



## OPEN ACCESS

## EDITED BY

Xuebo Zhang,  
Northwest Normal University, China

## REVIEWED BY

Zechao Liu,  
Shijiazhuang Tiedao University, China  
Wenbin Liu,  
Guangzhou University, China

## \*CORRESPONDENCE

Lei Xiao  
xiaolei@wzu.edu.cn

## SPECIALTY SECTION

This article was submitted to  
Ocean Observation,  
a section of the journal  
Frontiers in Marine Science

RECEIVED 17 July 2022

ACCEPTED 15 August 2022

PUBLISHED 30 September 2022

## CITATION

Yu S, Xiao L and Sun W (2022)  
Parameter measurement of aircraft-  
radiated noise from a single  
acoustic sensor node in three-  
dimensional space.  
*Front. Mar. Sci.* 9:996493.  
doi: 10.3389/fmars.2022.996493

## COPYRIGHT

© 2022 Yu, Xiao and Sun. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use,  
distribution or reproduction is  
permitted which does not comply with  
these terms.

# Parameter measurement of aircraft-radiated noise from a single acoustic sensor node in three-dimensional space

Shuwen Yu<sup>1</sup>, Lei Xiao<sup>2\*</sup> and Weitao Sun<sup>3</sup>

<sup>1</sup>Information Technology Center, Wenzhou University, Wenzhou, China, <sup>2</sup>Intelligent Information Systems Institute, Wenzhou University, Wenzhou, China, <sup>3</sup>The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, China

A line spectrum presents the form of a narrow-band time-varying signal due to Doppler effect when the single hydrophone node observes flight-radiated noise. The modulation law of the time-varying signal contains a large number of feature information of moving targets, which can be used for detection and classification. This paper studies the possibility of using instantaneous frequency measurements from the hydrophone node to improve the precision of the flight parameter estimates when the source spectrum contains a harmonic line of constant frequency. First of all, we build up and analyze the underwater sound field excited by the aircraft using the ray theory model; then the Doppler shift in the two isospeed media, which is caused by the aircraft, is established; finally, a robust time–frequency transform describes the time–frequency distribution of the received signal, and a geometric approach solves the flying parameters.

## KEYWORDS

underwater acoustic detection, Doppler effect, instantaneous frequency estimation, frequency estimation, white noise

## 1 Introduction

In most turboprop aircrafts, engines and propellers, as a power plant, produce periodic sounds (Chirico et al., 2018; Aygun et al., 2022) since they originate in mechanical rotation mechanisms. However, as these pressure waveforms are not purely sinusoidal, their spectrum consists of a fundamental frequency and higher-order Fourier components at multiples of the fundamental. The frequency of the first harmonic is between 15 and 100 Hz, and as many as a dozen higher harmonics may be detectable up to about 1 kHz. In terms of energy, the frequency of fundamental is a significant component of the source (Boashash, 2015; Zhang et al., 2022). The relative motion between the aircraft and hydrophone results in Doppler shift in the received

signal. Under the condition of two isospeed sound propagation media, the Doppler shift depends on the altitude and velocity of the aircraft, the depth of the hydrophone, and many more. Therefore, obtaining the precise Doppler shift curve is the premise of solving the cruising parameters.

The modeling analysis and the theoretical research of underwater soundings by scholars are mainly in the two fields of a single hydrophone or an array of hydrophones. In the case of individual hydrophones, Ferguson used the ray theory to analyze the acoustic propagation model of air–water media and derived the over-the-top flight parameters of an air target based on the Doppler shift formula in the two-dimensional domain. With the continuous advancement of underwater acoustic detection equipment, vector hydrophones with “ $\infty$ ” type directivity have been applied to detect aerial targets. Compared with traditional scalar hydrophones, vector hydrophones can obtain not only the scalar field information of the signal of interest but also the vector field information. Using a reasonable mathematical model to process vector information, it is possible to determine the target orientation and other flight parameters. In terms of hydrophone arrays, Ferguson and Lo use wide-aperture hydrophone arrays to obtain the azimuth information of aerial target and perform parameter estimation and positioning of helicopters flying in the air, realizing the expansion of target motion information from a two-dimensional domain to three-dimensional domains. Peng Zhaohui *et al.* proposed a robust and reliable line spectral detection method based on hydrophone arrays in view of the interference problem of underwater detection of sound sources of aerial motion. In these studies, the underwater sounding algorithm based on the traditional single hydrophone can only estimate the sound source flight parameter problem in the two-dimensional plane, that is, assume that the helicopter just flew over the top of the hydrophone, without taking into account information such as yaw distance in the three-dimensional domain. Single-vector hydrophones can estimate azimuth information, but there are directivity blind spots which are more susceptible to background noise interference in practical applications. Although the hydrophone array can estimate the azimuth information, it has a high cost and is inconvenient to arrange and not suitable for small-volume underwater detection platforms.

The obvious directivity of acoustic ray presents in the seawater sound field excited by an aircraft. For the-same-level plane in the seawater, the maximum sound pressure lies in position directly below the aircraft, and the greater horizontal offset distance is, the less sound pressure is. For the same vertical line directly below the aircraft in the seawater, the greater the depth causes the less gradient of attenuation pressure. For the same vertical line non-directly below the aircraft in the seawater, with increasing depth, sound pressure increases until reaching a maximum at a certain depth, then with the further increasing of depth, the sound pressure decreases at its lowest level. Generally,

time–frequency transform can project the time-varying signal in the time domain into the joint function of time and frequency in the two-dimensional plane. Currently, there are two methods of time–frequency analysis (Fang et al., 2016), and one is linear decomposition using kernel function, such as short-time Fourier transform, wavelet transform, and similar methods (Zhou et al., 2020; Aksenovich, 2020; Vandana and Arokiasamy, 2022). Similar to the main component of the signal, the linear combination of kernel function can represent it sparsely. The second one is quadratic time–frequency distribution, such as the Wigner–Ville distribution (Zcza et al., 2021; Kumar and Prasad, 2022). Without the interference of the window function, this method avoids constraints between time resolution and frequency resolution. The time–bandwidth product can reach the lower bound of the uncertainty principle, so the time–frequency representation can be improved to obtain accurate results (Zhang et al., 2021).

Compared with the incident rays, the energy of refracted rays has 30-dB energy attenuation; therefore, the time-varying signal that contained information about the parameters of movement is weak. Additionally, with the signal from background noise pollution, it is difficult to achieve the desired results by conventional time–frequency analysis. Therefore, in this paper, we propose a robust time–frequency analysis method to detect Doppler shift.

Furthermore, for the characteristics of the hydroacoustic signal excited by the airborne uniform flight motion target, a single hydrophone is used to solve the three-dimensional parameter estimation problem of the airborne fast-flight target state. Firstly, the discrete line spectrum of the helicopter is used as the characteristic sound source, and its three-dimensional Doppler propagation model in air–water medium is established. Then, based on the asymmetry of the Doppler frequency shift curve and its first- and second-order derivatives, the three-dimensional parameter estimation method of underwater sounding is derived. Finally, by analyzing the measured signals received by a single hydrophone, the paper verifies the estimation performance of the algorithm in the ocean noise scenario.

## 2 The model of a flight path along a straight line

The study only cares about the case that the direct refraction ray arrives at the hydrophone and considers that both aircraft altitude  $h$  and hydrophone depth  $d$  are far longer than the acoustic wavelength. Therefore, the aircraft-radiated noise is assumed to be a point source (Zhang et al., 2020). According to the ray theory, rays with the incident angle  $\theta_i$  less than the critical angle  $\theta_C$  can refract into the seawater, and only one of them can arrive at the hydrophone node.

## 2.1 Point source transmission model

The three-dimensional model of the aircraft flying over the hydrophone node in a straight line is depicted in Figure 1. A stationary hydrophone node is marked as  $H$  and is located at a constant depth  $d$ . An airborne source with a constant frequency  $f_0$  is marked as  $S$  as it travels past the hydrophone node in a straight line  $l_s$  at a constant velocity  $v$ , a constant altitude  $h$ , and slant range  $w(t)$ . Additionally, the source velocity  $v$  is assumed to be subsonic, *i.e.*,  $v < c_a < c_w$ , where  $c_a$  and  $c_w$  are the velocity of sound in air and water. The moving source  $S$  is projected to the horizontal plane where the stationary point  $H$  is located and named as  $S'$ , and the projection corresponding to its trajectory is named as  $l'_s$ . The projection of the straight-line distance  $r(t)$  between  $S$  and  $H$  is  $w(t)$ . Specially, the projection of straight-line distance  $r_{\min}(t)$  is  $w_{\min}$  when the source is at the closest point of approach (CPA).  $S$  is in any position on the straight line  $l_s$ ; only one sound ray with the incident angle  $\theta_i(t)$  refracts at the point  $T$  of the air–seawater interface and then arrives at  $H$  in the direction of the refraction angle  $\theta_r$ . In this case, the velocity  $v$  of  $S$  can be decomposed into two mutually perpendicular components; among them, one component is in the horizontal plane with the straight line  $l_s$ , *i.e.*,

$$v_1(t) = v \cos \alpha(t), \tag{1}$$

where  $\alpha(t)$  is the slant angle between  $l_s$  and  $\overline{SH}$ . Furthermore,  $v_1(t)$  can be decomposed into two mutually perpendicular components as well, and one component is consistent with the direction of the incident sound ray arriving at the hydrophone node, *i.e.*,

$$v_{11}(t) = v_1(t) \cos \beta(t), \tag{2}$$

where  $\beta(t)$  is the line-of-sight angle.  $\beta(t)$  and  $\theta_i(t)$  are complementary angles to each other on the vertical plane as determined by  $S$ ,  $H$ , and  $v_{11}$ .

The Doppler shift  $f_d(t)$  is completely calculated by the four source transmission parameters  $\{f_0, v, \alpha(t), \beta(t)\}$ . Furthermore, the angles  $\alpha(t)$  and  $\beta(t)$  are solved from the geometric relations shown in Figures 2 and 3, respectively.

## 2.2 Doppler shift model

The relative position varies between  $S$  and  $H$  as presented in Figure 2. Note that the source velocity  $v$  and the moment  $t$  is either negative or positive, depending on whether  $H$  is on the left- or right-hand side of the CPA as  $S$  moves along its

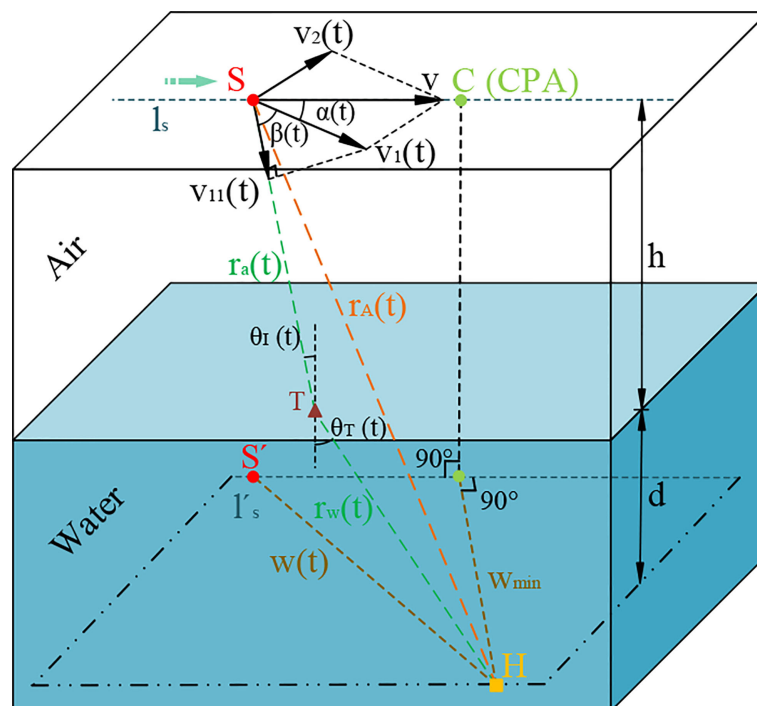


FIGURE 1 General geometrical configuration for a hydrophone node in the water and the linear trajectory of a source in the air as it transits past the hydrophone node in a straight line at constant velocity  $v$ , constant altitude  $h$ , and certain slant range  $w$ .

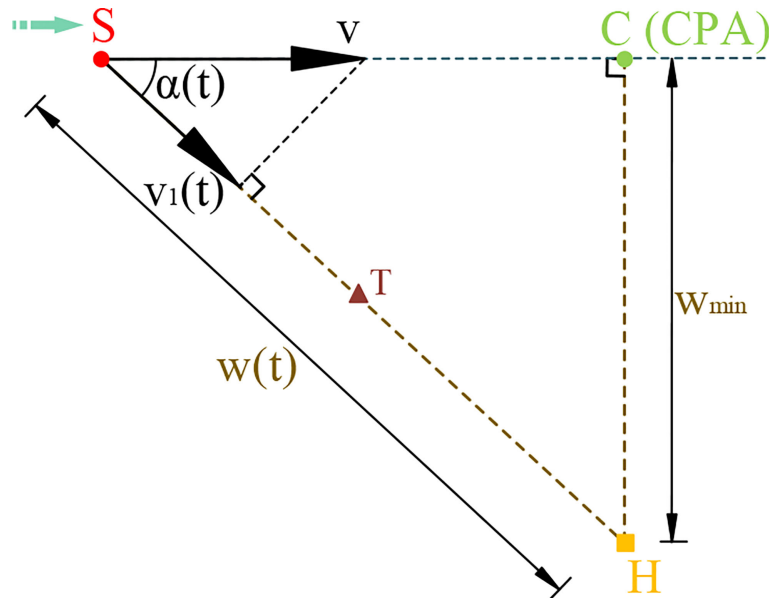


FIGURE 2 Top view of Figure 1. The slant range between the source and the hydrophone at time  $t$  is given by  $w$ . The distance from the hydrophone to the closest point of approach of the source is denoted by  $w_{min}$ .

trajectory. According to the Pythagorean theorem, the slant range  $w(t)$  at time  $t$  can be derived as

$$w(t) = \sqrt{v^2 t^2 + w_{min}^2}, \tag{3}$$

Then, according to the trigonometric function, the slant angle  $\alpha(t)$  is given by

$$\alpha(t) = \arcsin \frac{w_{min}}{\sqrt{v^2 t^2 + w_{min}^2}}. \tag{4}$$

According to the geometric relations in Figure 3, there exists

$$\beta_I(t) = \frac{\pi}{2} - \theta_I(t). \tag{5}$$

The propagation path of a sound ray with a specific incidence angle  $\theta_I(t)$  is presented in Figure 3. The refraction angle  $\theta_T(t)$  can be solved by the cosine theorem with the apex angle  $\angle STH$  in  $\triangle STH$  as follows:

$$\cos \{ \pi - [\theta_I(t) - \theta_T(t)] \} = \frac{r_a^2(t) + r_w^2(t) - r^2(t)}{2r_a(t)r_w(t)}, \tag{6}$$

where the propagation distance  $r_a(t)$  and  $r_w(t)$  is given by

$$r_a(t) = \frac{h}{\sin \theta_I(t)}, \quad r_w(t) = \frac{d}{\sin \theta_T(t)}, \tag{7}$$

and the straight-line distance between S and H is expressed as

$$r^2(t) = (d + h)^2 + v^2 t^2 + w_{min}^2. \tag{8}$$

Substituting Eq. (7) and Eq. (8) into Eq. (6),  $\theta_T(t)$  with respect to  $t$  can be written in the following form:

$$\sin \theta_T(t) \sqrt{1 - n^2 \cos^2 \theta_T(t)} [(d + h)^2 + v^2 t^2 + w_{min}^2 - 2dh] - 2dh n \cos^2 \theta_T(t) - \frac{d^2 \sin \theta_T(t)}{\sqrt{1 - n^2 \cos^2 \theta_T(t)}} - \frac{h^2 \sqrt{1 - n^2 \cos^2 \theta_T(t)}}{\sin \theta_T(t)} = 0 \tag{9}$$

While S moves toward or away from the CPA, the relationship of various parameters with the moment  $t$  is shown in Table 1.

During the propagation of sound, it has traveled over time as

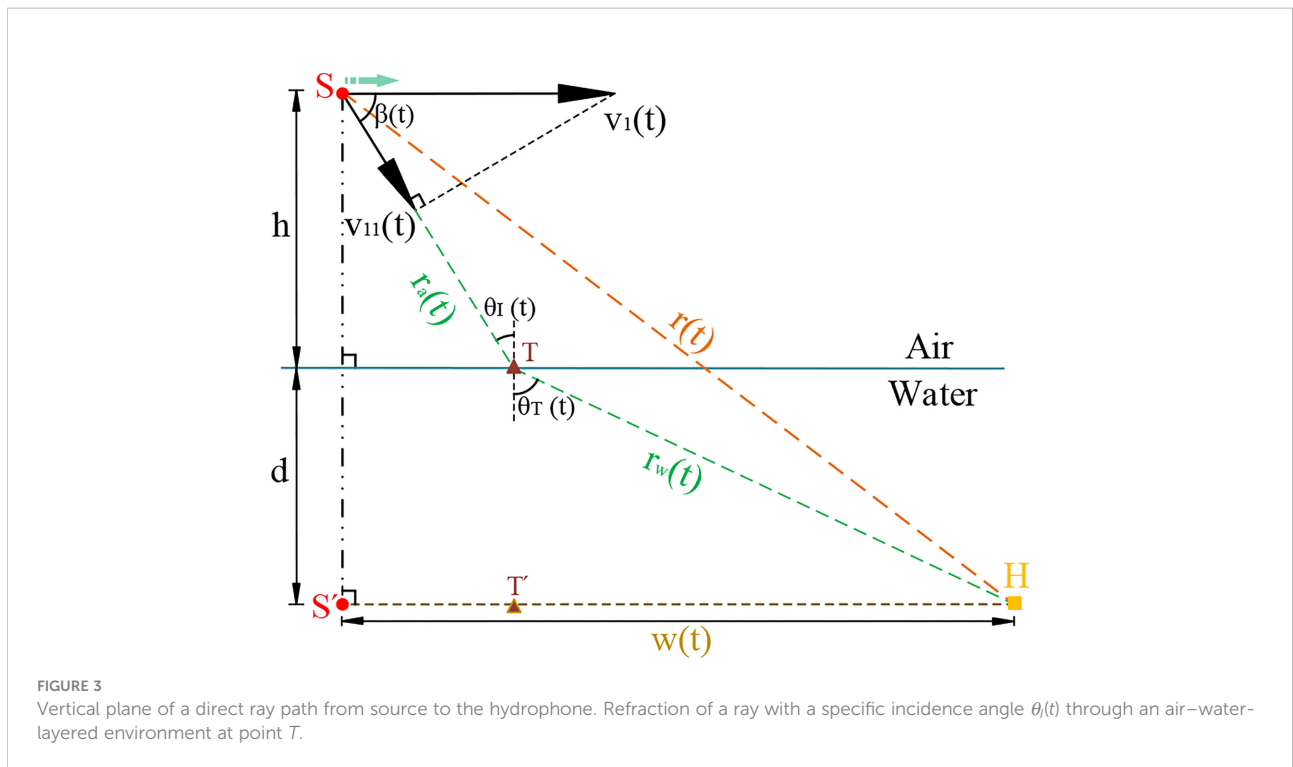
$$\tau(t) = \frac{h}{c_a \sin \theta_I(t)} + \frac{d}{c_w \sin \theta_T(t)}. \tag{10}$$

In particular, the propagation time traveled from the CPA to the hydrophone node is given by

$$\tau_{min}(t) = \frac{h}{c_a \sin \theta_{Ic}} + \frac{d}{c_w \sin \theta_{Tc}}. \tag{11}$$

where

$$w_{min} = \frac{h}{\tan \theta_{Ic}} + \frac{d}{\tan \theta_{Tc}}, \quad \frac{c_a}{c_w} = \frac{\sin \theta_{Ic}}{\sin \theta_{Tc}}. \tag{12}$$



For the moving source with constant frequency  $f_0$ , its Doppler frequency  $f_d(t_r)$  received by the hydrophone node is given by

$$f_d(t_r) = \frac{f_0}{1 - \frac{v}{c_w} \cos \gamma_T(t_r)}, \quad (13)$$

where  $t_r = t + \tau(t)$ ,  $\cos \gamma_T(t_r) = \cos \alpha(t_r) \cos \beta_T(t_r)$ , and  $\gamma_T(t_r)$  is the incidence angle in three-dimensional space at the moment  $t$ . When the parameters  $\{v, h, d, w_{\min}\}$  are known,  $\alpha(t_r)$  and  $\beta(t_r)$  can be obtained from Eq. (4) and Eq. (9). Therefore, for the straight-line flight, the Doppler frequency information  $f_d(t_r)$  is completely specified by the five flight parameters, i.e.,  $\{f_0, v, h, d, w_{\min}\}$ .

In this case, the five flight parameters are changed respectively, as shown in Figures 4A–E. In order to observe the influence of the different parameters on the change trend of the Doppler frequency intuitively, the Doppler frequency  $f_d(t)$  is depicted from the perspective of the source  $S$ . All five pictures demonstrate that there exists  $f_d(t) > f_0$  or  $f_d(t) < f_0$  during the

process of the source  $S$  approaching the CPA or moving away from the CPA.

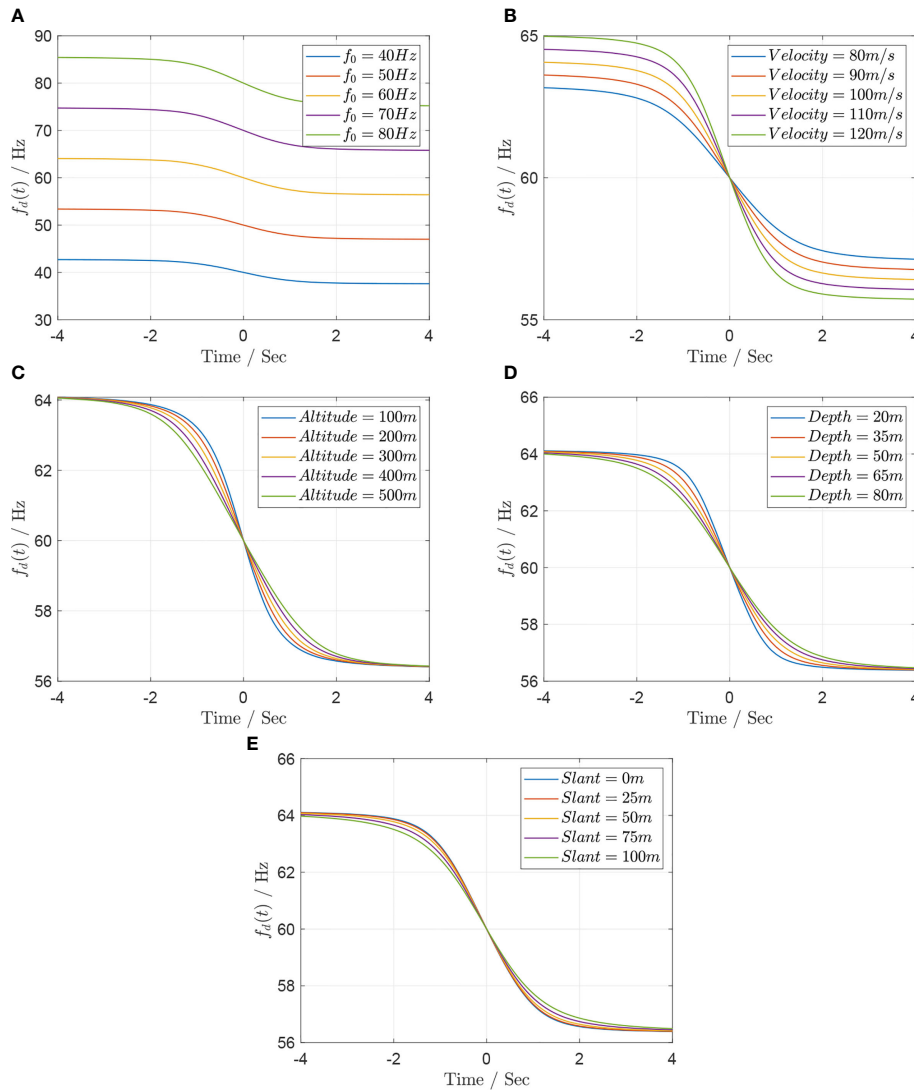
For each value of  $f_0$  ranging from 40 to 80 Hz at an increment of 10 Hz, Figure 4A shows that the Doppler frequency  $f_d(t)$  translates along the positive direction of the  $f_d$  axis with the increase of the source frequency  $f_0$ .

For each value of  $v$  ranging from 80 to 120 m/s at an increment of 10 m/s, Figure 4B shows that the maximum of the frequency shift  $|f_d(t) - f_0|$  is relative with the velocity  $v$ . At the same time  $t$ , the greater the velocity is, the greater the frequency shift and its gradient are.

For each value of  $h$  ranging from 100 to 500 m at an increment of 100 m, this is illustrated in Figure 4C. For each value of  $d$  ranging from 20 to 80 m at an increment of 15 m, this is depicted in Figure 4D. For each value of  $w_{\min}$  ranging from 0 to 100 m at an increment of 25 m, this is presented in Figure 4E. Figures 4C–E show that the altitude  $h$ , the depth  $d$ , and the minimum slant  $w_{\min}$  can affect the gradient of the Doppler frequency  $f_d(t)$ .

TABLE 1 Rule of parameter variation.

Position/parameters	From infinity to the CPA	CPA	From the CPA to infinity
$t$	$-\infty \rightarrow 0$	0	$0 \rightarrow +\infty$
$\alpha(t)$	$0 \rightarrow \pi/2$	$\pi/2$	$\pi/2 \rightarrow 0 \rightarrow \pi/2 - \arctan h/w_{a \min} \rightarrow \arcsin c^d/cw$
$\theta_i(t)$	$\arcsin c_d/cw \rightarrow \pi/2 - \arctan h/w_{a \min}$	$\pi/2 - \arctan h/w_{a \min}$	
$\theta_r(t)$	$\pi/2 \rightarrow \pi/2 - \arctan d/w_{w \min}$	$\pi/2 - \arctan d/w_{a \min}$	$\pi/2 - \arctan d/w_{w \min} \rightarrow \pi/2$



**FIGURE 4**  
**(A)** Doppler shift curve  $f_d(t)$  versus frequency  $f_0 \in [30, 70]$  Hz every 10 Hz, assuming  $v = 100$  m/s,  $h = 300$  m,  $d = 50$  m, and  $w_{min} = 50$  m. **(B)** Doppler shift curve  $f_d(t)$  versus frequency  $v \in [80, 120]$  m/s every 10 m/s, assuming  $f_0 = 60$  m/s,  $h = 300$  m,  $d = 50$  m, and  $w_{min} = 50$  m. **(C)** Doppler shift curve  $f_d(t)$  versus frequency  $h \in [100, 500]$  m every 100 m, assuming  $f_0 = 60$  m/s,  $v = 100$  m/s,  $d = 50$  m, and  $w_{min} = 50$  m. **(D)** Doppler shift curve  $f_d(t)$  versus frequency  $d \in [20, 80]$  m every 15 m, assuming  $f_0 = 60$  m/s,  $v = 100$  m/s,  $h = 300$  m, and  $w_{min} = 50$  m. **(E)** Doppler shift curve  $f_d(t)$  versus frequency  $w_{min} \in [0, 100]$  m every 25 m, assuming  $f_0 = 60$  m/s,  $v = 100$  m/s,  $h = 300$  m, and  $d = 50$  m.

### 2.3 Estimation of flight parameters

Assuming that the sound velocity  $c_a$  and  $c_w$  in air and water, the hydrophone depth  $d$ , and its received Doppler frequency  $f_d(t)$  are known, the section proposes a solution to estimate the sound source parameters, i.e.,  $f_0$ ,  $v$ ,  $h$  and  $w_{min}$ , using a single hydrophone.

Submitting Eq. (4) into Eq. (13), there exists

$$f_d(t) = f_0 \left[ 1 - \frac{v^2 t \cos \beta_T(t)}{c_w \sqrt{v^2 t^2 + w_{min}}} \right]^{-1}. \quad (14)$$

The velocity  $v$  and constant frequency  $f_0$  of source can be obtained, respectively, as follows:

$$v = c_w \frac{f_d(-\infty) - f_d(+\infty)}{f_d(-\infty) + f_d(+\infty)},$$

$$f_0 = 2 \frac{f_d(-\infty)f_d(+\infty)}{f_d(-\infty) + f_d(+\infty)}. \quad (15)$$

The first- and second-order derivative of  $f_d(t)$  is derived in Appendix A. The four flight parameters  $\{h, w_{min}, \beta_T(t)\}$  are

$w_{\min}, \beta_T(t)$  are obtained by solving a set of equations as follows:

$$\frac{df_d(t)}{dt} \Big|_{t=\tau_{\min}} = f'_d(f_0, v, h, w_{\min}) = f'_d(\tau_{\min}), \quad (16)$$

and

$$\frac{d^2f_d(t)}{dt^2} \Big|_{t=\tau_{s1}} = f''_d(f_0, v, h, w_{\min}). \quad (17)$$

Assume that the source with  $f_0=68\text{Hz}$  and  $v=123\text{m/s}$  flies along a straight line at  $h=200\text{m}$  and the shortest projection  $w_{\min}$  between  $H$  and  $S'$  is  $30\text{m}$ . The time–frequency curve of the signal received by the hydrophone with a depth of  $50\text{m}$  is shown in Figure 5.

Figure 5 shows that  $f_d(-\infty)$  and  $f_d(+\infty)$  can be approximated as  $f_d=73.8\text{Hz}$  at  $t=-4\text{s}$  and  $f_d=63.4\text{Hz}$  at  $t=4\text{s}$ , respectively. According to Eq. 15,  $v$  and  $f_0$ , respectively, is given by

$$v = c_w \frac{f_d(-4) - f_d(+4)}{f_d(-4) + f_d(+4)},$$

$$f_0 = 2 \frac{f_d(-4)f_d(+4)}{f_d(-4) + f_d(+4)}. \quad (18)$$

Next, we solve the following equations to obtain  $h, w_{\min}$ , and  $\beta_T(t)$  at the moment  $t$

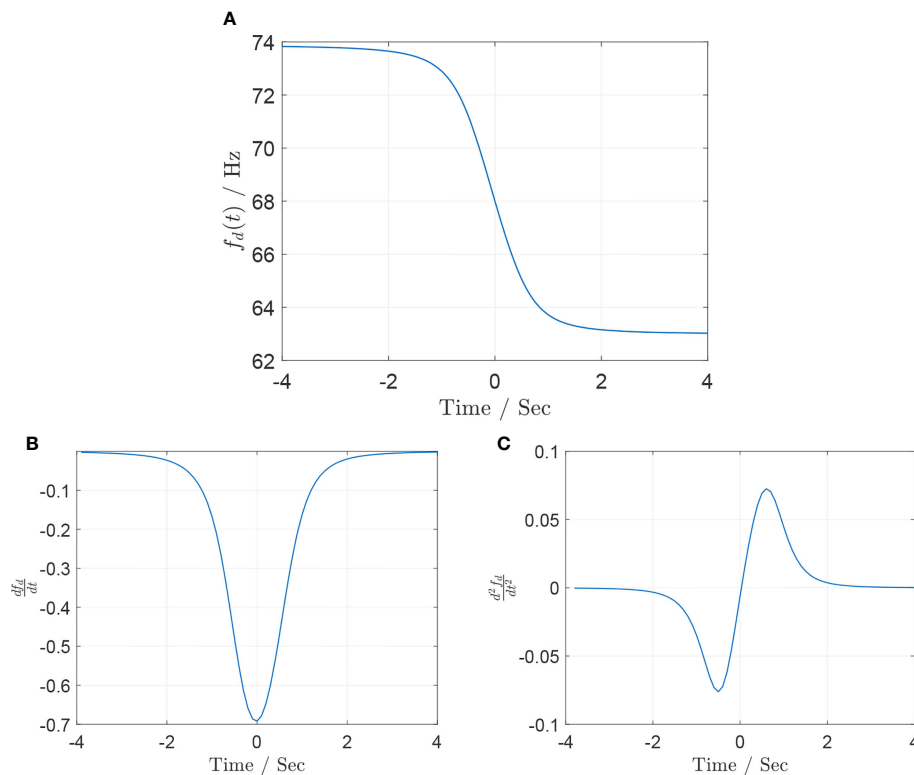
$$f'_d(0) = -6.921,$$

$$f''_d(-0.5) = -7.614,$$

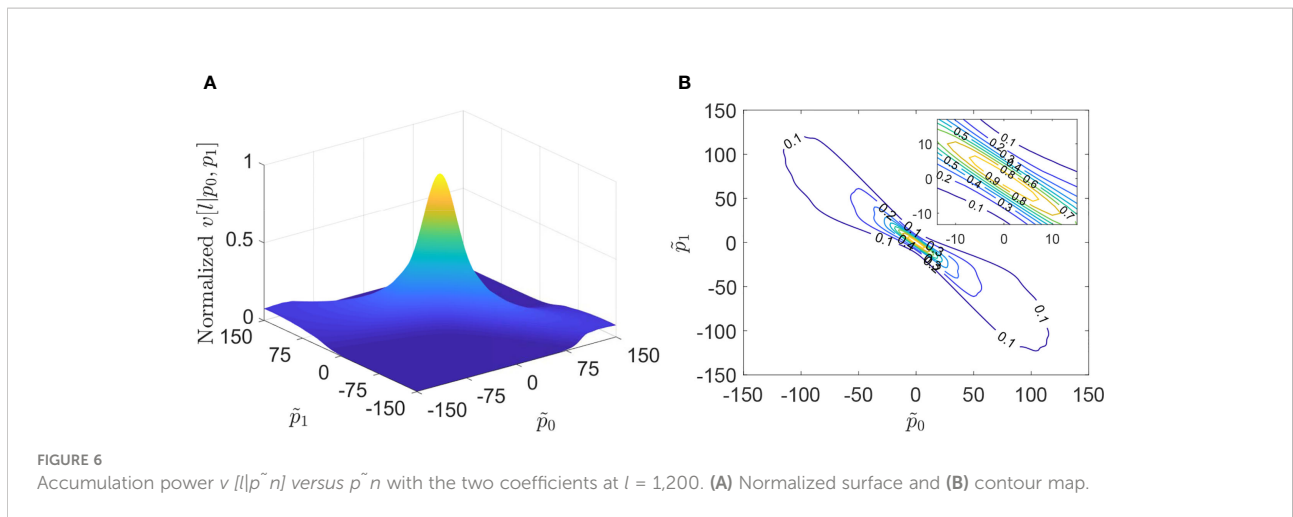
$$f''_d(0.6) = 7.253. \quad (19)$$

### 3 Time–frequency analysis method

Under strong noise conditions, most traditional methods are not able to ensure the robustness and accuracy of time–frequency analysis simultaneously. Therefore, this section



**FIGURE 5** (A) Doppler shift curve  $f_d(t)$ , assuming  $f_0 = 68\text{ m/s}$ ,  $v = 123\text{ m/s}$ ,  $h = 200\text{ m}$ ,  $d = 50\text{ m}$ , and  $w_{\min} = 30$ . (B) Doppler shift curve  $f'_d(t)$ , assuming  $f_0 = 68\text{ m/s}$ ,  $v = 123\text{ m/s}$ ,  $h = 150\text{ m}$ ,  $d = 20\text{ m}$ , and  $w_{\min} = 30$ . (C) Doppler shift curve  $f''_d(t)$ , assuming  $f_0 = 68\text{ m/s}$ ,  $v = 123\text{ m/s}$ ,  $h = 150\text{ m}$ ,  $d = 20\text{ m}$ , and  $w_{\min} = 30$ .



proposes a robust time–frequency analysis algorithm using a memory-dependent derivative.

### 3.1 The proposed algorithm

According to the theory of Ville, an analytic signal  $x \in \mathbb{C}^L$  with the time-varying IF can be given by

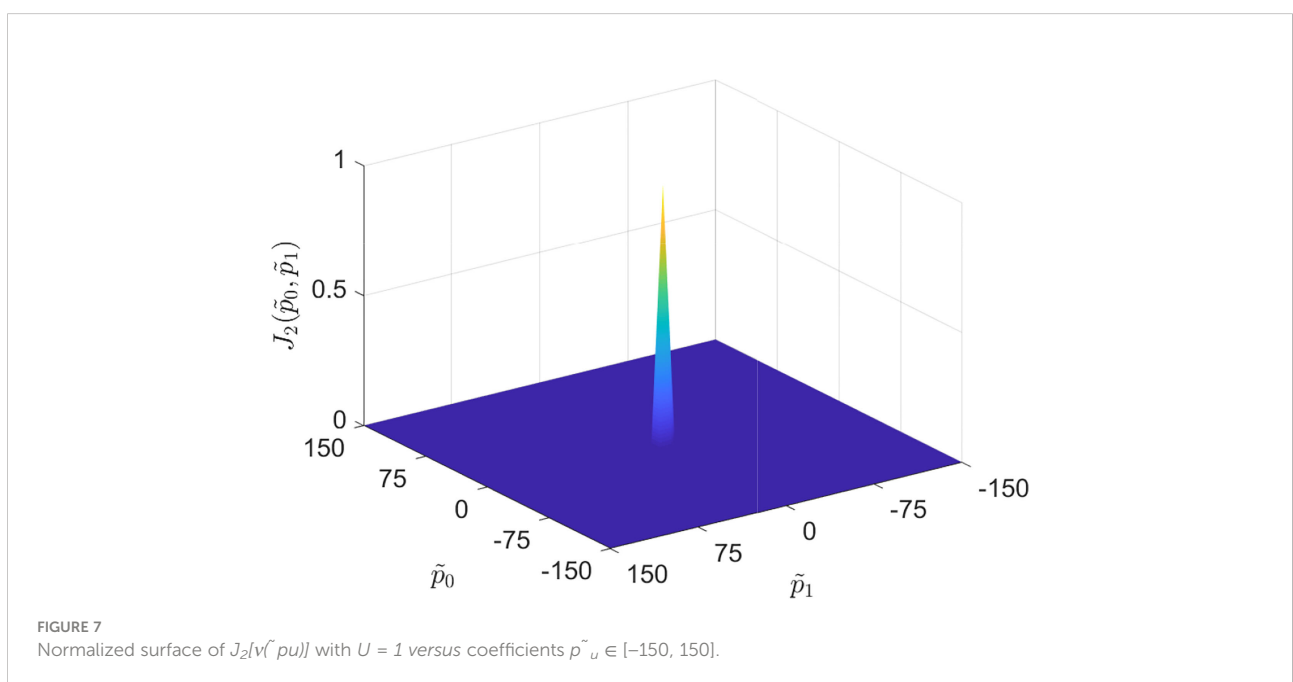
$$x[l] = ae^{i\sum_{m=0}^l \varphi[m]}, \tag{20}$$

where  $a \in \mathbb{R}_+$  and  $\varphi[m]$  is the amplitude and IF, respectively, and  $i = \sqrt{-1}$ . It should be noted that the FM signal may have a wide spectrum band since  $\varphi[m]$  can vary in a wide range.

According to the Weierstrass approximation theorem, any continuous function in a closed and bounded interval can be uniformly approximated by a polynomial to any degree of accuracy. so  $\varphi[m]$  can be expanded to  $U$ -order ( $U \leq L-1$ ) Taylor series at the moment  $m=0$ , i.e.,

$$\varphi[m] = \sum_{u=0}^U \frac{\varphi^{(u)}[0]}{u!} \left(\frac{m}{f_s}\right)^u + R_u \left[\left(\frac{m}{f_s}\right)^{u+1}\right], \tag{21}$$

where  $f_s$  is sample frequency,  $\varphi^{(u)}[0]$  is the  $u$ -th order difference coefficient about  $\varphi[0]$ , and  $R_u(m)$  is the Maclaurin remainder. There exists  $R_u(m)=0$  when  $U=L-1$ . Additionally, the maximum order  $U$  is related to the approximation accuracy to FM function. Specifically,  $U=0$  for the continuous wave and  $U=1$  for the linear FM signal.





Now, by inserting Eq. (21), omitting Maclaurin remainder  $R_u(m)$ , into Eq. (20), the reformulated formula is listed as follows:

$$x[l] = ae^{i \sum_{m=0}^l \sum_{u=0}^U p_u \left(\frac{m}{f_s}\right)^u}, \tag{22}$$

where  $p_u = \frac{\varphi^{(u)}[0]}{u!}$  is the coefficient of the  $u$ -th power term. In Eq. (22),  $e^{i p_u \frac{m}{f_s}}$  and  $e^{\sum_{m=0}^l \sum_{u=0}^U p_u \left(\frac{m}{f_s}\right)^u}$  is named as harmonic component and modulated element, respectively. Therefore, given a set of estimated parameters, a time–frequency estimation result is capable of producing a highly accurate approximation to a wide range of signals whose IF trajectories can be any continuous function of time. The goal of  $\varphi[m]$  estimation in Eq. (20) is transformed to obtain  $p_u$  in Eq. (22) since IF can be recovered from its Taylor series.

In the signal estimation field, noise has always been considered as a serious cause of error. Although the most common noise can be modeled by well-known Gaussian distributions, some observed signals experience many types of i.i.d. noise with unknown distribution, for instance, underwater acoustic noise, low-frequency atmospheric noise, and other types of man-made noise (Liu et al., 2018). The signal  $x$  is corrupted by the i.i.d. noise with unknown distribution  $z \in \mathbb{C}^L$ ; then, there is an observed signal  $y$  (Kozick and Sadler, 2000), i.e.,

$$y[l] = x[l] + z[l]. \tag{23}$$

Currently, there exists a class of i.i.d. noise, called symmetric  $\alpha$ -stable (S $\alpha$ S) noise, that can be used to model various types of real noise. The different noise statistical characteristics can be represented by changing parameters, where the characteristic exponent  $\alpha$  controls the heaviness of the tails of the density function and the dispersion parameter  $\gamma$  determines the spread of the density around its location parameter. Thus, the probability density function of S $\alpha$ S gives a much better approximation to the i.i.d. noise with various distribution types.

These conventional methods can be considered as the maximum likelihood estimates resulting from corresponding minimization problems for Gaussian the noise environment. Thus, they are very sensitive to the i.i.d. noise with unknown distribution, and their robustness against a model mismatch is not ensured. In such cases, these methods under the Gaussian assumption, e.g., parametric time–frequency analysis, may exhibit a serious performance degradation for the i.i.d. noise with arbitrary distribution. The choice of a more robust solution that can be used for various i.i.d. noises is thus salutary.

A fundamental issue discussed in this paper is the question of how to improve the resolution and robustness of the time–frequency estimation result in the i.i.d. noise with arbitrary distribution. The need for robustness and high resolution in signal analysis is self-evident, yet, at the present time, the traditional assessment of such properties remains largely *ad*

*hoc*. The goal has been to develop a theory and techniques for a robust and high-resolution time–frequency estimation.

### 3.2 Power of phase difference accumulation

The phase difference is a common tool used in polynomial phase estimation, and it is performed by multiplying each moment with a kernel function element-wise (Vucijak and Saranovac, 2010). If FM is approximated by the  $U$ -th order Taylor series in Eq. (22) and the selected kernel function is also the  $U$ -th order Taylor series with the estimated  $\hat{p}_u$ ,

$$h[l|\hat{p}_u] = e^{-i \sum_{m=0}^l \sum_{u=0}^U \hat{p}_u \left(\frac{m}{f_s}\right)^u}, \tag{24}$$

so the phase difference between  $x$  and  $h$  is given by

$$r[l|\tilde{p}_u] = x[l|\tilde{p}_u]h[l|\hat{p}_u] = ae^{i \sum_{m=0}^l \sum_{u=0}^U \tilde{p}_u \left(\frac{m}{f_s}\right)^u}, \tag{25}$$

where  $\tilde{p}_u = p_u - \hat{p}_u$  is the coefficient error, and  $\varphi[m|\tilde{p}_u] = \sum_{u=0}^U \tilde{p}_u \left(\frac{m}{f_s}\right)^u$  is defined as the IF error. As  $p_u$  is a fixed unknown value,  $r$  has also been parameterized with  $\tilde{p}_u$  to simplify the entire expression.

Coherence accumulation to phase difference can improve the SNR, which has obtained successful applications (Vucijak and Saranovac, 2010; Wang et al., 2017). For i.i.d. noise with arbitrary distribution, the averaging value of coherence accumulation will fluctuate slightly; however, the coherence accumulation of phase difference in Eq. (25) will keep increasing as time increases, especially when  $\tilde{p}_u$  approximates to zero. Inspired by this fundamental rule, a new cost function is constructed to find the optimal solution for all coefficients  $\tilde{p}_u$ .

We define the accumulation power of phase difference accumulation (APP) by removing the average value as

$$v[l|\tilde{p}_u] = \frac{1}{l} \sum_{m=0}^l \{ c[m|\tilde{p}_u] - c_M[m|\tilde{p}_u] \} \{ c[m|\tilde{p}_u] - c_M[m|\tilde{p}_u] \}^*, \tag{26}$$

where  $*$  represents the complex conjugation, and

$$\begin{aligned} c[m|\tilde{p}_u] &= \sum_{n=0}^m r[n|\tilde{p}_u], \\ c_M[m|\tilde{p}_u] &= \frac{1}{m} \sum_{n=0}^m c[n|\tilde{p}_u], \end{aligned} \tag{27}$$

are the phase accumulation of  $r[m|\tilde{p}_u]$  and the average of  $c[m|\tilde{p}_u]$  over the first  $m \leq l$  sample numbers, respectively. Obviously,  $v[l|\tilde{p}_u]$  is the accumulation power of phase accumulation  $r[m|\tilde{p}_u]$  over the interval  $[0, l]$ .

**THEOREM 1.** The power  $v[l|\tilde{p}_u]$  versus  $|\tilde{p}_u|$  is the monotone function.

**PROOF.** We have provided the proof in the [Appendix](#) section. In Eq. (26),  $\frac{dv[l|\tilde{p}_u]}{d\tilde{p}_u} \geq 0$  holds for  $\tilde{p}_u \geq 0$  and  $\frac{dv[l|\tilde{p}_u]}{d\tilde{p}_u} \leq 0$  holds for  $\tilde{p}_u \leq 0$ , so the maximal value of  $v[l|\tilde{p}_u]$  will be obtained at  $\tilde{p}_u = 0$ .

An example will be exhibited with the sampling frequency  $f_s=1,200$  H and the sample number  $L=1,200$ , the order number  $U=1$ , and the coefficient  $\tilde{p}_u \in [-150, 150]$  for the power  $v[l|\tilde{p}_u]$ . Without loss of generality, the mesh plot of  $v[l|\tilde{p}_u]$  versus coefficient  $\tilde{p}_u$  at  $l=1,200$  is demonstrated in [Figure 6A](#), and the contour curve of  $v[l|\tilde{p}_u]$  is shown in [Figure 6B](#). It shows that  $v[l|\tilde{p}_u]$  increases monotonically as coefficient  $\tilde{p}_u$  tends to zero. Specifically, the maximum value of  $v[l|\tilde{p}_u]$  lies at  $ldep_u = 0$ .

Theorems 1 and [Figure 6](#) demonstrate that the power  $v[l|\tilde{p}_u]$  at certain moment  $l$  is a monotonic function of  $|\tilde{p}_u|$  and can obtain the maximum at  $|\tilde{p}_u| = 0$  if the FM signal and  $h$  are coherent. Since  $v[l|\tilde{p}_u]$  is a positive time series as  $l$ , different types of norm, like  $\ell_1$ -norm,  $\ell_2$ -norm, and  $\ell_q$ -norm, can be used to measure the size of  $v[l|\tilde{p}_u]$ . From a statistical perspective,  $\ell_2$ -norm is suitable for Gaussian noise and  $\ell_1$ -norm is robust to outliers. Since  $v[l|\tilde{p}_u]$  could be corrupted by  $S\alpha S$  noise with impulses and the estimated IF can be determined by maximizing  $v[l|\tilde{p}_u]$  at any moment,  $\ell_1$ -norm will be chosen. Therefore, the problem of coefficient  $\tilde{p}_u$  estimation can be readily solved by  $\ell_1$ -norm maximization as follows:

$$\max_{\tilde{p}_u} \|v[l|\tilde{p}_u]\|_1 \quad s.t. \quad \tilde{p}_u = p_u - \hat{p}_u. \quad (28)$$

Moreover, Theorem 1 guarantees that the accurate coefficient is available in the continuous domain, and the previous work in time-frequency estimation is fundamentally limited by the Heisenberg-Gabor inequality for a better resolution. The estimated IF can be approximated as  $\hat{\phi}[m|\hat{p}_u] = \sum_{u=0}^U \hat{p}_u (\frac{m}{f_s})^u$ .

### 3.3 IF estimation based on multiple log-sum

Under the condition of short window size  $L$  or high noise power, the estimated IF trajectory with  $\ell_1$ -norm shows poor accuracy and is easily interrupted by the accumulation of noises, although the maximal value in (28) is still obtained at the true  $p_u$  from the statistical perspective. In view of the nonlinear amplification for DOA estimation in [Wang et al. \(2008\)](#) and [Wang and Kay \(2010\)](#), the multiple logarithmic sum (MLS) is proposed to achieve the sharper peak for target IF and deeper null for non-target area, which also improves the accurate and robust performance for IF estimation.

Log-sum function has often been used to solve for maximization and is shown to be more convex and smoother than the  $\ell_1$ -norm ([Zhang et al., 2019](#); [Qiao et al., 2020](#)), and this function can exhibit nonlinear amplification in some fixed

intervals. Alternating convexity-promoting functions such as log-sum function can be used to replace the  $\ell_1$ -norm, and a surrogate function with a sharper peak for target frequency and deeper null for side lobes is achieved. Thus, we consider the use of the log-sum function to improve the convexity of the objective function. Replacing the  $\ell_1$ -norm in Eq. (28) with the multiple log-sum function over the entire sample number  $L$  will result in

$$\max_{\tilde{p}_u} J_1(\tilde{p}_u) = \sum_{l=0}^{L-1} \log \{v[l|\tilde{p}_u] + 1\} \quad s.t. \quad \tilde{p}_u = p_u - \hat{p}_u. \quad (29)$$

where “1” is added to ensure that  $\log \{v[l|\tilde{p}_u] + 1\}$  is non-negative, and it does not change the location of the optimal solution.  $J_1(\tilde{p}_u)$  versus  $|\tilde{p}_u|$  is also monotonic since  $\log$  is a monotonic function and  $v[l|\tilde{p}_u]$  versus  $|\tilde{p}_u|$  is a monotonic decreasing function as proved in Theorem 1.  $J_1(\tilde{p}_u)$  in Eq. (29) is rewritten as follows:

$$J_1(\tilde{p}_u) = \log \prod_{l=0}^{L-1} \{v[l|\tilde{p}_u] + 1\} = \log v(\tilde{p}_u), \quad (30)$$

so  $J_1(\tilde{p}_u)$  can also be considered as the function of  $v(\tilde{p}_u)$  and rewritten as  $J_1[v(\tilde{p}_u)]$ .

Hence, for  $J_1[v(\tilde{p}_u)]$ , the accumulation of the frequency domain is transformed into the integral of power in the logarithm space. Benefiting from the monotonicity of  $v(\tilde{p}_u)$  versus  $v(t|\tilde{p}_u)$  as proved in Theorem 2, the sum of  $J_1[v(\tilde{p}_u)]$  over  $v(\tilde{p}_u)$  can significantly amplify the power  $v(t|\tilde{p}_u)$  in a non-linear way. Without loss of generality, only the accumulation over the non-positive interval of  $\tilde{p}_u$  is considered since  $v[l|\tilde{p}_u]$  is an even function. According to Theorems 1 and 6, the power  $v[l|\tilde{p}_u]$  tends to zero when  $\tilde{p}_u \rightarrow 0$ . Thus, the lower integral limit can be obtained as follows:

$$\begin{aligned} \lim_{\tilde{p}_u \rightarrow -\infty} \log v(\tilde{p}_u) \\ = \lim_{\tilde{p}_u \rightarrow -\infty} \log \prod_{l=0}^{L-1} \{v[l|\tilde{p}_u] + 1\} = C_0. \end{aligned} \quad (31)$$

Consequently, the variable upper limit integral of  $J_1[v(\tilde{p}_u)]$  over the continuous domain  $v(\tilde{p}_u)$  is given by

$$\begin{aligned} J_2[v(\tilde{p}_u)] &= \int_{C_0}^{v(\tilde{p}_u)} (\log \kappa + 1) d\kappa \\ &= v(\tilde{p}_u) \log v(\tilde{p}_u) + C_1, \end{aligned} \quad (32)$$

where a fixed constant “1” added in the integral is added to obtain a simple analytical solution. Because the linear accumulation of  $J_2[v(\tilde{p}_u)]$  with the constant term does not change the extreme value point, it does not affect the final solution coefficient  $\tilde{p}_u$ . Besides that, Eq. (32) with respect to  $v(\tilde{p}_u)$  is convex according to Chapter 3 in [Boyd et al. \(2004\)](#).

**THEOREM 2.**  $J_2[v(\tilde{p}_u)]$  versus  $|\tilde{p}_u|$  is monotone decreasing.

**PROOF.** According to the chain rule of derivation, the first-order derivative of  $J_2[v(\tilde{p}_u)]$  about  $|\tilde{p}_u|$  is given by

$$\frac{dJ_2[v(\tilde{p}_u)]}{d|\tilde{p}_u|} = \frac{dJ_2}{dv} \sum_{l=0}^{L-1} \left( \frac{dv}{dv[l]} \frac{dv[l|\tilde{p}_u]}{d|\tilde{p}_u|} \right).$$

Firstly, it follows from Theorem 1 that  $v[l|\tilde{p}_u]$  is a monotonically decreasing function of  $|\tilde{p}_u|$ , and there exists  $\frac{dv[l|\tilde{p}_u]}{d|\tilde{p}_u|} \leq 0$ .

Secondly, the first-order derivative of the composite function  $v\{v[l]\}$  about  $v[l]$  from Eq. (30) is given by

$$\begin{aligned} \frac{dv\{v[l]\}}{dv[l]} &= \left\{ \prod_{l=0}^{L-1} \{v[l|\tilde{p}_u] + 1\} \right\}_{v[l]} \\ &= \sum_{l=0}^{L-1} \prod_k \{v[k|\tilde{p}_u] + 1\}, \end{aligned}$$

where  $k \in \{0, L-1\} \setminus \{l\}$ .  $\frac{dv\{v[l]\}}{dv[l]}$  is non-negative since  $v[l|\tilde{p}_u]$  is non-negative. Thus, it indicates that  $v$  *bracev* $[l]$  versus  $v[l]$  is monotonically increasing.

Thirdly, the first-order derivative of  $J_2(v)$  is calculated as follows:

$$\frac{dJ_2[v]}{dv} = \frac{1}{v(\tilde{p}_u)} + 1,$$

which is positive since  $v(\tilde{p}_u) \geq 0$ . Thus,  $J_2(v)$  is monotonically increasing.

To summarize the above-mentioned points, in the case of  $\frac{d\phi}{dt} \leq$ , the first-order derivative of  $J_2[v(\tilde{p}_u)]$  about  $|\tilde{p}_u|$  is non-positive, *i.e.*,

$$\frac{dJ_2[v(\tilde{p}_u)]}{d|\tilde{p}_u|} \leq 0.$$

Therefore, it is proven.

After the nonlinear amplification with convexity-preserving operation and the accumulation in the frequency domain,  $J_2[v(\tilde{p}_u)]$  versus  $\tilde{p}_u$  has a sharper peak and a deeper null level than  $J_1[v(\tilde{p}_u)]$  in (29).  $J_2[v(\tilde{p}_u)]$  with two coefficients is depicted in Figure 7, and the same simulation condition as in Figure 6 is kept. The deep null level of  $J_2[v(\tilde{p}_u)]$  at the unmatched phase is beneficial to find the optimal coefficients and suppress the i.i.d. noise. Since  $J_2[v(\tilde{p}_u)]$  is convex, any gradient algorithm can be used to solve this maximization problem.

In summary, the time-frequency estimation by the APP-MLS algorithm can be converted into the maximization problem as follows:

$$\begin{aligned} \max_{\tilde{p}_u} J_3(\tilde{p}_u) &= v(\tilde{p}_u) \log v(\tilde{p}_u) \text{ s.t. } \tilde{p}_u \\ &= p_u - \hat{p}_u. \end{aligned} \tag{33}$$

We propose the APP-MLS algorithm to find the optimal solution for (37), and the basic idea of the IF estimation by the APP-MLS algorithm is to determine the coefficients  $\hat{p}_u$  of polynomial kernel to estimate the time-frequency trajectory. The details will be introduced in the next section. Due to the convexity

of  $J_3(\tilde{p}_u)$ , any gradient algorithm can be used to solve this maximization problem. In this paper, the optimization solution  $\hat{p}_u$  is obtained by maximizing Eq. (33) by the AdaBelief method, which is more efficient for the parameter tuning process. The intuition for AdaBelief is to adapt the step size according to the “belief” in the current gradient direction. This method uses the exponential moving average of the noisy gradient as the prediction of the gradient at the next time step; if the observed gradient greatly deviates from the prediction, we distrust the current observation and take a small step; if the observed gradient is close to the prediction, we trust it and take a large step. After localizing the  $\hat{p}_u$  in one time interval, the estimated time-frequency trajectory  $\hat{\phi}[m|\hat{p}_u] = \sum_{u=0}^U \hat{p}_u \left(\frac{m}{f_s}\right)^u$  can be obtained.

In practical application, the whole observed data will be partitioned into many overlapped slides with the window size  $\tau$  and the overlap size  $S$  between two close windows; then, the number  $K$  of slides can be obtained. The polynomial order  $U$  and the initial coefficient  $\hat{p}_u$  are given according to the rough character of the signal. In every slide, therein included are the following steps: firstly, the phase difference between the truncated signal and the kernel function is calculated, and its accumulated power can be obtained. Furthermore, by maximizing the multiple log-sum function with the AdaBelief algorithm, the estimated coefficient  $\hat{p}_u$  can be obtained after several iterations. Finally, the estimated IF trajectory can be obtained within the overlap size  $S$ . The detailed procedure is illustrated in Figure 8.

During the application of the APP-MLS algorithm, the estimation accuracy of the IF trajectory depends on the order  $U$ , the window size  $W$ , and the overlap size  $S$  ( $0 \leq S < W$ ). Specifically, as the order  $U$  increases, the Taylor series can approximate the estimated frequency function  $\phi[m]$  with a high accuracy within a fixed window size  $W$ . The longer window size  $W$  leads to larger deviations from the true IF trajectory at a fixed order  $U$ , especially at the window end. The window size  $W$  also affects the robustness of the algorithm, *i.e.*, the longer the window size, the better robustness to suppress noise; it should not be too short in practical application. Additionally, the higher-order  $U$  and longer window size  $W$  will increase the computation complexity, and their choice should be considered carefully.

## 4 Example

The 120-s aircraft noise data are used to validate the efficiency of the estimation model of flight parameters. These data were measured in an experiment where an aircraft traveled along a straight path at a constant speed. The speed of sound in air and seawater was 340 and 1,520 m/s, respectively. A hydrophone is placed at a depth of 90 m in the ocean, and its output is sampled at 44.1 kHz for 2 min. To reduce the computational burden, the data herein are downsampled to 1,200 Hz.

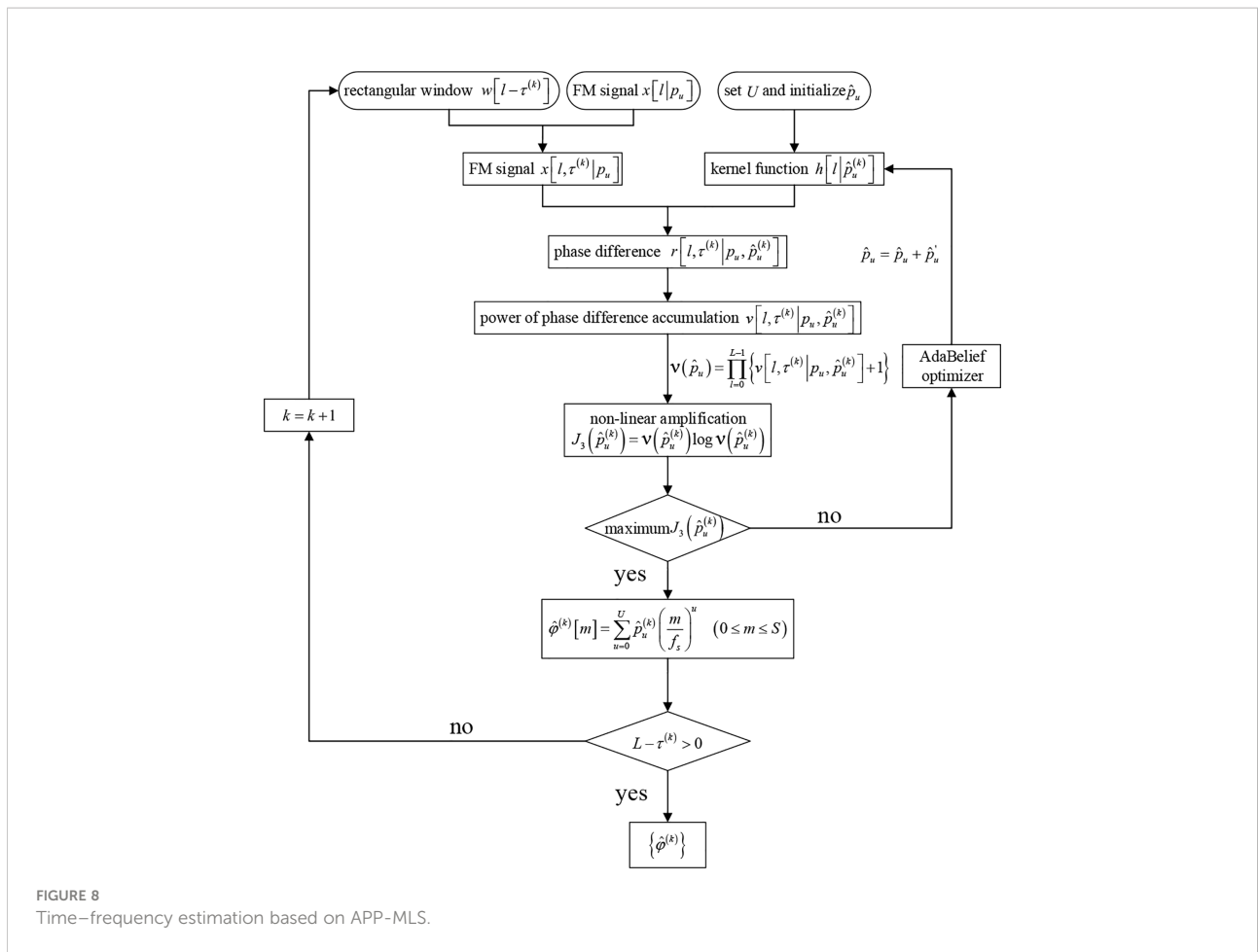


FIGURE 8 Time-frequency estimation based on APP-MLS.

The Acoustical Society of America provided a piece of audio data<sup>1</sup>, where a hydrophone is placed at a depth of 90m in the ocean and records the underwater acoustic data from a turboprop aircraft with a constant frequency of 68 Hz flying straightly over the hydrophone at a fixed altitude. The sampling rate is 44.1 kHz, and the recorded time is 120 s, as shown in Figure 9. As the aircraft approximates and leaves the overhead of the hydrophone, the relative Doppler of the received acoustic signal has an apparent and quick change, and the IF trajectory can determine the flying parameters of the aircraft. The detailed analysis and principle can be referred to Urlick R J (Urlick et al, 1972) and the accurate estimation of the IF trajectory is very important. According to Tre DiPassio’s 2019 Challenge Problem Solution, the original propeller frequency is estimated when the plane is directly overhead the hydrophone at a timestamp denoted as 54 s. To better show the estimated IF results for radiated noise, we intercepted the data from 49 to 59 s when the aircraft is nearly over the hydrophone.

1 <https://acousticstoday.org/iscpasp2019>

The maximum frequency bin of each frame is used to estimate the fundamental frequency present in that frame. Figure 10 shows these estimates as well as the amplitude computed in each frame. The majority of the audio solely contain underwater noise. However, at around 54 s, a few seconds of a plane flying overhead can be heard when listening to the file. The presence of the plane is shown in the highlighted section of Figure 10. There is a spike in the amplitude above the noise floor, paired with a more defined instantaneous frequency track. Figure 11 shows only the frames where the plane flying overhead is audible commensurate to the amplitude spike in Figure 10.

In order to extract instantaneous frequencies from the provided file, a short-time Fourier transform (STFT) is utilized in Figure 11. The audio is split into frames with a length of 8,192 samples per frame, windowed with an 8,192-point Hamming window. Each frame is converted to the frequency domain using the fast Fourier transform (FFT) algorithm run on the frame data. In total, 75% overlap is used between the frames, and zero padding is added such that the frequency resolution (bin size) of the FFT in each frame is approximately 0.1 Hz. Consistent with the Nyquist Theorem, the maximal frequency present in the

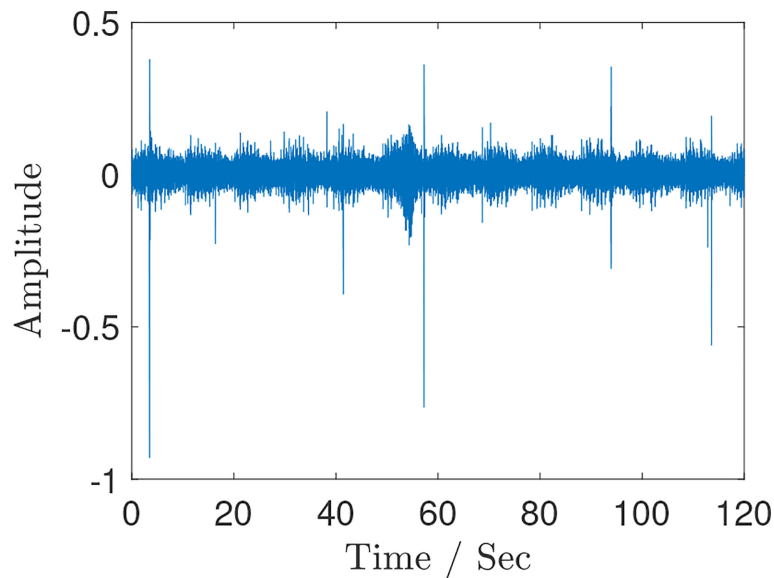


FIGURE 9 Helicopter waveform in each frame of the given audio file.

given audio is half the sampling frequency or 22,050 Hz. Therefore, the full-scale 22,050 Hz will be represented in the number of frequency bins equal to half of the FFT length due to the redundancy added when computing the FFT. The following equation thus shows how the desired resolution is achieved:

$$\text{Resolution} = \frac{f_s}{\text{FFTLLEN}} \quad (34)$$

Solving Eq. (38) for the desired resolution of around 0.1 Hz in each frame shows that 53-times zero padding should be used.

The estimates from each frame in the region with audible plane motion are smoothed to create an instantaneous frequency

track. This smoothing is done with a Savitzky–Golay filter. This type of filter is used to smooth sets of data points in the time domain using a least-squares fitting approach.

The estimated IF trajectory by the APP-MLS will be compared with STFT, PCT, and SST. Because of the high data sampling rate and the frequency of interest being located in the low-frequency range, the data were downsampled by 200 times. In order to estimate the IF better and compare fairly, some parameters of different methods will be set. The window size is set to  $W=128$ , and the overlap rate is set as 75%. Moreover, the iteration number of PCT is seven, the synchronous squeeze factor of SST is 200, and the maximum order of APP-MLS is  $U=2$ . Besides this, in Tre

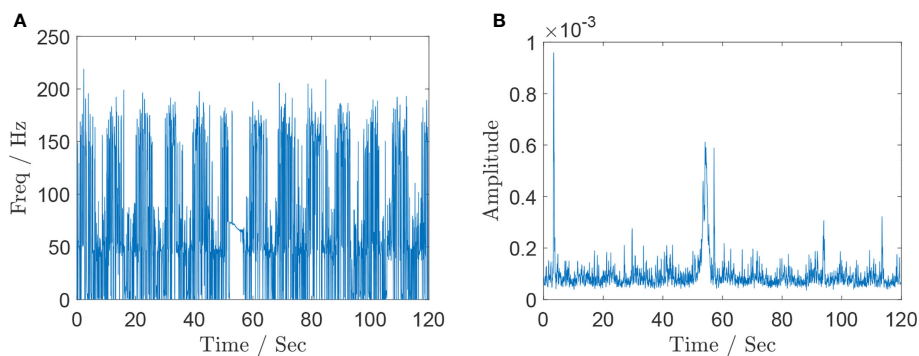
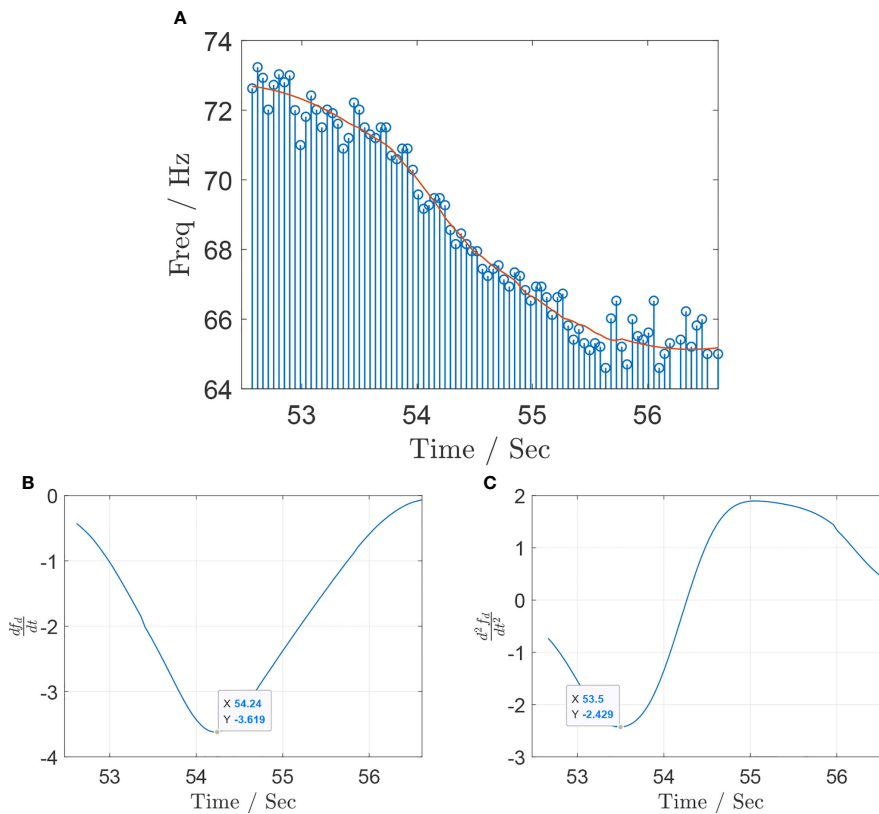


FIGURE 10 Instantaneous frequency estimation results for the aircraft acoustic data recorded by hydrophone with (A) the estimated fundamental frequency in each frame of the given audio file and (B) the estimated amplitude in each frame of the given audio file.



**FIGURE 11** Instantaneous frequency estimation results for the aircraft acoustic data recorded by hydrophone with (A) raw and smoothed frequency track at the hydrophone position. (B) First-order derivative of the Doppler shift curve  $f_d(t)$ . (C) Second-order derivative of the Doppler shift curve  $f_d(t)$ .

DiPassio’s 2019 Challenge Problem Solution<sup>2</sup>, the author gives details about how to extract data from provided audio, which is given as follows: firstly, the audio is split into frames with a length of 128 samples per frame, windowed with a 128-point Hamming window. Secondly, each frame is converted to the frequency domain using the fast Fourier transform algorithm run on the frame data. Then, 75% overlap is used between the frames, and 16 times more zero padding is added such that the frequency resolution of these methods in each frame is approximately 0.1 Hz. Thirdly, the maximum frequency bin of each frame is used to estimate the fundamental frequency present in that frame. Finally, this smoothing is done with a Savitzky–Golay filter to smooth the IF curve.

Figure 12 presents the estimated results of the four methods. Although the changing trend of IF estimation by STFT can be observed as shown in Figure 12A, there exist severe fluctuations outside the interval 53–57 s when the aircraft deviates far from the head of hydrophone. Figures 12B, C show that the interference within the frequency range of interest by PCT and

SST has been improved because the adaptive kernel function and the synchrosqueezing transform are used respectively. However, they are also distorted by the background noise with unknown probability distribution, and their estimated IF information is still not available effectively when the aircraft is far from the hydrophone. Benefitting from the gridless principle and nonlinear amplification, the IF trajectory yielded by APP-MLS can effectively suppress underwater acoustic noise even for low SNR. The IF result within the range of interest shown in Figure 12D is more accurate than the other three methods. The APP-MLS algorithm shows the robust IF estimation performance in the lower SNR environment, and its result is much closer to the true IF than those of the other methods.

The experimental frequency values from Figure 11 range from 64.17 to 73.45 Hz, for a total range of 9.28 Hz. Substituting these values into Eq. (15) will yield as follows:

$$\begin{cases} f_0 = 2 \frac{f_d(-\infty)f_d(+\infty)}{f_d(-\infty)+f_d(+\infty)} \approx 2 \times \frac{73.45 \times 64.17}{73.45+64.17} = 68.5 \text{ Hz} \\ v = c_w \frac{f_d(-\infty)-f_d(+\infty)}{f_d(-\infty)+f_d(+\infty)} \approx 1520 \times \frac{73.45-64.17}{73.45+64.17} = 101 \text{ m/s} \end{cases} \quad (35)$$

so the aircraft frequency is 68.5 Hz, and the speed of the aircraft is 101 m/s.

<sup>2</sup> <https://www.tredipassio.com/asa2019?rq=asa>

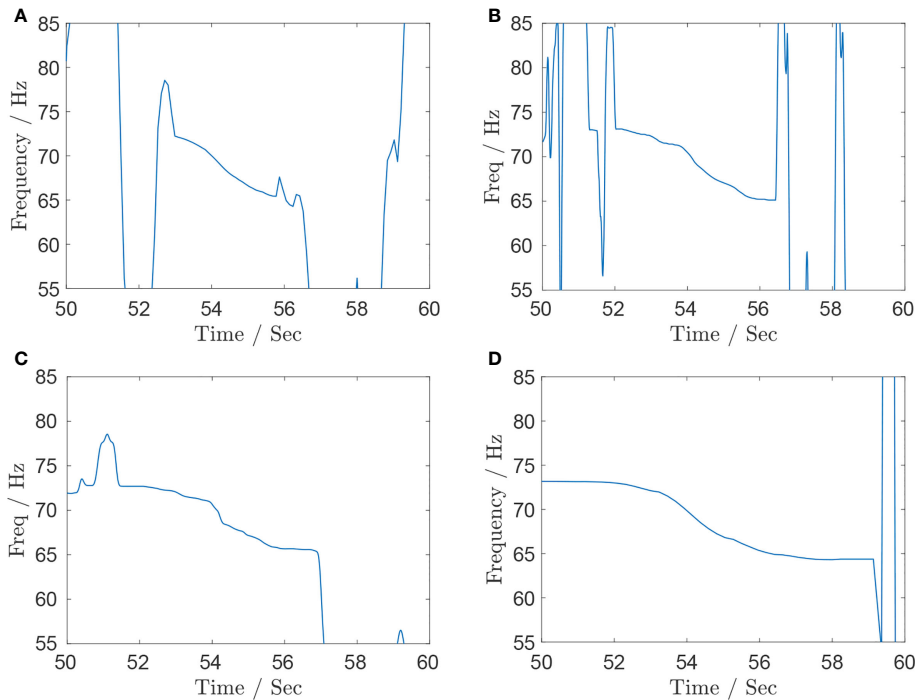


FIGURE 12 Instantaneous frequency estimation results for the aircraft acoustic data recorded by hydrophone with (A) STFT, (B) PCT, (C) SST, and (D) APP-MLS.

According to the first- and second-order derivative of  $f_d(t)$ , the rest of the two flight parameters  $\{h, w_{\min}\}$  are obtained by solving a set of equations as follows:

$$\frac{df_d(t)}{dt} = f'_d(f_0, v, h, w_{\min}) = -3.619, \quad (36)$$

and

$$\frac{d^2f_d(t)}{dt^2} = f''_d(f_0, v, h, w_{\min}) = -2.429. \quad (37)$$

Then, we obtain  $h=150.1m$  and  $w_{\min}=1.4m$ .

## 5 Conclusion

This paper proposes the use of both time–frequency trajectory and three-dimensional space model to estimate the flight parameters, including  $\{f_0, v, \alpha(t), \beta(t)\}$ . This solution explores a computational approach using the geometric relationships of three-dimensional space and using the result of time–frequency estimation to create a predictive model for the instantaneous frequency observed at the hydrophone position by using Doppler effect calculations. The accuracy of this model will be evaluated *via* comparison with given experimental data. The model will then be used to extract the propeller frequency, speed, height, and slant of the aircraft source for a given hydrophone recording.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

## Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

## Funding

This Project are supported by the Key Laboratory of Intelligent Image Processing and Analysis(Wenzhou, China, No.2021 HZSY0071) and the Key Project of Zhejiang Provincial Natural Science Foundation under Grant (No.LZ20F020004).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Aksenovich, T. V. (2020). "Comparison of the use of wavelet transform and short-time fourier transform for the study of geomagnetically induced current in the autotransformer neutral," in *2020 international multi-conference on industrial engineering and modern technologies (FarEastCon)*. pp. 1–5. doi: 10.1109/FarEastCon50210.2020.9271210
- Aygun, H., Kirmizi, M., and Turan, O. (2022). Propeller effects on energy, exergy and sustainability parameters of a small turboprop engine. *Energy* 249, 123759. doi: 10.1016/j.energy.2022.123759
- Boashash, B. (2015). *Time-frequency signal analysis and processing: A comprehensive reference* (Massachusetts, Washington, DC: Academic Press).
- Boyd, S., Boyd, S. P., and Vandenberghe, L. (2004). *Convex optimization* (University Printing House, Cambridge: Cambridge university press).
- Chirico, G., George, N., Barakos and Nicholas Bown. (2018). Propeller installation effects on turboprop aircraft acoustics. 424, 238–62. doi:10.1016/j.jsv.2018.03.003
- Fang, J., Wang, F., Shen, Y., Li, H., and Blum, R. S. (2016). Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach. *IEEE Trans. Signal Process.* 64, 4649–4662. doi: 10.1109/TSP.2016.2572041
- Kozick, R. J., and Sadler, B. M. (2000). Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures. *IEEE Trans. Signal Process.* doi: 10.1109/78.887045
- Kumar, A., and Prasad, A. (2022). Wigner-ville distribution function in the framework of linear canonical transform. *J. Pseudo Differential Operators Appl.* 13, 1–17. doi: 10.1007/s11868-022-00471-w
- Liu, S., Zhang, Y. D., and Shan, T. (2018). Detection of weak astronomical signals with frequency-hopping interference suppression. *Digital Signal Process.* 72, 1–8. doi: 10.1016/j.dsp.2017.09.003
- Qiao, C., Shi, Y., Diao, Y. X., Calhoun, V. D., and Wang, Y. P. (2020). Log-sum enhanced sparse deep neural network. *Neurocomputing* 407, 206–220. doi: 10.1016/j.neucom.2020.04.118
- Urlick, R. J. (1972). Noise Signature of an Aircraft in Level Flight over a Hydrophone in the Sea[J]. *J. Acoust. Soc. Am.* 52 (1A), 993–999.
- Vandana, A. R., and Arokiasamy, M. (2022). Comparison of radar micro doppler signature analysis using short time fourier transform and discrete wavelet packet transform. doi: 10.1007/978-981-16-9012-9\_35
- Vucijak, N. M., and Saranovac, L. V. (2010). A simple algorithm for the estimation of phase difference between two sinusoidal voltages. *IEEE Trans. Instrum. Meas.* 59, 3152–3158. doi: 10.1109/TIM.2010.2047155
- Wang, H., and Kay, S. (2010). Maximum likelihood angle-doppler estimator using importance sampling. *IEEE Trans. Aerospace Electronic Syst.* 46, 610–622. doi: 10.1109/TAES.2010.5461644
- Wang, H., Kay, S., and Saha, S. (2008). An importance sampling maximum likelihood direction of arrival estimator. *IEEE Trans. Signal Process.* 56, 5082–5092. doi: 10.1109/TSP.2008.928504
- Wang, P., Orlik, P. V., Sadamoto, K., Tsujita, W., and Gini, F. (2017). Parameter estimation of hybrid sinusoidal fm-polynomial phase signal. *IEEE Signal Process. Lett.* 24, 66–70. doi: 10.1109/LSP.2016.2638436
- Zcza, B., Szq, A., Xian, J. A., Pyh, A., Xys, A., and Ayw, A. (2021). Linear canonical wigner distribution of noisy lfm signals via variance-snr based inequalities system analysis. *Optik* 237, 166712. doi: 10.1016/j.ijleo.2021.166712
- Zhang, W., Fu, L., and Liu, Q. (2019). Nonconvex log-sum function-based majorization-minimization framework for seismic data reconstruction. *IEEE Geosci. Remote Sens. Lett.* PP 1–5, 1776–1780. doi: 10.1109/LGRS.2019.2909776
- Zhang, X., Wu, H., Sun, H., and Ying, W. (2021). Multireceiver sas imagery based on monostatic conversion. *IEEE J. Selected Topics Appl. Earth Observations Remote Sens.* 14, 10835–10853. doi: 10.1109/JSTARS.2021.3121405
- Zhang, X., Yang, P., Huang, P., and Ying, W. (2022). Wide-bandwidth signal-based multireceiver sas imagery using extended chirp scaling algorithm. *IET Radar Sonar Navigation.* doi: 10.1049/rsn.2.12200
- Zhang, X., Ying, W., Yang, P., and Sun, M. (2020). Parameter estimation of underwater impulsive noise with class b model. *IET Radar Sonar Navigation.* doi: 10.1049/iet-rsn.2019.0477
- Zhou, C., Zhang, G., Yang, Z., and Zhou, J. (2020). A novel image registration algorithm using wavelet transform and matrix-multiply discrete fourier transform. *IEEE Geosci. Remote Sens. Lett.* PP 1–5, 8002605. doi: 10.1109/LGRS.2020.3031335



## Appendix

$$\frac{df_d}{dt} = \frac{f_0 c_w v^2}{\sqrt{v^2 t^2 + w_{\min}^2}} \left\{ \frac{d\beta_T(t)}{dt} t (v^2 t^2 + w_{\min}^2) \sin \beta_T \right. \quad (38)$$

$$\left. - (2v^2 t^2 + w_{\min}^2) \cos \beta_T(t) + t \right\} .$$

$$\frac{d\beta}{dt} = (A_1 + A_2 - A_3 + A_4 A_5 + A_3 A_6 A_7^{-1} - A_8 A_9 - A_{10})^{-1} A_{11} . \quad (39)$$

where

$$\begin{aligned} A_1 &= \cos \beta(t) \sqrt{1 - n^2 \cos^2 \beta(t)} , \\ A_2 &= n^2 [1 - n^2 \cos^2 \beta(t)]^{-1/2} \sin^2 \beta(t) \cos \beta(t) , \\ A_3 &= 2n \sin \beta(t) \cos \beta(t) , \\ A_4 &= \cos \beta(t) \sqrt{1 - n^2 \cos^2 \beta(t)} / (2dh) , \\ A_5 &= d^2 / (1 - n^2 \cos^2 \beta(t)) + h^2 / \sin^2 \beta(t) \\ &\quad - [(d+h)^2 + v^2 t^2 + w_{\min}^2] , A_6 = d^2 / (1 - n^2 \cos^2 \beta(t)) + h^2 / \sin^2 \beta(t) \\ &\quad - [(d+h)^2 + v^2 t^2 + w_{\min}^2] , A_7 = 2dh \sqrt{1 - n^2 \cos^2 \beta(t)} , A_8 = \\ &\quad \sin \beta(t) \sqrt{1 - n^2 \cos^2 \beta(t)} / (2dh) , \\ A_9 &= 2d^2 h^2 \sin^2 \beta(t) \cos \beta(t) / (1 - n^2 \cos^2 \beta(t))^2 , \\ A_{10} &= h \sqrt{1 - n^2 \cos^2 \beta(t)} / [d \sin \beta(t)] , \\ A_{11} &= v^2 t \sin \beta(t) \sqrt{1 - n^2 \cos^2 \beta(t)} / (2dh) , \end{aligned} \quad (40)$$

$$\frac{d^2 f_d}{dt^2} = \frac{c_w f_0 v^2 (G_1 + G_2 + G_3 - G_4)}{2G_5} . \quad (41)$$

$$\begin{aligned} G_1 &= -2t \frac{d^2 \beta}{dt^2} \sin \beta(t) (v^2 t^2 + w_{\min}^2)^2 [v^2 t \cos \beta(t) - 1] \\ G_2 &= t \left( \frac{d^2 \beta}{dt^2} \right)^2 (v^2 t^2 + w_{\min}^2)^2 \{ v^2 t [\cos 2\beta(t) - 3] + 2 \cos \beta(t) \} \\ G_3 &= 4 \frac{d^2 \beta}{dt^2} \sin \beta(t) (v^2 t^2 + w_{\min}^2) \\ &\quad (v^2 t w_{\min}^2 \cos \beta(t) + 2v^2 t^2 + w_{\min}^2) \\ G_4 &= 4v^4 t^3 \cos \beta(t) + 3v^4 t^2 w_{\min}^2 \cos 2\beta(t) \\ &\quad + 2v^2 w_{\min}^4 \cos 2\beta(t) + 3v^4 t^2 w_{\min}^4 + 2w_{\min}^2 G_5 = \\ &\quad (v^2 t^2 + w_{\min}^2)^{3/2} + [v^2 t \cos \beta(t) - 1]^3 \end{aligned} \quad (42)$$

$$\frac{d^2 \beta}{dt^2} = [\cos \beta(t) (B_{12} + B_{13}) (B_{14} + nB_{15} - 2B_{16})]^{-1} \quad (43)$$

$$8\beta(t) \left\{ B_1 - \beta(t) \left[ \begin{array}{l} B_2 - B_3 + n(B_4 + B_5 + B_6) \\ + 2B_7 B_8 - B_9 + B_{10} - B_{11} \end{array} \right] \right\} .$$

Where

$$B_1 = 2v^2 t \cos \beta(t) [n^2 \cos 2\beta(t) - 1] [n^2 \cos 2\beta(t) + n^2 - 2]^2 / [\sin^2 \beta(t)]^2 ,$$

$$B_2 = n^6 \cos^6 \beta(t) \sin^{-1} \beta(t)$$

$$\left\{ \begin{array}{l} [\cos 4\beta(t) - 4 \cos 2\beta(t)] (d^2 + h^2 + v^2 t^2 + w_{\min}^2) + \\ 3d^2 + 7h^2 + 3v^2 t^2 + 3 + w_{\min}^2 \end{array} \right\} ,$$

$$B_3 = 2d^2 n^2 (11n^2 - 7) \sin^{-1} \beta(t) ,$$

$$B_4 = 5n^3 \sin \beta(t) (d^2 + h^2 + v^2 t^2 + w_{\min}^2) ,$$

$$B_5 = n \sin \beta(t) \left[ \begin{array}{l} 9d^2 (3n^2 - 2) + h^2 (27n^2 - 14) + \\ (23n^2 - 14) (v^2 t^2 + w_{\min}^2) \end{array} \right] ,$$

$$B_6 = 4dh (n^2 - 2) \sqrt{1 - n^2 \cos^2 \beta(t)} ,$$

$$B_7 = n^4 \cos 4\beta(t) + 2n^2 (n^2 - 2) \cos 2\beta(t) + n^4 - 2n^2 + 4 ,$$

$$B_8 = 2dhn^3 \cos^2 \beta(t) \sin^{-2} \beta(t) \sqrt{1 - n^2 \cos^2 \beta(t)} ,$$

$$B_9 = 2dhn^3 \sin^{-2} \beta(t) [\cos 4\beta(t) + 1] \sqrt{1 - n^2 \cos^2 \beta(t)} ,$$

$$B_{10} = 26h^2 n^4 \sin^{-3} \beta(t) ,$$

$$B_{11} = h^2 [(1 - 8n^2) \cos 2\beta(t) + 12n^4 - 4n^2 + 3] ,$$

$$B_{12} = [n^2 \cos 2\beta(t) \sin^2 \beta(t)] / 2 + n^2 - 2 ,$$

$$B_{13} = v^2 \sin^3 \beta(t) [n^2 \cos 2\beta(t) + n^2 - 2] ,$$

$$B_{14} = \cos 2\beta(t) [-d^2 n^2 (n^2 - 16) + h^2 (-9n^4 + 16n^2 + 8) (v^2 t^2 + w_{\min}^2)] ,$$

$$B_{15} = [n^3 \cos 6(t) - 6n \cos 4\beta(t)] (d^2 + h^2 + v^2 t^2 + w_{\min}^2) - 16\sqrt{2} d h \sin^3 \beta(t) [-n^2 \cos 2\beta(t) - n^2 + 2]^{3/2} ,$$

$$B_{16} = 5d^2 n^2 + h^2 (4n^4 - 11n^2 + 12) + (n^2 + 4) (v^2 t^2 + w_{\min}^2) , \quad (44)$$