

Deep reinforcement learning for near-deterministic preparation of cubic-phase gates and direct preparation of quartic-phase gates in photonic quantum computing

Amanuel Anteneh¹, Léandre Brunel^{1,*}, Carlos González-Arciniegas^{1,†} and Olivier Pfister^{1,2,‡}

¹Department of Physics, *University of Virginia*, 382 McCormick Road, Charlottesville, Virginia 22903, USA

²Charles L. Brown Department of Electrical and Computer Engineering, *University of Virginia*, 351 McCormick Road, Charlottesville, Virginia 22903, USA



(Received 11 June 2025; accepted 10 February 2026; published 2 March 2026)

Cubic-phase states are a sufficient gate resource for universal quantum computing over continuous variables. We present results from numerical experiments in which deep neural networks are trained via reinforcement learning to control a quantum optical circuit for generating cubic-phase states by leveraging quantum interference in phase space, with an unprecedented average success rate of 96%. The only non-Gaussian resource required is photon-number-resolving measurements. We also show that the exact same resources enable the direct generation of a quartic-phase gate, with no need for a cubic gate decomposition.

DOI: [10.1103/wkfp-tf74](https://doi.org/10.1103/wkfp-tf74)

Introduction. Continuous-variable quantum computing (CVQC) utilizes bosonic field (qumode) encoding in lieu of native qubits. Qumode encoding still enables hybrid bosonic qubit encoding (for example, via GKP (Gottesman, Kitaev, and Preskill) states [1]) and also offers the exceptional scalability of quantum optics, be it in free space [2–8] or, more recently, on chip [9–12]. As Lloyd and Braunstein pointed out, CVQC is universal if one has access to at least cubic Hamiltonian evolution in the quantum fields [13]. It is also important to note that, besides its aforementioned outstanding scalability, CVQC can be made fault tolerant [14]. Last but not least, CVQC offers a unique platform for the quantum simulation of quantum field theory [15,16].

An example of cubic Hamiltonian evolution is the cubic-phase gate $\exp(i\gamma Q^3)$, where γ is a real parameter and $Q = (a + a^\dagger)/\sqrt{2}$ is the field's amplitude quadrature, a being the photon annihilation operator. Cubic-phase gates have a non-Gaussian Wigner function since Gaussian gates are generated by Hamiltonians at most quadratic in the fields [17]. When combined with other cubic and quadratic gates, the cubic-phase gate can then yield evolution at any order of the fields, by virtue of the Baker-Campbell-Hausdorff formula [13]. Gate decompositions in terms of cubic-phase gates have been studied in detail [18–20]: Realizing a quartic gate $\exp(i\delta Q^4)$, for example, requires 29 gates, 15 of which cubic [19]. In this paper, we show that quartic-phase gates can be generated directly, with no need for a cubic decomposition.

In the context of measurement-based quantum computing, which is the foundation of modern photonic QC [21–26], cubic-phase gates can be applied to an arbitrary state $|\psi\rangle$, to Gaussian corrections left, by teleporting $|\psi\rangle$ through a cubic-phase state [1],

$$|\gamma\rangle = \sqrt{2\pi} e^{i\gamma Q^3} |0\rangle_p = \int_{-\infty}^{+\infty} ds e^{i\gamma s^3} |s\rangle_q, \quad (1)$$

where $|s\rangle_{p,q}$ are the respective eigenstates, of eigenvalue $s \in \mathbb{R}$, of the phase quadrature $[P = i(a^\dagger - a)/\sqrt{2}]$ and amplitude quadrature operators. The cubic-phase state $|\gamma\rangle$ is therefore a key resource for universal CVQC.

Thinking “classically,” one may seek to prepare cubic-phase states by generating cubic waves from a cubic field Hamiltonian. This deterministic generation of cubic-phase states can be done by squeezing a non-Gaussian trisqueezed state, easily obtained in the transmon-qubit cavity microwave fields that experience very large nonlinearities [27]. That direct approach is difficult to translate to the optical domain due to the weakness of third-order optical nonlinearities. This could be circumvented in the temporally and spatially confined, guided ultrafast optical regime, where nonlinearities are more efficient due to the higher peak field amplitudes [28].

Thinking more “quantum mechanically,” one may forgo using a third-order nonlinearity and instead seek to generate nonlinear waves by leveraging their corpuscular nature, e.g., using photon-number-resolving (PNR) measurements. This was the original—though probabilistic—proposal of GKP, employing displaced PNR detection on half of a two-mode-squeezed state [1]. It was shown in a later study that squeezing on the order of 17 dB and probabilistically detected photon numbers on the order of 50 would be needed to reach QC-useful values of γ [29]. While PNR measurements have since been demonstrated at the 100-photon level [30], this remains a demanding specification. Other, nonprobabilistic methods were proposed to reach smaller values of γ [31,32].

In this paper, we generalize the GKP protocol to an iterative process with an added in-loop displacement, leveraging

*Present address: Welinq, 40 Rue des Boulangers, 75005 Paris, France.

†Present address: Xanadu, 777 Bay Street, Toronto, Canada.

‡Contact author: olivier.pfister@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

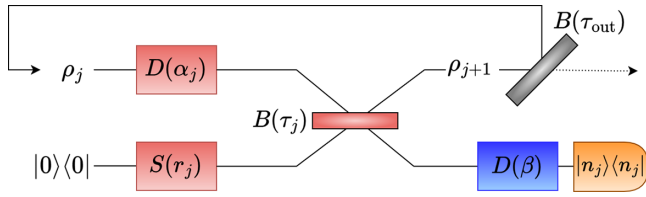


FIG. 1. Quantum optical circuit for cubic-phase state preparation. The photon number from the PNR measurement (orange box) conditions the density matrix that is input to the neural network. Real parameters α_j , r_j , and τ_j (red boxes) are set by the ML agent at each iteration j . $\tau_{\text{out}} = 0$ until the preparation process is over ($\tau_j = 0$), which triggers $\tau_{\text{out}} = 1$. The displacement $D(\beta)$ applied prior to the PNR measurement is fixed to $\beta = i2.5$.

interference in quantum phase space to quasideterministically (which is crucial for quantum computing scalability) shape the Wigner function to a high level of cubicity. When driven by a deep neural network trained by reinforcement learning, the process reaches an unprecedented 96% success rate for the generation of a cubic-phase state with $\gamma = 0.2$, using squeezing no higher than 10 dB, low displacements, and much lower PNR measurement values. We then give an even simpler protocol for directly, though probabilistically, generating a quartic-phase state and gate without relying on the cumbersome canonical cubic-gate decomposition. We expect this protocol to form the basis for another, similar machine learning (ML) algorithm.

Reinforcement learning for quantum state preparation. Quantum optical circuit: The quantum optical circuit we employed is displayed in Fig. 1.

$B(\tau_j)$ is a variable beamsplitter with transmittivity τ_j and $B(\tau_{\text{out}})$ is a switchable mirror with $\tau_{\text{out}} = 0$ when $j < m$ and $\tau_{\text{out}} = 1$ when $j = m$. The squeezed state can be realized by an optical parametric oscillator below threshold [33] and the loop must be interferometrically stabilized [34]. The PNR detector can be a superconducting transition edge sensor [30,35] or a superconducting nanowire single-photon detector in the PNR regime [36,37]. Displacements $D(\alpha_j)$ and $D(\beta)$ can be realized by interfering the quantum mode at a highly unbalanced beamsplitter with a weakly transmitted, frequency-degenerate coherent state.

Reinforcement learning concept and prior results: Reinforcement learning is a branch of ML that seeks to train a learning agent to act optimally in a given environment [38]. The optimality of an action is informed by a reward signal that the learner receives upon performing the action. The complexity of this learning task arises from the fact that actions taken at one point in time will affect future rewards: For example, selecting an action with a less-than-optimal immediate reward can lead to scenarios where larger rewards are available in the future. Deep reinforcement learning (DRL) is a subfield of reinforcement learning that utilizes deep neural networks as the learning agent and has achieved state of the art performance on many control tasks [39–41]. A lot of the prior work on quantum state preparation has treated the task as an optimization problem in which the parameters of the state preparation circuit are directly optimized to minimize some loss function, e.g., the negative of the fidelity of the output state of the circuit with some target state. These methods

typically compute the gradients of the circuit parameters with respect to the loss function, using the method of automatic differentiation, and iteratively update the circuit parameters, using the gradient descent algorithm, until a configuration of the circuit parameters is found which minimizes the loss [42]. Utilizing this method a measurement-free approach was taken in Refs. [43,44] for the preparation of various non-Gaussian quantum states of interest, the cubic-phase state among them. In the case of the former reference, the non-Gaussian resources were Kerr nonlinearities whereas in the latter the selective number-dependent arbitrary phase gate supplied the necessary non-Gaussian effects. However, both of these unitaries are challenging to realize in the optical domain.

Due to this difficulty, postselected PNR-measurement-based optical approaches have also been explored to the same end [45,46]. A major drawback to these approaches is their reliance on postselection of a specific photon number detection pattern, which leads to both low success rates and unfavorable optimization time. The latter stems from the fact that, in the worst case, one must optimize over the set of all possible photon number detection patterns to find patterns that result in states with both high enough fidelity and probability of success.

Our approach does not rely on postselection. Rather, it trains the agent to discover states preparation strategies that are adaptive and robust to the inherent randomness of the PNR measurement results. This also comes with the benefit of not needing to examine all possible detection patterns as the agent can learn the underlying distribution of possible measurements outcomes, given a particular configuration of the environment, from past experience. We successfully demonstrated this proposition for the generation of squeezed cat states, with a different optical circuit [47].

Reinforcement learning framework and implementation: In reinforcement learning, the interaction between the agent and the environment is modeled using the Markov decision process (MDP) formalism. An MDP is characterized by a set of states $\mathcal{S}$ that describe all possible states of the environment, a set of actions $\mathcal{A}$ that the agent can select from to interact with the environment, a scalar reward function $R(s)$ that determines the reward for reaching state $s \in \mathcal{S}$, and a transition function $\mathcal{T}(s'|s, a)$ that determines the probability that the environment will transition to state s' after the agent selects action $a \in \mathcal{A}$ when the environment is in state s . Finally, to be a valid MDP the decision process must also be Markovian, which means that the next state s' that the environment transitions to is solely dependent on the current state s and the action taken by the agent while in that state.

The quantum circuit of Fig. 1 can be viewed as a stochastic environment in which the state of the circuit at time step $j + 1$ only depends on the state of the circuit and the action taken at time step j . The state space $\mathcal{S} \subseteq \mathbb{R}^{N^2}$ consists of the set of all possible flattened density matrices, where N is the Hilbert space truncation. Specifically, we define the state at time step j , s_j , as the vector $s_j = [\text{Re}(\rho_j^u), \text{Im}(\rho_j^u), \text{diag}(\rho_j)]$, where ρ_j^u denotes all entries of ρ_j that are above its diagonal. The action space $\mathcal{A} \subseteq \mathbb{R}^3$ is the set of actions the agent can select. We represent actions as a vector $a_j = [\tau_j, r_j, \alpha_j]$, where τ_j is the beamsplitter transmittivity, r_j is the squeezing parameter of the input squeezed vacuum state, and α_j is the magnitude of

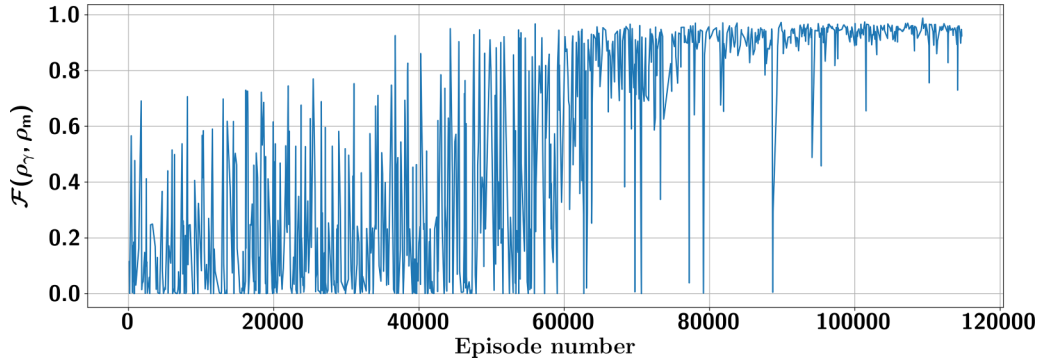


FIG. 2. Fidelity of the terminal state, ρ_m , with the target state, ρ_γ , per episode during lossless training.

the displacement applied to the input density matrix ρ_j . In our case, the transition function probabilities $\mathcal{T}(s_{j+1}|s_j, a_j)$ are given by the diagonal elements of the reduced density matrix of the measured mode prior to the PNR measurement.

We defined the reward function $R(s_j)$ at step j as

$$R(s_j) = \text{Tr}_{10}^{\lambda}(\rho_j) \mathcal{F}^{\lambda}(\rho_j, \rho_\gamma), \quad (2)$$

where ρ_γ is the target density matrix, $\mathcal{F}(\rho_j, \rho_\gamma) = \text{Tr}(\rho_j \rho_\gamma)$ is the pure-state fidelity, and where we set $\lambda = 55$ to tune the penalty on states with fidelities below unity [48]. The trace of density matrix ρ_j was included in $R(s_j)$ to steer the simulation clear of unphysical results stemming from computation errors caused by Hilbert space truncation, here at 31 photons. The target state $\rho_\gamma = |\gamma, r, \alpha\rangle \langle \gamma, r, \alpha|$ was a realistic version of Eq. (1) with an added displacement:

$$|\gamma, r, \alpha\rangle = D(\alpha) e^{i\gamma \hat{Q}^3} S(r) |0\rangle, \quad (3)$$

$|0\rangle$ being the vacuum state, $S(r) = \exp[\frac{r}{2}(\hat{a}^{\dagger 2} - \hat{a}^2)]$ the squeezing operator of parameter r , and $D(\alpha) = \exp(\alpha \hat{a}^{\dagger} - \alpha^* \hat{a})$ the displacement operator by α . The displacement was intended to maximize the state's occupation probability (>0.99) in the 31-photon finite Hilbert space, as per Ref. [44]. The target state's Wigner function was $|\gamma, r, \alpha\rangle = |-0.2, -0.7, i1.25\rangle$ (its Wigner function is plotted in Ref. [49]). We then proceeded to model the interaction between the agent and the quantum circuit as an MDP. Since each episode consisted of a fixed number m of time steps, this constituted a finite-horizon MDP [38].

Our reinforcement learning agent used proximal policy optimization (PPO) [50]. We selected this algorithm due to its on-policy nature, higher robustness to hyperparameter choice compared to other reinforcement learning algorithms [50], and successful use in previous quantum control tasks [40,41]. The on-policy attribute is particularly important for reducing the risk of training instability in our implementation of PPO, which utilizes both deep neural networks and bootstrapping. This is known as the “deadly triad,” which refers to the possibly detrimental, joint action of off-policy learning, universal function approximation, and bootstrapping [38]. A summary of PPO and a more detailed introduction to the relevant reinforcement learning concepts can be found in the Supplemental Material of Ref. [47]. To speed up the training process, we ran 40 environments in parallel for faster training data collection. The initial input state ρ_0 for each episode was a 10 dB squeezed state ($r = 1.15$). The maximum magnitude

of squeezing and displacement the agent can apply at each step were set to $|r_{\max}| = 1.15$ and $|\alpha_{\max}| = 2.25$, respectively. We fixed the displacement before the PNR measurement to $\beta = i2.5$.

Simulation results for cubic-phase state generation. The agent was trained over 5.7 million time steps with episodes of length $m = 50$. Figure 2 shows the fidelity of the state at the end of each episode with the target state over the course of training. The plot indicates that the agent has converged to a policy that maximizes the fidelity of the final state for each episode. Due to the stochastic nature of the environment, there still remain some episodes, which terminate with a state that has less than optimal fidelity. Once trained, we evaluated the agents performance by running five sets of 200 evaluation episodes. Each set of episodes utilized a different random-number seed for simulating PNR detection.

The results from the evaluation are collected in Fig. 3. We observed interesting emergent behaviors from the trained neural network:

(1) After reaching high fidelities [Fig. 3(a)], the agent set the beamsplitter's transmittivity τ_j to 0 and did not change it for the remainder of the episode. Instances of $\tau_j = 0$ at low fidelity did not remain unchanged.

(2) After reaching high fidelities and setting τ_j to 0, the agent frequently selected nonzero values of α_j for a few more steps before eventually setting α_j to 0 as well for the remaining time steps of the episode. We interpreted this behavior as the agent recognizing that the current input state to the loop, ρ_j , was the target cubic-phase state up to corrective displacements which it applied after setting the transmittivity to 0. Note that these “final” displacements sometimes exhibited oscillatory behaviors in the lossy case (see Ref. [49]).

(3) The agent occasionally set τ_j to 1, i.e., “reset” the environment to restart from the initial squeezed state input. This indicates that the agent learned to identify when a particular input state ρ_j is unlikely to be transformed into the target state with further iterations of the loop and thus found it more beneficial to simply reset the circuit and try again. From Fig. 3(c), we can see that in the vast majority of cases no more than ten iterations of the loop were needed before the agent generated a state it was satisfied with.

(4) The agent could also set $\tau_j = 0$ at very low fidelity but not stay at this value in the subsequent steps. These corresponded to some of the failure episodes.

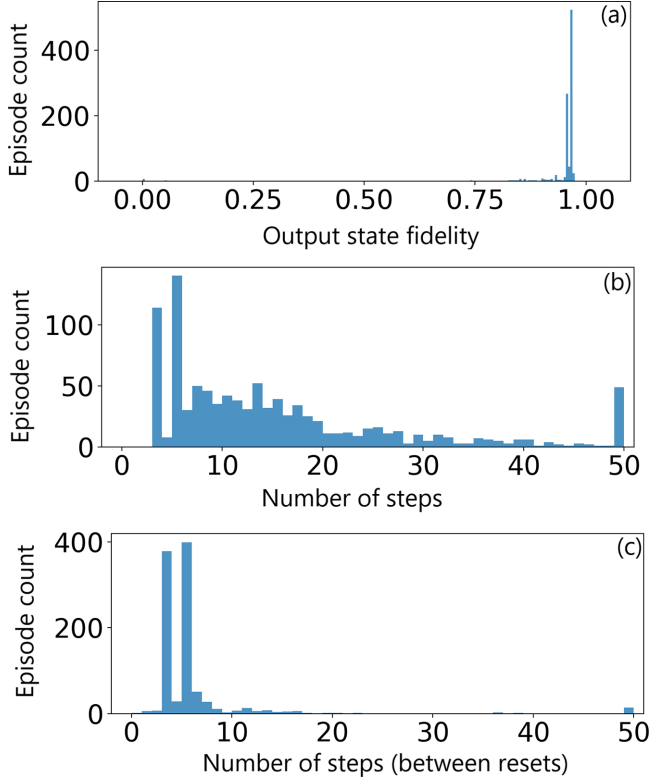


FIG. 3. Histograms for 1000 lossless generation episodes ($m = 50$): (a) output state fidelity; (b) steps per episode, with resets ($\tau_j = 1$), before $\tau_j = 0$; and (c) steps per episode, between resets ($\tau_j = 1$), before $\tau_j = 0$.

(5) About 50 episodes out of 1000 reached the maximum step number of 50 [Fig. 3(c)]. Out of these, 19 were failures, the rest were successful events in which the agent would keep “adjusting” the state with tiny, inconsequential displacements. We believe that more training and hyperparameter tuning should help these cases converge faster.

(6) We counted only 5 instances out of 1000 where the agent appeared “lost,” i.e., stuck at low fidelity at every step with no observable improvement.

We conducted a detailed visual examination of all the resulting Wigner functions from all 1000 episodes (thus crudely assessing minimum negativity), as well as checked the final fidelities, maximum step number, and that $\text{Tr}[\rho_j] \geq 0.98$ at all steps after the last reset. We found that a *bona fide* cubic-phase state was generated 96% of the time, with some variations of α and γ with respect to the target state, sometimes resulting in lower fidelities [Fig. 3(a)]. These variations were not judged detrimental for two reasons: First, they constituted *bona fide* cubic-phase states; second, all output states can be uniquely tagged by their record of measurement and parameter values, which enables the creation of lookup tables that can uniquely identify the generated state at every turn. We address experimental implementation using lookup tables in Sec. IV of the Supplemental Material [49]. Because of all this, all outputs are usable. Note finally that concatenating two cubic gates of parameters γ_1 and γ_2 generates, in principle, a gate of parameter $\gamma_1 + \gamma_2$ and storing cubic-phase states in quantum memories (or delay lines) can allow “breeding” larger ones.

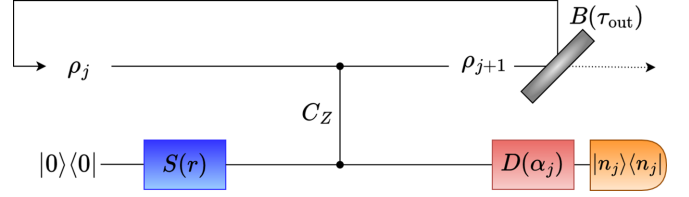


FIG. 4. Quantum circuit for quartic state preparation. $\rho_0 = S(r)|0\rangle\langle 0|S^\dagger(r)$.

We also evaluated the performance of the agent when photon loss was present in the circuit and present the results in Ref. [49].

Quartic-phase state generation. It is known that the quartic-phase gate $\exp(i\delta Q^4)$ of parameter δ can be realized by assembling 29 gates, including 15 cubic-phase gates [19]. One question is whether the methods presented above would not be applicable to the *direct* generation of a quartic-phase state. Our full numerical ML simulations of this case are still a work in progress as they require considerable run-time and error optimization compared to the cubic-phase case. (One main reason for this is the necessity to raise the photon number cutoff from 31 to at least 60, if not higher.) However, we have identified a fundamental, if postselected, quantum optical algorithm that can form the basis of subsequent, deterministic ML implementation. We present this algorithm here.

Quantum algorithm: The intuition is to “stamp” the Wigner function with a displaced Fock state at nearly opposite azimuths in phase space, so as to generate the required cubic-polynomial contour of a quartic-phase state. This can be effected by the circuit presented in Fig. 4. Here, a big difference is that the squeezing is held constant, the only adjustable parameter being the phase of the displacement before the PNR. Posing the PNR displacements $\alpha_{1,2} = |\alpha_{1,2}| \exp(i\phi_{1,2})$, we took $|\alpha_{1,2}| = 4.5$, $\phi_1 = \pi$, and $\phi_2 \simeq 0$ a nonzero adjustable parameter. ($\phi_2 = 0$ is too symmetric and produces zero quarticity.) The constant squeezing in Fig. 4 was set to $r = 1.38$ (12 dB).

Postselected preliminary results: We define the target quartic-phase state as

$$|\delta, s\rangle = e^{i\delta Q^4} S(s)|0\rangle, \quad (4)$$

where δ is the “quarticity” and s is a squeezing parameter. To determine which quartic-phase state was generated by our two-step circuit, we used the Nelder-Mead optimization algorithm to find the values of δ and s that maximized the fidelity between $|\delta, s\rangle$ and the output state from our circuit for each set of postselected photon number measurements. For the simulations in this section, we used a considerably higher Hilbert space truncation of 60 photons. Figure 5 displays one of the best results obtained for the output state and its corresponding optimized quartic-phase state. In an unsurprising confirmation of our original intuition, the best results were obtained for $n_1 = n_2$, which is the reason why this algorithm requires postselection. A gallery of results obtained for $\phi_2 = -0.1, -0.2, -0.3$ rad can be viewed in Ref. [49]. One of the outcomes expected from our upcoming ML extension of this algorithm will clearly be the ability to cope with $n_1 \neq n_2$. The gist of the results in Fig. 5 and Ref. [49] is that engineering

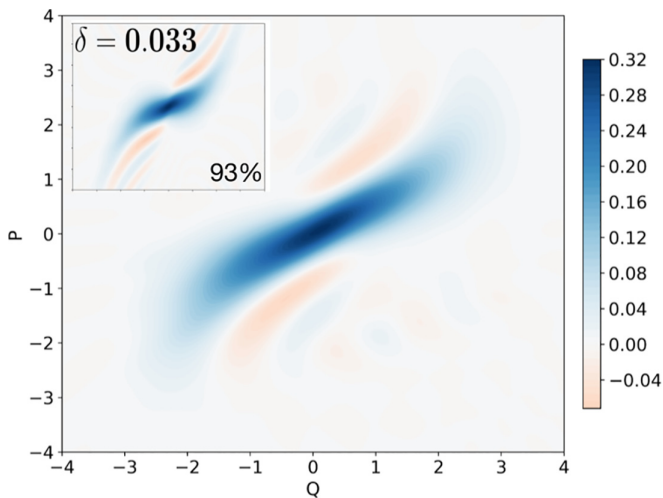


FIG. 5. Resulting Wigner function of two rounds of PNR detection on the circuit in Fig. 4, for $\phi_2 = -0.2$ rad and five photons detected in each round. The top left inset depicts the best-fit Wigner function with its fidelity and quarticity.

quantum interference in phase space [51] is a tenet of quantum state generation, here turning circular Fock state contours into the characteristic cubic-polynomial ripples of a quartic-phase state. This seemed to happen more effectively at higher photon numbers, where the fidelity with a genuine quartic-phase state

was reasonably high, which validated the algorithm. We did verify that this fidelity was always larger than the fidelity to a displaced squeezed state. The value of ϕ_2 seemed to play a role as well and its possible dependence on $n_1 - n_2$ is another landscape that we look forward to ML exploring.

Conclusion and prospective work. We proposed a method based on deep reinforcement learning for the near-deterministic preparation of cubic-phase states. Besides being QC-efficient, this method requires less squeezing and photon number detection capability than proposals so far and can be implemented with components that are readily available in current experimental setups, the only non-Gaussian resource being PNR measurements [30]. We also gave an algorithm for the direct, although probabilistic, generation of quartic-phase gates. The extension of this algorithm to the near-deterministic generation of quartic-phase states, using DRL simulations, is a work in progress. These simulations require a much larger Hilbert space that entails a significant growth of needed computational resources.

Acknowledgments. O.P. was supported by U.S. National Science Foundation grants OSI-2531569 [NQVL-QCAP NQVL:QSTD:Design: Quantum Computing Applications of Photonics (QCAP)], PHY-2514971, and ECCS-2530171. We thank University of Virginia Research Computing for providing access to the Rivanna computing cluster.

Data availability. The code that supports the findings of this article is openly available at [52].

- [1] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, *Phys. Rev. A* **64**, 012310 (2001).
- [2] M. Pysher, Y. Miwa, R. Shahrokshahi, R. Bloomer, and O. Pfister, Parallel generation of quadripartite cluster entanglement in the optical frequency comb, *Phys. Rev. Lett.* **107**, 030505 (2011).
- [3] M. Chen, N. C. Menicucci, and O. Pfister, Experimental realization of multipartite entanglement of 60 modes of a quantum optical frequency comb, *Phys. Rev. Lett.* **112**, 120505 (2014).
- [4] S. Yokoyama, R. Ukai, S. C. Armstrong, C. Sornphiphatphong, T. Kaji, S. Suzuki, J. Yoshikawa, H. Yonezawa, N. C. Menicucci, and A. Furusawa, Ultra-large-scale continuous-variable cluster states multiplexed in the time domain, *Nat. Photon.* **7**, 982 (2013).
- [5] J.-i. Yoshikawa, S. Yokoyama, T. Kaji, C. Sornphiphatphong, Y. Shiozawa, K. Makino, and A. Furusawa, Invited article: Generation of one-million-mode continuous-variable cluster state by unlimited time-domain multiplexing, *APL Photonics* **1**, 060801 (2016).
- [6] W. Asavanant, Y. Shiozawa, S. Yokoyama, B. Charoensombutamon, H. Emura, R. N. Alexander, S. Takeda, J.-i. Yoshikawa, N. C. Menicucci, H. Yonezawa, and A. Furusawa, Generation of time-domain-multiplexed two-dimensional cluster state, *Science* **366**, 373 (2019).
- [7] M. V. Larsen, X. Guo, C. R. Breum, J. S. Neergaard-Nielsen, and U. L. Andersen, Deterministic generation of a two-dimensional cluster state, *Science* **366**, 369 (2019).
- [8] C. Roh, G. Gwak, Y.-D. Yoon, and Y.-S. Ra, Generation of three-dimensional cluster entangled state, *Nat. Photon.* **19**, 526 (2025).
- [9] Z. Yang, M. Jahanbozorgi, D. Jeong, S. Sun, O. Pfister, H. Lee, and X. Yi, A squeezed quantum microcomb on a chip, *Nat. Commun.* **12**, 4781 (2021).
- [10] M. Jahanbozorgi, Z. Yang, S. Sun, H. Chen, R. Liu, B. Wang, and X. Yi, Generation of squeezed quantum microcombs with silicon nitride integrated photonic circuits, *Optica* **10**, 1100 (2023).
- [11] Z. Wang, K. Li, Y. Wang, X. Zhou, Y. Cheng, B. Jing, F. Sun, J. Li, Z. Li, B. Wu, Q. Gong, Q. He, B.-B. Li, and Q.-F. Yang, Large-scale cluster quantum microcombs, *Light Sci. Appl.* **14**, 164 (2025).
- [12] X. Jia, C. Zhai, X. Zhu, C. You, Y. Cao, X. Zhang, Y. Zheng, Z. Fu, J. Mao, T. Dai, L. Chang, X. Su, Q. Gong, and J. Wang, Continuous-variable multipartite entanglement in an integrated microcomb, *Nature (London)* **639**, 329 (2025).
- [13] S. Lloyd and S. L. Braunstein, Quantum computation over continuous variables, *Phys. Rev. Lett.* **82**, 1784 (1999).
- [14] N. C. Menicucci, Fault-tolerant measurement-based quantum computing with continuous-variable cluster states, *Phys. Rev. Lett.* **112**, 120504 (2014).
- [15] K. Marshall, R. Pooser, G. Siopsis, and C. Weedbrook, Quantum simulation of quantum field theory using continuous variables, *Phys. Rev. A* **92**, 063825 (2015).
- [16] R. A. Briceño, R. G. Edwards, M. Eaton, C. González-Arciniegas, O. Pfister, and G. Siopsis, Toward coherent quantum computation of scattering amplitudes with a measurement-based photonic quantum processor, *Phys. Rev. Res.* **6**, 043065 (2024).
- [17] R. L. Hudson, When is the Wigner quasi-probability density non-negative? *Rep. Math. Phys.* **6**, 249 (1974).

- [18] S. Sefi and P. van Loock, How to decompose arbitrary continuous-variable quantum operations, *Phys. Rev. Lett.* **107**, 170501 (2011).
- [19] T. Kalajdzievski and J. M. Arrazola, Exact gate decompositions for photonic quantum computing, *Phys. Rev. A* **99**, 022341 (2019).
- [20] N. Budinger, A. Furusawa, and P. van Loock, All-optical quantum computing using cubic phase gates, *Phys. Rev. Res.* **6**, 023332 (2024).
- [21] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, Universal quantum computation with continuous-variable cluster states, *Phys. Rev. Lett.* **97**, 110501 (2006).
- [22] A. Furusawa and P. van Loock, *Quantum Teleportation and Entanglement: A Hybrid Approach to Optical Quantum Information Processing* (John Wiley & Sons, New Jersey, 2011).
- [23] O. Pfister, Continuous-variable quantum computing in the quantum optical frequency comb, *J. Phys. B: At. Mol. Opt. Phys.* **53**, 012001 (2020).
- [24] S. Bartolucci, P. Birchall, H. Bombin, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant, F. Pastawski, T. Rudolph, and C. Sparrow, Fusion-based quantum computation, *Nat. Commun.* **14**, 912 (2023).
- [25] J. E. Bourassa, R. N. Alexander, M. Vasmer, A. Patil, I. Tzitrin, T. Matsuura, D. Su, B. Q. Baragiola, S. Guha, G. Dauphinais, K. K. Sabapathy, N. C. Menicucci, and I. Dhand, Blueprint for a scalable photonic fault-tolerant quantum computer, *Quantum* **5**, 392 (2021).
- [26] P. Renault, P. Yard, R. C. Pooser, M. Eaton, and H. A. Zaidi, End-to-end switchless architecture for fault-tolerant photonic quantum computing, *Quantum* **9**, 1796 (2025).
- [27] Y. Zheng, O. Hahn, P. Stadler, P. Holmvall, F. Quijandría, A. Ferraro, and G. Ferrini, Gaussian conversion protocols for cubic phase state generation, *PRX Quantum* **2**, 010327 (2021).
- [28] R. Yanagimoto, T. Onodera, E. Ng, L. G. Wright, P. L. McMahon, and H. Mabuchi, Engineering a Kerr-based deterministic cubic phase gate via Gaussian operations, *Phys. Rev. Lett.* **124**, 240503 (2020).
- [29] S. Ghose and B. C. Sanders, Non-Gaussian ancilla states for continuous variable quantum computation via Gaussian maps, *J. Mod. Opt.* **54**, 855 (2007).
- [30] M. Eaton, A. Hossameldin, R. J. Birrittella, P. M. Alsing, C. C. Gerry, H. Dong, C. Cuevas, and O. Pfister, Resolution of 100 photons and quantum generation of unbiased random numbers, *Nat. Photon.* **17**, 106 (2023).
- [31] K. Marshall, R. Pooser, G. Siopsis, and C. Weedbrook, Repeat-until-success cubic phase gate for universal continuous-variable quantum computation, *Phys. Rev. A* **91**, 032321 (2015).
- [32] P. Marek, R. Filip, H. Ogawa, A. Sakaguchi, S. Takeda, J.-I. Yoshikawa, and A. Furusawa, General implementation of arbitrary nonlinear quadrature phase gates, *Phys. Rev. A* **97**, 022329 (2018).
- [33] L.-A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, Generation of squeezed states by parametric down conversion, *Phys. Rev. Lett.* **57**, 2520 (1986).
- [34] F. E. Becerra, J. Fan, G. Baumgartner, J. Goldhar, J. T. Kosloski, and A. Migdall, Experimental demonstration of a receiver beating the standard quantum limit for multiple nonorthogonal state discrimination, *Nat. Photon.* **7**, 147 (2013).
- [35] A. E. Lita, A. J. Miller, and S. W. Nam, Counting near-infrared single-photons with 95% efficiency, *Opt. Express* **16**, 3032 (2008).
- [36] C. Cahall, K. L. Nicolich, N. T. Islam, G. P. Lafyatis, A. J. Miller, D. J. Gauthier, and J. Kim, Multi-photon detection using a conventional superconducting nanowire single-photon detector, *Optica* **4**, 1534 (2017).
- [37] M. Endo, T. Sonoyama, M. Matsuyama, F. Okamoto, S. Miki, M. Yabuno, F. China, H. Terai, and A. Furusawa, Quantum detector tomography of a superconducting nanostrip photon-number-resolving detector, *Opt. Express* **29**, 11728 (2021).
- [38] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction* (The MIT Press, Cambridge, 2018).
- [39] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, Human-level control through deep reinforcement learning, *Nature (London)* **518**, 529 (2015).
- [40] S. Borah, B. Sarma, M. Kewming, G. J. Milburn, and J. Twamley, Measurement-based feedback quantum control with deep reinforcement learning for a double-well nonlinear potential, *Phys. Rev. Lett.* **127**, 190403 (2021).
- [41] V. V. Sivak, A. Eickbusch, H. Liu, B. Royer, I. Tsioutsios, and M. H. Devoret, Model-free quantum control with reinforcement learning, *Phys. Rev. X* **12**, 011059 (2022).
- [42] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning* (MIT Press, Cambridge, 2016), Vol. 1.
- [43] J. M. Arrazola, T. R. Bromley, J. Izaac, C. R. Myers, K. Brádler, and N. Killoran, Machine learning method for state preparation and gate synthesis on photonic quantum computers, *Quantum Sci. Technol.* **4**, 024004 (2019).
- [44] M. Kudra, M. Kervinen, I. Strandberg, S. Ahmed, M. Scigliuzzo, A. Osman, D. P. Lozano, M. O. Tholén, R. Borgani, D. B. Haviland, *et al.*, Robust preparation of Wigner-negative states with optimized SNAP-displacement sequences, *PRX Quantum* **3**, 030301 (2022).
- [45] I. Tzitrin, J. E. Bourassa, N. C. Menicucci, and K. K. Sabapathy, Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes, *Phys. Rev. A* **101**, 032315 (2020).
- [46] Y. Yao, F. Miatto, and N. Quesada, Riemannian optimization of photonic quantum circuits in phase and fock space, *SciPost Phys.* **17**, 082 (2024).
- [47] A. Anteneh, L. Brunel, and O. Pfister, Machine learning for efficient generation of universal photonic quantum computing resources, *Optica Quantum* **2**, 296 (2024).
- [48] R. Porotti, A. Essig, B. Huard, and F. Marquardt, Deep reinforcement learning for quantum state preparation with weak nonlinear measurements, *Quantum* **6**, 747 (2022).
- [49] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/wkfp-tf74> for more details on the machine learning model tuning process and experimental imperfections in the state preparation circuit.
- [50] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, Proximal policy optimization algorithms, [arXiv:1707.06347](https://arxiv.org/abs/1707.06347).
- [51] W. P. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH Verlag Berlin GmbH, Berlin, 2001).
- [52] <https://github.com/amanuelanteneh/rl-loop>.