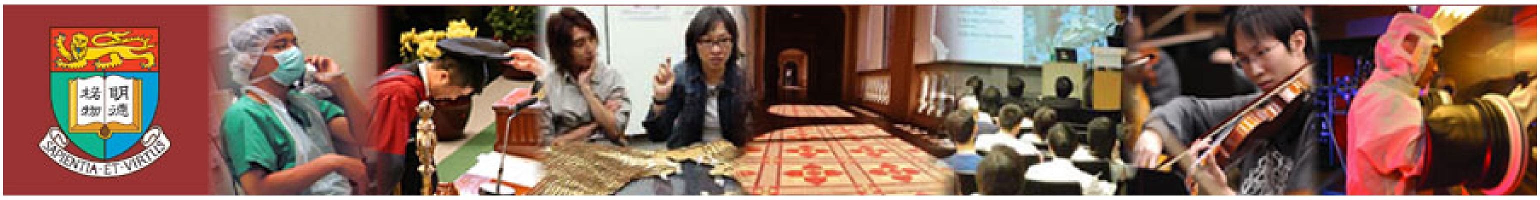

A New Recursive Algorithm for Time-Varying Autoregressive (TVAR) Model Estimation and Its Application to Speech Analysis

Y. J. Chu, S. C. Chan, Z. G. Zhang and K. M. Tsui
Electrical and Electronic Engineering Department, The University of Hong Kong, Hong Kong, P. R. China
E-mail: { yjchu; scchan; zgzhang; kmtsui}@eee.hku.hk

***Abstract*—This paper proposes a new state-regularized (SR) and QR decomposition based recursive least squares (QRRLS) algorithm with variable forgetting factor (VFF) for recursive coefficient estimation of time-varying autoregressive (AR) models. It employs the estimated coefficients as prior information to minimize the exponentially weighted observation error, which leads to reduced variance and bias over traditional regularized RLS algorithm. It also increases the tracking speed by introducing a new measure of convergence status to control the FF. Simulations using synthetic and real speech signals show that the proposed method has improved tracking performance and reduced estimation error variance than conventional TVAR modeling methods during rapid changing of AR coefficients.**

I. INTRODUCTION

Autoregressive (AR) modeling of signals is frequently employed in applications such as speech analysis, denoising and enhancement, vocal identification, spectrum estimation, biomedical engineering, vibration analysis, etc. In practice, the signal under consideration may be time-varying (TV) and hence the model coefficients are also time-varying. For instance, in AR modeling of speech signals, significant non-stationarity is encountered during transitions from one phoneme to the other. In conventional AR modeling of speech signal, the time variation is usually assumed to be small in a short time interval and hence the speech signal is usually windowed, inside which the AR coefficients are assumed to be fixed and estimated using covariance or autocorrelation methods [1]. This is also known as the classical piecewise-constant AR linear prediction (LP), which implies that within each analysis interval, the signal is assumed to be stationary. In reality, however, the vocal tract is continually changing, either slowly or rapidly.

To partially reconcile the time-varying nature of the vocal tract, one can divide the signal into smaller segments. However, it may lead to ill-posed problems [2] due to insufficient number of samples. Therefore, it is more natural to use TVAR model with time-varying AR coefficients for better approximation of real vocal tracts [3][4]. A way to implement this idea is to expand each AR model coefficient using a basis function, such as a polynomial. The coefficients of the basis functions are then obtained by using least squares error estimation such as time-varying LP (TVLP) [4]. However, when the signal changes rapidly, the accuracy will be severely degraded. This usually leads to instability problem especially when the true poles are close to the unit circle. On the other hand, several online or recursive estimation methods [5][6] were proposed to handle the TVAR modeling of speech recently. In [5], particle filtering was used for modeling and enhancement of speech signals, while in [6], the TVAR model is identified by using a maximum *a posteriori* estimation. They are more suitable for time-varying system identification and tracking than TVLP. However, they are usually very computationally expensive.

Another useful approach is to employ efficient adaptive filtering algorithms, such as the RLS algorithm, to estimate the AR coefficients recursively. To deal with the TV nature of the AR coefficients, the VFF approach is usually employed to achieve fast tracking and low steady state mean square error [7] [8]. However, in case of rapidly changing AR models, the VFF RLS algorithm may lead to large estimation variance if a small FF is used. To address this issue, a new state-based regularized RLS algorithm with VFF is proposed. First, a novel state regularization, which employs current estimated coefficients as prior information to minimize the observation errors, is developed. The concept is intimately connected to the Kalman filter and the least mean squares algorithm, except that infinite number of measurements is employed. This SR prevents significant fluctuation of the estimated TVAR coefficients due to small FF or insufficient observations. Moreover, based on the VFF framework in [9], we propose a VFF scheme using a new measure of convergence status to update the FF. This further improves the tracking speed, when the vocal tracts change rapidly. The proposed algorithm can also be implemented by the QRD structure which has the advantages of low roundoff error and efficient hardware realization over the direct implementation. The arithmetic complexity is $O(p^2)$, where p is the order of the AR model. Its effectiveness is demonstrated by computer simulations and comparison with classical TVLP algorithm in [4].

II. TVAR MODEL AND THE QRRLS ALGORITHM

In the TVAR model, the nonstationary discrete-time signal $x(n)$ at time instant n is modeled as:

This study was partially supported by a grant from the Research Grants Council of the Hong Kong SAR, China.

978-1-4673-0219-7/12/$31.00 ©2012 IEEE

$$x(n)=\sum_{i=1}^{p}a_i(n)x(n-i)+g(n)=\boldsymbol{a}^T(n)\boldsymbol{x}(n)+g(n)\,, \qquad (1)$$

where p is the order of the model, $\boldsymbol{a}(n)=[a_1(n),a_2(n),\cdots,a_p(n)]^T$ is the coefficient vector, $\boldsymbol{x}(n)=[x(n-1),...,x(n-p)]^T$ is the signal vector, and $g(n)$ is the modeling error which is usually assumed to be a zero-mean white Gaussian process with variance σ_η^2.

In this work, we propose to track recursively the TVAR coefficients using a novel VFF SR QRRLS adaptive filtering algorithm. Let the weight vector of the adaptive filter be $\boldsymbol{w}(n)=[w_1(n), w_2(n),\ldots,w_p(n)]^T$, which serves as an estimate of the unknown TVAR coefficients $\boldsymbol{a}(n)$. The adaptive filter aims to minimize the estimation error $e(n) = x(n) - \boldsymbol{x}^T(n)\boldsymbol{w}(n)$ recursively. In the RLS algorithms, the following least squares cost function is minimized

$$J(n)=\sum_{i=0}^{n}\lambda_{n-i}(n)e^2(i)\,, \qquad (2)$$

where $\lambda_{n-i}(n)=\lambda(n)\lambda_{n-i-1}(n-1), 0\le i<n$ serves the purpose of an exponential window which puts less weight to errors at distant past. Here, $\lambda_0(n)=1$ and $\lambda(n)$ is the FF used at the time index n, which usually satisfies $0<\lambda(n)<1$. For example, $\lambda_{n-i}(n)$ can be chosen as λ^{n-i}, where λ is a constant in conventional RLS algorithm or updated adaptively as in VFF algorithms.

By setting the first partial derivative of $J(n)$ with respect to (w.r.t.) $\boldsymbol{w}(n)$ to zero, one finds that the optimal weight vector satisfies the following normal equation:

$$\boldsymbol{R}_{XX}(n)\boldsymbol{w}_{opt}(n)=\boldsymbol{p}_X(n)\,, \qquad (3)$$

where $\boldsymbol{R}_{XX}(n)=\sum_{i=0}^{n}\lambda_{n-i}(n)\boldsymbol{x}(i)\boldsymbol{x}^T(i)$ and $\boldsymbol{p}_X(n)=\sum_{i=0}^{n}\lambda_{n-i}(n)x(i)\boldsymbol{x}(i)$ are the covariance matrix of $\boldsymbol{x}(n)$ and the cross-correlation vector of $x(n)$ and $\boldsymbol{x}(n)$, respectively. Applying the matrix inversion lemma to (3), the following RLS algorithm can be obtained [10]

$$\boldsymbol{P}(n)=\lambda^{-1}(n)(\boldsymbol{I}-\boldsymbol{k}(n)\boldsymbol{x}^T(n))\boldsymbol{P}(n-1)\,, \qquad (4a)$$

$$\boldsymbol{k}(n)=\frac{\boldsymbol{P}(n-1)\boldsymbol{x}(n)}{\lambda(n)+\boldsymbol{x}^T(n)\boldsymbol{P}(n-1)\boldsymbol{x}(n)}\,, \qquad (4b)$$

$$\boldsymbol{w}(n+1)=\boldsymbol{w}(n)+(x(n)-\boldsymbol{w}^T(n)\boldsymbol{x}(n))\boldsymbol{k}(n)\,, \qquad (4c)$$

where $\boldsymbol{P}(n)$ is the recursive update of $\boldsymbol{R}_{XX}^{-1}(n)$ and $\boldsymbol{I}$ is an identity matrix with appropriate dimension. Eqn. (4) can also be efficiently implemented using a QR-based algorithm [11] as summarized in Table I with the first update only. This QRRLS algorithm is mathematically equivalent to but has a higher numerical stability than the direct implementation. The arithmetic complexity is of order $O(p^2)$.

III. STATE-REGULARIZED VFF QRRLS ALGORITHM

A. The State-Regularized QRRLS Algorithm

In speech analysis and related applications, the input to the adaptive filter may not be persistently exciting when the input data is insufficient. This may also occur when a small FF is used to track rapidly changing vocal tracts. Consequently, the covariance matrix $\boldsymbol{R}_{XX}(n)$ may be ill-conditioned and a large estimation error variance will result. To address this problem, a regularization term on adaptive filter coefficients, $\kappa(n)\|\boldsymbol{w}(n)\|_2^2$, is usually imposed on the objective function in (2) to limit the variation in $\boldsymbol{w}(n)$. The solution, instead of (3), will be modified to

$$(\boldsymbol{R}_{XX}(n)+\kappa(n)\boldsymbol{D})\boldsymbol{w}(n)=\boldsymbol{p}_X(n)\,, \qquad (5)$$

where $\kappa(n)$ is a regularization parameter and $\boldsymbol{D}$ is a positive definite matrix. We note that the rank-one update of $\boldsymbol{R}_{XX}(n) = \boldsymbol{R}_{XX}(n-1)+\boldsymbol{x}(n)\boldsymbol{x}^T(n)$ in conventional QRRLS can be efficiently implemented by updating the Cholesky factor $\boldsymbol{R}(n)$ of $\boldsymbol{R}_{XX}(n)$ recursively using the QRD as shown in recursion (i) of Table I. The update of the term $\kappa(n)\boldsymbol{D}$ in (5) is however complicated since it is of full rank. According to [11], an L_2 regularized can be applied successively using QRD. In particular, QRD is executed once for the data vector $[\boldsymbol{x}^T(n), x(n)]$ and once for the regularization vector $[\sqrt{\mu(n)}\boldsymbol{d}_l, 0]$ at each time instant, where $\boldsymbol{d}_l$ is the l-th row of the regularization matrix $\boldsymbol{D}$ and $\mu(n)=\kappa(n)p$. If the vector is sequentially applied, then $l=(n \bmod p)+1$. Therefore, the complexity is twice that of the RLS algorithm. If the regularization parameter $\kappa(n)$ is made variable at each iteration, it yields the variable L_2 regularized QRRLS algorithm. In [11], the regularization parameter $\kappa(n)$ was proposed to balance between bias and variance errors as:

$$\kappa(n)=\bar{\sigma}_x^2\lambda^{-1}(n)\sqrt{1-\lambda(n)}\sqrt{\gamma(\sigma_\eta^2/\sigma_x^2(n))/\|\boldsymbol{w}_0\|_2^2}\,, \qquad (6)$$

where $\gamma=1/(\frac{1}{p}(2+\frac{(1-\lambda(n))p}{\lambda(n)})+\frac{(1-\lambda(n))p}{\lambda^2(n)})$, $\bar{\sigma}_x^2$ is the averaged input power over the whole duration while $\sigma_x^2(n)$ is the short term averaged input power which can be estimated by using a FF, and $\|\boldsymbol{w}_0\|_2^2$ is the norm of the system channel which is usually assumed to be known a priori. By using L_2 regularization, the ill-conditioned problem can be improved significantly. It can be seen from (5) that L_2 regularization introduces a bias to the true solution, especially when a large $\kappa(n)$ is used. To solve this problem, let us rewrite (3) as

$$(\boldsymbol{R}_{XX}(n)+\kappa(n)\boldsymbol{I})\boldsymbol{w}(n)=\boldsymbol{p}_X(n)+\kappa(n)\boldsymbol{w}(n)\,, \qquad (7)$$

where $\boldsymbol{D}$ has been chosen as an identity matrix. First, it can be seen that the optimal solution to (7) is identical to that of (3).

TABLE I THE R-QRRLS/VR-QRRLS ALGORITHM

Initialization:

$\boldsymbol{R}(0)=\sqrt{\delta}\boldsymbol{I}$, δ is a small positive constant;

$\boldsymbol{u}(0)=\boldsymbol{0}$, $\boldsymbol{w}(0)=\boldsymbol{0}$ are null vectors.

Recursion:

Given $\boldsymbol{R}(n\text{-}1)$, $\boldsymbol{u}(n\text{-}1)$, $\boldsymbol{w}(n)$, $\boldsymbol{x}(n)$ and $x(n)$, compute at time n:

(i). The first update:

$$\begin{bmatrix}\boldsymbol{R}^{(1)}(n) & \boldsymbol{u}^{(1)}(n)\\ \boldsymbol{0}^T & c^{(1)}(n)\end{bmatrix}=\boldsymbol{Q}^{(1)}(n)\begin{bmatrix}\sqrt{\lambda(n)}\boldsymbol{R}(n-1) & \sqrt{\lambda(n)}\boldsymbol{u}(n-1)\\ \boldsymbol{x}^T(n) & x(n)\end{bmatrix}$$

The second update (regularized RLS algorithms only):

a) For the L_2 regularization

$$\begin{bmatrix}\boldsymbol{R}(n) & \boldsymbol{u}(n)\\ \boldsymbol{0}^T & c(n)\end{bmatrix}=\boldsymbol{Q}(n)\begin{bmatrix}\boldsymbol{R}^{(1)}(n) & \boldsymbol{u}^{(1)}(n)\\ \sqrt{\mu(n)}\boldsymbol{d}_l & 0\end{bmatrix},$$

b) For the state-based regularization

$$\begin{bmatrix}\boldsymbol{R}(n) & \boldsymbol{u}(n)\\ \boldsymbol{0}^T & c(n)\end{bmatrix}=\boldsymbol{Q}(n)\begin{bmatrix}\boldsymbol{R}^{(1)}(n) & \boldsymbol{u}^{(1)}(n)\\ \sqrt{\mu(n)}\boldsymbol{d}_l & \sqrt{\mu(n)}w_l(n-1)\end{bmatrix}$$

where $\boldsymbol{Q}^{(1)}(n)$ and $\boldsymbol{Q}(n)$ are calculated by Givens rotation to obtain the left hand side of each equation above and $\mu(n)=\kappa(n)p$. For the QRRLS algorithm, $\boldsymbol{R}^{(1)}(n)=\boldsymbol{R}(n)$.

(ii). $\boldsymbol{w}(n)=\boldsymbol{R}^{-1}(n)\boldsymbol{u}(n)$ (back-substitution).

Secondly, the matrix $\boldsymbol{R}_{XX}(n)+\kappa(n)\boldsymbol{I}$ for a sufficiently large $\kappa(n)$ is positive definite and hence invertible. Therefore, the regularization in (7) is unbiased and depends on the state $\boldsymbol{w}(n)$. To iteratively solve (7), the weight vector $\boldsymbol{w}(n)$ on the right hand side is approximated by its values in the previous iteration, i.e. $\boldsymbol{w}(n-1)$. Hence, the algorithm is asymptotically unbiased. The relationship between this simplified version of (7) and the LMS algorithm can be seen by considering a single measurement, $\boldsymbol{x}(n)$, at time instant n. In this case, $\boldsymbol{R}_{XX}(n) \approx \boldsymbol{x}(n)\boldsymbol{x}^T(n)$, $\boldsymbol{p}_X(n) \approx \boldsymbol{x}(n)x(n)$, and hence it is unable to obtain a unique solution to (7). However, the relaxation $\boldsymbol{w}(n) \approx \boldsymbol{w}(n-1)$ allows us to utilize the prior information obtained up to the $(n$-1)-th iteration and the current information. Actually, Eqn. (7) after some manipulations, can be rewritten as

$$\boldsymbol{w}(n) = \boldsymbol{w}(n-1) + \frac{\boldsymbol{x}(n)e_p(n)}{\kappa(n) + \boldsymbol{x}^T(n)\boldsymbol{x}(n)}, \quad (8)$$

where $e_p(n)=x(n)-\boldsymbol{x}^T(n)\boldsymbol{w}(n-1)$. One immediately recognizes that Eqn. (8) is a variable step-size LMS algorithm with the step-size $\mu(n) = (\kappa(n) + \boldsymbol{x}^T(n)\boldsymbol{x}(n))^{-1}$, which can also be viewed as a normalized LMS algorithm with unity step-size. Thus, the relaxed form of (7) will reduce to the LMS algorithm when one measurement is used. This partially explains the improved performance of the LMS algorithm when the input changes considerably. On the other hand, in order to avoid an excessive bias during tracking, the number of measurements used in the RLS algorithm should be reduced by using a small FF. In this case, the RLS algorithm, without using the prior state information, $\boldsymbol{w}(n-1)$, may become unstable. Moreover, in a stationary environment, a large FF can be used together with the QRD to utilize all the available measurements. Since the regularization in (7) involves changing the covariance matrix to $\boldsymbol{R}_{XX}(n)+\kappa(n)\boldsymbol{I}$ and the cross correlation vector to $\boldsymbol{p}_X(n)+\kappa(n)\boldsymbol{w}(n-1)$, which is a function of the previous state vector, we shall call it "state regularization". In fact, the relaxation now becomes

$$\boldsymbol{I} \cdot \boldsymbol{w}(n) \approx \boldsymbol{w}(n-1), \quad (9)$$

after removing the 1st term on both sides of (7). One can view (9) as a state equation, $\boldsymbol{w}(n)=\boldsymbol{w}(n-1)+\boldsymbol{g}_w(n)$ with $\boldsymbol{g}_w(n)$ being the state noise, which requires the current weight vector to stay close to the previous weight vector in case the number of relevant measurements is limited at nonstationary environment.

Finally, we note that the variable regularization parameter described in (6) can be used together with the proposed state regularized QRD to overcome the problem due to non-persistent excitation. To solve the relaxed form of (7) recursively, we need to find the QRD of the matrix $(\boldsymbol{R}_{XX}(n)+\kappa(n)\boldsymbol{I})$. This can be implemented in a similar way as the L_2 regularized QRRLS. To implement the last term at the right hand side of (7), we found, instead of appending $[\sqrt{\mu(n)}\boldsymbol{d}_l, 0]$ to the second QRD, the state regularization can be approximately implemented by appending the row vector $[\sqrt{\mu(n)}\boldsymbol{d}_l, \sqrt{\mu(n)}w_l(n-1)]$ to the second QRD successively, where $w_l(n-1)$ is the *l-th* element of $\boldsymbol{w}(n-1)$. The regularization parameter $\kappa(n)$ can also be updated as in (6). This yields the proposed state-regularized QRRLS (SR-QRRLS) algorithm, as shown in Table I ((i).b).

B. SR-QRRLS Algorithm with Variable Forgetting Factor

As mentioned earlier, the FF plays an important role in the performance of RLS algorithms [7][8]. Intuitively, the FF controls how the measurements are used in estimation. Less number of measurements should be used if the estimation variance increases due to nonstationary inputs or systems. Here, we propose a measure of the estimation variance of $\boldsymbol{w}(n)$ to determine the number of measurements and hence FF to be used. It is known from classical performance analysis of the LMS algorithm for Gaussian inputs [12] $E[\boldsymbol{x}(n)e(n)] = E[\boldsymbol{x}(n)(\boldsymbol{x}^T(n)(\boldsymbol{w}_0(n)-\boldsymbol{w}(n))+\eta(n))] = \boldsymbol{R}_{XX}E[\boldsymbol{v}(n)]$, where $\boldsymbol{v}(n) = \boldsymbol{w}_0(n) - \boldsymbol{w}(n)$ is the weight error vector. Therefore, a good measure of convergence status is the norm of its time average:

$$\sigma_{xe}^2(n) = \lambda_e \sigma_{xe}^2(n-1) + (1-\lambda_e)\boldsymbol{x}_e^T(n)\boldsymbol{x}_e(n), \quad (10)$$

where λ_e is a forgetting factor and $\boldsymbol{x}_e(n)$ is averaged from $\boldsymbol{x}(n)e(n)$ over a time window of length T_s so as to suppress the effect of background noise on $\sigma_{xe}^2(n)$. By adopting the approach in [9], we propose to estimate the exponential window size of the algorithm, $L(n)$, at each time instant from the measure in (10) as follows:

$$L(n) = round\{L_L + [1 - g(\overline{G}_N(n))](L_U - L_L)\}, \quad (11)$$

where $\overline{G}_N(n) = \sigma_{xe}^2(n)/\overline{\sigma}_0^2$ with $\overline{\sigma}_0^2$ the average of the first T_{s0}-th estimates of $\sigma_{xe}^2(n)$ at the beginning of adaptation, the operator $round\{.\}$ rounds its argument to the nearest integer, L_L and L_U are respectively lower and upper bounds of $L(n)$, and $g(x)=\min\{x,1\}$ is a clipping function which keeps its positive argument x within the interval [0, 1]. From (11), the factor then can be estimated as:

$$\lambda(n) = 1 - 1/L(n). \quad (12)$$

Eqns. (10)-(12) yield the proposed VFF scheme for SR-VFF- QRRLS. It can be seen that if the system changes, $\sigma_{xe}^2(n)$ comparable to or larger than $\overline{\sigma}_0^2$ is experienced, and a small FF will be used to obtain fast tracking speed while at the steady state, $\overline{G}_N(n)$ is usually small and a larger FF will be employed to obtain a smaller MSE. Therefore, $\overline{\sigma}_0^2$ serves as a reference for $\sigma_{xe}^2(n)$ to control the FF through (11) and (12). To reduce the effect of input power on $\overline{\sigma}_0^2$, we assume that an approximate nominal signal level $\hat{\sigma}_x^2$, is available. By recording the signal level during the computation of $\overline{\sigma}_0^2, \hat{\sigma}_{x0}^2$, a correction factor $\delta = \hat{\sigma}_x^2/\hat{\sigma}_{x0}^2$ for $\overline{\sigma}_0^2$ can be used.

IV. SIMULATION RESULTS

A. Simulation Results with Synthetic Signal

To evaluate the proposed recursive algorithm for TVAR modeling of speech signals, a synthetic data generated by all-pole filters with known time-varying coefficients is first used so as to compare with the ground truth. It is synthesized by filtering a white Gaussian noise through a second-order digital resonator. The sampling frequency is 8 kHz and the speech segment is 0.1 second. The proposed algorithm and the classical TVLP using covariance power method without windowing [4] are examined.

The first set of test involves the response to a "formant-like" signal with step changes in the center frequency of the generated TVAR model, which is shown in the top left panel of Fig. 1. The TVLP uses a 4-order TVAR model with each coefficient being a quadratic power series (q=2). From experiments, the following parameters for VFF scheme are recommended so as to achieve satisfactory performance under a wide range of conditions: a short window length T_s = 20 and λ_e = 0.9 are used to achieve a quick response when the system changes rapidly and a longer window T_{s0} = 50 is used to estimate a more reliable reference for the convergence status $\bar{\sigma}_0^2$ and the correction factor δ is chosen to be 5. L_L and L_U are chosen as 2 and 50, respectively, so that the minimum and maximum FFs are around 0.5 and 0.98. The performance in this setting w.r.t. the coefficients and pole trajectories, is also presented in Fig. 1. It can be seen from the coefficient trajectories that the SR-VFF-QRRLS algorithm has much better tracking performance than the conventional TVLP method. It also shows the TVLP result usually deviates substantially from the true values (at the 0.02 and 0.09 second) in order to fit fast changing tracts, which may cause the problem of instability (i.e. poles are outside the unit circle), as shown in the top right panel of Fig. 1. The SR-VFF-QRRLS algorithm, on the other hand, uses the SR to reduce the variance of the estimated coefficients, and thus helps to avoid poles of the AR model from getting out of unit circle, except during fast tracking when extremely large FF is used.

B. Simulation Results with Real Speech

A similar simulation is carried out by using a speech segment "tea" from [13] with a length of 0.1 second, which is shown in the top left panel of Fig. 2. It was downsampled to 2 kHz in order to focus on low frequency formants. A 12-order model is applied to both algorithms and power series with q=4 is used for TVLP. Other settings are identical to that in the previous example, except λ_e = 0.9 are used to track fast power changing in real data and the regularization parameter needs to be multiplied with a constant factor around 5 to reflect the increase in noise power. Results are presented in Fig. 2. The spectrogram only shows the first and second tracts. The estimated tracts of the proposed algorithm agree well with the real one while the conventional TVLP algorithm suffers from slow tracking and significant deviation. As for the pole trajectories, the TVLP have many poles near or even outside the unit circle, which may cause instability.

V. CONCLUSION

A new SR-VFF-QRRLS algorithm has been presented for the TVAR modeling of speech signal. It stabilizes the update using previous estimated filter coefficients and selects the number of measurements adaptively by means of VFFs. Improved tracking performance and reduced variance over TVLP modeling can be achieved. Its effectiveness for speech modeling has been demonstrated by computer simulations and real speech signals.

REFERENCES

[1] J. I. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 32, pp. 561-582, April 1975.

[2] D. Rudoy and T. Georgiou, "Regularized parametric models of nonstationary processes," in *Proc. the 19th International Symposium on Mathematical Theory of Networks and Systems* (*MTNS2010*), Budapest, Hungary, pp. 5-9, July 2010.

[3] L. A. Liporace, "Linear estimation of nonstationary signals," *J. Acoust. Soc. Amer*., vol. 58, no. 6, pp. 1268-1295, Dec. 1975.

[4] M. G. Hall, A. V. Oppenheim, and A. S. Willsky, "Time-varying parametric modeling of speech," *Signal Process.*, vol. 5, pp. 267-285, 1983.

[5] J. Vermaak, C. Andrieu, A. Doucet, and S. J. Godsill, "Particle methods for Bayesian modeling and enhancement of speech signals," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 3, pp. 173-185, Mar. 2002.

[6] T. Hsiao, "Identification of time-varying autoregressive systems using maximum *a posteriori* estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3497-3509, Aug. 2008.

[7] S. H. Leung and C. F. So, "Gradient-based variable forgetting factor RLS algorithm in time-varying environments," *IEEE Trans. Signal Process.,* vol. 53, no. 8, pp. 3141-3150, Aug. 2005.

[8] C. Paleologu, J. Benesty, and S. Ciochina, "A robust variable forgetting factor recursive least-squares algorithm for system identification," *IEEE Signal Process. Lett.,* vol.15, pp. 597-600, 2008.

[9] B. Liao, Z. G. Zhang, and S. C. Chan, "A new robust Kalman filter-based subspace tracking algorithm in an impulsive noise environment," *IEEE Trans. Circuits Syst. II*, vol. 57, no. 9, pp. 740-744, Sep. 2010.

[10] S. Haykin, *Adaptive Filter Theory*. 4th ed., Prentice Hall, 2001.

[11] S. C. Chan, Y. J. Chu, and Z. G. Zhang, "A new regularized QRD recursive least M-estimate algorithm: performance analysis and applications," in *Proc. ICGCS 2010*, Shanghai, Jun. 21-23, 2010.

[12] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteritic of the LMS adaptive filter," in *Proc. IEEE*, vol. 64, pp. 1151-1162, 1976.

[13] T. Quatieri, *Discrete-time Speech Signal Processing*, Prentice Hall, 2001.

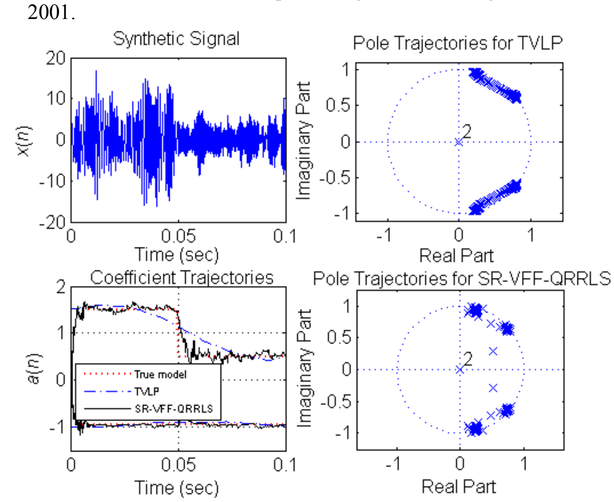

Fig. 1. Convergence performance for a "formant-like" synthetic signal (p=4).

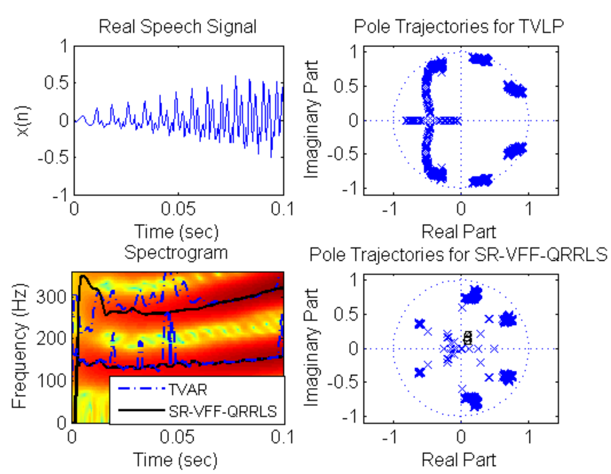

Fig. 2. Convergence performance for real speech signal (p=12).