Quantum Generative Adversarial Networks in a Continuous-Variable Architecture to Simulate High Energy Physics Detectors

Su Yeon Chang,[1,2] Sofia Vallecorsa,[1] Elías F. Combarro,[3] and Federico Carminati[1]

[1]*OpenLab, CERN, Geneva, Switzerland*
[2]*Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[3]*University of Oviedo, Oviedo, Spain*
(Dated: 28th January 2021)

Deep Neural Networks (DNNs) come into the limelight in High Energy Physics (HEP) in order to manipulate the increasing amount of data encountered in the next generation of accelerators. Recently, the HEP community has suggested Generative Adversarial Networks (GANs) to replace traditional time-consuming *Geant4* simulations based on the Monte Carlo method. In parallel with advances in deep learning, intriguing studies have been conducted in the last decade on quantum computing, including the Quantum GAN model suggested by IBM. However, this model is limited in learning a probability distribution over discrete variables, while we initially aim to reproduce a distribution over continuous variables in HEP.

We introduce and analyze a new prototype of quantum GAN (qGAN) employed in continuous-variable (CV) quantum computing, which encodes quantum information in a continuous physical observable. Two CV qGAN models with a quantum and a classical discriminator have been tested to reproduce calorimeter outputs in a reduced size, and their advantages and limitations are discussed.

I. INTRODUCTION

Modern society confronts an overwhelming amount of data, which should be treated with high precision and limited resources. Especially, in High Energy Physics (HEP), next High Luminosity Large Hadron Collider (HL-LHC) phase will collect and analyze an increasing amount of data with complex physics and small statistical error and it is therefore highly demanding in terms of computational resources. Traditional Monte Carlo based simulation is however very time-consuming, thus new approaches using Deep Neural Networks (DNNs) have been proposed to replace it.

Generative Adversarial Networks (GANs) [1] arise as one of the solutions for fast simulation. Based on its two neural networks, generator and discriminator, trained alternatively, the GAN model has been widely explored thanks to its performance to generate images with complex structures at much higher speed. In HEP, the variations of GAN, such as CaloGAN [2], have achieved to show similar performance to that of full Monte-Carlo based simulation, but with less amount of time taken for the computation. Meanwhile, quantum computing has emerged as another important pillar in modern studies attracting the attention of many researchers due to its potential to execute some tasks with an exponentially reduced amount of resources in time and space compared to a classical processor [3].

Advances in both deep learning and quantum computing have made it possible to extend the GAN paradigm to the quantum world, with the introduction of Quantum GAN (qGAN) models [4, 5]. Unfortunately, the most well-known qGAN model proposed by IBM [5], with a classical discriminator and a quantum generator in a qubit architecture, is limited to reproducing the probability distribution over discrete variables, while the calorimeters outputs that we would like to treat are continuous.

Continuous-variable (CV) quantum computation offers an insight to solve the problem. Unlike the well-known *qubit* architecture, which uses discrete fundamental information-carrying units, CV quantum computing uses *qumodes* which are continuous by their nature [6]. Taking a complementary position to the qubit architecture, the CV architecture has shown its efficiency in a number of papers that studies its applications, including CV neural networks (CVNNs) [7].

Our work introduces a new prototype of qGAN, implemented in CV architecture. The model has been tested to simulate continuous detector outputs, with a reduced size, and its results are analyzed and discussed in this paper.

II. CONTINUOUS-VARIABLE (CV) QUANTUM COMPUTING

Continuous-variable (CV) architecture is a paradigm of quantum computing, where the information is encoded in a continuous physical observable, for instance the strength of the electromagnetic field. In this model, the fundamental information-carrying unit is represented by a *qumode*, which is a quantum state of bosonic modes in the quantized electromagnetic field. The quantum state of N qumodes is often expressed as a superposition of position basis, $\{|\mathbf{x}\rangle\}_{\mathbf{x}\in\mathbb{R}^N}$ or Fock basis, $\{|n\rangle\}_{i\in\mathbb{N}}$:

$$|\psi\rangle = \int_{\mathbb{R}^N} \psi(\mathbf{x})\,|\mathbf{x}\rangle\,\mathrm{d}\mathbf{x} = \sum_{i=0}^{\infty} \langle n|\psi\rangle\,|n\rangle. \qquad (1)$$

By combining a set of common CV gates, such as displacement $D(\alpha)$, rotation $R(\phi)$, squeezing $S(z)$, beam-
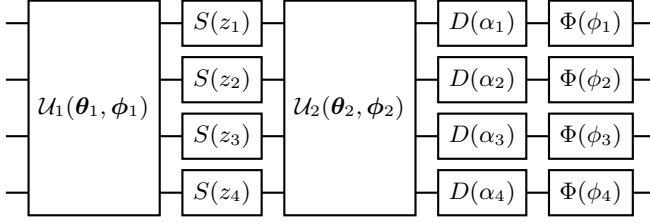
FIG. 1: *Structure of a CV neural network layer.* [7]

splitters $BS(\theta, \phi)$, etc, it is possible to build CV Neural Networks (CVNN), parameterized quantum circuits which perform exactly the same mathematical transformation as a fully connected layer.

Consider that a N-dimensional vector $\mathbf{x}$ is represented by an eigenstate $|\mathbf{x}\rangle$ of the position operator $\hat{x}$ with $|\mathbf{x}\rangle = \bigotimes_{i=1}^{N} |x_i\rangle = D(\mathbf{x}) |0\rangle^{\otimes N}$ where $|0\rangle$ is the vacuum state.

Then, the sequence of CV gates shown on FIG. 1 performs the following transformation $\mathcal{L}$ which is equivalent to a single classical fully connected layer

$$\mathcal{L} |\mathbf{x}\rangle = \Phi \circ \mathcal{D} \circ \mathcal{U}_2 \circ \mathcal{S} \circ \mathcal{U}_1 |\mathbf{x}\rangle \propto |\phi(W\mathbf{x} + \mathbf{b})\rangle, \quad (2)$$

where $\mathcal{U}$ is an interferometer [8], W the weight matrix, $\mathbf{b}$ the bias and Φ the non-linear activation function. More detailed explanations on CVNN can be found in Ref.[7]

III. CV QGAN

This section proposes two different prototypes of continuous-variable Quantum GAN (CV qGAN) models: one with both quantum generator and discriminator

(*Fully quantum* model), and the second with quantum generator and classical discriminator (*Hybrid* model). FIG. 2 summarizes the structure of the two models.

In both cases, N qumodes quantum generators are initialized with the first qumode displaced by a random noise $z \in \mathcal{N}(0, 1)$, while keeping all the other qumodes to be vacuum :

$$|initial\rangle = |z\rangle \otimes |0\rangle^{\otimes N-1} = D(z) |0\rangle \otimes |0\rangle^{\otimes N-1}. \quad (3)$$

It then applies CVNN layers shown on FIG. 1 to the initial state to transform it into the targeted state. The difference between the two prototypes comes from the measurement at the end of the generator. In fully quantum case, the quantum generator is directly connected to the quantum discriminator, without any measurement between them. The expectation value of the position operator at qumode N, $\langle x \rangle$ is measured at the end of the discriminator while discarding all the other qumodes. Another classical sigmoid function is then applied to return a predicted label, $\ell \in [0, 1]$, indicating the authenticity of the sample.

On the other hand, the hybrid model constructs a fake image by measuring the position expectation values at all N qumodes, $\langle \mathbf{x} \rangle$, at the end of the generator. This fake image is passed to the classical discriminator, to return a prediction used to update the generator parameters.

This CV qGAN has been tested only for three qumodes, i.e., $n = 3$, due to the limitation in simulation computing time and memory. The original calorimeter outputs of size 25×25 are averaged over the longitudinal direction, and then binned into 3 pixels, as shown on FIG. 4a.

The model is implemented in *Strawberryfields* [9, 10] for CV quantum computing and *Pennylane* [11, 12] for automatic quantum gradient descent.
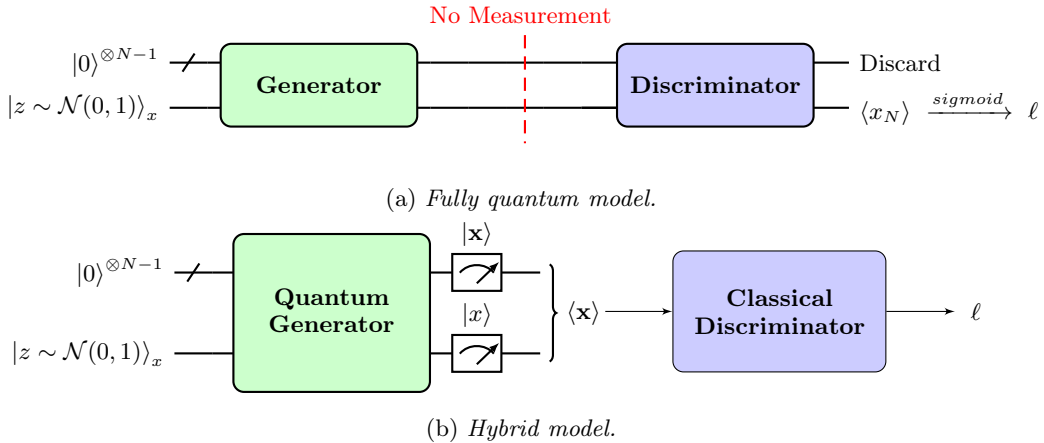

(a) *Fully quantum model.*

(b) *Hybrid model.*

FIG. 2: *Schematic diagrams of continuous-variable quantum GAN models, manipulating classical data embedded in quantum states.*

As shown on FIG. 3, the generator and the discriminator loss functions converge towards the expected value in all three cases. However, the mean images produced by the fully quantum and hybrid models with $d_g = 5$ differ from the targeted values, showing a limitation in the simulations. Unlike the previous two cases, the result for the hybrid model with $d_g = 3$ manifests a convergence of the mean image towards the targeted output.
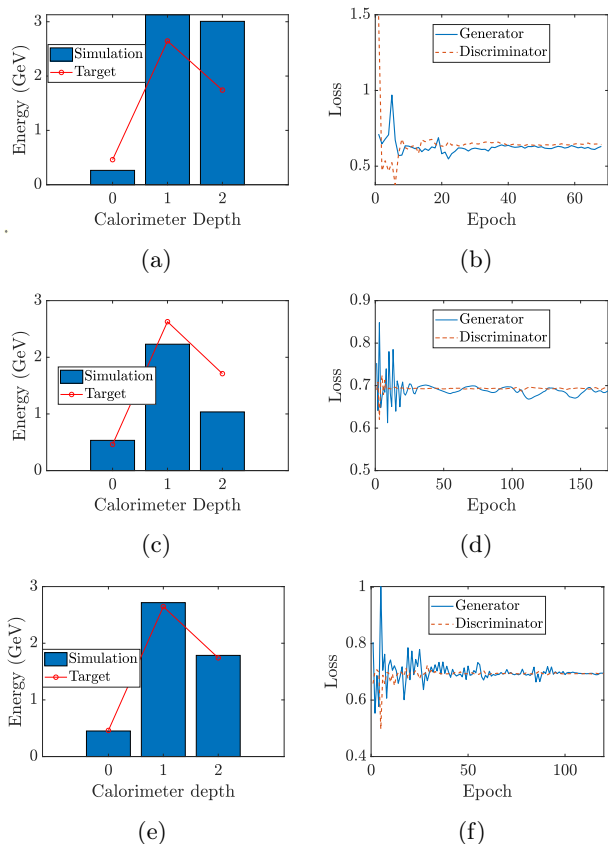

FIG. 3: *Mean images (a,c,e) and losses (b,d,f) obtained from the CV qGAN model simulations. Figures (a,b) display the result using the fully quantum model with $d_g = 5$, figures (c,d) the result using the hybrid model, with $d_g = 5$ and figures (e,f) with $d_g = 3$.*

The limitation is also revealed throughout the whole set of produced images, shown on FIG. 4. Notably, the images generated by the fully quantum model on FIG. 4b have significantly different shapes from the real images shown on FIG. 4a. The most critical problem is that at least half of the images contain negative *energy* values, which are certainly unphysical. The situation is slightly better in the hybrid model with $d_g = 5$ displayed on FIG. 4c. In spite of the difference in energy levels, the shape of images is similar to the real one, with a peak at $x = 1$. Furthermore, most of the images are above zero except three or four samples. It is also notable that the CV qGAN model can fall into a common GAN failure observed in classical cases. Despite the convergence

in mean images for the hybrid model with $d_g = 3$, the individual images shown on FIG. 4d exhibit an obvious mode collapse failure [13] where the generator produces only a small variety of samples.
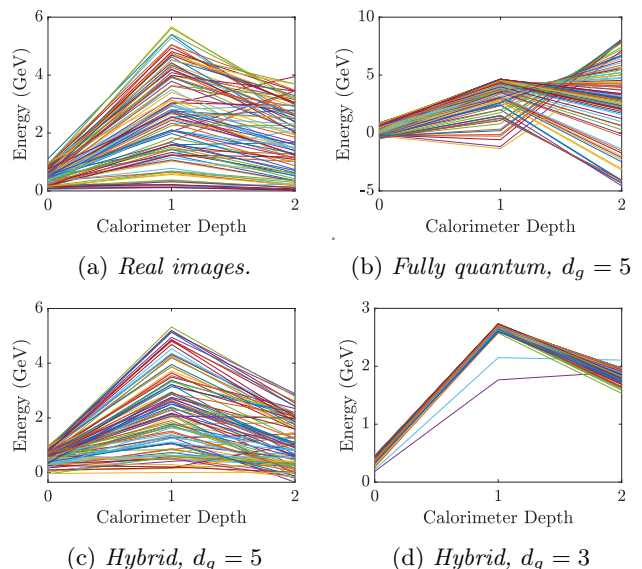

FIG. 4: *Comparison of the real training samples (a) with the images generated by the fully quantum CV qGAN with $d_g = 5$ (b) and the hybrid CV qGAN, with $d_g = 5$ (c) and $d_g = 3$ (d).*

Increasing latent space dimension To complete this work, hybrid CV qGAN model is tested with latent space dimension 3, i.e., by initializing all three qumodes in the generator with random noises and compared with a classical case as shown on FIG. 5. Both models use a classical discriminator with the same structure. But CV qGAN consists of the quantum generator with $d_g = 8$, while the classical one employs a classical generator having a similar size as the discriminator.

FIG. 5d and FIG. 5f reveal convergence in mean image as well as individual images reproduced by CV qGAN. In addition, unphysical behavior exhibited in the previous simulations is no more observed, as all the samples have positive energy. It is possible to find few samples with a peak at $x = 2$ on FIG. 5f, which exist in the real samples on FIG. 4a Remarkably, CV qGAN can imitate the performance of classical GAN with 264 parameters for the generator, which are 170 times less compared to 44947 parameters used in the classical case. Although it is impossible to make a direct comparison between the models due to large oscillations in loss functions shown on FIG. 5b, the study still gives an insight into the size of CV generator recommended for successful training.

Furthermore, the simulations have demonstrated that CV qGAN reveals convergence before 100 epochs while classical GAN started converging after 1000 epochs, exhibiting an advantage in terms of computational com-
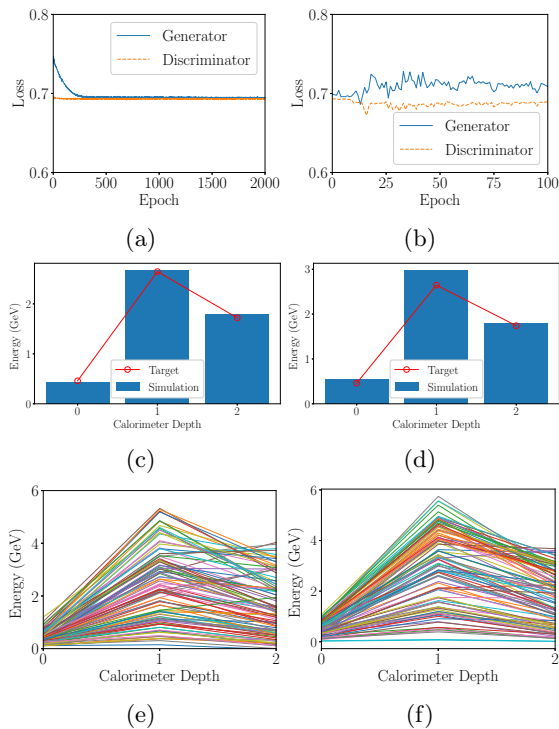
FIG. 5: *Losses (a,b), mean images (c,d) and image samples (e,f) obtained from the classical GAN (a,c,e) and the hybrid CV qGAN (b,d,f) simulations with latent space dimension 3.*

plexity with respect to the time to convergence. Unfortunately, the time needed to simulate the quantum generator is still the limiting factor to reach full convergence (350 min per epoch).

IV. CONCLUSIONS AND FUTURE WORK

In this work, we suggest a new prototype of qGAN using continuous-variable quantum computing. Thanks to an equivalence between classical and CV neural networks, it is possible to build a CV qGAN, having an architecture similar to the classical GAN. The current CV qGAN model has limitation in generating correctly the targeted image samples and it exhibits a mode collapse failure, which is often observed in the classical GAN. Stability and convergence are improved by increasing the latent space dimension, a result that is consistent with the finding reported in other works concerning the importance of the initialisation step [5] .

To further improve the model, we will explore techniques to speed the training (and in particular the gradient computation step) while, at the same time increasing the number of qumodes (and therefore the model representational power). The current simulations require a tremendous amount of time (several days) for a small number of qumodes (<5). Different entanglement configuration will be investigated together with mechanisms reproducing regularisation that are usually applied to classical networks, in order to improve the qGAN model and reproduce the real image data accurately in future studies.

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

[2] M. Paganini, L. de Oliveira, and B. Nachman, Calogan: Simulating 3d high energy particle showers in multilayer electromagnetic calorimeters with generative adversarial networks, Phys. Rev. D **97**, 014021 (2018).

[3] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[4] S. Lloyd and C. Weedbrook, Quantum generative adversarial learning, Phys. Rev. Lett. **121**, 040502 (2018).

[5] C. Zoufal, A. Lucchi, and S. Woerner, Quantum generative adversarial networks for learning and loading random distributions, npj Quantum Information **5**, 103 (2019).

[6] S. Lloyd and S. L. Braunstein, Quantum computation over continuous variables, Phys. Rev. Lett. **82**, 1784 (1999).

[7] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, Continuous-variable quantum neural networks, Phys. Rev. Research **1**, 033063 (2019).

[8] Z. Jiang, M. D. Lang, and C. M. Caves, Mixing nonclassical pure states in a linear-optical network almost always generates modal entanglement, Phys. Rev. A **88**, 044301 (2013).

[9] N. Killoran, J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook, Strawberry Fields: A Software Platform for Photonic Quantum Computing, Quantum **3**, 129 (2019).

[10] Strawberryfields, https://strawberryfields.readthedocs.io/ (2019), accessed: 20-01-2021.

[11] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, M. S. Alam, S. Ahmed, J. M. Arrazola, C. Blank, A. Delgado, S. Jahangiri, K. McKiernan, J. J. Meyer, Z. Niu, A. Száva, and N. Killoran, Pennylane: Automatic differentiation of hybrid quantum-classical computations (2018), arXiv:1811.04968 [quant-ph].

[12] Pennylane, https://pennylane.readthedocs.io/ (2018), accessed: 20-01-2021.

[13] I. J. Goodfellow, NIPS 2016 tutorial: Generative adversarial networks, CoRR **abs/1701.00160** (2017), arXiv:1701.00160.