

RECONSIDERING QUANTIZATION IN MIR

Jordan Lenchitz

College of Music, Florida State University
jlenchitz@fsu.edu

ABSTRACT

This paper presents a critique of the ubiquity of boilerplate quantizations in MIR research relative to the paucity of engagement with their methodological implications. The wide-ranging consequences of reflexivity on the future of scholarly inquiry combined with the near-universal contemporary recognition of the need to broaden the scope of MIR research invite and merit critical attention. To that end, focusing primarily on twelve-tone equal-tempered pitch and dyadic rhythm models, we explore the practical, cultural, perceptual, historical, and epistemological consequences of these pervasive quantizations. We analyze several case studies of meaningful and successful past research that balanced practicality with methodological validity in order to posit several best practices for both future intercultural studies and research centered on more narrowly constructed corpora. We conclude with a discussion of the dangers of solutionism on the one hand and the self-fulfilling prophecies of status quoism on the other as well as an emphasis on the need for intellectual honesty in metatheoretical discourse.

1. INTRODUCTION

Fostering cultural diversity in MIR research is not merely a question of “adapting” existing methodologies developed on the basis of certain *a priori* assumptions to repertoires or tasks that perhaps challenge said assumptions. In particular, it should also entail critical reflection on the benefits and drawbacks of those assumptions in studies of all repertoires, even (if not especially) those on the basis of which such assumptions were made in the first place. Twelve-tone equal-tempered quantizations of pitch and dyadic quantizations of rhythm represent arguably the most ubiquitous such assumptions in contemporary MIR research, and yet despite their prevalence their foundational role underlying many diverse methodological approaches all too often passes unannounced. This paper therefore presents an analysis not of any musical information in particular but rather of the apparatuses we use to retrieve it from musics in the hopes of fostering productive future methodological conversations.

Quantization may be defined as the organization of dimensional information into discrete sets of values (otherwise known as categorization in perception science), and it is cognitively essential for creating music. That being said, the centrality of the process of quantization to music

does not justify the perpetuation of reliance on any “default” quantizations. Indeed, certain quantizations are ubiquitous, all-too-often unstated *a priori* assumptions underlying a substantial majority of MIR methodologies. Are they compromises? Almost invariably yes they are. Do they represent pragmatic choices? Quite possibly they do, depending on the context. But pragmatic compromises or otherwise, the foundational position of these quantizations endows them with consequences that merit consideration in the context of any methodological decision-making and especially if meaningful progress is to be made in expanding the purview of MIR and its applications.

This paper is structured as follows: after this introduction providing the rationale for a reconsideration of quantization in MIR, its consequences practical, cultural, perceptual, historical, and epistemological are each explored in turn with an emphasis on pitch and rhythm in the second section. The third section provides illustrative case studies that demonstrate strategies for balancing pragmatism with methodological validity and posits best practices for both intercultural and intracultural research. The fourth and final section argues for the necessity of avoiding both solutionism on the one hand and status quoism on the other and emphasizes the importance of intellectual honesty in metatheoretical discourse.

2. CONSEQUENCES OF QUANTIZATION

2.1 Practical Consequences

The most obvious practical consequence of the ubiquity of certain quantizations in MIR is the widespread availability of platforms built around them and the corresponding noticeable absence of alternatives. In and of itself this is unsurprising and not necessarily a drawback; and yet, it can bring about the existence of unfortunate self-fulfilling prophecies. Consider the case of the justifiably popular Python audio analysis package *librosa* [1], in which one may specify the number of notes per octave but the assumption that they are equally spaced is not so easily changed. As maintainer Brian McFee put it with respect to pitch on the Music Information Retrievers Slack in July 2020, equal temperament “is a compromise, but one I’m willing to live with for the time being; extending to support just intonation in a fully consistent way would be a huge undertaking, much bigger than just adding notation support.” We are inclined to agree with McFee in his assessment that implementing meaningful support for tuning systems other than equal temperaments would be a very nontrivial task. At the same time, however, such tasks tend to be welcomed by MIR researchers as motivation for innovation. Why, then, does this remain unaddressed?



© J. Lenchitz. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** J. Lenchitz, “Reconsidering Quantization in MIR”, in *Proc. of the 22nd Int. Society for Music Information Retrieval Conf.*, Online, 2021.

Imagine a hypothetical MIR researcher who desires to do work on musics in non-equal-tempered pitch systems. Methodological convenience is seldom acknowledged as a motivating factor in the planning and implementation of research plans and yet psychologically it has an effect on what individuals decide to do and not to do. The existence vs. nonexistence of tools structured so tightly around certain pitch quantizations, in turn, has implications for such convenience or lack thereof. Adding the need to produce a novel notation and/or data format on top of the already major task of encoding a new corpus in a machine-readable format as prerequisites for exploring avenues for actual information retrieval can make the scale of a prospective project quickly balloon in size and therefore be discouraging to its actual execution. And even for projects that do make it to completion, there remain the potential for obstacles in the peer-review process. Excessive specialization brings about the possibility for manuscripts to be reviewed by homogeneous niches whose shared assumptions may prevent work from being disseminated to the broader community and from being built upon by subsequent endeavors, and such dynamics of review admit no simple fixes.

Beyond non-symbolic tools such as *librosa*, it additionally bears mentioning that symbolic tools have also embraced the compromise of equal-tempered quantization of pitch at just about every level of their functionality. To name just a few, *Humdrum* [2] and *music21* [3] are equally “not fully compatible” [4] with musics which do not consist of twelve pitches logarithmically equally spaced over 2:1 octaves in terms of pitch structure. These tools are all of course excellent in countless use cases and this is not by any means to suggest they be abandoned. Rather, this acknowledgement of their shared assumption with respect to the modeling of pitch leads to a fuller and more useful understanding of their limitations in considering what one might choose in terms of platform for a study if this exact assumption is not desired to be made for any given research plan. In particular, the pitch quantizations inherent in the notation encodings upon which such symbolic tools rely could be taken to motivate non-symbolic approaches in cases where one’s target musics are not served as well by these encodings’ inherent assumptions.

2.2 Cultural Consequences

Anthropologists, ethnomusicologists, linguists and other scholars distinguish between emic and etic understandings of culturally-specific phenomenon, which can be approximately understood as insider and outsider perspectives [5]. Quantizations in MIR are overwhelmingly based on etic views of culturally-specific musical phenomenon which harm the ecological validity of methodologies. Twelve-tone equal-tempered pitch quantization is regrettable in this regard because of the number of musical cultures in the world that use or have used more or fewer than twelve notes per octave (e.g. Indian *rāga*, Turkish *makam*, or Indonesian Gamelan musics), twelve non-equally-spaced notes per octave (e.g. much of the history of Western musics, as discussed further in subsection 2.4) or non-octave based pitch structures (e.g. tritave-based works by Wendy Carlos et al.). Tools that are ill-suited to handle the musical

realities of these and similar culturally-specific phenomenon lead either to their exclusion from consideration or (and arguably worse) facilitate their inclusion but in ways that do them a disservice.

Dyadic rhythm quantization has similarly negative consequences, especially in light of the differing approaches to meaning and teleology across musical cultures. Micro-timings pose a practical challenge in any case but depending on the culture in question, such minute expressive discrepancies from most quantizations could serve as the primary determinants of culturally-specific meaning. And with respect to teleology, cyclic vs. linear conceptions of rhythm and time ought to factor in to how and why we decide to treat rhythm in both symbolic and signal processing-based approaches (e.g. because structural repetitions might rely on subtle changes near the beginnings or ends of units of repetition part and parcel with large-scale formal processes). Yet by and large these implications are either swept under the rug or taken to mean that work will instead focus elsewhere.

At the beginning of this paper we alluded to the fact that fostering cultural diversity in MIR is not merely a matter of “adapting” existing methodologies to a broader variety of musics. The cultural consequences of quantization highlight this fact in that it would be quite simple if not trivial to apply say a twelve-tone chroma feature to Turkish *makam* or a dyadic rhythm quantization to Burundi Whispored Inanga but any results so obtained could be meaningless to practitioners of those musical cultures. It is not enough, furthermore, to involve emic perspectives in a research plan if its computational approaches do not also take them into account. Consideration of not only the primary audiences of MIR research but also all of the stakeholders in its endeavors leads to the realization that the deeper issues of cultural diversity thus lie in methodology as well as in repertoire.

2.3 Perceptual Consequences

Many if not all MIR methodologies attempt to relate to humans’ perceptions of musics in some way in order to yield results relevant to the experiences of listeners. And yet, the quantizations of pitch and rhythm that underlie many of these methodologies exhibit to a nontrivial extent an arbitrariness divorced from perceptual realities. In the domain of pitch, for instance, the number of notes in the twelve-tone equal-tempered scale from 16 Hz to 16 kHz is only 120 whereas the number of perceptible pitch steps in the same range is approximately 1400 [6]. If a methodology sought to explore perceptual quantization among individuals with absolute pitch and socialized in musical cultures featuring twelve notes per octave, the decision to limit its pitch quantization accordingly would follow naturally. Most methodologies, however, do not intend to ask and attempt to answer such questions and yet they employ this quantization of pitch all the same.

With respect to rhythm quantization, even SOTA or near-SOTA results similarly bring caveats in terms of their perceptual implications. Automatic transcription does not claim to be a representation of listeners’ experiences of the music in question, and yet its use in analyzing performed divergences from symbolic rhythmic notation attests to a

modeling of performers' conscious or subconscious decision-making in some sense. Automatic transcriptions of performances by the legendary Canadian musician Glenn Gould, for instance, produce rational approximations of his keyboard ornamentation necessary to fit into the prevailing dyadic rhythmic grid [7]. Such results are valid in the sense that they more or less closely match the raw interonset interval (IOI) data up to the capability of the algorithms in question. Still, from a perceptual perspective, the idea that either listeners or performers are themselves quantizing the dizzying diversity of IOIs present into some similarly messy rational dyadic mental representation is rather unfortunate.

This is not to say that MIR models that do not claim to be perceptual models should be rejected on perceptual grounds. On the contrary, such models should be considered in terms of their claims and their intentions. At the same time, however, recognition of the perceptual consequences of quantization in the modeling of musics could easily lead to more reflexivity in scholarly discourse, the absence of which seems to reflect an unspoken consensus. Consensus is one straightforward means of making tangible progress on solving difficult problems but when it is merely implicit and unspoken it is less likely to serve as a solid foundation for such progress. The perceptual implications of quantizations in MIR therefore would do well to figure more in scholarly discourse both for the sake of increasing methodological validity and also to promote engagement with the state of the art in the cognitive sciences—aims which have indeed been documented for as long as ISMIR has existed [8].

2.4 Historical Consequences

If we narrow our scope temporarily to studies of Western European classical musics, we find that common quantizations of pitch and rhythm have striking consequences with respect to historical ecological validity. Revisionist histories of tuning do not change the historical reality that twelve-tone equal temperament is a relatively recent phenomenon in practice. In terms of repertoire, one example par excellence is J. S. Bach's *The Well-Tempered Clavier* (WTC), a collection of keyboard works whose very title explicitly testifies to the composer's intention to explore all 24 unequal major and minor keys which were usable compared to meantone and yet had different affective profiles due to minor but nontrivial intervallic size differences. Any information retrieved from a symbolic representation of the WTC imposing the enharmonic equivalence and logarithmically equal note spacing of twelve-tone equal temperament on its pitch content would therefore do a disservice to the historical circumstance surrounding the creation of the work and divorce results obtained from its meaning at the time of its creation.

If we look more recently in music history, there do exist plenty of corpora for which a twelve-tone equal-tempered pitch model is quite appropriate. Composers working since the advent of twelve-tone equal-tempered tuning who rely on this pitch logic in the structure of their musics are a natural fit for this particular quantization, as are the generations of popular musicians who inherited this structure more or less wholesale. In a similar vein to what Cella [9]

has observed in the case of contemporary classical music, one possible explanation for the application of this quantization to musics earlier than for which it is best-suited may lie in the 'follow the money' reality of many MIR projects. That is, given that the largest audiences today are served by tools centered around the quantization of pitch most common in modern popular musics, one would not be surprised for there to be less investment in tools which would best handle earlier pitch quantizations. If dynamics of pitch quantization were fully considered before planning and executing research on historical repertoires, we might study some understudied repertoires more and overstudied repertoires less.

Rhythm quantization, in turn, is analogously problematic from the perspective of impacting historical ecological validity. The case of French Baroque *notes inégales* may serve as a representative example. In this case the widespread divergence of performance practice from symbolic representations of the music leads to situations where MIR studies based on symbolic data can make valid claims about the symbolic data that are nonetheless not reflective of the music as performed by historically-informed practitioners and would therefore not align with non-symbolic studies of the same music. Moreover, rhythm quantization in cases of performance practice poses a much less daunting (though still nontrivial) challenge for well-intentioned researchers than pitch. Whether or not something is challenging is seldom the sole determining factor in the processes of methodological development, and yet the relative ease of addressing *notes inégales* suggests that the broader trend at play here is a minimization of concern for historical realities in the implementation of such studies rather than practicality.

2.5 Epistemological Consequences

At once the most concerning and least-addressed drawback of the omnipresence of certain quantizations in MIR is their tendency to foster confirmation bias. If our hypotheses are to withstand critical scrutiny, they must be not only falsifiable but also tested in such a way that the results are not preordained by our methodological decision-making. A useful lesson can be drawn from the example of studies of cognitive and academic benefits of music training in children, where a recent meta-analysis suggests that confirmation bias might well have influenced the validity of nearly four decades of results [10]. The lesson to learn from these experimental studies is that when research is designed with the expectation to find a certain result, it should not be surprising nor necessarily meaningful to produce that result.

For example, the use of twelve-tone equal-tempered chroma features in signal-based analysis can quite possibly produce statistically significant results with respect to those bins. This does not, however, necessarily tell us anything meaningful about the underlying signal because such bins are merely rounding the chroma information actually present to the nearest 100 cents. The same principle applies to dyadic rhythm quantization of IOI information, where the results obtained can be elegant and convincing with respect to this quantization but again do not neces-

sarily reveal a deeper truth about the signal in question depending on the nature of its unquantized IOIs. Multidimensional attempts to quantize timbre, in turn, have had success in synthesis applications [11] that should not however be interpreted as a guarantee that the same approaches will lead to ecologically valid results in analysis.

Another epistemological drawback of such quantizations is their tendency to foster selection bias. As discussed earlier in the context of practical consequences, when mainstream tools exist that are suited for certain musics and not others, it should not come as a surprise that a significant number of researchers tend to apply those tools to those musics at the expense of others. Given the present recognition of the need to counter sampling bias in MIR research, all factors that contribute to it accordingly deserve acknowledgment and discussion. Methodologically careful intercultural studies as well as intracultural research have been performed and continue to be performed by members of the MIR community, and this is not to neglect their contributions (which will be addressed in the following section). Rather, to understand the role of these quantizations in contributing to selection bias is to admit that despite their ubiquity, they are far from “neutral” positions one might assume them to be.

3. CASE STUDIES AND BEST PRACTICES

If the preceding critiques of boilerplate quantizations have perhaps presented a morose picture of the state of contemporary MIR, the following case studies and recommended best practices should serve as a more optimistic change of pace. There are undoubtedly other case studies we have omitted that would serve just as well as examples and other best practices we do not recommend here. This section is therefore best understood as an attempt to highlight directions for future work to complement the aforementioned discussions of the consequences of “default” quantizations.

3.1 Case Studies

Although it is rather uncommon, theoretical work does exist that has explored pitch with an eye toward information retrieval through continuous rather than quantized perspectives. Callender [12] has investigated continuous harmonic spaces using a Fourier-based approach to symbolic data that enables the application of “harmonic intuitions to all possible chords of pitches and pitch classes in all possible tuning systems.” Wakefield [13] has examined the mathematical and computational implications of “joint time-chroma distributions” that could be used to produce unquantized chroma features on the signal-processing side. Neither of these theoretical contributions has been the subject of much follow-up work which is regrettable considering that they suggest possibilities for methodologies approaching the extraction of pitch information without assuming any particular quantization of pitch in advance. At the same time, that continuous alternatives to the received wisdom of certain quantizations have largely passed unnoticed in MIR is to an extent to be expected since those most committed to working with specific sets

of discrete values are not incentivized to explore options beyond them.

Approaches to quantization starting either from a continuous perspective or closer to it also merit mention here. Moelants, Cornelis, and Leman [14] implemented a methodology in which “pitch is first analyzed on a continuous scale” and “peak analysis is then applied on these data to extract the actual scale used.” Among the advantages of this approach are the fact that it is applicable to many signal-based MIR methods and that it is easily generalizable across repertoires. Six and Cornelis [15] used a granular “resolution of 1200 cents” to cover more than the pitches of twelve-tone equal temperament in order to “form musically meaningful representations” of non-Western musical traditions. In both of these cases the methodology was planned and implemented with maximal ecological validity as one desired outcome. Whether starting with continuous data or relying on a granular quantization for the sake of more meaningful analysis, both of these examples demonstrate that it is not only possible but quite doable to make such decisions in one’s own research agenda when the musical situations at hand call for it.

Two differing approaches to working with histogram bins can serve as worthwhile examples of quantizations not starting from continuous perspectives but still centering culturally-specific knowledge in their computational implementations. Panteli [16] in a comparative study of Cypriot pitch patterns increased their histogram resolution by a factor of three compared to the octave partitions specified in Byzantine and Turkish theoretical sources “for better precisions and tuning robustness” while remaining tied to the traditions’ emic perspectives. Bozkurt [17] found in the analysis of traditional *makam* music in Turkey performed by a living master that the histogram matched neither twelve-tone equal temperament nor the official standard tuning system of the music as codified in print sources and accordingly designed a tuning application based on such recordings rather than any frequency presets. Working within some discrete universe to start but modifying it to suit the investments of multiple stakeholders can balance tensions inherent in intercultural work.

One research project that serves as the umbrella for many approaches to quantization worthy of emulation is “CompMusic: Computational Models for the discovery of the world’s music” [18] coordinated by Serra. Recognizing the drawbacks of the hegemony of Western-centered paradigms in MIR, this project not only includes members from each of the cultures being studied but also targets practitioners of these specific traditions in its development of interactive systems. Another example of an organizational umbrella fostering potential departures from common quantization norms was the Music Encoding Initiative (MEI) and their MEI Incubator [19], which gave practitioners a “common space to ‘grow’ their customizations and share them with other members of the community.” Both of these case studies are also helpful reminders of the fact that much of the work needing to be done to overcome limitations of certain quantizations lies in the domain of encoding and formatting corpora for processing.

Lastly, two more recent examples involving the study of Indian art music (IAM) are further demonstrative of other possibilities for sensibly handling quantization in a

way most germane to the repertoire in question. The methodology of Ranjani et al. [20] involved “non-uniform quantization intervals...selected from pitch and time scales prevalent in IAM [and] accommodating pitch and inter-note-interval variations on [a] pitch-time grid,” accounting for the repertoire in the methodology rather than the other way around. Viraraghavan et al. [21], in turn, proposed a transcription methodology involving a Viterbi algorithm that outperformed uniform quantization (which would be the default in MIR of Western musics) in terms of adherence to *rāga*. The commonalities between these two contributions accordingly exceed their corpus and their aims in that both made conscious methodological decisions that decentered assumptions made in the study of Western musics in accordance with the questions they sought to ask and answer.

3.2 Best Practices

There are several potential best practices to be drawn from the brief survey of case studies presented above. Chiefly among these is the recognition of the continuous reality of pitch as a prerequisite for an implementation of its quantization. Our position is not that MIR researchers believe as Vincenzo Galilei did that pitch is discrete nor that all quantizations are unconscious reflexes but rather that intentionality is key in these methodological underpinnings. Continuous models used to obtain discrete ones (as in [14]) are applicable to any musics and increased granularity (as in [15]) need not be limited to non-Western repertoires. Discrete models informed by the theoretical and performance output of practitioners within the cultures in question [16-17] can foster increased usability of results beyond scholarly applications. Especially for the study of expressive intonation, experimental and microtonal musics, and signal-processing based on performance rather than symbolic analysis, these approaches with or without the orientation of explicitly continuous thinking at all stages of execution [12-13] are likely to increase ecological validity and offer pathways into understudied repertoires.

The institutional case studies [18-19], in turn, highlight the utility of large-scale cooperation as well as the potential impact of targeted strategic planning. By aiming big and serving both as a proof of concept and a timely contribution to the cutting-edge, CompMusic demonstrates the potential for reverse-engineering research structure from concrete culturally-specific goals. And though more open-ended than CompMusic and focusing on symbolic rather than signal representations, the MEI Incubator’s acknowledgement of the importance of explicitly carving space for digging into underexplored territory while minimizing duplication of labor can be taken as another replicable best practice. The deliberate emphasis of both of these initiatives on maximal dissemination of their output reinforces the importance of democratization of knowledge for empowering practitioners who may not yet realize they have options beyond those they might consider by default.

Finally, the case studies involving IAM [20-21] suggest that repertoire driving methodology and not the converse is a successful means for optimizing the ecological validity of research. Another best practice they suggest is the incorporation of emic knowledge of diverse musical cultures

in the research process. One need not be a practitioner of any particular musical repertoire in order to be able to access pertinent ethnomusicological research on it. On the contrary, the majority of such scholarship is written by experts who understand that they have firsthand experience with a given repertoire or culture but their audience does not. MIR as a whole has been generally good about interfacing [22] with the cognitive sciences, the library sciences, and the computational disciplines pertinent to what it does but has a spottier track record historically with respect to engaging with relevant ethnomusicology beyond studies of recordings and transcriptions [23]. In our opinion, remedying this would represent an additional best practice for future work.

4. RECOMMENDATIONS

4.1 Against Solutionism

Our first recommendation is for MIR to avoid the solutionist impulse it might feel in response to critiques of certain quantizations. Indeed, attempting to quickly address perceived drawbacks of methodologies without fully probing the nature of the issues at hand “is likely to have unexpected consequences that could eventually cause more damage than the problems they seek to address” [24]. We have deliberately refrained from positing any alleged panacea to the prevalence of these quantizations firstly because none exist but more importantly because that would be antithetical to our broader aim of fostering a conversation involving as many stakeholders with differing perspectives as possible. If we are to make meaningful structural (i. e. rather than cosmetic) progress as a discipline with respect to diversity and inclusion, we need to be having substantive conversations about what we seek to change and why we seek to change it before we dive into well-intentioned attempts to implement such changes.

4.2 Against Status Quoism

Our second recommendation is for MIR to reject the inertia that has enable these quantizations to remain so central and unchallenged to its methodologies for so long. Questioning received wisdom is necessary and crucial to the long-term success of any research enterprise, and the ubiquity and utility of any quantization should not exempt it from critique. In many ways status quoism is the dual of solutionism, and to reject one while embracing the other would be hypocritical to say the least. Practically and institutionally speaking, we recognize that status quoism is a stance rewarded by the mechanisms of publication, recognition, career advancement, etc. in MIR and adjacent disciplines. It therefore merits conscious effort on our part as researchers to actively and explicitly posit the need for change while simultaneously acknowledging that changes must be thoughtfully and meticulously planned before they are implemented if they are to last.

4.3 Intellectual Honesty in Metatheoretical Discourse

Our third and final recommendation is for increased intellectual honesty in metatheoretical discourse. It is easy to perceive critiques of methodologies one uses or has used

as critiques of one’s entire research agenda and this can only have negative consequences. By displacing positive emotional investment in one’s research into negative responses to methodological criticism, productive conversations often are extinguished before they can begin and shouting into the void on all sides can be encouraged. We therefore call above all for dialogue, open-mindedness, and lucidity in terms of understanding the potential for a meaningful research agenda to be built on assumptions that one may not have previously questioned and for improvements to be made possible by exploring those assumptions. Reflection and even more importantly self-reflection are crucial to the present and future of inquiry, and quantization just as much as anything else ought to be the subject of such consideration.

5. ACKNOWLEDGEMENTS

The author is supported by a Legacy Fellowship from Florida State University and would like to thank the anonymous reviewers for their invaluable feedback on the paper.

6. REFERENCES

- [1] B. McFee et al., “librosa: Audio and music signal analysis in python,” in *Proc. of the 14th Python in Science Conf.*, Austin, TX, 2015, pp. 18–25.
- [2] D. Huron. *The Humdrum Toolkit: Software for music researchers*. Software package, 1993.
- [3] M. Cuthbert and C. Ariza, “music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data,” in *Proc. of the 11th Int. Society for Music Information Retrieval Conf.*, Utrecht, Netherlands, 2010, pp. 637–642.
- [4] A. Blake, “Computational analysis of quarter-tone compositions by Charles Ives and Ivan Wyschnegradsky,” M.A. thesis, College of the Arts, Kent State Univ., Kent, Ohio, 2020. Accessed on: July 26, 2021. [Online]. Available: http://rave.ohiolink.edu/etdc/view?acc_num=kent1588259095079479
- [5] M. Morris et al., “Views from inside and outside: integrating emic and etic culture and justice judgement,” *Academy of Management Review*, vol. 24, no. 4, pp. 781–796, October 1999.
- [6] H. Olson, *Music, Physics and Engineering*. New York: Dover Publications, 1967.
- [7] G. Boenn, “Rhythm quantization,” in *Computational Models of Rhythm and Meter*. Cham, Switzerland: Springer, 2018, pp. 147–173.
- [8] D. Huron, “Perceptual and cognitive applications in music information retrieval,” in *Proc. of the 1st Int. Society for Music Information Retrieval Conf.*, Plymouth, MA, 2000, pp. 83–92.
- [9] C.-E. Cella, “Music Information Retrieval and Contemporary Classical Music: A Successful Failure,” *Trans. of the Int. Society for Music Information Retrieval*, vol. 3, no. 1, pp. 126–136, September 2020.
- [10] G. Sala and F. Gobet, “Cognitive and academic benefits of music training with children: A multi-level meta-analysis,” *Memory & Cognition*, vol. 48, pp. 1429–1441, July 2020.
- [11] A. Bitton, P. Esling, and T. Harada, “Vector-quantized timbre representation,” 2007, *arXiv:2007.06349*.
- [12] C. Callender, “Continuous harmonic spaces,” *J. Music Theory*, vol. 51, no. 2, pp. 277–332, Fall 2007.
- [13] G. Wakefield, “Mathematical representation of joint time-chroma distributions,” in *Proc. of the SPIE Conf. on Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, Denver, CO, 1999, pp. 637–645.
- [14] D. Moelants, O. Cornelis, and M. Leman, “Exploring African tone scales,” in *Proc. of the 10th Int. Society for Music Information Retrieval Conf.*, Kobe, Japan, 2009, pp. 489–494.
- [15] J. Six and O. Cornelis, “Tarsos – A platform to explore pitch scales in non-Western and Western music,” in *Proc. of the 12th Int. Society for Music Information Retrieval Conf.*, Miami, FL, 2011, pp. 169–174.
- [16] M. Panteli, “Pitch Patterns of Cypriot Folk Music between Byzantine and Ottoman Influence,” M. S. thesis, Dept. of Information and Communication Technologies, Univ. Pompeu Fabra, Barcelona, Spain, 2011. Accessed on: July 26, 2021. [Online]. Available: <http://mtg.upf.edu/node/2333>
- [17] B. Bozkurt, “A System for Tuning Instruments Using Recorded Music Instead of Theory-Based Frequency Presets,” *Computer Music Journal*, vol. 36, no. 3, pp. 43–56, Fall 2012.
- [18] X. Serra, “A Multicultural Approach in Music Information Research,” in *Proc. of the 12th Int. Society for Music Information Retrieval Conf.*, Miami, FL, 2011, pp. 151–156.
- [19] A. Hankinson, P. Roland, and I. Fujinaga, “The Music Encoding Initiative as a Document-Encoding Framework,” in *Proc. of the 12th Int. Society for Music Information Retrieval Conf.*, Miami, FL, 2011, pp. 293–298.
- [20] H. G. Ranjani et al., “A compact pitch and time representation for melodic contours in Indian art music,” *J. Acoustical Society of America*, vol. 145, no. 1, pp. 597–603, January 2019.
- [21] V. S. Viraraghavan et al., “State-based transcription of components of Carnatic music,” in *Proc. of the 45th Int. Conf. on Acoustics, Speech, and Signal Processing*, Online, 2020, pp. 811–815.
- [22] W. de Haas and F. Wiering, “Hooked on Music Information Retrieval,” *Empirical Musicology Review*, vol. 5, no. 4, pp. 176–185, October 2010.

- [23] A. C. Gedik and A. Holzappel, “The Meaning of Music in Ethnomusicology and Music Information Retrieval: Obstacles Against Computational Ethnomusicology,” presented at the Conference on Interdisciplinary Musicology–CIM18, Poznań, Poland, 2018.
- [24] E. Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: PublicAffairs, 2013.