden layer with 256, 128 nodes, respectively, and a single node output layer. After the first two layers follows a Leaky ReLU function [32] with $\alpha = 0.2$ and a sigmoid function [33] is applied after the output layer. In addition, a gradient penalty [34] for real images is added to help stability and convergence of the model, with the parameters $\lambda = 7$, $k = 0.01$ and $c = 1$. The dual-PQC GAN is trained using the AMSGRAD optimizer with initial learning rate of 10^{-4} for PQC1 and discriminator and 10^{-3} for PQC2.
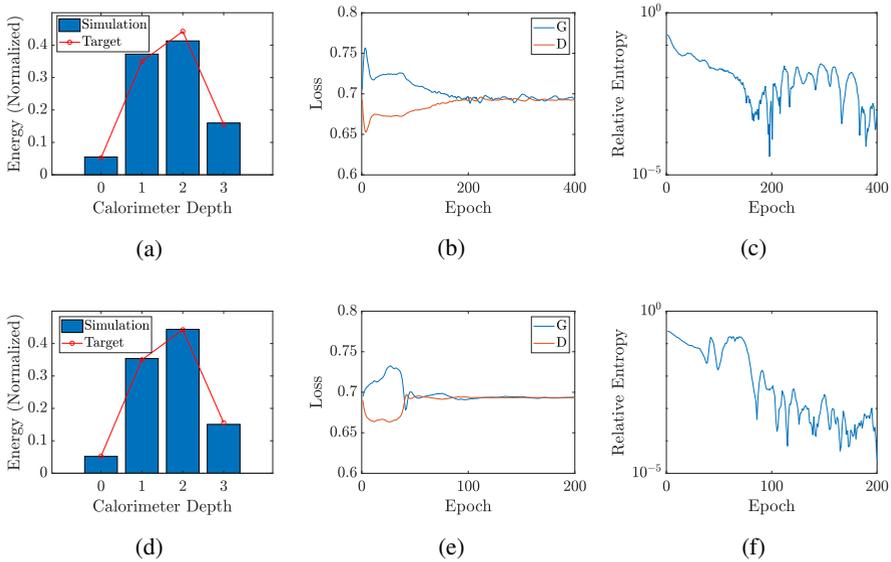
Figure 4: *Results of dual-PQC GAN simulations with $n = 2$, $n_1 = n_2 = 4$, $d_{g,1} = 2$ and $d_{g,2} = 6$ (a, b, c) / $d_{g,2} = 16$ (d, e, f). The mean image amplitudes calculated by Eq.(6) (a,d), the generator and the discriminator losses (b, e), and the progress in relative entropy for the mean image (c,f) are illustrated.*

Fig. 4 displays progress in the loss functions and relative entropy, as well as the average of generated images at the end of the training, weighted with their probability, given by:

$$\mathcal{I}_{mean} = \sum_{i=0}^{2^n - 1} p_g^i \cdot \mathcal{I}_i. \tag{6}$$

Both cases of $d_{g,2} = 6$ and $d_{g,2} = 16$ exhibit convergence in mean energy distribution towards the target, as well as convergence of generator and discriminator losses. Furthermore, the relative entropy between real and generated mean images reaches below 10^{-4} as shown on

Fig. 4c and Fig. 4f. Note that after the convergence in loss function, the relative entropy does not cease decreasing in case of $d_{g,2} = 16$, while it starts to oscillate with large amplitude in case of $d_{g,2} = 6$, reflecting certain degree of instability in the simulation.

As an analysis on the mean images is not enough to validate the GAN result, it is necessary to evaluate the individual images produced by PQC2 as well as the probability distribution generated by PQC1. The images generated by PQC2 with $d_{g,2} = 16$, displayed on Fig. 5c, are similar to the mean real images shown in Fig. 3b with the peaks at $x = 2$ or $x = 3$. On the other hand, those generated by PQC2 with $d_{g,2} = 6$ on Fig. 5a contain two images, $\mathcal{I}_0$ and $\mathcal{I}_2$, which are far from the real image profile.

The probability distribution generated by PQC1 can explain this discrepancy. As shown on Fig. 5b, the weights for $\mathcal{I}_0$ and $\mathcal{I}_2$ are negligible compared to those for $\mathcal{I}_1$ and $\mathcal{I}_3$, meaning that their labels, produced by the classical discriminator, are suppressed in the loss function given by Eq.(4). Therefore, the mean images and the generator loss could converge to the correct value, despite considerable errors in the produced images themselves. Contrarily, in case of $d_{g,2} = 16$, the probability on Fig. 5d implies that all 4 images have non-negligible weights, thus leading to consistency between the quality for the mean and individual images. This result certainly highlights the importance of choosing a correct structure of dual-PQC GAN model in order to prevent any bias during training.

(a)	(b)	(c)	(d)
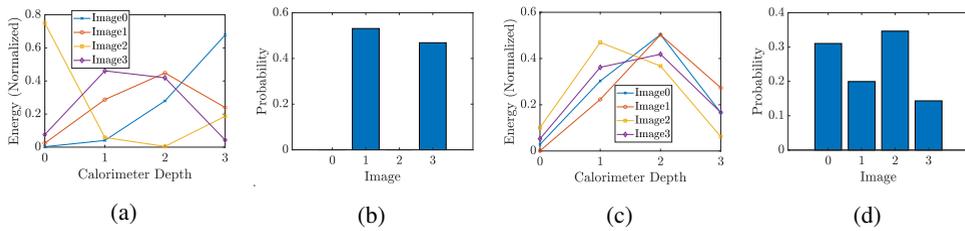
Figure 5: *Images generated by PQC2 (a,c) and its distribution obtained from PQC1 (b,d) with* $n = 2$, $n_1 = n_2 = 4$ *and* $d_{g,1} = 2$, *but with different* $d_{g,2}$ *: 1)* $d_{g,2} = 6$ *(a,b), 2)* $d_{g,2} = 16$ *(c,d). Although the mean images shown on Fig. 4 are close to the real value in both cases, PQC2 with* $d_{g,2} = 6$ *cannot imitate correctly the real data (c.f. Fig. 3b), while the one with* $d_{g,2} = 16$ *achieves to reproduce very similar images.*

Finally, the quality of individual images is evaluated by calculating the relative entropy between the real images Set i, shown on Fig. 3b, and the generated images $\mathcal{I}_i$ for $i = 0, 1, 2, 3$, as displayed on Fig. 6. Considering the lowest relative entropy values across the last 4 epochs (200 epochs are run in total), a bijection between real images and generated images can be constructed : $\mathcal{I}_0 \rightarrow$ Set0, $\mathcal{I}_1 \rightarrow$ Set3, $\mathcal{I}_2 \rightarrow$ Set1 and $\mathcal{I}_3 \rightarrow$ Set2. This result gives a quantitative proof that PQC2 can reproduce four different sets of images in the real training data. Despite this optimistic affirmation, the relative entropy large instability is the main point that should be improved in future studies.

Unfortunately, the number of qubits in the quantum generator scales not only with the number of pixels in one image but also with the number of images that can be produced, while in original GAN, the system size mainly scales with image size. This fact represents the largest limitation of the model and it requires further improvement. We are currently investigating a solution feeding extra noise to the remaining PQC2 $n_2 - n_1$ qubits, while PQC1 keeps generating a probability distribution over the zero-noise image.
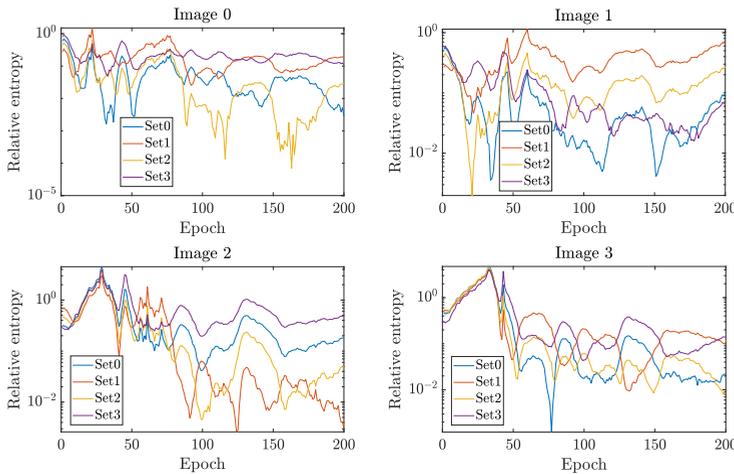
Figure 6: *Relative entropy of generated images $\mathcal{I}_0, ..., \mathcal{I}_3$ with respect to average of real image classes, shown on Fig. 3b. The images are generated via dual-PQC GAN with n = 2, $n_1 = n_2 = 4$, $d_{g,1} = 2$ and $d_{g,2} = 16$. According to minimum relative entropy at the end of the training, one-to-one correspondence can be established between real and generated images.*

6 Conclusion

This work presents a dual-PQC GAN model, a prototype of quantum GAN with two quantum generators, sharing the role of a single generator. One of the generators is responsible for reproducing the distribution over images, while the other for building amplitude distributions over pixels on a single image. The results obtained prove that this model can generate individual image samples and their probability distribution, similar to the training set.

It is also worth noting that, as the number of possible images (or classes of images if noise is added) grows *exponentially* with n_1, the fact that we sample from a finite set of images (or classes of images if noise is added) rather than a continuum of typical images (as in the corresponding classical case) is unlikely to seriously compromise performance.

An interesting direction for future research, that we intend to pursue, is testing the dual-PQC GAN in noisy simulators, and on real quantum hardware with lower depth for the quantum generator. Another intriguing question would be how to reproduce an arbitrary number of outputs with the dual-PQC GAN model. One possible way is to introduce a *complex-valued* noise to the remaining qubits in the PQC2. Using a fixed sampler instead of a trained quantum circuit for PQC1 to pass an entangled quantum state to PQC2 is another possible approach. Throughout future studies, we look forward to building a more advanced quantum GAN model to imitate the performance of classical GAN.

References

[1] M. Granieri, A. Estrada, *Down Converter Characterization in a Synthetic Instrument Context*, in *2006 IEEE Autotestcon* (2006), pp. 20–25, ISSN 1558-4550

[2] J. Allison et al., Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **835**, 186 (2016)

[3] G. Amadio et al., Journal of Physics: Conference Series **664**, 072006 (2015)

[4] I. Goodfellow et al., in *Advances in Neural Information Processing Systems 27*, edited by Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, K.Q. Weinberger (Curran Associates, Inc., 2014), pp. 2672–2680

[5] M. Paganini, L. de Oliveira, B. Nachman, Phys. Rev. D **97**, 014021 (2018)

[6] S. Vallecorsa, Journal of Physics: Conference Series **1085**, 022005 (2018)

[7] F. Arute et al., Nature **574**, 505 (2019)

[8] N. Moll et al., Quantum Science and Technology **3**, 030503 (2018)

[9] A. Gilliam, S. Woerner, C. Gonciulea, *Grover adaptive search for constrained polynomial binary optimization* (2019), `arXiv:1912.04088`

[10] P. Choy et al., *Cryptography in Quantum Computing*, in *Proceedings of the Future Technologies Conference (FTC) 2019* (Springer International Publishing, Cham, 2020), pp. 377–393, ISBN 978-3-030-32520-6

[11] A. Radford, L. Metz, S. Chintala, *Unsupervised representation learning with deep convolutional generative adversarial networks* (2015), `arXiv:1511.06434`

[12] A. Odena, C. Olah, J. Shlens, *Conditional Image Synthesis with Auxiliary Classifier GANs*, in *Proceedings of the 34th International Conference on Machine Learning*, edited by D. Precup, Y.W. Teh (PMLR, International Convention Centre, Sydney, Australia, 2017), Vol. 70 of *Proceedings of Machine Learning Research*, pp. 2642–2651

[13] T. Karras et al., CoRR **abs/1710.10196** (2017), `1710.10196`

[14] F. Carminati, G. Khattak, V. Loncar, T.Q. Nguyen, M. Pierini, R.B.D. Rocha, K. Samaras-Tsakiris, S. Vallecorsa, J.R. Vlimant, Journal of Physics: Conference Series **1525**, 012064 (2020)

[15] S. Lloyd, C. Weedbrook, Phys. Rev. Lett. **121**, 040502 (2018)

[16] C. Zoufal, A. Lucchi, S. Woerner, npj Quantum Information **5**, 103 (2019)

[17] J. Romero, A. Aspuru-Guzik, *Variational quantum generators: Generative adversarial quantum machine learning for continuous distributions* (2019), `1901.00848`

[18] H. Situ, Z. He, Y. Wang, L. Li, S. Zheng, Information Sciences **538**, 193 (2020)

[19] P.L. Dallaire-Demers, N. Killoran, Physical Review A **98** (2018)

[20] L. Hu et al., Science Advances **5** (2019)

[21] M. Benedetti et al., New Journal of Physics **21**, 043023 (2019)

[22] N. Shrivastava et al., *Opticalgan : Generative adversarial networks for continuous variable quantum computation* (2019), `1909.07806`

[23] A. Anand et al., *Experimental demonstration of a quantum generative adversarial network for continuous distributions* (2020), `2006.01976`

[24] Y. Du, M.H. Hsieh, D. Tao, *Efficient online quantum generative adversarial learning algorithms with applications* (2019), `1904.09602`

[25] M.S. Rudolph et al., *Generation of high-resolution handwritten digits with an ion-trap quantum computer* (2020), `2012.03924`

[26] J.R. McClean, J. Romero, R. Babbush, A. Aspuru-Guzik, New Journal of Physics **18**, 023023 (2016)

[27] M. Benedetti, E. Lloyd, S. Sack, M. Fiorentini, Quantum Science and Technology **4**, 043001 (2019)

[28] I. Kerenidis, A. Prakash, Phys. Rev. A **101**, 022316 (2020)

[29] P. Rebentrost, M. Schuld, L. Wossnig, F. Petruccione, S. Lloyd, New Journal of Physics **21**, 073023 (2019)

[30] A. Harrow, J. Napp, *Low-depth gradient measurements can improve convergence in variational hybrid quantum-classical algorithms* (2019), `arXiv:1901.05374`

[31] T. Kanungo et al., IEEE Trans. Pattern Anal. Mach. Intell. **24**, 881–892 (2002)

[32] A.L. Maas, *Rectifier Nonlinearities Improve Neural Network Acoustic Models* (2013)

[33] J. Han, C. Moraga, *The influence of the sigmoid function parameters on the speed of backpropagation learning*, in *From Natural to Artificial Neural Computation*, edited by J. Mira, F. Sandoval (Springer Berlin Heidelberg, Berlin, Heidelberg, 1995), pp. 195–201, ISBN 978-3-540-49288-7

[34] K. Roth, A. Lucchi, S. Nowozin, T. Hofmann, in *Advances in Neural Information Processing Systems 30*, edited by I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Curran Associates, Inc., 2017), pp. 2018–2028