Volume 28, issue 11 1 August 2007 ISSN 0167-8655

# Pattern Recognition Letters

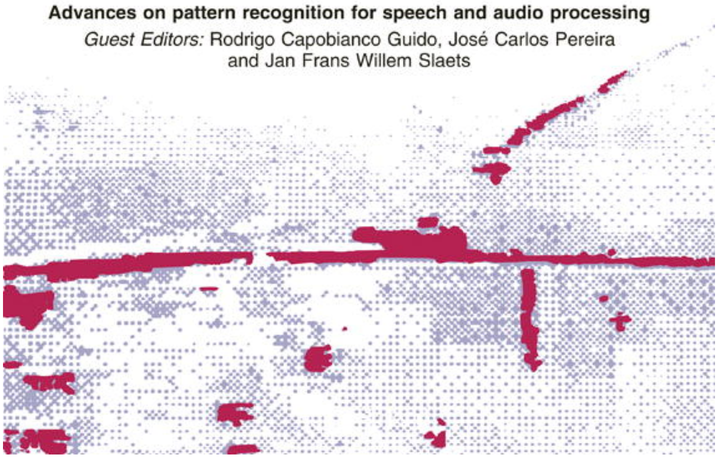
An official publication of the
International Association for Pattern Recognition

IAPR

**Special Issue**

**Advances on pattern recognition for speech and audio processing**

*Guest Editors:* Rodrigo Capobianco Guido, José Carlos Pereira
and Jan Frans Willem Slaets



Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 1343–1350
Pattern Recognition Letters
www.elsevier.com/locate/patrec

# On cepstral and all-pole based spectral envelope modeling with unknown model order

Axel Röbel, Fernando Villavicencio *, Xavier Rodet

*IRCAM – CNRS – UMR STMS, Analysis-Synthesis Team, 1, Place Igor-Stravinsky, 75004 Paris, France*

Available online 27 March 2007



**Abstract**

In this work, we investigate spectral envelope estimation for harmonic signals. We address the issue of model order selection and propose to make use of the fact that the spectral envelope is sampled by means of the harmonic structure of the signal in order to derive upper bounds for the estimator order. An experimental study is performed using synthetic test signals with various fundamental frequencies and different model structures to evaluate the performance of the envelope models. Experimental results confirm the relation between optimal model order and fundamental frequency.


*Keywords:* Spectral envelope modeling; All-pole modeling; Cepstral analysis; Source-filter deconvolution; Speech analysis; Speaker characterization; Speech synthesis

## 1. Introduction

Estimation of the *spectral envelope*, which is a smooth function passing through the prominent peaks of the spectrum, is an important task in signal processing applications. The spectral envelope is generally considered as one of the determining factors for the timbre of a sound. In terms of the well known source-filter model, which models sound creation by means of a white excitation signal passing through a filter, the spectral envelope is the transfer function of the filter. Accordingly, the task consists in estimating the resonator filter from the signal. Spectral envelope estimation methods can be used for applications as signal characterization, classification and modification. While signal characterization and classification applications generally do not require a very precise estimation of the spectral envelope, the quality of voice or timbre conversion systems depends on the quality of the envelope estimate.

In the case of white noise excitation signals there are various straightforward estimation techniques (Kay, 1988). If, however, the excitation signal is periodic (as for pitched instruments or voiced speech), the estimation is difficult due to the fact that the distinction between the spectral envelope and the excitation signal is ambiguous. In cases like these the peaks defining the spectral envelope are the harmonics of the fundamental frequency. Therefore, the spectral envelope should be a transfer function that, if inverted, renders the sequence of spectral peaks of the residual signal as flat as possible, without including the harmonic structure of the excitation signal.

Some problems that hinder the estimation are the proper selection of the filter model (AR, MA, or ARMA) and the proper selection of the model order. The estimation of AR or all-pole models by means of linear prediction (LP), that was described in (Makhoul, 1975), is a technique that is still used quite often for the estimation and parametric representation of the spectral envelope of speech signals. LP modeling can be considered a state of the art procedure if the excitation signal is white noise. For harmonic

* Corresponding author. Present address: Analysis and Synthesis Team, Institute for Music-Acoustic Research and Coordination, 1, Place Igor Stravinsky, 75004, Paris, France. Tel.: +33 14478 4862; fax: +33 14478 1540.

*E-mail addresses:* roebel@ircam.fr (A. Röbel), villavicencio@ircam.fr (F. Villavicencio).

*URLs:* http://www.ircam.fr\anasyin\roebel (A. Röbel), http://www.ircam.fr\anasyin\villavicencio (F. Villavicencio), http://www.ircam.fr\anasyin\rodet (X. Rodet).

0167-8655/$ - see front matter 
doi:10.1016/j.patrec.2006.11.021

excitation signals, however, the LP technique is known to be biased. For these excitation signals the discrete all-pole (DAP) technique that was presented in (El-Jaroudi and Makhoul, 1991) can be used to considerably reduce the bias. Note that compared to the LP method the computational costs and the algorithmic complexity of the DAP algorithm are significantly increased. For the order selection problem there exists only a physically motivated reasoning (O'Shaughnessy, 1987). The fact that the filter is observed after having been sampled by the harmonic structure has not yet been taken into account.

ARMA envelope models are most easily obtained through cepstrum based techniques. The cepstrum is a DFT representation of the log amplitude spectrum and it can be shown that ARMA transfer functions can be represented by means of the cepstrum (Smith, 2005). There are different techniques for cepstrum based envelope estimation. In (Imai and Abe, 1979) an attractive cepstrum-based spectral envelope estimator, named *true-envelope* (TE), is presented. This iterative technique allows efficient estimation of the spectral envelope (Roebel and Rodet, 2005) without the shortcomings of the discrete cepstrum (Cappé and Moulines, 1996; Galas and Rodet, 1990). The resulting estimation can be interpreted as a band limited interpolation of the major spectral peaks.

In the following article an experimental comparative study of envelope estimation techniques is presented. The goal of this investigation is to derive a simple and effective strategy allowing us to select an appropriate model order, and to investigate the performance of different models with respect to the filter properties. For experimental investigation the LP, DAP and TE techniques will be used. The experimental setup is especially relevant for tasks that require the estimation of the residual or excitation signal of pitched signals, such as voice morphing or timbre modification. For these tasks, in contrast to formant detection, a uniform approximation of the envelope is generally advantageous because an error in the excitation signal, whether due to a formant or an anti-formant, may become perceptually important once the envelope has been modified. With respect to the order selection problem we will demonstrate that for the DAP and TE estimators, a reasonable model order can generally be derived from the fundamental frequency of the excitation signal.

The article is organized as follows. The cepstrum based *True-Envelope* algorithm is introduced in Section 2. LP and DAP all-pole based models are described in Section 3. In Section 4, we present the experimental framework and we describe the results in Section 5. Section 6 summarizes the article.

## 2. Efficient cepstrum-based spectral envelope estimation

### 2.1. The True-Envelope estimator

There are a number of approaches for estimating the spectral envelope by means of cepstral smoothing. The discrete cepstrum is the most well know, but, is rather demanding computationally. It requires a pre-selection of the spectral peaks. The *True-Envelope* (TE) estimator was originally proposed in (Imai and Abe, 1979). Recently, a procedure has been proposed that allows significant reduction of computational costs to a level comparable with the Levinson recursion such that real time processing can be achieved (Roebel and Rodet, 2005). Note however, that reduction of the computational cost comes with slightly reduced precision. Therefore, we will not use the real time version of the TE estimator for the following experiments. The true-envelope estimator will be used as representative of the cepstrum based spectral estimators.

TE estimation is based on cepstral smoothing of the amplitude spectrum. Let $X(\omega_k)$ be the $K$-point DFT of the signal frame $x(n)$ and $C_i(\omega_k)$ the cepstrally smoothed spectrum at iteration $i$. The algorithm then iteratively updates the smoothed input spectrum $A_i(\omega_k)$ with the maximum of the original spectrum and the current cepstral representation

$$A_i(\omega_k) = \max(\log(|X(\omega_k)|), C_{i-1}(\omega_k)) \tag{1}$$

and applies the cepstral smoothing to $A_i(\omega_k)$ to obtain $C_i(\omega_k)$. The procedure is initialized setting $A_0(\omega_k) = \log(|X(\omega_k)|)$ and starting the cepstral smoothing to obtain $C_0(\omega_k)$. As depicted in Fig. 1 the estimated envelope grows steadily. The algorithm stops if for all $\omega_k$ the relation $A_i(\omega_k) < C_i(\omega_k) + \theta$ is true with $\theta$ being a user supplied threshold. For the current experiments $\theta = 0.01$ dB was used. Given the fact that the cepstral order is limited the TE estimator creates a band limited function that passes through the prominent spectral peaks. The peaks that are considered prominent are automatically selected according to the cepstral order. The explicit peak selection that is necessary for the DAP estimator as well as for the discrete cepstrum, is not required.

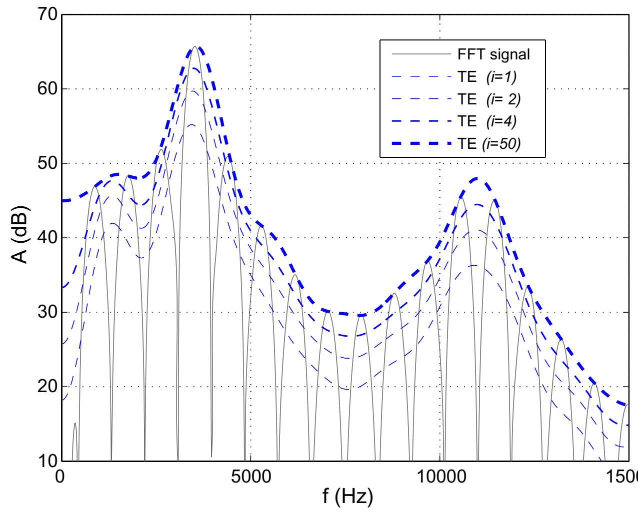



Fig. 1. True-Envelope estimator iteratively approaching the ARMA test spectrum (model order $O = 25$, sample rate $F_s = 44{,}100$ Hz, fundamental period $P_0 = F_s/F_0 = 50$).

### 2.2. Order selection

A major advantage of cepstral envelope estimation techniques is that a reasonable estimate of the optimal cepstral order can be provided. If the observed signal has a fundamental frequency $F_0$, the harmonic excitation spectrum samples the filter transfer function with a sample period given by $F_0$. Therefore, one may deduce that the information in the original filter that exceeds the related Nyquist bandlimit in the cepstral domain is lost. Assuming a sample rate of $F_s$ the related Nyquist frequency bin number in the discrete cepstrum is $F_s/(2F_0)$. This fact allows us to provide a simple way of selecting a nearly optimal cepstral order given that the maximum frequency difference between two spectral peaks that carry envelope information is known. If the difference between those peaks is $\Delta_F$ then the cepstral order used in the TE estimator should be

$$\widehat{O}_{\mathrm{TE}} = F_s/(2\Delta_F) = \alpha_{\mathrm{TE}} F_s/\Delta_F, \quad \alpha_{\mathrm{TE}} = 0.5. \tag{2}$$

Due to the fact that sinusoidal peaks are not ideal impulses, the sampling of the spectral envelope performed by the excitation signal will not be ideal. Moreover, for very smooth transfer functions a lower order may already be sufficient. Therefore, the optimal order (the order that provides an envelope estimate with minimum error) will generally not only depend on $F_0$ but also on the specific properties of the envelope spectrum. Nevertheless, as will be shown in the experimental section, the order selection according to (2) is appropriate for a wide range of situations, and the resulting estimation error is generally rather close to the optimum.

### 2.3. Pre-smoothed True-Envelope estimation

Initial experiments with the TE estimator revealed a problem with the order selection described above, which is due to the fact that in real world signals the spectral envelope is not sampled regularly. The main problem here is the fact that the spectral peak at 0 Hz and possibly $F_s/2$ is generally missing so that for harmonic excitation with fundamental frequency $F_0$ the maximal frequency difference between the supporting peaks will be $\Delta_F = 2F_0$. To be able to increase the model order we propose a two step estimation. First a TE model with order $O = F_s/(4F_0)$ is estimated. From the estimated envelope we derive an estimate of the appropriate sinusoidal amplitude at positions 0 Hz and $F_s/2$. For both positions we create artificial spectral peaks with the estimated amplitude whenever the original amplitude is smaller than the estimated amplitude. In the second step, due to the artificial spectral peaks, $\Delta_F$ is reduced to $F_0$. Therefore, according to (2), we may select $\widehat{O}_{\mathrm{TE}} = \alpha_{\mathrm{TE}} F_s/F_0$. As a result, all available details in the envelope spectrum can be resolved. The two step estimation procedure will be denoted as the TE method below.

## 3. All-pole based modeling

### 3.1. The baseline linear prediction model (LP)

The main reason for using linear prediction for speech envelopes modeling is that the vocal tract filter can be approximated by an all-pole model (Markel and Gray, 1976). LP is well-adapted for modeling speech spectra, and in particular the formants that characterize voiced speech. The LP model is obtained by means of minimization of the residual signal of a MA linear predictor, or equivalently, by means of maximization of the flatness of the residual spectrum (Kay, 1988).

Assume $X(\omega_k)$ to be the $K$-point DFT of the signal frame $x(n)$. The coefficients of the LP all-pole filter, $a_k$ are the solutions of the linear equation

$$-\sum_{k=1}^{p} R_{\mathrm{LP}}(i-k) a_k = R_{\mathrm{LP}}(i), \quad 1 \leqslant i \leqslant p, \tag{3}$$

where

$$R_{\mathrm{LP}}(i) = \frac{1}{K} \sum_{k=0}^{K} |X(\omega_k)|^2 \mathrm{e}^{\mathrm{j}\omega_k i} \tag{4}$$

is the autocorrelation sequence (acs) of the signal segment. It can be shown that the first $p$ samples of the autocorrelation sequence of the filter impulse response will match the corresponding samples of the acs of the signal.

It is well known that LP can be used to correctly estimate the spectral envelope for white noise excitation signals as long as the order of the model is sufficiently large. For harmonic excitation signals, the selection of the LP model order is more critical because with increasing order the LP model will not fit the envelope but the complete spectrum (including the harmonic structure). The usual approach to specifying the appropriate model order is based solely on the physical properties of the filter transfer function (O'Shaughnessy, 1987). While it is well known that for increasing pitch the LP model has systematic errors, no attempt was made to connect the model order to the fundamental frequency.

### 3.2. Discrete all-pole modeling

If the excitation signal is periodic the LP model introduces considerable bias into the estimation of the envelope parameters. This is because the spectral envelope is subsampled by the harmonics of the excitation signal. As a result, the acs of the input signal is an aliased version of the acs of the impulse response of the spectral envelope filter. As explained in the previous section the LP model matches the filter acs with the acs of the signal. Therefore, the LP filter estimated from a periodic signal is biased whenever the acs of the envelope filter is not time limited.

For harmonic signals, this systematic error appears in the fitted envelope as a bias of the envelope peaks towards the pitch harmonics. It is easy to see that the aliasing effect will increase with increasing fundamental frequency of the excitation signal as well as with decreasing smoothness of the spectral envelope.

The aim of the discrete all-pole model (DAP) (El-Jaroudi and Makhoul, 1991) is to solve the aliasing problem described above. The basic idea exploited with the DAP model is to fit the all-pole model using only the finite set of spectral locations that are related to the harmonic positions of the fundamental frequency. The objective function used with the DAP model is a discrete version of the Itakura-Saito error measure. If the observed and estimated speech envelopes are denoted as $S(\omega)$ and $\widehat{S}(\omega)$ respectively and the frequencies of the harmonics are given by $\omega_m$ the error measure is

$$E_{\mathrm{DAP}} = \sum_m \frac{S(\omega_m)}{\widehat{S}(\omega_m)} - \log\left(\frac{S(\omega_m)}{\widehat{S}(\omega_m)}\right) - 1. \tag{5}$$

Note that the frequency positions $\omega_m$ are not required to obey harmonic relations. Adaptive minimization of $E_{\mathrm{DAP}}$ yields the DAP estimate. As only the relevant spectral positions of the input signal are used, the systematic error of the DAP estimator is significantly lower than that of the LP model. To obtain a unique solution, however, a sufficient number of points must be used. As reported in (El-Jaroudi and Makhoul, 1991) the number $M$ of spectral peaks in the frequency range between 0 and the Nyquist frequency should exceed the model order $O$ of the all-pole model. Given the fact that for a harmonic signal with fundamental frequency $F_0$ we get $O < M = (F_s/2)/F_0$ the model order $O$ appears to be limited in a way similar to that discussed for the TE method. Nevertheless, the DAP model order is generally selected following the same guidelines as for the LP model. In Fig. 2, we show an example of LP and DAP estimation fitting the same spectrum used for the TE estimator example and equal model order.

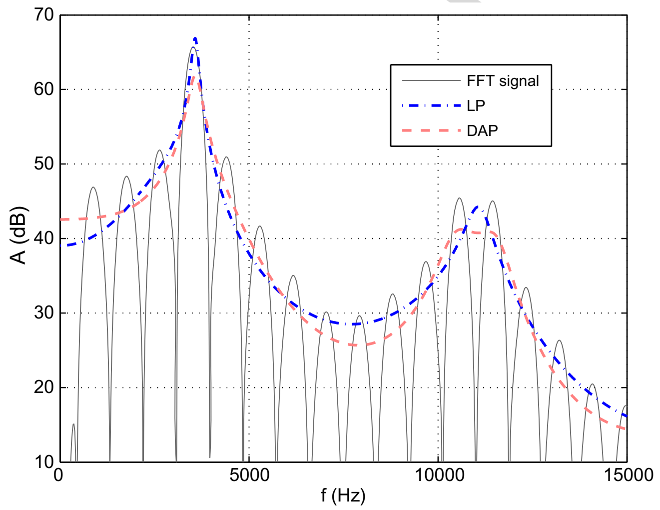



Fig. 2. LP and DAP estimations of the ARMA test spectrum (model order $O = 25$, sample rate $F_s = 44{,}100$ Hz, fundamental period $P_0 = 50$).

## 4. Experimental evaluation

In the following experiments, we try to establish relations between the envelope characteristics, the fundamental frequency, the envelope model and the model order. To be able to quantify the estimation error the test signals are synthetic ARMA signals with stationary envelope and excitation. In a real world situation a number of additional factors will affect the estimation and the estimation error will therefore generally be much higher. However, results concerning the systematic errors due to model type and order are expected to remain valid.

### 4.1. Synthetic ARMA test signals

To prevent an excessively large number of control parameters for the synthetic signals the ARMA filter envelopes of these signals are limited to rather low order. The transfer functions consist of two pairs of complex poles and two pairs of complex zeros. Because the smoothness of the spectral envelope is mostly related to the pole and zero radii we select fixed angular locations of poles and zeros. The first pole pair has angle $\pm\pi/6$ and the second $\pm\pi/2$. For the zeros the angles of $\pm 2\pi/6$ and $\pm 4\pi/6$ are used. The set of radii used to form the pole $r_p$ and zero $r_z$ locations is given by

$$r = \log(2.013 + k0.04) \quad \text{with } k = 0, 1, \ldots, 17. \tag{6}$$

The resulting radii cover the interval [0.7, 0.99]. For a sample rate of 44,100 kHz this represents 3 dB-bandwidths in the range 132–5014 Hz. These values cover the range of formant bandwidths that are common for the spectral envelopes of speech and musical instruments. For all experiments we use a Hanning window that covers exactly four periods of the fundamental frequency of the excitation signal. To prevent systematic errors that may arise due to the fact that the spectral bins do not sample harmonic peaks exactly at their local maximum, we use a DFT size that is a power of 2 at least eight times larger than the analysis window.

To restrict the dimensionality of the problem a two-dimensional grid of radii was used. The first dimension controls the radius of the four poles while the second dimension controls the radius of the zeros. The complete two-dimensional grid allows us to study transfer functions that are dominated by AR or MA filter characteristics, as well as an important number of ARMA filters. In Fig. 3, we show three limiting cases of special interest: ARMA with maximum radii ($r_p = r_z = 0.99$), AR dominated ($r_p = 0.99$, $r_z = 0.7$) and ARMA with minimal radii ($r_p = r_z = 0.7$). The input signals for the analysis are constructed by means of superposition of all the harmonics of a given fundamental frequency ($F_0$) that are below half the sample rate $F_s/2$. The set of fundamental frequencies that will be used covers the range $F_s/500 < F_0 < F_s/50$. No partial is added at 0 Hz or $F_s/2$. The harmonic excitation signal is then filtered using the respective ARMA filter.

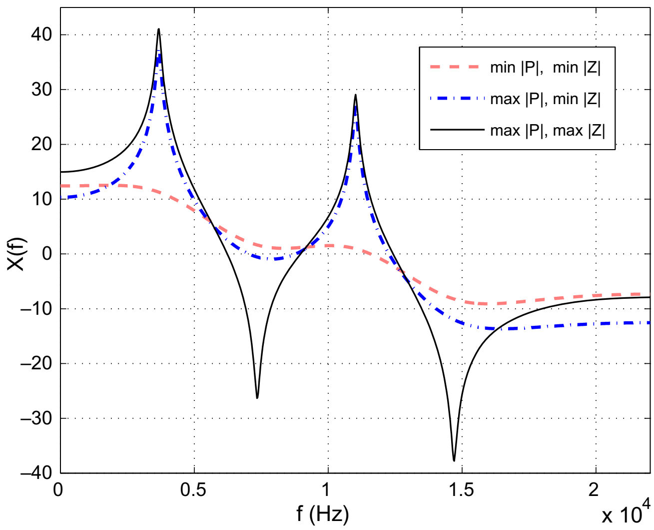



Fig. 3. Examples of the evaluation transfer functions. The solid line corresponds to a transfer function with pole/zero positions with maximal radii. The dashdotted line shows a pole dominated transfer function with maximum radii for the poles and minimum radii for the zeros, and the dashed line shows an ARMA transfer function with pole/zero positions with minimum radii.

The fundamental frequencies to be used in the experiments were selected such that the related fundamental periods cover the range $P_0 = [50, 500]$.

### 4.2. Evaluation criterion

The LP, DAP and TE algorithms are used to obtain estimates of the spectral envelope using a grid of orders covering the range $O = [5, F_s/F_0]$ and using an order increment of 5. Order values according to the proposed selection criteria are also included. To evaluate the estimation error we use the root mean square error of the log amplitude:

$$E_M = \sqrt{\frac{1}{K}\sum_{k=0}^{K-1}(\log|S(k)| - \log|\widehat{S}(k)|)^2}, \qquad (7)$$

where $S(k)$ and $\widehat{S}(k)$ are the $K$-point DFT of the filter transfer function and the estimated envelope, respectively. Using only the magnitude spectrum could appear to be a problem because phase values are not taken into account. Note however that for minimum phase filters the phases are unambiguously determined by the log amplitude spectrum. Therefore, the phase spectra would not add further information. The error measure (7) appears especially useful for algorithms that try to achieve timbre modification by means of deconvolution of the spectral envelope, because for these applications any error, whether it is underestimation or overestimation of the envelope is equally important. For applications that try to achieve formant location, another error measure would be preferable.

## 5. Experimental results

### 5.1. Order selection

In Figs. 4 and 5, two examples of the average error according to (7) are displayed as a function of the parameter $\alpha$, which represents the model order relative to the number of samples contained in the fundamental period $P_0 = F_s/F_0$ of the excitation signal. Two sets of transfer functions are used to calculate the averages of the estimation error. The first set contains all transfer functions described above, and the second set fixes the zero radius to the lowest value ($r_z = 0.7$), such that only the subset of transfer functions that is closest to an AR model is taken into account.

The comparison of the results obtained in the various experiments reveal that the error generally decreases with increasing $P_0$. This is related to the fact that with a larger period more harmonics are used to sample the envelope, which facilitates the estimation. It is interesting to note that the TE estimator does saturate at a level of about 0.01 dB which is achieved for $P_0 = 400$. This is easy to understand given the fact that the TE convergence criterion that we use stops the iterative adaptation whenever the estimated envelope has approached the observed spectral peaks to less than 0.01 dB. If we locate the order that provides the minimum error we find that for both sets of transfer functions the average optimal order of the TE estimator is always close to the expected position $\alpha_{TE} = 0.5$ that represents the Nyquist frequency limit. Note that if the number of spectral peaks is large ($P_0 = 400$) the optimal order corresponds to an $\alpha_{TE}$ that is slightly lower than 0.5. This can be attributed to the fact that the model starts becoming overly complex given the set of transfer functions that is being used. Even in this case, however, the order selected according to (2) achieves an error which remains close to the optimal error.

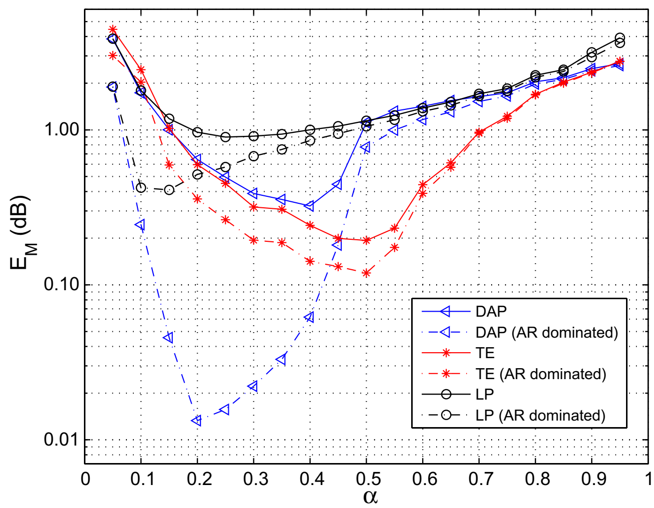



Fig. 4. Average envelope estimation errors for ARMA and AR-dominated transfer function sets as a function of $\alpha$ ($P_0 = 100$).

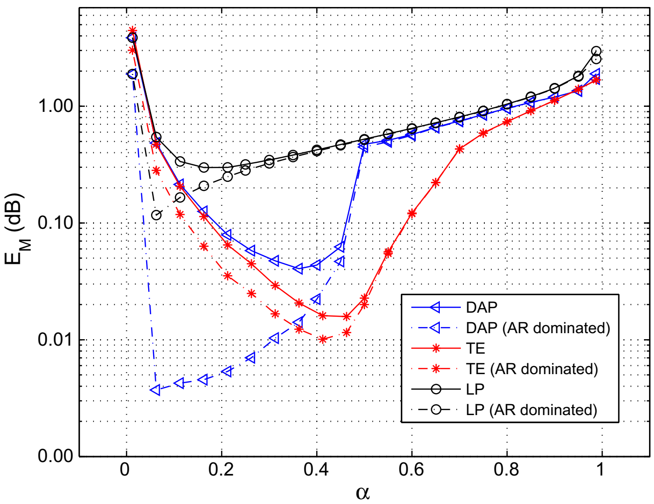



Fig. 5. Average envelope estimation errors for ARMA and AR-dominated transfer function sets as a function of $\alpha$ ($P_0 = 400$).

A similar behavior can be observed for the whole set of ARMA envelopes and the DAP method. Here, however, the optimal position $\alpha_{\mathrm{DAP}} = 0.4$ is slightly below the theoretical limit. For the AR dominated subset of transfer functions we find that the optimal DAP order is much smaller. The strong dependency of the optimal DAP order on the characteristics of the transfer function reveals a problem with the DAP estimator. The experimental results show that for orders that are close to the limit required for a unique solution ($\alpha < 0.5$) the objective function (5) does not provide enough information to bind the AR model, and therefore, the error begins increasing for model orders well below $\alpha < 0.5$. For the underdetermined case ($\alpha \geqslant 0.5$) DAP performs similarly to LP. This can be understood as the LP solution is used to initialize the iterative DAP procedure. We can conclude that for the underdetermined case the LP solution is already part of the solution manifold of the objective function (5). While the value of the optimal order of the DAP model depends strongly on the characteristics of the filter transfer function we may nevertheless conclude that when no a priori information about the system structure is available, which is generally true for real world situations, the DAP order should be selected using (2) with $\alpha_{\mathrm{DAP}} = 0.4$. Note that a particular disadvantage of the DAP estimator is the fact that the estimated filter may be unstable (El-Jaroudi and Makhoul, 1991).

When considering the LP model we find that the minimum error for the complete set of filters is found for values of $\alpha$ in the range [0.05, 0.25]. The lower order of the LP method makes sense as the LP estimator uses the complete spectrum to adapt its parameters such that a lower order suffices in order to prevent adaptation to the harmonic structure. Again due to the fact that the complete spectrum affects the LP estimate, it appears questionable whether the optimal value for $\alpha$ found in these experiments can be generalized.

### 5.2. Order selection evaluation

So far the order selection criterion has been validated using only two values of the fundamental frequency. In Figs. 6 and 7, the experimentally obtained optimal orders for the ARMA and AR-dominated set of transfer functions and a large range of fundamental frequencies are compared with the orders selected according to the proposal described above. Figs. 8 and 9 compare the modeling error for the same set of experiments and orders.

The results displayed in these figures show that the relations that have been discussed for the fundamental periods $P_0 = 100$ and $P_0 = 400$ in the previous section are valid for

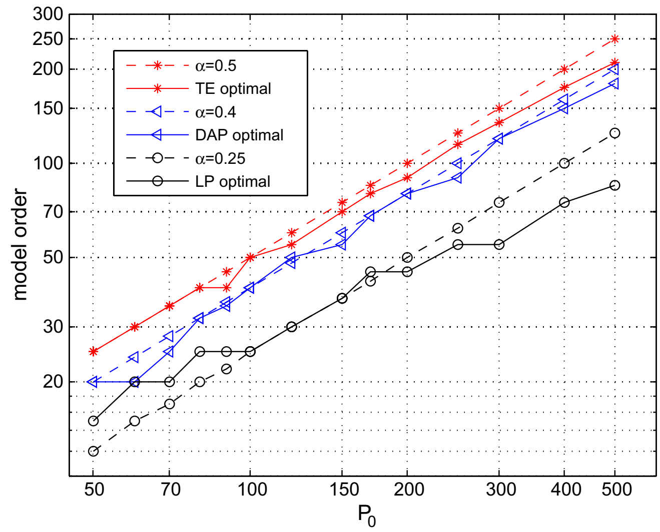



Fig. 6. Comparison of the experimental optimal order and the proposed order selection on the ARMA set.

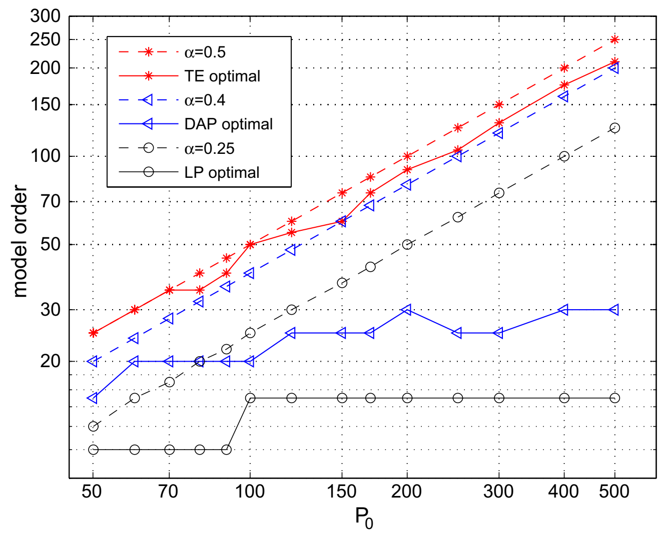



Fig. 7. Comparison of the experimental optimal order and the proposed order selection on the AR-dominated set.

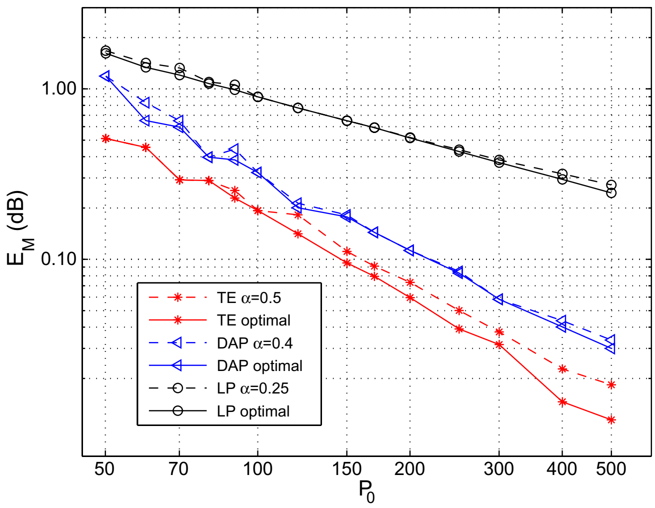


Fig. 8. Average modeling error using experimental optimal order and the proposed order selection (ARMA set).

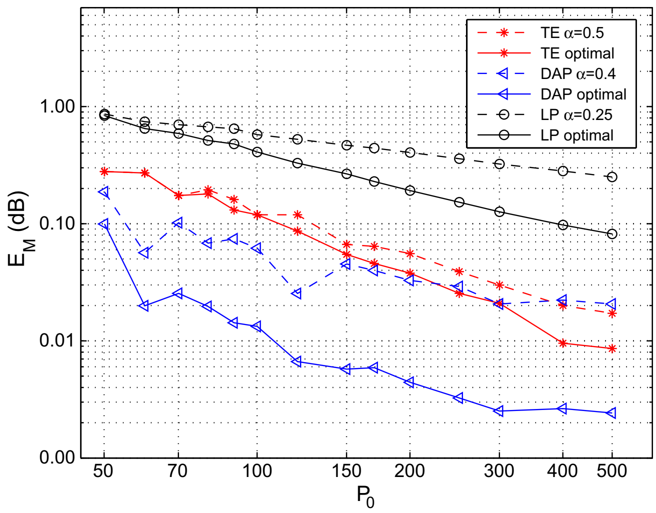


Fig. 9. Average modeling error using experimental optimal order and the proposed order selection (AR-dominated set).

the complete range of fundamental frequencies. As a result of the experiments described so far we conclude that the model order for TE and DAP estimation should be selected as a function of the fundamental frequency using $\alpha_{\mathrm{TE}}$ and $\alpha_{\mathrm{DAP}}$ in (2) as long as no a priori information about the underlying transfer function is available. For LP, no conclusive suggestion can be made.

### 5.3. Model selection

Having obtained a simple means to select proper orders for at least the DAP and TE estimator we aim to investigate the relation between the envelope characteristics and the model that performed best.

Comparing the average estimation error of the different estimators we find that LP always displays the worst

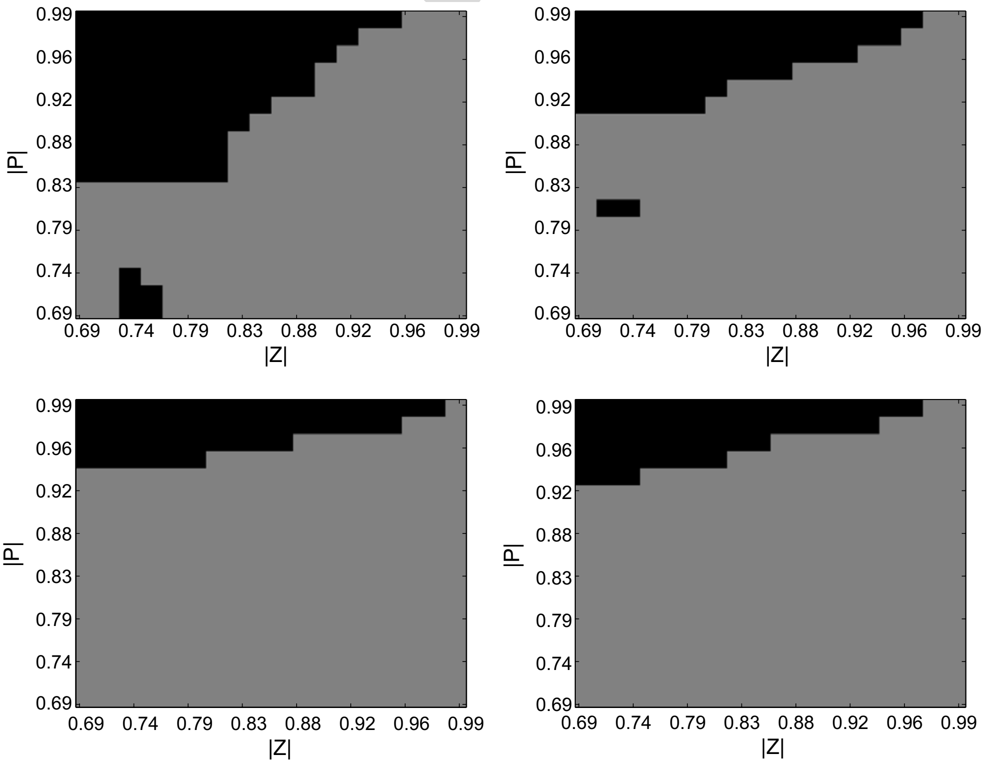


Fig. 10. Best estimator for each individual transfer function and fundamental periods $P_0 = 100$ (left, top), $P_0 = 200$ (right, top), $P_0 = 300$ (left, bottom), and $P_0 = 400$ (right, bottom) samples using the experimental optimal model orders. Black (DAP) and gray (TE) areas denote the model that performed best in the experiments. LP (white) is never the best model.

performance. TE performs better for the complete ARMA set and is outperformed by DAP only for the AR dominated set of transfer functions and only if the experimentally derived optimal order is used. If no information about the optimal order is available, DAP is only slightly better, even for the case of AR dominated transfer functions.

In Fig. 10, we display the model that had the smallest estimation error for four fundamental periods and for each filter transfer function used in the experiment. The estimators are color coded. Black represents DAP, gray TE, and white LP. Note that for Fig. 10 the order selection scheme was not used.

The figures indicate that the TE method is the best for all fundamental periods in the majority of cases. DAP performs better only if the pole radii are significantly larger then the zero radii. The LP model never outperformed the other methods. The advantage of DAP with respect to AR models is especially prominent for small periods (high $F_0$). As was made clear in the previous section, the advantage of DAP for AR dominated envelopes diminishes when the optimal order cannot be established because the characteristics of the target envelope are unknown.

## 6. Conclusions

The article has presented an experimental comparison of envelope estimation techniques for pitched excitation signals. The main goal of the investigation was to establish experimental evidences for a simple scheme that derives a proper model order from the fundamental frequency of the observed signal. A slight modification of the true-envelope estimator was required to be able to achieve optimal performance with the suggested order selection scheme. The experiments indicate that for the modified TE estimator, the Nyquist frequency is a proper indicator for model order selection. Accordingly, the appropriate order for the TE estimator (or related cepstral estimation methods) is equal to 0.5 times the number of samples per period of the fundamental frequency. In a similar manner for the DAP estimator, the model order should be limited to 0.4 times the number of samples per period of the fundamental frequency. While it is generally true that the best model order depends not only on the fundamental frequency but also on the model structure and the specific transfer function, the experimental findings support the conclusion that an appropriate choice can often be made only based upon the fundamental frequency of the signal.

Direct comparison of the estimators demonstrated that the LP model is clearly the worst estimator. A comparison of DAP and TE estimators revealed that the choice of the optimal estimator depends on the fundamental frequency and the envelope characteristics. For arbitrary envelopes the TE estimator seems to be the better choice.

## Acknowledgements

Fernando Villavicencio's work is supported by a grant (Reg. 143775) from the National Council for Science and Technology of Mexico (CONACYT).

## References

Cappé, O., Moulines, E., 1996. Regularization techniques for discrete cepstrum estimation. IEEE Signal Process. Lett. 3 (4), 100–102.

El-Jaroudi, A., Makhoul, J., 1991. Discrete all-pole modeling. IEEE Trans. Signal Process. 39 (2), 411–423.

Galas, T., Rodet, X., 1990. An improved cepstral method for deconvolution of source filter systems with discrete spectra: Application to musical sound signals. In: Proc. Internat. Computer Music Conf. (ICMC), pp. 82–84.

Imai, S., Abe, Y., 1979. Spectral envelope extraction by improved cepstral method. Electron. Comm. (in Japan) 62 (4), 10–17.

Kay, S., 1988. Modern Spectral Estimation: Theory and Application, Prentice-Hall Signal Processing Series, USA.

Makhoul, J., 1975. Linear prediction: A tutorial review. Proc. IEEE 63 (4), 561–580.

Markel, J.D., Gray, A.H., 1976. Linear Prediction of Speech. Springer Verlag.

O'Shaughnessy, D., 1987. Speech Communication. In: Addison-Wesley Series in Electrical Engineering: Digital Signal Processing. Addison-Wesley Publishing Company.

Roebel, A., Rodet, X., 2005. Efficient spectral envelope estimation and its application to pitch shifting and envelope preservation. In: Proc. DAFX'05, Spain.

Smith, J.O., 2005. Introduction to Digital Filters. Prentice-Hall, USA, September 2005 Draft.