Computational Analysis of Solo Versus Ensemble Performance in String Quartets: Intonation and Dynamics

Panagiotis Papiotis,[*1] Marco Marchini,[*2] Esteban Maestre[#*3]

[*]*Music Technology Group, Universitat Pompeu Fabra, Spain*

[#]*Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, U.S.A.*

[1]panos.papiotis@upf.edu, [2]marco.marchini@upf.edu, [3]esteban@ccrma.stanford.edu

ABSTRACT

Musical ensembles, such as a string quartet, are a clear case of music performance where a joint interpretation of the score as well as joint action during the performance is required by the musicians. Of the several explicit and implicit ways through which the musicians cooperate, we focus on the acoustic result of the performance – in this case in terms of dynamics and intonation - and attempt to detect evidence of interdependence among the musicians by performing a computational analysis. We have recorded a set of string quartet exercises whose challenge lies in achieving ensemble cohesion rather than correctly performing one's individual task successfully, which serve as a 'ground truth' dataset; these exercises were recorded by a professional string quartet in two experimental conditions: *solo*, where each musician performs their part alone without having access to the full quartet score, and *ensemble*, where the musicians perform the exercise together following a short rehearsal period. Through an automatic analysis and post-processing of audio and motion capture data, we extract a set of low-level features, on which we apply several numerical methods of interdependence (such as Pearson correlation, Mutual Information, Granger causality, and Nonlinear coupling) in order to measure the interdependence -or lack thereof- among the musicians during the performance. Results show that, although dependent on the underlying musical score, this methodology can be used in order to automatically analyze the performance of a musical ensemble.

I. INTRODUCTION

Joint performance in a musical ensemble is a very interesting case of social collaboration, since communication between the musicians is mostly implicit. It is generally difficult to assess to which degree the actions of one musician are driven by his/her own personal choices rather than by external influences. Broadly speaking, one can expect that the joint performance of a musical ensemble will sound different from an artificially synchronized collection of solo performances – and the reasons behind this expectation as well as the processes behind music performance that make it valid have been a popular research question in many fields.

However, little work has been carried out in quantifying *how well* an ensemble is collaborating, given the musical score they are tasked with performing. This question goes beyond the concept of temporal synchronization, and can be potentially crucial for research purposes as well as educational or artistic applications. This article presents our ongoing work towards building a methodology capable of answering the above question – how strong is the interdependence among the members of a musical ensemble. By focusing on the acoustic result of the performance, it is our goal to assess the degree to which an ensemble is truly collaborating in shaping the final performance (as opposed to each musician simply performing their part in time with the others). Of course, answering this question by itself will not shed much light in the inner workings of joint music performance; however, it can be argued that such a step can assist in understanding such a complex phenomenon more thoroughly.

A. Related work

A significant amount of research has been carried out on the topic of ensemble performance synchronization. One strong trend of research deals with the theoretical modeling of the synchronization mechanisms, while there is also a large amount of empirical work based on experimental data. Naturally, synergies between the above represent a large amount of the existing literature.

A theoretical description of joint musical performance can be found in (Keller, 2008), where the author focuses on three main sub-processes: auditory imagery, where the musician has his/her own anticipation of their own sound as well as the overall sound of the ensemble, prioritized integrative attention, where the musician divides his/her attention between their own actions and the actions of others, and adaptive timing, where the musician adjusts the performance to maintain temporal synchrony. The final process, essentially an error correction model where discrepancies between timing representations are detected, has its mathematical foundation in phase and period synchronizations. This description has been inspiring in designing an experimental framework where by manipulating the conditions of a controlled experiment we expect to measure differences in interdependence in different musical facets, such as *intonation*, *articulation*, *dynamics*, *rhythm* and/or *timbre*.

An example where such experimental conditions are manipulated can be found in (Goebl & Palmer, 2009). Several interaction paradigms of a leader-follower relationship have there been tested for a piano duet, showing differences between providing solely audio, audio-visual or solely visual feedbacks for tempo synchronization.

Other works focus their analysis on selected musical excerpts requiring specific skills. In (Moore & Chen, 2010), gesture data of two string quartet musicians performing a challenging synchronized fast sequence of note is analyzed. Their results show that the repetitive up-down bowing pattern of the two musicians exhibits the properties of an alternating renewal process related to the metrical structure of the piece.

Another approach that is contextually similar has been carried out in (Glowinski *et al*, 2010). Their work focuses more on the social interactions between the musicians (specifically dominance) by studying the musicians' head movements, suggesting that the leader in a musical ensemble is the musician whose movement is the least complex.

Less work has been carried out in studying specific aspects of the produced sound. An important contribution can be found in for the case of singing voice in (Kalin, 2005), where the difference between singing solo versus singing in a barbershop quartet is investigated; his findings demonstrate that singers tend to separate their formants one from the other when singing together, as this helps to hear oneself better and thus facilitate intonation. Regarding the specific case of intonation adjustments, (Mason, 1960) discussed about the effect of joint performance on the choice of tuning temperament from a musicological point of view.

However, to our knowledge, there exists no literature involving a computational approach to intonation adaptations or dynamics fluctuations.

B. Outline

In this article, we focus on validating a methodology for assessing the strength of interdependence in a string quartet, in terms of intonation and dynamics. Such a validation is carried out by applying the same methodology to recordings of joint performance and artificially synchronized solo recordings, and comparing the results of that methodology for both cases.

Our methodology can be described as follows: first, a series of analysis and post-processing steps is performed on recorded music performance data originating from a set of specifically designed experiments. Through this procedure, we extract a set of low-level features describing the aspect of performance we want to focus on (dynamics or intonation). Finally, we apply a set of numerical methods for quantifying interdependence, in order to assess the overall strength of the interdependence among the ensemble.

The rest of this article is structured as follows: in chapter 2, we describe our methodology for experimental data acquisition and analysis. In chapters 3 and 4, we present our interdependence analysis methodology and the obtained results for the case of Intonation and Dynamics, respectively. Finally, in chapter 5 we briefly discuss the implications of the obtained results as well as the future directions towards which we can expand our approach.

II. METHODOLOGY

Our experiments are based on an exercise handbook for string quartets (Heimann, 1995), specifically designed to assist in improving the ensemble's capabilities for collaborative expression. These exercises consist of short, simple musical tasks whose challenge lies in achieving overall synchrony rather than correctly performing one's individual task successfully. The material is divided into six categories, with each category containing a number of exercises dealing with a different aspect of ensemble performance: *Intonation*, *Dynamics*, *Unity of Execution*, *Rhythm*, *Phrasing*, and *Tone production/Timbre*; depending on the exercise, there can also be instructions/annotations on what is the specific goal that must be achieved by the quartet.

During the recordings we simultaneously capture two types of data: the sound produced by the musicians via piezoelectric pickups and conventional microphones, as well as the sound-producing gestures of each musician via a wired motion tracking system. Although only audio features are used in this particular study, the low-level features extracted from the motion capture data are used in the data post-processing steps.

Following signal acquisition, the acquired data is post-processed by performing a note level score-performance alignment as well as a temporal matching between the different experimental conditions.

Finally, we employ a number of computational methods for assessing the strength of the interdependence among the four musicians.

A. Experimental material

For the present work, we focus our study on one *Intonation* exercise and two *Dynamics* exercises; Figures 1, 2, and 3 show a small excerpt from each exercise.

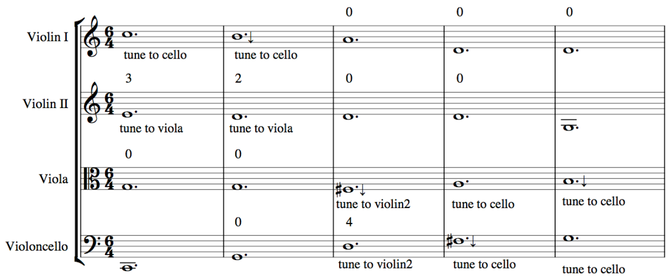

Figure 1. Short excerpt from the Intonation exercise

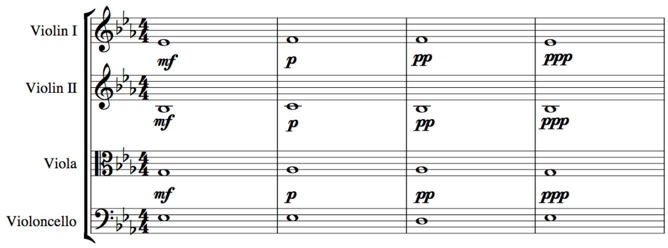

Figure 2. Short excerpt from the first Dynamics exercise (dynamics1)

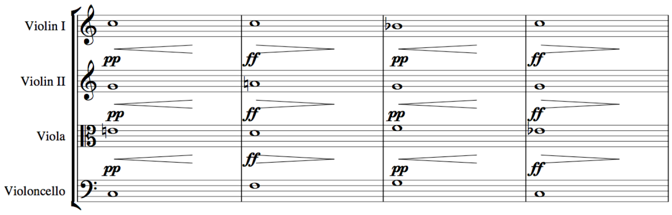

Figure 3. Short excerpt from the second Dynamics exercise (dynamics2)

We record the musicians' performance in two experimental conditions: *solo* and *ensemble*. In the first condition (*solo*), each musician must perform their part alone without having access to the full ensemble score nor the instructions that accompany the exercise. In this way we wish to eliminate any type of external influence on the performance, be it restrictions imposed by other voices of the ensemble or instructions by the composer that are not in relation to the individual score of the performer. Following the solo recordings of each quartet member, the group of musicians is provided with the full ensemble score plus the composer instructions; they are then left to rehearse for a short period (~10 minutes) until they are able to fulfill the requirements of the exercise. Following the rehearsal, the quartet is recorded in the second experimental condition (*ensemble*) performing the exercise as a group.

In the case of exercises with annotations on the score, as in the Intonation exercise studied in this article (see Fig. 1), we split the *ensemble* condition in two sub-conditions: *ensemble1*, where the quartet rehearses and performs the exercise without the annotations, and *ensemble2*, where the quartet rehearses the exercise again with the annotations added to the score and then records the final take.

B. Data acquisition

Both audio and motion capture data are acquired simultaneously and synchronized in real time using a master clock generator as well as a linear timecode audio signal.

Individual audio for each musician is captured through the use of piezoelectric pickups attached to the bridge of the instrument. The overall sound of the ensemble is captured using a cardioid medium diaphragm condenser microphone, as well as a binaural stereo recording dummy head; the binaural stereo recording is then used to set the gain of each individual pickup signal so that the audio level of each instrument corresponds to the overall acoustic result. For each individual pickup signal, we extract two audio features: the *fundamental frequency (F0)* as an estimation of pitch (using the YIN (de Cheveigne & Kawahara, 2002) algorithm), and the *root mean square* audio energy as an estimation of audio intensity.

Besides the audio signals, instrumental - i.e. *sound-producing* - gestures are also acquired through the use of a wired MOCAP system, as detailed in (Maestre, 2009). By analyzing the raw MOCAP data, a set of bowing features is extracted; such features include bow transversal velocity, bow pressing force, bow-bridge distance, bow tilt & inclination, et cetera.

C. Data post-processing

For every recording, a semi-automatic alignment between the performance and the music score is performed using a dynamic programming approach, a variation of the well-known Viterbi algorithm. This approach focuses into three main regions of each note: the note body and two transition segments (onset and offset). Different costs are computed for each segment, using features extracted by the audio (RMS audio energy, Fundamental frequency) as well as the bowing features described above. Finally, the optimal note segmentation is obtained so that a total cost (computed as the sum of the costs corresponding to the complete sequence of note segments) is minimized. This method, which can be seen in more detail in (Maestre, 2009), has so far provided robust results that only in few occasions require manual correction.

Since in the solo case the musicians were not performing simultaneously, there is no temporal coordination to ensure that the recordings are temporally matched; even if the four signals are manually shifted so that the first note onsets coincide, given time they will start to drift apart up to the point where the same sample corresponds to different points in the score. In order to focus on one aspect of the performance at a time, it is necessary to eliminate the above phenomenon by artificially synchronizing the recordings.

Our solution is to perform a non-destructive temporal warping of the *solo* recordings of each experiment in order to impose the note onsets and offsets of their respective *ensemble* recordings; the warping is not applied directly to the recorded sound, but rather to the features that are extracted from it. This is achieved by resampling the *solo* feature between each note onset and the note offset to match the duration of the equivalent *ensemble* note; then, this segment is temporally shifted so that the temporal position of the note onset and offset matches that of the *ensemble* note.

D. Methods for assessing interdependence

Following the post-processing steps, we obtain a set of four time series (one per each musician). Since our goal is to assess the strength of the interdependence among the musicians, we utilize four different numerical methods of measuring interdependence on our time series: *Pearson correlation*, *Mutual Information*, *Nonlinear coupling coefficient*, and *Granger Causality*.

Each of the above methods is based on a different methodology and originates from different fields of research. It is not practical to include an in-depth formulation of each method in the current article; for an excellent review we redirect the reader to (Pereda *et al.*, 2005). However, for the sake of coherence, we provide a summary of each method in the following paragraphs; the first two methods that are presented are *symmetrical* and the last two *directional*; the difference between the two being that a directional method can also assess the direction of influence between two interacting systems besides the strength of the interdependence.

The *Pearson correlation coefficient* is the most common method utilized for quantifying the linear dependence between two variables. The method's output is a value in the [-1,1] range, with a value of zero denoting a complete absence of linear dependence while positive or negative values denote direct or inverse correlation, respectively. Four our case of four interacting time series, we calculate the average correlation value of the whole ensemble by taking the mean of all six correlation values, one for every possible unordered pair between the time series.

Mutual Information, a method originating from Information Theory, measures the difference between two types of joint entropy; the joint entropy of the two variables as measured from the data, and the joint entropy of the two variables as if they were independent. The method's output is equal to zero for independent time series, or a non-bounded positive value representing the amount of information that one gains about one time series by knowing the outcome of the other. For our case, the overall Mutual Information value for the four time series is calculated in the same way as for the Pearson Correlation.

Granger Causality, a statistical concept originating from the field of Econometrics, follows the formulation that if one time series is causal to another, its past values should help in predicting future values of the time series it is causes. Our estimation for the overall causality of the ensemble uses the *total causal density* measure, a [0,1] bounded value with zero denoting a complete lack of causal interactivity.

The *Nonlinear coupling coefficient*, a measure originating from the field of Computational Neuroscience, consists of a variety of nonlinear interdependence measures that quantify the signature of directional couplings between two time series; it is assumed that the processes behind the time series are characterized by separate deterministic dynamics which both exhibit an independent self-sustained motion. Past values of both time series are used to reconstruct their dynamics in order to assess the coupling strength. Of these measures, we use the measure L, which was recently shown to be of higher sensitivity and specificity for directional couplings than previous approaches. For a more in depth explanation of the method as well as its mathematical formulation, we direct the reader to (Chicharro & Andrzejak, 2009) where the method was originally introduced. Since this method is directional, we calculate the average coupling strength of the whole ensemble by taking the mean of all twelve coupling values, one for every possible ordered pair between the time series.

III. INTONATION

The intonation exercise studied in this article consists of consecutive four-note chords, with each note performed by one musician. A short excerpt of the exercise can be seen in Figure 1; the reader can observe that there are annotations on the score, which instruct the musicians on how they must adjust their intonation as well as on whom they must adjust it to, alternating between an equal temperament system to just intonation. An upward-facing arrow signifies slightly sharper intonation, while a downward-facing arrow signifies a slightly flatter one.

Selective adjustments to each musician's intonation are fundamental in order to achieve harmonic consonance for the overall sound; these adjustments can be attributed to many factors, starting from the musicians' own perception of correct intonation besides external influences. Given the complexity of the problem, a first goal is to develop a methodology that is capable of quantifying the degree to which the intonation adjustments in a group of musicians are a result of interdependence among them; a previous attempt at addressing this problem can be found in (Papiotis *et al.*, 2011).

A. Extraction of intonation adjustments

The pitch contour of each recorded excerpt is obtained by applying a logarithmic transformation to the extracted fundamental frequency (akin to *pitch cent* conversion). In order to extract the adjustments of each musician's intonation, we consider the score as 'reference pitch', i.e. perfect, non-adjusted intonation according to the *equal temperament* system; this representation of the performance-aligned score is then subtracted from the obtained pitch contour (or the warped pitch contour for the *solo* condition). This final feature is our estimation of intonation adjustment, with positive values indicating a note played slightly sharper and negative values indicating a note played slightly flatter than equal tempered intonation, respectively.

B. Overall intonation statistics

As an overview of the differences between the three experimental conditions (*solo*, *ensemble1* for the recordings without annotations and *ensemble2* for the recordings with annotations), we calculated the standard deviation of the intonation adjustments of each musician; the result can be seen in Table 1 :

Table 1. Standard deviation (SD) of intonation adjustments per musician, in all three experimental conditions.

Musician	SD (in octaves), intonation		
	ensemble2	**ensemble1**	**solo**
V1	0.0081	0.0076	0.0059
V2	0.0083	0.0068	0.0070
VLA	0.0082	0.0059	0.0070
CLO	0.0063	0.0047	0.0049

It can be observed that the standard deviation is slightly increased in the *ensemble2* condition in comparison to the *solo* condition, with *ensemble1* located between the above conditions. This suggests that after being provided with the score annotations and practicing, the musicians tended to move away from an equal temperament tuning system; we have not focused on this phenomenon for the time being, although it is an important future step.

C. Interdependence results

We apply the four previously introduced interdependence methods on the intonation adjustments data, in the manner that was described in fourth section of chapter II.

Correlation values per musician pairs as well as average correlation values can be seen in Table 2:

Table 2. Correlation coefficient of intonation adjustments per musician, for all three experimental conditions.

Musician pair	Pearson correlation, intonation		
	ensemble2	**ensemble1**	**solo**
V1, V2	0.315	0.297	0.098
V1, VLA	-0.045	0.037	-0.350
V1, CLO	-0.210	-0.365	-0.086
V2, VLA	-0.212	0.021	0.090
V2, CLO	-0.421	-0.508	-0.010
VLA, CLO	-0.100	-0.145	-0.099
Average	**-0.112**	**-0.110**	**-0.059**

One can definitely observe similarities among the two ensemble set-ups as well as dissimilarity between the solo and ensemble set-ups; it can also be observed that the correlation values for the two ensemble set-ups is sporadically higher than the solo set-up. Nevertheless, this is not particularly consistent (such as the case of the violin1/viola pair in the *solo* condition); this keeps us still cautious against making an assumption regarding interdependence from the correlation values.

For the rest of the interdependence methods, we performed a sliding window analysis of interdependence; there are three main reasons behind this choice. First, as we are studying variables that change with time it is natural that the amount of interdependence will also vary; something which can potentially reveal the role of the musical score in the performance, as the collaborative task that must be jointly carried out by the musicians. Second, by windowing the signals by the average note length, we can reduce possible non-linearities and non-stationarities in out data, thus making the interdependence measures more reliable. Finally, we can deal with a smaller amount of data at a time, which removes the need to downsample our signals in order to cope with memory requirements. The analysis window was set equal to the average note length.

The mean value of interdependence across all analysis windows for each method following this analysis can be seen in Table 3 :

Table 3. Nonlinear coupling (NC), Mutual Information (MI) and Causal Density (CD) of intonation adjustments, for all three experimental conditions.

Method	Interdependence strength, intonation		
	ensemble2	ensemble1	solo
NC	0.334	0.307	0.166
MI	0.574	0.521	0.224
CD	0.034	0.025	0.024

It can be seen that the nonlinear methods are successful at separating the solo and two ensemble set-ups. We ran a 1-way ANOVA test to quantify the separation between the three set-ups: the results showed that both the Nonlinear coupling coefficient as well as Mutual Information could successfully separate the interdependence means of one set-up from the other at a confidence level larger than 95%, with the solo set-up showing significantly less interdependence than the two ensemble set-ups for both interdependence methods. One the other hand, Granger causality failed to provide significant separation for the three experimental set-ups; although we hypothesize that this is due to the nonlinearity of our data, it is necessary to investigate towards the validation of this hypothesis in future work.

IV. DYNAMICS

For the case of dynamics, two different exercises have been studied (referred to as *dynamics1* and *dynamics2* from here on). Short excerpts of these exercises can be seen in Figures 2 and 3. Both exercises consist of consecutive four-note chords; in *dynamics1*, the quartet is tasked with simultaneously changing the dynamics value from one bar to the other, while keeping the dynamics steady for the duration of the bar. In *dynamics2*, the quartet must perform a series of *crescendos* and *decrescendos* to and from different dynamics values for every bar.

A. Estimation of dynamics intensity

For each one of the bridge pickup recordings, the Root Mean Square (RMS) Energy of the audio signal is computed in the time domain over a sliding window of 40 milliseconds. The values of RMS energy are converted to a logarithmic scale, in order to obtain an estimation of loudness that is closer to human auditory perception (as per Fechner's law). Finally, the log-RMS energy of the signal is smoothed using a median filter with a window of 300 milliseconds.

It has been reported (Cheng, 2008) that consecutive values of musical dynamics (such as the transition from *pianissimo* to *piano* of from *forte* to *fortissimo*) are equally distanced by a margin of roughly 10 dB – which is consistent with our use of log-RMS as an descriptor for dynamics intensity. However, it is understood that musical dynamics belong on a relative scale rather than being tied to absolute values of loudness.

B. Overall dynamics statistics

As an overview of the differences between the *solo* and *ensemble* experimental conditions, we have calculated two statistics: the mean standard deviation of the dynamics of each musician, as well as the mean absolute difference between performers, in terms of dynamics.

1) Mean standard deviation of dynamics.

For every note of a recording, we calculate the standard deviation of the log-RMS feature; the mean standard deviation is obtained by averaging this value over all notes. Table 4 shows the obtained values.

Table 4. Mean standard deviation (MSD) of dynamics per musician, for both exercises and experimental conditions.

Musician	MSD (in dB), dynamics1		MSD (in dB), dynamics2	
	ensemble	solo	ensemble	solo
V1	2.7906	**2.3865**	3.8656	**2.3026**
V2	3.5746	**1.7824**	3.2393	**1.6585**
VLA	2.9497	**2.2292**	3.8071	**2.3192**
CLO	1.8997	**1.4412**	3.2745	**1.6118**

It can be observed that for both exercises, the mean standard deviation of the *solo* condition is consistently lower than that of the *ensemble* condition. This hints at a more steady behaviour for the *solo* case, presumably due to the lack of external perturbations from the other musicians.

2) Mean absolute difference between performers.

For every note of a recording, we calculate the absolute difference of log-RMS values between every possible pair of performers; the mean of these values is calculated as the mean absolute difference among the performers, as an estimation of how consistently similar are the values of dynamics among the ensemble. Table 5 shows the obtained values.

Table 5. Mean absolute difference (MAD) between performers, for both exercises and experimental conditions.

Musician pair	MAD (in dB), dynamics1		MAD (in dB), dynamics2	
	ensemble	solo	ensemble	solo
V1, V2	8.2125	5.8849	9.3396	4.5758
V1, VLA	14.5180	17.0905	13.7721	18.7440
V1, CLO	18.6416	19.5397	23.4995	21.2992
V2, VLA	6.6795	17.7792	4.3054	14.9096
V2, CLO	10.8925	20.3437	12.9974	17.2552
VLA, CLO	4.7863	3.1215	8.2695	4.4603
Average	**9.5525**	**17.4348**	**11.1685**	**16.0824**

It can be observed that for both exercises, the mean absolute difference of the *solo* condition is generally higher than that of the *ensemble* condition; which hints at a lack of cooperation between performers in order to achieve similar dynamics throughout the exercise.

3) Comments.

The above statistics, while indicative of the differences between the two experimental conditions *solo* and *ensemble*, are by no means sufficient for decidedly separating them; moreover, their potential to provide insight into the strength of interdependence among the performers is limited. The next section will attempt to approach this problem.

C. Interdependence results

Correlation values per musician pairs as well as average correlation values can be seen in Table 6:

Table 6. Correlation coefficient of log-RMS per musician, for both exercises and experimental conditions.

Musician pair	Pearson correlation, dynamics1		Pearson correlation, dynamics2	
	ensemble	**solo**	**ensemble**	**solo**
V1, V2	0.9267	0.8806	0.9427	0.9615
V1, VLA	0.9207	0.8719	0.9042	0.8895
V1, CLO	0.9291	0.8658	0.9081	0.8595
V2, VLA	0.9094	0.8880	0.9210	0.9346
V2, CLO	0.9002	0.8844	0.9247	0.9140
VLA, CLO	0.9345	0.9584	0.9212	0.9400
Average	**0.9237**	**0.8825**	**0.9211**	**0.9243**

It can be seen by the above table that correlation is not capable of showing a significant difference between the *ensemble* and *solo* conditions. Besides that, it is clear that all of the studied features are very correlated.

For the rest of the interdependence methods, we performed a sliding window analysis of interdependence, in this particular case using a short (1 second) as well as a long (5 seconds) analysis window. The mean value of interdependence for each method across all analysis frames can be seen in Table 7:

Table 7. Nonlinear coupling (NC), Mutual Information (MI) and Causal Density (CD) of log-RMS, for both exercises and experimental conditions.

Method	Interdependence strength, dynamics1		Interdependence strength, dynamics2	
	ensemble	**solo**	**ensemble**	**solo**
NC, short window	0.8276	0.7438	0.9333	0.8589
NC, long window	0.7441	0.6212	0.8257	0.6943
MI, short window	0.7149	0.6398	0.9123	0.7680
MI, long window	0.7801	0.7463	1.2372	0.9763
CD, short window	0.0046	0.0011	0.0063	0.0018
CD, long window	0.0059	0.0044	0.0041	0.0040

It can be observed that the interdependence strength is consistently higher for all *ensemble* conditions – with some reservation for the case of Granger Causality where the separation between *ensemble* and *solo* is small. However, a 1-way ANOVA analysis of the interdependence strength of each analysis window failed to provide significant separation between the *ensemble* and *solo* conditions at a 95% confidence level. Another observation is that the larger analysis window provides better separation between the two experimental conditions for the nonlinear coupling coefficient; a logical outcome given the fact that the calculation of the nonlinear coupling coefficient relies on past values in order to assess the coupling strength.

It is evident by looking at the two last tables is that values of interdependence are generally high, compared to those obtained for the intonation analysis; an important difference between the two cases being that in the intonation case the analyzed feature is not pitch itself but pitch adjustments, which to a certain degree eliminates the effect of the score on the interdependence strength. For the current case of dynamics, merely the existence of synchronized *crescendi* as well as simultaneous changes in dynamics is bound to increase the overall interdependence strength, making significant separation between *solo* and *ensemble* difficult. The following section deals with our attempt to reduce the effect of the score on the interdependence strength.

1) Score-Independent analysis of interdependence

Separating the effect of the score is a feasible task for the case of intonation, since the 'reference pitch' of each note is already known from the score. In the case of dynamics however, there is no absolute value for dynamics intensity and therefore no objective reference through which 'dynamics adjustments' can be estimated and removed from our features.

In order to reduce the effect of the underlying score as much as possible, we employed a rough 'dynamics adjustments' extraction method which is outlined as follows: for every note in the musical score, we subtract a linear trend from the log-RMS feature; this way, the note-to-note changes in dynamics are greatly reduced, making temporal fluctuations of dynamics within each note's boundaries much more prevalent. It must be noted however, that this affects the studied features in a biased way since the removed linear trend does not necessarily coincide with the 'reference' dynamics value for each note, which remains undefined. Mean interdependence values for the above scenario are shown in Table 8:

Table 8. Score-independent Nonlinear coupling (NC), Mutual Information (MI) and Causal Density (CD) of log-RMS, for both exercises and experimental conditions.

Method	Interdependence strength, dynamics1 (score-independent)		Interdependence strength, dynamics2 (score-independent)	
	ensemble	**solo**	**ensemble**	**solo**
NC	0.6427	0.4698	0.7302	0.5186
MI	0.3076	0.0896	0.6272	0.0288
CD	0.0209	0.0202	0.0159	0.0163

Two observations can be made from the above table. First, the overall strength of interdependence has been reduced. Second, the separation between *ensemble* and *solo* is much larger, again with the exception of Granger Causality. An 1-way ANOVA analysis of the interdependence strength for each note did show significant separation at a 95% confidence level, for the Nonlinear coupling as well as the Mutual Information interdependence methods.

The above can point to two main conclusions; first, that the musical score is indeed a very important factor in a musical ensemble's interdependence. Second, that the non-linear interdependence methods are not only capable at detecting higher interdependence strength for the *ensemble* condition as compared to the *solo* condition, but also at quantifying the overall strength of interdependence.

V. DISCUSSION

In this article we have presented a methodology for assessing the strength of interdependence in a string quartet, in terms of intonation and dynamics. A set of methods of assessing interdependence have been tested and presented; the results show that it is possible to, solely by studying one of these aspects at a time, distinguish between joint performance and artificially synchronized solo performances.

For the case of intonation, we have shown that the methods suited for nonlinear interactions (Mutual Information, Nonlinear coupling coefficient) are capable of detecting increased interdependence in the *ensemble* experimental condition as compared to the *solo* condition. Moreover, the addition of annotations on the score denoting a change from an equal-tempered system to just intonation caused a slight increase in interdependence, which helps in shedding some light on how the music score affects the ensemble's behavior.

For the case of dynamics, we have shown that the same interdependence methods are capable of showing higher levels of interdependence for the *ensemble* condition; it was also demonstrated that by reducing the importance of the score over the temporal fluctuations of dynamics, the separation between *ensemble* and *solo* was clearer, while the overall strength of interdependence was reduced.

However, despite the above encouraging results, there are still many directions in which we can advance and improve our methodology. For example, although we can successfully detect the overall interdependence strength, we have so far not addressed its temporal evolution in connection to the score.

Another direction in which we wish to advance is in characterizing inter-ensemble relationships and their fluctuations along the piece; one obvious next step would be to attempt to extract information on leadership, although the social aspect of a leader may not coincide with what is measured using our tools.

Although this article is focused on testing and validating a methodology, such a methodology is of little use without application. One can envision the application of the above both in revealing the inner workings of collaborative performance, as well as aiding in its realization; we intend to investigate towards bringing such a scenario closer to reality.

ACKNOWLEDGMENTS

The work presented on this document has been partially supported by the EU-FP7 FET SIEMPRE project and an AGAUR research grant from Generalitat de Catalunya.

The authors would like to thank Marcelo Wanderley, Erika Donald and Alfonso Perez Carrillo for their support in carrying out the experiments, as well as the CIRMMT and BRAMS labs at Montreal, Quebec, Canada for hosting them.

REFERENCES

Chicharro, D. & Andrzejak, R. G. (2009). Reliable detection of directional couplings using rank statistics. *Physical Review E - Statistical, Nonlinear and Soft Matter Physics*, 80(2 Pt 2):026217.

Cheng, W. (2008). *Maria Callas, the legend who lived for her art*. Retrieved April 18, 2012, from
http://www.aamhl.org/dBchart.htm

de Cheveigne, A. & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, 111(4):1917–1930.

Glowinski, D., Coletta, P., Volpe, G., Camurri, A., Schenone, A., & Dist, B.-l. (2010). Multi-Scale Entropy Analysis of Dominance in Social Creative Activities. *Complexity*, pages 1035–1038.

Goebl, W & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, 26(5), pages 427–438, 2009.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3):424–438.

Kalin, G. (2005). Formant frequency adjustment in barbershop quartet singing. *PhD thesis,* KTH.

Keller, P. (2008). Joint action in music performance. *EMERGING COMMUNICATION*, 10:205.

Maestre, E. (2009). Modeling instrumental gestures: an analysis/synthesis framework for violin bowing. *PhD thesis,* Universitat Pompeu Fabra.

Mason, J. A. (1960). Comparison of Solo and Ensemble Performances with Reference to Pythagorean, Just, and Equi-Tempered Intonations. *Journal of Research in Music Education*, 8(1):31–38.

Heimann, M. (1995). Exercises for the string quartet. ESTA Denmark

Moore, G. P. & Chen, J. (2010). Timings and interactions of skilled musicians. *Biological cybernetics*, 103(5):401–14.

Papiotis, P., Maestre, M., Marchini, M, & Perez, A. (2011). Synchronization of intonation adjustments in violin duets; towards an objective evaluation of musical interaction. *In Proceedings of the 14th conference in Digital Audio Effects (DAFx-11)*, Paris, France.

Pereda, E., Quiroga, R. Q., & Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology*, 77(1-2):1–37.

Seth, A. K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Journal of Neuroscience Methods*, 186(2):262–273.

Widmer, G. & Goebl, W. (2004). Computational Models of Expressive Music Performance: The State of the Art. *Journal of New Music Research*, 33(3):203–216.