A Non-Iterative Kalman Filtering Algorithm with Dynamic Gain Adjustment for Single-Channel Speech Enhancement

Stephen So, Aidan E. W. George, Ratna Ghosh, and Kuldip K. Paliwal
Signal Processing Laboratory, Griffith School of Engineering, Griffith University, Brisbane, Australia
Email: {s.so, a.george, k.paliwal}@griffith.edu.au, ratnand@gmail.com

***Abstract*—In this paper, we present a non-iterative Kalman filtering algorithm that applies a dynamic adjustment factor on the Kalman filter gain to alleviate the negative effects of estimating speech model parameters from noise-corrupted speech. These poor estimates introduce a bias in the first component of the Kalman gain vector, particularly during the silent (non-speech) regions, resulting in a significant level of residual noise in the enhanced speech. The proposed dynamic gain adjustment algorithm utilises a recently developed metric for quantifying the level of robustness in the Kalman filter. Objective and human subjective listening tests on the NOIZEUS speech database were performed. The results showed that the output speech from the proposed algorithm has improved quality over the non-iterative Kalman filter that uses noisy model estimates and is competitive with the MMSE-STSA method.**

***Index Terms*—speech enhancement, Kalman filter, robustness metric**

I. INTRODUCTION

The presence of corrupting noise in speech signals is a problem that degrades the performance of many speech applications such as speech coding, speaker recognition and speech recognition. Speech enhancement algorithms aim to suppress the level of corrupting noise in a speech signal so that the quality and intelligibility of the speech are improved. Speech enhancement is useful in many applications where corruption by noise is undesirable and unavoidable. Several speech enhancement methods have been reported in the literature, which include spectral subtraction [1], MMSE-STSA (minimum mean square error, short-term spectral amplitude) estimation [2], Wiener filtering [3], subspace methods [4], and Kalman filtering (in both acoustic [5] and modulation domains [6], [7].

The Kalman filter is an unbiased linear Minimum Mean Squared Error (MMSE) estimator, where the state vectors of a dynamic system are estimated using a linear combination of noise-corrupted observations and predicted state vectors. The predicted state vectors are computed as the output of a speech production model. Ideally, the speech model parameters, which are also known as Linear Prediction Coefficients (LPCs), should be estimated from the clean speech, which we refer to as the *oracle case*. However, in practice, the LPCs need to be estimated from the noise-corrupted speech. The presence of noise in the speech causes the LPC estimates to be poor, which adversely affects the performance of the Kalman filter. This is especially noticeable during the silent regions (i.e. when there is no speech), where the residual noise level is significantly high. Methods for dealing with poor estimates have been investigated previously, such as iterative Kalman filtering [8] where the LPCs are re-estimated using Kalman filter enhanced speech from a previous step or estimating parameters from speech that has been enhanced by another speech enhancement algorithm [9]. These methods have the disadvantage of computational delay due to the iterative process as well as the suffering from musical noise due to the variance of the parameter estimates. In a previous study [10], we had identified how the bias in poor model parameters manifested itself during the operation of the Kalman filter and presented a simple technique of Chebychev windowing to reduce this bias. This study also showed that the temporal trajectory of the Kalman filter gain component is a useful indicator of Kalman filter performance.

In this paper, we propose an algorithm for reducing the bias in the non-iterative Kalman filter gain that is dynamic in nature, as opposed to the static adjustment provided by tapered windowing. This novel algorithm modifies the Kalman filter gain by utilising a robustness metric that was recently proposed by Saha, *et al.* [11] in the instrumentation literature. Objective tests and human subjective listening tests were performed using the NOIZEUS speech corpus [12] to evaluate the proposed algorithm for white noise and compare its performance against conventional Kalman filtering (both oracle and non-oracle) as well as the MMSE-STSA (short-time spectral amplitude) method.

II. PROPOSED KALMAN FILTER FOR SPEECH ENHANCEMENT

A. *The Kalman Recursive Equations*

If the clean speech is represented as $x(n)$ and the noise signal as $v(n)$, then the noise-corrupted speech $y(n)$, which is the only observable signal in practice, is expressed as:

Manuscript received June 12, 2015; revised September 8, 2015.

doi: 10.18178/ijsps.4.4.263-268

$$y(n) = x(n) + v(n) \tag{1}$$

In the Kalman filter that is used for speech enhancement [5], $v(n)$ is a zero-mean, white Gaussian random noise that has a variance of σ_v^2 and is uncorrelated with $x(n)$. A pth order linear predictor (also referred to as an autoregressive model) is used to model the speech signal:

$$x(n) = -\sum_{k=1}^{p} a_k x(n-k) + w(n) \tag{2}$$

where a_k are the LPCs and $w(n)$ is the excitation signal, which is assumed to be a white and Gaussian random noise. By combining the speech production model with the noise corruption (1), we can represent the process as a state-space equation:

$$\boldsymbol{x}(n) = \boldsymbol{A}\boldsymbol{x}(n-1) + \boldsymbol{d}w(n) \tag{3}$$

$$y(n) = \boldsymbol{c}^T\boldsymbol{x}(n) + v(n) \tag{4}$$

where $\boldsymbol{A}$ is the state transition matrix containing the model parameters, $\boldsymbol{x}(n) = [x(n)\ x(n\text{-}1)\ \ldots\ x(n\text{-}p\text{+}1)]^T$ is the state vector, $\boldsymbol{d} = [1\ 0\ \ldots\ 0]^T$ and $\boldsymbol{c} = [1\ 0\ \ldots\ 0]^T$ are the measurement vectors for the excitation noise and observation, respectively [13].

The Kalman filter recursively computes an unbiased and linear MMSE estimate $\hat{\boldsymbol{x}}(n|n)$ of the state vector at time n given the noise observation $y(n)$, by using the following equations [13]:

$$\hat{\boldsymbol{x}}(n|n-1) = \boldsymbol{A}\hat{\boldsymbol{x}}(n-1|n-1) \tag{5}$$

$$\boldsymbol{P}(n|n-1) = \boldsymbol{A}\boldsymbol{P}(n-1|n-1)\boldsymbol{A}^T + \sigma_w^2\boldsymbol{d}\boldsymbol{d}^T \tag{6}$$

$$\boldsymbol{K}(n) = \boldsymbol{P}(n|n-1)\boldsymbol{c}[\sigma_v^2 + \boldsymbol{c}^T\boldsymbol{P}(n|n-1)\boldsymbol{c}]^{-1} \tag{7}$$

$$\hat{\boldsymbol{x}}(n|n) = \hat{\boldsymbol{x}}(n|n-1) + \boldsymbol{K}(n)[y(n) - \boldsymbol{c}^T\hat{\boldsymbol{x}}(n|n-1)] \tag{8}$$

$$\boldsymbol{P}(n|n) = [\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{c}^T]\boldsymbol{P}(n|n-1) \tag{9}$$

During the operation of the Kalman filter, the noise-corrupted speech $y(n)$ is windowed into non-overlapped and short (e.g. 20ms) frames and the LPCs and excitation signal variance σ_w^2 are estimated, with the latter given by:

$$\sigma_w^2 = R_{xx}(0) + \sum_{k=1}^{p} a_k R_{xx}(k) \tag{10}$$

where $R_{xx}(k)$ is the kth autocorrelation coefficient of the speech signal $x(n)$. The LPCs and excitation variance remain constant during the Kalman filtering of speech samples in the frame, while the Kalman gain vector $\boldsymbol{K}(n)$, *a posteriori* error covariance $\boldsymbol{P}(n|n)$ and state vector estimate $\hat{\boldsymbol{x}}(n|n)$ are continually updated on a sample-by-sample basis.

B. The Effect of Biased Speech Parameter Estimates

In Kalman filtering for speech enhancement, the enhanced speech sample is found by computing $\boldsymbol{c}^T\boldsymbol{x}(n)$, which extracts the first scalar component of the state vector.

We can simplify the analysis of the Kalman recursive equations by rewriting (8) in scalar form:

$$\hat{x}(n|n) = \boldsymbol{c}^T\hat{\boldsymbol{x}}(n|n) \tag{11}$$

$$= \boldsymbol{c}^T\hat{\boldsymbol{x}}(n|n-1) + \boldsymbol{c}^T\boldsymbol{K}(n)[y(n) - \boldsymbol{c}^T\hat{\boldsymbol{x}}(n|n-1)] \tag{12}$$

$$= [1 - K(n)]\hat{x}(n|n-1) + K(n)y(n) \tag{13}$$

In order to examine the effect of biased speech parameter estimates on the performance of the Kalman filter, we will first examine the computation of the Kalman gain $K(n)$. It can be shown that the inverse bracketed expression (also known as the innovation covariance) in (7) can be re-written in terms of scalar quantities:

$$[\sigma_v^2 + \boldsymbol{c}^T\boldsymbol{P}(n|n-1)\boldsymbol{c}]^{-1} = \frac{1}{\alpha^2(n)+\sigma_w^2+\sigma_v^2} \tag{14}$$

where $\alpha^2(n) = \boldsymbol{c}^T\boldsymbol{A}\boldsymbol{P}(n-1|n-1)\boldsymbol{A}^T\boldsymbol{c}$ represents the transmission of *a posteriori* error variances by the speech model from the previous time sample. Therefore, using similar scalar substitutions, the Kalman gain can be expressed as:

$$K(n) = \frac{\alpha^2(n)+\sigma_w^2}{\alpha^2(n)+\sigma_w^2+\sigma_v^2} \tag{15}$$

For the special case of no measurement noise present (i.e. $\sigma_v^2 = 0$), the Kalman gain is unity and according to (13), the output speech sample is equal to the observed speech $y(n)$. When we are in the silent regions where there is no speech present (i.e. $\sigma_w^2 = 0$ and $\alpha^2(n) = 0$), then $K(n)$=0 so no corrupting noise from $y(n)$ is passed to the output.

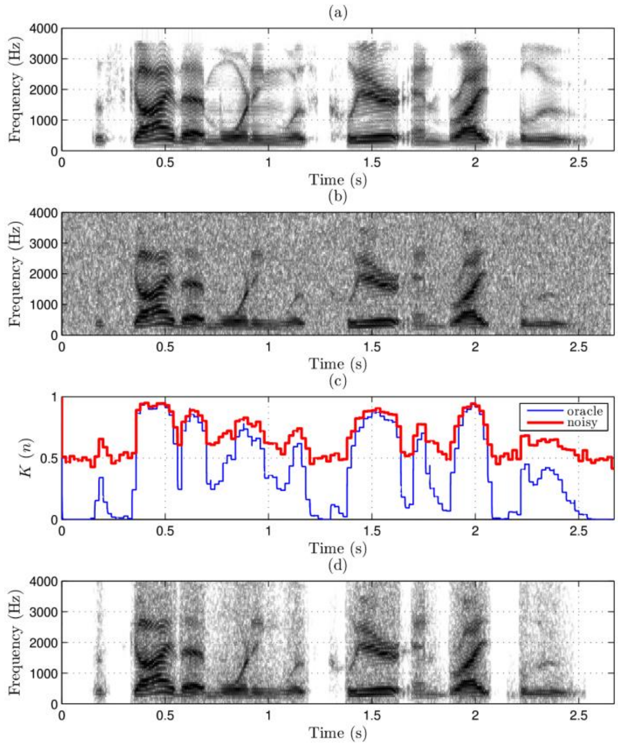
When the LPC parameters ($\{a_k\}$ and σ_w^2) are estimated from noise-corrupted speech, these estimates become biased ($\{\tilde{a}_k\}$ and $\tilde{\sigma}_w^2$). Assuming that the noise $v(n)$ is zero-mean, white and additive, using (10), it can be shown that the biased excitation variance is approximately:

$$\tilde{\sigma}_w^2 \approx \sigma_w^2 + \sigma_v^2 \tag{16}$$

After substituting this into (15), we obtain the resulting Kalman gain:

$$K(n) \approx \frac{\tilde{\alpha}^2(n)+\sigma_w^2+\sigma_v^2}{\tilde{\alpha}^2(n)+\sigma_w^2+2\sigma_v^2} \tag{17}$$

When we consider the case of the Kalman filter operating in the silent regions (i.e. $\sigma_w^2 = 0$ and $\tilde{\alpha}^2(n) = 0$), then (18) predicts that the Kalman gain would be biased at approximately 0.5.



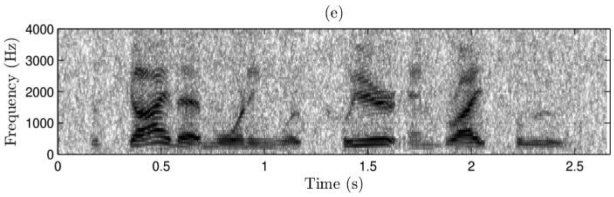


Figure 1. Spectrograms showing the performance of the Kalman filter for different LPC estimates: (a) clean speech (*'The sky that morning was clear and bright blue'*); (b) speech corrupted with white noise at 10 dB SNR (PESQ = 2.07); (c) time trajectories of the Kalman gain for the oracle (thin blue line) and noisy case (thick red line); (d) enhanced speech from oracle case (PESQ = 2.77); (e) enhanced speech from noisy case (PESQ = 2.25).

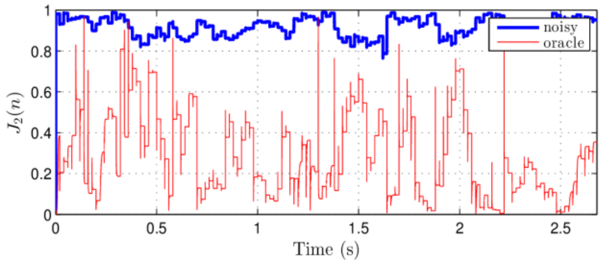


Figure 2. Temporal trajectories of the robustness metric $J_2(n)$ when LPCs are estimated from clean speech (thin red) and noise-corrupted speech (thick blue).

Fig. 1(c) shows the Kalman gain trajectories across time for a typical speech file from the NOIZEUS database (sp10) for both the oracle and noisy cases. It can be seen that in the oracle case, $K(n)$ goes to zero in the silent regions and this results in no residual noise, since according to (13), $K(n)$ controls the contribution of the noisy observation $y(n)$ in the final output. In the noisy case (i.e. where the LPCs are estimated from noise-corrupted speech), it can be seen that $K(n)$ is biased at approximately 0.5 in the silent regions, therefore causing a portion of the noisy observation signal $y(n)$ to pass through to the output.

C. *Using the Robustness Metric to Reduce the Bias in the Kalman Filter Gain*

Robustness refers to the ability of the Kalman filter to deal with uncertainties in the parameters of its dynamic model. In other words, a robust Kalman filter would mitigate the uncertainty or variance in the model by offsetting its contribution in favour of the observation data, when forming the *a posteriori* state vector estimate. In speech enhancement, the uncertainty or variance of the model relates to the ability of the low-order speech production model to estimate the clean speech. An example of this scenario occurs for voiced speech that has harmonic structure, where the low-order speech model is unable to capture the long-term correlation information. The mean square prediction error, which is also the excitation variance σ_w^2, becomes high for voiced speech and hence, the Kalman filter enters into a robust mode of operation, where the observation data $y(n)$ is favoured (i.e. $K(n)$ is high). Another cause for the Kalman filter to operate in the robust mode is the use of biased estimates of LPCs $\{a_k\}$.

The robustness metric $J_2(n)$ proposed in [11] allows us to quantify the level of robustness in the Kalman filter. Rewriting the $J_2(n)$ equation using the scalar Kalman filter variables:

$$J_2(n) = \frac{\sigma_w^2}{\alpha^2(n)+\sigma_w^2} \tag{19}$$

We can see that $J_2(n)$ is computed as the proportion of the excitation variance (in the numerator) with respect to the total *a priori* prediction error (in the denominator). The larger the excitation variance, the closer $J_2(n)$ will approach unity. Fig. 2 shows a plot of $J_2(n)$ for two cases of LPCs estimated from clean speech (oracle) and noisy speech. We can see that the robustness metric for the oracle case varies depending on the voiced speech content, while it is quite high in the noisy case, where it is close to unity in the silent regions. The 'artificially' high robustness of the Kalman filter in the silent regions is detrimental to the overall quality of the enhanced speech since the noisy observation is favoured in order to mitigate a high variance model.

We note that all the quantities in (19) are squared quantities hence $J_2(n)$ will vary between 0 and 1. Therefore, we propose a Kalman filter gain adjustment factor that is based on the value of the robustness metric $J_2(n)$:

$$K'(n) = K(n)[1 - J_2(n)] \tag{20}$$

where $K'(n)$ is the adjusted Kalman filter gain component. To see how this affects the Kalman filter operation, we substitute (19) (with biased terms) and (17) into (20):

$$K'(n) = \frac{\tilde{\alpha}^2(n)}{\tilde{\alpha}^2(n)+\tilde{\sigma}_w^2+\sigma_v^2} \tag{21}$$

When comparing with (17), it can be seen that the biased excitation variance $\tilde{\sigma}_w^2$ is not present in the numerator of (21). Therefore, during the silent regions (where $\tilde{\alpha}^2(n) = 0$), the modified Kalman filter gain $K'(n)$ is 0, rather than 0.5 for the unmodified case.

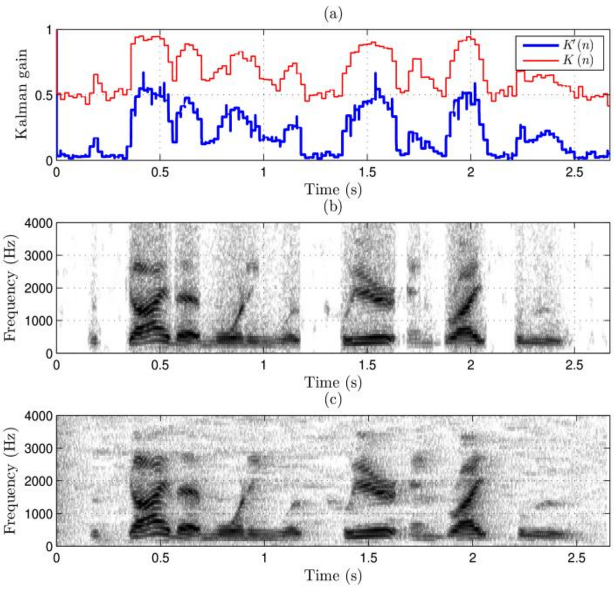


Figure 3. Kalman filter gain trajectories for the noisy and proposed method for 10~dB SNR [in (a)] and spectrograms of enhanced speech from: (b) proposed method (PESQ = 2.67); (c) MMSE-STSA method (PESQ = 2.52).

Fig. 3(a) shows the modified and unmodified Kalman filter gain components, $K'_0(n)$ and $K_0(n)$ respectively. The spectrogram of the enhanced speech from the proposed algorithm is shown in Fig. 3(b), while the spectogram of speech enhanced using the MMSE-STSA method [2] is shown in Fig. 3(c). We can see from Fig. 3(b) that the residual noise in the silent regions has been reduced in the proposed algorithm. A comparison of Fig. 3(b) and Fig. 3(c) shows that the speech from the proposed method suffers less from background residual noise, particularly in the silent regions, as compared to MMSE-STSA. The proposed method also scores a higher PESQ than the MMSE-STSA method on this particular speech file.

From Fig. 1(d), we can see that there is residual noise in the speech regions of the Kalman oracle method. This is due to the presence of the noise in the innovation signal, which is passed through to the output due to a high Kalman filter gain value $K(n)$ (as shown in Fig. 1(c)). In comparison, we can see that the level of residual noise in the speech regions is lower in the proposed method (in Fig. 3(b)), since the modified Kalman filter gain $K'(n)$ has been lowered (as shown in Fig. 3(a)). Therefore, the proposed modification in (21) has the added benefit of reducing noise in the speech regions as well, though this will have the effect of introducing some speech distortion as well.

III. Experimental Setup

In our experiments, we use the NOIZEUS speech corpus, which is composed of 30 phonetically balanced sentences belonging to six speakers [12]. The corpus is sampled at 8 kHz. For our objective experiments, we generate a stimuli set that has been corrupted by additive white Gaussian noise at four SNR levels (0, 5, 10 and 15 dB). The objective evaluation was carried out using the PESQ (perceptual evaluation of speech quality) measure [12].

The treatment types used in the evaluations are listed below (where p is the order of the LPC analysis):

1) Original clean speech (**Clean**);
2) Speech corrupted with white Gaussian noise (**Noisy**);
3) Kalman filter with LPCs estimated from clean speech, 20ms frames, p=10 (**Kalman oracle**);
4) Kalman filter with LPCs estimated from noise-corrupted speech, 20ms frames, p=10 (**Kalman normal**);
5) Proposed Kalman filter with dynamic gain adjustment, 20ms frames, p=10 (**Kalman proposed**);
6) MMSE-STSA method [2] (**MMSE**).

In addition, a subjective evaluation was performed on the white Gaussian noise case at 10 dB SNR and is in the form of blind AB listening tests [14] to determine subjective preference, using the testing methodology outlined in [7]. The speech utterance was from a female speaker (sp27) (*'Bring your best compass to the third class'*). A total of 17 English speaking listeners participated in the subjective listening tests.

IV. Results and Discussion

Table I shows the average PESQ results for each of the enhancement methods. It can be seen that the Kalman oracle method achieves the highest PESQ since it uses LPC estimates computed from the clean speech for its speech model. Therefore, it can be interpreted as the ideal case and serves as an informal upper bound of speech quality. The PESQ scores for the Kalman normal method are only marginally better when compared with those with no enhancement. The PESQ scores of the proposed Kalman filter with dynamic gain adjustment can be seen to have improved considerably when compared with the Kalman normal method. The scores are also very similar to those of the Kalman oracle method, which suggests that the proposed method is quite successful at reducing the negative effects of LPC parameter bias. It also suggests that the speech distortion associated with the Kalman gain suppression in the speech regions *has negligible impact on quality*.

TABLE I. Average PESQ Results Comparing the Different Speech Enhancement Methods with the Proposed Method for Speech Corrupted by White Noise

Method	Input SNR (dB)			
	0	5	10	15
No enhancement	1.57	1.83	2.13	2.47
Kalman oracle	2.12	2.36	2.64	2.96
Kalman normal	1.69	2.00	2.33	2.66
Kalman proposed	**1.97**	**2.32**	**2.63**	**2.93**
MMSE-STSA	1.96	2.33	2.64	2.94

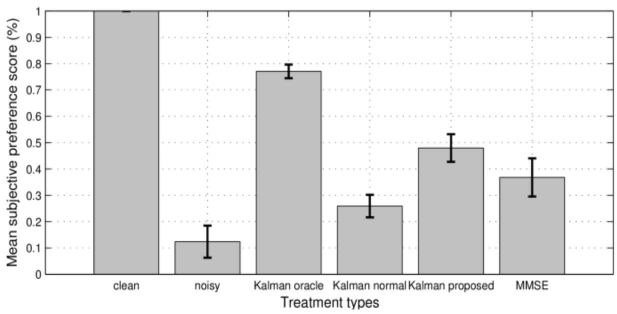


Figure 4. Mean subjective preference scores with 95% confidence intervals for all treatment types.

Fig. 4 shows the mean subjective preference scores with 95% confidence intervals for all treatment types. We can see that the Kalman oracle method was highly preferred over all the other methods, followed by the Kalman proposed method and MMSE-STSA. In order to determine the statistical significance of the subjective test scores, a one-way ANOVA (analysis of variance) test (at the 5% significance level) was performed. The results of the ANOVA test are shown in Table II. The ANOVA test rejected the null hypothesis which stated that all the mean subjective scores were equal (at the 5% significance level). Tukey's Honestly Significant Difference test (shown in Table III) was applied to determine the statistical significance of pairwise comparisons. The Kalman proposed method was shown to be *significantly different* to all the other methods.

TABLE II. ANOVA TABLE FOR THE SUBJECTIVE PREFERENCE SCORES (α=0.05): THE NULL HYPOTHESIS H_0 (ALL MEANS ARE EQUAL) IS REJECTED (F_α=2.34)

Source	*df*	*SS*	*MS*	*F*-statistic
Treatment	5	919.80	183.96	204.86
Error	96	86.21	0.90	
Total	101	1006.00		

TABLE III. DIFFERENCES BETWEEN MEAN SUBJECTIVE SCORES AND SIGNIFICANCE OF PAIRWISE COMPARISONS USING TUKEY'S HONESTLY SIGNIFICANT DIFFERENCE TEST. DIFFERENCES LARGER THAN 0.67 ARE SIGNIFICANT (MARKED WITH *) AT THE α=0.05 LEVEL

	Noisy	Kalman oracle	Kalman normal	**Kalman proposed**	MMSE
Clean	8.76*	2.29*	7.41*	5.21*	6.32*
Noisy		6.47*	1.35*	3.56*	2.44*
Kalman oracle			5.12*	2.91*	4.03*
Kalman normal				2.21*	1.09*
Kalman proposed					1.12*

V. CONCLUSION AND FURTHER WORK

In this paper, we have proposed the use of a robustness metric to dynamically adjust the Kalman filter gain to alleviate the negative effects of poor estimates of the speech model parameters. Using temporal Kalman filter gain trajectories and spectrograms, the proposed method was shown to reduce the level of residual noise in the enhanced speech. Objective tests showed the proposed Kalman filter to have better enhancement performance over the non-oracle algorithm. Human subjective listening tests supported the quality improvements of the proposed method, with ANOVA and post-hoc analysis confirming these improvements to be statistically significant. Therefore, we believe that the use of the robustness metric shows good promise in refining the Kalman filtering-based speech enhancement in practice. The present study considered the enhancement of speech when it was corrupted by white noise only. Further work will be done for the coloured noise case.

REFERENCES

[1] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-27, no. 2, pp. 113-120, 1979.

[2] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, pp. 1109-1121, Dec. 1984.

[3] N. Wiener, *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*, New York: Wiley, 1949.

[4] Y. Ephraim and H. L. V. Trees, "A signal subspace approach for speech enhancement," *IEEE Trans. Speech Audio Process.*, vol. 3, pp. 251-266, Jul. 1995.

[5] K. K. Paliwal and A. Basu, "A speech enhancement method based on Kalman filtering," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Apr. 1987, pp. 177-180.

[6] S. So, K. K. Wojcicki, and K. K. Paliwal, "Single-Channel speech enhancement using Kalman filtering in the modulation domain," in *Proc. Interspeech*, Sep. 2010.

[7] S. So and K. K. Paliwal, "Modulation-Domain Kalman filtering for single-channel speech enhancement," *Speech Commun.*, vol. 53, no. 6, pp. 818-829, Jul. 2011.

[8] J. D. Gibson, B. Koo, and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1732-1742, Aug. 1991.

[9] S. So, K. K. Wojcicki, J. G. Lyons, A. P. Stark, and K. K. Paliwal, "Kalman filter with phase spectrum compensation algorithm for speech enhancement," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Apr. 2009, pp. 4405-4408.

[10] S. So and K. K. Paliwal, "Suppressing the influence of additive noise on the Kalman filter gain for low residual noise speech enhancement," *Speech Commun.*, vol. 53, no. 3, pp. 355-378, Mar. 2011.

[11] M. Saha, R. Ghosh, and B. Goswami, "Robustness and sensitivity metrics for tuning the extended Kalman filter," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 4, pp. 964-971, Apr. 2014.

[12] P. Loizou, *Speech Enhancement: Theory and Practice*, 1st ed., CRC Press LLC, 2007.

[13] S. So, K. K. Wojcicki, and K. K. Paliwal, "Fast converging iterative Kalman filtering for speech enhancement using long and overlapped tapered windows with large side lobe attenuation," in *Proc. Interspeech*, Sep. 2010.

[14] P. Sorqvist, P. Handel, and B. Ottersten, "Kalman filtering for low distortion speech enhancement in mobile communication," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1997, pp. 1219-1222.

Stephen So received his BEng degree with First Class Honours in Microelectronic Engineering at Griffith University, Australia in 1999 and his PhD in image and speech signal processing at Griffith University in 2005. Currently he is a Senior Lecturer in the Griffith School of Engineering, Gold Coast campus, Griffith University, Australia. Dr. So is a Senior Member of the IEEE and is a member of the Editorial Board of Digital Signal Processing (published by Elsevier). His research interests are in image and speech coding, speech enhancement, distributed speech recognition and machine learning.

Aidan E. W. George received the BEng degree in mechatronic engineering (with advanced studies) from Griffith University, Australia, in 2013. Since 2014, he has been working towards a PhD in the field of speech enhancement under the supervision of Dr. Stephen So and Prof. Kuldip Paliwal. His research interests include speech enhancement, speaker recognition, and machine learning.

Ratna Ghosh was born in Ranchi, India. She obtained her B.Tech in Instrumentation Engineering from Indian Institute of Technology, Kharagpur, India in 1990. Thereafter, she completed her Ph.D. in the area of control systems in the Electrical Engineering Department from the same institute in 1997.
She has worked in Indian Airlines as a Graduate Engineer Trainee during 1990-1991. After her PhD, she has been teaching in the Instrumentation and Electronics Engineering Department in Jadavpur University (JU), Kolkata, India since 1997. Presently, she is an Associate Professor in the Department. Her main areas of interest are filtering and estimation, signal processing, control systems, biomedical instrumentation.

Kuldip K. Paliwal received the B.S. degree from Agra University, Agra, India, in 1969, the M.S. degree from Aligarh Muslim University, Aligarh, India, in 1971 and the Ph.D. degree from Bombay University, Bombay, India, in 1978. He has been carrying out research in the area of speech processing since 1972. He has worked at a number of organizations including Tata Institute of Fundamental Research, Bombay, India,

Norwegian Institute of Technology, Trondheim, Norway, University of Keele, U.K., AT&T Bell Laboratories, Murray Hill, New Jersey, U.S.A., AT&T Shannon Laboratories, Florham Park, New Jersey, U.S.A., and Advanced Telecommunication Research Laboratories, Kyoto, Japan. Since July 1993, he has been a professor at Griffith University, Brisbane, Australia, in the School of Microelectronic Engineering. His current research interests include speech recognition, speech coding, speaker recognition, speech enhancement, face recognition, image coding, pattern recognition and artificial neural networks. He has published more than 300 papers in these research areas.

Prof. Paliwal is a Fellow of Acoustical Society of India. He has served the IEEE Signal Processing Society's Neural Networks Technical Committee as a founding member from 1991 to 1995 and the Speech Processing Technical Committee from 1999 to 2003. He was an Associate Editor of the IEEE Transactions on Speech and Audio Processing during the periods 1994-1997 and 2003–2004. He is in the Editorial Board of the IEEE Signal Processing Magazine. He also served as an Associate Editor of the IEEE Signal Processing Letters from 1997 to 2000. He was the General Co-Chair of the Tenth IEEE Workshop on Neural Networks for Signal Processing (NNSP2000). He has co-edited two books: "Speech Coding and Synthesis" (published by Elsevier), and "Speech and Speaker Recognition: Advanced Topics" (published by Kluwer). He has received IEEE Signal Processing Society's best (senior) paper award in 1995 for his paper on LPC quantization. He served as the Editor-in-Chief of the Speech Communication journal (published by Elsevier) during 2005-2011.