QUANTIFYING COOPERATION IN CHOIR SINGING: RESPIRATORY AND CARDIAC SYNCHRONISATION

Apit Hemakom, Valentin Goverdovsky, Lisa Aufegger, Danilo P. Mandic

Imperial College London
{apit.hemakom12, goverdovsky, l.aufegger, d.mandic}@imperial.ac.uk

ABSTRACT

Cooperative tasks require coordinated joint actions among the participants, to the extent that a failure in an individual's action may have catastrophic consequences on the task of the group as a whole. One such activity is choir singing, where highly synchronised performance of the individual singers is a prerequisite to successful performance. The aim of this work is to provide a quantitative measure of the level of cooperation, established through the degrees of synchronisation between singers' physiological responses. To this end, we employ two new measures, the intrinsic phase synchrony and intrinsic coherence, which quantify synchronisation in respiration and heart rate variability (HRV) of: (i) five members of a choir and the conductor during a rehearsal and a real performance, and (ii) five members of the audience attending the performance. Both the proposed techniques successfully reveal degrees of synchronisation of singers' physiological signals which can be used as physically meaningful measures of the level of cooperation.

Index Terms— NA-MEMD, intrinsic multiscale analysis, intrinsic phase synchrony, coherence, choir singing

1. INTRODUCTION

Cooperative human activities require high degree of mental and physical synchronisation among multiple participants, to the extent that synchrony underpins performance level in activities such as rowing, marching and choir singing. Choir singing is particularly interesting as participants do not use any tools; it can be performed with or without musical instruments, whereby the conductor, essentially, plays 'human instruments'—the soprano, alto, tenor and bass. The normal respiratory rate in adults varies between 12 and 18 breaths per minute [1], yet despite this natural variation, breathing in unison among individuals is a prerequisite in choir singing, where the singers' breathing rhythm is dictated by the tempo and demands of a musical score. These demands and physical constraints give rise to both direct and indirect synchrony and causality in cardiac and respiratory activity at multiple levels – a subject of this study.

A challenge in quantifying synchrony among choir members is that, depending on the score and tempo, breathing rhythms of singers can be either: (i) voluntarily controlled by themselves, in order to perform long or short inhalation or exhalation, or (ii) involuntarily controlled by the autonomic nervous system (ANS). The ANS comprises the sympathetic (SNS) and parasympathetic nervous (PNS) subsystems, whereby the SNS, in addition to controlling the respiration in stressful situations, also accelerates other functions, such as the arterial blood pressure and heart rate [2, 3, 4]. This is achieved by dilating bronchioles in the lungs, and by regulating neuronal and hormonal responses to stimulate the body. The PNS, on the other hand, slows down physiological functions when the body is at rest.

The interplay between the SNS and PNS, among other factors, manifests itself in variations of the timing of the cardiac cycle – heart rate variability (HRV) – in response to both external and internal factors. Changes in HRV are commonly evaluated in two frequency bands: (i) the low frequency (LF) band, 0.04-0.15 Hz, which is linked to the interaction of the SNS and PNS, and (ii) the high frequency (HF) band, 0.15-0.4 Hz, which primarily reflects the activity of the PNS [5]. In addition, it is well understood that breathing modulates HRV via a phenomenon referred to as the respiratory sinus arrhythmia (RSA), whereby the heart rate accelerates during inspiration and decelerates during expiration. The RSA is usually attributed to the activity of the PNS, so that the HF component of HRV is dominated by the changes in heart rate induced by breathing.

Both the respiration and the electrical activity of the heart, measured via the electrocardiogram (ECG), typically exhibit nonlinear and nonstationary characteristics, and require specialised signal processing techniques which offer physically meaningful signal representation; one such technique is the empirical mode decomposition (EMD) algorithm [6] . Empirical mode decomposition is an adaptive, data-driven, method for the analysis of nonlinear and nonstationary time series. It employs the so-called sifting process to decompose a given signal into its multiple narrow-band amplitude/frequency modulated (AM/FM) components, which are referred to as intrinsic mode functions (IMFs) and are used as bases for signal representation.

Unlike conventional projection based time-frequency algorithms, such as the short-time Fourier transform and the discrete wavelet transform, the IMFs – the adaptive basis functions within EMD – enable physically meaningful interpretation of instantaneous phase and frequency, and a highly localised time-frequency representation via the Hilbert transform [7, 8]. Applications of EMD range from biosignal analysis [9, 10], through to mechanical systems [11] and seismology [12].

Due to the empirical nature of EMD, its direct componentwise application to multivariate signals may result in: (i) IMFs with different oscillatory components across multiple data channels for a given IMF index – a phenomenon known as *mode mixing*, and (ii) multiple IMFs containing similar oscillatory modes for a given data channel – a phenomenon referred to as *mode splitting*. To mitigate these problems in multivariate scenarios, several extensions of EMD have been proposed, which include the bivariate EMD (BEMD) [13], trivariate EMD [14], multivariate EMD (MEMD) [15, 16] and noise-assisted MEMD (NA-MEMD) [17]. The general multivariate MEMD has found applications in brain-computer interface [18, 19], image processing [20, 21], nuclear engineering [22] and system characterisation [23].

Such a decomposition into multiple multivariate data-driven bases offers unique opportunities; for example, our recent work [23] proposed a framework referred to as *intrinsic multiscale analysis* which combines MEMD with standard data-association measures, such as phase synchrony (PS), sample entropy (SE) and correlation, in order to quantify intra- and inter-component dependences of a complex system such as multiple synchronies and causalities.

The degree of synchronisation between data channels is typically measured using correlation, coherence and phase synchrony. Similar to correlation, coherence is a measure of linear synchronisation between the two signals, say $x_i(t)$ and $x_j(t)$. It accounts for both the amplitude and phase information, and yields a data association metric which is a function of frequency, f, given by

$$COH_{ij}(f) = \left| \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}} \right|^2, \qquad (1)$$

where $S_{ij}(f)$ is the cross-spectral power density between $x_i(t)$ and $x_j(t)$, and $S_{ii}(f)$ and $S_{jj}(f)$ are respectively the power spectral densities of $x_i(t)$ and $x_j(t)$. The COH values range from 0 to 1, with 0 indicating a non-coherent relationship and 1 the perfect coherence [24, 25].

Unlike the coherence, the phase synchrony metric quantifies only the phase relationship between $x_i(t)$ and $x_j(t)$, without accounting for amplitude information; it is defined in terms of the deviation from perfect synchrony via the phase synchronisation index (PSI) [23], given by

$$\rho(t) = \frac{S_{max} - S}{S_{max}}, \qquad (2)$$

where $S = -\sum_{n=1}^{N} p_n \ln p_n$ is the Shannon entropy of the distribution of phase differences $\phi_i(t - \frac{W}{2} : t + \frac{W}{2})$ within a window of length W, with N being the number of bins within the distribution of phase differences and p_n the probability of $\phi_i(t - \frac{W}{2} : t + \frac{W}{2})$ within the nth bin [26]. The maximum entropy S_{max} has been found to be $S_{max} = 0.626 + 0.4 \ln(W - 1)$ [23]. The PSI values range from 0 to 1, with 1 indicating the perfect phase locking and 0 non-phase-synchronous relationship.

The intrinsic phase synchrony (IPS) was originally proposed in the intrinsic multiscale analysis framework in [23], in order to generalise standard phase synchrony by equipping it with the ability to operate at the intrinsic scale level. It employs MEMD to decompose a given multivariate signal into narrowband intrinsic oscillations (IMFs), which makes it possible to quantify the temporal locking of the phase information in IMFs using the standard phase synchronisation index (PSI). The work in [23] also introduced an intrinsic correlation metric which measures phase and amplitude relationships between instantaneous amplitudes and frequencies. This measure can be further extended to quantify phase and amplitude relationships between intrinsic modes in the data via the IMFs (as a function of scale), a procedure to which we refer to as the *intrinsic coherence* (ICoh).

The aim of this study is to build upon the enhanced discrimination capability of the intrinsic phase synchrony and intrinsic coherence data association metrics, in order to characterise the scale-wise dependencies in the respiratory and HRV signals of: (i) choir during a rehearsal and a real performance; (ii) the conductor in both of these situations; (iii) a subset of audience during the real performance. The so-enabled investigation of the manifold couplings in human physiological responses during a performance promises new, objective measures of the degree of human cooperation, together with new avenues for multidisciplinary research on the *quantified self*.

2. DATA ACQUISITION AND PRE-PROCESSING

Respiratory and ECG signals were recorded from a conductor and a subset of five members of an 18-member choir during 5-minute periods of a low-stress rehearsal and a high-stress public performance. During the performance, physiological responses were also recorded from five members of the audience. Respiration of each participant was recorded using a custom-made respiration belt placed around the chest. For all participants, the ECG was recorded with three electrodes placed on the skin, just below the collar bone. The respiration belt and the electrodes were connected to an 8-channel portable biosignal data logger powered by a rechargeable coin cell battery. The data logger sampled the signals at 1 kHz and saved the respiratory and ECG data onto a micro-SD card; the respiratory signals were then downsampled to 10 Hz, and the trend was removed. The data logger also recorded timestamps onto the micro-SD card in order to guarantee the synchroni-

sation of the devices between rehearsal and performance. The HRV was estimated from the ECG data by band-pass filtering between $8\,Hz$ and $30\,Hz$, and the subsequent R-peak detection to obtain the RR-interval (i.e. HRV) time series with a sampling frequency of $4\,Hz$ [27].

3. SYNCHRONY ANALYSIS

The PSI and COH indices of the respiratory and HRV signals were estimated in six categories as follows:

1. Among the five members of the choir during the rehearsal—within-group estimation.
2. Between the conductor and the five members of the choir during the rehearsal—between-group estimation.
3. Among the five members of the choir during the performance—within-group estimation.
4. Between the conductor and the five members of the choir during the performance—between-group estimation.
5. Among the five members of the audience during the performance—within-group estimation.
6. Between the conductor and the five members of the audience during the performance—between-group estimation.

For the categories (1)-(4), the respiratory (or HRV) signals of the conductor (channel 1) and the five members of the choir (channels 2-6) during both the rehearsal and the performance were used to form 6-channel data which was decomposed using NA-MEMD with 10 adjacent WGN channels. For the categories (5) and (6), channels 2-6 contained data recorded from the five members of the audience, and the decomposition was carried out in the same manner as for the categories (1)-(4). It should be noted that in all the categories NA-MEMD was applied to 6-channel respiratory data and 6-channel HRV data separately.

The IMFs produced by the NA-MEMD with indices 3-7 of the 6-channel multivariate HRV signal of choir members contained the physically meaningful frequency range $0.04\,Hz$ to $0.4\,Hz$, that is, exactly the LF/HF frequency band of HRV. For convenience we have identified and used the same band in IMF indices 5-9 of respiration. The full band of interest in both the HRV and respiration data was produced by summing up the corresponding IMFs, in order to obtain the desired scale in data. The PSI and COH among the members of the choir (and the audience) were obtained by averaging PSI and COH values calculated from every combined-IMF pair of the data channels 2-6 (choir only), while PSI and COH between the conductor and each member of the choir (and the audience) were obtained by averaging the PSI and COH values between the combined-IMF of channel 1 (conductor) and combined-IMF of channels 2-6 (choir). The PSI indices between the combined-IMFs of the noise channels were also estimated in order to provide the PSI of random signals as a benchmark. For the coherence analysis, the power spectral

densities were estimated using the MVAR model of order 3 for the HRV signal of both the choir and the audience, and of orders 1 and 2 respectively for the respiratory signals of the choir and the audience.
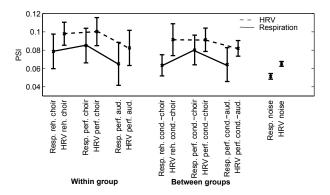

Fig. 1. The PSI of the respiratory and HRV signals within the same subject group and between the subject groups, where for brevity Resp.=Respiration, reh.=rehearsal, perf.=performance, aud.=audience and cond.=conductor.

4. RESULTS OF THE ANALYSIS

Phase synchrony. Fig. 1 shows the PSIs of the respiratory and HRV data for the 6 considered categories, estimated from 50 realisations of NA-MEMD. For each trial of NA-MEMD, the Z-test at a significance level of 0.05 was performed, in order to reveal statistical differences in the respiratory and HRV PSI values between the categories.

Observe that in the performance conditions, both the respiratory and HRV PSIs among the choir members significantly increased compared to the rehearsal. This indicates that the breathing rhythms, and by virtue of RSA the cardiac activities too, of the choir members exhibited stronger dynamic coupling which was reflected in an increase in the synchrony of their physiological responses.

Conjecture #1: The intrinsic synchrony of physiological responses of choir members is a signature of increased coordination between performers and their enhanced mental awareness in various types of performance. This is for the first time that we have been able to quantify the involvement of physiological mechanisms which are responsible for the change in the balance between the SNS and PNS.

Observe that the values of PSI indices of the respiratory and HRV signals among the audience were lower than those for the choir, both during the rehearsal and performance.

Conjecture #2: The selection of participants from the audience did not follow specific selection criteria (e.g. musical background) which may explain the more random nature of their physiological responses.

Fig. 1 also shows that the respiratory and HRV PSIs between the conductor and the members of the choir were lower

than PSIs among the members of the choir, in both the rehearsal and performance scenarios.

Conjecture #3: The physiological responses of the conductor were modulated by both the piece of music performed (as indicated by a degree of coupling with the choir) and the physical activity involved in the act of conducting.

Fig. 1 highlights that PSIs between human subjects were significantly higher than noise PSIs, thus indicating that human PSIs were not random; also the PSIs of HRV exhibited similar patterns to those present in the respiratory signal—a consequence of RSA.

Coherence analysis. Fig. 2 illustrates the coherence analysis of respiration and HRV, estimated for the 6 categories considered. Fig. 2(a) shows respiratory coherence among the members of the choir during both the rehearsal and performance, together with those among the members of the audience.

Observe the peaks in the frequency range $0.2\,$Hz to $0.35\,$Hz which indicate the coherence in breathing at normal rates. The members of the audience breathed faster than the choir, resulting in coherence at higher frequencies. Conversely, owing to long exhalation while singing, the choir exhibited high coherences in the low frequency region of $0.04\,$Hz to $0.05\,$Hz (ellipse 1).

Remark #1: The coherence of the choir during the performance was similar to that during the rehearsal, while the coherence between between the conductor and the members of the choir was markedly higher (see Fig. 2(b), ellipse 2). This indicates a higher degree of cooperation during the performance, when the stakes are high and both the conductor and the choir feel the pressure and thus the urge to work extra hard in order to produce a spot on performance.

Fig. 2(c) shows HRV coherences among the members of the choir during the rehearsal and performance, and those among the members of the audience during the performance.

Remark #2: The higher COH values among the choir in the HF band ($0.15\,$Hz to $0.4\,$Hz) of HRV during the performance, compared to the rehearsal, can be attributed to the RSA (ellipses 3 & 4). This indicates a higher degree of cooperation which is reflected in a more pronounced synchronisation between individuals' cardiac activity mediated by respiration via RSA, additionally notice how virtually no coherence is detected in the cardiac activity of the members of the audience.

5. CONCLUSIONS

This study has employed intrinsic phase synchrony and intrinsic coherence to quantify phase and amplitude relationships in the respiratory and cardiac signals of the choir, the conductor and the audience, in order to investigate degrees of cooperation during a social task. The results have shown that each group is represented by a distinctive degree of joint synchronisation of participants' physiological responses, caused by the act of performance. While the choir has demonstrated a
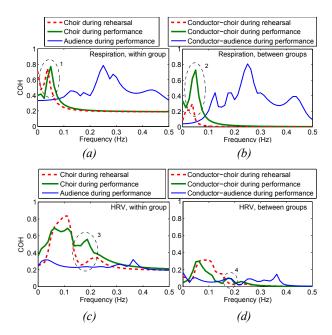

Fig. 2. The COH values of the respiratory and HRV data. *(a)* Respiratory COH within the same group. *(b)* Respiratory COH between groups. *(c)* COH of HRV within the same group. *(d)* COH of HRV between groups.

marked increase in coordination from the rehearsal to the performance, less agreement has been observed between the subject groups (choir, conductor, audience), with the lowest coordination observed for the audience. This can be attributed to each group experiencing the performance in different physical and mental ways, as exemplified by the lower synchronisation between the groups. On the signal processing side, it has been shown that intrinsic phase synchrony has captured phase relationship of both physiological signals in all situations effectively, yielding a meaningful and straightforward to interpret data association metric. We have also illuminated the coherence effects between the sympathetic and parasympathetic nervous systems in the participants, primarily mediated by respiration; these could not be found using intrinsic phase synchrony, however, the coherence is less amenable to physical interpretation. Both the considered intrinsic measures have designated a quantitative approach to assessing joint endeavours, and have paved the way for mathematical characterisation of cooperative physiological systems across human activities. Both the intrinsic phase synchrony and intrinsic coherence, however, are time-varying measures – therefore average PSI and COH are provided. Future work will focus on the quantification of time-varying dependencies of the respiratory and ECG signals.

6. REFERENCES

[1] K.E. Barrett, S.M. Barman, S. Boitano, and H.L. Brooks, *Ganong's Review of Medical Physiology, 24th ed.*, McGraw-Hill Education / Medical, 2012.

[2] M.P. Hlastala and A.J. Berger, *Physiology of Respiration, 2nd ed.*, Oxford University Press, 2001.

[3] W. Janig, *Integrative action of the autonomic nervous system*, Cambridge University Press, 2008.

[4] D.U. Silverthorne, *Human physiology: An integrated approach, 4th ed.*, Pearson / Benjamin Cummings, 2009.

[5] G.E. Billman, "The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance," *Frontiers in Physiology*, vol. 4, pp. 26–1 – 26–5, 2013.

[6] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N. Yen, C.C. Tung, and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, pp. 903–994, 1998.

[7] L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, 1995.

[8] N.E. Huang, M.C. Wu, S.R. Long, S.S.P Shen, W. Qu, P. Gloersen, and K.L. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *Proceedings of the Royal Society A*, vol. 459, pp. 2317–2345, 2003.

[9] H. Liang, Q.H. Lin, and Chen J.D.Z., "Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 10, pp. 1692–1701, 2005.

[10] T.M. Rutkowski, D.P. Mandic, A. Cichocki, and A.W. Przybyszekski, "EMD approach to multichannel EEG data - the amplitude and phase components clustering analysis," *Journal of Circuits, Systems and Computers*, vol. 19, no. 1, pp. 215–229, 2010.

[11] M. Amarnath and I.R.P. Krishna, "Empirical mode decomposition of acoustic signals for diagnosis of faults in gears and rolling element bearings," *IET Science, Measurement & Technology*, vol. 6, no. 4, 2012.

[12] Y. Zhou, W. Chen, J. Gao, and Y. He, "Application of Hilbert-Huang transform based instantaneous frequency to seismic reflection data," *Journal of Applied Geophysics*, vol. 82, pp. 68–74, 2012.

[13] G. Rilling, P. Flandrin, P. Goncalves, and J.M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 936–939, 2007.

[14] N. Rehman and D.P. Mandic, "Empirical mode decomposition for trivariate signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1059–1068, 2010.

[15] N. Rehman and D.P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of the Royal Society A*, vol. 466, no. 2117, pp. 1291–1302, 2010.

[16] D.P. Mandic, N. Rehman, Z. Wu, and N.E. Huang, "Empirical mode decomposition based time-frequency analysis of multivariate signals," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 74–85, 2013.

[17] N. Rehman and D.P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2421–2426, 2011.

[18] C. Park, D. Looney, N. ur Rehman, A. Ahrabian, and D.P. Mandic, "Classification of motor imagery BCI using multivariate empirical mode decomposition," *IEEE Transaction of Neural Systems and Rehabilitation Engineering*, vol. 21, no. 1, pp. 10–22, 2013.

[19] B. Xu, Y. Fu, G. Shi, X. Yin, Z. Wang, and H. Li, "Phase information for classification between clench speed and clench force motor imagery," *Sensors and Tranducers Journal*, vol. 170, no. 5, pp. 234–240, 2014.

[20] N. ur Rehman, B. Ehsan, S.M.U. Abdullah, M.J. Akhtar, D.P. Mandic, and K.D. McDonald-Maier, "Multi-scale pixel-based image fusion using multivariate empirical mode decomposition," *Sensors*, vol. 15, no. 5, pp. 10923–10947, 2015.

[21] S.M.U. Abdullah, N. ur Rehman, M.M. Khan, and D.P. Mandic, "A multivariate empirical mode decomposition based approach to pansharpening," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3974–3984, 2015.

[22] A. Prieto-Guerrero, G. Espinosa-Paredes, and G.A. Laguna-Sánchez, "Stability monitor for boiling water reactors based on the multivariate empirical mode decomposition," *Annals of Nuclear Energy*, vol. 85, no. 1, pp. 453–460, 2015.

[23] D. Looney, A. Hemakom, and D.P. Mandic, "Intrinsic multiscale analysis: A multivariate EMD framework," *Proceedings of the Royal Society A*, vol. 471, 2014.

[24] R.Q. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: A case study on electroencephalographic signals," *Physical Review E*, vol. 65, no. 4, pp. 041903–1 – 041903–14, 2002.

[25] L. Faes and G. Nollo, *Multivariate Frequency Domain Analysis of Causal Interactions in Physiological Time Series, Biomedical Engineering, Trends in Electronics, Communications and Software*, InTech, 2011.

[26] P. Tass, M.G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H.-J. Freund, "Detection of n:m phase locking from noisy data: application to magnetoencephalography," *Physical Review Letters*, vol. 81, no. 15, pp. 3291–3294, 1998.

[27] T. Chanwimalueang, W. von Rosenberg, and D.P. Mandic, "Enabling R-peak detection in wearable ECG: Combining matched filtering and Hilbert transform," *Proc. of IEEE International Conference on Digital Signal Processing*, vol. 1, pp. 134–138, 2015.