Assessing the QoME of NMP via Audio Analysis Tools

Konstantinos Tsioutas[a] and George Xylomenos[b]
Department of Informatics, Athens University of Economics and Business, Athens, Greece

Keywords: NMP, QoME, Tempo Analysis.

Abstract: Analyzing the Quality of Musicians' Experience (QoME) in Network Music Performance (NMP) typically involves having musicians perform NMP sessions and then assessing their experience via questionnaires. Such subjective studies produce results with wide variances, making the extraction of solid conclusions difficult. For this reason, we complemented a subjective study on the effects of delay in the QoME of NMP with an analysis of the audio captured during the study using automated tools. Specifically, we used signal processing techniques to analyze the captured audio, in order to detect tempo evolution during each performance and examine its correlation with delay. Our results indicate that musicians in real NMP settings are more tolerant to delay than previously thought, holding a steady tempo even with one way delays of 40 ms.

1 INTRODUCTION

The *Quality of Musicians' Experience* (QoME) in *Network Music Performance* (NMP), that is, the performance of music when musicians are connected over a network, is a complex function which depends on many variables, including audio delay and audio quality (Tsioutas et al., 2020). As in other human-to-human communication applications, NMP has strict delay requirements: while regular video conferencing can tolerate up to 100 ms of one way delay, in NMP delays of more than 20–30 ms are considered problematic; this delay limit is called the *Ensemble Performance Threshold* (EPT) (Schuett, 2002).

Such delays are hard to achieve over the wide area, even with very high speed networks, as increasing transmission speed only reduces transmission delay; propagation delay depends on the distance travelled, while queueing delay depends on router load and the number of hops in the network path. It is even harder to achieve such delays with residential DSL connections, as many musicians found out during the recent CoVid-19 pandemic.

Although numerous studies have investigated the effect of delay in the ability of musicians to synchronize, in an attempt to determine the tolerance of NMP to delay, there are many pieces missing to understand the big picture of QoME in NMP. Apparently, QoME is strongly connected with the perception of various audio phenomena and affected by audio features such as music rhythm, music tempo and audio spectral features (Rottondi et al., 2015). However, most studies with actual musicians had a small number of participants (8 – 12), hence it is hard to derive reliable conclusions from them.

As part of our work on the subjective evaluation of the effects of audio delay on the QoME of NMP, we have performed a large number of controlled experiments, where pairs of musicians play a musical piece under different conditions, completing a questionnaire at the end of each performance (Tsioutas et al., 2021). The analysis of these questionnaires reveals that not only different musicians perceive the same conditions in quite different ways, even the responses from the same musicians are not consistent with the underlying parameters; for example, their *assessment* of delay does not follow the *actual* delay in the experiments. The results from these subjective evaluations thus exhibit a very high variance, which makes drawing concrete conclusions harder.

For this reason, in this study we have chosen to follow a different path. Having recorded audio (and video) from our NMP experiments, we decided to use automated tools to extract information related to QoME, thus complementing the subjective questionnaires. Specifically, we used a signal analysis toolkit to determine how the performance tempo varies as delay is increased, a phenomenon often observed in previous studies where the tempo tended to slow down as delay grew; we applied this to actual music perfor-

[a] https://orcid.org/0000-0003-1136-3803
[b] https://orcid.org/0000-0003-2552-687X

Tsioutas, K. and Xylomenos, G.
Assessing the QoME of NMP via Audio Analysis Tools.
DOI: 10.5220/0010604100230030
In *Proceedings of the 18th International Conference on Signal Processing and Multimedia Applications (SIGMAP 2021)*, pages 23-30
ISBN: 978-989-758-525-8

mances however, not simple hand clapping synchronization exercises.

The outline of the rest of the paper is as follows. In Section 2, we briefly present related work on assessing the effects of delay on QoME and the performance tempo, in particular. Section 3 describes the technical setup of our experimental scenarios and the measurement process. In Section 4 we present a quantitative analysis of the NMP sessions that shows how tempo is affected by delay during the performances. We summarize our findings and discuss future work in Section 5.

2 RELATED WORK

A large amount of research touches upon QoME evaluation for NMP, looking at it from different perspectives. Most studies are subjective, that is, the participants perform one or more experimental sessions and then answer surveys evaluating their experience while audio delay or other variables are manipulated.

The pioneer study of (Schuett, 2002) investigated the effects of delay to synchronization, proposing the term *Ensemble Performance Threshold* (EPT) for the one way delay below which synchronization is possible, reporting that it lies between 20 and 30 ms. Another study reached the same conclusion, experimenting with musicians who clapped their hands, indicating that when the latency was below 11 ms, the tempo accelerated (Chafe and Gurevich, 2004).

Conversely, an experiment with two musicians who performed a clapping session without any external tempo reference (i.e., there was no starting metronome or countdown) found that the tempo of the hand claps slowed down as the delay was increased (Driessen et al., 2011).

In a study of eleven pairs of musically experienced subjects that attempted to synchronize their hand claps, the tempo was found to decrease more rapidly for higher delays, and the relation between delay and tempo slowdown was found to be approximately linear (Farner et al., 2009).

Finally, in an experiment where seventeen pairs of subjects performed hand clapping sessions under variable time delays, the authors reported that for delays shorter than 11.5 ms, 74% of the performances sped up, while at delays of 14 ms and above, 85% of the performances slowed down. No correlation with the starting tempo was found in the range sampled (Gurevich et al., 2004).

Turning to actual musical performances, in (Barbosa et al., 2005) four musicians played bass, percussion, piano and guitar. The authors found that regardless of the instrumental skills or the musical instrument, all musicians were able to tolerate more delay when slower tempos were used.

Another experiment investigated how the attack period of notes (that is, the time in which a note reaches it maximum amplitude) affects the tempo, using two musicians performing cello and violin. The analysis of the recordings showed that the tempo was generally higher with fast attack times than with slow attack times. In both cases, it decreased with increasing latency (Barbosa and Cordeiro, 2011).

Using a conductor is common in orchestral settings, therefore in (Olmos et al., 2009) six singers and a conductor performed over a network. For the most part, the singers that participated in the experiment managed to cope with the various delays. The singers mentioned that to a certain extent, they were able to establish emotional connections with each other.

In (Bartlette et al., 2006) two pairs of musicians were asked to performed two Mozart duets. Although the musicians chose different strategies to handle the latency, both duets were strongly affected by latency at and above 100 ms. At these levels, the musicians rated the performances as neither musical nor interactive, and they reported that they played as individuals and listened less and less to one another.

In a very comprehensive study of NMP (Rottondi et al., 2015), the authors asked seven pairs of musicians to participate in the experiments. Each repetition was characterized by different settings in terms of reference tempo, network latency, and jitter. The authors reported that the perceived delay was strongly affected by the timbral and rhythmic characteristics of the combination of instruments and parts. They also noted that the noisiness of the instrument has an impact on the perceived delay, for example, performers of percussive instruments reacted more strongly to delay changes. They concluded that the possibility of enjoying an NMP session is not only a function of delay, but also of the role and the timbral characteristics of the involved musical instruments, as well as the rhythmic complexity of the musical pieces performed.

In (Carôt et al., 2009) five professional drummers performed with five professional bass players. The authors reported that the overall delay thresholds ranges between a minimal delay of 5 ms and a maximal delay of 65 ms. They also noted that the players did not exhibit a common latency acceptance value, indicating that tolerance to delay is a subjective issue.

Finally, in (Delle Monache et al., 2018) ten volunteers participated in pairs, performing mandolin, accordion, guitar, percussion, harp, flute, alto sax. Delay had a negative effect to the involvement of the musicians with the process. The authors also reported

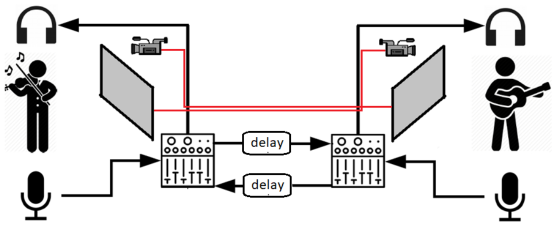

Figure 1: Experimental Setup.

that a general distress was caused by latency, but a willingness to find ways to cope with it also emerged from the answers.

To summarize, the hand clapping studies indicate that synchronization is hard when delay exceeds 30 ms, with the tempo slowing down as delay rises beyond this threshold. However, studies with real musical performances show quite diverse results, indicating that in real NMP sessions musicians do adapt to higher delays, often by slowing down their tempo, depending on the type of music performed and the instruments used.

3 EXPERIMENTAL SETUP

For our experiments, we used two visually and aurally isolated rooms on the same floor of the main AUEB building. Musicians performed with their counterparts in separate rooms, while listening to them through headphones and seeing them through a 32” video monitor.

As shown in Figure 1, an eight channel mixing console was used in each room for the necessary audio routing, monitoring and recording. Audio was captured by condenser microphones and closed type headphones were used by the musicians to listen to each other. A video camera was capturing and sending a composite video signal through the existing network cabling to the video monitor of the other room (red lines in the figure). The network cables were patched directly to each other, without passing through any network equipment; we basically used one pair of the UTP cables to transmit the composite video signal in analog mode. We used composite video in order to achieve the lowest possible visual delay between musicians; with the analog signal we did not have to wait for entire frames to be captured before transmission and received before display.

We experimentally measured the round trip delay by placing a smartphone with a running chronometer in front of the camera in one room, and turning the video camera to the video monitor in the other room, essentially reflecting the transmitted image back to the first room. We then recorded with another smartphone’s camera both the chronometer and its reflected image, and analyzed the video in a video editor, finding out that the round trip delay was 30 ms, therefore the one way delay was 15 ms.

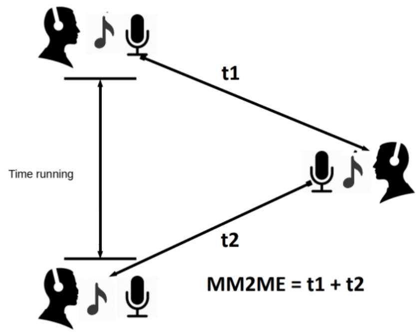

Figure 2: My Mouth to My Ear delay.

The two mixing consoles were also connected through the existing network cabling, using direct cable patching, hence the audio signal was also transmitted in analog form from one room to the other. The reason for connecting them directly was to be able to achieve perfectly fixed audio delays even below 10 ms, which is impossible when computers and network devices intervene in the signal path; the propagation delays were less than 1 ms due to the small cabling distance between the rooms. To manipulate audio delay we used AD-340 audio delay boxes by Audio Research, via which we were able to set delay in each direction to the desired value. We attached a PC with a Motu 828X external audio interface (not shown) to the auxiliary output of the mixing console to record each session for later analysis, without introducing delays in the signal path.

Unlike most NMP studies which use *Mouth to Ear* (M2E) delay, which is the end-to-end delay between the microphone at one end and the speaker at the other end, in our work we use the *My Mouth to My Ear* (MM2ME) delay proposed in (Tsioutas et al., 2020). As shown in Figure 2, MM2ME is the two-way counterpart to M2E, over which it has three advantages. First, when musicians play together, each musician plays one note and unconsciously expects to listen to the other musicians’ note to play his next one, and so on. Second, measuring MM2ME delay accurately is much easier than measuring the M2E delay, as it can be done at one endpoint, by simply reflecting the transmitted sound at the other endpoint; in contrast, M2E needs to be measured at both endpoints, thus requiring perfectly synchronized clocks, something very hard to achieve (Carôt et al., 2020). Third, MM2ME takes into account the possible asymmetry between the two directions of a connection, which is

Table 1: Instruments played by the musicians.

Duet 1	Duet 2	Duet 3	Duet 4	Duet 5	Duet 6	Duet 7	Duet 8	Duet 9	Duet 10	Duet 11
Folk	Folk	Rock	Rock	Funk	Funk	Rock	Rock	Classic	Folk	Folk
Piano	Piano	El Gtr	El bass	Organ	El bass	El bass	El Gtr	Flute	Ac Gtr	Laoud
Santuri	Oud	El Gtr	El Gtr	El Gtr	Perc	Ac Gtr	Violin	Violin	Buzuki	Violin

Table 2: MM2ME delays in order of use.

Repetition	1	2	3	4	5	6	7	8	9	10
MM2ME delay (ms)	10	25	35	30	20	0	40	60	80	120

the rule with residential DSL endpoints.

We conducted experimental sessions with 22 individual musicians (11 duets); to the best of our knowledge, this is the largest NMP study with actual musical performances (as opposed to hand claps) to date. The musicians performed with a variety of instruments, including piano, acoustic guitar, electric guitar, electric bass, violin and flute, as well as traditional instruments including the lute, toumberleki, santuri and oud, in a musical style of the choice. Table 1 shows the musical style and the instruments for each duet. Each pair of musicians played a musical part of their choice, with a duration of up to 60 sec, following their own tempo and repeating it ten (10) times, using a different MM2ME delay setting for each repetition. We kept the duration low, to avoid tiring the musicians, since they had to repeat the piece multiple times and subjectively assess their experience in the end. Table 2 shows the delays used and the sequence with which they were applied to each repetition; half of that delay was set in each direction via the audio delay device. No metronome or other synchronization aids were used.

Note that a delay of 0 ms is unnaturally low: two musicians 2 meters apart from each other experience a one way delay of 5.83 ms (11.66 ms MM2ME) based on the speed with which sound propagates through the atmosphere; a duet in the same room would therefore experience a natural MM2ME delay of 10 to 20 ms, depending on their positions in the room.

After the end of each repetition, each musician was asked to answer an electronic questionnaire on a tablet (see (Tsioutas et al., 2021) for details). Musicians were not informed about which variable was manipulated each time, or about the purpose of the experiment, and we randomly set the order in which the audio delay values were set for each repetition. The main goal was to conduct an experiment that would allow us to evaluate multiple variables without bias or noise in the answers. The audio (and video) of each performance was recorded, and was used for the analysis discussed in the following sections.

4 TEMPO ANALYSIS

As mentioned in Section 2, experiments where participants tried to synchronize their hand claps over the network have indicated that as the delay between the endpoints grows, the participants compensate by slowing down their tempo. Since hand claps have a simple audio signature, it is easy to note such slowdowns by simply looking at the waveform of the recordings. The same observation was made in some experiments with real musicians, even though the exact correlation between the delay and the tempo was harder to quantify with the more complex sonic imprint of actual musical pieces.

Although our original goal when setting up an NMP study was to perform a subjective evaluation of the effect of delay on QoME via questionnaires, we had recorded the audio tracks of each session for later analysis. This allowed us to assess whether the tempo does indeed grow as delay is increased in our more realistic setup, with actual musicians playing real instruments and real musical pieces. Of course, since each duet selected their own musical piece and tempo, we had to recover all relevant information from the actual recordings. That is, unlike in (Rottondi et al., 2015), we did not know what the intended tempo of each performance was.

To this end, we analysed the audio recordings using the MIRToolbox (Lartillot et al., 2013). To determine the tempo at a period of time, we start with the *event density*, which estimates the average number of note onsets per second as follows:

$$E = \frac{O}{T} \quad (1)$$

where E is event density, O is the number of note onsets and T is the duration of the musical piece. The MIRToolbox estimates how the music tempo, measured in *Beats per Minute* (BPM), varies over time, by detecting the note onsets via signal processing of the audio. The analysis is not perfect, as it depends on each instrument's sonic signature and manner of playing, but it is revealing, especially for instruments with

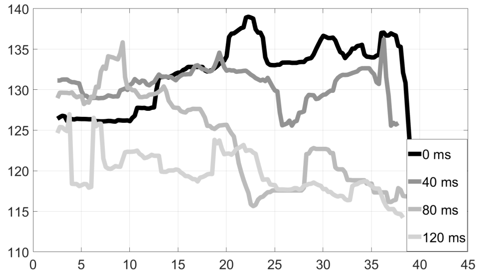

Figure 3: Tempo variation over time for various delay values: Duet 1, Piano-Rhythm-Folk.

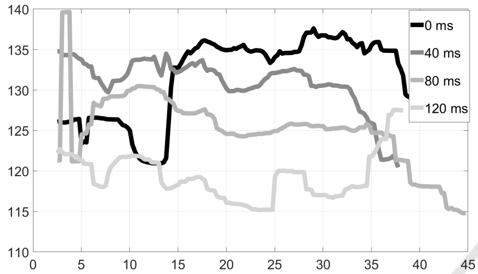

Figure 4: Tempo variation over time for various delay values: Duet 1, Santouri-Solo-Folk.

very clear sonic signatures, for example percussion, or with performances where the instrument plays a rhythmic pattern. We performed this analysis for the audio recording of each side of an NMP session.

These results are not easily amenable to numerical summarization, since musicians adapt their playing over time as they listen to each other; as a result, each performance leaves a unique time-varying imprint. However, when presented visually, they show interesting trends. The following figures show how the tempo (in BPM) varies over time (in seconds) for different musical instruments; each figure shows one such curve for each delay value, corresponding to one performance by a single musician, with progressively lighter curves corresponding to increasing MM2ME delays. To reduce visual clutter, we only show results at 40 ms intervals, that is, with 0, 40, 80 and 120 ms MM2ME delays.

Figures 3 and 4 show the delay variation for each instrument of duet 1. We can see that with a delay of 0 ms, which is unnaturally low, as explained above, both musicians actually speed up their tempo in the first part of the performance, as reported in previous studies. As the delay grows, the tempo slows down, but the musicians have a hard time keeping a steady tempo at all delay values, as evidenced from the ups and downs in the curves.

While in duet 1 the musicians have trouble keeping a steady rhythm, in duet 2, Figures 5 and 6 show a different situation: the instrument playing the rhythm part, in this case the piano, is visibly affected by delay, since as the delay grows, the tempo drops; however, the tempo is steady in all but the highest delay value. The instrument playing the solo part however, in this case the oud, shows larger tempo variations, even though the tempo does generally drop with growing delay. Of course, due to the method we are using to detect the tempo (note onsets), solo parts where mu-

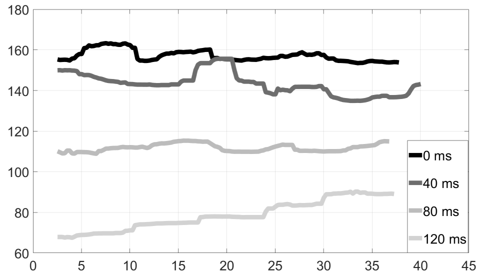

Figure 5: Tempo variation over time for various delay values: Duet 2, Piano-Rhythm-Folk.

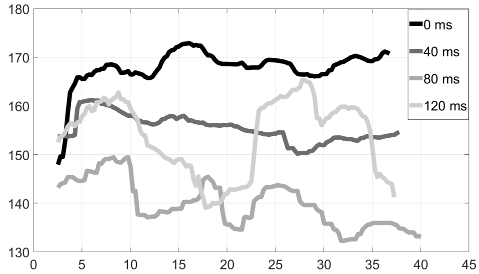

Figure 6: Tempo variation over time for various delay values: Duet 2, Oud-Solo-Folk.

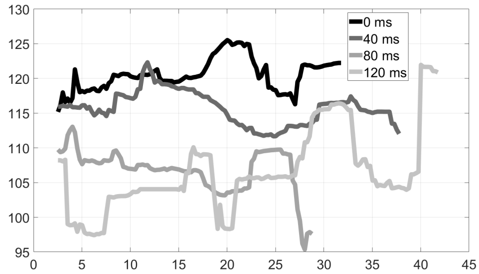

Figure 7: Tempo variation over time for various delay values: Duet 3, Electric Guitar-Solo-Rock.

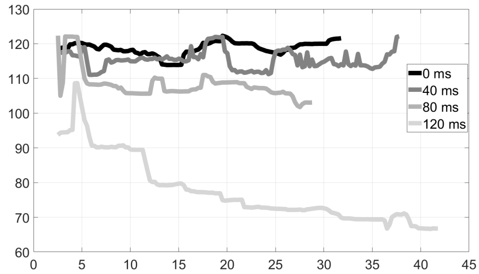

Figure 8: Tempo variation over time for various delay values: Duet 3, Electric Guitar-Solo-Rock.

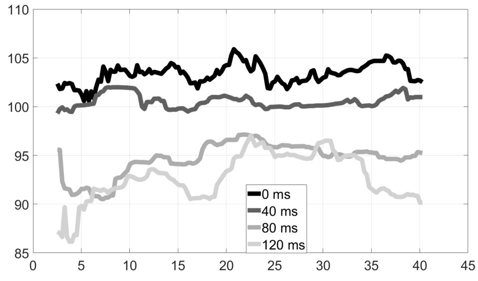

Figure 9: Tempo variation over time for various delay values: Duet 5, Organ-Solo-Funk.

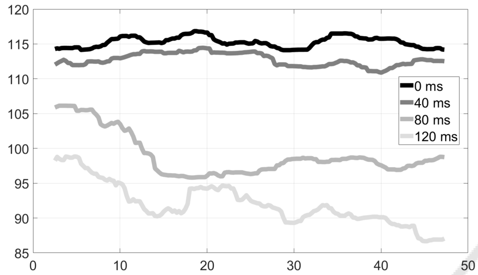

Figure 10: Tempo variation over time for various delay values: Duet 6, Percussion-Rhythm-Rock.

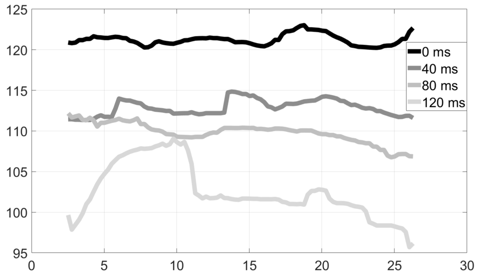

Figure 12: Tempo variation over time for various delay values: Duet 7, Acoustic Guitar-Solo-Rock.

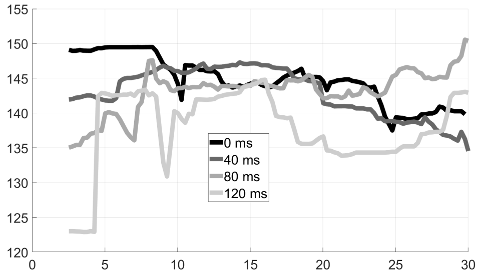

Figure 13: Tempo variation over time for various delay values: Duet 8, Guitar-Rhythm-Rock.

sicians play more freely and improvise are harder to characterize precisely in terms of tempo.

In duet 3 where both musicians have a solo role, we can see in Figures 7 and 8 that both exhibit tempo variations, however, the musicians do manage to keep a relatively steady tempo, except for the highest delay value of 120 ms. Again, the tempo tends to drop with higher delays. Note that since both musicians have improvisational roles, they end their performance at different time points for each delay value - they do finish at the same time, though.

The difficulty of keeping a steady tempo at higher delays is also apparent in Figure 9 which shows the solo instrument of duet 5 (organ); again, tempo drops with higher delays, and has wild variations at a delay of 120 ms. With the percussion instrument of duet 6, arguably the most rhythmic of instruments and the easiest in terms of automated tempo detection, as shown in Figure 10, the beat is noticeably slower for higher delays, and hard to keep steady when delay reaches 120 ms.

There are also cases where both sides of a duet can keep the same rhythm, as with duet 7, shown in Figures 11 and 12: the rhythm is steady with delays of up to 80 ms; there is a very slight reduction in tempo from 40 to 80 ms, but at 120 ms the tempo either slows down continuously or varies wildly.

Duet 8 is unusual, in that the rhythm instrument (guitar), shown in Figure 13 has an unsteady tempo, while the solo instrument (violin), shown in Figure 14 has a very steady tempo, despite the visible slowdown at delays of 80 and 120 ms. The reason for this is the very different expertise levels of the musicians: the violinist was a 45 year old professional musician, while the guitarist was a 23 year old amateur one. Hence the violinist's solo tempo was found to be more stable than the guitarist's, even though it was the guitarist who was supposed to keep a stable

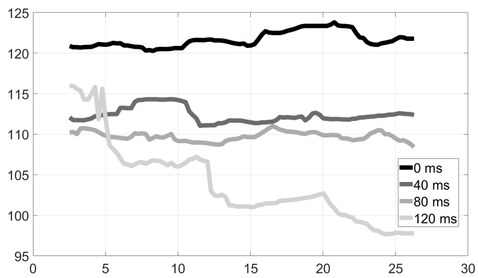

Figure 11: Tempo variation over time for various delay values: Duet 7, Bass-Rhythm-Rock.

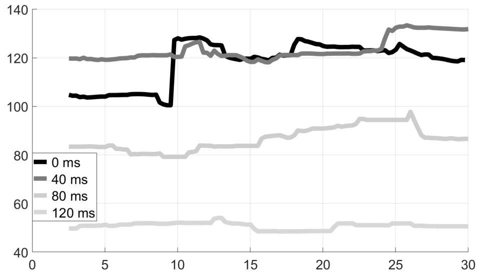

Figure 14: Tempo variation over time for various delay values: Duet 8, Violin-Solo-Rock.

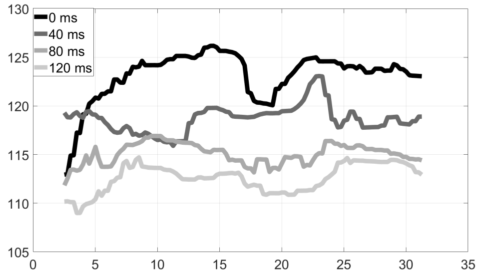

Figure 15: Tempo variation over time for various delay values: Duet 10, Guitar-Rhythm-Folk.

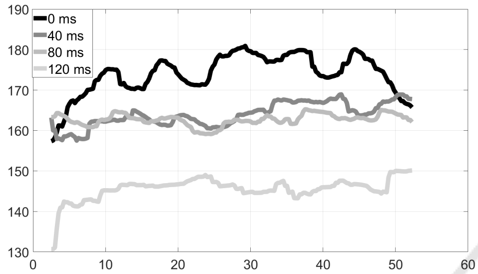

Figure 16: Tempo variation over time for various delay values: Duet 11, Lute-Rhythm-Folk.

rhythm with the guitar. This is an indication that more experienced musicians may manage to partially compensate for delay by adapting their performance.

Finally, the rhythm instruments of duet 10 and duet 11, shown in Figures 15 and 16 further verify the previous observations of tempo speedup at the unusually low delay of 0 ms, good tempo stability at 40 and 80 ms, albeit at a slight reduction of tempo at 80 ms, and higher variations at 120 ms.

From these figures, we can make the following general observations:

1. At the (unnaturally) low delay of 0 ms, musicians tend to speed up their tempo in the beginning of the session.
2. As delays rise beyond 40 ms, musicians adapt by slowing down the tempo of their performance.
3. Instruments performing rhythmic parts are more clearly affected by delay, as shown by their more visibly delineated curves.
4. Percussion instruments, which generally have a rhythmic role, are the most sensitive to delay.
5. In most cases, musicians manage to keep a steady tempo at delays of up to 80 ms.
6. At a delay of 120 ms performances break down, exhibiting either continuously slowing or wildly varying tempos.

These observations verify findings from past work that musicians who perform percussive instruments suffer more from delay than others. Indeed, the hand clap experiments, where the rhythmic patterns are very clear, fall in the same category. Of course, these instruments, with their easy to detect sonic signatures and their clear temporal patterns, are ideal for this type of analysis. We can further observe that this is true for instruments having a rhythmic role in a duet. Although solo instruments seem to follow more irregular tempos, we must keep in mind that this may be an artefact of our audio analysis which relies on a steady production of note onsets; with improvisational parts, performers are expected to more often deviate from the base rhythmic pattern.

The most interesting observation however is that the limits to tolerance can vary considerably; most musicians could achieve a stable tempo at MM2ME delays of 80 ms, corresponding to an one way delay of 40 rather than 20–30 ms, higher than what was previously considered the limit to synchronization, even though this may come at the the cost of a minor slow down in the performing tempo.

Finally, it should also be noted that we performed an ANOVA analysis for repeated measures of the average tempo scores for each session and for delays of 0, 40, 80 and 120 ms (MM2ME) and the p value was computed equal to 0.007 ($p < 0.05$). This indicates a strong statistical significance in the delay/tempo relationship, that is, the calculated tempos were found to be statistically correlated with the delay values, that is, higher delays did lead to slower tempos.

5 CONCLUSIONS AND FUTURE WORK

We conducted a set of NMP experiments, where the delay between a pair of musicians was varied in a controlled manner for each session, with the audio and video from the sessions being recorded for later analysis. In the experiments reported in this paper, 22 musicians participated as pairs, playing a diverse set of musical instruments in different musical styles.

The analysis performed on the recorded audio indicates that musicians tend to slow down their tempo as delays grow, an effect made very clear with percussive instruments and quite clear with instruments playing rhythmic parts. However, in most cases they can synchronize and maintain a stable tempo with MM2ME delays of up to 80 ms (equivalent to 40 ms one way delays), indicating that the delay tolerance of actual musicians performing in NMP scenarios is higher than previously thought, that is, the EPT is closer to 40 rather than 20–30 ms. Indeed, musicians, especially more experience ones, try to adapt to higher delays by slowing down their tempo. This conclusion is also supported by the analysis of the QoME questionnaires, reported in (Tsioutas et al., 2021).

Our work continues with a deeper analysis of the audio data collected, focusing on issues such as the dependence of tempo variations on other factors, such as the style of music performed and the target tempo of each piece. By grouping similar performances together, we hope to be able to derive quantitative expressions of the relationship between delay and tempo, depending on those factors.

Similarly, we are currently analyzing the video data gathered during the sessions via machine learning techniques, and specifically facial emotion recognition, in an attempt to quantify the emotional response of the participants in an NMP session to delay.

ACKNOWLEDGEMENTS

We would like to thank all the participating musicians for their patience during the experiments, as well as the fellows who helped with setting up and carrying out the experiments. We would also like to thank our colleague Dr. Ioannis Doumanis (Lecturer at the University of Central Lancashire) for help in designing the experiments.

REFERENCES

Barbosa, A., Cardoso, J., and Geiger, G. (2005). Network latency adaptive tempo in the public sound objects system. In *International Conference on New Interfaces for Musical Expression*, pages 184–187.

Barbosa, A. and Cordeiro, J. (2011). The influence of perceptual attack times in networked music performance. *Audio Engineering Society International Conference*.

Bartlette, C., Headlam, D., Bocko, M., and Velikic, G. (2006). Effect of network latency on interactive musical performance. *Music Perception: An Interdisciplinary Journal*, 24(1):49–62.

Carôt, A., Hoene, C., Busse, H., and Kuhr, C. (2020). Results of the fast-music project – five contributions to the domain of distributed music. *IEEE Access*, 8:47925–47951.

Carôt, A., Werner, C., and Fischinger, T. (2009). Towards a comprehensive cognitive analysis of delay-influenced rhythmical interaction. In *International Computer Music Conference*.

Chafe, C. and Gurevich, M. (2004). Network time delay and ensemble accuracy: Effects of latency, asymmetry. In *Audio Engineering Society Convention 117*.

Delle Monache, S., Buccoli, M., Comanducci, L., Sarti, A., Cospito, G., Pietrocola, E., and Berbenni, F. (2018). Time is not on my side: Network latency, presence and performance in remote music interaction. In *Colloquium on Musical Informatics (CIM)*, pages 20–23.

Driessen, P. F., Darcie, T. E., and Pillay, B. (2011). The effects of network delay on tempo in musical performance. *Computer Music Journal*, 35(1):76–89.

Farner, S., Solvang, A., Sæbø, A., and Svensson, U. P. (2009). Ensemble hand-clapping experiments under the influence of delay and various acoustic environments. *Journal of the Audio Engineering Society*, 57(12):1028–1041.

Gurevich, M., Chafe, C., Leslie, G., and Tyan, S. (2004). Simulation of networked ensemble performance with varying time delays: Characterization of ensemble accuracy. In *International Computer Music Conference*.

Lartillot, O., Cereghetti, D., Eliard, K., Trost, W. J., Rappaz, M.-A., and Grandjean, D. (2013). Estimating tempo and metrical features by tracking the whole metrical hierarchy. In *The 3rd International Conference on Music & Emotion*, Jyväskylä, Finland. University of Jyväskylä, Department of Music.

Olmos, A., Brulé, M., Bouillot, N., Benovoy, M., Blum, J., Sun, H., Lund, N. W., and Cooperstock, J. R. (2009). Exploring the role of latency and orchestra placement on the networked performance of a distributed opera. In *12th Annual International Workshop on Presence*, pages 1–9, USA. ISRP.

Rottondi, C., Buccoli, M., Zanoni, M., Garao, D., Verticale, G., and Sarti, A. (2015). Feature-based analysis of the effects of packet delay on networked musical interactions. *Journal of the Audio Engineering Society*, 63:864–875.

Schuett, N. (2002). The effects of latency on ensemble performance. Bachelor Thesis, CCRMA Department of Music, Stanford University.

Tsioutas, K., Xylomenos, G., and Doumanis, I. (2021). An empirical evaluation of QoME for NMP. In *IFIP International Conference on New Technologies, Mobility and Security*.

Tsioutas, K., Xylomenos, G., Doumanis, I., and Angelou, C. (2020). Quality of musicians' experience in network music performance: A subjective evaluation. In *Audio Engineering Society Convention 148*.