HAL open science

Algorithms and applications for estimating the standard deviation of AWGN when observations are not signal-free

Dominique Pastor, Asmaa Amehraye

▶ To cite this version:

Dominique Pastor, Asmaa Amehraye. Algorithms and applications for estimating the standard deviation of AWGN when observations are not signal-free. Journal of Computers (JCP), 2007, 2 (7), pp.1 - 10. hal-02136597

HAL Id: hal-02136597
https://hal.science/hal-02136597v1

Submitted on 22 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Algorithms and Applications for Estimating the Standard Deviation of AWGN when Observations are not Signal-Free

Dominique Pastor, Asmaa Amehraye
GET - ENST Bretagne, CNRS TAMCIC (UMR 2872), Technopôle de Brest Iroise,
CS 83818, 29238 BREST Cedex, FRANCE
Email: {dominique.pastor, asmaa.amehraye}@enst-bretagne.fr

***Abstract*— Consider observations where random signals are randomly present or absent in independent and additive white Gaussian noise (AWGN). By using a recently established limit theorem, we introduce a new estimator for the estimation of the noise standard deviation when the signals are less present than absent and have unknown probability distributions.**

The bias, the consistency and the minimum attainable mean square estimation error of the estimator we propose are still unknown. However, the experimental results that are presented are very promising. First, when the Minimum-Probability-of-Error decision scheme for the non-coherent detection of modulated sinusoidal carriers in independent AWGN is tuned with the outcome of our estimator instead of the true value of the noise standard deviation, the Binary Error Rate tends to the optimal error probability when the number of observations is large enough. Second, given some speech signal corrupted by independent AWGN, our estimator can be used to estimate the noise standard deviation so as to adjust the standard Wiener filtering of the noisy speech. The objective performance measurements obtained by so proceeding are very close to those achieved when the Wiener filtering is tuned with the true value of the noise standard deviation.

***Index Terms*— Binary hypothesis testing, decision, estimation, likelihood theory, multivariate normal distribution, speech denoising.**

I. Introduction

In many signal processing applications, observations can be modelled by d-dimensional random vectors that result from the random presence of signals in independent and additive white Gaussian noise (AWGN).

Often, very little is known about the signals or about most of their describing parameters. The probability distributions of the signals are sometimes partially or definitely unknown. This issue is typically met in Electronic (Warfare) Support Measure (ESM) systems faced with non-cooperant communications. This is also the case with sonar systems that receive signals resulting from noise generated by motors and hull vibrations transmitted through a fluctuating environment. Echoes received by a radar system are not always easy to model because they derive from a convolution between a known transmitted pulse and an unknown environment.

In addition to our lack of prior knowledge about the signals, the noise standard deviation may be unknown as well. In order to process the observations, it can then be necessary to estimate it.

For instance, Constant False Alarm Rate (CFAR) systems, standardly used in radar processing, perform an estimate of the noise standard deviation in order to detect radar targets with a false alarm rate close to some prespecified value. Basically, CFAR systems select observations considered as signal-free and called "reference cells". The estimation of the noise standard deviation is then achieved on the basis of these reference cells. The design of a CFAR system is no easy task (see [1], [2]).

A second example is the denoising of speech signals corrupted by independent AWGN. Indeed, most standard filtering techniques must be adjusted with an estimate of the noise standard deviation when the latter is unknown. A popular approach involves using a Voice Activity Detector (VAD). The VAD decides, with some error rate, that some periods of time are signal-free and the noise standard deviation is then estimated on the basis of these periods of time. Another approach is to achieve an estimate of the noise standard deviation by computing the smallest eigenvalues of the noisy speech autocorrelation matrix. Depending on the matrix size, the computation of the eigenvalues may prove unstable.

In contrast with such approaches mostly based on the physics of the problem, we hereafter propose an estimator relying on statistics only. This estimator is based on a recent theoretical result established in [3]. A simplified version of this result will however be sufficient to derive our estimator.

II. Paper organization

After introducing some notations and pieces of terminology, section III presents the theoretical result, namely proposition 3.1, from which the estimator proposed in this paper is derived. Proposition 3.1 and the more general

This paper is an extended version of "Estimating the Standard Deviation of Some Additive White Gaussian Noise on the Basis of Non Signal-Free Observations" by D. Pastor, which appeared in the Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006, Toulouse, France, May, 2006. © 2006 IEEE.

result [3, Theorem 1] concern any sequence of mutually independent d-dimensional real random observations where signals, with unknown probability distributions but with norms larger than or equal to some known value, are randomly present or absent with unknown probabilities of presence less than or equal to one half in independent AWGN.

Our estimator is then presented in section IV. It applies to sets of non signal-free observations where signals have unknown probability distributions and unknown probabilities of presence less than or equal to one half. No prior knowledge on a lower bound for the norms of the signals is any longer required.

The bias, the consistency and the mean square error of this estimator are still unknown. Nevertheless, several experimental results are presented in section V. These experimental results concern two-dimensional observations where the signals are uniformly distributed on the circle centred at the origin with known radius α. The two components of every signal can be regarded as the in-phase and quadrature components of a modulated sinusoidal carrier with amplitude α and phase uniformly distributed in $[0, 2\pi]$.

An application to speech denoising is also presented. Given some speech signal corrupted by independent AWGN, section VI describes how the estimator introduced in this paper makes it possible to estimate the noise standard deviation and adjust the standard Wiener filtering of the noisy speech signal.

Section VII concludes this paper by suggesting several theoretical developments, extensions to speech processing as well as applications in radar processing and electronic warfare.

III. Theoretical background

In what follows, only one probability space $(\Omega, \mathcal{M}, \mathrm{P})$ is considered and every random vector or variable is assumed to be defined for every $\omega \in \Omega$ by setting this random vector or variable to 0 on any negligible subset where it could be undefined. As usual, we write (a-s) for "almost surely".

Let $\mathcal{S}$ stand for the set of all the sequences of d-dimensional real random vectors. Given a positive real value σ_0, an element $X = (X_k)_{k\in\mathbb{N}}$ of $\mathcal{S}$ will be called a d-*dimensional white Gaussian noise* (WGN) with standard deviation σ_0 if the random vectors X_k, $k = 1, 2, \ldots$, are mutually independent and identically Gaussian distributed with null mean vector and covariance matrix $\sigma_0^2 \mathbf{I}_\mathrm{d}$ where $\mathbf{I}_\mathrm{d}$ stands for the $\mathrm{d} \times \mathrm{d}$ identity matrix. For the sake of brevity, we will henceforth write *independent* instead of *mutually independent*.

We define the *minimum amplitude* of an element $S = (S_k)_{k\in\mathbb{N}}$ of $\mathcal{S}$ as the supremum $\mathsf{a}(S)$ of the set of those $\alpha \in [0, \infty]$ such that, for every natural number k, $\|S_k\|$ is larger than or equal to α (a-s):

$$\mathsf{a}(S) = \sup\left\{\alpha \in [0, \infty] : \forall k \in \mathbb{N}, \|S_k\| \geq \alpha \text{ (a-s)}\right\}. \tag{1}$$

If f is some map of $\mathcal{S}$ into $\mathbb{R}$, we say that the limit of $f(S)$ is $\ell \in \mathbb{R}$ when $\mathsf{a}(S)$ tends to ∞ and write that $\lim_{\mathsf{a}(S)\to\infty} f(S) = \ell$ if, for any positive real value η, there exists some $\alpha_0 \in (0, \infty)$ such that, for every $\alpha \geq \alpha_0$ and every $S \in \mathcal{S}$ with $\mathsf{a}(S) \geq \alpha$, we have $|f(S) - \ell| \leq \eta$.

Given $\nu \in [0, \infty)$, let $L^\nu(\Omega, \mathbb{R}^\mathrm{d})$ stand for the set of those d-dimensional real random vectors Y such that $\mathbb{E}[\|Y\|^\nu] < \infty$. In the sequel, we will encounter the set $\ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^\mathrm{d}))$ of those elements $S = (S_k)_{k\in\mathbb{N}}$ of $\mathcal{S}$ such that $S_k \in L^\nu(\Omega, \mathbb{R}^\mathrm{d})$ for every $k \in \mathbb{N}$ and $\sup_{k\in\mathbb{N}} \mathbb{E}[\|S_k\|^\nu]$ is finite.

In what follows, ${}_0\mathrm{F}_1$ is the generalized hypergeometric function (see [4, p. 275]). Given $\rho \in [0, \infty)$, $\xi(\rho)$ is the unique positive solution for x in the equation

$${}_0\mathrm{F}_1(\mathrm{d}/2; \rho^2 x^2/4) = \exp(\rho^2/2). \tag{2}$$

Given $q \in [0, \infty)$, Υ_q stands for the map defined for every $x \in [0, \infty)$ by

$$\Upsilon_q(x) = \int_0^x t^{q+\mathrm{d}-1} \exp(-t^2/2) dt. \tag{3}$$

The following result is a particular case of a more general limit theorem established in [3]. The statement considered in the present paper is sufficient with respect to our purpose. Given any random vector Y and any real number τ, $I(\|Y\| \leq \tau)$ stands for the indicator function of the event $\{\|Y\| \leq \tau\}$.

Proposition 3.1: Let $Y = (Y_k)_{k\in\mathbb{N}}$ be some element of $\mathcal{S}$ such that, for every $k \in \mathbb{N}$, $Y_k = \varepsilon_k S_k + X_k$ where $S = (S_k)_{k\in\mathbb{N}}$, $X = (X_k)_{k\in\mathbb{N}}$ and $\varepsilon = (\varepsilon_k)_{k\in\mathbb{N}}$ are an element of $\mathcal{S}$, some d-dimensional WGN with standard deviation σ_0 and a sequence of random variables valued in $\{0, 1\}$ respectively.

Assume that

- **(A1)** for every $k \in \mathbb{N}$, S_k, X_k and ε_k are independent;
- **(A2)** the random vectors Y_k, $k \in \mathbb{N}$, are independent;
- **(A3)** the probabilities of presence $\mathrm{P}(\{\varepsilon_k = 1\})$, $k \in \mathbb{N}$, are less than or equal to one half and the random variables ε_k, $k \in \mathbb{N}$, are independent;
- **(A4)** there exists some $\nu \in (0, \infty]$ such that $S \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^\mathrm{d}))$.

Given two non-negative real numbers r and s such that $0 \leq s < r \leq \nu/2$, some natural number m and any pair (σ, T) of positive real numbers, define the random variable $\Delta_m(\sigma, T)$ by

$$\Delta_m(\sigma, T) = \left| \frac{\sum_{k=1}^m \|Y_k\|^r I(\|Y_k\| \leq \sigma T)}{\sum_{k=1}^m \|Y_k\|^s I(\|Y_k\| \leq \sigma T)} - \sigma^{r-s} \frac{\Upsilon_r(T)}{\Upsilon_s(T)} \right|. \tag{4}$$

Then, σ_0 is the unique positive real number σ such that, for every $\beta_0 \in (0, 1]$,

$$\lim_{\mathsf{a}(S)\to\infty} \left\| \overline{\lim_m} \Delta_m(\sigma, \beta\xi(\mathsf{a}(S)/\sigma)) \right\|_\infty = 0 \tag{5}$$

uniformly in $\beta \in [\beta_0, 1]$.

In the foregoing statement, U models a sequence of observations where, for every given $k \in \mathbb{N}$, S_k stands for some possible random signal and ε_k models the possible occurrence of S_k in the background of AWGN modelled by X.

In proposition 3.1, neither the probabilities of presence nor the signals are required to be identically distributed. Moreover, it is worth mentioning that the convergence criterion exhibited by this result derives from a corollary of Kolmogorov's classical strong limit theorem and not from usual generalizations of the central limit theorem such as the Lindeberg and the Lyapounov theorems.

Note that, for $\tau \in [0, \infty)$, the ratio

$$\sum_{k=1}^{m} \|Y_k\|^r I(\|Y_k\| \leq \tau) / \sum_{k=1}^{m} \|Y_k\|^s I(\|Y_k\| \leq \tau)$$

is defined everywhere for the following reason. Let $x_1, x_2, \ldots, x_m$ be m real numbers. If there exists at least one natural number $k \in \{1, \ldots, m\}$ such that $x_k \neq 0$, the finiteness of the ratio $\sum_{k=1}^{m} |x_k|^r / \sum_{k=1}^{m} |x_k|^s$ is trivial. Since $r > s$, the definition of this ratio is then extended by continuity by setting $\sum_{k=1}^{m} |x_k|^r / \sum_{k=1}^{m} |x_k|^s = 0$ if $(x_1, \ldots, x_m) = (0, \ldots, 0)$.

Proposition 3.1 concerns positive solutions of Eq. (5) only because $\sigma = 0$ trivially satisfies this equation regardless of the specific convergence involved. Straightforwardly, Eq. (5) is also satisfied for all $\sigma \in [0, \infty)$ when $r = s \geq 0$. This explains why it is assumed that $r > s \geq 0$.

IV. The Essential Supremum, the Modified and the Complex Estimates

In this section, the notations used so far are kept with exactly the same meaning. On the basis of the previous theoretical result, we present several algorithms for the estimation of the noise standard deviation. We begin with an algorithm already introduced in [3]. This algorithm requires prior knowledge of the minimum amplitude of the signals. We then propose a method that does not need such prior knowledge. This method is mainly heuristic. We particularize it to the case of two-dimensional real random vectors, or equivalently, complex random variables. This particular case will serve to address the applications treated in sections V and VI.

A. The Essential Supremum Estimate

According to Eq. (5), given $\beta_0 \in (0, 1]$ and some positive real value η, there exists $\alpha_0 \in (0, \infty)$ that satisfies the following property: for every real number α larger than or equal to α_0, every element S of $\ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^{\mathrm{d}}))$ whose minimum amplitude is larger than or equal to α, and every $\beta \in [\beta_0, 1]$, the absolute value of the difference between the random variable

$$\frac{\sum_{k=1}^{m} \|Y_k\|^r I(\|Y_k\| \leq \sigma\beta\xi(\alpha/\sigma))}{\sum_{k=1}^{m} \|Y_k\|^s I(\|Y_k\| \leq \sigma\beta\xi(\alpha/\sigma))}$$

and

$$\sigma^{r-s} \frac{\Upsilon_r(\sigma\beta\xi(\alpha/\sigma))}{\Upsilon_s(\sigma\beta\xi(\alpha/\sigma))}$$

does not exceed η (a-s) when the sample size or number of observations m is large enough.

In many applications, the signals S_k, $k \in \mathbb{N}$, have "finite energy" in the sense that the second-order moments $\mathbb{E}[\|S_k\|^2]$, $k \in \mathbb{N}$, are finite. In what follows, this "finite energy" hypothesis is made. Therefore, with respect to assumption **(A4)**, we consider the case $\nu = 2$ and assume that $S \in \ell^\infty(\mathbb{N}, L^2(\Omega, \mathbb{R}^{\mathrm{d}}))$. Since the values r and s must verify the inequalities $0 \leq s < r \leq 1$, our rather natural choice is $r = 1$ and $s = 0$.

Given m observations $Y_1, \ldots, Y_m$ for which the minimum amplitude $\mathsf{a}(S)$ is known to be larger than or equal to some known α, if we set $L \in \mathbb{N}$ and $\beta_\ell = \ell/L$ for every $\ell \in \{1, \ldots, L\}$, the foregoing suggests estimating σ_0 by a possibly local minimum of

$$\sup_{\ell \in \{1, \ldots, L\}} \left| \frac{\sum_{k=1}^{m} \|Y_k\| I(\|Y_k\| \leq \beta_\ell \sigma \xi(\alpha/\sigma))}{\sum_{k=1}^{m} I(\|Y_k\| \leq \beta_\ell \sigma \xi(\alpha/\sigma))} - \sigma\Theta(\beta_\ell \xi(\alpha/\sigma)) \right| \tag{6}$$

where Θ is defined for every $x \in [0, \infty)$ by

$$\Theta(x) = \frac{\Upsilon_1(x)}{\Upsilon_0(x)} = \frac{\int_0^x t^{\mathrm{d}} \exp(-t^2/2) dt}{\int_0^x t^{\mathrm{d}-1} \exp(-t^2/2) dt}. \tag{7}$$

Following the terminology proposed in [3], the estimate obtained by minimizing Eq. (6) is an Essential Supremum Estimate (ESE) of the noise standard deviation. This name follows from the fact that the essential supremum norm plays an important role in proposition 3.1 and its generalization stated in [3, Theorem 1].

In [3], experimental results are given for the case of independent signals that are two-dimensional random vectors uniformly distributed on a circle centred at the origin with known radius α and that have their probabilities of presence less than or equal to one half. Such random signals model modulated sinusoidal carriers whose amplitudes equal α and whose phases are uniformly distributed in $[0, 2\pi]$. The empirical bias and the empirical Mean-Square Error (MSE) of the ESE remain reasonably good for $\alpha \in \{0, 0.25, 0.5, 0.75, \ldots, 5\}$ and any probability of presence in $\{0.1, 0.2, 0.3, 0.4, 0.5\}$. One conclusion of [3] is therefore that the asymptotic conditions on which [3, Theorem 1] and proposition 3.1 rely, are not so constraining and can certainly be relaxed in practice.

B. Modified and Complex Essential Supremum Estimates

The computation of the ESE requires prior knowledge of a lower bound for the amplitudes of the signals. This lower bound is not always known. Since [3] suggests that the convergence stated by (5) is quite fast, we hereafter

set α to 0 in (6) and, thus, since $\xi(0) = \sqrt{\mathrm{d}}$ (see [5]), estimate σ_0 by a possibly local minimum $\widetilde{\sigma_0}$ of

$$\sup_{\ell \in \{1,\ldots,L\}} \left\{ \left| \frac{\sum_{k=1}^{m} \|Y_k\| I(\|Y_k\| \leq \beta_\ell \sigma \sqrt{\mathrm{d}})}{\sum_{k=1}^{m} I(\|Y_k\| \leq \beta_\ell \sigma \sqrt{\mathrm{d}})} - \sigma \Theta(\beta_\ell \sqrt{\mathrm{d}}) \right| \right\}.$$

Any minimization routine for scalar bounded non-linear functions is suitable. For instance, the experimental results presented in the next sections were obtained with the `MATLAB` routine `fminbnd.m` based on parabolic interpolation (see [6]). This routine is also that employed in [3] to compute the ESE.

The search interval $[\sigma_{\min}, \sigma_{\max}]$ is computed as follows. Sort the observations $Y_1, \ldots, Y_m$, $k = 1, \ldots, m$, by increasing norm. Let $Y_{[k]}$, $k = 1, \ldots, m$, be the resulting sequence. The right endpoint of the search interval is then $\sigma_{\max} = \|U_{[m]}\|/\sqrt{\mathrm{d}}$. As far as the left endpoint is concerned, choose a real number Q close to 1 but less than or equal to $1 - \frac{m}{4(m/2-1)^2}$. A typical choice is $Q = 0.95$, provided that $m \geq 24$. Set $h = 1/\sqrt{4m(1-Q)}$ and $k_{\min} = m/2 - hm$. The left endpoint is then $\sigma_{\min} = \|Y_{[k_{\min}]}\|/\sqrt{\mathrm{d}}$. The reader is asked to refer to [3] for justifications regarding the construction of this search interval.

Regarding the applications considered in the sequel, simulations of the same type as those proposed below show that $\widetilde{\sigma_0}$ is a reasonably good estimate of σ_0. However, we can propose a better estimate of σ_0. This new estimate, henceforth called the Modified ESE and denoted by $\widehat{\sigma_0}$, is computed on the basis of $\widetilde{\sigma_0}$ by setting

$$\widehat{\sigma_0} = \gamma \sqrt{\frac{\sum_{k=1}^{m} \|Y_k\|^2 I(\|Y_k\| \leq \widetilde{\sigma_0}\sqrt{\mathrm{d}})}{\sum_{k=1}^{m} I(\|Y_k\| \leq \widetilde{\sigma_0}\sqrt{\mathrm{d}})}} \tag{8}$$

where γ is some constant chosen empirically with respect to the application. For instance, this constant is different for the two applications addressed in the sequel.

The rationale is the following. Under the assumptions of proposition 3.1 and in the particular case where $\beta = 1$, Eq. (5) means that, when the amplitudes of the signals are larger than or equal to some sufficiently large value α and the sample size is large enough, the random variable

$$\frac{\sum_{k=1}^{m} \|Y_k\|^r I(\|Y_k\| \leq \sigma \xi(\alpha/\sigma_0))}{\sum_{k=1}^{m} \|Y_k\|^s I(\|Y_k\| \leq \sigma \xi(\alpha/\sigma_0))}$$

approximates $\sigma^{r-s} \Upsilon_r(\xi(\alpha/\sigma_0))/\Upsilon_s(\xi(\alpha/\sigma_0))$. For the computation of $\widetilde{\sigma_0}$, we have chosen $r = 1$, $s = 0$, and considered that the convergence stated by Eq. (5) holds true even for small values of α.

Assume now that $S \in \ell^\infty(\mathbb{N}, L^4(\Omega, \mathbb{R}^{\mathrm{d}}))$, which remains reasonable although stronger than the "finite energy" assumption since $\ell^\infty(\mathbb{N}, L^4(\Omega, \mathbb{R}^{\mathrm{d}})) \subset \ell^\infty(\mathbb{N}, L^2(\Omega, \mathbb{R}^{\mathrm{d}}))$. Then, we can choose $r = 2$ and $s = 0$. Still assuming that Eq. (5) is valid whatever the value of α, we consider that σ_0^2 can be approximated, in a certain mathematical sense that remains to be specified, by the random variable

$$\frac{\Upsilon_0(\xi(\alpha/\sigma_0))}{\Upsilon_2(\xi(\alpha/\sigma_0))} \times \frac{\sum_{k=1}^{m} \|Y_k\|^2 I(\|Y_k\| \leq \sigma \xi(\alpha/\sigma_0))}{\sum_{k=1}^{m} I(\|Y_k\| \leq \sigma \xi(\alpha/\sigma_0))}$$

By setting $\alpha = 0$ and taking into account that $\xi(0) = \sqrt{\mathrm{d}}$, the foregoing leads to approximate σ_0 as proposed in Eq. (8) with $\gamma = \sqrt{\Upsilon_0(\sqrt{\mathrm{d}})/\Upsilon_2(\sqrt{\mathrm{d}})}$. When $\mathrm{d} = 2$, the case addressed in the sequel, this constant equals 1.0937. According to the experimental results presented below, it may be necessary to adjust this value with respect to the application.

In practice, observations are often complex random variables, or equivalently, two-dimensional real random vectors. For instance, such observations can be the complex values provided by the standard I and Q decomposition encountered in most receivers in radar, sonar and telecommunication systems. Complex observations can also be simply the outcome of a Discrete Fourier Transform (DFT) as in some radar and sonar systems but also in a great variety of other applications. In section VI, we will consider an application to speech processing where the DFT plays a crucial role.

Because of the important role played by two-dimensional real random vectors and complex random variables in many applications, the Modified ESE when $\mathrm{d} = 2$ will be called the Complex ESE (C-ESE). In the two-dimensional case, note that the expression of Θ simplifies so that

$$\Theta(x) = \frac{\int_0^x t^2 \exp(-t^2/2) dt}{1 - \exp(-x^2/2)}. \tag{9}$$

V. Application to the non-coherent detection of modulated sinusoidal carriers in AWGN

The bias, the consistency and the minimum attainable mean square estimation error of the Modified and Complex ESEs are still unknown. However, we can undertake some experiments making it possible to assess the performance of such estimates. The experiments presented in this section are motivated by the following facts.

We keep the notations used so far. Given any real number x, by thresholding test with threshold height $h \in \mathbb{R}$, we hereafter mean the test $\mathcal{T}_h = I(\|\cdot\| \geq h)$ that assigns 1 to $y \in \mathbb{R}^{\mathrm{d}}$ if $\|y\| \geq h$ and 0, otherwise.

Given any non-negative real number α, let $\widehat{\mathcal{T}}$ be the statistical test $I(\|\cdot\| \geq \widehat{\sigma_0}\xi(\alpha/\widehat{\sigma_0}))$. Given $k \in \mathbb{N}$, the decision of this test is that ε_k is 1 if $\|Y_k\| \geq \widehat{\sigma_0}\xi(\alpha/\widehat{\sigma_0})$ and that ε_k is 0, otherwise.

If the Modified ESE $\widehat{\sigma_0}$ is a reasonably good estimate of the noise standard deviation, the performance of test $\widehat{\mathcal{T}}$ can be expected to approach that of the thresholding test $\mathcal{T}_{\sigma_0\xi(\alpha/\sigma_0)}$ with threshold height $\sigma_0\xi(\alpha/\sigma_0)$.

To detect the presence of any signal with norm larger than or equal to α and prior less than or equal to one half, it follows from [5, Theorem VII.1] that the probability of error $\mathcal{P}_e\{\mathcal{T}_{\sigma_0\xi(\alpha/\sigma_0)}\}$ of $\mathcal{T}_{\sigma_0\xi(\alpha/\sigma_0)}$ satisfies the following inequalities

$$\mathcal{P}_e\{\mathcal{L}\} \leq \mathcal{P}_e\{\mathcal{T}_{\sigma_0\xi(\alpha/\sigma_0)}\} \leq V(\alpha/\sigma_0). \qquad (10)$$

In these inequalities, $\mathcal{P}_e\{\mathcal{L}\}$ stands for the probability of error of the Minimum-Probability-of-Error (MPE) decision scheme $\mathcal{L}$, that is the likelihood ratio test with the smallest possible probability of error among all possible hypothesis binary tests. In Eq. (10), the map V is defined for every $x \in [0, \infty)$ and the reader can refer to [5] for the general expression of this map.

The inequalities in (10) become equalities in the least favourable situation where the signal is uniformly distributed on the sphere centred at the origin with radius α and has prior equal to one half (see [5, Theorem VII.1]). Therefore, in this least favourable case, if the ESE $\widehat{\sigma_0}$ is a good estimate of σ_0, the probability of error $\mathcal{P}_e\{\widehat{\mathcal{T}}\}$ of $\widehat{\mathcal{T}}$ should not significantly exceed $V(\alpha/\sigma_0)$. If the probability of presence of every S_k equals one half, $\mathcal{P}_e\{\widehat{\mathcal{T}}\}$ should even be close to $V(\alpha/\sigma_0)$.

We do not know the theoretical value of $\mathcal{P}_e\{\widehat{\mathcal{T}}\}$. Hence, we estimate this probability of error by the Binary Error Rate (BER) obtained by a Monte-Carlo simulation and compare this BER to $V(\alpha/\sigma_0)$. This Monte-Carlo simulation is carried out in the case of two-dimensional real random observations ($\mathrm{d} = 2$). We then choose some $p \in (0, 1/2]$ and every S_k has a probability of presence equal to p and is uniformly distributed on the circle centred at the origin with radius α. Therefore, the two components of every S_k can be regarded as the in-phase and quadrature components of a sinusoidal carrier. In other words, we consider the "non-coherent detection of modulated sinusoidal carriers", a problem particularly relevant for telecommunication and radar processing (see [7, p. 65]).

In the two-dimensional case, the mathematical expressions of ξ and V simplify. According to [3], for $\mathrm{d} = 2$, we have

$$\xi(\rho) = (1/\rho)\mathrm{I}_0^{-1}(\exp(\rho^2/2)) \qquad (11)$$

and

$$V(\rho) = \frac{1}{2}\exp(-\rho^2/2)\int_0^{\xi(\rho)} t\exp(-t^2/2)\mathrm{I}_0(\rho t)dt + \frac{1}{2}\exp(-\xi(\rho)^2/2), \qquad (12)$$

where I_0 is the zeroth-order modified Bessel function of the first kind (see [8, Eq. 9.6.47, p. 377]).

For the computation of the C-ESE and on the basis of some preliminary tests such as those described below, the constant γ in Eq. (8) is set to 1, the integer part of the constant suggested in the preceding section.

The BER of $\widehat{\mathcal{T}}$ is then computed as follows. Independent trials of m observations each are carried out until two conditions are fulfilled. First, at least M trials must be performed. Inasmuch as the decision about the presence or the absence of signals is made on the observations used for estimating σ_0, the accuracy of the estimate affects m decisions at one go. This effect is then reduced by fixing a minimum number of trials. Second, trials are performed until the total number N_e of errors made by test $\widehat{\mathcal{T}}$ for detecting the presence or the absence of signals is above or equal to some specified number N. If j is the first trial number larger than or equal to M for which the total number of errors N_e becomes larger than or equal to N, the BER of test $\widehat{\mathcal{T}}$ is then defined as the ratio $N_e/(j \times m)$.

The simulation is achieved with $\sigma_0 = 1$. The pre-specified number of errors is $N = 1000$ and the minimum number of trials is $M = 1000$. We choose $L = m$ and $Q = 0.95$ on the basis of preliminary trials. The comparison between the BER of $\widehat{\mathcal{T}}$ and $V(\alpha/\sigma_0)$ is achieved for $p \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ and $A \in \{0, 0.25, 0.5, 0.75, \ldots, 5\}$. The results are those of figures 1, 2, 3 and 4 for different values of m. As expected, the performance of $\widehat{\mathcal{T}}$ yields performance close to that of $\mathcal{T}_{\sigma_0\xi(\alpha/\sigma_0)}$.

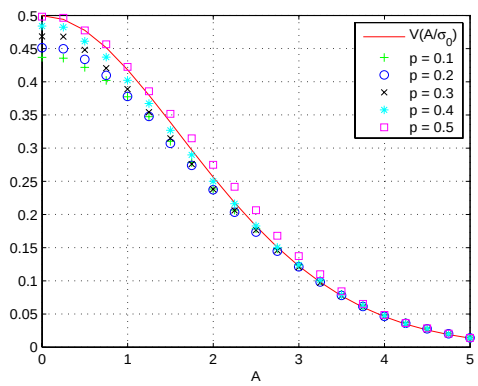


Figure 1. $V(A/\sigma_0)$ vs BERs of $\widehat{\mathcal{T}}$ for $m = 100$

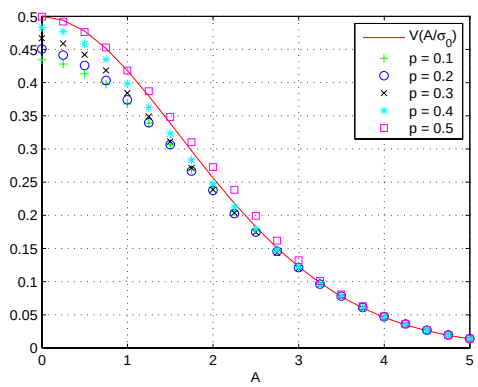


Figure 2. $V(A/\sigma_0)$ vs BERs of $\widehat{\mathcal{T}}$ for $m = 200$

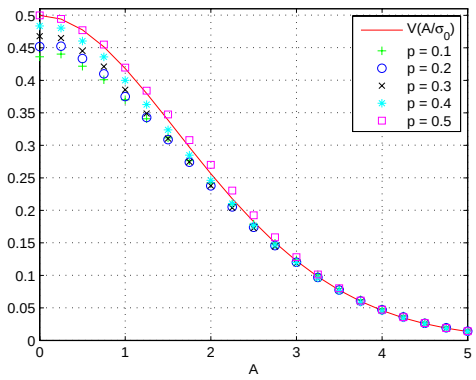


Figure 3. $V(A/\sigma_0)$ vs BERs of $\widehat{\mathcal{T}}$ for $m = 400$

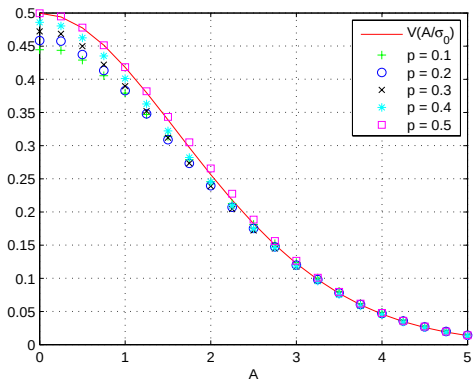


Figure 4. $V(A/\sigma_0)$ vs BERs of $\widehat{\mathcal{T}}$ for $m = 800$

Again for signals that have the same probability of presence p less than or equal to one half and that are uniformly distributed on the circle centred at the origin with radius α, we calculated the absolute value $|\text{BIAS}|$ and the empirical MSE of the C-ESE computed over $M = 1000$ trials. Tables I, II and III below display some results obtained for different values of α and p.

The experimental results presented above were achieved without prior knowledge on the amplitude of the signals. Nevertheless, they are significantly close to those given in [3] where the estimation of the noise standard deviation is performed via the ESE, which does require such prior knowledge.

VI. Application to speech denoising

In this section, we describe how the C-ESE can be used by the standard Wiener filter aimed at denoising speech signals corrupted by independent AWGN with unknown standard deviation. We briefly remind the reader of the basics concerning the standard Wiener filtering of noisy speech signals (see subsection VI-A). Then, in subsection VI-B, under the assumption that noise is white and Gaussian, we explain how to use the C-ESE to estimate the noise standard deviation and adjust the filtering. We complete this section by experimental results (see subsection VI-C).

A. Wiener filtering

Let $s(t)$, $t = 0, \ldots, T-1$, be the samples of some speech signal and suppose that these T samples are corrupted by some noise $x(t)$, $t = 0, 1, \ldots, T-1$, so that the samples of the observed signal are

$$y(t) = s(t) + x(t), t = 0, \ldots, T-1. \tag{13}$$

It is usual to split the T available samples $y(t)$, $t = 0, 1, \ldots, T-1$, into frames of $N = 2^q$ samples each where q is some integer such that $NF_s \approx 20$ms, F_s being the sampling frequency. Frames are generally constructed so that two consecutive ones overlap by one half. The samples of each frame can be weighted. In what follows, for the sake of simplifying the notations, the description of the filtering is made without taking such weighting into account.

For the kth frame, let $s_k(t), x_k(t)$ and $y_k(t)$, $t = 0, 1, \ldots, N-1$, stand for the N samples of the speech signal, noise and the observed noisy speech signal, respectively. We thus have $y_k(t) = s_k(t) + x_k(t)$.

Now, let $Y_k(\nu)$, $S_k(\nu)$ and $X_k(\nu)$, $\nu = 0, \ldots, N-1$, denote the DFT coefficients of $y_k(t)$, $s_k(t)$ and $x_k(t)$, $t = 0, 1, \ldots, N-1$, respectively. For every $\nu = 0, 1, \ldots, N-1$, we have $Y_k(\nu) = S_k(\nu) + X_k(\nu)$.

The standard Wiener filtering of y is based on Malah's decision-directed approach (see [9]). Each frequency component $S_k(\nu)$ is then estimated by $\widehat{S}_k(\nu) = W_k(\nu)Y_k(\nu)$ where $W_k(\nu)$ is the so-called Wiener gain function. This gain is given by

$$W_k(\nu) = \mathcal{R}_k(\nu)/(1 + \mathcal{R}_k(\nu)) \tag{14}$$

where

$$\mathcal{R}_k(\nu) = (1-w)h\left(\zeta_k(\nu) - 1\right) + w\frac{|\widehat{S}_{k-1}(\nu)|^2}{\mathbb{E}\left[|X_k(\nu)|^2\right]} \tag{15}$$

is the so-called decision-directed estimate of the *a priori* Signal to Noise Ratio (SNR) $\mathbb{E}\left[|S_k(\nu)|^2\right]/\mathbb{E}\left[|X_k(\nu)|^2\right]$. In Eq. (15), $\widehat{S}_{k-1}(\nu) = W_{k-1}(\nu)Y_{k-1}(\nu)$ is the νth spectral component of the Wiener denoised speech signal in frame $k-1$; $h(x) = x$ if $x \geq 0$ and $h(x) = 0$ otherwise; $\zeta_k(\nu) = |Y_k(\nu)|^2/\mathbb{E}\left[|X_k(\nu)|^2\right]$ is the *a posteriori* SNR; the weighting factor w will be set to 0.98 as recommended in [9]. The decision-directed approach described by (15) takes into account the current frame, with weight $(1-w)$, and the result of the processing of the previous frame, with weight w.

B. Estimation of the noise standard deviation via the C-ESE when noise is white and Gaussian

When noise is white and Gaussian with standard deviation σ_0, the noise frequency components are independent and identically distributed so that each $X_k(\nu)$ is complex Gaussian distributed with

$$X_k(\nu) \sim \mathcal{CN}(0, 2NA^2\sigma_0^2) \tag{16}$$

where A is some constant depending on the normalization chosen with regard to the Parseval identity. More

TABLE I.

ABSOLUTE VALUE OF THE EMPIRICAL BIAS AND EMPIRICAL MSE OF THE C-ESE WHEN SIGNALS ARE UNIFORMLY DISTRIBUTED ON THE CIRCLE CENTRED AT THE ORIGIN WITH RADIUS α AND HAVE A PROBABILITY OF PRESENCE p EQUAL TO 0.1.

α		0.5	1	1.5	2	2.5	3	3.5	4
m = 100	|BIAS|	0.077	0.07	0.053	0.039	0.04	0.049	0.048	0.064
	MSE	0.036	0.039	0.037	0.032	0.035	0.036	0.034	0.04
m = 200	|BIAS|	0.065	0.051	0.03	0.041	0.054	0.052	0.054	0.056
	MSE	0.023	0.023	0.02	0.022	0.026	0.024	0.022	0.023
m = 400	|BIAS|	0.069	0.058	0.05	0.045	0.054	0.061	0.068	0.062
	MSE	0.016	0.017	0.016	0.015	0.015	0.016	0.017	0.016
m = 800	|BIAS|	0.082	0.07	0.062	0.054	0.059	0.067	0.07	0.074
	MSE	0.014	0.013	0.012	0.011	0.011	0.012	0.012	0.012

TABLE II.

ABSOLUTE VALUE OF THE EMPIRICAL BIAS AND EMPIRICAL MSE OF THE C-ESE WHEN SIGNALS ARE UNIFORMLY DISTRIBUTED ON THE CIRCLE CENTRED AT THE ORIGIN WITH RADIUS α AND HAVE A PROBABILITY OF PRESENCE p EQUAL TO 0.3.

α		0.5	1	1.5	2	2.5	3	3.5	4
m = 100	|BIAS|	0.071	0.024	0.02	0.041	0.039	0.014	0.005	0.01
	MSE	0.037	0.038	0.042	0.046	0.05	0.041	0.038	0.035
m = 200	|BIAS|	0.059	0.005	0.028	0.043	0.033	0.013	0.002	0.017
	MSE	0.024	0.023	0.024	0.028	0.028	0.021	0.018	0.018
m = 400	|BIAS|	0.065	0.017	0.026	0.026	0.02	0.002	0.012	0.014
	MSE	0.017	0.015	0.016	0.017	0.014	0.012	0.01	0.01
m = 800	|BIAS|	0.07	0.033	0.009	0.024	0.002	0.009	0.017	0.023
	MSE	0.013	0.01	0.01	0.01	0.008	0.006	0.005	0.006

TABLE III.

ABSOLUTE VALUE OF THE EMPIRICAL BIAS AND EMPIRICAL MSE OF THE C-ESE WHEN SIGNALS ARE UNIFORMLY DISTRIBUTED ON THE CIRCLE CENTRED AT THE ORIGIN WITH RADIUS α AND HAVE A PROBABILITY OF PRESENCE p EQUAL TO 0.5.

α		0.5	1	1.5	2	2.5	3	3.5	4
m = 100	|BIAS|	0.056	0.02	0.124	0.193	0.223	0.161	0.101	0.07
	MSE	0.035	0.04	0.068	0.098	0.128	0.099	0.059	0.041
m = 200	|BIAS|	0.038	0.042	0.145	0.207	0.209	0.15	0.095	0.08
	MSE	0.021	0.027	0.052	0.083	0.088	0.057	0.029	0.021
m = 400	|BIAS|	0.045	0.028	0.129	0.197	0.177	0.135	0.113	0.105
	MSE	0.014	0.016	0.036	0.063	0.054	0.031	0.02	0.017
m = 800	|BIAS|	0.059	0.012	0.112	0.17	0.16	0.141	0.137	0.145
	MSE	0.012	0.01	0.024	0.043	0.037	0.026	0.021	0.024

specifically, if the sequence $U(\nu)$, $\nu = 0, 1, \ldots, N-1$, is the DFT of the discrete sequence of (possibly complex) values $u(t)$, $t = 0, 1, \ldots, N-1$, this constant is such that $\sum_{\nu=0}^{N-1} |U(\nu)|^2 = NA^2 \sum_{t=0}^{N-1} |u(t)|^2$ and generally chosen in $\{1/\sqrt{N}, 1, 1/N\}$.

To perform the standard Wiener filtering according to Eq. (15), an estimate of the noise standard deviation must be calculated so as to estimate $\mathbb{E}[|X_k(\nu)|^2] = 2NA^2\sigma_0^2$. A basic and popular solution to perform this estimate involves using a Voice Activity Detector (VAD). The estimate of σ_0 is then computed on the basis of the samples of the time frames that the VAD has detected as noise alone. Subspace approaches can also be used to estimate σ_0 by computing the smallest eigenvalues of the noisy speech autocorrelation matrix. In this case, the model order is difficult to choose and the computation of the eigenvalues may prove unstable.

We now present how to use the C-ESE for the estimation of σ_0. We begin by splitting the T available samples $y(t)$, $t = 0, 1, \ldots, T-1$, into frames of $N = 2^q$ successive samples each. In contrast to the previous subsection, the frames are here constructed so that they do not intersect; it is even conceivable to choose frames separated by some time lapse. We proceed thus for reasons given below.

Let K stand for the number of frames so constructed. The notations used above to denote the observed signal, the speech signal and noise in the kth frame are again $y_k(t)$, $s_k(t)$ and $x_k(t)$, $t = 0, 1, \ldots, N-1$, respectively. Consider now the set of observations $Y_k(\nu)$, $k \in \{1, \ldots, K\}$, $\nu \in \{0, \ldots, N-1\}$, where, again as above, $Y_k(\nu)$, $S_k(\nu)$ and $X_k(\nu)$, $\nu = 0, 1, \ldots, N-1$, stand for the respective DFTs of the signals $y_k(t)$, $s_k(t)$ and

$x_k(t)$, $t = 0, 1, \ldots, N-1$. Because of the Hermitian symmetry of the DFT, we can restrict our attention to half of this set, namely the complex random variables $Y_k(\nu)$, $k \in \{1, \ldots, K\}, \nu \in \{0, \ldots, N/2-1\}$ where N is assumed to be even.

Of course, we still have $Y_k(\nu) = S_k(\nu) + X_k(\nu)$ for every $k = 1, 2, \ldots, K$ and every $\nu = 0, 1, \ldots, N-1$. However, depending on the Signal to Noise Ratio (SNR) and the type of speech signal present during frame k, some speech time-frequency components can be neglected in comparison to noise and other speech time-frequency components. For instance, high frequency components of voiced speech signals can become negligible with regard to noise and the low-frequency components of the same speech signals. For unvoiced fricative speech signals, low-frequency components can also be significantly smaller than high-frequency ones and smaller than those due to noise.

According to these remarks, we propound to model the presence and the absence of the speech time-frequency component $S_k(\nu)$ by a discrete random variable $\varepsilon_k(\nu)$ valued in $\{0, 1\}$ and write that the observation is $Y_k(\nu) = \varepsilon_k(\nu) S_k(\nu) + X_k(\nu)$. With respect to this model, $\mathrm{P}(\{\varepsilon_k(\nu) = 1\})$ is the probability that some speech component is present at frequency ν during frame k. This probability of presence may be larger than one half for low frequency components. However, for high frequency components, this probability of presence becomes less than or equal to $1/2$ and even relatively small. Summarizing, we can consider that the speech time-frequency components are less present than absent. If we now split the observation set $Y_k(\nu)$, $k \in \{1, \ldots, K\}$, $\nu \in \{0, \ldots, N/2-1\}$, into subsets of m observations each and if the elements of each subset are chosen randomly amongst the available observations, the observations of each subset can reasonably be expected to be independent all the more since these observations are computed on the basis of non-intersecting frames. We thus assume that the model underlying proposition 3.1 and the construction of the C-ESE are satisfied for each subset of m observations.

According to Eq. (16), for each subset of m observations, the C-ESE will return an estimate of $A\sigma_0\sqrt{N}$ since each complex random vector $X_k(\nu)$ can be regarded as a centred two-dimensional random vector whose covariance matrix equals $A\sigma_0\sqrt{N}\mathbf{I}_2$. Therefore, our final estimate of $A\sigma_0\sqrt{N}$ is simply obtained by averaging all these estimates returned by the C-ESE. Dividing this average by $A\sqrt{N}$ yields an estimate of σ_0.

C. Performance evaluation

We randomly select twenty-five sentences in the TIMIT database and twenty-five sentences in the TIDIGITS database. These two databases are standard in speech processing. The TIMIT and the TIDIGITS databases are composed of read speech and connected digit sequences, respectively. The sentences selected in these databases are downsampled to 8 kHz before adding white Gaussian noise. We estimate the noise standard deviation as described above via the C-ESE on the basis of frames with $N = 256$ samples each. Such a frame thus corresponds to 32ms of noisy speech signals. As specified above, we use non-overlapping frames for the estimation of the noise standard deviation. They are not weighted either. On the basis of preliminary tests on signals other than the sentences used to achieve the experimental results presented below, we set the constant γ in (8) to $\sqrt{2}$.

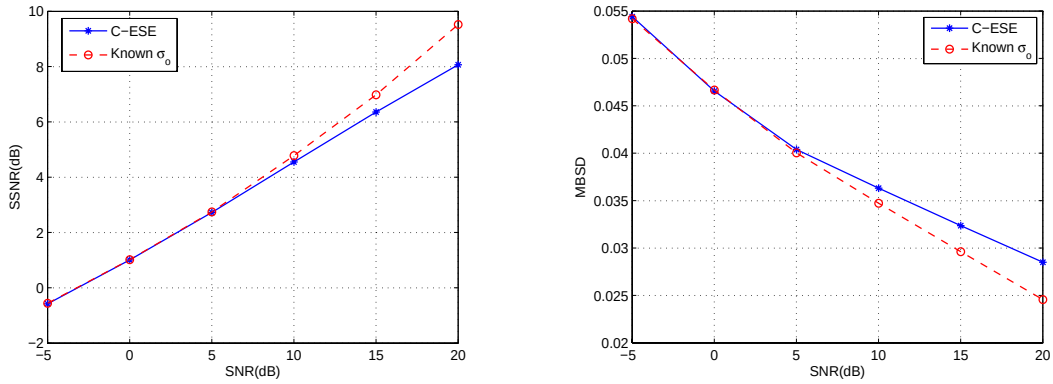
As far as the Wiener filtering is concerned, there is a 50% overlap between two adjacent frames and each frame is weighted by a Hanning window.

We evaluate the quality of the filtered speech signals by calculating the standard Segmental Signal to Noise Ratio (SSNR) (see [10]) and the Modified Bark Spectral Distortion (MBSD) (see [11]). The SSNR is the average of the SNR values on short segments. The SSNR is not relevant enough to measure the distortion of the denoised speech signals. This is the reason why we use the MBSD. The MBSD is an improved version of the Bark Spectral Distortion (BSD) proposed in [12]. Basically, it extends the BSD by incorporating the speech masking threshold so as to perform the measure without taking into account inaudible distortions. The MBSD proves to be highly correlated with subjective speech quality assessment [11].

The average SSNR and MBSD obtained over the twenty-five sentences of the TIMIT database are presented in figures 5(a) and 5(b). Figures 6(a) and 6(b) are the average SSNR and MBSD calculated over the twenty-five sentences of the TIDIGITS database. In all these figures, the solid curves are the performance measurements achieved with the filtering defined by equations (14) and (15) when the C-ESE is used instead of the true value of the noise standard deviation. The dashed curves are the results obtained when the filtering is achieved with the exact value of the noise standard deviation. The results are similar for the two sets of sentences considered during these experiments. Clearly, the Wiener filtering adjusted with the noise standard deviation estimate yields results that are significantly close to those obtained when the noise standard deviation is known. In addition, figure 7(a) (resp. figure 7(b)) compares the average value of the C-ESE obtained by processing the 25 sentences of the TIMIT database (resp. of the TIDIGITS database) to the true value of the noise standard deviations at the different SNR's tested during the experiments. These figures illustrate the good behaviour of our algorithm. As mentioned below, they are preliminary to further studies undertaken to get better insight into the behaviour of the C-ESE with respect to applications in speech denoising.

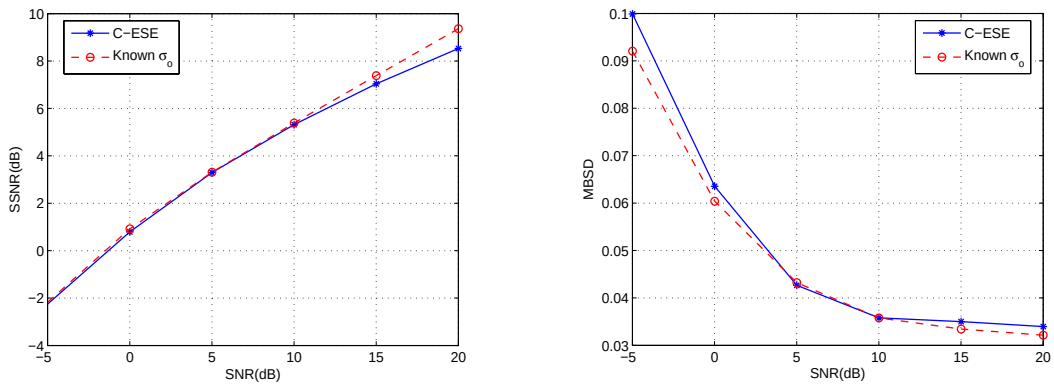
VII. Perspectives and extensions

We have presented algorithms for estimating the standard deviation of some AWGN when observations derive from signals less present than absent in this background. According to experimental results, this algorithm is very promising. An application to speech denoising has been



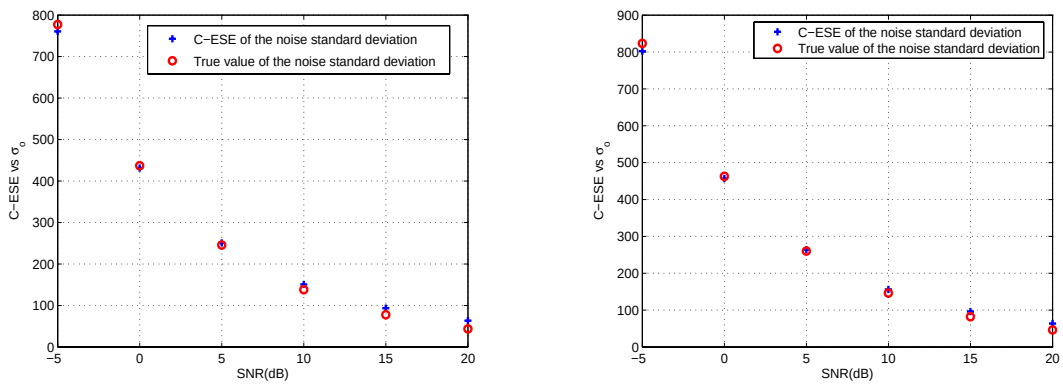
(a) SSNR (b) MBSD

Figure 5. (a) SSNR and (b) MBSD improvement for 25 speech signals randomly chosen in the TIMIT database and additively corrupted by independent AWGN with various SNRs.



(a) SSNR (b) MBSD

Figure 6. (a) SSNR and (b) MBSD improvement for 25 speech signals randomly chosen in the TIDIGITS database and additively corrupted by independent AWGN with various SNRs.



(a) Average over 25 sentences of the TIMIT database (b) Average over 25 sentences of the TIDIGITS database

Figure 7. For each tested SNR, comparison between the true value of the noise standard deviation and the average value of the C-ESE computed over (a) the 25 noisy sentences of the TIMIT database and (b) the 25 noisy sentences of the TIDIGITS database.

described where the estimator avoids the use of a DAV or subspace approaches that can be difficult to adjust.

New theoretical developments should be achieved in forthcoming work. In particular, further study of the Modified and Complex ESEs are required to get better insight into the behaviour of these estimates, in particular with regard to the constant γ in Eq. (8) since this constant is chosen empirically depending on the application.

More exhaustive experiments are in progress so as to better characterize the statistical behaviour of the C-ESE in the case of noisy speech signals. The analysis of the bias, the consistency and the mean square estimation error of the estimators proposed in this paper should also be addressed from a theoretical point of view.

As an extension of the application to speech denoising described above, on-going work concerns the application of the C-ESE to perceptual speech denoising in AWGN and coloured noise. As far as coloured noise is concerned, the use of the C-ESE could be extended along the following lines. The processing would involve assuming that the coloured noise power spectral density (psd) is a step function. Given a set of DFT bins where the noise psd is assumed to be constant, this constant value could then be estimated by computing the C-ESE on the basis of the values returned in these frequency bins by the DFTs of non-intersecting time-frames. By using non-intersecting frames, we still expect that the time-frequency components can reasonably be considered as independent.

The design of CFAR systems for the detection of radar targets also seems to be a rather natural field of application. Application in ESM for the interception of non-cooperant communications could also be investigated. Two other areas of applications are proximity sensing and distributed detection systems (see [13]). Proximity sensing aims at informing a robot that it is approaching an object. For the design of distributed detection systems, the absence of prior knowledge about the statistics of the observations is still an issue.

REFERENCES

[1] G. Minkler and J. Minkler, *The Principles of Automatic Radar Detection in Clutter, CFAR*. Magellan Book Company, Baltimore, 1990.

[2] N. Levanon, *Radar Principles*. John Wiley and Sons, New York, 1988.

[3] D. Pastor, "A theoretical result for processing signals that have unknown distributions and priors in white gaussian noise," *to appear in Computational Statistics & Data Analysis (CSDA)*.

[4] N. N. Lebedev, *Special Functions and their Applications*. Prentice-Hall, Englewood Cliffs, 1965.

[5] D. Pastor, R. Gay, and A. Groenenboom, "A sharp upper-bound for the probability of error of the likelihood ratio test for detecting signals in white gaussian noise," *IEEE Transactions on Information Theory*, vol. 48, pp. 228–238, 2002.

[6] W. H. Press, S. A. Teukolsky, and B. P. Flannery, *Numerical recipes in C, The Art of Scientific Computing, 2nd Edition*. University Press, Cambridge, 1992.

[7] H. V. Poor, *An Introduction to Signal Detection and Estimation, 2nd Edition*. Springer-Verlag, New-York, 1994.

[8] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions, Ninth printing*. Dover Publications Inc., New York, 1972.

[9] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 32, pp. 1109–1121, 1984.

[10] S. Quackenbush, R. Barnwell, and M. Clements, *Objective Measures of Speech Quality*. Prentice-Hall, Englewood Cliffs, 1988.

[11] W. Yang, M. Dixon, and R. Yantorno, "Modified bark spectral distortion measure which uses noise masking threshold," in *Proceedings of the 1997 IEEE Workshop on Speech Coding For Telecommunications*, september 1997.

[12] S. Wang, A. Sekey, and A. Gersho, "An objective measure for predicting subjective quality of speech coders," *IEEE Journal on Selected Areas in Communications*, vol. 10, pp. 819–829, 1992.

[13] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer-Verlag New-York Inc., 1997.

Dominique Pastor was born in Cahors, France. He graduated from Ecole Nationale Supérieure des Télécommunications de Bretagne (ENST Bretagne) in 1986. He received his PhD degree in Signal Processing and Telecommunications from the University of Rennes I, in 1997. From 1986 to 2000, he was an engineer with Thales Group. From 2000 to 2002, he was a consultant with Altran Technologies Netherlands. Since 2002, he is an Associate Professor at ENST Bretagne. His current research interests include statistics and wavelet theory with applications to speech processing and radar processing.

Asmaa Amehraye was born in Rabat, Morocco. She is currently a Ph.D candidate at ENST Bretagne, France, and Mohamed V University, Morocco. She received the two-year university diploma in physics and chemistry in 2000, the BA degree in computer sciences, electronics, electrical engineering and automatic in 2002 and the MA in research in Computer sciences and telecommunications in 2004, all from Mohamed V University of Morocco. Her research interests include speech enhancement and applications.