## SPEECH ENHANCEMENT IN A BAYESIAN FRAMEWORK

Gaafar M. K. Saleh

Mahesan Niranjan

Cambridge University, Engineering Department Cambridge CB2 1PZ gs113@eng.cam.ac.uk niranjan@eng.cam.ac.uk

## ABSTRACT

We present an approach for the enhancement of speech signals corrupted by additive white noise of Gaussian statistics. The speech enhancement problem is treated as a signal estimation problem within a Bayesian framework. The conventional all-pole speech production model is assumed to govern the behaviour of the clean speech signal. The additive noise level and all-pole model gain are automatically inferred during the speech enhancement process. The strength of the Bayesian approach developed in this paper lies in its ability to perform speech enhancement without the usual requirement of estimating the level of the corrupting noise from "silence" segments of the corrupted signal. The performance of the Bayesian approach is compared to that of the Lim & Oppenheim framework, to which it follows a similar iterative nature. A significant quality improvement is obtained over the Lim & Oppenheim framework.

#### 1. INTRODUCTION

Interest in the field of speech enhancement has grown rapidly in the past three decades. As such, various approaches to speech enhancement have been proposed and developed [3, 5, 11]. The general aim of speech enhancement is to remove the corrupting noise from the signal in order to improve its perceptual aspects of quality and intelligibility.

One of the primary concerns in developing an applicationspecific speech enhancer is the nature of the noise corrupting the speech signal. The wide variety of situations in which corrupting noise can occur give rise to a large number of possibilities with regard to the nature of the noise and the type of assumptions required in developing a relevant enhancement methodology. Noise may, for example, be broadband or narrow band. It may be additive: as in the case of background noise generated in the interior of a car, or an aircraft cockpit, convolutional; for example microphone noise, reverberative, impulsive,etc. In this paper, we make the assumption that speech is degraded by additive white Gaussian noise which is uncorrelated to the clean speech signal. We focus on improving the overall quality of the noisy speech by treating the speech enhancement problem as a signal estimation one in the time domain. The conventional all-pole speech production model is assumed to govern the behaviour of the clean speech signal [16]. The signal estimation problem is then formulated in a Bayesian framework where the variance of the additive noise and the gain of the all-pole model are related to *hyperparameters* within the Bayesian framework [13, 15, 14]. The optimisation of these hyperparameters within an iterative framework leads to an estimate of the clean speech signal and the coefficients of the all-pole, or equivalently Linear Prediction (LP), model.

The iterative nature of the Bayesian scheme that we propose in this paper is similar to that proposed by Lim & Oppenheim [10] (see also [8]). Lim & Oppenheim's system assumes an underlying all-pole model for the speech signal and, starting from the noisy speech signal, obtains alternate estimates for the all-pole model coefficients and the clean speech signal in an iterative manner. The main difference between the Bayesian scheme developed here and Lim & Oppenheim's framework is that whereas Lim & Oppenheim's framework requires the detection of non-speech activity segments in the noisy speech in order to estimate the energy of the corrupting noise and subsequently the all-pole model gain, our proposed method does not. The detection of non-speech activity segments in noisy speech is also a feature of speech enhancement algorithms which are based on Hidden Markov Modelling [4], spectral subtraction [1, 2], and signal subspace decomposition [6, 9].

This paper proceeds by first defining the speech enhancement problem. A brief overview of the Lim & Oppenheim approach then follows. The proposed speech enhancement scheme is then theoretically developed. Results of the application of the Bayesian scheme are then given and contrasted to Lim and Oppenheim's approach for stationary noise environments.

#### 2. PROBLEM STATEMENT

Our aim is to recover the speech signal, s, from an observed signal, y, which can be described as :

$$\mathbf{y} = \mathbf{s} + \mathbf{v} \tag{1}$$

where v is a zero-mean independent identically distributed (i.i.d) white Gaussian noise signal with variance  $\sigma_v^2$ . The signals v and s are assumed to be uncorrelated.

It is assumed that the clean speech signal, s, is described by an LP model of order k such that :

$$s(n) = \mathbf{s}^{T} (n-1, n-k) \mathbf{a} + Gu(n)$$
<sup>(2)</sup>

where

$$\mathbf{s}(n-1, n-k) = [s(n-1)s(n-2)\dots s(n-k)]$$
(3)

and

$$\mathbf{a} = [a(1) a(2) \dots a(k)] \tag{4}$$

This work was supported by the Educational and Physical Sciences Research Council (EPSRC) and Cambridge University Engineering Department.

u(n) is a zero-mean unity variance i.i.d. Gaussian process and G is the gain term of the corresponding all-pole modelling filter.

## 3. LIM & OPPENHEIM'S APPROACH

The practical application of the scheme relies on a sub-optimal optimisation procedure. The procedure can be viewed as a two-stage iterative process which, at the *i*<sup>th</sup> stage, derives an estimate for the clean speech,  $\hat{s}^i$ , which relies on the estimate of the all-pole model parameters,  $a^i$ . The estimated speech is then used to obtain an estimate of the all-pole parameters  $a^{i+1}$ . The procedure then continues until some convergence criterion is met.

In order to obtain an estimate of the clean speech, s, its posterior probability given the observations, y and the LP coefficients, a is expressed as :

$$p\left(\mathbf{s}|\mathbf{a}, \mathbf{y}, G, \sigma_v^2\right) = \frac{p\left(\mathbf{y}|\mathbf{a}, \mathbf{s}, G, \sigma_v^2\right) p\left(\mathbf{s}|\mathbf{a}, G, \sigma_v^2\right)}{p\left(\mathbf{y}|\mathbf{a}, G, \sigma_v^2\right)} \quad (5)$$

Since the denominator of equation (5) does not depend on s, and a is considered to be fixed within the process of estimating s, we require the maximisation of the product of the two terms appearing in the numerator of (5) with respect to s. The posterior probability of s can thus be written as :

$$p\left(\mathbf{s}|\mathbf{a},\mathbf{y},G,\sigma_{v}^{2}\right) \propto p\left(\mathbf{y}|\mathbf{a},\mathbf{s},G,\sigma_{v}^{2}\right) p\left(\mathbf{s}|\mathbf{a},G,\sigma_{v}^{2}\right)$$
(6)

The conditional probability of the noisy observations y is written as :

$$p\left(\mathbf{y}|\mathbf{a},\mathbf{s},G,\sigma_{v}^{2}\right) = \left(\frac{1}{2\pi\sigma_{v}^{2}}\right)^{\frac{N}{2}}\exp\left(\frac{1}{\sigma_{v}^{2}}E_{\mathbf{y}}\right)$$
(7)

where N is the length of the speech segment that is being analysed and  $E_{\mathbf{y}} = \frac{1}{2} (\mathbf{y} - \mathbf{s})^T (\mathbf{y} - \mathbf{s})$ . The likelihood  $p(\mathbf{s}|\mathbf{a}, G, \sigma_v^2)$ , on the other hand, is:

$$p(\mathbf{s}|\mathbf{a},G) = \left(\frac{1}{2\pi G^2}\right)^{\frac{N}{2}} \exp\left[-\frac{1}{2G^2}E_{\mathbf{s}}\right]$$
(8)

where  $E_s$  is the total squared prediction error:

$$E_{s} = \frac{1}{2} \sum_{n=n_{0}}^{n_{1}} \left( s\left(n\right) - \sum_{i=1}^{k} a_{i} s\left(n-i\right) \right)^{2}$$
(9)

and  $n_0, n_1$ , are the limits over which the prediction error is computed.

From equations (6), (7), and (8), we note that the posterior probability of s can be written as:

$$p\left(\mathbf{s}|\mathbf{a},\mathbf{y},G,\sigma_{v}^{2}\right) \propto \exp\left(\frac{1}{\sigma_{v}^{2}}E_{\mathbf{y}}+\frac{1}{G^{2}}E_{\mathbf{s}}\right)$$
 (10)

$$\propto \exp\left(-M_s\right)$$
 (11)

Setting the derivative of  $M_s$  with respect to s to zero, we obtain a *maximum a posteriori* estimate of the clean speech :

$$\mathbf{s}_{\mathsf{MAP}} = \left(\frac{1}{\sigma_v^2}\mathbf{I} + \frac{1}{G^2}\mathbf{B}\right)^{-1} \left(\frac{1}{\sigma_v^2}\mathbf{y}\right) \tag{12}$$

where I is the identity matrix and  $\mathbf{B}$  is a matrix of correlations of the LP coefficients [17]. Lim & Oppenheim construct a Wiener

filter from the estimated all-pole model in order to approximate the action of the estimator  $s_{MAP}$  appearing in equation (12) (see [10]). The Weiner filter requires estimates of the noise variance  $\sigma_v^2$  and the all-pole model gain, G.  $\sigma_v^2$  is estimated from nonspeech activity segments of noisy speech while G is obtained by utilising Parseval's theorem. In the LP coefficient estimation stage, an autocorrelation estimator is used on the estimated speech.

#### 4. THE BAYESIAN FRAMEWORK

We make the two following definitions in terms of the variance of the corrupting noise,  $\sigma_v^2$ , and the gain of the all-pole model, G:

$$\zeta = \frac{1}{\sigma_v^2} \tag{13}$$

$$\eta = \frac{1}{G^2} \tag{14}$$

where  $\zeta$  and  $\eta$  are termed hyperparameters [12, 14].

The posterior probability of the clean speech, s, given the noisy observations, y, is expressed as :

$$p(\mathbf{s}|\mathbf{y}) = \int p(\mathbf{s}|\mathbf{a},\zeta,\eta,\mathbf{y}) p(\mathbf{a},\zeta,\eta|\mathbf{y}) \, da \, d\zeta \, d\eta \qquad (15)$$

The evidence approximation [7, 13] is utilised in order to evaluate the integrand given in equation (15). Specifically, it is assumed that the posterior probability of  $[\mathbf{a}, \zeta, \eta]$  exhibits a strong peak around the most probable values of these parameters,  $[\mathbf{a}, \zeta, \eta]_{mp}$ , compared to the variation of  $p(\mathbf{s}|\mathbf{a}, \zeta, \eta, \mathbf{y})$  in that region. As such, the integral representing the posterior probability,  $p(\mathbf{s}|\mathbf{y})$ , is approximated as :

$$p(\mathbf{s}|\mathbf{y}) \approx p\left(\mathbf{s}|\left[\mathbf{a},\zeta,\eta\right]_{mp},\mathbf{y}\right)$$
 (16)

Now, the posterior probability of  $\mathbf{a}, \zeta, \eta$  can be written as :

$$p(\mathbf{a},\zeta,\eta|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{a},\zeta,\eta)p(\mathbf{a},\zeta,\eta)}{p(\mathbf{y})}$$
(17)

The prior  $p(\mathbf{a}, \zeta, \eta)$  is assumed to be uniform.  $p(\mathbf{y}|\mathbf{a}, \zeta, \eta)$  can thus be used to infer the optimal, or most probable, values of  $[\mathbf{a}, \zeta, \eta]$ . The term  $p(\mathbf{y}|\mathbf{a}, \zeta, \eta)$  is the normalising term appearing in the following expression and is called the *evidence*:

$$p(\mathbf{s}|\mathbf{a},\zeta,\eta,\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{s},\zeta) p(\mathbf{s}|\mathbf{a},\eta)}{p(\mathbf{y}|\mathbf{a},\zeta,\eta)}$$
(18)

The numerator terms in equation (18), which appeared when discussing the Lim & Oppenheim system in a different form as equations (7) and (8) are re-written as follows :

$$P(\mathbf{y}|\mathbf{s},\zeta) = \frac{1}{Z_y} \exp\left(-\zeta E_{\mathbf{y}}\right)$$
(19)

$$p(\mathbf{s}|\mathbf{a},\eta) = \frac{1}{Z_s} \exp\left(-\eta E_{\mathbf{s}}\right)$$
(20)

where  $Z_y = \left(\frac{2\pi}{\zeta}\right)^{\frac{N}{2}}$  and  $Z_s = \left(\frac{2\pi}{\eta}\right)^{\frac{N}{2}}$  The evidence can now be written as :

$$p(\mathbf{y}|\mathbf{a},\zeta,\eta) = \frac{1}{Z_y Z_s} \int \exp(-(\zeta E_{\mathbf{y}} + \eta E_{\mathbf{s}}) d\mathbf{s}$$
(21)

Defining  $M(\mathbf{s}) = \zeta E_{\mathbf{y}} + \eta E_{\mathbf{s}}$ , the integral in equation (21) can be evaluated after performing a second order Taylor expansion of  $M(\mathbf{s})$  around  $\mathbf{s}_{mp}$ , the most probable value of  $\mathbf{s}$ :

$$M(\mathbf{s}) = M(\mathbf{s}_{mp}) + \frac{1}{2} [\mathbf{s} - \mathbf{s}_{mp}]^T \mathbf{C} [\mathbf{s} - \mathbf{s}_{mp}]$$
(22)

where C is the Hessian matrix given by :

$$\mathbf{C} = \nabla \nabla M(\mathbf{s}) \tag{23}$$

$$= \zeta \mathbf{I} + \eta \mathbf{B} \tag{24}$$

and **B** is given by  $\mathbf{B} = \nabla \nabla E_s$ , which is analogous to the matrix of correlations of the LP coefficients which was encountered earlier in equation (12) of the previous section [17].

Setting  $\nabla M$  (s) to zero, the most probable estimate of s for any  $\zeta$ ,  $\eta$ , **B**, is obtained in a similar fashion to equation (12):

$$\mathbf{s}_{mp} = (\zeta \mathbf{I} + \eta \mathbf{B})^{-1} \zeta \mathbf{y}$$
(25)

Substituting equation (22) into (21), we use the standard Gaussian integral to obtain an analytical expression for the evidence :

$$p(\mathbf{y}|\mathbf{a},\zeta,\eta) = \frac{\exp{-M(\mathbf{s}_{mp})}}{Z_s Z_y} (2\pi)^{\frac{N}{2}} \det^{-\frac{1}{2}} \mathbf{C}$$
(26)

We can now write the log evidence from equation (26) as :

$$\log p(\mathbf{y}|\mathbf{a}, \zeta, \eta) = -\zeta E_{\mathbf{y}}^{mp} - \eta E_{\mathbf{s}}^{mp} - \frac{N}{2} \log 2\pi$$
$$-\frac{1}{2} \log \det \mathbf{C} + \frac{N}{2} \log \zeta$$
$$+\frac{N}{2} \log \eta \qquad (27)$$

where  $E_{\mathbf{y}}^{\text{mp}}$  and  $\eta E_{\mathbf{s}}^{\text{mp}}$  denote the values of the respective functions evaluated at  $\mathbf{s}_{\text{mp}}$  as expressed in equation (25).

## 4.1. Derivatives of the Evidence

The properties of the log evidence can now be exploited in order to obtain the estimate of the speech signal  $s_{mp}$  at the most probable a,  $\zeta$ , and  $\eta$ . This is done by setting the derivatives of the log evidence with respect to  $\zeta$ ,  $\eta$ , and a to zero such as to examine the conditions that are satisfied at the maximum of the log evidence surface. The conditions satisfied at the maximum of the log evidence are governed by the following relationships (Saleh, 1996):

$$2\zeta E_{\mathbf{y}}^{\mathrm{mp}} = \theta \tag{28}$$

$$2\eta E_{\rm s}^{\rm mp} = N - \theta \tag{29}$$

$$\frac{\partial E_{\mathbf{s}}^{mp}}{\partial \mathbf{a}} = \frac{1}{2} \sum_{i=1}^{N} \frac{x_i^T \frac{\partial \eta \mathbf{B}}{\partial \mathbf{a}} x_i}{\lambda_i + \zeta}$$
(30)

where  $\lambda_i$  are the eigenvalues of  $\eta \mathbf{B}$ , and  $\theta$  is the number of "well-determined" speech samples:

$$\theta = \sum_{i=1}^{N} \frac{\lambda_i}{\lambda_i + \zeta} \tag{31}$$

 $\theta$  is analogous to the number of well-determined model parameters  $\gamma$ , which occurred in (MacKay, 1992) within the context of parameter estimation for interpolation models to indicate the number of model parameters well determined by the data.

#### 5. IMPLEMENTATION CONSIDERATIONS

We implemented the Lim & Oppenheim iterative framework in the time domain through the evaluation of equation (12) on the speech estimation stage, rather thanutilising a Weiner filter. Since the original Lim & Oppenheim system was defined for an autocorrelation estimator, the analysis of a speech frame was carried after multiplying it by a Hamming window. The resulting enhanced speech is therefore a reconstruction of the windowed speech.An overlap-and-add method was thus maintained, whereby successive analysis frames were advanced by half the window size,  $\frac{N}{2}$ , and the overlapped segments added.

The Bayesian scheme was implemented iteratively in the same manner as the Lim & Oppenheim system whereby the estimation of the clean speech and the LP coefficients is carried out alternately on two separate stages. The optimisation of the log evidence term in equation (27) is performed by iteratively solving equations (28), (29) and (31) in order to obtain optimal values for  $\zeta$  and  $\eta$  which are substituted into equation (25) to obtain a clean speech estimate. The estimation of the LP coefficients is then carried out from the speech estimate just obtained using the conventional covariance estimator. The LP coefficients are utilised again within equations (28), (29) and (31) to obtain a clean speech estimate, and so on. This procedure leads to an estimate of the joint likelihood of the speech and the LP coefficients given the observed noisy data,  $p(\mathbf{a}, \mathbf{s}|\mathbf{y})$ . In order to circumvent any effects of discontinuities between the frames, an overlap of 2k samples between successive analysis frames was maintained.

#### 6. RESULTS AND DISCUSSION

The Bayesian scheme, and Lim & Oppenheim framework, were applied to the following utterance, sampled at 10 KHz, and spoken by male and female speakers :

# "Your sum shouldn't come to seven three seven point zero seven six."

A frame size of duration 25.6 ms was maintained in the analysis, which was performed using LP models of order 14. For both speakers, the utterance was corrupted with additive white Gaussian noise resulting in input signal-to-noise (SNR) ratios of 5, 10, 15, and 20 dB. The input SNR is defined as :

$$SNR_{\psi} = 10 \log_{10} \frac{\sum_{n} s^{2}(n)}{\sum_{n} d^{2}(n)}$$
(32)

where s(n) and d(n) refer to the clean speech and the noise respectively.

Perceptually, the noise removal effect of enhancement was clear after either 3 or 4 iterations for both methods, although a continuous "musical" tone was noted to be present in the enhanced speech. The "musical" tone is a feature of the majority of current model-based speech enhancement algorithms and is attributed to narrowband residual signals with time-varying frequencies which result from the approximation of the LP spectra of the clean speech. For iterations higher than 4, informal listening tests showed that there was no audible improvement in the quality of enhanced speech.

Output SNR measures were obtained after 3 and 4 iterations, and the results of both methods were compared under the same input SNR conditions. The output  $SNR_o$ , is defined as :

SNR<sub>o</sub> = 10 log<sub>10</sub> 
$$\frac{\sum_{n=1}^{N} s^{2}(n)}{\sum_{n=1}^{N} [s(n) - \hat{s}(n)]^{2}}$$
 (33)

Bayesian iterative	Male $SNR_o$ (dB)		Female SNR <sub>o</sub> (dB)	
scheme	Iteration		Iteration	
Input SNR	3	4	3	4
5	11.08	10.63	11.52	11.18
10	14.91	14.73	14.93	14.79
15	18.79	18.59	18.80	18.60
20	22.60	22.14	22.32	21.83
	Male $SNR_o$ (dB)		Female SNR <sub>o</sub> (dB)	
Lim & Oppenheim	Male S	$NR_{o}$ (dB)	Female	$SNR_o$ (dB)
Lim & Oppenheim system	Male S Iter	NR <sub>o</sub> (dB) ration	Female Ite	SNR <sub>o</sub> (dB) eration
Lim & Oppenheim system Input SNR	Male S Iter 3	NR <sub>o</sub> (dB) ration 4	Female Ite	SNR <sub>o</sub> (dB) eration 4
Lim & Oppenheim system Input SNR 5	Male S Iter 3 9.67	NR <sub>o</sub> (dB) ration 4 9.16	Female Ite 3 9.97	SNR <sub>o</sub> (dB) eration 4 9.55
Lim & Oppenheim system Input SNR 5 10	Male S Iter 3 9.67 12.94	NR <sub>o</sub> (dB) ration 4 9.16 12.39	Female Ite 3 9.97 13.86	SNR <sub>o</sub> (dB) eration 4 9.55 13.43
Lim & Oppenheim system Input SNR 5 10 15	Male S Iter 3 9.67 12.94 16.52	NR <sub>o</sub> (dB) ration 4 9.16 12.39 16.02	Female 3 9.97 13.86 17.00	SNR <sub>o</sub> (dB) eration 4 9.55 13.43 16.61

Table 1: Output SNR measures of enhanced speech obtained with the Bayesian iterative scheme, and the Lim & Oppenheim system in stationary noise environments.

Table I summarises the results for the Bayesian scheme and Lim & Oppenheim's system. In all cases, the output SNR is consistently higher at 3 iterations, than at 4. The decrease in output SNR values from 3 to 4 iterations is generally low and varies from 0.14 dB to 0.55 dB. It is also clear from Table I that the Bayesian iterative scheme achieves output SNR values which are higher than those obtained with the Lim & Oppenheim system for all the input SNR values considered.

The SNR improvement (between the input and output SNRs) generally becomes smaller as the input SNR gets larger for both systems. For the input SNR value of 20 dB, negative improvement of 0.7-0.9 dB in output SNR is obtained with the Lim & Oppenheim system. This is in contrast to "positive" improvement of around at least 2 dB for the Bayesian scheme. Informal listening tests by a number of subjects confirmed a preference towards the Bayesian scheme over the Lim & Oppenheim system in all the cases which are considered in Table I.

## 7. CONCLUSIONS

A Bayesian approach for speech enhancement whereby the corrupting additive noise can be optimised, as a hyperparameter, from the noisy speech was introduced. The Bayesian scheme was contrasted to the Lim & Oppenheim system. Both systems rely on the estimation of an underlying all-pole model and follow the same iterative scheme of alternately estimating the all-pole model parameters and the clean speech. Informal listening tests indicated a consistent preference towards the Bayesian scheme in comparison to Lim & Oppenheim's system when enhancement was performed on speech corrupted with additive white noise. The improvements in SNR were also consistently greater for the Bayesian scheme than Lim & Oppenheim's system.

The results given in this paper demonstrate the huge potential of Bayesian methods in speech enhancement. The ability to infer the corrupting noise variance from the noisy speech signal, based on the underlying all-pole model, is a useful step towards the development of a unified comprehensive Bayesian view of the all-pole modelling of noisy speech.

### 8. REFERENCES

- M. Berouti, R. Schwartz, and J. Makhoul. Enhancement of speech corrupted by acoustic noise. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 208–211, April 1979.
- [2] S. F. Boll. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics*, *Speech, and Signal Processing*, ASSP-27(2):113–120, 1979.
- [3] S. F. Boll. Speech enhancement in the 1980s: Noise suppression with pattern matching. In Sadaoki Furui and M. Mohan Sondhi, editors, *Advances in Speech Signal Processing*, pages 309–325. Marcel Dekker, Inc, 1992.
- [4] Y. Ephraim. A Bayesian estimation approach for speech enhancement using Hidden Markov Models. *IEEE Transactions on Signal Processing*, 40(4):725–735, April 1992.
- [5] Y. Ephraim. Statistical-model-based speech enhancement systems. *Proceedings of the IEEE*, 80(10):1526–1555, October 1992.
- [6] Y. Ephraim and H.L. Van Trees. A signal subspace approach for speech enhancement. *IEEE Transactions on Speech, and Audio Processing*, 3(4):251–265, 1995.
- [7] S. F. Gull. Developments in maximum entropy data analysis. In J. Skilling, editor, *Maximum Entropy and Bayesian Methods*, pages 53–71. Kluwer Academic Publishers, 1989.
- [8] J.H.L Hansen and M.A. Clements. Constrained iterative speech enhancement with application to speech recognition. *IEEE Transactions on Signal Processing*, 39(4):795–805, April 1991.
- [9] S. Jensen, P. Hansen, S. Hansen, and J. Sorensen. Reduction of broad-band noise in speech by truncated QSVD. *IEEE Transactions on Speech and Audio Processing*, 3(4):439– 448, November 1995.
- [10] Jae S. Lim and Alan V. Oppenheim. All-pole modelling of degraded speech. *IEEE Tansactions on Acoustics, Speech,* and Signal Processing, ASSP-26(3):197–210, June 1978.
- [11] Jae S. Lim and Alan V. Oppenheim. Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE*, 67(12):1586–1604, December 1979.
- [12] D. V. Lindley and A. F. M. Smith. Bayes estimates for the linear model. *Journal of the Royal Statistical Society*, 34B:1– 18, 1972.
- [13] D.J.C. MacKay. Bayesian interpolation. Neural Computation, 4(3):415–447, 1992.
- [14] D.J.C. MacKay. Hyperparameters: Optimise, or itegrate out ? In G. Heidbreder, editor, Maximum Entropy and Bayesian Methods, Santa Barbara 1993. Dordrecht:Kluwer, 1994.
- [15] D.J.C. MacKay. Probable networks and plausible predictions - a review of practical Bayesian methods for supervised neural networks. *Network-Computation in Neural Systems*, 6:469-505, 1995.
- [16] John Makhoul. Linear Prediction: A tutorial review. Proceedings of the IEEE, 63(4):561–580, April 1975.
- [17] G. Saleh. Bayesian Inference in Speech Processing. PhD thesis, Cambridge University Engineering Department, December 1996.