Methods — Compatibility Graph Construction

Segment-Level Feature Representation

Each beat-aligned audio segment is represented by a fixed-dimensional feature vector derived from its short-time Fourier transform (STFT). All audio is processed at a fixed sampling rate of 22,050 Hz, and identical STFT parameters are used across all segments to ensure comparability and reproducibility.

Given the complex STFT $X(f,t)$ of a segment, we compute the power spectrogram:

$$P(f,t) = |X(f,t)|^2. \tag{1}$$

A mel filterbank with 40 mel bands is applied to the power spectrogram to obtain a mel-spectrogram $M(m,t)$. The final feature vector $\mathbf{v} \in \mathbb{R}^{40}$ is obtained by temporal averaging:

$$v_m = \frac{1}{T} \sum_{t=1}^{T} M(m,t), \tag{2}$$

where T denotes the number of STFT frames in the segment.

All feature vectors are non-negative and ℓ_2-normalized prior to similarity computation. This feature definition is frozen for all experiments.

Feature Normalization and Degeneracy Detection

Feature vectors are normalized to unit ℓ_2 norm. Vectors with near-zero norm are explicitly rejected to prevent degeneracy. Feature norm statistics are validated prior to normalization to ensure numerical stability.

Compatibility Scoring

Segment compatibility is defined using cosine similarity between ℓ_2-normalized feature vectors.

To discourage degenerate mashups that remain entirely within a single source track, a multiplicative penalty is applied when two segments originate from the same parent song:

$$\mathrm{score}(A,B) = 0.7 \times \cos(\mathbf{v}_A, \mathbf{v}_B). \tag{3}$$

This penalty discourages trivial self-transitions while preserving musically continuous transitions.

Graph Construction and Symmetrization

A weighted k-nearest neighbor (kNN) graph is constructed using similarity scores, with $k = 7$.

Since kNN relations are not necessarily symmetric, the raw adjacency matrix A_{raw} is symmetrized as:

$$A_{ij} = \max(A_{\mathrm{raw},ij}, A_{\mathrm{raw},ji}). \tag{4}$$

This ensures a symmetric adjacency matrix, allowing the graph to be interpreted as a Hermitian Hamiltonian with real eigenvalues in subsequent quantum walk simulations.

Classical Baseline Path

A deterministic greedy walk is used as a classical baseline. Starting from a fixed segment, the next segment is selected as the highest-weight unused neighbor.

Immediate repetition is disallowed. The walk terminates when no valid neighbors remain or when a fixed path length is reached.

This baseline provides a reference for evaluating quantum-inspired transition dynamics.