Published in final edited form as:
Speech Commun. 2007 February ; 49(2): 134–143.

A Laplacian-based MMSE estimator for speech enhancement

Bin Chen and **Philipos C. Loizou***

Abstract

This paper focuses on optimal estimators of the magnitude spectrum for speech enhancement. We present an analytical solution for estimating in the MMSE sense the magnitude spectrum when the clean speech DFT coefficients are modeled by a Laplacian distribution and the noise DFT coefficients are modeled by a Gaussian distribution. Furthermore, we derive the MMSE estimator under speech presence uncertainty and a Laplacian statistical model. Results indicated that the Laplacian-based MMSE estimator yielded less residual noise in the enhanced speech than the traditional Gaussian-based MMSE estimator. Overall, the present study demonstrates that the assumed distribution of the DFT coefficients can have a significant effect on the quality of the enhanced speech.

Keywords

MMSE estimator; speech enhancement; Laplacian speech modeling; speech absence probability

I. INTRODUCTION

Single-channel speech enhancement algorithms based on minimum mean-square error (MMSE) estimation of the short-time spectral magnitude have received a lot of attention in the past two decades [1–3]. A key assumption made in the MMSE algorithms is that the real and imaginary parts of the clean Discrete Fourier Transform (DFT) coefficients can be modeled by a Gaussian distribution. This Gaussian assumption, however, holds asymptotically for long duration analysis frames, for which the span of the correlation of the signal is much shorter than the DFT size. While this assumption might hold for the noise DFT coefficients, it does not hold for the speech DFT coefficients, which are typically estimated using relatively short (20–30 ms) duration windows. For that reason, several researchers [4–9] have proposed the use of non-Gaussian distributions for modeling the real and imaginary parts of the speech DFT coefficients. In particular, the Gamma or the Laplacian probability distributions can be used to model the distributions of the real and imaginary parts of the DFT coefficients. Several have computed histograms of the real and imaginary parts of the DFT coefficients from a large corpus of speech and confirmed that the Gamma and Laplacian distributions provide a better fit to the experimental data than the Gaussian distribution [6][4]. This was also confirmed quantitatively in [7] by using the Kullback divergence to measure the ability of the Gamma probability density function (pdf) to fit the experimental data. A smaller Kullback divergence was found for the Gamma pdf when compared to the Gaussian pdf, suggesting that the Gamma pdf provides a better fit to the experimental data than the Gaussian pdf.

The use of Gamma or Laplacian distributions, however, complicates the derivation of the MMSE estimate of the magnitude spectrum. This is partly because the magnitude and phases

* Corresponding author: Department of Electrical Engineering, University of Texas at Dallas, Richardson, Texas 75083-0688. Email: loizou@utdallas.edu Phone: (972) 883-4617 Fax: (972) 883-2710.

of the DFT coefficients are no longer independendent when the real and imaginary parts of the DFT coefficients are modelled by a Laplacian (or Gamma) distribution. For that reason, alternative solutions were explored in [4–8]. For instance, in [6] the authors approximated the pdf of the magnitude of the DFT coefficients with a parametric function, and used that to derive a MAP estimator of the magnitude spectrum. The MAP estimator was pursued over the MMSE estimator since the resulting integrals were too difficult to evaluate in closed form. In [5], the estimators of the real and imaginary parts of the DFT coefficients were derived separately assuming Gamma and Laplacian distributions for the speech DFT coefficients. The two estimators combined yielded a complex-valued estimator for the signal DFT coefficients. Experimental results showed that those estimators provided consistently better results than the Wiener estimator.

In [9] we derived an approximate MMSE estimator of the speech magnitude spectrum based on a Laplacian model for the speech DFT coefficients and a Gaussian model for the noise DFT coefficients. This estimator was derived under the assumption that the magnitude and phases of the complex DFT coefficients were independent. Acknowledging that this assumption does not necessarily hold, we derive in this paper the true MMSE estimator of the speech magnitude spectrum based on Laplacian modeling. The derived estimator is implemented using numerical integration techniques, and compared to the approximate MMSE estimator [9]. To further improve the amplitude estimation, we also incorporate speech presence uncertainty into the Laplacian based estimator. The performance of the proposed estimator is compared to the conventional MMSE estimator [1] as well as the Laplacian estimator proposed in [5].

The paper is organized as follows. In sections II and III, we derive the Laplacian-based MMSE estimators and in section IV we derive the MMSE estimator under signal presence uncertainty. In Section V, we evaluate the performance of the proposed estimators, and in Section VI we present the conclusions.

II. LAPLACIAN BASED SHORT-TIME SPECTRAL AMPLITUDE ESTIMATOR

Let $y(n) = x(n) + d(n)$ be the sampled noisy speech signal consisting of the clean signal $x(n)$ and the noise signal $d(n)$. Taking the short-time Fourier transform of $y(n)$, we get:

$$Y(\omega_k) = X(\omega_k) + D(\omega_k) \tag{1}$$

for $\omega_k = 2\pi k/N$ where $k = 0, 1, 2, \ldots N-1$, and N is the frame length. The above equation can also be expressed in polar form as

$$Y_k e^{j\theta_y(k)} = X_k e^{j\theta_x(k)} + D_k e^{j\theta_d(k)} \tag{2}$$

where $\{Y_k, X_k, D_k\}$ denote the corresponding magnitude spectra and $\{\theta_y(k), \theta_x(k), \theta_d(k)\}$ denote the corresponding phase spectra of the noisy, clean and noise signals respectively. The MMSE estimator of the magnitude spectrum X_k is obtained as follows [1]:

$$\begin{aligned}\hat{X}_k &= E\{X_k \mid Y(\omega_k)\}, \quad k = 0, 1, 2, \ldots N-1 \\ &= \frac{\int_0^\infty \int_0^{2\pi} X_k p(Y(\omega_k) \mid X_k, \theta_k) p(X_k, \theta_k) d\theta_k dX_k}{\int_0^\infty \int_0^{2\pi} p(Y(\omega_k) \mid X_k, \theta_k) p(X_k, \theta_k) d\theta_k dX_k}\end{aligned} \tag{3}$$

where $E\{\cdot\}$ denotes the expectation operator, $\theta_k \triangleq \theta_x(k)$ for convenience, $p(X_k, \theta_k)$ is the joint pdf of the magnitude and phase spectra, $\lambda_d(k)$ denotes the noise variance and $p(Y(\omega_k)|X_k, \theta_k)$ is given by

$$p(Y(\omega_k) \mid X_k, \theta_k) = \frac{1}{\pi\lambda_d(k)} \exp\left\{ -\frac{1}{\lambda_d(k)} \mid Y(\omega_k) - X(\omega_k)\mid^2 \right\} \tag{4}$$

Following the procedure in [12], it is easy to show for a Laplacian distribution that $p(X_k, \theta_k)$ is given by:

$$p(X_k, \theta_k) = \frac{X_k}{2\sqrt{\lambda_x(k)}} \exp\left[-\frac{X_k}{\sqrt{\lambda_x(k)}}(\mid \cos\theta_k \mid + \mid \sin\theta_k \mid) \right] \tag{5}$$

Substituting (4) and (5) from Appendix B into (3), we get:

$$\hat{X}_k = \frac{\int_0^\infty X_k^2 \exp\left(-\frac{X_k^2}{\lambda_d(k)} \right) \int_0^{2\pi} \exp\left[\frac{2X_k Y_k \cos\theta_k}{\lambda_d(k)} - \frac{X_k}{\sqrt{\lambda_x(k)}}(\mid \cos\theta_k \mid + \mid \sin\theta_k \mid) \right] d\theta_x dX_k}{\int_0^\infty X_k \exp\left(-\frac{X_k^2}{\lambda_d(k)} \right) \int_0^{2\pi} \exp\left[\frac{2X_k Y_k \cos\theta_k}{\lambda_d(k)} - \frac{X_k}{\sqrt{\lambda_x(k)}}(\mid \cos\theta_k \mid + \mid \sin\theta_k \mid) \right] d\theta_x dX_k} \tag{6}$$

After substituting (33) and (34), we can express the above equation in terms of the *a priori* and posteriori SNRs as follows:

$$\hat{X}_k = \frac{\int_0^\infty X_k^2 \exp\left(-\frac{\gamma_k X_k^2}{Y_k^2} \right) \int_0^{2\pi} \exp\left[\frac{2\gamma_k X_k \cos\theta_k}{Y_k} - \frac{X_k\sqrt{\gamma_k}}{Y_k\sqrt{\xi_k}}(\mid \cos\theta_k \mid + \mid \sin\theta_k \mid) \right] d\theta_x dX_k}{\int_0^\infty X_k \exp\left(-\frac{\gamma_k X_k^2}{Y_k^2} \right) \int_0^{2\pi} \exp\left[\frac{2\gamma_k X_k \cos\theta_k}{Y_k} - \frac{X_k\sqrt{\gamma_k}}{Y_k\sqrt{\xi_k}}(\mid \cos\theta_k \mid + \mid \sin\theta_k \mid) \right] d\theta_x dX_k} \tag{7}$$

where $\xi_k \triangleq \lambda_x(k)/\lambda_d(k)$ and $\gamma_k \triangleq Y_k^2 / \lambda_d(k)$ denote the *a priori* and posteriori SNRs respectively [1]. The above equation gives the Laplacian MMSE estimator of the spectral magnitudes, and we will be referring to this estimator as the LapMMSE estimator. The closed form solution of the above estimator is unknown to the authors, and therefore alternative solutions were sought. To derive such solutions, we needed to make some assumptions about the relationship between the magnitude and the phase pdfs, and this is examined next.

III. Derivation of approximate Laplacian MMSE estimator

It is known that complex zero mean Gaussian random variables have magnitudes and phases which are statistically independent [12]. Furthermore, the phases have a uniform distribution. This is not the case, however, with the complex Laplacian distributions that are used in this paper for modeling the speech DFT coefficients. Further analysis of the joint pdf of the magnitudes and phases, $p(X_k, \theta_k)$, however, evealed that the pdfs of the magnitudes and phases are nearly statistically independent, at least for a certain range of magnitude values. To show that, we derive the marginal pdfs of the magnitudes and phases and examine whether $p(X_k, \theta_k) \approx p(X_k)p(\theta_k)$. The marginal pdf of the phases is derived from (5) and is given by:

$$\begin{aligned} p(\theta_k) &= \frac{X_k}{2\sqrt{\lambda_x(k)}}\int_0^\infty \exp\left[-\frac{X_k}{\sqrt{\lambda_x(k)}}\left|\cos\theta_k\right| + \left|\sin\theta_k\right|\right] dX_k \\ &= -\frac{X_k}{2\left|\cos\theta_k\right| + 2\left|\sin\theta_k\right|}\exp\left[-\frac{X_k}{\sqrt{\lambda_x(k)}}(\cos\theta_k + \sin\theta_k)\right]_0^\infty \\ &\quad -\frac{\sqrt{\lambda_x(k)}}{2 + 4\left|\cos\theta_k \sin\theta_k\right|}\exp\left[-\frac{X_k}{\sigma}(\cos\theta_k + \sin\theta_k)\right]_0^\infty \\ &= \frac{\sqrt{\lambda_x(k)}}{2 + 4\left|\cos\theta_k \sin\theta_k\right|} \end{aligned} \tag{8}$$

The density $p(X_k)$ of the spectral magnitudes is given by (see derivation in Appendix A):

$$p(X_k) = \frac{2X_k}{\lambda_x(k)}\left[\frac{\pi}{4}I_0\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right) + 2\sum_{n=1}^{\infty}\frac{1}{n}I_n\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right)\sin\left(\frac{\pi n}{4}\right)\right]u(X_k) \tag{9}$$

where $I_n(\cdot)$ denotes the modified Bessel function of nth order and $u(x)$ is the step function. Figures 1 and 2 show plots of the joint density $p(X_k, \theta_k)$ as well as plots of the product of the magnitude and phase pdfs. Figure 1 shows the joint density $p(X_k, \theta_k)$, Fig. 2 shows $p(X_k)p(\theta_k)$ and Fig. 3 shows the absolute difference between the densities displayed in Figs. 1 and 2. As can be seen, the difference between the two densities is large near $X_k \approx 0$, but is near zero for $X_k > 2$. The plot in Figure 3 demonstrates that the magnitudes and phases are nearly independent, at least for a specific range of magnitude values ($X_k > 2$, $\theta_k \in [-\pi, \pi]$). We can therefore make the approximation that $p(X_k, \theta_k) \approx p(X_k)p(\theta_k)$.

We further analyzed the phase pdf, $p(\theta_k)$, to determine the shape of the distribution and examine whether it is similar to a uniform distibution. Figure 4 shows the plots of $p(\theta_k)$ superimposed to a uniform distribution. The density $p(\theta_k)$ is clearly not uniform, but it oscillates near the $1/(2\pi)$ value of the uniform distribution for $\theta_k \in [-\pi, \pi]$. Despite this difference, we approximated $p(\theta_k)$ with a uniform distribution, i.e., $p(\theta_k) \approx 1/(2\pi)$ for $\theta_k \in [-\pi, \pi]$.

After taking into consideration the above two assumptions (statistical independence between X_k and θ_k, and a uniform distribution for the phases), we approximated the joint density in (3) with $p(X_k, \theta_k) \approx \frac{1}{2\pi}p(X_k)$, where $p(X_k)$ is the density of the spectral magnitudes. Finally, after substituting (9) and (4) into (3) and using [11, Eq. 6.633.1] we get an expression for the MMSE estimator in closed form (see derivation in Appendix B):

$$\hat{X}_k = \frac{A_k + B_k}{C_k + D_k} \tag{10a}$$

where

$$A_k = \frac{\left(\frac{Y_k^2}{\gamma_k}\right)^{\frac{3}{2}}}{2}\sum_{m=0}^{\infty}\frac{\Gamma(m+\frac{3}{2})}{m!\,\Gamma(m+1)}\left(\frac{\gamma_k}{2\xi_k^2 Y_k^2}\right)^m \cdot F(-m, -m; 1; 2\xi_k^2 Y_k^2) \tag{10b}$$

$$B_k = \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sin\left(\frac{\pi n}{4}\right)\frac{\left(\frac{2\gamma_k}{Y_k}\right)^n\left(\frac{\gamma_k}{Y_k^2}\right)^{-\frac{n+3}{2}}}{2^{n+1}\Gamma(n+1)}\cdot\sum_{m=0}^{\infty}\frac{\Gamma(m+\frac{1}{2}n+\frac{3}{2})}{m!\,\Gamma(m+1)}\left(\frac{\gamma_k}{2\xi_k^2 Y_k^2}\right)^m \tag{10c}$$

$$\cdot F(-m,\ -m;\ n+1;\ 2\xi_k^2 Y_k^2)$$

$$C_k = \frac{Y_k^2}{2\gamma_k}\sum_{m=0}^{\infty}\frac{1}{m!}\left(\frac{\gamma_k}{2\xi_k^2 Y_k^2}\right)^m \cdot F(-m,\ -m;\ 1;\ 2\xi_k^2 Y_k^2) \tag{10d}$$

$$D_k = \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sin\left(\frac{\pi n}{4}\right)\frac{\left(\frac{2\gamma_k}{Y_k}\right)^n\left(\frac{\gamma_k}{Y_k^2}\right)^{-\frac{n}{2}-1}}{2^{n+1}\Gamma(n+1)}\cdot\sum_{m=0}^{\infty}\frac{\Gamma(m+\frac{1}{2}n+1)}{m!\,\Gamma(m+1)}\left(\frac{\gamma_k}{2\xi_k^2 Y_k^2}\right)^m \tag{10e}$$

$$\cdot F(-m,\ -m;\ n+1;\ 2\xi_k^2 Y_k^2)$$

where ζ_k and γ_k are the *a priori* and *posteriori* signal-to-noise (SNR) ratios respectively, $\gamma(\cdot)$ is the gamma function and $F(a, b, c; x)$ is the Gaussian hypergeometric function [11, Eq. 9.100]. Equation (10a) gives the approximate Laplacian MMSE estimator of the spectral magnitudes, and we will be referring to this estimator as the ApLapMMSE estimator.

IV. Derivation of amplitude estimator under speech presence uncertainty

In this section, we derive the MMSE magnitude estimator under the assumed Laplacian model and uncertainty of speech presence. This is motivated by the fact that speech might not be present at all times and at all frequencies. We could therefore consider a two-state model for speech events that assumes that either speech is present at a particular frequency bin (hypothesis H_1) or that is not (hypothesis H_0). Intuitively, this amounts to multiplying the estimator by a term that provides an estimate of the probability that speech is present at a particular frequency bin. Following [1][10], this new estimator is given by:

$$\hat{X}_k = E(X_k \mid Y(\omega_k),\ H_1^k)P(H_1^k \mid Y(\omega_k)) \tag{11}$$

where H_1^k denotes the hypothesis that speech is present in frequency bin k, and $P(H_1^k \mid Y(\omega_k))$ denotes the conditional probability that speech is present in frequency bin k given the noisy speech (complex) spectrum $Y(\omega_k)$. The conditional probability $P(H_1^k \mid Y(\omega_k))$ can be computed using Bayes' rule [1][10]:

$$P(H_1^k \mid Y(\omega_k)) = \frac{\Lambda(Y(\omega_k),\ q_k)}{1+\Lambda(Y(\omega_k),\ q_k)} \tag{12}$$

where $\Lambda(Y(\omega_k), q_k)$ is the generalized likelihood ratio defined by:

$$\Lambda(Y(\omega_k),\ q_k) = \frac{1-q_k}{q_k}\frac{p(Y(\omega_k) \mid H_1^k)}{p(Y(\omega_k) \mid H_0^k)} \tag{13}$$

where $q_k = P(H_0^k)$ denotes the *a priori* probability of speech absence for frequency bin k.

Under hypothesis H_0, $Y(\omega_k) = D(\omega_k)$, and given that the noise is complex Gaussian with zero mean and variance $\lambda_d(k)$, it follows that $p(Y(\omega_k) \mid H_0^k)$ will also have a Gaussian distribution with the same variance, i.e.,

$$p(Y(\omega_k) \mid H_0^k) = \frac{1}{\pi\lambda_d(k)} \exp\left(-\frac{Y_k^2}{\lambda_d(k)}\right) \tag{14}$$

Under hypothesis H_1, $Y(\omega_k) = X(\omega_k) + D(\omega_k)$, and $p(Y(\omega_k) \mid H_1^k)$ will have the form:

$$p_{Y(\omega_k)}(y) = p(z_r, z_i) = p_{Z_r(k)}(z_r)\, p_{Z_i(k)}(z_i) \tag{15}$$

where $Z_r(k) = \mathrm{Re}\{Y(\omega_k)\}$, $Z_i(k) = \mathrm{Im}\{Y(\omega_k)\}$, and $p_{Z_r}(k)(z_r)$ and $p_{Z_i}(k)(z_i)$ are given by (see Appendix C):

$$p_{Z_r(k)}(z_r) = \frac{\sqrt{\gamma_k}\exp\left(\frac{1}{2\xi_k}\right)}{2\sqrt{2\xi_k}Y_k}\left[\exp\left(-\frac{\sqrt{\gamma_k}z_r}{Y_k\sqrt{\xi_k}}\right) + \exp\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k}\right) + \exp\left(-\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k}\right)\mathrm{erf}\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k} - \frac{1}{\sqrt{\xi_k}}\right) - \exp\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k}\right)\mathrm{erf}\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k} + \frac{1}{\sqrt{\xi_k}}\right)\right] \tag{16}$$

where erf(·) is the error function. A similar form of the above (16) can be also found in [17]. After substituting (12), (14), and (15) into (11) we get the final estimator that incorporates speech-presence uncertainty.

V. IMPLEMENTATION AND PERFORMANCE EVALUATION

A. Implementation

Evaluation of $p(X_k)$ in (9) involves an infinite number of terms, however, computer simulations indicated that retaining only the first 40 terms in (9), gave a good approximation of $p(X_k)$. This is demonstrated in Figure 5, which shows $p(X_k)$ estimated using numerical integration techniques and also approximated by truncating the summation in (9) using the first 40 terms.

As shown in (10a), the derived ApLapMMSE estimator is highly nonlinear and computationally complex. The implementation of (10a) proved to be challenging due to the infinite number of terms involved in the summation. In practice, scaling techniques can be used to implement (10a) to avoid possible overflows or underflows. Instead, we chose to use numerical integration techniques [13] to evaluate the integrals in (3). More specifically, after making the assumptions of independence and uniform phase distribution, we used numerical intergration techniques to evaluate the estimator of the magnitude spectrum as follows:

$$\hat{X}_k = \frac{\int_0^\infty X_k^2 \exp\left(-\frac{\gamma_k X_k^2}{Y_k^2}\right) I_0\left(\frac{2\gamma_k X_k}{Y_k}\right) p_X(X_k)\, dX_k}{\int_0^\infty X_k \exp\left(-\frac{\gamma_k X_k^2}{Y_k^2}\right) I_0\left(\frac{2\gamma_k X_k}{Y_k}\right) p_X(X_k)\, dX_k}$$

where $p_X(X_k)$ is given by (23). The above integrals were used to evaluate the ApLapMMSE estimator (MATLAB implementation of the above estimator is available in [16]). Numerical

integration techniques were also used to evaluate the integrals involved in the LapMMSE estimator in (7).

The proposed estimators were applied to 20-ms duration frames of speech using a Hamming window, with 50% overlap between frames. The "decision-directed" approach [1] was used in the proposed estimators to compute the *a priori* SNR ζ_k, with $\alpha = 0.98$. The enhanced signal was combined using the overlap and add approach. The *a priori* probability of speech absence, q_k, was set to $q_k = 0.3$ in (13).

B. Performance Evaluation

Twenty sentences from the TIMIT database were used for the objective evaluation of the proposed LapMMSE estimator, 10 produced by female speakers and 10 produced by male speakers. The TIMIT sentences were downsampled to 8 kHz. Speech-shaped noise constructed from the long-term spectrum of the TIMIT sentences as well as F-16 cockpit noise were added to the clean speech files at 0, 5 and 10 dB SNR. White noise was also used to corrupt the sentences at 0, 5 and 10 dB SNR. An estimate of the noise spectrum was obtained from the initial 100-ms segment of each sentence. The noise spectrum estimate was not updated in subsequent frames.

Objective measures were used to evaluate the performance of the proposed estimators implemented with and without speech presence uncertainty (SPU) and denoted as LapMMSE-SPU and LapMMSE respectively. Similarly, the approximate Laplacian estimators implemented with and without speech presence uncertainty were indicated as ApLapMMSE-SPU and ApLapMMSE respectively. For comparative purposes we evaluated the performance of the traditional (Gaussian-based) MMSE estimator [1] with and without incorporating speech presence uncertainty which were indicated as MMSE-SPU and MMSE respectively. We also evaluated the performance of the complex-valued MMSE estimator derived in [5] based on Laplacian speech priors. Note that in [5], the estimator $E[X(\omega_k)|Y(\omega_k)]$ was derived by combining the estimators of the real and imaginary parts of the DFT coefficients.

The segmental SNR, log likelihood ratio (LLR) and PESQ (ITU-T P.862) measures were used for objective evaluation of the proposed estimators. The segmental SNR was computed as:

$$SNRseg = \frac{10}{M}\sum_{k=0}^{M-1}\log_{10}\frac{\sum_{n=Nk}^{Nk+N-1}x^2(n)}{\sum_{n=Nk}^{Nk+N-1}(x(n)-\hat{x}(n))^2} \tag{17}$$

where M is the total number of frames, N is the frame size, $x(n)$ is the clean signal and $\hat{x}(n)$ is the enhanced signal. Since the segmental SNR can become very small and negative during periods of silence, we limited the *SNRseg* values to the range of [−10 dB, 35 dB] as per [14]. In doing so, we avoid the explicit marking and identification of speech-absent segments. Note that we report the *SNRseg* values for completeness and for comparitive purposes with prior studies. The *SNRseg* measure has been found in [15] to yield a low correlation ($\rho = 0.31$) with subjective quality ratings. The LLR and PESQ measures have been found to yield much stronger correlations with subjective quality ratings [15].

The log likelihood ratio (LLR) for each 20-ms frame was computed as:

$$LLR = \log\left(\frac{a_e^T R_x a_e}{a_x^T R_x a_x}\right) \tag{18}$$

where a_x and a_e are the prediction coefficients of the clean and enhanced signals respectively, and R_x is the autocorrelation matrix of the clean signal. The mean LLR value was computed across all frames for each sentence. Since the mean can be easily biased by a few outlier frames, we computed the mean based on the lowest 95% of the frames as per [14].

Tables 1 and 2 list the segmental SNR values and log-likelihood ratio values obtained by the various estimators at different SNRs, and Table 3 lists the PESQ values. Table 4 shows the evaluation of the proposed estimator with sentences corrupted with stationary white noise. As can be seen, higher segmental SNR values and higher PESQ values were obtained consistently by the proposed estimators (LapMMSE and ApLapMMSE). Statistical analysis (paired samples t-tests) indicated that the SNRseg values (Table 1) obtained with the Laplacian estimators (ApLapMMSE and ApLapMMSE-SPU) were significantly ($p < 0.005$) higher higher than the SRNseg values obtained with the Gaussian estimators (MMSE and MMSE-SPU) in all SNR conditions. The PESQ values obtained with the Laplacian estimators (ApLapMMSE and ApLapMMSE-SPU) were also found to be significantly higher ($p < 0.05$) than the values obtained by the Gaussian estimators (MMSE and MMSE-SPU) in all SNR conditions except one (0 dB speech-shaped noise). The PESQ values obtained with the ApLapMMSE-SPU estimator were not significantly ($p > 0.05$) higher than the values obtained by the MMSE-SPU estimator in 0dB speech-shaped noise. Further analysis (paired samples t-tests) indicated that the SNRseg values (Table 4) obtained with the ApLapMMSE-SPU estimator were significantly ($p < 0.005$) higher than the SRNseg values obtained with the Gaussian MMSE-SPU estimator. The Laplacian estimator proposed in [5] also performed better than the Gaussian MMSE estimator. The difference in performance between the LapMMSE and ApLapMMSE estimators was very small and statistically non-significant, suggesting that our assumptions about the independence of magnitudes and phase were reasonable and did not cause any significant performance degradation.

The pattern of results was similar with the log-likelihood ratio objective measure (Table 2). Smaller LLR values were obtained by the proposed Laplacian estimators compared to the Gaussian-based MMSE estimator for all SNR conditions. Statistical analysis (paired samples t-tests) confirmed that the LLR values obtained with the ApLapMMSEestimator were significantly ($p < 0.005$) lower than the LLR values obtained with the Gaussian MMSE estimator. Comparison between the LLR values obtained with the MMSE-SPU and the ApLapMMSE-SPU estimators indicated that the ApLapMMSE-SPU estimator performed significantly ($p < 0.005$) better in all SNR conditions but two. The difference in performance between the two estimators in 10 dB speech-shaped noise and 10 dB fighter noise was not statistically significant ($p > 0.05$).

Informal listening tests indicated that speech enhanced by the Laplacian MMSE estimators had less residual noise. This was confirmed by visual inspection of spectrograms of the enhanced speech signals. Figure 6 shows the spectrograms of the TIMIT sentence “The kid has no manners, boys” enhanced by the LapMMSE-SPU, ApLapMMSE-SPU and MMSE-SPU estimators. The sentence was originally embedded in +5 dB S/N speech-shaped noise. As can be seen, the sentence enhanced by the Laplacian MMSE estimators had less residual noise without introducing perceptible distortion in the speech signal. The quality of speech enhanced by the ApLapMMSE and LapMMSE estimators was nearly identical, consistent with the objective evaluation of these estimators (Tables 1–3).

VI. SUMMARY AND CONCLUSIONS

An MMSE estimator was derived for the speech magnitude spectrum based on a Laplacian model for the speech DFT coefficients and a Gaussian model for the noise DFT coefficients. An estimator was also derived under speech presence uncertainty and a Laplacian model

assumption. Results, in terms of objective measures, indicated that the proposed Laplacian MMSE estimators yielded better performance than the traditional MMSE estimator, which is based on a Gaussian model [1]. Overall, the present study demonstrated that the assumed distribution of the DFT coefficients can have a significant effect on speech quality.

Acknowledgements

This research was supported in part by a grant from NIDCD/NIH. The authors would like to thank Prof. Ali Hooshyar for all his help and suggestions regarding numerical integration. Thanks also go to the reviewers for providing valuable suggestions that helped improve the manuscript.

References

1. Ephraim Y, Malah D. Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator. IEEE Trans Acoustics, Speech, Signal Proc Dec;1984 32:1109–1121.
2. Ephraim Y, Malah D. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. IEEE Trans Acoust, Speech, Signal Process 1985;ASSP-23(2):443–445.
3. Cohen I, Berdugo B. Spech enhancement for non-stationary noise environments. Signal Processing 2001;81:2403–2418.
4. Martin, R. Speech Enhancement Using MMSE Short Time Spectral Estimation with Gamma Distributed Priors. Proc IEEE ICASSP; May 2002; p. 504-512.
5. Martin, R.; Breithaupt, C. Speech enhancement in the DFT domain using Laplacian speech priors. Proc. 8th Intern. Workshop on Acoustic Echo and Noise Control (IWAENC); Kyoto, Japan. September 2003; p. 8-11.p. 87-90.
6. Lotter, T.; Vary, P. Noise Reduction by Maximum a Posteriori Spectral Amplitude Estimation With Supergaussian Speech Modeling. Intern. Workshop. Acoust. Echo Noise Control; Kyoto, Japan. September 2003;
7. Breithaupt, C.; Martin, R. MMSE Estimation of Magnitude-Squared DFT Coefficients with SuperGaussian Priors. Proc. IEEE ICASSP; April 2003; p. 896-899.
8. Porter, J.; Boll, S. Optimal estimators for spectral restoration of noisy speech. Proc IEEE ICASSP; 1984. p. 18A.2.1-18A.2.4.
9. Chen, B.; Loizou, P. Speech Enhancement Using a MMSE Short Time Spectral Amplitude Estimator with Laplacian Speech Modeling. Proc IEEE ICASSP; 2005. p. 1097-1100.
10. McAulay RJ, Malpass ML. Speech enhancement using a soft-decision noise suppression filter. IEEE Trans Acoust, Speech, Signal Processing 1980;ASSP-28:137–145.
11. Gradshteyn, IS.; Ryzhik, IM. Table of Integrals, Series and Products. 6. Academic Press; 2000.
12. Papoulis, A.; Pillai, U. Probability, Random Variables and Stochastic Processes. 4. McGraw-Hill; 2001.
13. Kwon, Y.; Bang, H. The Finite Element Method Using Matlab. 2. CRC Press; New York: 2000. p. 35
14. Hansen, J.; Pellom, B. An effective quality evaluation protocol for speech enhancement algorithms. Proc ICSLP; 1998. p. 2819-2822.
15. Hu, Y.; Loizou, P. Evaluation of objective measures for speech enhancement. Proc. of INTERSPEECH; Philadelphia, PA. September 2006;
16. Chen, B. Speech Enhancement Using a MMSE Short Time Spectral Amplitude Estimator with Laplacian Speech Modeling. PhD Thesis, Dept. of Electrical Engineering; University of Texas-Dallas. 2005.
17. Martin R. Speech Enhancement based on Minimum Mean Square Error Estimation and Supergaussian Priors. IEEE Trans SAP 2005;13(5):845–856.

VIII. APPENDIX A

In this Appendix, we derive the PDF of the random variable $X = \sqrt{X_r^2 + X_i^2}$, where X_r, X_i are the real and imaginary parts respectively of the DFT components of the clean speech signal, assumed to have a *Laplacian* probability density function of the form:

$$p_{X_r}(x_r) = \frac{1}{2\sigma} \exp\left(-\frac{|x_r|}{\sigma}\right) \tag{19}$$

where σ is the standard deviation. Let $Y_1 = X_r^2$ and $Y_2 = X_i^2$. Then, we know [12] that $p_{Y_1}(y_1) = \frac{1}{\sqrt{y_1}} p_{x_1}(\sqrt{y_1})$ and so:

$$p_{Y_1}(y_1) = \frac{1}{\sqrt{y_1}} \frac{1}{2\sigma} \exp\left(-\frac{\sqrt{y_1}}{\sigma}\right), \quad p_{Y_2}(y_2) = \frac{1}{\sqrt{y_2}} \frac{1}{2\sigma} \exp\left(-\frac{\sqrt{y_2}}{\sigma}\right) \tag{20}$$

Let $Z = X_r^2 + X_i^2 = Y_1 + Y_2$ ($Y_1 > 0$, $Y_2 > 0$), then $p_Z(z) = p_{Y_1}(y_1) * p_{Y_2}(y_2)$, where '*' indicates convolution. Thus

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} p_{Y_1}(z - y_2) p_{Y_2}(y_2) dy_2 \\ &= \frac{1}{4\sigma^2} \int_0^z \frac{1}{\sqrt{y_2(z - y_2)}} \exp\left(-\frac{\sqrt{z - y_2} + \sqrt{y_2}}{\sigma}\right) dy_2 \end{aligned} \tag{21}$$

After substituting $y_2 = z \sin^2 t$, we get:

$$p_Z(z) = \frac{1}{2\sigma^2} \int_0^{\pi/2} \exp\left(-\frac{\sqrt{2z}\cos(t - \pi/4)}{\sigma}\right) dt \tag{22}$$

Now, let $X = \sqrt{Z}$, then we know [12, p. 133] that $p_X(x) = 2x p_Z(x^2) u(x)$, where $u(x)$ is the step function. Substituting $\theta = t - \pi/4$ in the above equation, we get:

$$\begin{aligned} p_X(x) &= \frac{2x}{\sigma^2} \int_0^{\pi/4} \exp\left(-\frac{\sqrt{2}x\cos\theta}{\sigma}\right) d\theta \\ &= \frac{2x}{\sigma^2} \int_0^{\pi/4} \exp(Ax\cos\theta) d\theta \quad x \geq 0 \end{aligned} \tag{23}$$

where $A = -\sqrt{2}/\sigma$. After making use of the generating function:

$$\begin{aligned} \exp\left[\frac{w}{2}\left(t + \frac{1}{t}\right)\right] &= e^{wt/2} e^{w/2t} \\ &= \left(\sum_{m=0}^{\infty} \frac{1}{m!} \frac{w^m t^m}{2^m}\right)\left(\sum_{n=0}^{\infty} \frac{1}{n!} \frac{w^n t^{-n}}{2^n}\right) \end{aligned} \tag{24}$$

with $t = \exp(j\theta)$ and $w = Ax$, we can express $\exp(Ax\cos\theta)$ in terms of Bessel functions:

$$\exp(Ax\cos\theta) = I_0(Ax) + 2\sum_{k=1}^{\infty} I_k(Ax)\cos k\theta \tag{25}$$

where $I_k(\cdot)$ denotes the modified Bessel function of kth order. Finally, after substituting (25) in (23) and integrating, we get:

$$p_X(x) = \frac{2x}{\sigma^2}\left[\frac{\pi}{4} I_0\left(-\frac{\sqrt{2}}{\sigma}x\right) + 2\sum_{k=1}^{\infty}\frac{1}{k} I_k\left(-\frac{\sqrt{2}}{\sigma}x\right)\sin\frac{\pi k}{4}\right]u(x) \tag{26}$$

where $u(x)$ is the step function.

IX. APPENDIX B

In this Appendix, we derive the approximate Laplacian MMSE estimator of the magnitude spectrum. We start from [1]:

$$\hat{X}_k = E[X_k \mid Y(\omega_k)] = \frac{\int_0^{\infty}\int_0^{2\pi} X_k\, p(Y(\omega_k) \mid X_k, \theta_k)\, p(X_k, \theta_k)\, d\theta_k\, dX_k}{\int_0^{\infty}\int_0^{2\pi} p(Y(\omega_k) \mid X_k, \theta_k)\, p(X_k, \theta_k)\, d\theta_k\, dX_k} \tag{27}$$

The $p(Y(\omega_k)|X_k, \theta_k)$ term is given by [1]:

$$p(Y(\omega_k) \mid X_k, \theta_k) = \frac{1}{\pi\lambda_d(k)} \exp\left\{-\frac{1}{\lambda_d(k)} \mid Y(\omega_k) - X(\omega_k)\mid^2\right\} \tag{28}$$

Assuming that X_k and θ_k are independent and θ_k is uniformly distributed in [0 2π], and after using (26), we get

$$p(X_k, \theta_k) \approx \frac{1}{2\pi} p(X_k) \tag{29}$$

$$= \frac{X_k}{4\lambda_x(k)} I_0\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right) + \frac{2X_k}{\pi\lambda_x(k)}\sum_{n=1}^{\infty}\frac{1}{n} I_n\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right)\sin\left(\frac{\pi n}{4}\right) \tag{30}$$

Now, substituting (28) and (29) into (27), we get:

$$\hat{X}_k = \frac{A_k' + B_k'}{C_k' + D_k'} \tag{31a}$$

where

$$A_k' = \int_0^{\infty} \frac{X_k^2}{4\lambda_x(k)} \exp\left(-\frac{X_k^2}{\lambda_d(k)}\right) I_0\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right) I_0\left(2X_k Y_k \Big/ \lambda_d(k)\right) dX_k \tag{31b}$$

$$B_k' = \frac{2}{\pi\lambda_x(k)}\sum_{n=1}^{\infty}\frac{1}{n}\sin\left(\frac{\pi n}{4}\right)\int_0^{\infty} X_k^2 \exp\left(-\frac{X_k^2}{\lambda_d(k)}\right) I_n\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right) I_0\left(2X_k Y_k \Big/ \lambda_d(k)\right) dX_k \tag{31c}$$

$$C_k' = \int_0^{\infty} \frac{X_k}{4\lambda_x(k)} \exp\left(-\frac{X_k^2}{\lambda_d(k)}\right) I_0\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right) I_0\left(2X_k Y_k \Big/ \lambda_d(k)\right) dX_k \tag{31d}$$

$$D_k' = \frac{2}{\pi\lambda_x(k)}\sum_{n=1}^{\infty}\frac{1}{n}\sin\left(\frac{\pi n}{4}\right)\int_0^{\infty} X_k \exp\left(-\frac{X_k^2}{\lambda_d(k)}\right) I_n\left(-\frac{\sqrt{2}}{\sqrt{\lambda_x(k)}}X_k\right) I_0\left(2X_k Y_k \Big/ \lambda_d(k)\right) dX_k \tag{31e}$$

After using [11, Eq. 6.633.1], we derive the MMSE estimator in closed form:

$$\hat{X}_k = \frac{A_k + B_k}{C_k + D_k} \tag{32a}$$

where

$$A = \frac{\lambda_d(k)^{\frac{3}{2}}}{2} \sum_{m=0}^{\infty} \frac{\Gamma(m+\frac{3}{2})}{m!\,\Gamma(m+1)} \left(\frac{\lambda_d(k)}{2\lambda_x(k)}\right)^m F\left(-m, -m; 1; 2\frac{\lambda_x(k)}{\lambda_d^2(k)} Y_k^2\right) \tag{32b}$$

$$B = \frac{8}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{\pi n}{4}\right) \frac{\left(\frac{2Y_k}{\lambda_d(k)}\right)^n \left(\frac{1}{\lambda_d(k)}\right)^{-\frac{n+3}{2}}}{2^{n+1}\Gamma(n+1)} \sum_{m=0}^{\infty} \frac{\Gamma(m+\frac{1}{2}n+\frac{3}{2})}{m!\,\Gamma(m+1)} \left(-\frac{\lambda_d(k)}{2\lambda_x(k)}\right)^m \cdot F\left(-m, -m; n+1; 2\frac{\lambda_x(k)}{\lambda_d^2(k)} Y_k^2\right) \tag{32c}$$

$$C = \frac{\lambda_d(k)}{2} \sum_{m=0}^{\infty} \frac{1}{m!} \left(\frac{\lambda_d(k)}{2\lambda_x(k)}\right)^m F\left(-m, -m; 1; 2\frac{\lambda_x(k)}{\lambda_d^2(k)} Y_k^2\right) \tag{32d}$$

$$D = \frac{8}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{\pi n}{4}\right) \frac{\left(\frac{2Y_k}{\lambda_d(k)}\right)^n \left(\frac{1}{\lambda_d(k)}\right)^{-\frac{n}{2}-1}}{2^{n+1}\Gamma(n+1)} \sum_{m=0}^{\infty} \frac{\Gamma(m+\frac{1}{2}n+1)}{m!\,\Gamma(m+1)} \left(\frac{\lambda_d(k)}{2\lambda_x(k)}\right)^m \cdot F\left(-m, -m; n+1; 2\frac{\lambda_x(k)}{\lambda_d^2(k)} Y_k^2\right) \tag{32e}$$

where $\Gamma(\cdot)$ is the gamma function and $F(a, b, c; x)$ is the Gaussian hypergeometric function [11, Eq. 9.100]. The above terms are given as a function of the signal and noise variances, but can also be expressed in terms of the *a priori SNR* ι_k ($\iota_k = \lambda_x(k)/\lambda_d(k)$) *and a posteriori* SNR γ_k ($\gamma_k = Y_k^2 / \lambda_d(k)$) using the following relationships:

$$\frac{1}{\lambda_d(k)} = \frac{\gamma_k}{Y_k^2} \tag{33}$$

$$\frac{1}{\lambda_x(k)} = \frac{\gamma_k}{\xi_k Y_k^2} \tag{34}$$

Finally, after substituting Equations (33) and (34) in (32b)–(32e), we get (10a)–(10e).

X. APPENDIX C

In this appendix, we derive the PDF of $Y(\omega_k) = X(\omega_k) + D(\omega_k)$, where $X(\omega_k) = X_r(\omega_k) + jX_i(\omega_k)$ and $D(\omega_k) = D_r(\omega_k) + jD_i(\omega_k)$. The pdfs of $X_r(\omega_k)$ and $X_i(\omega_k)$ are assumed to be Laplacian and the pdfs of $D_r(\omega_k)$ and $D_r(\omega_k)$ are assumed to be Gaussian with variance $\sigma_d^2/2$ and zero mean. We assume that $p_{X_r}(x_r) = \frac{1}{2\sigma_x} \exp\left(-\frac{|x_r|}{\sigma_x}\right)$, and

$p_{D_r}(d_r) = \frac{1}{\sqrt{\pi}\sigma_d} \exp\left(-\frac{d_r^2}{\sigma_d^2}\right)$. For simplicity, we neglect the frequency index ω_k in the following derivation.

Let $Z_r = X_r + D_r$ and $Z_i = X_i + D_i$, then $Y = Z_r + jZ_i$. The pdf of Z_r can be computed by the convolution of the Laplacian and Gaussian densities, and is given by:

$$\begin{aligned} p_{Z_r}(z_r) &= \int_{-\infty}^{\infty} p_{X_r}(z_r - d_r) p_{D_r}(d_r) dd_r \\ &= \int_{-\infty}^{z_r} \frac{1}{2\sigma_x\sigma_d\sqrt{\pi}} \exp\left(-\frac{\sigma_x d_r^2 - \sigma_d^2 d_r + \sigma_d^2 z_r}{\sigma_x\sigma_d^2}\right) dd_r + \\ &\int_{z_r}^{\infty} \frac{1}{2\sigma_x\sigma_d\sqrt{\pi}} \exp\left(-\frac{\sigma_x d_r^2 - \sigma_d^2 d_r + \sigma_d^2 z_r}{\sigma_x\sigma_d^2}\right) dd_r \end{aligned} \tag{35}$$

After using [11, Eq. 2.338.1] we get:

$$p_{Z_r}(z_r) = p_{Z_r}(z_r) = \frac{\exp\left(\frac{\sigma_d^2}{\sigma_x^2}\right)}{2\sqrt{2}\sigma_x} \left[\exp\left(-\frac{z_r}{\sigma_x}\right) + \exp\left(\frac{z_r}{\sigma_x}\right) + \exp\left(-\frac{z_r}{\sigma_x}\right) \mathrm{erf}\left(\frac{z_r}{\sigma_x} - \frac{\sigma_d}{\sigma_x}\right) - \exp\left(\frac{z_r}{\sigma_x}\right) \mathrm{erf}\left(\frac{z_r}{\sigma_x} + \frac{\sigma_d}{\sigma_x}\right) \right] \tag{36}$$

where *erf*(·) is the error function. Finally, after expressing the above equation in terms of ι_k and γ_k using Equations (33) and (34), we get:

$$p_{Z_r(k)}(z_r) = \frac{\sqrt{\gamma_k}\exp\left(\frac{1}{2\xi_k}\right)}{2\sqrt{2\xi_k}Y_k} \left[\exp\left(-\frac{\sqrt{\gamma_k}z_r}{Y_k\sqrt{\xi_k}}\right) + \exp\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k}\right) + \exp\left(-\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k}\right) \mathrm{erf}\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k} - \frac{1}{\sqrt{\xi_k}}\right) - \exp\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k}\right) \mathrm{erf}\left(\frac{\sqrt{\gamma_k}z_r}{\sqrt{\xi_k}Y_k} + \frac{1}{\sqrt{\xi_k}}\right) \right] \tag{37}$$

The probability density for the imaginary part, i.e., $p_{Z_i}(z_i)$, has exactly the same form as that of $p_{Z_r}(z_r)$. Assuming independence between Z_r and Z_i we get the following expression for the conditional density $p(Y(\omega_k) \mid H_1^k)$ at frequency bin ω_k:

$$p(Y(\omega_k) \mid H_1^k) = p(z_r z_i) = p_{Z_r(k)}(z_r) p_{Z_i(k)}(z_i) \tag{38}$$

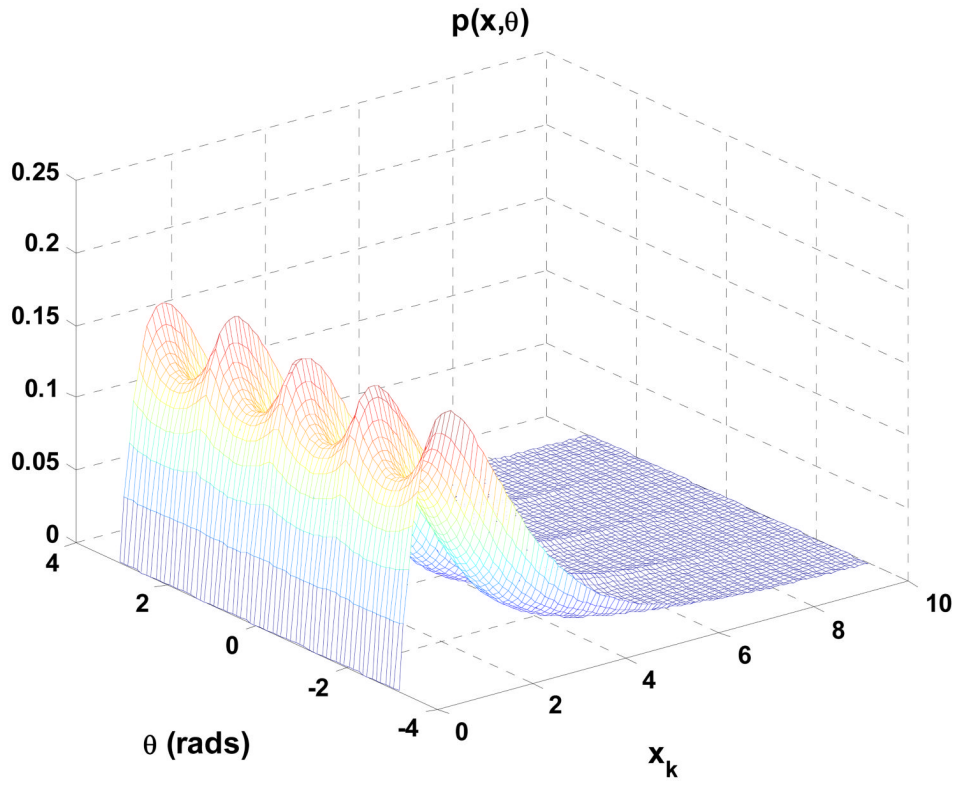

Fig. 1.
Plot of the joint density $p(x_k, \theta)$ of a zero mean complex Laplacian random variable ($\sigma^2 = 1$).

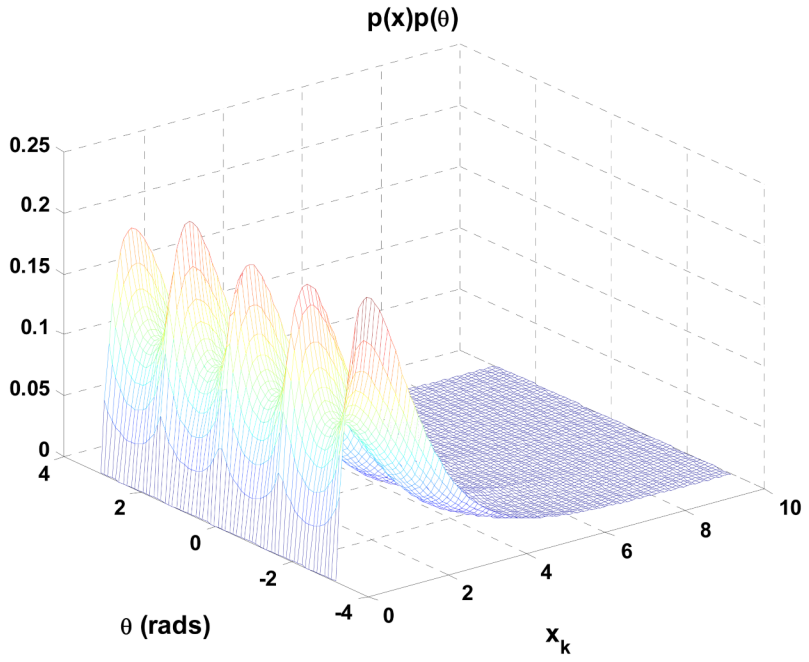

Fig. 2.
Plot of $p(x_k)p(\theta)$, where $p(x_k)$ is given by (9) and $p(\theta)$ is given by (8) [$\lambda_x(k) = 1$].

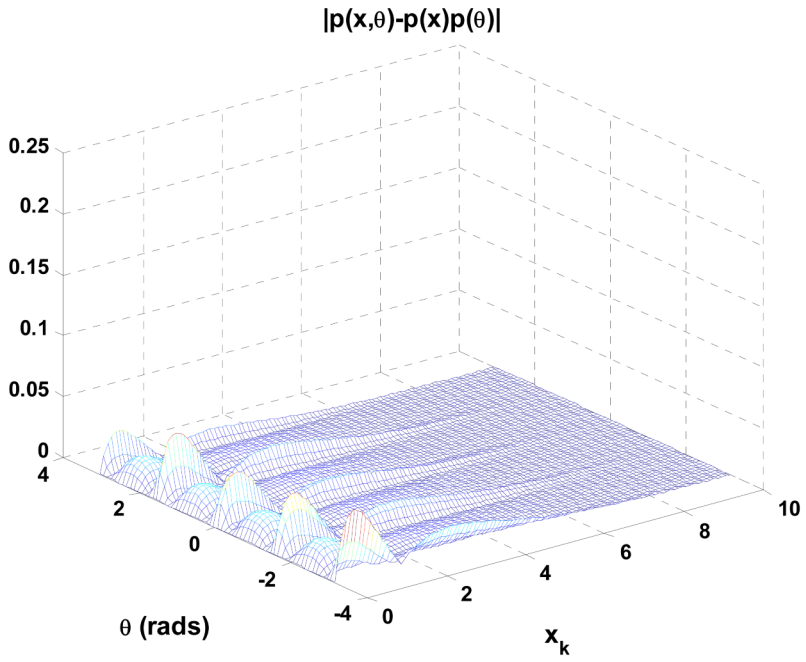

Fig. 3.
Plot of the absolute difference between the densities shown in Figs. 1 and 2.

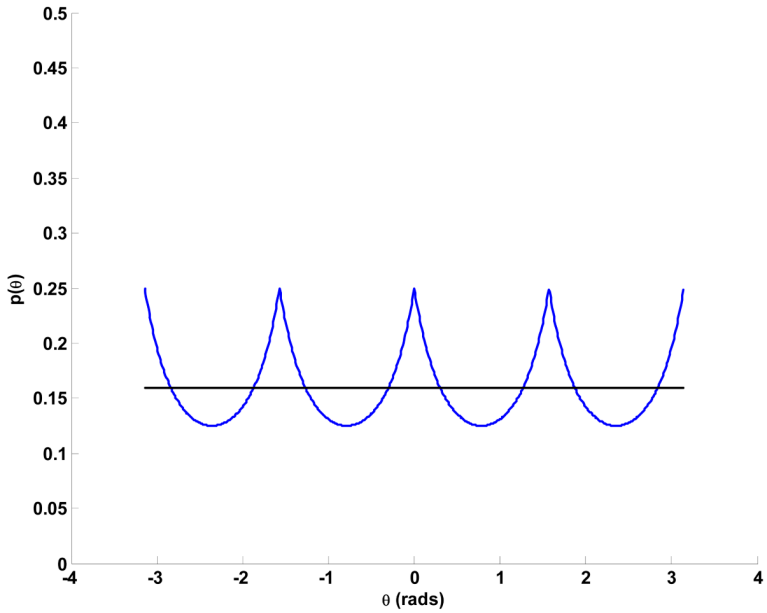

Fig. 4.
Plot of p(θ) (solid lines) superimposed to a uniform distribution (dashed lines).

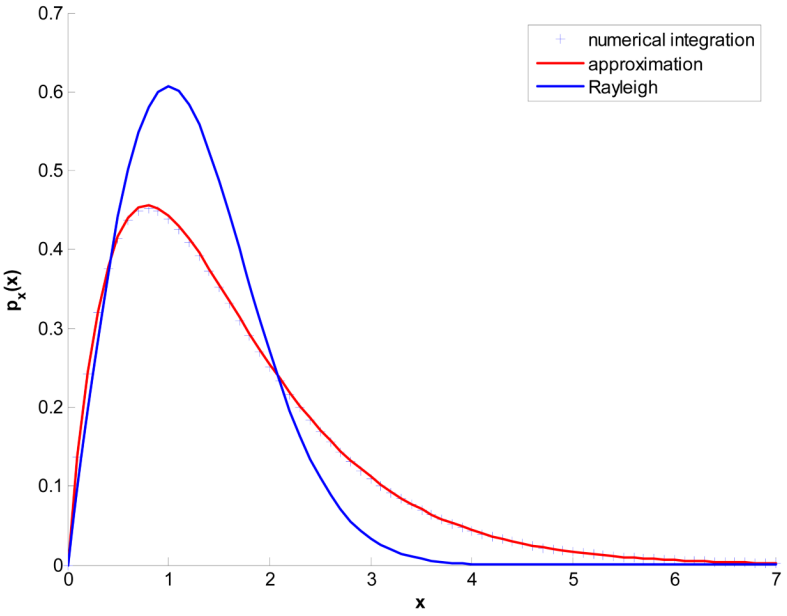

Fig. 5.

The pdf of the magnitude of the DFT coefficients assuming the real and imaginary parts are modeled by a Laplacian distribution ($\sigma^2 = 1$). The plot indicated by '+' shows the pdf computed by numerical integration of Eq. (23). The plot indicated by the solid line shows the pdf approximated by truncating the infinite summation in (6) with the first 40 terms. The Rayleigh distribution (dashed lines), used in the Gaussian-based MMSE estimator [1], is superimposed for comparative purposes.

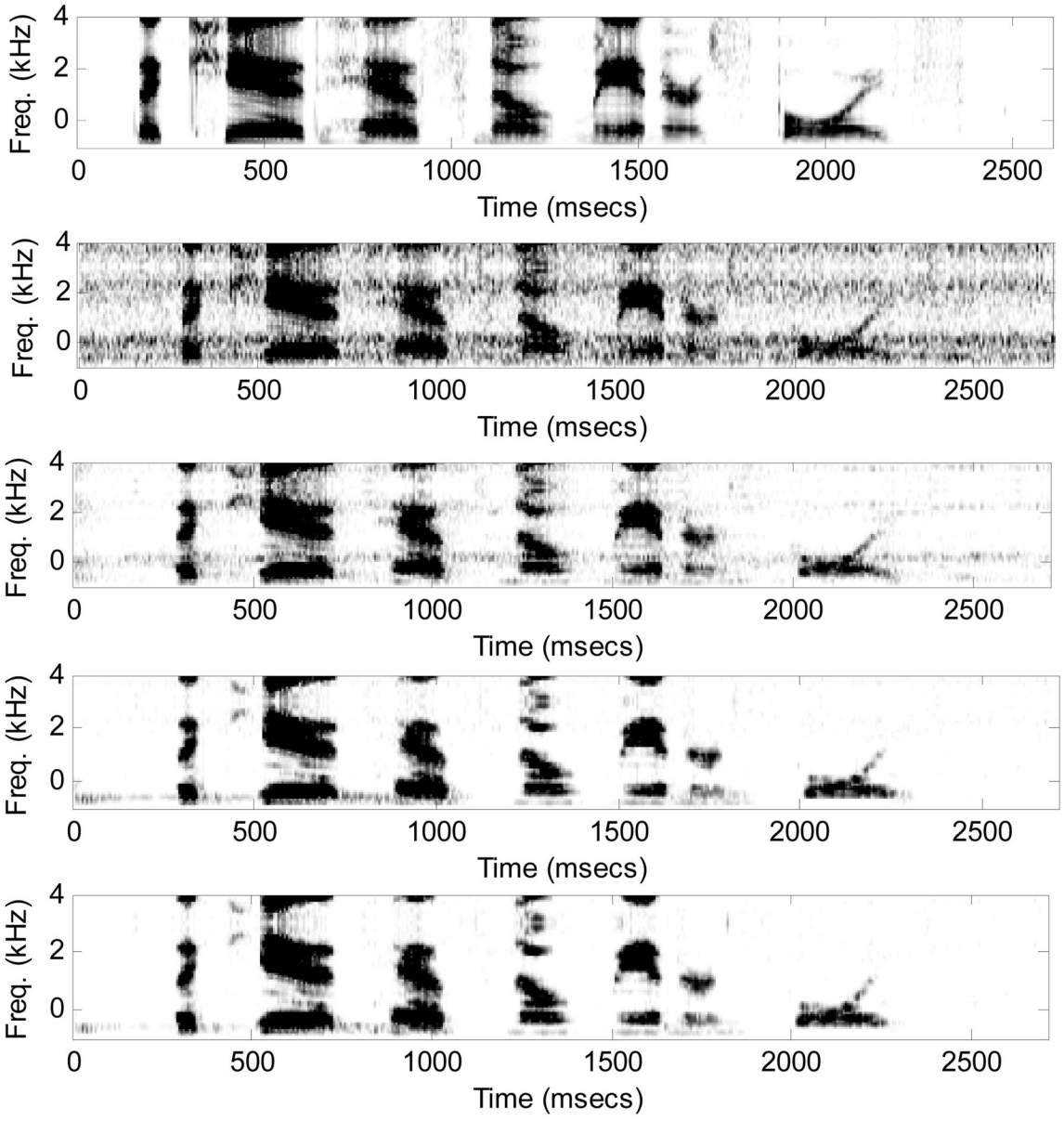

Fig. 6.

Spectrograms of a TIMIT sentence enhanced by the Gaussian and Laplacian MMSE estimators. From top to bottom, are the spectrograms of the signal in quiet, signal in noise, signal enhanced by the Gaussian MMSE estimator [1], signal enhanced by the LapMMSE estimator and signal enhanced by the ApLapMMSE estimator. All estimators incorporated signal-presence uncertainty.

TABLE I

Comparative performance, in terms of segmental SNR, of the Gaussian-based MMSE and Laplacian-based MMSE estimators. Numbers in top row indicate the input SNRseg/global input SNR.

Estimator	Speech-Shaped			F-16 Fighter		
	−12.4/0 dB	−7.4/5 dB	−2.35/10 dB	−12.4/0 dB	−7.4/5 dB	−2.35/10 dB
MMSE	0.863	3.712	5.733	1.524	4.127	6.687
LapMMSE	1.521	4.706	7.641	2.265	5.509	7.883
ApLapMMSE	1.492	4.784	7.482	2.179	5.432	7.854
MMSE-Lap [5]	1.891	4.124	6.658	3.022	4.812	7.251
MMSE-SPU	1.019	4.256	5.856	1.845	4.857	6.884
LapMMSE-SPU	1.915	5.218	7.785	2.684	5.998	7.995
ApLapMMSE-SPU	1.881	5.532	7.695	2.662	5.852	7.897

TABLE II

Comparative performance, in terms of log-likelihood ratio, of the Gaussian-based MMSE and Laplacian-based MMSE estimators.

Estimator	Speech-Shaped			F-16 Fighter		
	0 dB	5 dB	10 dB	0 dB	5 dB	10 dB
MMSE	1.433	1.11	0.852	1.094	0.86	0.676
LapMMSE	0.789	0.615	0.589	0.602	0.477	0.467
ApLapMMSE	0.783	0.618	0.603	0.598	0.479	0.479
MMSE-Lap [5]	1.017	0.772	0.610	0.870	0.693	0.554
MMSE-SPU	1.113	0.754	0.625	0.85	0.584	0.496
LapMMSE-SPU	0.756	0.598	0.584	0.577	0.464	0.463
ApLapMMSE-SPU	0.763	0.602	0.591	0.582	0.467	0.469

TABLE III

Comparative performance, in terms of PESQ, of the Gaussian-based MMSE and Laplacian-based MMSE estimators.

Estimator	Speech-Shaped			F-16 Fighter		
	0 dB	5 dB	10 dB	0 dB	5 dB	10 dB
MMSE	1.874	2.21	2.544	2.056	2.345	2.686
LapMMSE	1.882	2.448	2.783	2.19	2.587	2.914
ApLapMMSE	1.995	2.404	2.79	2.11	2.543	2.917
MMSE-Lap [5]	2.013	2.46	2.512	2.154	2.361	2.510
MMSE-SPU	1.952	2.315	2.634	2.03	2.346	2.70
LapMMSE-SPU	2.095	2.47	2.82	2.151	2.561	2.948
ApLapMMSE-SPU	1.997	2.439	2.81	2.146	2.563	2.943

TABLE IV

Comparative performance, in terms of segmental SNR, of the Gaussian-based MMSE and Laplacian-based MMSE estimators in white noise.

Estimator	White Noise		
	0 dB	5 dB	10 dB
MMSE-SPU	2.915	5.198	6.98
ApLapMMSE-SPU	3.587	5.945	7.875