Inter-component Phase Processing of Quasipolyharmonic Signals

Vasili I. Vorobiov, Daniil A. Kechik, Siarhei Y. Barysenka*

Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Abstract

The paper presents a generalization of theoretical and experimental research in the field of inter-component phase signal processing based on instantaneous phase estimates of multiple or rational frequency harmonic components. We propose to model harmonic phase of each component of quasipolyharmonic signal with consideration of relative delays that occur on different frequencies during the signal propagation. Based on the proposed harmonic phase model, it is argued the inter-component phase relations carry the information about parameters of these relative delays. We introduce the general expression for the inter-component phase relations estimates, showing their temporal constancy and invariance to the time-frequency shifts and fluctuations of the harmonic amplitudes. These properties correspond to the findings obtained for signal propagation experiments with prior knowledge of harmonic phases. Applicability of proposed estimates for processing of natural signals is justified by results of past speech processing research (including speaker identification and speech enhancement) and novel experiments on condition monitoring of industrial machines. By employing the proposed harmonic phase model, we discuss why the earlier research on speech structure using higher-order spectra techniques did not reveal the non-linear nature of speech. We carry out simple experiments on condition monitoring of industrial machines to demostrate the potential of distinguishing between different configurations of shaft misalignment based on the distribution of standard deviation of inter-component phase relations.

Keywords: phase estimation, inter-component phase relations, harmonic phase model, speech, fault diagnosis, shaft misalignment

1. Introduction

Inter-component signal processing techniques are based on the parameters estimation of multiple or rational frequency harmonic components. This paper focuses on inter-component *phase* signal processing techniques based on the *instantaneous phase* estimation of harmonic components.

Predominantly, the inter-component phase estimation has been carried out for synthetic signals. The research of modulated ultrasound waves propagation in dispersing medium revealed a novel (at that moment) method for ultrasound dispersion measurement, introducing the notion of the phase invariant [? ?]. Methods for phase difference estimation of two multiple frequency harmonic components have been widely used in radar-based object recognition systems [? ? ?]. Subsequently, these methods have been applied for the phase estimation of multiple frequency signals in underwater acoustics [?]. In telecommunications, there are examples of successful synthesis of bi-spectral organized signals for the information encoding in the phase of the third-order spectrum [?].

Considerably less attention has been dedicated to the inter-component phase estimation of natural acoustic signals, e.g. speech and acoustical vibration of rotary machines. For these signals, the instantaneous phase estimation is challenging due to the reasons outlined below.

*Corresponding author

Email address: viv314@gmail.com, ya.dan.kechik@ya.ru, siarhei.barysenka@gmail.com (Vasili I. Vorobiov, Daniil A. Kechik, Siarhei Y. Barysenka)

- Instantaneous phase is wrapped by 2π and requires a phase unwrapping procedure to be applied before processing [? , Chapter 2.3]. The phase unwrapping algorithms are non-trivial for the low SNR environments.
- It is challenging to set a threshold to distinguish between the phase of a clean signal and the noise [?].
- Estimation of instantaneous phase requires prior knowledge of the fundamental frequency (called *pitch* in speech signal processing), which increases computational complexity of processing algorithms [?].
- Some processing techniques (e.g., analysis in higher-order spectra domain) restrict the minimum length of signal realization to the value that is greater than the length of speech fragments to be analyzed [?]. In some algorithms, ensemble averaging is required, which vanishes the natural variability of instantaneous parameters of speech fragments [?]. In the meantime, these restrictions can be overcome in signal processing of acoustical vibration of rotary machines [?].
- In early research it has been argued that Fourier phase does not contribute to the speech intelligibility, as a result the phase processing has been considered of a low impact on the quality of speech recovered from noise [?].

In past decade, the phase estimation problem has received increased interest among the speech processing researchers [? ? ? ? ? ? ?]. In contrast, the phase estimation in acoustical vibration signal processing still remains a largely unexplored topic. In a recent study [?] the authors proposed an algorithm that attenuates unwanted components in vibration signal spectrum based on editing of phase spectrum solely, and, to the best of their knowledge, it was the first attempt of phase editing in the field of condition monitoring. As noted by the authors, their method is based on a similar technique earlier proposed for speech enhancement.

Motivated by developing direction of phase processing in natural acoustic signal processing applications, and emerging opportunity to leverage the advances of phase processing techniques on a cross-disciplinary level, this paper aims to generalize the research in the field of inter-component phase signal processing, carried out independently in acoustics and radio frequency signal processing.

The rest of the paper is organized as follows. In section 2, it is proposed to extend harmonic phase model with parameters describing the relative delays occurred on different frequencies during the quasipolyharmonic signal propagation. In section 3, based on the proposed harmonic phase model, it is argued the inter-component phase relations (ICPR) carry the information about parameters of these relative delays. In this section we introduce the general expression for ICPR estimates, showing their temporal constancy and invariance to the time-frequency shifts and fluctuations of the harmonic amplitudes. Section 4 outlines the results of experimental research on ICPR carried out for synthetic signals, and these results are consistent with the properties discussed in section 3. Section 5 demonstrates the advances of inter-component phase processing for speech processing applications. In this section we employ the proposed harmonic phase model to discuss why the earlier research on speech structure using higher-order spectra techniques did not reveal the non-linear nature of speech. Section 6 demostrates novel results on the potential of distinguishing between different configurations of shaft misalignment based on the distribution of ICPR standard deviation. Section 7 draws a conclusion on the work.

2. Signal model

In this paper, the *quasipolyharmonic signal model* is employed to describe the techniques and properties of the inter-component phase processing of acoustic signals. A modeled signal $x(t)$ is represented by a linear combination of quasiharmonic components with the frequencies that are multiples of the fundamental frequency F_0:

$$x(t) = \sum_{h=1}^{H} x(h,t) = \sum_{h=1}^{H} A_x(h,t) \cos\left(2\pi h F_0 t + \Phi_x(h,t)\right) \\ = \sum_{h=1}^{H} A_x(h,t) \cos \Psi_x(h,t), \tag{1}$$

where H denotes the number of quasiharmonic components, h denotes the index of a component, $A_x(h,t)$ denotes the amplitude of a component, $\Phi_x(h,t)$ denotes the harmonic phase of a component, $\Psi_x(h,t)$ denotes the instantaneous phase of a component. The functions $A_x(h,t)$ and $\Phi_x(h,t)$ are considered to be slowly varying in time compared to the harmonic frequency hF_0.

Instantaneous fluctuations $h\Delta f_0(t)$ of harmonic frequency hF_0, including the time period instability of a component, contribute to the harmonic phase $\Phi_x(h,t)$ of a respective component:

$$\Phi_x(h,t) = 2\pi h \Delta f_0(t) t + h\phi_x + \theta_x(h), \tag{2}$$

where $h\phi_x$ and $\theta_x(h)$ denote the phase shifts occurred due to the receiver's location in space and configuration of the wave propagation medium, respectively[1].

The choice of the quasipolyharmonic signal model is due to the following reasons:

- the quasiharmonic components of any combination have their frequencies related to each other as a rational number;
- it is possible to model the time fluctuations of amplitude $A_x(h,t)$ and phase $\Phi_x(h,t)$ of the natural acoustic signals.

The mixture of a clean signal $x(t)$ and the noise $\nu(t)$ is modeled as their sum:

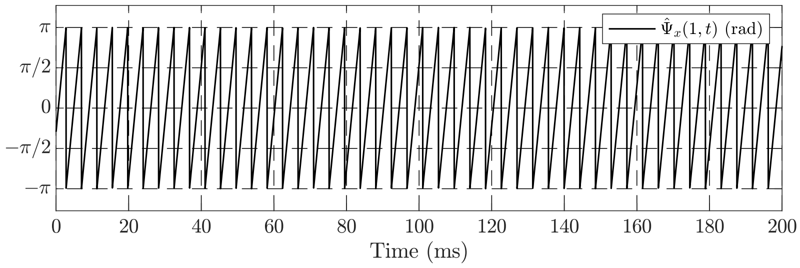
$$y(t) = x(t) + \nu(t) = \sum_{h=1}^{H} A_x(h,t) \cos \Psi_x(h,t) + \nu(t) \\ = \sum_{h=1}^{H} A_y(h,t) \cos \Psi_y(h,t) + r(t), \tag{3}$$

where $A_y(h,t)$ and $\Psi_y(h,t)$ denote the instantaneous parameters of a component degraded by the noise on frequencies $h(F_0 + \Delta f_0(t))$; and $r(t)$ denotes the residual noise exposure on all frequencies other than $h(F_0 + \Delta f_0(t))$. Estimates of amplitude $\hat{A}_x(h,t)$ and instantaneous phase $\hat{\Psi}_x(h,t)$ are used to obtain the restored signal $\hat{x}(t)$:

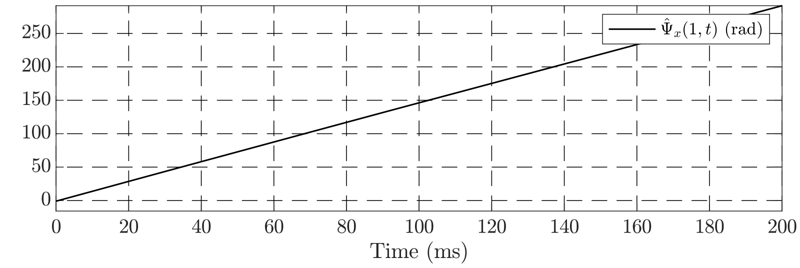
$$\hat{x}(t) = \sum_{h=1}^{H} \hat{A}_x(h,t) \cos \hat{\Psi}_x(h,t). \tag{4}$$

To visually demonstrate the difference between instantaneous, unwrapped instantaneous and harmonic phases of a natural signal, the estimates of instantaneous phase $\hat{\Psi}_x(1,t)$ and harmonic phase $\hat{\Phi}_x(1,t)$ of the quasiharmonic component with index $h = 1$ obtained from the voiced speech are depicted on Figure 1. Instantaneous phase estimate (Fig. 1a) is calculated from the filtered pitch component using the Hilbert transform. Harmonic phase estimate (Fig. 1c) is obtained by subtracting the linear phase term $2\pi F_0 t$ from the unwrapped instantaneous phase estimate (Fig. 1b). Phase unwrapping is done using the one-dimensional phase unwrapping algorithm [redacted]

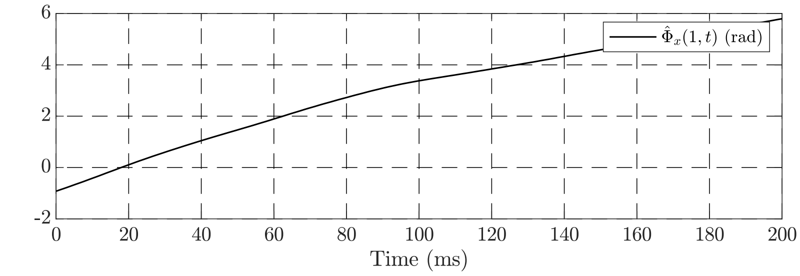
[1] Section 4 follows up with the results of experimental research on how the configuration of the wave propagation medium contributes to the phase values $\Phi_x(h,t)$.



a — instantaneous phase estimate



b — unwrapped instantaneous phase estimate



c — harmonic phase estimate

Figure 1: Estimates of instantaneous phase $\hat{\Psi}_x(1,t)$ and harmonic phase $\hat{\Phi}_x(1,t)$ of the pitch component obtained from the voiced speech fragment.

The phase information is not directly accessible due to wrapped pattern of instantaneous phase $\hat{\Psi}_x(1,t)$ presented on Fig. 1a. After unwrapping, instantaneous phase trajectory depicted on Fig. 1b gets closer to linear. However, it is dominated by the linear phase term $2\pi F_0 t$ determined solely by fundamental frequency estimate, and thus harmonic phase information still remains inaccessible. Removal of linear phase term $2\pi F_0 t$ unveils the harmonic phase trajectory $\hat{\Phi}_x(1,t)$ shown on Fig. 1c, enabling it for processing. Note the range of harmonic phase estimate is significantly lower compared to corresponding unwrapped instantaneous phase estimate, which allows application of smoothing algorithms over time axis to remove unwanted variance of harmonic phase introduced by noise.

3. Estimates of inter-component phase relations and their properties

Estimates of inter-component phase relations (ICPR estimates) are calculated as a linear combination of the instantaneous phase estimates $\hat{\Psi}_x(H(p),t)$:

$$\hat{\Theta}(t) = \sum_{p=1}^{P} S(p)K(p)\hat{\Psi}_x(H(p),t), \text{ given } \sum_{p=1}^{P} S(p)K(p)H(p) = 0, \tag{5}$$

where P denotes the number of the estimated instantaneous phase functions considered for the calculation; $K(p) \in \mathbb{Q}_{>0}$ denotes the constant multiplier for the instantaneous phase estimate at index p; $H(p)$ denotes the index h of the quasiharmonic component in (4) corresponding to the instantaneous phase estimate at index p; $S(p) \in \{-1, +1\}$ denotes the sign of the $K(p)$ multiplier. The phase unwrapping procedure should be taken on $\hat{\Psi}_x(H(p), t)$ prior calculation.

In order to describe the properties of ICPR estimates, let's formulate the expression for the instantaneous phase $\Psi_x(h, t)$ considering the equations (1) and (2):

$$\Psi_x(h, t) = 2\pi h(F_0 + \Delta f_0(t))t + h\phi_x + \theta_x(h). \tag{6}$$

Let's assume the SNR level is high enough, such that $\hat{\Psi}_x(h, t) \to \Psi_x(h, t)$. Substitution of the equation (6) into (5) gives the following expression for the ICPR estimate:

$$\begin{aligned}
\hat{\Theta}(t) &= \sum_{p=1}^{P} S(p)K(p)\hat{\Psi}_x(H(p), t) \\
&= 2\pi(F_0 + \Delta f_0(t))t \sum_{p=1}^{P} S(p)K(p)H(p) + \phi_x \sum_{p=1}^{P} S(p)K(p)H(p) \\
&+ \sum_{p=1}^{P} S(p)K(p)\theta_x(H(p)) \\
&= \sum_{p=1}^{P} S(p)K(p)\theta_x(H(p)) = const,
\end{aligned} \tag{7}$$

since the multiplier $\sum_{p=1}^{P} S(p)K(p)H(p)$ evaluates to zero according to the condition in (5).

The expression (7) allows to formulate the following properties of ICPR estimates assuming the high SNR levels:

- **Temporal constancy** assuming the phase shifts $\theta_x(h)$ occurred due to the configuration of the wave propagation medium are time independent. ICPR estimates are zero in a medium without dispersion, where $\theta_x(h) = 0$.

- **Time-shift independence** [? ?] (cancellation of the phase shift $h\phi_x$). ICPR estimates are independent to the reference phase in contrast to instantaneous phase estimates, that require prior estimation of the reference phase.

- **Frequency-shift independence** [?] (cancellation of $F_0 + \Delta f_0(t)$). ICPR estimates are not affected by the fundamental frequency fluctuations and the ADC frequency instability. Cancellation of the linear phase term $2\pi F_0 t$ reduces the dynamic range of the estimates.

- **Independence to the instantaneous amplitude fluctuations** [?] appeared due to the fluctuations of a measuring instrument frequency response. This property is ensured by the absence of $A_x(h, t)$ in the equations (5) and (7). At the same time, the level of the instantaneous amplitudes must be above the threshold that ensures the reliable harmonic tracking for the instantaneous phase estimation.

Aforementioned properties allow to estimate ICPR for the non-stationary signals and reduce the environmental effects on the final results [? ?]. The precision is gained due to the relative nature of the measurements and the cancellation of the interfering environmental components. ICPR estimates convey the information about the configuration of the wave propagation medium defined by the relative delays occurred on different frequencies due to the dispersion of the medium [?], reflections [?] and multipath propagation [?].

4. Inter-component phase relations of synthetic signals

4.1. Phase invariant of the modulated oscillation

A series of the experiments to study ICPR has been carried out for modulated signals [**? ? ? ?**]. Lets define the common expression for modulated signals for better representation of results. Amplitude modulated signals can be expressed as sum of carrier and side harmonics:

$$\begin{aligned} x_{AM}(t) &= (1+\sum_{p=1}^{P} A_m \cos\Phi_m(t) A_c \cos\Phi_c(t) \\ &= A_c\cos\Phi_c(t) + A_m A_c \sum_{p=1}^{P}\cos(\Phi_c(t)-\Phi_m(t)) + A_m A_c \sum_{p=1}^{P}\cos(\Phi_c(t)+\Phi_m(t)), \end{aligned} \tag{8}$$

where $\Phi_c(t) = \omega_c t + \phi_0$ and $\Phi_m(t) = \omega_m t + \phi_m$ are full phases of carrier and modulation signals, A_c and A_m are their amplitudes, ω_c and ω_m are their frequencies, ϕ_0 and ϕ_m are their phases respectively. Frequency modulated signals are usually approximated using decomposition of exponent e^x with imaginary argument x that is periodic at $-\pi \le x \le \pi$:

$$\begin{aligned} x_{FM}(t) &= A_c Re\{e^{j\Phi_c(t)} e^{jm\sin\Phi_m(t)}\} = A_c Re\{e^{j\Phi_c(t)} \sum_{k=-\infty}^{\infty} J_k(m) e^{jk\Phi_m(t)}\} \\ &= A_c \sum_{k=-\infty}^{\infty} J_k(m)\cos(\Phi_c(t)+k\Phi_m(t)) = A_c\cos\Phi_c(t) \\ &+ A_c\sum_{k=1}^{\infty} J_k(m)\cos(\Phi_c(t)+k\Phi_m(t)) + A_c\sum_{k=1}^{\infty} J_{-k}(m)\cos(\Phi_c(t)-k\Phi_m(t)) \\ &= A_c\cos\Phi_c(t) + A_c\sum_{k=1}^{\infty} J_k(m)\cos(\Phi_c(t)+k\Phi_m(t)) \\ &+ A_c\sum_{k=1}^{\infty}(-1)^k J_k(m)\cos(\Phi_c(t)-k\Phi_m(t)) = A_c\cos\Phi_c(t) \\ &+ A_c\sum_{k=1}^{\infty} J_k(m)\cos(\Phi_c(t)+k\Phi_m(t)) + A_c\sum_{k=1}^{\infty} J_k(m)\cos(\Phi_c(t)-k\Phi_m(t)+k\pi). \end{aligned} \tag{9}$$

Oscillations with amplitude or single-tone angle modulation **??** may be represented by the quasipolyharmonic signal model (1) in case if the harmonic amplitudes are identically zero $A_x(h,t) = 0$ for $h \notin [H_{\min}, H_{\max}]$[2], and the harmonic phases $\Phi_x(h,t)$ expressed as follows:

$$\Phi_x(h,t) = \begin{cases} \Phi_c(t) - \Phi_m(h,t) + \xi(h) & \text{if } h < H_c, \\ \Phi_c(t) & \text{if } h = H_c, \\ \Phi_c(t) + \Phi_m(h,t) & \text{if } h > H_c; \end{cases} \tag{10}$$

$$\xi(h) = \begin{cases} \pi(H_c - h) & \text{for single-tone angle modulation,} \\ 0 & \text{for amplitude modulation,} \end{cases} \tag{11}$$

where $\Phi_c(t)$ and $\Phi_m(h,t)$ denote the phase of the carrier signal and the modulating signal respectively, $H_{\min}$ and $H_{\max}$ denote the indices of the components with the lowest and the highest frequencies in the

[2]It is important to note the harmonic amplitudes $A_x(h,t)$ may be identically zero for some indices $h \in [H_{\min}, H_{\max}]$ if the corresponding components are absent in the spectrum of the modulated signal.

spectrum of the modulated signal, $H_c = \frac{1}{2}(H_{\min} + H_{\max})$ denotes the index of the carrier component, $h \in [H_{\min}, H_{\max}] \subseteq [1, H]$.

Let's consider an ICPR estimate (5), where functions $S(p)$ and $K(p)$ are defined as follows:

$$S(p) = \begin{cases} +1 & \text{if } H(p) \neq H_c, \\ -1 & \text{if } H(p) = H_c; \end{cases} \tag{12}$$

$$K(p) = \begin{cases} 1/2 & \text{if } H(p) \neq H_c, \\ \frac{1}{2\Delta H}(H_{\max} - H_{\min}) & \text{if } H(p) = H_c, \end{cases} \tag{13}$$

where $\Delta H = H_c - H_{\text{lower}} = H_{\text{upper}} - H_c$, defining H_{lower} and H_{upper} as the indices of the non-zero amplitude components closest to H_c in the lower and the upper spectrum bands of the modulated signal.

For simplicity, we consider the spectrum of the modulated signal consists of three components with indices H_{lower}, H_c and H_{upper}. In this scenario, $H_{\min} = H_{\text{lower}}$, $H_{\max} = H_{\text{upper}}$, $2\Delta H = H_{\text{upper}} - H_{\text{lower}}$, and finally, $K(p) = 1$ for $H(p) = H_c$ in (13). Then, ICPR estimate is defined as follows:

$$\hat{\Theta}(t) = \frac{\hat{\Psi}_x(H_{\text{lower}}, t) + \hat{\Psi}_x(H_{\text{upper}}, t)}{2} - \hat{\Psi}_x(H_c, t). \tag{14}$$

The expression (14) has all the properties of ICPR estimates described in the section 3. If the phase functions $\Phi_c(t)$ and $\Phi_m(h, t)$ defined in the form of (2), the expression (14) transforms to the following:

- for the amplitude modulated signal

$$\hat{\Theta}(t) = \frac{\theta_x(H_{\text{lower}}) + \theta_x(H_{\text{upper}})}{2} - \theta_x(H_c) = const; \tag{15}$$

- for the single-tone angle modulated signal

$$\hat{\Theta}(t) = \frac{\theta_x(H_{\text{lower}}) + \theta_x(H_{\text{upper}}) + \pi}{2} - \theta_x(H_c) = const. \tag{16}$$

As follows from the expressions (15) and (16), ICPR estimates of the form (14) carry the information not only about the parameters of the wave propagation medium, but also about the modulation method.

The results obtained during the modulated wave propagation experiments are consistent with the properties described above. In 1953, V. A. Zverev proposed a method of elastic wave dispersion measurement based on the effect of the modulation method change during the wave propagation in the dispersing medium [**?** **?**]. Zverev introduced the notion of *phase invariant* of the modulated oscillation, which corresponds to the negative ICPR estimate defined by equation (14). The value of the phase invariant is independent to the time reference point for any carrier frequency. It was noted the value of the phase invariant changes during the propagation of the modulated wave in the dispersive medium due to the additional phase shift appearing for the different components. The modulation method alternately changes from the amplitude modulation to the phase modulation. The proposed method of measurement features gained precision compared to the achievable precision of the previously developed methods due to the relative nature of the measurements.

Further, the phase invariant notion stands for the following ICPR estimate

$$PI(H_1, H_2, H_3, t) = \frac{\hat{\Psi}(H_1, t) + \hat{\Psi}(H_3, t)}{2} - \hat{\Psi}(H_2, t), \tag{17}$$

calculated for three components from signal model (4) with frequencies $\{f_1, f_2, f_3\}$, where the following condition holds:

$$\begin{cases} f_1 = H_1 F_0, & \text{where } H_1 = 1, 2, ... \\ f_2 = H_2 F_0, & \text{where } H_2 = H_1 + 1, H_1 + 2, ... \\ f_3 = H_3 F_0, & \text{where } H_3 = 2H_2 - H_1. \end{cases} \tag{18}$$

The applications of the phase invariant estimates include the modelling of the phase-dependent processes during the acoustic waves propagation in a nonlinear medium without dispersion [**?**], and the research of the EHF radio-frequency path properties [**?**].

4.2. Phase quasi-invariant

Methods for phase difference estimation of two multiple frequency harmonic components with frequencies $\{f_0, Hf_0\}$ (where $H \in \mathbb{N}$, $H \neq 1$) have been widely used in radar-based object recognition systems [**? ? ?**]. Subsequently, these methods have been applied for the phase estimation of multiple frequency signals in underwater acoustics [**?**]. For ultra wide band digital signal processing, theoretical generalization of inter-component phase estimation methods for a pair of quasiharmonic components $\{H_1 f_0, H_2 f_0\}$ (where $H_1, H_2 \in \mathbb{N}$, $H_1 < H_2$) was introduced in [**?**]. One of the estimates proposed there was later called *phase quasi-invariant* [**?**].

A phase quasi-invariant estimate is defined by the following expression

$$PQI(H_1, H_2, t) = \hat{\Psi}(H_1, t) - \frac{\hat{\Psi}(H_2, t) \cdot H_1}{H_2}, \tag{19}$$

where $H_1 F_0$ and $H_2 F_0$ (given $H_1 < H_2$) denote the frequencies of two components from signal model (4). All the properties of ICPR estimates described in the section 3 hold for the expression (19).

The phase quasi-invariant estimate was employed in performance evaluation of radar-based object recognition systems under the multipath propagation conditions [**?**]. In sonar experiments with dualpath propagation, it was shown the value of phase quasi-invariant can change only abruptly if mutual position of source and destination points changes [**?**].

The range of the phase quasi-invariant depends on the ratio between frequencies $H_1 F_0$ and $H_2 F_0$ [**? ? ? ? ?**]. Let's show this by considering the multi-valued function $\hat{\Psi}(h, t)$ from (19):

$$\hat{\Psi}(h, t) = Princ\big\{\hat{\Psi}(h, t)\big\} + 2\pi N, \tag{20}$$

where $N \in \mathbb{Z}$ and $Princ\big\{\hat{\Psi}(h, t)\big\}$ denotes the principal value of $\hat{\Psi}(h, t)$ in the interval $[-\pi; \pi)$. Plugging (20) into (19) yields:

$$\begin{aligned} PQI(H_1, H_2, t) &= Princ\big\{\hat{\Psi}(H_1, t)\big\} + 2\pi N_1 - \frac{H_1}{H_2} \cdot \left(Princ\big\{\hat{\Psi}(H_2, t)\big\} + 2\pi N_2 \right) \\ &= Princ\left\{ \hat{\Psi}(H_1, t) - \frac{\hat{\Psi}(H_2, t) \cdot H_1}{H_2} \right\} + \underbrace{2\pi N_1}_{\Delta_1} - \underbrace{\frac{2\pi H_1}{H_2} N_2}_{\Delta_2}, \end{aligned} \tag{21}$$

where $N_1, N_2 \in \mathbb{Z}$. Since $|\Delta_2| < |\Delta_1|$ for any $N_1 = N_2$, the decrease of the phase quasi-invariant range compared to the range of (20) is determined by $|\Delta_2|$, not $|\Delta_1|$. Hence, the component with Δ_1 vanishes in (21):

$$PQI(H_1, H_2, t) = Princ\left\{ \hat{\Psi}(H_1, t) - \frac{\hat{\Psi}(H_2, t) \cdot H_1}{H_2} \right\} + \frac{2\pi H_1}{H_2} N. \tag{22}$$

The range $\left[\frac{-\pi H_1}{H_2}, \frac{\pi H_1}{H_2}\right)$ denotes the *unambiguous definition range* of $PQI(H_1, H_2, t)$ that decreases when ratio H_1/H_2 decreases.

4.3. Phase of higher-order spectra

The theory of higher-order spectra has been studied since 1960s pioneering in mathematical statistics [**?**]. The most beneficial features of this theory for signal processing applications were outlined in [**?**]:

- suppression of Gaussian noise and the noise with symmetric probability density function in signal parameters estimation, signal classification and signal detection applications;
- magnitude and phase response recovery of signals and systems due to availability of phase information in the third-order spectra and higher;

- quadratic phase coupling detection as an indicator of non-linear processes occurred during signal generation.

Some definitions from the higher-order spectra theory necessary for the further context are outlined below.

A third-order spectrum, or *bi-spectrum*, for a finite energy real deterministic signal $x(n)$ defined as follows [?]:

$$M_3^x(\omega_1, \omega_2) = X(\omega_1)X(\omega_2)X^*(\omega_1 + \omega_2), \tag{23}$$

where $X(\omega)$ and $X^*(\omega)$ denote the Fourier transform of $x(n)$ and its complex conjugation, respectively.

The *bi-phase* of bi-spectrum (23) is defined as $\angle M_3^x(\omega_1, \omega_2)$. In the following, the bi-phase is considered for a harmonically related triplet of components with frequencies:

$$\begin{cases} f_1 = H_1 F_0, & \text{where } H_1 = 1, 2, ... \\ f_2 = H_2 F_0, & \text{where } H_2 = H_1 + 1, H_1 + 2, ... \\ f_3 = H_3 F_0, & \text{where } H_3 = H_1 + H_2. \end{cases} \tag{24}$$

The bi-phase estimate for a triplet (24) from model (4) is given by:

$$BiPh(H_1, H_2, H_3, t) = \hat{\Psi}(H_1, t) + \hat{\Psi}(H_2, t) - \hat{\Psi}(H_3, t) \tag{25}$$

and has all the properties of ICPR estimates described in the section 3.

The methods of synthesis of bi-spectral organized signals for information encoding, transmission and decoding applications have shown robust performance for signals with amplitude and phase distortion occurred due to multipath propagation [?]. Bi-spectrum signal processing methods were applied for solving the inverse non-linear acoustics problem of reconstructing the initial signal spectrum using the measured spectral and bispectral characteristics of the received signal on short tracks [?].

Having bi-phase evaluated to zero for harmonically related triplet serves as an indicator of quadratic phase coupling if the magnitude of bi-spectrum evaluates to non-zero [?]. According to (7), ICPR estimates convey the information about the configuration of the wave propagation medium, hence the bi-phase estimate is not identically zero in a dispersive medium. In this case, the approach of quadratic phase coupling detection proposed in [?] may not indicate the non-linear effects occurred during the process of signal generation[3].

5. Inter-component phase relations of speech signals

5.1. Estimation of inter-component phase relations in voiced speech

The research of phase structure of voiced speech introduced the estimates of *relative phase shift, RPS* [?] and *phase distortion, PD* [?] in speech signal processing.

Originally, RPS was introduced as a representation of the phase information in harmonic speech models [?]:

$$RPS(h, t) = \hat{\Psi}(h, t) - h\hat{\Psi}(1, t), \tag{26}$$

noting their temporal constancy and cancellation of linear phase shift. The follow-up studies have evaluated the RPS performance for synthetic speech detection in speaker verification systems [?].

In the work [?] dedicated to explore various phase representations for voice production modelling, it was noted the variance of RPS estimate (26) increases towards high frequencies because of increased index h in the term $h\hat{\Psi}(1, t)$, which makes RPS representation not convenient for directly characterizing the source shape in mid and high frequencies. To overcome this limitation, PD estimate was proposed, which doesn't have aforementioned issue:

$$PD(h, t) = RPS(h + 1, t) - RPS(h, t) = \hat{\Psi}(h + 1, t) - \hat{\Psi}(h, t) - \hat{\Psi}(1, t). \tag{27}$$

[3]The topic of non-linear effects occurring in the speech signal generation process is discussed in section 5.3.

As shown in our earlier works [? ?], the RPS and PD estimates represent special cases of phase quasi-invariant and bi-phase estimates, respectively:

$$PQI(1,h,t) = -\frac{RPS(h,t)}{h}, \tag{28}$$

$$BiPh(1,h,h+1,t) = -PD(h,t). \tag{29}$$

It was also noted in [? ?] the RPS defines the instantaneous phase relation between the fundamental frequency component $H_1 = 1$ and its higher harmonics, whereas the phase quasi-invariant doesn't limit H_1 to any specific harmonic number.

Figure 2 depicts the spectrogram (Figure 2a) and ICPR estimates—$PQI(2,4,t)$ (Figure 2b), $BiPh(1,3,4,t)$ (Figure 2c) and $PI(3,4,5,t)$ (Figure 2d)—of sustained sequence of vowels */aeiou/* uttered by female speaker. The instantaneous phase estimates $\hat{\Psi}_x(h,t)$ were obtained by applying Hilbert transform to the quasiharmonic components filtered on frequencies corresponding to harmonic numbers $h \in [1,5]$. The fundamental frequency estimate was obtained using Halcyon algorithm [?].

It can be seen from Figure 2 the ICPR estimates are close to constant or exhibit low variance on the whole length of a single vowel (for example, see the fragment 2–3.5 seconds representing */io/* vowels on Figure 2b–d).

Fluctuations of ICPR estimates are caused by the filtering procedure affected by harmonics magnitude. For example, on Figure 2b–d the component $h = 4$ has low level of magnitude for the fragments corresponding to vowels */i/* (interval of 2–3 seconds) and */u/* (interval of 3.75–4.25 seconds), which leads to aforementioned fluctuations on the corresponding time intervals.

The ICPR estimates change abruptly on transitions from one vowel to another (for example, see $PQI(2,4,t)$ and $PI(3,4,5,t)$ on Figure 2b and 2d in the middle of 2–3.5 seconds interval) even though instantaneous phase of corresponding harmonics are continuous and have no gaps. This fact indicates the phase-frequency response of the vocal tract changes on transitions from one vowel to another during articulation.

The results of various research show the ICPR estimates carry the information about the voice features of a speaker, and therefore applicable in automatic speaker recognition systems [? ? ?].

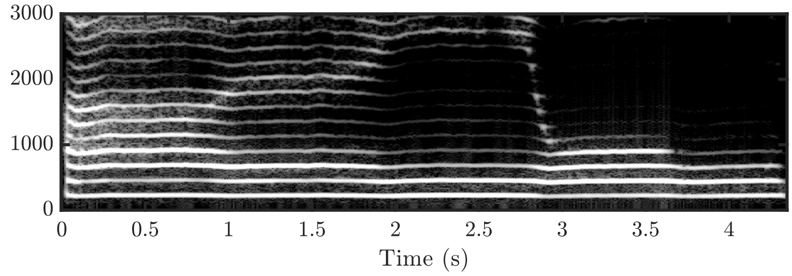
5.2. Speech enhancement

In past decade, research of phase-aware speech processing methods for speech enhancement applications has received increased interest due to the reconsidered vision on how phase impacts the perception of speech [? , Chapter 4], [?]. Modification of phase spectrum alone, without processing of magnitude part of the spectrum, was shown to improve quality and intellegibility of speech degraded by noise.

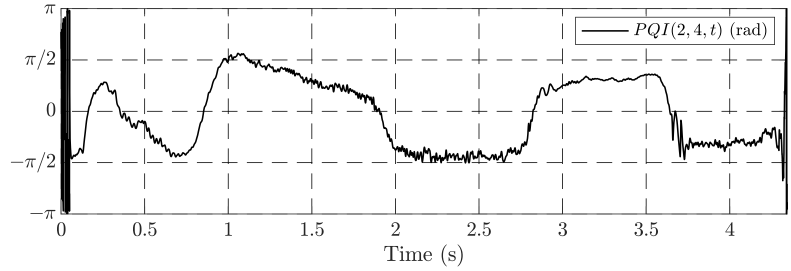
In our earlier works [? ?] it is assumed that joint estimation of the phase of several components can increase the accuracy of the phase estimation of individual components, thus improving the quality of noisy speech. In these works the noise reduction algorithms were proposed that employ temporal smoothing on phase invariant, phase quasi-invariant, and bi-phase estimates in voiced speech fragments to reduce the variance of these estimates introduced by noise. The instantaneous phase of components used during phase enhanced speech synthesis are calculated based on the smoothed inter-component phase relations estimates. A similar approach based on smoothing of PD estimates (27), (29) was proposed in [?].

The performance of algorithms proposed in [? ?] was evaluated using various metrics: perceptual evaluation of speech quality (PESQ) [?], short-time objective intelligibility (STOI) [?] and unwrapped root mean square estimation error (UnRMSE) [?]. The proposed algorithms showed improved perceived quality, speech intelligibility and phase estimation accuracy in most of considered noise scenarios, including non-stationary noise. In some noise scenarios these algorithms improved quality and intelligibility jointly, which is difficult to achieve for majority of conventional speech enhancement algorithms [?].

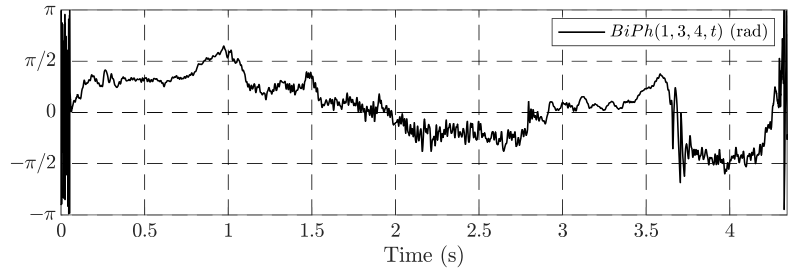
Figure 3 depicts the spectrograms of a clean speech (Figure 3a), noisy speech (Figure 3b) and enhanced by PQI $\bar{h} = 2$ algorithm [? ?] speech (Fig 3c). Enhanced speech spectrogram shows decrease in the noise impact in the lower frequency band, contributing to restored harmonic structure. Other speech records processed by algorithms [? ?] available in [?].



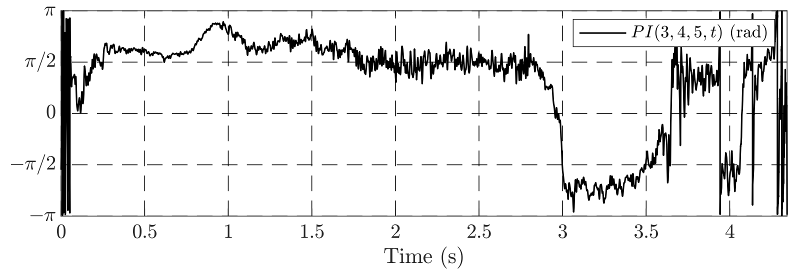
a — spectrogram. Ordinate — hertz



b — phase quasi-invariant estimate, $PQI(2, 4, t)$

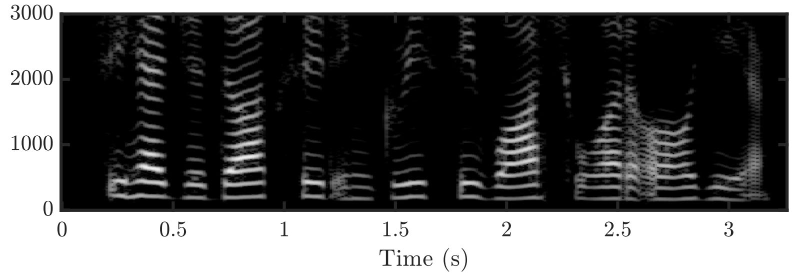


c — bi-phase estimate, $BiPh(1, 3, 4, t)$

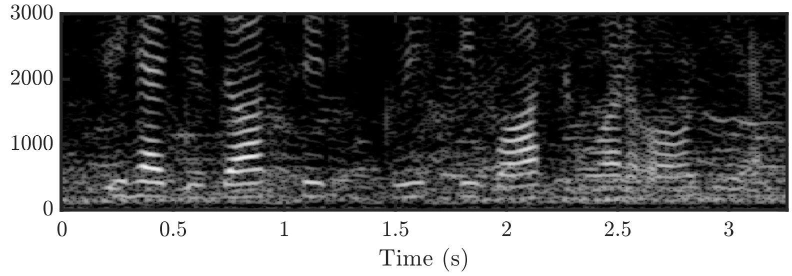


d — phase invariant estimate, $PI(3, 4, 5, t)$

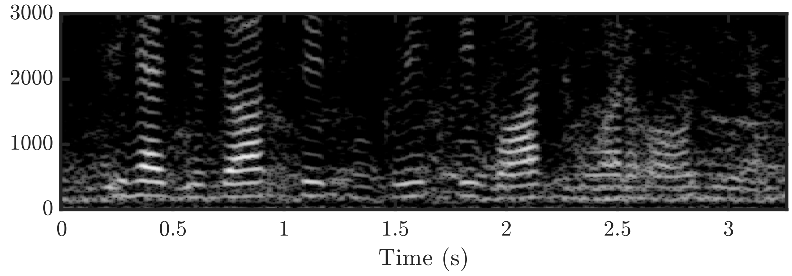
Figure 2: ICPR estimates for a voiced speech fragment. Female speaker uttering sustained sequence of vowels */aeiou/*.



a — spectrogram of clean speech



b — spectrogram of noisy speech



c — spectrogram of enhanced speech

Figure 3: Speech enhancement by PQI $\bar{h} = 2$ algorithm proposed in [**?** **?**]. Clean speech — utterance *She had your dark suit in greasy wash water all year* from TIMIT speech database [**?**]. Noise signal — babble noise from NOIZEUS database [**?**]. SNR = 5 dB. Ordinate — hertz

5.3. Quadratic phase coupling of components in speech production

Detection of non-linear coupling of speech signal components is a relevant task in speech production fundamentals research, and analysis in the higher-order spectra domain is recognized as a conventional technique of detecting the non-linearities in data (see Section 4.3).

The works [**?** **?**] summarized the features of voiced speech discoverable by means of bi-spectral processing in laryngeal pathology detection and speaker recognition applications. In these works, authors focused on detection of shape asymetry of pulse train excited by vocal cords, which is considered as a symptom of pathology.

Some results on employing bi-spectral processing for enhancement of speech degraded by Gaussian noise were presented in [**?**], where authors report improved speech quality for SNR levels not exceeding 6 dB. Other works [**?** **?** **?** **?**] unveiled the benefits of bi-spectral processing for voiced/unvoiced decision, speech reconstruction from noisy observation and speaker identification.

The work [**?**], dedicated to explore the non-linearities in speech signals by means of bi-spectrum, admitted the results of quadratic coupling detection presented in earlier works [**?** **?**] had been misinterpreted by their authors. In particular, function arguments maximizing the bi-spectrum magnitude are interpreted in [**?** **?**] as frequencies of components produced by quadratic coupling, however, the test for zero bi-phase at these frequencies is not attempted. This interpretation (without attempting zero bi-phase test) is valid only if phase is uniformly distributed on $[-\pi, \pi)$ interval. It is known the uniform distribution of phase does not hold for voiced speech [?], therefore, zero bi-phase test is requred to reliably detect the quadratic coupling case [? ?]. The results reported in [?] took zero bi-phase test into consideration, however, did not identify voiced speech to exhibit quadratic phase coupling of components. Therefore it was stated the non-linear effects does not occur in voiced speech production, unlike other studies that assert the opposite (e.g. [?]).

The fact of non-zero bi-phase estimate of voiced speech is consistent with the proposed harmonic phase model $\Phi_x(h, t)$ (2) that takes into account configuration of the wave propagation medium, and with the bi-phase estimate of a voiced speech fragment depicted in Figure 2c. Configuration of speech propagation medium defined by physical properties of vocal tract adds additional phase shifts, contributing to final bi-phase value according to (7). In other words, even if voiced speech is a non-linear product of individual components, the propagation of resulting wave in the vocal tract contributes with additional delays, captured by bi-phase. As a result, the value of bi-phase may differ from zero, obscuring the non-linear nature of voiced speech that may occur on earlier stages of speech production.

6. Inter-component phase relations of vibration signals

6.1. Motivation

Defect development monitoring of rotary machines is based on analysis of their acoustical vibration. A rotary machine unit generates vibration excited by periodical forces arising from mass imbalance, shock processes and unit geometry tolerance. Oscillations caused by these forces produced on the fundamental frequency and its multiples due to non-linear and parametric effects arising during state degradation of machine, allowing to model the vibration signal using (1) for components with frequencies up to 1–2 kHz [?].

Conventional approach for condition monitoring aims to extract features from the frequency domain representation of vibration signal. Faults are detected by observing the relations between magnitudes of fundamental frequency components (representing, for example, shaft rotation frequency, or ball spin frequency of a rolling bearing, etc.) and its multiples [?]. Some defects, e.g. shaft imbalance and shaft misalignment, produce components at the same frequencies, interferring with the ones generated by normal machine operation. In these scenarios trending of parameters over time is required to resolve ambiguity, bringing analysis to the time and spatial domains. Thus, estimation of ICPR may provide additional information to improve reliability of condition monitoring techniques.

In our earlier work [?] the ICPR estimates of rolling bearings vibration data were analyzed. The span of ICPR trajectories was observed in the wide interval of $[-\pi, \pi)$, which does not meet the expectation (7). The standard deviation of ICPR estimates showed some correspondence to the defective bearing states, but it was not reliable in general due to high variance of ICPR estimates. At that moment, theoretical explanation of these observations had not been found. In retrospective, we assume the reason of high variance is the low SNR level of analyzed data (ranging from -5 dB to 10 dB), preventing to perform the reliable phase estimation.

In the next sections, we examine the potential of employing the ICPR estimates for shaft parallel misalignment monitoring. The experiments are carried out on the original dataset of vibration signals representing different shaft misalignment configurations [?]. We address the following questions:

- How does SNR affect temporal constancy of ICPR estimates? Does this property hold for vibration signals in general for high SNR levels greater than 10 dB?
- How does the standard deviation of ICPR estimates distribute between normal and defective state of equipment? Is there a potential to distinguish between different operating conditions using the distribution of standard deviation?

6.2. Shaft misalignment experiment: Setup and methods

The original shaft parallel misalignment dataset [?] used in this work consists of 140 fragments of vibration acceleration data recorded at the test stand. The test stand has an elastic support (unlike the majority of industrial machines) and a shaft imbalance. Two shafts are joined by the flexible jaw coupling. Each recorded fragment consists of 215 722 samples of vibration acceleration data sampled at $F_s = 25\,597$ Hz, which amounts to approximately 8.43 seconds of data per fragment. The dataset description is given in Table 1.

Table 1: Description of original shaft parallel misalignment dataset [?] used in this work. Each recorded fragment consists of 215 722 samples of vibration acceleration data sampled at $F_s = 25\,597$ Hz, which amounts to approximately 8.43 seconds of data per fragment.

Misalignment type	Count	Axis	File names
No misalignment	20	Horizontal	`Misalignment_norm/1.mat...20.mat`
0.5 mm horizontal	24	Horizontal	`Misalignment_horz_parall/1.mat...24.mat`
1.15 mm horizontal	24	Horizontal	`Misalignment_horz_parall/25.mat...48.mat`
0.34 mm vertical	24	Vertical	`Misalignment_vert_parall/1.mat...24.mat`
0.8 mm vertical	24	Vertical	`Misalignment_vert_parall/25.mat...48.mat`
1.09 mm vertical	24	Vertical	`Misalignment_vert_parall/49.mat...72.mat`

Estimation of SNR and standard deviation of ICPR estimates is performed using the following procedure[4]:

- Original acceleration signals are processed using numerical integration and double-integration to obtain velocity and displacement signals, respectively. Vibration velocity and displacement signals are considered to better suit for analysis of low frequency components.
- Shaft rotation frequency $F_0 \approx 24.68$ Hz is refined using spectral interference method [? ?] for each data set record.
- Three harmonics of shaft rotation frequency are extracted using Fourier filtration to further calculate their instantaneous phases. The filter bandwith of 2 Hz is selected to decrease the effect of spectral leakage and maximize harmonic SNR levels.
- Instantaneous phases of selected components are obtained using Hilbert transform. $PQI(1, 2, t)$, $PQI(1, 3, t)$ and their standard deviation are computed[5].
- SNR level of each component is estimated assuming the ideal harmonics-in-noise model: the noise power density is considered to be uniformly distributed, and harmonics are represented by weighted Dirac deltas. Estimated average of the noise power is then subtracted from the estimated power of a harmonic (both expressed in decibels) to determine the harmonic SNR value. The lowest harmonic SNR value among the components included in particular ICPR estimate is considered as its SNR.

6.3. Shaft misalignment experiment: No misalignment vs. horizontal parallel misalignment

In this section, we analyze the vibration data from "No misalignment", "0.5 mm horizontal" and "1.15 mm horizontal" buckets of the data set described in Table 1.

Figure 4 depicts the magnitude spectra of individual vibration velocity signals from each bucket and corresponding $PQI(1, 2, t)$ estimates. For "No misalignment" condition (Figure 4a) the magnitude of the

[4] The implementation of some algorithms (except proprietary solutions) is available for download [?].

[5] First and last 5% of $PQI(1, 2, t)$ and $PQI(1, 3, t)$ samples were excluded from standard deviation computation. This is to avoid the influence of artifacts produced by Gibbs phenomenon on the final statistics demonstrated in the next sections. In the meantime, we have considered using the full length data for standard deviation computation (not presented in this paper): no notable impact on the final statistics has been observed.

second harmonic of shaft rotation frequency ($2F_0$ = 49.36 Hz) is lower than the one for other defective conditions (Figure 4b and Figure 4c), which is considered as a defect indicator in conventional condition monitoring approach. The $PQI(1, 2, t)$ trajectory for "No misalignment" condition (Figure 4a) spans over the wider range of values (the span exceeds 0.6 rad) compared to the range of $PQI(1, 2, t)$ trajectories for other defective conditions (Figure 4b and Figure 4c), where the span does not exceed 0.11 rad for each trajectory.

Figure 5 presents the overall statistics for selected buckets of data. Left panel of Figure 5 depicts the standard deviation $\sigma[PQI(1, 2, t)]$ of computed $PQI(1, 2, t)$ estimates vs. the corresponding SNR levels. It can be observed "No misalignment" data exhibit higher standard deviation levels compared to those from "0.5 mm horizontal" and "1.15 mm horizontal" buckets. It is possible to visually distinguish two groups of points corresponding to presence or absence of parallel misalignment in the data despite the SNR variability. The probability density functions of $\sigma[PQI(1, 2, t)]$ depicted on the right panel of Figure 5 show the probability of standard deviation values to be less than 0.05 rad is 0.75 and 0.83 for "0.5 mm horizontal" and "1.15 mm horizontal" buckets, respectively, whereas for "No misalignment" data the probability of standard deviation values to fall into this range is 0.

Aforementioned results obtained for "No misalignment", "0.5 mm horizontal" and "1.15 mm horizontal" buckets of data are consistent with the proposed expectation (7) for ICPR estimates to exhibit low variance for high SNR levels. For the given data set, distribution of the standard deviation $\sigma[PQI(1, 2, t)]$ justifies the ability of proposed method to distinguish between presence or absence of parallel misalignment. The explanation of this fact is that the lower SNR levels of the second harmonic contribute to degraded harmonic phase estimation accuracy, resulting in higher variance of $PQI(1, 2, t)$. The absence of parallel shaft misalignment is reflected in a lower SNR level of the second harmonic, contributing to the higher variance of $PQI(1, 2, t)$ compared to the one obtained in the presence of misalignment.

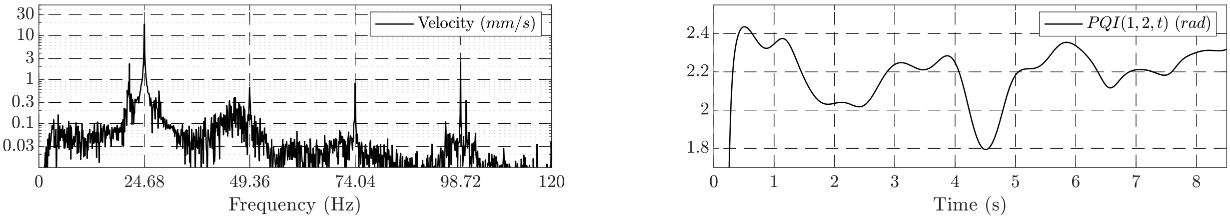
6.4. Shaft misalignment experiment: various configurations of vertical parallel misalignment

In this section, we analyze the vibration data from "0.34 mm vertical", "0.8 mm vertical" and "1.09 mm vertical" buckets of the data set described in Table 1.

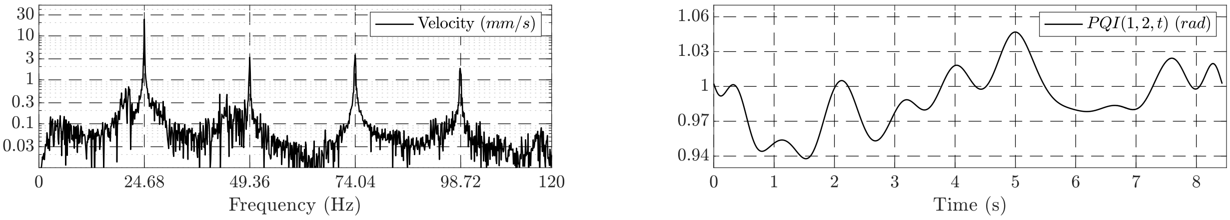
Figure 6 depicts the magnitude spectra of individual vibration displacement signals from each bucket and corresponding $PQI(1, 3, t)$ estimates. The ICPR between the component on the shaft rotation frequency and its third multiple (captured by $PQI(1, 3, t)$) is considered for this experiment as the increased level of the third harmonic magnitude is a feature of defective condition according to conventional condition monitoring approach. The magnitude of the third harmonic ($3F_0$ = 74.04 Hz) is either approaching to (Figure 6a) or greater than (Figure 6b and Figure 6c) the magnitude of the shaft rotation frequency component (F_0 = 24.68 Hz). The $PQI(1, 3, t)$ trajectories of these individual signals span over different ranges for various misalignment configurations, however, their span vary in the ranges that are close to each other.

Figure 7 presents the overall statistics for selected buckets of data. Left panel of Figure 7 depicts the standard deviation $\sigma[PQI(1, 3, t)]$ of computed $PQI(1, 3, t)$ estimates vs. the corresponding SNR levels. In the whole, the standard deviation of the $PQI(1, 3, t)$ estimates for each bucket does not exceed 0.07 rad. An important observation is that the data from "0.8 mm vertical" and "1.09 mm vertical" buckets is grouped in the separate areas, which creates an opportunity to distinguish between various misalignment configurations using the distribution of the standard deviation of ICPR. The peaks of the distributions for "0.8 mm vertical" and "1.09 mm vertical" buckets are located in different bins, which is depicted on the right panel of Figure 7. The distribution for "0.34 mm vertical" bucket overlaps with the distributions for other buckets, thus making this bucket indistinguishable from others by means of $PQI(1, 3, t)$ distribution.

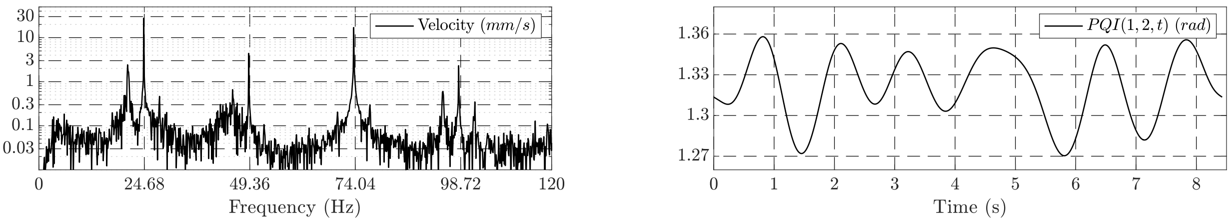
The fact of different distributions of the standard deviation $\sigma[PQI(1, 3, t)]$ obtained for various misalignment configurations could be explained by the impact of non-synchronous vibration components (such as additive harmonic noise) or intermodulation that modify harmonic phases of considered components due to non-linear effects arising under different misalignment configurations. In this case, according to the proposed harmonic phase model (2), the $\theta_x(h)$ parameter may change independently on different frequencies, contributing to the difference of values of ICPR estimates (7).



a — no parallel misalignment



b — 0.5 mm parallel misalignment



c — 1.15 mm parallel misalignment

Figure 4: Magnitude spectrum and $PQI(1,2,t)$ estimates of vibration velocity signal (horizontal direction) for various shaft misalignment configurations. $PQI(1,2,t)$ calculated for components at shaft rotation frequency $F_0 = 24.68$ Hz and its multiple $2F_0 = 49.36$ Hz.

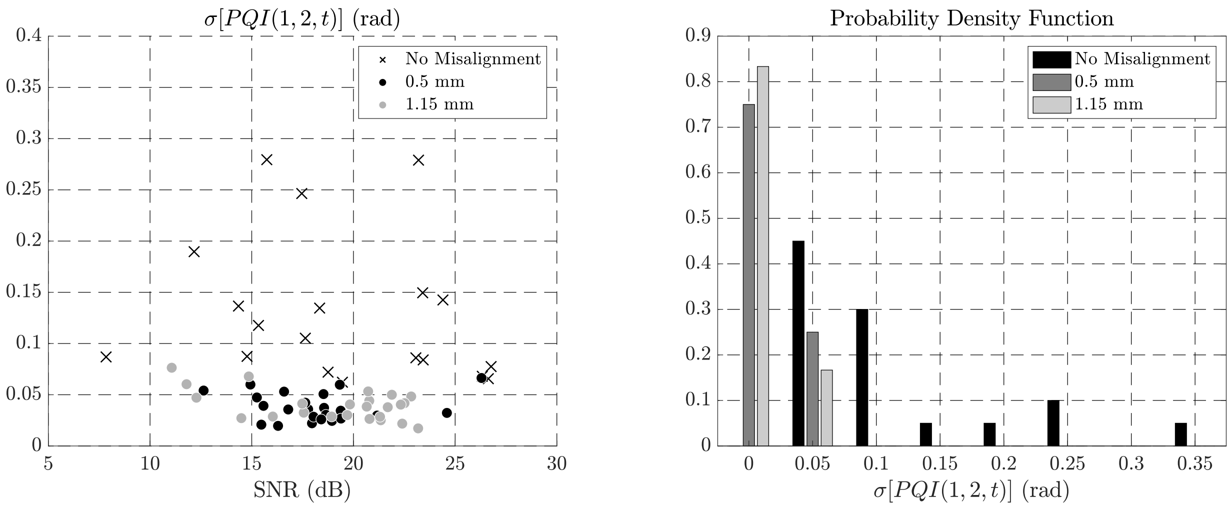
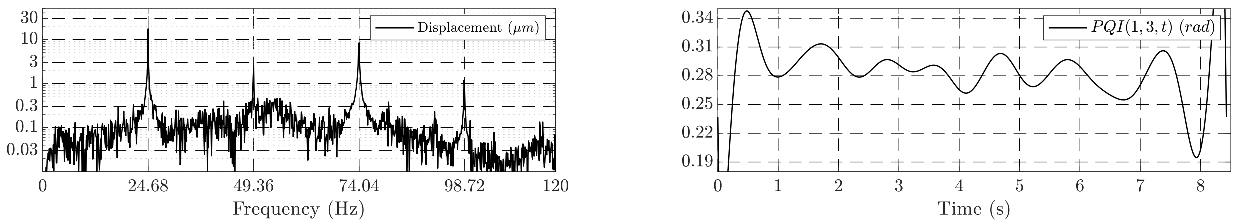
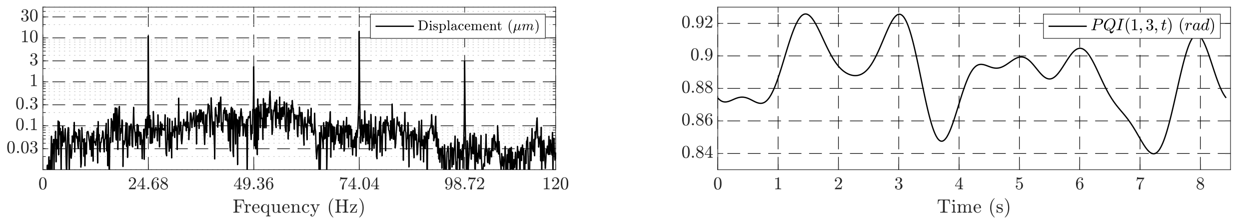


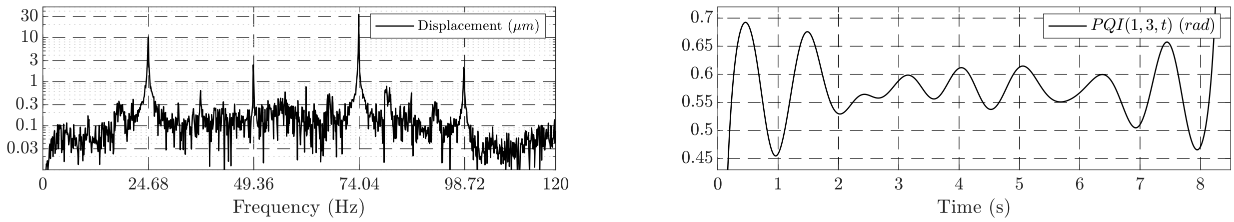
Figure 5: Left panel — standard deviation $\sigma[PQI(1,2,t)]$ vs. SNR (horizontal direction) for various shaft misalignment configurations. Right panel — probability density function of $\sigma[PQI(1,2,t)]$ corresponding to the same data points depicted in left panel.



a — 0.34 mm parallel misalignment



b — 0.8 mm parallel misalignment



c — 1.09 mm parallel misalignment

Figure 6: Magnitude spectrum and $PQI(1,3,t)$ estimates of vibration displacement signal (vertical direction) for various shaft misalignment configurations. $PQI(1,3,t)$ calculated for components at shaft rotation frequency $F_0 = 24.68$ Hz and its multiple $3F_0 = 74.04$ Hz.

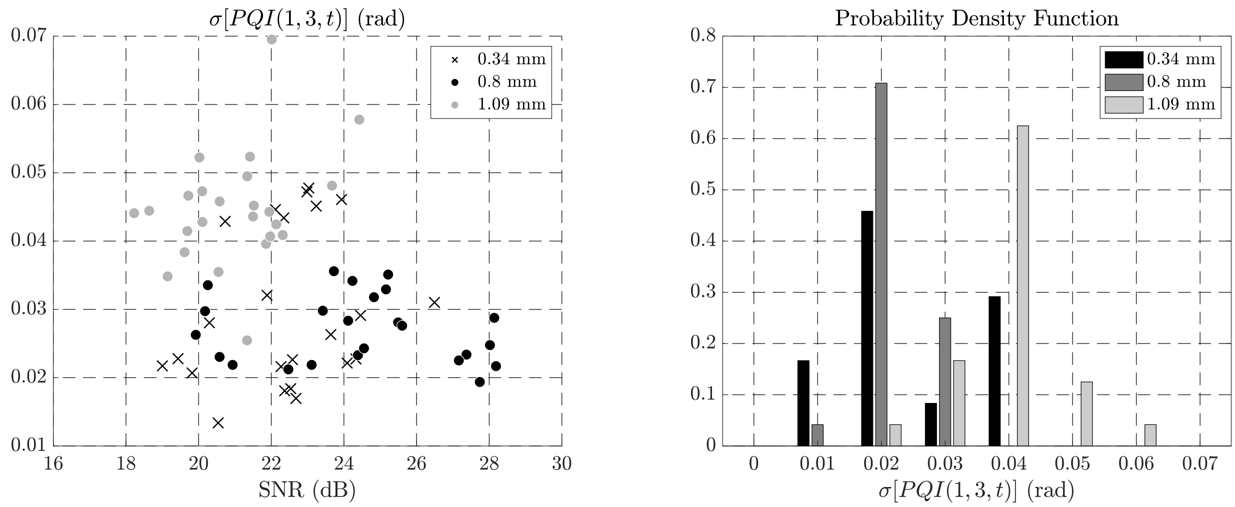


Figure 7: Left panel — standard deviation $\sigma[PQI(1,3,t)]$ vs. SNR (vertical direction) for various shaft misalignment configurations. Right panel — probability density function of $\sigma[PQI(1,3,t)]$ corresponding to the same data points depicted in left panel.

6.5. *Discussion*

6.5.1. *Impact of SNR on temporal constancy of inter-component phase relations estimates*

The results of experiments carried out in sections 6.3 and 6.4 are obtained for the data featuring the high SNR levels ranging from 8 dB to 28 dB. Among 140 data points representing standard deviation values of $PQI(1, 2, t)$ and $PQI(1, 3, t)$, more than 90% of points have standard deviation less than 0.1 rad, whereas remaining less than 10% of points have standard deviation not exceeding 0.4 rad. In contrast, in our previous work [?] we analyzed vibration signals with lower SNR levels ranging from −5 dB to 10 dB, where we obtained higher levels of standard deviation of ICPR: among 110 signals in total, only about 1% had standard deviation of ICPR less than 0.1 rad, and about 34% — less than 0.4 rad.

Based on these observations, we conclude the SNR levels below 10 dB drastically affect the temporal constancy property of ICPR estimates due to degradation of phase estimation accuracy. Therefore, additional signal enhancement step is required in these noise conditions prior the phase estimation. For vibration signals with SNR levels exceeding 10 dB, the ICPR estimation can be attempted without prior signal enhancement, providing the ICPR trajectories approaching the constant values as expected by (7).

6.5.2. *Impact of equipment operating condition on the standard deviation of inter-component phase relations estimates*

The results presented in section 6.3 indicate clear difference in distribution of standard deviation of $PQI(1, 2, t)$ calculated for defective (shaft misalignment) and normal (no misalignment) operating conditions. The second harmonic of shaft rotation frequency component has lower level of magnitude in normal state compared to defective state, contributing to lower SNR level and degraded harmonic phase estimation accuracy. This fact yields the decreased variance of $PQI(1, 2, t)$ for defective state compared to normal state. We assume the phenomenon of varying standard deviation of ICPR may indicate not only presence or absence of shaft misalignment defect in particular, but any other defect in general if it introduces increased harmonic SNR of components captured by particular ICPR.

Distributions of standard deviation of $PQI(1, 3, t)$ for different shaft misalignment configurations evaluated in section 6.4 show the potential of distinguishing between different misalignment configurations by comparing statistical characteristics of these distributions. Variance of ICPR for different misalignment configurations can be due to impact of non-synchronous vibration components or intermodulation effects that modify harmonic phase of components independently to each other within different range of values, depending on the level of misalignment.

7. Conclusion

In this paper we explored how inter-component phase signal processing techniques could be leveraged in natural acoustic signal processing applications, while existing applications already include synthetic acoustic signal processing and radio frequency signal processing.

We argue the inter-component phase relations carry the information about parameters of relative delays occurred on different frequencies during the quasipolyharmonic signal propagation. To demonstrate that, we proposed to extend harmonic phase model with parameters describing these relative delays, introduced the general expression for inter-component phase relations estimates, and derived their properties: temporal constancy and invariance to the time-frequency shifts and fluctuations of the harmonic amplitudes. These properties turn out to be consistent with the results of past experiments obtained for synthetic acoustic and radio frequency signals, and speech signals.

By employing the proposed harmonic phase model, we discussed why the earlier research on speech structure using higher-order spectra techniques did not reveal the non-linear nature of speech. We argue that even if voiced speech is a non-linear product of individual components that theoretically should produce a peak of bi-spectrum magnitude and identically zero bi-phase estimate, the propagation of resulting wave in the vocal tract adds additional delays. As a result, the value of bi-phase may differ from zero, obscuring the non-linear nature of voiced speech that may take place on earlier stages of speech production.

We examined the potential of employing inter-component phase relations estimates in vibration signal processing, namely, for shaft parallel misalignment monitoring applications. We report the inter-component phase relations estimates exhibit lower variance for signals corresponding to parallel shaft misalignment defect, and higher variance for signals corresponding to normal state with no misalignment. This is due to lower level of magnitude of shaft rotation component harmonics observed in normal state compared to defective state, which contributes to lower SNR level and degraded harmonic phase estimation accuracy. Furthermore, distributions of standard deviation of inter-component phase relations estimates for different shaft misalignment configurations demonstrate the potential of distinguishing between different misalignment configurations by comparing statistical characteristics of these distributions. This can be due to impact of non-synchronous vibration components or intermodulation effects that modify harmonic phase of components independently to each other within different range of values, depending on the level of misalignment.

The results reported in this work demonstrate that inter-component phase signal processing techniques offer a subtle framework for extracting fundamentally new information from acoustic signals of various nature. Possible areas of future research include bioacoustics, ocean acoustics and geology acoustics.

8. Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

The authors would like to thank VibroBox Minsk Research and Development office https://www.vibrobox.com for providing the shaft misalignment data set used in this work.