

Discovering Tonal Profiles with Latent Dirichlet Allocation

Music & Science
Volume 4: 1–15

© The Author(s) 2021
DOI: 10.1177/20592043211048827
journals.sagepub.com/home/mns



Fabian C. Moss¹  and Martin Rohrmeier¹

Abstract

Music analysis, in particular harmonic analysis, is concerned with the way pitches are organized in pieces of music, and a range of empirical applications have been developed, for example, for chord recognition or key finding. Naturally, these approaches rely on some operationalization of the concepts they aim to investigate. In this study, we take a complementary approach and discover latent tonal structures in an unsupervised manner. We use the topic model Latent Dirichlet Allocation and apply it to a large historical corpus of musical pieces from the Western classical tradition. This method conceives topics as distributions of pitch classes without assuming *a priori* that they correspond to either chords, keys, or other harmonic phenomena. To illustrate the generative process assumed by the model, we create an artificial corpus with arbitrary parameter settings and compare the sampled pieces to real compositions. The results we obtain by applying the topic model to the musical corpus show that the inferred topics have music-theoretically meaningful interpretations. In particular, topics cover contiguous segments on the line of fifths and mostly correspond to diatonic sets. Moreover, tracing the prominence of topics over the course of music history over ~600 years reflects changes in the ways pitch classes are employed in musical compositions and reveals particularly strong changes at the transition from common-practice to extended tonality in the 19th century.

Keywords

Topic modelling, Latent Dirichlet Allocation, tonal pitch classes, tonality, corpus studies

Introduction

Central questions in music theory and music information retrieval (MIR) concern the discovery of latent structures in musical pieces that abstract from the musical surface, that is, the observed notes or audio signal. Two common tasks in MIR that address these questions with respect to Western classical music are chord recognition and key finding, which, respectively, entail the retrieval of the harmonic content of a (usually short) segment of music (Müller, 2015; Korzeniowski and Widmer, 2018b; Micchi et al., 2020; McLeod and Rohrmeier, 2021) and the classification of the entire piece or larger segments into a musical key, for example, one of the 24 major and minor keys (Faraldo et al., 2016; Korzeniowski and Widmer, 2018a; Temperley, 1999; Weiß et al., 2020).

While chord recognition focuses on local and key finding on global levels, music theory suggests the existence of several intermediate levels that are frequently conceptualized as being hierarchically nested (e.g., Hauptmann, 1853; Schenker, 1935; Salzer, 1952; Lerdahl and Jackendoff, 1983; Lerdahl, 2001; Rohrmeier, 2011, 2020; Rohrmeier

and Moss, 2021). A number of psychological studies provide evidence for the perceptual reality of hierarchical organization in music (Krumhansl, 2004; Tillmann and Bigand, 2004; Koelsch et al., 2013; Farbood, 2016; Herff et al., 2021) but the exact relation between theoretically postulated and perceived hierarchies is not yet fully understood.

Moreover, a clear distinction between different levels might not always be possible, which suggests that discrete categories (such as chords and keys) might not entirely capture all aspects of tonal organization.¹ Circumventing the notion of a fixed number of discrete levels, Sapp (2001) introduces *key scapes* that map a key estimate provided by

¹ Digital and Cognitive Musicology Lab, Digital Humanities Institute, Collège des Humanités, École Polytechnique Fédérale de Lausanne, Lausanne, CH, Switzerland

Corresponding author:

Fabian C. Moss, Digital and Cognitive Musicology Lab, Digital Humanities Institute, Collège des Humanités, École Polytechnique Fédérale de Lausanne, Lausanne, CH, Switzerland.
Email: fabian.moss@epfl.ch



the Krumhansl–Kessler algorithm (Krumhansl, 1990) to all possible segments of a piece. His analyses show that stable tonal regions form over the course of tonal pieces, which he interprets as points on a continuous spectrum of increasing specificity, namely ‘key of piece’, ‘strong keys’, ‘weak keys’, ‘tonicizations’, ‘cadences’ and ‘chords’ (Sapp, 2005). This approach was extended by Lieck and Rohrmeier (2020) to infer prototypical modulation plans of musical pieces, and by Viacoz et al. (in press), who combine key scapes with the application of the discrete Fourier transform to pitch-class distributions. Instead of assuming a fixed number of levels or relying on pre-defined concepts, such as chords, local keys, and global keys, this rather continuous view of the hierarchical organization suggests using unsupervised methods to infer latent musical structures.

Several unsupervised and/or statistical methods have been successfully employed in music research for the analysis of musical structures across a large range of musical styles and genres. For instance, a number of studies have analysed Rock music with respect to harmony (De Clercq and Temperley, 2011; Temperley and de Clercq, 2013; Tan et al., 2019) and form (de Clercq, 2017). Others have studied harmony or improvised melodic solos in Jazz (Broze and Shanahan, 2013; Pfeiderer et al., 2017), harmony and form in Brazilian Choro (Moss et al., 2020b), or pitch-class distributions in atonal music (Ballance, 2020). Western classical music has been studied most extensively, and analyses range from studies of harmony based on distributions of pitch-classes or chord labels, or features extracted from audio (Albrecht and Shanahan, 2013; White, 2013; Moss et al., 2019; Rohrmeier and Cross, 2008; Hedges and Rohrmeier, 2011; Bellmann, 2012; van Balen, 2016). An increasing number of studies is recently addressing questions on the historical development of Western music on data-based accounts (Serrà et al., 2012; Burgoyne et al., 2013; Rodrigues Zivic et al., 2013; Albrecht and Huron, 2014; Gauvin, 2015; Mauch et al., 2015; Huang et al., 2017; Weiß et al., 2018; Nakamura and Kaneko, 2019; Moss, 2019; Yust, 2019; Anzuoni et al., 2021; Harasim et al., 2021).

In natural language processing as well as the digital humanities, a widely used approach for the unsupervised discovery of latent structures in textual sources is commonly subsumed under the term *topic modelling* (Steyvers and Griffiths, 2007; Piper, 2018), one of its most prominent variants being Latent Dirichlet Allocation (LDA; Blei et al., 2003). Taking a textual example, one would expect that a document with the topic ‘politics’ contains many names of politicians, institutions, states, or political events such as elections, wars, and so forth. It is rather unlikely that such a text would contain the names of composers, musical pieces, or music-theoretical terms such as ‘augmented-sixth chord’, ‘symphony’, ‘Passacaglia’, and the like. Topic models turn this argument around and assume that the hypothetical text is about politics precisely *because* it contains many words from political topics. That is, topics are defined by the frequency of co-occurrence of certain words. This is sometimes called the *distributional hypothesis* (Harris, 1954). Therefore, topic

models define a composite model that describes documents (e.g., texts), which are represented as ‘bags of items’ (e.g., words) as mixtures of several latent topics which, in turn, are defined as distributions over the same items.

This notion of a topic can directly be translated to the musical case if one considers pieces to be the documents and pitch classes to play the role of words in these documents. The vocabulary is then the set of all pitch classes that appear in any document in a corpus, and topics correspond to distributions over these pitch classes that represent which pitch classes frequently co-occur in the pieces. Accordingly, a piece containing only one or a few topics can thus be considered to be tonally more coherent than one that has a larger number of topics, in the same manner that texts are thematically more coherent when they talk about fewer topics.²

This translation of the LDA topic model to the case of music, however, is only structural and not semantic, in that the algorithmic structure is identical but the interpretation of basic units and topics is highly domain-dependent. Crucially, applying LDA to music does *not* entail that pitch classes play a similar or even analogous role as words do for language, for the simple reason that pitch classes bear very little semantic meaning on their own. While the semantics of words also are not completely independent of their context (Baayen, 2001), musical meaning is fundamentally contextualized (Huron, 2006; Koelsch, 2012; Schlenker, 2017). In the following sections, we thus do not claim that pitch classes in music possess similar functions to words in language but that the topics defined by the LDA model attain their meaning purely through the collocation of pitch classes in musical pieces, and that they need to be interpreted accordingly. Arguably, somewhat larger units in music – such as melodic fragments, motives, or short harmonically consistent segments – constitute more meaningful basic units. They are, however, not explicitly represented in the music and subject to manual or automated analysis, which in itself is a difficult task associated with some degree of ambiguity. By relying on ‘atomic’ pitch classes as basic units, the interpretation of the topics is thus less straightforward but circumvents the problem of inferential uncertainties of higher-order musical patterns. Moreover, the particular representation of pitch classes that we employ entails some notion of membership of a certain scale, which aids in the interpretations of the topics obtained (see below for more details).

LDA for symbolic representations of music

Whereas topic models are ubiquitous in the digital humanities, in particular in the context of literary studies (e.g., Blei, 2012b; Underwood, 2012; Jockers and Mimno, 2013; Goldstone and Underwood, 2014; Rhody, 2012; Piper, 2018), there are to date only a few applications of the LDA topic model to music that are not based on textual data (such as metadata or lyrics) but that take the note content of musical pieces into account.

Mauch et al. (2015) use LDA for harmonic and timbral descriptors of features derived from audio recordings of

Pop music and study the historical evolution of these topics between 1960 and 2010. Lieck et al. (2020) introduced the Tonal Diffusion Model (TDM), which is conceptually related to LDA but also incorporates information about interval relations between pitch classes given by the topology of tonal space to model the distributions of pitch classes in pieces by Bach, Beethoven and Liszt. It essentially explains the occurrence of tonal pitch classes by tracing them back to a common tonal centre through paths on the *Tonnetz* (Cohn, 1998) that are given by interval combinations of ascending and descending perfect fifths, major thirds and minor thirds, each also associated with certain inferred weights. This allows them, for a given piece, to infer the most likely combination of interval weights that gave rise to the piece’s pitch-class distribution according to the particular assumptions of their model.

Most closely related to the present study is the work by Hu and Saul (2009a, 2009b). They interpret musical keys as topics and use LDA to infer tone profiles (in the sense of distributions over the 12 enharmonically equivalent pitch classes) from pieces. They evaluate the inferred topics against the tone profiles provided by Krumhansl and Kessler (1982) and use their topics to trace key changes over the temporal course of a musical piece. Unfortunately, they do not report the numerical values of their profiles so that a direct comparison with the other profiles is impossible. The present approach differs from theirs in several regards.

First, they use a relatively small sample of manually selected pieces.³ Their dataset approximately spans a historical range of 260 years, whereas the number of pieces and composers, as well as the extent of the historical range of the corpus supporting the present study, is much larger (see ‘Corpus and representation of notes’).

Furthermore, in Hu and Saul’s (2009b) approach, the basic units of the model are not individual pitch classes but short segments of music that contain a certain amount of temporal information, which is not the case here. This allows them to trace changes of key (modulations) in the pieces and evaluate these assignments against the predictions by the Krumhansl–Kessler key-finding algorithm (Krumhansl, 1990). Our approach, in contrast, traces the presence and absence of latent topics on a historical time scale and interprets this as reflections of underlying changes in compositional procedures that bring about changes in tonality.

The two topics found by Hu and Saul (2009a) resemble the major and minor profiles by Krumhansl and Kessler, but their major profile in particular rather emphasizes the notes of a major triad than those of a major key with high weights for pitch classes 0, 4 and 7 (possibly an artefact of the segment level that they introduce) with the other in-scale pitch classes having very low weights, and pitch class 10 being stronger than pitch class 11 (the leading tone). This again suggests a more fine-grained conception of latent musical structure than just chords and keys.

Most importantly, their data is encoded in Musical Instrument Digital Interface (MIDI) format, which only allows for the representation of 12 distinct pitch classes

that do not allow to distinguish enharmonically equivalent notes. In contrast, the dataset used here encodes the exact spelling of the notes (but without octave information), which is sometimes called the *tonal* pitch-class (Temperley, 2000; Moss et al., in press) representation, leading to a larger vocabulary of pitch-class types and enabling enharmonic distinctions, for example, between $F\sharp$ and $G\flat$ or between C and $B\sharp$. While one can obtain the former representation from the latter by simply grouping all enharmonically equivalent pitch classes, the reverse direction requires some form of inference, and a number of algorithms have been proposed for this task, usually termed ‘pitch spelling’ (Temperley, 2001; Stoddard et al., 2004; Chew and Chen, 2005; Cambouropoulos, 2003; Meredith, 2006; Foscarin et al., 2021).

Corpus and representation of notes

The corpus used in this study is the *Tonal Pitch-Class Counts Corpus* (TP3C; Moss et al., 2020a). It consists of 2,012 musical pieces by 75 composers encoded in MusicXML format, spanning a range of almost 600 years and containing more than 2.7 million notes in total. It is available at <https://github.com/DCMLab/TP3C/> and it was assembled from a range of different sources, such as scores from the *Electronic Locator of Vertical Interval Successions* (ELVIS) project,⁴ and the Humdrum `**kern` scores of the *Center for Computer Assisted Research in the Humanities* (CCARH)⁵, as well as files from public repositories such as the *Choral Public Domain Library* (CPDL),⁶ or the community page of the *MuseScore* notation software,⁷ whereas others have been transcribed at the *Digital and Cognitive Musicology Lab* (DCML).⁸ A full list of the pieces and sources used is given in (Moss, 2019). The MusicXML encoding allows representing notes as tonal (spelled), as opposed to enharmonically equivalent, pitch classes. Using the chromatic circle or the circle of fifths would thus discard potentially valuable information, for instance, enharmonic differences between notes (e.g., between C and $B\sharp$). The line of fifths (Temperley, 2000, Figure 1) is more appropriate in this case. We consequently generalise the notion of tone profiles (distributions over 12 chromatic pitch classes) and define *tonal profiles* to be distributions of tonal pitch classes on the line of fifths and thus on a potentially infinite number of distinct notes. For practical reasons, we restrict the vocabulary of tonal pitch classes to those contained in the line-of-fifths segment from $F\flat$ to $B\sharp$, since no piece in the corpus contains a note outside of this range.

Note, in particular, that we represent pieces as *absolute* pitch-class distributions, that is, we do not transpose them

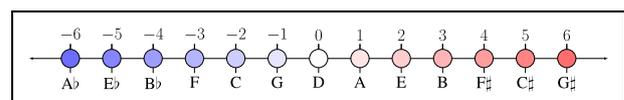


Figure 1. Schematic depiction of the tonal pitch classes on the line of fifths mapped to integers in \mathbb{Z} .

to a common center, for example, the tonic of a piece. The main rationale for this decision is that defining or inferring a tonic to which all pitch classes can be related is not equally feasible for all historical periods in our corpus. It is certainly appropriate for common-practice compositions (roughly from the Baroque to the early Romantic periods) to draw on the concept of a global tonic, but earlier Renaissance pieces based on modality as well as later late-Romantic or Modernist pieces in the idiom of extended tonality may employ different notions of tonal centers.⁹ Moreover, the general assumption of transpositional equivalence has been called into question, in particular on historical reasoning (Quinn and White, 2017; Rom, 2011). It is important to bear this in mind, in particular when interpreting our results. Since we base our study on pitch-class distributions from untransposed pieces, the inferred topics likewise will reflect underlying absolute pitch-class distributions and thus, for instance, allow us to draw conclusions about the absolute prevalence of certain scales, keys, or modes (in the modal music sense) in the corpus. It would be, in principle, possible to adapt the present study to the case of relative pitch classes by first estimating for each piece its tonal centre and transposing all of its pitch classes accordingly. The results and their interpretation would necessarily change since one would essentially ask a different research question, and we plan to explore this avenue in our future work.

Consider the distribution of tonal pitch classes from the first movement of Charles Valentin Alkan’s *Concerto for Solo Piano*, op. 39, no. 8 (1857). The distribution of tonal pitch classes in this piece on the line of fifths is shown in Figure 2. The colours emphasize the ordering of the line of fifths as well as the distance from the central D by the intensity of the colors and the direction towards more flat or sharp tonal pitch classes by the blue and red hues, respectively (see Figure 1). Note that this piece contains more than 12 different tonal pitch classes and spans a range from F♭ to G♯♯ (23 fifths). Hence, representing it by enharmonically equivalent pitch classes would unnecessarily discard valuable information and obliterate potentially important enharmonic differences.

The multi-modal shape of the distribution suggests modelling its generative process as a weighted mixture of a small number of simpler distributions. This modelling

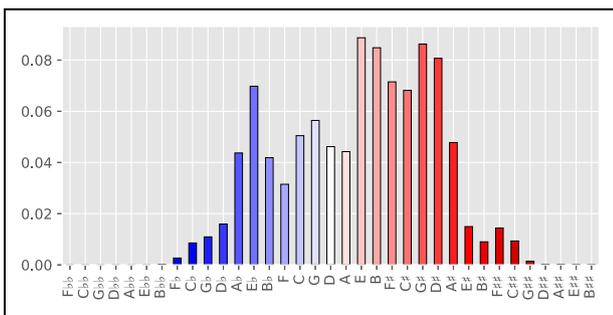


Figure 2. Distribution of tonal pitch classes of the first movement of Alkan’s *Concerto for Solo Piano*, op. 39, no. 8 (1857).

assumption entails understanding the overall distribution of tonal classes in a piece as a mixture of different tonal profiles. The components might, for example, correspond to the tonal pitch-class distributions of sections of a piece that are in different keys, and which may moreover contain chromaticism and enharmonicism. LDA naturally expresses these relations of weighted latent components (topics) that underlie the observed notes in a musical piece.

Overview

In the following section (‘Topic modelling with LDA’), we first illustrate the generative process for the tonal pitch-class distribution of a musical piece according to the assumptions and structure of the LDA model by creating an artificial corpus of pieces according to arbitrary parameter settings of the model (‘Generating pieces’). Subsequently, we apply this procedure to the TP3C corpus, inferring the latent topics using Gibbs sampling (‘Inferring topics’), and discussing their music-theoretical interpretations both qualitatively (‘Tonal profiles obtained by topic modelling’) and quantitatively (‘Topic similarities’). We further observe changes in the prominence of topics over the course of music history (‘Historical evolution of topics’) and conclude with a general discussion of the approach, outlining potential avenues for future research (‘Conclusion’).

Topic modelling with LDA

Generating pieces

The LDA model establishes probabilistic relations between topics and documents¹⁰ in a corpus and specifies a generative process that is assumed to underlie the distribution of notes in pieces (Blei, 2012a). This generative model can thus also be used to create new documents, given a certain setting of the parameters of the model. It is important to note that ‘generating’ does not mean that LDA attempts to simulate the process of the composition of a piece. We will first illustrate the model by artificially generating an artificial corpus with $D = 20$ documents and $K = 3$ topics. Subsequently, we will use the LDA model to infer $K = 7$ topics from the 2,012 pieces in the TP3C corpus and also compare the results for different values of K . In general, the more topics one assumes to exist in a corpus, the more specific they will be, whereas a small value of K leads to broader, more general topics. The vocabulary size is $V = 35$ in both cases since we consider tonal pitch classes from F♭ to B♯♯ on the line of fifths.

We start by creating a distribution of topic weights for each document in the corpus. These distributions $\theta_{1:D}$ determine how prominent the K topics are in each of the D documents.¹¹ The distribution θ_d over all topics for document d is modelled as a sample from a Dirichlet distribution with a K -dimensional hyperparameter α

$$\theta_d \sim \text{Dir}(\alpha), \quad \alpha \in \mathbb{R}_+^K \quad (1)$$

and the probability of topic $k \in \{1, \dots, K\}$ in document d in the artificial corpus is given by $\theta_{d,k}$. We set $\alpha = \mathbf{1}_K$ for all D documents in the artificial corpus, where $\mathbf{1}_K$ is the K -dimensional vector containing only 1s.

The sampled probabilities of the three topics in all 20 documents are shown in Figure 3. Topic 1 is shown in red, topic 2 in orange, and topic 3 in gray. For example, the topic distribution in document 1 is relatively balanced with $\theta_1 = (.27, .42, .31)$, while the one for document 15 is $\theta_{15} = (.51, .03, .46)$ and is almost exclusively concentrated on topics 1 and 3. The topic distributions vary considerably due to the setting of α that specifies a uniform distribution over the $K - 1$ -simplex to the effect that all configurations for $\theta_d, d \in \{1, \dots, D\}$, are equally probable. Note that, at this stage in the generative process, we do not know yet how the K topics are composed, that is, what the tonal pitch-class distributions look like that they represent.

This happens in the next step, where each of the K topics is associated with a distribution over the V tonal pitch classes that expresses how important each tonal pitch class is for this topic. Accordingly, the topics $\phi_{1:K}$ are sampled from a Dirichlet distribution with V -dimensional hyperparameter β

$$\phi_k \sim \text{Dir}(\beta), \quad \beta \in \mathbb{R}_+^V \quad (2)$$

For all K topics of the artificial corpus, we set $\beta = \mathbf{1}_V$, where $\mathbf{1}_V$ is the V -dimensional vector containing only 1s. Both α and β are fixed hyperparameters for the whole corpus. The distribution of tonal pitch classes for the $K = 3$ topics is shown in Figure 4. As mentioned before, a small number of topics K leads to a broader range of topics that are not very specific. It can be seen that all three artificial topics cover the whole range of tonal pitch classes and thus each of them contains the entire vocabulary. Yet comparing the three topics also reveals subtle differences. For example, the probability for tonal pitch class D to occur is .045 for topic 1, .164 for topic 2 and .003 for topic 3. These are represented by the white parts of the bars in Figure 4. The probabilities of some of the other tonal pitch classes vary even more. Note, however, that the structure of these topics is not meaningful because of the arbitrary choice of a uniform hyperparameter β .

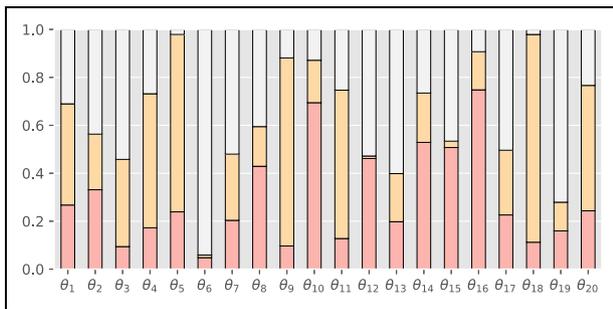


Figure 3. Topic weights for $K = 3$ topics and a corpus of $D = 20$ artificial documents sampled from a Dirichlet distribution with hyperparameter $\alpha = (1, 1, 1)$.

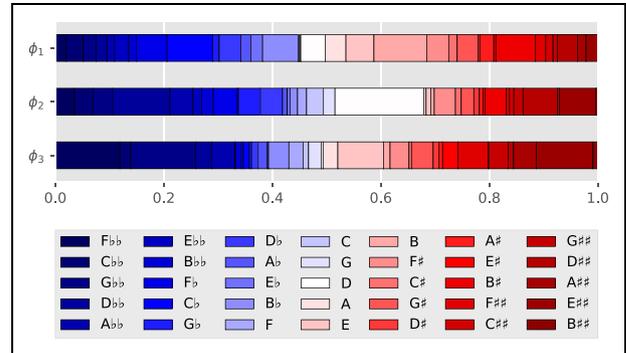


Figure 4. Tonal pitch-class distributions for $K = 3$ artificial topics.

So far, we have sampled distributions over the K topics $\theta_{1:D}$ for each of the D documents (Figure 3) and distributions over the V tonal pitch classes $\phi_{1:K}$ for each of the K topics (Figure 4). The final step in the generative process of the LDA model consists in sampling the actual tonal pitch classes for the D documents, given the distributions over topics $\theta_{1:D}$ and the distributions of tonal pitch classes for all topics $\phi_{1:K}$. For each of the D documents in the artificial corpus we sample its length, the total number of notes N_d in the d th document, from a Poisson distribution with hyperparameter $\lambda = 100$. This step is only necessary for the generation of new documents since the number of notes is deterministically given in actual documents in a corpus.

Next, we use the topic weights θ_d to sample a topic assignment $z_{d,n}$ for the n th note in the d th document from a categorical distribution using the corresponding parameter θ_d

$$z_{d,n} \sim \text{Cat}(\theta_d) \quad (3)$$

Given this topic assignment $z_{d,n}$, the actual n th tonal pitch class $w_{d,n}$ in the d th document w_d is sampled from a categorical distribution with parameter $\phi_{z_{d,n}}$

$$w_{d,n} \sim \text{Cat}(\phi_{z_{d,n}}) \quad (4)$$

The tonal pitch-class distribution for the d th document is then given by w_d , the collection of all notes sampled by the generative LDA process. The model is summarized by the joint probability¹² over all the random variables specified in equations (1)–(4):

$$\begin{aligned} & p(\phi, \theta, z, w \mid \alpha, \beta) \\ &= \prod_{d=1}^D \left(p(\theta_d \mid \alpha) \prod_{n=1}^{N_d} p(w_{d,n} \mid z_{d,n}, \phi_{z_{d,n}}) p(z_{d,n} \mid \theta_d) \right) \\ & \quad \times \prod_{k=1}^K p(\phi_k \mid \beta), \end{aligned} \quad (5)$$

where $\phi = \phi_{1:K}$, $\theta = \theta_{1:D}$, $z = z_{1:D}$, and $w = w_{1:D}$.

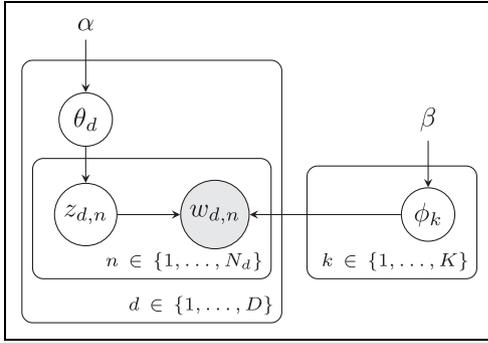


Figure 5. Graphical model for Latent Dirichlet Allocation (LDA) describing the relations between random variables $w_{d,n}$ (n th note in the d th document; observed), $z_{d,n}$ (topic assignment of this note; latent), θ_d (topic distribution in the d th document; latent), ϕ_k (tonal pitch-class distribution of topic k ; latent) and Dirichlet hyperparameters α and β .

A graphical representation of the LDA model in so-called *plate notation* (Bishop, 2006; Koller and Friedman, 2009) is shown in Figure 5. Circular nodes represent the observed (shaded) and latent (white) random variables of the model. The two hyperparameters α and β are shown without circles. Arrows between the variables represent the probabilistic dependencies between them. The boxes surrounding sets of variables are called plates. They stand for repeated sampling procedures of the random variables they contain.

The artificial corpus of all $D = 20$ documents generated with the hyperparameters α and β as defined above is shown in Figure 6. Because we set the parameter for the Poisson distribution to sample the lengths of the pieces in the artificial corpus to $\lambda = 100$, the expected length for pieces in this corpus is 100 notes. Since all $K = 3$ topics have non-zero probabilities for each of the $V = 35$ tonal pitch classes in the vocabulary (the line-of-fifth segment from $F\flat\flat$ to $B\sharp\sharp$; see Figure 4), it is not surprising that each piece in the artificial corpus contains every tonal pitch class at least once. To obtain a distribution of tonal pitch classes in the d th document, one can normalize $w_{d,d}$, the count of tonal pitch

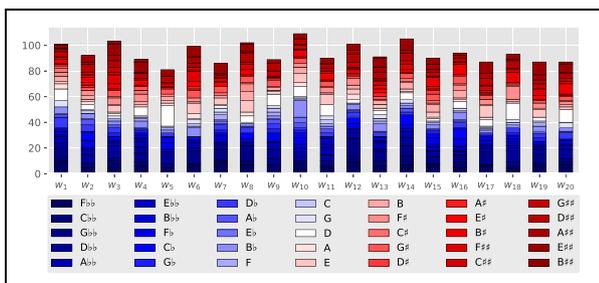


Figure 6. Counts of tonal pitch classes in $D = 20$ documents in a randomly generated corpus with $K = 3$ artificial topics.

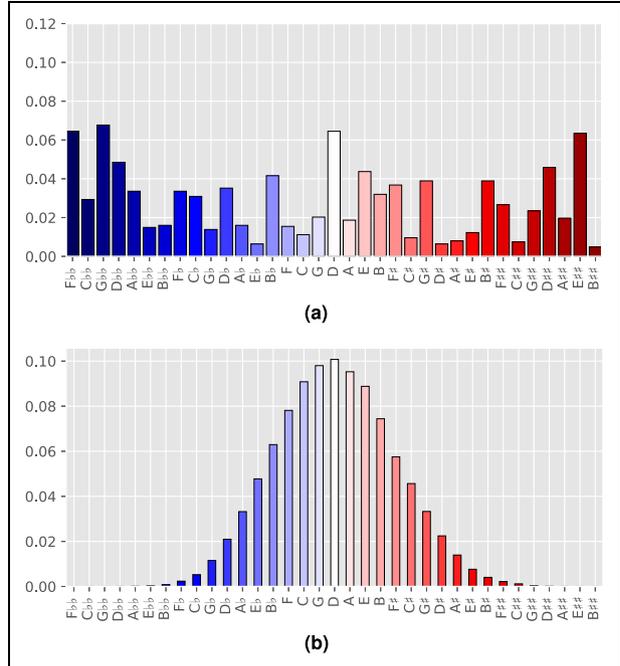


Figure 7. Tonal pitch-class distribution in a corpus of artificially sampled pieces (top) and the corpus (bottom). (a) Overall tonal pitch-class distribution in artificially generated corpus with $K = 3$ topics and $D = 20$ documents. (b) Overall tonal pitch-class distribution in the corpus with $D = 2012$ documents.

classes in the d th document, by dividing it by N_d , the total number of notes in this document (not shown).

Comparing the overall distribution of tonal pitch classes of all documents in the artificial corpus (Figure 7a) with that of the corpus of musical pieces used in this study (Figure 7b) shows a number of important differences: the former is spread out across the whole line of fifths, with the three most common tonal pitch classes being $G\flat\flat$, D , and $E\sharp\sharp$. It shows no recognizable structure due to the uniform parameter settings used in the illustration of the generative process above. In fact, increasing the size of the artificial corpus by generating more and more pieces will approximate a uniform distribution of tonal pitch classes. The actual distribution of tonal pitch classes of the observed data in the TP3C, however, seems to exhibit (discrete) Gaussian-like behaviour centred on D , a result of the pitch classes being largely organized along the line of fifths (Hentschel et al., 2021; Moss et al., in press). It is evident that the overall tonal pitch-class distribution in the artificial corpus, as well as the tonal pitch-class distributions of individual documents (Figure 6), have little to do with the actual tonal pitch-class distributions in musical pieces and the corpus, and the meaningfulness of the generated topics is questionable at best. This should not be surprising since notes in pieces by Western classical composers are not chosen uniformly at random (with notable exceptions of some 20th-century aleatoric composers such as Iannis Xenakis or John Cage). In the next section, we turn to the

Table 1. Tonal pitch-class distribution for $K = 7$ topics. Values < 0.001 have been omitted.

	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7
Fbb	—	—	—	—	—	—	—
Cbb	—	—	—	—	—	—	—
Gbb	—	—	—	—	—	—	—
Dbb	—	—	—	—	—	—	—
Abb	0.001	—	—	—	—	—	—
Ebb	0.004	—	—	—	—	—	—
Bbb	0.011	—	—	—	—	—	—
Fb	0.029	—	—	—	—	—	—
Cb	0.065	—	—	—	—	—	—
Gb	0.143	—	—	—	—	—	0.001
Db	0.21	—	—	—	—	—	0.032
Ab	0.149	—	—	—	—	—	0.168
Eb	0.131	—	—	0.043	—	—	0.241
Bb	0.149	—	0.001	0.217	—	—	0.132
F	0.097	—	0.184	0.162	—	0.001	0.126
C	0.007	—	0.267	0.073	—	0.089	0.192
G	—	—	0.055	0.166	0.007	0.341	0.1
D	0.003	—	0.033	0.225	0.11	0.267	0.001
A	0.001	0.003	0.2	0.103	0.225	0.029	—
E	—	0.062	0.224	—	0.201	0.04	0.005
B	—	0.132	0.031	—	0.126	0.183	0.001
F#	—	0.154	—	0.007	0.142	0.05	—
C#	—	0.142	0.001	0.003	0.128	—	—
G#	—	0.156	0.004	—	0.054	—	—
D#	—	0.148	—	—	0.005	0.002	—
A#	—	0.095	—	—	0.002	—	—
E#	—	0.052	—	—	0.001	—	—
B#	—	0.028	—	—	—	—	—
F##	—	0.015	—	—	—	—	—
C##	—	0.008	—	—	—	—	—
G##	—	0.002	—	—	—	—	—
D##	—	0.001	—	—	—	—	—
A##	—	—	—	—	—	—	—
E##	—	—	—	—	—	—	—
B##	—	—	—	—	—	—	—

topics has no particular meaning and that some weights for tonal pitch classes in the topics are too small to be displayed. The numerical values for the probabilities of the tonal pitch classes in the respective topics $\phi_{1:7}$ are given in Table 1.

It turns out that the shapes of the tonal pitch-class distributions for $K = 7$ topics fall into two groups: topics 1 and 2 (Figures 9a and 9b) are located towards the extremes of the line of fifths and cover relatively wide ranges; the remaining topics 3 to 7 (Figures 9c–9g) cover somewhat narrower ranges on the line of fifths and moreover feature some tonal pitch classes that are more distinct in terms of their relative frequencies, for example, F, C, A and E in Topic 3; Bb, F, G and D in Topic 4; A and E in Topic 5; G, D and B in Topic 6; and Ab, Eb and C in Topic 7. Notably, those pitch classes are related by perfect fifths, major and minor thirds, the constituent intervals of triads that figure prominently in theories of tonal hierarchies (Lerdahl, 2001; Lerdahl and Krumhansl, 2007; Lieck et al., 2020). As mentioned above, this is to be expected

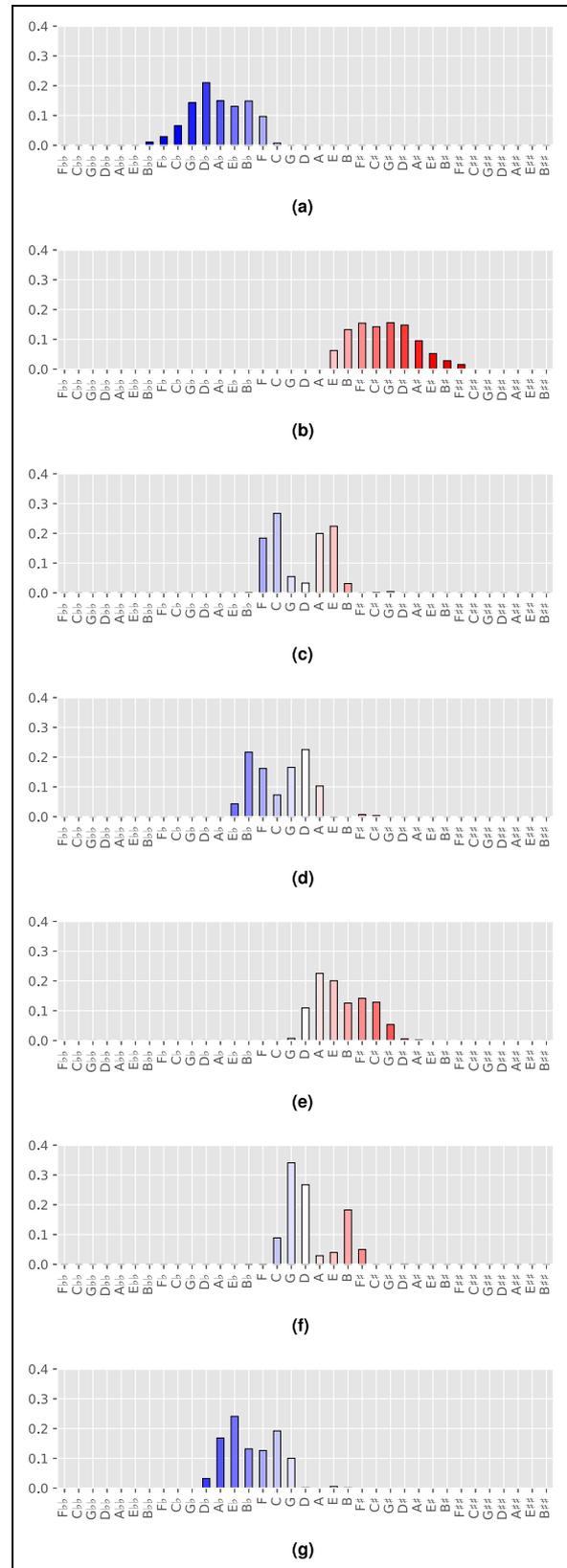


Figure 9. The note distributions for the $K = 7$ topics. (a) Topic 1 of 7. (b) Topic 2 of 7. (c) Topic 3 of 7. (d) Topic 4 of 7. (e) Topic 5 of 7. (f) Topic 6 of 7. (g) Topic 7 of 7.

in a corpus of mostly tonal compositions that are by-and-large governed by a single global key. Furthermore, they assign substantial probability mass to contiguous line-of-fifths segments that form diatonic scales, for example, C major/A minor in the case of Topic 3, or F major/D minor in the case of Topic 6.

This means that the best explanation according to the LDA model for the tonal pitch-class distributions in the corpus consists of several diatonic sets in the middle range of the line of fifths (mostly without accidentals) plus two topics representing the two extremes, flats (Topic 1) and sharps (Topic 2). We can interpret this as a corroboration of the validity of the results since Western classical pieces are commonly organized around one or a few keys that are related by fifths.

Along the same lines, Topics 1 and 2 can be interpreted as chromaticism that for instance occurs when composers write chromatic passing notes in otherwise diatonic passages. Moreover, they are responsible for pitch-class distributions stemming from less frequently employed keys relatively far from the center of the line of fifths, for example, E♭ minor or D♯ major. If this is the case, then one would expect that these two topics are particularly prevalent in the 18th and 19th centuries, in which music theorists observe an increasing usage of more distantly related keys and chromaticism (Aldwell et al., 2010). Looking at the topic assignments for individual pieces confirms these considerations. For instance, more than 93.5% of the tonal pitch classes in J. S. Bach’s prelude in C major (BWV 846) are assigned to topics 3 (46.5%), 4 (6.9%) and 6 (40.1%), respectively. Note in particular that the two strongest topics 3 and 6 emphasize pitch classes of the subdominant and dominant triads in C major, respectively. In contrast, the assignments for Alkan’s piece (Figure 2) are more diverse with the three most prevalent topics 2 (44.8%), 7 (20.3%), and 5 (14.9%) covering only 80% of all pitch classes.

The topic distributions for other values of $K \in \{3, 7, 12, 24\}$ are shown in the “Supplemental material”. There, too, all topics consist of gapless segments on the line of fifths. This is not trivial since the model is agnostic to the interval relations between the tonal pitch classes as expressed by the ordering on the line of fifths but only ‘sees’ notes in pieces as an unordered collection (bag). As mentioned before, increasing the number of topics also increases their specificity. Comparing how the topics change with larger values of K shows that the growing specificity is largely reflected in emphasizing fifth and third relations between the most prominent notes. A fact that is predicted by several hierarchical models of tonal space (Krumhansl, 1990; Lerdahl, 2001; Rohrmeier, 2011, 2020).

Topic similarities

Since topics are defined as distributions over tonal pitch classes, one can define appropriate measures to assess the similarity between them. For two discrete probability

distributions P and Q over a discrete support \mathcal{X} , the *Kullback–Leibler divergence* (Cover and Thomas, 2006) is defined as

$$\text{KL}(P \parallel Q) = - \sum_{x \in \mathcal{X}} P(x) \log_2 \left(\frac{Q(x)}{P(x)} \right) \quad (7)$$

Here, \mathcal{X} is the line-of-fifths segment from F♭♭ to B♯♯. The *Jensen–Shannon divergence* is a symmetrized version of the Kullback–Leibler divergence and is given by

$$\text{JS}(P \parallel Q) = \frac{\text{KL}(P \parallel M) + \text{KL}(Q \parallel M)}{2} \quad (8)$$

where $M = (P + Q)/2$ is the mean distribution of P and Q . Since $\text{JS}(P \parallel Q)$ measures the average divergence from the mean of the two distributions, having zero values in the distributions is unproblematic because M is zero if and only if both $P(x)$ and $Q(x)$ are zero for $x \in \mathcal{X}$. The square root of the Jensen–Shannon divergence defines a distance metric

$$d_{\text{JS}}(P, Q) = \text{JS}(P \parallel Q)^{\frac{1}{2}} \quad (9)$$

the *Jensen–Shannon distance* (Endres and Schindelin, 2003) that we can use to compare the topics with each other. Finally, because the distance d_{JS} is bounded between 0 and 1, we can express the *Jensen–Shannon similarity* between two topics ϕ_k and $\phi_{k'}$ as

$$s_{\text{JS}}(\phi_k, \phi_{k'}) = 1 - d_{\text{JS}}(\phi_k, \phi_{k'}) \quad (10)$$

where a similarity value of 1 implies that the two topics are identical.

The similarities between all pairs of the $K = 7$ topics are shown in Figure 10. The tonal pitch-class distributions that are most similar to each other are topics 1 and 7, two ‘flat’ topics, with $s_{\text{JS}}(\phi_1, \phi_7) = .53$.

Because the probabilities of tonal pitch classes in these two topics are different, their similarity is relatively low in absolute terms while still being the largest among all

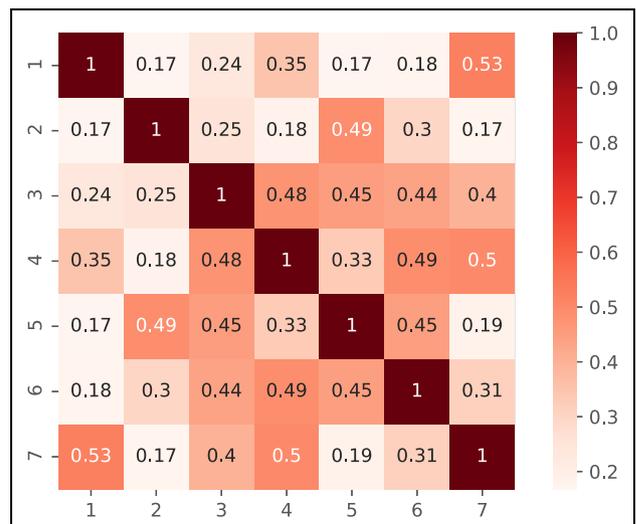


Figure 10. Jensen–Shannon similarities for $K = 7$ topics.

pairs of topics. The two most distinct topic pairs are topics 1 and 2, topics 1 and 5 and topics 2 and 7, each with a similarity score of .17, a consequence of the fact that they have almost no tonal pitch classes in common. Note that non-negative but very small probabilities are not visible in Figure 10 and Table 1, and that the Jensen–Shannon similarity involves the calculation of the average between two distributions, leading to relatively large similarities even if topics are as disparate as topics 1 and 2. The topic similarities for other values of $K \in \{3, 7, 12, 24\}$ are shown in the ‘Supplemental material’.

Historical evolution of topics

We now investigate how the prevalence of certain topics in musical pieces changes historically. Recall that in the LDA model, each note $w_{d,n}$ in each of the d pieces is associated with a topic assignment $z_{d,n}$ that represents the most likely topic this note was generated from. Consequently, we can count these topic assignments for each document to obtain θ_d , a distribution over topics for each document. The average distribution over the $K = 7$ topics for all pieces in the corpus is shown in Figure 11. The black error bars represent the standard error of the mean. The most prominent topics in the corpus are topic 3 (.192), topic 4 (.18) and topic 5 (.194), all of which are relatively close to the centre of the line of fifths.

Using topic modelling in the context of historical studies entails certain assumptions. As mentioned before, LDA is based on the bag-of-notes model and thus does not know the order of notes within a piece. Beyond that, it also does not have a concept for the order of pieces in the corpus, although some recent variants of the model attempt to incorporate chronological information (e.g., Blei and Lafferty, 2006; Zhu et al., 2016; Beykikhoshk et al., 2018). Under the basic LDA model, all pieces in the corpus are treated equally to infer the overall topics,

regardless of the time of their composition. This loosely corresponds to a synchronous perspective that a music theorist has when analysing pieces. Since dates of composition or publication dates of the pieces in the corpus are known, we are in a position to compare the topic distributions in the pieces diachronically to consider historical changes in these distributions.

To trace the *topic evolution* in the corpus, we calculate first the average topic distribution of all pieces for a given year for which we have pieces in the corpus. We moreover assume that the average topic distribution does not change if we do not have data for a year within the time range. Subsequently, we calculate a moving average with a window size of 35 years over the distributions that returns smoothed values for each topic while at the same time ensuring that the distributions per year always sum to 1. As mentioned above, we expect to see changes in tonality to be reflected in the historical development of the latent topics, in particular the increase in the usage of less common keys, as well as chromaticism and enharmonicism over the course of the 18th and 19th centuries.

The topic evolution over time for $K = 7$ topics is shown in Figure 12. The topic evolution plots for other values of $K \in \{3, 7, 12, 24\}$ are displayed in the ‘Supplemental material’. It can be seen that, while the overall most common topics 3 to 6 are prevalent in the earlier centuries, the other topics – in particular topics 1 and 2 – gain traction over the historical timeline and seem to stabilize in the 19th century, bearing witness to the fact that the tonal pitch-class vocabulary spreads out on the line of fifths (Moss et al., in press). The black line shows the smoothed normalized entropy of the topic distributions in each year. If in any year the topic distribution was uniform, the normalized entropy in this year would be equal to 1 before smoothing. The figure shows that the distributions tend to approach uniformity over time, which is reflected in the increasing entropy. In other words, the diverse topics become, on

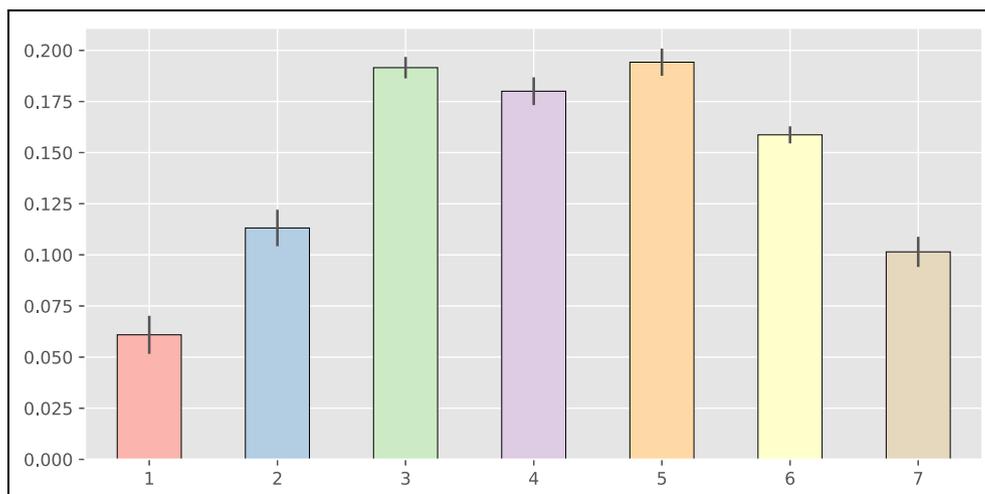


Figure 11. Average distribution of topics for all documents in the corpus for $K = 7$ topics. Error bars show the standard error of the mean.

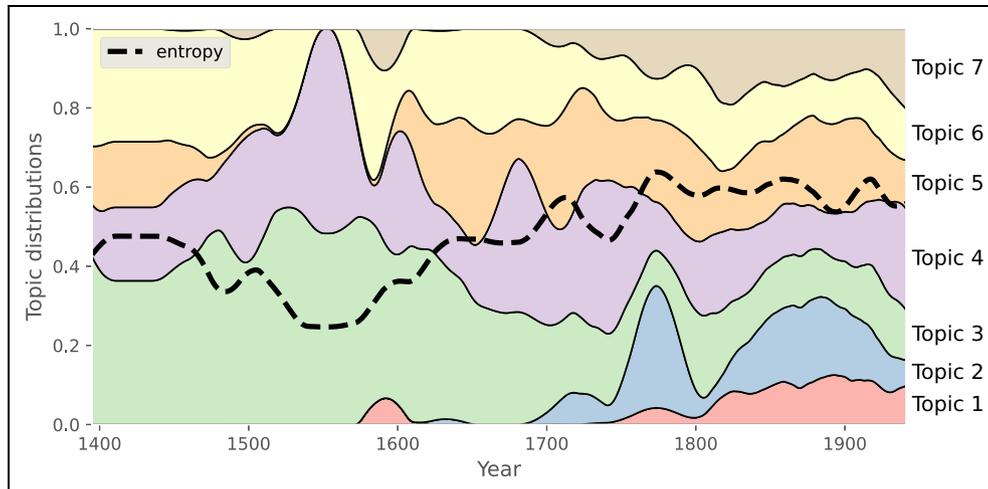


Figure 12. Topic evolution for $K = 7$ topics.

average, more and more similar to one another with respect to their frequencies of occurrence in pieces of a certain historical period. Note that this is not a mere corollary of the fact that more topics are used after ~ 1700 since the entropy is normalized for the number of topics in each year. The evolution of the latent topics does indeed corroborate the expectations for the historical development of tonality by showing that, over time, the number of used topics increases and is particularly strong in the 19th century.

The ‘Supplemental material’ shows the tonal pitch-class distributions, topic similarities, as well as average topic distributions and topic evolution plots for other values of $K \in \{3, 7, 12, 24\}$. As mentioned before, a larger number of topics increases their specificity. Indeed, a closer look at the tonal pitch-class distributions for a larger number K of topics indicates that topics become more pronounced and multimodal. This means that the probability of tonal pitch classes in these topics seems to be not only influenced by the arrangement of notes along the line of fifths, and that other intervals, in particular major and minor thirds, become more and more prevalent the larger the number of topics gets. This confirms our initial observation that latent tonal structures in music may be construed hierarchically, and the degree to which they resemble chords or keys, for instance, depends on the level of granularity.

Conclusion

In this study, we applied the LDA topic model to a historical corpus of Western classical musical pieces and showed that it can be used to infer music-theoretically meaningful topics. The obtained latent topics are well-interpretable from a music-theoretical perspective and correspond to contiguous segments on the line of fifths. The found topics roughly fall into two classes: topics that correspond approximately to diatonic collections of notes, and topics that represent chromaticism and less common keys. Subsequently, we observed the prominence of the topics in the corpus

diachronically and found that the number of topics increases over the course of history and that the two ‘chromatic topics’ are particularly prevalent in the 19th century.

This approach opens up a number of possible extensions. Drawing on the results from this study, it seems promising to extend the basic model by including our findings, for example, to restrict topics to line-of-fifths segments or to introduce a variable that decides whether the tonal/spelled or enharmonically equivalent pitch-class representation is more appropriate. Moreover, while the pitch dimension is often in the centre of music-theoretical analysis, incorporating other features to the basic LDA model, in particular, the duration of notes, formal sections of pieces, or rhythmical and metrical properties, as well as the incorporation of information about the sequential order of notes are promising directions for future research.

Acknowledgements

We are grateful to Markus Neuwirth, Christoph Finkensiep, and Gabriele Cecchetti for comments on an earlier version of this manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Swiss National Science Foundation within the project ‘Distant listening: The development of harmony over three centuries (1700–2000)’ (GA No. 182811). We thank Claude Latour for supporting this research through the Latour Chair in Digital Musicology at EPFL.

ORCID iD

Fabian C. Moss  <https://orcid.org/0000-0001-9377-2066>

Supplemental Material

Supplemental material is available at https://github.com/DCMLab/lda_tpcs.

Action Editor

David Meredith, Aalborg University, Department of Architecture, Design and Media Technology.

Peer Review

Emilios Cambouropoulos, Aristotle University of Thessaloniki, School of Music Studies.

Christof Weiß, Friedrich-Alexander-Universität Erlangen-Nürnberg, International Audio Laboratories Erlangen.

Notes

1. The distinction between a tonicization and a modulation might, for instance, not always be unambiguous in music analysis.
2. Note, however, that this definition is unrelated to other musicological definitions of a ‘topic’, for instance in Dickensheets (2012) or Johnson (2017), which are both situated in the context of topic theory (Ratner, 1980; Mirka, 2014b) that considers topics such as ‘Minuet’ or ‘March’ (called ‘types’), and topics such as ‘military’ or ‘hunting’ (called “styles”; Mirka, 2014a).
3. They use 235 pieces by six composers, namely Bach (1685–1750), Vivaldi (1678–1741), Mozart (1756–1791), Beethoven (1770–1827), Chopin (1810–1849), and Rachmaninoff (1873–1943). The data was gathered from <http://www.classicalmusicmidipage.com>.
4. <https://elvisproject.ca>
5. <http://kern.ccarh.org>
6. <http://cpdl.org>
7. <http://musescore.com>
8. <https://dcml.epfl.ch>
9. See also Tymoczko (2011) and the discussions in Harasim et al. (2021); Moss et al. (in press).
10. Since most applications of LDA work with textual data, we will use the terms ‘document’, ‘piece’ and ‘composition’ interchangeably.
11. The notation follows Blei (2012a), where $x_{1:N}$ is a shorthand notation for $\{x_1, \dots, x_N\}$.
12. See equation (1) in Blei (2012a).

References

- Albrecht, J., & Huron, D. (2014). A statistical approach to tracing the historical development of major and minor pitch distributions, 1400–1750. *Music Perception*, 31(3), 223–243.
- Albrecht, J., & Shanahan, D. (2013). The use of large corpora to train a new type of key-finding algorithm: An improved treatment of the minor mode. *Music Perception*, 31(1), 59–67.
- Aldwell, E., Schachter, C., & Cadwallader, A. (2010). *Harmony and voice leading* (4th ed.). Cengage Learning.
- Anzuoni, E., Ayhan, S., Dutto, F., McLeod, A., Moss, F. C., & Rohrmeier, M. (2021). A historical analysis of harmonic progressions using chord embeddings. In D. A. Mauro, S. Spagnol, & A. Valle (Eds.), *Proceedings of the 18th Sound and Music Computing Conference* (pp. 284–291). <https://doi.org/10.5281/zenodo.5038910>
- Baayen, R. H. (2001). *Word frequency distributions. Text speech and language technology*. Springer.
- Ballance, J. (2020). Pitch-class distributions in the music of Anton Webern. In CHR 2020: Workshop on Computational Humanities Research (pp. 214–224). Amsterdam, The Netherlands.
- Bellmann, H. G. (2012). *Categorization of tonal music style. A quantitative investigation* [PhD thesis, Griffith University, Queensland, Australia].
- Beykikhoshk, A., Arandjelović, O., Phung, D., & Venkatesh, S. (2018). Discovering topic structures of a temporally evolving document corpus. *Knowledge and Information Systems*, 55(3), 599–632.
- Binder, J. M. (2016). *Alien reading: Text mining, language standardization, and the humanities*. In M. K. Gold & L. F. Klein (Eds.), *Debates in the digital humanities 2016* (pp. 201–217). University of Minnesota Press. <https://doi.org/10.5749/j.ctt1cn6thb.21>
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Blei, D. M. (2012a). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.
- Blei, D. M. (2012b). Topic modeling and digital humanities. *Journal of Digital Humanities*, 2(1), 8–11.
- Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. In Proceedings of the 23rd International Conference on Machine Learning (pp. 113–120). Pittsburgh, Pennsylvania. <https://doi.org/10.1145/1143844.1143859>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Broze, Y., & Shanahan, D. (2013). Diachronic changes in jazz harmony: A cognitive perspective. *Music Perception*, 31(1), 32–45.
- Burgoyne, J. A., Wild, J., & Fujinaga, I. (2013). Compositional Data Analysis of Harmonic Structures in Popular Music. In J. Yust, J. Wild, & J. A. Burgoyne (Eds.), *Mathematics and computation in music* (Vol. 7937, pp. 52–63). Springer; Berlin Heidelberg. https://doi.org/10.1007/978-3-642-39357-0_4.
- Cambouropoulos, E. (2003). Pitch spelling: A computational model. *Music Perception*, 20(4), 411–429.
- Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Proceedings of the 22nd International Conference on Neural Information Processing Systems*, 288–296. Red Hook, NY, USA.
- Chew, E., & Chen, Y.-C. (2005). Real-time pitch spelling using the spiral array. *Computer Music Journal*, 29(2), 61–76.
- Cohn, R. (1998). Introduction to neo-Riemannian theory: A survey and a historical perspective. *Journal of Music Theory*, 42(2), 167–180.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed). Wiley.
- de Clercq, T. (2017). Interactions between harmony and form in a corpus of rock music. *Journal of Music Theory*, 61(2), 143–170.

- De Clercq, T., & Temperley, D. (2011). A corpus analysis of rock harmony. *Popular Music*, 30(1), 47–70.
- Dickensheets, J. (2012). The topical vocabulary of the nineteenth century. *Journal of Musicological Research*, 31, 97–137.
- Endres, D., & Schindelin, J. (2003). A new metric for probability distributions. *IEEE Transactions on Information Theory*, 49(7), 1858–1860.
- Faraldo, A., Gómez, E., Jordá, S., & Herrera, P. (2016). *Key estimation in electronic dance music*. In N. Ferro, F. Crestani, M.-F. Moens, J. Mothe, F. Silvestri, G. M. Di Nunzio, C. Hauff, & G. Silvello (Eds.), *Advances in information retrieval* (pp. 335–347). Springer International Publishing.
- Farbood, M. M. (2016). Memory of a tonal center after modulation. *Music Perception*, 34(1), 71–93.
- Foscarin, F., Audebert, N., & Fournier-S’niehotta, R. (2021). PKSpell: Data-driven pitch spelling and key signature estimation. In International Society for Music Information Retrieval Conference (ISMIR), Online, India.
- Gauvin, H. L. (2015). “The times they were A-changin’”: A database-driven approach to the evolution of harmonic syntax in popular music from the 1960s. *Empirical Musicology Review*, 10(3), 215–238.
- Goldstone, A., & Underwood, T. (2012). What can topic models of PMLA teach us about the history of literary scholarship?. *Journal of Digital Humanities*, 2(1). <http://journalofdigitalhumanities.org/2-1/what-can-topic-models-of-pmla-teach-us-by-ted-underwood-and-andrew-goldstone/>
- Griffiths, T. (2002). Gibbs sampling in the generative model of Latent Dirichlet Allocation. Technical report, Stanford University.
- Harasim, D., Moss, F. C., Ramirez, M., & Rohrmeier, M. (2021). Exploring the foundations of tonality: Statistical cognitive modeling of modes in the history of Western classical music. *Humanities and Social Sciences Communications*, 8(1), 1–11.
- Harris, Z. S. (1954). Distributional structure. *WORD*, 10(3), 146–162.
- Hauptmann, M. (1853). *Die natur der harmonik und der metrik*. Breitkopf und Härtel.
- Hedges, T., & Rohrmeier, M. (2011). *Exploring Rameau and beyond: A corpus study of root progression theories*. In C. Agon, M. Andreatta, G. Assayag, E. Amiot, J. Bresson, & J. Mandereau (Eds.), *Mathematics and computation in music* (pp. 334–337). Springer
- Hentschel, J., Neuwirth, M., & Rohrmeier, M. (2021). The annotated mozart sonatas: Score, harmony, and cadence. *Transactions of the International Society for Music Information Retrieval*, 4(1), 67–80.
- Herff, S. A., Harasim, D., Cecchetti, G., Finkensiep, C., & Rohrmeier, M. A. (2021). Hierarchical syntactic structure predicts listeners’ sequence completion in music. In Proceedings of the Annual Meeting of the Cognitive Science Society (vol. 43, pp. 903–909).
- Hu, D. J., & Saul, L. K. (2009a). *A probabilistic topic model for unsupervised learning of musical key-profiles*. In J. S. Downie & R. C. Veltkamp (Eds.), *Proceedings of the 10th international society for music information retrieval conference (ISMIR 2009)* (pp. 441–446) Kobe, Japan.
- Hu, D. J., & Saul, L. K. (2009b). *A probabilistic topic model for music analysis*. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Proceedings of the 22nd conference on neural information processing systems (NIPS 2009)* (pp. 1–4), Vancouver, Canada.
- Huang, P., Wilson, M., Mayfield-jones, D., Coneva, V., & Frank, M. (2017). The evolution of Western tonality: A corpus analysis of 24, 000 songs from 190 composers over six centuries. *SocArXiv*.
- Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. The MIT Press.
- Jockers, M. L., & Mimno, D. (2013). Significant themes in 19th-century literature. *Poetics*, 41(6), 750–769.
- Johnson, T. (2017). Tonality as topic: Opening a world of analysis for early twentieth-century modernist music. *Music Theory Online*, 23(4). <https://mtosmt.org/issues/mto.17.23.4/mto.17.23.4.johnson.html>
- Koelsch, S. (2012). *Brain and music*. John Wiley & Sons.
- Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), 15443–15448.
- Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. The MIT Press.
- Korzeniowski, F., & Widmer, G. (2018a). Genre-agnostic key classification with convolutional neural networks. In Proceedings of the 19th International Society for Music Information Retrieval Conference (pp. 264–270). Paris, France.
- Korzeniowski, F., & Widmer, G. (2018b). Improved chord recognition by combining duration and harmonic language models. In Proceedings of the 19th ISMIR Conference (pp. 10–17). Paris, France.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford University Press.
- Krumhansl, C. L. (2004). The cognition of tonality—as we know it today. *Journal of New Music Research*, 33(3), 253–268.
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89(4), 334–368.
- Lerdahl, F. (2001). *Tonal pitch space*. Oxford University Press.
- Lerdahl, F., & Jackendoff, R. S. (1983). *A generative theory of tonal music*. The MIT Press.
- Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception*, 24(4), 329–366.
- Lieck, R., Moss, F. C., & Rohrmeier, M. (2020). The tonal diffusion model. *Transactions of the International Society for Music Information Retrieval*, 3(1), 153–164.
- Lieck, R., & Rohrmeier, M. (2020). Modelling hierarchical key structure with pitch scapes. In Proceedings of the 21st International Society for Music Information Retrieval Conference. Montreal, Canada.
- MacKay, D. (2003). *Information theory, inference, and learning algorithms*. Montreal, Canada: Cambridge University Press.
- Mauch, M., MacCallum, R. M., Levy, M., & Leroi, A. M. (2015). The evolution of popular music: USA 1960–2010. *Royal Society Open Science*, 2(5), 150081.
- McLeod, A., & Rohrmeier, M. (2021). A modular system for the harmonic analysis of musical scores using a large vocabulary. In International Society for Music Information Retrieval Conference (ISMIR). (pp. 435–442).

- Meredith, D. (2006). The ps13 pitch spelling algorithm. *Journal of New Music Research*, 35(2), 121–159.
- Micchì, G., Gotham, M., & Giraud, M. (2020). Not all roads lead to rome: Pitch representation and model architecture for automatic harmonic analysis. *Transactions of the International Society for Music Information Retrieval*, 3(1), 42–54.
- Mirka, D. (2014a). Introduction. In D. Mirka (Ed.), *The Oxford handbook of topic theory* (pp. 1–60). Oxford University Press.
- Mirka, D. (ed.) (2014b). *The Oxford handbook of topic theory*. Oxford University Press.
- Mohr, J. W. (2013). Introduction–topic models: What they are and why they matter. *Poetics*, 41(6), 545–569.
- Moss, F. C. (2019). *Transitions of tonality: A model-based corpus study*. [Doctoral dissertation, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland].
- Moss, F. C., Neuwirth, M., Harasim, D., & Rohrmeier, M. (2019). Statistical characteristics of tonal harmony: A corpus study of beethoven’s string quartets. *PLoS One*, 14(6), e0217242.
- Moss, F. C., Neuwirth, M., & Rohrmeier, M. (2020a). Tonal Pitch-Class Counts Corpus (TP3C) (Version v1.0.0).
- Moss, F. C., Neuwirth, M., & Rohrmeier, M. (in press). The line of fifths and the co-evolution of tonal pitch-classes. *Journal of Mathematics and Music*.
- Moss, F. C., Souza, W. F., & Rohrmeier, M. (2020b). Harmony and form in Brazilian choro: A corpus-driven approach to musical style analysis. *Journal of New Music Research*, 49(5), 416–437.
- Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Springer International Publishing.
- Nakamura, E., & Kaneko, K. (2019). Statistical evolutionary laws in music styles. *Scientific Reports*, 9(1), 15993.
- Pfleiderer, M., Frieler, K., Abeßer, J., & Zaddach, W. G. and B, B., (eds), (2017). *Inside the Jazzomat: New perspectives for jazz research*. Schott. .
- Piper, A. (2018). *Enumerations: Data and literary study*. University of Chicago Press.
- Quinn, I., & White, C. W. (2017). Corpus-derived key profiles are not transpositionally equivalent. *Music Perception*, 34(5), 531–540.
- Ratner, L. G. (1980). *Classic music: Expression, form, and style*. Schirmer.
- Rhody, L. M. (2012). Topic modeling and figurative language. *Journal of Digital Humanities*, 2(1). <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-figurative-language-by-lisa-m-rhody/>
- Rodrigues Zivic, P. H., Shifres, F., & Cecchi, G. A. (2013). Perceptual basis of evolving western musical styles. *Proceedings of the National Academy of Sciences of the United States of America*. 110(24), 10034–10038. <https://doi.org/10.1073/pnas.1222336110>
- Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1), 35–53.
- Rohrmeier, M. (2020). The syntax of jazz harmony: Diatonic tonality, phrase structure, and form. *Music Theory and Analysis (MTA)*, 7(1), 1–63.
- Rohrmeier, M., & Cross, I. (2008). *Statistical properties of tonal harmony in Bach’s chorales*. In K. Miyazaki (Ed.), *Proceedings of the 10th international conference on music perception and cognition (ICMPC 2008)* (pp. 619–627) Sapporo, Japan.
- Rohrmeier, M. A., & Moss, F. C. (2021). A formal model of extended tonal harmony. In Proceedings of the 22nd International Society for Music Information Retrieval Conference, Onlinepp. 569–578.
- Rom, U. B. (2011). *Tonartbezogenes denken in mozarts werken unter besonderer berücksichtigung des instrumentalwerks*. [Doctoral dissertation, Technische Universität Berlin, Berlin, Germany].
- Salzer, F. (1952). *Structural hearing: Tonal coherence in music*. Dover.
- Sapp, C. S. (2001). Harmonic Visualizations of Tonal Music. *International Computer Music Conference Proceedings*, 2001. San Francisco: International Computer Music Association.
- Sapp, C. S. (2005). Visual hierarchical key analysis. *Computers in Entertainment*, 3(4), 1–19.
- Schenker, H. (1935). *Neue musikalische Theorien und Phantasien III. Der freie Satz*. Universal Edition.
- Schlenker, P. (2017). Outline of music semantics. *Music Perception*, 35(1), 3–37.
- Schmidt, B. M. (2012). Words alone: Dismantling topic models in the humanities. *Journal of Digital Humanities*, 2(1), 49–65.
- Serrà, J., Corral, L., Bogañá, M., Haro, M., & LI Arcos, J. (2012). Measuring the evolution of contemporary western popular music. *Scientific Reports*, 2(521), 1–6.
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (eds.), *Latent Semantic Analysis: A Road to Meaning* (pp. 427–448). Lawrence Erlbaum Associates Publishers.
- Stoddard, J., Raphael, C., & Utgoff, P. E. (2004). Well-tempered spelling: A key-invariant pitch spelling algorithm. In Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004). 1–6. Barcelona, Spain. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.100.9663>
- Tan, I., Lustig, E., & Temperley, D. (2019). Anticipatory syncopation in rock: A corpus study. *Music Perception*, 36(4), 353–370.
- Tang, J., Meng, Z., Nguyen, X., Mei, Q., & Zhang, M. (2014). Understanding the limiting factors of topic modeling via posterior contraction analysis. In *International Conference on Machine Learning*, Proceedings of Machine Learning Research (PMLR) Beijing, China. 190–198. <https://proceedings.mlr.press/v32/tang14.html>
- Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical dirichlet processes. *Journal of the American Statistical Association*, 101(476), 1566–1581.
- Temperley, D. (1999). What’s key for key? the krumhansl-schmuckler key-finding algorithm reconsidered. *Music Perception*, 17(1), 65–100.
- Temperley, D. (2000). The line of fifths. *Music Analysis*, 19(3), 289–319.
- Temperley, D. (2001). *The cognition of basic musical structures*. The MIT Press.
- Temperley, D., & de Clercq, T. (2013). Statistical analysis of harmony and melody in rock music. *Journal of New Music Research*, 43(2), 187–204.

- Tillmann, B., & Bigand, E. (2004). The relative importance of local and global structures in music perception. *The Journal of Aesthetics and Art Criticism*, 62(2), 211–222.
- Tymoczko, D. (2011). *A geometry of music: Harmony and counterpoint in the extended common practice*. Oxford University Press.
- Underwood, T. (2014). The quiet transformations of literary studies: What thirteen thousand scholars could tell us. *New Literary History*, 45(3), 359–384.
- van Balen, J. M. H. (2016). *Audio description and corpus analysis of popular music* [Doctoral Dissertation, Utrecht University, Utrecht, Netherlands].
- Viaccoz, C., Harasim, D., Moss, F. C., & Rohrmeier, M. (in press). Wavescapes: A visual hierarchical analysis of tonality using the discrete Fourier transformation. *Musicae Scientiae*.
- Wallach, H. M., Murray, I., Salakhutdinov, R., & Mimno, D. (2009). Evaluation methods for topic models. In Proceedings of the 26th Annual International Conference on Machine Learning (ICML 2009) (pp. 1105–1112). New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/1553374.1553515>
- Weiß, C., Mauch, M., & Dixon, S. (2018). Investigating style evolution of western classical music: A computational approach. *Musicae Scientiae*, 23(4), 486–507.
- Weiß, C., Schreiber, H., & Mueller, M. (2020). Local key estimation in music recordings: A case study across songs, versions, and annotators. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 28, 2919–2932. <https://doi.org/10.1109/TASLP.2020.3030485>
- White, C. (2013). *Some statistical properties of tonality, 1650–1900*. [Dissertation, Yale University].
- Yust, J. (2019). Stylistic information in pitch-class distributions. *Journal of New Music Research*, 48(3), 217–231.
- Zhu, M., Zhang, X., & Wang, H. (2016). A LDA based model for topic evolution: Evidence from information science journals. In Proceedings of the 2016 International Conference on Modeling, Simulation and Optimization Technologies and Applications (MSOTA 2016) (pp. 49–54). Xiamen, China: Atlantis Press.