

AUTOMATIC FIGURED BASS ANNOTATION USING THE NEW BACH CHORALES FIGURED BASS DATASET

Yaolong Ju¹ Sylvain Margot¹ Cory McKay² Luke Dahn³ Ichiro Fujinaga¹

¹Schulich School of Music, McGill University, Canada

²Department of Liberal and Creative Arts, Marianopolis College, Canada

³School of Music, The University of Utah, USA

yaolong.ju@mail.mcgill.ca, sylvain.margot@mail.mcgill.ca,
cory.mckay@mail.mcgill.ca, luke.dahn@utah.edu, ichiro.fujinaga@mcgill.ca

ABSTRACT

This paper focuses on the computational study of *figured bass*, which remains an under-researched topic in MIR, likely due to a lack of machine-readable datasets. First, we introduce the Bach Chorales Figured Bass dataset (BCFB), a collection of 139 chorales composed by Johann Sebastian Bach that includes both the original music and figured bass annotations encoded in MusicXML, **kern, and MEI formats. We also present a comparative study on automatic figured bass annotation using both rule-based and machine learning approaches, which respectively achieved classification accuracies of 85.3% and 85.9% on BCFB. Finally, we discuss promising areas for MIR research involving figured bass, including automatic harmonic analysis.

1. INTRODUCTION

Figured bass is a type of music notation that uses numerals and other symbols to indicate intervals to be played above a bass note, and which can provide insight on underlying harmonies [1]. It was commonly used in Baroque music, and served as a guide for performance, especially for the instruments improvising the *basso continuo* accompaniment (e.g., harpsichord, organ, lute, etc.). Fig. 1 shows an example of figured bass, as well as how a harpsichordist might realize the figured bass as an improvised accompaniment. Such realizations are not typically explicitly included in scores, as the musical tradition of the time left them to be improvised based on the skills and taste of the continuo player. There are three aspects of figured bass annotations (FBAs) that should be highlighted:

- (1) The Neue Bach Ausgabe edition [2] uses FBAs consisting of numbers with backslashes through them to indicate raised intervals (e.g., m. 3.3 and m. 3.4). Forward slashes indicate lowered intervals (e.g., ♯).
- (2) FBAs followed by continuation lines indicate that the harmony of the preceding figure is prolonged (e.g., m. 1.4, m. 4.2, and m. 4.4).
- (3) Multiple FBAs over a stationary bass (e.g., 4–3 in m. 5) usually indicate a suspension being resolved.

Figured bass also serves pedagogical and theoretical purposes: not only does it provide contrapuntal information on how to conduct the resolution of dissonances, it also offers insights into the chords and harmonic rhythm intended by composers. Figured bass can therefore provide a preliminary description of harmonic structure, and serves as a promising basis for approaching harmonic analysis.

As a useful analytical tool for studying Baroque compositional and performance practices, figured bass has been an important topic in music pedagogy [3], music theory, and musicology [4]. The computational study of figured bass, however, has drawn little attention over the years. We have only found two papers on automatic figured bass annotation, both using a rule-based approach: Barthélemy and Bonardi treated figured bass as a harmonic reduction and devised rules to identify and remove ornamental notes, permitting them to cluster the remaining chord tones as figures [5]; Wead and Knopke, in contrast, manually designed a decision tree to determine the figured bass for a given bass line [6]. Unfortunately, with no open-source code and a lack of quantitative results, it is impossible to objectively evaluate or compare the performances of these models. Furthermore, we are not aware of any previous applications of machine learning to figured bass, nor of any existing digital dataset with figured bass annotations (FBAs). These limitations have likely limited the computational study of figured bass to date.

There are four main contributions of this paper:

- (1) We introduce the new Bach Chorales Figured Bass (BCFB) dataset, which consists of 139 chorales composed by Johann Sebastian Bach (Section 2). These chorales came from larger choral works composed by Bach: the cantatas, passions, motets, and the Christmas Oratorio. We chose this repertoire due to its key role in modern music pedagogy and its general historical importance.
- (2) In order to facilitate the future creation of more figured bass datasets, we include our methodology for digitizing FBAs in an efficient and effective way.
- (3) We present a comparative study of automatic figured bass annotations of BCFB, using both rule-based and machine learning approaches (Section 3). The results are discussed with reference to specific musical examples (Section 4).
- (4) We highlight possible applications of figured bass annotation, especially in connection with converting figured bass to chord labels, which could benefit research on automatic harmonic analysis (Section 5).



The figure shows a musical score for three instruments: Flute, Harpsichord, and Continuo. The Continuo line is in the bass clef with a 4/4 time signature and a key signature of two flats. Below the Continuo line are figured bass annotations: 6 - 5 #4 5 6 5 6 5 - 6 - 4 3. The Harpsichord line is in the grand staff (treble and bass clefs) and shows an improvisation based on the Continuo line. The Flute line is in the treble clef and plays a melodic line. The score is numbered 1 at the beginning.

Figure 1. A sample musical passage we composed, where figured bass annotations (FBAs) are shown below the continuo line, and where we added the harpsichord line as an example of what a continuo player might improvise based on the figured bass. Figures indicate intervals above the continuo line that could be played in the improvisation. For example, the “6” in the first measure corresponds to the pitch class “G”, which is a 6th above the bass “B \flat ”. An actual improvisation would likely also typically contain the pitch class “D” (a 3rd above the bass “B \flat ”) in this slice, but this is not explicitly indicated in the figures. This is an example of how FBAs do not always specify all the notes to be played by the continuo player, and usually omit some obvious figures (see Section 3.2.2 for details).

2. BACH CHORALES FIGURED BASS DATASET

To the best of our knowledge, there is no prior publicly available digital figured bass dataset. We therefore present the Bach Chorales Figured Bass dataset (BCFB), a corpus we constructed containing FBAs in MusicXML, ****kern**, and MEI (Music Encoding Initiative) formats. It consists of all 139 J. S. Bach four-voice chorales that include his own figured bass, based on the *Neue Bach Ausgabe* (NBA) critical edition [2]. NBA is chosen as the source of BCFB because it is the most up-to-date scholarly critical edition, prepared with exacting methods of source criticism.

2.1. Finding Chorales with FBAs

To find all the chorales attributed to Bach, we constructed a reference table (<https://bit.ly/303jzfs>) with all 420 chorales indexed by BWV catalogue numbers, and cross-referenced them with the NBA. We checked whether original FBAs are accessible for each of these chorales, and found 139 settings meeting this criterion. We then made an expanded reference table, consisting of the: BWV number,¹ Breitkopf number (when relevant),² title of the work of origin (e.g. cantata, passion, etc.), date of the first performance, text setting, location of the score in the NBA edition, and other musicological metadata for each cho-

rale. This table is designed to facilitate musicological research, which, along with the BCFB dataset, is available at: <https://bit.ly/2OoWC16>.

2.2. Digitization

We began the creation of BCFB by assembling existing symbolic encodings of the relevant Bach chorales from the KernScores repository, which contained 109 of the 139 NBA chorales with FBAs.³ We automatically translated these 109 ****kern** files into MusicXML using *music21* (v. 5.1.0),⁴ and made changes to match the musical content⁵ of the NBA edition before adding Bach’s figured bass. We manually encoded the remaining 30 figured chorales found in the NBA edition. We chose MusicXML as our master file format since it is widely supported by music notation software. We used the *MuseScore* 3 score editor⁶ for both editing musical content and adding FBAs.⁷

2.3. Converting to Other Symbolic File Formats

BCFB includes encodings in two other symbolic file formats⁸ beyond MusicXML: ****kern**, and MEI.⁹ Existing software was used to automatically convert the original MusicXML to the other two formats. Through a series of experiments with a variety of alternatives, we found that converting figured bass from MusicXML to ****kern** using *musicxml2hum*¹⁰ worked well, except for the continuation lines, and the conversion from ****kern** to MEI using *Verovio*¹¹ was perfect. However, direct conversion of figured

¹ Bach-Werke-Verzeichnis (BWV) catalogue number, which indexes all the compositions attributed to J. S. Bach.

² The Breitkopf edition contains 371 four-voice J. S. Bach chorales, and indexes them differently from BWV.

³ KernScores (kern.ccarh.org) is maintained by Stanford’s Center for Computer Assisted Research in the Humanities, and includes 371 four-part chorales encoded in the Humdrum ****kern** representation (www.humdrum.org).

⁴ <https://web.mit.edu/music21>

⁵ Including: adding a continuo line and/or instrumental voices; transposing; changing the meter, pitch, and duration of certain notes; etc. We did not encode the textual content specified in the NBA.

⁶ <https://musescore.org>

⁷ <https://musescore.org/en/handbook/figured-bass>

⁸ This diversity of symbolic formats offers researchers the opportunity to use the format most convenient to their preferred software, because if only one format were offered, which might not be supported by a given piece of preferred research software, then it would need to be converted to the format supported by the software. This could lead to a potential loss of figured bass information or to other conversion errors [7].

⁹ We also used *music21* (v. 5.1.0) to generate MIDI files from the master MusicXML, but they do not include FBAs.

¹⁰ <https://github.com/craigshapp/humlib>

¹¹ <https://github.com/rism-ch/verovio>

Figure 2. Measures 3 and 4 from BWV 117.04 *Sei Lob und Ehr dem höchsten Gut*. The original FBAs are annotated underneath the bass voice part. Note that not all slices are necessarily figured and not all the intervals in a sonority are necessarily specified in FBAs. We artificially added the final bottom staff, which collapses all sonorities into one octave so as to more directly reveal the pitch-class content. The number of semitones above the bass implied by the original FBAs have also been added underneath this bottom staff. We can also translate the number of semitones back to FBAs by examining the actual notes in the score and then calculating and labelling the intervals from the bass note.

bass from MusicXML to MEI (using *Verovio*) was problematic, as accidentals, slashes, and continuation lines were not converted. We therefore first converted from MusicXML to `**kern`, and then manually added the continuation lines to the `**kern` files using a text editor. The resulting `**kern` files were then converted to MEI files.

3. AUTOMATIC FIGURED BASS ANNOTATION

We automatically generated¹² FBAs that closely resembles those of Bach for two main reasons: to learn about Bach’s figured bass habits, which are of musicological interest, and to provide figured bass for those Bach chorales for which no FBAs exist. We used both rule-based and machine learning algorithms to perform this automatic figured bass annotation:¹³ the rule-based approach has the potential to model Bach’s style of writing figures in ways that are easily human-interpretable, and machine learning has the potential to model patterns in Bach’s style that might be difficult to codify into precise, direct rules. We therefore explored the efficacy of both approaches.

3.1. Data

We used 120 chorales out of the full 139 chorales in BCFB to train and test our models. We excluded 12 interlude chorales¹⁴ because they are significantly different from the

other largely homorhythmic chorales, and we excluded five other chorales¹⁵ that are barely figured. Finally, we excluded BWV 8.06¹⁶ and BWV 161.06 because they feature irregular textures, such as having an obbligato continuo and/or instrumental part.

3.2. Rule-based algorithms

3.2.1. Initial Simple Rule-based Algorithm

We began by implementing a simple rule-based algorithm that labels all the intervals above the bass in the generated FBAs. First, the music is segmented into a series of *note onset slices* [8,9]. A new slice is formed whenever a new note onset occurs in any musical voice, and each slice consists of the vertical set of notes sounding at that moment. Take the first slice of Fig. 2 as an example: since the pitch classes above the bass “G” are “G”, “D”, and “B,” the FBA generated is $8/5/3$.

We then compared the generated FBAs against Bach’s original FBAs, and found that the percentage of exact matches was only 3%. This is partly because Bach did not explicitly label all the intervals above the bass in his FBAs; it is often assumed that both he and other Baroque composers employed FBAs that include what are in effect abbreviations that omit obvious intervals [1,10]. For example, consider m. 3.2.5¹⁷ of Fig. 2: although the pitch classes “A” and “D” are present in this slice above the bass “F#”, only “D” is explicitly specified¹⁸ by the figure “6”; “3” is not explicitly indicated, but is nonetheless implied.

3.2.2. Evaluation Metric

To allow for the equivalence in musical content of different figured bass notation conventions, as discussed above, we created an evaluation metric that treats figures that are musically equivalent as notationally equivalent. The purpose of this metric is to realistically evaluate the generated figured bass when it does not match Bach’s figured bass exactly. The equivalence rules are inspired by Arnold [11]:

- A “3” can be omitted (Fig. 3a, 3b, 3d, and 3e) unless there is a 4th in the sonority,¹⁹ or unless the 3rd is the resolution of a 4–3 suspension (Fig. 3c).
- A “5” (Fig. 3a, 3c, and 3d) can be omitted, unless one of the following conditions is true: there is a 6th (Fig. 3b) in the sonority, the 5th is the resolution of a 6–5 suspension, or the 5th has an accidental (Fig. 3e).
- An “8” (Figs. 3c and 3d) can be omitted, unless one of the following conditions is true: there is a 9th (Fig. 3b) in the sonority, the 8th is the resolution of a 9–8 suspension, or the 8th has an accidental.
- A “6” can be omitted if the sonority forms a “6/4/3” or a “6/4/2” chord, as shown in Fig. 3f.

¹² The generated FBAs use flats and sharps to respectively indicate lowered and raised intervals, and do not contain continuation lines.

¹³ The code is available at: <https://bit.ly/2P8Qbjv>.

¹⁴ These chorales have elaborate instrumental interludes between phrases (BWV 24.06, 76.07, 100.06, 105.06, 113.01, 129.05, 167.05, 171.06, 248.09, 248.23, 248.42, and 248.64).

¹⁵ BWV 16.06, 48.07, 149.07, 195.06, and 447.

¹⁶ “.06” in “BWV 8.06” means this chorale is the sixth movement of the “BWV 8” cantata.

¹⁷ “m. 3.2.5” means the third measure, the second and half beat.

¹⁸ Since “D” forms a 6th interval above the bass “F#”.

¹⁹ “Sonority” means the set of pitch classes present in a note onset slice. For example, the added bottom staff of Fig. 2 shows the sonorities of the four voices for each slice. “4th” means that a wrapped interval of a 4th can be found between the bass and an upper voice, regardless of whether it is labelled in the figured bass.

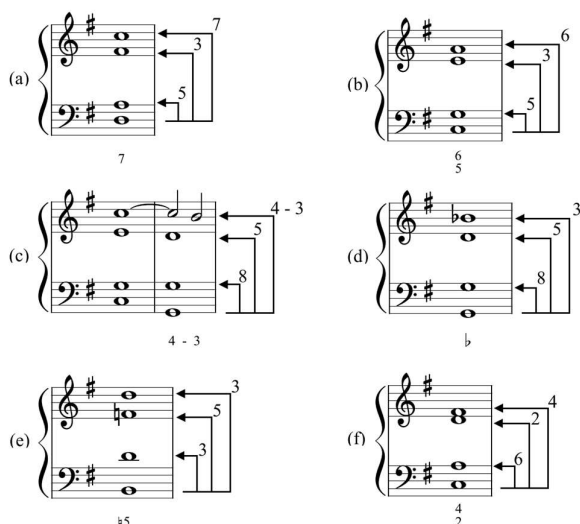


Figure 3. Common examples of standard figured bass abbreviations taken into account by the evaluation metric explained in Section 3.2.2. In each of the six examples (a)–(f), all the intervals above the bass are shown to the right of the notes connected with arrows, and typical abbreviated FBAs for the chords are shown below the notes. For example, (a) consists of a dominant 7th chord in root position, and includes notes that are a 3rd, 5th, and 7th above the bass: the figured bass consists of only a “7”, with the “3” and “5” omitted, as is often the practice in FBA.

In order to see how the evaluation metric based on these rules operates in practice, consider Fig. 3a as an example: we can see that “5” and “3” can be omitted, which means “7”, “7/3”, “7/5”, and “7/5/3” are all considered equivalent. Therefore, if the ground truth and the generated figured bass respectively consist of any pairing of “7”, “7/3”, “7/5”, or “7/5/3”, then the generated figured bass will be considered correct by the metric.

3.2.3. Improved Rule-based Algorithm

Using this evaluation metric, the simple rule-based algorithm described in Section 3.2.1 has an effective agreement of 64.5% with Bach’s FBAs. We found that when they disagreed, the generated FBAs tend to have more figures than those of Bach. To improve this, we manually developed additional rules for omitting certain figures, permitting us to better predict Bach’s style of annotation.²⁰

First, we examine each note onset slice and omit the figure for a given note in an upper voice if both of the following two conditions are met: (1) the note is labelled in the previous slice, and (2) the pitch class of the bass in the current slice remains the same as in the previous slice.

Then, we consider slices on fractional beats (e.g., beat 2.5 and 3.5), looking for ornamental notes, such as passing tones, neighbour tones, escape tones, and anticipations, which are all approached or departed by step. If such a note

is in an upper voice, its corresponding number is removed from the figure; if such a note is in the bass, the slice is left entirely unfigured.

After the addition of these rules, the model was able to achieve 85.3% agreement with Bach’s figures, a large improvement over the 64.5% agreement achieved with the simple method and equivalency rules.

Although it would have been possible to invest more time manually analyzing Bach’s figured bass to develop still more rules to improve agreement, none were readily obvious from the perspective of music theory or performance practice, and we wanted to avoid overfitting our rules. So, we turned our attention to machine learning, to see if it could be employed to model Bach’s FBA style with equal or better results.

3.3. Machine learning algorithms

In order to efficiently perform automatic figured bass annotation with machine learning methods—something that has never been explored in the literature—we transformed the FBAs into interval-class vectors.

3.3.1. Transformation from Figured Bass to Interval Classes

Recall that figured bass indicates intervals above the bass note. Thus, for each slice, we convert the figures to an interval-class vector.²¹ An interval class, similar to a pitch class, is a set of intervals wrapped by octaves. For example, an interval class of a major second includes a major ninth and all other octave expansions of a major second. As with a pitch-class vector, an interval-class vector contains 12 elements, representing intervals in semitone increments. In our case, each FBA is converted to an interval-class vector that includes all the notes above the bass that are sounding in the current slice. In cases where the figured bass does not specify the exact interval in semitones, such as the “6”, which could be either a major 6th or a minor 6th, we rely on the score to determine the exact interval using heuristics-based post-processing. We similarly rely on the score to later convert interval-class vectors back to figures: for example, an interval of three semitones can be interpreted as either a minor third (figure “b3”) or an augmented second (figure “#2”). We can decide which is appropriate by considering the pitch spelling in the original score.²² This representation of the figured bass is used for both input and output of the machine learning algorithms.

3.3.2. Input Features

The three feature vectors used as input to the machine learning algorithms are: (1) interval classes (see Section 3.3.1); (2) onsets, which specify which notes above the bass have onsets within the slice, as opposed to being held from a previous slice; and (3) metrical context, which specifies whether a slice occurs on the downbeat of a measure, on another beat (e.g., beat 2, 3, or 4 in 4/4), or on

²⁰ The rules were proposed by observing the generated FBAs and were evaluated against the ground truth FBAs from BCFB. We selected the rules that yielded higher accuracy.

²¹ We are using the term “interval class” here to refer to *ordered* interval class. Since we wish to calculate the intervallic relationship *from* the bass *to* and upper voice (in that order), we wish to distinguish between intervals and their inversional equivalents (e.g. minor 7^{ths} are distinguished

from major 2^{nds}). Thus, ordered interval classes range from 0 to 11, while unordered interval classes range from 0 to 6 only.

²² For example, to distinguish “2” and “9” when unwrapping interval-class vectors, we need to find the actual note above the bass and compare its pitch to the pitch of the bass. If they are one octave apart, the generated figure will be “9”, and “2” otherwise.

a fractional beat (e.g., beat 3.5). These binary vectors are specified for each slice. The following example demonstrates each of these feature vectors for m. 4.2.5 of Fig. 2 (the bit-length of each feature is indicated in parentheses):

- *Interval classes (12)*: The bass note is “A” (held), with pitch classes of “C#”, “G”, and “E” (held) above it, which are respectively four, ten, and seven semitones away from it. The feature vector is thus: [0,0,0,0,1,0,0,1,0,0,1,0].²³
- *Onsets (12)*: “C#” and “G”, which are respectively four and ten semitones above the bass, are the pitch classes with onsets on this slice, so the feature vector will be [0,0,0,0,1,0,0,0,0,0,1,0].
- *Metrical context (3)*: Because the slice is on beat 2.5 of a 4/4 measure, the feature vector will be [0,0,1] (i.e., it is a fractional beat).

Each slice, therefore, is represented as a 27-dimensional (12+12+3=27) binary vector. To provide a context for each slice, the machine learning algorithms are also provided with the two 27-dimensional vectors for the previous and following slices (zero-padded for first and the last slices). Thus, the total length of the input vector for each slice is 81 (27×3=81).

3.3.3. Machine Learning Algorithms

We experimented with two machine learning algorithms: Decision Trees (DT)²⁴ and Deep Neural Networks (DNN).²⁵ Both algorithms used the input features specified above. Their output each consisted of a 12-dimensional binary vector specifying the number of semitones above the bass,²⁶ as discussed in Section 3.3.1.

3.3.4. Experimental Setup

Ten-fold cross-validation was used for evaluation. For the DNN experiments, we divided the data into training (80%), validation (10%),²⁷ and testing (10%) folds. For the DT experiments, the data was divided into training (90%, the union of the DNN training and validation sets) and testing (10%, matching the DNN test sets) partitions.

3.3.5. Results

The Decision Trees and Deep Neural Networks respectively achieved classification accuracies of 84.3±0.5% and 85.9±0.6% on BCFB.²⁸ These accuracies are calculated based on the evaluation metrics proposed in Section 3.2.2.

4. DISCUSSION

It is useful to examine the types of errors that our model made, in order to better understand its performance and how it can be improved. We will focus this discussion on the two musical examples shown in Fig. 4, as they are representative of the kinds of errors our model made. One

common error made by our model was to miss figures that indicate the resolution of a suspension, such as the 9–8 shown in Fig. 4(a), m. 8.4. This may be because the features we used did not contain sufficient voice-leading information to detect such suspensions.

Two further types of disagreement between our model and Bach’s figures are shown in Fig. 4(b). At m. 2.3 our model generated “#”, but the ground truth had no label. In fact, the generated “#” is technically correct, as the D is explicitly sharpened in the soprano. Turning to m. 3.2, our model’s prediction included a “#7,” unlike the ground truth. Perhaps this suggests that Bach might have considered the corresponding “D#” to be a passing tone? Or perhaps the D# was understood as a “diatonic” note in this Dorian chorale tune? At any rate, the “#7” in the generated figures should not necessarily be considered wrong. Both these figures are in fact theoretically acceptable answers.

Such differences between the ground truth and the predicted figures are intriguing, as they hint at contrapuntally or harmonically meaningful information present in Bach’s figures that is not explicit in the four vocal lines. Or, perhaps they are of negligible meaning? It is impossible to know with the information we have now, but future comparisons with models trained not just on Bach but on the figures of many Baroque composers could potentially reveal fascinating insights on compositional style.

We also observed interesting variability in the types of figures Bach used under seemingly similar musical contexts. Three examples are shown in Fig. 4:

- *Accidentals*: in Fig. 4(b), Bach did not label the first “#” at m.2.3, but did label the second one at m. 3.3.
- *Suspensions*: Bach sometimes labelled suspensions (e.g., m. 8.4 of Fig. 4(a)), and sometimes omitted them (e.g., m. 1.4 of BWV 194.06 [not shown]).
- *The same chord*: Bach sometimes labelled a 6/4/2 chord as a 4/2 chord (e.g., m. 2.1 of Fig. 4(b)), and sometimes as a 6/4/2 chord (e.g., m. 10.4 of BWV 13.06 [not shown]).

So, one cannot expect 100% agreement to be achievable, given such variability. Bach, like everyone, was sometimes inconsistent with himself, which imposes an artificial performance ceiling on our models [12]. Also, these types of variabilities can be of great interest to music theorists and musicologists, and offer significant potential for future research.

²³ The first dimension indicates a unison (or collapsed octaves).

²⁴ We used the “DecisionTreeClassifier” function from the “scikit-learn” library, under default settings. This function is an optimized version of CART (Classification And Regression Tree).

²⁵ We used a feedforward network with three hidden layers, each with 300 hidden units. Adaptive Moment Estimation was used as an optimizer, with a binary cross-entropy-based loss function. These hyperparameters were tuned using the validation set.

²⁶ For example, the output vector for m. 4.2.5 of Fig. 2 will be [0,0,0,0,1,0,0,0,0,0,0,0], considering only pitch class “C#” is indicated by the FBA “#” and is four semitones above the bass.

²⁷ The validation set was also used for the selection of the best DNN model using early stopping.

²⁸ Uncertainty values show standard error across cross-validation folds.

Figure 4 consists of two parts, (a) and (b), each showing musical notation and figured bass. Part (a) is for measure 8 of BWV 108.06, and part (b) is for measures 2 and 3 of BWV 145.05. Each part includes four staves of musical notation (treble and bass clefs) and a fifth staff of figured bass. Below the notation, the ground truth and generated figured bass figures are listed, along with a results row indicating correctness.

Measure	Ground truth	Generated figures	Results
8 (a)	6 6 5 9 8	6 6 9 5	✓ ✓ ✓ ✓ ✓ ✓ ✗
2 (b)	2 6 6 5 #	2 6 # 6 #7 # 4 5	✓ ✓ ✗ ✓ ✓ ✗ ✓

Figure 4. An illustration of figured bass generated by our best-performing model for measure 8 of BWV 108.06 *Es ist euch gut, daß ich hingehe*, and measures 2 and 3 of BWV 145.05 *Ich lebe, mein Herze, zu deinem Ergötzen*, which are labelled (a) and (b) here, respectively. We artificially added the fifth (bottom) staff, which collapses all sonorities into one octave so as to more directly reveal the pitch-class content. As discussed in Section 3.3.1, our model predicts interval classes, and the figured bass is generated based on the intervals between the bass note and each predicted interval class. The agreement of each prediction with Bach’s FBAs are shown as well: “✓” means that the generated figured bass exactly matches Bach’s FBAs (the ground truth), “✓” in red means they are considered correct by our evaluation metric that treats musically equivalent figures as equivalent (see Section 3.2.2). An example of the latter can be found at m. 2.1 of (b) where the generated figures can be reduced to “2/4” from “2/4/6” (since the “6” can be omitted, as discussed in Section 3.2.2). “✗” means the generated figures are considered to be errors in our evaluations.

5. CONCLUSION AND FUTURE RESEARCH

This paper presents the Bach Chorales Figured Bass dataset (BCFB), which consists of 139 four-voice Bach Chorales with figured bass annotations encoded in MusicXML, **kern, and MEI. This dataset, and others that can be constructed using methodologies similar to those we propose, offer important potential for use in future computational studies in domains such as music theory, musicology, pedagogy, and performance practice.

This paper further shows how BCFB can be used as the basis for developing and evaluating both rule-based and machine learning models for predicting figured bass; our models achieved classification accuracies of 85.3% and 85.9% on BCFB, respectively. A potential reason the machine learning models did not outperform the rule-based model may be the relatively small size of BCFB.

Such automatically generated figured bass could help performers improvise *basso continuo* accompaniment for the remaining unfigured Bach chorales, or inform the design of pedagogical software for teaching Baroque theory or composition. Of particular interest, figured bass can potentially benefit automatic harmonic analysis research. Existing methods tend to either identify chords directly from the music [13–15], or identify and remove non-chord tones from the score and then generate chord labels from the remaining chord tones [16,17]. A new approach would be to

first generate figured bass automatically from the music and then convert the figures to chord labels; this would allow a chord classifier to take advantage of knowledge implicitly learned from Bach’s ground truth FBAs by a figured bass annotator during its training.

There are several refinements that could potentially improve the quality of the figured bass our approaches generate. The first is to add voice-leading information (how one voice moves horizontally), which may reduce some of the errors discussed in Section 4. The second would be to improve our rule-based model (e.g., by analyzing automatically trained decision trees), which in turn could provide further insights into Bach’s approach to figuring bass, and perhaps provide musicological insight on how his methods changed over time or by the context of performance.

Another potential extension to this research would be to incorporate FBAs from other pieces by Bach, such as his chamber music, or from pieces by other Baroque composers, which are usually figured throughout the Baroque period. Once we have a variety of figured bass datasets for different genres and composers, we may then be able to train models that generalize better. Also, by comparing Bach’s FBAs to FBAs by other composers, we may gain meaningful insights into Bach’s unique compositional style and discover a sense of the degree of stylistic variability with which composers approached figured bass.

6. ACKNOWLEDGEMENTS

We would like to thank the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Fonds de recherche du Québec-Société et culture (FRQSC) for their generous funding. We would also like to acknowledge the contributions of our many collaborators on the Single Interface for Music Score Searching and Analysis (SIMSSA) project, especially Julie Cumming and Samuel Howes.

7. REFERENCES

- [1] P. Williams and D. Ledbetter, “Figured Bass,” *Oxford Music Online*, Oxford University Press, 2001.
- [2] J. S. Bach, A. Dürr, and W. Neumann, *Neue Ausgabe Sämtlicher Werke*. Bärenreiter, 1954–2007.
- [3] C. P. E. Bach, *Essay on the True Art of Playing Keyboard Instruments*, 2 vols. (Berlin, 1753 and 1762). Translated by W. J. Mitchell, New York: W. W. Norton, 1949.
- [4] D. Remeš, *Realizing Thoroughbass Chorales in the Circle of J. S. Bach (2 vols.)*. Wayne Leupold Editions, 2019.
- [5] J. Barthélemy and A. Bonardi, “Figured Bass and Tonality Recognition,” in *Proc. of the 2nd International Society for Music Information Retrieval Conference*, 2001.
- [6] A. Wead and I. Knopke, “A Computer-Based Implementation of Basso Continuo Rules for Figured Bass Realizations,” in *Proc. of International Computer Music Conference*, pp. 188–191, 2007.
- [7] N. Nápoles, G. Vigliensoni, and I. Fujinaga, “Encoding matters,” in *Proc. of the 5th International Conference on Digital Libraries for Musicology*, pp. 69–73, 2018.
- [8] P. Kröger, A. Passos, M. Sampaio, and G. De Cidra, “Rameau: A System for Automatic harmonic Analysis,” in *Proc. of International Computer Music Conference*, pp. 273–281, 2008.
- [9] Y. Ju, N. Condit-Schultz, C. Arthur, and I. Fujinaga, “Non-chord Tone Identification Using Deep Neural Networks,” in *Proc. of the 4th International Workshop on Digital Libraries for Musicology*, pp. 13–16, 2017.
- [10] G. Chew and R. Rastall, “Notation. 4. Mensural Notation from 1500. (viii) Scores; Harmonic and Descriptive Notations,” *Grove Music Online*, 2001.
- [11] F. T. Arnold, *The Art of Accompaniment from a Thorough-bass: As Practised in the XVIIth & XVIIIth Centuries*. Oxford University Press, 1931.
- [12] A. Flexer and T. Lallai, “Can We Increase Inter- and Intra-Rater Agreement in Modeling General Music Similarity?,” in *Proc. of the 20th International Society for Music Information Retrieval Conference*, pp. 494–500, 2019.
- [13] T.-P. Chen and L. Su, “Harmony Transformer: Incorporating Chord Segmentation into Harmony Recognition,” in *Proc. of the 20th International Society for Music Information Retrieval Conference*, pp. 259–267, 2019.
- [14] K. Masada and R. Bunescu, “Chord Recognition in Symbolic Music: A Segmental CRF Model, Segment-Level Features, and Comparative Evaluations on Classical and Popular Music,” *Trans. Int. Soc. Music Inf. Retr.*, Vol. 2, No. 1, pp. 1–13, 2019.
- [15] M. T. Granroth-Wilding, “Harmonic Analysis of Music Using Combinatory Categorical Grammar,” PhD thesis, *The University of Edinburgh*, 2013.
- [16] N. Condit-Schultz, Y. Ju, and I. Fujinaga, “A Flexible Approach to Automated Harmonic Analysis: Multiple Annotations of Chorales by Bach and Prætorius,” in *Proc. of the 19th International Society for Music Information Retrieval Conference*, pp. 66–73, 2018.
- [17] Y. Ju, S. Howes, C. McKay, N. Condit-Schultz, J. Calvo-Zaragoza, and I. Fujinaga, “An Interactive Workflow for Generating Chord Labels for Homorhythmic Music in Symbolic Formats,” in *Proc. of the 20th International Society for Music Information Retrieval Conference*, pp. 862–869, 2019.