



HAL
open science

Investigation of phase coupling phenomena in sustained portion of musical instruments sound

Shlomo Dubnov, Xavier Rodet

► **To cite this version:**

Shlomo Dubnov, Xavier Rodet. Investigation of phase coupling phenomena in sustained portion of musical instruments sound. *Journal of the Acoustical Society of America*, 2003, 1 (113), pp.348-359. hal-01161397

HAL Id: hal-01161397

<https://hal.science/hal-01161397>

Submitted on 8 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Investigation of phase coupling phenomena in sustained portion of musical instruments sound

Shlomo Dubnov^{a)}

Ben Gurion University, Beer-Sheva 84105, Israel

Xavier Rodet^{b)}

Analysis/Synthesis Team, IRCAM, Paris 75004, France

(Received 5 February 2002; revised 9 August 2002; accepted 16 August 2002)

This work investigates aperiodicities that occur in the sustained portion of a sound of musical instrument played by a human player, due to synchronous versus asynchronous deviations of the partial phases. By using an additive sinusoidal analysis, phases of individual partials are precisely extracted and their correlation statistics and coupling effects are analyzed. It is shown that various musical instruments exhibit different phase coupling characteristics. The effect of phase coupling is compared to analysis by means of higher order statistics and it is shown that both methods are closely mathematically related. Following a detailed analysis of phase coupling for various musical instruments it is suggested that phase coupling is an important characteristic of a sustained portion of sound of individual musical instruments, and possibly even of instrumental families. Interesting differences in phase deviations were found for the flute, trumpet and cello. For the cello, the effect of vibrato is examined by comparing the analysis of a closed string sound played with a natural vibrato to analysis of an open string sound that contains no vibrato. Following, a possible model for phase deviations in the cello is presented and a simulation of phase fluctuations for this model is performed. © 2003 Acoustical Society of America. [DOI: 10.1121/1.1518981]

PACS numbers: 43.60.Cg, 43.75.De, 43.75.Ef [NHF]

I. INTRODUCTION

Acoustical musical instruments, which are considered to produce a well-defined pitch, emit waveforms that are never exactly periodic. These aperiodicities, which occur in the sustained portion of musical instrument, supposedly originate in some not-well-known fundamental mechanism of their sound production that depends both on the manner of playing the instrument and the instrument specific physical properties.

Among the many mechanisms of possible deviations from periodicity (Beauchamp, 1974; McIntyre *et al.*, 1981; Schumacher, 1992; Rodet, 1993, 1995; Vettori, 1995; Weinreich, 1997; Vergez and Rodet, 2000), we analyze two contrasting conditions which appear to be important for sustained portion of sound in musical instruments: synchronous phase deviations of proportional magnitude¹ that preserve the relative phase relations between the partials and thus do not change the shape of the waveform except for compressing or stretching in time, versus independent or phase asynchronous deviations that change the shape of the signal due to the changing phase relations between the different frequency components. By using a sinusoidal analysis, phases of individual partials are precisely extracted and their correlation statistics and coupling effects are analyzed. It is shown that various musical instruments exhibit different phase coupling characteristics.

One must note that in this work we analyze sounds as played by a human player, with all the attendant vibrato and

pitch and amplitude variability. The dependence on the playing method is analyzed in detail for the cello, where naturally played sound on a closed string with vibrato is compared to a naturally played open string cello sound that has no vibrato.

In earlier works (Dubnov *et al.*, 1998, 1997, 1995), we have shown that the particular aspect of phase synchronous versus asynchronous phase fluctuations is strongly related to nonlinear properties of the time series model of the signal. These properties are measured by higher order statistics (HOS) or polyspectra (Mendel, 1991; Nikias and Mendel, 1993) and were shown to be important for characterization of musical instruments in the sustained portion of the sound.

It should be noted that statistical properties of a signal due to phase variations can not be easily revealed by standard spectral analysis methods due to the fact that second-order statistics and the power spectrum are “phase blind,” i.e., they are not sensitive to phase variations. In the current work we employ sinusoidal analysis in order to estimate precisely the phase behavior of each partial in sustained portion of musical instruments. We analyze the relative phase fluctuations among different partials using a measure called quadratic phase coupling (QPC). The precise definition of QPC will be given in the next section.² The QPC analysis is compared and mathematically related to higher order statistical analysis that is applied directly to the signal.

In the paper we first consider the problem of detection of QPC and define a detector function that receives values in the range [0, 1], with value one indicating perfect phase coupling and value zero indicating that partials have completely uncoupled phases. We show that phase coupling is a major property of musical instruments with a clearly distinct be-

^{a)}Electronic mail: dubnov@bgumail.bgu.ac.il

^{b)}Electronic mail: rodet@ircam.fr

havior for the case of brass versus string instruments. Specifically, we extend the earlier HOS research (Dubnov *et al.*, 1997) by introducing a discrete bispectral measure that is shown to be equivalent to the phase coupling detector under some mild shift invariance assumptions.

Finally, a detailed analysis of cello instrument is given, comparing the QPC for naturally played sound on a closed string that includes vibrato with a sound of an open string with no vibrato. For the case of the cello, and possibly string instruments in general, we suggest a source-filter model that can generate some phase fluctuations. Simulations with this model create a signal with uncoherent phase deviations that look like the fluctuations that were observed in the original signal.

II. PHASE SYNCHRONOUS VERSUS ASYNCHRONOUS APERIODICITIES

The additive sinusoidal model (Rodet, 1997; Serra and Smith, 1989) of an almost perfectly periodic signal is mathematically expressed by

$$x(t) = \sum_{l=-L, l \neq 0}^L A_l(t) e^{i\theta_l(t)}, \quad (1)$$

with the phase $\theta_l(t)$ defined as

$$\theta_l(t) = l \cdot \omega_0 \cdot t + \phi_l(t), \quad (2)$$

where ω_0 is the fundamental frequency and A_l , ϕ_l are the amplitudes and phases of the l th sinusoidal component (partial).

Among the many possible mechanisms of deviations from periodicity that may occur in the sustained portion of a pitched sound, we analyze two extreme cases:

- (1) application of a synchronous and proportional random modulation to the phases of each partial, such as $\phi_k(t) = l \cdot \phi_1(t)$, and
- (2) application of random and asynchronous phase modulations to each partial.

A. Synthetic example

In order to demonstrate the effect of synchronous versus asynchronous deviations we constructed a synthetic signal consisting of a sum of eight equal-amplitude cosine functions at harmonic frequencies, with fundamental frequency of 220 Hz. The signal was generated at 8000-Hz sampling rate.

For the purpose of generating the phase noise, an initial random vector was created using a random Gaussian noise generator, producing a signal at a sampling rate of 160 Hz. Then a phase noise signal at the signal sampling rate was generated by resampling of the original random vector at a rate that is 50 times higher than the original random vector, thus creating a random phase noise signal at 8000-Hz sampling rate.

Next, two sets of phase noise signals were created, for the synchronous and asynchronous cases, respectively. In the first case, a single-phase noise signal $\phi_1(n)$ was used to

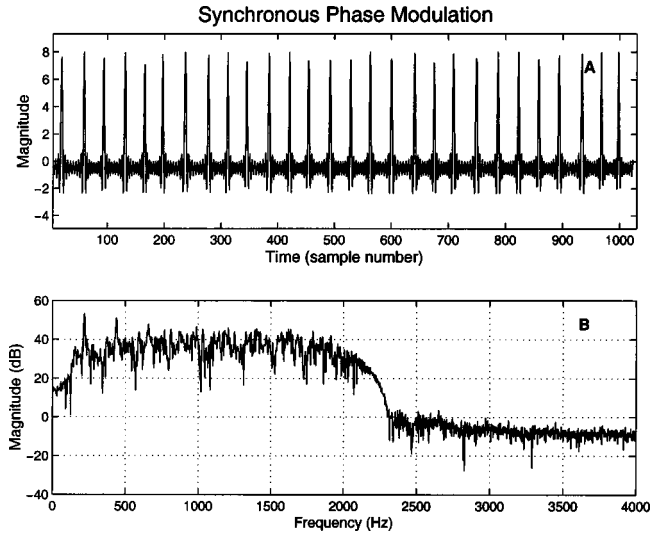


FIG. 1. Synthetic signal that was produced by applying synchronous modulation to phases of the harmonic partials. This type of modulation preserves the shape of the waveform, except for period to period time stretching or shortening, i.e., the period between successive peaks is modulated. In frequency this amounts to spread in bandwidth of the partials.

create the separate phase noise signals of the eight harmonics, with phase noise at harmonic $k = [1, \dots, 8]$, being $\phi_k(n) = k \phi_1(n)$. In the asynchronous case, eight independent noises were created for the eight partials. In order to match the noise variance in the synchronous and asynchronous cases, the standard deviation of the noise in the asynchronous case at partial k was set to be k times the standard deviation of the noise at partial 1. One additional free parameter is the standard deviation of the phase noise in the first partial. In the simulation presented below (Figs. 1 and 2) we used the value of 0.6 for standard deviation of the first partial.

Figures 1 and 2 demonstrate the signal waveform and magnitude spectra in the case of synchronous and asynchronous phase deviations, respectively. The top figure (a) shows

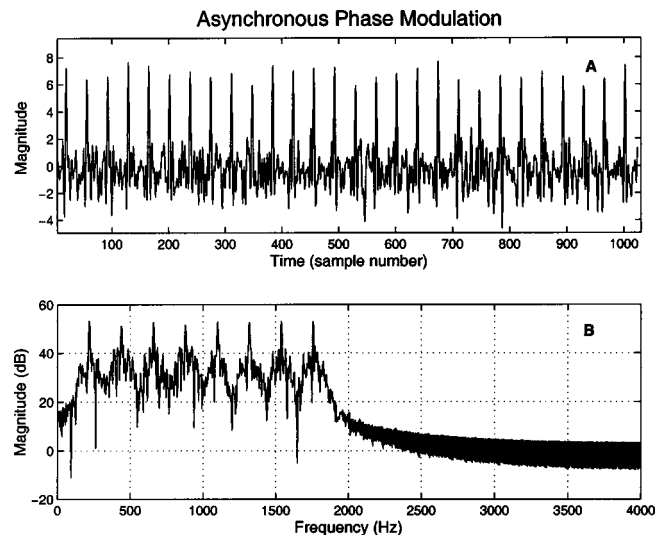


FIG. 2. Synthetic signal that was produced by applying asynchronous modulation to phases of the harmonic partials. This type of modulation does not preserve the shape of the waveform. In frequency this causes a relatively smaller spread in bandwidth of the partials.

the time domain signal and the bottom figure (b) represents the Fourier analysis of a segment of 1024 samples of the signal, with a rectangular analysis window. In the synchronous case, the small variation of the signal period, although not altering the basic waveform, significantly spread the spectral peaks of the higher partials, effectively turning partials higher than 3 or 4 into noise. For the asynchronous case we observe that the original, pulselike shape of the waveform is significantly distorted, giving a visual impression of pulses submerged in a high level of noise. This noise effect can be seen also in the signal's spectrum. One should note that the spectral peaks for the asynchronous case remain significantly pronounced above noise level.

B. Additive sinusoidal analysis procedure

The additive sinusoidal model represents the audio signal as a sum of sinusoids (partials) with time-varying amplitudes and frequencies. Mathematical expression of this model was given earlier in Eqs. (1) and (2) in Sec. II. For review of sinusoidal modeling for musical signals see Rodet (1997). The signal parameters are estimated in the following manner:

- (1) Window audio signal segment. The windowed segment will be termed in the following "analysis frame," or simply a frame.
- (2) Estimate instantaneous pitch (this can be done using a variety of methods). Since a precise pitch is crucial for quality modeling, an extra pitch refinement step is performed.
- (3) Searching for peaks of the local periodogram, i.e., peaks in amplitudes of the short-time Fourier transform (STFT).
- (4) Amplitudes and the instantaneous frequencies are estimated at the precise frequency of the spectral peaks. In our analysis we use a quadratic interpolation of three points around each spectral extrema in order to determine the precise peak location.
- (5) Advance the signal segment to the following frame and repeat previous analysis steps. It is referred in the following as the "analysis time step."
- (6) Finally, a nearest-neighbor matching in time is performed, in order to relate different partials in succession of analysis frames.

The analysis in the paper was performed using a 20-ms analysis window with analysis time step of 4 ms. For the purpose of phase correlation and phase-coupling analysis, only the sustained portion of the analysis was retained. There were some cases where "death" or "birth" of harmonics was observed. In such a case, the instances with almost zero partial amplitudes were removed from the analysis since the phase in such a case was meaningless.

C. Statistical analysis of a sinusoidal model

Taking the phases of a signal derived from sinusoidal analysis we look at the instantaneous harmonicity among different groups of partials. For a triplet of harmonically related partials j , k , and $l=j+k$, a "synchronous" phase be-

havior means that the respective phases ϕ_j , ϕ_k , ϕ_l obey the following relation $\phi_j + \phi_k - \phi_l = 0$, i.e., that any deviations that occurs for ϕ_j and ϕ_k sum up to occur identically in ϕ_l , up to a constant additive factor of the initial phase of each partial. In the case when phase coupling does not occur, a difference in phase deviations occurs between the phases. This phase coupling difference signal can be either bounded (and possibly periodic), or increase by accumulating over time, so that it eventually passes through all values in $[0, 2\pi]$ range.

We introduce a measure $d_3(j, k)$ that evaluates the effect of stochastic phase deviations that occur between partials j and k ,

$$d_3(j, k) \stackrel{\text{def}}{=} \frac{1}{N_F} \sum_{n=1}^{N_F} e^{i(\phi_j(n) + \phi_k(n) - \phi_l(n))}, \quad (3)$$

where N_F is the number of frames available from sinusoidal analysis. The integral of the exponent of the phase coupling difference has the following property: in the case of a perfect coupling, the argument in the exponent is identically 0 and $d_3(j, k)$ equals 1. In the case when the phase coupling difference oscillates in a limited range, the resulting $d_3(j, k)$ will converge to some value between 1 and 0. If the error "spreads" over the whole $[0, 2\pi]$ range, the value of d_3 approaches 0.

D. Phase correlation and phase-coupling analysis

In this section two statistical measures of phase deviations among partials of a sinusoidal model are considered—correlation and coupling. In order to be able to calculate correlation between phases, a careful unwrapping of the phases must be done. One should note that unwrapping is **not** necessary for phase coupling analysis, since the phase differences that appear in the exponent are not sensitive to jumps of 2π . The unwrapping is performed in the following manner:

- (i) Phase derivative of the k th partial's phase is calculated and divided by 2π . The resulting signal is an estimate of an instantaneous frequency of the k th partial.
- (ii) A difference signal between the measured instantaneous frequency and theoretical instantaneous frequency (estimated as a k times multiplication of the fundamental frequency) is calculated.
- (iii) Points of 2π jumps are detected by rounding the difference signal to a nearest integer. The piecewise constant signal is then integrated and multiplied by 2π to create a "phase correction" signal.
- (iv) The "phase correction" signal is added to the original phase to create the unwrapped phase.

This procedure results in unwrapped phase values that represent phase fluctuations around an ideal theoretical value that is derived from the fundamental frequency. This step also eliminates phase deviations due to effects such as vibrato or slight pitch changes. This unwrapped phase is used to calculate the phase correlation. The QPC measure is calculated according to Eq. (3).

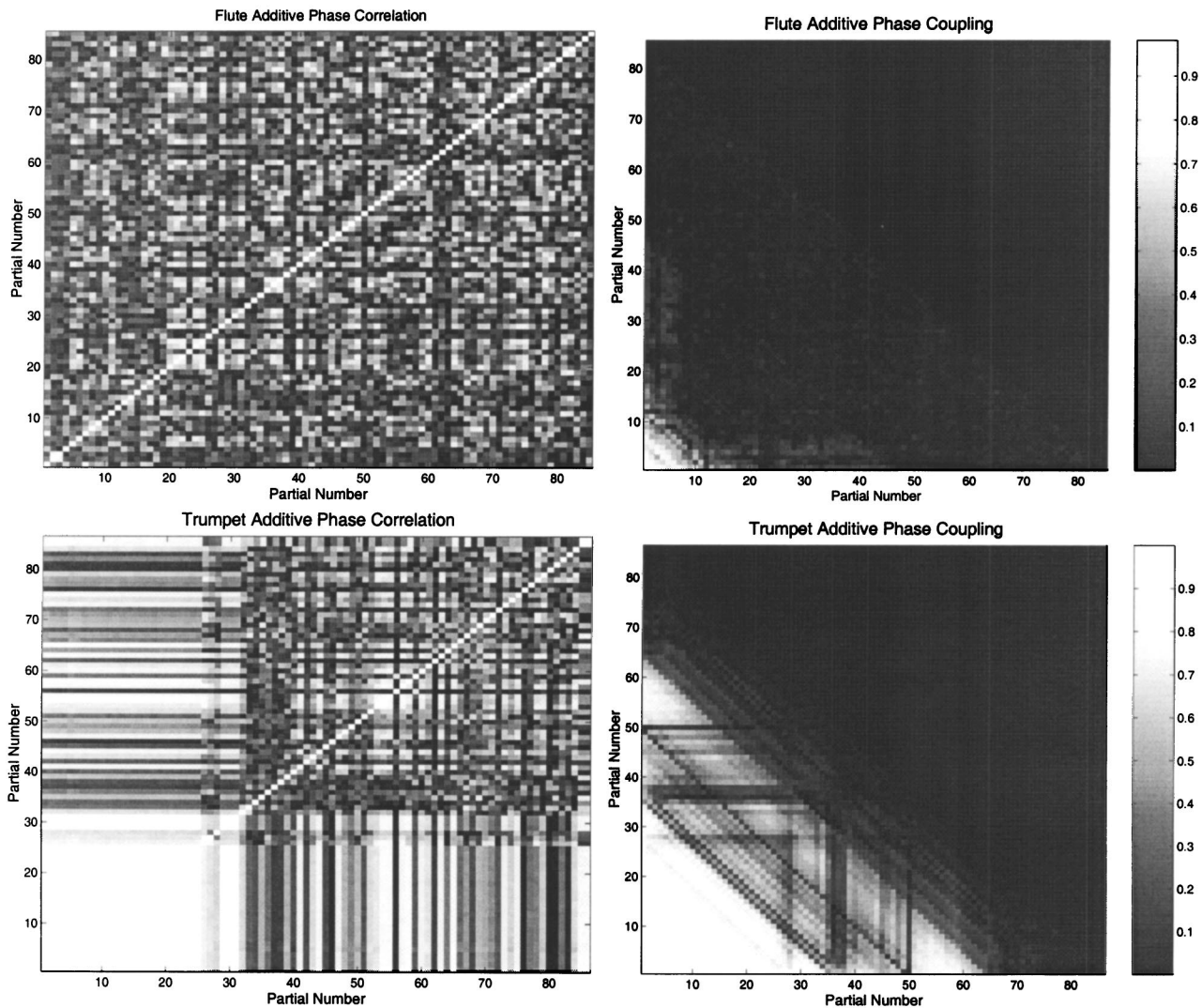


FIG. 3. Flute (top) and trumpet (bottom) correlation and QPC analysis. Phase correlation is plotted on the left side, QPC on the right side. The axes correspond to partial numbers, each point on the graph representing the value of correlation or QPC analysis between the two partials. The color scale is white corresponding to 1, which represents high correlation/coupling, and black corresponding to 0, no correlation or coupling. One should note also that QPC analysis is meaningful only at the lower left triangular part of the figure, since there exists no partial beyond Nyquist frequency. The right top part should be ignored (it is colored black).

In order to examine the difference between phase correlation and coupling, we calculated the matrix of pairwise correlations among the phases and compared them to the two-dimensional matrix of quadratic phase coupling (QPC). Figure 3 presents the phase correlation and QPC for flute and trumpet sounds. As can be readily seen from Fig. 3, the two instruments exhibit a very different correlation and QPC behavior. In the trumpet signal we observe a common behavior of the phase correlation and phase-coupling measure. The flute exhibits a very little QPC but still a rather significant correlation. This figure demonstrates the difference between the two analysis methods: QPC requires a precise instantaneous canceling of the relative phase deviations among triplets of partials, while the correlation coefficient depends on normalized inner product between phase functions.

The differences between correlation and QPC behavior of the two instruments can be further understood by considering the behavior of QPC estimates as a function of time. In the trumpet signal, the phase deviations among the different

partials occur in an almost exact instantaneous correspondence with each other,³ as demonstrated in Fig. 4. This causes the QPC in the trumpet to remain high. In the flute signal pairwise correlations occur to a large extent, but the phase deviations are mostly uncoupled. This lack of instantaneous phase coupling in the flute signal causes the QPC estimate to converge to zero. The behavior of the QPC estimate as function of time for the trumpet and the flute signals are shown in Fig. 5.

III. RELATION TO HIGHER ORDER STATISTICAL ANALYSIS

Higher order statistical (HOS) analysis investigates the behavior of higher order cumulants of stochastic signals. For third-order analysis we consider triple correlations of the signal, which equals the third-order cumulant assuming a zero mean signal,

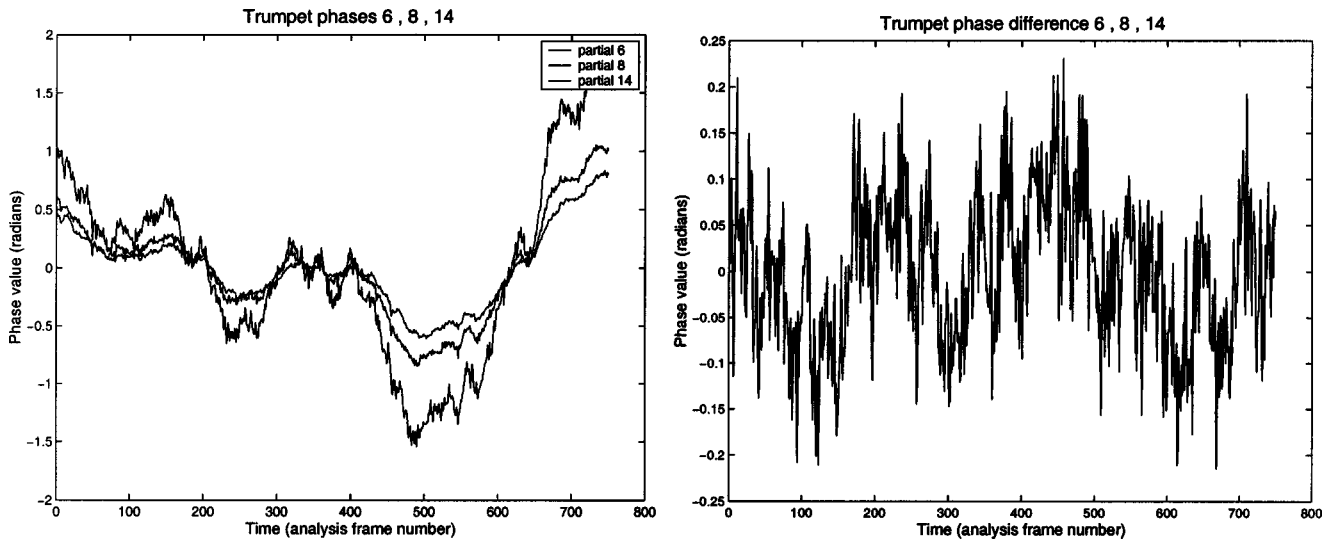


FIG. 4. Phase behavior of partials 6,8,14 of the trumpet. Phases of the partials appear on the left, phase triplet difference appears on the right (see text).

$$c_3(n,m) \stackrel{\text{def}}{=} \langle x(t)x(t+n)x(t+m) \rangle, \quad (4)$$

where $\langle \rangle$ means an ensemble average or time average for ergodic signals. Transforming the third-order cumulant into frequency domain gives the bispectrum

$$B_x(\omega_1, \omega_2) \stackrel{\text{def}}{=} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} c_3(n,m) e^{-i(\omega_1 n + \omega_2 m)}. \quad (5)$$

Note that bispectrum is calculated by applying a 2D Fourier transform to the third-order cumulant function. It can be shown, moreover, that B_x is equivalently expressed as an average over the Fourier transform $X(\omega)$ of $x(t)$. (This result is known as the 2D version of the Parseval theorem.)

$$B_x(\omega_1, \omega_2) = \langle X(\omega_1)X(\omega_2)X^*(\omega_1 + \omega_2) \rangle. \quad (6)$$

Let us consider now a sinusoidal signal \hat{x} that has equal amplitudes for all partials. We shall term this signal “white.”

We can express this signal as a sum of complex exponentials with unit amplitude

$$\hat{x}(t) = \frac{1}{2} \sum_{l=-L, l \neq 0}^L e^{i\theta_l}$$

with $\theta_l = -\theta_{-l}$, where θ_l is defined in Eq. (2).

In the Appendix we prove that, under not too severe assumptions, a discrete version of $B_{\hat{x}}(\omega_1, \omega_2)$ can be defined, which approximately equals to $d(j,k)$. Assuming that the frequency resolution of the Fourier analysis is M , we get the following lemma:

Lemma: For pseudo-periodic signal $\hat{x}(t) = \frac{1}{2} \sum_{l=-L, l \neq 0}^L e^{i\theta_l(t)}$ with $\theta_l(t) = l\omega_0 t + \phi_l(t)$, the discrete bispectrum obeys

$$B_{\hat{x}}\left(\frac{2\pi}{L}j, \frac{2\pi}{L}k\right) = d_3(j,k) \quad (7)$$

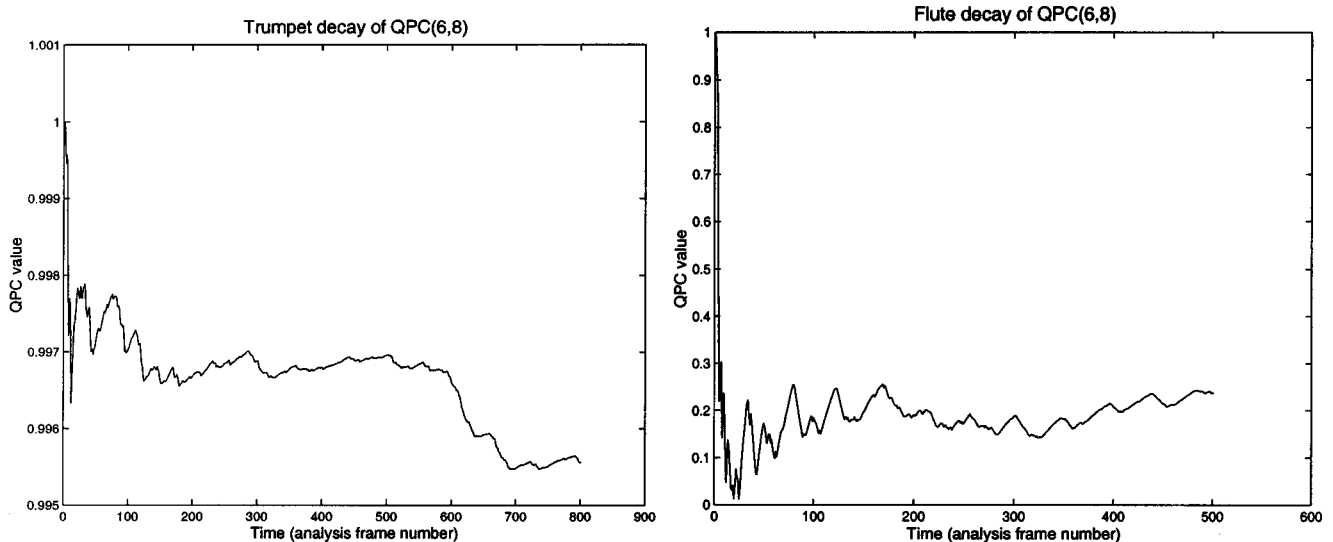


FIG. 5. QPC of the trumpet (left) and flute (right) partials for triplet (6,8,14). The graphs show the decay of QPC as a function of time (analysis frame number).

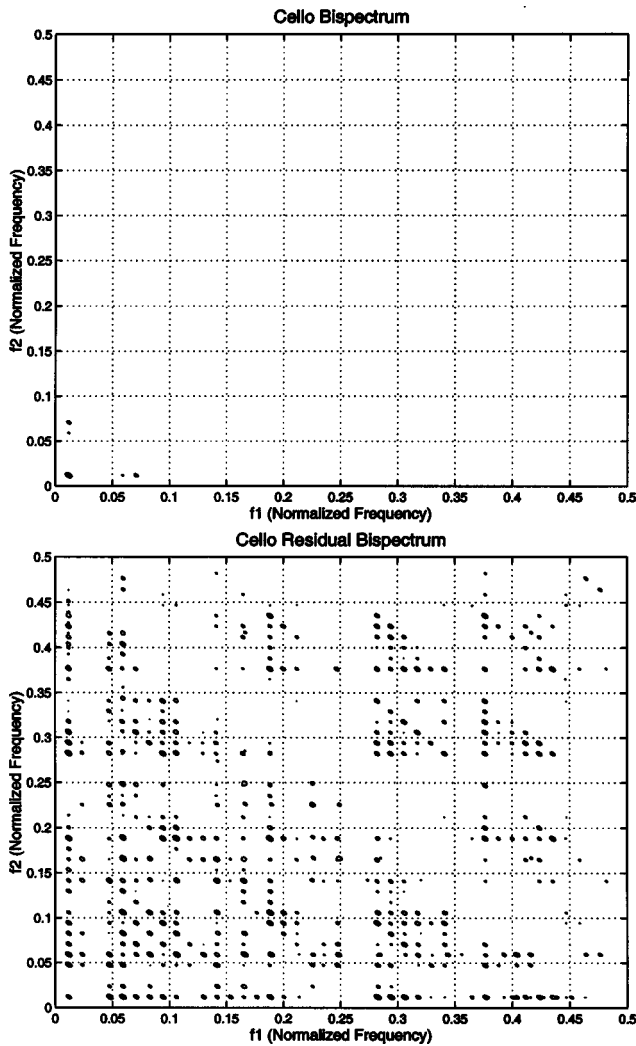


FIG. 6. Bispectrum of a cello original signal (top) and the bispectrum of the residual Cello sound (bottom). See text for more details.

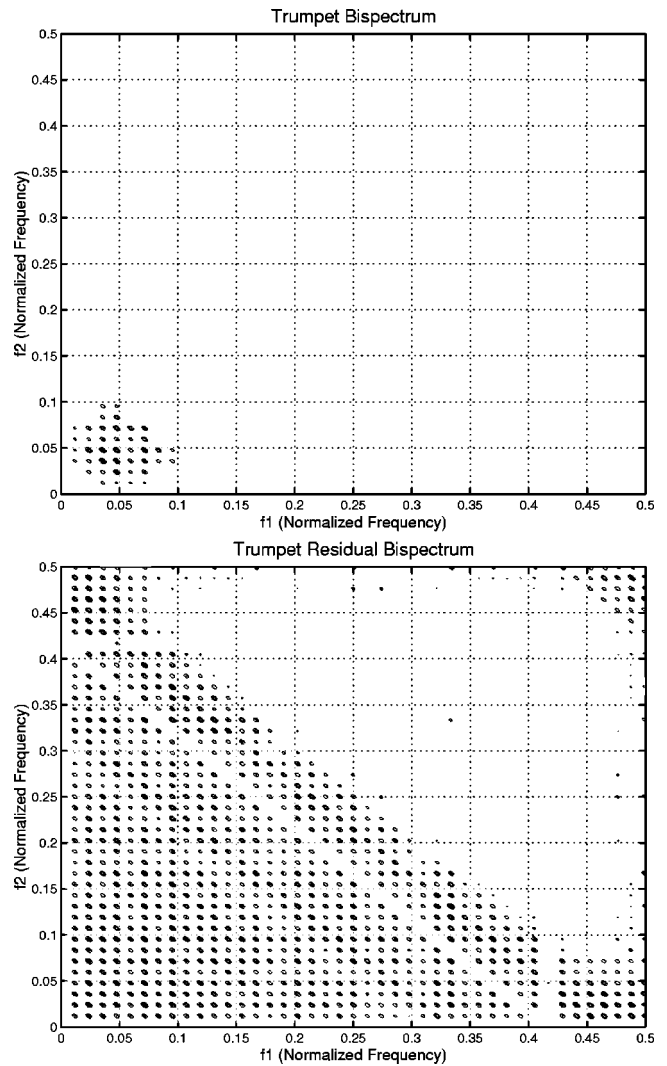


FIG. 7. Bispectrum of a trumpet original signal (top) and the bispectrum of the trumpet residual sound (bottom). See text for more details.

for phase deviation $\phi_l(t)$ being a white random process.

Proof: See the Appendix.

A. Comparison to bispectral analysis

In order to compare the sinusoidal phase coupling analysis to bispectral methods of detection of QPC, all amplitudes of the partials must be made approximately equal. This is required in order to remove the effect of the spectral envelope, which contains information about the amplitudes of harmonics, on the bispectrum. The estimation of the spectral envelope was done using linear prediction analysis (LPC) analysis (Markel and Gray, 1976), using filter of order 16. The process of equalizing the partial amplitudes was done by passing the original signal through an inverse LPC filter. Having performed the equalization step, we remove an initial segment of the signal that contains the transient behavior of the filter and look at the HOS properties of the remaining inversely filtered result, or the so-called *residual* or “white” signal. We chose to do a LPC equalization due to its simplicity, and also due to the fact that other methods such as the bicoherence estimator, which divide FFT bins by their absolute values for eliminating the amplitude dependence, often

suffer from numerical problems, due to nearly zero amplitudes of the FFT bins that do not fall on the partials.

The bispectral analysis of the cello and trumpet residual signals are given in Figs. 6 and 7, respectively. The strong presence of the high harmonics in the residual significantly affects the bispectral analysis. Notice that cello residual has only a few peaks. The analysis is performed on a 2000-sample-long signal segment, sampled at frequency of 22 050 Hz, i.e., a 90 ms segment. The method of analysis employed is a so-called “direct” Bispectrum estimation method that uses the FFT (this is in contrast to “indirect” method that uses third order cumulant matrix). In the direct method, the original frame of analysis is subdivided into eight segments with 50% overlap. Analysis resolution is 1024 (FFT size). In Figs. 6 and 7 the axes correspond to normalized frequency, with $f=1$ corresponding to sampling frequency of 22 050 Hz. When comparing bispectral analysis to additive phase coupling analysis, one should note the following differences in the analysis plots.

- (i) While phase coupling analysis is performed on a full CD-bandwidth signal, i.e., signal with sampling frequency of 44 100 Hz, the bispectral analysis is per-

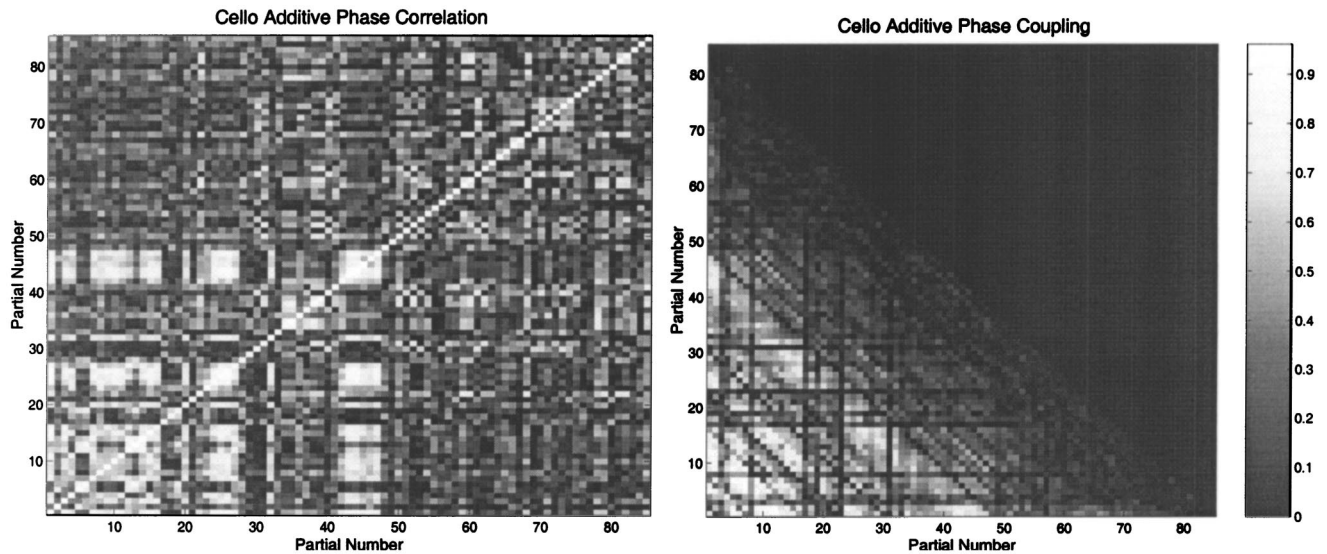


FIG. 8. Cello phase correlation and QPC analysis.

formed on a subsampled signal at 22 050 Hz. The reason for using a lower bandwidth in the bispectral analysis stems from practical considerations of memory and time requirements of the bispectral estimation algorithm. Moreover, since most of higher order statistical information appears at lower frequencies, there is no need to consider the high frequencies.

- (ii) Bispectral analysis is done on a shorter segment compared to additive phase coupling analysis. In the additive analysis the QPC estimate is averaged over a large segment that contains several hundred analysis frames, with frame step of 4 ms. This amounts to an analysis segment of an order of magnitude of few seconds. In order to compare this to the shorter averaging times of the bispectral analysis, one can evaluate the decay of the QPC as function of time in the appropriate graphs in the previous sections (90 ms in bispectral analysis corresponds to 22.5 analysis frames in QPC).
- (iii) The resolution of bispectral analysis is approximately 43 Hz. This resolution allows one to clearly see the separate peaks that correspond to the harmonics. This is in contrast to the phase-coupling graph where the values of the axes correspond to partial numbers. Thus, the partial numbers correspond, in frequency, to sampling of the bi-frequency plane at a resolution that correspond to fundamental frequency.

It is interesting and important to note the similarity between the bispectral analysis and the QPC analysis of Fig. 3. Apparently, although the QPC method and the bispectral method are different, they discover similar phenomena.

IV. THE MYSTERY OF THE CELLO: MODELING OF ASYNCHRONOUS PERIODIC PHASE DEVIATIONS

Cello produces its sound through a nonlinear bow-excitation process that excites a complex resonance body with many closely spaced resonances. Since our analysis is

applied to the final sound, we cannot separate the bow-excitation properties from the effect of the body resonances. Moreover, we would like to examine the difference between natural vibrato sound and open-string sound that has no vibrato. Accordingly, we analyzed two different cello sounds: a closed-string C4 pitch sound that contained natural vibrato and an open-string A3 sound that had no vibrato.

Applying phase analysis to string instruments reveals a particularly interesting phenomenon. Repeating the same analysis as done previously for the flute and trumpet, one finds out (see Fig. 8) that cello played with vibrato exhibits very little phase coupling among higher partials.

Considering the phase behavior of the various partials of the cello, it seems that the phase deviations have a random behavior. A closer look at specific triplets of partials reveals actually quite a smooth phase behavior, at least for the lower partials. The phase deviations are periodic, with a period corresponding approximately to the vibrato rate. This lack of coupling is different from the random phase deviations of the flute and possibly is caused by a different acoustical mechanism. In Fig. 9 the phases, phase differences, and phase coupling analysis for two triplets of cello partials are shown.

Comparing these results to analysis of an open-string cello sound (Fig. 10) reveals another interesting phenomenon. An open-string sound has a significantly higher QPC, showing strongly coupled harmonics up to partial number 60. Moreover, looking at the phase behavior (Fig. 11), one sees that the deviations of the phases among the different partials have similar trajectories, giving high phase correlations as well. These results suggest that the phase coupling and phase correlation in the cello are closely linked to the vibrato.

In order to understand the origin of the very asynchronous behavior of the phases in string instruments (Weinreich, 1997; Schumacher, 1992; McIntyre *et al.*, 1981; Beauchamp, 1974; Rodet, 1993, 1995) we have tried to model the periodic temporal variations of the phases of the cello using a source-filter model. The bow-excitation and the resonance

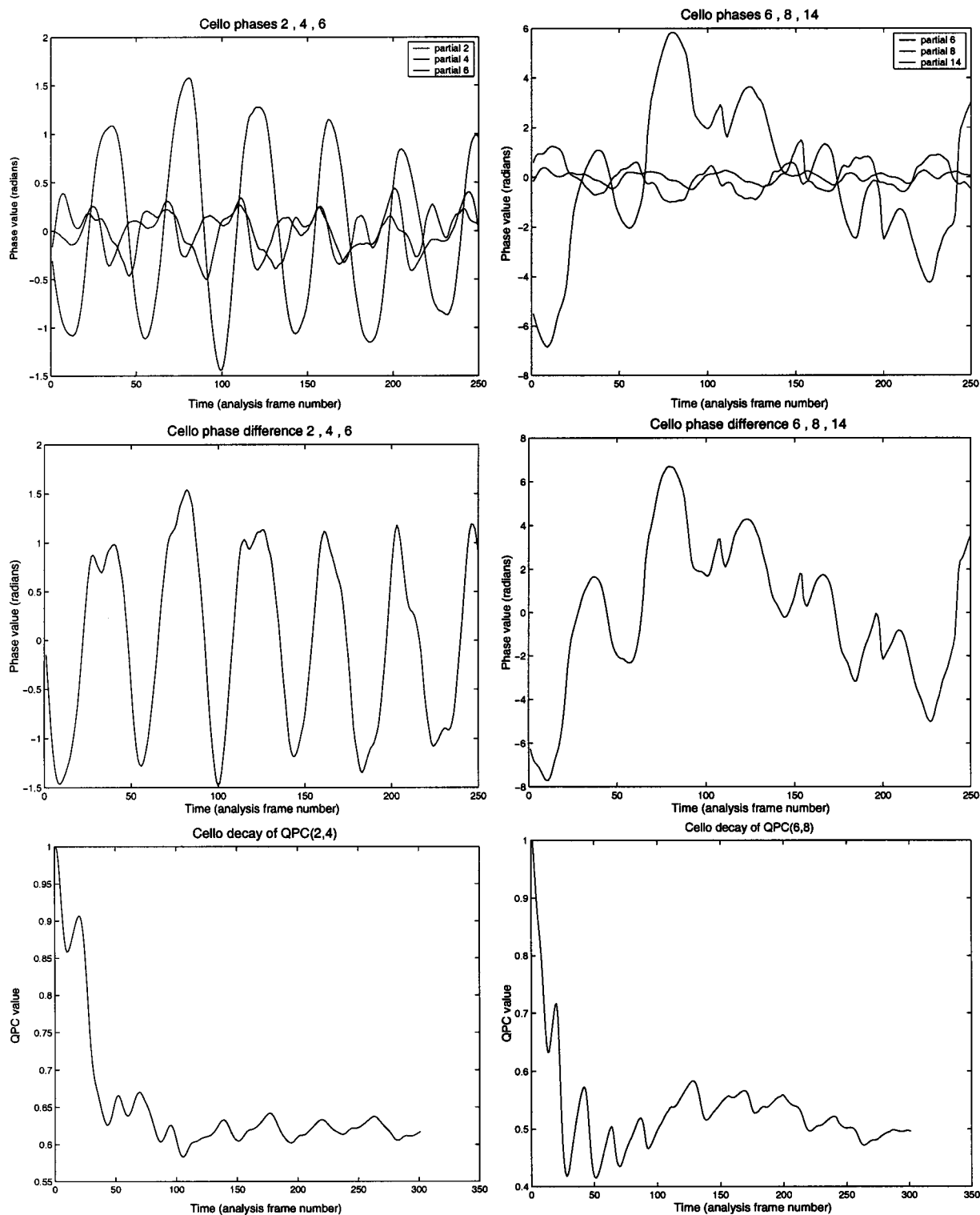


FIG. 9. Cello phases, phase differences, and QPC for two triplets of partials, numbers 2,4,6 and 8,6,14. The graphs in the left column correspond to analysis of partials 2,4,6. Right column graphs show analysis of partials 6,8,14. The different rows correspond to phases (top), phase difference (middle), and QPC analysis (bottom).

body effect are approximately modeled by a harmonic excitation that passes through a bank of closely spaced resonators. In this model the origin of the uncoupled phase behavior could be created as a result of an interaction between a time-varying excitation signal and the resonator body. Accordingly, we model it by a slowly varying frequency excitation signal, i.e., a harmonic signal with time-varying fun-

damental frequency $f_0(t)$ that simulates a vibrato, consequently filtering it by a linear filter that represents an instrument body resonance. If the instruments' body resonances, at the frequencies of the partials, are comprised of very close and narrow peaks, two situations might occur:

- (1) Relative phase shifts, up to π , could exist between dif-

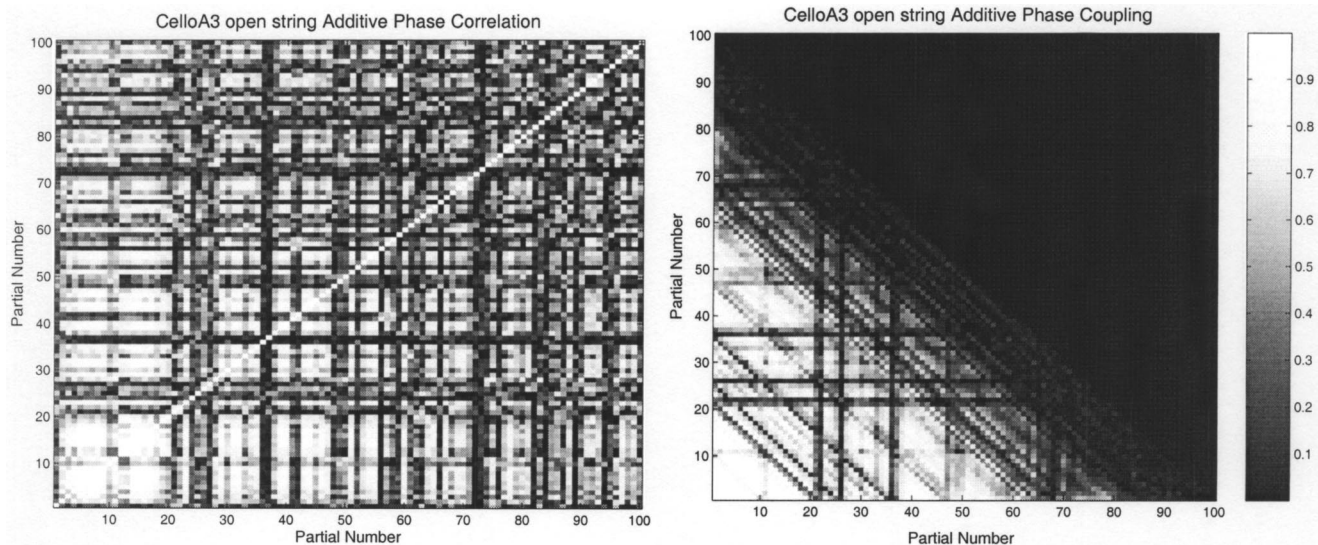


FIG. 10. Cello A3 open string phase correlation and QPC analysis.

ferent partials, due to differences in phase response of the body resonance at their corresponding frequencies.

- (2) A large phase deviation would occur for a single partial with a varying frequency, when it passes through a narrow peak of a body filter resonance.

The first case could be simulated by having harmonically related, time-varying partials, that move on the opposite slopes of very narrow resonators (McAdams and Rodet, 1988). For the second case, a single time-varying sinusoid might excite almost simultaneously a pair of very closely spaced and narrow filters. This causes two peaks in the spectrum to be present for a single partial (!) when the excitation occurs between the two resonances. In terms of phase, a jump would occur again in the middle region, i.e., when moving away from one and approaching the other resonance, it reaches a point where a phase difference is nearly 2π .

We expected, at least for the second case, that this effect would be visible in the spectrum as well. Figure 12 presents a high resolution analysis of a cello sound around its fifth partial. The sound was recorded with a close microphone to eliminate room effects (Studio Online, IRCAM). Surprisingly enough, we find that instead of having one peak, we have two very close peaks with their average frequency lying at the expected harmonic. The additive analysis in such a case would capture this as a single partial with a widely varying phase.

A. Modeling of the phase deviation

Simulation of this phenomena was done using a perfectly harmonic excitation, whose fundamental frequency variation was obtained from analysis of the fundamental frequency of an original cello signal. The body resonances were simulated using a bank of closely spaced (30 Hz) and narrow-band (BW=10 Hz) filters. A single resonator filter can be designed as a filter having two poles in conjugate location in the Z plane, located at radius R and angle θ , with corresponding transfer function

$$H(z) = \frac{1}{1 - 2R \cos(\theta)z^{-1} + R^2z^{-2}}, \quad (8)$$

so that θ corresponds to the center frequency of the resonator and R is determined according to the desired bandwidth⁴ B , with the relation between bandwidth and R being $R \approx 1 - B/2$. Inspecting the phase response of a resonator shows an abrupt phase change when going through the center peak frequency.

When a time-varying harmonic sinusoidal excitation (harmonic excitation with simulated vibrato) passes through a bank of such filters, nonsynchronous phase deviations occur between the different sinusoidal components as each component passes through a different set of filters (one must note also that the rate of frequency sweep in every partial is different and proportional to the partial number, i.e., higher frequencies move faster and “pass” on their way through more resonator filters). Figure 13 demonstrates the phases and the QPC of the simulated signal for partials pair (2,4) obtained by sinusoidal analysis. As can be seen from the figure, this behavior is very much reminiscent of the real cello behavior. Although the simulation does not constitute a “proof” that the actual cello resonance behaves in this manner, both the analysis and listening impression seem to support this assumption.

B. Discussion

In our analysis we were dealing with sounds of instruments that were played by a human player (McGill University Master Samples), thus containing vibrato and pitch variations that are typical to human playing, even when playing a single note. In the flute there was a significant vibrato, while in the trumpet there was almost no vibrato. The two instruments were analyzed at the same pitch (C4). For cello we have investigated two cases, a natural playing closed-string note (C4) that contained a vibrato and an open-string note (A3) that contained no vibrato. Our findings in all cases

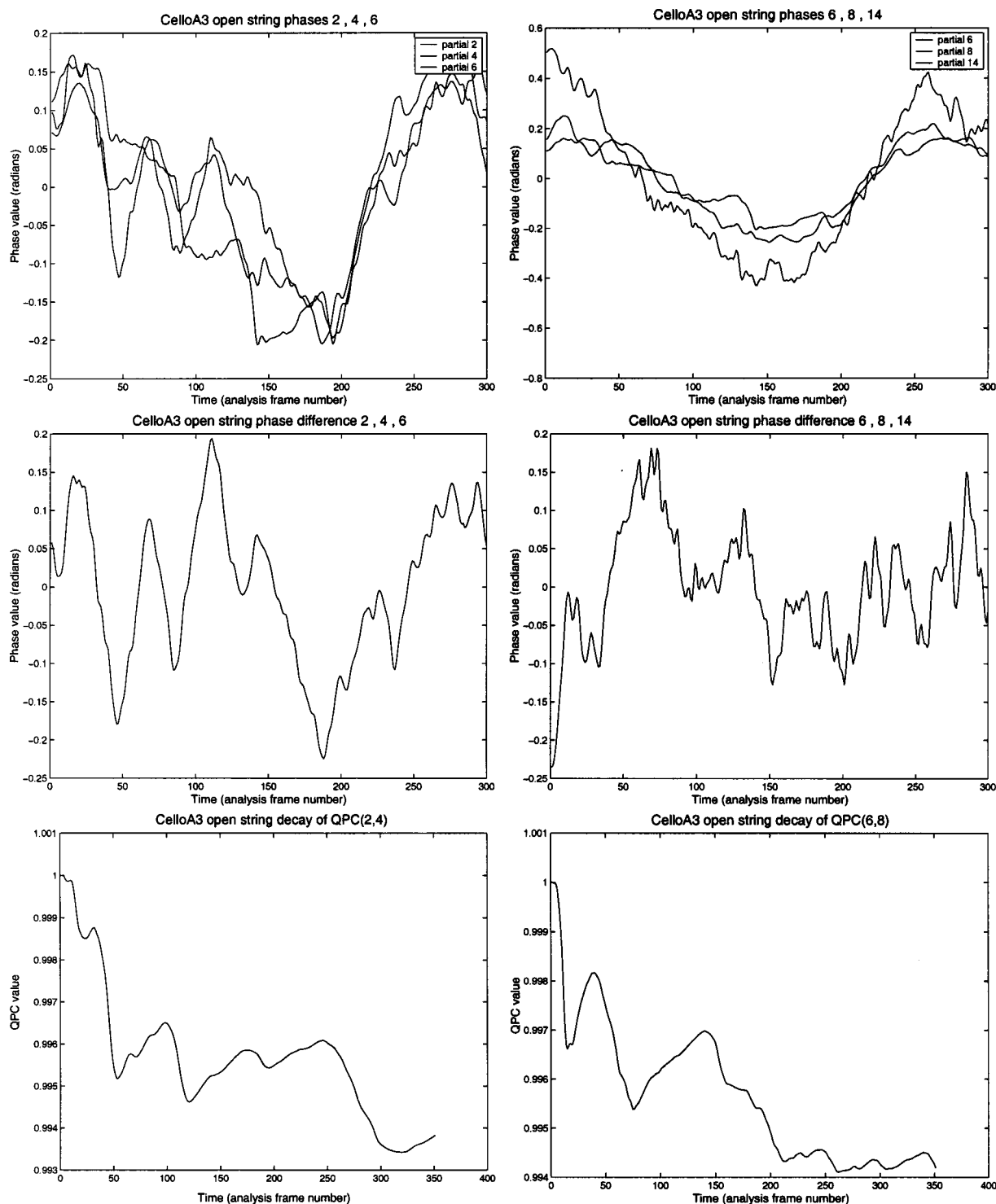


FIG. 11. Cello A3 open string: the phases, phases differences, and QPC for two triplets of partials numbers 2,4,6 and 8,6,14. The graphs in the left column correspond to analysis of the phases for partials 2,4,6. Right column graphs show analysis of partials 6,8,14. The different rows correspond to phases (top), phase difference (middle), and QPC analysis (bottom).

are consistent with what might be expected from physical considerations. In the flute vibrato is produced by variations in blowing pressure. The production of harmonics is influenced partly by jet-propagation and largely by variations in the exact intersection plane of the embouchure edge and the jet (Fletcher and Rossing, 1995). In the trumpet, blowing pressure is nearly steady and the primary excitation is caused by motion of the lips at a frequency dependent on the natural

resonance frequency. This generates higher partials by non-linear variation of lip opening and the flow velocity. In the cello, vibrato is produced by varying the length of the string, which interacts also with the nonlinear bow-excitation process. Our findings for the flute and the trumpet seem to be consistent with these physical interpretations. Moreover, for the cello we see a clear distinction between open-string and close-string with vibrato sounds. The lower QPC in the vi-

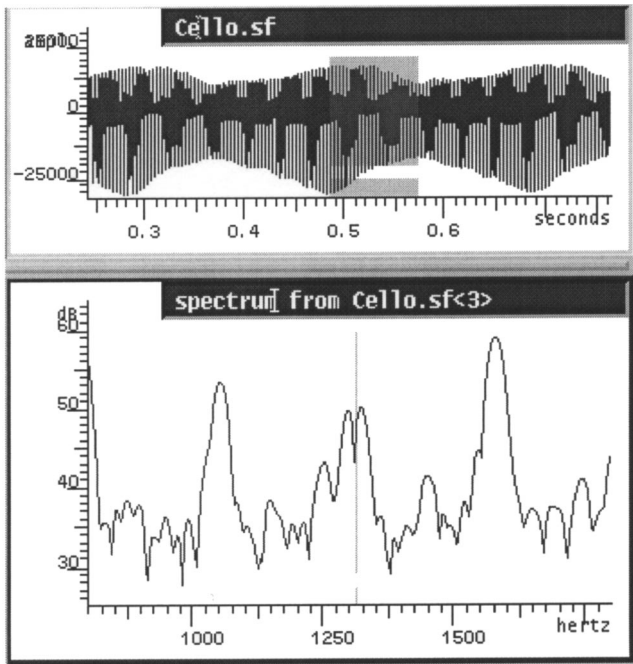


FIG. 12. The occurrence of double spectral peaks instead of a single partial in a cello signal recorded with a close microphone to eliminate room effects. The top graph shows the signal and the segment analyzed. Bottom graph shows the window Fourier transform of the corresponding signal segment (0.49–0.58 s).

brato sound could be related to the phase asynchronous effect due the closely spaced resonances of the cello body, occurring in a signal with varying fundamental frequency, as our simulation experiment suggests.

V. CONCLUSION

In this paper we have investigated the nature of a periodicities that occur in the sustained portion of harmonic instrumental sounds. Specifically, we focused on phenomena of phase deviations that occur in different partials of the sound.

Comparative study of the phase fluctuations was performed for several instruments. The results suggest that phase coupling, which is basically a nonlinear phenomenon, is a significant feature that distinctly characterizes for instance cello and flute versus trumpet sound.

Finally, a model for the mechanism that might cause the phase uncoupling effect in vibrato sounds of string instruments was suggested. A simulation of this phenomena was presented and the findings are shown to be in support of this model.

ACKNOWLEDGMENTS

We would like to thank Neville Fletcher for pointing out the relation between our analysis and the physics of musical instruments. We also thank the reviewers for the many most useful and important comments.

APPENDIX: PROOF OF THE RELATIONS BETWEEN ADDITIVE PHASE COUPLING AND HIGHER ORDER STATISTICS

We want now to establish the relationship between $B(\nu_1, \nu_2)$ and $d_3(i, k)$ in general, which is formulated also in terms of $c_3(n, m)$ for n and $m \neq 0$.

The definition of $B(\nu_1, \nu_2)$ is

$$B(\nu_1, \nu_2) \stackrel{\text{def}}{=} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} c_3(n, m) e^{-j(\nu_1 n + \nu_2 m)},$$

and $c_3(n, m)$ is defined as

$$c_3(n, m) \stackrel{\text{def}}{=} \langle x(t)x(t+n)x(t+m) \rangle$$

for zero mean x and independent of t due to third-order stationarity. We take $t=0$ for simplicity of the derivations

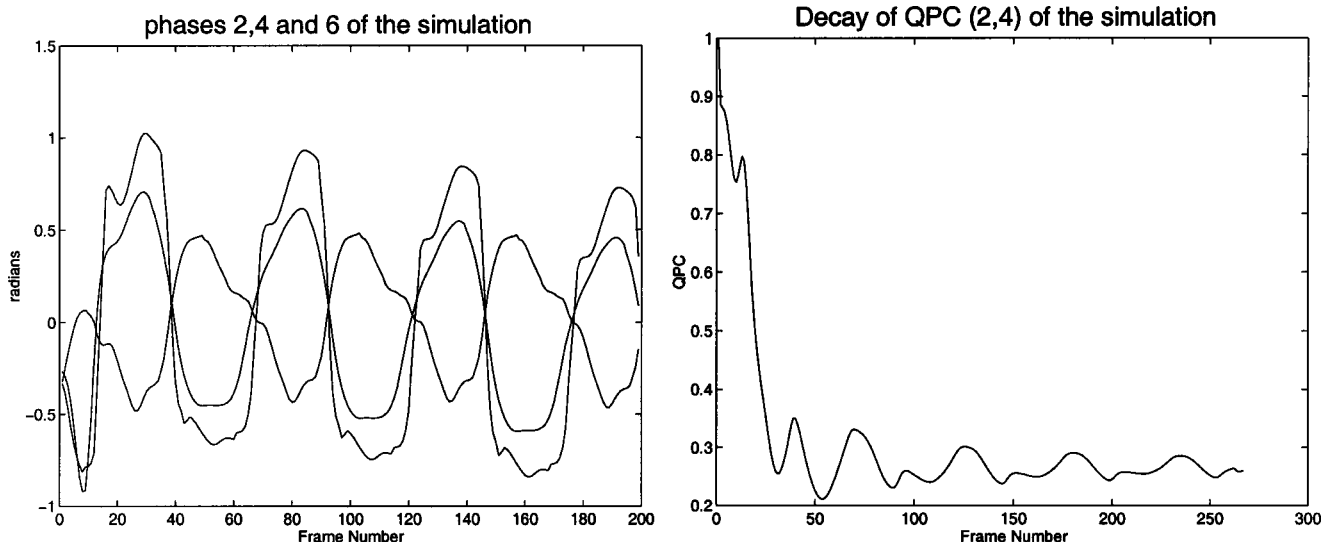


FIG. 13. Phases 2,4,6 and QPC decay of the simulation signal.

$$c_3(n, m) = \left\langle \sum_{i=-N, i \neq 0}^N e^{j(\omega_0 i \tau + \phi_i(0))} \sum_{k=-N, k \neq 0}^N e^{j(\omega_0 k(n+\tau) + \phi_k(n))} \sum_{l=-N, l \neq 0}^N e^{j(\omega_0 l(m+\tau) + \phi_l(m))} \right\rangle$$

$$= \sum_{i, k, l} \langle e^{j\omega_0(i+k+l)\tau} e^{j\omega_0(kn+lm)} \langle e^{j(\phi_i(0) + \phi_k(n) + \phi_l(m))} \rangle \rangle = \sum_{k, l} e^{j\omega_0(kn+lm)} \langle e^{j(\phi_k(n) + \phi_l(m) - \phi_{k+l}(0))} \rangle.$$

The right-hand term is called now a “lagged” d_3 , which contains the lag arguments

$$d_3(k, l; n, m) \stackrel{\text{def}}{=} \langle e^{j(\phi_k(n) + \phi_l(m) - \phi_{k+l}(0))} \rangle.$$

In such a case, we have $c_3(n, m)$ as a function of d_3 ,

$$c_3(n, m) = \sum_{k, l=-N, \neq 0}^N e^{j\omega_0(kn+lm)} d_3(k, l; n, m).$$

Let us now consider the discretization of the bispectral expression $B(\nu_1, \nu_2)$. Assuming that $c_3(n, m)$ is of finite support of size M , let us define a discrete frequency bispectrum

$$B_D(p, q) \stackrel{\text{def}}{=} \frac{1}{M^2} \sum_{m, n=-M/2+1}^{M/2} c_3(n, m) e^{-j((2\pi/M)pn + (2\pi/M)qm)}.$$

The relations between B_D and d_3 are

$$B_D(p, q) = \frac{1}{M^2} \sum_{m, n=-M/2+1}^{M/2} \left(\sum_{k, l=-N, \neq 0}^N e^{j\omega_0(kn+lm)} \times d_3(k, l; n, m) \right) e^{-j((2\pi/M)pn + (2\pi/M)qm)}.$$

From physical considerations we assume that d_3 is independent of m, n , i.e., $d_3(k, l; n, m) = d_3(k, l)$. Rewriting the above expression,

$$B_D(p, q) = \frac{1}{M^2} \sum_{k, l=-N, \neq 0}^N \sum_{m, n=-M/2+1}^{M/2} e^{j(\omega_0 k - p2\pi/M)n} \times e^{j(\omega_0 l - q2\pi/M)m} d_3(k, l),$$

and in the case that ω_0 falls into one of the FT bins, i.e., $\omega_0 = r2\pi/M$ for some r , we get

$$B_D(p, q) = \frac{1}{M^2} \sum_{k, l=-N, \neq 0}^N M^2 \delta(rk, p) \delta(rl, q) d_3(k, l) = d_3(p/r, q/r).$$

¹The proportion in magnitude of deviations among different partials corresponds to their partial number, i.e., k th partial will have k times the magnitude of the deviation of the fundamental. The reason for this proportion will become clear in later sections.

²The term “quadratic” comes from the fact that partials that are created by nonlinear interaction, such as multiplication, between lower partials have phase deviations that are a sum of phase deviations of their constituent lower partials. Since we consider statistical deviations among triplets of harmonically related partials, the corresponding measure is called quadratic phase coupling (QPC).

³One must note that correlation analysis is sensitive to the phase unwrapping. There are situations where a sound may exhibit high coupling with almost no correlation. This occurs when a signal has very few phase deviations,

resulting in small and uncorrelated phase difference signal.

⁴Bandwidth is determined as the width of magnitude response at half-power points relative to the peak value.

- Beauchamp, J. W. (1974). “Time-variant spectra of violin tones,” *J. Acoust. Soc. Am.* **56**.
- Dubnov, S., and Tishby, N. (1998). “Testing for Gaussinity and Non Linearity in the Sustained Portion of Musical Sounds,” *Recherches et Applications en Informatique Musicale*, edited by M. Chemillier and F. Pachet, Editions HERMES, pp. 212–224.
- Dubnov, S., Tishby, N., and Cohen, D. (1995). “Hearing Beyond The Spectrum,” *J. New Music Res.* **24**(4), 342–368.
- Dubnov, S., Tishby, N., and Cohen, D. (1997). “Polyspectra as Measures of Sound Texture and Timbre,” *J. New Music Res.* **26**(4), 277–314.
- Fletcher, N. H., and Rossing, T. D. (1995). *The Physics of Musical Instruments*, 3rd ed. (Springer Verlag, New York).
- Hinich, M. J. (1982). “Testing for Gaussinity and Linearity of a Stationary Time Series,” *J. Time Series Anal.* **3**(3), 169–176.
- McAdams, S., and Rodet, X. (1988). “The role of FM-induced AM in dynamic spectral profile analysis,” in *Basic Issues in Hearing*, edited by H. Duifhuis, J. Horst, and H. Wit (Academic, London), pp. 359–369.
- McGill University Master Samples. Faculty of Music, McGill University, 555 Sherbrooke St. West, Montreal, Quebec.
- McIntyre, M. E., Schumacher, R. T., and Woodhouse, J. (1981). “Aperiodicity in bowed string motion,” *Acustica* **49**, 13–32.
- Markel, J. D., and Gray, Jr., A. H. (1976). *Linear Prediction of Speech* (Springer-Verlag, Berlin).
- Mendel, J. M. (1991). “Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory,” *Proc. IEEE* **79**(3), 278–305.
- Nikias, C. L., and Mendel, J. M. (1993). “Signal Processing with Higher-Order Spectra,” *IEEE Signal Process. Mag.* July.
- Priestley, M. B. (1989). *Non-Linear and Non-Stationary Time Series Analysis* (Academic, New York).
- Rodet, X. (1993). *Nonlinear Oscillations in Sustained Musical Instruments: Models and Control* (Euromech, Hamburg).
- Rodet, X. (1995). “Oscillations, Chaos and Control of Nonlinear Dynamical Models,” NOLTA, Las Vegas.
- Rodet, X. (1997). “Musical Sound Signals Analysis/Synthesis: Sinusoidal + Residual and Elementary Waveform Models,” in *Proc. of the IEEE Time-Frequency and Time-Scale Workshop (TFTS’97)*, Coventy, UK.
- Schumacher, R. T. (1992). “Analysis of aperiodicities in nearly periodic waveforms,” *J. Acoust. Soc. Am.* **91**, 438–451.
- Serra, X., and Smith, J. O. (1989). “Spectral Modeling Synthesis: A sound analysis/synthesis system based on deterministic plus stochastic decomposition,” *Comput. Music J.* **14**(4), 12–24.
- Studio Online. IRCAM, Paris, <http://sol.ircam.fr>
- Vergez, C., and Rodet, X. (2000). “Dynamical Systems and Physical Models of Trumpet-like Instruments: Analytical Study and Asymptotical Properties,” *Acustica-acta acustica* **86**, 147–162.
- Vettori, P. (1995). “Fractional ARIMA Modeling of Microvariations in Additive Synthesis,” *Proceedings of the XI Colloquium on Musical Informatics*, Bologna.
- Weinreich, G. (1997). “Directional tone color,” *J. Acoust. Soc. Am.* **101**, 2389–2406.