Principal component analysis for emergent acoustic signal detection with supporting simulation results

Elizabeth Hoppe[a)] and Michael Roan
Department of Mechanical Engineering, Virginia Polytechnic Institute and State University,
114G Randolph Hall, Blacksburg, Virginia 24061

A method is introduced that uses principal component analysis (PCA) to detect emergent acoustic signals. Emergent signal detection is frequently used in radar applications to detect signals of interest in background clutter and in cognitive radio to detect the primary user in a frequency band. The method presented differs from other standard techniques in that the detection of the signal of interest is accomplished by detecting a change in the covariance between two channels of data instead of detecting the change in statistics of a single channel of data. For this paper, PCA is able to detect emergent acoustic signals by detecting when there is a change in the eigenvalue subspace of the covariance matrix caused by the addition of the signal of interest. The algorithm's performance is compared to an energy detector and the Neyman-Pearson theorem. Acoustic simulations were used to verify the performance of the algorithm. Simulations were also used to examine the effectiveness of the algorithm under various signal-to-interferer and signal-to-noise ratios, and using various test signals.
© 2011 Acoustical Society of America. [DOI: 10.1121/1.3628324]

PACS number(s): 43.60.Bf, 43.60.Cg [EJS] Pages: 1962–1973

I. INTRODUCTION

Change detection and emergent signal detection are widely studied problems in a variety of signal processing fields, including radar and sonar, biomedical, and speech detection applications.[1–4] One branch of change detection research involves development of algorithms for detecting important changes in a signal, such as a change in the mean or variance.[3–5] This is accomplished by processing signals that emanate from a system, such as vibration signals measured on the outside of a gear box.[6] A crack in a gear tooth manifests itself as a detectable change in the statistics of the signal. In a similar application, change detection is used to passively monitor the structure of a naval ship.[7] In another application, change detection is used for automatic electrocardiogram (ECG) diagnosis.[8] The statistics of the ECG signal are monitored to detect changes in the rhythm of the heart, such as fibrillation and hypertrophy.[9,10]

Another branch of change detection research aims to determine whether a target signal is present within noise.[11–13] This category of algorithms is often used in passive sonar applications.[2] Algorithms of this type include wildlife monitoring, such as the detection of whale calls in the ocean.[14,15] Similar algorithms are also widely employed in radar applications to detect signals of interest in background clutter.[16,17] This type of change detection is used to determine when new signals appear in (or leave) the environment. This is useful for algorithms that require an estimate of the number of sources (e.g., the MUSIC algorithm[18]). A subset of this branch of change detection is voice activity detection (VAD).[19–22] The algorithm presented in this paper

falls under this second branch of change detection research and seeks to detect an emergent acoustic signal in noise.

Statistical hypothesis testing is one of the primary tools used to detect the presence of an emergent signal in a set of measurements.[23] In binary detection there are two hypotheses, and for this paper, these hypotheses are the signal of interest is either present or absent. In hypothesis testing, a sufficient statistic (such as the log-likelihood ratio) is used to select one of these hypotheses based on a small sample of data.[24] This paper presents a unique approach to emergent acoustic signal detection based on the signals recorded from an array of two sensors. This new method does not require prior knowledge of the probability distribution of either the signal or noise, but relies on detecting an increase in the covariance of the mixed signals when the signal of interest is present.

The algorithm developed in this paper uses principal component analysis (PCA) to calculate a test statistic based on the smallest eigenvalue of the covariance matrix to determine whether the target signal is present. Where PCA has been used previously in change detection is specifically for the VAD problem.[25,26] In Ref. 25, the authors use PCA to project the signal into a space that has the highest signal-to-noise ratio (SNR). They are finding a projection direction that minimizes the noise present in the signal, which allows for improved voice detection. In Ref. 26, PCA analysis is used on the discrete cosine transform (DCT) coefficients of a signal to separate the independent noise from the voice. In these previous works, PCA is not used as an emergent acoustic signal detector, but rather a noise suppression approach.

II. BACKGROUND

A. Principal component analysis

PCA is a tool used for analyzing multivariate data that has seen use across a wide variety of application areas such as

a)Author to whom correspondence should be addressed. Electronic mail: ehoppe@vt.edu

classification, pattern recognition, and noise reduction.[27–30] PCA transforms mutually correlated variables into a set of uncorrelated variables that preserve the variation in the original data set. This allows PCA to separate sources that are uncorrelated. This correlation transformation is only dependent on the second order statistics of the signals and has a simple, closed form solution.[31] The transformed variables are ordered so that the first few components contain most of the original variation information, and the last components contain little to no additional variance information.

Suppose that $\mathbf{x}$ is a vector of p random variables that are mutually correlated. The first principal component (PC) is a linear combination, $y_1 = \mathbf{w}_1^T \mathbf{x}$, such that the variance of y_1 is a maximum. In other words, the first PC is calculated by finding a weight vector, $\mathbf{w}_1$, which is the optimal solution to the LaGrange problem

$$\max[\text{var}(y_1)] \quad \text{subject to} \quad \|\mathbf{w}_1\| = 1, \tag{1}$$

where the norm of $\mathbf{w}_1$ is the Euclidean norm. The norm of $\mathbf{w}_1$ is constrained to be equal to one to prevent the variance from growing without limits. The variance of the PC can be rewritten as

$$\text{var}(y_1) = E\{y_1^2\} = E\{(\mathbf{w}_1^T \mathbf{x})^2\} = \mathbf{w}_1^T E\{\mathbf{x}\mathbf{x}^T\}\mathbf{w}_1$$
$$= \mathbf{w}_1^T \mathbf{C_x} \mathbf{w}_1 \tag{2}$$

where $E\{\cdot\}$ is the expectation and $\mathbf{C_x}$ is the covariance matrix of $\mathbf{x}$. Substituting Eq. (2) into Eq. (1) and differentiating the LaGrange function produces the equation

$$\mathbf{C_x}\mathbf{w}_1 - \lambda\mathbf{w}_1 = (\mathbf{C_x} - \lambda)\mathbf{w}_1 = 0, \tag{3}$$

which shows that the weight vector $\mathbf{w}_1$ is the unit length eigenvector of the covariance matrix. The kth PC of $\mathbf{x}$ is $y_k = \mathbf{w}_k^T \mathbf{x}$, and has a variance given by

$$\text{var}(y_k) = \mathbf{w}_k^T \mathbf{C_x} \mathbf{w}_k = \mathbf{w}_k^T \lambda_k \mathbf{w}_k = \lambda_k \mathbf{w}_k^T \mathbf{w}_k = \lambda_k, \tag{4}$$

where λ_k is the kth largest eigenvalue of $\mathbf{C_x}$. The second PC is another linear combination, $y_2 = \mathbf{w}_2^T \mathbf{x}$, such that $\mathbf{w}_2$ is orthogonal to $\mathbf{w}_1$ (uncorrelated with the first PC). Since eigenvectors are always orthogonal, the weight vectors, $\mathbf{w}_i$, are the set of eigenvectors of the data covariance matrix.[32]

For this paper, PCA will be used in an application where only two channels of data are measured. Therefore, the data set $\mathbf{x}$ is a $2 \times n$ matrix (also assumed zero mean). This allows for a closed form derivation of the PCs in order to test the hypothesis that PCA can be used as a simple and robust VAD. The sample covariance matrix is calculated element by element according to

$$c_x(i,j) = E\{\mathbf{x_i}^T \mathbf{x_j}\} = \frac{1}{n}\sum_{k=1}^{n} x_i(k)x_j(k). \tag{5}$$

For the data set $\mathbf{x}$, the covariance matrix will be a 2×2 matrix, and have two eigenvalues. The closed form solution to the eigenvalue problem is straightforward to derive, and is

$$\lambda_{1,2} = 0.5[c_x(1,1) + c_x(2,2)]$$
$$\pm 0.5\sqrt{\left[c_x(1,1)^2 - c_x(2,2)^2\right]^2 + 4c_x(1,2)^2}. \tag{6}$$

B. Emergent acoustic signal detection using principal component analysis

PCA separates uncorrelated sources by determining directions that have maximum variance. Uncorrelated sources are recovered by projecting the mixed signals along these directions. The addition of emerging signals alters the eigenvalue subspace, and therefore, the projection directions. By examining the changes in the individual PCs, the change in the subspace can be uncovered, and PCA can be used to detect emergent acoustic signals. The advantage of the PCA method is that it is nonadaptive and has an easily derived closed form solution.

Independent component analysis (ICA) theory states that the mixture of two or more sub- or super-Gaussian signals is more Gaussian than either of the individual unmixed source signals.[33] This paper exploits this property to develop an emergent acoustic signal detection technique. If the interferer only case has a certain probability distribution function (pdf), the addition of the signal of interest to that interferer will cause an increase in the variance of that pdf. The advantage here is that the individual distributions do not need to be known in advance, and the signal and noise could have identical pdfs without affecting the method. If only one channel of data were available, the detection would be based a change in the variance. For two channels of data, the detection is based on a change in the variance of the joint pdf.

Figure 1 shows a plot of the amplitude of the first sensor versus the second sensor. Figure 1(a) is the case where only the interferer is present. Figure 1(b) is the case where both signal and interferer are present. The lines in Fig. 1 lie along the directions of the PCs. In Fig. 1(a), one direction (i.e., one PC) contains most of the variance. This direction corresponds to the interferer, and since the interferer is the only signal present, there is little variance in any other direction. In Fig. 1(b), the directions of the PCs have changed, and the variance along both PC directions has increased. This increase is especially pronounced along the second PC, labeled "PC 2" in Fig. 1. This increase indicates the presence of one additional signal. For this paper, the smallest PC will be considered the significant statistic. The smallest PC is chosen because of a unique attribute that occurs in the simultaneous mixing case discussed in Sec. III A.

1. Relationship to the MUSIC algorithm for DOA estimation

This new PCA based detection algorithm is based on some of the same principles as the multiple signal
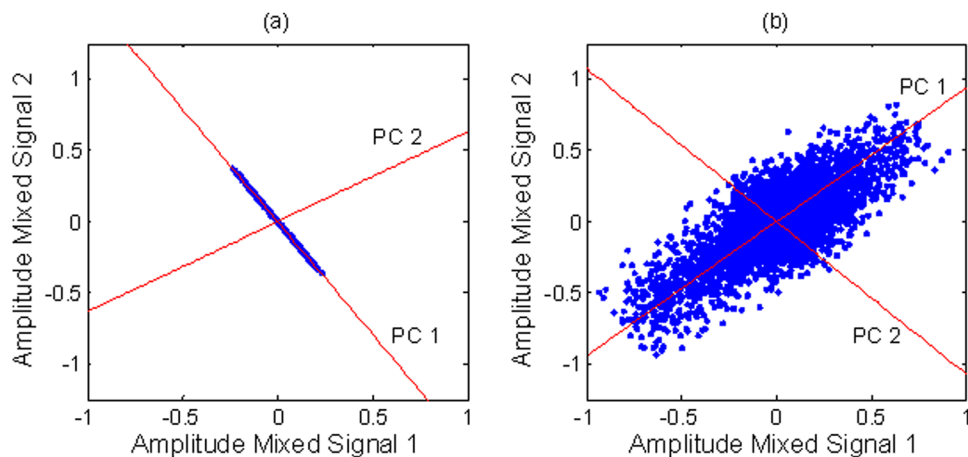
FIG. 1. (Color online) A visualization of the joint pdf for two cases is presented: only one signal present and both signals present. A PCA analysis is performed on the data, and the directions of the PCs are plotted as lines. (a) The joint pdf for the case where only the interferer is present. (b) The joint pdf for the case where both the interferer and the signal of interest are present.

classification (MUSIC) algorithm for direction of arrival (DOA) estimation. The MUSIC algorithm is used to determine the parameters of multiple wavefronts arriving at an array.[18] The D waveforms are received at an array with M elements according to the model

$$X = AF + W, \tag{7}$$

where is X is vector of the received signal, A is the array manifold matrix, F is the vector of incident signals, and W is the vector of noise. The $M \times M$ covariance matrix of X is

$$S = XX^* = AFF^*A^* + WW^* = APA^* + \lambda S_0, \tag{8}$$

where the asterisk represents the complex conjugate and the signal and noise are assumed to be uncorrelated. When the number of incident waveforms is less than the number of array elements ($D < M$), APA^* is singular and

$$|APA^*| = |S - \lambda S_0| = 0. \tag{9}$$

This equation is only satisfied if λ is an eigenvalue of S in the metric of S_0.[18] When determining the eigenvalues of S, the minimum eigenvalue is repeated N times. This eigenvalue spans the noise subspace. The D remaining eigenvalues span the signal subspace. In the MUSIC algorithm, the DOA of the sources are estimated by picking the D highest peaks of the function

$$P_{MU}(\theta) = \frac{1}{a^*(\theta)E_N E_N^* a(\theta)}, \tag{10}$$

where E_N is the $M \times N$ matrix whose columns are the N noise eigenvectors. Therefore, the MUSIC algorithm uses the noise subspace (defined by the noise eigenvectors) to estimate parameters of the incident wavefronts. Instead of using the eigenvectors to describe the noise subspace, the new method uses eigenvalues to detect changes in the size of the signal subspace.

C. Energy detection test

The energy detector test is a useful tool for detecting a signal within white, Gaussian noise[34] and will be used as a baseline for comparison in this paper. The H_0 hypothesis is that only the interfering source is present, while H_1 states that both the signal of interest and the interferer are present (no noise). For this derivation, the signal of interest and interferer are additively mixed according to

$$\begin{aligned} H_0: \quad & x[n] = w[n], & n = 0, 1, \ldots, N-1, \\ H_1: \quad & x[n] = s[n] + w[n], & n = 0, 1, \ldots, N-1, \end{aligned} \tag{11}$$

where $w[n]$ is the interferer and $s[n]$ is the signal of interest. The interferer is assumed to be white, Gaussian noise, with a mean of zero and a variance of σ^2. The signal of interest is also assumed to be white, Gaussian noise, with a mean of zero and a variance of σ_s^2. The Neyman-Pearson theorem says to declare H_1 true if

$$L(x) = \frac{p(x; H_0)}{p(x; H_1)} > \gamma, \tag{12}$$

where γ is a threshold value chosen by the user.[35] The function $L(x)$ is called the likelihood ratio because it determines the likelihood of H_1 versus H_0 for each value of x. Using the assumption that both the interferer and the signal of interest have Gaussian distributions, the likelihood ratio is written as[35]

$$p(x; H_0) = \mathcal{N}(0, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1} x^2[n]\right]$$

$$p(x; H_1) = \mathcal{N}(0, \sigma_s^2 + \sigma^2) = \frac{1}{[2\pi(\sigma_s^2 + \sigma^2)]^{N/2}}$$

$$\times \exp\left[-\frac{1}{2(\sigma_s^2 + \sigma^2)}\sum_{n=0}^{N-1} x^2[n]\right]$$

$$L(x) = \left(\frac{\sigma_s^2 + \sigma^2}{\sigma^2}\right)^{N/2} \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2} - \frac{1}{(\sigma_s^2 + \sigma^2)}\right)\right.$$

$$\left. \times \sum_{n=0}^{N-1} x^2[n]\right] > \gamma, \tag{13}$$

E. Hoppe and M. Roan: Signal detection using principal components

where $N(\mu, \sigma^2)$ represents a normal distribution with a mean of μ and a standard deviation of σ^2. Taking the natural log of Eq. (13) gives the log-likelihood ratio to be

$$l(x) = \frac{N}{2}\ln\left(\frac{\sigma_s^2 + \sigma^2}{\sigma^2}\right) - \frac{1}{2}\left(\frac{1}{\sigma^2} - \frac{1}{(\sigma_s^2 + \sigma^2)}\right)\sum_{n=0}^{N-1}x^2[n]$$

$$> \ln(\gamma)\frac{1}{2}\left(\frac{\sigma_s^2}{\sigma^2(\sigma_s^2 + \sigma^2)}\right)\sum_{n=0}^{N-1}x^2[n]$$

$$> \frac{N}{2}\ln\left(\frac{\sigma_s^2 + \sigma^2}{\sigma^2}\right) - \ln(\gamma). \tag{14}$$

Therefore, the H_1 hypothesis is declared true if

$$T(x) = \sum_{n=0}^{N-1}x^2[n] > \gamma', \tag{15}$$

where γ' is a user defined threshold value.

1. Neyman-Pearson limit for the energy detector

In this paper, the problem is to choose between the null hypothesis, H_0, and the alternative hypothesis, H_1. When the H_1 hypothesis is actually true, and the test declares H_1 to be true, it is called a detection. When the H_0 hypothesis is actually true, and the test declares H_1 to be true, it is called a false alarm. To design an optimal detector, the goal is to minimize the number of false alarms, while maximizing the number of detections.[34] The probability of false alarm (P_{FA}) constraint is written as

$$P_{FA} = \int_\gamma^\infty p(x; H_0)dx = \alpha, \tag{16}$$

where $p(x; H_0)$ is the pdf of signal x under the H_0 hypothesis, γ is the unknown threshold value, and α is the specified maximum allowable P_{FA} value. The goal is to chose the threshold, γ, such that the P_{FA} value does not exceed α, while maximizing the P_D, which is written as

$$P_D = \int_\gamma^\infty p(x; H_1)dx, \tag{17}$$

where $p(x; H_1)$ is the pdf of signal x under the H_1 hypothesis. For this paper, the P_{FA} is the probability that a frame of audio does not contain the signal of interest, but the test algorithm declares that the signal of interest is present. Furthermore, the P_D is the probability that a frame of audio contains the signal of interest, and the test algorithm declares that the signal of interest is present. The algorithm declares the signal of interest is present when the test statistic exceeds the user defined threshold. By sweeping the detection threshold through a range of values, and calculating P_D and the P_{FA} for each threshold, the receiver operating characteristic (ROC) curve is generated.[35] ROC curves are used to examine the performance of a test algorithm and to compare multiple test algorithms.

The Neyman-Pearson limit determines the theoretical best performance of a binary hypothesis test by using Eqs. (16) and (17) to determine the maximum P_D for a specific P_{FA} value.[35] Since the statistic $T(x)$ in Eq. (15) is the sum of squares of N independent and identically distributed Gaussian random variables, this test statistic has a χ^2 distribution given by

$$\frac{T(x)}{\sigma^2} \sim \chi_N^2 \quad \text{under } H_0,$$
$$\frac{T(x)}{\sigma_s^2 + \sigma^2} \sim \chi_N^2 \quad \text{under } H_1, \tag{18}$$

The P_D is found by

$$P_D = \Pr\{T(x) > \gamma'; H_1\} = \Pr\left\{\frac{T(x)}{\sigma_s^2 + \sigma^2} > \frac{\gamma'}{\sigma_s^2 + \sigma^2}; H_1\right\}$$
$$= Q_{\chi^2}\left(\frac{\gamma'}{\sigma_s^2 + \sigma^2}\right), \tag{19}$$

while the P_{FA} is found by

$$P_{FA} = \Pr\{T(x) > \gamma'; H_0\} = Q_{\chi^2}\left(\frac{\gamma'}{\sigma^2}\right), \tag{20}$$

where $Q_{\chi 2}(n)$ is the probability that the χ^2 distribution has a value greater than n. Equation (20) can be used to determine an appropriate threshold value to maximize P_D for a given P_{FA}. In addition, Eqs. (19) and (20) can be used to generate the optimum ROC curve for the energy detector.

III. EMERGENT ACOUSTIC SIGNAL DETECTION MODELS AND SUFFICIENT STATISTICS

In this work, there are three types of signals to be considered: the signal of interest, an interferer that is considered to be a point source, and background noise. The interferer is considered to always be present. Two general cases, with and without the background noise, will be considered. The significant statistic is derived for the model where the signal of interest and the interferer are mixed additively. The interferer for all models is assumed to be a stationary signal. In the simulations, the interferer is a white, Gaussian noise source. In order to accommodate a variety of signal of interest types, the only assumption made about the signal of interest is that it has zero mean. Therefore, the mixed signals, for both the signal of interest present and absent case, are assumed to have zero mean.

In all instances, it is assumed that the interferer and background noise are uncorrelated with the signal of interest. The PCA emergent signal detection algorithm is based on separating uncorrelated sources. Each uncorrelated source corresponds to a subspace eigenvector direction. One direction corresponds to the signal of interest and one to the interferer. If the interferer is correlated with the signal of interest, then the presence of the signal of interest will add variance along the interferer's subspace direction, as opposed to adding variance along the signal of interest's subspace direction. Therefore, the new algorithm will not be able to detect a change in the smallest eigenvalue corresponding to the signal of interest's subspace and the algorithm will fail to detect the presence of the signal of interest. Similarly, the detection of the signal of interest is affected by the presence of

multiple sources and interferers. The models presented in this section are not applicable to that situation. To determine if PCA is useful in that situation, new models would have to be developed and examined. For this paper, the focus is for the instance of one signal of interest and one interferer.

A. Simultaneous mixing model

Two forms of signal mixing are introduced in this section. The first is simultaneous mixing, where there is no time of arrival delay between the first and second microphones. This is the instantaneous mixing case that is the underlying assumption in many of the ICA models.[31,33] The second case reflects the more realistic case, and allows for a time delay of arrival (TDOA) between the two sensors. While adding another sensor increases the amount of information available, it comes with the drawback that the signals are not time aligned between the two sensors. Following sections will show how the performance changes due to this time misalignment.

In this section, the simultaneous mixing model is presented. Let s_1 be the signal of interest with variance σ_{s1}, and s_2 be an interfering signal with variance σ_{s2}. The goal is to detect the presence of the signal of interest s_1. H_1 represents the case where the signal of interest, s_1, and the interferer, s_2, are both present and H_0 represents the case where only the interferer is present. Models including diffuse background noise appear in later sections. For the simultaneous mixing model, the two hypotheses are represented by the models

$$H_1 : x_1(k) = as_1(k) + bs_2(k), \quad x_2(k) = cs_1(k) + ds_2(k),$$
$$H_0 : x_1(k) = bs_2(k), \quad x_2(k) = ds_2(k), \tag{21}$$

where k is the sample number, and a, b, c, and d are the linear mixing coefficients. The smallest eigenvalue of the covariance matrix is calculated for each hypothesis. In each case, the covariance matrix elements are calculated using Eq. (5). For brevity, only the derivation for the (1,2) element will be shown. The (1,2) element of the covariance matrix for the H_1 hypothesis is calculated as

$$c_x(1,2) = \frac{1}{n}\sum_{k=1}^{n} x_1 x_2 = \frac{1}{n}\sum_{k=1}^{n}(as_1 + bs_2)(cs_1 + ds_2)$$
$$= \frac{1}{n}\sum_{k=1}^{n}\left[acs_1^2(k) + (ad + bc)s_1(k)s_2(k) + bds_2^2(k)\right]. \tag{22}$$

Using the properties of the summation, Eq. (22) can be rewritten as

$$c_x(1,2) = ac\left[\frac{1}{n}\sum_{k=1}^{n}s_1^2(k)\right] + (ad + bc)\left[\frac{1}{n}\sum_{k=1}^{n}s_1(k)s_2(k)\right]$$
$$+ bd\left[\frac{1}{n}\sum_{k=1}^{n}s_2^2(k)\right]$$
$$= ac\sigma_{s1} + (ad + bc)\sigma(s_1 s_2) + bd\sigma_{s2}, \tag{23}$$

where $\sigma(s_1 s_2)$ is the covariance between the signal of interest and the interferer. For the H_1 hypothesis, the covariance matrix is

$$\mathbf{C_x} = \begin{bmatrix} a^2\sigma_{s1} + 2ab\sigma(s_1 s_2) + b^2\sigma_{s2} & ac\sigma_{s1} + (ad + bc)\sigma(s_1 s_2) + bd\sigma_{s2} \\ ac\sigma_{s1} + (ad + bc)\sigma(s_1 s_2) + bd\sigma_{s2} & c^2\sigma_{s1} + 2cd\sigma(s_1 s_2) + d^2\sigma_{s2} \end{bmatrix}. \tag{24}$$

Substituting the values from Eq. (24) into Eq. (6) gives the smallest eigenvalue of the covariance matrix for the H_1 hypothesis as

$$\lambda_1 = 0.5\left[(a^2 + c^2)\sigma_{s1} + 2(ab + cd)\sigma(s_1 s_2) + (b^2 + d^2)\sigma_{s2}\right]$$
$$- 0.5\sqrt{\begin{array}{l}[(a^2 - c^2)\sigma_{s1} + 2(ab - cd)\sigma(s_1 s_2) + (b^2 - d^2)\sigma_{s2}]^2 \\ + 4[ac\sigma_{s1} + (ad + bc)\sigma(s_1 s_2) + bd\sigma_{s2}]^2\end{array}}. \tag{25}$$

For the H_0 hypothesis, the smallest eigenvalue is

$$\lambda_1 = 0.5(b^2 + d^2)\sigma_{s2} - 0.5\sqrt{[(b^2 + d^2)\sigma_{s2}]^2} = 0. \tag{26}$$

Equations (25) and (26) show that the smallest eigenvalue can be used to detect the presence of the signal of interest when it is mixed with an interferer. When the signal of interest is present, the smallest eigenvalue is a function of the variances of the signal of interest and the inter-

ferer, the covariance between the two signals, and the elements of the linear mixing matrix. However, when the signal of interest is not present, the eigenvalue becomes zero. This theoretical result shows, interestingly, that for the simultaneous mixing case with one interferer, the signal of interest can be detected for any signal-to-interferer ratio (SIR) because the elements of the linear mixing matrix do not influence the result for the H_0 hypothesis. This result is based on knowing the signal variances and covariance exactly. In practice, only the sample variances and covariances are known, so the smallest eigenvalue for the H_0 hypothesis may not always reduce to precisely zero, and the SIR will have a small influence on the results. In addition, while the method theoretically works for infinitely large values of SIR, in practice, for large SIR values, the signal of interest could drop below the quantization level of the analog to digital convertor and the method would break down. In the case of simultaneous mixing, the hardware would be the limiting factor in determining the SIR values at which the algorithm breaks down.

E. Hoppe and M. Roan: Signal detection using principal components

B. Mixing model including TDOA

The previous section derived the smallest eigenvalue in the case of simultaneous mixing. This model, while widely used for ICA derivations, is not suitable for most recorded array signals because time delays between array elements are introduced by the source/receiver geometry.[36] This section applies the new PCA technique to a model that includes these delays. This model is applicable to a variety of situations, for example, passive sonar. In this example, a surface ship engine could be the interferer, and the object is to detect the presence of another submarine or ship engine. Similarly, if the goal is to detect animals, a steady hum from a generated could be used as the interferer, and the calls of the animals are the signals of interest. If the animal of interest is a human, this model is applicable to the VAD problem. In that case, the fan noise from the HVAC system could be the interfere and the voice is the signal of interest.

Similar to the previous section, let s_1 be the signal of interest with variance σ_{s1}, and s_2 be an interfering signal with variance σ_{s2}. The two hypotheses are

$$H_1 : x_1(k) = as_1(k) + bs_2(k),$$
$$x_2(k) = cs_1(k + \tau_1) + ds_2(k + \tau_2),$$
$$H_0 : x_1(k) = bs_2(k), \quad x_2(k) = ds_2(k + \tau_2), \tag{27}$$

where k, a, b, c, and d are the same as in the previous section, and τ_1 and τ_2 are the time delays for the signal of interest and interferer, respectively. As before, the covariance matrix is constructed element by element, and the (1,2) element calculation for the H_1 hypothesis is given next as an example. The (1,2) element is calculated as

$$c_x(1,2) = ac\left[\frac{1}{n}\sum_{k=1}^{n} s_1(k)s_1(k + \tau_1)\right] + ad\left[\frac{1}{n}\sum_{k=1}^{n} s_1(k)s_2(k + \tau_2)\right]$$
$$+ bc\left[\frac{1}{n}\sum_{k=1}^{n} s_1(k + \tau_1)s_2(k)\right] + bd\left[\frac{1}{n}\sum_{k=1}^{n} s_2(k)s_2(k + \tau_2)\right]. \tag{28}$$

Equation (28) generates several new covariance terms that arise due to delays between the signals. A superscript τ denotes a delayed signal. Equation (28) is rewritten as

$$c_x(1,2) = ac\sigma(s_1 s_1^\tau) + ad\sigma(s_1 s_2^\tau) + bc\sigma(s_1^\tau s_2) + bd\sigma(s_2 s_2^\tau), \tag{29}$$

where $\sigma(\cdot)$ is the covariance between the signals in parentheses. For the simulations in this paper, the interfering signal, s_2, is assumed to be white, Gaussian noise. Other interferers could be used, but stationary, white, Gaussian noise is chosen for simplicity:

$$\mathbf{C_x} = \begin{bmatrix} a^2\sigma_{s1} + 2ab\sigma(s_1 s_2) + b^2\sigma_{s2} & ac\sigma(s_1 s_1^\tau) + ad\sigma(s_1 s_2^\tau) \\ +ac\sigma(s_1 s_1^\tau) + ad\sigma(s_1 s_2^\tau) & bc\sigma(s_1^\tau s_2) + bd\sigma(s_2 s_2^\tau) \\ + bc\sigma(s_1^\tau s_2) + bd\sigma(s_2 s_2^\tau) & c^2\sigma(s_1^\tau s_1^\tau) + 2cd\sigma(s_1^\tau s_2^\tau) + d^2\sigma_{s2} \end{bmatrix}. \tag{30}$$

The smallest eigenvalue for the H_1 hypothesis is

$$\lambda_1 = 0.5\left[a^2\sigma_{s1} + c^2\sigma(s_1^\tau s_1^\tau) + 2ab\sigma(s_1 s_2) + 2cd\sigma(s_1^\tau s_2^\tau) + (b^2 + d^2)\sigma_{s2}\right]$$
$$- 0.5\sqrt{\begin{array}{c}\left[a^2\sigma_{s1} - c^2\sigma(s_1^\tau s_1^\tau) + 2ab\sigma(s_1 s_2) - 2cd\sigma(s_1^\tau s_2^\tau) + (b^2 - d^2)\sigma_{s2}\right]^2 \\ + 4\left[ac\sigma(s_1 s_1^\tau) + ad\sigma(s_1 s_2^\tau) + bc\sigma(s_1^\tau s_2) + bd\sigma(s_2 s_2^\tau)\right]^2\end{array}}. \tag{31}$$

The smallest eigenvalue for the H_0 hypothesis is

$$\lambda_1 = 0.5(b^2 + d^2)\sigma_{s2} - 0.5\sqrt{(b^2 - d^2)^2\sigma_{s2}^2 + 4b^2d^2\sigma(s_2 s_2^\tau)^2}. \tag{32}$$

Unlike in the previous section, the eigenvalue for the H_0 hypothesis does not reduce to zero in the absence of the signal of interest. The covariance terms introduced by the delays in the model do not cancel. However, switching from the H_0 to the H_1 hypothesis still causes an increase

in the smallest eigenvalue, which allows for the detection of the signal of interest, but complicates the threshold calculation.

C. One signal, one interferer with uncorrelated background noise

The model in this section builds on the previous section by using the same delayed signal scheme, but adding uncorrelated noise to the two mixed signals. This is a more realistic scenario that includes the effect of noise that is not present in both sensors, such as sensor noise. As before, s_1 is the signal of interest with variance σ_{s1}, and s_2 is the interfering signal with variance σ_{s2}. The two hypotheses are written as

$$
\begin{aligned}
H_1 \; &: \; \begin{array}{l} x_1(k) = as_1(k) + bs_2(k) + n_1(k), \\ x_2(k) = cs_1(k+\tau_1) + ds_2(k+\tau_2) + n_2(k), \end{array} \\
\\
H_0 \; &: \; \begin{array}{l} x_1(k) = bs_2(k) + n_1(k), \\ x_2(k) = ds_2(k+\tau_2) + n_2(k). \end{array}
\end{aligned}
\tag{33}
$$

In this case, the noise signals, n_1 and n_2, represent a combination of diffuse background noise, sensor noise, and electrical noise, and are uncorrelated between sensors. The covariance matrix is constructed element by element as before. Once again, the interferer signal, s_2, is assumed to be short-term stationary, and the variance of the delayed signal, $s_2(k+\tau_2)$, is assumed to be equal to the variance of signal s_2. For brevity, the covariance matrix has been omitted, and the smallest eigenvalue for the H_1 hypothesis is

$$
\begin{aligned}
\lambda_1 = 0.5 &\left[\begin{array}{l} a^2\sigma_{s1} + c^2\sigma(s_1^\tau s_1^\tau) + (b^2+d^2)\sigma_{s2} + \sigma_{n1} + \sigma_{n2} + 2ab\sigma(s_1 s_2) + 2cd\sigma(s_1^\tau s_2^\tau) \\ + 2a\sigma(s_1 n_1) + 2b\sigma(s_2 n_1) + 2c\sigma(s_1^\tau n_2) + 2d\sigma(s_2^\tau n_2) \end{array} \right] \\
- 0.5 &\sqrt{ \left[\begin{array}{l} a^2\sigma_{s1} - c^2\sigma(s_1^\tau s_1^\tau) + (b^2-d^2)\sigma_{s2} + \sigma_{n1} - \sigma_{n2} + 2ab\sigma(s_1 s_2) - 2cd\sigma(s_1^\tau s_2^\tau) \\ + 2a\sigma(s_1 n_1) + 2b\sigma(s_2 n_1) - 2c\sigma(s_1^\tau n_2) - 2d\sigma(s_2^\tau n_2) \end{array} \right]^2 + 4\left[\begin{array}{l} ac\sigma(s_1 s_1^\tau) + ad\sigma(s_1 s_2^\tau) + bc\sigma(s_1^\tau s_2) + bd\sigma(s_2 s_2^\tau) + a\sigma(s_1 n_2) \\ + b\sigma(s_2 n_2) + c\sigma(s_1^\tau n_1) + d\sigma(s_2^\tau n_1) + \sigma(n_1 n_2) \end{array} \right]^2 },
\end{aligned}
\tag{34}
$$

where $\sigma(\cdot \; \cdot)$ is the covariance between the signals in parentheses, and as before, delayed signals are indicated with a superscript τ. For the H_0 hypothesis, the smallest eigenvalue is

$$
\begin{aligned}
\lambda_1 = 0.5 &\left[(b^2+d^2)\sigma_{s2} + \sigma_{n1} + \sigma_{n2} + 2b\sigma(s_2 n_1) + 2d\sigma(s_2^\tau n_2) \right] \\
- 0.5 &\sqrt{ \left[(b^2-d^2)\sigma_{s2} + \sigma_{n1} - \sigma_{n2} + 2b\sigma(s_2 n_1) - 2d\sigma(s_2^\tau n_2) \right]^2 + 4\left[bd\sigma(s_2 s_2^\tau) + b\sigma(s_2 n_2) + d\sigma(s_2^\tau n_1) + \sigma(n_1 n_2) \right]^2 }.
\end{aligned}
\tag{35}
$$

As in the previous section, the eigenvalue for the H_0 hypothesis does not reduce to zero because the new covariance terms introduced by the delays in the model and the uncorrelated noise do not cancel. Fortunately, switching from the H_0 to the H_1 hypothesis causes an increase in the smallest eigenvalue that may be thresholded for detection.

D. Threshold selection guideline for a given false alarm rate

The previous sections showed how the smallest eigenvalue of the covariance matrix changes with the addition of the signal of interest. In other words, the smallest eigenvalue of the covariance matrix is the test statistic that is used to distinguish between the H_1 and H_0 hypotheses. This new PCA based detector decides on the H_1 hypothesis if the smallest eigenvalue of the covariance matrix exceeds a threshold, γ, or if

$$
T(\mathbf{X}) = \lambda_1 > \gamma.
\tag{36}
$$

The detection performance can be found by using Eqs. (11) and (12). In the energy detector described in Sec. II C, the pdf of the test statistic under each hypothesis is known to be χ^2, which allows for a closed form solution of the threshold value (for a specific false alarm rate). The pdf of the smallest eigenvalue cannot be derived easily. Training data will be used to estimate the pdf of the test statistic under each hypothesis to provide a guideline for threshold selection.

For this paper, a normal distribution will be used as the model for the pdf of both the H_1 and H_0 hypotheses, so

$$p(\lambda_1; H_1) = N\left(\mu_{\lambda1,H1}, \sigma^2_{\lambda1,H1}\right) \text{ where } \mu_{\lambda1,H1} = E\left[\lambda_1^{H1}\right],$$

$$\sigma^2_{\lambda1,H1} = E\left[\left(\lambda_1^{H1} - \mu_{\lambda1,H0}\right)\right],$$

$$p(\lambda_1; H_0) = N\left(\mu_{\lambda1,H0}, \sigma^2_{\lambda1,H0}\right) \text{ where } \mu_{\lambda1,H0} = E\left[\lambda_1^{H0}\right],$$

$$\sigma^2_{\lambda1,H1} = E\left[\left(\lambda_1^{H0} - \mu_{\lambda1,H0}\right)\right], \tag{37}$$

where λ_1^{H0} is the test statistic calculated when the signal of interest is present [Eq. (31) or (34)] and λ_1^{H1} is the test statistic calculated when the signal of interest is not present [Eq. (32) or (35)]. Training data where the signal of interest is not present is collected and used to estimate the mean and variance of the test statistic for the H_0 hypothesis. For a given P_{FA} value, the threshold, γ, is chosen to satisfy the equation

$$P_{FA} = \int_\gamma^\infty p(\lambda_1; H_0)dx = \int_\gamma^\infty \frac{1}{\sqrt{2\pi\sigma^2_{\lambda1,H0}}}$$

$$\times \exp\left[-\frac{1}{2\sigma^2_{\lambda1,H0}}\left(\lambda_1 - \mu_{\lambda1,H0}\right)\right]dx. \tag{38}$$

If training data for the signal of interest present case is available, it can be used to estimate the mean and variance of the test statistic for the H_1 hypothesis. Using these estimates, the P_D for the threshold γ is given as

$$P_D = \int_\gamma^\infty p(\lambda_1; H_1)dx = \int_\gamma^\infty \frac{1}{\sqrt{2\pi\sigma^2_{\lambda1,H1}}}$$

$$\times \exp\left[-\frac{1}{2\sigma^2_{\lambda1,H0}}\left(\lambda_1 - \mu_{\lambda1,H1}\right)\right]dx. \tag{39}$$

IV. SIMULATIONS AND RESULTS

This section performs a parametric study using simulations to examine the algorithm's performance. All the simulations assume an anechoic in-air environment, with the speed of sound of 343 m/s, and a sampling frequency of 44100 Hz. In all cases, the interferer is white, Gaussian noise. Unless otherwise stated, the signal of interest is a series of white, Gaussian noise bursts where the burst duration

is a random value. The data signal is split into 100 samples per frame, and the significant statistic for each frame is calculated as the smallest eigenvalue of the covariance matrix [using Eq. (6)].

A. One signal, one interferer—simultaneous mixing

The first simulation uses the simultaneous mixing model from Sec. III A. This model is only applicable in geometries where the signal of interest and the interferer are the same distance from both sensors. Despite its limited applicability, it is included for its ease of model derivation and threshold selection. For this simulation, the SIR was set to 0 dB and the two signals are mixed using a linear mixing matrix as in Eq. (21). Figure 2(a) shows the signal of interest before mixing, Fig. 2(b) the mixed signal, and Fig. 2(c), the significant statistic. During periods when the signal of interest is absent, the significant statistic drops to nearly zero. In Eq. (26), the significant statistic is exactly equal to zero, but since the sample covariance matrix is used in the simulations instead of the true covariance matrix, the simulations results do not always reduce to exactly zero. For this simulation, any threshold above zero can accurately differentiate between the frames that contain the signal of interest and the frames that do not contain the signal of interest.

B. One signal, one interferer—delayed mixing

While the previous simulation's model is only applicable for specific geometries, the model in Sec. III B is applicable to any source/sensor geometry. For this simulation, the signal of interest and the interferer are mixed according to the model of Eq. (27) with the delays calculated based on the geometry shown in Fig. 3(a), and an SIR of −8 dB. An SIR of −8 dB was the first point at which false alarms began to appear with the PCA technique. The same signal of interest from the previous section was used for this simulation. When compared to Fig. 2(c), during times when the signal of interest is present, the significant statistics has risen from clustering around 0.0175 to around 0.10. During times when the signal of interest is not present, the significant statistic has risen from nearly zero to clustering around 0.05. For this particular case, a threshold value around 0.075 would be
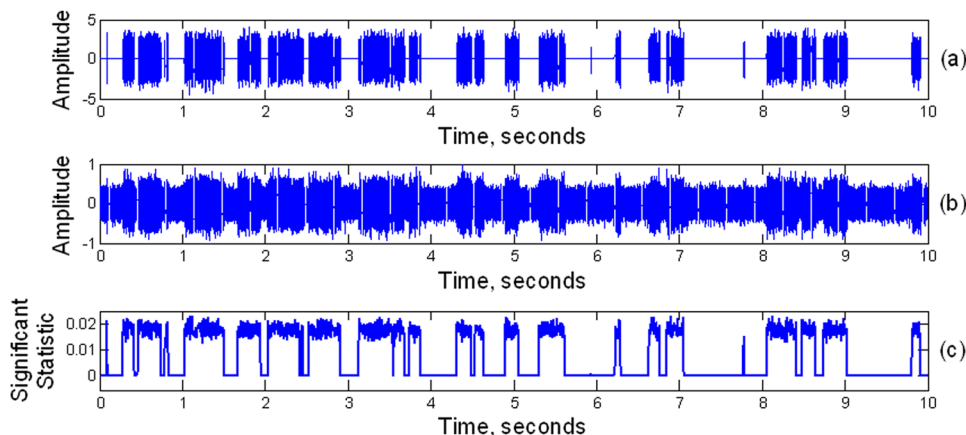

FIG. 2. (Color online) (a) The original signal of interest used for all simulations. (b) The mixed signal, X, containing the signal of interest and the white, Gaussian interferer mixed simultaneously according to Eq. (21). (c) The significant statistic generated using PCA. Using this significant statistic and a user chosen threshold, each frame is categorized as signal present or signal absent. For the simultaneous mixing scheme, when the signal of interest is absent, the significant statistic drops to nearly zero.
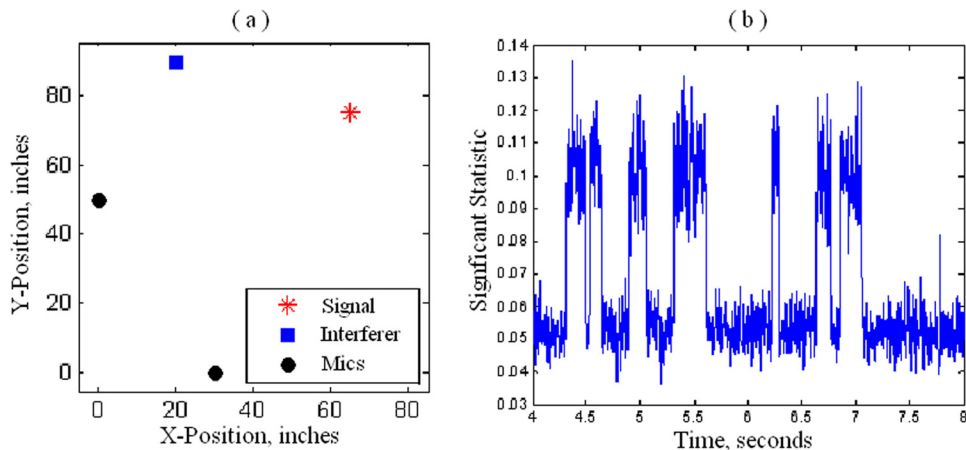
FIG. 3. (Color online) (a) The positioning of the signals and microphones is shown. The TDOAs calculated based on the source/receiver geometry presented are used as the time delays in Eq. (27). (b) The significant statistic generated using PCA. The signal of interest is the same as in Sec. IV A, which allows for comparison with Fig. 2(c).

able to correctly predict the presence of the signal of interest for most frames. This confirms the analysis in Sec. III B that despite not reducing to zero under H_0, this method is still able to detect the presence of the signal of interest (for this particular SIR).

Since the P_{FA} is no longer zero, a ROC curve is used to describe the performance of the algorithm for a variety of P_{FA} conditions. Figure 4 shows the ROC curves generated for the case of -8 dB SIR. The solid line is the ROC curve for the new PCA technique. The P_D is greater than 96% for all values of P_{FA}. Using one second of training data where the signal of interest is absent, the pdf of the significant statistic under the H_0 condition can be modeled as a normal distribution with a mean of 0.0528, and a variance of 0.0056. Using these values and Eq. (16), a P_{FA} of 0.05 is achieved with a threshold value of 0.0620. By using one second of training data where the signal of interest is present, the theoretical best P_D value is calculated from Eq. (17) to be 99.99%. Using
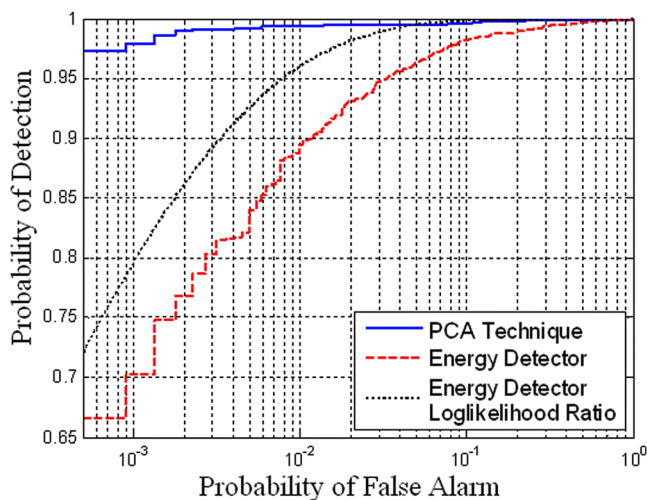

FIG. 4. (Color online) The ROC curves generated in the delayed mixing simulation corresponding to the geometry in Fig. 3(a). The solid line shows the results using the new PCA technique, while the dashed line shows the results using the energy detection test introduced in Sec. II D. The dotted line shows the Neyman−Pearson theoretical best result for the energy detector generated by using Eqs. (19) and (20).

the simulated data, for a 5% P_{FA}, the new PCA technique has a 99.49% P_D. The energy detection test presented in Sec. II B was implemented to provide a comparison with the new technique. The dashed line in Fig. 3(b) shows the ROC curve for the energy detector. For a 5% P_{FA}, the energy detector test has a P_D of 96.47%. The dotted curve in Fig. 4 shows the Neyman-Pearson theoretical best result for the energy detector generated using Eqs. (19) and (20).

1. Varying signal-to-interferer ratio

For the previous simulation, the two signals were at -8 dB SIR. In order to evaluate the new technique's robustness to low SIR, cases were generated where the SIR is varied between 0 dB and -15 dB. The SIR is given by

$$ \text{SIR} = 10 \log_{10} \frac{P_{\text{signal}}}{P_{\text{int}}}, \tag{40} $$

where P_{signal} and P_{int} are the powers of the signal of interest and interferer, respectively. For all cases, the delays for both signals remain constant and are the same delays used in the previous simulation. The series of plots in Fig. 5 illustrate the impact of SIR on performance. For each SIR, the ROC curve is calculated and the P_D at 1%, 5%, and 10% P_{FA} is recorded. For all three reference P_{FA} values, the P_D decreases as SIR decreases. The solid line gives the performance of the new PCA technique. For the 5% P_{FA} reference point, the new algorithm has near perfect performance for SIRs above -8 dB. After that point, the performance decreases as SIR decreases. The dashed lines in Fig. 5 give the performance of the energy detector from Sec. II D. For all SIR and all reference P_{FA}, the new technique outperforms the energy detector. Both the new technique and the energy detector show similar trends in their degradation with respect to SIR. In real world applications, the SIR varies with the environment. For the VAD case where the microphone is very close to the speaker of interest, SIR values above 10 dB are common. For a conference room scenario, the SIR is around 6 to 10 dB. In the more difficult scenario of passive sonar, the SIR values are typically closer to 0 dB and, in the extreme, could be as low as -50 dB.
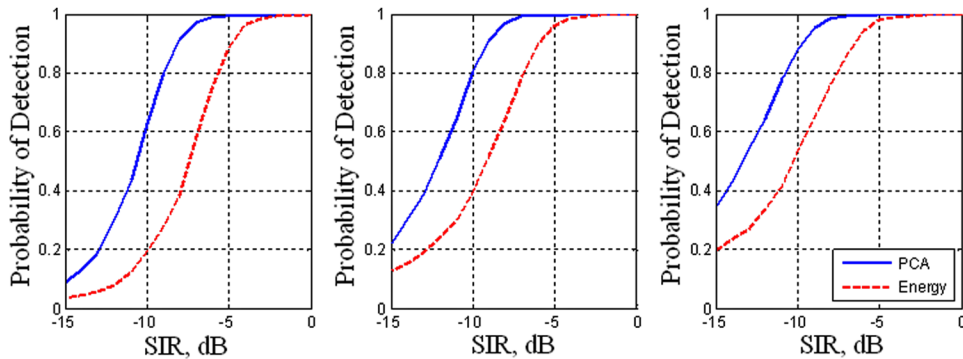
FIG. 5. (Color online) The effect of changing SIR on the algorithm's performance. Each plot frame represents a different reference P_{FA} value. From left to right, the reference values are 1%, 5%, and 10% P_{FA}. The signals are arranged in the geometry indicated in Fig. 3(a). For all cases, the new technique has a better performance than the energy detector.

2. Varying the type of signal of interest

In the simulations above, the signal of interest was a burst of white, Gaussian noise. This simulation investigates the algorithm's performance for a variety of signal types. The five signal types used in this section are named "Gauss burst," "tones," "bandpass," "uniform pdf," and "impulse." The "Gauss burst" is the previously described series of burst of random, white, Gaussian noise. "Tones" is a series of tonal bursts at random frequencies (up to the Nyquist frequency). "Bandpassed" is a series of bursts of 500 Hz wide filtered white noise. The center frequencies for these bursts are the same as the random frequencies used for the tone signal. The "uniform pdf" case is a series of bursts of a uniformly distributed signal with zero mean. Finally, the "impulse" signal is a series of impulses. In order to replicate the ringdown characteristic of sensors, the impulse decays exponentially.

Figure 6 shows the performance of the new technique for each of the signal trials. Figure 6(a) shows the ROC curves generated when the SIR is −5 dB and no additional noise is added to the system. Figure 6(b) shows the performance for the various signal types as SIR decreases. For the varying SIR case, the ROC curve is calculated at each SIR and the P_D at 5% P_{FA} is recorded. In both plots, the performance for the tone, Gaussian burst, and uniform pdf cases are

nearly identical. The bandpassed case has a lower performance. At 0 dB SIR case, the new technique's performance for the bandpassed case is nearly perfect, and then decreases for all decreasing SIR. The impulse case has the worst performance of all signal types examined. The P_D never achieves a value above 45% for all SIRs. Its performance is only slightly better than a random guess. Impulses have very small variances; in this example the sample variance of one impulse is 0.0169, while the Gaussian interferer has a sample variance of 0.999. The change in variance of the joint pdf due to the addition of the impulse to the Gaussian interferer is negligible. Since this method relies on detecting the increase in the variance of the joint pdf, it fails when detecting impulses.

3. New PCA method for voice activity detection

The previous sections presented the general case of detecting an unknown acoustic signal in the presence of an interferer. The signal of interest for the previous sections was a series of random tone bursts (other signals were also investigated). This section addresses the specific case where the unknown signal is a voice signal. For this simulation, the voice signal and the interferer are mixed according to the model of Eq. (27) with the delays calculated based on the geometry shown in Fig. 3(a). The voice
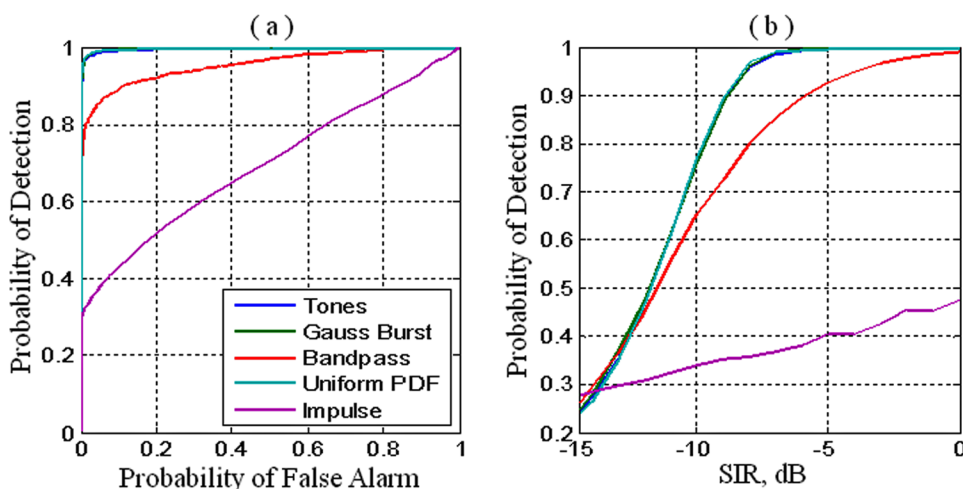

FIG. 6. (Color online) The results of changing the form of the signal of interest on the algorithm's performance. The ROC curve for the −5 dB SIR is shown in (a). For the reference P_{FA} values of 5%, the P_D as a function of SIR is shown in (b). For the tone, random signal, and uniform pdf cases, the performance lines lie of top of one another, indicating nearly identical performance.
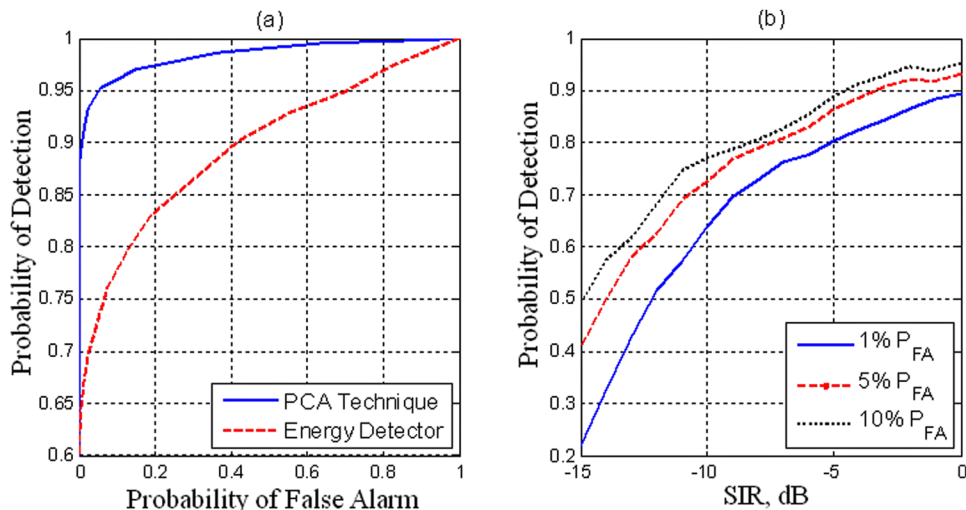
FIG. 7. (Color online) (a) The ROC curves generated in the delayed mixing VAD simulation corresponding to the geometry in Fig. 3(a). The solid line shows the results using the new PCA technique, while the dashed line shows the results using the energy detection test introduced in Sec. II D. (b) The effect of changing SIR on the algorithm's performance. The ROC curves generated for VAD simulations with 0, -5, and -10 dB SIR.

signal and interferer have equal power. Figure 7(a) shows the ROC curve for the new PCA technique. For a 5% P_{FA}, this new technique has a 94.31% P_D. The dashed curve in Fig. 7(a) shows the ROC curve for the energy detector. For a 5% P_{FA}, the statistical VAD algorithm has a P_D of 70.69%. Figure 7(b) illustrates the impact of changing the SIR on the new technique's performance as a VAD algorithm. As expected, decreasing SIR causes a decrease in the algorithm's performance. The performance of the technique as a VAD algorithm differs from the trend seen in the previous tonal case. This is due to the nonstationary characteristics of speech. The sample variance of the tone in the tonal case is nearly constant for all frames, while the voice signal has a variance that is a function of time. As a result, the pdf of the significant statistic under the H_1 and H_0 hypotheses overlap more in the VAD case than in the tonal case, which leads to lower P_D values for a given P_{FA}.

C. One signal, one interferer with noise

For this series of simulations, the signal of interest and the interferer are mixed according to the model of Eq. (33), which includes the addition of uncorrelated noise. The SNR is

$$ \text{SNR} = 10 \log_{10} \frac{P_{\text{signal}}}{P_{\text{noise}}}, \tag{41} $$

where P_{signal} and P_{noise} are the powers of the signal of interest and noise, respectively. This simulation evaluates the new technique's robustness to decreasing SNR. The two signals have equal power and the delays were the same as calculated in Sec. IV B. The series of plots in Fig. 8 illustrate the impact of SNR on the technique's performance. For each SNR, the ROC curve is calculated and the P_D at 1%, 5%, and 10% P_{FA} is recorded. For all three reference P_{FA} values, the new algorithm maintains nearly perfect performance for high SNR, and then the performance decreases as SNR decreases. The trend seen in this simulation is similar to the trend seen in the varying SIR case. The dashed lines in the plots of Fig. 8 show the performance of the energy detector, which is worse for all reference P_{FA} and SNR. For each reference P_{FA} value, the PCA technique and the energy detector show similar trends in the performance degradation as a function of SNR. For real world applications, the performance of this algorithm will depend highly the noise in the data acquisition system, such as the sensitivity of the microphone systems used, the quantization from analog to digital, and the electrical noise in the system.
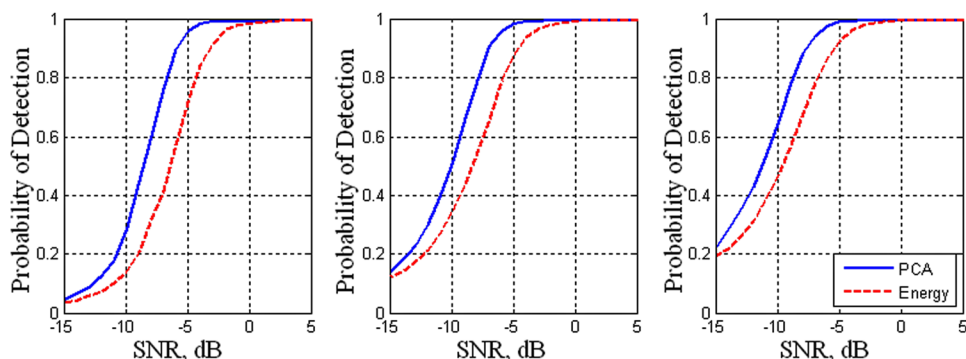

FIG. 8. (Color online) The algorithms performance as a function of varying SNR. Each plot represents a different reference P_{FA} value. From left to right, the reference values are 1%, 5%, and 10% P_{FA}. The signals are arranged in the geometry indicated in Fig. 3(a). For all cases, the new technique has a better performance than the energy detector.

V. CONCLUSIONS

This paper introduced a new method for emergent acoustic signal detection based on PCA. The new method has the capability to detect a signal consisting of bursts of white, Gaussian noise in the presence of a Gaussian interferer for a variety of conditions. It was also shown that a variety of other signals could be detected using this new technique, including tone bursts, bandpassed signals, and signals with a uniform pdf. Existing algorithms detect the emergent signal by detecting changes in the pdf of the mixed signals. The new method uses PCA to detect the change in variance along the projection direction corresponding to an uncorrelated source. Simulations showed the capabilities of the new algorithm in detecting emergent acoustic signals in a variety of conditions. Even with high SIR and in the presence of uncorrelated background noise, the new algorithm was able to detect the emergent acoustic signal of interest. For both varying SIR and SNR, the new algorithm outperformed the standard energy detector. The method introduced is easy to implement and requires no *a priori* information about the signal of interest.

For future work, the effect of multiple sources and multiple interferers must be investigated. In addition, experiments are necessary to verify the algorithm's applicability to real world situations, especially the affect of reverberation on the algorithm. The multiple paths from the source to the microphones cause delayed versions of the signal of interest to appear in several subsequent time frames. The echoes of the signal of interest may not reach the microphone during the same processing frame as the direct path. Since the reverberation is coherent with the signal of interest, it will increase the eigenvalue along the signal of interest's subspace. Therefore, the significant statistic is increased by the presence of echoes. For frames where the direct path of the signal of interest is present along with the echoes, this increase in the significant statistic causes an increase in the P_D. However, in frames where the direct path of the signal of interest is not present and the echoes are from a previous frame, the increase in the significant statistic causes an increase in the P_{FA}. Further model development is necessary to quantify the effect of reverberation on the algorithm's performance.

[1] R. W. Crow and S. C. Schwartz, "Quickest detection for sequential decentralized decision systems," IEEE Trans. Aerosp. Electron. Syst. **32**(1), 267–283 (1996).

[2] H. Kim, B. L. Rozovskii, and A. G. Tartakovsky, "A nonparametric multichart CUSUM test for rapid detection of DOS attacks in computer networks," Int. J. Comput. Inf. Sci. **2**(3), 149–158 (2004).

[3] D. Ramakrishnan and J. Krolik, "Adaptive radar detection in double nonstationary autoregressive Doppler spread clutter," IEEE Trans. Aerosp. Electron. Syst. **45**(2), 484–501 (2009).

[4] R. E. Bethel, G. J. Paras, E. J. Hatfalvi, and J. M. Skora, "Rate aided multisignal time delay detection and tracking," IEEE Trans. Aerosp. Electron. Syst. **31**(3), 1019–1042 (1995).

[5] M. Basseville, "Edge detection using sequential methods for change in level – Part II: Sequential detection of change in mean," IEEE Trans. Acoust., Speech, Signal Process. **29**(1), 32–50 (1981).

[6] D. A. Abraham, "Analysis of a signal starting time estimator based on the page test statistic," IEEE Trans. Aerosp. Electron. Syst. **33**(4), 1225–1234 (1997).

[7] K. G. Sabra and S. Huston, "Passive structural health monitoring of a high-speed naval ship from ambient vibrations," J. Acoust. Soc. Am. **129**(5), 2991–2999 (2001).

[8] E. Tatara and A. Cinar, "Interpreting ECG data by integrating statistical and artificial intelligence tools," IEEE Eng. Med. Biol. Mag. **21**(1), 36–41 (2002).

[9] J. Park, S. Lee, and M. Jeon, "Atrial fibrillation detection by heart rate variability in Poincare plot," Biomed. Eng. Online, **8**, 38 (2009).

[10] R. K. Rogers, S. P. Collins, M. C. Kontos, M. Zuber, P. Arand, and A. D. Michaels, "Diagnosis and characterization of left ventricular hypertrophy by computerized acoustic cardiography, brain natriuretic peptide, and electrocardiography," J. Electrocardiol. **41**(6), 518–525 (2008).

[11] R. J. Stahl and P. K. Willet, "Quickest detection of a tonal burst," IEEE Trans. Signal Process. **45**(8), 2037–2047 (1997).

[12] T. Oskiper and H. V. Poor, "Quickest detection of a random signal in background noise using a sensor array," EURASIP J. Appl. Signal Process. **1**, 13–24 (2005).

[13] B. Chen and P. Willett, "Detection of hidden Markov model transient signals," IEEE Trans. Aerosp. Electron. Syst. **36**(4), 1253–1268 (2000).

[14] D. K. Mellinger, S. W. Martin, R. P. Morrissey, L. Thomas, and J. J. Yosco, "A method for detecting whistles, moans, and other frequency contour sounds," J. Acoust. Soc. Am. **129**(6), 4055–4061 (2011).

[15] A. Thode, J. Skinner, P. Scott, J. Roswell, J. Straley, and K. Folkert, "Tracking sperm whales with a towed acoustic vector sensor," J. Acoust. Soc. Am. **128**(5), 2681–2694 (2010).

[16] E. Conte, A. De Maio, and C. Galdi, "Signal detection in compound-Gaussian noise: Neyman-Pearson and CFAR detectors," IEEE Trans Signal Process. **48**(2), 419–428 (2000).

[17] P. M. Schultheiss, "Passive sonar detection in the presence of interference," J. Acoust. Soc. Am. **43**(3), 418–425 (1968).

[18] R. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas Propag. **34**(3), 276–280 (1986).

[19] A. Davis, S. Nordholm, and R. Togneri, "Statistical voice activity detection using low-variance spectrum estimation and an adaptive threshold," IEEE Trans. Audio, Speech, Lang. Process. **14**(2), 412–424 (2006).

[20] J. M. Gorriz, J. Ramirez, J. C. Segura, and C. G. Puntonet, "An effective cluster-based model for robust speech detection and speech recognition in noisy environments," J. Acoust. Soc. Am. **120**(1), 470–481 (2006).

[21] E. Nemer, R. Goubran, and S. Mahmoud, "Robust voice activity detection using higher-order statistics in the LPC residual domain," IEEE Trans Speech Audio Process. **9**(3), 217–231 (2001).

[22] J. Ramirez, J. M. Gorriz, and J. C. Seguar, "Statistical voice activity detection based on integrated bispecturm likelihood ratio tests for robust speech recognition," J. Acoust. Soc. Am. **121**(4), 2946–2958 (2007).

[23] H. L. Van Trees, *Optimum Array Processing, Detection, Estimation, and Modulation Theory, Part III* (Wiley, New York, 2001), Chap. 8.

[24] T. D. Wickens, *Elementary Signal Detection Theory* (Oxford University Press, Oxford, 2002), Chaps. 1 and 2.

[25] D. Ying, Y. Shi, F. Soong, J. Dang, and X. Lu, "A robust voice activity detection based on noise eigenspace projection," Lect. Notes Comput. Sci. **4274**, 76–86 (2002).

[26] J. M. Górriz, J. Ramírez, C. G. Puntonet, E. W. Lang, and K. Stadlthanner, "Independent component analysis applied to voice activity detection," Lect. Notes Comput. Sci. **3991**, 234–241 (2006).

[27] W. Puyati and A. Walairacht, "Efficiency improvement for unconstrained face recognition by weightening probability values of modular PCA and wavelet PCA," *10th International Conference on Advanced Communication Technology* (ICACT, 2008 Gangwon-Do). Vol. 2, pp. 1449–1453.

[28] B. Qiu, V. Prinet, E. Perrier, and O. Monga, "Multi-block PCA method for image change detection," *Proceedings of the 12th International Conference on Image Analysis and Processing* (2003 Mantova), pp. 385–390.

[29] D. Brauckhoff, K. Salamatian, and M. May, "Applying PCA for traffic anomaly detection: Problems and solutions," *IEEE INFOCOM 2009* (2009 Rio de Janeiro), pp. 2866–2870.

[30] D. X. Tien, K. Lim, and L. Jun, "Comparative study of PCA approaches in process monitoring and fault detection," *30th Annual Conference of IEEE Industrial Electronics Society* (IECON, 2004, Roanoke), Vol. 3, pp. 2594–2599.

[31] T.-W. Lee, *Independent Component Analysis, Theory and Applications* (Kluwer Academic, Boston, MA, 1998), Chap. 2.

[32] I. T. Jolliffe, *Principal Component Analysis. Springer Series in Statistics* (Springer, New York, 2002), Chaps. 1, 3, 5, and 6.

[33] J. V. Stone, *Independent Component Analysis: A Tutorial Introduction* (MIT Press, Cambridge, MA, 2004), Chaps. 6 and 7.

[34] H. Urkowitz, "Energy detection of unknown deterministic signals," Proc. IEEE **55**, 523–531 (1967).

[35] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, (Prentice-Hall, Upper Saddle River, NJ, 1998), Chaps. 3, 5, and 8.

[36] E. Hoppe and M. Roan, "Non-linear, adaptive array processing for acoustic interference suppression," J. Acoust. Soc. Am. **125**(6), 3835–3843 (2009).