
Learning Quantum Data Distribution via Chaotic Quantum Diffusion Model

Quoc Hoan Tran*, Koki Chinzei, Yasuhiro Endo, Hirotaka Oshima
Quantum Laboratory, Fujitsu Research, Fujitsu Limited
Kawasaki, Kanagawa 211-8588, Japan

Abstract

Generative models for quantum data pose significant challenges but hold immense potential in fields such as chemoinformatics and quantum physics. Quantum denoising diffusion probabilistic models (QuDDPMs) enable efficient learning of quantum data distributions but require high-fidelity random unitary circuits, which demand precise control and exhibit susceptibility to noise on analog hardware. We propose the chaotic quantum diffusion model, a framework that produces projected ensembles generated by chaotic Hamiltonian time evolution, offering a flexible, hardware-compatible diffusion scheme. Requiring only global, time-independent control, our approach reduces implementation overhead across diverse analog quantum platforms while achieving accuracy comparable to that of QuDDPMs. This method enhances trainability and robustness, broadening the applicability of quantum generative models.

1 Introduction

Generative models aim to produce diverse data by learning their underlying distributions, crucial for understanding datasets and synthesizing new data. Quantum circuits, which can generate classically intractable distributions, suggest that generative models using quantum systems may outperform classical models in specific real-world tasks. Consequently, models like quantum generative adversarial networks (QuGANs) [1–3], quantum variational autoencoders (QVAEs) [4, 5], tensor networks [6], and diffusion-based quantum models [7, 8] have been developed for classical data generation. However, their advantages over classical models remain unproven.

Recent advancements underscore the critical role of quantum machine learning (QML) in processing quantum data derived from quantum systems [9]. Unlike classical data such as text and images, which do not require quantum effects for processing, quantum data pose unique challenges due to their quantum mechanical properties. Generating quantum data is fundamental to advancing our understanding of quantum phenomena and harnessing quantum technologies for chemistry, biology, and materials science applications. Moreover, quantum data generation relies on quantum computing resources, making classical generative models insufficient.

Quantum generative models such as QuGANs and QVAEs can be used to prepare a fixed single quantum state [10–12], but are inefficient in generating ensembles of quantum states [13] due to the need for training deep variational quantum circuits (VQCs). The quantum denoising diffusion probabilistic model (QuDDPM) [14] provides a promising approach that resolves the training issues like barren plateaus [15] inherent in the variational quantum algorithms. Inspired by classical denoising diffusion probabilistic models, QuDDPM learns a map from a noisy and unstructured distribution of quantum states to the structured distribution \mathcal{E}_0 , utilizing quantum scrambling for

*tran.quochoan@fujitsu.com

the forward diffusion and VQCs with projective measurements for the backward denoising process. QuDDPM employs sufficient circuit layers to ensure expressivity while introducing intermediate training steps that smoothly interpolate between the target distribution and noise, mitigating barren plateaus and enabling efficient training. However, the diffusion process in QuDDPM requires high-fidelity scrambling random unitary circuits (RUCs), demanding implementation challenges of precise spatio-temporal control. We propose the chaotic quantum diffusion model to eliminate the need for RUCs, enabling a flexible implementation across quantum hardware.

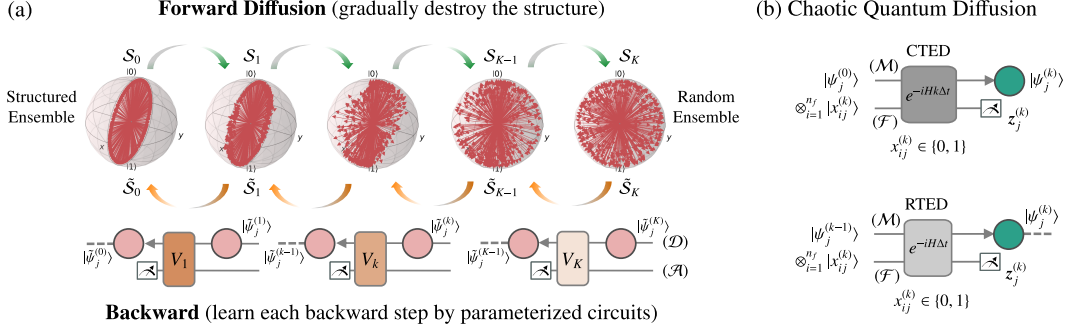


Figure 1: The general scheme of quantum denoising diffusion probabilistic model (a) and the implementation of chaotic quantum diffusion in our proposal (b).

2 Learning quantum data distribution

We address the problem of learning quantum data distributions and explain the idea of the conventional QuDDPM model. Consider a training dataset \mathcal{S} of N independent quantum states drawn from an unknown probability distribution \mathcal{E}_0 . A generative model is characterized by a parameterized probability distribution \mathcal{E}_θ , from which samples can be drawn. If \mathcal{E}_θ is realized through parameterized quantum circuits, θ corresponds to the adjustable parameters of these circuits. The training goal is to obtain a distribution \mathcal{E}_θ that closely approximates \mathcal{E}_0 . We minimize the distance $D(\mathcal{E}_\theta, \mathcal{E}_0)$ such as the Maximum Mean Discrepancy (MMD) and Wasserstein distance between these distributions. Since $D(\mathcal{E}_\theta, \mathcal{E}_0)$ cannot be computed directly, we take a dataset \mathcal{S}_θ sampling from \mathcal{E}_θ and compute the distance $D(\mathcal{S}, \mathcal{S}_\theta)$. In the inference phase, we fix the trained θ and generate new states $|\psi\rangle \sim \mathcal{E}_\theta$.

2.1 Forward diffusion process

In the forward diffusion process, K random unitary gates $U_1^{(j)}, \dots, U_K^{(j)}$ are applied to each sample $|\psi_j^{(0)}\rangle$ ($j = 1, \dots, N$) in the training dataset \mathcal{S} . This evolves the ensemble $\mathcal{S}_k = \{|\psi_j^{(k)}\rangle = \prod_{l=1}^k U_l^{(j)} |\psi_j^{(0)}\rangle\}_j$ toward a Haar-random states ensemble over the Hilbert space. We define this scheme as RUCs diffusion (RUCD). Assuming the execution time for each $U_k^{(j)}$ is τ_u , the execution time to generate each \mathcal{S}_k is $\tau_u N k$. Therefore, the RUCD requires NK random unitary gates with $\tau_u N \sum_{k=1}^K k = \tau_u N K(K+1)/2$ execution time to generate all \mathcal{S}_k .

2.2 Backward denoising process

The backward process starts with an ensemble $\tilde{\mathcal{S}}_K = \{|\tilde{\psi}_j^{(K)}\rangle\}_j$ sampled from Haar random states, and reduces noise step by step. As depicted in Fig. 1(a), the denoising step applies a parameterized unitary $V_k = V(\theta_k)$ to the data system \mathcal{D} (input $|\tilde{\psi}_j^{(k)}\rangle$) and n_a ancilla qubits in the ancilla system \mathcal{A} ($|0\rangle_{\mathcal{A}}$), followed by projective measurements in the computational basis on \mathcal{A} , yielding the state $|\tilde{\psi}_j^{(k-1)}\rangle$ in \mathcal{D} . Assuming the measurement outcome $z_j^{(k)}$ is obtained on the ancilla, this operation is formulated as

$$\Phi_j^{(k)}(|\tilde{\psi}_j^{(k)}\rangle) = \frac{(\mathbf{I}_{\mathcal{D}} \otimes \Pi_{\mathcal{A}}) V_k |\tilde{\psi}_j^{(k)}\rangle}{\sqrt{\langle \tilde{\Psi}_j^{(k)} | V_k^\dagger (\mathbf{I}_{\mathcal{D}} \otimes \Pi_{\mathcal{A}}) V_k |\tilde{\Psi}_j^{(k)} \rangle}} = |\tilde{\psi}_j^{(k-1)}\rangle \otimes |z_j^{(k)}\rangle_{\mathcal{A}}, \quad (1)$$

where $\Pi_{\mathcal{A}} = |\mathbf{z}_j^{(k)}\rangle\langle\mathbf{z}_j^{(k)}|_{\mathcal{A}}$ and $|\tilde{\Psi}_j^{(k)}\rangle = |\tilde{\psi}_j^{(k)}\rangle \otimes |\mathbf{0}\rangle_{\mathcal{A}}$.

Training involves K cycles with a layerwise scheme. At the cycle $(K-k+1)$ (with $k = K, \dots, 1$), the forward diffusion with $U_1^{(j)}$ to $U_{k-1}^{(j)}$ generates the noisy ensemble $\mathcal{S}_{k-1} = \{|\psi_j^{(k-1)}\rangle\}_j$. Parameters of $V(\theta_k)$ are optimized to make the denoised ensemble $\tilde{\mathcal{S}}_{k-1} = \{|\tilde{\psi}_j^{(k-1)}\rangle\}_j$ approximate \mathcal{S}_{k-1} via the minimization of the cost function $D(\mathcal{S}_{k-1}, \tilde{\mathcal{S}}_{k-1})$. After optimization, the parameters θ_k are fixed for use in the next cycle to optimize θ_{k-1} . This layerwise training approach divides the original training problem into K manageable sub-tasks, ensuring convergence for incremental distribution transitions and mitigating issues such as barren plateaus [15].

The cost function in QuDDPM measures the similarity between two quantum state ensembles using a symmetric, positive definite quadratic kernel $\kappa(|\mu\rangle, |\phi\rangle)$. This kernel can be defined by the state fidelity computed via the SWAP test. We consider two cost functions D_{MMD} and D_{Wass} , corresponding to MMD distance and 1-Wasserstein distance based on the state fidelity.

The MMD distance between two state ensembles $\mathcal{X} = \{|\mu_i\rangle\}$ and $\mathcal{Y} = \{|\psi_j\rangle\}$ is defined as

$$\mathcal{D}_{\text{MMD}}(\mathcal{X}, \mathcal{Y}) = \bar{\kappa}(\mathcal{X}, \mathcal{X}) + \bar{\kappa}(\mathcal{Y}, \mathcal{Y}) - 2\bar{\kappa}(\mathcal{X}, \mathcal{Y}), \quad (2)$$

where $\bar{\kappa}(\mathcal{X}, \mathcal{Y}) = \mathbb{E}_{|\mu\rangle \in \mathcal{X}, |\phi\rangle \in \mathcal{Y}}[\kappa(|\mu\rangle, |\phi\rangle)]$. The 1-Wasserstein distance is further presented as an enhancement in the situation where the MMD distance is not feasible to distinguish two state ensembles. Given normalized κ (i.e., $\kappa(|\phi\rangle, |\phi\rangle) = 1 \forall |\phi\rangle$), the pairwise cost matrix $\mathbf{C} = (C_{i,j}) \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{Y}|}$ is computed as $C_{i,j} = 1 - \kappa(|\mu_i\rangle, |\psi_j\rangle)$. The 1-Wasserstein distance is calculated via the formulation of the optimal transport problem into a linear programming procedure to find the optimal transport plan $\mathbf{P} = (P_{i,j}) \in \mathbb{R}^{|\mathcal{X}| \times |\mathcal{Y}|}$:

$$\mathcal{D}_{\text{Wass}}(\mathcal{X}, \mathcal{Y}) = \min_{\mathbf{P}} \sum_{i,j} P_{i,j} C_{i,j}, \text{ s.t. } \mathbf{P}\mathbf{1}_{|\mathcal{Y}|} = \mathbf{a}, \quad \mathbf{P}^\top \mathbf{1}_{|\mathcal{X}|} = \mathbf{b}, \quad \mathbf{P} \geq 0. \quad (3)$$

Here, $\mathbf{1}_{|\mathcal{X}|}$ and $\mathbf{1}_{|\mathcal{Y}|}$ are all-ones vectors with size $|\mathcal{X}|$ and $|\mathcal{Y}|$, respectively, and $\mathbf{a} \in \mathbb{R}^{|\mathcal{X}|}$ and $\mathbf{b} \in \mathbb{R}^{|\mathcal{Y}|}$ are the probability vectors histogram corresponding to \mathcal{X} and \mathcal{Y} . Normally, we set uniform histograms as $\mathbf{a} = \frac{1}{|\mathcal{X}|} \mathbf{1}_{|\mathcal{X}|}$ and $\mathbf{b} = \frac{1}{|\mathcal{Y}|} \mathbf{1}_{|\mathcal{Y}|}$.

3 Proposal: Chaotic Quantum Diffusion

RUCD demands significant computational overhead in designing sequences of random gates, rendering it unsuitable for many quantum systems, particularly analog platforms such as Rydberg atom arrays and ultracold atoms in optical lattices. These platforms excel in simulating continuous-time dynamics via global Hamiltonians but lack the fine-grained, time-dependent control needed for arbitrary gate sequences. Implementing RUCD on analog hardware would require digitizing the process through Trotterization or pulse shaping, which introduces approximation errors and increases susceptibility to noise from environmental couplings.

We address this challenge by adopting the *projected ensemble* framework [16–18], which utilizes a single chaotic many-body wave function to generate a random ensemble of pure states on a subsystem. In this framework, projective measurements are performed on the larger subsystem of a bipartite state undergoing quantum chaotic evolution. This process yields a set of pure states on the smaller subsystem, accompanied by their respective Born probabilities. Collectively, these states form the projected ensemble, which converges to a state design when the measured subsystem is sufficiently large. A chaotic Hamiltonian in this context is characterized by non-integrability, ergodic dynamics, spectral properties resembling those of random matrix theory, and strong entanglement generation. This approach offers significant advantages over RUCD, as it requires only global and time-independent control. Consequently, it is more accessible and adaptable to a wide range of quantum systems, including analog platforms, where precise gate sequences are challenging to implement.

3.1 Projected Ensemble

We consider a many-body system partitioned into a subsystem \mathcal{M} (with n_m qubits) and its complement \mathcal{F} (with n_f qubits). Let a generator state $|\Phi\rangle$ be chosen uniformly at random from the

Hilbert space on the total system $\mathcal{M} + \mathcal{F}$. We perform local measurements on \mathcal{F} , typically in the computational basis. This yields different pure states $|\Phi_{\mathcal{M}}(z_{\mathcal{F}})\rangle$ on \mathcal{M} , each corresponding to a distinct measurement outcome $z_{\mathcal{F}}$, which are bitstrings of the form, for example, $z_{\mathcal{F}} = 001 \dots 010$. The collection of these states, together with probabilities $p(z_{\mathcal{F}})$, forms the projected ensemble on \mathcal{M} : $\{|\Phi_{\mathcal{M}}(z_{\mathcal{F}})\rangle, p(z_{\mathcal{F}})\}_{z_{\mathcal{F}}}$. The projected ensemble provides a full description of the total system state as $|\Psi\rangle = \sum_{z_{\mathcal{F}}} \sqrt{p(z_{\mathcal{F}})} |\Phi_{\mathcal{M}}(z_{\mathcal{F}})\rangle \otimes |z_{\mathcal{F}}\rangle$. For a case of infinite-temperature thermalization, with sufficiently large $n_f = \Omega(kn_m)$ and the generator state $|\Phi\rangle$ obtained by quenched time evolution of chaotic Hamiltonians, the projected ensemble approximates k -design of Haar-random states [16, 17]. The generator state chaotic dynamics ensure that quantum information is scrambled, resulting in universal correlations and randomness within the subsystem. Building on this framework, we propose two schemes in which the diffusion is implemented through chaotic Hamiltonian evolution [Fig. 1(b)].

3.2 Cumulative time evolution diffusion (CTED)

For each $|\psi_j^{(0)}\rangle \in \mathcal{S}_0$ on \mathcal{M} , we implement a diffusion process to obtain the k -step state $|\psi_j^{(k)}\rangle \in \mathcal{S}_k$. The initial state $|\mathbf{x}_j^{(k)}\rangle$ in \mathcal{F} is randomly sampled from the ensemble $\{|\mathbf{x}\rangle, q(\mathbf{x})\}_{\mathbf{x} \in \{0,1\}^{n_f}}$ of computational basis states with probability $q(\mathbf{x})$. The input $|\psi_j^{(0)}\rangle \otimes |\mathbf{x}_j^{(k)}\rangle$ of the system $(\mathcal{M} + \mathcal{F})$ is then evolved cumulatively under the same unitary $e^{-iHk\Delta t}$ with $\forall j$, where H is a fixed chaotic Hamiltonian [17]. Subsequently, a random local measurement is performed on \mathcal{F} , yielding the measurement record $z_j^{(k)}$, and the resulting state $|\psi_j^{(k)}\rangle$ on \mathcal{M} . Since the execution time for $e^{-iHk\Delta t}$ is $k\tau_c$. This CTED requires K unitaries and $\tau_c N \sum_{k=1}^K k = \tau_c N K(K+1)/2$ execution time to generate all \mathcal{S}_k .

In Fig. 2, we depict the schematic circuit to compute the state fidelity between the forward state $|\psi_j^{(k-1)}\rangle$ and the denoised state $|\tilde{\psi}_j^{(k-1)}\rangle$ can be computed using the SWAP test. The SWAP test consists of two Hadamard gates and a controlled-swap gate applied on $2n_m + 1$ qubits. The probability of measure 0 on the SWAP register is $p(0) = \frac{1}{2} + \frac{1}{2} \left| \langle \psi_j^{(k-1)} | \tilde{\psi}_j^{(k-1)} \rangle \right|^2$, which directly derives the fidelity. A post-selection protocol is applied to the projective measurements in the forward and backward processes, ensuring $|\psi_j^{(k-1)}\rangle$ and $|\tilde{\psi}_j^{(k-1)}\rangle$ remain consistent across SWAP test measurements.

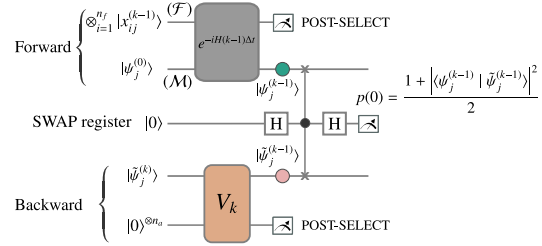


Figure 2: The schematic circuit to compute the fidelity between the forward state $|\psi_j^{(k-1)}\rangle$ and the denoised state $|\tilde{\psi}_j^{(k-1)}\rangle$ using the SWAP test.

3.3 Repeated time evolution diffusion (RTED)

Different from the CTED, we consider the input of $(\mathcal{M} + \mathcal{F})$ as $|\psi_j^{(k-1)}\rangle \otimes |\mathbf{x}_j^{(k)}\rangle$, where $|\psi_j^{(k-1)}\rangle \in \mathcal{S}_{k-1}$, then evolve this state under the unitary $e^{-iH\Delta t}$ with $\forall j$. This process is then repeated for $k = 1, \dots, K$. Assuming the execution time for $e^{-iH\Delta t}$ is τ_r , to generate each \mathcal{S}_k , we need to repeat $e^{-iH\Delta t}$ for k times, making the execution time is $k\tau_r$. Therefore, this RTED requires only one unitary but $\tau_r N \sum_{k=1}^K k = \tau_r N K(K+1)/2$ execution time to generate all \mathcal{S}_k .

4 Results

We conduct numerical experiments on generating *cluster* and *circular* ensembles of quantum states with $n_m = 4$ qubits. The cluster ensemble is a set of quantum states near the zero state, in which the quantum state has the form of $|\psi\rangle = \frac{1}{\sqrt{1+\epsilon^2 \sum_j |r_j|^2}} \left(|0\dots 0\rangle + \epsilon \sum_{j=1}^{2^{n_m}-1} r_j |j\rangle \right)$ with the scale parameter $\epsilon = 0.06$. Here, $|j\rangle$ represents the computational basis states excluding $|0\dots 0\rangle$, and $r_j = x_j + iy_j$ are complex numbers with $x_j, y_j \sim \mathcal{N}(0, 1)$. For the circular ensemble, the quantum state has the form of $|\psi(\beta)\rangle = \cos(\beta/2) |0\dots 0\rangle + \sin(\beta/2) |1\dots 1\rangle$ of n_m qubits, where $\beta \in [0, 2\pi)$.

For CTED and RTED, we consider a complement system \mathcal{F} with $n_f = 4$ qubits. The dynamics are governed by a one-dimensional mixed-field Ising Hamiltonian with open boundary conditions, defined on the total system comprising $n_m + n_f$ sites as $H = \sum_{j=1}^{n_m+n_f} (h^x S_j^x + h^y S_j^y) + J \sum_{j=1}^{n_m+n_f-1} S_j^x S_{j+1}^x$. Here, S_j^μ (with $\mu = x, y, z$) denotes the spin-1/2 operators at site j , J represents the interaction strength, and h^x and h^y are the strengths of the longitudinal and transverse magnetic fields, respectively. In the presence of a non-zero longitudinal field ($h^x \neq 0$), the Hamiltonian exhibits ergodic behavior [17], with its eigenvalues and eigenvectors conforming to the predictions of the Eigenstate Thermalization Hypothesis [19, 20]. To model the non-integrable regime, we adopt $h^x = 0.8090$, $h^y = 0.9045$, and $J = 1.0$. The time evolution is discretized with a time step of $\Delta t = 0.01$ over $K = 50$ diffusion steps.

For RUCD, we adopt the fast scrambling circuits from Ref. [14], applying $\prod_{l=1}^k U_k^{(j)}$ to each initial state $|\psi_j^{(0)}\rangle$. Here, $U_k^{(j)} = \Omega_k(s_k^{(j)})W_k(g_k^{(j)})$, where $W_k(g_k^{(j)})$ implements single-qubit rotations as $W_k(g_k^{(j)}) = \bigotimes_{q=1}^{n_m} e^{-ig_{k,3q-1}^{(j)} \frac{Z_q}{2}} e^{-ig_{k,3q-2}^{(j)} \frac{Y_q}{2}} e^{-ig_{k,3q-3}^{(j)} \frac{Z_q}{2}}$, and an entangling layer $\Omega_k(s_k^{(j)})$ applies ZZ rotations across all qubit pairs as $\Omega_k(s_k^{(j)}) = \prod_{q_1 < q_2} \exp \left[\frac{-is_k^{(j)}}{2\sqrt{n_m}} Z_{q_1} Z_{q_2} \right]$. The ranges of uniformly random rotation angles $g_{k,q}^{(j)}$ and $s_k^{(j)}$ are tuned as $[-\pi/8, \pi/8]$ and $[0.4, 0.6]$, respectively. In the backward process with n_a ($n_a = 4$) ancilla qubits, each V_k is constructed using a Hardware Efficient Ansatz on $n = n_m + n_a$ qubits with L ($L = 10$) layers as $V_k(\theta_k) = \prod_{l=1}^L \tilde{\Omega}_k \tilde{W}_k(\theta_k)$, where $\tilde{W}_k(\theta_k) = \bigotimes_{q=1}^n e^{-i\theta_{k,2q-1} \frac{Y_q}{2}} e^{-i\theta_{k,2q-2} \frac{X_q}{2}}$ and $\tilde{\Omega}_k = \bigotimes_{q=1}^{\lfloor (n-1)/2 \rfloor} CZ_{2q,2q+1} \bigotimes_{q=1}^{\lfloor n/2 \rfloor} CZ_{2q-1,2q}$. The training to optimize each V_k involves 500 samples over 2000 epochs with MMD and 1-Wasserstein distance cost functions for cluster ensemble and circular ensemble, respectively.

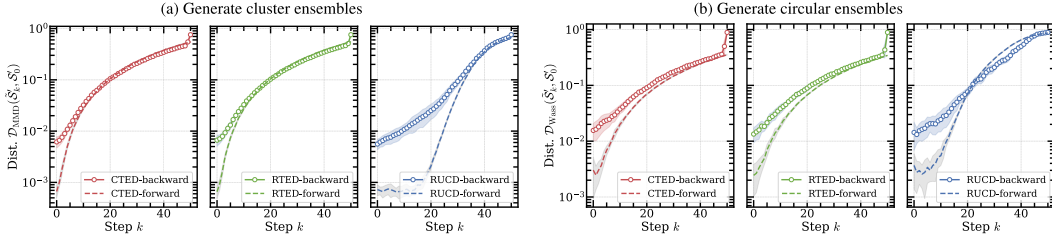


Figure 3: MMD and 1-Wasserstein distances between the generated ensemble \tilde{S}'_k and the true ensemble S'_0 , sampled from (a) cluster and (b) circular datasets of quantum states, for CTED, RTED, and RUCD. Lines and shaded areas represent the mean and standard deviation over twenty trials.

In Fig. 3, we illustrate the variation in MMD and 1-Wasserstein distances between the generated ensemble \tilde{S}'_k and the true ensemble S'_0 for the backward process across the CTED, RTED, and RUCD. For the forward processes, we compute the distance between the diffused ensemble and the true ensemble S'_0 . At $k = 0$, the backward processes of all methods yield comparable distances to the true ensemble, with CTED and RTED exhibiting greater implementation efficiency. The forward diffusion behaviors of CTED and RTED diverge from that of RUCD, whereas their backward processes more closely align with their respective forward dynamics compared to RUCD. This discrepancy likely stems from the design of the backward, which resembles a projected ensemble scheme.

5 Summary

We have proposed a quantum diffusion model with two schemes, CTED and RTED, using chaotic Hamiltonian evolution. These schemes match RUCD's accuracy in learning quantum data distributions while eliminating its complex spatio-temporal control, making them ideal for analog quantum systems. Future work will optimize forward chaotic dynamics for better backward process trainability, develop efficient loss functions, dequantize to resolve variational parameter training bottlenecks, and design conditioned generative diffusion models for applications like molecular data generation.

References

- [1] Seth Lloyd and Christian Weedbrook. Quantum generative adversarial learning. *Phys. Rev. Lett.*, 121: 040502, Jul 2018. doi: 10.1103/PhysRevLett.121.040502. URL <https://link.aps.org/doi/10.1103/PhysRevLett.121.040502>.
- [2] Christa Zoufal, Aurélien Lucchi, and Stefan Woerner. Quantum generative adversarial networks for learning and loading random distributions. *npj Quantum Information*, 5(103), November 2019. doi: 10.1038/s41534-019-0223-2. URL <http://dx.doi.org/10.1038/s41534-019-0223-2>.
- [3] He-Liang Huang, Yuxuan Du, Ming Gong, Youwei Zhao, Yulin Wu, Chaoyue Wang, Shaowei Li, Futian Liang, Jin Lin, Yu Xu, Rui Yang, Tongliang Liu, Min-Hsiu Hsieh, Hui Deng, Hao Rong, Cheng-Zhi Peng, Chao-Yang Lu, Yu-Ao Chen, Dacheng Tao, Xiaobo Zhu, and Jian-Wei Pan. Experimental quantum generative adversarial networks for image generation. *Phys. Rev. Appl.*, 16:024051, Aug 2021. doi: 10.1103/PhysRevApplied.16.024051. URL <https://link.aps.org/doi/10.1103/PhysRevApplied.16.024051>.
- [4] Amir Khoshaman, Walter Vinci, Brandon Denis, Evgeny Andriyash, Hossein Sadeghi, and Mohammad H Amin. Quantum variational autoencoder. *Quantum Science and Technology*, 4(1):014001, September 2018. doi: 10.1088/2058-9565/aada1f. URL <http://dx.doi.org/10.1088/2058-9565/aada1f>.
- [5] Huaijin Wu, Xinyu Ye, and Junchi Yan. QVAE-mole: The quantum VAE with spherical latent variable learning for 3-d molecule generation. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL <https://openreview.net/forum?id=RqvesBxqDo>.
- [6] Michael L. Wall, Matthew R. Abernathy, and Gregory Quiroz. Generative machine learning with tensor networks: Benchmarks on near-term quantum computers. *Phys. Rev. Res.*, 3:023010, Apr 2021. doi: 10.1103/PhysRevResearch.3.023010. URL <https://link.aps.org/doi/10.1103/PhysRevResearch.3.023010>.
- [7] Andrea Cacioppo, Lorenzo Colantonio, Simone Bordoni, and Stefano Giagu. Quantum diffusion models, 2023. URL <https://arxiv.org/abs/2311.15444>.
- [8] Marco Parigi, Stefano Martina, and Filippo Caruso. Quantum-noise-driven generative diffusion models. *Advanced Quantum Technologies*, page 2300401, 2024. ISSN 2511-9044. doi: 10.1002/qute.202300401.
- [9] Editorial. Seeking a quantum advantage for machine learning. *Nat. Mach. Intell.*, 5(8):813–813, Aug 2023. ISSN 2522-5839. doi: 10.1038/s42256-023-00710-9. URL <http://dx.doi.org/10.1038/s42256-023-00710-9>.
- [10] Murphy Yuezheng Niu, Alexander Zlokapa, Michael Broughton, Sergio Boixo, Masoud Mohseni, Vadim Smelyanskiy, and Hartmut Neven. Entangling quantum generative adversarial networks. *Phys. Rev. Lett.*, 128:220505, Jun 2022. doi: 10.1103/PhysRevLett.128.220505. URL <https://link.aps.org/doi/10.1103/PhysRevLett.128.220505>.
- [11] Leeseok Kim, Seth Lloyd, and Milad Marvian. Hamiltonian quantum generative adversarial networks. *Phys. Rev. Res.*, 6:033019, Jul 2024. doi: 10.1103/PhysRevResearch.6.033019. URL <https://link.aps.org/doi/10.1103/PhysRevResearch.6.033019>.
- [12] Quoc Hoan Tran, Shinji Kikuchi, and Hirotaka Oshima. Variational denoising for variational quantum eigensolver. *Phys. Rev. Res.*, 6:023181, May 2024. doi: 10.1103/PhysRevResearch.6.023181. URL <https://link.aps.org/doi/10.1103/PhysRevResearch.6.023181>.
- [13] Kerstin Beer and Gabriel Müller. Dissipative quantum generative adversarial networks, 2021. URL <https://arxiv.org/abs/2112.06088>.
- [14] Bingzhi Zhang, Peng Xu, Xiaohui Chen, and Quntao Zhuang. Generative quantum machine learning via denoising diffusion probabilistic models. *Phys. Rev. Lett.*, 132:100602, Mar 2024. doi: 10.1103/PhysRevLett.132.100602. URL <https://link.aps.org/doi/10.1103/PhysRevLett.132.100602>.
- [15] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.*, 9(1):4812, 2018. doi: 10.1038/s41467-018-07090-4. URL <https://www.nature.com/articles/s41467-018-07090-4>.
- [16] Joonhee Choi, Adam L. Shaw, Ivaylo S. Madjarov, Xin Xie, Ran Finkelstein, Jacob P. Covey, Jordan S. Cotler, Daniel K. Mark, Hsin-Yuan Huang, Anant Kale, Hannes Pichler, Fernando G. S. L. Brandão, Soonwon Choi, and Manuel Endres. Preparing random states and benchmarking with many-body quantum chaos. *Nature*, 613(7944):468–473, 2023. ISSN 0028-0836. doi: 10.1038/s41586-022-05442-1. URL <https://www.nature.com/articles/s41586-022-05442-1>.

- [17] Jordan S. Cotler, Daniel K. Mark, Hsin-Yuan Huang, Felipe Hernández, Joonhee Choi, Adam L. Shaw, Manuel Endres, and Soonwon Choi. Emergent quantum state designs from individual many-body wave functions. *PRX Quantum*, 4(1):010311, 2023. doi: 10.1103/prxquantum.4.010311. URL <https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.4.010311>.
- [18] Wai-Keong Mok, Tobias Haug, Adam L. Shaw, Manuel Endres, and John Preskill. Optimal conversion from classical to quantum randomness via quantum chaos. *Phys. Rev. Lett.*, 134:180403, May 2025. doi: 10.1103/PhysRevLett.134.180403. URL <https://link.aps.org/doi/10.1103/PhysRevLett.134.180403>.
- [19] Mark Srednicki. Chaos and quantum thermalization. *Phys. Rev. E*, 50:888–901, Aug 1994. doi: 10.1103/PhysRevE.50.888. URL <https://link.aps.org/doi/10.1103/PhysRevE.50.888>.
- [20] Luca D’Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Advances in Physics*, 65(3):239–362, 2016. doi: 10.1080/00018732.2016.1198134. URL <https://doi.org/10.1080/00018732.2016.1198134>.