Estimating keys and modulations in musical pieces

Laurent Feisthauer, Louis Bigo, Mathieu Giraud, Florence Levé

Estimating keys and modulations in musical pieces

Laurent Feisthauer[1] **Louis Bigo**[1] **Mathieu Giraud**[1] **Florence Levé**[2,1]
[1] Université de Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France
[2] Université de Picardie Jules-Verne, MIS, F-80000 Amiens, France
{laurent,louis,mathieu,florence}@algomus.fr

ABSTRACT

Modulations, the moments where key change, are structurally important in tonal music. Analyzing music, especially studying large-scale music structure of a piece, often implies to look for modulations. State-of-the-art key-finding algorithms generally aim at identifying keys rather than studying the way they change. Here, we introduce new ways to model modulations with the help of features based on musicological knowledge, as well as an algorithm estimating the *tonal plan* of a piece. We study the concept of current diatonic pitch set and introduce a heuristic to detect dominant-to-tonic progressions. We design three proximity measures to assess how close the music is from each key. These measures are then combined by an algorithm that identifies an optimal tonal plan. We report results on a corpus including 38 movements from Mozart's string quartets, obtaining a 84.8% prediction of correct keys with insight on where the modulations occur.

1. INTRODUCTION

1.1 Tonality and Modulation

How do pitches relate to each other during a musical piece? Are some pitches more stable than others? Do some pitches attract some others? Tonality consists of a hierarchical arrangement of pitches, referred to in this context as *scale degrees*. This hierarchy implies specific relations between these pitches like stabilities, attractions as well as harmonic directedness.

The principle of tonality appeared during the baroque era, in the 17th century, emerging from modality. Synthesizing and expanding concepts from earlier theorists (Zarlino, Mersenne, Sauveur), Rameau developed a comprehensive theory on tonality, in his *Traité de l'harmonie réduite à ses principes naturels* [1]. The term "tonality" however appeared first in the *Dictionnaire historique des musiciens, artistes et amateurs, morts ou vivans* written by French music theorist Choron in 1810 [2]. Choron states that "modern tonality" only includes two modes, the Major one and the minor one. He also indicates the scales associated with these modes. Theories on tonality were considerably deepened across the XVIIIth and XIXth century by music theorists who were often composers themselves, such as Riemann. Tonality is still largely used nowadays in Western music, notably in popular music.

Laurent Feisthauer, Louis Bigo, Mathieu Giraud, Florence Levé, *Estimating keys and modulations in musical pieces*, Sound and Music Computing Conference (SMC 2020), Torino, 2020, pp. 323–330

Tonal musical pieces usually start and end in the same *key*, referred to as the *main key* or *global key*. However, composers tend to switch to alternative keys in order to create contrast, emotions, or simply give a new direction to the musical rhetoric. The moment when the key changes during the piece is called a *modulation*.

Modulating is generally considered to be a delicate musical task. The arrival of the new key must be "prepared" in order to smooth the perception of the change while conserving its novelty. Some theorists have proposed "good" and "bad" ways to modulate. Among other composers, Max Reger wrote 100 melodic and harmonic progressions that modulate, mainly from $C\ Major$ or $A\ minor$ to each of the other keys.

Identifying keys and modulations is generally one of the first tasks of the music analyst when confronted with a new score. Indeed, key changes are often associated with section boundaries defining the musical structure. The *sonata form* is a well-known example of this phenomenon where the two themes forming the *exposition* are in a different key. The modulation between these two themes, often accompanied by a specific cadence called the *medial caesura*, provides an outline of its structure [3].

1.2 MIR Research on Key Finding in Symbolic Scores

Researchers in the Music Information Retrieval (MIR) community have designed a number of key finding algorithms in symbolic scores. In 1971, the approach of Christopher Longuet-Higgins [4] consisted in progressively eliminating the keys which do not include the notes of the beginning of the score until one last key remained. Some decision rules are added to help with the most difficult cases. This algorithm was able to retrieve the key of the fugues of Bach's *Well-tempered Clavier*. A similar approach has later been used by Vos and van Geenen [5].

One of the most popular approach is the use of *pitch profiles*, first introduced by Krumhansl and Kessler in 1982 [6]. A pitch profile is an histogram that states the prevalence of the different pitches in a key. Later, Krumhansl and Schmuckler proposed a key finding algorithm that estimates the correlation between the pitch profile and the histogram of pitches extracted from the musical score [7], whereas Lerdahl introduced tonal pitch space [8] that can be both used as a pitch profile for key finding algorithms – but also as a way to compare chords and keys.

The use of different pitch profiles became a standard approach to estimate the global key of a piece. Such methods were experimented upon and improved by Temperley [9], Aarden [10], Bellman [11], Haas et al. [12], Sapp [13], Albrecht and Shanahan [14], and more recently by Nápoles López et al. [15].

Machine learning approaches, especially those including neural networks, have gained considerable popularity these last years in MIR researches including key finding. For instance, both Chen et al. and Micchi et al. [16–18] estimate local keys, among other musical attributes, from a per beat tonal analysis of Beethoven Piano Sonatas.

These algorithms are fairly accurate when it comes to detecting global keys and possibly long term local keys. However, they are not designed to precisely identify *where the modulations occur* and what actually makes the modulations, with the notable exception of work by Chew [19]. Chew establishes a distance from a set of pitches to a key using the spiral array introduced in [20]. The tonal plan is then modeled as a sequence of boundaries dividing the score in successive keys. It minimizes the distance with pitches included in the corresponding segments. The complexity of the approach increases with the number of modulations.

1.3 Contents

We present here an original approach to *model key change* with the help of musicological knowledge-based features used by an algorithm estimating the tonal plan of a piece.

We introduce the *current diatonic pitch set* as well as *a heuristic to detect $V \rightarrow I$ progressions*. Based on these concepts, we design two proximity measures to assess how close the music is from a given key. We also design a measure assessing the *smoothness* of a modulation from one key to another (Section 2). We then propose a way to combine these three measures to identify keys and modulations (Section 3). We finally present our results on a corpus of Mozart's string quartets. We analyze further results on the *Presto* of the third quartet (K 157), which we denote K157.3, and give perspectives for future research on key and modulation detection (Sections 4 and 5).

2. THREE MEASURES FOR A MODULATION

What aspects of the music make the listener – and the analyst – feel that the key is modulating? A first signal is when one of the notes does not "match" the current tonality. This note can be generally identified on the score by the presence of an *accidental*, generally notated with a flat, a sharp or a natural sign, or because it is a cancellation of a previous accidental. Such an accidental changes the scale currently heard. However, it does not always mean that the key has changed: It can also be a *nonchord tone* that ornaments the melodic line, such as in passing tones, neighbor tones, retardations, appoggiaturae, escape tones and anticipations.

A confirming signal of modulation is often found in some harmonic progression, such as an arrival on the new tonic or, even stronger, a dominant-tonic ($V \rightarrow I$) progression in the new key [21].

Finally, depending on the period, a composer will tend to favor certain keys relative to the main key, raising expectations of some modulations in the experienced listener or analyst. For the classical period, the keys that a composer tends to employ are generally associated with the scale degrees of the main key, prioritizing the dominant, or the relative major for a minor main key [21]. This is indeed the case for the Mozart quartets studied here (see, at the end of this paper, Table 3 and discusion).

We introduce thus three proximity measures to determine at a beat b "how far" we are from a key k. The first two measures look at previous pitches and chords, and the third one assesses the plausibility of modulation:

- How different are the pitches in the score around b from the pitches of the key k? (*pitch compatibility*)
- Did a stable harmonic sequence in key k occur in the near past of b ? (*harmonic anchoring*)
- Does a key k have a peculiar affinity or relationship to other potential previous key k' ? (*relationship proximity*)

These measures are designed for scores with *full pitch spelling* information. We strongly believe that when it comes to tonal music, it provides information not only about pitch but also about its function – "sharpening" or "flattening" a pitch is a thoughtful choice by the composer, important for the analysis. On data without such information, pitch spelling algorithms [22] could be applied first but the pitches which are the most difficult to spell are precisely the challenging ones for key and modulation estimation.

2.1 Pitch Compatibility

2.1.1 Current Diatonic Pitch Set

Each note sounding in a score confirms or modifies our perception of the current tonality. Suppose that we are used to hearing $C\natural$. When one hears a $C\sharp$, the perception of the diatonic scale may change. If we now hear only $C\sharp$, we are likely to think we are in a tonality including this pitch. Of course, approaches based on pitch profiles consider statistics on $C\natural$ and $C\sharp$, but they do not take into account the *directedness* of music: one or a few $C\sharp$ can alter our perception, no matter how many $C\natural$ there were before.

To model this phenomenon, we introduced the *current diatonic pitch set*, also called current diatonic scale, in [23]. We define $\mathcal{N}$ as the set of the seven pitch names:

$$\mathcal{N} = \{C, D, E, F, G, A, B\}$$

Figure 1. mm27-29 of Mozart's first movement of String Quartet No. 13 in $D\ minor$ (K173.1). This is a high modulating passage, to an unexpected $E\ minor$, just after the secondary thematic zone. Each bar presents a $V\!\rightarrow\!I$ progression, in $G\ minor$, $C\ Major$, then $B\ Major$ – even if they are only tonicizations and not full modulations.

The current diatonic pitch set $\mathcal{CS}(b)$ is a vector with 7 values built by associating to each of the pitch names in $\mathcal{N}$ the last encountered accidental ($\flat\flat, \flat, \natural, \sharp, \sharp\sharp$) on b or right before. It can thus have theoretically 5^7 different values, even if many of these values are not musically relevant. If one of the 7 pitches was not used before b, we consider the accidental suggested by the key signature of the piece. In Figure 1, the current diatonic pitch set at beat 2 of the measure 27 is thus

$$\mathcal{CS}(2) = \{C\sharp, D, E\flat, F\sharp, G, A, B\flat\}$$

(the $C\sharp$ being heard in a previous measure that is not shown here).

The current diatonic pitch set is evaluated at each beat. If two accidentals are encountered for the same pitch name between two beats, only the last one is considered.

2.1.2 Usual Diatonic Pitch Sets

Which pitch set do we expect in a given key? We define $\mathcal{S}(k)$ as follows:

- for the major keys, pitches of the usual major scale, as $\{C, D, E, F, G, A, B\}$ for $C\ Major$,
- for the minor keys, pitches of the minor harmonic scale, as $\{C, D, E\flat, F, G, A\flat, B(\natural)\}$ for $C\ minor$.

2.1.3 Diatonic Pitch Set Distance and Pitch Compatibility Measure

Given any two diatonic pitch sets $\mathcal{S}$ and $\mathcal{S}'$, we define the distance $d_{diat}(\mathcal{S}, \mathcal{S}')$ as the number of differently altered notes between them:

$$d_{diat}(\mathcal{S}, \mathcal{S}') = |\{n \in \mathcal{N} \text{ with } \mathcal{S}[n] \neq \mathcal{S}'[n]\}|$$

We evaluate the pitch compatibility between the notes of the score until beat b and some key k by comparing the alterations of the current diatonic pitch set $\mathcal{CS}(b)$ with the alterations of $\mathcal{S}(k)$, the usual diatonic pitch set of key k.

For each beat b and candidate key k, we will use as a proximity measure the value $d_{diat}(\mathcal{CS}(b), \mathcal{S}(k))$. For example, in Figure 1:

$$d_{diat}(\mathcal{CS}(2), \mathcal{S}(G\ minor)) = |\{C\}| = 1$$

$$d_{diat}(\mathcal{CS}(2), \mathcal{S}(C\ minor)) = |\{C, F, A, B\}| = 4$$

The measure captures that having recently heard an $F\sharp$ makes a $G\ minor$ more relevant than a $C\ minor$.

Figure 2 shows an example of an actual modulation. At the cadence in $D\ minor$, we have

$$\mathcal{CS}(b) = \{C\sharp, D, E, F, G, A, B\flat\},$$

and the same $\mathcal{CS}(b)$ is still found at measure 8. The $C\natural$ at measure 10 yields a scale which is exactly $\mathcal{S}(F\ Major)$:

$$\mathcal{CS}(c) = \{C, D, E, F, G, A, B\flat\},$$

2.2 Tonality Anchoring

2.2.1 Detection of $V\!\rightarrow\!I$ progressions

The most stable harmonic progression for a key is the one from the dominant (fifth scale degree, V) to the tonic (first scale degree, I). Such $V\!\rightarrow\!I$ progressions are a good confirmation that a modulation occurs. However, a $V\!\rightarrow\!I$ progression can occur on another scale degree than the first one (what is called a *tonicization*, as in Figure 1). Due to tonicization, modulations can thus not be solely based on $V \rightarrow I$, but such progressions are nevertheless important contributions to modulations.

Our heuristic to detect a $V \rightarrow I$ progression in key k is when there are at least two of the three following voice leadings:

① the third of V (also known as the *leading tone*) going to the tonic of I,

② the seventh of V going to the third of I,

③ the root from V going to the root from I.

In Figure 2, this heuristic identifies a $V\!\rightarrow\!I$ progression in $F\ Major$ with V occurring on the second beat of measure 10 and I on the first beat of measure 11.

Note that this heuristic produces false positives for *homonym* keys. In Figure 2 again, the $V\!\rightarrow\!I$ on the cadence on measure 8 is on a cadence in $D\ minor$, but the heuristic also falsely detects it as a $V\!\rightarrow\!I$ in $D\ Major$, because two of the three voice leadings characteristic of $D\ Major$ are found. The heuristic will also falsely consider $I \rightarrow IV$ movements as $V\!\rightarrow\!I$.

2.2.2 Tonality Anchoring Measure

We propose the proximity for some key k to be *harmonically minimal* if, on some beat b, a progression from the dominant to the tonic is occurring in this key. Otherwise, the proximity value should be as big as its distance from a $V\!\rightarrow\!I$ progression. Thus we define the measure $c_{V\rightarrow I}(b, k)$ using the distance in beats from the last $V\!\rightarrow\!I$ progression in the key k:

$$c_{V\rightarrow I}(b, k) = \begin{cases} 0 & \text{if the harmony at } b \text{ is the V or the I} \\ & \quad \text{of a } V\!\rightarrow\!I \text{ progression in } k \\ \min[c, c_{V\rightarrow I}(b-1, k) + 1] & \text{otherwise} \end{cases}$$

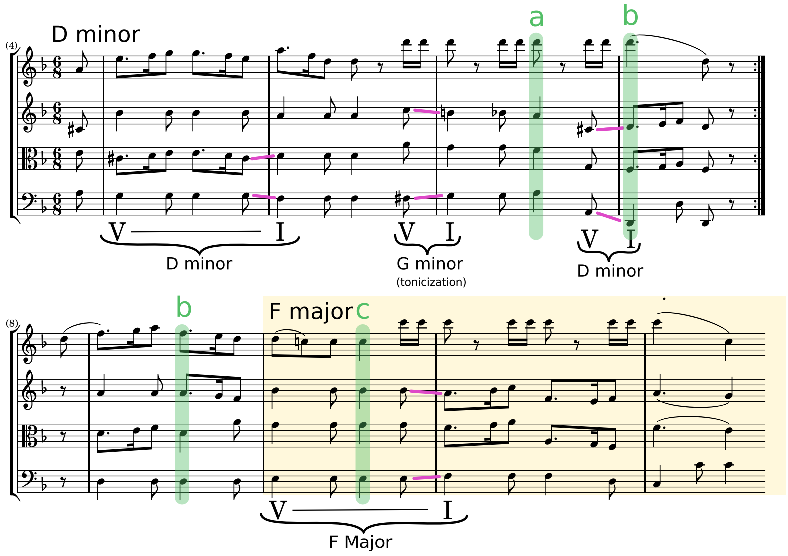

Figure 2. Measures 4 to 12 of the fourth movement of Mozart's String Quartet No. 15 (K421).

The movement is in $D\ minor$. There is a modulation in $F\ Major$ starting at measure 10. The $V \rightarrow I$ movements confirm both of these keys. However, there is also a $V \rightarrow I$ movement of a tonicization at the end of measure 6. The voice leadings used by the proposed heuristic are highlighted. The diatonic pitch sets a, b, and c are discussed in the text: The $C\natural$ of measure 10 changes the perception of the current diatonic pitch set and contributes to the modulation in $F\ Major$.

To avoid values that are too high, we bound this value by a constant $c = 20$. Note that the value 0 is given at the moment the V occurs, because it triggers the modulation.

2.3 Tonality Proximity

Music theorists have developed several systems drawing relationships between keys. The best known is probably the *circle of fifths*, introduced by Diletski [24] and refactored to its current form by Heinichen [25]. By assuming that the enharmonics are in fact the same notes (equal temperament hypothesis; $E\flat = D\sharp$) and starting from a given note, the sequence of 12 fifths in the same direction (ascending or descending) allows to go through all the notes of the chromatic scale then to return to the starting note.

While working on Mozart's string quartets, we noticed that Mozart often switched between Major and minor modes of the same key. The circle of fifths is not designed for this behavior, whereas it is present in the table of relationships between keys introduced in 1817 by Gottfried Weber in *Versuch einer geordneten Theorie der Tonsetzkunst* [26] (see Figure 3). This Weber's table is actually one of the spaces that Lerdhal deduces from his tonal pitch space framework [8].

We define $d_W(k, k')$, our measure of proximity between two keys k and k', as the Euclidean distance between those two keys in Weber's table:

$$d_W(k, k') = \min\left(\sqrt{|x_{k'} - x_k|^2 + |y_{k'} - y_k|^2},\ w\right)$$

where (x_k, y_k) and $(x_{k'}, y_{k'})$ are the coordinates of both keys in the table and $w = 10$ is a bounding constant.

As an example, $d_W(D\ minor, F\ Major) = 1$ whereas $d_W(D\ minor, C) = \sqrt{2}$. It is more common in the classical era to modulate from $D\ minor$ towards $F\ Major$, as in Figure 2, than towards $C\ Major$.

Note that other distances could also be defined on the Weber's table. Moreover, this table seems to be relevant mostly for the classical period. Modulations are more daring from the romantic era, favoring enharmonic and chromatic modulations. It would then be interesting to favor theories that allow switching from one key to another in very few chords, as for example those of Max Reger [27] or Tchaïkovski [28], as well as to compare several tonality spaces.

E♭	c	C	a	A	f♯	F♯	d♯	D♯	b♯
A♭	f	F	d	D	b	B	g♯	G♯	e♯
D♭	b♭	B♭	g	G	e	E	c♯	C♯	a♯
G♭	e♭	E♭	c	C	a	A	f♯	F♯	d♯
C♭	a♭	A♭	f	F	d	D	b	B	g♯
F♭	d♭	D♭	b♭	B♭	g	G	e	E	c♯
B♭♭	g♭	G♭	e♭	E♭	c	C	a	A	f♯

Figure 3. Extract of the *Table of the relationship of keys* from [26]. Major keys are uppercase, minor keys are lowercase.

3. ESTIMATING THE TONAL PLAN AND THE MODULATIONS BY COMBINING THE THREE MEASURES

We now describe a dynamic programming algorithm that uses the three measures described above to estimate the tonal plan of a piece. Computing a tonal plan $k_1 \ldots k_B$ of a piece of B beats requires to select a key k_b for each beat $b \in [1, B]$. The tonal plan returned by the algorithm optimally minimizes a combination of the three measures for the whole piece. We use here a simple weighted combination of the three measures. We associate to any tonal plan $k_1 \ldots k_B$ a cost:

$$
\begin{aligned}
D(k_1 \ldots k_B) \;\; = & \sum_{b\in[1,B]} [\alpha \cdot c_{V\to I}(b, k_b)/c \\
& \qquad + \beta \cdot d_{diat}(\mathcal{CS}(b), \mathcal{S}(k_b))/7] \\
& + \sum_{b\in[2,B]} \gamma \cdot d_W(k_{b-1}, k_b)/w
\end{aligned}
$$

where the divisions by c, 7, and w normalize each of the three measures to obtain values between 0 and 1. The weights α, β, and γ further ponder the relative importance of the three measures.

The optimal plan $k_1 \ldots k_B$ that minimizes $D(k_1 \ldots k_B)$ can be computed by dynamic programming, as optimizing $k_1 \ldots k_b$ can be done by combining an optimal partial plan $k_1 \ldots k_{b-1}$ with a key k_b. We consider that there are 42 possible keys that correspond to all the triplets:

$$\{C, D, E, F, G, A, B\} \times \{\sharp, \natural, \flat\} \times \{Major,\ minor\}$$

The algorithm thus builds an array D of size $B \times 42$. For a beat b and a candidate key k, the value $D(b, k)$ estimates the likelihood from this key k on the beat b, assuming that the tonal plan calculated until there is optimal. It is computed by D(1,k) = 0, and, when $b \geq 2$:

$$
\begin{aligned}
D(b, k) \;\; = & \;\; \alpha \cdot c_{V\to I}(b, k)/c \\
& + \beta \cdot d_{diat}(\mathcal{CS}(b), \mathcal{S}(k))/7 \\
& + \min_{k'} [\gamma \cdot d_W(k, k')/w \; + \; D(b-1, k')]
\end{aligned}
$$

When the $B \times 42$ values of $D(b, k)$ are computed, we look for the optimal plan $k_1 \ldots k_B$ in the table by backtracking:

- for the last beat b_B, choose the key k_B that minimizes $D(b_B, k_B)$;
- for the preceding beat b_{B-1}, choose the key k_{B-1} ; that minimizes $d_W(k_{B-1}, k') + D(b_{B-1}, k_{B-1})$
- repeat until the first beat b_1.

4. EVALUATION AND DISCUSSION

4.1 Corpus and Implementation

The corpus gathers 38 movements of Mozart's String Quartets, with manual annotation of keys and cadences. It is an enhancement of the corpus used and described in previous works on sonata forms [23, 29]. The corpus contains **kern files with full pitch spelling information, retrieved from KernScores [30], together with reference annotation data including keys, and is available at `www.algomus.fr/data`.

We implemented the algorithms within the Python music21 framework [31]. The beat granularity used is the quarter note for binary time signatures and the dotted quarter note for ternary time signatures. We consider this to be sufficient to model most of the harmonic rhythm in this repertoire.

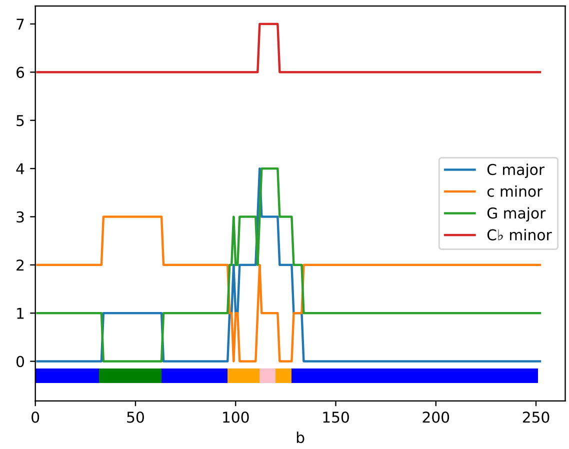

Figure 4. Pitch compatibility measure $d_{diat}(\mathcal{CS}(b), \mathcal{S}(k))$ along K157.3 for keys $k \in \{C\ Major, G\ Major, C\ minor, C\flat\ minor\}$. The thick bar at the bottom shows the keys in the reference annotation.

4.2 Discussion on the Measures

To discuss the relevance of the proposed measures, we focus on the third movement of the string quartet No 3, here called K157.3. This movement is in rondo form and its main key is $C\ Major$. The modulations to other keys (the dominant $G\ Major$ and the parallel $C\ minor$) are easily identifiable and observable.

Figures 4 and 5 show, for each beat, the d_{diat} and $c_{V\to I}$ measures for a set of selected keys as well as the combined optimal measure. The selected keys are:

- $C\ Major$, the main key of the movement,
- $G\ Major$, the dominant key of $C\ Major$ (modulation at measures 17 through 32),
- $C\ minor$, the parallel key of $C\ Major$ (modulation at measures 49-56 and 61-64),
- $C\flat\ minor$, an intentional distant key to confirm the pertinence of each measure.

The distance d_W on keys is not computed on each beat in an independent way as it requires the hypothetical last best key in order to take into account how likely the resulting transition is. We thus do not evaluate independently d_W, but together with D in the next section.

4.2.1 Current Diatonic Pitch Sets and Pitch Compatibility Measure

Figure 4 shows that the pitch compatibility measure successfully estimates $d_{diat}(b, C\flat\ minor)$ to be unlikely anywere in the piece. There is indeed not much compatibility between pitches of $C\flat\ minor$ and the pitches actually played in the movement.

For a majority of the beats, the key of the reference annotation is the only one to be assigned the minimal value

key	ref	pred	TP	FN	FP	F_1
C $Major$	28	34	24	4	10	0.77
C $minor$	4	26	2	2	24	0.13
G $Major$	4	5	3	1	2	0.67
G $minor$	0	3	0	0	3	·
$E\flat$ $Major$	2	0	0	2	0	·
others	0	48	0	0	48	·
($C\flat$ $minor$)	0	0	0	0	0	·

Table 1. $V \to I$ progressions in K157.3, in the reference annotation (ref), predicted by the heuristic (pred), among which false negatives and positives (FN and FP), and associated F_1 measure.

(0) by the pitch compatibility measure. The best $d_{diat}(b, k)$ value describes the correct key for 243 out of the 252 beats. The current diatonic pitch set alone, therefore, appears to be sufficient to detect the key at beat b.

Starting from beat 100, only $d_{diat}(b, C\ minor)$ provides a value of zero for a few beats. In this region, Mozart progressively modulates by gradually introducing notes of the upcoming key $C\ minor$. The current diatonic pitch sets retrieved in this region do not exactly fit with any of the keys but still favor this key of $C\ minor$ in a few beats. Note that between beats 112 and 119, the key of the reference annotation $E\flat\ Major$ is correctly detected, although the corresponding curve is not shown in the figure for clarity.

4.2.2 $V \to I$ Progressions

The reference annotation on K157.3 contains 38 $V \to I$ movements. The heuristic detects 116 $V \to I$ movements, with 29 true positives (Table 1). As expected, most of the 87 false positives are on parallel keys and on $I \to IV$ progressions. For example, spurious $V \to I$ are detected in $F\ Major$, $F\ minor$, $A\flat\ Major$ and in $A\flat\ minor$ while they are in fact $I \to IV$ in $C\ Major$ and $E\flat\ Major$. Focusing on the two most present keys in the piece, $C\ Major$ and $G\ Major$, there are finally few false positives (12/39).

The heuristic thus manages to detect relevant signals for modulations, even if it detects other signals too.

4.2.3 Tonality Anchoring Measure

The Figure 5 displays the evolution of the tonality anchoring measure $c_{V \to I}(b, k)$ with the same selection of four keys along K157.3. As for the pitch compatibility measure, the maximum value $c = 20$ of $c_{V \to I}(b, C\flat\ minor)$ for $b > 20$ indicates that tonality anchoring measure successfully excludes distant keys. Indeed, $c_{V \to I}(0, k) = 0$ for all keys and progressively converge towards c unless a $V \to I$ movement is detected.

The values of $c_{V \to I}(b, C\ Major)$ and $c_{V \to I}(b, C\ minor)$ are often very close due to the false positives in the $V \to I$ detection. The consequence is the poor performance of $c_{V \to I}(b, k)$ by itself, predicting correctly only 66 of the 252 beats of K157.3.

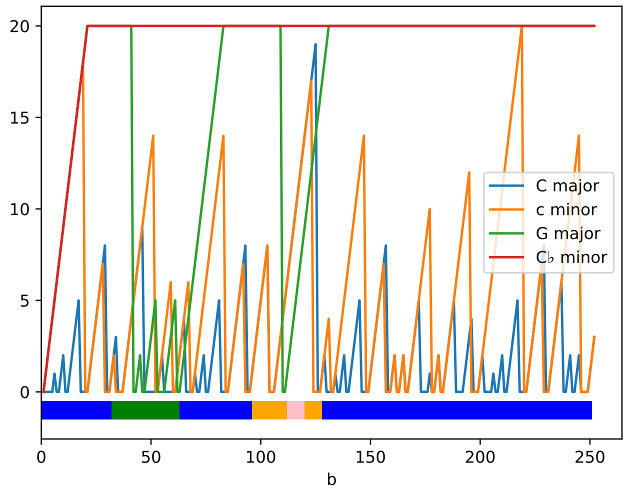

Figure 5. Tonality anchoring measure $c_{V \to I}(b, k)$ along K157.3

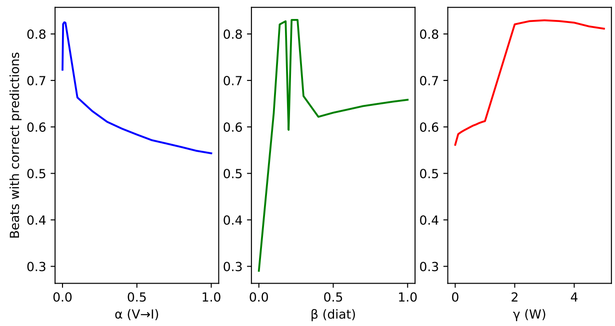

Figure 6. Average accuracies of key detection over the 38 Mozart's string quartets when one of the $\alpha/\beta/\gamma$ coefficient is constant.

4.3 Key and Modulation Detection

We now evaluate the algorithm by combining the three measures. The best coefficients α, β, and γ were tracked by evaluating combinations of values between 0 and 4 (with increments of about 0.002 for α, 0.02 for β, and 0.5 for γ) on an annotated corpus of 38 Mozart's string quartets. Although no training set was separated from the corpus, we felt that overfitting was a minor risk due to the size of the corpus and the very small number of parameters.

Figure 6 shows the average prediction accuracy of key detection when one coefficient is constant and the two other changes. Table 2 shows that although the current diatonic pitch set measure alone ($\beta = 1$) yields good results, adding the others measures improves the local key detection by 17.5%. The optimal values of $\alpha = 0.016$, $\beta = 0.3$, and $\gamma = 4.0$ reflect that, even after bounding and normalization, $c_{V \to I}$ is generally high and does not significantly help the detection here, probably due to the false positives (see the discussion in the previous section).

Table 3 further details these results, gathering keys according to how their tonic relates to the scale degrees of the main key of each movement. The algorithm shows a higher F_1 measure (and so better results) when it detects

method	α	β	γ	correct
only d_{diat}	0	1	0	67.3
only $c_{V \to I}$	1	0	0	16.3
no modulation	0	0	1	50.0
best coefficients	0.016	0.3	4	84.8

Table 2. Correct key prediction (percentage of beats) on the corpus. When $\alpha = \beta = 0$, then $D(b, k)$ is minimal when each beat is in the main key: This baseline algorithm considers that no modulations occur.

Main key in major mode (I)

key	ref	pred	TP	FP	FN	F_1
I	5801	5999	5451	548	350	0.92
V	2851	2846	2525	321	326	0.89
vi	469	410	330	80	139	0.75
IV	386	382	241	141	145	0.63
i	204	217	159	58	45	0.76
v	152	129	81	48	71	0.58
ii	130	18	9	9	121	0.12
♭III	102	84	63	21	39	0.68
others	238	248	139	109	99	–
total	10333	10333	8998	1335	1335	0.87

Main key in minor mode (i)

key	ref	pred	TP	FP	FN	F_1
i	1268	1038	1012	26	256	0.88
I	544	543	462	81	82	0.85
♭III	421	654	385	269	36	0.72
V	262	234	213	21	49	0.86
iv	247	133	118	15	129	0.62
v	177	230	155	75	22	0.76
others	253	340	117	223	136	–
total	3172	172	2462	710	710	0.78

Table 3. Key detection accuracy over the 30 Mozart string quartets movements in a major key (top), and in the 8 movements in a minor key (bottom), with the best coefficients. The results are gathered by keys. Each key is identified by how its tonic relates to the scale degrees of the main key of each movement. Major key are uppercase while minor key are lowercase.

the main key – and that concerns most of the beats in the corpus. The algorithm shows also better results on pieces in major main keys than in minor main keys. This is a known issue in key detection and is mostly explained by the floating 6th and 7th scale degrees of the minor scale. It is also confirmed by the better performance for estimating major keys than minor keys. However, the algorithm has more trouble finding subdominant related keys (ii, iv and IV), and further research could investigate this behaviour. Note finally that keys mostly used by Mozart in the corpus are the ones that are directly aside the main key in the Weber's table (with the exception of V when the main key is in minor mode).

As an example, we examine the prediction on K157.3 in detail. Figure 7 plots the $D(b, k)$ measure (relatively to the best measure between all keys) with the optimal values of α, β, and γ found above. It confirms again that $C\flat\ minor$ is not expected. The proposed method manages to find the main key of $C\ Major$ and all modulations ($G\ Major$, $C\ minor$, and – not shown – $E\flat\ Major$), at most within 2 beats of the place the modulation occurs in the reference analysis. The key detection works here on 248 of the 252 beats, improving the prediction with d_{diat} alone. Indeed, the computation of D, including d_W, favors some stability in the predicted keys, preventing the algorithm from switching keys for only 1 or 2 beats when a nonchord tone do appear.

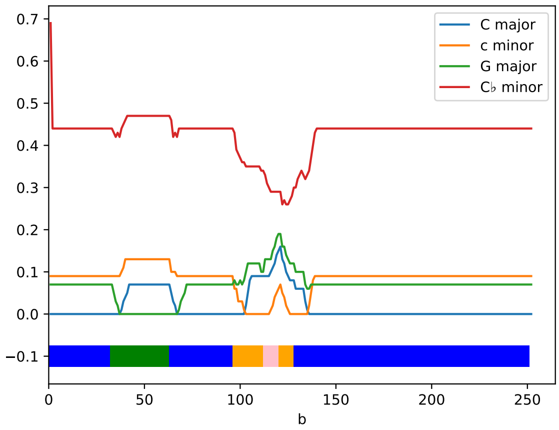

Figure 7. Key Detection along K157.3 by combining the three measures with the optimal α, β, γ coefficients. The graph shows the difference $D(k, b) - \min_{k'} D(k', b)$.

5. CONCLUSIONS

When analyzing music, looking at how keys change and modulations are performed is at least as interesting as keys themselves. We have introduced techniques to model key changes based on essential signs of modulation, such as the current diatonic pitch set and $V \to I$ progressions. Combined with a tonality proximity measure, these techniques successfully estimate the tonal plan, with more than 80% correct predictions on some classical string quartets – without any computation of a pitch profile. In particular, the d_{diat} measure alone provides a very good estimation of the keys and their change.

Perspectives include a more accurate $V \to I$ detection and more complete benchmarks, including comparisons with alternative key finding algorithms and other theories on tonality, as well as an evaluation of the detected position of each modulation. The combination of the measures is still very simple – improved combinations, possibly involving machine learning, could improve the raw results.

Altogether, these new models on tonality could help in (semi-)automatic harmonic analysis while further providing insights on the signals involved in the modulations.

Acknowledgments

We thank Timothy de Reuse, Gianluca Micchi and Néstor Nápolés López for their comments on this study.

6. REFERENCES

[1] J.-P. Rameau, *Traité de l'harmonie réduite à ses principes naturels*. Paris: J. B. C. Ballard, 1722.

[2] A. Choron and F.-J.-M. Fayolle, *Dictionnaire historique des musiciens artistes et amateurs, morts ou vivans*. Paris: Valade ou Lenormant, 1810.

[3] J. Hepokoski and W. Darcy, *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth-Century Sonata*. Oxford University Press, 2006.

[4] H. Longuet-Higgins and M. Steedman, *On Interpreting Bach*. MIT Press, 1987, pp. 221–241.

[5] P. G. Vos and E. W. V. Geenen, "A parallel-processing key-finding model," *Music Perception: An Interdisciplinary Journal*, vol. 14, no. 2, pp. 185–223, 1996.

[6] C. L. Krumhansl and E. J. Kessler, "Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys," *Psychological Review*, vol. 89, no. 2, pp. 334–368, 1982.

[7] C. L. Krumhansl, *Cognitive Foundations of Musical Pitch*. Oxford University Press, 1990.

[8] F. Lerdahl, "Tonal pitch space," *Music Perception*, vol. 5, pp. 315–350, 01 1988.

[9] D. Temperley, "What's key for key ? the Krumhansl-Schmuckler key-finding algorithm reconsidered," *Music Perception*, vol. 17, no. 1, pp. 65–100, 1999.

[10] B. J. Aarden, "Dynamic melodic expectancy," Ph.D. dissertation, The Ohio State University, 2003.

[11] H. Bellmann, "About the determination of key of a musical excerpt," in *International Symposium on Computer Music Modeling and Retrieval (CMMR 2006)*. Springer, 2006, pp. 76–91.

[12] W. B. de Haas, M. Robine, P. Hanna, R. C. Veltkamp, and F. Wiering, "Comparing approaches to the similarity of musical chord sequences," in *Exploring Music Contents*. Springer, 2011, pp. 242–258.

[13] C. S. Sapp, "Computational methods for the analysis of musical structure," Ph.D. dissertation, Stanford University, 2011.

[14] J. Albrecht and D. Shanahan, "The use of large corpora to train a new type of key-finding algorithm: An improved treatment of the minor mode," *Music Perception: An Interdisciplinary Journal*, vol. 31, no. 1, pp. 59–67, 2013.

[15] N. Nápoles López, C. Arthur, and I. Fujinaga, "Key-finding based on a Hidden Markov Model and key profiles," in *Digital Libraries for Musicology (DLfM 2019)*, 2019, p. 33–37.

[16] T. Chen and L. Su, "Functional harmony recognition of symbolic music data with multi-task recurrent neural networks," in *International Society for Music Information Retrieval Conference (ISMIR 2018)*, 2018, pp. 90–97.

[17] ——, "Harmony transformer: Incorporating chord segmentation into harmony recognition," in *International Society for Music Information Retrieval Conference (ISMIR 2018)*, 2019, pp. 259–267.

[18] G. Micchi, M. Gotham, and M. Giraud, "Not all roads lead to Rome: Pitch representation and model architecture for automatic harmonic analysis," *Transactions of the International Society for Music Information Retrieval*, vol. 3, no. 1, pp. 42–54, 2020.

[19] E. Chew, "The spiral array: An algorithm for determining key boundaries," in *Music and Artificial Intelligence*. Springer, 2002, pp. 18–31.

[20] ——, "Towards a mathematical model of tonality," 2000.

[21] N. Rimski-Korsakov, *Practical Manual of Harmony*. St. Petersburg: A Büttner, 1886.

[22] D. Meredith and G. A. Wiggins, "Comparing pitch spelling algorithms," in *International Conference on Music Information Retrieval (ISMIR 2004)*, 2004.

[23] L. Feisthauer, L. Bigo, and M. Giraud, "Modeling and learning structural breaks in sonata forms," in *International Society for Music Information Retrieval Conference (ISMIR 2019)*, 2019.

[24] N. Diletsky, *Idea grammatikii musikiyskoy*. Moskow, 1679.

[25] J. D. Heinichen, *Der Ganeral-Bass in der Composition*. Dresden, 1728.

[26] G. Weber, *Versuch einer geordneten Theorie der Tonsetzkunst*. Mainz: B. Schott's Sohnen, 1817.

[27] M. Reger, *Supplement to the Theory of Modulation*. C. F. Kahnt Nachfolger, 1904.

[28] P. I. Tchaïkovski, *Guide to the Practical Study of Harmony*. Moscow: P. Jurgenson, 1871.

[29] P. Allegraud, L. Bigo, L. Feisthauer, M. Giraud, R. Groult, E. Leguy, and F. Levé, "Learning Sonata Form Structure on Mozart's String Quartets," *Transactions of the International Society for Music Information Retrieval*, vol. 2, no. 1, pp. 82–96, 2019.

[30] C. S. Sapp, "Online database of scores in the Humdrum file format," in *International Conference on Music Information Retrieval (ISMIR 2005)*, 2005, pp. 664–665.

[31] M. S. Cuthbert and C. Ariza, "music21: A toolkit for computer-aided musicology and symbolic music data," in *International Society for Music Information Retrieval Conference (ISMIR 2010)*, 2010, pp. 637–642.