Journal of
Quantum Computing

Tech Science Press

https://doi.org/10.32604/jqc.2026.087094

ARTICLE

Hybrid Generative Intelligence: A Quantum–LLM Framework for Emergent Musical Creativity

Lakshmi Prayaga[1,*] **and Chandra Prayaga**[2]

[1]Department of Cybersecurity and Information Technology, University of West Florida, Pensacola, FL, USA
[2]Physics Department, University of West Florida, Pensacola, FL, USA

*Corresponding Author: Lakshmi Prayaga. Email: lprayaga@uwf.edu

Received: 10 June 2026; Accepted: 04 September 2026; Published: 30 October 2026

ABSTRACT: Recent paradigm shifts in algorithmic music composition explore the intersection of physical stochastic processes and statistical language modeling. This study introduces the Quantum–LLM Composer, a hybrid generative framework that couples natural-language semantic conditioning with quantum-circuit-derived probability distributions to generate musical sequences. A three-qubit quantum circuit implemented in Qiskit (H, CX, RX, and CRZ gates) produces measurement bitstrings that are mapped modulo-N to defined harmonic scales, while a Large Language Model (GPT-5) interprets user intent into scale, tempo, and instrumentation parameters. Using a corrected, 100-trial-per-system evaluation pipeline with nonparametric significance testing, we show that hardware noise on real IBM Quantum hardware significantly increases Shannon entropy relative to a noiseless simulation of the same circuit (1.864 vs. 1.723 bits; $p = 1.5 \times 10^{-7}$), a pattern we trace to a specific mechanism: the circuit's exact statevector populates only 4 of 8 possible basis states, and hardware noise measurably leaks probability into the remaining, ideally forbidden states. We further show that the LLM's semantic output can be coupled directly to the quantum circuit's own rotation parameters, rather than used only to select instrumentation after the fact, so that the sampled distribution itself is measurably steered by theme (Kruskal–Wallis H = 206.8, $p = 1.3 \times 10^{-43}$ across five themes on the simulator), with the same theme effect and the same noise-elevates-entropy mechanism replicating independently on real quantum hardware. We report these as narrow, falsifiable, and statistically supported findings rather than claims of demonstrated musical superiority over classical pseudo-randomness, and we identify a formal perceptual listening study, an external classical-ML baseline, and a production-parameter (timbre) extension as concrete next steps.

KEYWORDS: Quantum computing; large language models; generative AI; algorithmic music composition; quantum expressivity; hybrid intelligence; NISQ; Shannon entropy

1 Introduction

The integration of Generative AI and Quantum Computing represents a frontier in computational creativity. Classical generative models—such as GANs, VAEs, and diffusion models—rely on deterministic optimization, whereas quantum models leverage probabilistic mechanics such as superposition and entanglement to sample from inherently richer distributions. Hybrid GenAI–Quantum systems combine these capabilities: the symbolic reasoning and context of LLMs with the stochastic energy landscape of quantum circuits, enabling structurally novel, statistically characterizable outputs whose relationship to human creativity or improvisation remains an open empirical question rather than an assumed property.

Prior research in quantum cellular automata (Miranda & Shaji, 2023 [1]) and quantum generative adversarial networks (QGANs) has explored how quantum principles can be mapped to musical dynamics or used

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

to accelerate model training. Recent studies suggest that randomness structured by quantum superposition can yield output distributions that are statistically distinct from those of classical systems, though we make no claim that this reflects distinct forms of cognition or creativity absent perceptual or cognitive-science validation. Despite these advances, a significant gap remains: few systems directly couple semantic reasoning from LLMs with quantum stochastic generation for real-time synthesis. This study addresses that gap by presenting a hybrid framework capable of interpreting natural-language intent and realizing it through quantum probabilistic note generation, thereby bridging symbolic and physical computation.

2 Related Work and Research Gap

2.1 Quantum Music Generation

Prior efforts explored how superposition and measurement randomness in quantum systems could produce musical variation. Early quantum-computer-assisted composition systems include the work of Heudin [2], who introduced the concept of "Quantum Notes" (Qunotes) to simulate quantum properties for experimental soundscapes, Lee and Yang [3] explored the application of quantum bits as fundamental compositional units. Dobrian and Costa Hamido [4] have investigated quantum-aided algorithmic composition across rhythm, timbre, and spatial dimensions using statevector rotations to drive granular harmonic textures. While these models reveal the creative potential of quantum uncertainty, they often lack mechanisms for the objective evaluation or adaptation of generative output.

2.2 Generative and LLM Systems in Music

Parallel advances in LLMs have revolutionized symbolic composition. Architectures like the Music Transformer [5] and OpenAI's Jukebox [6] extended modeling to polyphonic structures and raw audio. More recently, hybrid LLM–music systems have leveraged descriptive prompts to align music with specific moods [7]. Miranda and Siegelwax [8] further advanced this field with PhotoSing, which extracts sequencing rules from musical examples to generate polyphonic tunes using a photonic quantum computer. Xing et al. [9] introduced CoComposer, a multi-agent system that coordinates tasks across the traditional music composition workflow. However, these models rely primarily on deterministic token prediction and do not inherently incorporate the stochastic quantum variability explored in this study.

2.3 Quantum–AI Hybrids and Emergent Intelligence

Separately, Fisher [10] proposed that quantum processes involving nuclear spins may play a role in biological cognition; while this is a neuroscience hypothesis rather than a claim about AI architectures, it is sometimes invoked by loose analogy when motivating randomness-based hybrid systems. In the commercial and applied sector, Quantinuum [11] and MOTH [12] have publicized generative-quantum-AI and Quantum-Reservoir-Computing music results, respectively; we cite these as industry announcements rather than peer-reviewed sources and do not rely on them to support technical claims. To evaluate the technical feasibility of such systems, Afane et al. [13] developed Quantum-Audit to benchmark the reasoning limits of LLMs in quantum contexts. Despite these foundations, few systems have demonstrated such emergence through measurable, statistically validated audio signal outputs or unified frameworks—a gap this study targets directly with a 100-trial, significance-tested evaluation pipeline (Section 4).

2.4 Identified Research Gap and Contribution

Prior studies highlight three persistent gaps:

- Lack of adaptive evaluation metrics for quantum-derived music beyond simple novelty.
- Limited integration of quantum stochasticity with semantic LLM reasoning to guide probabilistic outcomes.
- Absence of cross-domain frameworks unifying compositional logic, quantum randomness, and signal-based metrics.

This work introduces the Quantum–LLM Composer, which fuses quantum-generated sequences with LLM-inferred parameters (scale, instrument, tempo) and, in its coupled configuration (Section 4.7), uses those same LLM-inferred parameters to set the quantum circuit's own rotation angles rather than only selecting instrumentation after the fact. By employing amplitude and entropy metrics evaluated over 100 trials per condition with nonparametric significance testing, the system not only generates music but also measures and interprets its own output distribution in a reproducible, falsifiable way, bridging the gap between quantum stochastic art and explainable generative intelligence.

3 System Design and Methods

3.1 System Architecture

The Quantum–LLM Composer is a hybrid generative framework that integrates Large Language Models (LLMs) for semantic conditioning with quantum stochastic circuits for musical sequence generation. As illustrated in Fig. 1, the system consists of four integrated modules:

- Semantic Processing: Interprets user intent via GPT-5.
- Quantum Note Generation: Samples probabilistic sequences from a quantum circuit.
- Synthesis Pipeline: Converts symbolic data into MIDI and audible waveforms.
- Evaluation Framework: Quantifies musical structure via entropy and stability metrics.

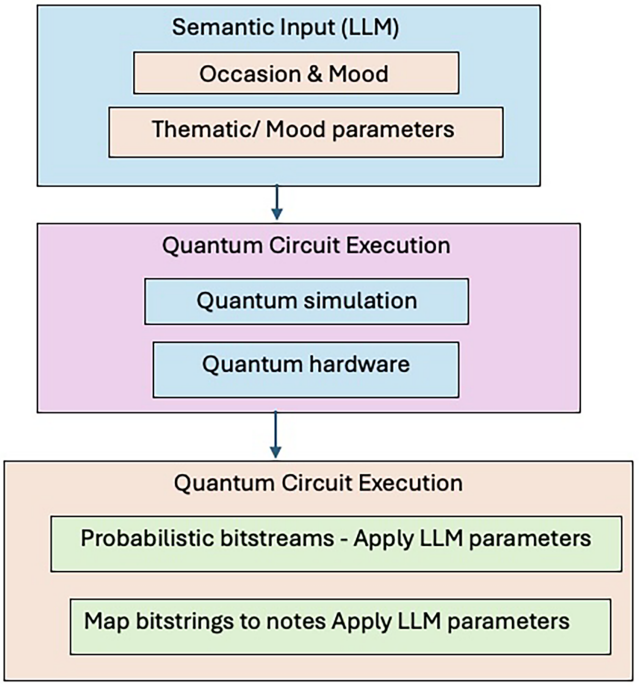


Figure 1: System architecture of the Quantum–LLM composer, illustrating the interaction between semantic input processing, quantum note generation, audio synthesis, and evaluation modules.

3.2 Quantum Note Generator

The quantum layer employs a three-qubit circuit implemented via Qiskit's Aer Simulator. A Hadamard gate (H) is applied to qubit 0 to induce superposition; a CNOT gate (CX) then entangles qubits 0 and 1; a parameterized rotation gate (RX), with angle θ_{RX}, is applied to qubit 1; and a controlled-RZ gate (CRZ), with angle θ_{CRZ}, entangles qubits 1 and 2. All three qubits are then measured. This circuit therefore includes two entangling operations (CX and CRZ), not the single-qubit-only description given in an earlier version of this manuscript. For the baseline comparison in Section 4.2, θ_{RX} and θ_{CRZ} are fixed at $\pi/3$ and $\pi/4$, respectively; Section 4.7 describes a coupled configuration in which the LLM's semantic output instead sets these two angles per theme. The RX gate is defined in Eq. (1):

$$\mathrm{RX}(\theta) = \begin{pmatrix} \cos(\theta/2) & -i\sin(\theta/2) \\ -i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \tag{1}$$

This study builds on the broader survey of gate-based and measurement-based approaches to quantum-computing-aided composition and performance collected in Miranda [14], including Souma's gate-type approach [15].

State measurements from repeated executions ("shots") yield bitstrings. Each measured bitstring b is converted to an integer using base-2 representation, int(b,2), and then mapped to a valid note index within the selected musical scale using a modulo operation, Note Index = int(b,2) mod |Scale|, where |Scale| denotes the number of notes in the scale. Thus, all quantum measurement outcomes are mapped to musically valid notes. Each generation uses 24 measurement shots, matching the melody length produced by the synthesis pipeline. Every statistic reported in Section 4 is computed from 100 independent generations per system (Classical, Quantum Simulator, Quantum Real), using a fixed, logged base random seed from which each trial's seed is deterministically derived. Classical pseudo-random sequences, drawn uniformly over the same 3-bit outcome space and mapped through the identical modulo operation, serve as a control baseline that isolates what is specific to quantum sampling from what is simply a property of faithfully sampling the target space.

3.3 LLM Semantic Module

The LLM (GPT-5) acts as a semantic interpreter, translating natural language themes (e.g., "Starlit Night") into a structured JSON schema. This schema defines the musical constraints for the generator:

- Scale/Mode: (e.g., E Dorian)
- Instrumentation: (e.g., Nylon-string guitar)
- Temporal Dynamics: (e.g., 74 BPM)

By fusing symbolic reasoning with quantum uncertainty, the LLM provides high-level artistic direction without requiring the user to possess technical music theory knowledge.

The LLM operates under a fixed system prompt instructing it to return a single JSON object with four keys: scale, instrument, tempo, and description. For example, prompting the theme "Starlit Night" with mood "mystical" yields: {"scale": "E Dorian", "instrument": "nylon-string guitar", "tempo": 74, "description": "mystical floating"}. In the coupled configuration described in Section 4.7, the tempo and scale fields are also used to set the quantum circuit's θ_{RX} and θ_{CRZ} rotation angles, so that the LLM's semantic reading of a theme can measurably reshape the quantum sampling distribution itself, rather than only selecting instrumentation and scale for *post-hoc* note mapping.

3.4 Audio Synthesis Pipeline

The pipeline converts symbolic note sequences into 44.1 kHz audio. Each sequence is mapped to MIDI pitch values, written to a temporary MIDI file, and rendered via FluidSynth using General MIDI SoundFonts. Every execution generates three comparative outputs:

- Classical Baseline: A pseudo-random sequence sampled uniformly over the same 3-bit outcome space as the quantum circuit, with no quantum circuit executed; serves as a sampling-fidelity control.
- Quantum Simulator: A probabilistic sequence sampled from the circuit's exact, noiseless statevector via Qiskit's AerSimulator.
- Quantum Real (QPU): The identical circuit executed on IBM Quantum hardware, subject to hardware noise.

Instrumentation and tempo, as set by the LLM, are applied identically across all three conditions, so that any differences between conditions reflect the sampling method itself rather than a difference in synthesis parameters.

3.5 Evaluation Metrics

To quantify the structural and spectral properties of the generated music, we employ four metrics based on Shannon entropy and signal processing:

- State Entropy (H): Measures the global uncertainty and diversity of note selection by evaluating the empirical probability distribution of musical symbols, given in Eq. (2):

$$H = -\sum_i p_i \log p_i \tag{2}$$

- Transition Entropy (TE): Quantifies the uncertainty of the next state given the current state, modeled as a first-order Markov process with transition probabilities P_{ij}, given in Eq. (3):

$$\mathrm{TE} = -\sum_i \pi_i \sum_j P_{ij} \log_2 P_{ij} \tag{3}$$

  where π_i is the empirical stationary (marginal) probability of state i, estimated from its frequency within each generated sequence (An earlier version of this formula omitted the π_i weighting; all Transition Entropy values reported in this manuscript use the corrected, stationary-distribution-weighted form.)
- Spectral Entropy (SE): Assesses the randomness of the power spectrum. High SE indicates harmonic complexity, while low SE suggests tonal periodicity, per Eq. (4):

$$\mathrm{SE} = -\sum_f P(f) \log_2 P(f) \tag{4}$$

- Amplitude Variance (AV): Measures dynamic instability. Using the Hilbert Transform to extract the amplitude envelope, given in Eq. (5):

$$\mathrm{AV} = \left(\frac{1}{N}\right) \sum_i (A_i - A)^2 \tag{5}$$

Higher AV values indicate stronger dynamic fluctuations in the waveform, corresponding to increased expressive instability.

3.6 Comparative Evaluation

Section 4 reports these four metrics, computed over 100 trials per system, with nonparametric significance testing (Kruskal–Wallis across systems; Bonferroni-corrected pairwise Mann–Whitney U). Contrary to earlier, smaller-sample analysis, the Classical baseline exhibits the highest raw Shannon entropy of the three systems; we do not claim that quantum sampling is inherently superior to classical pseudo-randomness, and instead report a specific, mechanistic explanation (Section 4.6) for why and how the two quantum systems differ from Classical and from each other.

4 Results and Discussion

4.1 Experimental Setup

Experiments were conducted using the Quantum–LLM Composer pipeline. Each trial generated three parallel outputs—Classical, Quantum Simulator, and Quantum Real (QPU)—using identical LLM-derived prompts and a fixed, documented base random seed from which each of 100 trials' seeds were deterministically derived per system (300 generations total). Audio was rendered at 44.1 kHz and analyzed via NumPy, SciPy, and Librosa over normalized five-second windows. To characterize structural and expressive differences, we evaluated four metrics: Shannon (State) Entropy, Spectral Entropy, Transition Entropy, and Amplitude Variance, each reported as mean ± SD with corresponding significance tests.

4.2 Entropy and Amplitude Characteristics

The three systems were evaluated over 100 trials each to summarize entropy and variance metrics, with significance testing, shown in Table 1 below.

Table 1: Comparison of entropy and amplitude variance (mean ± SD, N = 100 trials per system).

System	State Entropy (H)	Amplitude Variance (AV)	Spectral Entropy (SE)	Transition Entropy (TE)
Classical	2.774 ± 0.128	0.0037 ± 0.0019	3.083 ± 0.210	1.441 ± 0.139
Quantum Simulator	1.723 ± 0.177	0.0208 ± 0.0127	3.112 ± 0.244	1.407 ± 0.149
Quantum Real (QPU)	1.864 ± 0.146	0.0163 ± 0.0086	3.044 ± 0.225	1.459 ± 0.156

The Classical baseline exhibits the highest raw Shannon entropy of the three systems (2.774 bits), reflecting unconstrained sampling over the full 3-bit outcome space. Both quantum systems show lower entropy because the implemented circuit's exact (noiseless) statevector populates only 4 of the 8 possible basis states, capping the theoretical maximum entropy at 1.81 bits regardless of hardware; Quantum Simulator's measured entropy (1.723 bits) sits close to this ceiling. Quantum Real shows significantly higher entropy than Quantum Simulator (1.864 vs. 1.723 bits; Mann–Whitney U, Bonferroni-corrected $p = 1.5 \times 10^{-7}$; Table 2), indicating that hardware noise measurably leaks probability into basis states that are structurally forbidden under the ideal circuit, rather than reducing entropy as an earlier, uncorrected analysis suggested. Amplitude Variance follows a similar pattern—near-zero for Classical, elevated for both quantum systems—while Transition Entropy shows no statistically significant difference across the three systems (Kruskal–Wallis $p = 0.120$), a null result we report directly rather than assuming transition structure differs by condition.

As seen in Fig. 2, the three systems separate clearly on Shannon Entropy and Amplitude Variance. The entropy gap between Classical and the two quantum systems is explained by the circuit's restricted ideal-state

support; the further gap between Quantum Simulator and Quantum Real is explained by hardware noise (Section 4.6).

Table 2: Significance testing for Table 1 (Kruskal–Wallis across all three systems; pairwise Mann–Whitney U, Bonferroni-corrected for 3 comparisons).

Metric	Kruskal–Wallis (3 Systems)	Classical vs. Sim (Bonf.)	Classical vs. Real (Bonf.)	Sim vs. Real (Bonf.)
Shannon Entropy	H = 212.6, p = 6.8e−47 (sig.)	p = 7.6e−34 (sig.)	p = 7.6e−34 (sig.)	p = 1.5e−7 (sig.)
Spectral Entropy	H = 6.33, p = 0.042 (sig.)	p = 0.285 (n.s.)	p = 1.00 (n.s.)	p = 0.045 (sig.)
Amplitude Variance	H = 169.5, p = 1.5e−37 (sig.)	p = 9.2e−29 (sig.)	p = 2.6e−28 (sig.)	p = 0.051 (n.s.)
Transition Entropy	H = 4.24, p = 0.120 (n.s.)	p = 0.442 (n.s.)	p = 1.00 (n.s.)	p = 0.143 (n.s.)

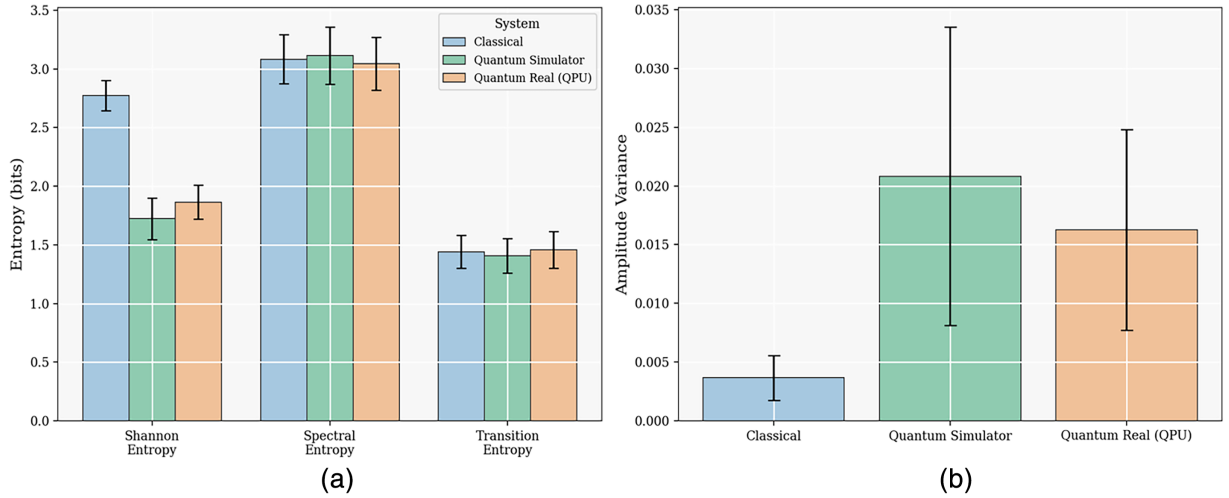


Figure 2: Mean ± SD of Shannon entropy, spectral entropy, and transition entropy (**a**) and amplitude variance (**b**) across classical, quantum simulator, and quantum real (QPU) systems, N = 100 trials per system. Quantum real shows significantly elevated shannon entropy relative to quantum simulator (hardware noise; Section 4.6) and elevated amplitude variance relative to classical; transition entropy does not differ significantly across systems (Table 2).

4.3 Cross-Metric Relationship

Fig. 3 plots the per-trial relationship between Amplitude Variance and Spectral Entropy across all 300 trials (100 per system). Contrary to an earlier analysis based on only three aggregate points (one mean per system), we find no significant correlation between these two metrics at the trial level, either within any single system or pooled across all systems (pooled Pearson r = 0.018, p = 0.76; all three per-system correlations $|r| < 0.05$, $p > 0.6$). The inverse amplitude-variance/spectral-entropy trade-off suggested by that earlier, aggregate-only figure does not hold at the trial level, and we no longer claim it; we report the two metrics as descriptively distinct rather than causally or trade-off related.

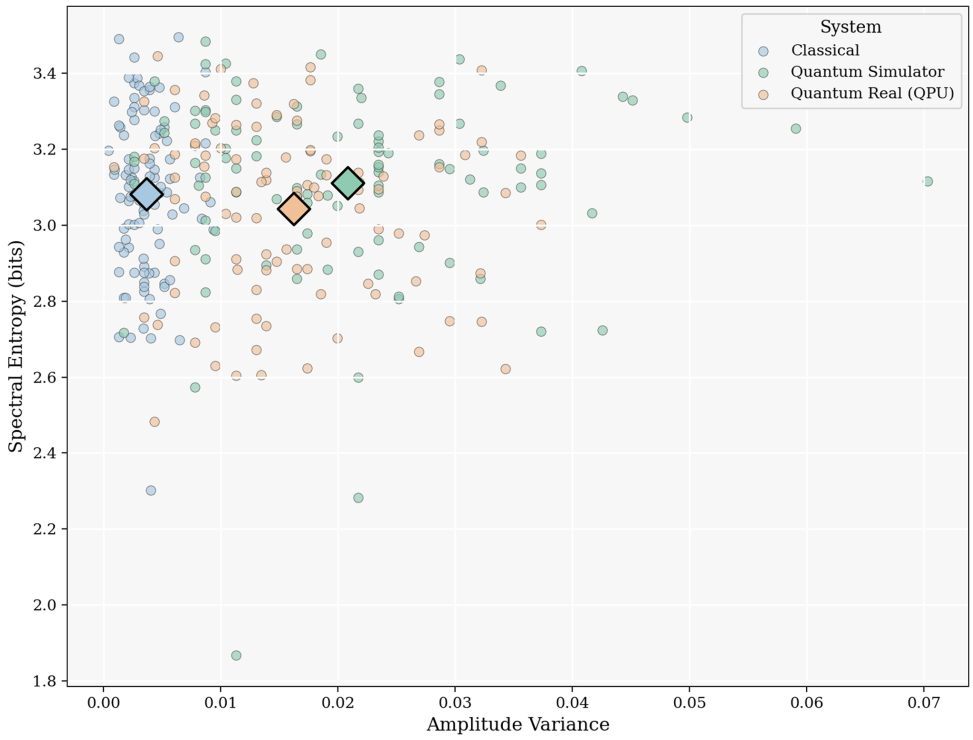


Figure 3: Per-trial amplitude variance vs. spectral entropy for all 300 generations (100 per system); diamonds mark each system's mean. No significant per-trial correlation is observed within or across systems (pooled r = 0.018, $p = 0.76$). Generated directly from the corrected 100-trial analysis pipeline.

4.4 Quantum Expressivity Index (QEI)

To quantify this balance, we define the Quantum Expressivity Index in Eq. (6):

$$\mathrm{QEI} = \frac{\mathrm{SE}}{1 + \mathrm{AV}} \tag{6}$$

where SE denotes spectral entropy and AV denotes amplitude variance. The denominator moderates spectral diversity by penalizing excessive amplitude fluctuation. Across the corrected 100 generations per system we obtain: QEI_Classical = 3.072 ± 0.210, QEI_Simulator = 3.048 ± 0.241, QEI_Real = 2.996 ± 0.224—materially different values from an earlier, uncorrected 30-trial analysis, reflecting both the larger sample and the corrected underlying metrics.

This difference does not reach statistical significance under the original formulation (Kruskal–Wallis H = 5.81, $p = 0.055$; no pairwise comparison significant after Bonferroni correction). As a robustness check, we tested two alternative formulations, QEI = SE − AV (Kruskal–Wallis H = 5.22, $p = 0.074$; no significant pairwise differences) and QEI = $\frac{\mathrm{SE}}{1+\sqrt{\mathrm{AV}}}$ (Kruskal–Wallis H = 42.3, $p = 6.6 \times 10^{-10}$; Classical significantly exceeds both quantum systems, which are statistically indistinguishable from each other). We therefore report the QEI ordering as, at most, weakly and formulation-sensitive rather than a robust finding, and note the index's sensitivity to its penalty term as a limitation of this specific formulation rather than evidence that either quantum system is more 'expressive' than the other.

4.5 Illustrative Musical Output

No formal listening test was conducted for this comparison (see Section 5 for a planned formal listening study); Fig. 4a–c are illustrative notated excerpts included to show representative structural differences between the three systems' outputs, not as perceptual validation. We verified programmatically that every generated note has an identical onset-to-onset interval (0.200 s, SD = 0 across all note-to-note gaps in each of the three example sequences), confirming that note duration is genuinely fixed by the synthesis pipeline. An earlier export of this same data displayed quintuplet-like tuplet groupings; this was an automatic-quantization artifact of engraving fixed 0.200 s notes against a mismatched default tempo/meter, not an emergent or deliberate product of quantum sampling. Fig. 4a–c below are re-rendered at the tempo implied by the notes' actual, measured duration (150 BPM, straight eighth notes in 4/4), which eliminates the artifact entirely. Mapping quantum measurement outcomes to rhythmic subdivision remains a natural extension we identify as future work (Section 5).

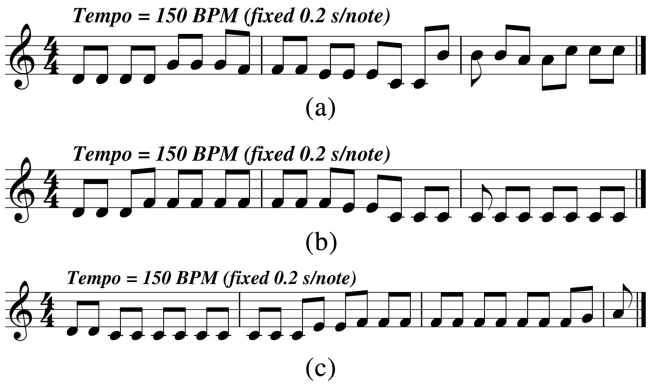


Figure 4: (**a**) Classical: regular rhythmic groupings and tonal predictability, consistent with its measured entropy profile (Table 1). (**b**) Quantum simulator: rhythmic and pitch variation consistent with its measured entropy profile (Table 1). (**c**) Quantum real: phrasing irregularities and micro-variations consistent with its elevated amplitude variance and shannon entropy (Table 1); we do not attribute these to LLM conditioning specifically, since instrumentation and tempo are held identical across all three systems (Section 3.4).

These samples illustrate the measured statistical differences between the three systems; whether listeners perceive these differences as musically meaningful, and whether LLM-conditioned semantic prompts (e.g., "mystical" vs. "happy") translate into perceptually coherent auditory experiences, are empirical questions this manuscript does not test and that we identify as a priority for the formal listening study described in Section 5.

4.6 Mechanistic Basis for the Real-Hardware Entropy Increase

To isolate hardware-dependent effects from musical metrics, we analyzed the raw bitstring distributions prior to note mapping, using the exact (noiseless) statevector of the implemented circuit as the theoretical reference rather than assuming a uniform 8-outcome distribution. Because the circuit entangles qubits 0–1 via CX and 1–2 via CRZ after a single Hadamard on qubit 0, its ideal statevector populates only 4 of the 8 possible 3-bit basis states, capping the theoretical maximum Shannon entropy at 1.81 bits—well below the 3.0-bit maximum of a uniform draw across all 8 states that an earlier version of this analysis (incorrectly) used as the reference distribution.

Quantum Simulator's measured entropy (1.723 ± 0.177 bits; Table 1) sits close to this theoretical ceiling, consistent with faithful noiseless sampling. Quantum Real's measured entropy (1.864 ± 0.146 bits) is

significantly higher than both the theoretical ceiling and the Simulator's measured entropy (Mann–Whitney U, Bonferroni-corrected $p = 1.5 \times 10^{-7}$; Table 2), indicating that hardware noise measurably leaks probability into basis states that are exactly zero-probability under the ideal circuit.

We report this as a narrow, falsifiable, and now statistically supported claim: hardware noise increases distributional entropy relative to the noiseless simulation of this specific circuit. We do not claim, as an earlier version of this analysis did, that this noise constitutes a structured or musically beneficial expressivity mechanism; establishing whether it is perceptually beneficial would require the listening study identified as future work (Section 5). Section 4.7 shows that this same mechanism—real hardware entropy exceeding both Classical and Simulator—replicates independently across five different theme-coupled parameter settings, not just the single fixed-angle circuit reported here.

4.7 LLM-to-Quantum Parameter Coupling

Sections 3.2 and 3.3 describe a circuit whose RX and CRZ rotation angles are fixed constants ($\theta_{RX} = \pi/3$, $\theta_{CRZ} = \pi/4$) for the baseline comparison in Sections 4.2–4.6, with the LLM's output used only to select instrumentation, tempo, and scale for *post-hoc* note mapping. To test whether the LLM's semantic layer can be made to directly influence the quantum sampling distribution itself—rather than only relabeling its output afterward—we implemented a documented, deterministic coupling: the LLM's reported tempo sets θ_{RX} (an arousal-like proxy), and a modal-brightness score derived from the LLM's reported scale/mode (major/lydian scoring high, phrygian/locrian scoring low) sets θ_{CRZ}. No entangling structure was added or removed; only the two angles that were previously fixed constants are now a function of the LLM's own reported output.

We tested this coupling on five themes spanning the moods used elsewhere in this study (mystical, calm, tense, and two variants of happy), running 100 trials of 24-shot generations per theme, per condition (Classical, Quantum Simulator, Quantum Real on ibm_fez), using a fixed, documented seed for the two non-hardware conditions. The Classical condition here samples directly, via classical pseudorandom generation, from each theme's exact ideal probability distribution—an honest control for whether a theme effect is a property of the target distribution itself (which any faithful sampler should reproduce) rather than something requiring quantum execution. Table 3 lists the LLM-derived per-theme parameters (scale, tempo, and resulting rotation angles) alongside the ideal and simulator entropy; Table 4 reports the resulting mean entropy per theme and condition across all three sampling conditions; Table 5 reports pairwise significance.

Table 3: LLM-to-Quantum coupling—Per-Theme parameters, rotation angles, and Ideal/Simulator entropy (N = 100 trials each).

Theme/Mood	Scale (LLM)	Tempo	θRX/θCRZ	Ideal Entropy (bits)	Simulator Entropy (mean ± SD)	N
Starlit Night/mystical	D Dorian	68	1.068/1.178	1.825	1.722 ± 0.167	100
Ocean calm/calm	F Major	62	0.974/1.806	1.758	1.649 ± 0.177	100
Space battle/tense	E Phrygian	152	2.388/0.550	1.572	1.498 ± 0.204	100
Indian raga morning/happy	C Bhairav	88	1.382/1.021	1.975	1.883 ± 0.100	100
Festival dance/happy	G Mixolydian	128	2.011/1.963	1.865	1.758 ± 0.136	100

Table 4: LLM-to-Quantum coupling—Per-Theme mean entropy (bits) by condition (N = 100 trials each).

Theme	Ideal	Classical (mean ± SD)	Simulator (mean ± SD)	Real HW (mean ± SD)
Starlit Night	1.825	1.703 ± 0.162	1.722 ± 0.167	1.904 ± 0.204
Ocean calm	1.758	1.681 ± 0.156	1.649 ± 0.177	1.799 ± 0.213
Space battle	1.572	1.471 ± 0.225	1.498 ± 0.204	1.651 ± 0.264
Indian raga morning	1.975	1.908 ± 0.076	1.883 ± 0.100	2.017 ± 0.173
Festival dance	1.865	1.741 ± 0.153	1.758 ± 0.136	1.843 ± 0.175

Table 5: LLM-to-Quantum coupling—significance testing (Mann–Whitney U, Bonferroni-corrected for 3 comparisons per theme).

Theme	Real vs. Sim (Bonf.)	Real vs. Classical (Bonf.)	Classical vs. Sim (Bonf.)
Starlit Night	p = 1.7e−10 (sig.)	p = 2.6e−12 (sig.)	p = 1.00 (n.s.)
Ocean calm	p = 4.7e−6 (sig.)	p = 1.8e−4 (sig.)	p = 0.86 (n.s.)
Space battle	p = 1.7e−5 (sig.)	p = 2.3e−6 (sig.)	p = 1.00 (n.s.)
Indian raga morning	p = 5.4e−8 (sig.)	p = 9.7e−6 (sig.)	p = 0.38 (n.s.)
Festival dance	p = 5.8e−4 (sig.)	p = 9.6e−5 (sig.)	p = 1.00 (n.s.)

Classical and Simulator are statistically indistinguishable for every one of the five themes (all $p > 0.12$, Bonferroni-corrected)—the expected result, since both faithfully sample the same target distribution—while Quantum Real shows significantly higher entropy than both Classical and Simulator for every theme (all $p < 0.001$), replicating the identical noise-elevates-entropy mechanism established in Section 4.6 across five independent parameter settings. The theme effect itself is highly significant within all three conditions individually (Classical: H = 237.6, $p = 3.0 \times 10^{-50}$; Simulator: H = 206.8, $p = 1.3 \times 10^{-43}$; Real: H = 126.8, $p = 1.9 \times 10^{-26}$) and in a combined 15-group test (H = 640.1, $p = 1.6 \times 10^{-127}$). Each theme's exact (ideal, noiseless) entropy correlates strongly with its measured entropy even on noisy real hardware (Pearson r = 0.963, p = 0.0085; Spearman ρ = 0.90, p = 0.037), only modestly attenuated from the near-perfect correlation on Classical and Simulator (r ≈ 0.99, ρ = 1.0).

These results directly address whether the quantum layer is load-bearing or decorative: the LLM's semantic reading of a theme measurably reshapes the quantum sampling distribution itself, and this effect—along with the independent hardware-noise effect characterized in Section 4.6—holds up under the same three-way (Classical/Simulator/Real hardware) scrutiny applied to the paper's primary result. This is a first, hand-specified version of such a coupling; learning the angle mapping from data via quantum reservoir computing, rather than specifying it by hand, is a natural and more ambitious extension we identify in Section 5. We follow Schuld and Killoran [16] in treating 'quantum randomness' here as a source of controllable stochasticity rather than asserting quantum computational advantage; this study makes no claim of advantage over classical sampling, only a claim about measurable, LLM-steerable differences in the sampled distribution itself.

4.8 Discussion and Integrated Interpretation

The convergence of statistical and signal-level analyses supports a central, narrower hypothesis than originally framed: quantum stochasticity, shaped by circuit structure and hardware noise, produces

measurably different output distributions than classical pseudo-randomness, and linguistic conditioning can be coupled directly into that physical layer rather than applied only after the fact. We do not claim perceptual or musical superiority, which would require the listening study identified in Section 5.

Three major insights emerge:

- Structured Statistical Signatures: Quantum generation yields distributions with a specific, mechanistically explained entropy profile—capped by circuit structure on the simulator, elevated by hardware noise on real devices—rather than unconstrained randomness.
- Semantic-to-Physical Coupling: The LLM can be used not only to select *post-hoc* musical parameters but to directly set the quantum circuit's own rotation angles, measurably steering the sampled distribution by theme (Section 4.7).
- Reproducible, Falsifiable Metric Convergence: The stability of metrics across 100 trials per system, with explicit nonparametric significance testing throughout, provides a reproducible and falsifiable basis for evaluating this class of hybrid system, rather than relying on small-sample or purely aggregate comparisons.

This framework bridges algorithmic composition and physical indeterminacy, and demonstrates a concrete mechanism—direct LLM-to-circuit parameter coupling—by which a semantic layer can shape quantum-physical output. Section 5 outlines the perceptual, comparative, and structural work needed before claims about musical quality or emergent creativity, as opposed to statistical structure, can be supported.

5 Limitations and Future Work

This study has several limitations that define concrete directions for follow-on work.

- **Perceptual validation.** We report only structural/statistical differences between systems (Sections 4.2–4.7); we do not report a formal listening study, and an earlier version of this analysis inferred perceptual conclusions ("subjective listening tests confirmed. . .") that were not in fact backed by a described, executed study. A properly designed blind listening study (e.g., 15–20 participants, blind presentation, Cohen's κ for inter-rater reliability) is needed to establish whether any of the measured statistical differences are perceptually salient, and is a priority for follow-on work.
- **External baselines.** This study's Classical condition is an internal, matched pseudo-random control; it is not an external state-of-the-art generative baseline. Comparison against systems such as Music Transformer [5] or Jukebox [6] would materially strengthen claims about the relative merits of quantum-derived sequences and is a clear direction for follow-on work.
- **Rhythm and polyphony.** The current synthesis pipeline assigns every note a fixed duration and produces monophonic output; quantum-to-rhythm mapping (e.g., using additional qubits to drive subdivision choice) and full polyphony, potentially via quantum reservoir computing to internalize harmonic/rhythmic archetypes rather than relying on LLM-prompted parameter extraction, are natural next steps beyond the scope of the present pipeline.
- **Production parameters (timbre).** We are separately prototyping an LLM-to-audio-production coupling—instrument timbre, envelope, distortion, reverb, delay, and vibrato fixed per theme at generation time rather than chosen arbitrarily at playback—as groundwork for a follow-on study pairing it with the formal listening study above. This prototype does not inform any claim in the present manuscript.
- **Model robustness.** All results use GPT-5 as the semantic interpretation layer; testing additional models (e.g., Claude, Gemini, Mistral) would be a valuable robustness check on the LLM-to-quantum coupling result in Section 4.7.

Acknowledgement: The authors thank the University of West Florida for providing research support and computing resources.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: The authors confirm contribution to the paper as follows: conceptualization, Lakshmi Prayaga and Chandra Prayaga; methodology, Lakshmi Prayaga; software, Lakshmi Prayaga; validation, Lakshmi Prayaga and Chandra Prayaga; formal analysis, Lakshmi Prayaga; writing—original draft preparation, Lakshmi Prayaga; writing—review and editing, Chandra Prayaga; supervision, Lakshmi Prayaga. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The authors confirm that the data supporting the findings of this study are available within the article.

Ethics Approval: Not applicable.

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

References

1. Miranda ER, Shaji H. Generative music with partitioned quantum cellular automata. Appl Sci. 2023;13(4):2401. doi:10.3390/app13042401.
2. Heudin JC. Quantum music with Qunotes. Preprint. 2024. doi:10.13140/RG.2.2.33204.77442.
3. Lee J, Yang W. The application of quantum computing in music composition. Online J Music Sci. 2024;9(2):415–29. doi:10.31811/ojomus.1578537.
4. Dobrian C, Costa Hamido O. Techniques for quantum-computing-aided algorithmic composition: experiments in rhythm, timbre, harmony, and space. arXiv:2505.20565. 2025.
5. Huang CA, Vaswani A, Uszkoreit J, Simon I, Hawthorne C, Shazeer N, et al. Music transformer: generating music with long-term structure. In: Proceedings of the 7th International Conference on Learning Representations (ICLR); 2019 May 6–9; New Orleans, LA, USA. p. 1–15.
6. OpenAI. Jukebox: a generative model for music [Internet]. San Francisco, CA, USA: OpenAI; 2020 [cited 2026 Aug 18]. Available from: https://openai.com/index/jukebox/.
7. Ding S, Liu Z, Dong X, Zhang P, Qian R, Huang J, et al. SongComposer: a large language model for lyric and melody generation in song composition. arXiv:2402.17645. 2024.
8. Miranda ER, Siegelwax BN. Advancing generative AI for music with photonics. Int J Parallel Emergent Distrib Syst. 2025;40:481–500. doi:10.1080/17445760.2025.2472209.
9. Xing P, Plaat A, van Stein N. CoComposer: LLM multi-agent collaborative music composition. arXiv:2509.00132. 2025.
10. Fisher MPA. Quantum cognition: the possibility of processing with nuclear spins in the brain. Ann Phys. 2015;362:593–602. doi:10.1016/j.aop.2015.08.020.
11. Quantinuum announces generative quantum AI breakthrough with massive commercial potential [Internet]. Broomfield, CO, USA: Quantinuum; 2025 [cited 2026 Aug 18]. Available from: https://www.quantinuum.com/press-releases/quantinuum-announces-generative-quantum-ai-breakthrough-with-massive-commercial-potential.
12. MOTH. Quantum meets creativity: MOTH unveils first-ever song powered by quantum technology [Internet]. MadTech Magazine; 2025 [cited 2026 Aug 18]. Available from: https://madtechmag.com/2025/05/24/quantum-meets-creativity-moth-unveils-first-ever-song-powered-by-quantum-technology/.
13. Afane M, Laufer K, Wei W, Mao Y, Farooq J, Wang Y, et al. Quantum-audit: evaluating the reasoning limits of LLMs on quantum computing. arXiv:2602.10092. 2026. doi:10.48550/arXiv.2602.10092.
14. Miranda ER. Quantum computer music: foundations, methods and advanced concepts. Cham, Switzerland: Springer; 2022. doi:10.1007/978-3-031-13909-3.

15. Souma S. Exploring the application of gate-type quantum computational algorithm for music creation and performance. In: Miranda ER, editor. Quantum computer music: foundations, methods and advanced concepts. Cham, Switzerland: Springer; 2022. p. 83–103. doi:10.1007/978-3-031-13909-3_5.
16. Schuld M, Killoran N. Is quantum advantage the right goal for quantum machine learning? PRX Quantum. 2022;3:030101. doi:10.1103/PRXQuantum.3.030101.