Description of Chord Progressions by Minimal Transport Graphs Using the System Contrast Model

Corentin Louboutin, Frédéric Bimbot

Description of Chord Progressions by Minimal Transport Graphs Using the System & Contrast Model

Corentin Louboutin
Université de Rennes 1 / IRISA
corentin.louboutin@irisa.fr

Frédéric Bimbot
CNRS / IRISA
frederic.bimbot@irisa.fr

ABSTRACT

In this paper, we model relations between chords by minimal transport *and we investigate different types of relations within chord sequences. For this purpose, we use the "System & Contrast" (S&C) model [1, 2], designed for the description of music segments, to infer non-sequential structures called* chord progression graphs *(CPG). Minimal transport is defined as the shortest displacement of notes, in semitones, between a pair of chords. The paper presents three algorithms to find CPGs for chords sequences: one is sequential, and two others are based on the S&C model. The three methods are compared using the perplexity as an efficiency measure. The experiments on a corpus of* 45 *segments taken from songs of multiple genres, indicate that optimization processes based on the S&C model outperform the sequential model with a decrease in perplexity over* 1.0*.*

1. INTRODUCTION

One of the topics of major interest in Music Information Retrieval (MIR) is to understand how elements are related to one another in a music piece. For this purpose, some studies use principles from formal language theories [3, 4, 5], some others formalize notions from conventional musicology [6, 7] and another branch in music information retrieval is mainly based on probabilistic models [8, 9].

Recently Bimbot et al. designed the "System & Contrast" model [1, 2] to describe music at the scale of phrases and *sections*, i.e. segments of 12 to 25 seconds, typically from songs. The S&C model is a multidimensional model which can be applied to melody, harmony, rhythm or any other musical dimension. The S&C model is based on the idea that relations between musical elements are not essentially sequential and that they can be infered on the basis of an economy principle. We focus here on the application of this model to the description of chord progression structures.

The study presented in this paper is based on the notion of *minimal transport* which is used to model the relation between two chords. It is defined as the set of connections between the notes of the two chords such that the sum of intervals (in semitones) resulting from the connections is minimal. As such, the notion of minimal transport can be seen as a computational approximation of "voice leading" as described by Cohn [10] or Tymoczko [11, 12]. However, minimal transport is here extended to also infer non-sequential structures which is a way to describe how chords are related, to one another, while relaxing the sequentiality hypothesis.

It was observed in Deruty et al. [13] that it is possible to create a multi-scale segment structure using the S&C model at different scales simultaneously. The present paper investigates the computational potential of this hypothesis for minimal transport graph search.

In Section 2.3, we define the notion of chord progression graph (CPG) and minimal transport graph (MTG), and we briefly recall the square form of the System & Contrast model. We then describe in Section 3 three optimization algorithms, one sequential and two based on the S&C model, to compute a minimal transport chord sequence. Finally, in Section 4, we present an experimental comparison of these three optimization methods, in terms of perplexity.

2. KEY CONCEPTS

2.1 Definitions

A *chord sequence* can be defined as the *in extenso* representation of all chords observed in a segment at specific metric positions and ordered by time. A chord is itself represented by the set of pitch classes (pc) of each note composing it.

A *chord progression graph* (CPG) is a pair (S, M) where S is a sequence of chords and M is the *model structure* of relations between the chords, that is the set of links between them. Two kinds of CPGs are considered in this paper:

- *sequential* CPGs which are based on the sequential description of the chord sequence. For these graphs, each link defines a relation between a chord and the chord appearing just after, in the chord sequence.

- *systemic* CPGs, based on the S&C model described in subsection 2.3 for which relations between chords are causal but not necessarily sequential.

While for sequential CPGs the *antecedent* of a chord is its immediate predecessor, it can be some other previous chord for systemic CPGs. In both cases we make the hypothesis that a given chord, S_i, depends only on one antecedent, $\Phi(S_i)$, itself of the chord type. Using a probabilist point of view, we can use Φ to define an approxima-

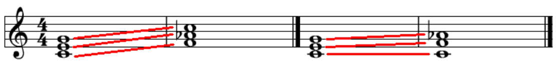

Figure 1. Example of transports between the two chords C and Fm. The first is the $\{(0,5),(4,8),(7,0)\}$ and the second is $\{(0,0),(4,5),(7,8)\}$

tion of $P(S_i|S_{i-1}\dots S_0)$:

$$P(S_i|S_{i-1}\dots S_0) \approx P(S_i|\Phi_M(S_i)) \tag{1}$$

where M denotes the model structure of the CPG.

For the sequential CPG, $\Phi_{Seq}(S_i) = S_{i-1}$, and this is equivalent to a first order Markov approximation. When Φ_M is deterministic, the CPG (S, M) can be denoted by the pair (S, Φ_M).

2.2 Minimal Transport

A chord P is represented by a set of m_p pitch classes p_i: $P = (p_i)_{0\leq i\leq m_p}$. A *transport* between P and an other chord $Q = (q_j)_{0\leq j\leq m_q}$ is a set :

$$T = \{(p_k, q_k)|p, q \in [\![0; 11]\!], k \in [\![0; n]\!]\} \tag{2}$$

where n is the number of connections or voices. Indeed a transport can be seen as a way to associate voices to notes in the two chords. We focus here on complete transports, i.e. each note is associated to at least one voice. Examples of such transports are given on Figure 1.

The optimality of a transport between two chords is defined by the *taxicab norm* or *smoothness* [14]. That is, for a transport T:

$$|T| = \sum_{(p,q)\in T} |d(p,q)| \tag{3}$$

where

$$d(p,q) = ((q - p + 5) \pmod{12}) - 5 \tag{4}$$

The term $d(p,q)$ is the shortest displacement in semitones from pitch class p to pitch class q (with $d(p,q) \in [\![-5; 6]\!]$).

In Figure 1 the second transport is minimal. A *minimal transport graph* (MTP) is an instantiation of a CPG (S, M) where all transports associated with M are fixed and their sum is minimal.

2.3 System & Contrast Model

The System & Contrast (S&C) model [2] is a (meta-)model of musical data based on the hypothesis that the relations between musical elements in a segment are not necessarily sequential. Initially designed for the description of phrase structure for annotation purposes [1], the S&C model has been further formalized as a generalization and an extension of Narmour's implication-realisation model. Its applications to various music genres for multidimensional and multiscale description has been explored in [13]. Our aim, here is to give a computational elaboration of this model.

The principles of the S&C model is that relations between elements in a musical segment create a *system* of matricial expectations which can be more or less strongly denied by the last element called *contrast*. The first element of the system is called *primer* and plays a particular role in the construction of the expectation system. The contrast acts as a closure to the segment. In this paper, we focus on *square* systems, i.e. systems of four elements.

2.3.1 Formalization

A sequence of four elements $(x_i)_{0\leq i\leq 3}$ can be arranged as a square matrix:

$$X = \begin{bmatrix} x_0 & x_1 \\ x_2 & x_3 \end{bmatrix} \tag{5}$$

Assuming two relations f and g between the primer x_0 and its neighbors in X, we have:

$$x_1 = f(x_0) \quad \text{and} \quad x_2 = g(x_0)$$

Note that these two relations may apply only to a subset of the properties characterizing the elements of the system.

The S&C model envisions the fourth element x_3 in relation to a *virtual* projected element $\hat{x_3}$ which would result from the combination of f and g: The disparity between $\hat{x_3}$ and the actual (observed) x_3 is modeled by a *contrast* function γ:

$$\hat{x_3} = f(g(x_0)) \tag{6}$$

$$x_3 = \gamma(\hat{x_3}) \tag{7}$$

The description of a S&C is the quadruplet (x_0, f, g, γ) which can be used as a compact representation of the segment. It can be viewed as a minimal description in the sense of the Kolmogorov complexity [15] in line with several other works in MIR [16, 17, 18].

For a chord sequence $(S_i)_{0\leq i\leq 3}$ modeled as a S&C, the antecedent function $\Phi_{S\&C}$ (see Eq. 1) is defined as follows:

$$\Phi_{S\&C} : \begin{array}{l} S_1 \longmapsto S_0 \\ S_2 \longmapsto S_0 \\ S_3 \longmapsto \hat{S}_3 \end{array} \tag{8}$$

Under the minimal transport approach, f, g and γ are complete transports.

2.3.2 Multiscale S&C

Musical phrases and sections generally contain time varying chord information which can be sampled at specific intervals, for instance downbeats. In this work, chord progressions are assumed to be composed of 16 elements, for instance:

$$Cm\ Cm\ Cm\ Bb\ Ab\ Ab\ Ab\ Gm\ F\ F\ F\ F\ Cm\ Cm\ Bb\ Bb$$

The S&C model can be used to model such sequences by extending it to a multiscale framework [13].

A multiscale CPG is a structure that combines elementary sub-CPGs built from square S&Cs. Figure 2 represents a view of the above chord sequence explained on several scales simultaneously as a hypercube or *tesseract*. Bolded chords are contrastive elements. In the first system $[Cm, Cm, Cm, Bb]$, Bb contrasts with the expectation: $Cm + Cm + Cm \to Cm$. In the second group, Gm denies $Ab{+}Ab{+}Ab \to Ab$. The last F is a non-contrastive chord in the $[F, F, F, F]$ group. Ultimately, the sequence concludes by $Cm + Cm + Bb \to Bb$.

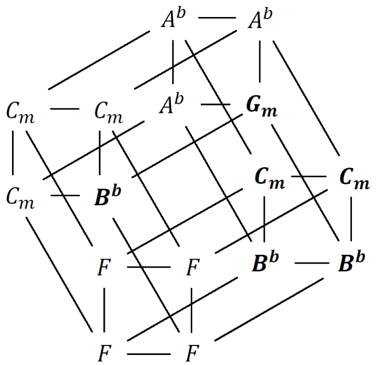

Figure 2. Tesseractic representation of the chord sequence of the chorus of *Master Blaster* by Stevie Wonder.

But many other systems can also be considered. For instance chords number $[0, 1, 4, 5]$ form a non-contrastive sub-system $[Cm, Cm, Ab, Ab]$, while chords $[8, 10, 12, 14]$ form a contrastive sub-system, $[F, F, Cm, Bb]$, etc. In fact, any quadruplet of adjacent vertices forming a square in the tesseract can be considered as a S&C. This results in a graph of implications which describes the chord sequence in a multiscale fashion.

3. APPLICATION TO CHORD PROGRESSION ANALYSIS

Finding the Minimum Transport Graph (MTG) on a chord sequence is an optimization problem. It consists in finding the global transport graph whose transport cost is minimal. In this section, we present three structure models designed for 16-chord sequences, and the corresponding optimization algorithms: namely the sequential model (Seq), the bi-scale model ($SysP$) and the dynamic scale model ($SysDyn$).

Each optimization process described below explores the space of all transport graphs corresponding to a CPG and chooses the solution with the minimal global cost.

3.1 Sequential Model

The sequential model corresponds to the conventional point of view where each chord is related to its direct predecessor, i.e. $\Phi_{Seq}(S_i) = S_{i-1}$.

As the number of possible transport graphs grows very fast ($\mathcal{O}(n!^l)$) with the length of the chord sequence l and the number of voices n (as defined in Equation 2), we divide the transport graph optimization in local optimizations of complexity $\mathcal{O}(n!^{\frac{l}{4}+1})$ on four sub-graphs. The first optimization is the search of the minimal transport graph corresponding to the CPG $([S_i]_{0\leq i\leq 3}, \Phi_{Seq})$. Then, the algorithm builds a second CPG using the last chord of the previous optimization and the four next chords of S, that is $([S_i]_{3\leq i\leq 7}, \Phi_{Seq})$, and searches for the corresponding MTP. This step is iterated on sequences $[S_i]_{7\leq i\leq 11}$ and $[S_i]_{11\leq i\leq 15}$.

Figure 3 represents the global model structure of the sequential model. Each transport links a chord with the next chord in the sequence and the graph is optimized by groups of four or five chords.

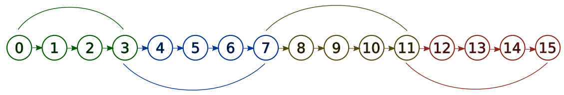

Figure 3. Representation of the transport links of the sequential model. Each colored sub-graph represents one optimization. Numbers are the chord indexes in the initial sequence.

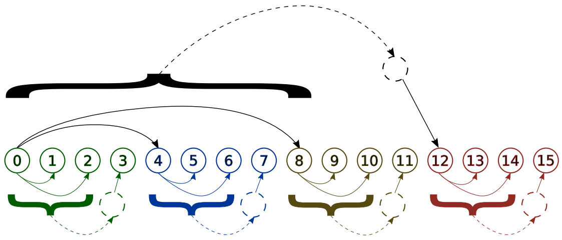

Figure 4. Representation of the transport links of the bi-scale model. In black, the upper scale system, while the colored systems are the lower scale systems. Numbers are the chord indexes in the initial sequence.

3.2 Static Bi-Scale Model

The second structure model is based on a multiscale vision of the S&C model as described in Section 2.3.2. A sequence of 16 chords can be structured in a S&C of four disjoint nested sub-CPGs with a S&C structure, in other words, a S&C of four S&Cs. Under this approach:

- An upper scale CPG models the systemic relations between the first elements of the four lower scale CPGs:
  $([S_0, S_4, S_8, S_{12}], \Phi_{S\&C})$.
- The four lower scale CPGs describe the structure of four disjoint parts of the segment:
  $([S_{4*i+j}]_{0\leq j\leq 3}, \Phi_{S\&C})_{0\leq i\leq 3}$.

The global bi-scale model is represented on Figure 4. The function of the upper scale CPG is to ensure the global coherence of the description.

As for Seq model, each optimization has to be computed separately to reach a reasonable computing time ($\mathcal{O}(n!^{\frac{l}{4}})$).

As relations in a S&C are matricial, they become tensorial at the multiscale level (see Figure 2), and it is therefore interesting to consider *permutations* of the initial sequence such that each of the five S&Cs corresponds to a square in the tesseract and each point of the tesseract appears only in one lower scale CPG (see Section 2.3.2).

Moreover, to ensure that each CPG can be described using a S&C model structure, chord indexes of each CPG have to correspond to a quadruplet forming an adjacent square in the tesseract view (see Figure 2). There are only 36 possibilities of such permutations which respect local causality inside each CPG.

As a system of 4 elements, $abcd$, is equivalent to its dual, $acbd$, due to the fact that both b and c are related commutatively to a in the MTG approach. Using this equivalence on the upper CPG, it is possible to reduce the 36 permutations to 30 equivalence classes [1]. For example, the permutation $[0, 1, 2, 3, 8, 9, 10, 11, 4, 5, 6, 7, 12, 13, 14, 15]$ is equivalent to the one represented on Figure 4. A bi-scale model structure associated with a permutation number x

[1] The list of permutations is given in [19]

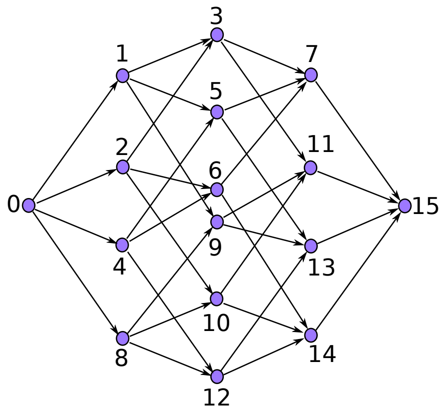

Figure 5. Projection of the tesseract where each chord of a same column has the same contrastive function. Numbers are the chord indexes in the initial sequence.

is denoted as $SysPx$ ($SysWPx$ in the case where $\hat{x_3}$ is replaced by x_0).

3.3 Dynamic Scale Model

3.3.1 Principle

This third structure model, denoted as $SysDyn$, is also based on the S&C model and the tesseractic representation of the chord sequence. But, while the arrangement of nested systems is fixed by the permutation in the bi-scale model, the dynamic model considers a wider range of combinations. Figure 5 represents the tesseract in a way such that, each column aligns the chords having the same contrastive depth in the sequence (i.e. they are contrastive elements for a same number of systems). The first column contains only the primer, the second column contains the secondary primers $(1, 2, 4, 8)$ (which are not contrastive elements of any system), then on the third column, the contrastive elements of only one system $(3, 5, 6, 9, 10, 12)$. Then, elements 7, 11, 13 and 14 can act as contrastive elements of three systems and the final element (15) is potentially contrastive in six systems.

The principle of the dynamic method is to optimize on the fly the sub-CPGs which contribute to the MTG of the overall chord progression. For instance, chord 11 is hypothesized as the contrast of sub-CPGs:

- $([S_1, S_3, S_9, S_{11}], \Phi_{S\&C})$
- $([S_2, S_3, S_{10}, S_{11}], \Phi_{S\&C})$
- $([S_8, S_9, S_{10}, S_{11}], \Phi_{S\&C})$

Among these three possibilities, the one yielding the minimal transport graph is selected dynamically as the local structure within the global description. Therefore, this requires a two level optimization process: one for the search of the best sub-CPG that "explains" a contrastive element and one for the transport graph of each sub-CPG.

3.3.2 Handling optimization conflicts

To prevent optimization conflicts when two different CPGs contain the same relation (e.g. (S_0, S_1) in $[S_0, S_1, S_2, S_3]$ and $[S_0, S_1, S_8, S_9]$), each transport is fixed at the optimization of the first CPG in which it appears. It implies that when optimizing the CPG, $([S_0, S_1, S_8, S_9], \Phi_{S\&C})$, the transport considered for (S_0, S_1) is fixed and is the one considered in the minimal transport graph associated with the CPG: $([S_0, S_1, S_2, S_3], \Phi_{S\&C})$.

Moreover this constraint also applies to the voices associated with each note. If a former optimization step has determined the voices relating two chords, the transport between these two chords is kept fixed for the forthcoming sub-CPGs optimizations. For example, once the optimization on CPG $([S_0, S_1, S_8, S_9], \Phi_{S\&C})$ has been achieved, the voices associated with pitch classes of S_1 and S_9 fixes the transport between these two chords for later optimizing the CPG $([S_1, S_5, S_9, S_{13}], \Phi_{S\&C})$. In the current implementation of the algorithm, the CPG optimizations are carried out in ascending order of the index of the contrastive element of the CPGs, which preserves causality.

4. EXPERIMENTS

4.1 Data

In this section, we present experimental results on the behaviour of the proposed models on a dataset of 45 structural sections from a variety of songs, reduced to 16 (downbeat sychronous) chord sequences, including artists such as Miley Cyrus, Edith Piaf, Abba, Pink Floyd, Django Reinhardt, Eric Clapton, Rihanna, etc [2].

4.2 Evaluation

4.2.1 NLL Score

As there exists no ground truth as of the actual structure of the chord sequences, we compare the different models with regards to their ability to predict the entire chord sequence in the CPG framework. This is done by calculating a *perplexity* [20] for each model derived from the *negative log-likelihood*, denoted as NLL_M.

The NLL_M of a transport graph is defined as the arithmetic mean of the NLL_M of each voice inferred by the transport graph. Let $X = (x)_{0 \leq i \leq n-1}$ be the sequence of pitch classes of a "voice", considering the first-order approximation defined in Section 2.1, the NLL_M associated with a CPG M, is defined as:

$$NLL_M(X) = -\frac{\log p(x_0) + \sum_{d \in D_M} \log p(d)}{|D_M| + 1} \tag{9}$$

where D_M is the set of pitch class displacements in semitones in the voice considering the CPG structure model, $|D_M|$ is the size of the set [3], and $p(d)$ is the estimated probability of the displacement d.

4.2.2 Probability Estimation

In this work, $p(d)$ is estimated as:

$$p(d) = \frac{1 + N(d, C_M)}{12 + \sum_{z=-5}^{6} N(z, C_M)} \tag{10}$$

where C_M is the description of the training corpus with the model M (using a leave-one-out cross-validation strategy),

[2] The full list of chord sequences is presented in [19].

[3] For $SysDyn$, if a displacement is used in two sub-CPGs, the displacement is counted twice for the likelihood.

	$\hat{x_3} = f(g(x_0))$	$\hat{x_3} = x_0$
Seq	3.84	
$BestSysP$	2.73	2.95
$SysP8$	3.17	3.64
$SysDyn$	2.80	

Table 1. Average perplexity obtained by the different models.

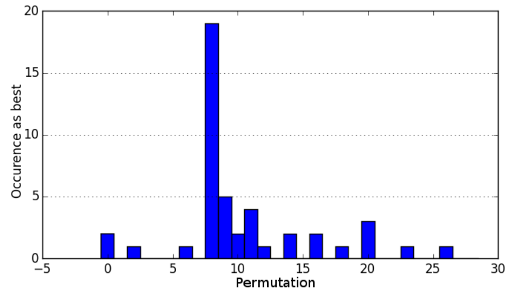

Figure 7. Histogram of the top ranking permutations, in terms of perplexity, across the 45 chord sequences.

and $N(d, C_M)$ the number of occurences of displacements d observed in C_M. $p(d)$ is an estimation of $P(x|\Phi_M(x))$ where $d(\Phi_M(x), x) = d$.

We hypothetize the *a priori* uniformity in the distribution of the initial notes and therefore estimate $p(x_0) = \frac{1}{12}$ which preserve the comparability between the models.

4.2.3 Perplexity

We convert a NLL into a perplexity value defined as:

$$PP_M(S) = 2^{NLL_M(S)} \tag{11}$$

which can be interpreted as the average probabilistic branching factor between successive notes in the graph.

4.3 Results

Figure 6 depicts a comparison between Seq, $BestSysP$, and $SysDyn$ models for each of the 45 chord sequences where $BestSysP$ is defined as the optimal permutation of the $SysP$ configuration for each song individually.

Table 1 summarizes the results with the three types of models. While the sequential model (Seq) provides a perplexity of 3.84, it is clearly outperformed by both the bi-scale model and the dynamic model, 2.73 and 2.80 respectively, i.e. more than 1.0 perplexity difference.

It is worth noting from Figure 7 that permutation 8 (represented on Figure 8) is the optimal permutation for 19 songs out of 45 (i.e. 42%). An explanation of the success of this permutation can be that it considers implicitly three types of scale relations: short, medium and long. The upper scale optimization maximizes the coherence of the first half of the chord sequence, while lower scale optimizations combine local and distant relations.

In the context of the bi-scale model, the role of the virtual element, $\hat{x_3}$ in $SysPx$ has been investigated experimentally by substituting it with the primer x_0 in the CPGs, in order to compare both of them as predictors of x_3. The second column of Table 1 shows a clear advantage of the virtual element which comforts the idea of its implicative role in the S&C model. However, there are 5 chord sequences for which x_0 is significantly better than $\hat{x_3}$. This may happen when the last element falls back on the primer

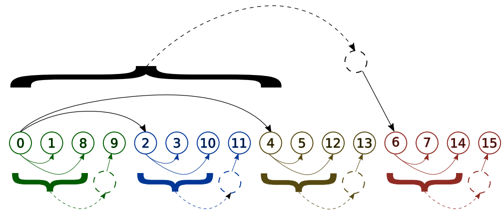

Figure 8. CPG of the bi-scale model, using permutation 8: $[0, 1, 8, 9, 2, 3, 10, 11, 4, 5, 12, 13, 6, 7, 14, 15]$.

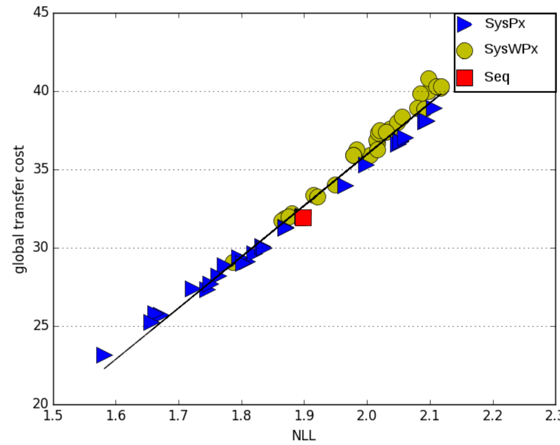

Figure 9. Graph of average transport cost, function of average NLL on 45 sequences corpus for Seq, $SysP0 \ldots 29$ and $SysWP0 \ldots 29$.

or if the virtual element contains pitch classes which do not belong to the "tonality" of the segment.

As optimizing the transport cost between chords minimizes the average pitch class displacement, there are only few intervals capturing most of the NLL. This raises the idea, as Figure 9 shows, that there is a correlation (0.990) between the global transport cost and the NLL. This may indicate that the distribution of the displacement distances is somehow exponentially decreasing. It would therefore be interesting to investigate how replacing the trained probability estimations by a Laplacian law would affect the results.

Finally, $SysDyn$ happens to perform equivalently well to $BestSysP$ but with a much faster computation time. The optimal model structure can be traced back *a posteriori*. Interestingly, a chord that is contrastive in a CPG can then be used in a new CPG to build the expectation for a subsequent contrastive "surprise". In a sense it can be seen as a similar notion to that of "resolution" in conventional musicology [21, 10]—with the difference that, here, the resolution is realised from a virtual chord.

In summary, this first set of results shows that considering non-sequential relations between chords seem relevant to provide an efficient description of chord progressions.

5. CONCLUSIONS AND PERSPECTIVES

The approach presented in this paper is based on minimal transport to model relations between chords. Three optimization algorithms have been presented and tested on a corpus of 45 sequences of 16 chords using perplexity as an efficiency measure. The two methods based on the S&C model substantially outperform the sequential approach.

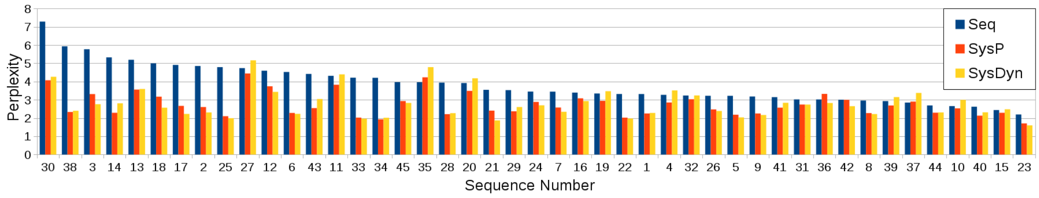

Figure 6. Perplexity obtained for each of the 45 chord sequences by: Seq (sequential model), $BestSysP$ (optimal bi-scale permutation for each song) and $SysDyn$ (dynamic model).

These results constitute a strong incentive to further consider the use of the S&C model in MIR.

The S&C model could also prove to be useful in musicology: in particular, the virtual element considered by the S&C model seems to play a relevant role. It may have a similar function to that of the augmented triad in Cohn's theory [10], that is, a passage chord which can be "invisible" in the observed sequence. Future studies could investigate how the definition of the virtual element affects the MTG optimization and how to constraint transports to comply with musicological rules.

Furthermore, we focused here only on the chord dimension of music, but the System & Contrast model can handle other dimensions such as melody, rhythm, etc, which will be a subject for future investigations.

Acknowledgments

This work has greatly benefited from initial scientific investigations carried out with Anwaya Aras, during her internship at IRISA, in 2013.

6. REFERENCES

[1] F. Bimbot, E. Deruty, G. Sargent, and E. Vincent, "Semiotic structure labeling of music pieces: concepts, methods and annotation conventions," in *Proc. ISMIR*, 2012.

[2] F. Bimbot, E. Deruty, G. Sargent, and E. Vincent, "System & Contrast : A Polymorphous Model of the Inner Organization of Structural Segments within Music Pieces," *Music Perception*, vol. 33, pp. 631–661, June 2016. Former version published in 2012 as Research Report IRISA PI-1999, hal-01188244.

[3] B. de Haas, J. P. Magalhaes, R. C. Veltkamp, and F. Wiering, "Harmtrace: Improving harmonic similarity estimation using functional harmony analysis.," in *Proc. ISMIR*, pp. 67–72, 2011.

[4] W. B. De Haas, M. Rohrmeier, R. C. Veltkamp, and F. Wiering, "Modeling harmonic similarity using a generative grammar of tonal harmony," *Proc. ISMIR*, 2009.

[5] M. Rohrmeier, "A generative grammar approach to diatonic harmonic structure," in *Proceedings of the 4th Sound and Music Computing Conference*, pp. 97–100, 2007.

[6] M. Giraud, R. Groult, and F. Levé, "Computational analysis of musical form," in *Computational Music Analysis*, pp. 113–136, Springer, 2016.

[7] M. Giraud, R. Groult, and F. Levé, "Subject and counter-subject detection for analysis of the well-tempered clavier fugues," in *From Sounds to Music and Emotions*, pp. 422–438, Springer, 2013.

[8] R. Scholz, E. Vincent, and F. Bimbot, "Robust modeling of musical chord sequences using probabilistic n-grams," in *ICASSP 2009*, pp. 53–56, IEEE, 2009.

[9] D. Conklin, "Multiple viewpoint systems for music classification," *Journal of New Music Research*, vol. 42, no. 1, pp. 19–26, 2013.

[10] R. Cohn, *Audacious Euphony: Chromatic Harmony and the Triad's Second Nature*. Oxford University Press, 2011.

[11] D. Tymoczko, "The geometry of musical chords," *Science*, vol. 313, no. 5783, pp. 72–74, 2006.

[12] D. Tymoczko, "Scale theory, serial theory and voice leading," *Music Analysis*, vol. 27, no. 1, pp. 1–49, 2008.

[13] E. Deruty, F. Bimbot, and B. Van Wymeersch, "Methodological and musicological investigation of the System & Contrast model for musical form description," Research Report RR-8510, INRIA, 2013. hal-00965914.

[14] J. N. Straus, "Uniformity, balance, and smoothness in atonal voice leading," *Music Theory Spectrum*, vol. 25, no. 2, pp. 305–352, 2003.

[15] P. M. Vitányi and M. Li, "Minimum description length induction, Bayesianism, and Kolmogorov complexity," *IEEE Trans. Information Theory*, vol. 46, no. 2, pp. 446–464, 2000.

[16] P. Mavromatis, "Minimum description length modelling of musical structure," *Journal of Mathematics and Music*, vol. 3, no. 3, pp. 117–136, 2009.

[17] D. Temperley, "Probabilistic models of melodic interval," *Music Perception*, vol. 32, no. 1, pp. 85–99, 2014.

[18] C. Louboutin and D. Meredith, "Using general-purpose compression algorithms for music analysis," *Journal of New Music Research*, 2016.

[19] C. Louboutin and F. Bimbot, "Tensorial Description of Chord Progressions - Complementary Scientific Material," IRISA, 2016. hal-01314493.

[20] P. F. Brown, V. J. D. Pietra, R. L. Mercer, S. A. D. Pietra, and J. C. Lai, "An estimate of an upper bound for the entropy of english," *Computational Linguistics*, vol. 18, no. 1, pp. 31–40, 1992.

[21] A. Forte, *Tonal harmony in concept and practice*. Holt, Rinehart and Winston, 1974.