Quantum Circuit Born Machines in Qiskit: Theory, Benchmarks, and Applications

*Developed by: **Natalie Hawkins** (Mentor, Quantum Algorithms and Applications, QuantumBits), **Jorge Plazas** (Mentee, Escuela Politécnica Superior, Universidad Francisco de Vitoria), and **Debshata Choudhury** (Mentee, B.Tech CSE-ASET Noida; BS Data Science-IIT Madras).*

QISKIT ADVOCATE 2025 2025 MENTORSHIP PROGRAM

INTRODUCTION

Quantum Circuit Born Machines are expressive quantum generative models that represent probability distributions with simple circuit architectures, making them well-suited for unsupervised learning and near-term quantum devices.
This project builds a ***Qiskit-based QCBM toolkit*** and applies it to benchmark experiments including ***three-dice distributions, Bars & Stripes, and molecular data augmentation***.

MAXIMUM MEAN DISCREPANCY (MMD)

Mathematical definition

Given distributions $p, q \in \mathbb{R}^N$ and kernel $K \in \mathbb{R}^{N\times N}$,

$$\mathrm{MMD}^2(p,q) = (p-q)^\top K(p-q)$$

where

$$K_{ij} = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\exp\big(-\gamma(i-j)^2\big)$$

Measures distribution mismatch using kernel-based Maximum Mean Discrepancy.

KL DIVERGENCE

Mathematical definition

$$D_{\mathrm{KL}}(p\|q) = \sum_i p_i \log\frac{p_i}{q_i}$$

with numerical stabilization $p_i, q_i \geq \varepsilon$.

Computes information-theoretic divergence between target and model distributions.

SINKHORN LOSS

Mathematical definition

Given cost matrix $C_{ij} = (x_i - x_j)^2$, solve

$$\min_{T\in\Pi(p,q)} \sum_{i,j} T_{ij}C_{ij} + \varepsilon\sum_{i,j} T_{ij}(\log T_{ij} - 1)$$

with Sinkhorn iterations:

$$K = \exp(-C/\varepsilon), \quad T = \mathrm{diag}(u)\,K\,\mathrm{diag}(v)$$

Evaluates entropy-regularized optimal transport distance between distributions.

THREE DICE QCBM

A QCBM learns the probability distribution of the sum of three dice mapped onto a ***16-state space***.

- ***4 qubits*** model all outcomes with high accuracy (***Fidelity ≈ 0.99, MMD ≈ 10^{-3}***).
- Empirical samples closely match the target distribution, validating stable learning.
- As shown in the figure, the QCBM learns the bell-shaped three-dice distribution, with highest probability on ***mid-range sums (10–11)*** and low mass on extreme outcomes, using ***4 qubits***.

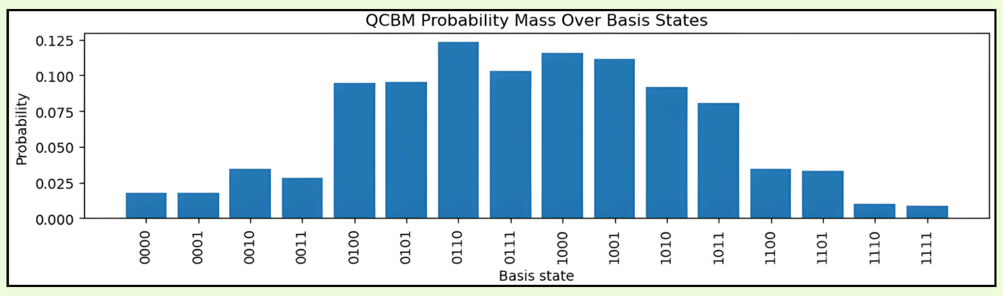

Fig. Learned Dice Distribution

BAS SAMPLE VALIDITY ANALYSIS

A QCBM is trained to learn the structured, multi-modal Bars & Stripes image distribution.

- ***4 qubits (2×2)*** successfully capture the dominant BAS patterns and their relative weights.
- Validity stabilizes at ***χ ≈ 0.76***, highlighting the challenge of perfectly modeling sparse, constraint-heavy distributions.
- As shown in the figure, valid BAS patterns are learned, with some invalid samples (red) indicating constraint leakage.

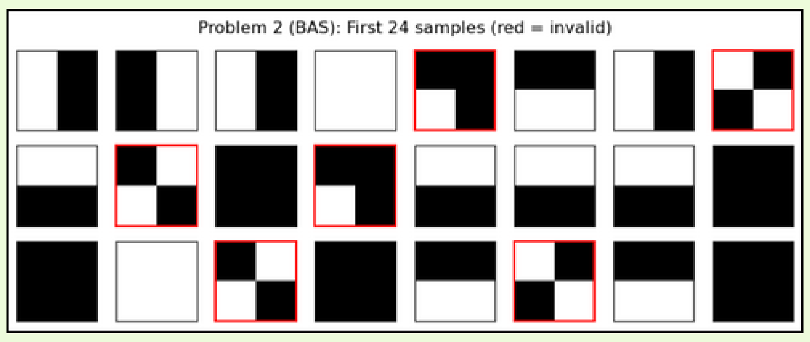

Fig. BAS Sample Validity

LATENT-SPACE QCBM FOR SCALABLE KRAS MOLECULE GENERATION

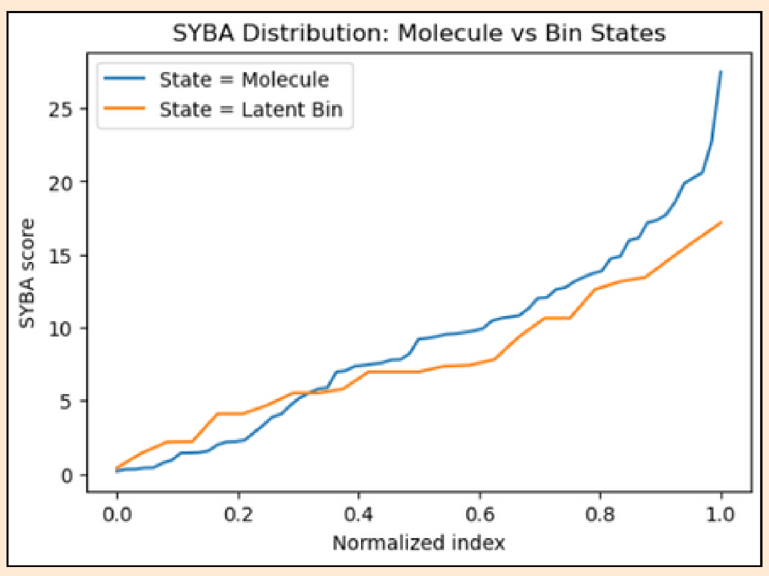

Fig. SYBA Score Comparison

We use a QCBM-inspired probabilistic sampler to generate feasible KRAS-like molecules by learning distributions in a chemically meaningful latent space, avoiding brute-force enumeration.

Pipeline

Known molecules → SELFIES → latent bins → QCBM-style learning → sampling → RDKit + SYBA filtering.

Key findings

- Molecule-as-state: Near-perfect fit ***(MMD ~10^{-26}, Sinkhorn ~10^{-9})*** but non-scalable ***(≈ 10 qubits)***.
- Latent-bin state: Good fit ***(MMD ~4.8×10^{-3}, Sinkhorn ~0.15)*** with half the qubits ***(≈ 5)***.
- Chemical realism: Dozens of sampled molecules pass feasibility filters with matching property distributions.
- No circuits run: QCBM used as a theoretical estimator, showing latent bins make quantum generative chemistry scalable.

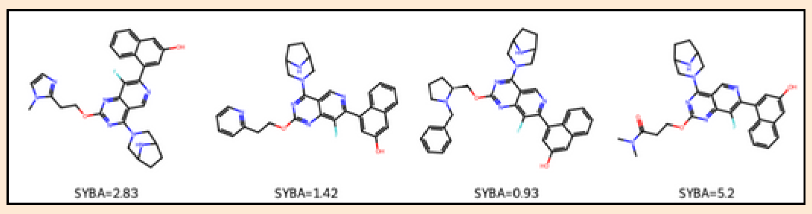

Fig. Molecule State Space

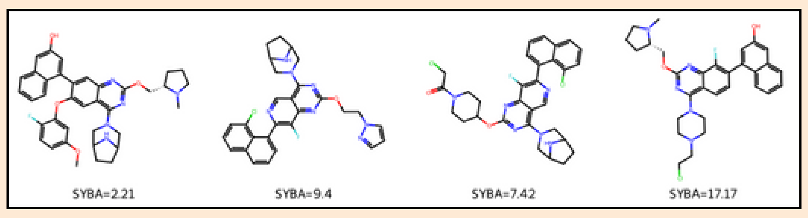

Fig. Latent Bin State Space

CONCLUSION

This work demonstrates that QCBMs can reliably learn both smooth and structured distributions while revealing clear scalability and validity trade-offs.
The developed Qiskit-based framework provides a practical foundation for studying quantum generative models and their near-term applicability.

References: [1] *Vakili et al., Nature Biotechnology (2025)*, [2] *Hamilton, Dumitrescu & Pooser, Physical Review A (2019)*, [3] *Liu & Wang, Physical Review A (2018).*

Github Repository:-
https://github.com/jorgeplazas/Qiskit-QCBMs