Complexity of Musical Patterns

Iris Yuping Ren[1]

[1]University of Warwick,

*To whom correspondence should be addressed; E-mail: yuping.ren@warwick.ac.uk.

Abstract: In this project we analyse music using pitch and time information by employing a network projection. We support our conclusion by topological data analysis(TDA) method. Through the application to a variety of music pieces, we characterize network observations and find non-trival community structures. Using TDA, a modern data-driven idea of persistent homology, we calculate the Betti-number bar-codes of individual musical networks. The results confirm the conclusions of the network-based analysis.

Keywords: Music, Network, TDA

Contents

1 Introduction

The relationships between music and mathematics have generated considerable interest since the Pythagoreans. Music emerges from the the non-trivial interactions of the sounds produced by the musicians. Thus, the interdisciplinary approach of complex systems science, using methods from mathematics, physics, statistics, computer science is a natural choice to produce quantitative analysis of musical pieces. In this paper, we detail the study of western classical music pieces by means of statistical physics and network science, methods widely used in studying social, biological, technological and economic systems (*1*), (*2*), (*3*).

And at the simplest level, musical pieces consist of the repetition of a discrete set of sounds, which differ in pitch and duration. For almost all musical works, especially in classical western music, each musical instrument or singer voice can emit a limited number of such sounds, as there are only 12 different pitches per octave.And the octave range of singers and instruments is limited. This consideration has motivated many relevant works from a pure mathematical perspective, generating many useful results for both musicians and mathematicians. For example, in (*4*), (*5*), and (*6*), group structures are investigated, while in (*7*), (*8*), (*9*), (*10*), and (*11*), geometrical and topological studies are presented.

In the simplest form, music is discretizable in terms of pitch and duration, which are the backbone of music. This allows us to use a network based approach to analyse music. Although there are many studies on music and networks, mostly concern social networks but not music itself (*12*), (*13*), and (*14*). One main exception is (*15*), which identifies frequently used harmonic motifs. However, our approach is different, in that it incorporates duration information and all the melody, rather than just the harmonic progression as analysed in (*15*). Thus, using a network method, not only can we exploit a vast body of existing theoretical results, but also can introduce a new contribution, since the classic network models of friendship, social relations, power grids and emails and phone communication are of a very different nature compared to the music networks we build.

Finally, to attain all the pitch and duration elements, we developed a new package of programs using Python. Our programs can efficiently extract almost all musical information from standard musicxml files.

The remainder of this report is organises as follow: we start with an introduction to the data structure; in section 3, we presnt the network analysis, followed by topological data analysis. Theoretical backgrounds will be explained before applications in each section. For sake of succinctness, only some figures are shown here. The rest of the visualization products can be accessed at https://sites.google.com/site/irisyupingren/documents

2 Data Structure

2.1 Musicxml structure

Musicxml is not a sequential encoding format. For example, backup and forward commands are used to encode single parts with multiple staves or multiple voice. This necessitates care in creating a parsing algorithm to convert musicxml to numerical data structure. The way we did it is first to create a list of elements including pitches, time in the music, duration, and dynamics, for each note in the musicxml sequentially. Second, sort them according to there time element. Only after obtaining a time series of notes, can we construct our sequential network thereafter.

Figure 1: *"Hello world"*

The following codes is a "Hello World!" (Figure 1) program of musicxml. The structure is self-explanatory: different layers of commands forming a tree of information.

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE score-partwise PUBLIC
   "-//Recordare//DTD MusicXML 3.0 Partwise//EN"
   "http://www.musicxml.org/dtds/partwise.dtd">
<score-partwise version="3.0">
  <part-list>
    <score-part id="P1">
      <part-name>Music</part-name>
    </score-part>
  </part-list>
  <part id="P1">
    <measure number="1">
      <attributes>
        <divisions>1</divisions>
        <key>
          <fifths>0</fifths>
        </key>
        <time>
          <beats>4</beats>
          <beat-type>4</beat-type>
        </time>
        <clef>
          <sign>G</sign>
          <line>2</line>
        </clef>
      </attributes>
      <note>
```

```
      <pitch>
        <step>C</step>
        <octave>4</octave>
      </pitch>
      <duration>4</duration>
      <type>whole</type>
    </note>
  </measure>
  </part>
</score-partwise>
```

And to justify our use of musicxml, some discussion on the superiority of musicxml and its future can be found in (*16*), (*17*), (*18*).

2.2 Music pieces

The musical pieces we used for analysis are:

- Beethoven String Quartet No.01(shown in 2) - No.09
- More than 1000 pieces of The Hymns and Carols of Christmas
- Miscellaneous: Nation Songs, Mozart quartet, Bach's Brandenburg Concerto No. 2 in F Major, BWV 1047, Bach's Air on G String, random pop and jazz pieces

The reason for choosing these pieces is twofold. First, a range of pieces with different genre and composers is presumably going to generate either universal results or different views from our analysis. Second, we had a limited resource of free musicxml online. However, Project Gutenberg Sheet Music Project (*19*) is still carrying on producing new pieces in this format.

2.3 Python output

Our program provides a variety of output for analysis purposes. Most of the numerical ones are in text format so as to be easily read to other programs. Some output could be used as debug

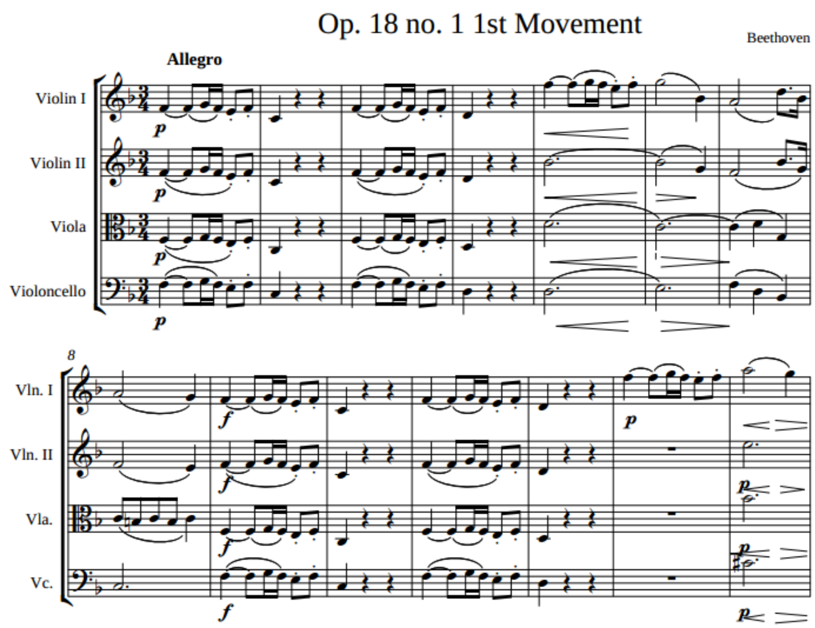

Figure 2: *String Quartet No.1, Op.18 No.1 (Beethoven, Ludwig van)*

information to check whether the process was performing properly, such as ”summary files” and ”backup files”. The fourth elements mentioned in the last subsection are in the main result output: the pitch information in every note is listed first, then the time in second, followed by the duration, and the dynamics.

After the completion of the network analysis, further processing information such as adjacency matrix and the nodes’ names are written into text files for further processing.

3 Network Analysis

A network theory viewpoint can reveal some unseen structures in connected and discretizable systems like music. In this section, we describe the network construction and analysis of music data.

3.1 Adjacency matrices

To build our networks, we consider the adjacency matrices of the music pieces. To map notes to nodes, we use the lexicographical order of pitch and duration. For a larger piece, such as 4 movements of a string quartet, this normally yields a few hundreds nodes; for smaller pieces, with a duration of 2-3 minutes, the nodes are usually less than a hundred. To standardize our approach, we label the nodes with the commonly accepted midi number for pitches(C_4=60, D_4=62, ect.), and assign them a duration in seconds, converted from the logic value in musicxml using tempo information. Rests are denoted by zeros, while grace notes and other musical decorations are ignored. The octaves are not moduled out by 12, as in (*10*), because we are not looking for particular harmonic topological structures, and from the point of the melody, the same note in different octaves are two different notes.

As for the entry in the adjacency matrices, that is, the edges in the network, we will add one whenever one node is followed by the other in the music. For example, if we have a scale C-D-E-F-G-A-B, each note lasting for 1 second, then there will be 7 nodes, the first one is (C,1s), (D, 1s) the second, (E,1s) the third, ect. Also, there will be 6 edges, which are (C,1s)-(D, 1s), (D, 1s)-(E, 1s), ect. When there is a chord in the musical sequence, we connect all the nodes in the chord to the next one, and all the previous nodes to all the nodes in the chord.

In the end, we obtain a network with the number of nodes equal to the number of all pitches multiplied by the number of all durations in the piece. As the music progress, it defines a walk in the network. Also, we consider our network layer per instrument,and build a multilayer network where the inter-layer links are defined by simultaneous playing. The visualizations of the matrices obtained are available in the subsection 5.

3.2 Melody clips

To visually inspect the network structure, we produce ”Melody clips”, obtained by grouping 5 or more notes in to a single ”supernode”. This allows for a qualitative representation of the network that avoid the cluster due to the large number of nodes. We give some examples of this visualization in subsection 5.

3.3 Network construction

Having adjacency matrices from music pieces and melody clips, we can create different graphs. They can be weighted, unweighted, directed, undirected, simple or multigraphs. Different network types yield different visualization and characteristic values.

3.4 Multi-layer network construction

The idea of the multi-layer network is to incorporate all the instruments in one music work at the same time. After building the network of each instrument, as detailed above, we can build the multi-layer network by adding the edges between layer if the nodes are connected in the music. For example, if an instrument is playing a C major scale, and another one, starting at the same time, is playing a D major scale both with notes lasting for 1 second, then the edges will be (C,1s)-(E,1s), (D,1s)-(F,1s), ect. Thus, the construction between different instruments is similar to the way we treated chords in one instrument.

3.5 Visualizations

Adjacency matrices

Figure 3 gives the adjacency matrix of Bach's Brandenburg Concerto No. 2 in F Major(violin solo). Weights indicate how often nodes are connected. Figure 4 is the adjacency matrix of Mozart's Eine kleine Nachtmusik, K.525 and Figure 5 is the visualization of Japan nation song. Already from these simple examples, we can see main differences between pieces. Also, their form of the matrices suggest the existence of a non-trivial community structure.

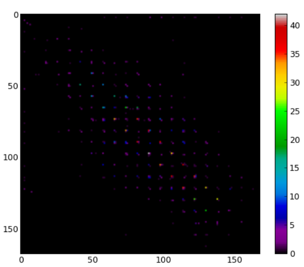

Figure 3: *Weighted adjacency matrix 1*

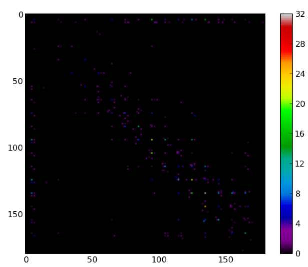

Figure 4: *Weighted adjacency matrix 2*

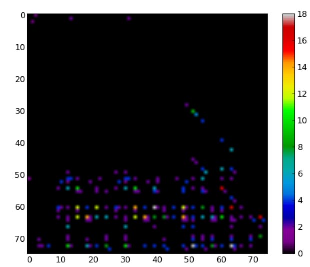

Figure 5: *Weighted adjacency matrix 3*

Multi-layer Networks

In Figure 6, we visualized the notes of Beethoven String Quartet No. 01 in F major Opus 18 on a sphere. It's easy to see the four 'clusters' of instruments already. But to really provide layers, we put the same instrument's notes into the same height as in Figure 8. But it can be a little hard to see what's going on inside, since all the connections just formed a big 'fibre'. In Figure 7, we visualized a smaller piece, to see the structure more clearly.

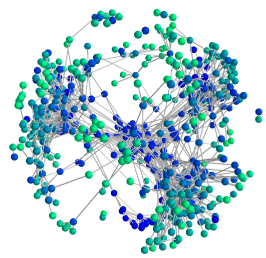

Figure 6: *3d network*

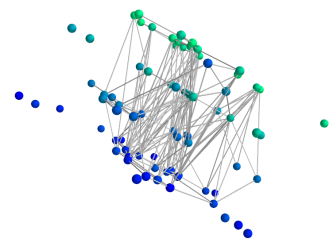

Figure 7: *Connect all layers in a small piece*

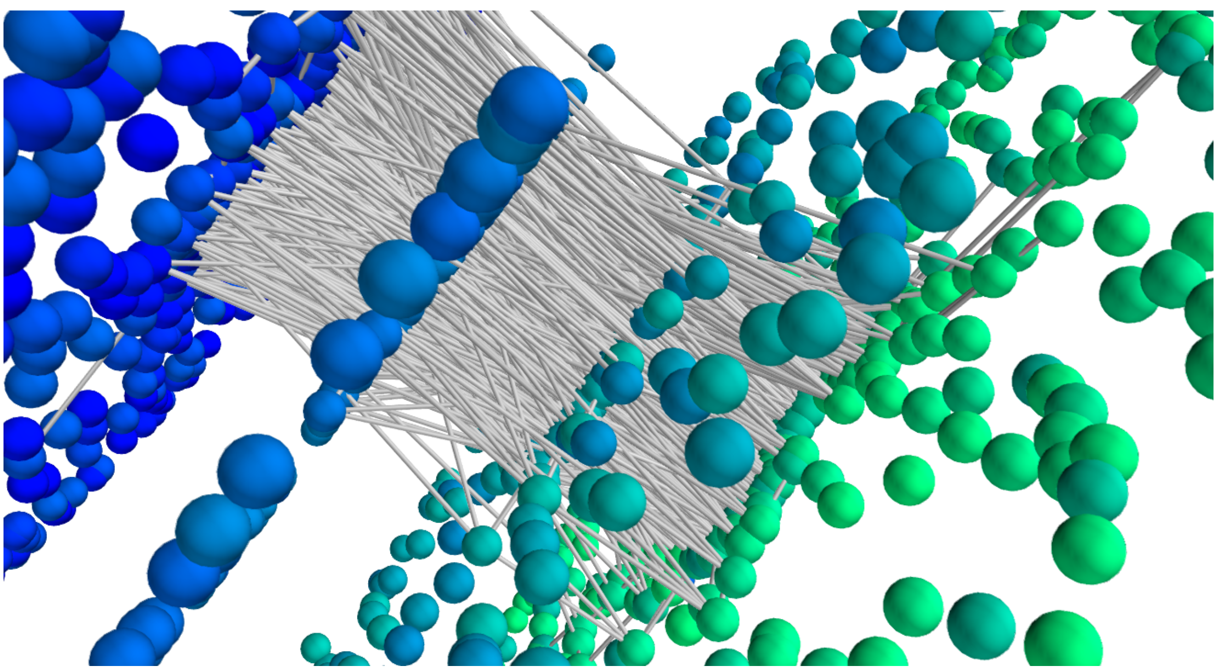

Figure 8: *Connect all layers in a big piece*

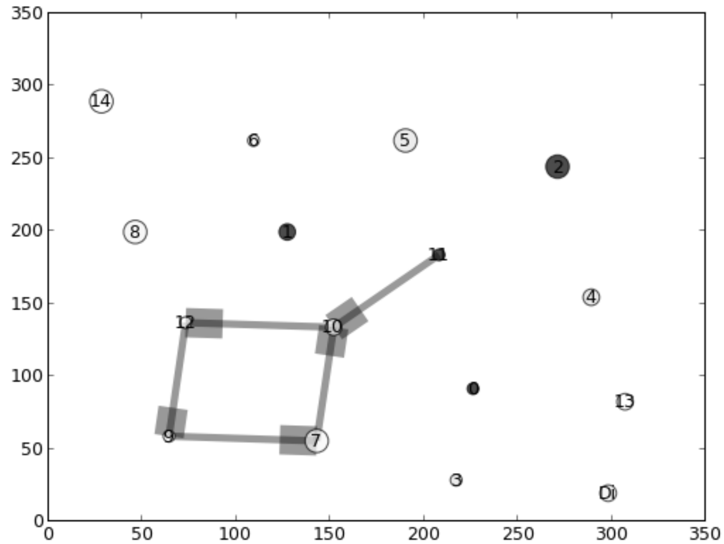

Figure 9: *Melody clip*

Melody clips

In Figure 9, we are showing the directed graph of a melody clip consist of 14 nodes (1 node indicating the type of the graph as programmed). The number denoting the first time such node appears in the network, the color is representing pitch of the note and the size is proportional to the duration.

3.6 Network statistics

3.6.1 Degree

Before discussing the degree distribution, in Figure 10, we can see the effect of our node labelling choice: higher pitch and shorter duration for higher parts, lower pitch and longer duration for lower parts. The peaks exist because not all notes are played and those are played have different frequency.

From the degree distribution such as the one shown in Figure 11, we can see that they are exponential. This suggests that the composer is exploring the whole space of sounds available to the instruments almost without preference. The small fluctuation in them is saying there are some patterns of degree after all. Notably, the exponential degree distribution has been rated as the ”more normal than normal” network degree distribution, some examples include the network of power-grid, email users relationships (*20*) and so on. Therefore, we can say that the pitch and duration recover the most essential parts to music, since we don't see this in mere pitch statistics nor duration alone (*21*).

However, from Figure 15 and Figure 16, we can see that it could be fitted with power-law. Power-law distributions are also known as long-tail distributions, which are deemed to have a good robustness with random-attack. Therefore it provides us with some explanation towards why we still can recognise music even if a player or singer is not mastering the music piece with perfect pitches and durations. This 'robustness' might be able to contribute to discover what is essential for music to be counted as 'a music' as well.

We've also plotted in-degree against out-degree in Figure 14, which is largely linear in normal scale, suggesting our conceptual construction of the network is working. Because in our model, a note in the middle of a piece is certain to have some nodes come into it (the notes are played just before it) and out of it (the notes are played just after it).

Moving on to multi-layer network's results, we can see similar phenomena, but in a more cumulated fashion, as shown in Figure 15.

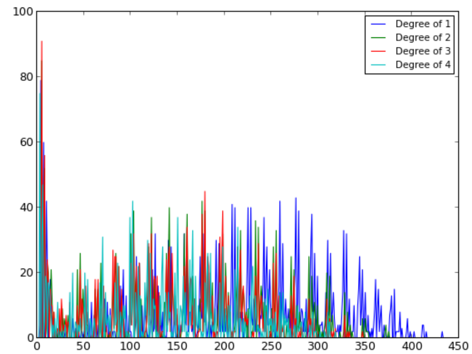

Figure 10: *Degree plot of four instrusments of Beethoven String Quartet No. 03 in D major Opus 18. Number one in legend denotes violin I, two denotes violin II, three denotes viola and four denotes cello. Stay the same whenever there are only four numbers in the legend through out the paper. When there are five numbers, the fifth will indicate the multi-layer network construction. When there are more than five, they will be indicating different piece numbers instead of instruments.*

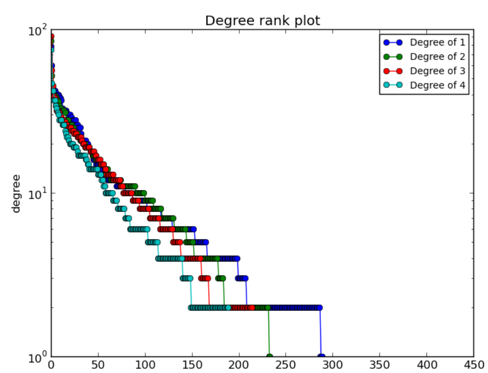

Figure 11: *Degree-rank histogram plot of Beethoven String Quartet No. 03 in D major Opus 18, indicating exponential structure*

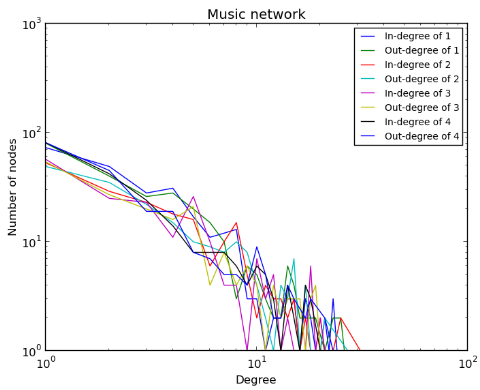

Figure 12: *In/out degree distribution of Beethoven String Quartet No. 05 in A major Opus 18, indicating power-law structure*

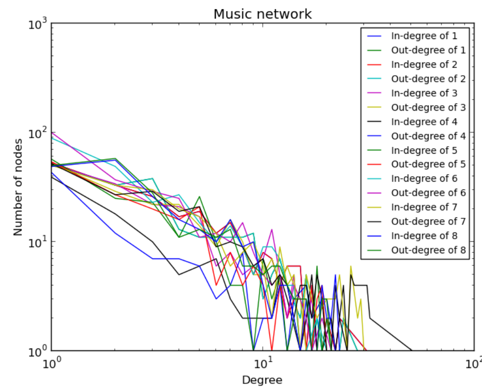

Figure 13: *In/out degree distribution of the viola part of eight pieces of Beethoven String Quartets, indicating power-law structure*

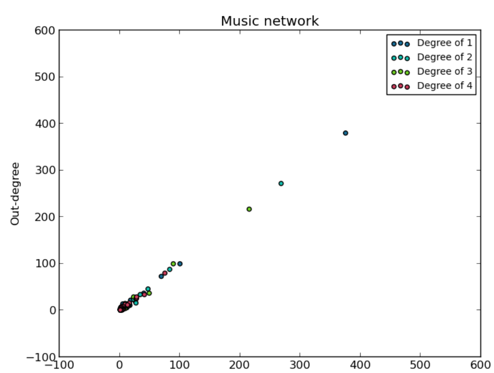

Figure 14: *In- vs. Out-degree. The construction of network is correct.*

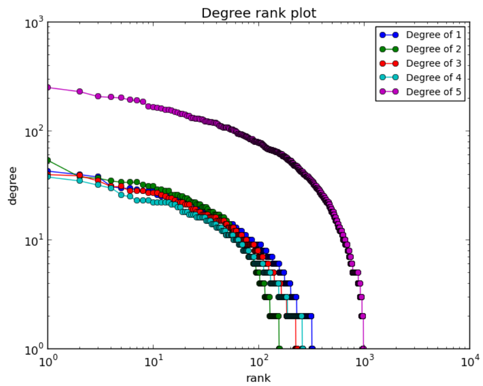

Figure 15: *Multi-layer network degree-rank histogram*

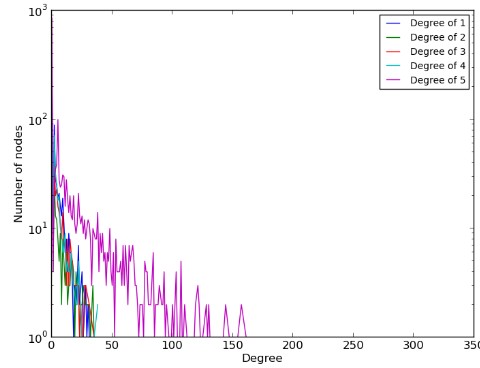

Figure 16: *Multi-layer netowrk degree distribution*

3.6.2 Centralities

Centralities addresses the question ”Which are the most important or central vertices in a network?” (*22*). Using this kind of theoretical tools can help us understand the different roles of different instruments and notes.

There are many kinds of centralities, the three we’ll be looking at are degree centrality, betweenness centrality and closeness centrality. Degree centrality of a node, the simplest kind, is just the degree of the node.

However, it doesn’t have the ability to detect many interesting structures, such as the bridging between groups or how good a position is to receive information, which are better represented by betweenness centrality and closeness centralities:

In an intuitive form, betweenness centrality measures how many pairs of individuals would have to go through a node in order to reach one another. In formula,

$$C_B(i) = \frac{\sum_{j<k} g_{jk}(i)}{g_{jk}}$$

,

where $g_{jk}(i)$ =the number of the shortest paths connecting j and k passing through i

and g_{jk} =total number of shortest paths

And is usually normalised by $\frac{(N-1)(N-2)}{2}$

Closeness centrality measures how close is a node to all other nodes in a graph. In formula,

$$C_C(i) = [\sum_{j=1}^{N} d(i,j)]^{-1}$$

It is usually normalised by $N - 1$

As we can see, the centralities have some similarities and a common goal, but they have every different meaning graphically. To extract information about the music structure, we study the relation between those centralities. The results of our analysis are shown in Figure 17 and Figure 18. We can clearly see the power-law trend between both degree centrality vs. closeness centrality and betweenness cetrality vs. closeness centrality.

Some implications of those phenomena are the following: the power-law between degree centrality and closeness centrality is most probably caused by the prominent nodes having relatively low closeness, which suggests that high-degree nodes are assortative star-like hubs. And after running a weighted rich-club effect test on several adjacency matrices, we found the ϕ coefficients (definition in (*23*)) are indeed larger than one, which implies the original network has a positive weighted rich-club effect (*23*). The second power-law is also implying the nodes at the prominence stand are star-like hubs amongst them, which is verifying our results of stability from degree distributions, and giving clues about certain patterns in the music exist. Finally, what should be pointed out is that this is very natural because of out construction of the network: in a big piece with 4000+ notes and about a few hundreds nodes where connections are generally not sparse, we should indeed expect such a phenomenon given that some notes are very important in music. To sum up, we can complete our conclusions about the robustness due to degree distribution: there are some extremely important notes. If an instrumentalist keeps missing them, the music would break down anyway. In similar fashion to targeted attack, this indicates that the music network is weak against the remotion of particular nodes.

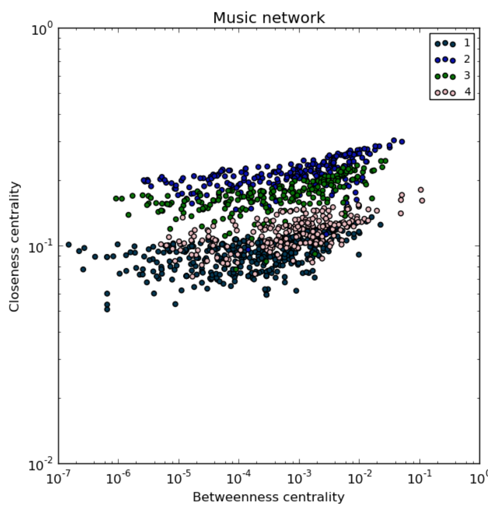

Figure 17: *Betweenness vs. Closeness centralities*

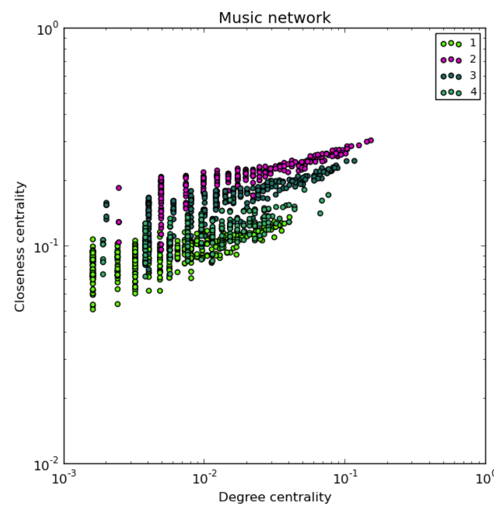

Figure 18: *Degree vs. Closeness centralities*

3.6.3 Clustering coefficients

Local clustering coefficients are defined as the proportion of links between the vertices within its neighbourhood against all the links that could exist between them. And it is a measure of the degree to which nodes in a graph tend to cluster together. The distribution of clustering coefficients for the music networks we used is shown in Figure 19.

Besides the scatter plot, the clustering coefficients are also plotted against degree, which is shown in Figure 20. We can see an obvious trend in the high degrees, where it shows the nodes tend to cluster less when the degree is high. The conclusion can be drawn from this is that the hierarchical structures are not very likely, since one would normally have a power-law clustering against degree plot. And it is indeed not natural to have hierarchical structures in pitches durations.

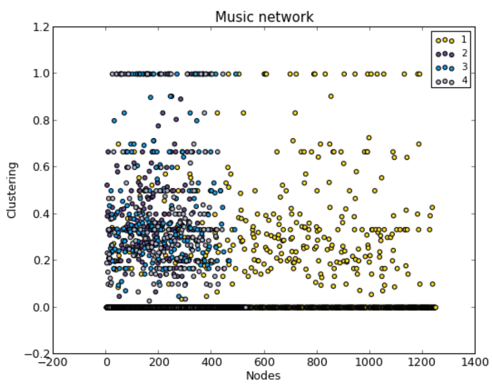

Figure 19: *String Quartet No. 07 in F major Opus 59*

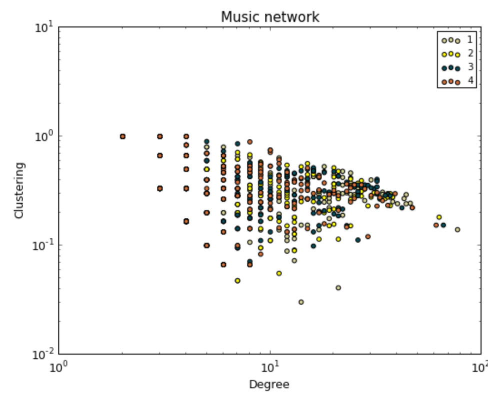

Figure 20: *String Quartet No. 05 in A major Opus 18*

3.6.4 Assortativity

The property of assortativity measures the extent to which nodes tend to link to nodes of similar degree. One way of analyzing assortativity is by plotting all nodes of a network by their degree and the average degree of their neighbors (*24*), as shown in Figure 21. The plot is showing a trend in the nodes with high degrees and less so for degree with low degrees. After comparing with other same plots available online (*24*), we found that it's similar to networks of Brightkite, Facebook, and Route views. This assortativity can also be infered from the rich-club effect we talked about earlier.

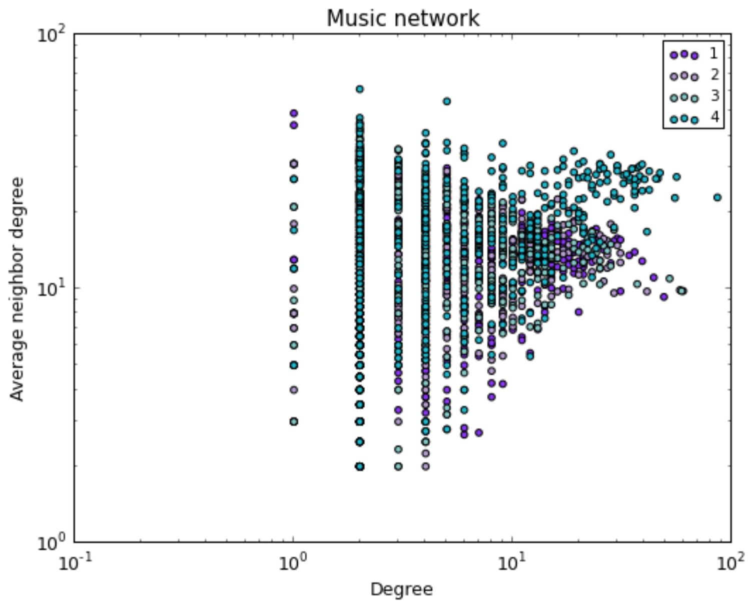

Figure 21: *String Quartet No. 01 in F major Opus 18, Degree related plot which reflects assortativity*

3.6.5 Connectivity

The connectivity we used is the average degree connectivity, which is the average nearest neighbor degree of nodes with degree k. For a weighted graph like our construction, the analogous measure is defined as below:

$$k^{w}_{nn,i} = \frac{1}{s_i} \sum_{j \in N(i)} w_{ij} kj$$

s_i =wighted degree of node i

w_{ij} =the weight of the edge that links i and j

$N(i)$ are the neighbors of node i

The result in Figure 22 is showing a loose strucure slightly leaning upwards. The implication should be we have a very diverse system in terms of the connection between different degrees in music.

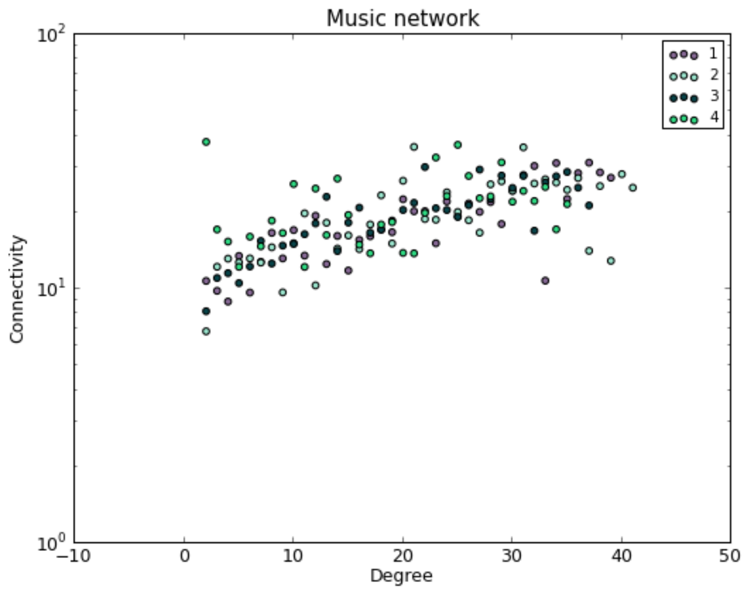

Figure 22: *String Quartet No. 09 in C major Opus 59, connectivity for each degree*

3.6.6 Community detection

There are many methods for community detection, we used the Louvain method described in (*25*). The results have been plot with nodes and edges in each community and in every community. The color is indicating different communities, as shown in Figure 23. We usually have around ten different communities in the Beethoven's string quartets.

We also plotted the summary statistics for each community, as displayed in Figure 24. A natural result could be thought of before the plotting is that we can only get some communities with distinct difference in the range of pitches and durations. We can see that there are some communities do have those differences. Nevertheless, we also have communities having similiar ranges in both pitches and durations. With a closer examination in the notes we have in different communities, we can see that they are actually some structures that are frequently used in the music. Although we can't fully recover the temporal information of the music, each community has the potential to create more diversed music in this aspect.

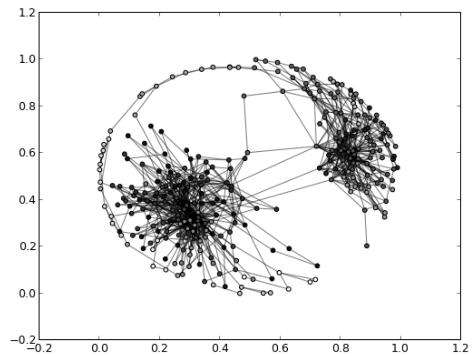

Figure 23: *Community detection on String Quartet No. 09 in C major Opus 59 Cello*

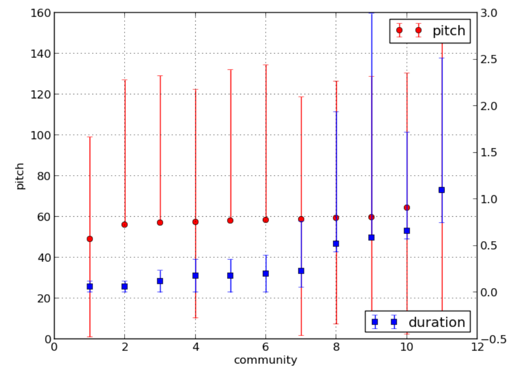

Figure 24: *Summary statistics for community detection String Quartet No. 03 in D major Opus 18 Viola. Errorbar ceiling stands for the max value and the floor is the minium value.*

Of course, more characteristic values could be calculated, such as modularity and other centrality measures, and they are indeed part of the plan for the next step of research.

4 Topological Data Analysis(TDA)

From (*20*), we know that complex networks with distinct degree distributions exhibit distinct persistent topological features. Since the degree distributions in the last section are quite similar in large picture for almost all the music, while the music themselves are very different from each other, it's possible that by using this method we can uncovered some structures that cannot be seen by merely observing degree distributions. But due to time constraints, we just calculated the bar-code for a few music pieces and haven't found any profound results yet. The bar-code is just shown here for future research.

4.1 Topology

This subsection will give a very short intuitive introduction on the mathematics related to the bar-code results. For more details, see (*26*), (*27*).

Simplicial complexes are simple objects, such as 0-simplices can be thought of as vertices, 1-simplices as edges, 2-simplices as triangles, and so on. Homology is described by $Betti_k$ numbers, which roughly speaking is the number of k dimensional holes. For example, $Betti_0$ is the number of connected components. Mathematically speaking, it's the rank of homology group , which is defined as:

$$H_k := ker\partial_k / im\partial_{k+1}$$

where $\partial_k : C_k \longrightarrow C_{k-1}$ is the boundary operator

in which C_k is group consists of the set of all k-chains and the operation of addition

k-chain is a sequence of simplices $\sigma, \sigma_1, \sigma_2, \cdots, \rho$ s.t. any 2 consecutive ones share a k-face

Another important concept is filtration, which acts on simplicial complexes and gives a collection of blowing-like new complexes. There's a parameter whose going to control the level of the blowing, which is called filtration time. This is also a commonly used concept in algebra and measure theory. In a network background, it can be thought of as the time we add a edge, because the effect of blowing-up is getting simplicial complexes connected to each other. Also due to the adding of the edges, the network is an evolving network now. The bar-code results which will be shown in the next section is a representation of this evolution, with the x axis being the filtration time, and the bar-intervals being the filtered simplicial complex at times when the addition of simplices takes place (*20*). And a Betti interval with long filtration interval can be interesting because it is exhibiting some long-existing topological structure.

4.2 Bar-code results

The results of bar-code calculation is used to reveal the birth and death of the algebraic topological invariants as the network evolves in time. The x-axis is the lifetimes of various stages of filtration, and the y-axis is the arbitrarily ordered homology generators H_k. So, the barcodes do not provide delicate topological structure but just the essential ones.

The bar-code results in Figure 25 are constructed by the notes in Bach's Brandenburg Concerto No. 2 in F Major, BWV 1047, where the nodes are added as the lexicographical order mentioned in network construction, and the edges are added if the entry in the adjacency matrix is not zero, in an uniform-speed fashion, as the nodes. Result in Figure 26 uses the same music piece, the same way to add nodes, but the filtration time for edges is the value stored in the adjacency matrix.

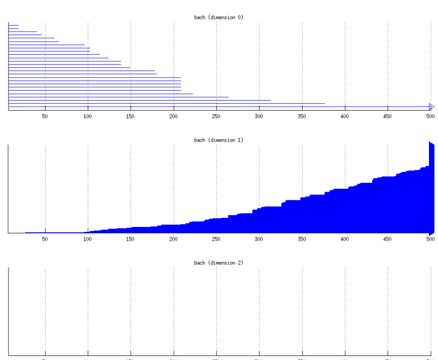

Figure 25: *Barcode result1*

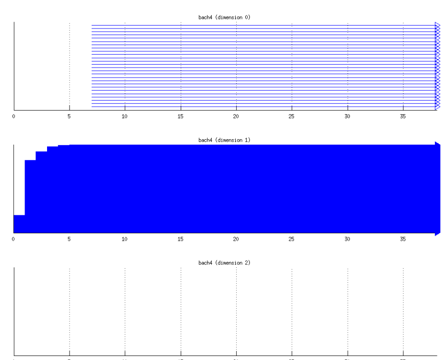

Figure 26: *Barcode result2*

The barcode results in Figure 27 and Figure 28 are the parallel results as above using Beethoven's String Quartet No. 01 in F major Opus 18.

It's clear that the filtration steps of both of them are limited to dimension 3 because of the sparsity of the network. And we can see distinctively that the topological features of this network is very different from a random network, whose bar-code result can be found in (*20*). And the fact that the existence of the persistent homology groups is reflecting the robustness of the network tells us we are having a very robust network as deduced in the Network Analysis section.

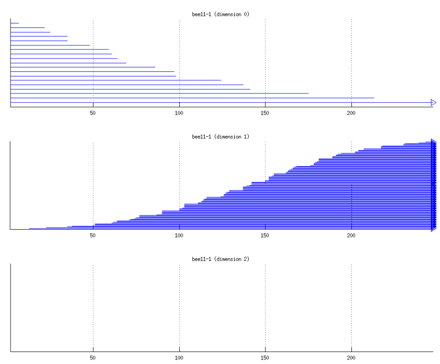

Figure 27: *Barcode result1*

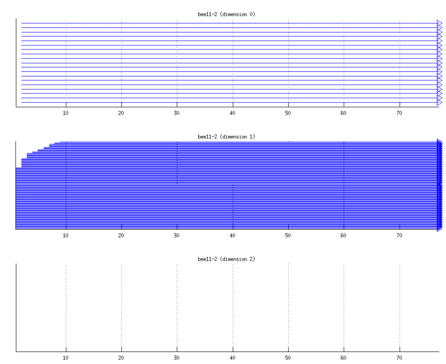

Figure 28: *Barcode result2*

4.3 Other comparison

In (*20*) and (*10*), we found the bar-code of an e-mail users' network and the result of a folk music obtained using different algebraic topological methods. And in both of those results, we see the high-dimensional noise phenomena in their bar-code.

We can therefore conjecture that the exponentiality of the systems resulted in this topological phenomenon. Moreover, in the e-mail users' case, the presence of the high dimensional homology groups is coming from the internal organization of the e-mail network, i.e., the community structure. This lends support to our adjacency matrices' visually-drawn conclusion on the community structure.

Turning to the relatively short filtration intervals of those high dimensional bar-codes, we can say that there are certain groups of notes which repeat appearing and disappearing during the evolution of the network. So it's also supportive to the conclusion we've drawn in the Network Analysis section.

5 Conclusion and Outlook

In conclusion, we have looked at a particular type of music network and studied its features using many different approaches as a case study, through which exponential and power-law structures have been found, which certifies the robustness of music network. The robustness against random attacks provide us clue as to why we can still recognize the music even with off-pitch and off-beat playing, which is essential for understanding when a music should be counted as 'the music'.

One can say it's a over simplified model for a musical continuum, and aspects in the performance of music are ignored, nevertheless, a wealth of phenomenon can arise from just these pitch and duration elements as we can see in this paper. In other works, researchers have considered acoustic perspectives, for instance by means of spectrograms (*28*), or considering music performance (*29*). Furthermore, another area of interest is algorithmic composition, explored via different means, such as genetic algorithms, cellular automata, chaos and self-similarity theory, articial neural networks (*30*), (*31*), (*32*).

In a forthcoming paper, we shall analyse more musical works, more quantities as mentioned in the second section, and make more in-depth studies in time series and TDA area. And there are many other possible future works, including but not limited to:

- Examine upon more implications of the network statistics, barcode results, and verify them by testing on more pieces of music using ensemble averaging.
- Incorporate dynamics information of the music
- Extend to the area of algorithmic composition
- Carry on the research to the field of machine learning and information geometry.

I believe I'll keep following on this area of study. Many thanks for reading!

References and Notes

1. Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

2. Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

3. Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308, 2006.

4. Alissa S Crans, Thomas M Fiore, and Ramon Satyendra. Musical actions of dihedral groups. *The American Mathematical Monthly*, pages 479–495, 2009.

5. Thomas M Fiore, Thomas Noll, and Ramon Satyendra. Incorporating voice permutations into the theory of neo-riemannian groups and lewinian duality. In *Mathematics and Computation in Music*, pages 100–114. Springer, 2013.

6. Thomas Noll. Musical intervals and special linear transformations. *Journal of Mathematics and Music*, 1(2):121–137, 2007.

7. Dmitri Tymoczko. *A geometry of music: harmony and counterpoint in the extended common practice*. Oxford University Press, 2011.

8. Kenneth Jinghwa Hsü and Andreas J Hsü. Fractal geometry of music. *Proceedings of the National Academy of Sciences*, 87(3):938–941, 1990.

9. Dmitri Tymoczko. The geometry of musical chords. *Science*, 313(5783):72–74, 2006.

10. R. Budney and W. Sethares. Topology of Musical Data. *ArXiv e-prints*, July 2013.

11. M. G. Bergomi and A. Portaluri. Modes in modern music from a topological viewpoint. *ArXiv e-prints*, September 2013.

12. Nicole B Ellison, Charles Steinfield, and Cliff Lampe. The benefits of facebook friends: social capital and college students use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4):1143–1168, 2007.

13. Nancy K Baym and Andrew Ledbetter. Tunes that bind? predicting friendship strength in a music-based social network. *Information, Communication & Society*, 12(3):408–427, 2009.

14. Nicholas J Bryan and Ge Wang. Musical influence network analysis and rank of sample-based music. In *ISMIR*, pages 329–334, 2011.

15. Shalev Itzkovitz, Ron Milo, Nadav Kashtan, Reuven Levitt, Amir Lahav, and Uri Alon. Recurring harmonic walks and network motifs in western music. *Advances in Complex Systems*, 9(01n02):121–132, 2006.

16. Christopher J. Russell. Midi vs. musicxml. http://techinmusiced.wordpress.com/2012/05/26/midi-vs-musicxml/, 2012.

17. Constantine Zavras. Xml music notation encoding standards:musicxml and mei. http://web.simmons.edu/ zavras/XML/Final$_{P}resentation.ppt$, 2012.

18. Craig Stuart Sapp. Musicxml. https://www.google.co.uk/url?sa=trct=jq=esrc=ssource=webcd=1ved=0CCsQFj 2005.

19. Project gutenberg sheet music project. http://www.gutenberg.org/wiki/Gutenberg:The$_{S}heet_{M}usic_{P}roject/$.

20. Danijela Horak, Slobodan Maletić, and Milan Rajković. Persistent homology of complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(03):P03034, 2009.

21. Tuomas Eerola and Petri Toiviainen. Midi toolbox: Matlab tools for music research. *University of Jyväskylä, Jyväskylä, Finland*, page 14, 2004.

22. Mark Newman. *Networks: an introduction*. Oxford University Press, 2010.

23. Tore Opsahl, Vittoria Colizza, Pietro Panzarasa, and José J Ramasco. Prominence and control: The weighted rich-club effect. *Physical review letters*, 101(16):168702, 2008.

24. Konect project. http://konect.uni-koblenz.de/plots/assortativity.

25. Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

26. Allen Hatcher. *Algebraic topology*. 2002.

27. Mark Anthony Armstrong. *Basic topology*. Springer, 1983.

28. James S Walker and Gary W Don. *Mathematics and Music: Composition, Perception, and Performance*. CRC Press, 2013.

29. Terry Allen and Camille Goudeseune. Topological considerations for tuning and fingering stringed instruments. *arXiv preprint arXiv:1105.1383*, 2011.

30. J. D. Fernandez and F. Vico. AI Methods in Algorithmic Composition: A Comprehensive Survey. *ArXiv e-prints*, February 2014.

31. Gerhard Nierhaus. *Algorithmic composition: paradigms of automated music generation*. Springer, 2009.

32. Jeremy Leach and John Fitch. Nature, music, and algorithmic composition. *Computer Music Journal*, pages 23–33, 1995.