Robust Estimation of Modulation Frequency in Impulsive Acoustic Data

HARI VISHNU, Member, IEEE
MANDAR CHITRE, Senior Member, IEEE
National University of Singapore, Singapore

Estimation of the modulation frequency of a source is important in various sonar signal processing applications. The conventional estimator of modulating frequency has been formulated under the assumption that the source and noise are Gaussian distributed. This is often inaccurate in several applications such as underwater acoustics where the signal or noise can be impulsive. We formulate novel modulation frequency estimators that are robust to impulses in the data, and outperform the conventional estimator in environments where the observed data are contaminated by impulses. We demonstrate their performance using recorded data. We characterize the performance of the methods in terms of their accuracy, harmonic distortion, and robustness to knowledge of noise statistics. We also derive the Cramer–Rao lower bound for the modulation frequency estimation problem in impulsive data.

Manuscript received November 3, 2015; revised May 26, 2016 and December 18, 2016; released for publication February 7, 2017. Date of publication March 2, 2017; date of current version August 7, 2017.

DOI. No. 10.1109/TAES.2017.2677621

Refereeing of this contribution was handled by K. Davidson.

Authors' address: H. Vishnu and M. Chitre are with the Acoustic Research Laboratory, National University of Singapore, Singapore 119222, E-mail: (harivishnu@gmail.com; mandar@arl.nus.edu.sg).

0018-9251 © 2017 IEEE

I. INTRODUCTION

Estimation of the modulating frequency of acoustic sources such as aircraft rotors and ship propellers is vital in sensing applications in passive sonar. This can be applied for identification of aircraft and sea vessels and characterization of their properties such as number of blades, shaft rotation frequency, and blade rate. This can also be used in detection of animals and divers [1]. Estimation of the modulation frequency is usually performed using a method known as "Detection of envelope modulation on noise" (DEMON) [2]. DEMON is a narrowband signal analysis algorithm based on the assumption that the sound generated by sources such as propellers can be described as the modulation of a carrier waveform by a modulating waveform. The carrier waveform is a random waveform representing broadband cavitation noise, and the deterministic modulating waveform represents the periodicity in the propeller rotation with a fundamental frequency f. For a helicopter rotor or ship propeller with b blades and rotor speed r, the fundamental modulating frequency f is equal to br.

A conventional DEMON estimator of f has been derived previously under the assumption that the cavitation noise and the ambient noise are independent and identically distributed (i.i.d.) random variables following a Gaussian probability density function (pdf) [3], [4]. Several modifications of the DEMON algorithm have been proposed, including ones which consider colored cavitation noise [5], [6], time-variation in the modulating frequency [7], [8], tracking of multiple sources in a decoupled way [9], estimation in a multipath environment [10] and use of a 3/2D-spectral analysis to extract propeller features from acoustic vector sensor data [11]. DEMON-based algorithms have been tested experimentally [8], [12], and performance bounds on estimation of the modulation parameters have been derived [13]. DEMON algorithms have also been used to detect the breathing pattern of divers from acoustic data [14], [15]. However, in all these works, the noise in the observed data is assumed to be Gaussian distributed.

In many applications such as underwater acoustics and room acoustics, the noise is impulsive and characterized by large outliers in the time-series [16]–[21]. The source is also impulsive in some cases such as helicopter sound signatures [22], [23]. In all these cases, the statistical properties of the signal or noise cause the observed data to be impulsive, and a Gaussian pdf is inappropriate to model such data. The performance of the conventional DEMON algorithm, which is formulated for the case of Gaussian-distributed data, degrades considerably in the presence of these impulses. In such a case, it is necessary to consider pdf models that can capture the impulsiveness of the data, such as the generalized Gaussian (GG) pdf [24].

We present two novel DEMON estimators of modulation frequency that are robust to the impulses in the data, and yield better performance in environments where the data contains impulses. These formulations are based on modeling the observed data using a GG pdf. These are generalized forms of the DEMON algorithm that are applicable for a wide range of environments with impulsive noise or signals, and include Gaussian-distributed data as a special case. We show that the performance of the robust DEMON estimators is superior to the conventional DEMON estimator, using simulated data as well as recorded data that are impulsive. We also discuss the problem of harmonic distortion of the estimators with theoretical expressions. Finally, we derive the Cramer–Rao lower bound (CRLB) for the estimation of broadband modulation parameters in impulsive data, and discuss the optimal performance achievable in such a scenario.

II. DATA MODEL

We consider the following data model. Let $x(t)$ denote the tth sample of the observed pressure time-series data. $x(t)$ contains a broadband additive noise $n(t)$, and a broadband signal $s(t)$ formed by the modulation of a carrier waveform $w(t)$ by a modulating waveform $m(f, t)$. These can be described by the equations

$$x(t) = s(t) + n(t) \tag{1}$$

where

$$s(t) = m(f, t)w(t). \tag{2}$$

The modulating waveform $m(f, t)$ is periodic with a frequency given by f. If $m(f, t)$ is periodic, $m^2(f, t)$ is also periodic, and hence, we can express it by a cosine series as

$$m^2(f, t) = \sum_{l=0}^{L} A_l \cos(lcft + l\theta) \tag{3}$$

where $c = \frac{2\pi}{f_s}$, f_s is the sampling frequency, A_l is the $l + 1$th coefficient in the cosine series expansion of $m^2(t)$, θ is the phase of the first fundamental frequency term in the expansion, and L is the number of coefficients considered. Since the left-hand side of (3) is always positive, the coefficients A_l are such that the right-hand side is also always positive. Generally this means that the coefficient A_0 is positive and larger than the successive coefficients A_l, $l > 0$. Before we begin the discussion on estimation, we must make the following assumptions about the signal in order to facilitate the derivation of the estimators:

A1) The carrier waveform and ambient noise are i.i.d. zero-mean sequences, independent of each other as well as the modulating waveform, with finite variances given by σ_w^2 and σ_n^2, respectively.

A2) The power of the signal is weak, i.e.,

$$\frac{m^2(f, t)\sigma_w^2}{\sigma_n^2} \ll 1. \tag{4}$$

This assumption is made in order to simplify the formulation while still ensuring that the estimator formulated is effective for weak signals, which is a more challenging scenario to deal with. If the signal is strong, we expect that the formulated estimator would still work well. This will also be verified in the results in Section V.

A3) N samples of data are observed, and the observation length is large compared to periodicity of the signal, i.e., $cfN \gg 1$.

The variance of the data $x(t)$ observed at the sample number t, is given by

$$\sigma_x^2(t) = m^2(f, t)\sigma_w^2 + \sigma_n^2. \tag{5}$$

The conventional DEMON estimator has been formulated by Nielsen [4] and Lourens and Du Preez [3]. To arrive at this estimator, apart from A1 to A3, they have used the additional assumption:

A4) The ambient noise and carrier waveform are both Gaussian distributed.

However, there are several cases where the ambient noise $n(t)$ or the signal $s(t)$ is impulsive in nature. The case of impulsive noise is particularly relevant since it arises in several applications such as underwater acoustics. In such cases, assumption A4 becomes inappropriate, and it is insufficient to model the pdf of the data as Gaussian. Pdf models that can capture the impulsiveness of the data are more suitable to model the data pdf for these cases. One such distribution is the GG distribution. The GG pdf is defined by an exponential parameter g, which characterizes the impulsiveness of the data (heaviness of tail of the pdf). The GG pdf of a zero-mean random variable $x(t)$ conditional on the frequency f with variance $\sigma_x^2(t)$ and exponent factor g is given by [24]

$$p(x(t)|f) = \frac{G}{\sigma_x(t)} \exp\left(-H \left|\frac{x(t)}{\sigma_x(t)}\right|^g\right) \tag{6}$$

where

$$G = \frac{g\,(\Gamma(3/g))^{0.5}}{2\,(\Gamma(1/g))^{1.5}} \tag{7}$$

and

$$H = \left(\frac{\Gamma(3/g)}{\Gamma(1/g)}\right)^{0.5g} \tag{8}$$

where $\Gamma(.)$ refers to the gamma function. The parameter g can take on values in the range $[0, \infty]$. When $g = 2$, the GG pdf reduces to a Gaussian pdf. When $g < 2$, the GG pdf is more impulsive than a Gaussian pdf. In order to derive an estimator of the modulation frequency robust to the impulsiveness of the data, we replace the assumption A4 used in [3] with the assumption A5:

A5) The data follow the GG pdf described in (6) and the value of g is known.

In practice, g may not be known, but it may be estimated from the data or from prior noise recordings if they are available. The choice of the parameter g is discussed at the end of Section VI.

III. ROBUST ESTIMATION OF MODULATION FREQUENCY

A. GG-DEMON Estimator

Considering that the data follow the GG pdf in (6), we can derive a nearly optimal estimator of the modulating frequency using a reasoning similar to that outlined in [3]. Modeling the observed data using a GG pdf allows us to obtain a generalized formulation of the estimator applicable to a wide range of scenarios with varying impulsiveness based on the value of g, and includes Gaussian-distributed data as a special case. We now proceed to derive this estimator, which we refer to as GG-DEMON. The log likelihood function $L(\mathbf{x}|f)$ of the data vector $\mathbf{x} = [x(1), x(2), \ldots, x(N)]^T$ conditional on the unknown modulating frequency f is given by

$$L(\mathbf{x}|f) = b_0 - 0.5\sum_{t=1}^{N} \log\left(\sigma_x^2(t)\right) - b_2 \sum_{t=1}^{N} |x(t)|^g \left(\frac{1}{\sigma_x^2(t)}\right)^{0.5g} \tag{9}$$

where "b_i"s denote constants independent of f, and $b_0 = N\log(G)$, $b_2 = H$. From (1) and (2), the term $\sum_{t=1}^{N} \log\left(\sigma_x^2(t)\right)$ can be written as

$$\sum_{t=1}^{N} \log\left(\sigma_x^2(t)\right) = \sum_{t=1}^{N} \log\left(\sigma_n^2\right) + \sum_{t=1}^{N} \log\left(1 + \frac{m^2(f,t)\sigma_w^2}{\sigma_n^2}\right)$$
$$\approx N\log\left(\sigma_n^2\right) + \sum_{t=1}^{N} \frac{m^2(f,t)\sigma_w^2}{\sigma_n^2} \approx N\log(\sigma_n^2) + NA_0\frac{\sigma_w^2}{\sigma_n^2}$$

using assumption A2 and A3. This term depends on the noise level and the dc component of the signal. But it is independent of the frequency since the dc level of a periodic signal does not change when the fundamental frequency of the signal changes [3]. Thus, the second term in (9) can be clubbed together with b_0 into a constant term b_1, and (9) can be written as

$$L(\mathbf{x}|f) = b_1 - b_2 \sum_{t=1}^{N} |x(t)|^g \left(\frac{1}{\sigma_x^2(t)}\right)^{0.5g}. \tag{10}$$

Invoking the weak signal assumption A2, we can write

$$\left(\frac{1}{\sigma_x^2(t)}\right)^{0.5g} \approx \frac{1}{\sigma_n^g}\left(1 - \frac{m^2(f,t)\sigma_w^2}{\sigma_n^2}\right)^{0.5g}$$
$$\approx \frac{1}{\sigma_n^g}\left(1 - \frac{g m^2(f,t)\sigma_w^2}{2\sigma_n^2}\right) \tag{11}$$

Substituting this in (10), we obtain

$$L(\mathbf{x}|f) \approx b_1 - b_3 \sum_{t=1}^{N} |x(t)|^g + b_4 \sum_{t=1}^{N} |x(t)|^g m^2(f,t). \tag{12}$$

It is clear that for the given data, the maximum likelihood estimate (MLE) of f, which maximizes $L(\mathbf{x}|f)$, is the one that maximizes the third term on the right of (12), since that is the only term dependent on f. Hence the estimator of f is

$$\hat{f} = \arg\max_f \left[\sum_{t=1}^{N} |x(t)|^g m^2(f,t)\right]. \tag{13}$$

Equation (13) gives us an intuition behind the estimator. The estimator consists of the correlation between the absolute values of the data vector samples raised to the exponent g, and the square of the modulated signal, which acts as a template. From (3),

$$\hat{f} = \arg\max_f \left[A_0 \sum_{t=1}^{N} |x(t)|^g + \sum_{t=1}^{N} |x(t)|^g \sum_{l=1}^{L} A_l \cos(lcft + l\theta)\right]. \tag{14}$$

Since the first term within brackets is independent of f, only the second term needs to be considered, leading to

$$\hat{f} = \arg\max_f \left[\sum_{t=1}^{N} |x(t)|^g \times \sum_{l=1}^{L} \left(A_l \cos(lcft)\cos(l\theta) - A_l \sin(lcft)\sin(l\theta)\right)\right]$$
$$= \arg\max_f \left[\sum_{t=1}^{N} |x(t)|^g \times \sum_{l=1}^{L} \left(B_l \cos(lcft) - C_l \sin(lcft)\right)\right] \tag{15}$$

where $B_l = A_l \cos(l\theta)$ and $C_l = A_l \sin(l\theta)$. If the terms A_l and θ are known, the above equation may be used for estimation. In most cases, however, these quantities are unknown; and hence, we would need to replace them with their MLEs in order to obtain an estimate of the template signal corresponding to the modulated waveform. Obtaining the MLEs of these terms would involve a multidimensional optimization, which is quite computationally complex to solve for. However, in place of the MLEs, we can use other unbiased estimators using an approach similar to [3], which makes the estimation simpler. We follow this approach to obtain estimators for the unknown coefficients.

For a positive integer value k, it can be shown that

$$E\left[\frac{1}{N}\sum_{t=1}^{N} x^2(t)\cos(kcft)\right]$$
$$= \frac{\sigma_n^2}{N}\sum_{t=1}^{N} \cos(kcft) + \frac{\sigma_w^2}{N}\sum_{t=1}^{N} m^2(t)\cos(kcft)$$
$$\approx 0.5\sigma_w^2 A_k \cos(k\theta) \approx 0.5\sigma_w^2 B_k, \tag{16}$$

by invoking assumption A3, where $E[.]$ represents the expectation operator. Similarly, it can be shown that

$$E\left[\frac{1}{N}\sum_{t=1}^{N} x^2(t)\sin(kcft)\right] \approx -0.5\sigma_w^2 C_k. \tag{17}$$

From (16) and (17), it is clear that we can use the unbiased estimators

$$\hat{B}_k = \frac{2}{N\sigma_w^2}\sum_{t=1}^{N} x^2(t)\cos(kcft),$$
$$\hat{C}_k = -\frac{2}{N\sigma_w^2}\sum_{t=1}^{N} x^2(t)\sin(kcft), \tag{18}$$

where $\hat{d}$ represents the estimate of some parameter d. Substituting (18) in (15) yields the expression for the DEMON-based estimator as

$$\hat{f} = \arg\max_f \sum_{l=1}^{L} \Re\left(\sum_{t=1}^{N} |x(t)|^g \exp(ilcft) \times \sum_{t=1}^{N} x^2(t)\exp(-ilcft)\right) \tag{19}$$

where $i = \sqrt{-1}$ and $\Re(z)$ denotes the real part of a complex number z. Note that when $g = 2$, the estimator in (19) reduces to the conventional estimator presented in [3]. The first and second summation terms within the brackets in (19) are equivalent to the conjugate fast-Fourier-transform (FFT) of the series $\left[|x(1)|^g, |x(2)|^g, \ldots, |x(N)|^g\right]^T$ and FFT of the series $\left[x^2(1), x^2(2), \ldots, x^2(N)\right]^T$, respectively.

In practice, the summation of L harmonic terms is avoided for the sake of simplicity. This also reduces errors introduced due to limited frequency resolution in the frequency search space considered, which increases at higher frequencies. Instead, the simplified estimator given below, which considers a single fundamental frequency term, is used:

$$\hat{f}_{GG} = \arg\max_f [P_{\mathrm{GG}}(f)] \tag{20}$$

where

$$P_{\mathrm{GG}}(f) = \Re\left(\sum_{t=1}^{N} |x(t)|^g \exp(icft) \times \sum_{t=1}^{N} x^2(t)\exp(-icft)\right). \tag{21}$$

The estimator in (20) is referred to as the GG-DEMON estimator. The output spectrum of the GG-DEMON estimator may contain multiple peaks corresponding to harmonics, in which case the location of the peak occurring first is taken to be the estimate of the modulating frequency.

In practical scenarios, the DEMON algorithm used may involve a few more steps. A bandpass-filtered version of the data vector **x** is often used in place of **x** [2]. This is because in practical scenarios, the signal and noise power are not uniformly distributed across all frequencies in the received data. The signal-to-noise-ratio (SNR) of the modulated signal is higher for certain frequency bands than others. Hence, the cutoff frequencies of the bandpass filter are selected in such a way that the SNR of the modulated signal is maximized. Another often-used processing step is that the received data are subsampled by a decimation factor [2]. This is because the modulation frequency to be estimated is often much lower than the sampling frequency. Hence, the search for the modulating frequency needs to be performed over a much smaller range of search frequencies. The decimation factor is chosen based on the expected size of the search range.

B. Harmonic Distortion and Modified GG DEMON

One of the drawbacks of the GG-DEMON estimator is that its output spectrum contains additional harmonics of the frequencies in the signal. This leads to presence of strong harmonic peaks in the GG-DEMON output. In the presence of noise, this may sometimes lead to the final estimate of the algorithm being a harmonic peak instead of the fundamental frequency.

We discuss the harmonic distortion of the GG-DEMON estimator by considering its response to a signal $u(t)$ with a single modulating frequency f_m, given as

$$u(t) = (1 + M\cos(cf_m t + \theta)) .w(t) \tag{22}$$

where M is a modulation index, $0 < M < 1$. With this signal, the response of the GG-DEMON estimator is given as

$$\begin{aligned} P_{\mathrm{GG}}(f) = E\Bigg[&\sum_{t=1}^{N} |w(t)|^g (1 + M\cos(cf_m t + \theta))^g \cos(cft) \\ &\times \sum_{t=1}^{N} w^2(t)(1 + M\cos(cf_m t + \theta))^2 \cos(cft) \\ &+ \sum_{t=1}^{N} |w(t)|^g (1 + M\cos(cf_m t + \theta))^g \sin(cft) \\ &\times \sum_{t=1}^{N} w^2(t)(1 + M\cos(cf_m t + \theta))^2 \sin(cft) \Bigg]. \end{aligned} \tag{23}$$

Now, we use the relation that $E[y_1 y_2] = E[y_1]E[y_2]$ for two independent random variables y_1 and y_2 to reduce (23) further. The product terms in (23) are not really independent and hence the final outcome of this analysis may have some error associated with it. However, this error is shown to be very low in our validation of these expressions using simulations at the end of this section. Additionally, we make the assumption

A6) $M\cos(cf_m t + \theta) \ll 1$.

Assumption A6 indicates that the results derived here may only be used for small values of the modulation index M. However, in practice, the error is shown to be very small in the comparison of the theoretical results against the simulation results shown at the end of this section, even for a value of M as large as 0.9.

Using the abovementioned relation and assumption A6, the expression for $P_{\mathrm{GG}}(f)$ in (23) can be reduced to

$$P_{\mathrm{GG}}(f) = k \Bigg(\sum_{t=1}^{N} [1 + gM \cos(cf_m t + \theta) + \frac{g(g-1)}{2} M^2 \cos^2(cf_m t + \theta)] \cos(cft) \times \sum_{t=1}^{N} [1 + 2M \cos(cf_m t + \theta) + M^2 \cos^2(cf_m t + \theta)] \cos(cft) + \sum_{t=1}^{N} [1 + gM \cos(cf_m t + \theta) + \frac{g(g-1)}{2} M^2 \cos^2(cf_m t + \theta)] \sin(cft) \times \sum_{t=1}^{N} [1 + 2M \cos(cf_m t + \theta) + M^2 \cos^2(cf_m t + \theta)] \sin(cft) \Bigg).$$

In the above equation, A6 was used to retain just three terms in the expansion of $(1 + M\cos(cf_m t + \theta))^g$ and neglect the rest, and

$$k = \frac{\sigma_w^{(g+2)}}{g} \left(\frac{\Gamma(\frac{1}{g})}{\Gamma(\frac{3}{g})} \right)^{0.5g}. \tag{24}$$

The output amplitude of the GG-DEMON estimator at the fundamental frequency $f = f_m$ is found to be

$$P_{\mathrm{GG}}(f_m) = 0.5gkM^2N^2. \tag{25}$$

Equation (25) was obtained by invoking assumption A3 in order to employ the approximation

$$\sum_{t=1}^{N} \cos^2(cf_m t + \theta) \approx 0.5N, \quad \sum_{t=1}^{N} \sin^2(cf_m t + \theta) \approx 0.5N.$$

The output of the GG-DEMON estimator at the second harmonic frequency, $f = 2f_m$ is

$$P_{\mathrm{GG}}(2f_m) = \frac{g(g-1)kM^4N^2}{32}. \tag{26}$$

From (25) and (26), the harmonic ratio R_{GG} for GG-DEMON, i.e., ratio of second harmonic to first harmonic is found to be

$$R_{\mathrm{GG}} = \frac{P_{\mathrm{GG}}(2f_m)}{P_{\mathrm{GG}}(f_m)} = \frac{M^2(g-1)}{16}. \tag{27}$$

Note that at $g = 1$, the amplitude of the harmonic is expected to be zero, and for $g < 1$, the second harmonic has negative values.

It is possible to reduce the harmonic distortion by using a modified version of the GG-DEMON estimator. We refer to this as the modified GG-DEMON estimator (MGG-DEMON), and it has the expression given by

$$\hat{f}_{\mathrm{MGG}} = \arg\max_f [P_{\mathrm{MGG}}(f)] \tag{28}$$

where

$$P_{\mathrm{MGG}}(f) = \Re\left(\sum_{t=1}^{N} |x(t)|^g \exp(icft) \times \sum_{t=1}^{N} |x(t)| \exp(-icft) \right). \tag{29}$$

Using a similar approach as done for the case of the GG-DEMON estimator, the response of the MGG estimator to $u(t)$ can be found to be

$$P_{\mathrm{MGG}}(f) = E\Bigg[\sum_{t=1}^{N} |w(t)|^g (1 + M\cos(cf_m t + \theta))^g \cos(cft) \times \sum_{t=1}^{N} |w(t)|(1 + M\cos(cf_m t + \theta)) \cos(cft) + \sum_{t=1}^{N} |w(t)|^g (1 + M\cos(cf_m t + \theta))^g \sin(cft) \times \sum_{t=1}^{N} |w(t)|(1 + M\cos(cf_m t + \theta)) \sin(cft) \Bigg].$$

The response of the MGG estimator at the second harmonic frequency is

$$P_{\mathrm{MGG}}(2f_m) = 0.$$

Thus the harmonic ratio R_{MGG} of the MGG estimator is zero. In fact, all harmonics in the output of the MGG estimator are zero. Hence, we see that when the data contains the single-frequency signal given by $u(t)$, using $\hat{f}_{\mathrm{MGG}}$ ensures that no additional harmonic peaks are introduced in the output.

We now verify these results through simulations. Consider a simulation scenario with a modulated source following the signal model in (22). The modulated signal consists of an i.i.d. Gaussian distributed carrier signal modulated by a single-frequency modulating waveform. We consider the source has a modulating frequency of $f_m = 50$ Hz and modulation index $M = 0.9$. The sampling rate is $f_s = 16$ kHz. Four seconds of the data are used in the estimation. We employ a bandpass filter with passband 1–4 kHz to filter the data prior to estimation.

In Fig. 1, we plot the output power spectral density (PSD) obtained using (a) conventional DEMON, (b) GG-DEMON with $g = 1.2$, and (c) MGG-DEMON with $g = 1.2$, using the simulation parameters mentioned above. The plots Fig. 1(b) and (c) are scaled linearly in such a way that their peak level matches that of conventional DEMON [see Fig. 1(a)], for easier comparison. The absence of the second harmonic in the output of MGG-DEMON can be clearly observed. Thus, the simulation demonstrates that when the modulated signal does not contain any harmonics, the MGG-DEMON output also exhibits no harmonics. This is also what our theoretical analysis predicts. It is also seen that the harmonic ratio of the output of GG-DEMON is lower than that for conventional DEMON by a factor $(g - 1)$ which is about 7 dB, as predicted by (27).

In Table I, we compare the harmonic ratio as found through simulations, against the predicted values of

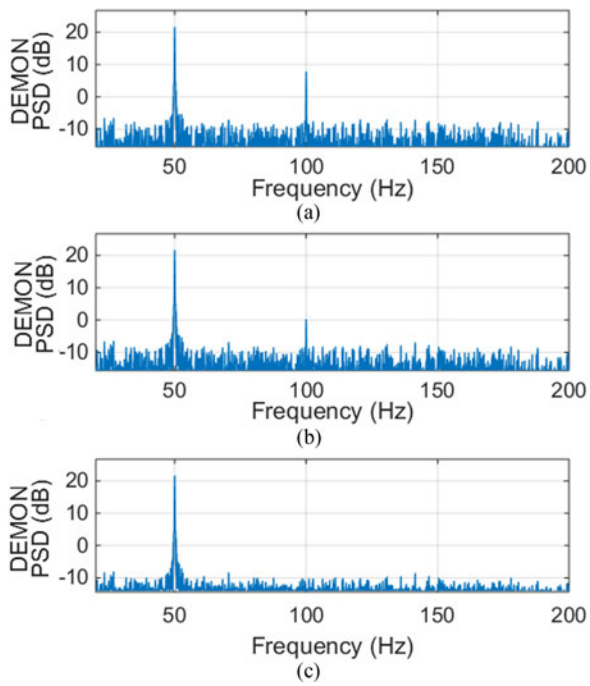

Fig. 1. Output PSD versus frequency for (a) conventional DEMON, (b) GG-DEMON ($g = 1.2$), and (c) MGG-DEMON ($g = 1.2$).

TABLE I
Comparison of Harmonic Ratios Calculated Through Simulation, Against Theoretically Predicted Values

Method	Harmonic ratio from simulations	Harmonic ratio predicted from theoretical analysis
Conventional DEMON (GG DEMON with $g = 2$)	0.049	0.05
Modified GG DEMON with $g = 2$	6×10^{-5}	0
GG-DEMON with $g = 1$	6×10^{-5}	0
GG-DEMON with $g = 1.5$	0.0253	0.0253
GG-DEMON with $g = 0.5$	-0.026	-0.0253

harmonic ratios for GG-DEMON and MGG-DEMON methods deployed with some values of the parameter g (which are not necessarily equal to the exponent parameter of the data pdf). The methods are assumed to use the same simulation parameters as that used for Fig. 1. Table I shows that there is a good match between simulation results and the theoretically predicted harmonic ratios presented in this section, thus validating the theoretical expressions.

IV. CRAMER–RAO LOWER BOUND

The CRLB establishes a lower bound on the variance of the estimates of unknown parameters. The CRLB for estimation of the modulation parameters was derived in [13] under the assumption that the data are Gaussian distributed. In this section, we discuss the CRLB for modulation parameters when the data are considered to be GG-distributed.

As in [13], we consider that the modulating source contains a single frequency component as described in (22). This model is more restrictive as compared to (3), but it allows us to simplify the derivation and obtain good insight into the problem of estimation of modulating frequency. Our intent is to find the CRLB of the unknown source parameters, which are the carrier power σ_w^2, modulating index M, modulating frequency f_m, and phase of modulating waveform θ.

In order to obtain the CRLB, we first compute the Fisher information matrix (FIM). The derivation of the expressions for the elements of the FIM is given in the Appendix.

It can be seen that the simplified expressions (45)–(63) for the elements of the FIM differ from those given for the case of a Gaussian pdf in [13] by only the term ($0.5g$). Hence, expressions (45)–(63) reduce to those given in [13] when $g = 2$. The block diagonal structure of the FIM allows us to easily obtain the CRLB of the parameters as

$$\begin{aligned}\mathrm{CRLB}(\sigma_w^2) &= \frac{J_{M,M}}{J_{\sigma,\sigma}J_{M,M} - J_{\sigma,M}^2}\\ &= \frac{2}{g}\cdot\frac{2\sigma_n^4(1+0.75M^2)}{N(1-0.75M^2+0.375M^4)}\end{aligned} \tag{30}$$

$$\begin{aligned}\mathrm{CRLB}(M) &= \frac{J_{\sigma,\sigma}}{J_{M,M}J_{\sigma,\sigma} - J_{\sigma,M}^2}\\ &= \frac{2}{g}\cdot\frac{\sigma_n^4(1+3M^2+0.375M^4)}{N\sigma_w^4(1-0.75M^2+0.375M^4)}\end{aligned} \tag{31}$$

$$\begin{aligned}\mathrm{CRLB}(\theta) &= \frac{J_{f,f}}{J_{f,f}J_{\theta,\theta} - J_{\theta,f}^2}\\ &= \frac{2}{g}\cdot\frac{4\sigma_n^4}{N\sigma_w^4M^2\left(1+0.25M^2\right)}\end{aligned} \tag{32}$$

and

$$\begin{aligned}\mathrm{CRLB}(f_m) &= \frac{J_{\theta,\theta}}{N^2(J_{\nu,\nu}J_{\theta,\theta} - J_{\theta,\nu}^2)}\\ &= \frac{2}{g}\cdot\frac{12\sigma_n^4}{N^3\sigma_w^4M^2(1+0.25M^2)\left(\frac{2\pi}{f_s}\right)^2}.\end{aligned} \tag{33}$$

As observed in [13], the CRLB of M, θ, and f_m are inversely proportional to the square of the broadband SNR which is given by $\frac{\sigma_w^2}{\sigma_n^2}$. However, the CRLB of σ_w^2 is only proportional to square of the broadband noise power. The CRLBs are also inversely proportional to the number of samples. Additionally, we observe from (30)–(33) that the CRLBs of all the parameters vary inversely with the parameter g. This means that as the data become more impulsive, the estimation of the modulation frequency becomes more challenging.

We consider an example scenario to study the variation of the CRLB. Consider a single-frequency modulated source with modulation index $M = 0.9$. The data SNR is $\frac{\sigma_w^2}{\sigma_n^2} = 0.01$ (i.e., -20 dB), and the sampling rate is $f_s =$ 4 kHz. A data length of three seconds is used in the estimation. In Fig. 2, we plot the variation of the CRLB of the parameter f_m with variation in the parameter g of the GG distributed data, for the simulation parameters considered in this example. The inverse variation of the CRLB with the parameter g can be observed in the figure.

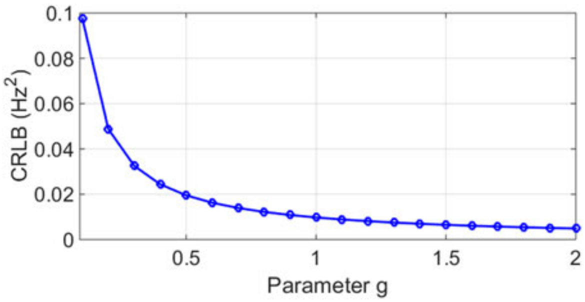

Fig. 2. Variation of CRLB of modulation frequency with parameter g of GG pdf. The SNR is -20 dB, and modulation index is 0.9.

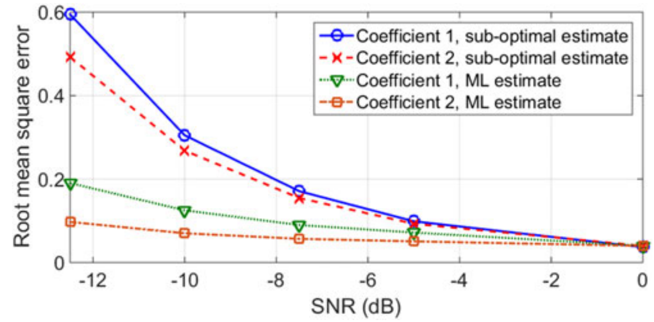

Fig. 4. RMSE of estimation of Fourier coefficients versus SNR using simulated data with MLE and simple suboptimal estimator in (18).

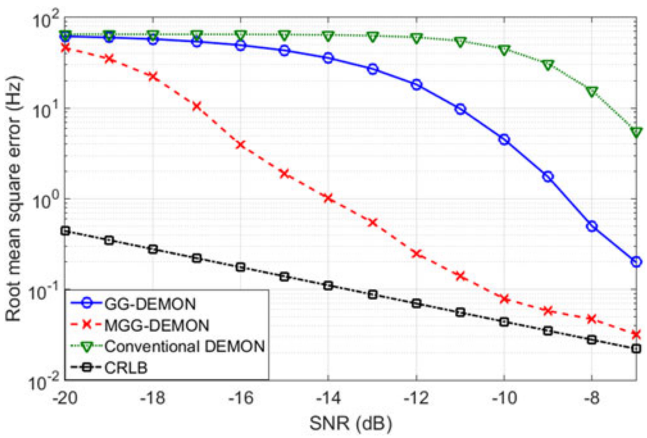

Fig. 3. RMSE of frequency estimation versus SNR using simulated data.

V. RESULTS WITH SIMULATED DATA

In this section, we study the performance of the DEMON algorithms using simulated data.

We first compare the performance of the DEMON algorithms. We consider a simulation scenario with a source whose SNR varies from -20 dB to -6 dB. The source has a modulating frequency of $f_m = 50$ Hz, a modulation index of $M = 0.9$, and no harmonic terms in the modulating waveform. The noise is GG-distributed with $g = 0.5$. The sampling frequency for the simulation is 8 kHz, and two seconds of the data are used in estimation of the modulation frequency. In Fig. 3, we compare the performance of the GG-DEMON and MGG-DEMON estimators, both using a value of $g = 0.5$, against the conventional DEMON estimator. The performance is compared in terms of the root-mean-squared-error (RMSE) of estimation of modulating frequency using these methods, computed by using 30 000 trials. The CRLB for the given simulation parameters is also plotted using (33).

Fig. 3 shows that the MGG-DEMON and GG-DEMON methods outperform the conventional DEMON method for the scenario considered. The RMSE of MGG-DEMON exceeds 1 Hz when the SNR of the data falls lower than -14 dB, whereas GG-DEMON exceeds 1 Hz of RMSE at an SNR of -8.5 dB. The suboptimality of MGG-DEMON and GG-DEMON in comparison to the CRLB, arises from their suboptimality in estimating the Fourier cosine series coefficients of the modulating waveform envelope. This leads to a performance drop at low SNR. At high SNR, the performance of all the methods approaches the CRLB.

During our discussion in Section III-A, we noted that the robust estimators are essentially a correlation of the data vector (raised to an exponent g) with the estimated envelope of modulated signal. When there is an error in the estimation of the cosine series coefficients, the estimate of the envelope of the modulated signal suffers a degradation too. Since the derived GG-DEMON and MGG-DEMON estimators use simple but suboptimal estimates of these coefficients to simplify the estimation, this leads to suboptimality in their performance compared to the MLE. To study this, we consider a simulation scenario with a modulated source whose SNR varies from -12 dB to 0 dB. The source has a single-frequency modulating waveform as given in (22) with $M = 0.9$ and $f_m = 100$ Hz, and the noise is GG-distributed with $g = 0.5$. The sampling frequency is 8 kHz, and two seconds of the data are used in estimation of the Fourier series coefficients. Since the GG-DEMON estimator uses the square of the modulated waveform as the template for correlation, its Fourier cosine series in (19) consists of two terms in the expansion. We compare the performance of estimation of the cosine series coefficients using maximum likelihood estimation based on multidimensional optimization, and the simple estimator described in (18), which is used by the robust estimators.

In Fig. 4, we plot the RMSE of estimation of the two Fourier cosine coefficients, computed using 5000 trials. Fig. 4 demonstrates the additional error in estimation of the Fourier coefficients incurred by the suboptimal methods. The difference in performance between the suboptimal estimator and the MLE, increases with decrease in SNR. The estimation error rises quickly when the SNR falls below -7.5 dB. This could be the reason for the large increase in error of the conventional, GG-DEMON, and MGG-DEMON estimators below this value of SNR, as seen in Fig. 3.

VI. RESULTS WITH RECORDED DATA

In this section, we compare the effectiveness of the DEMON methods using recorded data. We use three sets of

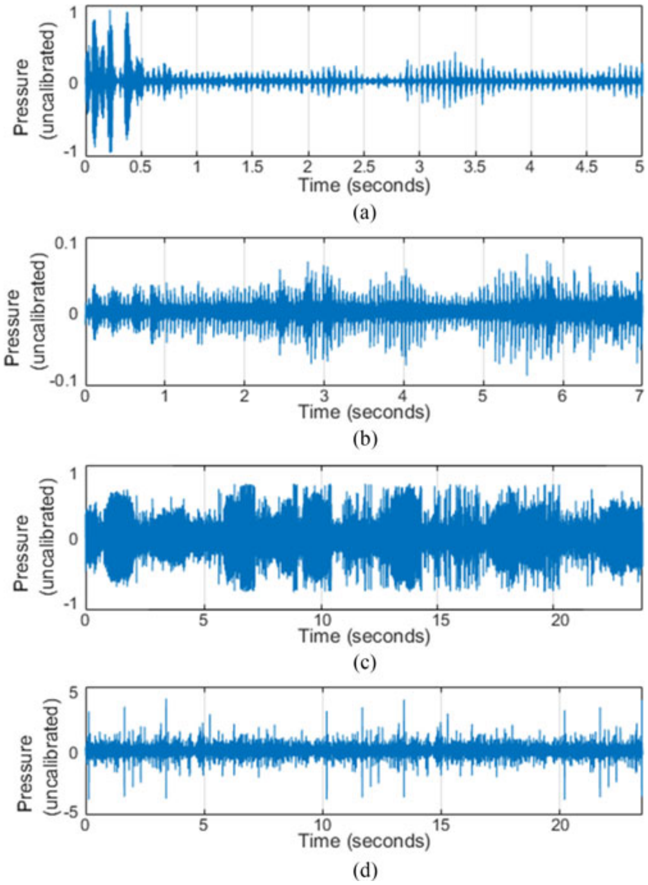

Fig. 5. (a) Time-series of dataset #1, (b) time-series of dataset #2, (c) spatially filtered time-series showing divers breathing signature in ROMANIS dataset, (d) time-series of dataset #3.

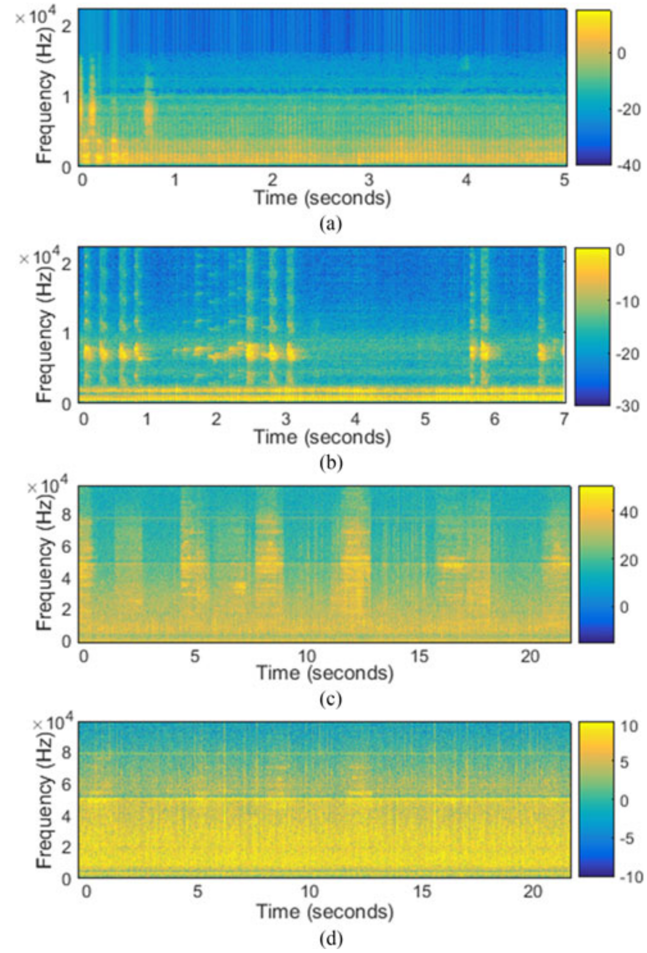

Fig. 6. Spectrograms of (a) dataset #1, (b) dataset #2, (c) spatially filtered time-series showing divers breathing signature in ROMANIS dataset, (d) dataset #3.

recorded data referred to as datasets #1, #2, and #3, respectively. Datasets #1 and #2 are of lengths five and seven seconds, respectively, and contain microphone recordings of a helicopter recorded at a sampling rate $f_s = 44.1$ kHz. In these datasets, the helicopter rotor rotation produced a modulated sound waveform with a modulating frequency of around 20.35 Hz. The time-series in these datasets was contaminated by noise sources such as gusts of wind and other external sources from the experiment area, and the time-signatures of the helicopters are impulsive as well.

Dataset #3 consists of data collected from Singapore waters using a planar hydrophone array called remotely operated mobile ambient noise imaging system (ROMANIS) developed at National University of Singapore [20]. ROMANIS has an operating frequency range of 25–75 kHz and sampling rate of 196 kSa/s. In 2010, the array was deployed near Selat Pauh Island, Singapore, for experiments. The recorded time-series was contaminated by snaps from snapping shrimp, which often dominate the high-frequency noise in this region. During the experiment, breathing sounds from two open-circuit divers in the water were recorded on the sensor array. The breathing signature is observed as a broadband sound modulated at the human breathing rate that lies in the range 0.15–0.4 Hz [15]. Since the recordings are of high SNR, we combine it with a recording of ambient noise collected on ROMANIS on the same day, to obtain a dataset with a broadband SNR of around -8 dB.

In Fig. 5, we show plots of the time-series of (a) dataset #1, (b) dataset #2, (c) 23 s of data obtained by spatially filtering recordings from 498 sensors of ROMANIS by steering it in the direction of the divers, showing the divers' breathing signature, and (d) dataset #3 consisting of 23 s of ROMANIS data from one sensor. In Fig. 6, we plot the spectrograms corresponding to the plots in Fig. 5. The spectrograms in Fig. 6(a) and (b) are computed using 2048 FFT points, and that in (c) and (d) using 8192 FFT points, using an overlap of 80% between successive windows.

In Figs. 5 and 6, subplots (a) and (b) show that the helicopters' "whopping" sound signatures can be observed as a series of impulses that manifests as broadband lines in the spectrogram spanning a large frequency range. The impulsiveness of the data arises because of the helicopter's signal as well as the ambient noise.

In Figs. 5(c) and 6(c), the signature of the two divers' exhalation can be observed as occasional bursts of energy that sustain for 1–1.3 s. This can be seen in the time-series and as a broadband signature in the spectrogram. The divers' breathing signature is most evident against the background

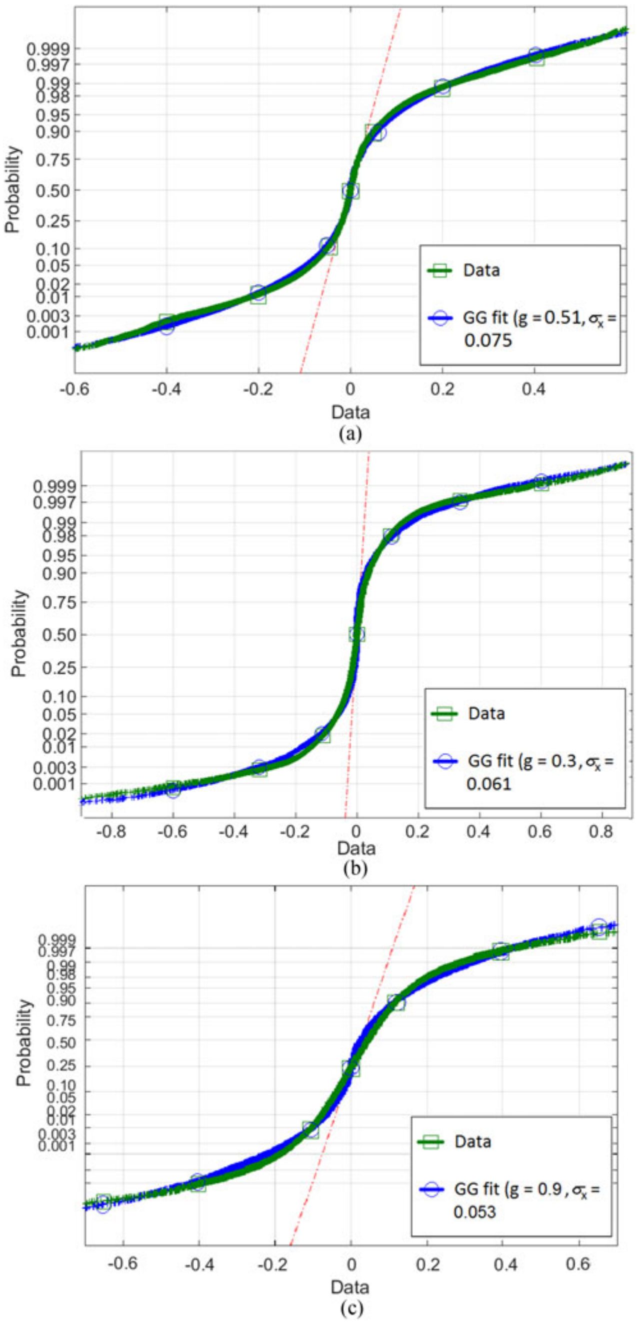

Fig. 7. Normal probability plot along with GG pdf fits for datasets (a) #1, (b) #2, and (c) #3.

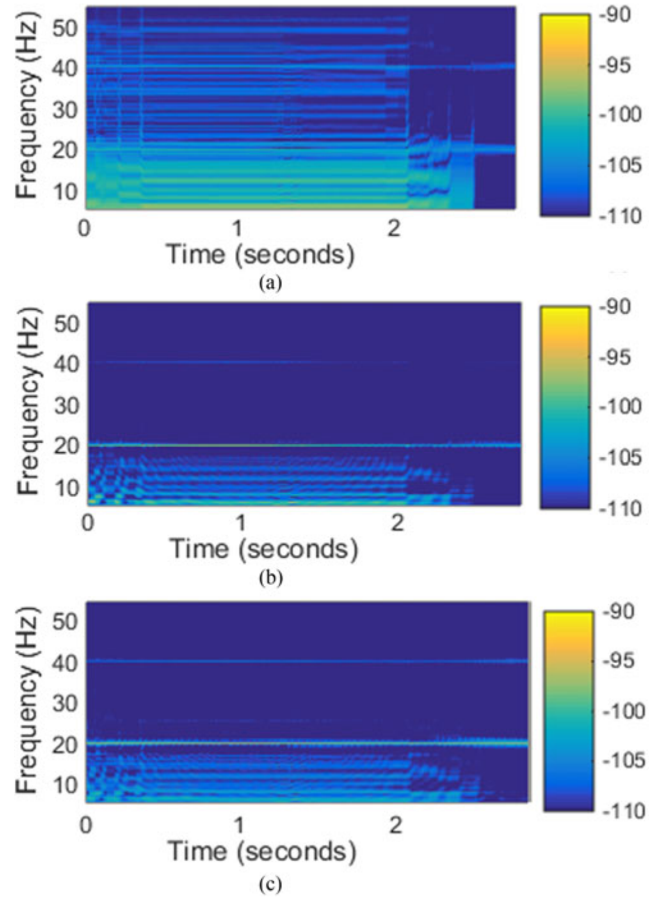

Fig. 8. DEMONgram output (in dB) with dataset #1, using (a) conventional DEMON, (b) GG-DEMON ($g = 0.18$), and (c) MGG-DEMON ($g = 0.2$).

noise in the high-frequency range, agreeing with what has been reported by other authors [15]. The 25–75-kHz frequency band will be utilized for the detection of the divers' breathing signature. In Fig. 5(d), the spikes arising due to snapping shrimp noise are evident. These reduce the SNR and hence in Fig. 6(d) the divers' acoustic signature can be barely distinguished against the background noise.

In order to study the impulsiveness of the data in these recordings, we plot the normal probability plots of these datasets and also fit a GG pdf to them. Fig. 7 shows normal probability plots comparing the distribution of the datasets (a) #1, (b) #2, and (c) #3 with their corresponding GG pdf fits. The red dashed line is a line joining the first and third quartiles of the data, and helps assess the Gaussianity of the data. If the distribution fit is a straight line, the data can be said to be Gaussian distributed. It can be seen that all three datasets have heavy tails in their distribution showing the presence of impulses in the data. This is not well-modeled by a Gaussian pdf, but the GG pdf is able to capture this property of the datasets well.

In sonar applications, sonar operators often monitor sources by observing the DEMONgram, which is a plot showing the evolution of the DEMON spectrum with time. Presence of a source would be detected in the form of a track at the modulating frequency. A DEMONgram is computed using data windows from the data time-series, which may contain overlapping samples. The overlap factor η is defined as the ratio of number of overlapping samples in each data window to the total number of samples, and lies between 0 and 1. The time axis of the DEMONgram indicates the start-time of the data window used to compute the DEMON output at that particular time-instant. We compare the performance of conventional DEMON ($g = 2$) against that of GG-DEMON and MGG-DEMON by plotting their DEMONgrams. For datasets #1 and #2, we use a bandpass filter with passband 500–3000 Hz to filter the data windows prior to estimation, and an overlap factor of $\eta = 0.999$. We decimate the data by a factor of 400 be-

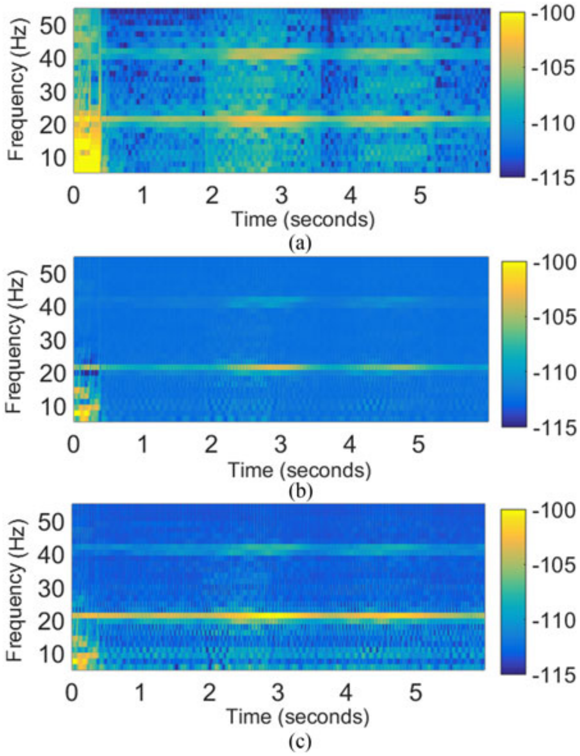

Fig. 9. DEMONgram output (in dB) with dataset #2, using (a) conventional DEMON, (b) GG-DEMON ($g = 0.2$), and (c) MGG-DEMON ($g = 0.2$).

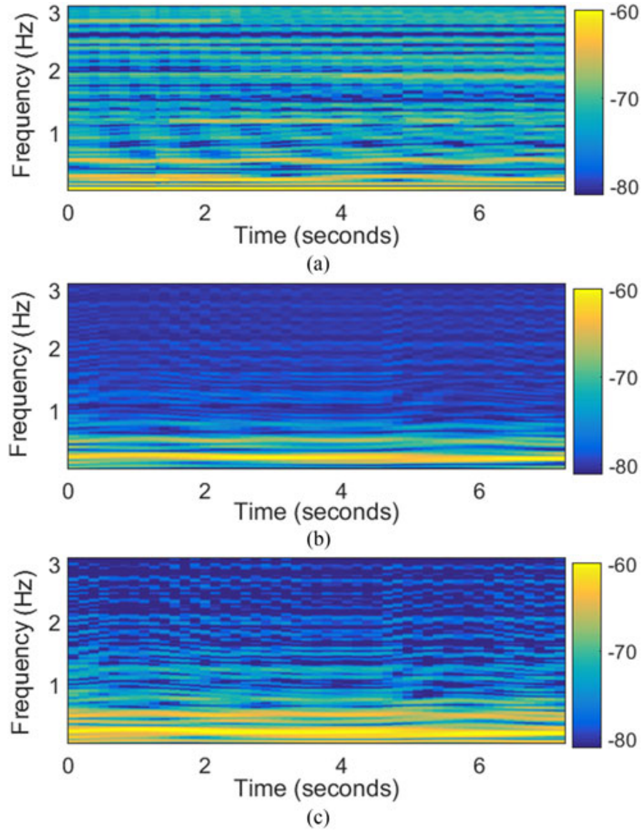

Fig. 10. DEMONgram output (in dB) with dataset #3, using (a) conventional DEMON, (b) GG-DEMON ($g = 0.7$), and (c) MGG-DEMON ($g = 0.93$).

fore applying a 1024-point FFT for frequency analysis. For dataset #3, we use $\eta = 0.99$, decimation factor 50 000 and a 1024-point FFT.

Fig. 8 shows a comparison of the DEMONgrams of dataset #1, computed using (a) conventional, (b) GG ($g = 0.18$), and (c) MGG ($g = 0.2$) DEMON. Later, we briefly discuss the sensitivity of the performance of these methods to the selection of the parameter g. A window size of 2 s is used for generating the plots in Fig. 8. Each plot shows the modulating frequencies on the y-axis and the starting time of each data window on the x-axis. We see that the conventional method in Fig. 8(a) is unable to track the source satisfactorily, whereas the GG-DEMON and MGG-DEMON methods are able to track the source better. The outputs of the GG-DEMON and MGG-DEMON yield clearly visible source tracks as compared to the conventional method.

Fig. 9 shows a comparison of the DEMONgrams of dataset #2, computed using (a) conventional, (b) GG ($g = 0.2$), and (c) MGG ($g = 0.2$) DEMON methods, using a window size of 1 s. For this dataset, the noise level was 1.8 dB lower than dataset #1 and we observe that the conventional DEMON method is able to track the source, in contrast to Fig. 8(a). However, its output is quite noisy. The GG-DEMON and MGG-DEMON methods, on the other hand, yield cleaner outputs than the conventional DEMON method and clearly show a track corresponding to the source frequency.

Fig. 10 shows a comparison of the DEMONgrams of dataset #3 computed using (a) conventional, (b) GG ($g = 0.7$), and (c) MGG ($g = 0.93$) DEMON. Since the breathing rate is very small, a large window size of 15 s is used to detect the sound. Conventional DEMON is able to track the divers' breathing signature at 0.25 Hz successfully in Fig. 10, but has a noisy output. The GG-DEMON and MGG-DEMON methods, on the other hand, yield cleaner outputs than conventional DEMON and allow us to distinguish the divers' breathing against the background noise in the DEMONgram.

From our discussion in Section III-B, we know that the harmonic distortion of the MGG-DEMON is zero when the signal does not consist of any harmonics. Note that in the DEMONgram outputs in Figs. 8(c), 9(c), and 10(c) however, the MGG-DEMON method exhibits harmonic peaks. This indicates that sources in the recorded data have more than one harmonic frequency in their modulating waveforms. In other words, they do not adhere to the simple source model in (22) and are better described using the more general model in (3).

We now compare the effectiveness of the DEMON algorithms in terms of their ability to detect a peak corresponding to the source modulating frequency in the DEMONgram. The comparison is made in terms of probability of detection (P_D) at a particular level of false alarm (P_{FA}), by plotting the receiver operating characteristics (ROC). The ROC is a plot of the P_D of the peak corresponding to the fundamental frequency of the modulating source, versus the variation in the P_{FA}. In the computation of P_D, we only con-

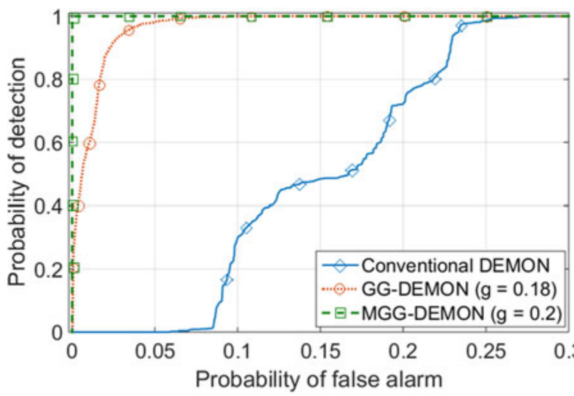

Fig. 11. ROC curve of DEMON methods with recorded dataset #1.

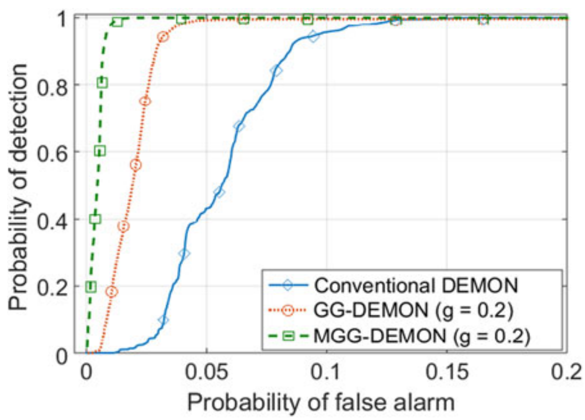

Fig. 12. ROC curve of DEMON methods with recorded dataset #2.

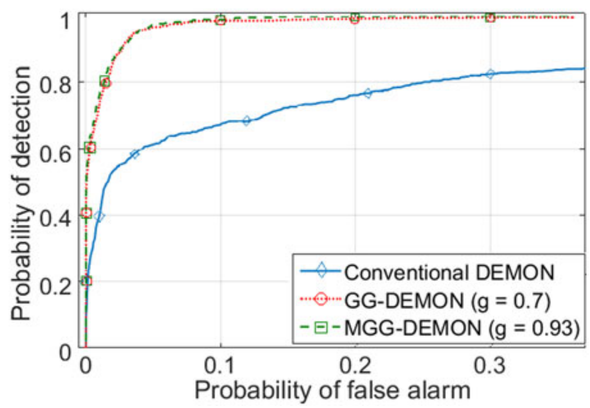

Fig. 13. ROC curve of DEMON methods with recorded dataset #3.

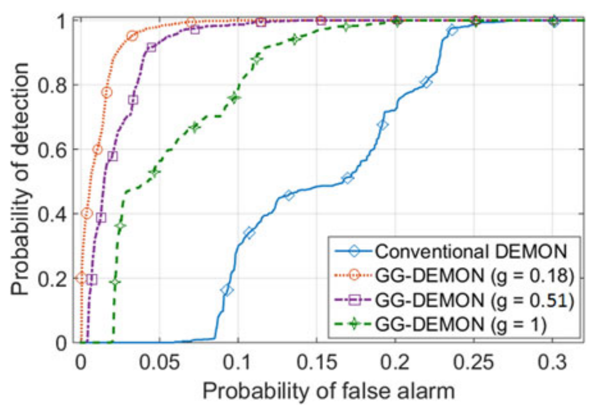

Fig. 14. ROC curve of DEMON methods with different values of g with recorded dataset #1.

sider the detection of fundamental frequency component and ignore the output at the harmonic frequencies in the analysis. This is under the assumption that a sonar operator would be able to discard the harmonic peaks once the fundamental modulating frequency peak is spotted. If the output peaks at the fundamental frequency are above the detection threshold, the source is considered to be detected at these points. Output peaks occurring anywhere other than the fundamental and harmonic frequencies, which are greater than the detection threshold, are considered false alarms.

In Fig. 11, we plot the ROC of the conventional DEMON, GG-DEMON, and MGG-DEMON methods, using dataset #1. The DEMON algorithm parameters used are the same as those used for Fig. 8. Fig. 11 indicates that the robust methods clearly outperform the conventional DEMON detector. The difference in performance is significant, especially at low P_{FA} values. We also observe that MGG-DEMON performs better than GG-DEMON and offers nearly perfect detection for the current scenario considered.

Similarly, in Figs. 12 and 13, we plot the ROC of the DEMON methods using dataset #2 and #3, respectively, with the same algorithm parameters used for Figs. 9 and 10. In Figs. 12 and 13 also, the improvement of GG-DEMON and MGG-DEMON over the conventional DEMON detector is evident. MGG-DEMON outperforms GG-DEMON in Fig. 12, whereas in Fig. 13 the performance of both is more or less equal.

We now consider the selection of the parameter g to be used in the DEMON algorithms. In many applications, since the most effective value of parameter g to be used in the DEMON algorithms is unknown, we can get a good idea on the practical values to be used by fitting a GG pdf to the available data beforehand. This can also be done on ambient noise data that are available prior to detection, in cases where the impulsiveness of the data arises from the noise. Even if ambient noise recordings are not available beforehand, it is clear that in order to achieve robustness in impulsive noise, the algorithms must use a value of $g < 2$. Hence, it is possible to use the detectors with a "blind" nominal value such as $g = 1$, which can ensure robustness and fairly good performance. We will compare the performance of the methods by considering the following four scenarios:

1) No robustification used: $g = 2$ (conventional detector).
2) Robust methods with nominal value of $g = 1$.
3) Robust methods, with the values of g obtained by fitting a GG pdf to the data.
4) Robust methods using values of g that seem to yield the best performance in the range ($0 < g < 2$).

Fig. 14 shows an ROC plot of the GG-DEMON method with the four cases considered when used with dataset #1. We can observe that using the GG-DEMON with $g = 1$

yields fairly better performance than the conventional detector. The best-fitting GG pdf to the data yielded a value of $g = 0.51$, and using this value of g yields a better performance than $g = 1$. The most effective value of $g = 0.18$ was smaller than the best-fit value. Varying the value of g leads to some variation in the performance. However, using any value of g less than or equal to 1 still yields performance much better than the conventional method.

VII. CONCLUSION

The widely used conventional DEMON algorithm is unequipped to tackle source estimation in the case where the data contain impulses. This arises regularly in cases where there is an impulsive signal or background noise, such as in helicopter detection or underwater acoustics. The estimation of the modulating frequency of a source becomes more challenging as the impulsiveness of the data increases. We investigated this by deriving the CRLB for the estimation of a modulating source's parameters in impulsive data.

The theoretical analysis shows that even the optimal achievable performance of estimation decreases with increase in the impulsiveness in data. Given this observation, it is no wonder that the conventional DEMON estimator suffers a heavy performance loss in highly impulsive data. This is because it is formulated for the case of Gaussian-distributed data and, thus, is far from optimal for an impulsive data scenario. This underscores, once again, the importance of formulating robust DEMON methods in order to deal with the more challenging problem of impulsive data.

We formulated two robust near-optimal parametric DEMON algorithms for estimation of the modulating frequency of a source in impulsive data. These were formulated under the assumption that the observed data can be modeled using a GG pdf. We theoretically derived the amplitude of the harmonics introduced by GG-DEMON and MGG-DEMON in their outputs during estimation of a modulated source, and validated this using simulations. We have thus demonstrated both theoretically and using simulations, that the robust methods exhibit lower harmonic distortion than the conventional method. The MGG-DEMON method discussed in this paper offers a solution with even lower harmonic distortion in its output than the GG-DEMON method.

The robust methods were compared with the conventional DEMON method, using simulations as well as recorded data. The results show that the robust methods are effective and deliver good performance even in cases where conventional DEMON fails. The GG-DEMON method performed worse than MGG-DEMON, possibly due to the former's higher harmonic distortion as shown by our theoretical analysis. Hence, GG-DEMON's output duplicates the input noise energy of any frequency in the data at multiple output harmonic frequencies. This aggravates the negative impact of ambient noise by affecting multiple frequencies. However, MGG-DEMON does not suffer from this duplication of noise energy due to its negligible harmonic distortion. Hence, the output spectrum of MGG-DEMON is cleaner than GG-DEMON. This gives it an edge over GG-DEMON in terms of performance of yielding a clear peak at the true source frequency.

The dependence of the performance of the robust methods to the selection of the exponential parameter g was studied. We showed that an effective value of g for implementing a detection algorithm can be obtained by fitting a GG pdf to the data. The most effective value of g may be lower than that obtained through data fit, as indicated by our results with recorded data. If no prior ambient noise recordings are available, even using a nominal value $g = 1$ can deliver performance much better than the conventional method. These pointers establish an approach to select the parameter g for the robust methods in a practical scenario.

The DEMON methods explored in this paper have been formulated for the case of i.i.d. noise. However, this is rarely observed in practical scenarios. Thus, an exciting and practical extension to our work would be to consider the case of colored noise, similar to that done for the case of Gaussian noise previously [5].

APPENDIX

A. DERIVATION OF FISHER INFORMATION MATRIX

In the derivation of the elements of the FIM, we replace the estimation of the unknown parameter f_m with the parameter Nf_m, which allows us to use a lemma from [25] to simplify the expressions. In order to derive meaningful expressions, it is necessary to employ assumptions A2 and A3, as well as an additional assumption A7, which is

A7) The modulation frequency falls in the range $0 < \frac{f_m}{f_s} < \frac{1}{2}$ and $\frac{f_m}{f_s} \neq \frac{1}{4}$ or $\frac{1}{3}$. Since $f_m \ll f_s$, these conditions are almost always satisfied.

The log-likelihood ratio of the N-sample GG-distributed data vector $\mathbf{x}$, conditional on the unknown modulation parameter vector $\alpha = [\sigma_w^2, M, \theta, Nf_m]^T$, is given by

$$\log p(\mathbf{x}|\alpha) = N \log G - 0.5 \sum_{t=1}^{N} \log \sigma_x^2(t) - H \sum_{t=1}^{N} \frac{|x(t)|^g}{(\sigma_x^2(t))^{g/2}}. \tag{34}$$

We will use the notation $\sigma = \sigma_w^2$ and $\nu = Nf_m$. Using (39), the (i, j)th element of the FIM J is given by

$$J_{i,j} = J_{\alpha(i),\alpha(j)} = -E\left[\frac{d^2 \log(p(\mathbf{x}|\alpha))}{d\alpha(i) d\alpha(j)}\right].$$

The derivation follows a procedure similar to that given in [13]. Here, we will demonstrate the derivation for one element of the FIM, $J_{\sigma,\sigma}$, and this can be extended to other elements. For the sake of brevity, let us define the term $z = cf_m t + \theta$ and $y(t) = 1 + M \cos(z)$. $J_{\sigma,\sigma}$ can be derived as

$$
\begin{aligned}
J_{\sigma,\sigma} &= E\left[\frac{d}{d\sigma}\left(\sum_{t=1}^{N}\frac{0.5y^2(t)}{\sigma_x^2(t)} - 0.5gH\sum_{t=1}^{N}\frac{|x(t)|^g y^2(t)}{(\sigma_x^2(t))^{0.5g+1}}\right)\right] \\
&= E\left[-\sum_{t=1}^{N}\frac{0.5y^4(t)}{\sigma_x^4(t)} + 0.5g(0.5g+1)H\sum_{t=1}^{N}\frac{|x(t)|^g y^4(t)}{(\sigma_x^2(t))^{0.5g+2}}\right] \\
&= -\sum_{t=1}^{N}\frac{0.5y^4(t)}{\sigma_x^4(t)} + 0.5g(0.5g+1)H\sum_{t=1}^{N}\frac{E[|x(t)|^g]y^4(t)}{(\sigma_x(t))^{g+4}}.
\end{aligned} \tag{35}
$$

For a GG-distributed variable [26], [27]

$$
E[|x(t)|^g] = \sigma_x^g(t)\left(\frac{\Gamma(1/g)}{\Gamma(3/g)}\right)^{0.5g}/g. \tag{36}
$$

From (36) and (8), we can find that

$$
gH.E[|x(t)|^g] = \sigma_x^g(t). \tag{37}
$$

Substituting (37) in (35)

$$
\begin{aligned}
J_{\sigma,\sigma} &= \left(\sum_{t=1}^{N}\frac{-0.5y^4(t) + 0.5(0.5g+1)y^4(t)}{\sigma_x^4(t)}\right) \\
&= 0.5g\sum_{t=1}^{N}\frac{y^4(t)}{2\sigma_x^4(t)}.
\end{aligned} \tag{38}
$$

This expression is similar to the intermediate one obtained in [13, eq. (A1)], with the additional term $(0.5g)$. This expression is exact, but it offers little insight into the structure of the CRLB of the parameter vector. But we can simplify this expression further in the following way. By invoking the weak signal approximation A2, we can approximate

$$
\sigma_x^2(t) \approx \sigma_n^2. \tag{39}
$$

We will also employ a Lemma by Zhou and Giannakis [25], which states that for $\omega \neq 0 \mod 2\pi$, and any continuous function $h()$

$$
\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}h\left(\frac{t}{N}\right)\exp(i\omega t) = 0. \tag{40}
$$

If assumption A3 is invoked, this implies that for $p = 0, 1, 2,$

$$
\frac{1}{N}\sum_{t=1}^{N}\left(\frac{t}{N}\right)^p \sin(\omega t) \approx 0 \tag{41}
$$

and

$$
\frac{1}{N}\sum_{t=1}^{N}\left(\frac{t}{N}\right)^p \cos(\omega t) \approx 0. \tag{42}
$$

We can expand $y^4(t)/N$ as

$$
\begin{aligned}
\frac{y^4(t)}{N} = \frac{1}{N}\sum_{t=1}^{N}&\big(1 + 4M\cos(z) + 6M^2\cos^2(z) \\
&+ 4M^3\cos^3(z) + M^4\cos^4(z)\big).
\end{aligned} \tag{43}
$$

The second and fourth terms are negligible if we use (42) with $p = 0$. The third and fifth terms can be found to be

$$
\frac{\sum_{t=1}^{N}6M^2\cos^2(z)}{N} \approx 3M^2, \quad \frac{\sum_{t=1}^{N}M^4\cos^4(z)}{N} \approx 0.375M^4. \tag{44}
$$

Substituting (39) and (44) in (38), we can obtain a simplified expression for $J_{\sigma,\sigma}$ as

$$
J_{\sigma,\sigma} \approx (0.5g)\frac{N(1 + 3M^2 + 0.375M^4)}{2\sigma_n^4}. \tag{45}
$$

Using a similar approach as above, we can obtain the exact expressions for the other elements of the FIM as

$$
J_{\sigma,M} = (0.5g)\sum_{t=1}^{N}\frac{\sigma_w^2 y^3(t)\cos(z)}{\sigma_x^4(t)} \tag{46}
$$

$$
J_{\sigma,\theta} = -(0.5g)\sum_{t=1}^{N}\frac{\sigma_w^2 M y^3(t)\sin(z)}{\sigma_x^4(t)} \tag{47}
$$

$$
J_{\sigma,\nu} = -(0.5g)\sum_{t=1}^{N}\left(\frac{2\pi t}{Nf_s}\right)\frac{\sigma_w^2 y^3(t)\cos(z)}{\sigma_x^4(t)} \tag{48}
$$

$$
J_{M,M} = 2(0.5g)\sum_{t=1}^{N}\frac{\sigma_w^4 y^2(t)\cos^2(z)}{\sigma_x^4(t)} \tag{49}
$$

$$
J_{M,\theta} = -2(0.5g)\sum_{t=1}^{N}\frac{\sigma_w^4 M y^2(t)\cos(z)\sin(z)}{\sigma_x^4(t)} \tag{50}
$$

$$
J_{M,\nu} = -2(0.5g)\sum_{t=1}^{N}\left(\frac{2\pi t}{Nf_s}\right)\frac{\sigma_w^4 M y^2(t)\cos(z)\sin(z)}{\sigma_x^4(t)} \tag{51}
$$

$$
J_{\theta,\theta} = 2(0.5g)\sum_{t=1}^{N}\frac{\sigma_w^4 M^2 y^2(t)\sin^2(z)}{\sigma_x^4(t)} \tag{52}
$$

$$
J_{\theta,\nu} = 2(0.5g)\sum_{t=1}^{N}\left(\frac{2\pi t}{Nf_s}\right)\frac{\sigma_w^4 M^2 y^2(t)\sin^2(z)}{\sigma_x^4(t)} \tag{53}
$$

$$
J_{\nu,\nu} = 2(0.5g)\sum_{t=1}^{N}\left(\frac{2\pi t}{Nf_s}\right)^2\frac{\sigma_w^4 M^2 y^2(t)\sin^2(z)}{\sigma_x^4(t)}. \tag{54}
$$

Expressions (46)–(54) are similar to [13, eqs. (A2)–(A20)]. Using (39), and (41) and (42) along with assumption A7, these expressions can be simplified to

$$
J_{\sigma,M} \approx \frac{1.5MN\sigma_w^2(1 + 0.25M^2)(0.5g)}{\sigma_n^4} \tag{55}
$$

$$
J_{\sigma,\theta} \approx 0 \tag{56}
$$

$$
J_{\sigma,\nu} \approx 0 \tag{57}
$$

$$
J_{M,M} \approx \frac{N\sigma_w^4(1 + 0.75M^2)(0.5g)}{\sigma_n^4} \tag{58}
$$

$$
J_{M,\theta} \approx 0 \tag{59}
$$

$$
J_{M,\nu} \approx 0 \tag{60}
$$

$$
J_{\theta,\theta} \approx \frac{\sigma_w^4 NM^2(1 + 0.25M^2)(0.5g)}{\sigma_n^4} \tag{61}
$$

$$J_{\theta,\nu} \approx \frac{\pi\sigma_w^4(N-1)M^2(1+0.25M^2)(0.5g)}{f_s\sigma_n^4} \approx \frac{\pi\sigma_w^4 N M^2(1+0.25M^2)(0.5g)}{f_s\sigma_n^4} \quad (62)$$

$$J_{\nu,\nu} \approx \frac{2\pi^2\sigma_w^4 M^2(2N-1)(N-1)(1+0.25M^2)(0.5g)}{3Nf_s^2\sigma_n^4}. \quad (63)$$

Equations (45) and (55)–(63) yield the simplified expressions for the elements of the FIM.

ACKNOWLEDGMENT

The authors would like to thank the ROMANIS team led by Dr. V. Pallayil at Acoustic Research Laboratory for painstakingly collecting the experimental data from the ROMANIS array during the trials in 2010. The authors would also like to thank the anonymous reviewers whose reviews helped improve the quality of the paper.

REFERENCES

[1] R. Stolkin, S. Radhakrishnan, A. Sutin, and R. Rountree
Passive acoustic detection of modulated underwater sounds from biological and anthropogenic sources
In *Proc. Oceans Conf. Rec.*, 2007, pp. 1–8.

[2] N. Moura, J. Seixas, and R. Ramos
Passive sonar signal detection and classification based on independent component analysis
In *Sonar Systems*, N. Kolev, Ed. Rijeka, Croatia: InTech, 2011, ch. 5, pp. 93–103.

[3] J. Lourens and J. du Preez
Passive sonar ML estimator for ship propeller speed
IEEE J. Ocean. Eng., vol. 23, no. 4, pp. 448–453, Oct. 1998.

[4] R. Nielsen
Sonar Signal Processing. Norwood, MA, USA: Artech House, 1991.

[5] P. Clark, I. Kirsteins, and L. Atlas
Multiband analysis for colored amplitude-modulated ship noise
In *Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Dallas*, 2010, pp. 3970–3973.

[6] I. Kirsteins, P. Clark, and L. Atlas
Maximum likelihood estimation of propeller noise modulation characteristics
In *Proc. Underwater Acoust. Meas., Kos*, 2011.

[7] S. Wisdom, L. Atlas, and J. Pittore
Extending coherence time for analysis of modulated random processes
In *Proc. IEEE Int. Conf. Acoust., Speech Signal Process*. Florence, Italy: IEEE, May 2014, pp. 340–344.

[8] R. Tao, Y. Feng, and Y. Wang
Theoretical and experimental study of a signal feature extraction algorithm for measuring propeller acceleration in a port surveillance system
IET Radar Sonar Navig., vol. 5, no. 2, pp. 172–181, 2011.

[9] L. Fillinger, A. Sutin, and A. Sedunov
Acoustic ship signature measurements by cross-correlation method
J. Acoust. Soc. Amer., vol. 129, no. 2, pp. 774–778, Feb. 2011.

[10] H. Amindavar and P. Moghaddam
Estimation of propeller shaft rate and vessel classification in multipath environment
In *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, 2000, pp. 125–128.

[11] S. Li and D. Yang
DEMON feature extraction of acoustic vector signal based on 3/2-D spectrum
In *Proc. 2nd IEEE Conf. Ind. Electron. Appl., Harbin*, 2007, pp. 2239–2243.

[12] K. Chung, A. Sutin, A. Sedunov, and M. Bruno
DEMON acoustic ship signature measurements in an Urban Harbor
Adv. Acoust. Vib., vol. 2011, no. I, pp. 1–13, 2011.

[13] R. Nielsen
Cramer-Rao lower bounds for sonar broad-band modulation parameters
IEEE J. Ocean. Eng., vol. 24, no. 3, pp. 285–290, Jul. 1999.

[14] R. Lennartsson, E. Dalberg, L. Persson, and S. Petrovic
Passive acoustic detection and classification of divers in harbor environments
In *Proc. OCEANS, Biloxi*, 2009, pp. 1–7.

[15] R. Stolkin *et al.*
Feature based passive acoustic detection of underwater threats
Proc. SPIE, vol. 6204, 2006, Art. no. 620408.

[16] D. Bertilone and D. Killeen
Statistics of biological noise and performance of generalized energy detectors for passive detection
IEEE J. Ocean. Eng., vol. 26, no. 2, pp. 285–294, Apr. 2001.

[17] F. Machell, C. Penrod, and G. Ellis
Statistical characteristics of ocean acoustic noise processes
In *Proc. Top. Non-Gaussian Signal Process.*, E. J. Wegman, S. C. Schwartz, and J. Thomas
Eds. New York, NY, USA: Springer-Verlag, 1989, pp. 29–57.

[18] W. W. L. Au and K. Banks
The acoustics of the snapping shrimp Synalpheus parneomeris in Kaneohe Bay
J. Acoust. Soc. Amer., vol. 103, no. 1, pp. 41–47, 1998.

[19] M. A. Chitre, J. Potter, and S.-H. Ong
Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise
IEEE J. Ocean. Eng., vol. 31, no. 2, pp. 497–503, Apr. 2006.

[20] M. Chitre, M. Legg, and T. Koay
Snapping shrimp dominated natural soundscape in Singapore waters
In *Contributions to Marine Science*, K.-S. Tan, Ed., National University of Singapore, Singapore, 2012, pp. 127–134.

[21] A. Willemsen and M. Rao
Characterization of sound quality of impulsive sounds using loudness based metric
In *Proc. Int. Congr. Acoust.*, 2010, pp. 3397–3404.

[22] F. Schmitz and Y. Yu
Helicopter impulsive noise: Theoretical and experimental status
J. Sound Vib., vol. 109, no. 3, pp. 361–422, Sep. 1986.

[23] Y. H. Yu
Rotor bladevortex interaction noise
Prog. Aerosp. Sci., vol. 36, no. 2, pp. 97–115, Feb. 2000.

[24] G. Arce
Nonlinear Signal Processing: A Statistical Approach. Hoboken, NJ, USA: Wiley 2005.

[25] G. Zhou and G. Giannakis
Harmonics in Gaussian multiplicative and additive noise: Cramer-Rao bounds
IEEE Trans. Signal Process., vol. 43, no. 5, pp. 1217–1231, May 1995.

[26] M. K. Varanasi and B. Aazhang
Parametric generalized Gaussian density estimation
J. Acoust. Soc. Amer., vol. 86, no. 4, pp. 1404–1415, 1989.

[27] J. Dominguez-Molina
A practical procedure to estimate the shape parameter in the generalized Gaussian distribution
Tech. Rep., 2003. [Online]. Available: http://random.cimat.mx/reportes/enlinea/I-01-18_eng.pdf

Hari Vishnu (S'04–M'08) received the B.Tech degree in electronics and electrical engineering from the National Institute of Technology, Calicut, India, in 2008 and the Ph.D. degree in computer engineering from the Nanyang Technological University, Singapore, in 2013. His Ph.D. research involved development of improved underwater signal processing algorithms. These included detection and localization, vector sensor processing, developing robust methods in non-Gaussian noise, acoustic modeling, and parameter estimation using Bayesian filtering.

Since 2013, he has been at the Acoustic Research Laboratory, National University of Singapore, Singapore. He is currently a research lead, and his research interests include active sonar, bio-acoustics, acoustic propagation modeling, machine learning, and robust signal processing.

Dr. Vishnu was a reviewer for several reputed international journals. He is currently the Secretary of the IEEE Singapore AUV Challenge Committee and the IEEE Oceanic Engineering Society, Singapore chapter.

Mandar Chitre (S'04–M'05–SM'11) received the B.Eng. and M.Eng. degrees in electrical engineering from the National University of Singapore (NUS), Singapore, in 1997 and 2000, respectively, the M.Sc. degree in bioinformatics from the Nanyang Technological University, Singapore, in 2004, and the Ph.D. degree in underwater communications from NUS, in 2006.

From 1997 to 1998, he was with the ARL, NUS, Singapore. From 1998 to 2002, he headed the technology division of a regional telecommunications solutions company. In 2003, he rejoined ARL, initially as the Deputy Head (Research) and is currently the Head of the laboratory. He also holds a joint appointment in the Department of Electrical and Computer Engineering, NUS, as an Assistant Professor. His current research interests include underwater communications, autonomous underwater vehicles, model-based signal processing, and modeling of complex dynamic systems.

Dr. Chitre has served on the technical program committees of the IEEE OCEANS, WUWNet, DTA, and OTC conferences and has served as a reviewer for numerous international journals. He was the Chairman of the student poster committee for the IEEE OCEANS 06 in Singapore, and the Chairman for the IEEE Singapore AUV Challenge 2013. He is currently the IEEE Oceanic Engineering Society Technology Committee Co-Chair of underwater communication, navigation, and positioning.