Doctoral Candidacy Research Bibliography

Qingyang Xi
N13070555
700 Grove Street Apt 4H,
Jersey City, NJ 07310
(617)678-2045
tom.xi@nyu.edu
Program: Music Technology
Advisor: Dr. Brian McFee

January 31, 2022

1 General Bibliography

1.1 Music Theory and Musicology

[1] S. Adler and P. Hesterman, *The study of orchestration.* WW Norton New York, NY, 1989, vol. 2.

> Comprehensive reference book on essential orchestral combinations, instrument characteristics, and limitations.

[2] J. P. Burkholder, D. J. Grout, and C. V. Palisca, *A history of western music*, 8th ed. New York: W.W. Norton & Co., 2010.

> Survey textbook on Western music history. Provides basic musicological context for different periods.

[3] A. C. Cadwallader and D. Gagné, *Analysis of tonal music: a Schenkerian approach.* New York: Oxford University Press, 2011.

> Introductory text to Schenkerian Analysis. Concepts such as the Urlinie, Ursatz are introduced here, along with other devices such as linear progressions and initial ascent.

[4] W. E. Caplin, *Classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven.* Oxford University Press, 2001.

> This undergraduate level reference textbook is a standard for musical forms in the common-practice era. Clear and detailed rules are stated about segmentation on different hierarchical levels, and the relationships between the resulting segments.

[5] E. Chew *et al.*, "Mathematical and computational modeling of tonality," *AMC*, vol. 10, no. 12, p. 141, 2014.

> Deeply invested in representing tonality in a graphical and computational way, Elaine Chew is a Titan in the MIR community, and this more recent work of hers provides a template of how to produce scholarly work at the intersection of Music Theory and MIR.

[6] R. Cohn, *Audacious Euphony: Chromatic Harmony and the Triad's Second Nature.* Oxford University Press, USA, 2012.

> A prominent Neo-Reimmaninan theorist, Cohn's work follows the foot step of Lewin's Transformational theory. Triadic Voice-Leading Spaces are introduced, providing a basis for a topology in the triad space.

[7] J. J. Fux and J. Edmunds, *The study of counterpoint from Johann Joseph Fux's Gradus ad parnassum.* WW Norton & Company, 1965, no. 277.

> This is a textbook and reference on species counterpoint. Written over 200 years ago, this text is in the style of a conversation between a master and a pupil. Nevertheless, this concise manual provides ample musical examples for species counterpoint.

[8] J. Hepokoski and W. Darcy, *Elements of sonata theory: Norms, types, and deformations in the late-eighteenth-century sonata.* Oxford University Press, 2006.

> Sonata Theory gives detailed guidance on structures related to the sonata allegro form. These forms are bigger and more complex than those that are introduced in Caplin.

[9] S. G. Laitz and C. A. Bartlette, *Graduate review of tonal theory : a recasting of common-practice harmony, form, and counterpoint.* New York: Oxford University Press, 2010.

> Graduate level review of music theory during the common-practice era. Concise reference for basic music theory concepts. Topics include harmony, form and counterpoint.

[10] F. Lerdahl *et al.*, *Tonal pitch space.* Oxford University Press, USA, 2005.

> Following GTTM, Lerdahl starts to develop relationships between different scale degrees into a network of relationships, which he calls the tonal pitch space. Paths through these relationship networks give rise to trajectory in the pitch space, and provide more avenue for analytical interpretations.

[11] F. Lerdahl and R. S. Jackendoff, *A Generative Theory of Tonal Music, reissue, with a new preface.* MIT press, 1983.

> This classic in Music Theory combines insights from Natural Language Processing. While working exclusively in the symbolic domain, the authors formalize principles for music reduction and segmentation. While influenced by the Schenkerian school, GTTM does not follow strict Schenkerian rules.

[12] D. Lewin, *Generalized musical intervals and transformations.* Oxford University Press, USA, 2011.

> the seminal work that gave birth to Transformational theory. 3 essential triadic relationships, Parallel, Relative, and Leittonwechsel, are elevated to the same level of importance as dominant and subdominant relationships

[13] J. N. Straus, *Introduction to post-tonal theory.* WW Norton & Company, 2016.

> This text introduces Set Theory, which is a standard tool for analyzing post-tonal music. While tonality is not of much concern to Straus, the techniques and concepts of transformation theory can be applied with pitch class sets.

[14] J. Yust, *Organized time: Rhythm, tonality, and form.* Oxford University Press, 2018.

this new and important work breaks music apart into three orthogonal elements: tonality and pitch, rhythm and meter, and large scale forms. This work studies structures in these elemental spaces individually, before considering the gestalt of the combination.

1.2 General Music Information Retrieval

[15] J. P. Bello, "Measuring structural similarity in music," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 7, pp. 2013–2025, 2011.

This work compares structure of different pieces of music by looking at their intermediate representations, and the recurrence plots associated with. Fascinating article on how to go about comparing different structures.

[16] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler, "A tutorial on onset detection in music signals," *IEEE Transactions on speech and audio processing*, vol. 13, no. 5, pp. 1035–1047, 2005.

While operating on a very different timescale than typical music structure analysis, onset detection is still music segmentation. Ideas introduced here do not involve machine learning.

[17] E. Chew, "Slicing it all ways: Mathematical models for tonal induction, approximation, and segmentation using the spiral array," *INFORMS Journal on Computing*, vol. 18, no. 3, pp. 305–320, 2006.

This work focus on segmentation based on ideas of Chew's spiral pitch array.

[18] J. T. Foote and M. L. Cooper, "Media segmentation using self-similarity decomposition," in *Storage and Retrieval for Media Databases 2003*, vol. 5021. International Society for Optics and Photonics, 2003, pp. 167–175.

The Foote method applied on the self similarity matrix of an audio file is still a relevant method for automatic music segmentation.

[19] P. Grosche and M. Muller, "Extracting predominant local pulse information from music recordings," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 6, pp. 1688–1701, 2010.

> This beat tracker is able to provide expressive beat tracking, with no machine learning component.

[20] E. J. Humphrey, J. Salamon, O. Nieto, J. Forsyth, R. M. Bittner, and J. P. Bello, "Jams: A json annotated music specification for reproducible mir research." in *ISMIR*, 2014, pp. 591–596.

> A digital format to store annotations that's developed here at MARL. Will be relying on this schema for my expanded structure annotations.

[21] A. Lerch, *An introduction to audio content analysis: Applications in signal processing and music informatics.* Wiley-IEEE Press, 2012.

> A introductory text on music informatics. Different from Muller's FMP, this book contains more data driven techniques.

[22] B. McFee and J. P. Bello, "Structured training for large-vocabulary chord recognition." in *ISMIR*, 2017, pp. 188–194.

> This is a robust chord recognition algorithm, that employs structured training to leverage intrinsic structures in the annotation labels. This work is also an example of multitask learning.

[23] M. Müller, *Fundamentals of music processing: Audio, analysis, algorithms, applications.* Springer, 2015.

> a reference textbook for basic concepts and topics in MIR. Methods discussed range from onset detection to music structure analysis. Contents generally free of machine learning.

1.3 Machine Learning and Data Science

[24] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to probability.* Athena Scientinis, 2000.

Intro level textbook for probability. The book often provides cheat-sheet like summary of important facts in probability, and is therefore a good reference.

[25] C. M. Bishop, *Pattern recognition and Machine Learning.* Springer, 2006.

Classical text book on Machine Learning. A good reference that covers a broad range of topics, including graphical models and more. While lacking on deep learning techniques, this book is still the go-to reference for machine learning.

[26] R. M. Bittner, B. McFee, and J. P. Bello, "Multitask learning for fundamental frequency estimation in music," *arXiv preprint arXiv:1809.00381*, 2018.

an excellent example of using multi-task learning to solve the f0 estimation problem.

[27] J. Lee, N. J. Bryan, J. Salamon, Z. Jin, and J. Nam, "Disentangled multidimensional metric learning for music similarity," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2020, pp. 6–10.

This work shows how to use metric learning for disentangling different aspects of music similarity. Metric learning can be a useful tool for building my topology, and providing distance functions.

[28] A. Roberts, J. Engel, and D. Eck, "Hierarchical variational autoencoders for music," 2017.

This article is my reference for variational autoencoders (VAE). The hourglass structure of VAE means its a natural compressor, exposing different levels of resolution at different levels. This can be a relevant idea if I decide to work with hierarchical structures.

1.4 Digital Signal Processing

[29] B. Logan, "Mel frequency cepstral coefficients for music modeling," in *In International Symposium on Music Information Retrieval.* Citeseer, 2000.

Introduces MFCC and Cesptral analysis. These tools are still widely effective and relevant now.

[30] R. G. Lyons, *Understanding Digital Signal Processing, Third Edition.* Prentice Hall, 2010.

A reference book for digital signal theory.

[31] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way.* Academic Press, 2008.

Mallat's wavelet tour is the go to reference for wavelet techiniques, including Constant Q Transform and Scattering Transform.

[32] K. C. Pohlmann, *Principles of digital audio.* McGraw-Hill, 2011.

General textbook that covers major areas of music technology.

[33] C. Schörkhuber and A. Klapuri, "Constant-q transform toolbox for music processing," in *7th sound and music computing conference, Barcelona, Spain*, 2010, pp. 3–64.

A detailed article that supports several popular implementations of CQT. A good intro to CQT without the wavelet tour.

[34] J. O. Smith, *Mathematics of the discrete Fourier transform (DFT): with audio applications.* Julius Smith, 2007.

A reference for discrete Fourier transform.

[35] S. A. Van Duyne and J. O. Smith, "Physical modeling with the 2-d digital waveguide mesh," in *Proceedings of the international computer music conference.* INTERNATIONAL COMPUTER MUSIC ACCOCIATION, 1993, pp. 40–40.

Digital Waveguide Mesh is a computationally cheap way of generating complex sounds. It was a popular method for physical modeling instruments or room acoustics.

[36] U. Zölzer, X. Amatriain, D. Arfib, J. Bonada, G. De Poli, P. Dutilleux, G. Evangelista, F. Keiler, A. Loscos, D. Rocchesso, *et al.*, *DAFX-Digital audio effects.* John Wiley & Sons, 2002.

A well of digital effects information, this book is a excellent resource for looking up the inner workings and code of all major audio effects.

1.5 Immersive Audio, Acoustics, and Studio Technologies

[37] G. Davis and G. D. Davis, *The sound reinforcement handbook.* Hal Leonard Corporation, 1989.

Detailed and practical guide on many aspects of live sound reinforcement, from basic electrical properties of loudspeakers to techniques of preventing feedback in live situations, this book is a go to reference on all things live sound.

[38] P. M. Hofman, J. G. Van Riswick, and A. J. Van Opstal, "Relearning sound localization with new ears," *Nature neuroscience*, vol. 1, no. 5, pp. 417–421, 1998.

This paper demonstrate the flexibility and plasticity of human auditory localization. The study shows that while modifying the pinna significantly alter auditory localization cues, human are able to quickly re-adapt to the modified cues with training.

[39] P. Horowitz and W. Hill, *The art of electronics.* Cambridge university press Cambridge, 2002.

The Bible of electronics. Not only is it a excellent textbook in going through all the finer details of analogue and digital electronics, it is also a dictionary of a vast number of useful circuitry. The definitive reference in Electronics.

[40] G. S. Kendall, "A 3-d sound primer: directional hearing and stereo reproduction," *Computer music journal*, vol. 19, no. 4, pp. 23–46, 1995.

This introductory article concisely summarizes the most important auditory cues for spacial hearing: the Interaural Intensity Difference and Interaural Time Difference. Some common issues like front back confusion with reproduction method based on these cues are also discussed.

[41] D. G. Malham and A. Myatt, "3-d sound spatialization using ambisonic techniques," *Computer music journal*, vol. 19, no. 4, pp. 58–70, 1995.

This article introduces the ambisonic recording format and technique, which progressively decomposed the sound scene into spherical harmonics, which can be leveraged for localized recording and reproduction.

[42] J. Meyer, *Acoustics and the performance of music: Manual for acousticians, audio engineers, musicians, architects and musical instrument makers*. Springer Science & Business Media, 2009.

This text is a concise overview of the fundamentals of acoustics. Detailed account of orchestral instrument's sound propagation and the effects of the physical space.

[43] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the audio engineering society*, vol. 45, no. 6, pp. 456–466, 1997.

VBAP is a extension of stereo panning that work towards placing monophonic audio sources anywhere on a sphere, interpolating between recorded HRTF positions.

[44] J. H. Rindel, "The use of computer modeling in room acoustics," *Journal of vibroengineering*, vol. 3, no. 4, pp. 219–224, 2000.

This article is combines two established methods (source image method and ray-tracing method) for generating small room acoustics. A concise review of artificial reverb generation.

[45] A. Roginska and P. Geluso, *Immersive sound: The art and science of binaural and multi-channel audio*. Taylor and Francis, Jan. 2017.

A survey of major immersive audio recording and reproduction techniques. Basic reference on all things immersive audio.

[46] T.-Y. Wu, "Listening with realism: Sound stage extension for laptop speakers," *New York University-Department of Music and Performing Arts Professions Steinhardt School, New York*, 2013.

This article implements the Recursive Ambiophonic Crosstalk Elimination (RACE) algorithm for achieving quasi-biaural effects over loudspeaker settings. Effective and simple algorithm, that's easy to implement and with aurally salient results.

1.6 Psychoacoustics and Music Cognition

[47] M. Bosi and R. E. Goldberg, *Introduction to digital audio coding and standards.* Springer Science & Business Media, 2002, vol. 721.

> This definitive guide on perceptual coding explains auditory masking and its application in music coding. With discussion not only limited to spectral masking, but also temporal masking.

[48] A. S. Bregman, *Auditory scene analysis: The perceptual organization of sound.* MIT press, 1994.

> Several primitive and elementary audio segmentation principles are thoroughly reviewed. Some of the primitives introduced in this seminal work has been implemented into source separation algorithms with good results.

[49] D. Deutsch, *The Psychology of Music.* Academic Press, 2012.

> This book provides in-depth looks on the psychology of music, in a rather technical setting. This is my general reference on music psychology.

[50] M. M. Farbood, "A parametric, temporal model of musical tension," *Music Perception*, vol. 29, no. 4, pp. 387–428, 2012.

> This extensive article models musical tension from a cognitive perspective with dynamic and temporal experiences. This work provides a thorough review of previous work on musical tension, and is a good starting reference.

[51] C. L. Krumhansl, *Cognitive foundations of musical pitch.* Oxford University Press, 2001.

> This work delves deep into the perception of pitch, with many experiments included. Discussion goes beyond common practice period and touches upon atonality.

[52] E. W. Large and M. R. Jones, "The dynamics of attending: How people track time-varying events." *Psychological review*, vol. 106, no. 1, p. 119, 1999.

Tackles the issues of how human perceive priodic events with varying rates. How a expressive musical performance doesn't lie on "the grid", nevertheless, the rhythmic relationships can still be clearly perceived.

[53] E. W. Large and C. Palmer, "Perceiving temporal regularity in music," *Cognitive science*, vol. 26, no. 1, pp. 1–37, 2002.

This work discusses human perception of rhythm and groove, also touches upon micro-timing.

[54] R. Parncutt and G. Hair, "A psychocultural theory of musical interval: Bye bye pythagoras," *Music Perception: An Interdisciplinary Journal*, vol. 35, no. 4, pp. 475–501, 2018.

This work examines human perception of musical intervals, and challenges the principle of interger-ratios set out by Pythagoras. This work provides new ways to establish interval relationships.

[55] W. F. Thompson and R. Parncutt, "Perceptual judgments of triads and dyads: Assessment of a psychoacoustic model," *Music Perception*, vol. 14, no. 3, pp. 263–280, 1997.

Perception study of the pitch space taken into account fundamentals, overtones, and subharmonics.

2 Specialized Bibliography

2.1 Automatic Music Structure Analysis

[56] D. P. Ellis, "Identifying'cover songs' with beat-synchronous chroma features," 2006.

One of the earily papers that defines the task of cover song identification. This is now a popular MIR task that's related to music structure.

[57] B. McFee and D. Ellis, "Analyzing song structure with spectral clustering." in *ISMIR*. Citeseer, 2014, pp. 405–410.

a clear demonstration of how to use spectral graph methods to tackle the problem of audio segmentation. It is well noted that the spectral methods generate segments that tend to have bigger variance in their duration.

[58] B. McFee and K. M. Kinnaird, "Improving structure evaluation through automatic hierarchy expansion," in *Proceedings of the 20th ISMIR Conference, Delft, Netherlands*, 2019.

This work address the issue that structures are often hierarchical, and comparison between structures that are on different levels of hierarchy is hard. This work provides a way for structures annotated on different granularity to become more easily comparable.

[59] M. Müller and N. Jiang, "A scape plot representation for visualizing repetitive structures of music recordings." in *ISMIR*. Citeseer, 2012, pp. 97–102.

the scape plot provides a novel visualization of the hierarchical structure of the audio file under analysis. There seems to be influence from GTTM in this work.

[60] O. Nieto, "Discovering structure in music: Automatic approaches and perceptual evaluations," Ph.D. dissertation, New York University, 2015.

A recent dissertation on music structure analysis from MARL. This is a starting reference for a survey of topics related to structure analysis. The work makes clear the different approaches the MIR community and music cognition community take to tackle this problem.

[61] O. Nieto and J. P. Bello, "Music segment similarity using 2d-fourier magnitude coefficients," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 664–668.

This article proposes a way of relating different musical segment via a standard metric. These metrics can be useful when building relational spaces with segments as members.

[62] O. Nieto and M. M. Farbood, "Identifying polyphonic patterns from audio recordings using music segmentation techniques," in *Proc. of the 15th International Society for Music Information Retrieval Conference*, 2014, pp. 411–416.

This article shows how segmentation can lead to pattern discovery in musical structures.

[63] O. Nieto, M. McCallum, M. E. Davies, A. Robertson, A. M. Stark, and E. Egozy, "The harmonix set: Beats, downbeats, and functional segment annotations of western popular music." in *ISMIR*, 2019, pp. 565–572.

This is another newer dataset that comes with segment annotations.

[64] O. Nieto, G. J. Mysore, C.-i. Wang, J. B. Smith, J. Schlüter, T. Grill, and B. McFee, "Audio-based music structure analysis: Current trends, open challenges, and applications," *Transactions of the International Society for Music Information Retrieval*, vol. 3, no. 1, 2020.

a survey article that summarize the current trend and challenge of audio-based music structure analysis from the point of view of he MIR community.

[65] A. Rosenberg and J. Hirschberg, "V-measure: A conditional entropy-based external cluster evaluation measure," in *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, 2007, pp. 410–420.

Different from other supervised learning tasks, the evaluation metric of segmentation is not immediately obvious. V-measure is a recognized metric for segmentation evaluation.

[66] J. Salamon, J. Serrà, and E. Gómez, "Melody, bass line, and harmony representations for music version identification," in *Proceedings of the 21st International Conference on World Wide Web*, 2012, pp. 887–894.

This is also a work on cover song detection, by looking at several features of the music concurrently.

[67] J. Serra, X. Serra, and R. G. Andrzejak, "Cross recurrence quantification for cover song identification," *New Journal of Physics*, vol. 11, no. 9, p. 093017, 2009.

an extension to Self Similarity Matrix, Cross Recurrence has been a classical idea that has survived the invasion of Deep Learning. Still a relevant technique to acquire.

[68] J. B. L. Smith, J. A. Burgoyne, I. Fujinaga, D. De Roure, and J. S. Downie, "Design and creation of a large-scale database of structural annotations." in *ISMIR*, vol. 11. Miami, FL, 2011, pp. 555–560.

> The SALAMI dataset is one of the most commonly used dataset for evaluating structure analysis; the dataset is described in this paper.

[69] C. J. Tralie and B. McFee, "Enhanced hierarchical music structure annotations via feature level similarity fusion," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, 2019, pp. 201–205.

> this SSM based work shows how the output structure of a MSA tool is deeply tied with the audio feature from which the SSM is constructed upon. This has pushed me to promote independently unfolding structures, as opposed to merging several structures into a single one, which was done in this work.

[70] K. Ullrich, J. Schlüter, and T. Grill, "Boundary detection in music structure analysis using convolutional neural networks." in *ISMIR*, 2014, pp. 417–422.

> an example of a deep learning method that tackles the question of MSA. MSA has been one of the sub-field in MIR that has seen some resistance to rapidly adapting deep learning techniques, and this article provides a baseline.

2.2 Music Theory

[71] C. Callender, "Continuous harmonic spaces," *Journal of Music Theory*, vol. 51, no. 2, pp. 277–332, 2007.

> This is a succession from Ian Quinn's dissertation. The concept that musical objects occupy relational spaces are further developed here. The topology of these musical spaces are at the heart of my research interest.

[72] C. Callender, I. Quinn, and D. Tymoczko, "Generalized voice-leading spaces," *Science*, vol. 320, no. 5874, pp. 346–348, 2008.

Setting up with five invariances, (octave shift, permutation, transposition, inversion, and cardinality change), this work explores how these new rules can influence the voice leading space and create new progressions.

[73] E. Chew, "The spiral array: An algorithm for determining key boundaries," in *International Conference on Music and Artificial Intelligence.* Springer, 2002, pp. 18–31.

Chew's organization of pitch space, with an heavy emphasis on tonality.

[74] F. Lerdahl and C. L. Krumhansl, "Modeling tonal tension," *Music perception*, vol. 24, no. 4, pp. 329–366, 2007.

This is a good introduction to Lerdahl's theory of tonal tension. Tension can be understood as one valence dimension that can potentially be captured by a appropriate graph/topology.

[75] M. Montiel and R. W. Peck, *Mathematical Music Theory: Algebraic, Geometric, Combinatorial, Topological and Applied Approaches to Understanding Musical Phenomena.* World Scientific Publishing, 2018.

This very recent book seems to indicate a increase in interest in using more mathematical and computational tools for musical analysis. Comprised of articles, this can be a reference for a survey of techniques that are commonly deployed in contemporary Music Theory.

[76] I. Quinn, "General equal-tempered harmony (introduction and part i)," *Perspectives of New Music*, pp. 114–158, 2006.

Quinn explores the space of harmony, and explore the relationships and distances of these objects in a systematic and quantitative way.

[77] N. A. Smith and L. Cuddy, "Perceptions of musical dimensions in beethoven's waldsiein sonata: An application of tonal pitch space theory," *Musicae Scientiae*, vol. 7, no. 1, pp. 7–34, 2003.

This will be a case study for me to refer to on a piece of music that I know well.

[78] W. M. Szeto and M. H. Wong, "A graph-theoretical approach for pattern matching in post-tonal music analysis," *Journal of New Music Research*, vol. 35, no. 4, pp. 307–321, 2006.

This work uses Set Theory and graph approaches to analyze Schoenberg's Op.11 No.1, also a piece that I know rather well.

[79] D. Tymoczko, *A geometry of music: Harmony and counterpoint in the extended common practice*. Oxford University Press, 2010.

Another important reference book that pays special attention to shapes and geometries in a musical relational space.

[80] ——, "The generalized tonnetz," *Journal of Music Theory*, pp. 1–52, 2012.

the Tonnetz, being one of the oldest and most important graph/topology, has alternative constructions. By systematically exploring the potential intervals that are represented by the edges, extended tonnetz can account for non-traditional harmonies outside of the triadic language.

[81] J. Yust, "Generalized tonnetze and zeitnetze, and the topology of music concepts," *Journal of Mathematics and Music*, vol. 14, no. 2, pp. 170–203, 2020.

Recent article that closely examines the graph/topology aspects of tonnetz. Many details about extended tonnetz that are not apparent are thoroughly discussed here. A great reference for working with Tonnetz in general.

2.3 Graph Theory and Topology

[82] M. G. Bergomi, A. Baratè, and B. Di Fabio, "Towards a topological fingerprint of music," in *International Workshop on Computational Topology in Image Context*. Springer, 2016, pp. 88–100.

This is a potential application for analyzing the structure with an emphasis on graphs. The shapes these graphs create can be used to identify musical pieces, and is called topological fingerprinting.

[83] F. Chung, "Laplacians and the cheeger inequality for directed graphs," *Annals of Combinatorics*, vol. 9, no. 1, pp. 1–19, 2005.

How to form Laplacians for directed graph, if I decided to apply spectral techniques with directed graphs. Reference in case the need arises.

[84] E. D. Demaine, F. Gomez-Martin, H. Meijer, D. Rappaport, P. Taslakian, G. T. Toussaint, T. Winograd, and D. R. Wood, "The distance geometry of music," *Computational geometry*, vol. 42, no. 5, pp. 429–454, 2009.

Similar to Groove Pizza, this paper talks about rythmic patterns as trajectroy on a circular graph, with the concept of maximually even set.

[85] F. C. Graham, *Spectral graph theory.* American Mathematical Soc., 1997, no. 92.

My main reference on spectral graph theory. Chapter 1-4 seems perticularly relavent.

[86] J. R. Hughes, "Using fundamental groups and groupoids of chord spaces to model voice leading," in *International Conference on Mathematics and Computation in Music.* Springer, 2015, pp. 267–278.

Introduced how algebraic topology tools (groups, category, homotopy), and how they can work with Music. Studies paths in different musical topologies.

[87] V. Lostanlen, "Eigentriads and eigenprogressions on the tonnetz," *arXiv preprint arXiv:1810.00790*, 2018.

This very condensed work shows how scattering transform can be applied to graphical structures like the tonnetz.

[88] C. Louboutin and F. Bimbot, "Modeling the multiscale structure of chord sequences using polytopic graphs," in *18th International Society for Music Information Retrieval Conference*, 2017.

This work looks at the relationship between harmonic events across different hierarchical time-spans. The comparison of harmonies with different time-strides is systematically considered.

[89] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

Consice introduction to spectral clustering. Close to 10k citations.

[90] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 1, pp. 4–24, 2020.

a reference that I'm keeping in the case that graph neural network becomes a useful tool.

[91] L. Yao and P. Bendich, "Graph spectral embedding for parsimonious transmission of multivariate time series," in *2020 IEEE Aerospace Conference*. IEEE, 2020, pp. 1–12.

A little bit outside of my comfort zone, this one introduces a technique called Laplacian Events Signal Segmentation, which is a unsupervised technique that segments time series into events.

2.4 Topological Data Analysis

[92] P. Bendich, E. Gasparovic, J. Harer, and C. J. Tralie, "Scaffoldings and spines: organizing high-dimensional data using cover trees, local principal component analysis, and persistent homology," in *Research in computational topology*. Springer, 2018, pp. 93–114.

A topological technique of organizing point cloud data, with applications on Music data. My reference for further investigation with TDA.

[93] V. Lostanlen, S. Sridhar, B. McFee, A. Farnsworth, and J. P. Bello, "Learning the helix topology of musical pitch," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 11–15.

This work demonstrate that natural sonic snippets have intrinsic organization in their own space, and can be discovered via unsupervised training. The technique of isomap employed in this work can be useful when trying to discover the topological structures of many member segments.

[94] C. Tralie, "High-dimensional geometry of sliding window embeddings of periodic videos," in *32nd international symposium on computational*

geometry (socg 2016). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

By studying sliding window embeddings (short continous segments) of periodic videos with topological tools, repeatable segments can be understood as loops on hypertorus.

[95] C. J. Tralie, “Geometric multimedia time series,” Ph.D. dissertation, Duke University, 2017.

My primary reference on TDA. This work examines the geometrical shape of multimedia time series (audio and video), and use TDA tools and idea to work on audio and video problems. Concepts such as persistent homology and persistence graphs are used heavily in this work.

[96] B. Xu, C. J. Tralie, A. Antia, M. Lin, and J. A. Perea, “Twisty takens: A geometric characterization of good observations on dense trajectories,” *Journal of Applied and Computational Topology*, vol. 3, no. 4, pp. 285–313, 2019.

Following up on (Tralie, 2016), this is a more general study of trajectory formed in different topologies by sliding window embeddings of time series.