Similarity measures for tonal models

Richard R. Randall
Department of Music
University of Massachusetts, Amherst.
randall@music.umass.edu.

Bilal Khan
Department of Mathematics
John Jay College of Criminal Justice

ABSTRACT

Multidimensional-scaling models of tonal hierarchy encode cognitive relationships between chords as euclidean distances. Fred Lerdahl's tonal pitch space (TPS) model approximates cognitive perceptual relations between chords by providing a combinatorial procedure for computing the distance value between two chords. Because of the influence of experimental data on the TPS model, we would expect a high correlation between experimental data and analyses of chord progressions generated by the TPS model. The value of such a comparison is clear. If the TPS model posits a hypothesized model of perception, then we would like to know if and by how much it differs from experimental data it claims to approximate. In this paper, we focus on the intra-regional relation descriptions of TPS. We achieve two important goals. First, we develop a similarity measure that allows us to accurately compare the TPS model with a model of perceived chord relations created by Bharucha, et al. Second, this paper applies the similarity measure to normalized canonical representations of each model, thereby avoiding comparisons affected by arbitrary design choices. The greatest obstacle inherent in quantifying a comparative procedure is negotiating the manner in which each constituent model codes its data. The similarity measure and the method of normalization are applicable to any model with formal properties described herein and has the potential to focus experimental design and strengthen the relationship between experimental data and analytic systems.

I. INTRODUCTION

The search for compelling representations of tonal hierarchy and its constituent harmonic relations has a long-standing history. Fred Lerdahl describes geometric approaches to this problem [1] that involve the collection and modelling of data from experiments in music cognition. The multidimensional-scaling models of Bharucha and Krumhansl [2], and Deutsch [3], for example, seek to encode cognitive relationships between chords *within a single key area* (region) as Euclidean distances. The work of Heinichen [4], Kellner [5], and Weber [6], is more speculative and, when formalized, develops geometric representations of relationships between different key areas (regions) through their placement within a multidimensional space.

The first class of representations informs us about perceived tonal relationships within a region, while the second presents geometric representations of inter-regional relationships. Both representations can be integrated into an analytical program, facilitating analysis and prescriptive understanding.

Fred Lerdahl's tonal pitch space (TPS) model [1] approximates the cognitive perceptual relation between chords by providing a combinatorial procedure for computing the distance value between two arbitrary chords. The procedure employed by the TPS model is informed by experimental data and plausible hypotheses about how we perceive tonal relations. The TPS model is able to describe relations both between chords within a region (e.g. Bharucha, et al. [2]) and between regions themselves (e.g. Heinichen, et al.); TPS thus bridges these two representational classes.

Because of the influence of experimental data on the TPS model, we would expect a high correlation between experimental data and analyses of intra-regional chord progressions generated by the TPS model. The value of such a comparison is clear. If the TPS model posits a hypothesized model of perception, then we would like to know if and by how much it differs from the experimental data it claims to approximate. In this paper, we shall focus on the intra-regional relation descriptions of TPS and their counterparts in Bharucha, et al.

The greatest obstacle inherent in quantifying a comparative procedure of tonal models is negotiating the manner in which each constituent model codes its data. This paper describes a similarity measure between canonical representations of each model, thereby avoiding comparisons affected by arbitrary design choices. This measure is applicable to any set model with formal properties described herein.

II. INTRA-REGIONAL MODELS

Given a key area (region) R, a *graphical* intra-regional model $G_R = (V, E, d)$ is a graph whose vertex set V consists of every chord in R. Two chords $c, c' \in V$ are said to be connected by an edge if $(c, c') \in E$, and in this event, the weight of this edge is determined by the positive-valued function $d : E \to \mathbb{R}^{>0}$.

Given a graphical intra-regional model $G_R = (V, E, d)$, there is a natural extension of d to a closely related function $d^* : V \times V \to \mathbb{R}^{\geqslant 0}$, defined as follows. Given two vertices c, c' in V, let $P_{c,c'}$ be the set of all (non self-intersecting) paths connecting c to c' in G_R. Each path p in $P_{c,c'}$ can be viewed as a sequence of edges $p = (e_1, e_2, \ldots, e_{|p|})$, where $e_i \in E$ $(i = 1, \ldots, |p|)$. We define $d(p) = \Sigma_{i=1}^{|p|} d(e_i)$ (and adopt the convention that $d(p) = 0$ for paths of length 0). In this manner, we extend the definition of d to all of $P_{c,c'}$. Now we define

$$d^*(c, c') \stackrel{def}{=} \min_{p \in P_{c,c'}} d(p).$$

Intuitively, $d^*(c, c')$ is the length of the shortest path from c to c' in G_R. Because we require a graphical intra-regional model to be a *connected* graph, $d^* : V \times V \to \mathbb{R}^{\geqslant 0}$ assigns a finite non-negative value to every pair of vertices in V.

It is easy to see that the above definition of d^* permits us to convert a graphical intra-regional model $G_R = (V, E, d)$ into a metric space $M_R = (V, d^*)$. We refer to this as the associated *metric* intra-regional model of R.

By the construction above, every graphical model of R gives rise to a *unique* metric model. The correspondence is not bijective, however, since several graphical models may give rise to the same metric model. The simplest example of this is to consider a triangle with edges weighted 2, 1, and 1, and a chain of three vertices with two edges weighted 1 and 1. Both graphs, though different, give rise to the same metric model.

In light of the previous remark, note that the set of all metric models of a region R is no larger than the set of all graphical models. Initially, then, we will restrict ourselves to metric models over R. Accordingly, let $\mathcal{M}_R$ be the *set consisting of all metric models derived from graphical models of R.*

The following notation is used as it will be useful in later arguments. Given a region R and a model $M_R = (V, d^*)$ from $\mathcal{M}_R$ we denote the *minimal* (resp. *maximal*) *separation* as $\sigma_R^{\min}(M_R)$ (resp. $\sigma_R^{\max}(M_R)$) and define these quantities by

$$\sigma_R^{\min}(M_R) \stackrel{def}{=} \min_{\substack{v_1, v_2 \in V \\ v_1 \neq v_2}} d^*(v_1, v_2),$$

$$\sigma_R^{\max}(M_R) \stackrel{def}{=} \max_{\substack{v_1, v_2 \in V \\ v_1 \neq v_2}} d^*(v_1, v_2).$$

Formal Intra-Regional Models

Lerdahl Tonal Pitch Space

Using a large body of empirical evidence, Lerdahl created an algebraic model for quantifying the distance between any two chords. He dwells mainly on triads, but considers other sonorities as well. Lerdahl's TPS model assigns a number to each chord pair in a single key or region. That number is the intra-regional chord-pair "distance" as defined by the chord distance rule.

> CHORD DISTANCE RULE: $\delta(x \to y) = j + k$, where $\delta(x \to y)$ = the distance between chord x and chord y; j = the number of applications of the chordal circle-of-fifths rule needed to shift x into y; and k = the number of distinctive pcs in the basic space of y compared to those in the basic space of x.[1]

First, a region is determined by affixing a major scale to the universal chromatic space. Second, triadic structures are overlaid on the diatonic space in a weighted fashion, reflecting the perceptual hierarchy of root, fifth, and third. This is the "basic space" of a triad. Finally, triad structures are shifted to different positions on the diatonic space. The distance (δ) between two triads X and Y is number of diatonic fifths moved plus the number of pcs (p) in the basic space of a chord X that are unique to X plus those that intersect with Y: if a $p \in X$ and Y (as in the case of 7), then only the "highest" occurrence is counted (shown by the underscore). If a $p \in X$ and $\notin Y$, then every occurrence of p is counted.

For a single region, the pairwise chord distances are given in Table I. Figure 1 shows an excerpt from the "Madamina" aria from Mozart's *Don Giovanni.* Figure 2 shows the progression comprising the initial eight-bar parallel period annotated with the intra-regional chord-pair distance as defined by the chord distance rule. Added up we create an index representing the total regional distance covered.

FIG. 1: Measures 85-92 from "Madamina" Aria.

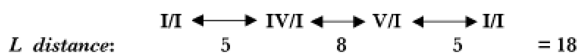

FIG. 2: Madamina Progression.

Experimental Intra-Regional Models

Bharucha-Krumhansl Space

Table II shows the results the 1983 experiment by J. Bharucha and C. Krumhansl in which all possible pairs of diatonic triads from a single major mode key were judged in terms of how well they sounded in succession [2]. The higher the number between two chord pairs, the more strongly they are associated. Bharucha and Krumhansl's experiments considered ordered sequences of chords.

The symmetrical regularity of Lerdahl's TPS model contrasts with the irregularity the findings of Bharucha and Krumhansl. Whereas the TPS model produces symmetric relations between chord pairs, Bharucha and Krumhansl found a significant ordering effect. In Lerdahl's basic intra-regional model, interval-cycle 7 plays a central organizing role. A problem is created when perceptually important relations are generalized and used to model perceptually less-important relations. For example, is the relationship between vii to IV the same as ii to V? There are some significant differences between the Bharucha and Krumhansl (BK) space and the Lerdahl (L) space. For example, in any order, the progression I to ii perceptually stronger than the progression I to iii. We must keep in mind that both models are products of different lines questioning. The relations they describe are the result of different methodologies designed for different reasons and therefore they show different kinds of information. Nevertheless, they describe relations between the same set of objects. Furthermore, Lerdahl claims that there is a correlation between the relations his model describes and the relations that have been (and could be) described by experimental results.

TABLE I: Theoretical Harmonic Relations from Lerdahl (2001).

First Chord	Second Chord						
	I	II	III	IV	V	VI	VII
I	0						
II	8	0					
III	7	8	0				
IV	5	7	8	0			
V	5	5	7	8	0		
VI	7	5	5	7	8	0	
VII	8	7	5	5	7	8	0

TABLE II: Perceived Harmonic Relations from Bharucha-Krumhansl (1983).

First Chord	Second Chord						
	I	II	III	IV	V	VI	VII
I	0	5.1	4.78	5.91	5.94	5.26	4.57
II	5.69	0	4.0	4.76	6.1	4.97	5.14
III	5.38	4.47	0	4.63	5.03	4.6	4.47
IV	5.94	5.0	4.22	0	6.0	4.35	4.79
V	6.19	4.79	4.47	5.51	0	5.19	4.85
VI	5.04	5.44	4.72	5.07	5.56	0	4.5
VII	5.85	4.16	4.16	4.53	5.16	4.19	0

In order to formally compare BK to L we approximate unordered harmonic relationships by considering the symmetrized magnitudes of the relationships they reported. To wit, the (i, j) entry in Table III is obtained by averaging the (i, j) and (j, i) entries of Table II. The values are proportionally inverted and symmetrized (averaged) so the cognitively closest pair is represented by the smallest distance.

III. DISTANCE MEASURES BETWEEN MODELS

This section presents a methodology for measuring the similarity between two intra-regional models of a given region R.

Fix a region R consisting of a set of chords V, where $|V| \geqslant 2$. Recall that $\mathcal{M}_R$ is the set of *all metric models derived from graphical models of* R. Let $M_R^1 = (V, d_1^*)$

TABLE III: Symmetrized Harmonic Relations derived from Bharucha-Krumhansl.

First Chord	Second Chord						
	I	II	III	IV	V	VI	VII
I	0						
II	0.185	0					
III	0.197	0.236	0				
IV	0.165	0.205	0.226	0			
V	0.165	0.184	0.211	0.174	0		
VI	0.194	0.192	0.215	0.212	0.186	0	
VII	0.192	0.215	0.232	0.215	0.200	0.230	0

and $M_R^2 = (V, d_2^*)$ be any two metric models in $\mathcal{M}_R$. Define:

$$\mu_R(M_R^1, M_R^2) \overset{def}{=} \max_{\substack{v_1, v_2 \in V \\ v_1 \neq v_2}} \left\{ \frac{d_2^*(v_1, v_2)}{d_1^*(v_1, v_2)} \right\},$$

$$\epsilon_R(M_R^1, M_R^2) \overset{def}{=} |\log[\mu_R(M_R^1, M_R^2)]|.$$

Clearly, $\epsilon_R : \mathcal{M}_R \times \mathcal{M}_R \rightarrow \mathbb{R}^{\geqslant 0}$. Intuitively, $\epsilon_R(M_R^1, M_R^2)$ measures the maximum distortion of pairwise distances that would be experienced by switching from M_R^1 to M_R^2. Representing the distance between two models a single value might appear inadequate. It is a narrow view of the relation between two models. However, a single distance measure between pairs of models allows us later to define a special kind of relation between sets of metric spaces.

Consider two degenerate examples: (i) Suppose M_R^2 simply reduces every pairwise distance in M_R^1 by a factor of k. Then $\mu_R(M_R^1, M_R^2) = 1/k$, so $\epsilon_R(M_R^1, M_R^2) = |\log(1/k)| = |\log 1 - \log k| = \log k$. (ii) Suppose M_R^2 simply expands every pairwise distance in M_R^1 by a factor of k. Then $\mu_R(M_R^1, M_R^2) = k$, so $\epsilon_R(M_R^1, M_R^2) = |\log(k)| = \log k$. The examples show that by incorporating absolute values into the definition, the measure ϵ_R is insensitive to whether distortion is expansive or contractive. Furthermore, because ϵ_R grows logarithmically with the extent of the distortion, it exhibits greater measurement sensitivity in situations where the distortion is low. The following section will describe equivalence classes based on the kinds of scaling causes distortion between models. In particular, we define the cases when distortion between two models results from specific functions on the distances of those models.

Having defined ϵ_R we now define the distance between models M_R^1 and M_R^2 to be

$$\delta_R(M_R^1, M_R^2) \overset{def}{=} \epsilon_R(M_R^1, M_R^2) + \epsilon_R(M_R^2, M_R^1).$$

Intuitively, $\epsilon_R(M_R^1, M_R^2)$ interprets distance between two models as the sum of two values: (i) the log of the maximum distortion of pairwise distances that would be experienced by switching from M_R^1 to M_R^2 and (ii) the log of the maximum distortion of pairwise distances that would be experienced by switching from M_R^2 to M_R^1. Clearly, $\delta_R : \mathcal{M}_R \times \mathcal{M}_R \rightarrow \mathbb{R}^{\geqslant 0}$. The next result provides a compelling argument for for our choice of δ_R as a similarity measure between intra-regional models of a given region R.

Theorem III.1. *$(\mathcal{M}_R, \delta_R)$ is a metric space.*

The theorem follows immediately from Propositions III.2, III.3, and III.4, which verify that each of the defining properties for a metric space hold in $(\mathcal{M}_R, \delta_R)$.

Proposition III.2. *(*Reflexivity*) Let $M_R^1 = (V, d_1^*)$ and $M_R^2 = (V, d_2^*)$ be any two metric models in $\mathcal{M}_R$. Then*

$$\delta(M_R^1, M_R^2) = 0 \;\Leftrightarrow\; M_R^1 = M_R^2.$$

Proof. $(\Leftarrow)$ If $M_R^1 = M_R^2$ then $d_1^* \equiv d_2^*$ as functions. Hence for all $v_1, v_2 \in V$, $d_1^*(v_1, v_2) = d_2^*(v_1, v_2)$. Hence $\epsilon_R(M_R^1, M_R^2) = |\log 1| = 0$. By a symmetric argument, $\epsilon_R(M_R^2, M_R^1) = 0$. Hence $\delta(M_R^1, M_R^2) = 0$.

$(\Rightarrow)$ Suppose $\delta(M_R^1, M_R^2) = 0$. Since $\delta_R(M_R^1, M_R^2)$ is the sum of two non-negative quantities, it follows that $|\log[\mu_R(M_R^2, M_R^1)]| = |\log[\mu_R(M_R^1, M_R^2)]| = 0$. Hence $\mu_R(M_R^2, M_R^1) = \mu_R(M_R^1, M_R^2) = 1$. It follows that for each $v_1, v_2 \in V$, $d_1^*(v_1, v_2) = d_2^*(v_1, v_2)$, and so $M_R^1 = M_R^2$. □

Proposition III.3. *(*Symmetry*) Let $M_R^1 = (V, d_1^*)$ and $M_R^2 = (V, d_2^*)$ be any two metric models in $\mathcal{M}_R$. Then*

$$\delta(M_R^1, M_R^2) \;=\; \delta(M_R^2, M_R^1).$$

Proof. Immediate, since

$$\begin{aligned} \delta_R(M_R^1, M_R^2) &= \epsilon_R(M_R^1, M_R^2) + \epsilon_R(M_R^2, M_R^1) \\ &= \epsilon_R(M_R^2, M_R^1) + \epsilon_R(M_R^1, M_R^2) \\ &= \delta_R(M_R^2, M_R^1) \end{aligned}$$

i.e. δ_R is symmetric in its arguments. □

Proposition III.4. *(*Triangle inequality*) Let $M_R^1 = (V, d_1^*)$, $M_R^2 = (V, d_2^*)$, and $M_R^3 = (V, d_3^*)$ be any three metric models in $\mathcal{M}_R$. Then*

$$\delta_R(M_R^1, M_R^3) \;\leqslant\; \delta_R(M_R^1, M_R^2) + \delta_R(M_R^2, M_R^3).$$

Proof. Let φ be vertices v_1 and v_2. Then

$$\frac{d_1^*(\varphi)}{d_3^*(\varphi)} = \left[\frac{d_1^*(\varphi)}{d_2^*(\varphi)}\right] \cdot \left[\frac{d_2^*(\varphi)}{d_3^*(\varphi)}\right]$$

$$\leqslant \mu_R(M_R^1, M_R^2) \cdot \mu_R(M_R^2, M_R^3).$$

Since $\mu_R(M_R^1, M_R^3)$ is the maximum value of $\frac{d_1^*(\varphi)}{d_3^*(\varphi)}$ (maximized over all distinct φ in V), we see that

$$\mu_R(M_R^1, M_R^3) \leqslant \mu_R(M_R^1, M_R^2) \cdot \mu_R(M_R^2, M_R^3).$$

Taking logarithms and appealing to convexity of absolute values, it follows that

$$\begin{aligned} & |\log\left[\mu_R(M_R^1, M_R^3)\right]| \\ \leqslant\ & |\log\left[\mu_R(M_R^1, M_R^2)\right] + \log\left[\mu_R(M_R^2, M_R^3)\right]| \\ \leqslant\ & |\log\left[\mu_R(M_R^1, M_R^2)\right]| + |\log\left[\mu_R(M_R^2, M_R^3)\right]|. \end{aligned}$$

It follows, by the definition of ϵ_R, that

$$\epsilon_R(M_R^1, M_R^3) \leqslant \epsilon_R(M_R^1, M_R^2) + \epsilon_R(M_R^2, M_R^3). \quad (1)$$

A symmetric argument yields that

$$\epsilon_R(M_R^3, M_R^1) \leqslant \epsilon_R(M_R^3, M_R^2) + \epsilon_R(M_R^2, M_R^1). \quad (2)$$

By combining corresponding sides in expressions (1) and (2), we conclude that $\delta_R(M_R^1, M_R^3) \leqslant \delta_R(M_R^1, M_R^2) + \delta_R(M_R^2, M_R^3)$, as claimed. □

IV. NORMALIZATIONS

Having defined a distance measure between models, we would like to apply it to study the relationship between specific models, such as Lerdahl's TPS model and the symmetrized Bharucha-Krumhansl model. As it stands, we can compare the two models as they are presented in Table I and Table III, by using the distance metric δ_R. The question is, how meaningful is this as a similarity measure?

Hypothesis IV.1 (Fundamental Hypothesis). *Each of the spaces (L and BK) is in fact a concrete realization of an abstract system, and the particular concrete realization incorporates arbitrary choices in the representational design.*

Why, for example, should our comparison be constrained by $BK's$ choice of a of scale of integers 1-7, when it may as well have been -10 to 70 or real numbers between 0 and 1. Lerdahl chose to make the root of a chord a certain weight, where he could have made it heavier or lighter. When we compare these two spaces, we must make every effort to ensure that we are comparing their most basic essence.

What follows is a precise quantitative interpretation of hypothesis IV.1. First is given a formal interpretation to the notion that models "incorporate arbitrary choices in their representational design". Given a model M_R and two real numbers α, β, where $\alpha \in (0, +\infty)$ and $\beta \in (-\infty, \sigma_R^{\min}(M_R))$, we denote the (α, β)-normalization of M_R as

$$\langle M_R \rangle_{\alpha,\beta} \stackrel{def}{=} (V, \langle d^* \rangle_{\alpha,\beta}),$$

where $\langle d^* \rangle_{\alpha,\beta} : V \times V \to \mathbb{R}^{\geqslant 0}$ is defined to be

$$\langle d^* \rangle_{\alpha,\beta}(v_1, v_2) \stackrel{def}{=} \begin{cases} \alpha[d^*(v_1, v_2) - \beta] & \text{if } v_1 \neq v_2 \\ 0 & \text{otherwise,} \end{cases}$$

for every v_1, v_2 in V. Intuitively, (α, β)-normalization of M_R represents a linear translation of all positive distances by $-\beta$ followed by a rescaling by a factor of α. Normalization parameters must be considered when assessing the similarity of two models, since the implicit $\alpha = 1$, $\beta = 0$ choices comes from arbitrary choices in the formal underpinnings or experimental design. These arbitrary choices are unimportant when models are considered *in isolation*, but when we want to measure similarity *between* models, the choices exert undue influence.

In particular, given two models $M_R^1 = (V, d_1^*)$ and $M_R^2 = (V, d_2^*)$, and specific

$$\begin{aligned} \alpha_1, \alpha_2 &\in (0, +\infty) \\ \beta_1 &\in (-\infty, \sigma_R^{\min}(M_R^1)) \\ \beta_2 &\in (-\infty, \sigma_R^{\min}(M_R^2)), \end{aligned}$$

it is difficult to make any general assertions (i.e. independent of our choices of α_1, α_2, β_1, β_2) regarding the relationship between the distance between the two original models and the distance between their normalizations. In other words the following question does not have a uniform answer independent of our choices of α_1, α_2, β_1, β_2:

> Is $\delta_R(M_R^1, M_R^2)$ less than, or equal to, or greater than
> $\delta_R(\langle M_R^1 \rangle_{\alpha_1,\beta_1}, \langle M_R^2 \rangle_{\alpha_2,\beta_2})$?

The previous discussions leads us to seek *normalization invariant* distance measures between models. Two models M_R^1 and M_R^2 said to be *normalization-equivalent* if one model is merely a linear normalization of the other, i.e.

$$M_R^1 \simeq M_R^2 \Leftrightarrow \langle M_R^1 \rangle_{\alpha_1,\beta_1} = M_R^2,$$

for some $\alpha_1 \in (0, +\infty), \beta_1 \in (-\infty, \sigma_R^{\min}(M_R^1))$. It is easy to see that $\simeq$ defines an equivalence relation on $\mathcal{M}_R$. The equivalence class of a model M_R under the equivalence relation is

$$[M_R] = \{\langle M_R \rangle_{\alpha,\beta} \mid \alpha \in (0, +\infty), \beta \in (\sigma_R(M_R), +\infty)\}.$$

Viewed as a subset of $\mathcal{M}_R$, $[M_R]$ is the set of all linear normalizations of the model M_R.

Lerdahl's model L is in fact only one arbitrary member of an infinite collection $[L]$ of related models, each of which corresponds to a different normalization of L. The position of L within that set $[L]$ reflects specific arbitrary choices in the representational design of L. Indeed, both L and BK are laden with arbitrary choices that influence the value of $\delta_R(L, BK)$– it is from this arbitrariness that we must divest ourselves. We must find a way to measure the distance between models in a way that is insensitive to the arbitrary design choices.

Returning to the formal scaffolding, let us define the set of all equivalence classes in $\mathcal{M}_R$ as

$$[\mathcal{M}_R] = \{[M_R] \mid M_R \in \mathcal{M}_R\}.$$

The goal is to define a *normalization-invariant similarity measure*, which is to say we would like to describe a distance on equivalence classes; i.e. a function $\lambda : [\mathcal{M}_R] \times [\mathcal{M}_R] \to \mathbb{R}^{\geqslant 0}$, which will endow the set of equivalence classes with a metric structure to yield a metric space $([\mathcal{M}_R], \lambda)$. This will be done by the method of canonical representatives.

V. CANONICAL REPRESENTATIVES

Lemma V.1. *Fix a model M. Fix a value for β between 0 and $-\sigma_{\min}(M)$. Now vary α. As α increases from $0 \to +\infty$, the diameter of $\langle M\rangle_{\alpha,\beta}$ increases monotonically from $0 \to +\infty$.*

Proof. Diameter$(\langle M\rangle_{\alpha,\beta}) = \alpha *$ Diameter$(\langle M\rangle_{1,\beta})$. □

Lemma V.2. *Fix a model M. For any fixed choice of β, there exists a unique value of α –called $\alpha(\beta)$– with the property that the diameter of $M_{\alpha(\beta),\beta}$ is 1.*

Proof. Existence follows from Lemma V.1, and uniqueness from the mean value theorem and the property of monotonicity. □

Definition V.3. *For any model M. Define $I(M) = \sigma_{\max}(M)/\sigma_{\min}(M)$. i.e. $I(M)$ is the dynamic range of distances, measured as the farthest pairwise separation over the smallest pairwise separation.*

Lemma V.4. *Fix a model M. As β is varied from $+\sigma_{\min}(M)$ towards $-\infty$, the value of $I(M_{\alpha(\beta),\beta})$ decreases monotonically from $+\infty$ to 0.*

Proof. Fix β_1. Note that $\sigma_{\max}M_{1,\beta_1} = \sigma_{\max}M - \beta_1$ and $\sigma_{\min}M_{1,\beta_1} = \sigma_{\min}M - \beta_1$. Since $\alpha(\beta_1) \cdot \sigma_{\max}M_{1,\beta_1} = 1$, we have that

$$\alpha(\beta_1) = \frac{1}{\sigma_{\max}M - \beta_1}.$$

It follows that

$$\sigma_{\min}M_{\alpha(\beta_1),\beta_1} = \frac{\sigma_{\min}M - \beta_1}{\sigma_{\max}M - \beta_1}, \tag{3}$$

and $\sigma_{\max}M_{\alpha(\beta_1),\beta_1} = 1$. By definition of I,

$$I(M_{\alpha(\beta_1),\beta_1}) = \frac{\sigma_{\min}M - \beta_1}{\sigma_{max}M - \beta_1}. \tag{4}$$

Now, if $\beta_2 \leqslant \beta_1$, then for a fixed M

$$\frac{\sigma_{\min}M - \beta_2}{\sigma_{max}M - \beta_2} \leqslant \frac{\sigma_{\min}M - \beta_1}{\sigma_{max}M - \beta_1}, \tag{5}$$

and hence

$$I(M_{\alpha(\beta_2),\beta_2}) \leqslant I(M_{\alpha(\beta_1),\beta_1}). \tag{6}$$

□

Proposition V.5. *Fix a model M, and a positive number x. There exists a unique value of β –called β_x– with the property that $I(M_{\alpha(\beta),\beta}) = x$.*

Proof. Existence follows from Lemma V.4 and uniqueness from the mean value theorem and properties of monotonicity. □

Given a graphical model M on a set of n vertices, we note that the number of binary bits needed to represent the distance between two specific vertices u and v is

$$2\log_2 n + \log_2 I(M). \tag{7}$$

The first term in (7) is explained by the fact that u and v must be specified uniquely, and since they are members of a set of n vertices, this can be achieved by assigning each of the n vertices a unique name from the set of $(\log n)$-bit binary numbers. The second term in (7) is understood by observing that if the edge (u, v) has weight w, this can be rescaled to $(w/\sigma_{\min}M)$, and the transmission of the rescaled quantity requires at most $\log_2 I(M)$ binary bits. To conclude, transmission of a model M (ignoring compression) requires on the order of

$$|E|[2\log_2 n + \log_2 I(M)] + c \tag{8}$$

where E is the set of edges in the model, and c is a constant number of bits required to encode the edge-weight rescaling factor $\sigma_{\min}M$. Note that if M' is a renormalization of M then the representation size of M' differs from the representation size of M to the extent that $I(M')$ differs from $I(M)$; since other aspects of expression (8) remain unchanged.

In comparing two models on n chords, we would like to ensure that both are subject to the same constraints in terms of descriptive size; it would be of dubious value to attempt to compare a model which requires a much larger representation size to one which can be represented using a much smaller number of bits. We quantify this as follows: Two models M_1 and M_2 on n chords are comparable in a meaningful way if $I(M_1) = I(M_2)$.

Lemma V.4 and Proposition V.5 demonstrate that given a model M we can always find a renormalization of it, M', which has diameter 1 an arbitrary value for $I(M')$. By extension, given two models M_1 and M_2, and an arbitrary positive real number x and we can find renormalizations M_1', M_2' such that $I(M_1') = I(M_2') = x$. The question remains as to what to take as a our suitable value of x? Here we return to examining expression (7), from which we see that each edge in a graphical model requires at least $\log_2 n$ bits *just to represent the identities of its endpoint vertices.* We propose the constraint that $I(M) = n$, since this implies that the representation size of each edge weight, being $\log_2 I(M) = \log_2 n$ bits, is asymptotically no more than the number of bits required to convey the identity of the edge itself, vis-a-vis the endpoint vertex names. This aesthetic choice leads us to the next corollary:

Corollary V.6. *Fix a model M on n chords. There exists a unique value of β–which we'll call $\beta_{\log n}$ – with the property that $I(M_{\alpha(\beta),\beta}) = n$*

Proof. Follows from Proposition V.5 taking $x = n$. □

Definition V.7 (Canonical Representative). *The canonical representative of a model M on n chords is defined to be*

$$\overline{M} \stackrel{def}{=} M_{\alpha(\beta_{\log n}),\beta_{\log n}}.$$

Thus, $\bar{M}$ is the unique normalization of M which has diameter 1 and for which the ratio between the longest and shortest pairwise distances is n.

Definition V.8. *A function on ordered pairs of equivalence classes $\lambda : [\mathcal{M}_R] \times [\mathcal{M}_R] \to \mathbb{R}^{\geqslant 0}$ is defined as follows:*

$$\lambda_R([M_R^1],[M_R^2]) \;=\; \delta_R(\overline{M}_R^1, \overline{M}_R^2)$$

.

Theorem V.9. $([\mathcal{M}_R], \lambda_R)$ *is a metric space.*

Proof. The distance between equivalence classes is defined in terms of δ_R distance between canonical representatives, and the representatives themselves inhabit a metric space $(\mathcal{M}_R, \delta_R)$. Since a subspace of a metric space is a metric space, the theorem follows. □

Canonical Representation of L and BK Spaces

To find the canonical representations of L and BK, we find M_R^* by multiplying each metric in M_R by $\frac{n-1}{\sigma_R^{\max}(M_R)-\sigma_R^{\min}(M_R)}$ where n equals the number of V in R. We then add $1-\sigma_R^{\min}(M_R^*)$ to each d in M_R^*. The result,$\overline{M}_R$, is the canonical representation of M_R.

Recall the definition of λ_R as a distance function such that $\lambda_R([M_R^1],[M_R^2]) = \delta_R(\overline{M}_R^1, \overline{M}_R^2)$ where $([M_R^1],[M_R^2])$ are equivalence classes and members of $[\mathcal{M}_R]$, the set of all equivalence classes. In the case of L and BK, $\lambda = 4.727$. Looking at Table IV we see that the largest points of divergence occurs at IV-V and iii-vii. In terms of BK, L greatly overrates the distance between iii and vii, and underrates the distance between IV and V. These discrepancies summarizes the difference between the two models. Lerdahl's algorithmic approach privileges all fifth related harmonies by giving them the lowest value (i.e. representing the "closest" cognitive relations). By doing so, he distorts some important relations, the most important being the IV-V progression. We see this clearly reflected in the interpretation the "Madamina" progression using the canonical forms of L and BK shown in Figure 3. The difference is clear. The overall distance covered by L(Madamina) = 9 and for BK(Madamina) = 6.212.

TABLE IV: Canonical Representatives for L ($\overline{M}_1$) and BK ($\overline{M}_2$)

	$\overline{M}_1$	$\overline{M}_2$	$\overline{M}_1 / \overline{M}_2$	$\overline{M}_2 / \overline{M}_1$
I-ii	7	2.690	2.602	0.384
I-iii	5	3.704	1.350	0.741
ii-iii	7	7.000	1.000	1.000
iii-IV	7	6.155	1.137	0.879
ii-IV	5	4.380	1.141	0.876
iii-V	5	4.887	1.023	0.977
iii-vi	1	5.225	0.191	5.225
iii-viio	1	6.662	0.150	**6.662**
I-IV	1	3.451	0.290	3.451
ii-V	1	2.606	0.384	2.606
ii-vi	1	3.282	0.305	3.282
ii-viio	5	5.225	0.957	1.045
I-V	1	1.000	1.000	1.000
I-vi	5	3.451	1.449	0.690
I-viio	7	3.282	2.133	0.469
IV-V	7	1.761	**3.976**	0.252
IV-vi	5	4.972	1.006	0.994
IV-viio	1	5.225	0.191	5.225
vi-vii	7	6.493	1.078	0.928
V-vi	7	2.775	2.523	0.396
V-vii	5	3.958	1.263	0.792

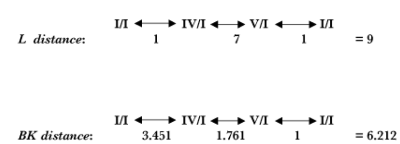

FIG. 3: Madamina Progression Under Canonical Forms of L and BK.

VI. CONCLUSION

This paper has achieved four important music-analytic goals. First, we showed how certain tonal models are also metric models, meeting the requirements of a metric space. Second, we defined a distance measure between pairs of tonal models. Third, we showed how to derive canonical representatives from tonal models. This allowed us to compare models without the interference of arbitrary design choices. Comparing canonical representatives allowed us to critique analytic claims made by L against BK. Finally, these two tonal models are members of two different equivalence classes, whose representatives come from a metric space of tonal models, and since every subset of a metric space is a metric space, the two models form a metric space.

[1] F. Lerdahl, *Tonal Pitch Space* (Oxford University Press, New York, 2001).

[2] J. Bharucha and C. Krumhansl, Cognition **13**, 63 (1983).

[3] D. Deutsch, in *The Psychology of Music*, edited by D. Duetsch (Academic Press, New York, 1999), chap. 10.

[4] J. D. Heinichen, *Der Generalbass in Der Composition* (Dresden, 1728).

[5] D. Kellner, *Treulicher Unterricht im General-Bass* (Christian Herold, Hamburg, 1737).

[6] G. Weber, *Versuch Einer Geordenten Theorie Der Tonsetzkunst* (B. Schotts Sohne, Mainz, 1821-1824).