CLASSICAL MUSIC CLUSTERING BASED ON ACOUSTIC FEATURES

Syed Arefinul Haque
Xindi Wang

PROBLEM STATEMENT

- Given a set of music with its acoustic features, find clusters in them
- Gain insights for music composition
- Help content-based music recommendation

BACKGROUND

- Music recommendation is mostly done by collaborative filtering
- Music features and distance for clustering
  - Spectrogram of the music piece (Fourier Transform)
  - Normalized compression distance [R. Cilibrasi P.M.B. Vitanyi,2004]

Isaacson, Eric J. "What You See Is What You Get: on Visualizing Music." *ISMIR*. 2005.
Cilibrasi, Rudi, and Paul MB Vitányi. "Clustering by compression." IEEE Transactions on Information theory 51.4 (2005): 1523-1545.

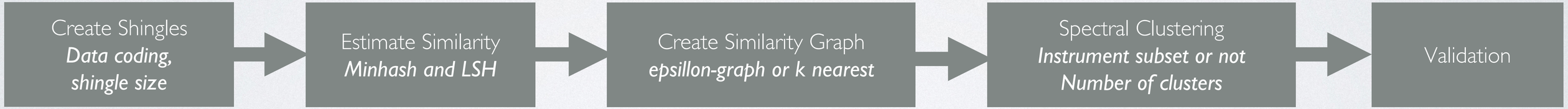
METHODS

- Shingles
- Chord Trajectory Matrix
- Spectral Clustering
- Validation: TF-IDF on meta-data, Silhouette coefficient

DATA AND EXPERIMENT SETUP

- MusicNet data, including 330 classical music pieces
- Extract notes sequence from Interval Tree
- Several ways of data coding:
  - 128 MIDI notes coding, 12 pitches coding, change coding and relative note coding

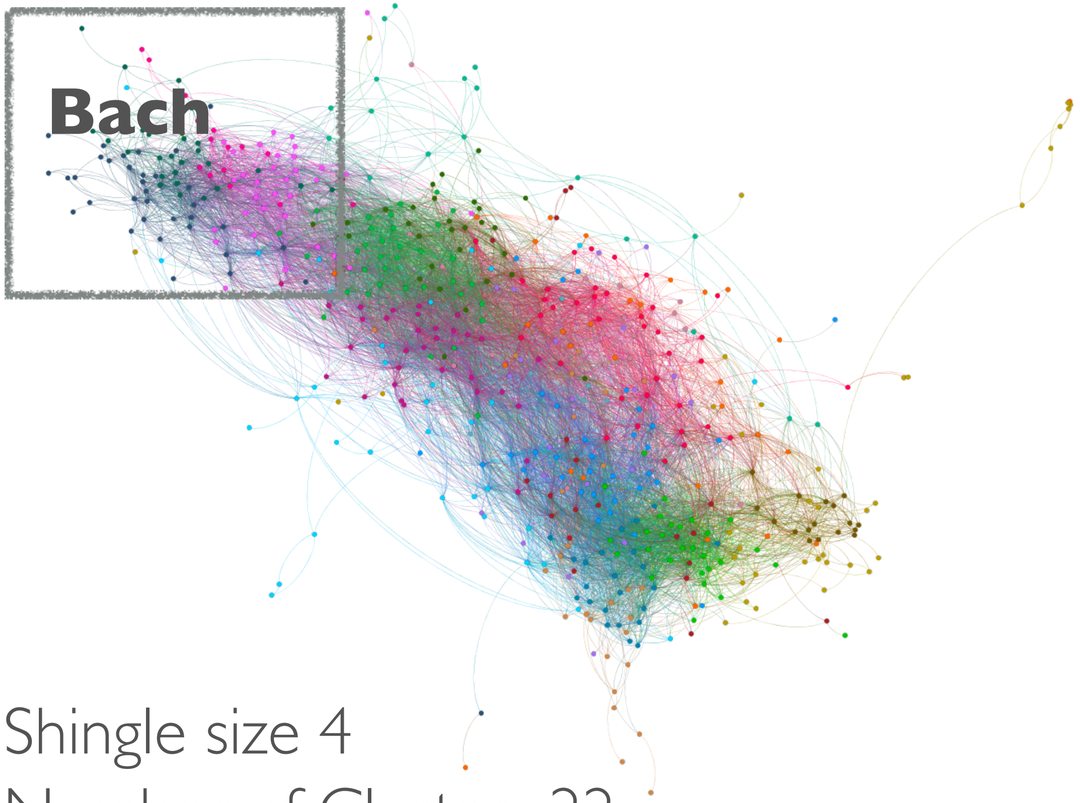
Shingles Flow Chart

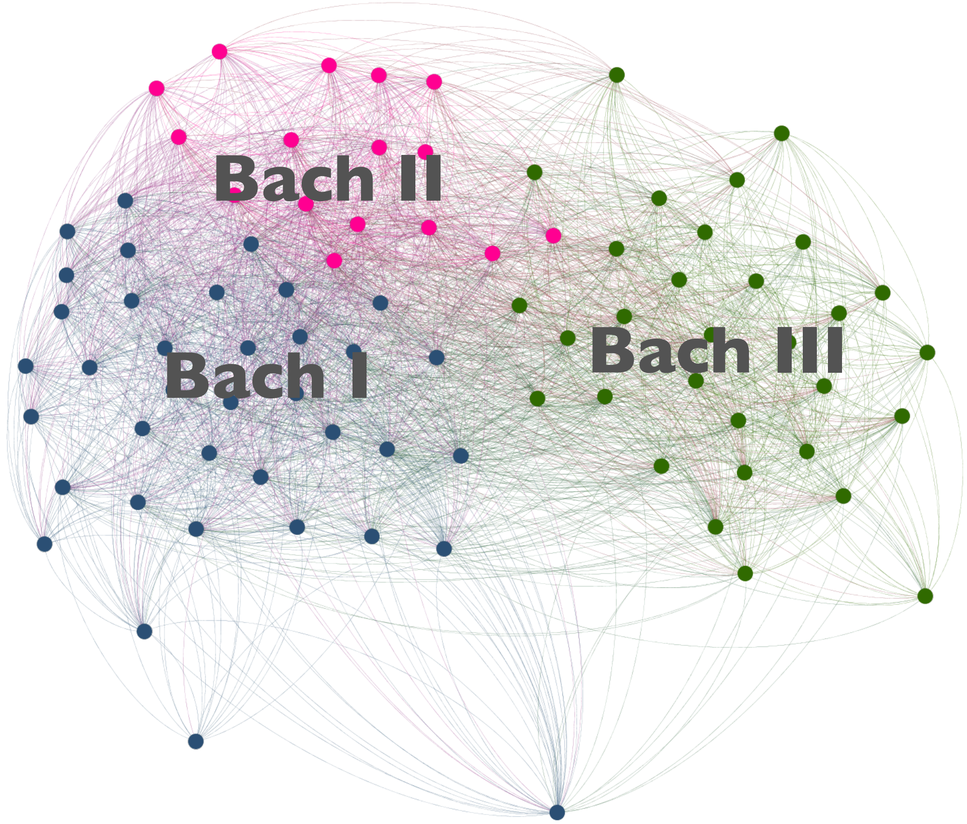


Chord Trajectory Flow Chart

METHOD 1 - SHINGLES

Cluster could reflect certain work of a composer



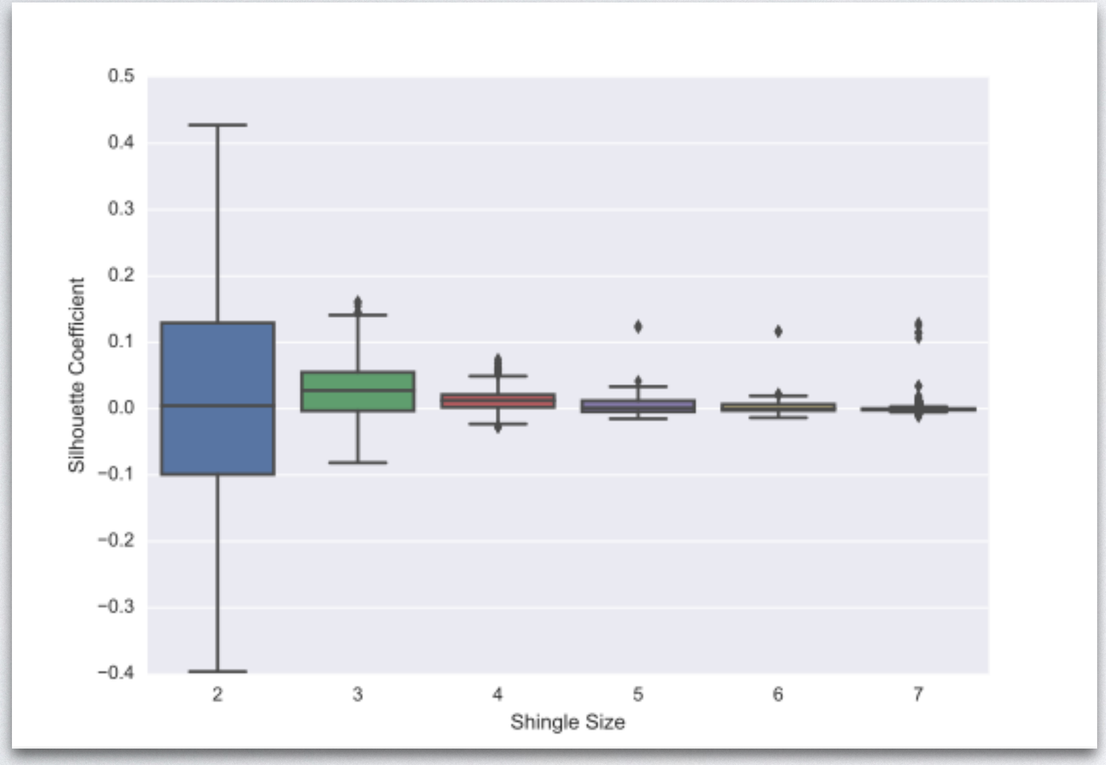


Shingle size 4
Number of Clusters 22
All instruments
Pitch Coding

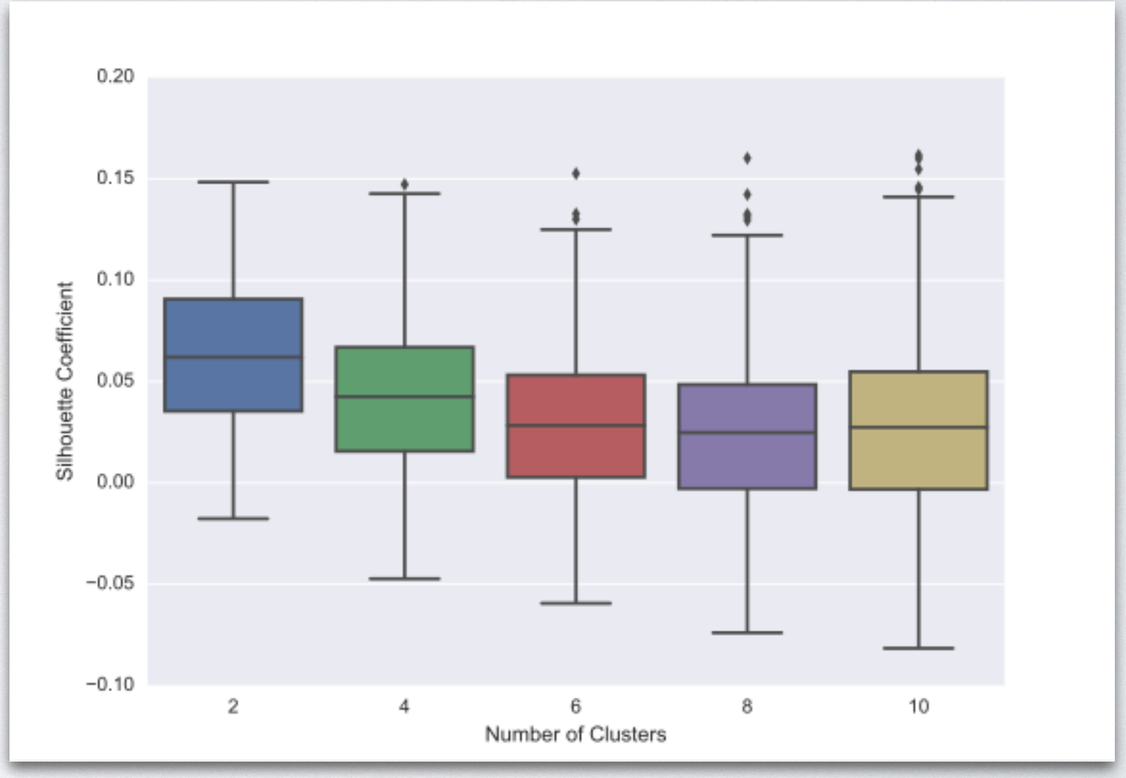
Bach I: prelude-fugue-Bach
Bach II: cello-suite-danza
Bach III: harpsichord-Bach-violin

METHOD 1 - SHINGLES

Silhouette Coefficient Analysis



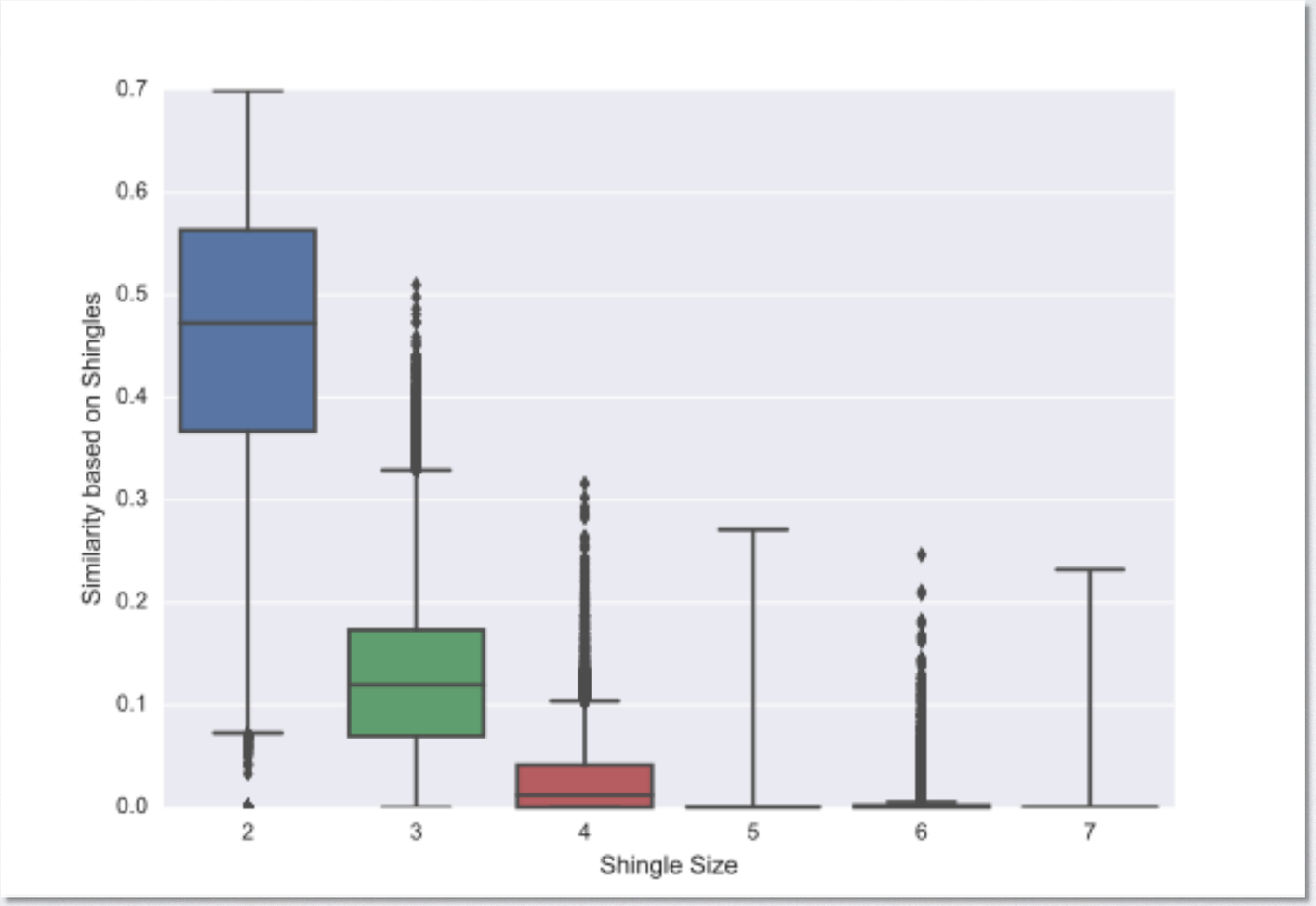
Fix # of clusters: 10



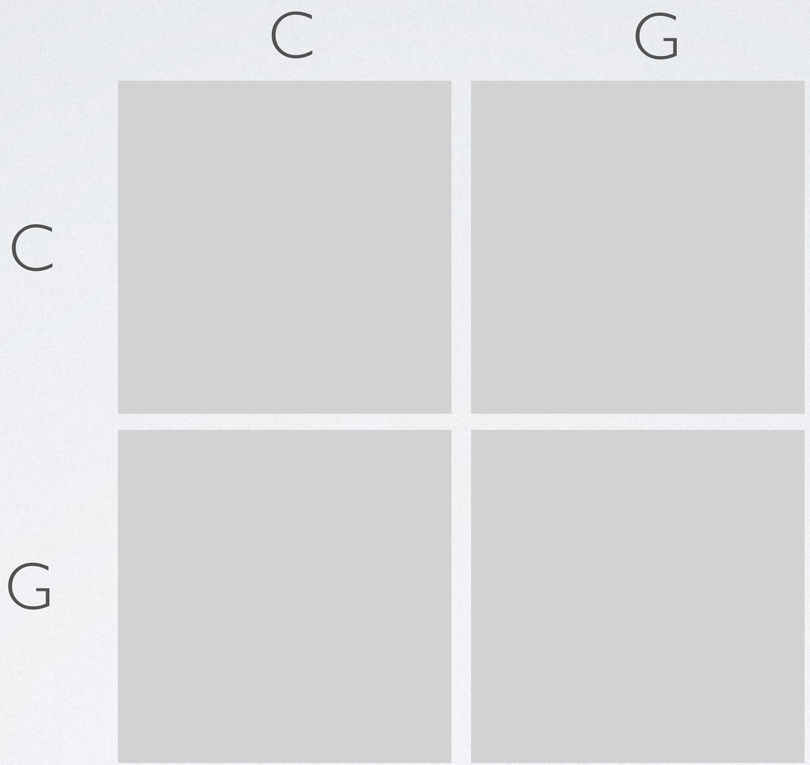
Fix Shingle Size: 3

METHOD 1 - SHINGLES

Silhouette Coefficient Analysis



CHORD TRAJECTORY MATRIX



C → G → G → C → G

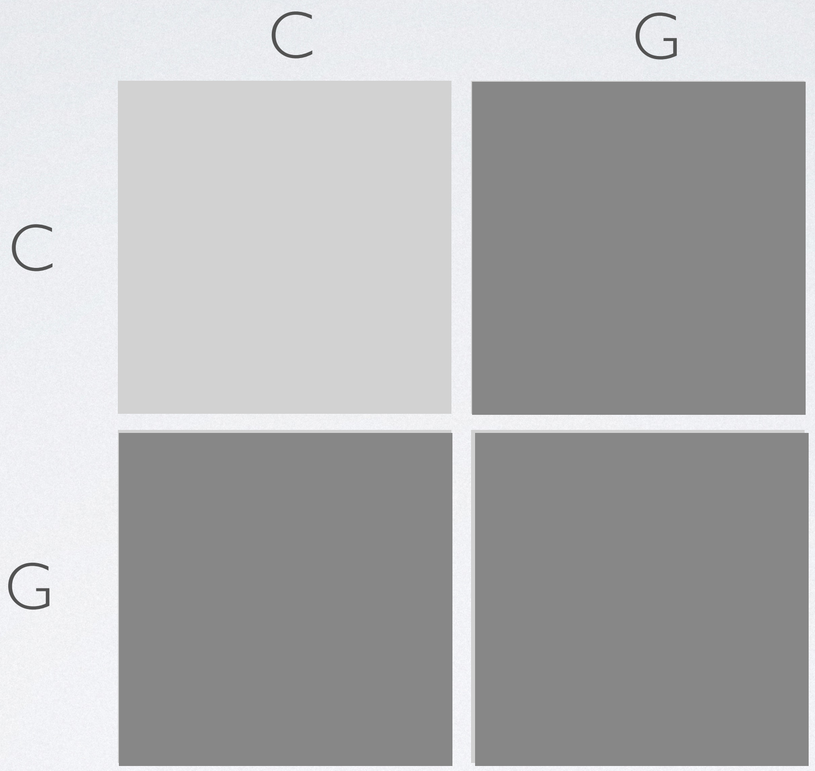
CHORD TRAJECTORY MATRIX

C → G → G → C → G

CHORD TRAJECTORY MATRIX

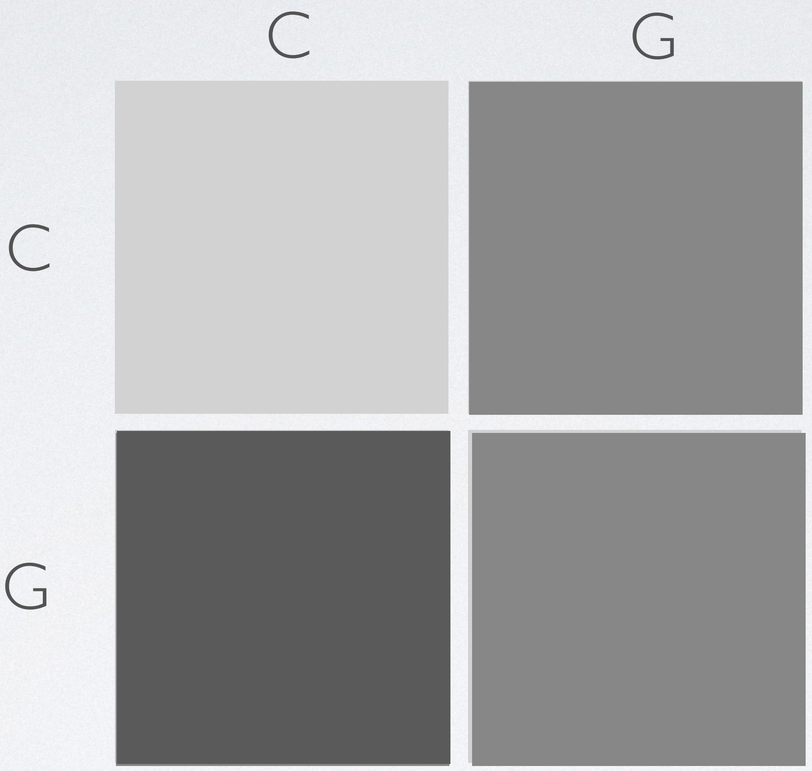
C → G → G → C → G

CHORD TRAJECTORY MATRIX



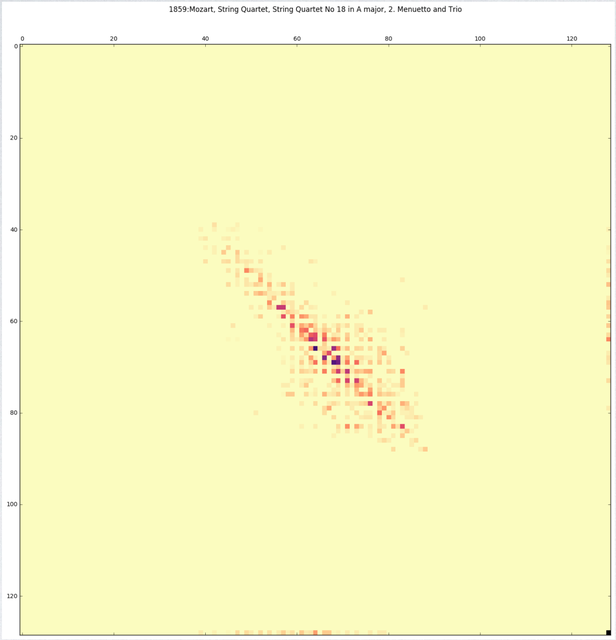
C → G → G → C → G

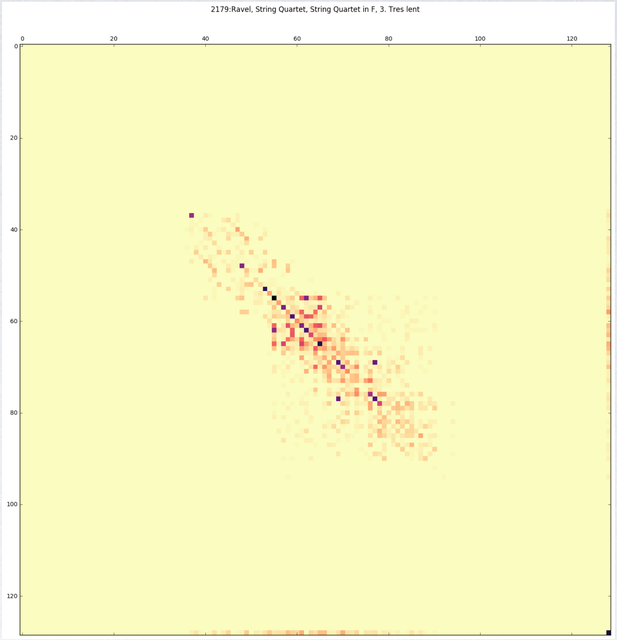
CHORD TRAJECTORY MATRIX

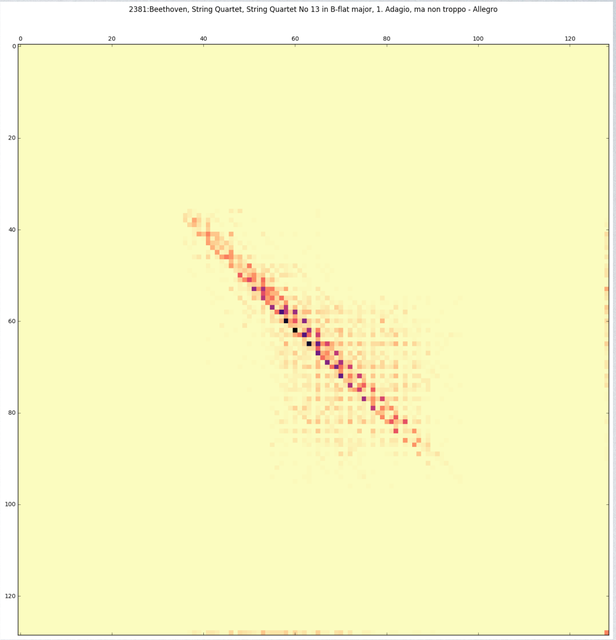


C → G → G → C → G

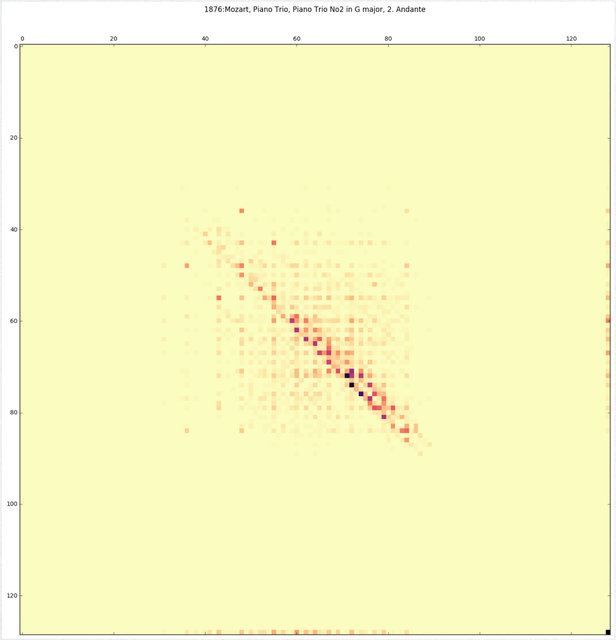
String Quartet

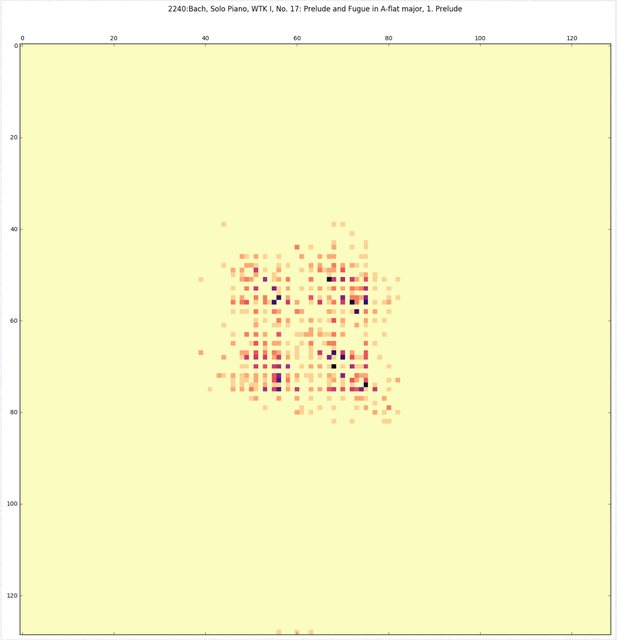


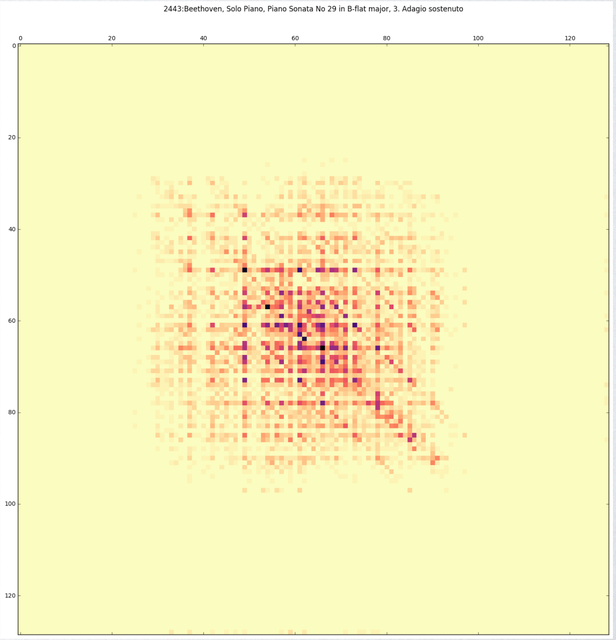




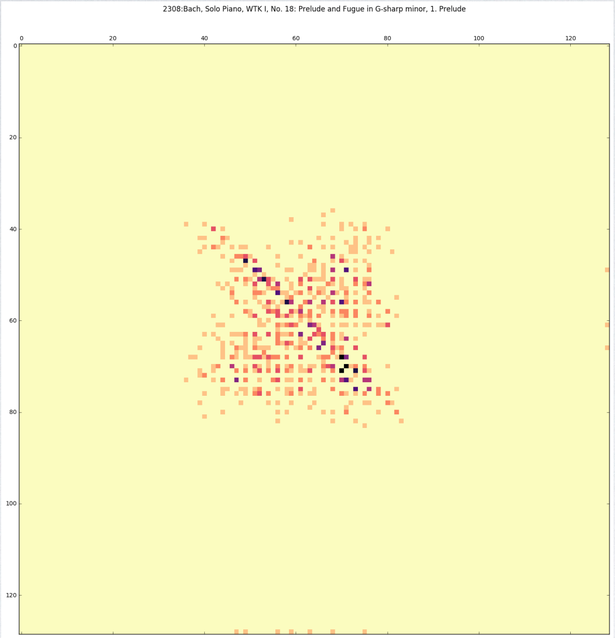
Piano

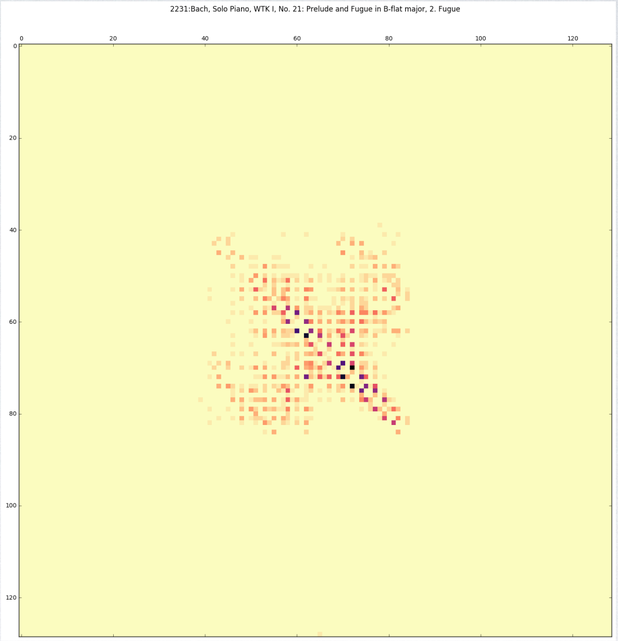


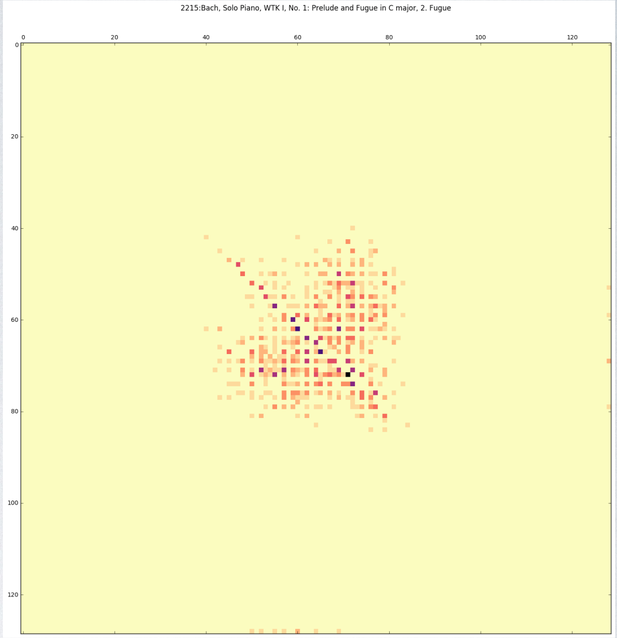




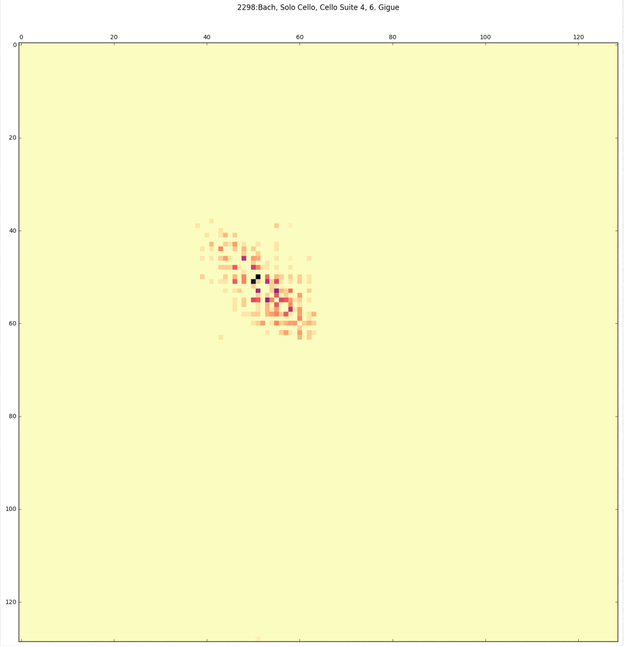
Bach Fugue

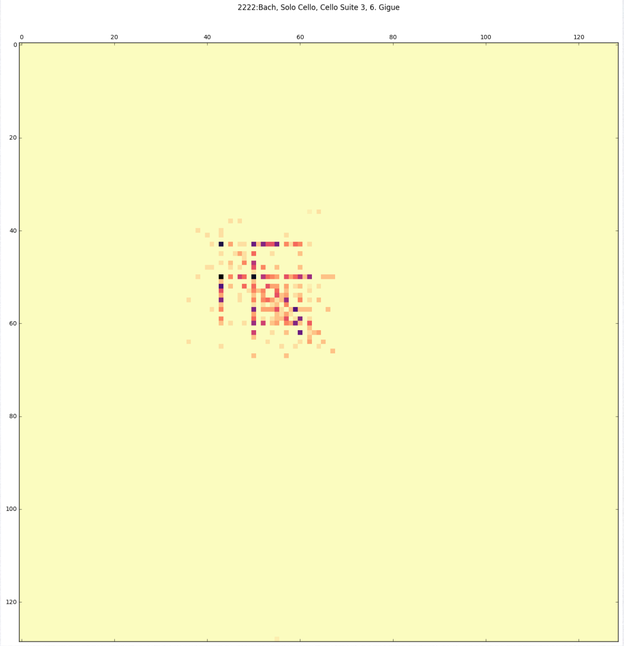


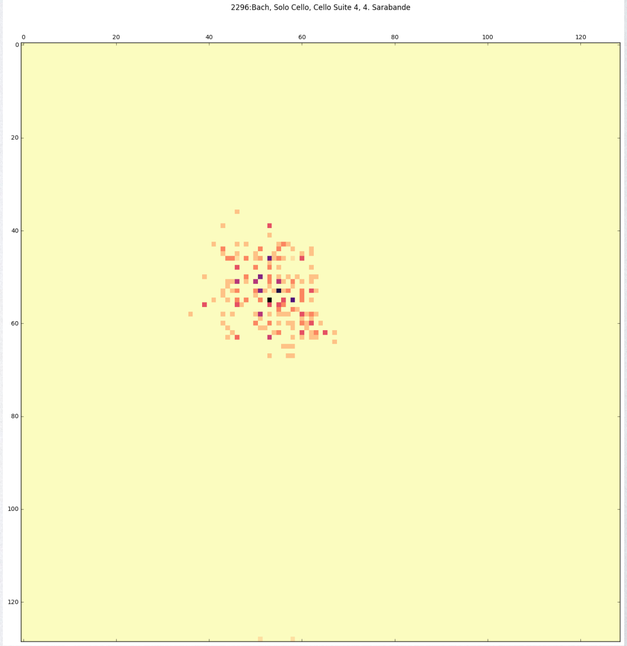




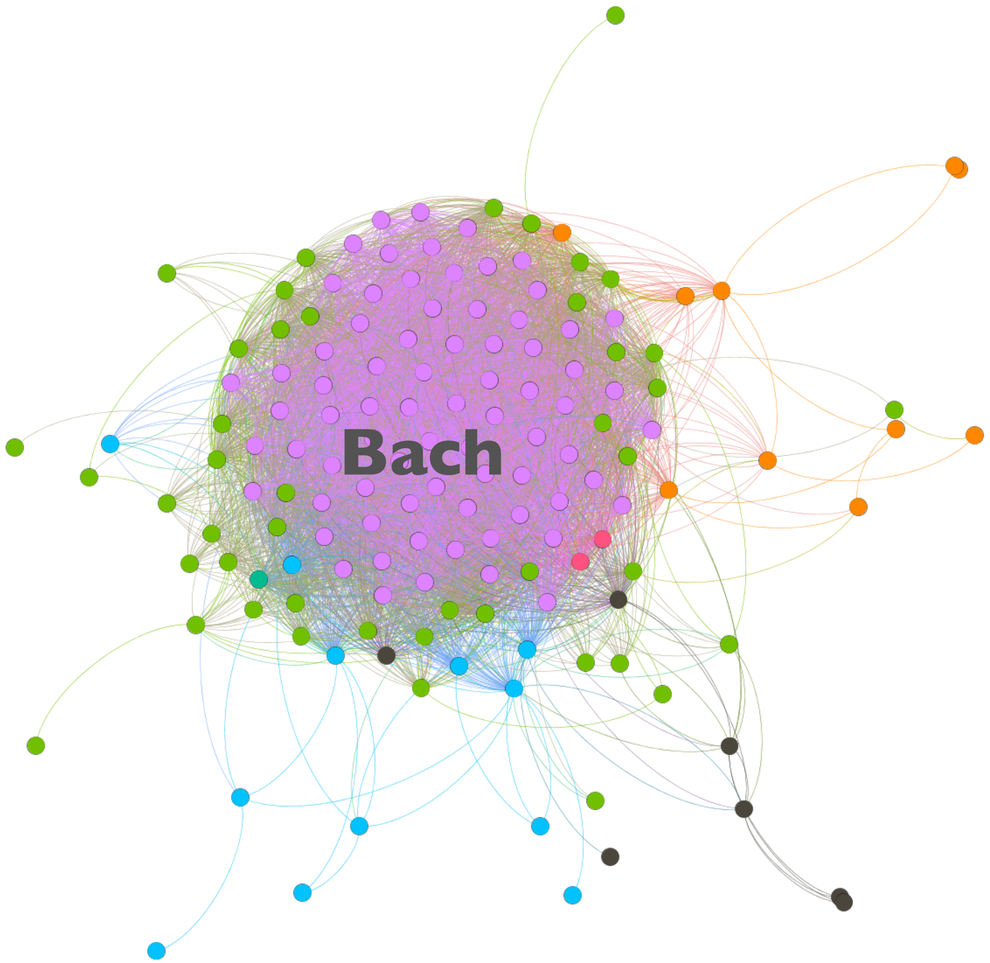
Bach Cello



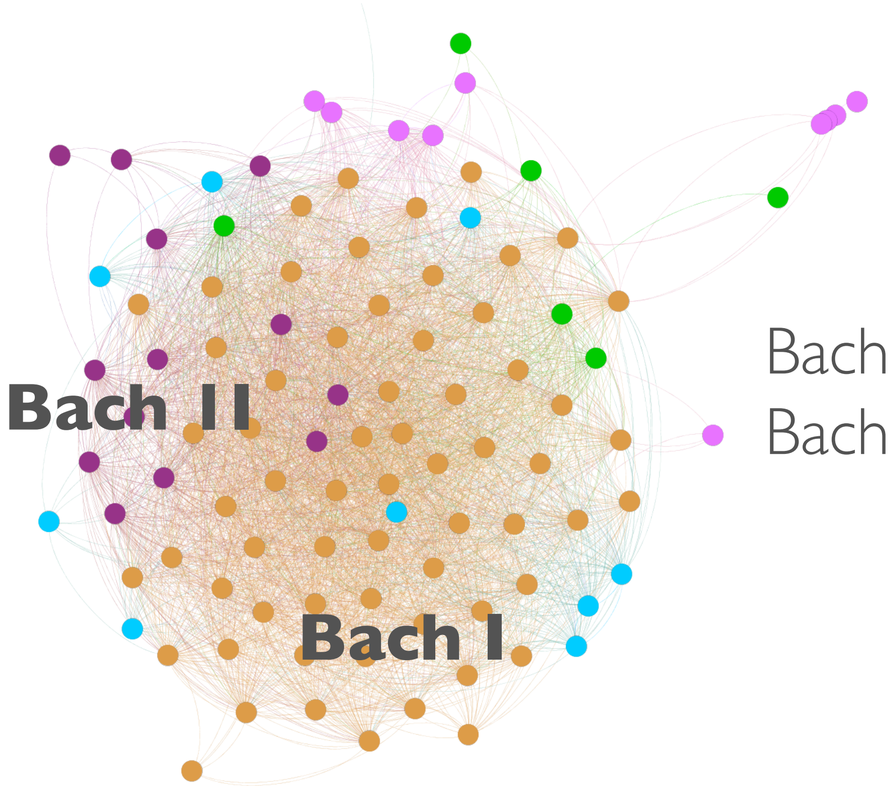




CHORD TRAJECTORY MATRIX



10 Clusters



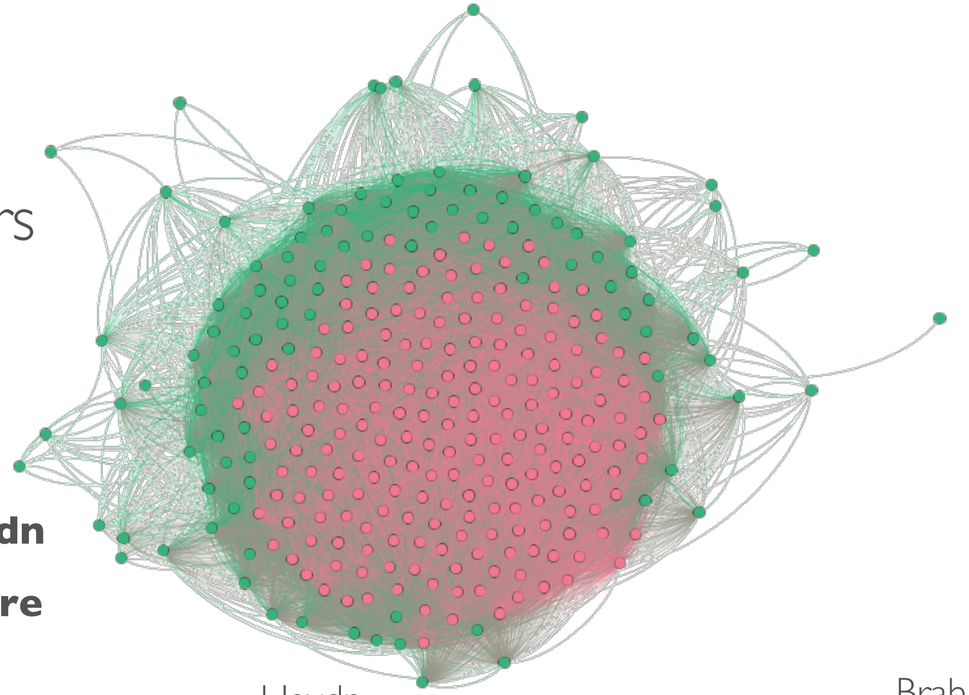
24 Clusters

Bach I: prelude-fugue-Bach
Bach II: suite-cello-Bach

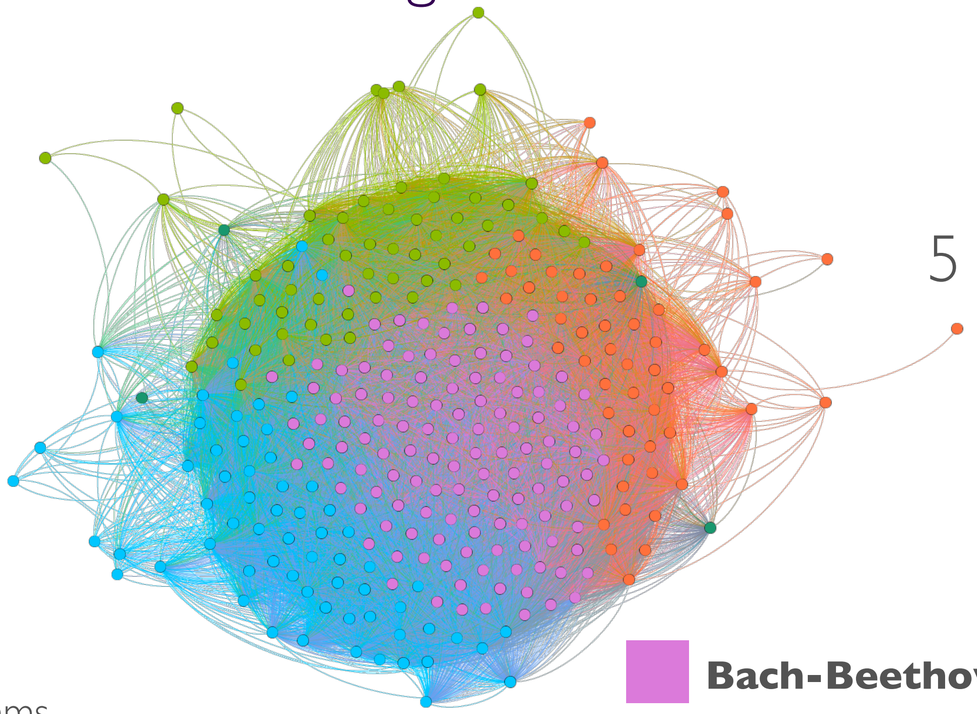
CHORD TRAJECTORY MATRIX

We could get a sense of how music changes across time

2 Clusters



5 Clusters



Bach-Haydn

Ravel-Faure

Bach-Beethoven-Cambini

Beethoven-Bach-Haydn

Beethoven-Cambini-Dvorak

Beethoven-Dvorak-Faure

Faure-Ravel-Schubert

Bach
1685-1750

Haydn
1732-1809
Cambini
1746-1825
Mozart
1756-1791

Beethoven
1770-1827
Schubert
1797-1828

Brahms
1833-1897
Dvorak
1841-1904
Faure
1845-1924

Ravel
1875-1937

Baroque → Classic → Romantic → Modern

SUMMARY & FUTURE WORK

- Take home message
  - Shingle size is critical for clustering
  - Chord Trajectory seems to be better than shingling based on our meta-data
  - We found clusters that aligns with composer's individual pieces and music develop periods
- Future work
  - Temporal information