

Data-driven Classification of Harmonic Progression Types in Popular Music using Clustering and Transition Network Analysis

Sungbum Kim ^{1,*}

¹ Kumoh National Institute of Technology; Professor; it89@kumoh.ac.kr

* Correspondence

<https://doi.org/10.5392/IJoC.2026.22.3.087>

Manuscript Received 21 May 2026; Received 25 June 2026; Accepted 6 July 2026

Abstract: Harmonic progressions are central to the popular music listening experience, yet corpus-based, data-driven classification of song-level harmonic progression patterns remains underexplored. This study analyzes 178 songs from the Billboard Hot 100 year-end charts for 2023–2025 using chord charts collected from Ultimate Guitar. The chord charts were converted into key-neutral chord-root sequences through Roman numeral analysis, and six indices were developed to capture harmonic diversity, repetition, concentration, and entropy-based information content. Using these standardized indices, K-means clustering with silhouette analysis identified a four-cluster solution with interpretable harmonic types and descriptively distinct mean profiles. A Genre × Type chi-square analysis revealed no significant association between broad genre categories and harmonic type. Chord transitions within each type were then modeled as directed networks, and high-frequency transition patterns were interpreted within a music-theoretical framework. The resulting networks exhibited type-specific structures consistent with distinct harmonic grammars. By integrating quantitative clustering with a network-based interpretation of complete songs, this study presents an empirical methodology that translates data-driven harmonic patterns into musically meaningful insights into contemporary popular music harmony.

Keywords: Music Information Retrieval; Popular Music Harmony; Chord Progression; Roman Numeral Analysis; K-means Clustering; Network Analysis

1. Introduction

Chord progressions are fundamental in shaping the musical character of individual songs in popular music. Contemporary popular music production is characterized by hook-centered structures, repetitive arrangements [1, 2], and the convergence of diverse genres [3, 4]. Within this context, patterns of harmonic repetition and the timing of harmonic change have been identified as important factors in shaping the listening experience [2], [5–7]. Empirical work on music perception has linked harmonic predictability and surprise to listener responses across multiple domains: probabilistic expectancy models predict tension, surprise, and emotional engagement from chord-to-chord transitions [6, 7], and repetition has been shown to enhance attention, perceived familiarity, and aesthetic appeal even in the absence of semantic understanding [2]. These findings collectively suggest that song-level differences in harmonic dynamism—the very property targeted by the present study—are not merely descriptive properties of the score but plausibly correspond to distinct listener experiences, even when other musical parameters (timbre, tempo, lyrics) are held constant. Even among songs sharing identical chord vocabularies, the frequency of harmonic change and the degree of repetition can differ substantially; some songs convey a static impression through the recurrence of a fixed pattern, while others produce a sense of dynamic harmonic development. Despite the perceptual importance of these song-level harmonic dynamics, large-scale empirical research on popular music harmony has rarely treated them as a unified analytical object. Many corpus studies have prioritized chord-vocabulary statistics or local transition patterns over song-level harmonic trajectories [8, 9]. Section-specific analyses, such as studies of chorus loops, can also underrepresent transitional

events that occur across full song forms [10]. Consequently, the song-level harmonic trajectory—which unfolds across the formal sections of a song (e.g., intro, verse, pre-chorus, chorus, bridge, and outro)—has remained underexplored as a coherent analytical unit, leaving open the question of whether complete songs in popular music can be systematically classified by their overall harmonic behavior. The present study addresses this gap by treating the full chord sequence of each song as the primary unit of analysis, with the goal of classifying song-level harmonic progression patterns in popular music into distinct, music-theoretically interpretable types. To enable consistent comparison across songs in different keys, this study employs Roman numeral analysis (RNA)—a key-neutral representation of each chord root's scale-degree position relative to the inferred tonic—rather than relying on absolute chord labels [9], [11]. For instance, the progressions C–F–G in C major and G–C–D in G major are both represented as the standardized progression I–IV–V. By applying RNA, the distribution of harmonic patterns and chord-to-chord transitions can be compared across songs in different keys within a unified analytical framework.

Building on this framework, this study addresses four research questions, each specified with the analytical rationale and procedure appropriate to it. First, prior corpus research on popular music harmony has documented non-uniform distributions of chord vocabulary, repetition, and transition behavior across songs [8, 9], suggesting that song-level harmonic profiles organize into a small number of coherent groups rather than a continuous distribution; RQ1 asks whether songs separate into multiple structurally distinct clusters interpretable as archetypal harmonic-development types (descriptive—evaluated through K-means cluster cohesion and silhouette analysis). Second, RQ2a asks whether the derived types exhibit distinct mean profiles across the six indices, particularly in harmonic diversity and repetition (descriptive by construction, since the indices used to evaluate the profiles are the same indices used to derive the clusters). Building on accounts of genre-specific stylistic conventions in harmonic practice [5], [8], RQ2b asks whether the distribution of types varies systematically across genres, a question addressed through an independent Genre $\times$ Type chi-square test (inferential). Third, building on theoretical accounts that treat chord transitions as conditional, syntax-like grammars [12, 13], RQ3 asks whether the most frequent chord transition patterns within each type are qualitatively distinguishable across types, reflecting music-theoretically coherent differences in harmonic syntax (evaluated through network analysis, a music-theoretical characterization independent of the index profiles).

To address these research questions, this study proceeds through four analytical stages. First, the chord progressions of complete songs collected from Ultimate Guitar are converted into RNA sequences, establishing a standardized representation that enables cross-song comparison. Second, six indices reflecting diversity, repetitiveness, rate of change, and information content are defined to operationalize dimensions of harmonic structure that have been studied in corpus-based and information-theoretic research on popular music [6], [9]. Third, an unsupervised clustering approach is applied to these indices to derive representative cluster types that capture the harmonic development patterns of popular music [12]. Fourth, the statistical association between the derived types and musical genres is examined, and frequently occurring chord transition patterns within each type are extracted and interpreted from a music-theoretical perspective [13]; network-based analysis is employed to enhance the interpretability of these quantitative findings.

In departing from the conventional section-centered mode of analysis, this study makes three distinct contributions to the study of popular music harmony. First, it extends existing corpus research—often confined to specific genres or small-scale samples [9], [14]—by standardizing complete-song data at scale into RNA-based sequences [11], thereby establishing a foundation for systematic comparison of songs differing in tonality and style within a unified analytical framework. Second, it operationalizes musically meaningful attributes—such as harmonic complexity and repetitiveness—as a unified set of six quantitative indices with explicit formulas, enabling harmonic development to be measured with greater precision and analytical rigor. Third, it interprets the derived clusters through the lens of music-theoretic harmonic grammar [9], [12], grounding the statistical typology in established analytical concepts rather than presenting it as a purely numerical classification. By combining data-driven quantitative analysis with music-theoretical interpretation, this study offers a framework that meaningfully connects large-scale empirical findings with the musical context of popular song.

2. Literature review

2.1 Harmonic schemata and structural characteristics in popular music

Harmonic progressions in popular music are not random; rather, they are organized around shared conventional frameworks—schemata—common to both creators and audiences [5], [9], [15]. Tagg (2014)

argues that many forms of contemporary popular music are built on the repeated cycling of short chord progressions (e.g., four-bar loops). Listeners derive a sense of harmonic stability and momentum from these repetitive patterns, which function as distinctive harmonic grammars characteristic of popular music [5].

Furthermore, popular music frequently employs distinctive chord patterns that depart from classical harmonic rules in order to emphasize particular expressive qualities. Even chords that deviate from tonal convention function within popular music as established harmonic idioms, serving to evoke specific moods or propel dramatic development [15]. In this way, popular music harmony is governed by a small number of dominant schemata and repetitive conventions. Accordingly, analyzing chord succession patterns in large-scale datasets provides an empirical means of examining musical conventions that might otherwise remain abstract [9].

2.2 Corpus-based research on popular music harmony

In contrast to many strands of traditional musicological scholarship, which have often emphasized score-based interpretation and the particularity of individual works [16], contemporary music research has evolved toward empirically grounded inquiry by incorporating cognitive psychology and statistical methods applied to large-scale corpora [7], [17]. Within the broader current of empirical musicology, harmonic research in popular music has particularly focused on statistically characterizing the distribution of harmonies, syntactic rules, and recurring patterns [8]. Within this paradigm, the field of computational musicology has made sustained efforts to construct high-quality datasets in order to ensure the reliability of analytical results.

The RWC database was developed as a standardized collection of research audio and annotations, thereby laying the groundwork for systematic comparative analysis in music information retrieval research [18]. Subsequently, the McGill Billboard Corpus provides an expert-annotated alternative to noisier web-sourced data because trained musicians transcribed its harmonic annotations directly from audio recordings [14]. Although such expert-curated corpus construction has contributed significantly to improving data accuracy, the manual nature of the process imposes structural limitations on scalability.

In response, recent research in Music Information Retrieval (MIR) has expanded its scope to large-scale web-based chord-sharing platforms such as Ultimate Guitar, accepting a degree of annotation noise in exchange for scale and ecological validity [19, 20]. The aim is to move beyond a select body of canonical works and to survey, from a macroscopic perspective, the chord progressions of tens of thousands of widely consumed songs, thereby deriving generalized patterns [20].

2.3 Standardization of popular music harmonic data

Comparing songs in different keys requires first encoding each chord's functional role within its key in a unified format, rather than relying on absolute chord names. Prior studies have noted that absolute chord names such as C–G–Am–F are encoded as different chord labels when transposed to other keys, thereby impeding the identification of statistical patterns [9], [11]. As an alternative, Roman numeral analysis (RNA)—a system that encodes each chord relative to the tonic—has become a standard analytical method for cross-key comparison. Converting divergent chord progressions into key-neutral data enables direct cross-song comparison [9], [11].

RNA represents these functional relationships as standardized symbols relative to a specified tonic, rendering harmonies amenable to computational processing [11]. Building on this methodological rationale, the present study converts the collected chord data into RNA sequences and treats them as the primary unit of analysis.

2.4 Probabilistic harmonic modeling and data-driven clustering

To quantitatively analyze harmonic progressions in popular music, scholars have introduced Markov models, which treat chord connections as stochastic processes. Under this framework, a harmonic progression is understood not as a mere sequence of chords but as a chain of conditional probabilities, in which each chord influences the likelihood of the next. Research has demonstrated that transition matrices extracted from large-scale corpora enable the mathematical modeling of harmonic grammar preferred by particular genres or styles [8], [13].

Beyond this, unsupervised learning approaches that classify songs based on similarity in harmonic transition patterns have also been actively pursued in computational musicology. K-means clustering and other unsupervised clustering algorithms are particularly effective for identifying groups of songs sharing similar harmonic progressions based on data-driven structural patterns, without relying on predefined genre labels [21, 22]. Prior studies have validated cluster solutions using statistical measures such as the Silhouette Coefficient [23] and have reported that the resulting clusters may reflect musically meaningful harmonic practices or style types [22].

2.5 Quantitative measures of harmonic complexity and repetition

A growing body of research has applied information-theoretic measures to musical analysis in order to quantify the structural properties of harmonic progressions [6], [13]. Most notably, Shannon entropy is widely employed as an index for measuring the unpredictability of harmonic distributions [13], [24]. Large-scale corpus studies have likewise documented quantitative trends in popular music's harmonic, timbral, and loudness properties [25], demonstrating that computational measures can reveal systematic patterns invisible to traditional analysis. Additionally, attempts to quantify repetition—a defining aesthetic principle of popular music—have attracted considerable attention. From a cognitive-psychological perspective, Margulis (2013) argues that repetition is a core mechanism sustaining listener engagement. Building on this account, the present study operationalizes repetition through measures such as unique chord-type count and the recurrence rate of specific loop patterns [2]. In this way, quantitative indices such as entropy and repetition rate translate the intuitive perceptual difference between "monotony" and "change" into concrete numerical values, enabling systematic comparison.

3. Research Methodology

3.1 Data

3.1.1 Data source and scale

Chord charts were collected from Ultimate Guitar (UG; <https://www.ultimate-guitar.com>), a large community-driven online platform where users contribute and rate chord annotations, tablatures, and related song-level information for popular music. For each song in the Billboard Hot 100 corpus, the corresponding chord chart was identified and collected from UG when available. When multiple chord sheets were available for the same song, the sheet with the highest aggregated user rating was selected. The collected chord charts were then preprocessed to extract chord sequences from song sections, remove non-chord annotations, standardize chord symbols, estimate keys, and convert chord roots into Roman numerals. This procedure produced a structured harmonic dataset suitable for cross-key comparison and transition analysis.

UG was selected as the primary data source for two reasons. First, chord annotation in popular music is inherently subjective even among trained experts [19], and community-vetted annotations may better reflect the chord interpretations actually used in popular music practice [19], [26]. Second, the platform's scale enables corpus analysis far beyond what expert-annotated datasets typically allow. The sample was drawn from the top 100 songs on each of the Billboard Hot 100 year-end charts for 2023, 2024, and 2025, yielding an initial pool of 300 chart entries. When multiple chord sheets were available for the same song on UG, the highest-rated sheet was selected.

The data-curation process proceeded in three stages. First, entries with usable UG chord annotations were retained, yielding 234 entries. Second, duplicate songs appearing across multiple year-end charts were consolidated, yielding 189 unique songs. Third, songs for which the chorus section could not be reliably identified or annotated (Chorus = 0) were excluded. This criterion was applied because the chorus commonly functions as a central and recurrent formal reference point in chart-oriented popular music [27]. By contrast, Intro and Outro were not used as exclusion criteria because they are auxiliary framing sections rather than obligatory core sections [27]. Verse = 0 and Bridge = 0 cases were not automatically excluded when an identifiable chorus and a usable song-level chord sequence were available. These cases were treated as reflecting possible section-label ambiguity in pop/rock form analysis, rather than as automatically indicating incomplete harmonic data [28]. The final analytical sample comprised 178 songs. Table 1 summarizes this data-curation process, and Table 2 provides a description of the resulting dataset. The collected chord charts were subsequently processed into Roman-numeral sequences, as described in Section 3.1.2.

Table 1. Data curation flow

Stage	Criterion	N retained	Rationale
0	Initial pool: top 100 songs × 3 year-end charts (2023, 2024, 2025)	300	Sampling frame defined by Billboard Hot 100 year-end charts
1	Usable UG chord annotation available	234	Excludes entries lacking a usable chord sheet on UG
2	Duplicate songs across year-end charts consolidated	189	Each unique song counted once even if charted in multiple years
3	Identifiable chorus section (Chorus ≠ 0)	178	Final analytical sample; chorus is the central recurrent reference point [27]

Table 2. Data description

Item	Description
Corpus	Billboard Hot 100 songs
Period	2023–2025
Final analytical sample	178 songs
Chord data source	Ultimate Guitar chord charts
Unit of analysis	Song
Main harmonic representation	Roman numeral chord-root sequence

3.1.2 Data construction and refinement process

Data construction and refinement were carried out in three stages. First, song metadata—including year, chart rank, song title, artist, and broad genre classification—were compiled and integrated into the analytical dataset. Second, chord progressions were extracted from UG chord sheets and arranged into continuous sequences covering the full length of each song. Third, chords were standardized by converting them into RNA notation based on the key of each song, enabling consistent comparison of functional progressions across different keys. Roman numeral sequences were constructed through a four-stage pipeline, as summarized in Table 3.

Table 3. Four-step RNA conversion pipeline

Step.	Input	Procedure	Output
1	UG chord sheet PDF	Extract text and isolate chord lines using regex heuristics	Chord-line sequence
2	Chord tokens	Remove slash-bass and quality suffixes	Chord-root sequence
3	Chord-root sequence	Score 24 candidate keys using weighted diatonic matching; if capo information is available, additionally transpose to estimate sounding key (used for mode classification only)	Inferred notated key; sounding key metadata (if applicable)
4	Chord root + notated key	Map pitch-class interval to Roman numeral	RN sequence

Note. Sounding key (Step 3, where applicable) is used exclusively for mode classification in Section 4.4.2 and does not affect the Roman-numeral output of Step 4

First, PDF text was extracted using the Python package *pdfplumber*, and section labels such as Intro, Verse, and Chorus were identified from header markers; chord lines were then isolated using regex-based heuristics designed to distinguish chord rows from lyrics, tablature, and metadata. Second, chord tokens were preprocessed to extract their root: slash-chord bass information was removed (e.g., G/B was reduced to G). Because subsequent analysis used only chord roots, quality-suffix variations (e.g., "7," "maj7," "sus4") were not retained in the Roman numeral representation. Third, the key of each song was inferred using a weighted scoring scheme that evaluated 24 candidate keys, comprising 12 major and 12 minor candidates [17], [29]. Chord roots were scored against each candidate's diatonic scale, with positive points for in-scale roots and penalties for out-of-scale roots, and an additional emphasis for tonic-matching roots. Chord roots in chorus and outro sections, as well as those near the song's end, received elevated weights to reflect their tonal salience. The candidate with the highest score was selected as the song's key.

When capo information was available in the chord chart, this notated key was additionally transposed to estimate the song's sounding key, which was used for mode classification in Section 4.4.2 but not for Roman-numeral conversion itself; because Roman-numeral labels depend only on the relative pitch-class interval between chord root and tonic, performing the conversion on the notated key and un-transposed chord roots yields the same relative representation as transposing both to the sounding pitch level.

Fourth, each chord root was mapped to a Roman numeral according to its semitone interval from the inferred tonic. Specifically, the algorithm calculated the pitch-class interval between each chord root and the inferred tonic using a 12-semitone system and mapped the resulting interval onto a Roman-numeral category, ranging from 0 (I) through 11 (VII), with chromatic alterations notated as *b*II, *b*III, *b*VI, and *b*VII [9], [11]. For example, C–F–G in C major and Em–C–G–D in E minor are encoded as I–IV–V and I–*b*VI–*b*III–*b*VII,

respectively, illustrating how songs in different keys and modes are unified within a single Roman-numeral framework.

To focus the analysis on root-interval relationships across songs rather than on chord quality, bass motion, or inversion, the present study adopted a simplified Roman numeral representation in which chord tokens are represented primarily by their root function relative to the inferred tonic. In the final analytical dataset, major/minor triad quality was not systematically encoded by Roman-numeral capitalization [30]. This convention prioritizes the syntactic positions of chord roots over voicing and quality details, consistent with prior corpus analyses that focus on root-progression patterns in popular music [4], [9]. Diminished chords were retained as an exception when explicitly marked in the source chord sheet, because they may function as leading-tone or passing sonorities that are not adequately captured by root-function coding alone. The complete Roman numeral (RN) sequence of each song served as the primary unit of analysis.

Because the conversion was rule-based rather than based on independent human coders, inter-rater reliability statistics such as Cohen's kappa were not applicable to the primary conversion process. Instead, the ambiguity of key estimation was assessed algorithmically by examining the score difference between the highest-scoring and second-highest-scoring key candidates. A larger difference indicates that the selected key was clearly preferred over competing alternatives, whereas a smaller difference indicates greater ambiguity. We therefore treat this score difference as a diagnostic indicator of key-estimation confidence rather than as an inter-rater reliability statistic. This diagnostic does not constitute an independent accuracy validation of the conversion output, but it provides an algorithmic indicator of key-estimation ambiguity. Nevertheless, we acknowledge that future studies could further validate the conversion pipeline using expert-coded benchmark datasets.

3.2 Variable definition

To quantify the harmonic characteristics of popular music, this study established six analytical variables grounded in three complementary lines of corpus-based and information-theoretic research on popular-music harmony. For diversity, two variables were adopted: UniqueRN — the number of distinct chord types in a song — follows the chord-vocabulary measures used in [8, 9]; UniqueBigram — the number of distinct chord-transition patterns — extends the same logic from chord tokens to transition tokens and parallels the transition-vocabulary measure used in [22]. For structural repetition, two variables were introduced: Top1Share — the proportion of the song accounted for by the single most frequent chord — captures the degree to which one chord dominates the song [8]; RepeatRate — the rate of consecutive chord repetition — follows the operationalization of harmonic repetition used in corpus analyses of popular and rock music [4], [25], [31].

For complexity, information-theoretic principles were applied through two variables: H_{RN} , which captures the entropy of the RN distribution [24], and H_{Bigram} , which captures the entropy of RN transition pairs, reflecting the overall diversity and unpredictability of the progression pattern [13]. H_{RN} (Variable 5) measures the uniformity of chord distribution in a song. A lower value signifies concentration on a specific chord (high predictability), whereas a higher value denotes an even (uniform) distribution across chords (low predictability). H_{Bigram} (Variable 6) reflects how predictable or formulaic the chord transitions of a given song are: higher values indicate greater unpredictability and transition complexity, while lower values indicate more formulaic or repetitive progressions.

These indices are grounded in the information-dynamics framework of music perception, which posits that listeners continuously form probabilistic expectations about upcoming musical events based on prior musical context [6]. Within this framework, measures of entropy (H_{RN} , H_{Bigram}) correspond to the degree of listener uncertainty or surprise, while measures of repetition (RepeatRate, Top1Share) reflect the extent to which harmonic events conform to established expectation patterns. This theoretical grounding ensures that the quantitative indices employed in the present study capture perceptually meaningful dimensions of harmonic experience rather than purely descriptive statistical properties. Intercorrelations among the six indices were examined prior to clustering; the full correlation matrix and the rationale for retaining all six indices rather than reducing their dimensionality are presented in Section 5.

Table 4. Six harmonic progression indices: formula and descriptions

No.	Variable Name	Formula	Description
1	UniqueRN	$ \{r\} $	The number of distinct RNs used in a song, representing the breadth of harmonic materials employed.
2	UniqueBigram	$ \{(r_t, r_{t+1})\} $	The number of distinct RN transitions occurring in a song, representing the diversity of progression patterns.
3	Top1Share	$\frac{\max_r(\text{count}(r))}{N}$	The proportion of the most frequently used RN relative to the total, indicating the degree to which a single chord or function dominates the song.
4	RepeatRate	$\frac{\sum_{t=1}^{N-1} 1(r_t = r_{t+1})}{N - 1}$	The proportion of instances in which the same chord is repeated consecutively.
5	H_RN	$-\sum_r p(r) \ln p(r)$	The Shannon entropy of the RN occurrence probability distribution, measuring the degree of unpredictability or uncertainty of chord functions.
6	H_Bigram	$-\sum_{(a,b)} p(a \rightarrow b) \ln p(a \rightarrow b)$	The Shannon entropy of RN transition-pair probabilities, measuring the unpredictability or complexity of the progression grammar.

Several alternative measures were considered and excluded from the final six-index set. Chord-quality entropy (treating major, minor, and seventh qualities as distinct tokens) and voicing-derived features (e.g., bass-note motion) were not retained because the present pipeline collapses each chord to its root function (Section 3.1.2); incorporating such quality- and voicing-sensitive measures would have required reliable automatic extraction of these harmonic details, which the present study intentionally omitted for reasons of robustness and reproducibility [11], and would have re-introduced dependence on the noisy quality suffixes filtered out during preprocessing. A candidate key-stability index (e.g., the proportion of out-of-key chords, conceptually related to tonal-hierarchy-based measures of key clarity [29]) was also considered, but was not retained because, by construction, a higher proportion of out-of-key chords corresponds to a flatter, more even chord distribution and is therefore expected to closely track H_RN, adding little independent interpretive contrast. Section-level measures (e.g., chorus-versus-verse harmonic divergence) were excluded because the present analytical unit is the complete song (Section 3.1.2). The six retained indices were instead organized around the three dimensions introduced above — diversity, structural repetition, and complexity — with each dimension represented by a token-level measure and a corresponding transition-level measure (UniqueRN/UniqueBigram, Top1Share/RepeatRate, and H_RN/H_Bigram, respectively); although the two measures within each pair are correlated, each retains a distinct music-theoretical interpretation (e.g., UniqueRN captures chord-vocabulary size, whereas H_RN captures the evenness of its distribution).

3.3 Analytical methods

3.3.1 Type derivation: Index-based clustering

To systematically classify harmonic types in popular music, this study applied K-means clustering—an unsupervised learning algorithm—using the six feature vectors defined in Section 3.2 as input [32]. Because each song was represented as a fixed-length, song-level feature vector, K-means was selected as the primary clustering method. The purpose of the clustering procedure was not to model chord progressions as temporal sequences, but to derive an interpretable song-level typology from aggregate harmonic indices.

Prior to clustering, each of the six indices was standardized to zero mean and unit variance, ensuring that no single index's raw measurement scale disproportionately influenced the Euclidean distance metric underlying the algorithm. K-means is methodologically aligned with this representation because it defines clusters around centroids in the same standardized feature space. Therefore, the resulting clusters can be directly interpreted through their mean feature profiles and subsequently linked to type-specific transition-network analysis.

K-means partitions a dataset into k clusters by minimizing the distance from each data point to its assigned centroid. To determine the optimal number of clusters (k), silhouette analysis was used. The Silhouette Coefficient is an index that simultaneously evaluates the cohesion of data points within a cluster and their separation from other clusters [23]. This study computed the mean Silhouette Coefficient across a range of candidate k values and selected the value of k at which it was maximized as the final number of cluster types.

To examine whether this analytical choice unduly constrained the findings, we also conducted a sensitivity check using alternative clustering methods applicable to the same standardized six-index feature space: Ward hierarchical clustering, spectral clustering, and Gaussian mixture models [33–35]. These methods were used not to redefine the main typology, but to assess whether they provided a clearly stronger or more interpretable alternative to the standardized K-means solution.

HMM-GMM was not implemented because it addresses a different modeling problem. HMM-GMM is designed for sequential generative modeling and typically requires time-indexed continuous observations, such as chord embeddings or duration-weighted harmonic-state vectors. By contrast, the present study clusters songs using six aggregate song-level harmonic indices. Applying HMM-GMM would therefore require a different chord-level sequential representation and a different research design.

Following the derivation of clusters, the high-frequency chord transitions characterizing each cluster were extracted and qualitatively interpreted with reference to established harmonic conventions in popular music [9], [30], [36], in order to characterize the music-theoretical profile of each cluster type.

3.3.2 Genre association and descriptive profiling of the derived types

To examine the association between genre and type distribution, a Genre $\times$ Type cross-tabulation was constructed, and a Pearson chi-square test was performed. In cases where expected cell frequencies fall below 5, the chi-square approximation may not be reliable [37]; the analysis was therefore supplemented a priori by a Monte Carlo chi-square test ($B = 100,000$) and a likelihood-ratio G-test [38]. The strength of the association was reported using Cramér's V as an effect-size measure [39].

The Monte Carlo chi-square test estimates the null distribution of the chi-square statistic by repeatedly generating random contingency tables that preserve the observed row and column totals, and computing the proportion of simulated tables yielding a statistic as extreme as the observed one. This procedure produces a valid p -value without relying on the asymptotic chi-square approximation and is therefore appropriate when expected frequencies are small [40]. The likelihood-ratio G-test uses the same null hypothesis of independence but bases its test statistic on the ratio of likelihoods between the observed and expected frequencies ($G = 2 \cdot \sum O_{ij} \cdot \ln(O_{ij}/E_{ij})$); it is asymptotically equivalent to the Pearson chi-square but is reported here as an additional convergent check. Cramér's V quantifies the strength of association on a 0–1 scale and is interpreted using conventional benchmarks (≈ 0.1 small, ≈ 0.3 medium, ≈ 0.5 large) [41]. In addition, a two-way ANOVA was performed for each of the six variables with Type and Genre as factors to test the main effects and the Type $\times$ Genre interaction.

3.3.3 Network analysis

This study treated the RNA chord sequence of each song as sequential data, defined each pair of consecutive chords as a transition, and modeled the data as a directed, weighted network [42, 43]. Network nodes represent individual RNA chords (I, IV, V, VI, bVII, etc.), while edges represent transitions in which chord B follows chord A ($A \rightarrow B$) [42], [44]. Network analysis was introduced to enable structural comparison of harmonic syntax across types, capturing not only chord occurrence frequencies but also chord-to-chord sequential dependencies. Since network centrality measures the structural importance of individual nodes, it was employed to identify the pivotal chords within the progressions of each type [42]. In particular, weighted degree centrality was applied to compute the sum of incoming transition frequencies (in-weight) and the sum of outgoing transition frequencies (out-weight) for each chord [45]. For visualization, an explicit cutoff rule was applied—retaining only the top six nodes by centrality and their representative transitions—to highlight the core transition structure and improve interpretability.

4. Results

4.1 Type derivation results based on six indices

K-means cluster analysis was conducted using the six feature indices defined in Section 3.2. All 178 songs were valid for clustering. A comparison of the four- to eight-cluster solutions showed that $k = 4$ produced the highest mean silhouette coefficient ($SC = 0.3232$), with $k = 5$ ($SC = 0.3214$), $k = 6$ ($SC = 0.2887$), $k = 7$ ($SC = 0.2718$), and $k = 8$ ($SC = 0.2786$). Accordingly, $k = 4$ was adopted as the final solution. Although a mean silhouette coefficient of 0.3232 falls within the range commonly interpreted as indicating weak structure [46], this guideline should be treated as a heuristic benchmark rather than a strict decision rule [47]. The four-cluster solution was retained because it produced the highest silhouette coefficient within the candidate range, yielded interpretable harmonic profiles, and was further supported by the domain-informed validation reported in Section 4.4.

Table 5. Mean profile by type (averages of six variables) and number of songs

Type	n	Unique RN	Unique Bigram	Top1 Share	Repeat Rate	H_RN	H_Bigram
1	72	9.347	15.639	0.296	0.090	1.891	2.447
2	47	5.553	9.426	0.395	0.162	1.383	1.868
3	47	13.362	23.383	0.217	0.071	2.287	2.928
4	12	4.083	6.500	0.692	0.617	0.794	1.249

A sensitivity analysis was conducted to examine how the clustering solution changed when alternative algorithms were applied to the same standardized six-index feature space. The alternative methods were selected to represent three distinct clustering principles: Ward hierarchical clustering as a hierarchical variance-minimizing approach, spectral clustering as a graph-based affinity approach, and Gaussian mixture modeling as a probabilistic model-based approach. The reported standardized K-means solution showed a silhouette coefficient of 0.323. Re-estimated K-means produced a nearly identical silhouette coefficient of 0.322, indicating that the centroid-based solution was reproducible within the same standardized feature space.

Among the alternative methods, Ward hierarchical clustering yielded a slightly higher silhouette coefficient of 0.328. However, this improvement over the reported K-means solution was marginal. Spectral clustering produced a lower silhouette coefficient of 0.295, and Gaussian mixture models also yielded lower values of 0.283. Thus, although alternative clustering methods produced somewhat different classifications, none provided a clearly stronger replacement for the standardized K-means solution in terms of internal separation and interpretability. The comparison results are summarized in Table 6. This analysis was conducted not to demonstrate the universal superiority of K-means over other clustering methods, but to assess the stability and interpretability of the standardized K-means solution as a sensitivity check.

Table 6. Sensitivity analysis of alternative clustering methods

Method	Methodological principle	Feature space	Silhouette coefficient	Interpretation
Reported standardized K-means	Centroid-based clustering	Standardized six-index feature space	0.323	Main analytical typology
Re-estimated standardized K-means	Centroid-based clustering re-run on the same feature space	Standardized six-index feature space	0.322	Within-method reproducibility check
Ward hierarchical clustering	Hierarchical clustering minimizing within-cluster variance increase	Standardized six-index feature space	0.328	Slightly higher SC, but only marginally
Spectral clustering	Graph-based clustering using affinity structure among songs	Standardized six-index feature space	0.295	Lower internal separation
Gaussian mixture model	Probabilistic model-based clustering	Standardized six-index feature space	0.283	Lower internal separation

Analysis of index characteristics across the four clusters revealed a clear structural polarization in harmonic dynamism. The six analytical indices—capturing harmonic vocabulary size, transition diversity, concentration on a dominant chord, consecutive repetition, and entropy—differentiated the four types into a dynamic-complex pole, a static-loop pole, and two intermediate types occupying distinct positions between them.

Type 1—Dynamic-Moderate ($n = 72$) is the largest cluster, comprising 40.4% of the total sample. UniqueRN (9.347), UniqueBigram (15.639), H_RN (1.891), and H_Bigram (2.447) occupied the upper-middle range, while RepeatRate (0.090) was among the lowest of all types. This profile represents the normative harmonic grammar of contemporary popular music: harmonically active and moving through a moderately broad chord vocabulary without reaching the maximal complexity of Type 3.

Type 2—Simple-Dynamic ($n = 47$) recorded a low UniqueRN (5.553) and UniqueBigram (9.426), and H_Bigram (1.868) was substantially reduced. RepeatRate (0.162) was higher than in Type 1 but still moderate, and Top1Share (0.395) indicated a tendency toward concentrated chord usage. This combination indicates that chord changes occur but the harmonic vocabulary remains limited—a small set of chords is cycled through, consistent with the four-chord-loop convention prevalent in popular music.

Type 3—Dynamic-Complex ($n = 47$) recorded the highest values across all diversity and complexity indicators: UniqueRN (13.362), UniqueBigram (23.383), H_RN (2.287), and H_Bigram (2.928). RepeatRate (0.071) was simultaneously among the lowest, and Top1Share (0.217) was the lowest of all types, confirming that harmonic change is both frequent and diverse with no single dominant chord.

Type 4—Static-Loop ($n = 12$) recorded the highest values of Top1Share (0.692) and RepeatRate (0.617) among all types, while H_RN (0.794) and H_Bigram (1.249) were the lowest. UniqueRN (4.083) and UniqueBigram (6.500) also reached their minimum values across all types. This profile indicates that a small number of chords, most prominently the tonic, govern the entire song, with the same chord recurring consecutively across extended passages.

To visualize the structural relationships among the four types, Figure 1 plots each cluster along two axes. Following the coordinate procedure described in the next paragraph, the horizontal axis represents harmonic change frequency (operationalized through reverse-coded RepeatRate and Top1Share), and the vertical axis represents harmonic vocabulary diversity (operationalized through UniqueRN, UniqueBigram, H_RN, and H_Bigram). Circle size is proportional to the number of songs in each cluster.

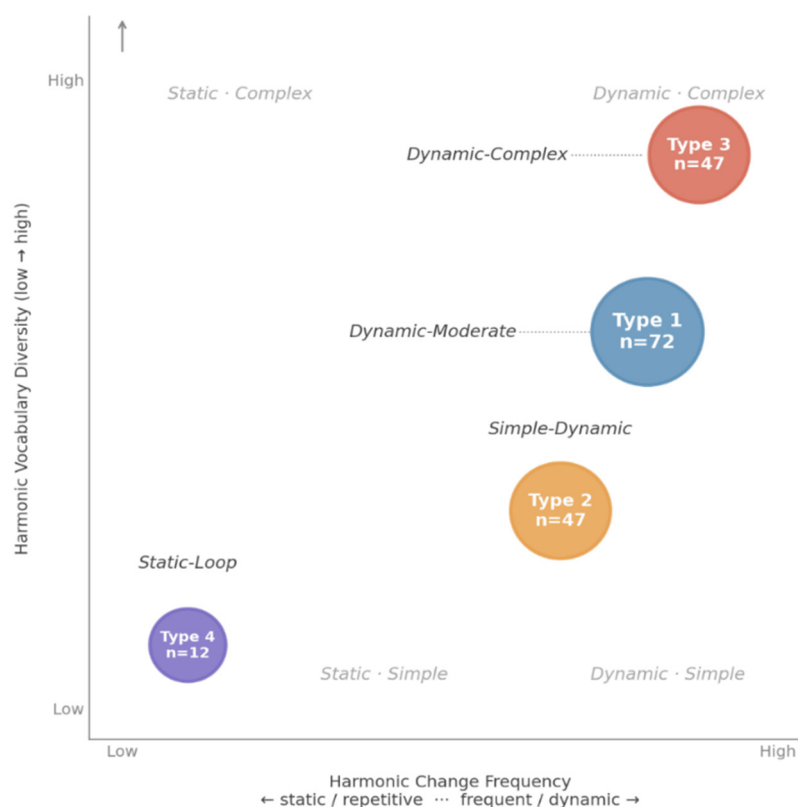


Figure 1. Two-dimensional classification of harmonic progression types

The positioning map translates the cluster-level mean profiles into two interpretable dimensions. Before plotting, all six profile indicators were rescaled using min-max normalization across the four types so that variables measured on different scales could be compared within a common range. The horizontal axis captures the degree of harmonic movement. Because both RepeatRate and Top1Share indicate structural recurrence or concentration, they were reverse-coded prior to aggregation; thus, lower repetition and weaker dependence on a dominant chord correspond to a higher position on the dynamic end of the axis. The vertical axis represents the breadth and dispersion of harmonic vocabulary. This dimension was derived from the normalized values of UniqueRN, UniqueBigram, H_RN, and H_Bigram, which together reflect the number of distinct harmonic units, the variety of chord-to-chord transitions, and the evenness of their distributions. Types located higher on this axis therefore exhibit greater harmonic variety and more dispersed progression patterns.

Based on these two dimensions, each cluster was assigned an interpretive label. Type 4, located in the lower-left region, was labeled Static-Loop because it combines low harmonic diversity with high recurrence. Type 2 was labeled Simple-Dynamic, since it shows frequent harmonic change but relies on a relatively limited harmonic vocabulary. Type 1 was labeled Dynamic-Moderate, indicating high movement with a moderate-to-high level of diversity, and represents the largest cluster in the sample ($n = 72$; 40.4% of all songs). Finally, Type 3 was labeled Dynamic-Complex, as it combines the highest harmonic vocabulary diversity with a strongly dynamic progression profile.

Table 7. Characteristic summary of four harmonic progression types

Type	n	X	Y	Characteristics	Label
1	72	89.9	63.9	Frequent harmonic change with moderate diversity; normative profile of contemporary popular music	Dynamic-Moderate
2	47	72.9	27.4	Frequent change within a constrained chord set; rapid cycling through a limited harmonic vocabulary	Simple-Dynamic
3	47	100.0	100.0	Highest diversity and complexity across all indicators; richest and most unpredictable harmonic grammar	Dynamic-Complex
4	12	0.0	0.0	Dominant self-repetition with minimal chord vocabulary; archetypal loop-stasis grammar	Static-Loop

4.2 Relationship between genre and type

To examine the association between genre and type, a 4×4 cross-tabulation was constructed and a Genre $\times$ Type chi-square test was performed. The Pearson chi-square result was not statistically significant ($\chi^2 = 3.68$, $df = 9$, $p = .931$). Of the 16 cells in the contingency table, seven (43.8%) had expected frequencies below 5, with the lowest being 0.67 (Rock/Indie $\times$ Type 4). Under these conditions, the standard chi-square approximation may not be reliable [37]; accordingly, a Monte Carlo chi-square test with fixed margins ($B = 100,000$) and a likelihood-ratio G-test were conducted as supplementary analyses, both of which yielded results consistent with the Pearson test (Monte Carlo $p = .937$; $G = 4.38$, $p = .885$). The effect size was very small (Cramér's $V = 0.083$), indicating that even with a larger sample the genre–type association would likely remain weak. To further examine whether specific genre pairs differed in harmonic-type distribution, we conducted post-hoc pairwise comparisons among genre categories with Holm correction [48] to control for multiple comparisons. No pairwise comparison was statistically significant after correction, indicating that no specific pair of broad genre categories showed a reliable difference in harmonic-type distribution.

Because the contingency table was sparse and the overall association was clearly non-significant, these pairwise results were treated as supplementary rather than primary evidence. To further assess whether any individual genre–type combination occurred more or less often than expected under a model of no association between genre and type, we inspected the adjusted standardized residuals [49] of the full 4×4 contingency table. These residuals provide a cell-level diagnostic comparing observed and expected frequencies for each genre–type combination. All adjusted residuals fell well within the conventional threshold of $|z| \approx 2$ (maximum $|z| = 1.30$, Country/Folk $\times$ Type 2), consistent with both the corrected pairwise comparisons and the overall null result. Taken together, these convergent results provide no statistically significant evidence that broad genre categories are concentrated in particular types within the present sample.

Table 8. Distribution of genre by type

Genre	Type				Row Total
	1	2	3	4	
Country/Folk	31	15	20	5	71
Pop/Dance	19	15	11	4	49
R&B/Soul/Hip-Hop	17	15	13	3	48
Rock/Indie	5	2	3	0	10
Column total (178)	72	47	47	12	178

4.3 Post clustering profile analysis

Section 4.3 is intended as a descriptive profile of the four harmonic types derived from the clustering procedure. Because the four harmonic types were derived directly from the six analytical indices, inferential tests comparing the types on these same variables involve a degree of circularity. Accordingly, the ANOVA results reported in this section should not be interpreted as independent statistical validation of the cluster solution. Instead, they are used descriptively to summarize the mean profiles of the derived types and to clarify which variables contributed most strongly to their interpretive differentiation. Independent validation of the musical meaningfulness of the types is provided in Section 4.4 through transition-network analysis, which examines chord-to-chord sequential structures not reducible to the aggregate index values alone.

A two-way ANOVA was conducted with Type and Genre as factors. Because the four Types were derived from the six indices through K-means clustering, Type-related F-statistics on these same indices are not independent confirmatory evidence; we therefore restrict the inferential ANOVA reporting to the Genre main effect and the Type $\times$ Genre interaction, which are independent of the clustering procedure, and summarize the Type-level differences descriptively through the cluster mean profiles already reported in Table 5 (Section 4.1).

The Genre main effect was not statistically significant for any of the six indices, and no Type $\times$ Genre interaction reached significance.

These findings suggest that the six indicators primarily capture structural differences in song-level harmonic progression style (Type) rather than differences attributable to genre labels, and indicate that such stylistic patterns are not confined to specific genres but may be observed broadly across genres.

Table 9. Post clustering profile analysis of the six indices

	P_Type	P_Genre	P_Interaction
Unique RN	—	.603	.318
UniqueBigram	—	.968	.483
Top1Share	—	.670	.789
RepeatRate	—	.488	.731
H_RN	—	.155	.317
H_Bigram	—	.596	.401

Note. Type-related inferential p-values are not reported because the four Types were derived from the same six indices through K-means clustering; reporting them would constitute circular evidence. The Type-level profiles are summarized descriptively in Table 5 (Section 4.1). P_Genre and P_Interaction remain interpretable as independent tests of factors not involved in the clustering step.

Because Type was derived via K-means clustering on these same six indices, pairwise differences between Types on these variables are expected by construction and would constitute the same circular evidence noted above; we therefore do not report Type-level pairwise contrasts. Descriptively, Types 1 (Dynamic-Moderate) and 3 (Dynamic-Complex) showed similarly low RepeatRate values (0.090 and 0.071, respectively), but differed substantially on the remaining five indices (UniqueRN, UniqueBigram, Top1Share, H_RN, H_Bigram). This descriptive observation is revisited in Section 4.4 through transition-network analysis.

4.4 Domain-informed validation through network analysis

The index-based analysis presented in Sections 4.1 through 4.3 quantified how much harmonic variety and repetition each cluster type exhibits—that is, the size of the chord vocabulary and the degree to which individual chords recur. However, harmonic progression is inherently a directed, sequential phenomenon: what matters is not only which chords are present, but how they move from one to another and with what probability. Index measures, being aggregate statistics, do not capture the directionality or conditional structure of chord transitions. Section 4.4 therefore complements the index analysis by examining chord transition networks expressed in RNA notation, in which nodes represent chords and weighted directed edges represent the conditional probability of each transition. This approach reveals aspects of harmonic grammar—specifically, the direction, hierarchy, and relative likelihood of chord movements—that remain invisible to index-based quantification alone.

4.4.1 Overview and visualization method

In the chord transition network expressed in RNA notation, weighted in-degree and weighted out-degree were nearly identical at the aggregate level, suggesting that chords broadly operate within a cyclical harmonic framework. Because selecting either weighted in-degree or weighted out-degree yields no substantial interpretive difference—with the top-ranked nodes remaining consistent—this study adopts weighted in-degree as the primary criterion, emphasizing how frequently each chord functions as a harmonic arrival point. Figure 2 presents a simplified RNA transition network for each of the four types. For each type, the six chords with the highest weighted in-degree, defined as the total incoming transition frequency, were selected as nodes. From each selected node, treated as a source, the single most representative outgoing transition was extracted based on the highest conditional probability, $p(\text{dst}|\text{src})$, and coded as a representative edge ($\text{RepEdge} = 1$) after applying a minimum transition-frequency threshold ($\text{min_freq} \geq 3$) to ensure reliability. Node size is proportional to weighted in-degree, and the numerical values above the arrows represent $p(\text{dst}|\text{src})$ (%).

4.4.2 Transition patterns by type

It should be noted that the corpus includes both major-key and minor-key songs, all encoded using the unified root-tonic Roman numeral framework described in Section 3.1.2. Before interpreting the type-specific transition networks, we therefore examined whether the observed network patterns were affected by the distribution of major- and minor-key songs across clusters. This diagnostic check was necessary because flat-side Roman numerals such as $\flat\text{III}$, $\flat\text{VI}$, and $\flat\text{VII}$ may have different theoretical meanings depending on mode [30], [36]: in major-key contexts, these chords often indicate modal mixture or borrowed harmonic color, whereas in natural minor-key contexts they may function as diatonic scale-degree chords.

The major/minor mode of each song was identified during the key-estimation stage. For each song, the algorithm evaluated 24 possible keys, consisting of 12 major and 12 minor candidates, and selected the key with the highest weighted score. When capo information was available in the chord chart, the initially inferred notated key was transposed according to the capo position to estimate the sounding key of the original song. Songs whose estimated sounding key was minor were classified as minor-key songs, whereas all others were classified as major-key songs. Table 10 reports the major/minor distribution within each harmonic type. All four types contained both major- and minor-key songs. Type 1 consisted of 41 major-key songs and 31 minor-key songs, Type 2 consisted of 32 major-key songs and 15 minor-key songs, Type 3 consisted of 25 major-key songs and 22 minor-key songs, and Type 4 consisted of 9 major-key songs and 3 minor-key songs. A Type $\times$ Mode chi-square test indicated that the association between harmonic type and mode was not statistically significant, $\chi^2(3) = 3.60$, $p = .309$, Cramér's $V = .142$. Therefore, the transition-network patterns in Figure 2 are unlikely to be driven primarily by mode imbalance across clusters.

Table 10. Major / minor distribution

Type	Major	Minor	Total
1	41	31	72
2	32	15	47
3	25	22	47
4	9	3	12
Total	107	71	178

As an additional robustness check, Appendix A presents mode-separated transition networks for major- and minor-key songs within each harmonic type. These networks were constructed using the same visualization rule as Figure 2: nodes were selected based on the top six in-weight centrality values, and one representative outgoing transition per source node was selected according to the highest conditional transition probability, with a minimum frequency threshold of three. The mode-separated networks generally preserve the broad type-specific tendencies observed in Figure 2, particularly the tonic-centered organization of the transition structures, while also revealing mode-dependent differences in the occurrence and interpretation of flat-side Roman numerals such as $\flat\text{III}$, $\flat\text{VI}$, and $\flat\text{VII}$. Therefore, the appendix analysis supports the robustness of the main type-specific network patterns while confirming the need for mode-sensitive interpretation. The Type 4 minor-key sub-network should be interpreted with particular caution because it is based on only three songs.

Accordingly, flat-side Roman numerals in the networks should be interpreted in a mode-sensitive manner rather than as functionally identical across all songs. In particular, $\flat\text{VII} \rightarrow \text{I}$ may reflect modal borrowing in major-key songs, whereas it may represent diatonic subtonic-to-tonic motion in minor-key songs. Thus, Type 3 is best characterized not simply by modal borrowing, but by a broader modal/minor-inflected approach to tonic arrival.

In Type 1 (Dynamic-Moderate), the largest cluster ($n = 72$), I was the dominant node with $\text{IV} \rightarrow \text{I}$ (71.4%) and $\text{III} \rightarrow \text{I}$ (46.3%) as the principal incoming transitions, supplemented by $\text{II} \rightarrow \text{III}$ (33.4%), $\text{I} \leftrightarrow \text{V}$ (24.5%), and $\flat\text{VII} \rightarrow \text{IV}$ (22.8%). The network is organized around I as the principal arrival point, but without the self-loop dominance observed in Type 4. This profile reflects a harmonic grammar centered on the $\text{IV} \rightarrow \text{I}$ relationship, supplemented by additional approach pathways (notably $\text{III} \rightarrow \text{I}$ and $\flat\text{VII} \rightarrow \text{IV}$). The distributed conditional probabilities and the multi-step approach paths to I are consistent with the moderate diversity and low repetition observed in the index profile.

In Type 2 (Simple-Dynamic, $n = 47$), I was the largest node, approached primarily by $\text{II} \rightarrow \text{I}$ (56.0%), $\text{IV} \rightarrow \text{I}$ (52.2%), and $\text{VI} \rightarrow \text{I}$ (30.1%), with the outgoing pathway $\text{I} \leftrightarrow \text{V}$ (18.0%). The intermediate node IV was itself approached by $\flat\text{III} \rightarrow \text{IV}$ (20.4%), indicating an additional approach pathway through the $\flat\text{III}$ chord ($\flat\text{III} \rightarrow \text{IV} \rightarrow \text{I}$). The network is more constrained than Type 1, and the absence of a self-loop on I, combined with the moderate conditional probabilities and a constrained chord vocabulary, reflects the Simple-Dynamic profile: chord changes occur but the harmonic vocabulary cycles through a limited set of relationships.

In Type 3 (Dynamic-Complex, $n = 47$), I was the largest node, with $\text{IV} \rightarrow \text{I}$ (61.1%) and $\flat\text{VII} \rightarrow \text{I}$ (45.0%) as the dominant incoming transitions, followed by $\text{II} \rightarrow \text{I}$ (28.8%), $\text{VI} \rightarrow \text{I}$ (28.0%), and $\text{I} \leftrightarrow \text{V}$ (20.3%).

As established above, $\flat\text{VII} \rightarrow \text{I}$ in this type is best interpreted in a mode-sensitive manner: in major-key songs, it reflects a modal approach widely employed in popular music harmony [30], [36], whereas in minor-key songs it represents a diatonic subtonic-to-tonic motion. Together with the more conventional $\text{IV} \rightarrow \text{I}$ and $\text{II} \rightarrow \text{I}$ transitions, this modal/minor-inflected approach to tonic arrival reflects a particularly diverse transition grammar in this type. The co-occurrence of multiple high-probability approach paths to I, rather than a single dominant transition, is consistent with the Dynamic-Complex profile: harmonic arrival is achieved through varied routes, producing the high entropy and unpredictability captured in H_{RN} and H_{Bigrm} .

In Type 4 (Static-Loop, $n = 12$), I was overwhelmingly the largest node, exhibiting a strong self-loop of $\text{I} \rightarrow \text{I}$ (39.2%), the only type in which self-repetition appears as a dominant edge. All remaining top nodes converged upon I with exceptionally high conditional probabilities: $\text{III} \rightarrow \text{I}$ (91.7%), $\text{IV} \rightarrow \text{I}$ (74.2%), $\text{VII} \rightarrow \text{I}$ (73.6%), $\flat\text{VII} \rightarrow \text{I}$ (69.8%), and $\text{V} \rightarrow \text{I}$ (55.9%). This star-shaped convergence pattern, in which virtually every displayed representative transition resolves immediately to I with near-deterministic probability, reflects the loop-stasis grammar characteristic of this type: harmonic motion is minimal, and I functions less as a point of tonal resolution than as a static anchor sustained throughout the song.

Taken together, the network analysis yielded three findings that extend beyond what the index-based clustering alone could establish. First, the self-loop $\text{I} \rightarrow \text{I}$ (39.2%) in Type 4 (Static-Loop) provides directional confirmation of the loop-stasis grammar implied by its high RepeatRate: not merely that chords repeat, but that the tonic itself cycles back to itself with near-deterministic frequency, a pattern absent as a dominant edge in all other types. Second, Type 3 (Dynamic-Complex) is distinguished from the other types by the prominence of $\flat\text{VII} \rightarrow \text{I}$ (45.0%), a characteristic modal approach in popular music harmony that, alongside the conventional $\text{IV} \rightarrow \text{I}$ (61.1%), reveals a harmonic grammar irreducible to its index profile. Third, although Types 1 (Dynamic-Moderate) and 3 (Dynamic-Complex) showed similarly low RepeatRate values in the descriptive cluster profiles (0.090 and 0.071, respectively), their transition structures diverge markedly: Type 1 is anchored by $\text{IV} \rightarrow \text{I}$ (71.4%) with ancillary $\text{III} \rightarrow \text{I}$ and $\flat\text{VII} \rightarrow \text{IV}$ motion, whereas Type 3 distributes harmonic arrival across

multiple high-probability approach paths (IV→I, ♭VII→I, II→I, VI→I), confirming that frequent harmonic change combined with different vocabulary breadths produces structurally distinct, not merely quantitatively different, progressions. Together, these findings provide a music-theoretical characterization that connects the four statistical types to musically interpretable harmonic grammars.

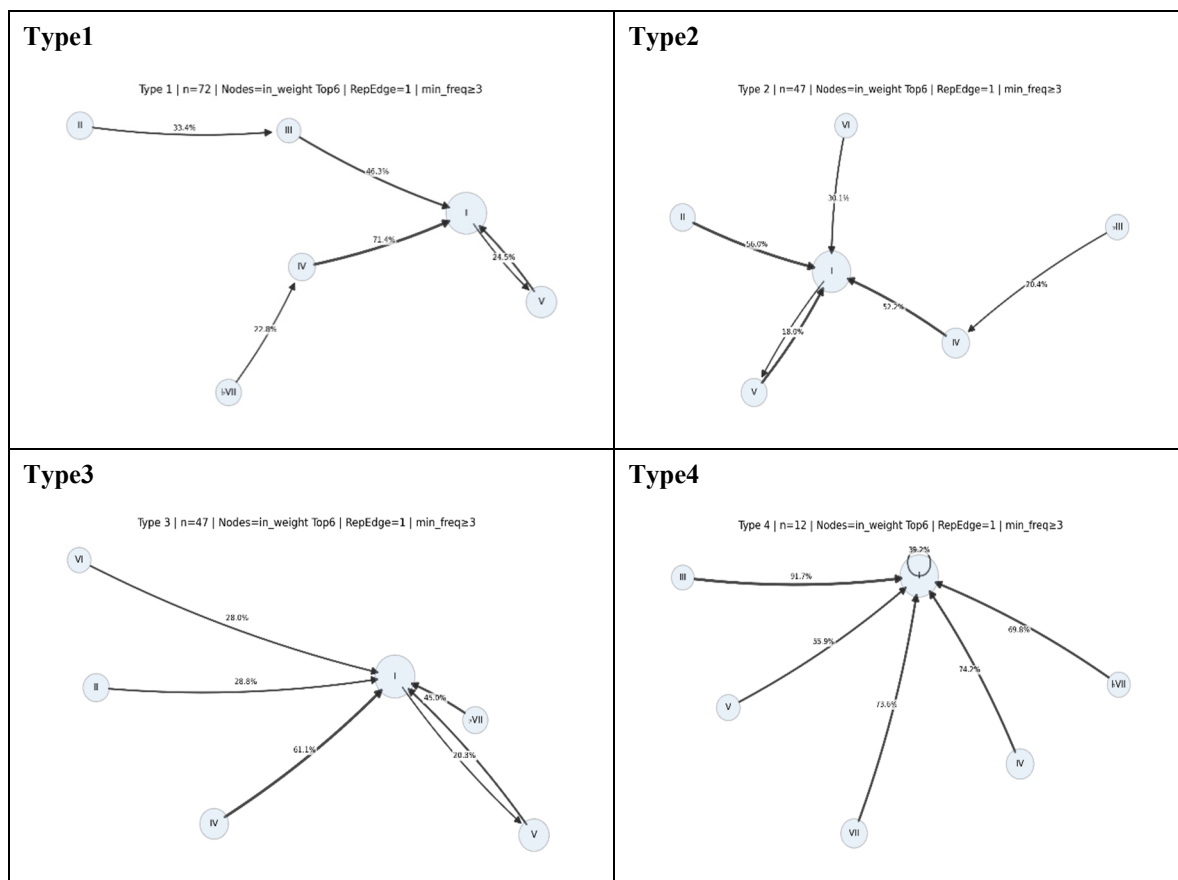


Figure 2. Directed transition-network diagrams for the four harmonic types

5. Discussion and conclusion

Using six derived indices and K-means clustering, this study identified four representative harmonic-development types and interpreted their most frequent chord transitions as distinct harmonic grammars: IV→I-anchored harmonic resolution (Type 1, Dynamic-Moderate), constrained-vocabulary cycling with multiple approach paths (Type 2, Simple-Dynamic), diverse multi-pathway approach including modal transitions (Type 3, Dynamic-Complex), and tonic-loop stasis (Type 4, Static-Loop).

The analysis confirmed that harmonic development in popular music can be classified into four descriptively distinguishable and interpretable types. Although no statistically significant variation in type distribution across musical genres was observed, each derived type exhibited distinct chord transition rules and structural characteristics that transcend genre boundaries.

With respect to the four research questions stated in the Introduction, the findings may be summarized as follows. RQ1—whether songs separate into multiple structurally distinct clusters within the feature space of the six indices—is confirmed descriptively and exploratorily: the K-means solution at $k = 4$ yielded four types with a mean silhouette coefficient of 0.3232. Although the silhouette coefficient indicates a modest degree of separation typical of clustering analyses on continuous feature spaces in music corpora, the four-cluster solution produced interpretable mean profiles across the six indices, while complementary music-theoretical characterization was provided by the type-specific transition-network patterns. RQ2a—whether the derived types exhibit distinct mean profiles across the six indices—is confirmed descriptively. Because the six indices were used to derive the clusters, inter-type differences on these same variables should not be treated as independent inferential evidence. Rather, the mean profiles show that the clustering solution produced interpretable contrasts in harmonic diversity, repetition, and entropy. RQ2b—whether the distribution of types

varies systematically across genres—is answered in the negative: the Genre $\times$ Type test revealed no statistically significant association between genre and type distribution ($\chi^2 = 3.68$, $df = 9$, $p = .931$; Monte Carlo $p = .937$; $G = 4.38$, $p = .885$; Cramér's $V = 0.083$). This null result, consistent across multiple convergent tests, suggests that the derived harmonic types represent cross-genre structural archetypes rather than genre-specific conventions. RQ3—whether the most frequent chord transition patterns are qualitatively distinguishable across types—is answered affirmatively: network analysis revealed qualitatively differentiated transition grammars for each type. The clearest contrast is between Type 4 (Static-Loop), which exhibits tonic anchoring through $I \rightarrow I$ (39.2%) and multiple high-probability resolutions into I , and Type 3 (Dynamic-Complex), which exhibits multi-pathway tonic arrival through $IV \rightarrow I$, $bVII \rightarrow I$, $II \rightarrow I$, and $VI \rightarrow I$. Type 1 (Dynamic-Moderate) shows $IV \rightarrow I$ -centered harmonic resolution (71.4%), and Type 2 (Simple-Dynamic) cycles through a constrained set of approach paths anchored by $II \rightarrow I$ and $IV \rightarrow I$. These patterns confirm that the cluster distinctions have meaningful musical content beyond their numerical profiles.

Two further observations warrant discussion. First, the partial similarity of certain dynamic profiles deserves attention. Types 1 and 3 both show low RepeatRate values in the descriptive cluster profiles (0.090 and 0.071, respectively). However, they differ clearly in harmonic vocabulary and entropy: Type 3 has substantially higher UniqueRN, UniqueBigram, H_RN , and H_Bigram values, and the network analysis further reveals that Type 1 is anchored primarily by $IV \rightarrow I$, whereas Type 3 distributes tonic arrival across multiple approach paths. This convergence between the descriptive and network analyses shows why repetition alone is insufficient to classify harmonic development and why the six-index framework, complemented by transition-network analysis, is necessary for separating different forms of harmonic dynamism. Second, the lower-than-expected explanatory power of genre with respect to harmonic type warrants attention. The non-significant chi-square result, supported by convergent supplementary tests and a very small Cramér's V (0.083), indicates that broad genre categories do not adequately predict harmonic type in the present sample.

Two complementary explanations are offered. First, increasing cross-genre hybridization in contemporary popular music (e.g., the incorporation of hip-hop production aesthetics into country and pop, or the adoption of R&B harmonic language in indie rock) may be progressively eroding the harmonic distinctiveness that once differentiated genres, rendering broad genre labels less informative as predictors of harmonic structure [3]. Second, and more critically from a methodological standpoint, the four genre categories employed in this study (Country/Folk, Pop/Dance, R&B/Soul/Hip-Hop, Rock/Indie) subsume considerable within-category heterogeneity. The R&B/Soul/Hip-Hop category, in particular, conflates traditions with divergent harmonic grammars: while Soul and R&B often rely on extended functional harmony ($ii-V-I$ progressions, chromatic voice-leading), contemporary Hip-Hop frequently employs modal stasis, sample-based harmonic loops, and pitch-shifted vocal samples that resist conventional RNA representation.

Future research should adopt finer-grained genre taxonomies—such as those provided by MusicBrainz or Discogs genre tags—to better capture intra-genre harmonic diversity [50, 51]. Additionally, the present sample ($N = 178$) is drawn exclusively from the Billboard Hot 100 (2023–2025) and is further restricted to songs containing an identifiable chorus section, which may systematically under-represent harmonically complex sub-genres (e.g., jazz-influenced R&B, progressive indie) as well as non-English popular music.

On the methodological front, three issues warrant disclosure. First, the adoption of a simplified RNA notation necessarily omits finer harmonic details—including chord quality, voicing, and bass line movement [11], [52]—but this was an intentional choice made to enhance the universality and reproducibility of the analysis and to reliably capture the core structural features of harmonic grammar. Second, the six analytical indices are not entirely independent: examining their intercorrelations revealed several strong associations (the full correlation matrix is provided in Appendix B). UniqueRN and UniqueBigram were highly correlated ($r = .931$), H_RN and H_Bigram were strongly correlated ($r = .924$), and Top1Share was negatively correlated with H_RN ($r = -.856$), reflecting that higher concentration on a dominant chord mathematically implies lower distributional entropy and that a larger chord vocabulary tends to co-occur with a larger transition vocabulary. A principal component analysis confirmed that the six indices share substantial common structure ($PC1 = 76.5\%$; $PC1+PC2 = 93.5\%$ of variance explained). We nevertheless retained all six indices rather than collapsing them into principal components for three reasons. First, each index captures a music-theoretically distinct construct that is interpretable on its own terms (e.g., the dominance of a single chord vs. the distributional uniformity of all chords); PCA components would obscure these interpretive distinctions for music-theoretical readers. Second, prior to clustering, each of the six indices was standardized to zero mean and unit variance (Section 3.3), ensuring that no single index's raw measurement scale could disproportionately influence the Euclidean

distance metric underlying the K-means algorithm. Third, the present study is exploratory and descriptive rather than predictive; PCA would be appropriate for predictive efficiency [53], whereas the goal here is to preserve the distinct musical meaning of each variable — including the contrast between "concentration" and "complexity" — as a basis for the music-theoretical interpretation given in Section 4.4. Future predictive applications may revisit the dimensionality of this feature space.

Third, the four-type classification should be interpreted as a standardized K-means-based analytical typology rather than as an algorithm-independent natural classification. Different clustering algorithms rely on different assumptions and partitioning principles; consequently, alternative methods may produce different song assignments and, in turn, somewhat different cluster profiles and transition networks. This limitation is important because the subsequent network analysis depends on the cluster labels used to define each type. Nevertheless, the purpose of the present study was not to establish a definitive taxonomy of all possible harmonic progression types. Rather, it was to demonstrate how aggregate harmonic indices and transition-network analysis can be combined to construct an interpretable song-level typology of contemporary popular music. Within this design, K-means provided a direct connection between the standardized six-index feature space, type-level mean profiles, and subsequent transition-network interpretation. Future research may extend this approach by comparing K-means-based typologies with sequence-based models, such as HMM-GMM, or with chord-embedding approaches. Such models would require chord-level sequential representations and could address different questions, such as the generative process of harmonic progression or the prediction of chord sequences over time.

In addition to the issues discussed above, this study has several limitations. First, as user-generated content (UGC), the UG data contain inherent noise: notational inconsistencies and chord simplifications introduced for ease of performance cannot be fully eliminated. Second, the final analytical sample was reduced to 178 songs after duplicate removal and the Chorus = 0 exclusion criterion; this strengthens analytical consistency but limits generalizability. Third, treating each song as a single undivided sequence means that fine-grained structural variation across sections (e.g., introduction, verse, chorus) could not be captured in detail. Fourth, interpretive subjectivity arising from K-means clustering and researcher-defined type labels cannot be fully excluded. Fifth, $\flat$ III, $\flat$ VI, and $\flat$ VII are encoded uniformly regardless of mode, which conflates modal-borrowing chords in major keys with diatonic chords in minor keys; although Appendix A confirms that the principal transition patterns are stable when major and minor songs are analyzed separately, future research adopting mode-aware encoding may further refine the type-specific harmonic grammars reported here.

Despite these limitations, this study makes a theoretical contribution by proposing an empirical analytical framework that classifies popular music harmony not by restricting analysis to a handful of representative progressions, but by treating entire songs from the perspective of "transitional grammar." In particular, by proposing six indices capable of comparing structural differences among songs without requiring complex harmonic theory, and by translating statistical findings into musical interpretation, the study helps narrow the gap between numerical data and musical context.

Building on the mode-sensitive distinction established above, the $\flat$ VII→I transition that distinguishes Type 3 merits more direct theoretical grounding. In major-key contexts, $\flat$ VII is borrowed from the parallel Mixolydian/minor mode and resolves to I through a characteristic plagal-style descent (an approach to the tonic from the chord one step below, bypassing the conventional dominant chord)—often termed the "Mixolydian cadence" in rock-harmony scholarship [9], [30], [36]; in minor-key contexts, the same $\flat$ VII→I motion is diatonic and represents a natural step-wise resolution. In both readings, $\flat$ VII→I bypasses the dominant V and substitutes a non-leading-tone approach to the tonic (i.e., it does not use the seventh scale degree that strongly pulls toward the tonic, as in classical harmony), producing the rock/pop-idiomatic, less tension-driven sense of arrival that distinguishes much popular-music harmony from common-practice tonal voice-leading (the chord-progression conventions of Western classical music) [13], [36]. The fact that Type 3 combines this transition with diverse additional approach paths (IV→I, II→I, VI→I)—rather than relying on a single dominant approach—is therefore best read as a marker of modally and harmonically broad arrival vocabularies, consistent with the high diversity and entropy values that define this cluster numerically.

For composition and arrangement practice, the four type-specific transition grammars supply ready-made templates that prioritize different listener experiences: Type 4's static-loop grammar provides a sustained tonal anchor suitable for hook-centered or texture-driven sections; Type 1's IV→I-anchored functional grammar supplies a clear sense of forward motion within a moderate vocabulary; Type 2's constrained cycling grammar offers a compact, four-chord-loop scaffold; and Type 3's multi-pathway grammar—including the $\flat$ VII→I

substitution—is appropriate when stylistic breadth or harmonic surprise is desired. These type-specific grammars further illustrate how harmonic type could complement the genre-, artist-, and tempo-based features discussed below in recommendation-system applications—for example, by distinguishing two pop songs that share a tempo and instrumentation but differ in whether the listener experiences static loop, functional resolution, or modally-mixed motion.

From a practical standpoint, the resulting harmonic types may serve as a classification tool for rapidly communicating the developmental characteristics of a song in composition and arrangement practice. In composition-support systems, type-specific transition grammars could be used to recommend likely next chords or to suggest alternative harmonic paths within a selected stylistic space. In music recommendation systems, harmonic type could serve as a structural feature that complements genre, artist, tempo, or audio-based similarity measures. In playlist curation, songs may be grouped not only by surface genre but also by underlying harmonic organization. In music education settings, the framework can function as a pedagogical instrument that enables students to understand chord progressions not through rote memorization but as structural patterns in motion. Thus, the four-type classification provides a bridge between computational harmonic analysis and practical applications in music creation, recommendation, curation, and education.

Beyond their descriptive value, the four type-specific transition grammars also lend themselves to falsifiable predictions about listener responses. The information-dynamics framework [6, 7] would predict that Type 4 (Static-Loop), with its near-deterministic resolutions and dominant I→I self-loop, should elicit lower moment-to-moment expectancy violation and stronger reliance on textural or rhythmic affect; that Type 3 (Dynamic-Complex), with its multiple-pathway approaches including bVII→I, should elicit higher expectancy variance and greater perceived harmonic surprise; and that Types 1 and 2, occupying intermediate positions, should elicit intermediate expectancy profiles. Empirically testing these predictions through listener studies—for example, by collecting moment-by-moment expectancy or affect ratings on representative songs from each type—would convert the present statistical typology into a properly perceptual one and would directly address whether the four grammars correspond to distinct listening experiences in the way the literature on harmonic expectation suggests.

In conclusion, this study has demonstrated that harmonic development in popular music can be empirically classified into four interpretable types using six RNA-based indices. The results show that popular-music harmony is organized not by a single universal progression rule but through multiple distinct transition grammars, ranging from static tonic anchoring (Type 4) to dynamic-complex multi-pathway movement (Type 3), with IV→I-anchored and constrained-vocabulary intermediates (Types 1 and 2).

Furthermore, the analytical sample is drawn from a narrow three-year window (Billboard 2023–2025) and disproportionately reflects mainstream commercial popularity; the claim that the derived types represent cross-genre structural archetypes is therefore strictly bounded by this sampling frame and does not, in itself, speak to historical trends in harmonic practice.

Future research should extend this framework by separately analyzing section-level (verse/chorus) sequences, incorporating richer chord-quality information, and testing larger, temporally more diverse, and stylistically broader datasets—including pre-2020 chart corpora and non-English popular music—to assess the historical and cross-cultural robustness of the four-type typology reported here. This study offers a new analytical point of departure for the empirical study of popular music harmony.

Acknowledgments: This research was supported by Kumoh National Institute of Technology (2024-2026).

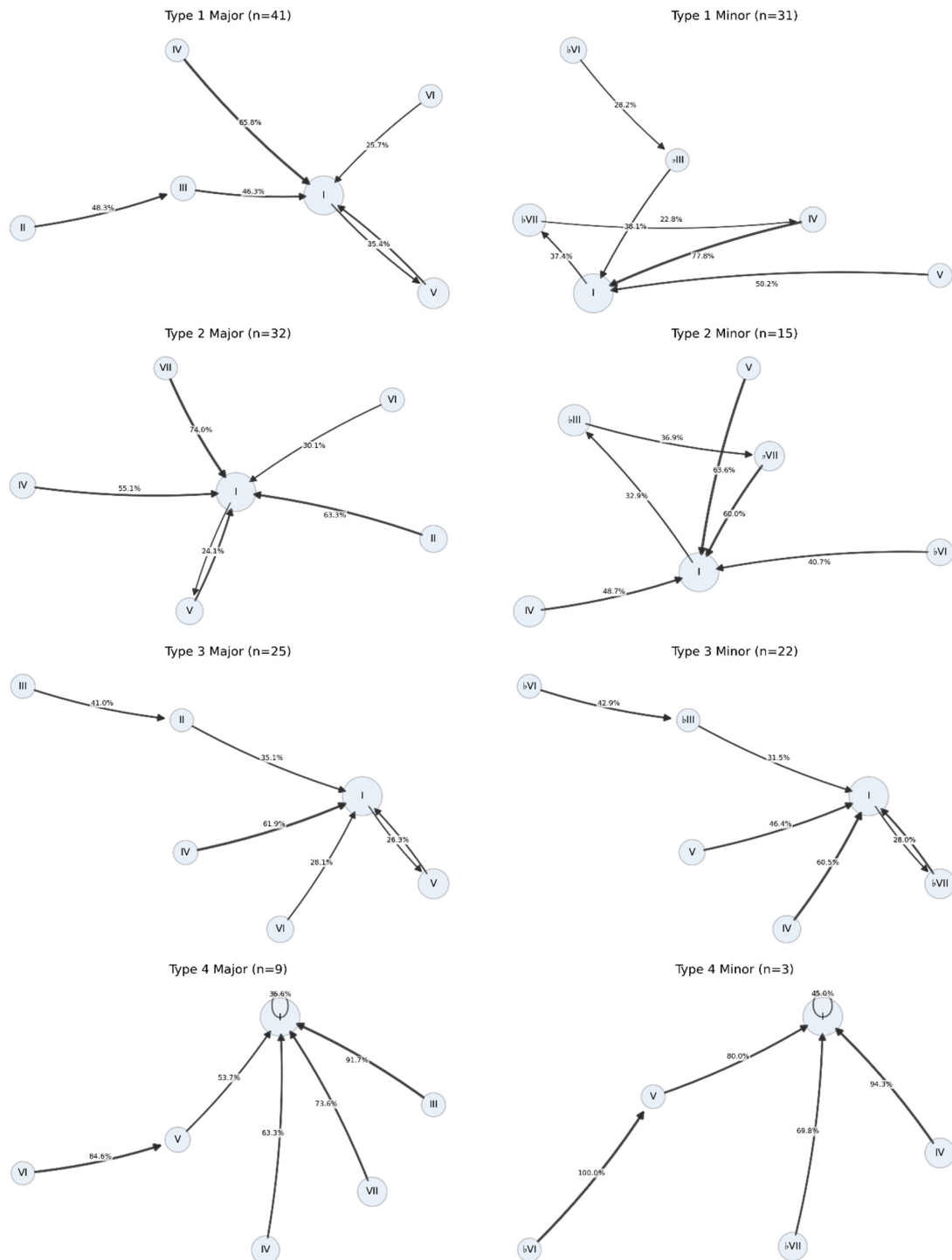
Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- [1] B. Duinker, "Song form and the mainstreaming of hip-hop music," *Current Musicology*, vol. 107 pp. 93–135, 2020, doi: <https://doi.org/10.52214/cm.v107i.7177>.
- [2] E. H. Margulis, *On repeat: How music plays the mind*, Oxford University Press, 2013, doi: <https://doi.org/10.1093/acprof:oso/9780199990825.001.0001>.
- [3] D. Brackett, *Categorizing sound: Genre and twentieth-century popular music*, Univ of California Press, 2016, doi: <https://doi.org/10.1525/california/9780520248717.001.0001>.

- [4] M. Mauch, R. M. MacCallum, M. Levy, and A. M. Leroi, "The evolution of popular music: USA 1960–2010," *Royal Society open science*, vol. 2, no. 5, 2015, doi: <https://doi.org/10.1098/rsos.150081>.
- [5] P. Tagg, *Everyday tonality II, Towards a Tonal Theory of What Most People Hear*, New York and Huddersfield: The Mass Media Music Scholars' Press (MMMSPP), 2014.
- [6] M. T. Pearce and G. A. Wiggins, "Auditory expectation: the information dynamics of music perception and cognition," *Topics in cognitive science*, vol. 4, no. 4, pp. 625–652, 2012, doi: <https://doi.org/10.1111/j.1756-8765.2012.01214.x>.
- [7] D. B. Huron, *Sweet Anticipation: Music and the Psychology of Expectation*, MIT Press, 2006, doi: <https://doi.org/10.7551/mitpress/6575.001.0001>.
- [8] D. Temperley and T. d. Clercq, "Statistical Analysis of Harmony and Melody in Rock Music," *Journal of New Music Research*, vol. 42, no. 3, pp. 187–204, 2013, doi: <https://doi.org/10.1080/09298215.2013.788039>.
- [9] T. De Clercq and D. Temperley, "A corpus analysis of rock harmony," *Popular music*, vol. 30, no. 1, pp. 47–70, 2011, doi: <https://doi.org/10.1017/S026114301000067X>.
- [10] T. De Clercq, "Interactions between harmony and form in a corpus of rock music," *Journal of Music Theory*, vol. 61, no. 2, pp. 143–170, 2017, doi: <https://doi.org/10.1215/00222909-4149525>.
- [11] D. Tymoczko, M. Gotham, M. S. Cuthbert, and C. Ariza, "The Romantext Format: A flexible and standard method for representing Roman numeral analyses," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, pp. 123–129, 2019, doi: <https://doi.org/10.5281/zenodo.3527756>.
- [12] M. Rohrmeier, "Towards a generative syntax of tonal harmony," *Journal of Mathematics and Music*, vol. 5, no. 1, pp. 35–53, 2011, doi: <https://doi.org/10.1080/17459737.2011.573676>.
- [13] M. Rohrmeier and I. Cross, "Statistical properties of tonal harmony in Bach's chorales," in *Proceedings of the 10th international conference on music perception and cognition*, vol. 6, no. 4: ICMPC Sapporo, pp. 619–627, 2008.
- [14] J. Burgoyne, J. Wild, and I. Fujinaga, "An Expert Ground Truth Set for Audio Chord Recognition and Music Analysis," in *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, FL, pp. 633–638, 2011.
- [15] D. Temperley, *The musical language of rock*, Oxford University Press, 2018, doi: <https://doi.org/10.1093/oso/9780190653774.001.0001>.
- [16] E. Clarke and N. Cook, *Empirical Musicology: Aims, Methods, Prospects*, Oxford University Press, 2004, doi: <https://doi.org/10.1093/acprof:oso/9780195167498.001.0001>.
- [17] D. Temperley, *Music and Probability*, MIT Press, 2007, doi: <https://doi.org/10.7551/mitpress/4807.001.0001>.
- [18] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC Music Database: Popular, Classical and Jazz Music Databases," in *[Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR 2002)]*(<http://ismir2002.ismir.net/>), vol. 2, pp. 287–288, 2002, doi: <https://doi.org/10.5281/zenodo.1416474>.
- [19] H. V. Koops, W. B. De Haas, J. A. Burgoyne, J. Bransen, A. Kent-Muller, and A. Volk, "Annotator subjectivity in harmony annotations of popular music," *Journal of New Music Research*, vol. 48, no. 3, pp. 232–252, 2019, doi: <https://doi.org/10.1080/09298215.2019.1613436>.
- [20] S. Kantarelis, K. Thomas, V. Lyberatos, E. Dervakos, and G. Stamou, "Chordonomicon: A dataset of 666,000 songs and their chord progressions," *arXiv preprint arXiv:2410.22046*, 2024, doi: <https://doi.org/10.48550/arXiv.2410.22046>.
- [21] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recogn. Lett.*, vol. 31, no. 8, pp. 651–666, 2010, doi: <https://doi.org/10.1016/j.patrec.2009.09.011>.
- [22] K. Shaffer et al., "A cluster analysis of harmony in the McGill Billboard dataset," *Empirical Musicology Review*, vol. 14, no. 3–4, pp. 146–162, 2020, doi: <https://doi.org/10.18061/emr.v14i3-4.5576>.
- [23] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987, doi: [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [24] L. Knopoff and W. Hutchinson, "Entropy as a measure of style: The influence of sample length," *Journal of Music Theory*, vol. 27, no. 1, pp. 75–97, 1983, doi: <https://doi.org/10.2307/843561>.
- [25] J. Serrà, Á. Corral, M. Boguñá, M. Haro, and J. L. Arcos, "Measuring the evolution of contemporary western popular music," *Scientific reports*, vol. 2, no. 1, p. 521, 2012, doi: <https://doi.org/10.1038/srep00521>.
- [26] D. Odekerken, H. V. Koops, and A. Volk, "Improving audio chord estimation by alignment and integration of crowd-sourced symbolic music," *Transactions of the International Society for Music Information Retrieval*, vol. 4, no. 1, pp. 141–155, 2021, doi: <https://doi.org/10.5334/tismir.81>.
- [27] J. Summach, "The structure, function, and genesis of the prechorus," *Music Theory Online*, vol. 17, no. 3, 2011, doi: <https://doi.org/10.30535/mto.17.3.2>.
- [28] T. De Clercq, "Embracing ambiguity in the analysis of form in pop/rock music, 1982–1991," *Music Theory Online*, vol. 23, no. 3, 2017, doi: <https://doi.org/10.30535/mto.23.3.4>.

- [29] C. L. Krumhansl, *Cognitive foundations of musical pitch*, Oxford University Press, 2001, doi: <https://doi.org/10.1093/acprof:oso/9780195148367.001.0001>.
- [30] C. Doll, *Hearing harmony: Toward a tonal theory for the rock era*, University of Michigan Press, 2017, doi: <https://doi.org/10.3998/mpub.3079295>.
- [31] T. De Clercq, "Tempo versus average rates of harmonic rhythm in popular music: A study of three corpora," *Musicae scientiae*, vol. 27, no. 3, pp. 521-540, 2023, doi: <https://doi.org/10.1177/10298649221091483>.
- [32] D. Arthur and S. Vassilvitskii, "k-means++: the advantages of careful seeding," presented at the Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, New Orleans, Louisiana, pp. 1027-1035, 2007.
- [33] J. H. Ward, Jr., "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236-244, 1963, doi: <https://doi.org/10.1080/01621459.1963.10500845>.
- [34] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Advances in Neural Information Processing Systems (NIPS)*, vol. 14, pp. 849-856, 2001.
- [35] G. J. McLachlan and D. Peel, *Finite Mixture Models*, Wiley Series in Probability and Statistics, New York: John Wiley & Sons, 2000, doi: <https://doi.org/10.1002/0471721182>.
- [36] N. Biamonte, "Triadic modal and pentatonic patterns in rock music," *Music Theory Spectrum*, vol. 32, no. 2, pp. 95-110, 2010, doi: <https://doi.org/10.1525/mts.2010.32.2.95>.
- [37] W. G. Cochran, "Some methods for strengthening the common χ^2 tests," *Biometrics*, vol. 10, no. 4, pp. 417-451, 1954, doi: <https://doi.org/10.2307/3001616>.
- [38] R. R. Sokal and F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research*, 2nd ed., San Francisco, CA: W. H. Freeman, 1981.
- [39] H. Cramér, *Mathematical Methods of Statistics*, Princeton, NJ: Princeton Univ. Press, 1946, repr. 1999.
- [40] A. C. Hope, "A simplified Monte Carlo significance test procedure," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 30, no. 3, pp. 582-598, 1968, doi: <https://doi.org/10.1111/j.2517-6161.1968.tb00759.x>.
- [41] J. Cohen, *Statistical power analysis for the behavioral sciences*, Routledge, 2013, doi: <https://doi.org/10.4324/9780203771587>.
- [42] M. E. J. Newman, *Networks*, 2nd ed. Oxford, U.K.: Oxford Univ. Press, 2018. doi: <https://doi.org/10.1093/oso/9780198805090.001.0001>.
- [43] E. Kim, J. Han, and G. Kwon, "Distinguishing Supportive Activities at Industry-Level Value Chain Analysis," *INTERNATIONAL JOURNAL OF CONTENTS*, vol.19, no.1, pp.1-10, Mar. 2023, doi: <https://doi.org/10.5392/ijoc.2023.19.1.001>.
- [44] J. F. Paiement, D. Eck, and S. Bengio, "A Probabilistic Model for Chord Progressions," in *Proc. 6th Int. Conf. Music Inf. Retrieval (ISMIR)*, London, UK, pp. 312-319, Sep. 2005.
- [45] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215-239, 1978, doi: [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7).
- [46] L. Kaufman and P. J. Rousseeuw, "Finding groups in data," Hoboken: Wiley Online Library, vol. 1, p. 371, 1990.
- [47] A. Struyf, M. Hubert, and P. Rousseeuw, "Clustering in an Object-Oriented Environment," *Journal of Statistical Software*, vol. 1, no. 4, pp. 1-30, Feb. 10, 1997, doi: <https://doi.org/10.18637/jss.v001.i04>.
- [48] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65-70, 1979.
- [49] S. J. Haberman, "The analysis of residuals in cross-classified tables," *Biometrics*, vol. 29, no. 1, pp. 205-220, 1973, doi: <https://doi.org/10.2307/2529686>.
- [50] J. C. Lena and R. A. Peterson, "Classification as culture: Types and trajectories of music genres," *American sociological review*, vol. 73, no. 5, pp. 697-718, 2008, doi: <https://doi.org/10.1177/000312240807300501>.
- [51] H. Schreiber, "Improving Genre Annotations for the Million Song Dataset," in *Proc. 16th Int. Soc. Music Inf. Retrieval Conf. (ISMIR)*, Málaga, Spain, pp. 241-247, Oct. 2015.
- [52] B. Pardo and W. P. Birmingham, "Algorithms for Chordal Analysis," *Computer Music Journal*, vol. 26, no. 2, pp. 27-49, 2002, doi: <https://doi.org/10.1162/014892602760137167>.
- [53] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Springer, 2002.

APPENDIX A. Mode-separated transition networks by harmonic type

The figure presents separate transition networks for major-key and minor-key songs within each harmonic type. The networks were constructed using the same visualization rule as Figure 2: nodes were selected based on the top six in-weight centrality values, and one representative outgoing transition per source node was selected according to the highest conditional transition probability, with a minimum frequency threshold of three. These networks are provided as a diagnostic robustness check to assess whether the type-specific transition patterns are affected by the mixture of major- and minor-key songs. Because some Type $\times$ Mode subgroups, especially Type 4 minor-key songs, contain a small number of cases, the networks should be interpreted as supplementary diagnostic evidence rather than as independent bases for redefining the harmonic types.

APPENDIX B. Correlation matrix of the six analytical indices

Variables	UniqueRN	UniqueBigram	Top1Share	RepeatRate	H_RN	H_Bigram
UniqueRN	1.000	0.931	-0.585	-0.377	0.887	0.858
UniqueBigram	0.931	1.000	-0.541	-0.311	0.829	0.907
Top1Share	-0.585	-0.541	1.000	0.784	-0.856	-0.726
RepeatRate	-0.377	-0.311	0.784	1.000	-0.604	-0.475
H_RN	0.887	0.829	-0.856	-0.604	1.000	0.924
H_Bigram	0.858	0.907	-0.726	-0.475	0.924	1.000



© 2026 The Author(s). Copyright remains with the author(s). This article is published by The Korea Contents Association under the Creative Commons Attribution 4.0 International License (CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/>