quantum reports

MDPI

Article

Entanglement as a Resource for Correlation: A Quantum Circuit Born Machine for System-Wide Flight-Delay Scenario Generation from Real BTS Data

Mariwan Ahmed * and **Ismael Abdulrahman**

Department of Technical Information Systems Engineering, Technical College of Computer and Informatics Engineering, Erbil Polytechnic University, Erbil 44001, Kurdistan Region, Iraq; ismael.abdulrahman@epu.edu.iq
* Correspondence: mariwan.ahmed@epu.edu.iq

Abstract

Network delays are correlated, so treating hub airports as independent can severely understate simultaneous delay events. This study formulates system-wide flight-delay scenario generation with a pure-state quantum circuit Born machine (QCBM), in which each qubit represents one hub and each computational-basis measurement produces one joint binary delay scenario. Within this unitary pure-state model, a non-entangling circuit yields a factorized distribution; entangling operations therefore provide the resource required to represent cross-hub dependence. Using 24 months of U.S. Bureau of Transportation Statistics data, the five-hub exact-probability finite-difference MMD simulation reduces correlation-matrix error from 0.380 for the independent model to 0.028 ± 0.010. A finite-shot SPSA simulation using 8192 shots per circuit evaluation reaches 0.060 ± 0.029 under ideal sampling and 0.066 ± 0.026 after 0.5% noise-injected training with noiseless post-training evaluation. The independent model underestimates the observed system-wide probability by 6.6-fold for five hubs and 116-fold for 10 hubs. However, second-order classical point estimates lie within the observed five-hub bootstrap interval, and classical models evaluated over the complete 32–1024-state spaces fit efficiently. No quantum advantage or physical-hardware demonstration is claimed.

Keywords: entanglement; flight delay; generative modeling; quantum circuit Born machine; quantum machine learning; scenario generation

1. Introduction

Delay is a networked operational risk in air transportation. A disturbance at one hub can propagate through shared aircraft, crews, and passengers, and widespread delays can cross a continental network within hours [1–3]. Robust scheduling therefore requires joint scenarios, not only point forecasts. Such scenarios allow plans to be stress-tested against correlated and tail-heavy network states.

This generative objective is distinct from the usual flight-delay classification, which aims for a point forecast rather than a joint distribution. This distinction matters for quantum machine learning (QML) [4–6]. A related submitted study applies a network-coupled quantum reservoir to flight-delay forecasting [7,8]. More generally, QML performance often depends more on the data and task than on the device [9]. Here, the task is generative: the target is a joint distribution. QCBMs are a natural candidate because their measured Born distributions can represent correlated, multimodal binary states [10,11]. The circuit

Academic Editors: Chao Zheng and Jim Freericks

Received: 15 August 2026
Revised: 29 August 2026
Accepted: 3 September 2026
Published: 10 September 2026

assigns one qubit to each hub and one measured bit string to each joint delay scenario. In the pure-state unitary QCBM studied here, a circuit without entangling operations remains a product state and produces independent bits. Cross-dependence between airports is representable with nonproduct amplitudes in entangling operations and, in practice, requires appropriate angle of rotation and appropriate alignment of the measurement-bases.

The methodological ingredients—Born-machine sampling, MMD training, and the data-processing inequality—are established. The contribution is their integration into a system-wide flight-delay scenario generator and their evaluation on 24 months of real BTS data. The paper interprets observed airport-pair dependence as an information requirement for this pure-state QCBM, and adapts the entangler to a static data-guided coupling map. It then evaluates classical baselines, multiple seeds, finite-shot training, simulated noise, threshold choice, topology, depth, hub count, and season conditioning. This is a feasibility and mechanism study; it does not demonstrate quantum advantage.

2. Background and Related Work

Delay propagation and scenarios. Airport delays propagate and synchronize across the network, producing system-wide events [1,2,12]. Statistical models [13] and graph neural networks [14,15] address delay characterization and network-wide prediction. They are mainly discriminative point-prediction methods. Copulas, Gaussian processes, and deep generative models support classical scenario generation [16], where the central challenge is to reproduce correlated, multimodal joint structure.

Classical generative families for correlated binary data. The target is $P(X)$ over $X \in \{0,1\}^N$. The independent product captures only first-order structure. Gaussian copulas [16,17] and fully visible pairwise Boltzmann machines capture second-order dependence [18]. RBMs add hidden units and can represent higher-order interactions [19]. VAEs, GANs, normalizing flows, autoregressive models, and graph generators become important at higher dimension [20,21]. At the fully observed 32–1024-state scale used here, however; exact maximum-likelihood families already fit nearly exactly (Section 5.1); deep models would not create a meaningful present-scale comparison.

Quantum generative models. QCBMs define distributions through the Born rule and can be trained with MMD [10,22,23]. Their expressivity and learnability have been studied in parameterized-circuit models [24]. Related quantum-generative work includes adversarial generators [25] and generative modeling through quantum correlations [11]. Formal comparisons and hardware demonstrations are cited with the corresponding analyses in Section 4.4. Recent work extends QGANs toward high-resolution images, dual-discriminator quantum-convolutional architectures, and QGAN generators based on quantum Born machines [26–28]. Those studies emphasize adversarial image or general-data generation. In contrast, this work trains a Born distribution directly with MMD for interpretable binary network scenarios and uses no discriminator.

3. Problem, Data, and Experimental Setup

3.1. Network Delay States

In a network of N hubs, each hour represents a delay state $X \in \{0, 1\}^N$, where $X_i = 1$ marks hub i as delayed. A hub counts as delayed in a given hour if its departure delay in that hour exceeds a threshold. The quantity of interest is the joint distribution $P_{data}(X)$ over 2^N possible states. Cross-hub correlation is strongly positive, and the shape is bimodal, clustering at the all-clear and all-delayed states. The term "system-wide risk" is written $P_{sys} = p$ (all N hubs delayed). Consequence modeling (cost, passenger impact, crew disruption) is outside the scope but is an important extension; even a small system-wide probability can matter operationally once weighted by cost.

3.2. Data, Labeling, and Setup

The pipeline aggregates BTS On-Time Performance records [29] (24 monthly files, 2024–2025) into a per-hub table of hourly departure delays. The code creates the table using the monthly raw data files or the original ones. In labeling the hubs, it classifies a hub as having delay if the average hour delay exceeds the median delay of the hub (about 6 to 12 min depending on the hub). The choice of this labeling was deliberate since it ensures that there is equality in the proportion of delays for each hub–about 50 percent. This is the cleanest setting for testing the entanglement-enabled correlation mechanism. The operationally standard fixed 15-minute threshold is reported in parallel (Section 5.4). It yields unbalanced base rates (0.25–0.41) and a stronger operational message (a 22.7-fold independence underestimate). The qualitative conclusions consequently do not depend on the labeling.

The core experiments use N = 5 hubs (ATL, DFW, DEN, ORD, CLT), yielding 13,038 hourly states across all 32 configurations and a mean cross-hub correlation of 0.380. The main circuit has L = 4 entangling layers and 75 trainable angles. Scaling and topology studies extend the network to the 10 busiest hubs by adding PHX, LAX, LAS, SEA, and MCO. Metrics include MMD, mean absolute pairwise-correlation error, KL divergence, total variation (TV), the co-occurrence distribution, and the system-wide probability P_{sys}. All experiments are simulations; no physical quantum processor was used. Figure 1 summarizes the end-to-end workflow, while Figure 2 summarizes the hourly state raster, labeling thresholds, and bimodal training target.

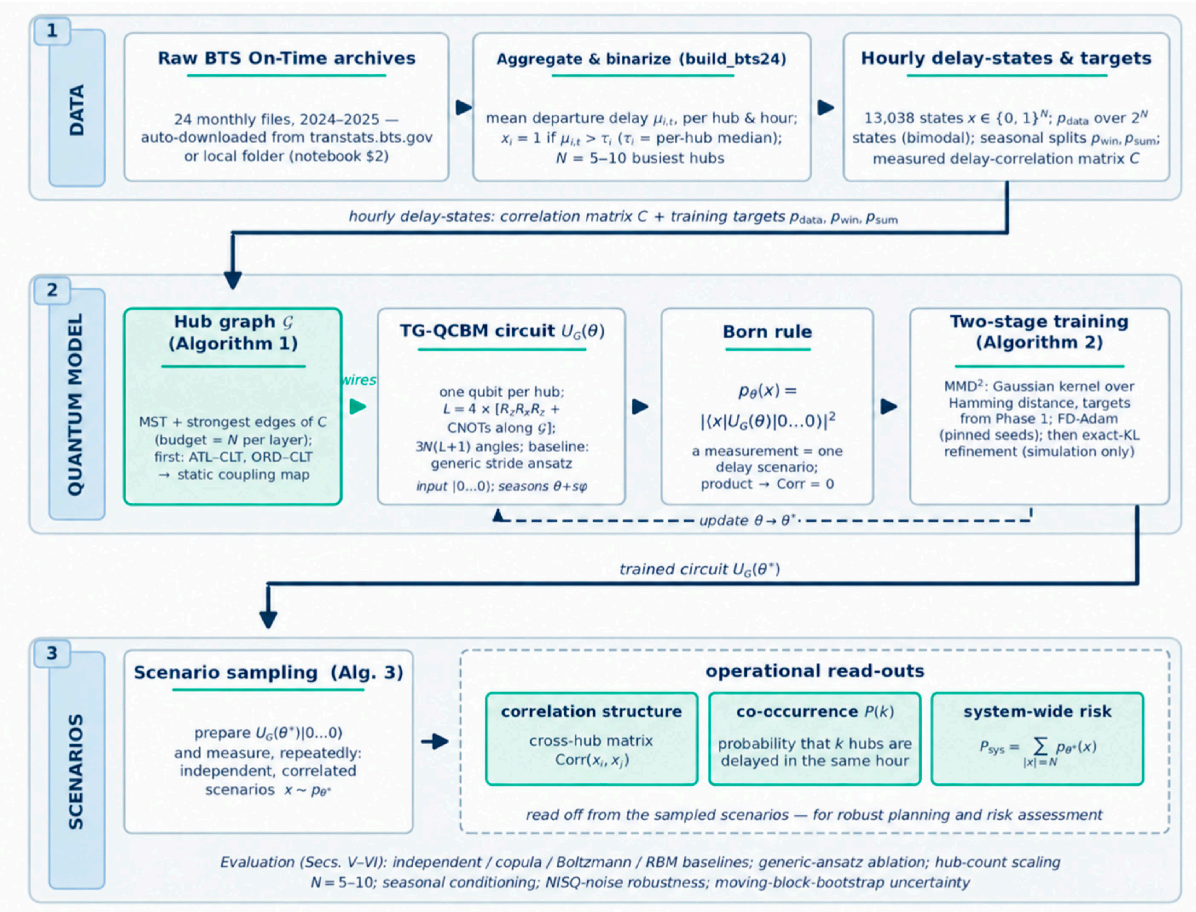

Figure 1. End-to-end methodology. Stage 1: BTS records are aggregated into hourly hub-delay states and an empirical correlation matrix. Stage 2: The Born distribution is trained either by exact-probability finite-difference MMD or by finite-shot SPSA MMD; exact-probability KL refinement is

https://doi.org/10.3390/quantum8030094

retained only as a simulation-side expressivity assessment. Algorithm 1 supplies the TG-QCBM wiring. Stage 3: Circuit measurements generate joint scenarios for correlation, co-occurrence, and system-wide probability analysis. θ^* represents the optimized circuit-parameter vector obtained through training.

Algorithm 1. Build the topology-guided entangler (TG-QCBM)

Input: hourly delay-states $\rightarrow$ corr. matrix C; hubs N; budget B = N edges
```
E ← []; sort pairs (i,j) by |Cij| descending
for (i,j) in pairs:  ▷ maximum spanning tree (connectivity)
  if find(i) ≠ find(j):  union(i,j); E.append((i,j))
for (i,j) in pairs:  ▷ strongest extras up to the budget
  if |E| = B: break
  if (i,j) ∉ E: E.append((i,j))
return CNOT layer along E ▷ identical every round (static coupling map)
```

Uncertainty is reported over independent parameter initializations: 10 seeds for the core summary, five for the main ablations, and three for SPSA. Consecutive hours are serially correlated, so the binomial standard error of P_{sys} (0.0035) is not used for inference. The moving block bootstrap using blocks of size 168 h in order to preserve the weekly structure, with 2000 samples, results in a standard error of 0.0196. The method implies an effective sample size of around 415 h and delivers a 95% confidence interval for 0.205 of [0.167, 0.241]. This interval quantifies sampling uncertainty in the observed P_{sys}; because the models are fitted to the same data, it is not a confidence interval for model–data differences.

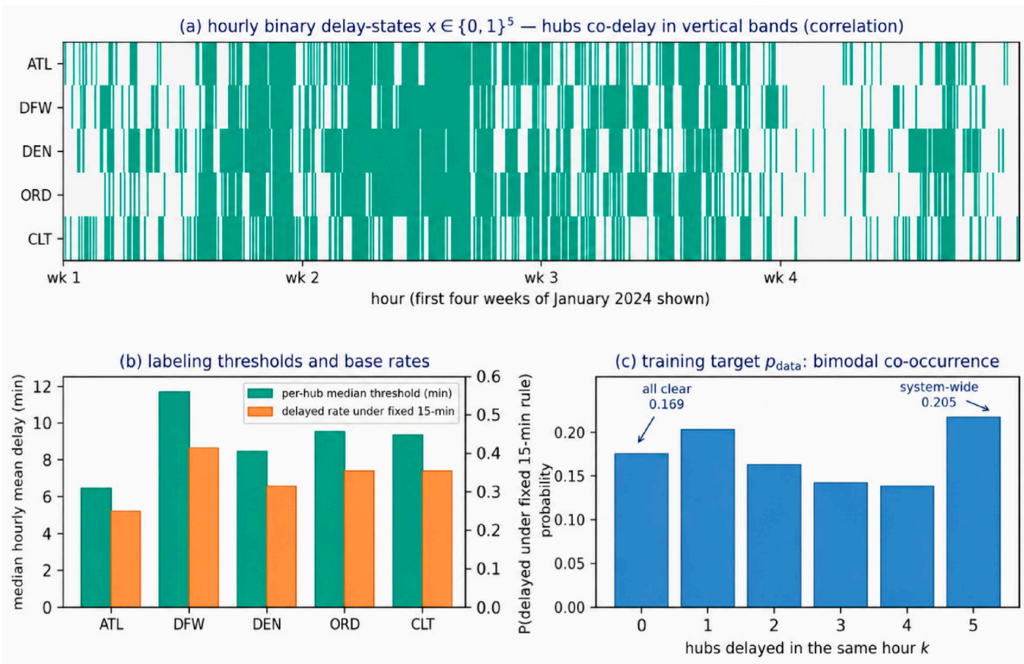

Figure 2. Experimental data and setup (24 months of BTS records, five busiest hubs). (**a**) Hourly binary delay-states, where correlated hub delays appear as vertical bands (first 4 weeks of January 2024). (**b**) Per-hub median labeling thresholds (6–12 min) and delayed rates under the fixed 15-min rule. (**c**) The training target is bimodal, with all-clear probability 0.169 and system-wide probability 0.205.

3.3. Simulation and Reproducibility Settings

The implementation uses custom NumPy state-vector and density-matrix routines in Python 3.9.23 with double-precision real and complex arithmetic. Exact-probability stages enumerate the complete Born distribution. The finite-shot SPSA stage exposes the optimizer only to sampled bit strings: each loss evaluation draws 8192 multinomial shots from the simulated circuit distribution. Table A1 consolidates probability access, initialization, optimizer settings, random seeds, and noise channels. Values were fixed through preliminary convergence checks and are not claimed to be globally optimal.

4. Method

4.1. Born Machine and Ansatz

A parameterized circuit $U(\theta)$, with trainable rotation angles θ, acts on N qubits initialized to $|0...0\rangle$. It defines a probability distribution through the Born rule, where the probability of an outcome is the squared amplitude [10,24],

$$P_\theta(x) = |\langle x \mid U(\theta) \mid 0 \dots 0\rangle|^2. \tag{1}$$

The strongly entangling ansatz [30–32] is the generic baseline (Figure 3). It applies one rotation block to every qubit in each rotation round. Every block belongs to the same *Rz*–*Ry*–*Rz* gate family:

$$G_{\ell,q}(\boldsymbol{\theta}_{\ell,q}) = R_z(\theta_{\ell,q,1})\, R_y(\theta_{\ell,q,2})\, R_z(\theta_{\ell,q,3}).$$

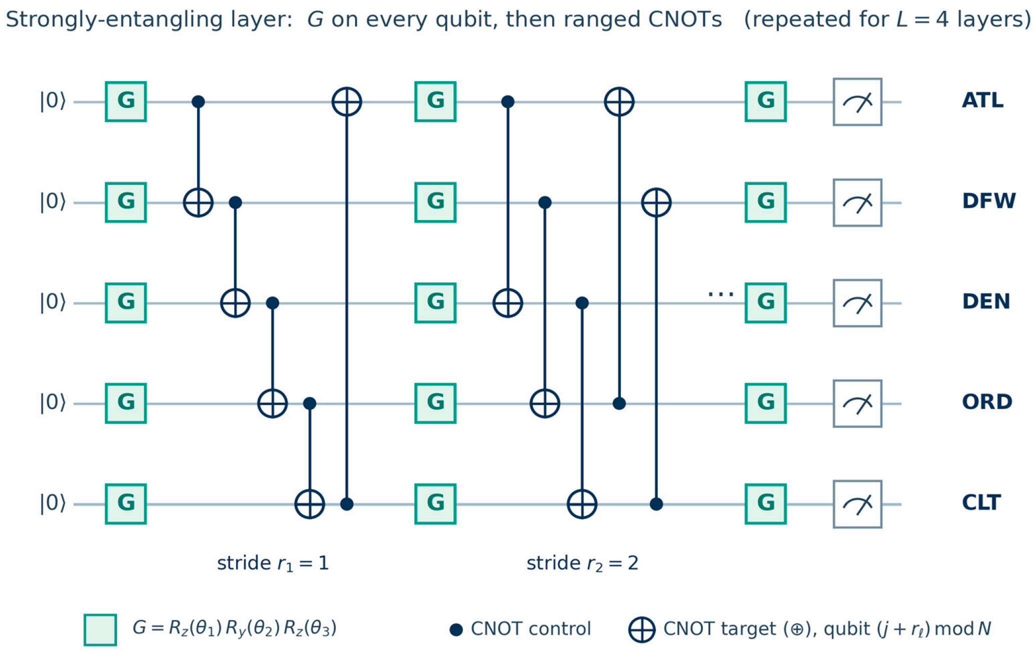

Figure 3. Strongly entangling QCBM ansatz with one qubit per hub. All $G_{\ell,q}$ blocks use the same Rz–Ry–Rz gate family, but every qubit q and rotation round ℓ has an independent angle triplet. Ranged CNOTs follow each of L = 4 entangling layers. For $N = 5$, $N_\theta = 3 \times 5 \times (4 + 1) = 75$ trainable angles. The figure shows the first two layers; the wrap-around gate is an ordinary CNOT.

Here, bold $\boldsymbol{\theta}_{\ell,q}$ denotes the three-angle vector for qubit q and rotation round ℓ. Every $G_{\ell,q}$ block has its own independently trained triplet; the notation denotes a shared gate family, not shared parameter values. Each rotation round before an entangling sublayer is followed by ranged CNOT gates, as defined in Equation (2).

$$U(\theta) = (\textstyle\prod_q G_{L,q}) \prod_{l=1}^{L} (\prod_{j=1}^{N} CNOT_{j,(j+r_l)\, mod\, N})(\prod_q G_{l,q}), \tag{2}$$

The circuit has $L + 1$ rotation rounds: one before each of the L entangling sublayers and one final round. Its trainable-parameter count is therefore $N_\theta = 3N(L + 1)$. For the main configuration, $N_\theta = 3 \times 5 \times (4 + 1) = 75$ angles. The entire list of optimized angles used in the exact-probability MMD experiment for pinned devices is provided in Appendix B (Table A2). Unlike feature-map QML approaches [33,34], the model requires no data-encoding feature map because the empirical distribution enters through the loss. The generic wrap-around and variable-stride CNOTs can require routing overhead on fixed-connectivity hardware, motivating the static topology-guided variant in Section 4.3.

4.2. Pure-State Foundation and Information Budget

The circuit starts in a pure product state and evolves unitarily. Without an entangling operation it remains a product state, so the measured bits are independent and every cross-hub correlation vanishes:

$$|\psi\rangle = |\psi_1\rangle \ldots |\psi_N\rangle \;\Rightarrow\; Corr\left(x_i, x_j\right) = 0, \quad i \neq j. \tag{3}$$

Equivalently, the Born probability factorizes:

$$P_\theta(x_1,\ldots,x_N) = \prod\nolimits^{N}{}_{i=1} \mid \langle x_i \mid \psi_i\rangle \mid^2 = \prod\nolimits^{N}{}_{i=1} p_{\theta,i}(x_i) \tag{4}$$

Entangling gates enable nonfactorizable amplitudes, making correlated bit distributions representable. They do not guarantee useful computational-basis correlation; the rotation angles and measurement-basis alignment also matter. The information-theoretic link follows from data processing:

Proposition 1 (information budget)**.** *For any Born-machine state and hubs i and j, the mutual information of the measured bits satisfies $I(X_i\colon X_j) \leq I_q(i\colon j)$. Proof sketch. Computational-basis measurement is a local quantum-to-classical channel. Quantum mutual information is monotone under local channels by the data-processing inequality, and the post-measurement value is the Shannon mutual information of the two bits. The inequality is standard; its role here is to convert observed airport-pair dependence into a model-specific information requirement.*

The scope is important. Quantum mutual information measures total correlation and is not an entanglement measure. A separable mixed state can have nonzero quantum mutual information and classically correlated measurements. Such mixtures are outside the pure-state unitary QCBM analyzed here. Within this model family, however, a non-entangling circuit remains a pure product state; entangling operations are therefore necessary, but not sufficient, for nonzero cross-hub dependence.

The bound is not tight, and its converse does not hold: neither quantum mutual information nor entanglement alone guarantees measured-bit correlation. The BTS data contain 0.063–0.088 nats of pairwise mutual information. For approximately symmetric binary pairs, $MI \approx r^2/2$, consistent with the observed correlations $r \approx 0.35$–0.42. A pure product circuit has $I_q = 0$ and cannot meet this requirement. The trained state has $I_q = 0.097$–0.203 nats per pair, and its measured-bit mutual information never exceeds the bound (Figure 4).

Remark 1 (basis alignment)**.** *Entanglement is necessary but not sufficient within the pure-state unitary model. A control circuit trained on the independent projection is almost as globally entangled as the BTS-trained circuit (Meyer–Wallach measure 0.94 versus 0.97), yet it measures to nearly independent bits. Proposition 1 constrains total information; training must also align that information with the computational basis.*

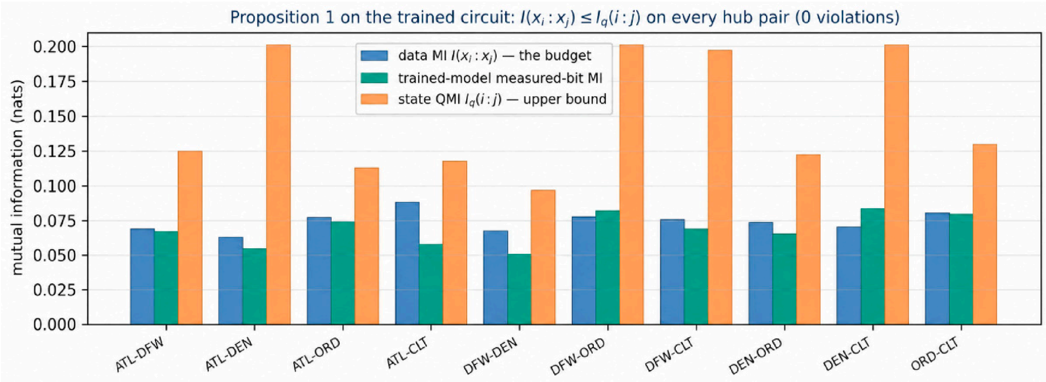

Figure 4. The information budget of Proposition 1 on the trained circuit, for all 10 hub pairs: the pairwise mutual information of the real BTS data (blue) sets the budget; the trained state carries more quantum mutual information (orange) on every pair, and the measured-bit mutual information of the model (green) never exceeds the bound (0 violations).

4.3. Topology-Guided Wiring (TG-QCBM)

The topology-guided variant derives a static entangler from measured airport dependence (Figures 5 and 6). Algorithm 1 constructs a maximum spanning tree and adds the strongest remaining edges until the budget reaches *N CNOTs* per layer, matching the generic ansatz. The same edge set is repeated in all layers, with the hub having the larger 24-month departure count as control. At equal gate budget, TG-QCBM and the generic ansatz are statistically indistinguishable for N = 5–10 (Section 5.3). The value of TG-QCBM is therefore architectural: it replaces changing long-range strides with an interpretable, fixed coupling map.

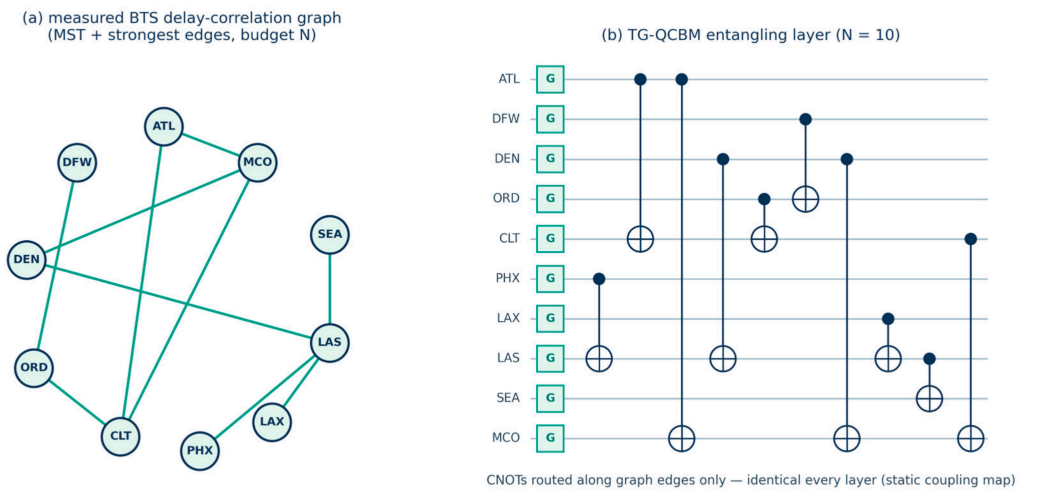

Figure 5. Topology-guided graph construction and CNOT wiring for N = 10. (**a**) The measured BTS correlation graph. (**b**) The topology-guided entangling sublayer repeats one directed CNOT per edge in every layer, with the hub having the larger 24-month departure count as control. Panel (**b**) isolates the entangling sublayer; the independently parameterized $G_{\ell,q}$ rotation blocks remain as defined in Figure 3.

Seasonal conditioning uses $\theta_{eff} = \theta + s \cdot \varphi$, where $s = 0$ denotes winter, $s = 1$ denotes summer, and φ is a trainable offset. This parameter modulation is related to data re-uploading [35]. One circuit is trained jointly on both regime distributions. More regimes and richer temporal variables remain extensions.

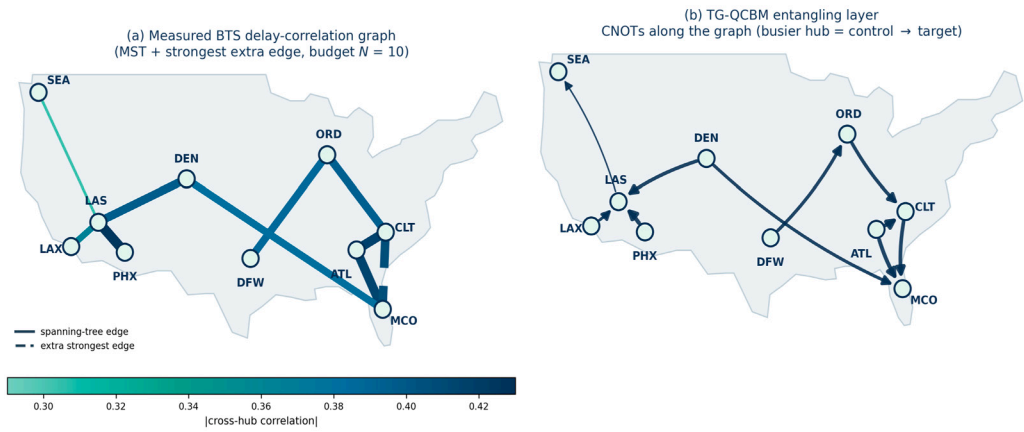

Figure 6. The same topology-guided wiring shown on the U.S. hub network, with each qubit drawn at its airport's true geographic location. (**a**) The measured BTS delay-correlation graph of Figure 5a placed on the map; edge shade and width encode the absolute cross-hub correlation. (**b**) The TG-QCBM entangling layer, drawn as one directed CNOT per edge (larger 24-month departure count = control; the arrow points to the target).

4.4. Training Regimes and Probability Access

Training uses MMD with a Gaussian kernel over Hamming distance, averaged over fixed bandwidths $\sigma \in \{0.5, 1, 2\}$ [10,22,23]. Although MMD can, in principle, be estimated from samples, the reported finite-difference experiment evaluates the objective from exact state-vector probabilities. Only the SPSA experiment evaluates its loss entirely from finite-shot circuit samples:

$$\mathcal{L}(\theta) = \left(P_\theta - P_{data}\right)^\top K\,(P_\theta - P_{data}), \tag{5}$$

in three regimes that the results keep strictly separate:

(i) Exact-probability finite-difference MMD simulation. A one-sided finite-difference gradient is evaluated from the complete Born distribution and optimized with Adam (Algorithm 2). The pinned run uses $h = 0.12$, learning rate 0.12, 160 iterations, and two restarts. The 10-seed summary uses one 130-iteration restart per seed. Initialization, Adam constants, and seed ranges appear in Table A1.

(ii) Finite-shot SPSA MMD simulation. The optimizer receives only bit-string samples, not exact probabilities. Each loss estimate uses an unbiased U-statistic from 8192 shots. SPSA runs for 6000 iterations with two circuit evaluations per iteration, $a_t = 0.6/(t + 0.1T)^{0.602}$, and $c_t = 0.2/t^{0.101}$, totaling 98,304,000 shots per run. Ideal sampling and the specified 0.5% simulated noise model are tested over three seeds. Related studies have demonstrated quantum generative circuits on hardware [36,37], but this workflow was evaluated only in simulation.

(iii) Exact-probability KL refinement (simulation-only expressivity assessment). Formal analyses address expressivity, learnability, and possible quantum-classical separations [38–40]. The MMD solution is fine-tuned on $KL(P_{data}\|p_\theta)$ using exact probabilities, $h = 0.05$, Adam learning rate 0.05, and 250 iterations for the pinned run. Under noiseless simulation, the result provides numerical evidence that the 75-parameter ansatz can approximate the five-hub target closely. It is not a formal certificate, a sample-based trainability result, or evidence of hardware performance.

The process of generating scenarios is equivalent to measuring the trained circuit (Algorithm 3). Preparing $U(\theta^*)|0\ldots0\rangle$ and measuring draws independent scenarios from P_{θ^*}, from which $P_{sys} = \Sigma_{|x|=N}\, P_{\theta^*}(x)$ and the co-occurrence distribution are read off. Near the observed $P_{sys} = 0.205$, one-standard-error precision of ± 0.01 needs $\approx 1.6 \times 10^3$ shots independent of N; a normal-approximation 95% half-width of ± 0.01 needs $\approx 6.3 \times 10^3$ shots.

Algorithm 2. Train QCBM by MMD

Input: empirical p_{data} over 2^N states; layers L; kernel K; steps T; learning rate η
Init angles $\theta \in R^{(L+1)\times N\times 3}$, uniform $[0, \pi]$
for t = 1 . . . T **do**
  $p_\theta \leftarrow$ BornProbs(θ) (exact) or 8192-shot estimate (SPSA regime)
  Loss $\leftarrow \mathrm{MMD}^2(p_\theta, p_{data})$; $\theta \leftarrow \theta - \eta\, \nabla_\theta$ Loss (FD or SPSA)
return θ* with lowest Loss

Algorithm 3. Generate delay scenarios

Input: trained angles θ*; number of scenarios S
for s = 1 . . . S:
  ψs ← U(θ*) |0. . .0⟩
  xs ← measure ψs (comp. basis)
return $\{x_1, \ldots, x_S\}$ ▷ correlated scenarios

4.5. Classical Baselines

Six classical model configurations are fitted to the same empirical distribution by complete-state-space likelihood evaluation or moment matching. These are the independent product of marginals ($N = 5$ parameters, with zero correlation by construction); the Gaussian copula, built through the tetrachoric construction (15 parameters); the fully visible pairwise Boltzmann machine (15 parameters) [41,42]; and RBMs with $M = 4$, 8, and 14 hidden units (29, 53, and 89 parameters) [43,44]. The largest RBM matches and exceeds the QCBM's 75 parameters, and is added so that no comparison hinges on the parameter budget. The copula and Boltzmann models capture all second-order structure by construction, while the RBMs and the QCBM can also express higher-order dependence. The QCBM (75) and RBM-14 (89) exceed the 31 degrees of freedom of this simplex, while RBM-4 (29) is close to that scale. Near-exact fits are therefore expected across families, and no claim of parameter economy is made for any model at this scale.

5. Results

5.1. Core Results: What Quantum Training Achieves vs. What the Circuit Can Represent

Table 1 separates classical exact fits, exact-probability finite-difference MMD simulation, finite-shot SPSA MMD simulation, and exact-probability KL refinement. This distinction is essential: only SPSA uses a loss computed entirely from finite-shot samples, while the KL stage is a simulation-only expressivity assessment.

Three results follow. First, the entanglement-enabled mechanism is robust within the pure-state QCBM. Exact-probability finite-difference MMD reduces correlation error from 0.380 for the independent model to 0.028 ± 0.010 (Figure 7). Finite-shot SPSA reaches 0.060 ± 0.029 under ideal sampling and 0.066 ± 0.026 with simulated 0.5% noise (Table 1). Second, likelihood remains much worse than for classical exact fits: KL is 0.175 ± 0.048 for FD MMD and about 0.26–0.29 for SPSA, versus at most 0.007 for the copula, Boltzmann, and RBM models. Third, exact-probability KL refinement reaches KL = 0.003, and correlation error = 0.011. This is numerical evidence of the ansatz's capacity for this target, not a hardware-performance result (Figure 8).

https://doi.org/10.3390/quantum8030094

Table 1. Generative models on 13,038 hourly BTS states (mean |corr| = 0.380). QCBM rows report mean ± standard deviation over 10 FD-MMD seeds and three SPSA/KL seeds. Observed P_{sys} = 0.205; moving-block 95% CI [0.167, 0.241]; effective sample size ≈ 415. The exact-KL row requires complete probabilities and is unavailable on hardware. The bootstrap interval describes uncertainty in observed P_{sys}, not a refitted model–data difference.

Model (Params)	Corr. Error	KL	TV	P_{sys}
Classical baselines (complete-state-space likelihood evaluation/moment matching)				
Independent (5)	0.380	0.473	0.346	0.031
Gaussian copula (15)	0.000	0.007	0.056	0.188
Pairwise Boltzmann (15)	0.000	0.007	0.056	0.185
RBM, M = 4 (29)	0.007	0.003	0.032	0.196
RBM, M = 8 (53)	0.006	0.003	0.031	0.193
RBM, M = 14 (89)	0.004	0.006	0.049	0.185
QCBM (75), reported simulation regimes				
Exact-probability FD MMD, 10 seeds	0.028 ± 0.010	0.175 ± 0.048	0.174 ± 0.023	0.214 ± 0.008
Finite-shot SPSA MMD, 8192 shots, ideal, 3 seeds	0.060 ± 0.029	0.259 ± 0.051	—	0.175 ± 0.006
Finite-shot SPSA MMD, 8192 shots, 0.5% noise-trained; noiseless evaluation, 3 seeds	0.066 ± 0.026	0.286 ± 0.038	—	0.180 ± 0.003
QCBM (75), exact-probability expressivity assessment				
MMD → exact-probability KL refinement	0.011	0.003 (3-seed 0.005 ± 0.002)	0.026	0.205

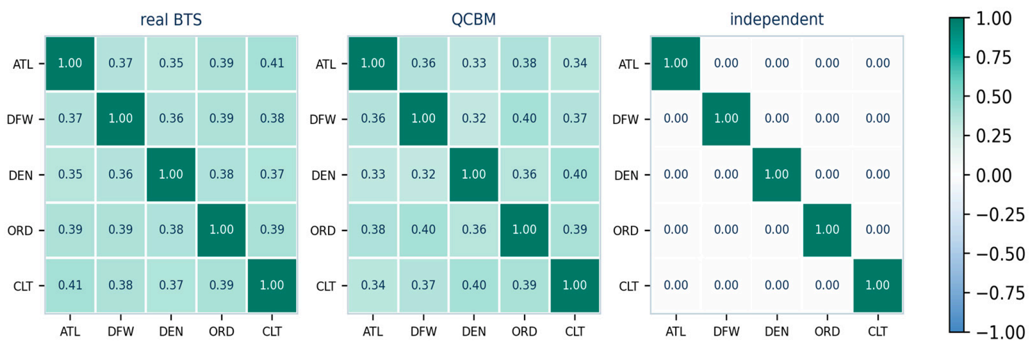

Figure 7. Real U.S. BTS cross-hub delay correlations (24 months, five busiest hubs). The QCBM (**center**) reproduces the observed structure (**left**); the independent baseline (**right**) captures none.

System-wide probability requires a qualified interpretation. The independent model gives 0.031 and underestimates the observed value by 6.6-fold. The Gaussian copula (0.188) and pairwise Boltzmann model (0.185) have point estimates within the moving-block 95% interval [0.167, 0.241] for the observed 0.205. Their point estimates are lower because P_{sys} is a fifth-order quantity not fixed by pairwise moments, but the available effective sample size does not resolve this difference. No statistically superior QCBM performance over the second-order baselines is established for this quantity.

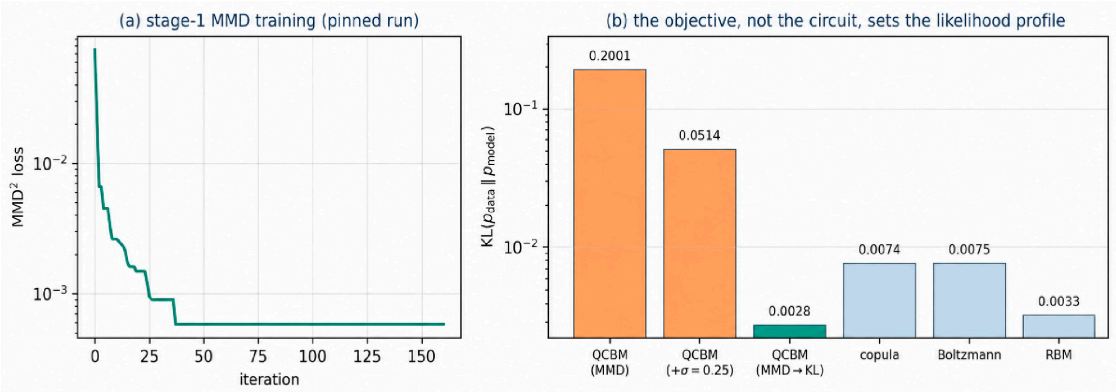

Figure 8. Two-stage objective study. (**a**) Stage-1 MMD^2 training (pinned run). (**b**) KL divergence (log scale) of the same 75-parameter circuit under three objectives: smooth-kernel MMD, MMD with an added σ = 0.25 bandwidth, and the simulation-only exact-KL refinement, each compared against the classical baselines. The objective, not the circuit, sets the likelihood profile.

5.2. Operational Scenario Analysis

Figure 9 evaluates the generated scenarios. The data are bimodal in the number of delayed hubs; the independent model is unimodal, whereas the QCBM reproduces the two dominant modes. The observed system-wide and all-clear probabilities are 0.205 and 0.169. The independent model assigns about 0.031 to each, underestimating them by 6.6- and 5.4-fold. The pinned MMD model gives 0.222 and 0.198, and exact-probability KL refinement gives 0.205 and 0.162 (Table 2).

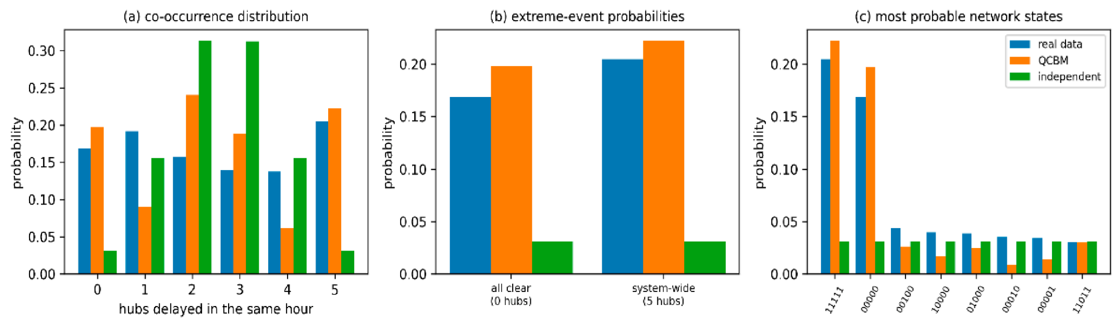

Figure 9. Scenario analysis on real BTS data. (**a**) Co-occurrence distribution: reality is bimodal; the QCBM captures it, the independent model does not. (**b**) Extreme-event probabilities. (**c**) Most likely network states.

Table 2. Real-data extreme-event probabilities (24 months; pinned run, cf. Table 1 for seed spread).

Event	Real	QCBM (MMD)	Indep.	Indep. Error	QCBM (Exact-KL)
All clear (0 hubs)	0.169	0.198	0.031	5.4× under	0.162
System-wide (5 hubs)	0.205	0.222	0.031	6.6× under	0.205

5.3. Ablations: Entangler, Depth, Hub Count, Topology, and Season

Each ablation tests a distinct hypothesis. Removing entanglers tests whether this pure-state model can leave the product-distribution family. Alternative connectivity patterns test whether the presence of connectivity matters more than a particular graph. Depth sweeps evaluate minimum capacity and saturation. Sweep tests with N = 5–10 analyze stability with the increase in the size of the observed state space from 32 to 1024 events, as

https://doi.org/10.3390/quantum8030094

opposed to offering an asymptotic benefit. Tests contrasting TG against arbitrary wiring ensure interpretable fixed connectivity against the backdrop of a constant CNOT cost. Since these grids are exploratory and are not corrected for multiple comparisons, they support mechanistic interpretation rather than confirmatory model selection.

Entangler topology and depth. All configurations of entangling circuits that have been tested exhibit higher cross-hub correlation recovery under their defined parameters compared to the non-entangled product circuit (see Table 3 and Figure 10). In fact, because Table 3 changes the type of rotations (*Ry* vs. *Rz*-*Ry*-*Rz*), the number of parameters (25 vs. 75) and the number of CNOTs per layer (4–10), this is more of a test of mechanisms than ranking of topologies. A matched-budget comparison is presented later in the topology-guided wiring analysis. Depth saturates quickly (Table 4): $L = 2$ reaches correlation error 0.028 with half the *CNOTs* of $L = 4$. Nonmonotonic results at $L = 3$ and $L = 6$ are consistent with optimizer variability rather than a depth advantage.

Table 3. Entangler pattern study (pinned single seed protocol; * = deployed configuration; observed mean ∣corr∣ = 0.380).

Entangler	CNOTs/Layer	Params	MMD	Corr. Error
Independent (none)	0	25 trained (5 effective)	0.040	0.380
Linear (chain)	4	25	0.002	0.048
Circular (ring)	5	25	0.002	0.048
Star	4	25	0.001	0.027
Full (all-to-all)	10	25	0.002	0.059
Pairwise (brick)	4	25	0.001	0.029
Strongly entangling *	5	75	0.001	0.024

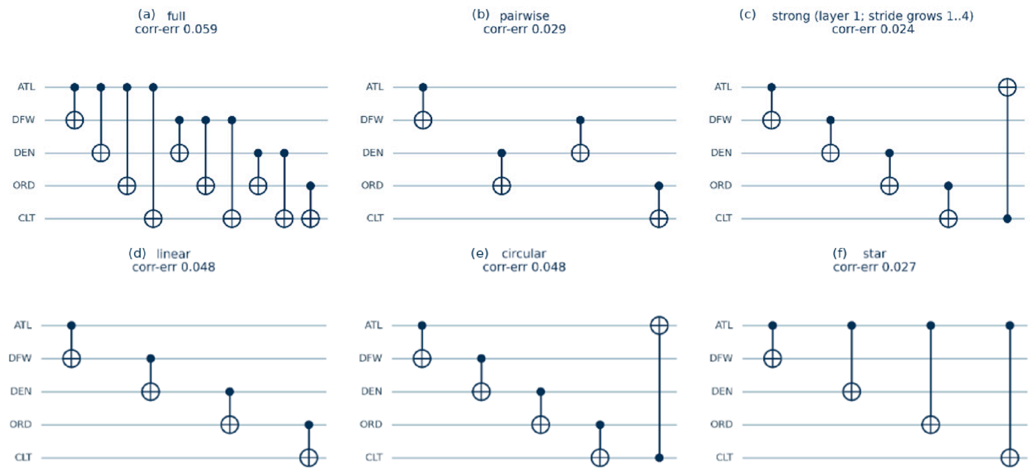

Figure 10. Entangler patterns on the real BTS data with their correlation errors (pinned protocol): (**a**) full, (**b**) pairwise, (**c**) strongly entangling patterns, (**d**) linear, (**e**) circular, and (**f**) star. All entangling variants recover correlation under their reported parameterizations; the experiment does not isolate topology at matched capacity.

Table 4. Depth study (pinned protocol; * = deployed).

L	Params	MMD	Corr. Error
1	30	0.0010	0.048
2	45	0.0003	0.028

Table 4. *Cont.*

L	Params	MMD	Corr. Error
3	60	0.0004	0.041
4 *	75	0.0006	0.024
5	90	0.0007	0.017
6	105	0.0007	0.030

Hub-count scaling. Table 5 and Figure 11 repeat the pipeline for N = 5–10 with five seeds per size. Mean correlation error remains 0.024–0.037, compared with 0.34–0.38 for the independent model. The observed system-wide probability decreases from 0.205 to 0.113; the independent product decreases much faster and underestimates it by 6.6-fold for five hubs and 116-fold for 10 hubs. These sizes remain classically tractable and do not establish computational advantage.

Table 5. Hub count scaling (generic ansatz; mean ± std over 5 seeds; per-hub median threshold). under = observed P_{sys}/independent P_{sys}.

N	Hours	m∣corr∣	QCBM Corr-Err	P_{sys} Real	P_{sys} QCBM	P_{sys} Indep	Under
5	13,038	0.380	0.032 ± 0.011	0.205	0.215 ± 0.007	0.031	6.6×
6	13,014	0.375	0.024 ± 0.002	0.183	0.192 ± 0.009	0.016	11.7×
7	13,014	0.352	0.024 ± 0.004	0.152	0.162 ± 0.007	0.008	19.4×
8	13,010	0.355	0.025 ± 0.003	0.141	0.140 ± 0.011	0.004	36×
9	13,010	0.339	0.026 ± 0.003	0.119	0.125 ± 0.004	0.002	61×
10	12,955	0.344	0.037 ± 0.011	0.113	0.109 ± 0.012	0.001	116×

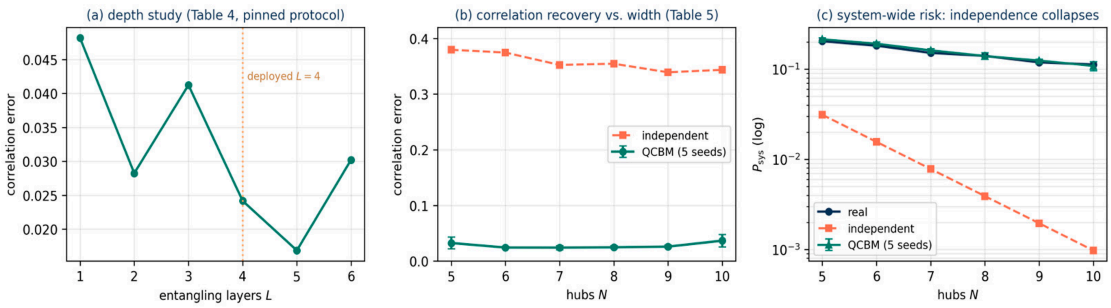

Figure 11. Depth and width studies. (**a**) Correlation error vs. depth L (pinned protocol); the deployed L = 4 sits in the saturated regime. (**b**) Correlation error vs. network size N (mean ± std over 5 seeds), against the independent baseline. (**c**) System-wide probability (log scale): the real value declines slowly, and the QCBM (5 seeds) tracks it, while the independent product collapses to a 116-fold underestimate at N = 10.

Parity rather than superiority: topology-guided wiring. Table 6 and Figure 12a compare TG-QCBM with the standard ansatz, using the the same gate and parameter budget over five seeds each. The average error rates are 0.031 for TG-QCBM and 0.028 for the ansatz, with a difference between them of +0.003 ± 0.009. For N = 10, the distribution overlaps at 0.029 ± 0.006 and 0.037 ± 0.011. TG-QCBM is selected based on the fixed and interpretable coupling map from data, and not because of a superior fidelity.

https://doi.org/10.3390/quantum8030094

Table 6. TG-QCBM vs. generic ansatz at equal gate budget (mean ± std over 5 seeds). Pooled paired difference (TG − generic): +0.003 ± 0.009.

N	TG Corr-Err	Generic Corr-Err
5	0.031 ± 0.006	0.032 ± 0.011
6	0.033 ± 0.003	0.024 ± 0.002
7	0.030 ± 0.005	0.024 ± 0.004
8	0.029 ± 0.005	0.025 ± 0.003
9	0.032 ± 0.005	0.026 ± 0.003
10	0.029 ± 0.006	0.037 ± 0.011

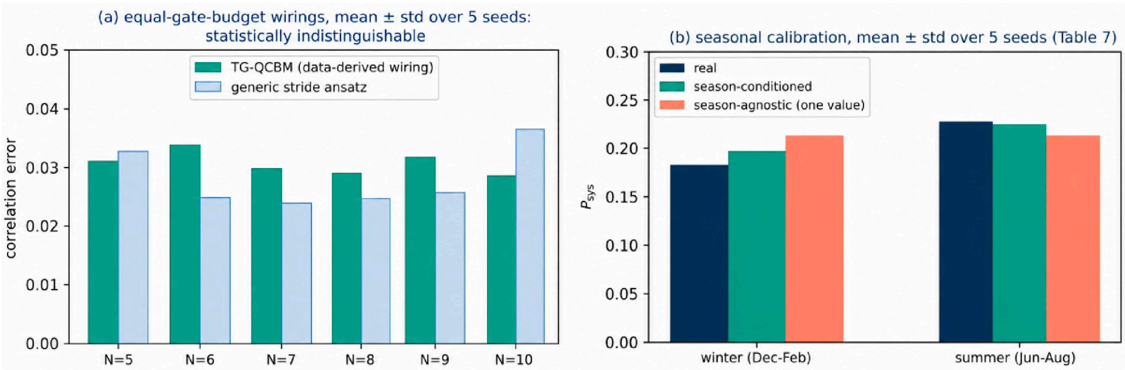

Figure 12. (**a**) TG-QCBM vs. the generic ansatz at equal gate budget (mean ± std over 5 seeds): the two are statistically indistinguishable, so the TG value is architectural (a static coupling map), not fidelity. (**b**) Seasonal calibration over 5 seeds: the agnostic circuit produces one system-wide probability for both regimes, while the conditional circuit separates them in agreement with the data.

Season conditioning: calibration, not accuracy. Winter and summer differ in mean correlation (0.345 versus 0.424), total variation (0.104), and observed P_{sys} (0.183 versus 0.228). Per-regime correlation errors of the conditioned and pooled circuits overlap across five seeds (Table 7, Figure 12b). The confirmed benefit is calibration: the pooled circuit emits one value, 0.214 ± 0.010, for both seasons. The conditioned circuit separates winter and summer at 0.197 ± 0.014 and 0.225 ± 0.011. The conditioned model uses 75 base plus 75 offset parameters (150 total), equal to the total for two separate 75-parameter circuits; its benefit is shared conditioning in one circuit family, not parameter reduction.

Table 7. Season-conditioned TG-QCBM (N = 5; winter = December–February, summer = June–August; mean ± standard deviation over five seeds). "cond." denotes one season-conditioned circuit; "agn." denotes one season-blind circuit; "sep." denotes two per-season circuits (upper bound).

Regime	m|corr|	P_{sys} Real	Err Cond.	Err Agn.	Err Sep.	P_{sys} Cond.	P_{sys} Agn.
winter	0.345	0.183	0.044 ± 0.002	0.038 ± 0.006	0.032 ± 0.002	0.197 ± 0.014	0.214 ± 0.010
summer	0.424	0.228	0.052 ± 0.004	0.057 ± 0.006	0.031 ± 0.008	0.225 ± 0.011	0.214 ± 0.010

5.4. Robustness: Threshold and NISQ Noise

Delay threshold. Under the operationally standard fixed 15-minute threshold (base rates 0.25–0.41), the cross-hub correlation is essentially unchanged (mean 0.368), and the QCBM again recovers it (0.038 vs. 0.368 for the independent model). The system-wide message strengthens, because lower base rates shrink the independent product: real P(all-5) = 0.093 vs. 0.004 independent, a 22.7-fold underestimate (Table 8).

Table 8. Robustness to the delay threshold (24 months; * = primary labeling).

Threshold	Base Rate	$m\lvert\mathrm{corr}\rvert$	QCBM Corr-Err	Indep Corr-Err	P(all-5): Real/Indep
Per-hub median *	0.50	0.380	0.024	0.380	0.205/0.031 (6.6×)
Fixed 15 min	0.34	0.368	0.038	0.368	0.093/0.004 (22.7×)

Simulated noise robustness. Figure 13a applies per-qubit Pauli depolarization after each entangling layer and after the final rotation round, followed by independent readout bit flips. The per-qubit depolarization and readout rates are set equal and swept from 0 to 8%; these layer-level rates are not device per-gate error rates. Correlation error rises from 0.024 without noise to 0.029 at 0.1%, 0.058 at 0.5%, 0.096 at 1%, and 0.159 at 2%. Figure 13b depicts the comparison of the training algorithms described: exact-probability finite difference MMD, optimal finite shots SPSA, noise-trained SPSA applied to the noise-free distribution, and the independent benchmark.

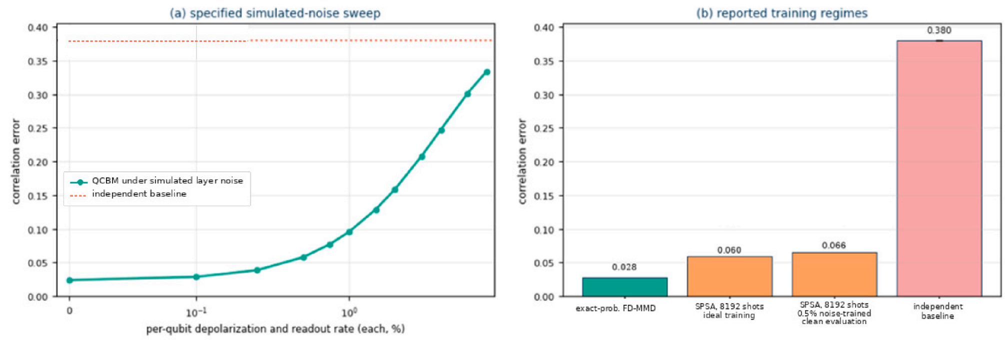

Figure 13. (**a**) Correlation error under the specified per-qubit depolarization and readout model, swept from 0 to 8% with additional resolution through 2%; these are layer-level simulation parameters, not device per-gate error rates. (**b**) Exact-probability FD-MMD and finite-shot SPSA results (mean ± standard deviation). The 0.5% SPSA bar represents noise-injected training followed by noiseless evaluation of the learned parameters. No physical quantum processor was used. The finite-shot SPSA study uses the effective training-time approximation specified in Table A1; its reported post-training metrics evaluate the resulting parameters on the noiseless Born distribution. These are simulations of selected error channels; they omit amplitude damping, phase damping, coherent error, crosstalk, drift, and device-specific compilation. These omitted effects are among the recognized challenges for NISQ and QML implementation [45,46].

6. Discussion

What is established. The independent product model cannot reproduce the observed cross-hub dependence and underestimates simultaneous-delay probability by 6.6–116-fold across N = 5–10. Within the pure-state unitary QCBM, entangling operations enable the required nonproduct distributions, while Proposition 1 supplies a total-information bound. Shallow circuits recover the correlation matrix under exact-probability and finite-shot simulations. TG wiring provides a static coupling map at fidelity parity, and seasonal conditioning improves regime-specific risk calibration.

What is not established. The experiments cover only 2^N = 32–1024 states. Exact classical models fit these distributions efficiently, and the QCBM has worse likelihood under the reported MMD regimes. Exact-KL parity is available only through simulation-side refinement. Second-order baseline point estimates lie within the bootstrap interval for the observed system-wide probability; a refitted model–data difference was not bootstrapped. Consequently, there is no present operational or computational reason to replace the classical models, and linear parameter growth alone is not evidence of quantum advantage.

Trainability. At fixed $L = 4$, the variance of the forward-difference derivative of the local cost $C = \langle Z_0 Z_1 \rangle$ with respect to the first-round ATL Ry angle (step 10^{-3}; 200 random initializations; RNG seed 0 at each N) decreases from about 0.015 at $N = 5$ to 0.0005 at $N = 10$. This limited diagnostic is consistent with an emerging barren-plateau trend [47] but does not establish asymptotic scaling. The depth sweep does not show the same monotonic decline. Scaling beyond the present study must address this trainability risk and replace exact distribution evaluation with sample-based estimators.

Potential optimizations. Testable directions include layer-wise circuit growth; warm starts from smaller networks or classically fitted marginals; identity-block initialization; adaptive MMD bandwidths; local or correlation-targeted objectives; shot-frugal mini-batch or linear-time MMD estimators; natural-gradient methods; measurement-error mitigation; device-native routing; CNOT-direction search; sparse topology selection; and parameter sharing across repeated network motifs. Each should be evaluated under matched budgets and multiple seeds rather than assumed to improve performance.

Cost. All of the classical baselines finished in roughly 1.6 s, the FD-MMD exact probability computation took around 5.7 s, and the SPSA simulation took about 7 s because of sampling shots computationally. Neither the processor architecture nor the timing-repeat method is stated here, and therefore, these wall-clock times are indicative but not portable benchmarks. In practice, on hardware, about 10^8 preparation-measurement operations per run would dominate the wall-clock time. When N is between 5 and 10, the QCBM is the resource-hungrier choice.

Scalability. The number of parameters and the number of two-qubit gates increase linearly with N(L + 1) and O(NL), respectively. Circuit depth depends on connectivity, routing, and scheduling of the gates. Exact kernel computation scales as O(4N); sampling allows us to avoid this scaling but comes with a cost of sampling noise. The practical benefit of the circuit for $N \geq 50$, compared to RBMs, VAEs, and graph generators, assuming equivalent resource usage and noise conditions, needs to be confirmed empirically.

7. Threats to Validity

Construct validity. Binarization does not provide any information about the delay amount and timings. Median labeling, by design, balances the marginals, but the fixed 15-min cut-off produces the same qualitative result (Table 8).

Internal validity. Absolute accuracy depends on pinned optimizers. The ablation grids are exploratory, are not multiplicity corrected, and should not be read as confirmatory model selection. Multi-seed uncertainty accompanies the central comparisons. All fidelity metrics are in-sample because fitting and evaluation use the same empirical distributions; temporal holdout generalization is untested.

Statistical validity. Serial dependence reduces the effective sample size to about 415 h, so inference for P_{sys} uses the moving-block bootstrap rather than a binomial interval.

External validity. The generator models exchangeable hourly snapshots and does not capture temporal propagation. No experiment was executed on a physical quantum processor. Finite-shot and noise-injected simulations omit connectivity-dependent compilation, coherent errors, calibration drift, crosstalk, queue latency, and correlated readout. Hardware execution, temporal encodings, cost-weighted risk, and larger-N classical comparisons remain future work.

8. Conclusions

This study formulates correlated flight-delay scenario generation with a pure-state QCBM. In this model family, a non-entangling circuit produces a factorized distribution, whereas entangling operations enable nonproduct measurement statistics. The

https://doi.org/10.3390/quantum8030094

data-processing inequality converts observed pairwise mutual information into a total-information requirement, but quantum mutual information is not itself an entanglement measure. On 24 months of BTS data, shallow circuits recover the correlation matrix under exact-probability FD-MMD and finite-shot SPSA simulations, including 0.5% noise-injected training with noiseless post-training evaluation. The independent product strongly underestimates simultaneous-delay probability, but second-order classical point estimates lie within the bootstrap interval for the observed value. Classical models evaluated over the complete 32–1024-state spaces also fit efficiently. Thus, no quantum advantage or hardware demonstration is claimed. TG-QCBM contributes a static, interpretable coupling map at fidelity parity; the unresolved question is whether any benefit emerges at larger scales under matched classical budgets and physical-device constraints.

Author Contributions: Conceptualization, I.A.; methodology, M.A.; software, M.A.; validation, M.A. and I.A.; formal analysis, M.A.; investigation, M.A. and I.A.; data curation, M.A.; writing—original draft preparation, M.A.; writing—review and editing, M.A.; visualization, M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The authors received no specific funding for this work.

Data Availability Statement: The data presented in this study are openly available in the U.S. Bureau of Transportation Statistics [42]. The implementation code and data-processing notebook are publicly available at https://itplus.company/repository/EaC.ipynb (accessed on 2 September 2026). The processed hourly dataset can be reconstructed from the public source data using the notebook.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Experimental and Reproducibility Settings

Table A1. Consolidated simulation and optimization settings.

Setting Used	Scope	Component
Custom NumPy implementation; Python 3.9.23; float64/complex128 arithmetic; no physical quantum processor.	All experiments	Software and precision
$N = 5$ core and $N = 5$–10 scaling; $L = 4$ unless varied; Rz–Ry–Rz blocks; $N\theta = 3N(L + 1)$, giving 75 angles at $N = 5$, $L = 4$.	Core/scaling	Circuit
$\theta \sim$ Uniform $[0, \pi)$. Core summary: seeds 0–9. Scaling, TG, and seasonal studies: seeds 0–4. SPSA: seeds 0–2; measurement RNG seed 1000 + s.	All QCBMs	Initialization and seeds
Gaussian kernel of Hamming distance, averaged over $\sigma \in \{0.5, 1.0, 2.0\}$; kernel ablation adds $\sigma = 0.25$.	MMD stages	MMD kernel
Exact state-vector probabilities; one-sided finite difference $h = 0.12$; Adam lr = 0.12, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$; 160 iterations; 2 restarts; retain best observed loss.	Figures/pinned tables	Pinned exact-probability FD MMD
Exact state-vector probabilities; same h and Adam settings; 130 iterations; 1 restart per seed; seeds 0–9.	Core error bars	Ten-seed FD MMD summary

Table A1. *Cont.*

Setting Used	Scope	Component
8192 shots per loss evaluation; 6000 iterations; 2 evaluations/iteration; $a_t = 0.6/(t + 0.1T)^{0.602}$; $c_t = 0.2/t^{0.101}$; 98,304,000 shots/run; seeds 0–2.	Ideal/noisy sampling	Finite-shot SPSA MMD
Starts from the MMD solution; forward difference h = 0.05; Adam lr = 0.05 with the same β values; 250 iterations for the pinned run. Three-seed check: 130 MMD + 150 KL iterations, 1 restart, seeds 0–2.	Expressivity assessment	Exact-probability KL refinement
Exact-probability FD MMD; L = 4; 120 iterations; 1 restart per seed; seeds 0–4. Pattern study: 2 restarts and 150 iterations (120 for full graph). Depth study: 2 restarts and 160 iterations.	Main ablations	Topology/width/season
After each entangling layer and after the final rotation round, apply per-qubit Pauli depolarization p_g; then apply independent readout bit flips p_r. Set $p_g = p_r$ and sweep 0–8%. These are layer-level simulation rates, not device per-gate errors.	Post-training robustness	Noise sweep
At $p_g = p_r = 0.005$, mix p with the uniform distribution using $\lambda = 1 - (1 - p_g)^{NL}$, then apply independent per-bit readout flips before shot sampling. Table 1 reports the resulting parameters under noiseless post-training evaluation.	Training-time noise	Noisy SPSA approximation
No early stopping; fixed iteration budgets. Retain the lowest-loss parameters encountered. Report mean ± population standard deviation across the stated seeds.	All optimizers	Stopping and reporting

Appendix B. Optimized Circuit Parameters

Table A2 reports the 75 optimized angles of the pinned exact-probability finite-difference MMD run used for Figure 7 and the pinned entries of Tables 1–3. These values reproduce correlation error 0.0242 and MMD 5.8×10^{-4}. The released notebook deterministically generates the exact-KL and SPSA parameters.

Table A2. Optimized angles θ* (radians; pinned exact-probability finite-difference MMD run).

Round	Hub	θ_1 (Rz)	θ_2 (Ry)	θ_3 (Rz)
1	ATL	1.6079	1.5806	0.5653
1	DFW	2.9803	−0.0713	1.1291
1	DEN	2.6003	1.1745	1.2722
1	ORD	0.0866	1.9055	1.5035
1	CLT	1.0359	2.4736	1.0008
2	ATL	1.1316	−0.7864	1.3469
2	DFW	0.4199	0.1523	2.5246

https://doi.org/10.3390/quantum8030094

Table A2. *Cont.*

Round	Hub	θ_1 (Rz)	θ_2 (Ry)	θ_3 (Rz)
2	DEN	−0.5836	0.7945	3.4324
2	ORD	1.2432	2.0394	1.2158
2	CLT	0.0638	0.6312	1.6840
3	ATL	0.7931	0.2624	1.3371
3	DFW	1.4888	0.7835	2.0874
3	DEN	−0.5686	1.1451	0.2926
3	ORD	−0.1010	1.4201	3.2891
3	CLT	1.2090	−0.1850	1.8002
4	ATL	1.4390	1.7114	0.9620
4	DFW	−1.1957	1.0576	2.2250
4	DEN	2.1386	0.3078	2.1939
4	ORD	0.2125	−0.3633	2.1582
4	CLT	2.3975	1.7156	0.2487
5	ATL	1.1069	−0.9538	2.0286
5	DFW	1.9471	2.5916	0.8855
5	DEN	−0.7772	1.5375	2.5292
5	ORD	4.2824	−1.5058	1.5149
5	CLT	1.7760	−0.0676	1.8520

References

1. Fleurquin, P.; Ramasco, J.J.; Eguíluz, V.M. Systemic delay propagation in the US airport network. *Sci. Rep.* **2013**, *3*, 1159. [CrossRef] [PubMed]
2. Pyrgiotis, N.; Malone, K.M.; Odoni, A. Modelling delay propagation within an airport network. *Transp. Res. Part C* **2013**, *27*, 60–75. [CrossRef]
3. AhmadBeygi, S.; Cohn, A.; Guan, Y.; Belobaba, P. Analysis of the potential for delay propagation in passenger airline networks. *J. Air Transp. Manag.* **2008**, *14*, 221–236. [CrossRef]
4. Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 195–202. [CrossRef] [PubMed]
5. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2018**, *2*, 79. [CrossRef]
6. Cerezo, M.; Arrasmith, A.; Babbush, R.; Benjamin, S.C.; Endo, S.; Fujii, K.; McClean, J.R.; Mitarai, K.; Yuan, X.; Cincio, L.; et al. Variational quantum algorithms. *Nat. Rev. Phys.* **2021**, *3*, 625–644. [CrossRef]
7. Khanal, P.; Adhikari, N.B. Hybrid classical–quantum ensemble learning for real-time flight delay prediction at Tribhuvan International Airport. *TELKOMNIKA Telecommun. Comput. Electron. Control* **2026**, *24*, 527–535. [CrossRef]
8. Fujii, K.; Nakajima, K. Harnessing disordered-ensemble quantum dynamics for machine learning. *Phys. Rev. Appl.* **2017**, *8*, 024030. [CrossRef]
9. Huang, H.-Y.; Broughton, M.; Mohseni, M.; Babbush, R.; Boixo, S.; Neven, H.; McClean, J.R. Power of data in quantum machine learning. *Nat. Commun.* **2021**, *12*, 2631. [CrossRef] [PubMed]
10. Benedetti, M.; Garcia-Pintos, D.; Perdomo, O.; Leyton-Ortega, V.; Nam, Y.; Perdomo-Ortiz, A. A generative modeling approach for benchmarking and training shallow quantum circuits. *npj Quantum Inf.* **2019**, *5*, 45. [CrossRef]
11. Gao, X.; Anschuetz, E.R.; Wang, S.-T.; Cirac, J.I.; Lukin, M.D. Enhancing generative models via quantum correlations. *Phys. Rev. X* **2022**, *12*, 021037. [CrossRef]
12. Ball, M.; Cynthia, B.; Martin, D. Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States. NEXTOR Technical Report. 2010. Available online: https://rosap.ntl.bts.gov/view/dot/6234 (accessed on 2 September 2026).
13. Rebollo, J.J.; Balakrishnan, H. Characterization and prediction of air traffic delays. *Transp. Res. Part C* **2014**, *44*, 231–241. [CrossRef]

14. Cai, K.; Li, Y.; Fang, Y.; Zhu, Y. A deep learning approach for flight delay prediction through time-evolving graphs. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 11397–11407. [CrossRef]
15. Bao, J.; Yang, Z.; Zeng, W. Graph to sequence learning with attention mechanism for network-wide multi-step-ahead flight delay prediction. *Transp. Res. Part C* **2021**, *130*, 103323. [CrossRef]
16. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer: New York, NY, USA, 2006. [CrossRef]
17. Sklar, A. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **1959**, *8*, 229–231.
18. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [CrossRef]
19. Smolensky, P. Information processing in dynamical systems: Foundations of harmony theory. In *Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations*; MIT Press: Cambridge, MA, USA, 1986; pp. 194–281. Available online: https://direct.mit.edu/books/monograph/4424/chapter/189414/Information-Processing-in-Dynamical-Systems (accessed on 2 September 2026).
20. Kingma, D.P.; Welling, M. Auto-encoding variational Bayes. In Proceedings of the International Conference on Learning Representations (ICLR), Banff, AB, Canada, 14–16 April 2014.
21. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Montréal, QC, Canada, 8–13 December 2014; pp. 2672–2680. Available online: https://papers.nips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html (accessed on 2 September 2026).
22. Liu, J.-G.; Wang, L. Differentiable learning of quantum circuit Born machines. *Phys. Rev. A* **2018**, *98*, 062324. [CrossRef]
23. Gretton, A.; Borgwardt, K.M.; Rasch, M.J.; Schölkopf, B.; Smola, A. A kernel two-sample test. *J. Mach. Learn. Res.* **2012**, *13*, 723–773.
24. Benedetti, M.; Lloyd, E.; Sack, S.; Fiorentini, M. Parameterized quantum circuits as machine learning models. *Quantum Sci. Technol.* **2019**, *4*, 043001. [CrossRef]
25. Dallaire-Demers, P.-L.; Killoran, N. Quantum generative adversarial networks. *Phys. Rev. A* **2018**, *98*, 012324. [CrossRef]
26. Deng, Y.-W.; Chen, M.; Gong, C.; Zhou, N. High resolution image generation based on quantum generative adversarial networks. *Comput. Vis. Image Underst.* **2026**, *271*, 104884. [CrossRef]
27. Gong, L.-H.; Chen, Y.-Q.; Zhou, S.; Zeng, Q.-W. Dual Discriminators Quantum Generation Adversarial Network Based on Quantum Convolutional Neural Network. *Adv. Quantum Technol.* **2025**, *8*, e2500224. [CrossRef]
28. Ding, Y.-F.; Li, Z.; Zhou, N. Quantum generative adversarial network based on the quantum Born machine. *Adv. Eng. Inform.* **2025**, *68*, 103622. [CrossRef]
29. U.S. Bureau of Transportation Statistics. Reporting Carrier On-Time Performance Data, 2024–2025. Available online: https://www.transtats.bts.gov/ONTIME/ (accessed on 2 September 2026).
30. Mitarai, K.; Negoro, M.; Kitagawa, M.; Fujii, K. Quantum circuit learning. *Phys. Rev. A* **2018**, *98*, 032309. [CrossRef]
31. Schuld, M.; Bocharov, A.; Svore, K.M.; Wiebe, N. Circuit-centric quantum classifiers. *Phys. Rev. A* **2020**, *101*, 032308. [CrossRef]
32. Combarro, E.F.; González-Castillo, S. *A Practical Guide to Quantum Machine Learning and Quantum Optimization*; Packt: Birmingham, UK, 2023. Available online: https://www.packtpub.com/en-us/product/a-practical-guide-to-quantum-machine-learning-and-quantum-optimization-9781804613832 (accessed on 2 September 2026).
33. Havlíček, V.; Córcoles, A.D.; Temme, K.; Harrow, A.W.; Kandala, A.; Chow, J.M.; Gambetta, J.M. Supervised learning with quantum-enhanced feature spaces. *Nature* **2019**, *567*, 209–212. [CrossRef] [PubMed]
34. Schuld, M.; Killoran, N. Quantum machine learning in feature Hilbert spaces. *Phys. Rev. Lett.* **2019**, *122*, 040504. [CrossRef] [PubMed]
35. Pérez-Salinas, A.; Cervera-Lierta, A.; Gil-Fuster, E.; Latorre, J.I. Data re-uploading for a universal quantum classifier. *Quantum* **2020**, *4*, 226. [CrossRef]
36. Zhu, D.; Linke, N.M.; Benedetti, M.; Landsman, K.A.; Nguyen, N.H.; Alderete, C.H.; Perdomo-Ortiz, A.; Korda, N.; Garfoot, A.; Brecque, C.; et al. Training of quantum circuits on a hybrid quantum computer. *Sci. Adv.* **2019**, *5*, eaaw9918. [CrossRef] [PubMed]
37. Rudolph, M.S.; Toussaint, N.B.; Katabarwa, A.; Johri, S.; Peropadre, B.; Perdomo-Ortiz, A. Generation of high-resolution handwritten digits with an ion-trap quantum computer. *Phys. Rev. X* **2022**, *12*, 031010. [CrossRef]
38. Du, Y.; Hsieh, M.-H.; Liu, T.; Tao, D. Expressive power of parametrized quantum circuits. *Phys. Rev. Res.* **2020**, *2*, 033125. [CrossRef]
39. Sweke, R.; Seifert, J.-P.; Hangleiter, D.; Eisert, J. On the quantum versus classical learnability of discrete distributions. *Quantum* **2021**, *5*, 417. [CrossRef]
40. Coyle, B.; Mills, D.; Danos, V.; Kashefi, E. The Born supremacy: Quantum advantage and training of an Ising Born machine. *npj Quantum Inf.* **2020**, *6*, 60. [CrossRef]
41. Ackley, D.H.; Hinton, G.E.; Sejnowski, T.J. A learning algorithm for Boltzmann machines. *Cogn. Sci.* **1985**, *9*, 147–169. [CrossRef]
42. Besag, J. Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Stat. Soc. B* **1974**, *36*, 192–236. [CrossRef]
43. Hinton, G.E. Training products of experts by minimizing contrastive divergence. *Neural Comput.* **2002**, *14*, 1771–1800. [CrossRef] [PubMed]

44. Le Roux, N.; Bengio, Y. Representational power of restricted Boltzmann machines and deep belief networks. *Neural Comput.* **2008**, *20*, 1631–1649. [CrossRef] [PubMed]
45. Bharti, K.; Cervera-Lierta, A.; Kyaw, T.H.; Haug, T.; Alperin-Lea, S.; Anand, A.; Degroote, M.; Heimonen, H.; Kottmann, J.S.; Menke, T.; et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **2022**, *94*, 015004. [CrossRef]
46. Cerezo, M.; Verdon, G.; Huang, H.-Y.; Cincio, L.; Coles, P.J. Challenges and opportunities in quantum machine learning. *Nat. Comput. Sci.* **2022**, *2*, 567–576. [CrossRef] [PubMed]
47. McClean, J.R.; Boixo, S.; Smelyanskiy, V.N.; Babbush, R.; Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **2018**, *9*, 4812. [CrossRef] [PubMed]