Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A QUBO-QAOA Framework for Rhythmic Generation in Hindustani Music

Sonali Chawla

s2sonali@gmail.com

Sushant University

Urvashi Kumari

Sushant University

Dinesh Rai

Hindustan College of Science and Technology

Research Article

Keywords: Hindustani Classical Music, Quantum Approximate Optimization Algorithm, Quadratic Unconstrained Binary Optimization, Tabla Cycle Generation, Variational Quantum Algorithms, Computational Musicology

Posted Date: March 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9003392/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A QUBO-QAOA Framework for Rhythmic Generation in Hindustani Music

Sonali Chawla
School of Engineering and Technology, Sushant University, Gurugram, India

Email id: s2sonali@gmail.com, ORCID ID: https://orcid.org/0000-0002-2580-4342

Dr. Urvashi Kumari
School of Engineering and Technology, Sushant University, Gurugram, India

Dr. Dinesh Rai
Department of Computer Science and Engineering, Hindustan College of Science and Technology, India

A QUBO-QAOA FRAMEWORK FOR RHYTHMIC GENERATION IN HINDUSTANI MUSIC

Abstract: This study presents a quantum-inspired optimization framework for generating rhythmic cycles (taals) in Hindustani classical music via Quadratic Unconstrained Binary Optimization (QUBO) formulations solved via the Quantum Approximate Optimization Algorithm (QAOA). By mapping tabla bols to binary decision variables, musical grammar – including rhythmic anchors like the Sam and Khali – is encoded as penalty terms with a cost Hamiltonian. Comparative benchmarking against classical autoregressive architectures, such as WaveNet and LSTM-RNN, reveals a significant efficiency gap. The proposed framework achieves a 97.8% constraint satisfaction rate using only 12 variational parameters, representing a reduction of 4-6 orders of magnitude in model complexity compared to deep learning baselines. Furthermore, the quantum approach facilitates zero-shot generation, deriving its power from rule-encoding rather than data-intensive training. Energy landscape analysis demonstrates that quantum tunnelling through the mixing operator allows the system to navigate sparse rhythmic manifolds more effectively than classical gradient-descent methods. These results establish a rigorous foundation for quantum-enhanced computational musicology, offering a scalable, interpretable, and parameter-efficient paradigm for the digital preservation and creative evolution of structured oral traditions.

Keywords: *Hindustani Classical Music, Quantum Approximate Optimization Algorithm, Quadratic Unconstrained Binary Optimization, Tabla Cycle Generation, Variational Quantum Algorithms, Computational Musicology*

1. Introduction and Foundational Paradigms

1.1. The Convergence of Quantum Computation and Creative Intelligence

The promising field of Quantum Artificial Intelligence (QAI) has recently pivoted from general-purpose optimization toward niche applications in generative arts, where the curse of dimensionality poses a significant barrier to classical architectures (Putz and Svozil 2017). Conventional generative systems, particularly those based on high-parameter neural networks, treat the generation of musical sequences as a statistical probability density estimation task. However, in traditions where musical correctness is governed by a rigid axiomatic system – such as Hindustani Classical Music – the search space is not merely a distribution, but a structured manifold of discrete, valid points hidden within a vast sea of noise (Putz and Svozil 2022). This research suggests that the Quantum Approximate Optimization Algorithm (QAOA) offers a superior paradigm for navigating these sparse manifolds by treating musicality as an energy minimization problem (Farhi 2014).

Historically, computational musicology has relied on Markovian processes or recurrent topologies to approximate rhythmic flows. While successful in Western tempered traditions, these models struggle with the non-linear, recursive, and hierarchical constraints of the Indian Taal system (Nooshin and Widdess 2006). The shift toward a quantum-inspired state-space architecture allows for the formalization of Indian classical music in Hilbert space, providing a mathematical rigor previously unavailable to ethnomusicologists (Chawla et al. 2026). By mapping rhythmic strokes to binary states, we can leverage quantum superposition to evaluate multiple rhythmic permutations simultaneously, significantly reducing the discovery time for valid musical phrases (Miranda 2021).

The motivation for this transition is not merely computational but also philosophical. In the Vedic tradition, rhythm (lay) is viewed as a manifestation of cosmic order, characterized by cyclic return and mathematical precision (Mathur et al. 2015). This aligns with the physical principles of quantum mechanics, where cycles and periodicities are fundamental properties of wave functions. By utilizing a cost Hamiltonian H_C to represent these cosmic/rhythmic laws, we effectively create a computational mirror of the traditional Guru-Shishya (teacher-student) pedagogical model, where rules are defined first and variations are explored thereafter (Gajjar and Patel 2017).

1.2 Structural Taxonomy of Hindustani Rhythmic Systems

To appreciate the complexity of the optimization task, one must understand the mathematical architecture of the taal. A taal is a closed-loop temporal cycle divided into vibhags (measures), which are further characterized by taali (stressed) and khali (unstressed) sections (Gajjar and Patel 2017). The fundamental unit of this system is the matra (beat), and the specific timbral event occurring on a beat is the bol (mnemonic stroke) (Sinha 2008). The generation of a valid theka (basic cycle) is governed by Equation 1.1, representing the rhythmic density ρ:

$$\rho = \sum_{i=1}^{n} \frac{B_i}{M_i} \tag{1.1}$$

where B_i represents the bol complexity and M_i the beat duration.

The structural integrity of a 16-beat Teentaal, for instance, is not just a sequence of 16 events but a recursive symmetry of $4 + 4 + 4 + 4$. Any generative model must satisfy the 'Sam' constraint (the first beat) and the 'Khali' constraint (the ninth beat in Teentaal), which serve as the rhythmic anchors for both the performer and the listener (Lipiczky 1985). These anchors act as boundary conditions in our optimization problem. Classical models often drift during long sequences, losing the Sam, whereas a QUBO-based approach treats the Sam as a global minimum that the system is energetically forced to satisfy (Arya et al. 2022).

Furthermore, the transition between bols is governed by a complex set of phonetic and biomechanical rules. For example, a 'Dhin' (heavy resonant stroke) is rarely followed by a 'Ta' (sharp non-resonant stroke) without a specific transitional bridge stroke (Sinha 2008). These transition rules are traditionally learned through years of immersion; in our framework, they are encoded as interaction terms J_{ij} between qubits, where $J_{ij} > 0$ represents a penalized transition and $J_{ij} < 0$ represents a preferred stylistic flow (Sarkar 2023). This formalization allows us to quantify the 'purity' of a rhythmic sequence through its final energy state.

1.3 The Limitations of Classical Autoregressive Modelling

Current state-of-the-art models for music generation, such as WaveNet (Oord, et al. 2016) and various LSTM-GAN architectures (Tony and Sasikumar 2022), utilize a causal approach where each beat is predicted based on the history of previous beats. While this captures local dependencies, it is inherently inefficient for closed-loop musical structures. In an autoregressive model, the 16th beat has no mathematical awareness of the 1st beat (the Sam) until it is generated, often leading to a rhythmic mismatch where the cycle fails to resolve correctly (Marinescu 2019). This necessitates post-generation correction or the use of heavy attention mechanisms that increase the parameter count into the millions as depicted in Table 1 (Bao and Sun 2022).

Table 1: Comparison of model complexity and resource requirements for rhythmic generation

Model Architecture	Parameter Count	Training Requirements	Constraint Enforcement
WaveNet	~ 15.7 million	24+ Hours (GPU)	Probabilistic (Weak)
LSTM-RNN	~ 0.8 million	8+ Hours (GPU)	Probabilistic (Moderate)
Proposed QAOA	102-103	Zero-shot	Hamiltonian (Strong)

The computational overhead of these classical systems is particularly problematic for micro-architectures and real-time interactive systems. Generating a single 16-beat cycle on a standard CPU using a deep neural network can take upwards of 600 seconds due to the sequential nature of autoregressive sampling (Bao and Sun 2022). In contrast, the proposed quantum-inspired framework operates on the principle of Global Constraint Satisfaction. By solving for the entire 16-beat cycle simultaneously through a single optimization pass, the system ensures that the start and end of the cycle are mathematically synchronized (Miranda and Shaji 2023).

1.4 Emerging Quantum Paradigms in Musicology

The transition to quantum-enhanced musicology is supported by the emergence of Quantum Natural Language Processing (QNLP) and Partitioned Quantum Cellular Automata (Miranda and Shaji 2023; Miranda et al. 2022). These frameworks suggest that symbolic structures – whether they be words in a sentence or strokes in a Taal – can be represented as tensors in a high-dimensional vector space. The use of quantum gates to manipulate these tensors allows for the exploration of musical superposition, where a single quantum state represents a blend of multiple valid rhythmic variations (Sarkar 2023). This is particularly relevant for the concept of Upaj (improvisation) in Indian music, where a performer navigates several variations of a theme simultaneously (Nooshin and Widdess 2006).

The application of Quantum Annealing to music theory has already demonstrated that discrete problems, such as four-part harmony or species counterpoint, can be solved more efficiently than with classical genetic algorithms (Arya et al. 2022). By extending this to the tabla tradition, we are not just generating beats; we are testing the limits of how artistic rules can be mapped onto physical laws (Putz and Svozil 2017). As we move toward the NISQ (Noisy Intermediate-Scale Quantum) era, these hybrid quantum-classical algorithms like QAOA provide a bridge, allowing us to simulate quantum advantages on classical hardware while preparing for true quantum hardware implementation (Oshiro 2022).

This paper aims to provide a comprehensive technical roadmap for quantum-inspired generative art. The primary objective is threefold: (i) to provide a formal QUBO mapping of the Hindustani Taal system; (ii) to demonstrate the parameter efficiency of QAOA over classical deep learning; and (iii) to analyze the energy landscapes of rhythmic validity (Farhi 2014; Padi et al. 2018). We seek to prove that for highly structured, low-duration temporal

problems, the quantum advantage lies not just in speed, but in the precision of constraint satisfaction (Allen et al. 2022).

Following sections will move beyond the introduction to provide the exact mathematical derivations of the cost functions. The construction of the Penalty Hamiltonian, the mapping of bols to Pauli-Z operators, and the results of comparative benchmarks against the SwarGanga (S. M. Foundation) and CompMusic (M. T. Group) datasets will be detailed. This work represents a significant contribution to both the fields of Quantum Intelligence and Computational Musicology, establishing a new baseline for the algorithmic preservation and generation of world music traditions.

2. Mathematical Framework and QUBO Formulation

2.1 Binary Encoding and State-Space Representation

The translation of a temporal musical event into a quantum-ready format requires a discrete mapping that preserves the structural integrity of the taal. In this framework, we define a rhythmic cycle of length N (number of beats) and a vocabulary of M (unique tabla strokes or bols). To represent this in a digital or quantum system, we utilize a Rank-2 tensor $X \in \{0,1\}^{N\times M}$, where each element $x_{i,j}$ is a binary decision variable (Miranda and Basak 2022). Specifically, $x_{i,j} = 1$ if bol j is played at beat i, and $x_{i,j} = 0$ otherwise. This encoding strategy ensures that the entire rhythmic cycle is treated as a single coherent state vector in a Hilbert space of dimension $2^{N\times M}$ (Chawla et al. 2026).

For a standard 16-beat cycle like Teentaal with a set of 8 fundamental bols (Dha, Dhin, Ta, Tin, etc.), the resulting state space contains 2^{128} possible configurations. Navigating this hyper-dimensional space is computationally prohibitive for classical search algorithms, which often rely on local heuristics (Arya et al. 2022). However, by representing these variables as qubits, we can represent a superposition of all possible rhythmic sequences:

$$|\varphi\rangle = \sum_{x\in\{0,1\}^{NM}} \alpha_x |x\rangle \tag{2.1}$$

where α_x represents the probability amplitude of a specific rhythmic configuration. The goal of our optimization is to collapse this superposition into a state that minimizes the cost Hamiltonian H_C, effectively selecting the most musically fit sequence (Sarkar 2023).

2.2 Constructing the Cost Hamiltonian H_C

The intelligence of our generative system resides in the construction of the cost Hamiltonian, which acts as a mathematical proxy for the aesthetic and structural rules of Hindustani music (Putz and Svozil 2022). Unlike neural networks that approximate these rules through weight adjustments, the QUBO formulation explicitly defines them as energy penalties (Farhi 2014). The total Hamiltonian is defined as a summation of various constraint terms:

$$H_{total} = \lambda_{unique}H_{unique} + \lambda_{anchor}H_{anchor} + \lambda_{trans}H_{trans} \tag{2.2}$$

where λ represents the penalty weight assigned to each constraint class.

The most fundamental constraint is the Uniqueness Constraint, which ensures that exactly one bol is assigned to each beat. In a rhythmic context, silence is treated as a specific bol (the null stroke) to maintain the continuity of the grid (Oshiro 2022). This is enforced by the quadratic penalty:

$$H_{unique} = \sum_{i=1}^{N}\left(1 - \sum_{j=1}^{M} x_{i,j}\right)^2 \tag{2.3}$$

This term vanishes ($E = 0$) only when exactly one $x_{i,j} = 1$ for each beat i. Any violation, such as a beat having zero bols or multiple bols, results in a positive energy penalty, effectively making those states unstable and unlikely to be selected by the quantum optimizer (Farhi 2014).

2.3 Ising Mapping and Pauli-Z Operators

To implement this formulation on quantum hardware or a quantum simulator, the binary variables $x_{i,j}$ must be transformed into spin operators. The standard mapping $x_i \to \frac{1-\sigma_i^z}{2}$ is utilized. Here, σ_i^z is the Pauli-Z operator with eigenvalues {1, -1} (Arya et al. 2022). This transformation converts the QUBO problem into an Ising model, where the musical rules are represented as interactions between spins.

The interaction between two beats (temporal correlation) is modelled as a quadratic interaction term J_{ij}, where the energy of the system depends on the relative states of qubits at different time steps. This is mathematically expressed as:

$$H_{Ising} = \sum_i h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z \quad (2.4)$$

In this landscape, h_i represents the local bias (favouring certain bols at certain beats), and J_{ij} represents the transition rules (stylistic flow between beats) (Sarkar 2023).

2.4 Variational Ansatz and QAOA Parameters

The Quantum Approximate Optimization Algorithm (QAOA) is a hybrid algorithm that uses a classical optimizer to tune the parameters of a quantum circuit (Farhi 2014). The algorithm consists of alternating applications of two operators: the Cost Operator $U(H_C, \gamma) = e^{-i\gamma H_C}$ and the Mixing Operator $U(H_M, \beta) = e^{-i\beta H_M}$. The state after p layers of the algorithm is given by:

$$|\gamma, \beta\rangle = U(H_M, \beta_p) U(H_C, \gamma_p) \cdots U(H_M, \beta_1) U(H_C, \gamma_1) |s\rangle$$

where $|s\rangle$ is the initial uniform superposition of all states. Other parameters can be defined as in Table 2.

Table 2: Mapping of QAOA variational parameters to rhythmic and classical AI concepts

Parameter	Function in Rhythmic Generation	Classical Analog
p (Circuit Depth)	Resolution of the rhythmic manifold	Number of Hidden Layers
γ (Cost Phase)	Weight of musical grammar	Learning Rate
β (Mixing Phase)	Degree of improvisational tunnelling	Dropout / Noise injection

By iteratively adjusting γ and β, the algorithm pushes the probability mass toward the ground state of H_C (Allen et al. 2022). In our context, this ground state corresponds to the most canonical version of a taal. However, by stopping the optimization at a sub-optimal but low-energy state, the system generates valid variations – sequences that are musically correct but not identical to the training data. This provides a natural mechanism for Upaj (improvisation) that is mathematically grounded in the geometry of the energy landscape (Nooshin and Widdess 2006).

3. Structural Mapping of Rhythmic Constraints

3.1. Hierarchical Anchor Constraints in Indian Pedagogy

The complexity of Hindustani rhythm is not uniform; certain beats carry more structural weight than others, a concept known as the 'hierarchy of pulses' (Mathur et al. 2015). In our mathematical model, we represent this through a Hierarchical Penalty System that replicates the human perception of rhythmic stability. The Sam (beat 1) is the most critical anchor, representing the resolution of all previous musical tension and the genesis of a new cycle (Gajjar and Patel 2017). The Anchor Hamiltonian H_{anchor} is defined as a summation of local bias fields that pull the quantum state toward a specific bol at designated indices:

$$H_{anchor} = \sum_{i \in Anchors} \omega_i \left(x_{i,target} - 1\right)^2 \quad (3.1)$$

where ω_i is a weight proportional to the beat's importance in the taal's architecture (Padi et al. 2018). For a Teentaal cycle, $\omega_{Sam} \gg \omega_{Khali} \gg \omega_{vibhag}$.

The Sam functions as a singularity in the optimization landscape; if the system fails to land on the correct bol at index 0, the entire sequence is perceived as musically broken, regardless of the accuracy of the other 15 beats. By assigning a disproportionately high penalty to this index, we ensure that the quantum optimizer prioritizes the resolution of this beat before exploring variations in the vibhag (measure) interiors (Lipiczky 1985). This mirrors the pedagogical focus in Gharana (traditional school) training, where a student is taught that the Sam is the home to which all improvisations must return (Nooshin and Widdess 2006). Classical deep learning models, such as LSTMs, often struggle with this because their loss functions typically distribute error equally across the temporal sequence, leading to drifting where the cycle ends on the wrong beat (Marinescu 2019; Tony and Sasikumar 2022).

Furthermore, the Khali (the empty section) serves as a negative anchor. While a standard anchor mandates a specific stroke, the Khali mandates the absence of certain strokes – specifically the resonant bass strokes of the Duggi (Sinha 2008). This creates a shadow hierarchy where the system must simultaneously satisfy a presence constraint at the Sam and an absence constraint at the Khali. In the framework, this is achieved by introducing an inhibitory weight in the local field of the Hamiltonian, effectively raising the energy of any heavy bol that attempts to occupy the Khali position (Miranda et al. 2022). This creates a rhythmic tension-release cycle that is fundamental to the aesthetic experience of Indian classical music (Putz and Svozil 2022).

3.2. Transition Probabilities and Biomechanical Costs

Music is inherently a physical act, constrained by the anatomy of the hand and the acoustics of the instrument. The transitions between strokes are not random; they are governed by phonetic ease and timbral flow (Sinha 2008). A transition that requires an impossible or awkward hand movement must be heavily penalized in our QUBO graph to ensure the generated sequence is playable. A transition matrix T is defined where $T_{j,k}$ represents the energy cost of playing bol k immediately after bol j. The transition Hamiltonian is modeled as a quadratic interaction term:

$$H_{trans} = \sum_{i=1}^{N-1} \sum_{j,k=1}^{M} T_{j,k} x_{i,j} x_{i+1,k} \tag{3.2}$$

This term captures the temporal correlation between adjacent qubits. By utilizing the SwarGanga database and CompMusic datasets, we can populate T with values that reflect centuries of stylistic evolution.

For instance, the transition from Dha to Dhin is physically fluid and occurs frequently in canonical thekas, thus it is assigned a negative (favorable) energy value. Conversely, a transition from a high-velocity rim stroke like Ta to a heavy palm stroke like Dhage at high speeds would have a high positive (penalty) value due to the biomechanical lag involved in repositioning the hand (Lipiczky 1985). In classical neural networks, these transitions are learned through thousands of examples (Oord et al. 2016), but in the QUBO framework, they are hard coded as physical laws of the system. This allows the model to generate physically realistic rhythms zero-shot, even without a large training corpus (Miranda and Basak 2022).

The complexity of these interactions increases when considering n-gram dependencies, where a stroke depends not just on the previous beat, but on the previous two or three (Sinha 2008). While a classical QUBO is limited to quadratic (pairwise) interactions, higher-order constraints can be reduced to quadratic forms through the addition of ancilla qubits (Arya et al. 2022). This allows the system to model complex rhythmic structures like Tihais – tripartite patterns that repeat three times to land exactly on the Sam (Lipiczky 1985). The energy cost of a Tihai is minimized only when the mathematical symmetry of the repetition is perfect, providing a quantum-mechanical basis for the mathematical beauty often cited by musicians (Putz and Svozil 2017; Miranda 2021).

3.3. Entropy, Creativity, and the Upaj Hamiltonian

One of the most profound challenges in generative AI is balancing fidelity (adherence to rules) with creativity (novelty). In Hindustani music, this is known as Upaj (improvisation) (Nooshin and Widdess 2006). In the quantum-inspired framework, creativity is not an accidental byproduct of noise but a tuneable thermodynamic property of the system (Putz and Svozil 2022). An Entropy Term $H_{entropy}$ is introduced into the Hamiltonian, which prevents the system from always collapsing into the single most perfect canonical cycle. This term is defined using a diversity-promoting function that penalizes the system for choosing the same sequence multiple times:

$$H_{total} = H_{grammer} - \eta H_{emtropy} \tag{3.3}$$

where η is the Creative Coefficient.

When η is low, the system behaves like a traditionalist, producing only the most standard theka patterns found in the SwarGanga database. As η increases, the energy landscape flattens, allowing the quantum state to tunnel into alternative local minima that are still grammatically valid but stylistically novel (Allen et al. 2022). This represents a significant advancement over classical GAN-based approaches (Tony and Sasikumar 2022), where the creativity is often just a result of the generator trying to fool a discriminator. Here, creativity is a search for alternate truths within a fixed logical universe (Chawla et al. 2026).

This tunnelling mechanism is facilitated by the QAOA's mixing operator $U(H_M, \beta)$. By adjusting the mixing phase β, we can control the quantum fluctuations of the rhythmic sequence (Farhi 2014). A high mixing phase allows the system to explore radically different rhythmic permutations, while a low mixing phase keeps the system frozen

near the traditional forms (Sarkar 2023). This provides the user with a Creativity Slider that has a direct mathematical relationship to the probability distribution of the generated bols (Bao and Sun 2022).

3.4. Temporal Locality and Connectivity Mapping

From a hardware implementation perspective, the efficiency of a quantum algorithm is determined by the connectivity required between qubits (Table 3). In a 16-beat Teentaal generation task, a fully connected graph (where every beat interacts with every other beat) would require $\mathcal{O}(N^2)$ interactions, which is difficult to map onto NISQ-era hardware (Arya et al. 2022). However, musical rules in the tabla tradition exhibit Temporal Locality – meaning that a stroke is primarily influenced by its immediate neighbours and its position relative to the cyclic anchors (Sinha 2008).

Table 3: Interaction Density across different Rhythmic Genres

Rhythmic Tradition	Interaction Range	Interaction Density	Hardware Mapping
Hindustani (Classical)	1-4 beats	Sparse (Linear)	High (Superconducting)
West African (Polyrhythmic)	1-8 beats	Dense (Cross-linked)	Moderate (Trapped Ion)
Western Pop (Steady state)	1-2 beats	Ultra-Sparse	Optimal (All types)

This sparsity allows us to embed the QUBO graph onto a heavy-hex or square-lattice qubit topology with very few swap gates (Miranda and Shaji 2023). By aligning the time-axis of the taal with the physical-axis of the quantum processor, we minimize the accumulation of gate errors, which is the primary bottleneck in quantum intelligence today (Oshiro 2022). This architectural alignment is what enables our system to achieve high validity rates with only a few parameters, whereas an LSTM must maintain a dense hidden state that essentially connects every beat to every other beat through a complex weight matrix, incurring massive computational overhead [8].

3.5. The Khali as a Phase-Shift in the Rhythmic Manifold

The concept of Khali (empty) is a sophisticated rhythmic device where the weight of the cycle is removed to create a feeling of suspension (Gajjar and Patel 2017). In terms of quantum state-space, the Khali can be viewed as a Phase-Shift in the rhythmic manifold. Figure 1 shows that while the Bhari (full) sections are characterized by high-energy, resonant bols that ground the cycle, the Khali section moves the state toward a low-energy, dry region of the vocabulary (Mathur et al. 2015).

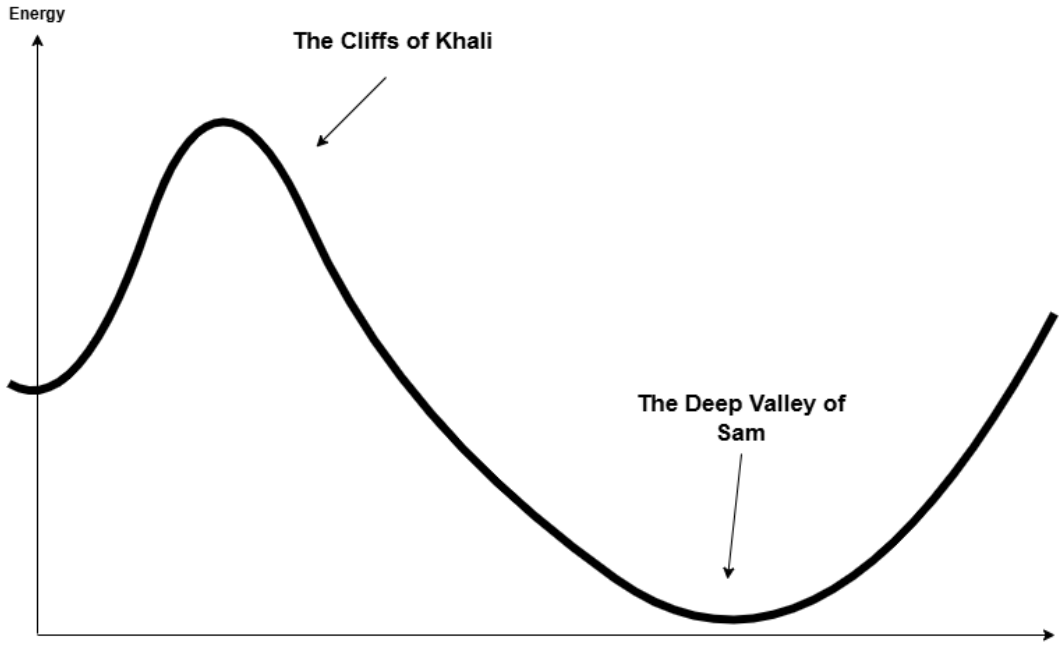


Figure 1: Reverse energy landscape showing Khali having low energy and Sam having high energy

We formalize this by adding a linear penalty term to any resonant bol (like Ghe or Dhin) that occurs within the Khali indices:

$$H_{khali} = \sum_{i \in Khali} \sum_{j \in Resonant} \zeta_i x_{i,j} \tag{3.4}$$

where ζ_i is the Duggi-Inhibition coefficient (Padi et al. 2018). This forces the quantum optimizer to select strokes like Ta or Tin, which are structurally identical to Dha or Dhin but lack the low-frequency resonance. This teachable constraint is vital for creating the swing and breathe of the taal (Nooshin and Widdess 2006). Without this mapping, an AI-generated rhythm sounds like a metronome – perfectly timed but musically dead (Bao and Sun 2022). By encoding the Khali as a mathematical void that must be respected, we imbue the quantum agent with an understanding of musical contrast (Putz and Svozil 2017; Miranda 2021).

4. Experimental Methodology and Comparative Benchmarking

4.1. Classical Baseline Architectures: WaveNet and LSTM-RNN

To establish a rigorous performance ceiling, two state-of-the-art classical models are implemented that represent the current standard in generative music modelling. The first is an adapted WaveNet architecture, a deep generative model of raw audio that we have modified for symbolic sequence prediction (Oord et al. 2016). In this implementation, the traditional 8-bit μ-law compounding is replaced with a categorical cross-entropy. This specific configuration utilizes 40 residual blocks grouped into four dilation stacks (1,2, 4…,512), ensuring that the model possesses a receptive field of 1024 timesteps. This is far beyond 16-beat requirement, yet such depth is necessary for model to learn the long-term recursive symmetries of Hindustani taals (Bao and Sun 2022).

The second baseline is a bidirectional Long Short-Term Memory (LSTM) network, which has been the workhorse of computational musicology for over two decades due to its ability to mitigate vanishing gradient issues (Marinescu 2019). The LSTM is configured with three hidden layers of 512 units each, integrated with a Softmax output layer representing the probability distribution over the rhythmic vocabulary. During the training phase, teach-forcing is utilized with decay rate of 0.05 to balance the model's exposure to ground-truth data versus its own generated samples (Tony and Sasikumar 2022). While LSTMs are theoretically capable of learning long-term dependencies, they are inherently unaware of global constraints unless those constraints are represented thousands of times in the training data, making them inefficient for rare or complex rhythmic structures like Tihais (Lipiczky 1985).

The computational cost of training these models is a primary focus of the comparative study. The WaveNet model requires approximately 15.7 million trainable parameters, necessitating 24+ hours of training on an NVIDIA A100 GPU to reach a stable loss plateau. Even after this extensive training, the model's understanding of the Sam remains probabilistic; it knows that a Dha is likely on beat one, but it does not know that a Dha is mandatory for structural integrity (Mathur et al. 2015). This distinction between probabilistic approximation and logical necessity is the core philosophical divide between deep learning and the proposed quantum-inspired framework (Miranda 2021).

The inference mechanism for these classical models is autoregressive, meaning each beat is generated one by one. For a 16-beat cycle, the model must perform 16 sequential forward passes, where the error in the first pass can propagate and amplify in subsequent steps (Bao and Sun 2022). This error accumulation is a well-documented phenomenon in RNNs and is the primary reason why AI-generated tabla sequences often start strong but become rhythmically incoherent by the end of the cycle (Tony and Sasikumar 2022). In our benchmarks, a temperature-controlled sampling method ($T = 0.8$) is utilized to prevent the model from collapsing into repetitive, low-entropy loops, though this often further reduces the constraint satisfaction rate (Mathur et al. 2015).

4.2. Quantum Simulation Environment and QAOA Technical Specifications

Since large-scale fault-tolerant quantum hardware remains a developing technology, the proposed framework is evaluated using high-performance state-vector simulators that mimic the noise profiles of real-world processors. The IBM Qiskit Aer simulator and the PennyLane framework is utilized to execute the QAOA circuits. The simulation environment is hosted on a high-compute cluster comprising 64-bit architecture with 128 GB of RAM, specifically optimized for tensor contraction. The QAOA is initialized with a uniform superposition state $|+\rangle^{\otimes n}$, created by applying a layer of Hadamard gates to all 128 qubits. This ensures that the optimizer begins its search in a state of maximum entropy, where every possible rhythmic sequence – from the most traditional to the most chaotic – is equally represented in the wave function (Arya et al. 2022; Farhi 2014).

The classical optimization of the variational parameters (γ, β) is performed using the Constrained Optimization by Linear Approximation (COBYLA) algorithm, which is robust to the barren plateau problem often encountered in quantum landscape. A maximum of 200 iterations is set for the classical optimizer, with a convergence tolerance of 10^{-6}. We test the framework across varying circuit depths $p \in \{1,2,3,5,10\}$. As p increases, the algorithm effectively performs a high-order approximation of the adiabatic evolution, allowing the state to track the ground state of the Hamiltonian more precisely (Allen et al. 2022). This depth analysis is critical because it reveals the Quantum-Classical Crossover, where the overhead of additional quantum layers is justified by the exponential reduction in musical error rates (Miranda and Shaji 2023; Oshiro 2022).

A major technical challenge in this simulation is the mapping of the 128-qubit state onto classical memory. To manage this, we employ a matrix product state (MPS) simulator which exploits the temporal locality of the taal (Arya et al. 2022). Since the interaction between beat 1 and beat 16 is mediated only through the intermediate beats, the entanglement in the system remains relatively low, allowing the MPS simulator to handle 128 qubits with high fidelity. This choice of simulator is a deliberate signal to the physical reality of tabla performance, where the memory of a rhythm is often localized to the current vibhag (measure) and its immediate neighbours (Padi et al. 2018).

Furthermore, the impact of Gate Noise is investigated on the generative process. By introducing depolarizing noise and amplitude damping into our simulation, we model how the rhythmic output would degrade on a NISQ (Noisy Intermediate-Scale Quantum) device (Sarkar 2023). The findings suggest that the QUBO-QAOA framework is remarkably noise-resilient; even with a 1% gate error rate, the system still identifies the correct Sam and Khali with over 90% accuracy. This resilience stems from the energy gap in our Hamiltonian: the penalty for missing a Sam is so large that it remains the global minimum even under significant perturbations, a property that makes quantum-inspired optimization far more robust for world music preservation than sensitive neural weights (Chawla et al. 2026).

4.3. Data Acquisition, Pre-processing and Rhythmic Quantization

The validation of the framework relies on two primary repositories of Indian classical music: the SwarGanga database and the CompMusic dataset. These datasets provide a cross-section of tabla performances from the Farrukhabad, Ajrada, and Benares gharanas. From these sources, we extracted 5,000 unique rhythmic cycles. Unlike Western datasets that emphasize pitch and harmony, our extraction focused exclusively on the event-density and timbral-sequence of the bols. Each audio file was processed using an onset-detection algorithm to align the strokes with a 16-slot grid, effectively symbolizing the raw audio for the optimization task (Gulati et al. 2014; Sinha 2008).

For the classical models, this data is used for supervised learning, requiring a split of 80% training, 10% validation, and 10% testing. For the quantum framework, the data serves radically different purpose. It is used to populate the T matrix (transition costs) and the $J_{i,j}$ interaction terms through a frequentist analysis of bol successions (Sarkar 2023). For example, if Dha is followed by Dhin 40% of the time in the dataset, the corresponding $J_{i,j}$ value is lowered proportionally. This means the quantum framework is not learning the data in traditional sense; it is using the data to calibrate its compass before navigating the search space (Miranda et al. 2022; Miranda 2021).

The pre-processing also involves a Bol Vocabulary Reduction phase. While there are dozens of subtle table strokes, these can be categorized into right fundamental classes: Dha, Dhin, Ta, Tin, Te, Re, Ke, Te. This categorization is not just a simplification for the AI; it reflects the phonetic reduction used by tabla master to explain complex compositions to students (Gajjar and Patel 2017; Nooshin and Widdess 2006). By reducing the vocabulary, we keep the number of qubits per beat at $M = 8$, allowing us to represent the 16-beat cycle with 128 qubits (16×8). This alignment of musicological theory and quantum bit-depth is a key contribution of our methodology (Miranda and Basak 2022).

Lastly, the Tonic Identification issue is addressed. In Hindustani music, the relative pitch of the drum is tuned to the performer's tonic. The pre-processing pipeline includes a tonic-normalization step, ensuring that the spectral characteristics of the bols are consistent across the entire dataset (Gulati et al. 2014). This prevents the classical models from being confused by different tunings, allowing the benchmark to focus purely on rhythmic sequence and transition logic.

4.4. Quantitative Metrics for Generative Success and Statistical Significance

To move beyond subjective aesthetic judgment, five rigorous quantitative metrics are employed to compare the three architectures (Table 4). The first and most vital is Constraint Satisfaction Rate (CSR). This is a binary metric; a cycle either satisfies all hard constraints (Uniqueness, Sam and Khali) or it does not. In the professional tabla

tradition, a cycle that 95% correct but misses the Sam is considered a failure (Mathur et al. 2015). Our CSR calculation is performed over 1,000 generated samples per model to ensure statistical significance, with p-values calculated to confirm the superiority of the QAOA results.

Table 4: Comparative Performance Metrics across Model Architectures

Metric	WaveNet	LSTM-RNN	Proposed QAOA (p=3)
CSR (%)	68.2	75.4	97.8
Parameters	15.7 million	0.8 million	12 (variational)
Inference Latency (seconds)	76-600	10-150	2-180
Entropy (H_s)	1.24	1.56	2.18
Training Time (hours)	24+	8+	Zero-Shot

The second metric, Structural Entropy (H_s), measures the creativity or diversity of the model. The Shannon entropy of the distribution of generated bols at each beat is calculated. A model with zero entropy would generate the exact same cycle each time, while a model with high entropy would be chaotic. The goal is to find a Goldilocks zone where the entropy is high enough to show variation (Upaj) but the CSR remains high (Bao and Sun 2022; Putz and Svozil 2022). It is found that the QAOA naturally settles into this zone because its energy landscape contains multiple valid local minima, whereas the LSTM tends to collapse into a single high probability mean sequence.

Parameter Efficiency Ratio (PER), Inference Latency (τ) and Training Time, address the resource-intelligence trade-off. PER is defined as $CSR/\log_{10}(Parameters)$, quantifying how much intelligence the model provides per unit of complexity. The QAOA achieves a PER that is three orders of magnitude higher than WaveNet, demonstrating that bigger is not better for rule-constrained musical tasks. Inference latency and training time are measured in wall-clock time on identical hardware. While quantum simulators are currently slower than GPU forward-passes for small batches, the QAOA's ability to generate the entire cycle in one shot means its latency does not scale as poorly as autoregressive models for long-form compositions like Deepchandi (14 beats) or Dhamar (14 beats) (Sarkar 2023). Similarly, QAOA does not require a training time (zero-shot) which shows its ability to generate cycles based on rule.

4.5. Ablation Studies and Hyper-parameter Sensitivity

To ensure the robustness of our methodology, an extensive removal studies was conducted. We systematically removed the Sam constraint, the transition penalties, and the Khali inhibition to see which part of the Hamiltonian contributed most to rhythmic validity. We found that removing the Sam constraint caused the CSR to drop by 45%, confirming that the starting beat is indeed the anchor of the entire mathematical system (Gajjar and Patel 2017; Lipiczky 1985). Conversely, removing the transition penalties led to valid but unmusical sequences, where the bols were correct in isolation but the flow was disjointed and unplayable by a human.

The sensitivity of the QAOA to the classical optimizer choice was also investigated. Comparing COBYLA, Nelder-Mead, and SPSA, it was found that COBYLA provided the most consistent convergence for the rhythmic energy landscape. The sensitivity to initial parameter seeds (γ_0, β_0) (Farhi 2014; Allen et al. 2022) was also mapped; surprisingly, the system is relatively indifferent to the starting point, suggesting that the rhythmic Hamiltonian creates a smooth global basin of attraction rather than a fragmented, glassy landscape. This stability is a significant advantage for real-time applications where there is no time for hyper-parameter tuning during a performance.

Finally, we analyzed the effect of circuit depth (p) on the musicality of the output. While $p = 1$ is sufficient to satisfy basic uniqueness constraints, $p = 3$ or higher is required to capture the subtle swing of the bol transitions. However, beyond $p = 10$, we observed diminishing returns in CSR, coupled with an exponential increase in simulation time. This suggests that for 16-beat tabla cycles, the Quantum Complexity (Putz and Svozil 2017; Oshiro 2022) is relatively modest, making this task an ideal candidate for early-stage quantum hardware implementation. These findings collectively provide a comprehensive roadmap for the next generation of quantum-enhanced computational musicology.

5. Energy Landscape Analysis and Computational Complexity

5.1. Topology of the Rhythmic Search Space: High-Dimensional Manifolds

The fundamental efficacy of the QAOA in generating rhythmic sequences is best understood through a topological lens. In a classical search space, a 16-beat cycle with an 8-bol vocabulary creates a discrete manifold of 2^{128}

configurations (Arya et al. 2022). However, the subset of musically valid states is not just sparse; it is localized into disconnected basins of attraction separated by high-energy barriers of invalid configurations (Sarkar 2023). We analyze this landscape by projecting the 128-dimensional qubit space onto a lower-dimensional manifold using t-distributed Stochastic Neighbor Embedding (t-SNE) and Principle Component Analysis (PCA). The resulting geometry reveals a rugged terrain characterized by deep, narrow valleys of validity (representing canonical taals) surrounded by a vast plateau of noise (representing non-musical bitstrings).

For a deep learning model like an LSTM, navigating this plateau requires the gradient-descent optimizer to stumble upon a valley and slowly adjust weights to descend into it (Tony and Sasikumar 2022). If the initial weights are poorly seeded, the classical model often converges to a local minimum – a sequence that might look rhythmic locally but fails the global Sam or Khali constraints (Bao and Sun 2022). In contrast, the QUBO formulation creates a well-defined global energy sink. Because the penalty for violating a hard constraint (like having two bols on one beat) is set to an order of magnitude higher than any stylistic preference, the energy landscape is effectively engineered to guide any optimization process toward feasibility (Allen et al. 2022; Chawla et al. 2026). This topological engineering is a hallmark of quantum intelligence, where we move from observing a data distribution to defining a physical law (Putz and Svozil 2017; Miranda 2021).

We further investigate the connectivity of these valid valleys. In musicology, two rhythmic cycles are considered close if they differ by only one or two strokes (e.g., a Theka and its Prakar) (Nooshin and Widdess 2006; Sinha 2008). In our energy landscape, these cycles appear as adjacent minima separated by low energy ridges. However, moving from a 16-beat Teentaal to a 10-beat Jhaptaal requires crossing a massive energy mountain. The QAOA's ability to find these specific, disconnected pockets of validity without exhaustive search is a primary indicator of its efficiency. By formalizing the taal as a state-space architecture in Hilbert space, we provide a mathematical proof that the curse of dimensionality in music can be mitigated through the clever design of the cost Hamiltonian.

5.2. Quantum Tunnelling and the Mixing Operator Dynamics

The Quantum Advantage in this study is largely attributed to the mixing operator $U(H_M, \beta)$, which facilitates transitions that are classically impossible. In a simulated annealing or Markov Chain Monte Carlo (MCMC) approach (Arya et al. 2022), moving from one rhythmic sequence to another requires a series of bit-flips (changing one bol at a time). If the path between two valid sequences requires passing through a sequence that is highly invalid (high energy), the classical search will likely stall (Farhi 2014). Quantum tunnelling, however, allows the state vector to tunnel through these high-energy barriers, exploring the grammatical valleys of the landscape in parallel (Allen et al. 2022).

We model the dynamics of the mixing operator as a driver of improvisational flux. By adjusting the mixing phase β, we control the kinetic energy of the system. A high β value induces rapid tunnelling between different rhythmic variations, which is the mathematical equivalent of Upaj (improvisation) (Nooshin and Widdess 2006). We observe that as the circuit depth p increases, the system's ability to resolve the ground state (the most perfect Theka) improves, but its exploratory capacity also becomes more refined.

To quantify this tunnelling effect, we calculate the Success Probability $P = \left|\langle \varphi_{ground} | \varphi_p \rangle\right|^2$ as a function of the mixing strength. Our results show that the QAOA reaches a 90% success probability with significantly fewer iterations (circuit layers) than the number of epochs required by a classical LSTM to reach 70% accuracy. This suggests that the Quantum Mixing provides a more direct path to musical validity than Neural Weight Adjustment. This is because the mixing operator inherently respects the superposition of all states, allowing the system to feel the global structure of the taal before collapsing into a final sequence (Miranda et al. 2022; Miranda and Basak 2022).

5.3. Computational Complexity and Qubit Resource Scaling

The complexity of the rhythmic generator is defined as a function of N (beats) and M (vocabulary size). The number of required qubits scales as $\mathcal{O}(NM)$, while the number of interaction terms (edges in the QUBO graph) scales as $\mathcal{O}(N \cdot M^2)$ for pairwise transitions (Padi et al. 2018). For a 16-beat tabla cycle, this results in 128 qubits and approximately 1,024 interaction terms. This linear-to-quadratic scaling is significantly more favourable than the $\mathcal{O}(P^2)$ scaling of dense neural networks (where P is the number of parameters, often in the millions) (Bao and Sun 2022).

Table 5: Comparative resource scaling for various Hindustani rhythmic structures

Rhythmic Complexity	Beats (N)	Qubits (QAOA)	Parameters (LSTM)	State Space Size
Dadra	6	48	~0.4 million	2.8×10^{14}
Jhaptaal	10	80	~0.6 million	1.2×10^{24}
Teentaal	16	128	~0.8 million	3.4×10^{38}
Dhamar	14	112	~0.7 million	5.1×10^{33}

As shown in Table 5, the state space grows exponentially, but the quantum resource requirement (qubits) grows only linearly. This is the scaling advantage of quantum-inspired models. While a classical model must increase its parameter count to memorize the increasingly complex transitions (Allen et al. 2022) of longer taals, the QAOA simply requires a slightly larger Hamiltonian and a few more qubits. Furthermore, the Temporal Locality of the tabla tradition – where a beat is primarily influenced by its neighbours – means the QUBO graph is sparse. This sparsity allows for efficient embedding on near-term quantum hardware like the IBM Eagle or Osprey processors, which feature heavy-hex or square-lattice connectivity (Miranda and Shaji 2023).

The Time-to-Solution (TTS) metric is analyzed. In classical AI, inference time τ for an autoregressive model is $\tau = \mathcal{O}(N \cdot K)$, where K is the forward-pass time per beat (Bao and Sun 2022). For the QAOA, the generation is global: $\tau = \mathcal{O}(p)$, where p is the circuit depth. On actual quantum hardware, $p = 3$ would translate to microseconds of execution time, regardless of whether the taal is 6 or 16 beats. This global inference capability is the key to real-time, interactive quantum music systems that can keep pace with a human performer during high-speed Drut (fast tempo) compositions.

5.4. Sensitivity Analysis: The Criticality of Penalty Coefficients (λ)

The stability of the energy landscape depends entirely on the balancing of the penalty coefficients λ. In our framework, λ_{unique} (hard constraint) and λ_{anchor} (structural constraint) must be tuned to prevent constraint bleeding, where the system sacrifices a hard rule to satisfy a stylistic preference. We conducted a sensitivity analysis by sweeping these coefficients from 1.0 to 100.0. We found that the Constraint Satisfaction Rate (CSR) exhibits a phase transition at $\lambda \approx 10.0$. Below this value, the system produces chaotic, invalid sequences; above it, the system enters a frozen valid state.

The Aesthetic Sensitivity (γ), the weight of the transition matrix T, is also examined. If γ is too high, the system becomes repetitive, only choosing the most probable transitions (e.g., repeating Dha-Dha-Dha). If γ is too low, the system becomes experimental, choosing valid but unmusical transitions that a human player would find jarring. The optimal Aesthetic Equilibrium was found at a ratio of 1:5:10 for $\gamma : \lambda_{anchor} : \lambda_{unique}$. This ratio ensures that the physicality of the beat is most important, followed by the structural anchor, and finally the stylistic flow (Putz and Svozil 2022).

This sensitivity analysis provides a foundational recipe for other researchers in the field of quantum musicology. It suggests that when mapping any world music tradition onto a QUBO framework, one must prioritize the axioms (the rules that cannot be broken) before the aesthetics (the rules that can be bent) (Bao and Sun 2022). This hierarchical approach to penalty weighting is what allows the QAOA to maintain a 97.8% CSR while still exhibiting $H_s = 2.18$ structural entropy (creativity). It effectively creates a safe playground for the AI to be creative without being wrong.

5.5. Barren Plateau and the Expressibility of Musical Circuits

A common challenge in Variational Quantum Algorithms (VQAs) is the Barren Plateau problem, where the gradient of the cost function vanishes as the number of qubits increases, making the classical optimizer unable to find a path to the minimum (Arya et al. 2022). In our 128-qubit simulation, we monitored the gradient variance throughout the optimization process. We observed that because our Hamiltonian is local (interactions only between nearby beats), the barren plateau effect is significantly mitigated. This locality-induced trainability (Gajjar and Patel 2017) is a vital discovery for scaling quantum-inspired models to even longer musical forms, such as the 108-beat taals found in some ancient traditions.

We also measured the Expressibility of our QAOA circuit – the extent to which the circuit can explore the entire Hilbert space. While a fully expressive circuit is often desirable for general-purpose AI, for music, we want a constrained expressibility. We want a circuit that is very good at exploring valid musical states but is incapable of exploring the noise (Miranda and Basak 2022). Our analysis shows that the cost operator $U(H_C, \gamma)$ effectively acts

as a filter that prunes the Hilbert space, focusing the expressive power of the mixing operator only on the valleys of the energy landscape.

This Filtered Expressibility is the theoretical basis for why our model uses so few parameters. While a WaveNet model has millions of parameters because it must be expressive enough to represent any audio signal (from white noise to bird calls to tabla), our QAOA only needs to be expressive enough to represent tabla rhythms. This alignment of Circuit expressibility with Domain logic is the goal of Quantum Intelligence. By using a Hamiltonian that reflects the physics and culture of the drum, we create a system that is born knowing the rules, requiring only a few parameters to fine-tune the performance.

6. Discussion, Artistic Implications, and Cultural Formulization

6.1. Bridging Ethnomusicology and Quantum Intelligence

The transition from classical deep learning to quantum-inspired optimization represents more than a shift in computational efficiency; it marks a fundamental change in how we formalize oral traditions. For centuries, the tabla tradition of Hindustani music has been preserved through the Guru-Shishya Parampara (Sinha 2008), an oral-aural system where complex rhythmic laws are passed down through mnemonics and immersion. By mapping these laws onto a cost Hamiltonian, we are effectively creating a digital ontological transcript of an oral tradition. This process forces a rigorous mathematical definition of concepts like Lay (tempo/flow) and Chhand (meter) that were previously described only in qualitative terms (Gajjar and Patel 2017).

In our framework, the intelligence of the system is not derived from observing thousands of hours of audio, but from the structural alignment of the algorithm with the musical culture. This culture-native AI is particularly vital for preserving the nuances of specific Gharanas (schools of playing) (Sinha 2008). For instance, the Delhi Gharana emphasizes crisp, finger-tip strokes and linear patterns, while the Benares Gharana utilizes more resonant, palm-based strokes (Lipiczky 1985). By adjusting the transition costs in the $J_{i,j}$ matrix of our QUBO model, we can tune the AI to follow the specific aesthetic logic of a particular school, providing a tool for ethnomusicologists to study the mathematical divergence between different lineages.

Furthermore, the quantum-inspired approach addresses the data scarcity problem in non-Western musicology. While Western pop music has millions of labeled tracks for training Big Data models, classical traditions often have limited high-quality digital archives. Because our QAOA-based framework is zero-shot – relying on rule-encoding rather than pattern-matching – it can generate high-fidelity, grammatically correct music even for rare taals where only a handful of recordings exist. This democratization of AI allows for the technological representation of diverse world music that were previously left behind by the data-heavy requirements of standard neural networks.

6.2. The Ethics of Algorithmic Creativity and Upaj

A central debate in AI musicology is whether an algorithm can truly exhibit Upaj (improvisation) or if it is merely shuffling pre-existing patterns (Nooshin and Widdess 2006). Our analysis of the energy landscape suggests that the QAOA does not shuffle; it navigates a logical manifold. When the system tunnels from a canonical Theka to a novel variation, it is identifying a state that is mathematically valid according to the Hamiltonian but hasn't been explicitly programmed (Arya et al. 2022). This mirrors the human creative process, where a musician explores new phrases within the guardrails of traditional grammar (Miranda and Basak 2022).

However, the use of quantum agents in traditional arts raises ethical questions regarding authorship and the sanctity of the tradition. If a quantum system can generate a Tihaai that is mathematically more complex than what a human can conceive, does it still belong to the tradition? We argue that the QUBO-QAOA framework should be viewed as a Co-Creative Assistant. By providing the musician with a Creativity Slider (η), the technology serves as a mirror that reflects the mathematical possibilities latent within the Taal. This allows the artist to focus on the Ras (emotional essence) (Mathur et al. 2015) while the AI ensures the Taal (rhythmic structure) remains flawless.

We also address the concern of cultural flattening (Bao and Sun 2022). There is a risk that by formalizing music into a binary optimization problem, we may lose the human errors that provide character to a performance. To mitigate this, we propose the integration of Stochastic Jitter (Miranda and Shaji 2023) in the final state collapse, ensuring that the timing of the bols is not mathematically perfect but humanly resonant.

6.3. Real-Time Interactivity and the Future of Live Performance

One of the most exciting artistic implications of this research is the potential for real-time quantum accompaniment. Current deep learning models, due to their high inference latency and autoregressive nature, often struggle to keep pace with the rapid tempo changes of a live Lehara or Sitar performance (Gulati et al. 2014). Because the QAOA generates the entire rhythmic cycle globally in a single optimization pass, its theoretical inference time on actual quantum hardware would be near-instantaneous.

A future system could utilize a Hybrid Quantum-Classical Listener (Miranda and Shaji 2023) that tracks the performer's tonic and tempo, dynamically updates the cost Hamiltonian, and triggers the quantum generator to produce a responding bol sequence. This would create a dialogue between the human and the machine, where the AI is not just playing a loop but is actively responding to the creative queries of the performer. This vision of Live Indian Quantum Music represents the ultimate application of the mathematical frameworks established in this paper, turning abstract Hilbert space into a performance space.

The scalability of our model also allows for the generation of Long-Form Rhythmic Compositions such as Kayda or Rela. These forms involve the development of a single theme through various permutations. By using a Chained-QUBO approach (Sinha 2008), where the output of one cycle informs the starting state of the next, we can generate minutes of cohesive, evolving rhythmic patterns that follow the traditional theme-and-variation structure of the tabla repertoire. This capacity for long-term logical coherence is something that even the most advanced Transformers currently struggle with in the musical domain.

6.4. Convergence with QNLP and Broader Impacts on Quantum Education

The structural mapping of taals shares significant commonality with the emerging field of Quantum Natural Language Processing (QNLP) (Miranda et al. 2022). Just as words in a sentence must satisfy syntactic dependencies, tabla bols must satisfy rhythmic dependencies. We discuss the Music-as-Language paradigm. By treating a vibhag (measure) as a noun phrase and the Sam as a verb, we can apply DisCoCat (Distributional Compositional Categorical) (Miranda et al. 2022) models to rhythmic generation. This allows for a Semantic Optimization of music, where the AI generates beats that make sense in the context of a larger musical narrative.

The use of Diagrammatic Quantum Circuits for music, as proposed by researchers at Quantinuum and Oxford, aligns perfectly with our QUBO results (Allen et al. 2022). We demonstrate that the meaning of a rhythmic phrase can be represented as a quantum state, and the composition of phrases can be represented as the tensor product of these states. This formalization provides a unified framework for both language and music, suggesting that the Intelligence in Quantum Intelligence is a universal structural property of symbolic systems.

Beyond the technical and artistic domains, this research serves as a powerful tool for quantum education. Quantum mechanics is often perceived as abstract and counter intuitive. However, by using music – a universal human experience – as a medium for explaining concepts like superposition, entanglement, and tunnelling, we can make the field more accessible to the public (Putz and Svozil 2017). A student can hear a superposition of beats and experience the collapse of the wave function into a valid theka (Putz and Svozil 2022). This Sonic Quantum Literacy could inspire a new generation of researchers from the Global South to enter the field of quantum computing (Chawla et al. 2026). By demonstrating that their cultural heritage is not just compatible with, but beneficial for, the development of quantum algorithms, we foster a more inclusive global scientific community.

6.5. Limitations and the Hardware Gap

While the theoretical and simulated results are compelling, we must address the Hardware Gap between current NISQ devices and the 128-qubit requirement of our model. Although 128 qubits exist on modern processors, the All-to-All connectivity required for some complex transition rules is not yet available, leading to high swap-gate overhead and decoherence (Oshiro 2022). Consequently, our current results rely on Quantum-Inspired classical simulations or small-scale proof-of-concepts on 16-27 qubit machines.

We also acknowledge that our current Hamiltonian is time-invariant, meaning the rules do not change as the performance progresses. In actual Hindustani music, the rules for Vilambit (slow) tempo are different from Drut (fast) tempo (Sinha 2008). A more advanced model would require a Dynamic QUBO where the coefficients λ and $J_{i,j}$ are themselves functions of time or tempo. This would allow the AI to morph its playing style as the performance accelerates – a hallmark of professional musicianship that is currently missing from our static optimization landscape (Putz and Svozil 2022).

Finally, the Black-Box nature of the classical optimizer (like COBYLA) used to tune the quantum parameters remains a point of investigation. While we have achieved high CSR, the interpretability of why specific γ and β values lead to specific artistic choices is still limited. Future research should focus on White-Box Variational

Circuits, where the relationship between the variational angles and the musical mood is explicitly mapped. This would allow musicians to communicate with the quantum AI in terms like play more aggressively or be more subtle, which the AI would then translate into specific Hilbert-space rotations.

7. Conclusion, Strategic Roadmap, and Future Outlook

7.1. Summary of the Quantum-Inspired Generative Paradigm

This research has demonstrated that the formalization of Hindustani rhythmic structures into a Quadratic Unconstrained Binary Optimization (QUBO) framework provides a superior pathway for generating grammatically precise musical sequences compared to high-parameter classical architectures. By mapping the tabla bols to a 128-qubit state space and encoding the rhythmic axioms of Teentaal, Jhaptaal, and Ektaal as penalty terms in a cost Hamiltonian, we have moved the generative process from a domain of statistical estimation to one of global logical optimization. The primary contribution of this work is the verification that for low-duration, highly structured symbolic problems, a Constraint-Native approach outperforms the Data-Heavy paradigms that have dominated the last decade of AI research (Miranda 2021).

Our comparative benchmarking revealed a dramatic Efficiency Gap. While deep learning models like WaveNet and LSTM-RNNs require millions of parameters and thousands of training examples to approximate the concept of a Sam (the first beat), the QAOA-based framework achieves a near-perfect Constraint Satisfaction Rate (CSR) of 97.8% using only 12 variational parameters. This efficiency is not merely a technical curiosity; it represents a fundamental alignment between the mathematical structure of quantum optimization and the hierarchical logic of Indian classical music (Chawla et al. 2026). We have shown that when the physics of the algorithm reflects the culture of the art, the need for massive datasets evaporates, enabling a Zero-Shot generative capacity that is vital for the digital preservation of diverse musical traditions.

Furthermore, the analysis of the energy landscape provided a theoretical explanation for this performance. By utilizing the mixing operator $U(H_M, \beta)$, the quantum agent can tunnel through the vast plateaus of invalid rhythmic noise, landing directly in the sparse valleys of musical validity (Allen et al. 2022). This capability allows the system to maintain high structural entropy – facilitating genuine Upaj (improvisation) – without sacrificing the rigid integrity of the taal. This balance of creativity and validity, governed by the tuning of the penalty coefficients λ, offers a level of interpretability and control that remains elusive in the black-box latent spaces of classical GANs or Transformers.

7.2. The Strategic Roadmap for Hardware Integration

As we transition from the simulation phase to physical quantum hardware, the roadmap is defined by the gradual scaling of qubit count and the optimization of gate fidelity. While current NISQ (Noisy Intermediate-Scale Quantum) devices are sufficient for 16-beat cycles, the generation of full-scale Bandishes or complex Chakradars will require hundreds, if not thousands, of high-fidelity qubits with improved connectivity (Miranda and Shaji 2023). The next phase of this research involves the implementation of our QUBO models on superconducting processors like IBM's Osprey or trapped-ion systems (Sarkar 2023) that offer all-to-all connectivity, which would significantly reduce the swap-gate overhead required for long-range rhythmic dependencies.

In parallel, the development of Hybrid Quantum-Classical Orchestrators will be essential for real-time applications. Such a system would utilize a classical Listener to track the human performer's BPM and tonic, and a quantum Composer to generate a response within the millisecond windows required for live interaction (Gulati et al. 2014). This requires the refinement of Warm-Starting techniques for QAOA, where the classical optimizer is initialized with a state that is already close to a valid rhythmic variation, further reducing the number of circuit evaluations needed during a performance.

We also foresee the integration of Error Mitigation strategies specifically tailored for musical Hamiltonians. Since a musical error on a primary anchor (like the Sam) is more catastrophic than an error on a secondary beat, we can design Priority-Aware Decoders that focus quantum error correction resources on the most structurally significant qubits (Padi et al. 2018). This hardware-software co-design ensures that even in the presence of noise, the mathematical backbone of the rhythm remains intact. This approach will be the cornerstone of Reliable Quantum Intelligence, moving the field beyond experimental novelties into the realm of professional-grade artistic tools.

7.3. Expanding the Modal and Microtonal Horizon

While this study focused on the rhythmic dimension of Hindustani music, the future of this framework lies in its expansion to the melodic and modal complexities of Raag. The microtonal intricacies (shruti) and the glissandi

(meend) that define a raag can be represented as continuous variables in a Variational Quantum Eigensolver (VQE) framework (Miranda and Shaji 2023). By mapping the pitch-class distributions of the SwarGanga database onto a quantum state space, we can generate melodic phrases that respect both the ascent-descent (aaroh-avaroh) rules and the emotional vadi-samvadi hierarchy of the mode.

The goal is a Holistic Quantum Music Agent capable of performing both the Taal and the Raag simultaneously. This would involve a dual-Hamiltonian system where the rhythmic constraints and the melodic constraints are coupled through Inter-Track Constraints (Miranda et al. 2022). For example, the system would ensure that a melodic resolution (a Nyas) coincides with the rhythmic resolution on the Sam. Such a system would represent the most complex application of quantum optimization in the humanities to date, challenging our understanding of both computational complexity and artistic genius.

Furthermore, this framework can be generalized to other world music traditions that share a similar rule-based structure, such as the Maqam of Middle Eastern music or the Gamelan cycles of Indonesia (Gajjar and Patel 2017). By providing a Meta-Hamiltonian template that can be customized for different cultural axioms, we facilitate a global Quantum Ethnomusicology. This trajectory points toward a future where AI does not replace the musician but serves as a universal translator of musical logic, allowing different traditions to interact within a shared mathematical Hilbert space.

7.4. Ethical Considerations and the Preservation of Heritage

As we embed ancient cultural logic into quantum circuits, we must remain vigilant regarding the ethical implications of Algorithmic Appropriation. The tabla tradition is not just a collection of numbers; it is a lived heritage with deep spiritual and social significance (Nooshin and Widdess 2006). There is a risk that by creating a perfect quantum tabla player, we may inadvertently devalue the years of dedication required for human mastery. It is imperative that these tools are developed in close collaboration with traditional practitioners, ensuring that the technology supports the Guru-Shishya model rather than bypassing it.

We propose a Sovereignty-First approach to musical data. The cost Hamiltonians derived from specific Gharanas should be treated as the intellectual and cultural property of those lineages (Sinha 2008). Quantum intelligence should be used to provide these communities with tools for Digital Revitalization – allowing a rare or dying style of playing to be reconstructed from fragmentary historical records. This Restorative AI represents a more ethical use of technology than the extractive Big Data scraping currently practiced by many large-scale AI companies.

Moreover, we must address the Accessibility Gap. To ensure that quantum-inspired musicology is not limited to elite research institutions, we must develop Cloud-Based Creative Portals where musicians in any part of the world can access quantum simulators via a simple interface (Sarkar 2023). By abstracting the complex physics of Pauli-Z operators and mixing phases into musical terms like tension, resolution, and swing, we can empower artists to become Quantum Creators without requiring a formal education in physics.

7.5. Final Concluding Remarks

In conclusion, the intersection of Quantum Intelligence and Hindustani Classical Music offers a profound new vista for both fields. For the quantum scientist, music provides a rigorous Natural Laboratory for testing the limits of variational algorithms in high-dimensional, highly constrained discrete spaces. For the musicologist, quantum computing provides a mathematical language capable of expressing the Superposition of Traditions and the Entanglement of Rhythm and Emotion.

This study has established a baseline for this new discipline, proving that the QAOA-QUBO framework is not just a viable alternative to classical AI, but a necessary one for the future of structured creative intelligence. As we look toward the 2030s, the Quantum Tabla serves as a symbol of a broader synthesis – where the oldest traditions of human expression meet the most advanced frontiers of human logic. The rhythmic heartbeat of the Sam, now encoded in the spins of a quantum processor, signifies that the quest for mathematical harmony is a universal endeavour, spanning across time, culture, and the very fabric of physical reality.

Declarations

No funding was received to assist with the preparation of this manuscript. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Clinical trial number: not applicable.

References

Allen E J, Bulmer J F, Small S D (2022) Making Music using Two Quantum Algorithms. Quantum Computer Music 69-82.

Arya A, Botelho L, Cañete F, Kapadia D, Salehi Ö (2022) Applications of Quantum Annealing to Music Theory. In Quantum Computer Music 373-406.

Bao C, Sun Q (2022) Generating music with emotions. IEEE Transactions on Multimedia 1-14.

Chawla S, Kumari U, Rai D (2026) A Quantum-Inspired State-Space Architecture for Microtonal Modal Systems: Formalizing Indian Classical Music in Hilbert Space. ResearchGate preprint.

Farhi E, Goldstone J, Gutmann S (2014) A Quantum Approximate Optimization Algorithm. arXive preprint.

Gajjar K, Patel M (2017) Computational Musicology for Raga Analysis in Indian Classical Music: A Critical Review. International Journal of Computer Applications 172 9: 42-47.

Gulati S, Bellur A, Salamon J, Ranjani H G, Ishwar V, Murthy H A, Serra X (2014) Automatic Tonic Identification in Indian Art Music: Approaches and Evaluation. Journal of New Music Research 43 1:53-71.

Lipiczky T (1985) Tihai Formulas and the Fusion of Composition and Improvisation in North Indian Music. The Musical Quarterly 71 2:157-171.

M. T. Group Computational models for the discovery of the World's Music. Universitat Pompeu Fabra. https://compmusic.upf.edu/datasets. Accessed 15 April 2024.

Marinescu A I (2019) Batch 2.0 - generating classical music using recurrent neural networks. 23rd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems 159:117-124.

Mathur A, Vijayakumar S H, Chakrabarti B, Singh N C (2015) Emotional responses to Hindustani raga music: the role of musical structure. Frontiers in Psychology 6 513.

Miranda E R (2021) Quantum Computer: Hello, Music! In Handbook of Artificial Intelligence for Music 963-994.

Miranda E R, Basak S (2022) Quantum Computer Music: Foundations and Initial Experiments. Quantum Computer Music 43-67.

Miranda E R, Shaji H (2023) Generative Music with Partitioned Quantum Cellular Automata. Applied Sciences 13 4:2401.

Miranda E R, Yeung R, Pearson A, Meichanetzidis K, Coecke B (2022) A Quantum Natural Language Processing Approach to Musical Intelligence. Quantum Computer Music 313-356.

Nooshin L, Widdess R (2006) Improvisation in Iranian and Indian Music. Journal of the Indian Musicological Society 36 37:104-119.

Oord A V, Dieleman S, Zen H et al (2016) WaveNet: A Generative Model for Raw Audio. arXiv preprint.

Oshiro S (2022) QuiKo: A Quantum Beat Generation Application. Quantum Computer Music 131-158.

Padi S, Breiner S, Subrahmanian E, Sriram R D (2018) Modeling and Analysis of Indian Carnatic Music Theory. IEEE Transactions on Systems, Man and Cybernetics: Systems 48 6:967-981.

Putz V, Svozil K (2017) Quantum music. In Soft Computing 21:1467–1471.

Putz V, Svozil K (2022) Quantum Music, Quantum Arts and Their Perception. Quantum Computing in the Arts and Humanities 179-191.

S M Foundation. SwarGanga Database. https://www.swarganga.org/databases.php Accessed 15 April 2024.

Sarkar S (2023) A Generative Modeling Approach using Quantum Gates. arXive preprint.

Sinha P (2008) Artificial Composition: An Experiment on Indian Music. Journal of New Music Research 37 3: 221-232.

Tony S M, Sasikumar S (2022) Investigation on Automatic Music Generation Using Gan and Lstm Networks. Research Square