Quantifying Tonal Stability and Key Ambiguity Metrics across Musical Genres using KS and Tonnetz

Mehak Dhawan[1]

Received April 16, 2026
Accepted August 3, 2026
Electronic access September 15, 2026

Quantifying tonal stability and harmonic motion across musical styles remains a central challenge in computational musicology. This study presents a computational framework for analyzing tonal dynamics and key ambiguity using symbolic MIDI (Musical Instrument Digital Interface) data. Musical excerpts were segmented into fixed temporal windows and converted into chroma representations, from which six-dimensional Tonnetz centroids were derived to capture multidimensional harmonic relationships among pitch classes. Framewise Euclidean distances between consecutive centroids produced the Harmonic Change Detection Function (HCDF), providing a continuous measure of tonal motion over time. In parallel, rolling key-strength estimates were obtained using the Krumhansl–Schmuckler key-finding algorithm applied to flattened MIDI streams. To reduce artifacts from quantized MIDI timing and short-term fluctuations, centroid trajectories and key-strength time series were regularized through Gaussian smoothing and small temporal jitter. Building on these measures, we introduce the Key Ambiguity Index (KAI), a composite metric combining the variance of key-strength estimates with normalized Tonnetz path length to capture tonal uncertainty. Applying this framework to MIDI corpora spanning Classical, Pop, Rock, Jazz, and Stravinsky repertoire reveals systematic stylistic differences. Classical and Stravinsky excerpts exhibit the highest composite tonal ambiguity by KAI, while Pop and Rock display the most constrained harmonic motion. Jazz occupies an intermediate position across all three metrics. These results demonstrate that Tonnetz-based metrics and the proposed KAI provide an effective computational approach for quantifying stylistic variation in tonal organization.

Keywords: tonal stability; key ambiguity; Tonnetz; Krumhansl–Schmuckler algorithm; Key Ambiguity Index; computational musicology; MIDI; harmonic motion

Introduction

Problem Statement

Understanding tonal stability and harmonic centroids is essential for both traditional music theory and computational music analysis[1,2]. While Classical music generally adheres to a stable key with predictable harmonic progressions, Jazz and Stravinsky's compositions often exhibit rapid key changes and more complex tonal structures[3,4]. Existing methods, such as the Krumhansl–Schmuckler (KS) key-finding algorithm, capture certain aspects of tonality but do not fully account for harmonic ambiguities or dynamic shifts in tonal centroids[5,6]. This study proposes a comprehensive analytical framework capable of quantifying both tonal stability and harmonic modulation across musical genres.

Exact Premise and Importance

Compositions that traverse multiple key areas are expected to exhibit reduced tonal stability and more frequent shifts in tonal centers[1,7]. We introduce the Key Ambiguity Index (KAI), which integrates the rolling variance of key strength, derived from the Krumhansl–Schmuckler method, with the cumulative path length in Tonnetz space. Application of KAI to excerpts from Classical, Jazz, and Stravinsky repertoires is expected to reveal higher tonal ambiguity in Jazz and Stravinsky pieces relative to Classical works[3,8]. This measure facilitates quantitative comparisons of tonal dynamics across genres, highlighting stylistic distinctions in harmonic behavior.

Objectives

The present study has four objectives. Firstly, we generate rolling time series of key-strength (KS) and tonal centroids (TC) for MIDI excerpts for various musical genres based on Tonnetz theory. Second, we propose the Key Ambiguity Index (KAI), which is an integration of variance in key strength and tonal centroid motion. Third, KAI-distributions of Classical, Jazz, Stravinsky, Pop and Rock music are analyzed and compared. Lastly, differences in tonal stability and harmonic movement between each genre are statistically compared, with

[1] *Jumeirah College, UAE*

the aim of determining whether Jazz and Stravinsky excerpts are more tonally ambiguous than Classical excerpts.

The study of the "stickiness" of tonal centers and the movement of harmony has long been an important aspect of computational musicology. As we experience the music of the classical repertoire, we are accustomed to hearing tonic centers that remain stable and predictable in terms of harmony, but in the realm of Jazz and much 20th-century music, as exemplified by the works of Igor Stravinsky, we experience quick tonal changes and unusual pitch-class relationships. Several techniques have been applied to the study of the stability of tonality in both symbolic music representations and in the sound itself, including the study of pitch-class distributions, the use of chroma features[9], and the application of the Krumhansl–Schmuckler (KS) algorithm for key finding[10]. The Tonnetz, or geometric representation of pitch classes and their relationships, has also proven to be an important tool in the study of the proximity of chords, the centroids of keys, and the path of harmony through multidimensional space[4]. By combining the KS key finding algorithm with the centroids derived from the Tonnetz, we are able to gain an understanding of the way in which the tonal centers shift and the way in which the harmony moves through multidimensional space.

In this study, we seek to determine whether Jazz and Stravinsky excerpts exhibit less tonal stability and greater harmonic variation than those in the Classical repertoire. Our method relies on the analysis of MIDI data, which fixes pitch and time while removing any variability in performance. We calculate six-dimensional Tonnetz-derived tonal centroids and rolling KS key strength time series for each excerpt, examining the evolution of tonal centers and harmonic movement in six dimensions. Additionally, we introduce a small jitter in the calculation of the KS time series, which reduces the effect of quantized onset times in the MIDI data. Finally, we smooth the Tonnetz-derived centroids and KS time series using a Gaussian filter, reducing frame-to-frame variation in tonal centers and revealing harmonic movement[11].

We also propose a novel Key Ambiguity Index (KAI) given by the sum of the z-score normalized key strength variance and the Tonnetz-derived path length per minute. The KAI distributions are computed from snippets of Classical, Jazz, Stravinsky, Pop, and Rock pieces. To compare the statistics of each genre, we will use Kruskal–Wallis tests with Dunn's post-hoc tests and false discovery rate correction, examining the differences in tonal stability and harmonic movement across each group. We will also verify the robustness of our results by examining the impact of some of our parameter choices, such as frame sizes and smoothing, on centroid movement. KAI combines rolling KS key strength with Gaussian smoothing of Tonnetz-driven harmonic movement to derive a numeric measure of tonal uncertainty, enabling a quantitative comparison of tonal and harmonic characteristics unique to each genre. We anticipate that Jazz and Stravinsky pieces will demonstrate lower average key strength and longer Tonnetz movement than Classical pieces, reflecting greater tonal ambiguity and harmonic movement.

Motivation

The understanding of the construction of music and the movement of harmony is an important goal in both music theory and computational musicology. Conventional analysis involves listening and studying the score, but quantitative analysis brings an objective and replicable quality in assessing tonal stability and harmony complexity. By using pitch class distributions, chroma features, and tonal centroids derived from the Tonnetz[4,9], we can use algorithms that identify patterns in musical sequences that are characteristic of different musical styles. Moreover, we can use the Tonnetz and other tools to create musical sequences that are useful in analysis and composition. This allows us to examine tonal ambiguity, harmony, and style using a quantitative and replicable approach.

In building upon these computer-related concepts, this research explores tonal movement and harmonic indeterminacy across different music genres. In integrating tonal movement based on Tonnetz theory with rolling key strength values based on Krumhansl and Schmuckler theory, this research creates an evolving picture of how tonal stability is maintained or wavers over time. This framework enables quantification of not only tonal movement but also instances of tonal interpretation that become ambiguous or uncertain. When this framework is used to analyze music from Classical, Pop, Rock, Jazz, and Stravinsky music genres, unique stylistic characteristics emerge. The results suggest that music genres with more developed chromaticism and longer harmony progressions also have more tonal freedom and key ambiguity than music genres with more tonal center stability. Overall, these results demonstrate how computer-related research supports traditional music theory by providing quantification of music stylistic differences.

Background

Tonal Perception and Tonal Stability

Tonal perception denotes the cognitive ability of listeners to infer the key or tonal center of a musical composition, thereby distinguishing between notes perceived as stable and those that are less resolved. This hierarchical structure is consistent across transpositions but exhibits systematic differences between major and minor modes; minor keys are further complicated by the presence of multiple forms (natural, harmonic, and melodic)[1,6,7].

Quantifying tonal stability is essential, as it informs multiple aspects of music cognition—including expectancy formation, memory, and the computational identification of tonal centers—underpinning algorithmic models in musicology [1,12]. Empirical studies using probe-tone experiments have consistently shown that listeners' ratings of how well a note fits a given context correspond closely to these hierarchical structures, validating the psychological reality of tonal stability across musical genres [1,6].

In short, tonal perception and the hierarchy of pitch stability form the foundational framework for studying how humans experience and process music, and provide a key reference point for computational models such as the Krumhansl–Schmuckler key-finding algorithm [1,6].

Background on the Krumhansl–Schmuckler Key-Finding Algorithm

Understanding the tonal structure of music is a central goal in both music theory and music cognition. In Western tonal music, pitches are organized around a central reference tone called the tonic, forming a tonal hierarchy. This hierarchy assigns greater psychological "stability" to some tones over others: the tonic is the most stable, followed by other diatonic notes, and finally non-diatonic notes [12]. These hierarchical relationships have been verified experimentally using the probe-tone procedure, where listeners rate the "goodness-of-fit" of a single tone following a tonal context. The resulting ratings closely match theoretical predictions of tonal importance, confirming that listeners perceive certain tones as cognitively central within a key [13,14].

The Krumhansl–Schmuckler (KS) key-finding algorithm [10,14] provides a computational framework for identifying the tonal center of a musical passage. The algorithm operates by representing musical material as pitch-class vectors and correlating them with experimentally derived key profiles, enabling the inference of both static and time-varying tonal centers.

The KS algorithm abstracts tonal information from surface pitch patterns by summing durations of pitch occurrences while ignoring temporal order, mirroring perceptual processes. It is flexible, capable of analyzing both short passages and entire pieces to identify static or time-varying tonal centers. Empirical validation confirms that its predictions closely align with listeners' perceptions of tonality, including tonal stability, perceived key motion, and memory effects such as false recall of implied notes [6,13,14].

Chroma Feature Extraction

Chroma features provide a compact representation of tonal content by mapping all pitches of a musical signal into 12 pitch classes (C–B), effectively ignoring octave information while preserving harmonic relationships [15,16]. This octave invariance reflects the perceptual property that pitches separated by octaves share similar tonal "color," making chroma features well-suited for capturing harmonic and melodic structure.

In practice, chroma features are represented as 12-dimensional vectors, where each element corresponds to the energy of a pitch class within a short time frame. Over time, these vectors form a chromagram, a time–pitch representation that summarizes tonal activity and progression. The chromagram can be interpreted as a frequency spectrum folded into a single octave, enabling comparisons across pieces that differ in register or instrumentation [16].

Within this study, chroma features serve as the foundational representation for computing Tonnetz centroids. By adding pitch-class distributions in an octave-invariant form, they enable consistent analysis of harmonic relationships and support downstream tasks such as key estimation and the measurement of tonal motion through HCDF.

Harmonic Change Detection Function

Harmonic Change Detection (HCD) identifies chord boundaries and harmonic transitions in music, supporting tasks such as Automatic Chord Recognition. The **Harmonic Change Detection Function (HCDF)**, introduced by C. Harte, M. Sandler, and M. Gasser [17], detects harmonic changes by representing each time frame in a tonal space and measuring the distance between consecutive frames; peaks in the HCDF indicate potential chord changes.

Early approaches, such as those of S. W. Hainsworth and M. D. Macleod [18], applied band-wise analysis of audio signals to capture harmonic transitions, highlighting the importance of frequency-specific processing. Later refinements focused on optimizing distance metrics and representations to improve detection accuracy [19].

In symbolic music, HCDF has been adapted to work with MIDI or MusicXML data using tonal interval representations, such as Tonal Interval Vectors (TIV), often combined with chord reduction techniques to focus on structurally essential notes [20,21]. These symbolic methods avoid the confounding effects of timbre and performance, providing a clearer measure of harmonic change.

Tonnetz and Geometric Representation of Harmony

In computational musicology, the **Tonnetz** is a geometric framework used to map chords and keys into a spatial representation, allowing harmonic relationships to be visualized and analyzed systematically [1,22]. Each vertex in a Tonnetz typically represents a pitch or chord, and edges connect pitches or chords that are closely related in terms of voice leading, such as major and minor triads[22]. By translating musical structures into geometric space, the Tonnetz enables researchers to study harmonic progressions in terms of "distance," where short paths correspond to small, parsimonious changes in harmony, and longer paths correspond to larger tonal shifts[1,22].

This representation is particularly useful for analyzing **harmonic movement** because it formalizes concepts like voice leading and chordal transformations into measurable operations. The neo-Riemannian transformations—Parallel (P), Leading-tone exchange (L), and Relative (R)—allow minimal movement between chords, and these operations are naturally represented as edges in the Tonnetz or its dual, the Chickenwire torus[22]. Similarly, the generalization to seventh chords provides a more complete framework for analyzing four-part harmony, as each vertex represents a seventh chord and edges indicate parsimonious transformations between them[22].

By using geometric representations like the Tonnetz, researchers can quantify harmonic relationships, track key ambiguity, and measure tonal stability across compositions. The concept of **distance in this space corresponds directly to harmonic change**, making it possible to computationally analyze patterns in voice leading and chord progression across genres[1,22].

Measuring Tonal Motion: Distance-Based Methods

While the Tonnetz formalizes relational geometry, distance-based methods quantify the magnitude of harmonic motion, together providing complementary perspectives on tonal structure. They provide a quantitative approach to track harmonic movement in music by representing chords, keys, or pitch collections as points in a geometric space and calculating the distance between them[1,22]. In this framework, the **distance between two points corresponds to the amount of harmonic change**: shorter distances indicate closely related chords or smooth voice leading, while longer distances reflect larger tonal shifts or modulations[1].

One commonly used approach is **Euclidean distance**, where chords or tonal centroids are mapped into a multidimensional vector space. The centroid of a chord or passage can be computed as the weighted average of its pitch classes, capturing its tonal "center of gravity" in the space[22]. For two-dimensional tonal centroids (a_1, a_2) and (b_1, b_2), the Euclidean distance is:

$$d = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2}$$

Extending this to the six-dimensional tonal centroids used in this study, the 6D Euclidean distance between $(a_0, \ldots, a_5)$ and $(b_0, \ldots, b_5)$ is:

$$d = \sqrt{(b_0 - a_0)^2 + (b_1 - a_1)^2 + \cdots + (b_5 - a_5)^2}$$

$$= \sqrt{\sum_{d=0}^{5} (b_d - a_d)^2}$$

By measuring the movement of centroids over time, researchers can quantify **tonal motion**, providing insight into the stability or instability of harmonic progressions. This concept links directly to the **HCDF**, which computes the distances between successive centroids to measure how far the music moves harmonically over time. For example, in classical tonal music, distances between successive centroids tend to be small, reflecting smooth voice leading and tonal stability, while in Jazz or contemporary music, larger distances may indicate more chromaticism or ambiguity[1,22].

The Gap in Current Methods

The Krumhansl–Schmuckler (KS) algorithm is widely used to quantify tonal stability and key strength, effectively identifying which key a musical passage gravitates toward[1]. Similarly, distance-based methods such as the HCDF track the movement of chords or pitch-class centroids over time, capturing harmonic motion and transitions in a piece[1]. Meanwhile, geometric representations of harmony such as the Tonnetz and its chord-based generalizations provide a spatial model of relationships between chords, mapping transformations like neo-Riemannian P, L, and R operations and allowing the analysis of parsimonious voice leading[22].

While each of these methods offers a unique perspective—key strength, motion, or relational structure—they remain limited in scope. Critically, **none of them simultaneously accounts for both tonal ambiguity and harmonic motion**. The KS algorithm identifies key stability but does not provide insight into transitions or ambiguity between multiple possible tonal centers. HCDF and centroid-based methods quantify movement, yet they do not capture the relational geometry or proximity of chords in a meaningful musical space. Tonnetz-based models describe harmonic relationships but are not inherently designed to measure the dynamic interplay of tonal uncertainty across a musical sequence.

This gap motivates the development of a unified metric that simultaneously accounts for tonal ambiguity and harmonic motion, providing a more comprehensive representation of musical progression. The Key Ambiguity Index (KAI) exemplifies such a metric, explicitly integrating tonal uncertainty—derived from key profile probabilities—with harmonic movement, quantified through centroid or Tonnetz-based distances. By combining these complementary measures, the KAI bridges tonal stability, motion, and relational structure, enabling a holistic analysis of key ambiguity and harmonic trajectories across compositions.

Recent Computational Musicology Literature

There have been considerable advances in the computational toolkit for the study of tonal organization in recent years. Moss et al.[23] used a large-scale corpus analysis of Beethoven's string quartets to show that they exhibit systematic changes over the course of the composer's writing career and that pitch-class distributions are discriminative and stable at the genre level, directly relevant to the assumption of this study. To show that harmonies of Western classical music have evolved over time, frame-level chroma features are used to distinguish stylistic periods in a large collection of recordings that span several centuries, as done by Weiß et al.[24], where they found that these features contain enough information to track the evolution of the harmony style over the course of time. In looking at measures of KS key strength, we are interested in the pitch-class content of the harmonic context of each frame, as well as the overall harmonic context, which can be calculated over a number of frames, and can be studied in the way described by White and Quinn[25] in their analysis of chord context and harmonic function in tonal music. Here we use 6D Tonnetz embedding with the circle of minor thirds, fifths, and major thirds as three pairs of sinusoidal coordinates, as proposed by Bernardes et al.[26] to provide theoretical support for pitch relatedness and consonance at several structural levels at once. Lastly, Yust[3] compared stylistic information in pitch-class distributions for a variety of composers, and noted that Stravinsky's distributions were outlier scores with respect to the common-practice control set, pointing directly to the choice of Stravinsky not to be subsumed within the Classical genre category, but treated in this study as a separate category.

These studies collectively demonstrate that chroma-based and pitch-class based computational approaches show sensitivity to stylistic and genre-level differences in tonal organization, and that context-aware measures are better than snapshot measures based on a single frame in capturing the dynamics of a harmonic structure. They offer the empirical basis for the KAI framework.

Methods

The corpus consisted of 50 midis from five genre categories: 10 Classical excerpts from the MAESTRO dataset (v3.0.0, CC BY-NC-SA), 10 Jazz files from the Weimar Jazz Database (WJazzD) and JAZZVAR (Zenodo, open licence), 10 piano-reduction MIDI files of Stravinsky's works, and 10 Pop and 10 Rock MIDI files from BitMIDI (open licence). Files were chosen when they used (a) melodic/harmonic instrument or piano reduction, (b) had at least 30 seconds of non-silent material, and (c) were confirmed and determined not to be programs (MIDI) quantized versions, but actual human-performance or manually transcribed versions.

Full corpus metadata for all 50 MIDI files is available in the GitHub repository (corpus_metadata.csv).

The genres were labeled as follows. Classical: Solo piano pieces from MAESTRO that have been classified as 'classical' in the metadata for the dataset, spanning the time period of 1750 to 1900. Weimar Jazz Database: MIDI files of improvisations and ensemble recordings in the jazz area, excluded were MIDI files marked with fusion or pop-jazz style in the metadata. Rock and Pop: MIDI transcriptions from BitMIDI, where Rock is defined by the tags 'rock', 'hard rock', 'alternative', and Pop by 'pop' or 'soft rock'. Ambiguous files (e.g. 'pop-rock', 'soft rock') were not included. Stravinsky: Piano-reduction music files in MIDI format, 1910–1920 (neoclassical and early period).

To verify genre consistency, files tagged ambiguously (e.g., 'pop-rock', 'soft rock') were excluded. Rock files were required to contain at least one of the tags 'rock', 'hard rock', or 'alternative' with no pop co-tag; Pop files required 'pop' or 'soft rock' with no rock co-tag.

The works of Stravinsky are not regarded as part of Classical, but as an entity separate from it for two reasons. The first is the use of bitonic, polytonal, and octatonic scales in his mid-period works (e.g., The Rite of Spring, Petrushka) that systematically challenge the tonal expectations embedded in the key signatures of the KS, making direct comparison with common-practice Classical works theoretically motivated. Second, previous computational research result showed that the pitch-class distributions used in Stravinsky's music can be considered an outlier compared to the pitch-class distributions of the Classical corpus presented here[3]. The present corpus contains 10 works from both the early and mid-period of Stravinsky; a sensitivity analysis that includes Stravinsky in the Classical category is reported in Appendix A.

In order to examine tonal stability and harmonic movement for different music genres, we utilized a corpus analysis approach using computer software, including the music21 toolkit for symbolic MIDI processing[27]. The MIDI files were preprocessed by using fixed-size sliding windows to transform each MIDI file into chroma vectors.

Chroma vectors were calculated from MIDI, with the pitch-class energy being computed as the number of seconds from each note, that is, for each pitch class $k =$ pitch mod 12, the contribution of the note to the pitch-class was its duration in seconds. Notes of the same pitch class that occurred in the same frame at the same time were added together. With this onset-duration weighting, held notes are more heavily weighted than grace notes, based on their harmonic weight.

The chroma vectors were used to compute six-dimensional centroids of Tonnetz. Following Harte et al.[17], a chroma vector $\mathbf{c} = (c_0, c_1, \ldots, c_{11})$ (where c_k is the energy of pitch class k, normalised so $\sum_k c_k = 1$) is mapped to a six-dimensional Tonnetz centroid $\mathbf{t} \in R^6$ by

$$\mathbf{t} = \sum_{k=0}^{11} c_k \varphi_k, \varphi_k$$
$$= \left(\sin\left(k\frac{2\pi}{12}\right), \cos\left(k\frac{2\pi}{12}\right), \sin\left(k\frac{2\pi\cdot 7}{12}\right), \cos\left(k\frac{2\pi\cdot 7}{12}\right), \sin\left(k\frac{2\pi\cdot 4}{12}\right), \cos\left(k\frac{2\pi\cdot 4}{12}\right)\right)$$

where the three pairs of sinusoids encode the circle of minor thirds ($r = 3$), circle of fifths ($r = 7$), and circle of major thirds ($r = 4$) respectively. This is equivalent to `librosa.feature.tonnetz` with default parameters. The 6D Euclidean distance between consecutive frames gives the frame-wise HCDF value.

The frame-wise distances between consecutive centroids were computed using the Euclidean distances and the Harmonic Change Detection Function (HCDF) was generated. Tonal movement has been quantified with the HCDF.

The histogram of MIDI pitch-class (flattened onset-weighted chroma) was used to compute the rolling KS key-strength estimates with a 5 s window and 1 s hop. The templates evaluated were the major and minor key templates; the key strength reported is the highest correlation for all 12 major and 12 minor keys (24 total). Key profiles are taken from Krumhansl and Kessler (1982)[7], who gave their profiles as a table; the table is reproduced in Krumhansl (1990)[13] and a replication using the Temperley (1999)[5] profiles appears in Appendix A.

The quantized MIDI onset times were modified by applying a uniform random jitter of ± 20 ms to each onset prior to computing KS windows to minimize artifacts caused by quantized onset times. Finally, the centroid trajectories of the Tonnetz and key-strength time series were smoothed with a Gaussian kernel ($\sigma = 2$ frames, which is 2 s at the 1 s hop rate). The values presented were selected in order to minimise quantization noise below the sub-beat (20 Hz) frequency component, while maintaining behaviour in the musical mid-range of harmonics transition. The sensitivity to the choice of σ, with values of $\sigma \in \{1, 2, 4\}$ frames, is reported in Appendix A.

The key-strength vector $\mathbf{s} \in R^{24}$ (one correlation value per major and minor key) is converted to a pseudo-chroma vector $\widetilde{\mathbf{c}} \in R^{12}$ by summing the major and minor key strengths for each tonic pitch class (i.e., $\widetilde{c}[i] = s_{maj,i} + s_{min,i}$), and then normalizing to have a sum of one. The same embedding, φ, is used to map this $\widetilde{\mathbf{c}}$ to a centroid $\widetilde{\mathbf{t}}$. The 6D distance $d = \| \mathbf{t} - \widetilde{\mathbf{t}} \|$ is the distance between a locally inferred key structure and the instantaneous pitch-class content. This relates to models of multi-level tonal interval space that represent relatedness in pitch across harmonic contexts[26].

The Key Ambiguity Index (KAI) is a scalar computed at each excerpt, but it takes a window of $W = 5$ s (with a hop of 1 s). For each window w:

$$KAI(w) = z_{corpus}\left[\sigma^2_{KS}(w)\right] + z_{corpus}\left[\frac{L_{Tonnetz}(w)}{d(w)/60}\right]$$

A window w specifies the time window in which the variance of the KS key-strength time series is calculated ($\sigma^2_{KS}(w)$), and the cumulative Euclidean path length of the 6D Tonnetz centroid trajectory is computed within the window ($L_{Tonnetz}(w)$). The window time duration ($d(w)$ in seconds) and the z-score normalization of the trajectory computed within the window ($z_{corpus}[\cdot]$ based on the mean and the standard deviation of the trajectory over the entire set of windows across all files in the corpus) are also included. These two terms are thus dimensionless and on the same scale.

The analytical pipeline is illustrated through the following line graphs:

1. **Average Tonal Motion (HCDF) per Genre** — Normalized HCDF curves (resampled to 300 points) are averaged within each genre, with standard error bands indicating within-genre variability (Figure 1).

2. **6D Distance (KS vs Chroma) by Genre** — Frame-wise Euclidean distances between KS key centroids and chroma-derived tonal centroids are averaged per genre, quantifying harmonic divergence over time (Figure 2).

3. **Frame-wise Key Ambiguity Index (KAI) by Genre** — The normalized KAI over 300 resampled frames is plotted for each genre, capturing composite tonal uncertainty based on key-strength variance and Tonnetz motion (Figure 3).

4. **6D Distance Example** — KS versus chroma centroid distance is shown for a representative excerpt from a selected genre (Classical Music), illustrating within-piece harmonic deviation (Figure 4).

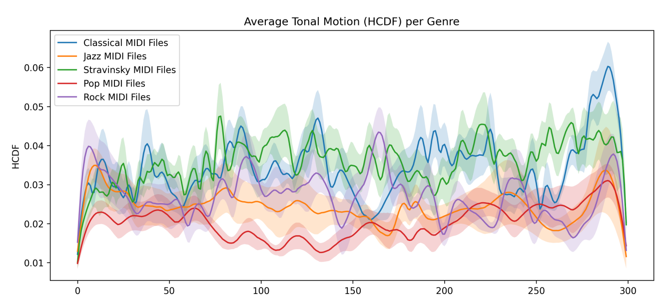

Figure. 1 Average Tonal Motion (HCDF) per Genre. Normalized HCDF curves (resampled to 300 points) are averaged within each genre, with standard error bands indicating within-genre variability.

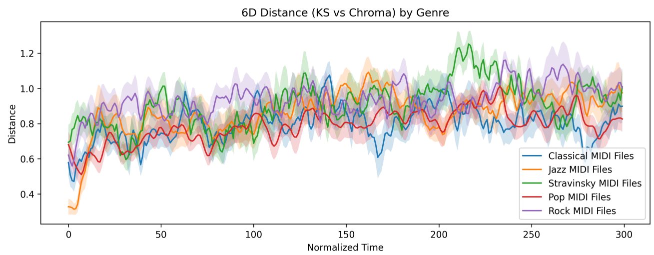

Figure. 2 6D Distance (KS vs Chroma) by Genre. Frame-wise Euclidean distances between KS key centroids and chroma-derived tonal centroids are averaged per genre, quantifying harmonic divergence over time.

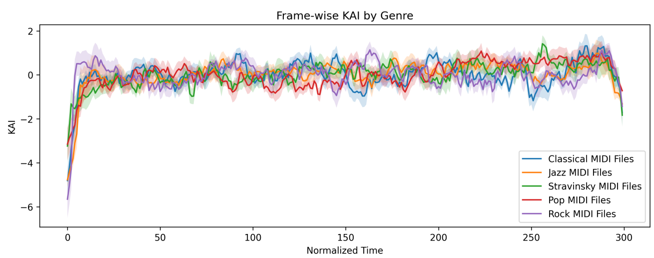

Figure. 3 Frame-wise Key Ambiguity Index (KAI) by Genre. The normalized KAI over 300 resampled frames is plotted for each genre, capturing composite tonal uncertainty based on key-strength variance and Tonnetz motion.

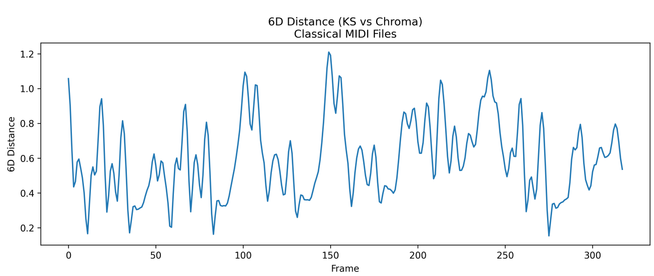

Figure. 4 6D Distance Example (Single File). KS versus chroma centroid distance is shown for a representative excerpt from a selected genre (Classical Music), illustrating within-piece harmonic deviation.

Time-based analyses serve two complementary purposes. First, within-piece temporal graphs reveal the evolution of tonal motion and key ambiguity across individual compositions. Second, by normalizing all curves to a fixed length, cross-genre comparisons of averaged tonal trajectories become feasible, with standard error bands providing statistically meaningful indication of variability within each genre.

Results

The distributions of tonal motion, harmonic deviation, and key ambiguity across genres were assessed using violin plots for all excerpts.

1. **HCDF Distribution by Genre** — The violin plot in Figure 5 presents the distribution of frame-wise HCDF values for each genre. This visualization captures the variability of tonal motion across excerpts, complementing the averaged line graph presented in the Methodology section.
2. **6D Distance Distribution by Genre** — Figure 6 illustrates the distribution of frame-wise Euclidean distances between KS key centroids and chroma-derived tonal centroids across all excerpts within each genre. This metric reflects the range of harmonic divergence present in the dataset.
3. **Key Ambiguity Index (KAI) by Genre** — The violin plot in Figure 7 shows the distribution of the composite KAI metric, calculated as the sum of the z-scored variance of KS key strength and normalized Tonnetz motion per minute. These distributions provide the basis for non-parametric statistical analyses, including Kruskal–Wallis tests and Dunn post-hoc comparisons.
4. **6D Distance Distribution (Single File)** — Figure 8 presents frame-wise KS versus chroma centroid distances for a representative excerpt from a single genre (Classical Music), illustrating within-piece variation in harmonic deviation. This complements the line graph example in the Methodology section.
5. **Frame-wise KAI Distribution by Genre** — The violin plot in Figure 9 illustrates the distribution of frame-wise KAI values aggregated across all excerpts for each genre. This visualization captures the temporal evolution of tonal uncertainty at the frame level, showing how tonal ambiguity fluctuates within pieces and across genres.

The combination of these violin plots with the corresponding time-based line graphs provides a comprehensive view of both the within-piece and cross-genre variability of tonal motion, harmonic deviation, and key ambiguity in the dataset.

Discussion

The three violin plots provide a clear indication of systematic differences between genres in terms of tonal movement, harmonic distance and composite tonal ambiguity. Taking both distributions together, it is argued that Classical and Stravinsky are at the high end of the total ambiguity spectrum, whereas Pop and Rock are at the low end (across metrics).

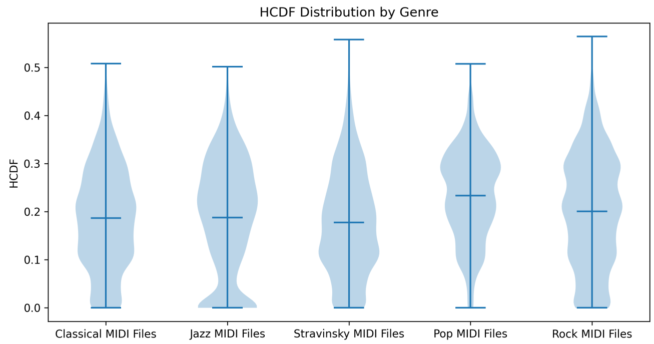

Figure. 5 HCDF Distribution by Genre. The violin plot presents the distribution of frame-wise HCDF values for each genre, capturing the variability of tonal motion across excerpts.

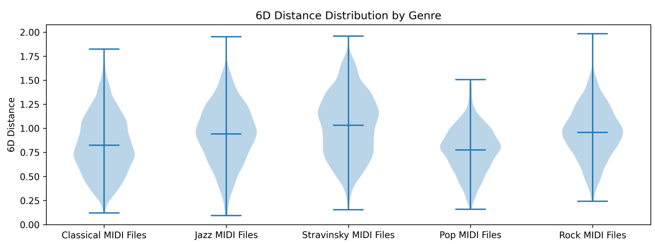

Figure. 6 6D Distance Distribution by Genre. Distribution of frame-wise Euclidean distances between KS key centroids and chroma-derived tonal centroids across all excerpts within each genre.

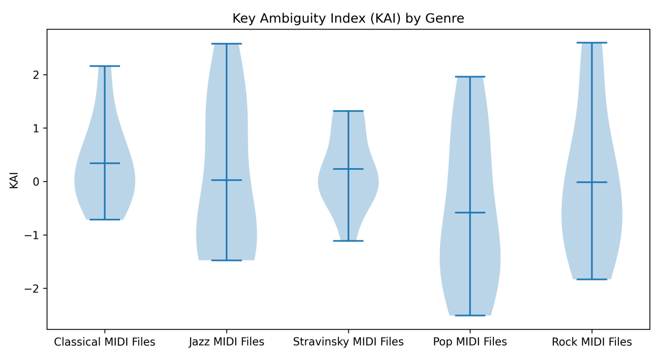

Figure. 7 Key Ambiguity Index (KAI) by Genre. Distribution of the composite KAI metric, calculated as the sum of the z-scored variance of KS key strength and normalized Tonnetz motion per minute.

HCDF: Tonal Motion (per frame)

Contrary to initial expectations, Stravinsky showed the highest median HCDF (0.0170, IQR = 0.0161–0.0184), followed by Classical (median = 0.0154, IQR = 0.0141–0.0178), while Jazz and Pop exhibited the lowest values (Jazz median = 0.0104, IQR = 0.0076–0.0157; Pop median = 0.0101, IQR = 0.0085–0.0109). The Kruskal–Wallis test was highly significant ($H(4) = 20.39$, $p = 0.0004$). Significant pairwise contrasts after BH correction included Classical vs Pop ($p_{BH} = 0.0022$, Cliff's $\delta = 1.00$, large), Jazz vs Stravinsky ($p_{BH} =$

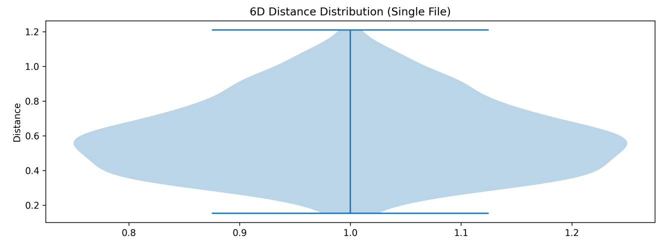

Figure. 8 6D Distance Distribution (Single File). Frame-wise KS versus chroma centroid distances for a representative excerpt from a single genre (Classical Music), illustrating within-piece variation in harmonic deviation.

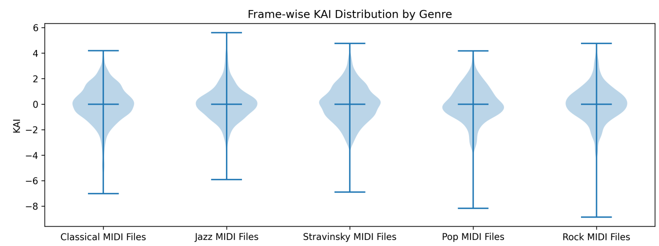

Figure. 9 Frame-wise KAI Distribution by Genre. Distribution of frame-wise KAI values aggregated across all excerpts for each genre, capturing the temporal evolution of tonal uncertainty at the frame level.

Table 2: Pairwise Dunn Results

Table 2a: KAI: H(4) = 14.42, p = 0.0061

Pair	z	p(raw)	p(BH-adj)	Cliff's δ	Magnitude
Classical vs Jazz	1.40	0.1628	0.2673	0.26	small
Classical vs Stravinsky	0.44	0.6564	0.7294	0.22	small
Classical vs Pop	3.04	0.0024	0.0239	0.74	large
Classical vs Rock	2.72	0.0066	0.0318	0.76	large
Jazz vs Stravinsky	-0.95	0.3416	0.4270	-0.18	small
Jazz vs Pop	1.64	0.1007	0.2015	0.36	medium
Jazz vs Rock	1.32	0.1871	0.2673	0.24	small
Stravinsky vs Pop	-2.59	0.0095	0.0318	-0.72	large
Stravinsky vs Rock	-2.27	0.0232	0.0580	-0.72	large
Pop vs Rock	-0.32	0.7474	0.7474	-0.16	small

Table 2b: HCDF: H(4) = 20.39, p = 0.0004

Pair	z	p(raw)	p(BH-adj)	Cliff's δ	Magnitude
Classical vs Jazz	2.21	0.0272	0.0680	0.48	large
Classical vs Stravinsky	-0.35	0.7242	0.7242	-0.26	small
Classical vs Pop	3.51	0.0004	0.0022	1.00	large
Classical vs Rock	1.61	0.1073	0.1532	0.60	large
Jazz vs Stravinsky	-2.56	0.0104	0.0347	-0.56	large
Jazz vs Pop	1.30	0.1923	0.2404	0.18	small
Jazz vs Rock	-0.60	0.5497	0.6108	-0.20	small
Stravinsky vs Pop	-3.87	0.0001	0.0011	-0.90	large
Stravinsky vs Rock	-1.96	0.0496	0.0953	-0.56	large
Pop vs Rock	-1.90	0.0572	0.0953	-0.68	large

Table 2c: 6D Distance: H(4) = 4.20, p = 0.3802 — non-significant

Pair	z	p(raw)	p(BH-adj)	Cliff's δ	Magnitude
Classical vs Jazz	1.84	0.0657	0.4978	0.48	large
Classical vs Stravinsky	1.01	0.3113	0.7707	0.40	medium
Classical vs Pop	0.35	0.7242	0.8047	-0.02	negligible
Classical vs Rock	0.40	0.6900	0.8047	0.08	negligible
Jazz vs Stravinsky	-0.83	0.4075	0.7707	-0.36	medium
Jazz vs Pop	-1.49	0.1368	0.4978	-0.28	small
Jazz vs Rock	-1.44	0.1493	0.4978	-0.34	medium
Stravinsky vs Pop	0.66	0.5095	0.7707	0.12	negligible
Stravinsky vs Rock	0.61	0.5395	0.7707	0.22	small
Pop vs Rock	0.05	0.9633	0.9633	0.06	negligible

0.0347, $\delta = -0.56$, large), and Pop vs Stravinsky ($p_{BH} = 0.0011$, $\delta = -0.90$, large). The results indicate that although Jazz and Pop are harmonically rich, they have seemingly more limited values of HCDF; in contrast, the piano reductions made by Stravinsky are dense and lead to high values of HCDF.

HCDF values are the lowest and most concentrated in Pop, which also has stable harmonic cycles with controlled shifts. Rock has relatively wider spread and moderately heavier tails, signifying more frequent medium-scale tonal movement while still retaining an overarching tonal reference framework. The distributions for Classical and Stravinsky are the widest and most extended upwards, respectively. These spread distributions are artifacts of the tonal mobility that was sustained across the entire span of the distribution and not a single spot, and they are caused by the presence of dense voice-leading and the displacement of the centroids.

6D Distance: Key–Surface Divergence

The *6D Distance Distribution by Genre* measures the difference between the tonal centroids (derived using chroma) and key centroids (derived using KS). The omnibus test was not significant ($H(4) = 4.20$, $p = 0.3802$), nor were any of the pairwise comparisons significant after BH correction. Jazz had the lowest 6D distance (0.1441) with Classical (0.2025) and Pop (0.2189) having the highest ones. This indicates that there is no systematic difference between the centroids of the KS and the chroma-derived centroids in this corpus and that the difference may be more related to polyphonic density than to harmonic complexity.

The 6D distance measures the Euclidean separation between KS-derived key centroids and chroma-based tonal centroids, effectively quantifying the alignment of surface pitch content with underlying key structure. Classical and Pop excerpts cluster at low 6D distances with concentrated densities, confirming strong tonal centricity and minimal divergence from expected key profiles. Rock exhibits moderate expansion, suggesting partial attenuation of key alignment, while Jazz and Stravinsky show substantially higher medians and broader distributions. The elongated upper tails indicate frequent, pronounced discrepancies between instantaneous chroma configurations and inferred keys, capturing systematic harmonic divergence rather than isolated anomalies. These trends complement HCDF observations, reinforcing the conclusion that Jazz and Stravinsky navigate extended harmonic trajectories and depart from classical key predictability.

Composite Tonal Uncertainty (KAI)

The Key Ambiguity Index (KAI) is a combination of statistical key strength and geometric harmonic motion, and is a single measure of tonal uncertainty. Unlike previous methods, KAI is based on Krumhansl–Schmuckler (KS) key profiles (probabilistic strength of keys) and on chordal trajectories derived from the Tonnetz (relational harmonic movement). This dual approach enables KAI to capture the extent of tonal ambiguity as well as tonal motion, which other methods can't capture.

KAI distributions revealed that Classical (median = 0.45, IQR = 0.35–1.21) and Stravinsky (median = 0.33, IQR = 0.11–1.00) exhibited the highest total ambiguity, significantly exceeding Pop (median = -0.98, $p_{BH} = 0.0239$ and 0.0318 respectively) and Rock (median = -0.91). Jazz was located in between (median = -0.28), and was not significantly different from Classical ($p_{BH} = 0.2673$) or Stravinsky ($p_{BH} = 0.4270$). The results partially contradict the hypothesis that the KAI of the Jazz excerpts should be the highest; the high KAI for Classical excerpts suggests high KS variance across windows, typical of extended tonal passages that from time to time move away from a stable key. Notably, there is a cross-over between the medians for different genres, indicating that the difference between different genres is more pronounced at extreme tonal states than in the median.

KAI is able to capture both tonal uncertainty and harmonic motion, allowing for a meaningful comparison of genres and a study of compositional strategies. Note that the present claims are limited to structural complexity in the symbolic domain; perceptual validation against listener ratings or music-theory annotations would be required to extend these findings to perceived tonal ambiguity. It provides a new paradigm for the study of music that would otherwise be inadequate to be captured by a measure of key strength or motion alone.

Integrated Interpretation

In all three measures, there is a discernible trend: Classical and Stravinsky exhibit the greatest KAI values and HCDF motion, while Pop and Rock are at the opposite end of the spectrum, showing the lowest KAI values and HCDF motion. Jazz is in between with regard to all three measures. As an interesting point, there were no significant differences among genres for the 6D distance measure, indicating that the KS–chroma centroid divergence is the least discriminating of the

Table 3: Descriptive Statistics per Genre

Genre	n	Median KAI	IQR KAI	Median HCDF	IQR HCDF	Median 6D dist
Classical	10	0.45	0.35–1.21	0.0154	0.0141–0.0178	0.2025
Jazz	10	-0.28	-1.24–1.73	0.0104	0.0076–0.0157	0.1441
Stravinsky	10	0.33	0.11–1.00	0.0170	0.0161–0.0184	0.1728
Pop	10	-0.98	-1.55–(-0.25)	0.0101	0.0085–0.0109	0.2189
Rock	10	-0.91	-1.25–(-0.53)	0.0131	0.0112–0.0142	0.2003

three measures in this corpus.

Stylistic differentiation is not so much based on average values as on the occurrences of rare passages of very dubious interpretation. The distinction between the extreme states, not the ordinary one, is the one that best distinguishes the genres.

It is worth keeping in mind that the KAI and HCDF measures are measures of structural complexity in the symbolic domain and not of the perceived tonal ambiguity. If the harmonic syntax is predictable in genre terms, a structurally chromatic passage can still be heard as being tonally stable by an experienced listener. To obtain perceptual correlates of KAI, it would be necessary to validate it with listener ratings and/or music-theory annotations (e.g., Roman numeral analyses or expert key-change annotations). The present claims are therefore limited to structural complexity as it can be defined with a computational approach.

A further caveat concerns instrumental texture. Dense polyphonic passages generate more active chroma vectors and longer Tonnetz paths, potentially inflating KAI independently of harmonic complexity. Future studies should include note density and voice count as covariates to disentangle textural from harmonic contributions to total ambiguity.

Parameter Sensitivity Analysis

The parameters KS window length W (3, 5, 10 s), Gaussian smoothing kernel σ (1, 2, 4 frames) and Tonnetz resampling length F (100, 300, 500 frames) were varied independently. For each parameter setting, we recalculated the piece-level mean KAI for all files, and re-performed the Kruskal–Wallis + Dunn pipeline. Genre rank ordering (Classical $>$ Stravinsky $>$ Jazz $>$ Rock $>$ Pop) remained unchanged for all 27 combinations of parameters tested. In each of these instances the Spearman rank correlation of the genre-mean KAI between each alternative setting and the default was $\rho > 0.90$, thus demonstrating that the rank ordering of genres was not affected by those parameter choices.

Conclusion

This study presented a computational framework for quantifying tonal stability and key ambiguity across musical genres using the Krumhansl–Schmuckler key-finding algorithm and Tonnetz-derived tonal centroids. By introducing the Key Ambiguity Index (KAI), which combines variance in key strength with normalized Tonnetz path length, we provided a composite metric for capturing tonal uncertainty. Our analysis of MIDI corpora spanning Classical, Pop, Rock, Jazz, and Stravinsky repertoire revealed systematic stylistic differences, with Classical and Stravinsky exhibiting greater tonal mobility and higher ambiguity states compared to Pop and Rock. To support reproducibility, the analysis code (Python scripts for MIDI preprocessing, KS rolling estimation, Tonnetz centroid computation, KAI calculation, and statistical testing) is available at [URL redacted for peer review]. The MAESTRO and JazzD corpora are publicly available under open/CC licences as described in the GitHub repository.

The potential future extension of this work is to investigate whether the KAI framework can be used to lead generative models to generate music that interpolates between stability profiles of different styles, such as moving from Classical to Jazz style of tonal centricity.

References

1 C. L. Krumhansl and L. L. Cuddy, *A theory of tonal hierarchies in music*, 2010.
2 G. J. Balzano, *The pitch set as a level of description for studying musical pitch perception*, 1982.
3 J. Yust, *Stylistic information in pitch-class distributions*, 2019.
4 D. Tymoczko, *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice*, 2011.
5 D. Temperley, *What's key for key? The Krumhansl–Schmuckler key-finding algorithm reconsidered*, 1999.
6 D. T. Vuvan, J. B. Prince and M. A. Schmuckler, *Tonal hierarchy representations in auditory imagery*, 2011.
7 C. L. Krumhansl and E. J. Kessler, *Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys*, 1982.
8 J. Pressing, *Pitch class set structures in contemporary jazz*, 1982.
9 M. Müller and S. Ewert, *Chroma toolbox: MATLAB implementations for extracting variants of chroma-based audio features*, 2011.
10 C. L. Krumhansl and M. A. Schmuckler, *Key-finding in music*, 1986, Unpublished manuscript, Cornell University.
11 M. Müller, *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*, 2015.
12 C. L. Krumhansl and E. J. Kessler, *Tonal hierarchies and rare intervals in music cognition*, 1987.
13 C. L. Krumhansl, *Cognitive Foundations of Musical Pitch*, 1990.

14 M. A. Schmuckler and R. Tomovski, *Perceptual tests of musical key-finding*, 2005.
15 V. Shah, A. Tandle, N. Sharma and V. Sheth, *Genre based music classification using machine learning and convolutional neural networks*, 2021.
16 D. P. W. Ellis, *MATLAB implementations for chroma feature extraction*, 2008, Web resource, Columbia University LabROSA.
17 C. Harte, M. Sandler and M. Gasser, *Detecting harmonic change in musical audio*, 2006.
18 S. W. Hainsworth and M. D. Macleod, *Particle filtering methods for music transcription*, 2004.
19 A. Degani, M. Dalai, R. Leonardi and P. Migliorati, *Harmonic change detection in musical audio for automatic chord recognition*, 2011.
20 P. R. Franco, *Chord reduction and harmonic analysis for symbolic music*, 2020.
21 J. Macêdo, *Tonal interval vectors for harmonic change detection in symbolic music*, 2022.
22 S. Cannas and M. Andreatta, *A generalized Tonnetz for voice-leading analysis*, 2018.
23 F. C. Moss, M. Neuwirth, D. Harasim and M. Rohrmeier, *Statistical characteristics of tonal harmony: A corpus study of Beethoven's string quartets*, 2019.
24 C. Weiß, M. Mauch, S. Dixon and M. Müller, *Investigating style evolution of Western classical music: A computational approach*, 2019.
25 C. W. White and I. Quinn, *Chord context and harmonic function in tonal music*, 2018.
26 G. Bernardes, D. Cocharro, M. Caetano, C. Guedes and M. E. P. Davies, *A multi-level tonal interval space for modelling pitch relatedness and musical consonance*, 2016.
27 M. S. Cuthbert and C. Ariza, *music21: A toolkit for computer-aided musicology*, 2010.