



# An Enhanced Data-Driven Array Shape Estimation Method Using Passive Underwater Acoustic Data

Qisong Wu \*<sup>®</sup>, Hao Zhang <sup>®</sup>, Zhichao Lai <sup>®</sup>, Youhai Xu, Shuai Yao and Jun Tao

Key Laboratory of Underwater Acoustic Signal Processing of Ministry of Education, Southeast University, Nanjing 210096, China; 220180930@seu.edu.cn (H.Z.); ghu@seu.edu.cn (Z.L.); 220190913@seu.edu.cn (Y.X.); shuai.yao@seu.edu.cn (S.Y.); juntao@seu.edu.cn (J.T.)

\* Correspondence: qisong.wu@seu.edu.cn

Abstract: Beamforming-based signal enhancement technologies in passive sonar array processing often have poor performance due to array distortion caused by rapid tactical maneuvers of the towed platform, oceanic currents, hydrodynamic effects, etc. In this paper, an enhanced datadriven shape array estimation formulation is proposed using passive underwater acoustic data. Beamforming based on a hypothetically ideal array is firstly employed to perform the detection of narrow-band components from sources of opportunity, and the corresponding phases of these detected narrow-band components are subsequently extracted to acquire time-delay differences. Then, a weighted outlier-robust Kalman smoother is proposed to acquire enhanced estimates of the time-delay differences, since the underlying properties of slowly changing time-delay differences in the hydrophone array and diverse signal to interference and noise ratios in multiple narrowband components are explored; and its Cramer-Rao Lower Bound is also provided. Finally, the hydrophone array shape is estimated based on the estimated time delay differences. The proposed formulation fully exploits directional radiated noise signals from distant underwater acoustic targets as sources of opportunity for real-time array shape estimation, and thus it requires neither the number nor direction of sources to be known in advance. The effectiveness of the proposed method is validated in simulations and real experimental data.

**Keywords:** array shape estimation; passive sonar system; radiated noise; time delay estimation; beamforming

# 1. Introduction

Hydrophone arrays are widely used in underwater acoustics to perform direction-ofarrival (DOA) estimation and beamforming-based signal enhancement in the passive sonar system. Most array signal processing techniques are based on the assumption that the shape of the array remains unchanged. Unfortunately, the shape of a hydrophone array, such as a towed array, is often perturbed due to fluctuations in ship tactical maneuvering, oceanic currents, nonneutral buoyancy, and so on, particularly in the demand of the long-towed arrays for the acquisition of a low-frequency signal source at long range. These would lead to uncertainty in array shape and deviation from the hypothetical one [1–4]. However, beamforming as a typical signal enhancement technique requires knowledge of accurate array positions to properly process the measurements, which means an accurate estimate of the array shape inevitably degrades the beamforming performance [2]. A theoretical derivation of the performance loss in a distorted array shape due to the tactical maneuvering is provided [5]. Therefore, it is requisite to implement an array shape estimation to improve the acoustic system performance.



Article

Citation: Wu, Q.; Zhang, H.; Lai, Z.; Xu, Y.; Yao, S.; Tao, J. An Enhanced Data-Driven Array Shape Estimation Method Using Passive Underwater Acoustic Data. *Remote Sens.* **2021**, *13*, 1773. https://doi.org/10.3390/ rs13091773

Academic Editor: Jaroslaw Tegowski

Received: 19 March 2021 Accepted: 26 April 2021 Published: 2 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

#### 1.1. Related Works

A huge amount of effort has been devoted to array shape estimation. In general, these methods can be classified into source-independent methods and source-dependent methods. For source-independent methods, one effective method is to place depth-measurement sensors and compasses at several points within the array. These additional sensors provide localized respectively vertical and horizontal information on the transverse displacements of the array, whereas it is less economical due to the use of additional sensors. Further work is to model the array shape by a low order polynomial and to use the available sensors' information to determine the coefficients of this polynomial [6]. A Kalman filter technique with a small angle approximation is used to estimate the array shape evolution over time [7], and an extended Kalman filter approach is also considered with the predicted state update following a higher-order nonlinear model [8].

As for source-dependent methods, acoustic hydrophone data would be used to estimate the array shape when there is no sensors' information within the array or there are too few working sensors, and the shape can be estimated by measuring the relative time or phase delays among sensors due to one or more sources with known or unknown directions [9–19]. Historically, the array shape model employed for acoustic-based estimation is often considered to be a horizontal axis with perturbations at regular intervals on the vertical axis. Under this formulation, the Cramer-Rao Lower Bound (CRLB) for the narrow-band plane-wave sources is provided for source localization with a single known source direction or a known angle between the first receiver and the second one [9]. The maximum-likelihood (ML) approach is used to jointly estimate both the array and source parameters for multiple narrow-band or broadband sources with the known number of sources through a two-stage iteration procedure [10]. Subspace-based eigenvector techniques form another way to estimate array shapes when a single source at a known frequency and direction is known [11]. The eigenvector method for estimating the positions of the array elements is based on the eigen-decomposition of the array covariance matrix to acquire the phase delays between adjacent array elements. The statistic of phase delays is analyzed in [12]. The CRLB of the subspace-based array shape estimation method is provided in [3]. The eigenvector method is similar to the method [13], and it can be also regarded as a special case of the maximum-likelihood method [14]. Reference [15] proposes an improved array shape estimation method by jointly exploiting the acoustic hydrophone data and sensor data.

#### 1.2. Contributions

In this paper, an enhanced data-driven array shape estimation formulation is proposed using passive underwater acoustic data. Instead of using the traditional method with a deterministic signal model and a known number of discrete sources, a novel datadriven approach is proposed by exploiting these directionals radiated noise sources, such as shipping and cargo as sources of opportunity for real-time array shape estimation. The conventional beamforming based on the hypothetical uniform linear array is firstly employed to acquire an improved waveform due to spatial gain from the hydrophone array. Since the radiated noise sources involve the amount of narrow-band components that originated from the vibration of mechanical components like diesel generator and air conditioning equipment, we perform the narrow-band frequency detection and extract their corresponding phases for the acquisition of noisy time-delay differences. A novel weighted outlier-robust Kalman smoother (WORKS) method is then proposed to acquire enhanced time-delay estimates by exploring the properties of slowly changing time-delay difference in hydrophone array as well as the diverse signal to interference and noise ratios (SINR) in multiple narrow-band components, and its CRLB is also analyzed and provided. Finally, the hydrophone array shape is obtained using the estimated time delay differences, and thus beamforming-based improved radiated noise signal is acquired. The proposed formulation fully exploits the directional radiated noise source signals as sources of opportunity for real-time array shape estimation, and thus it requires neither the number

nor direction of sources to be known in advance. Furthermore, the proposed WORKS approach has the capability of acquiring the enhanced time-delay difference estimation by treating the weights associated with each observed noisy time-delay difference data in a probabilistic approach.

The contributions of the paper lie in three aspects: (1) An enhanced data-driven array shape estimation formulation using passive underwater acoustic data is provided. The proposed formulation fully exploits the directional radiated noise source signals as sources of opportunity for real-time array shape estimation, and thus it requires neither the number nor direction of sources to be known in advance. (2) A novel weighted outlier-robust Kalman smoother method is proposed to acquire enhanced time-delay estimates by the exploitation of slowly changing time-delay differences in hydrophone array and has the capability of alleviating the effects of the outliers resulted from the diverse noise levels in the multiple frequencies. (3) The CRLB of the proposed method is analyzed.

#### 1.3. Organization and Notation

The paper is organized as follows: The underwater acoustic radiated noise signal model is introduced in Section 2. A novel array shape estimate technique is described in Section 3, and the weighted robust-outlier Kalman smoother method is also proposed in this section. Performance bound is also provided in Section 3. Simulations and experiments demonstrate the robustness of the proposed algorithms compared to conventional methods in Section 4. Concluding remarks are given in Section 5.

Notation: We use lower-case (upper-case) bold characters to denote vectors (matrices).  $p(\cdot)$  denotes the probability density function (pdf), and  $\mathcal{N}(x|a, b)$  denotes that random variable x follows a Gaussian distribution with mean a and variance b. In addition,  $(\cdot)^T$  and  $(\cdot)^H$  denote transpose and conjugate transpose, respectively,  $\mathbf{I}_N$  denotes the  $N \times N$  identity matrix.  $\nabla_{\mathbf{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_N} \end{bmatrix}$  denotes the gradient operator respect to the variable vector  $\mathbf{x} = [x_1, \cdots, x_N]^T \in \mathcal{R}^{N \times 1}$ , and  $\mathbb{E}(x)$  represents the expectation operator of the random variable x.

## 2. Underwater Acoustic Radiated Noise Model in the Distorted Array

In this section, the underwater acoustic radiated noise model in a distorted hydrophone array is introduced. We assume that a thin flexible cable with M omnidirectional hydrophone sensors mounted at fixed inter-element spacing d is employed. Due to the transverse motion of ship tactical maneuvering, oceanic currents, and hydrodynamic effects, it would result in transverse displacements of the array from a straight line, as shown in Figure 1. The positions of the first two sensors define a baseline for the coordinate system and consequently are always assumed to reside on the x axis. It is observed that a deformed array occurs when any one of the inner sensor y positions are no longer zero.



Figure 1. An example of a distorted hydrophone array.

The deformation is caused by bending at the sensor junctions due to the soft and elastic external shell of hydrophone array, which implies that the inter-sensor distances remain constant, and thus we have,

$$d = \sqrt{(x_{m+1} - x_m)^2 + (y_{m+1} - y_m)^2}, \quad m = 0, \cdots, M - 1,$$
(1)

where the location of the *m*th sensor is denoted by  $(x_m, y_m)$  in the x - y Cartesian plane.

With respect to *K* far-field sources of opportunity impinging on the hydrophone array and the source with the bearing or direction of  $\theta_k$ , the range difference between the first two sensors becomes, as shown in Figure 1,

$$\Delta r_{k,0} = r_{k,1} - r_{k,0} \approx d\cos\theta_k = (x_1 - x_0)\cos\theta_k,\tag{2}$$

where  $r_{k,1}$  and  $r_{k,0}$  denote the distances from the *k*th source to the second and reference sensors, respectively, and the result above is based on the far-field hypothesis. However, the locations of sensors deviate from the nominal ones, when the hydrophone array shape is distorted or deformed. Assuming the deviation angle formed by the line between the *m*th sensor and (m + 1)th sensor along *x* axis is denoted as  $\phi_m$ , the range difference becomes,

$$\Delta r_{k,m} \approx d\cos(\theta_k - \phi_m) = (x_{m+1} - x_m)\cos\theta_k + (y_{m+1} - y_m)\sin\theta_k.$$
(3)

For the first two sensors as a baseline,  $\phi_0 = 0$ . The range difference between the *m*th sensor and the reference can be written by,

$$r_{k,m} - r_{k,0} = \sum_{i=0}^{m-1} \Delta r_{k,i} = \sum_{i=0}^{m-1} d\cos(\theta_k - \phi_i) = \sum_{i=0}^{m-1} (x_{i+1} - x_i) \cos\theta_k + (y_{i+1} - y_i) \sin\theta_k.$$
(4)

The corresponding time-delay difference is

$$\tau_{k,m} = \frac{\sum_{i=0}^{m-1} \Delta r_{k,i}}{v} = \sum_{i=0}^{m-1} \Delta \tau_{k,i},$$
(5)

where v is the speed of acoustic wave in water, and  $\Delta \tau_{k,i}$  denotes the time-delay difference between the (i + 1)th and *i*th sensors with respect to the source of opportunity. It should be noted that the accuracies of the array shape estimates will be confined to the accuracies of each estimated time-delay difference for adjacent sensors, giving rise to error accumulations of the time-delay difference according to Equation (5).

Assuming the *k*th radiated noise source signal as  $s_k(t)$  with the bearing of  $\theta_k$ , the received signal  $y_m(t)$  in the *m*th sensor can be written as

$$y_m(t) = \sum_{k=1}^{K} \zeta_{k,m} s_k(t - \tau_{k,m}) + e_m(t),$$
(6)

where  $\zeta_{k,m}$  and  $\tau_{k,m}$ , represent respectively the amplitude and time delay from the source to the *m*th sensor, and  $e_m(t)$  is an additive noise.

According to the statistical model of the underwater acoustic radiated noise [20,21], the corresponding time-domain  $s_k(t)$  is expressed by,

$$s_k(t) = p_{kl}(t) + (1 + p_{km}(t))p_{kc}(t),$$
(7)

where  $p_{kl}(t)$  is a narrow-band signal arised from the vibration of machinery,  $p_{km}(t)$  represents a modulation function due to the propeller movement, and  $p_{kc}(t)$  denotes a stationary hydrodynamic noise.

The narrow-band signal  $p_{kl}(t)$ , which is often generated from the vibration of mechanical components like diesel generator and air conditioning equipment, can be denoted by the superposition of multiple sinusoid signals as,

$$p_{kl}(t) = \sum_{l=1}^{L} \alpha_{kl} \sin(2\pi f_{kl} t),$$
(8)

where  $\alpha_{kl}$  is an amplitude of the *l*th single-tone frequency from the source, *L* is the number of the sinusoid signals, and  $f_{kl}$  is the corresponding frequency.

In the real sonar system, the narrow-band signals in the underwater acoustic radiation noise often have high source levels owing to inevitable mechanical vibration, and thus they are easier to be detected and recognized, compared to the broadband hydrodynamic noise signals [20,21]. Figure 2 shows a typical spectrum example of radiated noise from a ship. It is observed that there exist several narrow-band frequencies that have higher powers and are prominent in their nearby frequency ranges, their phases involve the information of time delays between the sources of opportunity and sensors in the deformed array. It possibly provides a way to estimate the distorted array shape.



Figure 2. Spectrum example of real radiated noise.

# 3. Array Shape Estimation

In this section, a novel data-driven array shape estimation formulation will be provided, and the weighted robust-outlier Kalman smoother method is proposed to acquire an enhanced time-delay difference estimation.

#### 3.1. Phase Extraction Based on Detected Narrow-Band Frequencies

Assume a passive sonar system with the sample rate of  $F_s$  and the time interval of T, and thus the total number of sample is  $N = F_s T$ . In accordance with Equation (6), the observed signal  $y_m(t)$  at the *m*th sensor is decomposed into N narrow-band components using Discrete Fourier Transform (DFT). The resulting narrow-band components corresponding to any one of the frequency points may be expressed as,

$$y_m(f_n) = \sum_{k=1}^K \zeta_{k,m}(f_n) s_k(f_n) \exp(-j2\pi f_n \tau_{k,m}) + e_m(f_n), \quad n \in (1, \cdots, N),$$
(9)

where  $f_n = nF_s/N$  is the *n*th frequency grid,  $\zeta_{k,m}(f_n)$  is the coefficient at the  $f_n$  frequency with a high value, when a strong sinusoidal noise signal at the frequency of  $f_n$  exists, and *K* is the number of the far-field sources of opportunity.

$$\mathbf{w}(f_n, \vartheta_q) = \left[1, \exp\left(\frac{-j2\pi f_n d\cos\vartheta_q}{v}\right) \cdots, \exp\left(\frac{-j2\pi f_n (M-1)d\cos\vartheta_q}{v}\right)\right]^T, \quad (10)$$

where  $\vartheta_q \in (0, 180)$ ,  $q \in 1, \dots, Q$ , and Q is the number of the pre-defined bearing grid. According to weights in Equation (10), beamforming is achieved by

$$\varrho(\boldsymbol{\vartheta}_k) = \sum_{n=1}^{N} \left| \mathbf{w}(f_n, \boldsymbol{\vartheta}_k)^H \mathbf{y}(f_n) \right|^2, \quad k \in 1, \cdots, Q,$$
(11)

where  $\mathbf{y}(f_n) = [y_1(f_n), \dots, y_M(f_n)]^T \in C^M$  is a received signal vector in the *n*th frequency point. According to the result in Equation (11), the  $\hat{\vartheta}_k$  is estimated based on the maximum-value detection criterion. The output signal based on the estimated  $\hat{\vartheta}_k$  is,

$$\tilde{y}_k(n) = \sum_{m=1}^M y_m \left[ n - \frac{(m-1)d\cos\hat{\vartheta}_k}{v} \right], \ n \in (1, \cdots, N),$$
(12)

where *M* is the number of the hydrophone sensors.

The operations above in Equations (9)–(12) are involved in the basic wideband frequency beamforming method. The weight vector in Equation (10) is optimal, and the signal to noise ratio (SNR) of the output signal  $\tilde{\mathbf{y}}_k = [\tilde{y}_k(1), \dots, \tilde{y}_k(N)]^T$  can be improved by Mtimes, when the array is nominally uniform linear in the Gaussian noise case. However, the array should be deformed severely and not remain a straight line due to the transverse motion of the towing vessel, oceanic currents, and hydrodynamic effects, weights in Equation (10) would be wrong. Therefore, the performance of beamforming would worsen dramatically, when the hydrophone array greatly deviates from the nominal one. Taking a uniform linear towed array of M = 128 hydrophone sensors with a spatial interval of 1.6 m for example, the shape of the distorted array approximately becomes a parabola and deviation angle  $\phi_m$  is  $\sin(\pi/(2M))$ , the average loss of the narrow-band signals would reach up to about 15 dB.

The output signal  $\tilde{\mathbf{y}}_k$  based on the nominal array shape would not be the optimal one in the distorted array cases because the time delays in the weight vector Equation (10) do not match the true ones. Considering the fact that an important component in radiated noise originates from the machinery noise and can be regarded as the superposition of multiple narrowband signals, we would utilize the phases of these narrow-band signals to estimate time delays in the distorted array instead of the time delays in the nominal array. Instead of directly performing the element-space narrow-band detection, we carry out the detection on the spectrum of  $P_{yk}(f)$ , which is acquired by performing the DFT operation on the output signal  $\tilde{\mathbf{y}}_k$ . By exploiting the background calibration and threshold-based detection techniques, the detected narrow-band frequencies radiated by the *k*th source of opportunity can be denoted as  $\hat{f}_{kl}$  with  $l \in 1, \dots, L_0$ , where  $L_0$  is the number of detected narrow-band frequencies. The corresponding unwrapped phase of the *l*th narrow-band frequency from the *k*th source to the *m*th sensor can be acquired by,

$$\psi_m(\hat{f}_{kl}) = 2\pi \hat{f}_{kl} \tau_{k,m}.\tag{13}$$

In a similar way, we obtain the phase of  $\psi_{m+1}(\hat{f}_{kl})$  in the (m+1)th hydrophone, and thus estimate the adjacent phase difference by

$$\Delta \hat{\psi}_{k,l,m} = \psi_{m+1}(\hat{f}_{kl}) - \psi_m(\hat{f}_{kl}), \quad l \in 1, \cdots, L_0.$$
(14)

$$\Delta \hat{\tau}_{k,l,m} = \frac{\psi_{m+1}(\hat{f}_{kl}) - \psi_m(\hat{f}_{kl})}{2\pi \hat{f}_{kl}}, \quad l \in 1, \cdots, L_0.$$
(15)

## 3.2. Proposed Weighted Outlier-Robust Kalman Smoother

Ideally, the estimated time-delay difference  $\Delta \hat{\tau}_{k,l,m}$  between the (m + 1)th hydrophone sensor the *m*th hydrophone sensor expressed by Equation (14) should be independent of narrow-band frequencies, and be identical among  $L_0$  narrow-band frequencies. The reason is that its phase difference is only determined by the spatial interval *d* between adjacent hydrophones and the bearing of impinging source. However, the time-delay differences  $\Delta \hat{\tau}_{k,l,m}$  with  $l \in 1, \dots, L_0$  are often diverse in the real case. Figure 3 shows an example in a public SWellEx-96 HLA dataset [22]. It is observed that the time-delay differences in the detected narrow-band frequencies are completely different across sensors. The possible reason is that there are diverse SINRs in these detected narrow-band frequencies, and these time-delay differences will deviate from the true values at these frequencies.



Figure 3. Time-delay differences varying with narrow-band frequencies in real data case.

One way to alleviate the fluctuations of  $\Delta \hat{\tau}_{k,l,m}$  is average these estimated time-delay differences and acquire [23]

$$\Delta \hat{\tau}_{k,m} = \frac{1}{L_0} \sum_{l=1}^{L_0} \Delta \hat{\tau}_{k,l,m}, \quad m \in 0, \cdots, M-2.$$
(16)

However, the average operator being a linear estimator is sensitive to outliers and easily deviates from the true value when there exists a small number of outliers. The  $\Delta \hat{\tau}_{k,l,m}$  acquired by the adjacent phase differences easily deteriorates in the low SINR cases, and thus become abnormal. In this case, these outliers would lead to a large deviation in the average operator. Therefore, the average operator is not an optimal way to acquire the time-delay estimation in the distorted array. To address this issue, we propose to use a weighted robust-outlier Kalman smoother.

Instead of averaging all the  $\Delta \hat{\psi}_{k,l,m}$  in multiple narrow-band frequencies, we stack  $\Delta \hat{\psi}_{k,l,m}$  with  $l = 1, \dots, L_0$  into a vector as  $\mathbf{z}_{k,m} = \left[\Delta \hat{\psi}_{k,1,m}, \dots, \Delta \hat{\psi}_{k,L_0,m}\right]^T$ , and regard  $\mathbf{z}_{k,m}$  as the observation data sample. In conformity with the analysis above, the observation equation is expressed by,

$$\mathbf{z}_{k,m} = \boldsymbol{\alpha}(f) \times \Delta \tau_{k,m} + \boldsymbol{\epsilon}_{m},\tag{17}$$

where  $\boldsymbol{\alpha}(f) = [2\pi f_1, \dots, 2\pi f_{L_0}]^T \in \mathcal{R}^{L_0 \times 1}$  is a vector,  $\Delta \tau_{k,m}$  is the true time-delay difference between the (m + 1)th and the *m*th sensors, and  $\boldsymbol{\epsilon}_m \in \mathcal{R}^{L_0 \times 1}$  is an additive noise.

According to Equation (17), the observed time-delay differences in multiple narrowband frequencies can be considered as the true time-delay difference plus noise. To alleviate the effects of outliers and diverse noise levels in the multiple frequencies, we introduce a novel Bayesian algorithm that treats the weights associated with each observation data samples probabilistically as proposed by [24,25]. However, unlike [24,25] which use a scalar to weight the observed data sample and all elements in observed data sample are equally treated, we introduce a vector weight  $\gamma_m$  for each observed data sample  $\mathbf{z}_{k,m}$  such that the variances of  $\mathbf{z}_{k,m}$  are weighted with  $\gamma_m$ , owing to diverse SINRs in the corresponding narrow-band frequencies, and have,

$$\mathbf{z}_{k,m}|\boldsymbol{\gamma}_{m},\mathbf{r}\sim\mathcal{N}\left(\mathbf{z}_{k,m}|\boldsymbol{\alpha}(f)\Delta\tau_{k,m},\boldsymbol{\Gamma}_{m}^{-1}\mathbf{R}\right),\tag{18}$$

where  $\mathbf{R} = \operatorname{diag}(\mathbf{r})$ , which is a covariance matrix for observation noise, is a diagonal matrix with  $\mathbf{r} = [r_1, \cdots, r_{L_0}]^T \in \mathcal{R}^{L_0 \times 1}$  on its diagonal.  $\Gamma_m = \operatorname{diag}(\gamma_m)$  is a diagonal matrix with  $\gamma_m = [\gamma_{1m}, \cdots, \gamma_{L_0m}]^T \in \mathcal{R}^{L_0 \times 1}$  on its diagonal. It should be noted that each elements in  $\mathbf{r}$  is weighted by the corresponding  $\gamma_m$ . A Gamma prior distribution for the weights in order to ensure they remain positive is imposed on the weight vector  $\gamma_m$  as,

$$\gamma_{lm} \sim \text{Gamma}(\gamma_{lm}|a_0, b_0), \quad l = 1, \cdots, L_0, \quad m = 0, \cdots, M - 2,$$
(19)

where Gamma(·) represents a Gamma distribution, and  $a_0$  and  $b_0$  are hyperparameters respectively, and often set to  $a_0 = b_0$ .

Considering the fact that external shell of hydrophone array is soft and elastic, it is reasonable to assume the time-delay differences  $\Delta \tau_{k,m}$  are slowly changing across sensors (These  $\Delta \tau_{k,m}$  with  $m = 0, \dots, M - 2$  should be identical, when the array is ideally straight), and thus we have the state equation as

$$\Delta \tau_{k,m+1} = \Delta \tau_{k,m} + \varepsilon_m, \quad m = 0, \cdots, M - 2, \tag{20}$$

where  $\Delta \tau_{k,m+1}$  is the time-delay difference between the (m + 2)th sensor and the (m + 1)th sensor, and  $\varepsilon_m$  is an additive noise. For the m = 0, the initialization of  $\Delta \tau_{k,0}$  is assumed to follow the Gaussian distribution of  $\Delta \tau_{k,0} \sim \mathcal{N}(\Delta \tau_{k,0}|0, \nu_0)$ . Without loss of generality, a Gaussian prior is considered to model this noise, and can be expressed by,

$$\varepsilon_m | \nu \sim \mathcal{N}(\varepsilon_m | 0, \nu),$$
 (21)

where  $\nu$  is the variance of state noise.

In accordance with these time-delay differences in the *M* sensors, we have observed data samples  $\Delta \tau_{k,0:M-2}$  at one time. To estimate the posterior distributions of the random variables and parameter values, we should consider the log evidence of the data samples observed, i.e., log  $p(\mathbf{z}_{k,0:M-2}, \Delta \tau_{k,0:M-2}, \gamma_{0:M-2})$ . It is a typical Kalman smoother.

We can treat this entire problem as an Expectation Minimization (EM) learning problem [26]. The expectation of the complete data likelihood should be taken with respect to the true posterior distribution of all hidden variables  $q(\Delta \tau_{k,0:M-2}, \gamma_{0:M-2})$ . However, since this is an analytically intractable expression, we make a factorial approximation of the true posterior as follows [27],

$$q(\Delta \tau_{k,0:M-2}, \gamma_{0:M-2}) = \prod_{m=0}^{M-2} q(\gamma_m) \prod_{m=0}^{M-3} q(\Delta \tau_{k,m+1} | \Delta \tau_{k,m}) q(\Delta \tau_{k,0}).$$
(22)

This factorization of  $\Delta \tau_{k,0:M-2}$  considers the influence of each  $\Delta \tau_{k,m}$  within its Markov chain. According to this factor approximation, all resulting posterior distributions over hidden variables become analytically tractable. We can derive the final EM update equations from standard manipulations of Gaussian and Gamma distributions and acquire the following,

# E-step : Forward recursion,

$$\mathcal{N}(\Delta \tau_{k,m} | \mu_m, v_m) \propto \mathcal{N}\left(\mathbf{z}_{k,m} | \boldsymbol{\alpha}(f) \Delta \tau_{k,m}, \boldsymbol{\Gamma}^{-1} \mathbf{R}\right) \mathcal{N}(\Delta \tau_{k,m} | \mu_{m-1}, \langle v \rangle + v_{m-1}), \quad (23)$$

$$\rightarrow \mu_m = v_m \left( \sum_{l=1}^{L_0} \frac{\langle \gamma_{lm} \rangle}{\langle r_l \rangle} z_{k,m,l} + \frac{2\pi f_l \mu_{m-1}}{\langle \nu \rangle + v_{m-1}} \right), \tag{24}$$

$$\rightarrow v_m = \left(\sum_{l=1}^{L_0} \frac{\langle \gamma_{lm} \rangle}{\langle r_l \rangle} + \frac{1}{\langle \nu \rangle + v_{m-1}}\right)^{-1},\tag{25}$$

$$\rightarrow \langle \gamma_{lm} \rangle = \frac{a_0 + 1/2}{\frac{1}{2\langle r_l \rangle} (z_{k,m,l} - 2\pi f_l \mu_m)^2 + b_0} \quad l \in \{1, \cdots, L_0\}.$$
 (26)

Backward recursion,

$$\hat{\gamma}(\Delta\tau_{k,m}) = \mathcal{N}(\Delta\tau_{k,m}|\hat{\mu}_m, \hat{v}_m) \propto \mathcal{N}(\hat{\mu}_{m+1}|\Delta\tau_{k,m}, \langle\nu\rangle)\mathcal{N}(\Delta\tau_{k,m}|\mu_m, v_m),$$
(27)

$$\rightarrow \hat{\mu}_m = \hat{v}_m \left( \frac{\mu_m}{v_m} + \frac{\mu_{m+1}}{\langle \nu \rangle} \right) = \mu_m + \frac{v_m}{v_m + \langle \nu \rangle} (\hat{\mu}_{m+1} - \mu_m), \tag{28}$$

$$\rightarrow \hat{v}_m = v_m + \frac{v_m^2}{\left(v_m + \langle v \rangle\right)^2} [\hat{v}_{m+1} - \left(v_m + \langle v \rangle\right)]. \tag{29}$$

M-step :

$$\langle r_l \rangle = \sum_{i=1}^m \frac{\langle \gamma_{li} \rangle}{m} (z_{k,i,l} - \langle \Delta \tau_{k,i} \rangle)^2, \tag{30}$$

$$\langle \nu \rangle = \sum_{i=1}^{m} \frac{1}{m} \left( \left\langle \Delta \tau_{k,i} \right\rangle - \left\langle \Delta \tau_{k,i-1} \right\rangle \right)^2. \tag{31}$$

Once these time-delay differences  $\{\mu_m\}_{m=0}^{M-1}$  have been obtained, we can estimate the array shape. Since the first two sensors define the the baseline, the estimated bearing  $\hat{\vartheta}_k$  of the *k*th noise source can be calculated based on the estimated time-delay  $\hat{\mu}_0$  according to Equation (3),

$$\hat{\vartheta}_k = \arccos\left(\frac{v\hat{\mu}_1}{d}\right).$$
 (32)

Therefore, the estimated deviation angle  $\hat{\phi}_m$  can be calculated as according to Equation (3),

$$\hat{\phi}_m = \arccos\left(\frac{v\hat{\mu}_1}{d}\right) - \arccos\left(\frac{v\hat{\mu}_m}{d}\right) \tag{33}$$

The location  $(\hat{x}_{m+1}, \hat{y}_{m+1})$  of (m+1)th sensor can be estimated by ,

$$\hat{x}_{m+1} = \hat{x}_m + v\hat{\mu}_m \cos(\hat{\phi}_m) + \sqrt{d^2 - v^2 \hat{\mu}_m^2 \sin(\hat{\phi}_m)}$$
(34)

$$\hat{y}_{m+1} = \hat{y}_m + v\hat{\mu}_m \sin(\hat{\phi}_m) + \sqrt{d^2 - v^2 \hat{\mu}_m^2 \cos(\hat{\phi}_m)},$$
(35)

where  $m = 2, \cdots, M - 1$ .

Finally, we perform the beamforming and acquire the improved output signal based on the estimated array shape as

$$\bar{y}_k(n) = y_1 + \sum_{m=1}^{M-1} y_m \left[ n - \sum_{i=1}^m \hat{\mu}_{i-1} \right],$$
(36)

where  $n = [1, \dots, N]$ .

3.3. Summary of Proposed Distorted Array Shape Estimation Scheme

The proposed distorted array shape estimation scheme is summarized in Table 1.

Table 1. Summary of Proposed Distorted Array Shape Estimation Scheme.

- Given the observed array data  $\{y_m(t)\}_{m=1}^M$ . (1) Calculate  $\tilde{\mathbf{y}}_k$  with respect to the *k*th source based on the hypothetical uniform linear array.
- (2) Acquire the detected narrow-band frequencies  $\hat{f}_{kl}$  with  $l \in 1, \dots, L_0$  based on  $y_m(t)$ .

(3) Calculate time-delay differences vector  $\mathbf{z}_{k,m} = \left[\Delta \hat{\psi}_{k,1,m}, \cdots, \Delta \hat{\psi}_{k,L_0,m}\right]^T$ ,  $m = 0, \cdots, M - 2$ . (4) Estimate the time-delay differences using proposed Weighted Robust-Outlier Kalman

Smoother:

a. Initialize hyperparameters  $a_0$  and  $b_0$ .

- b. Initialize covariance for observation noise **R**, variance of state noise  $\nu$ .
- c. Estimate time-delay differences  $\hat{\mu}_m$  using EM algorithm Equations (23)–(31).
- (5) Perform array shape estimates  $(\hat{x}_m, \hat{y}_m)$  with  $m = 2, \dots, M-1$  using Equations (34) and (35).
- (6) Perform the beamforming and acquire the improved output signal  $\bar{y}_k(t)$ .

## 3.4. Performance Bound Analysis

To quantify the estimation performance of the proposed method, we perform the CRLB analysis in this section. Considering the problem of estimating the time-delay difference vector  $\Delta \tau_k = [\Delta \tau_{k,0}, \cdots, \Delta \tau_{k,M-2}]^T$  from a set of observed data vectors of  $\mathbf{z}_{k} = \left[\mathbf{z}_{k,0}^{T}, \cdots, \mathbf{z}_{k,M-2}^{T}\right]^{T}$  with  $\mathbf{z}_{k,m} = \left[\Delta \hat{\psi}_{k,1,m}, \cdots, \Delta \hat{\psi}_{k,L_{0},m}\right]^{T}$ , we can acquire the joint probability distribution function (pdf)  $p(\Delta \tau_{k}, \mathbf{z}_{k})$  from the *k*th source as based on Equations (18) and (20),

$$p(\Delta \boldsymbol{\tau}_{k}, \boldsymbol{z}_{k}) = \int_{\boldsymbol{\gamma}_{m}} \prod_{m=0}^{M-2} \mathcal{N}\Big(\boldsymbol{z}_{k,m} | \boldsymbol{\alpha}(f) \Delta \boldsymbol{\tau}_{k,m}, \boldsymbol{\Gamma}_{m}^{-1} \mathbf{R}\Big) \mathcal{N}(\Delta \boldsymbol{\tau}_{k,m+1} | \Delta \boldsymbol{\tau}_{k,m}, \boldsymbol{\nu}) \\ \times \mathcal{N}(\Delta \boldsymbol{\tau}_{k,0} | \boldsymbol{0}, \boldsymbol{\nu}_{0}) \operatorname{Gamma}(\boldsymbol{\gamma}_{m} | \boldsymbol{a}_{0}, \boldsymbol{b}_{0}) d\boldsymbol{\gamma}_{m}.$$
(37)

The Fisher information matrix (FIM)  $J(\Delta \tau_k)$  for the parameter vector  $\Delta \tau_k$  is defined as follows,

$$\mathbf{J}(\Delta \boldsymbol{\tau}_k) = -\mathbb{E}\Big\{\nabla_{\Delta \boldsymbol{\tau}_k} \big[\nabla_{\Delta \boldsymbol{\tau}_k} \ln p(\Delta \boldsymbol{\tau}_k, \mathbf{z}_k)\big]^T\Big\}.$$
(38)

When the derivatives and expectation exist, we acquire an alternative formula for the FIM as,

$$\mathbf{J}(\Delta \boldsymbol{\tau}_k) = \mathbb{E}\Big\{ \big[ \nabla_{\Delta \boldsymbol{\tau}_k} \ln p(\Delta \boldsymbol{\tau}_k, \mathbf{z}_k) \big]^T \nabla_{\Delta \boldsymbol{\tau}_k} \ln p(\Delta \boldsymbol{\tau}_k, \mathbf{z}_k) \Big\}.$$
(39)

Let  $\Delta \hat{\tau}_k = \Delta \hat{\tau}_k(\mathbf{z}_k)$  be an estimated of the parameter vector  $\Delta \boldsymbol{\tau}_k$ . The mean-square error matrix (MSEM) is defined as

$$\Xi_k(\Delta \hat{\tau}_k) = \mathbb{E}\Big[ (\Delta \tau_k - \Delta \hat{\tau}_k) (\Delta \tau_k - \Delta \hat{\tau}_k)^T \Big].$$
(40)

In fact, the MSEM is bounded by an inverse of the FIM, and we acquire [28],

$$\Xi_k(\Delta \hat{\boldsymbol{\tau}}_k) \ge \mathbf{J}^{-1}(\Delta \boldsymbol{\tau}_k). \tag{41}$$

The inverse of  $\mathbf{J}^{-1}(\Delta \tau_k)$  is just the well-known posterior or Bayesian Cramer–Rao bound [29], and it will be denoted as  $C(\Delta \tau_k)$ . According to the proposed generative model

in Section 3.2, the logarithm of joint pdf of the state and measurement histories  $p(\mathbf{z}_k, \Delta \boldsymbol{\tau}_k)$  can be expressed as,

$$\ln p(\mathbf{z}_{k}, \Delta \boldsymbol{\tau}_{k}) = \sum_{m=0}^{M-2} \ln p(\mathbf{z}_{k,m} | \Delta \tau_{k,m}) + \sum_{m=1}^{M-2} \ln p(\Delta \tau_{k,m} | \Delta \tau_{k,m-1}) + \ln p(\Delta \tau_{k,0}).$$
(42)

Motivated by [28], we follow a similar recursive approach for the smoothing CRLBs, and express CRLBs for the filtering as in the forward recursion,

$$C_{m|m}^{-1} = C_{m|m-1}^{-1} + S_m^m, (43)$$

$$C_{m+1|m}^{-1} = G_{m+1}^{m+1} - G_{m+1}^{m+1,m} (G_{m+1}^m + C_{m|m}^{-1})^{-1} G_{m+1}^{m+1,m},$$
(44)

$$C_{0|-1}^{-1} = G_0^0, (45)$$

where  $m = 0, \dots, M - 2$ . We acquire the CRLBs of the smoother in the backward recursion [28],

$$C_{m|M-2}^{-1} = C_{m|m}^{-1} + G_{m+1}^{m} - G_{m+1}^{m+1,m} (G_{m+1}^{m+1} + C_{m+1|M-2}^{-1} + C_{m+1|m}^{-1})^{-1} G_{m+1}^{m+1,m}.$$
 (46)

where,

$$G_{m+1}^{m} = \mathbb{E}\left\{-\nabla_{\Delta\tau_{k,m}}\left[\nabla_{\Delta\tau_{k,m}}\ln p(\Delta\tau_{k,m+1}|\Delta\tau_{k,m})\right]\right\},\tag{47}$$

$$G_{m+1}^{m,m+1} = \mathbb{E}\left\{-\nabla_{\Delta\tau_{k,m+1}}\left[\nabla_{\Delta\tau_{k,m}}\ln p(\Delta\tau_{k,m+1}|\Delta\tau_{k,m})\right]\right\},\tag{48}$$

$$G_{m+1}^{m+1} = \mathbb{E}\left\{-\nabla_{\Delta\tau_{k,m+1}}\left[\nabla_{\Delta\tau_{k,m+1}}\ln p(\Delta\tau_{k,m+1}|\Delta\tau_{k,m})\right]\right\},\tag{49}$$

$$S_m^m = \mathbb{E}\Big\{-\nabla_{\Delta\tau_{k,m}} \left[\nabla_{\Delta\tau_{k,m}} \ln p(\mathbf{z}_{k,m} | \Delta\tau_{k,m})\right]\Big\}.$$
(50)

According to Equations (18) and (19) in the generative model,  $p(\mathbf{z}_{k,m})$  would follow a multivariate Student distribution by integrating the weight vector  $\gamma_m$ , and can be expressed by,

$$\mathbf{z}_{k,m} | \Delta \tau_{k,m}, \sim \operatorname{st}(\boldsymbol{\alpha}(f) \Delta \tau_{k,m}, \mathbf{R}, v), \tag{51}$$

where st(·) denotes the student-t distribution, and the freedom  $v = 2a_0$ .

According to [28], we derive the recursive formula for the CRLB of the proposed WORKS method as in the forward recursion,

$$C_{m|m}^{-1} = C_{m|m-1}^{-1} + \frac{v+L}{v+L+2} \boldsymbol{\alpha}(f)^{H} \mathbf{R}^{-1} \boldsymbol{\alpha}(f),$$
(52)

$$C_{m+1|m}^{-1} = \nu^{-1} - \nu^{-1} (\nu^{-1} + C_{m|m}^{-1})^{-1} \nu^{-1},$$
(53)

$$C_{0|-1}^{-1} = \nu_{k,0}.$$
(54)

and we acquire the corresponding recursive formula for the CRLB as in the backward recursion,

$$C_{m|M-2}^{-1} = C_{m|m}^{-1} + \nu^{-1} - \nu^{-1}(\nu^{-1} + C_{m+1|M-2}^{-1} + C_{m+1|m}^{-1})^{-1}\nu^{-1},$$
(55)

where  $m = 0, \cdots, M - 2$ .

## 4. Simulations and Experiments

In this section, both simulations and experiments will verify the effectiveness and correctness of the proposed method. Several typical source-dependent approaches, such as subspace-based eigenvector method [11] and average operation [23], are used for comparisons in simulations and experiments.

# 4.1. Simulations

Suppose that a uniform linear towing array has M = 128 hydrophone sensors with a spatial interval of 1.6 m, and the bearing of a radiated noise source is 60°. The deviation angle  $\phi_m$  is set to be  $\sin(\pi/(2M))$ , and thus the shape of the distorted array approximately becomes a parabola, as the black line shown in Figure 4a. The time duration is T = 5 s and the sampling rate  $F_s = 4$  kHz. According to the characteristics of radiated noise [20,21], a three parameters model with  $f_c = 500$  Hz,  $f_m = 200$  Hz and  $\kappa = 0$  is used to model the continuous spectrum component. In addition, seven narrow-band (sinusoid) signals with frequencies of 59 Hz, 97 Hz, 125 Hz, 163 Hz, 198 Hz, 232 Hz and 280 Hz are considered. Without loss of generality, additive Gaussian noise and additional interferences in three narrow-band frequencies with 59 Hz, 97 Hz, and 163 Hz are considered with the SINR of -10 dB.



**Figure 4.** Performance comparisons. (**a**) Beamforming result based on the hypothetical array. (**b**) Estimation of array shape. (**c**) Comparison of the estimates of time-delay differences. (**d**) Time-delay differences in these detected narrow-band signals. (**e**) Spectrum of the improved output signal.

Figure 4b shows the beamforming result based on the hypothetical uniform linear array. The maximum value in the beam energy diagram is 55°, which is deviated from the true bearing of 60°, due to the time-delay mismatches in the distorted array. Only five narrow-band frequencies with 59 Hz, 97 Hz, 198 Hz, 163 Hz, and 232 Hz are acquired after the narrow-band detection based on the beamforming with the hypothetical array shape, and the narrow-band frequencies with 125 Hz and 280 Hz are missing. The corresponding time-delay differences of these detected narrow-band frequencies with 59 Hz, 97 Hz, 163 Hz, 198 Hz, and 232 Hz are shown in Figure 4d. It is observed that the time-delay differences are generally different, and show significant fluctuations particularly in the frequencies. Several outliers appear in the estimated time-delay differences. The time-delay differences should be independent of the considered narrow-band frequencies and be identical for the five narrow-band frequencies.

The ground truth of time-delay differences is shown in Figure 4c with the black line, while the average of these phase differences is shown in Figure 4c with the blue line. The estimation of time-delay differences in the subspace method is shown in Figure 4c with the green line. It is found that the estimates of time-delay differences in both the subspacebased method and the average operation method exhibit relatively large deviations from the ground truth and lead to quite similar estimates of the time-delay differences. The possible reason is that both the subspace-based method and the average operation approach are based on the narrow-band phase differences between adjacent elements, and both of them utilize the average operators for the estimates of time-delay differences in essence. The estimated time-delay differences in the proposed WORKS method, which significantly inhibits outliers, are quite closer to the true values, as shown in Figure 4c with the red line. According to the detected narrow-band frequencies, the array shape estimation is acquired by Equations (34) and (35), as shown in Figure 4a. The estimated array shape in the proposed method is closer to the real one, compared to those of the other two methods, even if gaps exist between them, especially in the latter part of the array. It is due to the error accumulations of time-delay differences in the array shape estimation, according to Equation (5).

The output spectrum based on the estimated array shape is provided in Figure 4e. It is reasonable that the output spectrum has higher gain, and is closer to the spectrum in the known array shape, compared to these in the subspace method and average operation method. It is also observed that the missing narrow-band frequency of 125 Hz becomes quite clear and can easily be detected in this enhanced spectrum. To quantify the superior performance, we use the amplitude error as the performance index. The amplitude error is defined as the absolute error between the amplitude of the detected narrow-band frequency and the amplitude in the corresponding frequencies in the spectrum based on the known array shape. The amplitude errors of the three methods are presented in Table 2. It is obvious that the proposed method is superior over the average operation and the subspacebased methods, owing to the ability to mitigate the effects of the outliers. It is also observed that the proposed method generally has the highest gain improvements among these three methods, and its amplitude errors in these detected narrow-band frequencies are generally less than 0.3 dB.

Table 2. Amplitude errors of detected narrow-band frequencies.

Frequency/Hz	59	97	125	163	198	232	280
Amplitude error in average method/dB	25.3	15.0	43.2	24.6	37.9	11.4	25.2
Amplitude error in subspace method/dB	24.9	15.7	49.7	24.8	32.8	11.3	22.6
Amplitude error in proposed method/dB	0.1	0.1	0.1	0.1	0.2	0.3	0.3

4.1.1. Variance Estimates of Time-Delay Difference

In this subsection, the variance estimates of time-delay differences are analyzed. All the parameters are the same as the simulations above. Figure 5 shows the variance results

of the estimated  $\Delta\tau$ . It is obvious that the proposed approach has the lowest estimated variance across 128 sensors, and its average variance is 0.023, which is quite closer to the CRBL of 0.006, compared to 0.842 and 0.821 in the average operation and the subspace-based methods, respectively. The main reason is that the proposed method is quite robust to these outliers in the estimated time-delay differences possibly caused by the low-SNR phases, whereas both average operation and the subspace-based method are sensitive to these outliers, and would lead to relatively poor performance.



Figure 5. Estimated time-delay difference variances.

4.1.2. Performance Comparison versus Bearings of the Radiated Noise Source

The effectiveness of the proposed method has been verified thanks to the aforementioned simulations. In the following simulation, the performance versus bearings of the noise source will be analyzed. The source bearings change from  $5^{\circ}$  to  $175^{\circ}$  by  $10^{\circ}$  steps. Other parameters are the same as in the previous simulations. It is observed that all three approaches seem robust to the bearings of targets and their average gain errors of narrowband frequencies in the end-fire bearing are comparable to those in the upright bearing. However, the average gain errors of results in the proposed method are much smaller than in the average operation and the subspace-based methods, as shown in Figure 6.



Figure 6. Amplitude error versus bearings of the radiated noise source.

4.1.3. Performance Comparison versus SINR

In this simulation, we focus on the performance comparison versus diverse SINRs. The narrow-band frequency interferences of 59 Hz, 97 Hz, and 163 Hz with different

amplitudes, which would cause diverse outliers in time-delay differences, are considered to form diverse SINRs from -3 dB to -14 dB with the step of -2 dB. Figure 7 shows average amplitude errors versus diverse SINRs. The proposed method has superior performance over the average operation and the subspace-based methods. It is observed that the average amplitude errors of the subspace-based method and the average operation approach gradually increase as the SINR decreases, whereas the proposed method is robust to different SINRs, which means the proposed method could mitigate the effects of different SINRs.



Figure 7. Average amplitude errors versus SINR in narrow-band signals.

#### 4.2. Experiments in Lake Trial Data

In this section, a set of lake trial data based on the distorted array is used to verify the effectiveness of the proposed method. In the experiment, the power amplifier as the acoustic source with 15 m in depth actively transmits the 83 Hz, 104 Hz, and 109 Hz low-frequency narrow-band signals and a chirp signal whose initial frequency is 90 Hz and bandwidth 8 Hz. The other ship equipped with a towing array of 64 hydrophone sensors with a uniform interval of 1.5 m, located 1 km away, is used to acquire the radiated noise data. There is a slight bow in the array. The data duration of *T* is 300 second, and the sample rate is 4 kHz.

The three narrow-band frequencies of 83 Hz, 104 Hz, and 109 Hz are acquired by using the narrow-band frequency detection. The corresponding time-delay differences are shown in Figure 8a. It is observed that some perturbations of time-delay differences exist among these detected frequencies, and they would result in large derivations from each other and lead to the appearances of outliers. The possible reason is that the SINRs in each detected narrow-band frequencies are quite different. The time-delay differences estimation in the average method, the subspace-based method, and the proposed WORKS method are shown in Figure 8b, respectively. The time-delay differences in the proposed method are quite smooth and slowly change, compared to thees in the average method and the subspace method. The output time-frequency spectrum based on the enhanced signals during 300 s are given in Figure 9a–d. It is found that the spectrum of enhanced signals based on the proposed method has higher amplitudes in both narrow-band signals and chirp signals, compared to these in the other two methods. The proposed method is superior to the average operation and the subspace-based methods. The quantitative performance concerning the average gains of three narrow-band frequencies are provided as shown in Figure 10. Note that the gains are calculated by the narrow-band frequency amplitude in the output signal over that in the conventional beamforming assuming a hypothetical uniform linear array due to the unknown distorted array shape. It is found the gains in the our proposed method are generally above 15 dB, even up to 30 dB, compared to about 5 dB gains in both the average operation and the subspace-based methods.



**Figure 8.** Performance comparisons. (a) Time-delay differences in 4 detected narrow-band signals. (b) Comparisons of the estimates of time-delay differences.



**Figure 9.** Performance comparisons of Short Time Fourier Transform using 20 s hamming window with a window shift of 10 s. (a) Proposed method. (b) Average method. (c) Subspace method. (d) CBF method.



Figure 10. Gains of the detected narrow-band frequencies.

#### 5. Conclusions

A robust and data-driven signal enhancement scheme is proposed to take into account the array distortion in this paper. The proposed scheme firstly extracts the phases of narrowband signals from the conventional beamforming result using narrow-band signal detection. Then, we proposed a novel weighted outlier-robust Kalman smoother to estimate timedelay differences to mitigate outliers in the time-delay differences. Finally, an enhanced output signal is achieved by beamforming based on the estimated time-delay differences. The proposed scheme is validated by both simulations and experiments. The proposed method is a data-driven approach that fully exploits the directional radiated noise signal due to distant underwater acoustic targets as sources of opportunity for real-time array shape estimation. It requires neither the number nor direction of sources to be known in advance.

**Author Contributions:** Conceptualization, Q.W. and H.Z.; methodology, Q.W., H.Z. and Z.L.; software, H.Z. and Y.X.; validation, H.Z., Z.L. and S.Y.; formal analysis, Q.W. and H.Z.; investigation, Q.W., Y.X. and S.Y. writing—original draft preparation, Q.W. and H.Z.; writing—Q.W. and J.T.; supervision, Q.W., Y.X., S.Y. and J.T.; funding acquisition, Q.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was supported in part National Natural Science Foundation under Grants No. 61701109 and 91938203, in part by Science and Technology on Sonar Laboratory under Grant No. 6142109180202, and in part by National Defense Basis Scientific Research program of China under Grant No. JCKY2019110C143, in part supported by National Natural Science Foundation of Jiangsu Province under Grant No BK20160701.

**Data Availability Statement:** The data presented in this paper are available after contacting the corresponding author.

**Acknowledgments:** We thank anonymous reviewers for their comments towards improving this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- 1. Schurman, I. Reverberation rejection with a dual-line towed array. IEEE J. Ocean. Eng. 1996, 21, 193–204. [CrossRef]
- Felisberto, P.; Jesus, S.M. Towed-array beamforming during ship's maneuvering. Proc. Inst. Elect. Eng. Radar Sonar Navig. 1996, 143, 210–215. [CrossRef]
- Park, H.Y.; Lee, C.; Kang, H.G.; Youn, D.H. Generalization of the subspace-based array shape estimations. *IEEE J. Ocean. Eng.* 2004, 29, 847–856. [CrossRef]
- 4. Hodgkiss, W. The effects of array shape perturbation on beamforming and passive ranging. *IEEE J. Ocean. Eng.* **1983**, *8*, 120–130. [CrossRef]

- Groen, J.; Beerens, S.P.; Doisy, Y. Shape and Doppler correction for a towed sonar array. In Proceedings of the Oceans'04 MTS/IEEE Techno-Ocean '04, Kobe, Japan, 9–12 November 2004; Volume 2, pp. 613–620.
- 6. Owsely, N.L. Shape estimation for a flexible underwater cable. *IEEE EASCON* **1981**, *14*, 16–19.
- Gray, D.; Anderson, B.; Bitmead, R. Towed array shape estimation using Kalman filters-theoretical models. *IEEE J. Ocean. Eng.* 1993, 18, 543–556. [CrossRef]
- 8. Newhall, B.; Jenkins, J.; Dietz, J. Improved estimation of the shape of towed sonar arrays. In Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Como, Italy, 18–20 May 2004; pp. 873–876.
- 9. Rockah, Y.; Schultheiss, P.M. Array shape calibration using sources in unknown locations—Part I: Far-field sources. *IEEE Trans. Acoust. Speech Signal Process.* **1987**, *35*, 286–299. [CrossRef]
- 10. Weiss, A.J.; Friedlander, B. Array shape calibration using sources in unknown locations-a maximum likelihood approach. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 1958–1966. [CrossRef]
- Gray, D.A.; Wolf, W.O.; Riley, J.L. An eigenvector method for estimating the positions of the elements of an array of receivers. In Proceedings of the Conference Signal Processing Applic, Adelaide, Australia, 17–19 April 1989; pp. 391–393.
- Smith, J.J.; Leung, Y.H.; Cantoni, A. Sptatistics of the phase delays between array receivers estimated from the eigendecomposition of the signal correlation matrix. In Proceedings of the Conference IEEE International Conference on Acoustics, Speech and Signal Processing, Adelaide, SA, Australia, 19–22 April 1994; pp. 325–329.
- 13. Bouvet, M. Beamforming of a distorted line array in the presence of uncertainties on the sensor positions. *J. Acoust. Soc. Am.* **1987**, *81*, 1833–1840. [CrossRef]
- Owsley, N.L.; Swope, G.R. Array shape determination using time delay estimation procedures. *IEEE EASCON* 1980, 13, 158–164.
   Odom, J.L.; Krolik, J.L. Passive towed array shape estimation using heading and acoustic data. *IEEE J. Ocean. Eng.* 2015, 40, 465–474. [CrossRef]
- 16. Ferguson, B.G. Sharpness applied to the adaptive beamforming of acoustic data from a towed array of unknown shape. *J. Acoust. Soc. Am.* **1990**, *88*, 2695–2701. [CrossRef]
- 17. Quinn, B.G.; Barrett, R.F.; Kootsookos, P.J.; Searle, S.J. The estimation of the shape of an array using a hidden Markov model. *IEEE J. Ocean. Eng.* **1993**, *18*, 557–564. [CrossRef]
- 18. Wahl, D.E. Towed array shape estimation using frequency wavenumber data. IEEE J. Ocean. Eng. 1993, 18, 582–590. [CrossRef]
- 19. Viberg, M.; Swindlehurst, A.L. A Bayesian approach to autocalibration for parametric array signal processing. *IEEE Trans. Signal Process.* **1994**, *42*, 3495–3507. [CrossRef]
- 20. Ross, D.; Kuperman, W.A. Mechanics of underwater noise. J. Acoust. Soc. Am. 1989, 86, 1626–1626. [CrossRef]
- 21. Wales, S.C.; Heitmeyer, R.M. An ensemble source spectra model for merchant ship-radiated noise. *J. Acoust. Soc. Am.* 2002, 111, 1211–1231. [CrossRef] [PubMed]
- 22. Information on the Swellex96 Experiment Data. Available online: http://swellex96.ucsd.edu/ (accessed on 15 July 2020).
- Wu, Q.; Xu, P.; Li, T.; Fang, S. Feature enhancement technique with distorted towed array in the underwater radiated noise. In Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference Proceedings. Institute of Noise Control Engineering, Hong Kong, China, 27–30 August 2017; Volume 255, pp. 3824–3830.
- 24. Ting, J.; Evangelos, T.; Stefan, S. A Kalman filter for robust outlier detection. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 29 October–2 November 2007; pp. 1514–1519.
- Ting, J.; Evangelos, T.; Stefan, S. Learning an outlier-robust Kalman filter. In Proceedings of the European Conference on Machine Learning, Warsaw, Poland, 17–21 September 2007; pp. 748–756.
- Dempster, A.; Laird, N.; Rubin, D. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B 1977, 39, 1–38.
- 27. Bishop, C.M. Pattern Recognition and Machine Learning; Springer: New York, NY, USA, 2006; pp. 423–517.
- Šimandl, M.; Královec, J.; Tichavsky, P. Filtering, predictive, and smoothing Cramer–Rao bounds for discrete-time nonlinear dynamic systems. *Automatica* 2001, 37, 1703–1716. [CrossRef]
- Tichavsky, P.; Muravchik, C.H.; Nehorai, A. Posterior Cramer–Rao bounds for discrete-time nonlinear filtering. *IEEE Trans. Signal Process.* 1998, 46, 1386–1396. [CrossRef]