Time Delay Stability Analysis of Pairwise Interactions Amongst Ensemble-Listener RR Intervals and Expressive Music Features

Mateusz Soliński[1], Courtney N Reed[1], Elaine Chew[1]

[1]Department of Engineering and School of Biomedical Engineering & Imaging Sciences
King's College London, London, UK

Abstract

Time Delay Stability (TDS) can reveal physiological function and states in networked organs. Here, we introduce a novel application of TDS to a musical setting to study interactions between RR intervals of ensemble musicians and a listener, and music properties.

Three musicians performed a movement from Schubert's Trio Op. 100 nine times in the company of one listener. Their RR intervals were collected during baseline (5 min, silence) and performances (~10 min each). Loudness and tempo were extracted from recorded music audio. Regions of stable optimal time delay were identified during baseline and music, shuffled data, and data pairs from incongruent recordings. Bootstrapping was employed to obtain mean TDS probabilities (calculated based on all performances).

A significant difference in mean TDS probability between music and baseline was observed for all musician pairs ($p < .001$) and for cello-listener ($p = .025$). A significant decrease in mean TDS probability was observed for piano-violin ($p < .001$), violin-tempo ($p = .045$), and cello-tempo ($p < .001$) for mixed pairs. The highest inter-musician TDS probabilities were observed in musically tense sections. This framework offers a promising way to track dynamic RR interval interactions between people engaged in a shared activity, and, here, between the people and music properties.

1. Introduction

The collaborative performance of scripted music live generates a network of mutual interactions between musicians and listeners, as well as individual physiological responses, moderated by instructions in the score. For listeners, prior studies have shown changes in autonomic nervous system (ANS) activity, observable in the heart rate (HR) and heart rate variability (HRV), as effects of variations in the rhythm and tempo [1] and specific music structures, such as progressive crescendos and rhythmic phrases [2]. In groups of professional musicians (soloists or orchestral), these physiological signals were found to change with music difficulty [3][4], familiarity with the music, tempo, and the time of day [5]. These studies determined average or individual changes in physiological signals to music-related factors. However, there is a lack of studies about *mutual* changes and the network of interactions between musicians, listeners, and music features.

We applied the time delay stability (TDS) method, based on the framework developed in [6], for monitoring these couplings in a musical setting. We analysed the stability of mutual changes (defined as approximately constant delay of the maximum cross-correlation between two time series) in RR interval series obtained from a musical trio, one listener, and music tempo and loudness during the performance of a classical music piece. TDS analysis was previously used to reveal changes in networks within different physiological systems during sleep [6]. The main goals of this study were to apply the TDS method to a system comprised of musical ensemble, a listener, and the performed music, and discover the physiological coupling measured by RR interval series and music features.

2. Methodology

2.1. Data collection

Data were collected during nine performances on five different days between December 2022 and March 2023. A trio of professional musicians (violinist, cellist, and pianist, each with over 20 years of performance experience) performed Schubert's Trio No. 2, Op. 100, *andante con moto* (referred to as "the Schubert"). The Schubert was selected for its clear musical structures. ECG signals were measured with a Polar H10 (Polar Electro Oy, Kempele, Finland) HR monitor with a sampling frequency of 130 Hz. The unit's firmware automatically detected the QRS complexes and calculated the RR interval series. Each signal was manually reviewed to eliminate artefacts and premature ventricular contractions. We also recorded audio signals with a Zoom H5 (Zoom, Tokyo, Japan) handheld recorder, placed ~2 metres from the musicians. At the be-

 ISSN: 2325-887X DOI: 10.22489/CinC.2023.367

ginning of each session, a 5-minute baseline measurement was taken in silence before the performance. On the days when the piece was performed twice or three times, at least a 5-minute break was taken between each play-through.

2.2. Data Processing

The recorded audio signal was used to extract expressive musical features and for signal synchronisation. We manually annotated the eighth note pulse in the recording using Sonic Visualizer [7] and aligned it using the the CHARM Mazurka Project's TapSnap algorithm[1]. Tempo was calculated as beats per minute (BPM). Perceptual loudness in sones was calculated using the `ma_sone` function of the Music Analysis Matlab Toolbox [8]. Due to intentional and human variability, performances had different expressions and duration. Musical features and corresponding RR interval time series were aligned across the nine performances for analysis by converting them from (real) time to the score-time domain [9]: We used the timestamps of the previously annotated musical beats (eighth notes). We interpolated all signals to a high sampling frequency (1000 Hz) and selected only the samples at each annotated timestamp. The resulting signals, therefore, had lengths equal to the number of eighth note timestamps (848) in the piece. All signals were then filtered using a 3rd-order low-pass Butterworth filter with normalised critical frequency $Wn = 0.125$ using the `butter` function in Python.

2.3. Time Delay Stability

We implemented the TDS algorithm following Bashan et al. [6], dividing all signals into a total of N_L overlapping segments $v = 1, \ldots, N_L$, where N is the length of the signal (848 eighth notes, the score-time unit), L the length of each segment (30 eighth notes), and hop size 10 eighth notes. Each segment was approximately half that reported in Bashan et al., as we used comparatively shorter signals, and the structure of the Schubert changed more rapidly than sleep phases overnight. The signals were normalised by the mean and standard deviation in each segment to reduce constant trends. Then, we calculated a cross-correlation for all possible pairs of signals for that segment using the `correlate` function from the Python `scipy` library. We determined a score-time delay, τ_v, that maximises the absolute value of the cross-correlation function for segment v. We consider two signals in segment v linked if the value τ_v is approximately constant in time. Following Bashan et al., period v is labelled as stable when τ_v changes no more than ± 1, in at least four out of five consecutive segments, i.e., $\sum_{s=v-2}^{v+2} I_s \geq 4$, where $I_s = 1$ if $|\tau_{s+1} - \tau_s| \leq 1$, 0 otherwise. All nine performances were analysed to estimate the TDS probability p_v in each segment (the number of stable segments v from each performance divided by the number of all performances).

[1] http://mazurka.org.uk/cgi-bin/tapsnap

2.4. Statistical Analysis

We compared TDS probabilities for all signal pairs (violin, cello, piano, and listener RR and music loudness and tempo) between music period and baseline and two types of surrogate data: 1) all samples in the signals were shuffled, and 2) signals from different performances (randomly selected). The mean value and 95% confidence intervals (CIs) of the probabilities for these sets were estimated using bootstrapping ($N = 1000$ iterations). For each signal pair, the method was: 1) sample with replacement 9 performances for TDS probability calculation (5 for the baseline); 2) calculate the mean TDS probability, calculated in segments for the current set of performances during music, baseline, and for the surrogate data; 3) repeat (1) and (2) N times and aggregate mean values to create distributions; then, 4) calculate the lower and upper 95% CI for the 2.5^{th} and 97.5^{th} percentile of the distribution of the means.

Bootstrapping was also used to calculate p-values for the null hypothesis: no difference between the means from the baseline vs. music, and surrogate vs. music (2-sided test). First, the baseline distribution, $X = X_b$ (or surrogate, $X = X_s$), was shifted by subtracting the difference between the mean of the baseline, $\mu = \mu_b$ (or surrogate, $\mu = \mu_s$) and music distributions, μ_m, from every element: $X'(i) = X(i) - |\mu_m - \mu_b|$. The p-value is the probability of observing the absolute values of shifted distribution that are larger than the μ_m. This is equivalent to use the signum, sgn, function: $p = 1/N\Sigma_{i=0}^{N}\ sgn(|X'(i)| - \mu_m)$. Significant p-values $< .05$ were those where the means of the distributions were significantly different.

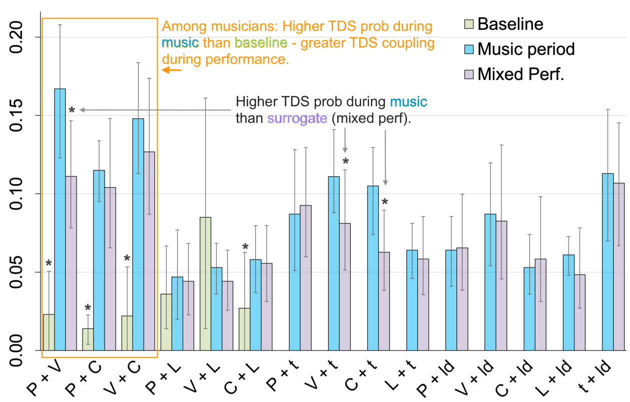

Figure 1: Mean TDS probability (y-axis, 95%CI) by signal pair, P: piano, V: violin, C: cello, L: listener, t: tempo, ld: loudness (*$p < .05$ difference with music period).

3. Results

Mean TDS probabilities for all signal pairs and for the surrogate data (mixing performances) are shown in Figure 1. The TDS probability in all segments for the pairs, considering musicians' interactions, is shown in Figure 2. We observed greater mean TDS probability for the music period compared to the baseline for the violin + piano ($p < .001$), cello + piano ($p < .001$), cello + violin ($p < .001$), and listener + cello ($p = .025$). Compared to surrogate data (mixing performances), the mean probability was greater for the piano + violin ($p < .001$), violin + tempo ($p = .045$), and cello + tempo ($p < .001$). The mean TDS probability for shuffled data was approximately constant for all pairs, equalling $0.038 - 0.041$ (lower and upper bounds of the 95%CI: $0.0185 - 0.0214$ and $0.0627 - 0.0684$). These values were significantly lower than those from the original data for all pairs, except the listener + any musician and cello + loudness.

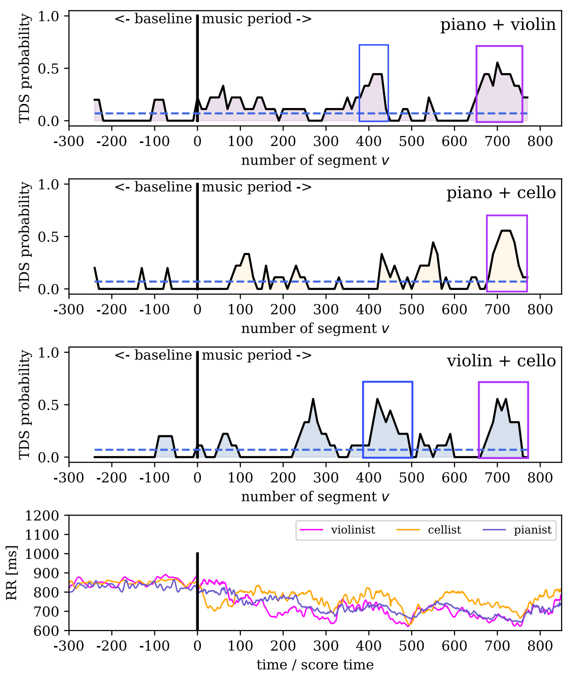

Figure 2: Segment-wise inter-musician TDS probability and RR interval series averaged over nine performances. Dashed blue lines indicate upper 95%CI obtained for shuffled data (random effects threshold). Purple and blue rectangles mark the areas with highest TDS probability.

4. Discussion

The results show larger average TDS probabilities while playing music in comparison to the baseline (Figure 1), which suggests an increase in physiological coupling during that study phase. That effect was observed for all musician pairs and also between the cellist and listener During baseline, all participants were sitting in the same room in silence; we assume any TDS to be due to random effects.

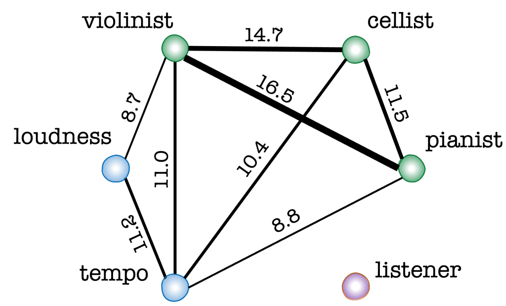

Figure 3: Network connectivity during the music period between pairs of signals. Signals were considered connected if the mean value of TDS probability was larger than 7% (larger than the upper bound of the shuffled data probabilities). Thicker lines indicates reflects greater number of performances where this connection was observed.

In turn, during music, the probability of TDS increases at certain moments of the music piece (Figure 2). The increase in TDS probability could be caused by specific expressive music structures played by the musicians. The highest TDS probabilities for all musician pairs are observed between 650-750 score-time units (see purple rectangles in Figure 2) during the final intense climactic section, where the musicians' RR intervals simultaneously decrease, just before the music dies down for the end. A peak in TDS probabilities is also observed for the piano + violin and violin + cello pairs between 400-450 score-time units (blue rectangles in Figure 2) in another tense section where the instruments jointly but quietly crescendo (build up in sound) then die down to transition to a new key.

The comparison of the TDS probability between the music period and surrogate data showed another aspect of temporal coupling: Having pairs constituted from different performances eliminates the influence of playing Schubert's piece jointly vs. independently. We observed a significant decrease in TDS probability for the mixed-performance piano + violin, tempo + violin and tempo + cello pairs. These results suggest that TDS analysis provides physiologically relevant information regarding the interaction between these ensemble pairs. In the music period, we found the violin and piano were the most coupled, suggesting that they or (gesturally) the parts they play were more in sync. We also observed that tempo seems to be more coupled with RR intervals than loudness (for violin and cello). Wright and Palmer [5] suggested tempo as a factor that increases the predictability of cardiac dynamics.

We considered the upper bound of the shuffled data probabilities, rounded up to 0.07 (dotted lines in Figure 2), as the threshold for the coupling to be beyond a random effect. Using that value, we created a network of connections between signals, depicting their level of interaction during the music piece (Figure 3). Most values obtained from the baseline and pairs with the listener (all but the violin + listener) are below this threshold. In comparison to the performer pair probabilities, TDS values for listener pairings are low. The listener is not active in the performance process; physiological coupling with the musicians might be weaker. This could be investigated with more listeners to explore group dynamics during music listening.

It is difficult to determine what effect strong coupling has on the performance. It could be related to the mechanics of musical performance or music structures themselves, rather than the performance quality. However, the results show TDS to be a promising direction for further studies. The results are limited by the study of only one music piece, albeit performed 9 times, by one music ensemble. The analysis should be repeated with other music pieces, musicians, and listeners to further evaluate individual responses and music structures that contribute to TDS. As well, the study examined a rehearsal context; a concert setting's likely impact on TDS should be investigated.

5. Conclusions

In this study, we applied TDS analysis to music data and cardiac signals obtained from musicians while performing and a listener. In contrast to the original TDS application, which focused on a single body during sleep, we studied a system comprising a group of people and music features. To the best of our knowledge, we have presented the first use of this method not only across people and music, but also in a musical environment. Our results suggest that some structures and characteristic sections of the music piece might increase the mutual coupling between musicians and between each musician and musical tempo. This approach allowed us to examine physiological interactions in a music setting. Future extensions include using network analysis between musicians to analyse signals from musicians with cardiovascular diseases by studying their TDS probabilities in comparison to healthy individuals.

Acknowledgments

This result is part of the COSMOS and HEART.FM projects, which have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement Nos. 788960 and 957532). Charles Picasso developed the code for data collection; Ian Pressland and Hilary Sturt were collaborative artists on the performances.

References

[1] Bernardi L. Cardiovascular, Cerebrovascular, and Respiratory Changes Induced by Different Types of Music in Musicians and Non-Musicians: the Importance of Silence. Heart 2005;92(4):445–452.

[2] Bernardi L, Porta C, Casucci G, Balsamo R, Bernardi NF, Fogari R, Sleight P. Dynamic Interactions Between Musical, Cardiovascular, and Cerebral Rhythms in Humans. Circulation 2009;119(25):3171–3180.

[3] Mulcahy D, Keegan J, Fingret A, Wright C, Park A, Sparrow J, Curcher D, Fox KM. Circadian Variation of Heart Rate is Affected by Environment: a Study of Continuous Electrocardiographic Monitoring in Members of a Symphony Orchestra. Heart 1990;64(6):388–392.

[4] Williamon A, Aufegger L, Wasley D, Looney D, Mandic DP. Complexity of Physiological Responses Decreases in High-Stress Musical Performance. Journal of The Royal Society Interface 2013;10(89):20130719.

[5] Wright SE, Palmer C. Physiological and Behavioral Factors in Musicians' Performance Tempo. Frontiers in Human Neuroscience 2020;14.

[6] Bashan A, Bartsch RP, Kantelhardt JW, Havlin S, Ivanov PC. Network Physiology Reveals Relations Between Network Topology and Physiological Function. Nature Communications 2012;3(1).

[7] Cannam C, Landone C, Sandler M. Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files. In Proceedings of the ACM Multimedia 2010 International Conference. Firenze, Italy, 2010; 1467–1468.

[8] Pampalk E. A Matlab Toolbox to Compute Music Similarity from Audio. In ISMIR International Conference on Music Information Retrieval. Barcelona, Spain, 2004; 1–4.

[9] Chew E, Callender C. Conceptual and Experiential Representations of Tempo: Effects on Expressive Performance Comparisons. In Mathematics and Computation in Music. MCM 2013, volume 7937 of Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2013; 76–87.

Address for correspondence:

Dr. Mateusz Soliński
King's College London
1 Lambeth Palace Road
London, SE1 7EU
United Kingdom
mateusz.solinski@kcl.ac.uk