SUPERGAUSSIAN GARCH MODELS FOR SPEECH SIGNALS

Israel Cohen

Department of Electrical Engineering, Technion - Israel Institute of Technology
Technion City, Haifa 32000, Israel
icohen@ee.technion.ac.il

ABSTRACT

In this paper, we introduce supergaussian *generalized autoregressive conditional heteroscedasticity* (GARCH) models for speech signals in the short-time Fourier transform (STFT) domain. We address the problem of speech enhancement, and show that estimating the variances of the STFT expansion coefficients based on GARCH models yields higher speech quality than by using the decision-directed method, whether the fidelity criterion is minimum mean-squared error (MMSE) of the spectral coefficients or MMSE of the log-spectral amplitude (LSA). Furthermore, while a Gaussian model is inferior to Gamma and Laplacian models when estimating the variances by the decision-directed method, a Gaussian model is superior when using the GARCH modeling method. This facilitates MMSE-LSA estimation, while taking into consideration the heavy-tailed distribution.

1. INTRODUCTION

Speech modeling in the short-time Fourier transform (STFT) domain underlies the design of many speech enhancement systems [1]. The Gaussian model, proposed by Ephraim and Malah [2], describes the individual STFT expansion coefficients of the speech signal as zero-mean statistically independent Gaussian random variables. It enables to derive useful minimum mean-squared error (MMSE) estimators for the short-term spectral amplitude (STSA), as well as the log-spectral amplitude (LSA) [2,3]. Porter and Boll [4] proposed to compute the optimal estimator directly from the speech data, rather than from a parametric model of the speech statistics. They argued that *a priori* speech spectra do not have a Gaussian distribution, but Gamma-like distribution. Martin [5] considered a Gamma speech model, under which the real and imaginary parts of the STFT coefficients are modeled as independent and identically distributed (iid) Gamma random variables. He assumed that distinct expansion coefficients are statistically independent, and derived their MMSE estimators. He showed that the Gamma model yields higher improvement in the segmental SNR than the Gaussian model.

Recently, we introduced a novel approach for statistically modeling speech signals in the STFT domain [6]. This approach is based on generalized autoregressive conditional heteroscedasticity (GARCH) modeling, which is widely-used for modeling the volatility of financial time-series such as exchange rates and stock returns [7]. Similar to financial time-series, speech signals in the STFT domain are characterized by heavy tailed distributions and volatility clustering. Specifically, when observing a time series of successive expansion coefficients in a fixed frequency bin, the expansion coefficients are clustered in the sense that large magnitudes tend to follow large magnitudes and small magnitudes tend to follow small magnitudes, while the phase is unpredictable.

This paper summarizes the main results of [8]. We present supergaussian GARCH models for speech signals in the STFT domain. We address the problem of spectral enhancement of noisy speech, and consider eight different speech enhancement algorithms, as summarized in Table 1. The statistical model is either Gaussian, Gamma or Laplacian; the spectral variance is estimated based on either the proposed GARCH models or the decision-directed method of Ephraim and Malah [2]; the fidelity criteria include MMSE of the STFT coefficients and MMSE of the LSA. We show that estimating the variance by the GARCH modeling method yields lower log-spectral distortion (LSD) and higher Perceptual Evaluation of Speech Quality (PESQ) scores (ITU-T P.862) than by using the decision-directed method. Furthermore, while a Gaussian model is inferior to Gamma and Laplacian models if the speech variance is estimated by the decision-directed method, a Gaussian model is superior in the case speech variance is estimated by using the GARCH modeling method. This facilitates MMSE-LSA estimation, while taking into consideration the heavy-tailed distribution. Speech spectrograms and informal listening tests confirm that the quality of the enhanced speech obtained by using the GARCH modeling method is better than that obtainable by using the decision-directed method.

In Sec. 2, we introduce the statistical models. In Sec. 3, we address the speech enhancement problem. In Sec. 4, we derive estimators for the spectral variances. Finally, in Sec. 5, we evaluate the performances of MMSE and MMSE-LSA estimators under Gaussian, Gamma and Laplacian models.

2. STATISTICAL MODELS

Let x and d denote speech and uncorrelated additive noise signals, and let $y = x + d$ represent the observed signal. Applying the STFT to the observed signal, we have in the time-frequency domain

$$Y_{tk} = X_{tk} + D_{tk} \tag{1}$$

where t is the time frame index ($t = 0, 1, \ldots$) and k is the frequency-bin index ($k = 0, 1, \ldots, K-1$). Let H_0^{tk} and H_1^{tk} denote, respectively, hypotheses of signal absence and presence in the noisy spectral coefficient Y_{tk}, and let $\lambda_{tk} \triangleq E\left\{|X_{tk}|^2 \mid H_1^{tk}\right\}$ denote the variance of a speech spectral coefficient X_{tk} under H_1^{tk}. Then, the variances $\{\lambda_{tk}\}$ are hidden from direct observation, in the sense that even under perfect conditions of zero noise, their values are not directly observable. Therefore, our approach is to assume that $\{\lambda_{tk}\}$ themselves are random variables, and to introduce *conditional* variances which are estimated from the avail-

Table 1. List of the Evaluated Speech Enhancement Algorithms.

Algorithm #	Statistical Model	Variance Estimation	Fidelity Criterion
1	Gaussian	GARCH	MMSE
2	Gamma	GARCH	MMSE
3	Laplacian	GARCH	MMSE
4	Gaussian	Decision-Directed	MMSE
5	Gamma	Decision-Directed	MMSE
6	Laplacian	Decision-Directed	MMSE
7	Gaussian	GARCH	MMSE-LSA
8	Gaussian	Decision-Directed	MMSE-LSA

able information (*e.g.*, the clean spectral coefficients through frame $t-1$, or the noisy spectral coefficients through frame t).

Let $\mathcal{X}_0^\tau = \{X_{tk} \,|\, t = 0, \ldots, \tau,\ k = 0, \ldots, K-1\}$ represent the set of clean speech spectral coefficients up to frame τ, and let $\lambda_{tk|\tau} \triangleq E\left\{|X_{tk}|^2 \,|\, H_1^{tk},\, \mathcal{X}_0^\tau\right\}$ denote the *conditional* variance of X_{tk} under H_1^{tk} given $\mathcal{X}_0^\tau$. Our statistical models in the STFT domain rely on the following set of assumptions:

1. The speech spectral coefficients $\{X_{tk}\}$ are generated by

$$X_{tk} = \sqrt{\lambda_{tk}}\, V_{tk} \tag{2}$$

where $\left\{V_{tk} \,|\, H_0^{tk}\right\}$ are identically zero, and $\left\{V_{tk} \,|\, H_1^{tk}\right\}$ are statistically independent complex random variables with zero mean, unit variance, and iid real and imaginary parts:

$$\begin{aligned} H_1^{tk}: &\quad E\{V_{tk}\} = 0\,,\ E\left\{|V_{tk}|^2\right\} = 1\,, \\ H_0^{tk}: &\quad V_{tk} = 0\,. \end{aligned}$$

2. The probability density function (pdf) of V_{tk} under H_1^{tk} is determined by the specific statistical model. Let $V_{Rtk} = \Re\{V_{tk}\}$ and $V_{Itk} = \Im\{V_{tk}\}$ denote, respectively, the real and imaginary parts of V_{tk}. Let $p\left(V_{\rho tk} \,|\, H_1^{tk}\right)$ denote the pdf of $V_{\rho tk}$ ($\rho \in \{R, I\}$) under H_1^{tk}. Then, for a Gaussian model

$$p\left(V_{\rho tk} \mid H_1^{tk}\right) = \frac{1}{\sqrt{\pi}} \exp\left(-V_{\rho tk}^2\right)\,, \tag{3}$$

for a Gamma model

$$p\left(V_{\rho tk} \mid H_1^{tk}\right) = \frac{\sqrt[4]{6}}{2\sqrt{2\pi|V_{\rho tk}|}} \exp\left(-\sqrt{\frac{3}{2}}\,|V_{\rho tk}|\right)\,, \tag{4}$$

and for a Laplacian model

$$p\left(V_{\rho tk} \mid H_1^{tk}\right) = \exp\left(-2\,|V_{\rho tk}|\right)\,. \tag{5}$$

3. The conditional variance $\lambda_{tk|t-1}$, referred to as the *one-frame-ahead conditional variance*, is a random process which evolves as a $\mathrm{GARCH}(1,1)$ process:

$$\lambda_{tk|t-1} = \lambda_{\min} + \mu\,|X_{t-1,k}|^2 + \delta\left(\lambda_{t-1,k|t-2} - \lambda_{\min}\right) \tag{6}$$

where

$$\lambda_{\min} > 0\,, \quad \mu \geq 0\,, \quad \delta \geq 0\,, \quad \mu + \delta < 1 \tag{7}$$

are the standard constraints imposed on the parameters of the GARCH model [7]. The parameters μ and δ are, respectively, the moving average and autoregressive parameters of the GARCH(1,1) model, and $\lambda_{\min}$ is a lower bound on the variance of X_{tk} under H_1^{tk}.

The first assumption implies that the speech spectral coefficients $\left\{X_{tk} \,|\, H_1^{tk}\right\}$ are conditionally zero-mean statistically independent random variables given their variances $\{\lambda_{tk}\}$. The real and imaginary parts of X_t under H_1^t are conditionally iid random variables given λ_{tk}.

3. SPECTRAL ENHANCEMENT OF NOISY SPEECH

In this section, we address the problem of spectral enhancement of noisy speech under the proposed statistical models. Let

$$d\left(X_{tk}, \hat{X}_{tk}\right) = \left|g(\hat{X}_{tk}) - \tilde{g}(X_{tk})\right|^2 \tag{8}$$

denote a distortion measure between X_{tk} and its estimate $\hat{X}_{tk}$, where $g(X)$ and $\tilde{g}(X)$ are specific functions of X (*e.g.*, X, $|X|$, $\log|X|$, $e^{j\angle X}$). Let $\hat{p}_{tk}$ denote an estimate for the signal presence probability, and let $\hat{\lambda}_{tk}$ denote an estimate for λ_{tk}. Then, the design of a particular estimator for X_{tk} requires the following specifications:

- Functions $g(X)$ and $\tilde{g}(X)$, which determine the fidelity criterion of the estimator.
- A conditional pdf $p\left(X_{tk} \mid \lambda_{tk}\,,\, H_1^{tk}\right)$ for X_{tk} under H_1^{tk} given its variance λ_{tk}, which determines the statistical model.
- Estimators $\hat{\lambda}_{tk}$ and $\widehat{\sigma_{tk}^2}$ for the speech and noise spectral variances, respectively.
- An estimator $\hat{p}_{tk}$ for the signal presence probability.

In this work we assume knowledge of the noise variance σ_{tk}^2, which in practice can be estimated by using the *Minima Controlled Recursive Averaging* approach [9]. Furthermore, to simplify the comparisons between the speech enhancement algorithms, we focus on implementations that assume speech presence (*i.e.*, $\hat{p}_{tk} = 1$) whenever $20\log_{10}|X_{tk}| > \epsilon$, where $\epsilon = \max_{tk}\{20\log_{10}|X_{tk}|\} - 50$ confines the dynamic range of the log-spectrum to 50 dB. In the other time-frequency bins, $\hat{p}_{tk}$ is set to zero. We consider MMSE estimators for the spectral coefficients under Gaussian, Gamma and Laplacian models [5, 10], and MMSE-LSA estimator under a Gaussian model [1, 3]. An MMSE estimator is obtained by using the functions

$$g(\hat{X}_{tk}) = \hat{X}_{tk}\,, \quad \tilde{g}(X_{tk}) = \begin{cases} X_{tk}\,, & \text{under } H_1^{tk}\,, \\ G_{\min}\,Y_{tk}\,, & \text{under } H_0^{tk}\,, \end{cases} \tag{9}$$

where $G_{\min} \ll 1$ represents a constant attenuation factor. An MMSE-LSA estimator is obtained by using the functions

$$g(\hat{X}_{tk}) = \log|\hat{X}_{tk}|, \quad \tilde{g}(X_{tk}) = \begin{cases} \log|X_{tk}|, & \text{under } H_1^{tk}, \\ \log\left(G_{\min}|Y_{tk}|\right), & \text{under } H_0^{tk}. \end{cases} \tag{10}$$

Estimators $\hat{X}_{tk}$, which minimize the expected distortion given $\hat{p}_{tk}$, $\hat{\lambda}_{tk}$ and Y_{tk}, are calculated from

$$\begin{aligned} g(\hat{X}_{tk}) &= E\left\{\tilde{g}(X_{tk}) \,\middle|\, \hat{p}_{tk}\,, \hat{\lambda}_{tk}\,, Y_{tk}\right\} \\ &= \hat{p}_{tk}\,E\left\{\tilde{g}(X_{tk}) \,\middle|\, H_1^{tk}\,, \hat{\lambda}_{tk}\,, Y_{tk}\right\} \\ &\quad +(1-\hat{p}_{tk})\,E\left\{\tilde{g}(X_{tk}) \,\middle|\, H_0^{tk}\,, Y_{tk}\right\}\,. \end{aligned} \tag{11}$$

Table 2. Log-Spectral Distortion and PESQ Scores Obtained by Using Different Variance Estimation Methods (GARCH Modeling Method vs. Decision-Directed Method), Statistical Models (Gaussian vs. Gamma vs. Laplacian) and Fidelity Criteria (MMSE vs. MMSE-LSA).

	Input SNR [dB]	GARCH modeling method				Decision-Directed method			
		Gaussian		Gamma	Laplacian	Gaussian		Gamma	Laplacian
		MMSE	MMSE-LSA	MMSE	MMSE	MMSE	MMSE-LSA	MMSE	MMSE
Log-Spectral Distortion	0	7.77	**4.85**	8.03	7.91	18.89	11.35	17.76	18.14
	5	5.78	**4.04**	6.93	6.45	17.29	11.03	15.73	16.26
	10	4.14	**3.27**	5.35	4.85	13.87	9.13	11.83	12.48
	15	2.50	**2.25**	3.23	2.92	9.19	6.05	6.95	7.59
	20	1.30	**1.28**	1.55	1.44	4.88	3.13	2.88	3.34
PESQ Scores	0	2.52	**2.55**	2.47	2.48	1.91	2.21	1.98	1.96
	5	2.97	**2.98**	2.90	2.91	2.30	2.61	2.38	2.36
	10	3.37	**3.38**	3.28	3.31	2.70	2.99	2.77	2.75
	15	3.67	**3.69**	3.59	3.62	3.09	3.31	3.17	3.15
	20	3.88	**3.89**	3.83	3.85	3.53	3.64	3.62	3.60

The speech spectral variance is estimated based on the proposed GARCH models, as described in following section.

4. VARIANCE ESTIMATION USING GARCH MODELS

The speech variance estimation follows the rational of Kalman filtering. We start with an estimate $\hat{\lambda}_{tk|t-1}$ that relies on the noisy observations up to frame $t-1$, and "update" the variance by using the additional information Y_{tk}. Then, the variance is "propagated" ahead in time to obtain a conditional variance estimate at frame $t+1$ from the information available at frame t. The propagation and update steps are iterated, to recursively estimate the speech variances as new data arrive.

Assuming an estimate $\hat{\lambda}_{tk|t-1}$ for the one-frame-ahead conditional variance of X_{tk} is available, an estimate for $\lambda_{tk|t}$ can be obtained by calculating its conditional mean under H_1^{tk} given Y_{tk} and $\hat{\lambda}_{tk|t-1}$. By definition, $\lambda_{tk|t} = |X_{tk}|^2 = X_{Rtk}^2 + X_{Itk}^2$. Hence,

$$\hat{\lambda}_{tk|t} = \sum_{\rho \in \{R,I\}} E\left\{ X_{\rho tk}^2 \;\middle|\; H_1^{tk}, \hat{\lambda}_{tk|t-1}, Y_{\rho tk} \right\}. \quad (12)$$

Defining the *a priori* and *a posteriori* signal-to-noise ratios (SNRs), respectively, by

$$\xi_{tk|t-1} \triangleq \frac{\lambda_{tk|t-1}}{\sigma_{tk}^2}, \quad \gamma_{\rho tk} \triangleq \frac{Y_{\rho tk}^2}{\sigma_{tk}^2}, \quad (13)$$

we can write for $Y_{\rho tk} \neq 0$

$$E\left\{ X_{\rho tk}^2 \;\middle|\; H_1^{tk}, \hat{\lambda}_{tk|t-1}, Y_{\rho tk} \right\} = G_{\mathrm{SP}}\left(\hat{\xi}_{tk|t-1}, \gamma_{\rho tk}\right) Y_{\rho tk}^2 \quad (14)$$

where the specific expression for $G_{\mathrm{SP}}(\xi, \gamma_\rho)$, representing the MMSE gain function in the spectral power domain, depends on the particular statistical model [8]. Equation (14) does not hold in the case $Y_{\rho tk} \to 0$, since $G_{\mathrm{SP}}(\xi, \gamma_\rho) \to \infty$ as $\gamma_\rho \to 0$, and the conditional variance of $X_{\rho tk}$ is generally not zero. However, we can define a function $f(\lambda, \sigma^2, Y_\rho^2)$ such that

$$E\left\{ X_{\rho tk}^2 \;\middle|\; H_1^{tk}, \hat{\lambda}_{tk|t-1}, Y_{\rho tk} \right\} = f\left(\hat{\lambda}_{tk|t-1}, \sigma_{tk}^2, Y_{\rho tk}^2\right) \quad (15)$$

for all $Y_{\rho tk}$ [8]. Substituting (15) into (12), we obtain the update step of the recursive estimation given by

$$\hat{\lambda}_{tk|t} = f\left(\hat{\lambda}_{tk|t-1}, \sigma_{tk}^2, Y_{Rtk}^2\right) + f\left(\hat{\lambda}_{tk|t-1}, \sigma_{tk}^2, Y_{Itk}^2\right). \quad (16)$$

To formulate the propagation step, we assume that we are given at frame $t-1$ an estimate $\hat{\lambda}_{t-1,k|t-2}$ for the conditional variance of $X_{t-1,k}$, which has been obtained from the noisy measurements up to frame $t-2$. Then a recursive MMSE estimate for $\lambda_{tk|t-1}$ can be obtained by calculating its conditional mean under $H_1^{t-1,k}$ given $\hat{\lambda}_{t-1,k|t-2}$ and $Y_{t-1,k}$:

$$\hat{\lambda}_{tk|t-1} = E\left\{ \lambda_{tk|t-1} \;\middle|\; H_1^{t-1,k}, \hat{\lambda}_{t-1,k|t-2}, Y_{t-1,k} \right\}. \quad (17)$$

Substituting (6) into (17) and employing (12), we obtain

$$\hat{\lambda}_{tk|t-1} = \lambda_{\min} + \mu\, \hat{\lambda}_{t-1,k|t-1} + \delta\left(\hat{\lambda}_{t-1,k|t-2} - \lambda_{\min}\right). \quad (18)$$

Equation (18) is the propagation step, since the conditional variance estimates are propagated ahead in time to obtain a conditional variance estimate at frame t from the information available at frame $t-1$. The propagation and update steps are iterated as new data arrive, following the rational of Kalman filtering.

5. EXPERIMENTAL RESULTS AND DISCUSSION

The performances of the MMSE spectral and LSA estimators were evaluated under Gaussian, Gamma and Laplacian models, while the speech variance is estimated by using either the GARCH modeling or the decision-directed method. The evaluation includes two objective quality measures, and informal listening tests. The first quality measure is log-spectral distortion, in dB, which is defined by

$$\mathrm{LSD} = \left[\frac{1}{|\mathcal{H}_1|} \sum_{tk \in \mathcal{H}_1} \left(20 \log_{10} |X_{tk}| - 20 \log_{10} |\hat{X}_{tk}|\right)^2 \right]^{\frac{1}{2}} \quad (19)$$

where $\mathcal{H}_1 = \{tk \mid 20 \log_{10} |X_{tk}| > \epsilon\}$ denotes the set of time-frequency bins which contain the speech signal, $|\mathcal{H}_1|$ denotes its

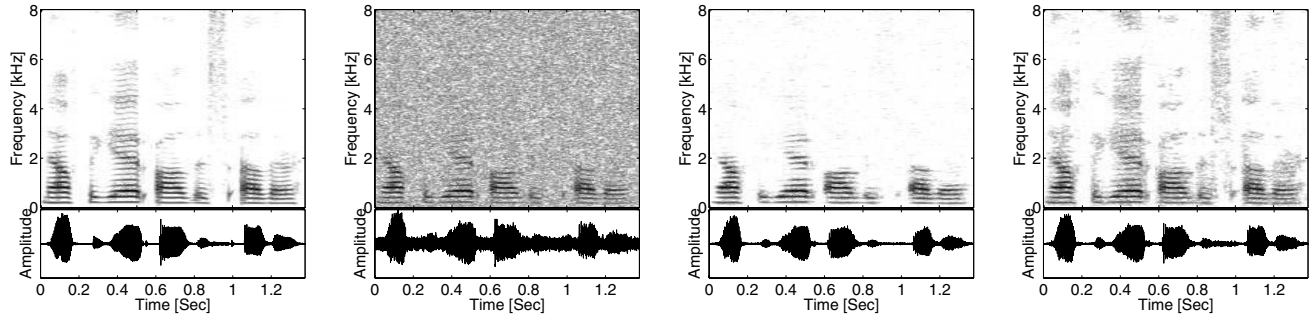

Fig. 1. Speech spectrograms and waveforms. (a) Original clean speech signal: "Now forget all this other."; (b) noisy signal (SNR $= 5$ dB, LSD $= 13.75$ dB, PESQ= 1.76); (c) speech reconstructed by using the decision-directed method, a Gaussian model and MMSE-LSA estimator (LSD $= 9.00$ dB, PESQ $= 2.57$); (d) speech reconstructed by using the GARCH modeling method, a Gaussian model and MMSE-LSA estimator (LSD $= 3.59$ dB, PESQ $= 2.88$).

cardinality, and $\epsilon = \max_{tk} \{20 \log_{10} |X_{tk}|\} - 50$ confines the dynamic range of the log-spectrum to 50 dB. The second quality measure is the PESQ score (ITU-T P.862).

The speech signals, taken from the TIMIT database, include 20 different utterances from 20 different speakers, half male and half female. The signals are sampled at 16 kHz, degraded by white Gaussian noise with SNRs in the range $[0, 20]$ dB, and transformed into the STFT domain using half overlapping Hamming analysis windows of 32 milliseconds length. Maximum-likelihood estimates of the model parameters (*i.e.*, $\hat{\mu}$, $\hat{\delta}$ and $\hat{\lambda}_{\min}$) are calculated independently for each speaker from the clean signal of that speaker, as described in [8]. Eight different speech enhancement algorithms are then applied to each noisy speech signal, as summarized in Table 1.

Table 2 shows the results of the LSD and PESQ scores obtained by using the different algorithms for various SNR levels. The results show that:

- MMSE-LSA estimators yield lower LSD and higher PESQ scores than MMSE spectral estimators, whether the variance is estimated by using the GARCH modeling method or the decision-directed method.
- An MMSE spectral estimator derived under a Gamma statistical model performs better than that derived under Gaussian or Laplacian models, but only if the speech variance is estimated by the decision-directed method. However, if the speech variance is estimated by using the GARCH modeling method, a Gaussian model is preferable to Gamma and Laplacian models.
- Speech variance estimation based on GARCH modeling yields lower LSD and higher PESQ scores than those obtained by using the decision-directed method.
- The best performance is obtained when using the GARCH modeling method, a Gaussian model and an MMSE-LSA estimator. The worst performance is obtained when using the decision-directed method, a Gaussian model and an MMSE spectral estimator.

A subjective study of speech spectrograms and informal listening tests confirm that the quality of the enhanced speech obtained by using the GARCH modeling method, the MMSE-LSA estimator and the Gaussian model is significantly better than that obtainable by using the decision-directed method. Figure 1 demonstrates the spectrograms and waveforms of a clean signal, noisy signal (SNR $= 5$ dB) and enhanced speech signals obtained by using the GARCH modeling and the decision-directed methods. It shows that weak speech components are better preserved by using the GARCH modeling method than by using the decision-directed method.

6. REFERENCES

[1] Y. Ephraim and I. Cohen, "Recent advancements in speech enhancement," in *The Electrical Engineering Handbook*, 3rd ed. CRC Press, to be published. [Online]. Available: http:// ece.gmu.edu/∼yephraim/ ephraim.html

[2] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-32, no. 6, pp. 1109–1121, December 1984.

[3] ——, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-33, pp. 443–445, April 1985.

[4] J. Porter and S. Boll, "Optimal estimators for spectral restoration of noisy speech," in *Proc. IEEE Internat. Conf. Acoust. Speech, Signal Process. (ICASSP)*, San Diego, California, 19–21 March 1984, pp. 18A.2.1–18A.2.4.

[5] R. Martin, "Speech enhancement using MMSE short time spectral estimation with gamma distributed speech priors," in *Proc. 27th IEEE Internat. Conf. Acoust. Speech Signal Process., ICASSP-02*, Orlando, Florida, 13–17 May 2002, pp. I–253–I–256.

[6] I. Cohen, "Modeling speech signals in the time-frequency domain using GARCH," *Signal Processing*, vol. 84, no. 12, pp. 2453–2459, December 2004.

[7] T. Bollerslev, R. Y. ChouKenneth, and F. Kroner, "ARCH modeling in finance: A review of the theory and empirical evidence," *Journal of Econometrics*, vol. 52, no. 1-2, pp. 5–59, April-May 1992.

[8] I. Cohen, "Speech spectral modeling and enhancement based on autoregressive conditional heteroscedasticity models," submitted to *Signal Processing*.

[9] ——, "Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging," *IEEE Trans. Speech and Audio Processing*, vol. 11, no. 5, pp. 466–475, September 2003.

[10] R. Martin and C. Breithaupt, "Speech enhancement in the DFT domain using Laplacian speech priors," in *Proc. 8th Internat. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Kyoto, Japan, 8–11 September 2003, pp. 87–90.