PROBABILISTIC MODULAR BASS VOICE LEADING IN MELODIC HARMONISATION

Dimos Makris
Department of Informatics,
Ionian University,
Corfu, Greece
c12makr@ionio.gr

Maximos Kaliakatsos-Papakostas
School of Music Studies,
Aristotle University of
Thessaloniki, Greece
maxk@mus.auth.gr

Emilios Cambouropoulos
School of Music Studies,
Aristotle University of
Thessaloniki, Greece
emilios@mus.auth.gr

ABSTRACT

Probabilistic methodologies provide successful tools for automated music composition, such as melodic harmonisation, since they capture statistical rules of the music idioms they are trained with. Proposed methodologies focus either on specific aspects of harmony (e.g., generating abstract chord symbols) or incorporate the determination of many harmonic characteristics in a single probabilistic generative scheme. This paper addresses the problem of assigning voice leading focussing on the bass voice, i.e., the realisation of the actual bass pitches of an abstract chord sequence, under the scope of a modular melodic harmonisation system where different aspects of the generative process are arranged by different modules. The proposed technique defines the motion of the bass voice according to several statistical aspects: melody voice contour, previous bass line motion, bass-to-melody distances and statistics regarding inversions and note doublings in chords. The aforementioned aspects of voicing are modular, i.e., each criterion is defined by independent statistical learning tools. Experimental results on diverse music idioms indicate that the proposed methodology captures efficiently the voice layout characteristics of each idiom, whilst additional analyses on separate statistically trained modules reveal distinctive aspects of each idiom. The proposed system is designed to be flexible and adaptable (for instance, for the generation of novel blended melodic harmonisations).

1. INTRODUCTION

In melodic harmonisation systems harmony is expressed as a sequence of chords, but an important aspect is also the relative placement of the notes that comprise chord sequence, which is known as the *voice leading* problem. As in many aspects of harmony, in voice leading there are certain sets of diverse conventions for different music *idioms* that need to be taken under consideration. Such rules have been hand-coded by music experts for the development of rule-based melodic harmonisation systems (see [15] for a review of such methods). Similarly, such hand-coded rules have been utilised as fitness criteria for evolutionary systems (see [4, 18] among others). However, the specification of rules that are embedded within these systems are very complex with many variations and exceptions. Additionally, the formalisation of such rules has not yet been approached for musical idioms that have not hitherto been thoroughly studied. Most of the works so far, have focused on either finding a satisfactory chord sequence for a given melody (performed by the soprano voice), or on completing the remaining three voices that constitute the harmony for a given melodic or bass line (known as the “four-part harmony” task) [5, 14, 18, 24]. Experimental evaluation of methodologies that utilise statistical machine learning techniques demonstrated that an efficient way to harmonise a melody is to add the bass line first [22]. To this end, the motivation behind the work presented in the paper at hand is further enforced by the findings in the aforementioned paper.

This study, is based on the following underlying melodic harmonisation strategy: 1) analyse a give melody in terms of segmentation, scale/pitch hierarchy, harmonic/embellishment notes, harmonic rhythm (this can be achieved automatically or, at this stage, manually), 2) assign abstract chords to the given melody from learned first-order chord transition tables, 3) select concrete pitches from abstract chords for the bass-line based on learned melody-to-bass-line movement (discussed in this paper), 4) select concrete pitches for inner voices (steady or varied number of notes per chord). This scheme would seem to be adequate for a large body of non-monophonic music, but not all. For instance, even the mere concept of chords (with inversions) is rather controversial in European music before the mid-eighteenth century and in other traditional polyphonic musics; more so, the idea of melody with chords and functional bass line is untenable in such music.

However, as the aim of this project is not individual fully-fleshed harmonic models of different idioms, but rather a general-as-possible method to ‘extract’ basic components of harmonic content in various harmonic textures, it is possible to employ the above strategy in any non-monophonic texture. It is known that outer voices tend to stand out perceptually (e.g. in [6]); additionally, note simultaneities can be encoded in a more abstract manner (e.g., GCT representation). Employing a computational methodology based on such generic concepts, can enable the construction of a 'generic' melodic harmoniser that can use harmonic components from various idioms, without claiming to emulate the idioms themselves.

This paper proposes a modular methodology for determining the bass voice leading, to be integrated in a melodic harmonisation system under development. The effectiveness of the proposed methodology that performs bass voice leading according to statistics describing the overall voicing layout (i.e. arrangement of pitches) of given chord sequences in the General Chord Type (GCT) [2] representation is examined. This methodology is extending the bass voice leading scheme presented in [12], by harnessing voicing layout information through additional voicing layout statistical, independently trained, *modules* concerning the chords that constitute the harmonisation. Those characteristics include distributions on the distance between the bass and the melody voice and statistics regarding the inversions and doublings of the chords in the given chord sequence. By training these modules on multiple diverse idioms, a deeper study is pursued within the context of the COINVENT project [20], which examines the development of a computationally feasible model for conceptual blending. Thereby, blending different modules from different idioms will expectedly lead to harmonisations with blended characteristics.

2. PROBABILISTIC MODULAR BASS VOICE LEADING

Given the fact that a melody is available in systems that perform melodic harmonisation, the methodology presented in [12] derives information from the melody voice in order to calculate the most probable movement for the bass voice, named as the *bass voice leading* (BVL). This approach, in combination with information regarding the *voice layout* (Section 2.2), is incorporated into a *larger modular probabilistic framework*. In the integrated modular melodic harmonisation system under development, the selection of chords (in GCT form [2]) is performed by another probabilistic module [10] not discussed in this paper. Therefore, the herein discussed modules have been developed to provide indications about possible movement of the bass as well as to define specific notes for the bass voice, providing a first step to complete information regarding specific voices from the chords provided by the chord selection module.

To this end, both the bass and the melody voice steps are represented by abstract notions that describe general quantitative information on pitch direction. In [12] several scenarios for voice contour refinement were examined, providing different levels of accuracy for describing the bass motion in different datasets. In the paper at hand, the selected methodology is the one with the greatest level of detail, i.e. the scenario where the melody and bass note changes are divided in seven steps, as exhibited in Table 1. While different range schemes could have been selected, the rationale behind the utilised one is that the perfect fourth is considered as a small leap and the perfect fifth as a big leap.

refinement	short name	range (semitones)
steady voice	`st_v`	$x = 0$
step up	`s_up`	$1 \leqslant x \leqslant 2$
step down	`s_down`	$-2 \leqslant x \leqslant -1$
small leap up	`sl_up`	$3 \leqslant x \leqslant 5$
small leap down	`sl_down`	$-5 \leqslant x \leqslant -3$
big leap up	`bl_up`	$5 < x$
big leap down	`bl_down`	$x < -5$

Table 1. The pitch step and direction refinement scale considered for the development of the utilised bass voice leading system.

2.1 The hidden Markov model module

The primary module for defining bass motion functions under the first order Markov assumption in combination with the fact that it depends on the piece's melody. To this end, the next step of the bass voice contour (bass direction descriptor) is dependent on the previous one and on the current melody contour (melody direction descriptor). This assumption, based on the fact that a probabilistic framework is required for the harmonisation system, motivates the utilisation of the *hidden Markov model* (HMM) methodology. According to the HMM methodology, a sequence of observed elements (melody direction descriptor) is given and a sequence of (hidden) states (bass direction descriptor) is produced as output. The "order" of the HMM utilised in the presented work, i.e. how many previous steps are considered to define the current, is 1. In melodic harmonisation literature different orders have been examined, e.g. [19], where it is shown that order 1 might not be the most efficient. In the context of the presented work, this investigation is part of future research.

The HMM training process extracts four probability values for each bass motion: 1) to begin the sequence, 2) to end the sequence, 3) to follow another bass motion (transition probability) and 4) to be present given a melody step (observation probability). The probabilities extracted by this process for each possible next bass motion is denoted with a vector of probabilities $\vec{p_m}$ (one probability for each possible bass motion step) and will be utilised in the product of probabilities from all modules in Equation 1.

2.2 The voicing layout information module

In order to assign a bass voice to a chord, additional information is required that is relevant to the chords of the harmonisation. The voicing layout statistics that are considered for the modules of the presented methodology are the *inversions* and the *doublings* of chords. The inversions of a chord play an important role in determining how eligible is a chord's pitch class to be a bass note, while the doublings indicate if additional "room" between the bass and the melody is required to fit doublings of specific pitch classes of the chords. For instance, the chord with pitch classes $[0, 4, 7]$ has three inversions, with each one having a bass note that corresponds to a different pitch class, e.g. $[60, 64, 67]$, $[64, 67, 72]$ or $[67, 72, 76]$, while, by considering the inversion prototype $[60, 64, 67]$ of the $[0, 4, 7]$ chord, there are four scenarios of single note doublings: $[60, 64, 67, 72]$, $[60, 64, 67, 76]$, $[60, 64, 67, 79]$ and $[60, 64, 67]$ (no-doubling scenario).

The voicing layout module of the harmonic learning system regarding chord inversions and note doublings, is trained through extracting relevant information from every (GCT) chord in pieces from a music idiom. Specifically, consider a GCT chord in the form $g = [r,\ \vec{t}]$, where r is the root of the chord in relation to the root of the key and $\vec{t}$ is the vector describing the type of the chord. For instance, the I chord in any key is expressed as $g = [0,\ [0, 4, 7]]$ in the GCT representation, where 4 denotes the major third and 7 the perfect fifth. This GCT type is a set of integers, $\vec{t} = [t_1,\ t_2,\ \ldots\ ,\ t_n]$, where n is the number of type elements, that can be directly mapped to relative pitch classes (PCs). The statistics concerning chord inversion are expressed as the probability that each type element in g is the bass note of the chord, or

$$p_{\mathrm{i}} = (v_1,\ v_2,\ \ldots\ ,\ v_n),$$

where v_i, $i \in \{1, 2, \ldots, n\}$, is the probability that the element t_i is the bass note. Similarly, probabilities about note doublings are expressed through a probability vector

$$p_{\mathrm{d}} = (d_1,\ d_2,\ \ldots\ ,\ d_n,\ s),$$

where d_i, $i \in \{1, 2, \ldots, n\}$, is the probability that the pitch class t_i gets doubled, while there is an additional value, s, that describes the probability that there is no doubling of pitch classes. Table 2 exhibits the extracted statistics for inversions and note doublings for the most often met chords of the major Bach Chorales.

2.3 The melody-to-bass distance module

An important aspect of voice layout has to do with absolute range of chords in the chord sequences of an idiom, i.e. the absolute difference between the bass voice and the melody. Different idioms encompass different constraints and characteristics concerning this voicing layout aspect, according to several factors, e.g., the utilised instruments' range. The proposed methodology addresses this voicing layout aspect by capturing statistics about the *region* that the bass voice is allowed to move according to the melody. Therefore, histograms are extracted that describe the frequency of all melody-to-bass intervals found in a training dataset, as illustrated by the bars in the example in Figure 1.

However, interval-related information in the discussed context are used only as approximate indicators about the expected pitch height of the bass voice, while the exact intervals (bars in Figure 1) are referring to specific intervals and, additionally, they are scale-sensitive, e.g. different scales potentially produce different distributions of melody-to-bass intervals. Therefore, the "expected" bass pitch height is approximated by a normal distribution that is adjusted to fit the distribution of the melody-to-bass intervals observed in the dataset. Figure 1 illustrates the normal distribution that is approximates the distributions of intervals for a collection of major Bach Chorales.

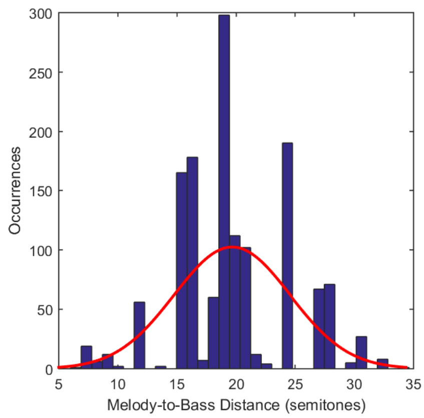


Figure 1. Histogram of pitch interval distances between melody and bass for a set of major Bach Chorales.

2.4 Combining all modules

The probabilities gathered from all the modules described hitherto are combined into a single value, computed as the product of all the probabilities from all the incorporated modules. To this end, for each GCT chord (C) in the composition every possible scenario of chord inversions, doublings and bass note pitch height, denoted by an index x, is generated. For each scenario (x), the product ($b_x(C)$) of all the modules discussed so far is computed, i.e. the bass motion ($p_{m_x}(C)$), the inversions ($p_{i_x}(C)$), doublings ($p_{d_x}(C)$) and melody-to-bass interval $p_{h_x}(C)$:

$$b_x(C) = p_{m_x}(C)\ p_{i_x}(C)\ p_{d_x}(C)\ p_{h_x}(C). \tag{1}$$

Therefore, the best scenario (x_{best}) for the bass voice of chord C is found by: $x_{\text{best}} = \arg\max_x(b_x(C))$. The bass note motion probability is obtained by the HMM module analysed in Section 2.1 and it takes a value given by the vector $\vec{p_m}$ according to the bass step it leads to.

3. EXPERIMENTAL RESULTS

The aim of the experimental process is to evaluate whether the proposed methodology efficiently captures the bass voice leading according to several factors related to the voice layout characteristics of each training idiom. Additionally, it is examined whether the separate trained modules, which constitute the overall system, statistically reveal aspects of each idiom that are more distinctive. A collection of eight datasets has been utilised for training and testing the capabilities of the proposed methodology, exhibited in Table 3.

These pieces are included in a music database with many diverse music idioms and it is developed for the purposes

GCT chord	relative PC	inversions	doublings
$[0, [0, 4, 7]]$	$[0, 4, 7]$	$[0.74,\ 0.23,\ 0.02]$	$[0.68,\ 0.15,\ 0.08,\ 0.09]$
$[7, [0, 4, 7]]$	$[7, 11, 2]$	$[0.78,\ 0.22,\ 0.00]$	$[0.83,\ 0.02,\ 0.09,\ 0.06]$
$[5, [0, 4, 7]]$	$[5, 9, 0]$	$[0.65,\ 0.34,\ 0.01]$	$[0.46,\ 0.30,\ 0.11,\ 0.13]$

Table 2. Probabilities for chord inversion (p_i) and note doublings (p_d) in the three most frequently used chords in the major Chorales of Bach.

Name (number)	Description
Bach Chorales (35)	a set of Bach chorales
Beatles (10)	set of songs from the band Beatles
Epirus (29)	traditional polyphonic songs from
Medieval (12)	fauxbourdon and organum pieces
Modal chorales (34)	15th-16th century modal chorales
Rembetika (22)	folk Greek songs
Stravinsky (10)	pieces composed by Igor Stravinsky
Tango (24)	pieces of folk tango songs

Table 3. Dataset description.

of the COINVENT project. For the presented experimental results, each idiom set includes from around 50 to 150 phrases. The Bach Chorales have been extensively utilised in automatic probabilistic melodic harmonisation [1, 7, 13, 16], while the polyphonic songs of Epirus [9, 11] and Rembetika [17] constitute datasets that have hardly been used in studies.

3.1 Cross-entropies for training and testing in all idiom combinations

The cross-entropy tests include the statistical modules that are independent of the GCT chords, i.e. HMM model and the melody-to-bass distance fitted distribution (will hereby be symbolised as mbd). Additionally, to examine the effect of the transition and the observation probabilities, the probabilities related to transitions of the bass (states transitions and will hereby be symbolised as tr) and the melody voice (observation transitions and will hereby be symbolised as mel) will be examined separately. The statistical combinations examined during the experimental evaluation process are: 1) the HMM model and the melody-to-bass distance fitted distribution probabilities (M^{all}), 2) only the bass voice transition probabilities from the HMM (M^{tr}), 3) only the melody observation probabilities from the HMM (M^{mel}) and 4) only the Melody-to-bass distance distributions (M^{mbd}).

Each idiom's dataset is divided in two subsets, a training and a testing subset, with a proportion of 90% to 10% of the entire idiom's pieces. The training subset of an idiom X is utilised to train the aforementioned modules, forming the trained model M_X, while the testing subset of the same idiom will be hereby denoted as D_X. For instance, the HMM trained with the Bach Chorales will be symbolised as M_{Bach} while its testing pieces will be symbolised as D_{Bach}. The evaluation of whether a model M_X predicts a subset D_X better than a subset D_Y is achieved through the cross-entropy measure. The measure of cross-entropy is utilised to provide an entropy value for a sequence from a dataset, $\{S_i, i \in \{1, 2, \ldots, n\}\} \in D_X$, according to the context of each sequence element, S_i, denoted as C_i, as evaluated by a model M_Y. The value of cross-entropy under this formalisation is given by

$$-\frac{1}{n}\sum_{1}^{n} \log P_{M_Y}(S_i, C_{i,M_Y}), \tag{2}$$

where $P_{M_Y}(S_i, C_{i,M_Y})$ is the probability value according to the examined scenarios of probabilities.

By comparing the cross-entropy values of a sequence X as predicted by two models, D_X and D_Y, we can assume which model predicts S better: the model that produces the *smaller* cross entropy value [8]. Smaller cross entropy values indicate that the elements of the sequence S "move on a path" with greater probability values. Tables 4 exhibits the cross-entropy values produced by the proposed model for the examined scenarios. The presented values are averages across 100 repetitions of the experimental process, with different random divisions in training and testing subsets (preserving a ratio of 90%-10% respectively for all repetitions). In every repetition the average cross entropy of all the testing sequences is calculated. The effectiveness of the combined proposed modules is indicated by the fact that most of the minimum values per row are on the main diagonal of the upper part of the matrix, i.e. where model M_X^{all} predicts D_X better than any other D_Y. A 10-fold cross-validation routine was also tested for splitting the dataset, however, replications of the experiment where different pieces in training and testing sets were used, gave considerably different results. The utilised experimental setup was providing similar results in several replications of the experiment.

It is evident that each module isolated does not produce lower values in the diagonal. Among the clearest isolated characteristics is the melody observations part of the HMM (M^{mel}), where 5 out of 8 diagonal elements are the lowest in their row. Thereby, these results indicate that the combination of all modules is a vital part for achieving better results.

3.2 Diversity in inversions and doublings of GCT chords

A straightforward comparison in statistics related to inversions and doublings between GCTs of different idioms is not possible for all idioms and all GCTs, since this information is harnessed on GCT sets that are in many cases different for different idioms. The differences in characteristics about voicing layout between different sets of GCTs that could be envisaged, relate to the *diversity* of the voicing layout scenarios that are used across different idioms.

	D_{Bach}	D_{Beattles}	D_{Epirus}	D_{Medieval}	D_{Modal}	$D_{\text{Rembetika}}$	$D_{\text{Stravinsky}}$	D_{Tango}
$M^{\text{all}}_{\text{Bach}}$	7.17	11.07	15.75	10.79	7.41	9.77	11.86	8.88
$M^{\text{all}}_{\text{Beattles}}$	9.75	7.82	15.97	14.86	9.77	8.27	7.64	9.01
$M^{\text{all}}_{\text{Epirus}}$	16.64	19.62	6.99	10.54	13.11	14.30	16.11	16.46
$M^{\text{all}}_{\text{Medieval}}$	10.96	17.56	7.68	7.47	8.49	12.46	16.18	12.63
$M^{\text{all}}_{\text{Modal}}$	9.27	15.94	15.04	10.96	8.39	10.89	15.32	10.72
$M^{\text{all}}_{\text{Rembetika}}$	8.73	8.56	13.65	11.79	8.22	7.11	7.80	8.29
$M^{\text{all}}_{\text{Stravinsky}}$	14.19	10.82	17.45	19.88	15.84	10.99	9.76	13.88
$M^{\text{all}}_{\text{Tango}}$	8.27	8.78	14.62	11.33	7.98	7.62	9.35	7.70
$M^{\text{tr}}_{\text{Bach}}$	2.09	2.61	3.16	2.25	2.24	2.99	2.97	2.62
$M^{\text{tr}}_{\text{Beattles}}$	3.51	2.33	2.47	3.30	2.88	1.82	2.28	2.20
$M^{\text{tr}}_{\text{Epirus}}$	5.39	3.17	2.04	4.90	4.31	2.06	2.64	3.78
$M^{\text{tr}}_{\text{Medieval}}$	2.73	2.92	1.97	2.33	2.33	2.49	2.74	3.11
$M^{\text{tr}}_{\text{Modal}}$	2.87	2.92	2.82	2.41	3.32	2.79	2.73	3.07
$M^{\text{tr}}_{\text{Rembetika}}$	4.11	2.66	1.90	3.53	3.21	1.67	1.88	2.62
$M^{\text{tr}}_{\text{Stravinsky}}$	5.44	3.98	2.51	4.51	4.73	2.63	3.50	4.50
$M^{\text{tr}}_{\text{Tango}}$	3.11	2.16	2.82	2.98	3.02	1.88	2.55	2.12
$M^{\text{mel}}_{\text{Bach}}$	1.79	2.14	2.28	1.95	1.85	2.34	2.44	2.15
$M^{\text{mel}}_{\text{Beattles}}$	2.34	1.92	2.09	2.26	1.93	1.65	1.87	1.86
$M^{\text{mel}}_{\text{Epirus}}$	2.72	2.43	1.42	2.21	2.43	1.72	1.74	2.59
$M^{\text{mel}}_{\text{Medieval}}$	2.54	3.32	2.15	2.13	2.50	2.36	2.51	3.04
$M^{\text{mel}}_{\text{Modal}}$	2.68	2.60	2.57	2.64	2.36	2.12	2.55	2.59
$M^{\text{mel}}_{\text{Rembetika}}$	2.81	2.13	1.86	2.39	2.20	1.37	2.17	2.00
$M^{\text{mel}}_{\text{Stravinsky}}$	3.77	3.12	2.29	3.85	3.39	2.83	2.53	3.77
$M^{\text{mel}}_{\text{Tango}}$	2.33	1.86	1.94	2.36	1.90	1.48	2.17	1.72
$M^{\text{mbd}}_{\text{Bach}}$	3.58	6.51	10.50	6.77	3.55	4.45	5.65	4.25
$M^{\text{mbd}}_{\text{Beattles}}$	4.90	4.24	12.17	10.13	5.63	4.72	3.90	5.38
$M^{\text{mbd}}_{\text{Epirus}}$	9.03	14.89	3.51	4.14	6.83	10.31	12.04	10.34
$M^{\text{mbd}}_{\text{Medieval}}$	6.10	13.05	3.77	3.93	4.57	7.72	10.82	7.15
$M^{\text{mbd}}_{\text{Modal}}$	4.44	11.53	10.35	6.48	3.47	6.18	9.70	5.63
$M^{\text{mbd}}_{\text{Rembetika}}$	3.79	4.80	10.59	6.80	3.92	4.11	4.32	4.20
$M^{\text{mbd}}_{\text{Stravinsky}}$	5.87	4.56	12.91	12.08	8.00	6.18	4.67	6.73
$M^{\text{mbd}}_{\text{Tango}}$	3.64	5.35	10.38	6.56	3.70	4.12	4.78	4.19

Table 4. Mean values of cross-entropies for all pairs of datasets, for all the combination of all probabilities, as well as in isolation concerning previous bass motion, melody motion and bass-to-melody distance.

Along these lines, the question would be: are there more diverse chord expressions regarding inversions and doublings – regardless of which chords (GCTs) – in the chorales of Bach, than in the modal chorales? The *diversity* in a discrete probability distribution (like the ones displayed in the examples of Table 2) is measured by the Shannon information entropy [21] (SIE). The SIE reflects the diversity in possibilities described by discrete probability distribution, with higher SIE values indicating a more random distribution with more diverse / less expectable outcomes. Therefore, by measuring the SIE values of all GCTs and comparing them for every pair of idioms, it can be concluded whether some idioms have richer possibilities for the voicing layouts of chords than others.

Table 5 exhibits the results of a test in the statistical significance in differences between the SIE values in every pair of idioms. The upper-diagonal elements concern inversions, while lower-diagonal elements doublings. A value of $+1$ indicates that the GCTs in the idiom of the row are statistically significantly more diverse in their voicing layout – according to the mean SIE values – than the ones in the idiom of the column. A -1 value indicates the opposite, while a 0 value indicates no statistically significant difference. The statistical significance is measured through a two–sided Wilcoxon [23] rank sum test, which is applied on the SIE values of all GCT voicing layout distributions for every idiom. The statistical significance test in statistics related to voice layout reveal that few datasets are significantly superior or inferior regarding their diversity.

3.3 Example compositions

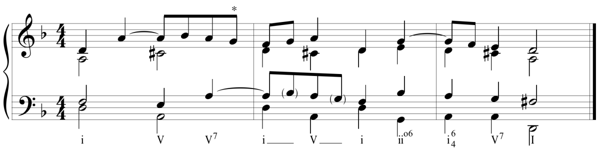
The proposed bass voice leading methodology was utilised in an “off-line” mode to produce two examples. The term “off-line” indicates the fact that the system was used to generate a single description for the bass voice leading on a given set of chords (in GCT representation [2] produced by a probabilistic chord-generation model [10]). This means that if no inversion of the predetermined chord can satisfy the requirements of the bass voice leading, then the system simply selected the most probable inversion of this chord, regardless of the bass voice leading indication. The bass voice for the generated examples was selected using the $argmax$ function mentioned in Section 2.4, which allows the reflection of some typical idiom characteristics, even though such an approach does not necessarily guaranty interestingness [3] (since the most “expected” scenario is fol-

	S_{Bach}	S_{Beattles}	S_{Epirus}	S_{Medieval}	S_{Modal}	$S_{\text{Rembetika}}$	$S_{\text{Stravinsky}}$	S_{Tango}
S_{Bach}	0	1	1	-1	1	1	1	1
S_{Beattles}	0	0	0	-1	0	0	0	0
S_{Epirus}	0	0	0	-1	0	0	0	0
S_{Medieval}	-1	-1	-1	0	1	1	1	1
S_{Modal}	0	0	0	1	0	1	1	0
$S_{\text{Rembetika}}$	0	-1	0	1	0	0	0	0
$S_{\text{Stravinsky}}$	0	0	1	0	1	1	0	0
S_{Tango}	0	0	0	1	0	0	0	0

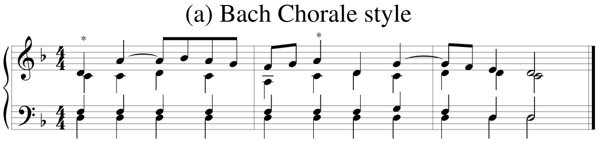
Table 5. Statistical significance of differences in the diversity of inversions (upper diagonal) and doublings (lower diagonal). Statistically significant superiority of diversity in the row dataset is exhibited with a $+1$, of the column dataset with -1, while 0 indicates no statistical significance in diversity differences.

lowed). The intermediate voices where manually adjusted by a music expert.

The presented examples (Figure 2) include two alternative harmonisations of a Bach Chorale melody with both the chord generation and the bass voice leading systems trained on sets of (a) the Bach Chorales and (b) polyphonic songs from Epirus. In the case of the Bach chorale, the system made erroneous bass voice assignments in the second bar that create consecutive anti-parallel octaves between the outer voices (due to the chord incompatibility problem discussed above) [1]. The harmonisation in the style of the polyphonic songs from Epirus indeed preserves an important aspect of these pieces: the drone note.



(a) Bach Chorale style



(b) Polyphonic Epirus songs style

Figure 2. Harmonisation examples in two different styles. Chord sequences in the GCT representation were previously produced by another probabilistic system.

4. CONCLUSIONS

This paper presented a modular methodology for determining the bass voice leading in automated melodic harmonisation given a melody voice and a sequence of chords. In this work it is assumed that harmony is not solely the expression of a chord sequence, but also of harmonic movement for all voices that comprise the harmonisation. The presented work focuses on generating the bass voice on a given sequence of chords by utilising information from the soprano /melody voice and other statistics that are related to the layout of the chords, captured by different statistical modules. Specifically, a hidden Markov model (HMM) is utilised to determine the most probable movement for the bass voice (hidden states), by observing the soprano movement (set of observations), while additional voicing layout characteristics of the incorporated chords are considered that include distributions on the distance between the bass and the melody voice and statistics regarding the inversions and doublings of the chords in the given chord sequence.

Experimental results evaluate that the learned statistical values from an idiom's data are in most cases efficient for capturing the idiom's characteristics in comparison to others. Additionally, similar tests were performed for each statistical module of the model in isolation, a process that revealed whether some characteristics of the examined idioms are more prominent than others. Furthermore, preliminary music examples indicate that the proposed methodology indeed captures some of the most prominent characteristics of the idioms it is being trained with, despite the fact that further adjustments are required for its application in melodic harmonisation.

[1] Another voice-leading issue occurs at the first beat of the 3rd bar, where the D in the 2nd voice is introduced as unprepared accented dissonance. Note that the parenthesised pitches in the 3rd voice (bar 2) were introduced manually (not by the system) to create imitation.

5. ACKNOWLEDGEMENTS

This work is founded by the COINVENT project. The project COINVENT acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 611553. The authors would like to thank Costas Tsougras for his assistance in preparing the presented musical examples.

6. REFERENCES

[1] Moray Allan and Christopher K. I. Williams. Harmonising chorales by probabilistic inference. In *Advances in Neural Information Processing Systems 17*, pages 25–32. MIT Press, 2004.

[2] Emilios Cambouropoulos, Maximos Kaliakatsos-Papakostas, and Costas Tsougras. An idiom-independent representation of chords for compu-

tational music analysis and generation. In *Proceeding of the joint 11th Sound and Music Computing Conference (SMC) and 40th International Computer Music Conference (ICMC)*, ICMC–SMC 2014, 2014.

[3] Tom Collins. *Improved methods for pattern discovery in music, with applications in automated stylistic composition*. PhD thesis, The Open University, 2011.

[4] Patrick Donnelly and John Sheppard. Evolving four-part harmony using genetic algorithms. In *Proceedings of the 2011 International Conference on Applications of Evolutionary Computation - Volume Part II*, EvoApplications'11, pages 273–282, Berlin, Heidelberg, 2011. Springer-Verlag.

[5] Kemal Ebcioglu. An expert system for harmonizing four-part chorales. *Computer Music Journal*, 12(3):43–51, 1988.

[6] David Huron. Voice denumerability in polyphonic music of homogeneous timbres. *Music Perception*, pages 361–382, 1989.

[7] Michael I. Jordan, Zoubin Ghahramani, and Lawrence K. Saul. Hidden markov decision trees. In Michael Mozer, Michael I. Jordan, and Thomas Petsche, editors, *NIPS*, pages 501–507. MIT Press, 1996.

[8] Dan Jurafsky and James H. Martin. *Speech and language processing*. Prentice Hall, New Jersey, USA, 2000.

[9] M. Kaliakatsos-Papakostas, A. Katsiavalos, C. Tsougras, and E. Cambouropoulos. Harmony in the polyphonic songs of epirus: Representation, statistical analysis and generation. In *4th International Workshop on Folk Music Analysis (FMA) 2014*, June 2011.

[10] Maximos Kaliakatsos-Papakostas and Emilios Cambouropoulos. Probabilistic harmonisation with fixed intermediate chord constraints. In *Proceeding of the joint 11th Sound and Music Computing Conference (SMC) and 40th International Computer Music Conference (ICMC)*, ICMC–SMC 2014, 2014.

[11] Kostas Liolis. *To Epirótiko Polyphonikó Tragoúdi (Epirus Polyphonic Song)*. Ioannina, 2006.

[12] Dimos Makris, Maximos Kaliakatsos-Papakostas, and Emilios Cambouropoulos. A probabilistic approach to determining bass voice leading in melodic harmonisation. In Tom Collins, David Meredith, and Anja Volk, editors, *Mathematics and Computation in Music*, volume 9110 of *Lecture Notes in Computer Science*, pages 128–134. Springer International Publishing, 2015.

[13] Leonard C. Manzara, Ian H. Witten, and Mark James. On the entropy of music: An experiment with bach chorale melodies. *Leonardo Music Journal*, 2(1):81–88, January 1992.

[14] Francois Pachet and Pierre Roy. Formulating constraint satisfaction problems on part-whole relations: The case of automatic musical harmonization. In *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI 98)*, pages 1–11. Wiley-Blackwell, 1998.

[15] Francois Pachet and Pierre Roy. Musical harmonization with constraints: A survey. *Constraints*, 6(1):7–19, January 2001.

[16] Jean-François Paiement, Douglas Eck, and Samy Bengio. Probabilistic melodic harmonization. In *Proceedings of the 19th International Conference on Advances in Artificial Intelligence: Canadian Society for Computational Studies of Intelligence*, AI'06, pages 218–229, Berlin, Heidelberg, 2006. Springer-Verlag.

[17] Risto Pekka Pennanen. The development of chordal harmony in greek rebetika and laika music, 1930s to 1960s. *British Journal of Ethnomusicology*, 6(1):65–116, 1997.

[18] Somnuk Phon-amnuaisuk and Geraint A. Wiggins. The four-part harmonisation problem: A comparison between genetic algorithms and a rule–based system. In *In proceedings of the AISB99 symposium on musical cretivity*, pages 28–34. AISB, 1999.

[19] Martin Rohrmeier and Thore Graepel. Comparing feature-based models of harmony. In *Proceedings of the 9th International Symposium on Computer Music Modelling and Retrieval*, pages 357–370, 2012.

[20] M. Schorlemmer, A. Smaill, K.U. Kühnberger, O. Kutz, S. Colton, E. Cambouropoulos, and A. Pease. Coinvent: Towards a computational concept invention theory. In *5th International Conference on Computational Creativity (ICCC) 2014*, June 2014.

[21] C. E Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5:3–55, January 2001.

[22] Raymond P. Whorley, Geraint A. Wiggins, Christophe Rhodes, and Marcus T. Pearce. Multiple viewpoint systems: Time complexity and the construction of domains for complex musical viewpoints in the harmonization problem. *Journal of New Music Research*, 42(3):237–266, September 2013.

[23] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.

[24] Liangrong Yi and Judy Goldsmith. Automatic generation of four-part harmony. In Kathryn B. Laskey, Suzanne M. Mahoney, and Judy Goldsmith, editors, *BMA*, volume 268 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2007.