NASA/CR-1998-206920
ICASE Report No. 98-11

A Novel Approach for Adaptive Signal Processing

Ya-Chin Chen
Imperial College, London, UK and ICASE

Jer-Nan Juang
NASA Langley Research Center

Institute for Computer Applications in Science and Engineering
NASA Langley Research Center
Hampton, VA

Operated by Universities Space Research Association

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center
under Contract NAS1-97046

February 1998

A NOVEL APPROACH FOR ADAPTIVE SIGNAL PROCESSING

YA-CHIN CHEN* AND JER-NAN JUANG†

Abstract. Adaptive linear predictors have been used extensively in practice in a wide variety of forms. In the main, their theoretical development is based upon the assumption of stationarity of the signals involved, particularly with respect to the second order statistics. On this basis, the well-known normal equations can be formulated. If high-order statistical stationarity is assumed, then the equivalent normal equations involve high-order signal moments. In either case, the cross moments (second or higher) are needed. This renders the adaptive prediction procedure non-blind. A novel procedure for blind adaptive prediction has been proposed and considerable implementation has been made in our contributions in the past year. The approach is based upon a suitable interpretation of blind equalization methods that satisfy the constant modulus property and offers significant deviations from the standard prediction methods. These blind adaptive algorithms are derived by formulating Lagrange equivalents from mechanisms of constrained optimization. In this report, other new update algorithms are derived from the fundamental concepts of advanced system identification to carry out the proposed blind adaptive prediction. The results of the work can be extended to a number of control-related problems, such as disturbance identification. The basic principles are outlined in this report and differences from other existing methods are discussed. The applications implemented are speech processing, such as coding and synthesis. Simulations are included to verify the novel modelling method.

Key words. blind equalization, blind adaptive prediction, equality and inequality constraints

Subject classification. Computer Science

1. Introduction. The progress in linear prediction (LP) techniques in recent years has collectively helped to advance many areas of research. Information supply services are now available in a wider range of application fields [1, 2, 3]. The LP technique was first used for speech analysis and synthesis by Itakura and Saito [4], and Atal and Schroeder [5], and has produced a large impact of speech research such as coding, recognition, enhancement and so on [6, 7, 8, 9].

The methodology of solutions in LP is essentially based on the minimization of squared error between real and estimated values. The conventional prediction can be carried out based on standard procedures only upon the assumption of a stationary discrete-time stochastic process as the model equations are based on second-order statistics assumed to be constant. In that case, the use of second order statistical properties of signals is needed. In practical applications, the observation interval can be quite short [10], (e.g. a speech sound or phoneme, assumed to be wide-sense stationary (WSS) for about 20 to 80 ms [11]). In order to compromise the assumption, these processes are considered to be "locally" WSS. However, the use of a short data record is not without its problems since any spectral estimate based upon a limited data set will exhibit a large variability due to inaccurate estimates of the second order statistical properties. Under these assumptions, the model so developed cannot respond to even moderately slowly varying signal attributes of

*Electrical & Electronic Engineering Department, Imperial College of Science, Technology, and Medicine, London SW7 2AZ, United Kingdom, (y-c.chen@ic.ac.uk). This research was supported by the National Aeronautics and Space Administration under NASA Contract Nos. NAS1-97046 and NAS1-19480 while the author was a Graduate Student in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23681-0001.

†Structural Dynamics Branch, NASA Langley Research Center, Hampton, Virginia 23681-0001, United States of America, (j.juang@larc.nasa.gov)

these statistical nature. So, a novel model which enables us to conquer these problems is essentially vital in practice.

To build up a predictor, ideally, we should make no a priori assumptions at all, and hence it is intended to be used for any sort of signal whether it is stationary or non-stationary, and of short data record or long data record. The crucial concept is to parameterize the model sample by sample, so that no training sequence is needed in the prediction procedure nor is the stationary statistical assumption necessary. In our early study [27], a procedure for blind adaptive prediction was proposed to overcome the difficulties inherent in the non-stationarity of the signals to be modelled. This approach is based upon a suitable interpretation of blind equalization methods that satisfy the constant modulus property and offers significant deviations from the standard prediction methods. We formulated the problem of signal prediction in the context of a linear model, in terms of keeping the error of representation at each estimate to within a predefined set of bounds rather than minimizing a functional on the assembly of a set of errors. By such means we can control the error per sample while retaining many of the original and desirable signal attributes.

The general way to solve equalization in blind equalizer is based upon stochastic gradient descent (SGD) algorithms. All of the existing algorithms, such as the Sato [12], Godard [13], and some SGD algorithms, are aimed at minimizing a cost function to converge to the equilibrium. Based on this as a starting point, the algorithms for implementing our blind adaptive prediction are derived from mechanisms of constrained optimization [27]. Other than the SGD algorithms, we derive three novel algorithms from the fundamental algebraic concepts of advanced system identification to carry out this proposed blind adaptive prediction. Also, the mechanisms of equality and inequality constraints are discussed. The results of the work can be extended to a number of control-related problems, such as disturbance identification.

1.1. Background. Blind equalizers are important devices in high data rate, bandlimited digital communication systems [14, 15]. Godard (1980) was the first one to propose a family of constant modulus blind equalization algorithms for use in two-dimensional digital communication system (e.g. M-ary phase shift keying). The most widely used blind equalization algorithm, the constant modulus algorithm (CMA), was so named by Treichler and Agee (1983) [16], independently of Godard's 1980 paper [13]. The CMA utilizes a novel quality assessment mechanism to solve adaptive filtering problems without the need for a "desired signal". Thus, the transmitted signals can be recovered when the channel is unknown [13, 16, 17]. This means that the transfer function, i.e., the convolution of the unknown channel and its equalizer, is the Kronecker delta with some delay and a constant gain. So, the equalizer, in fact, is the inverse of the unknown channel. A general diagram of blind equalization is shown in Fig 1.1. The simple way to achieve the implementation, as found in literature, is to define a cost function and approach the optimization by means of gradient descent method. Although CMA has been proposed, simulated, developed, and successfully applied, formal description of its behavioural properties is far from complete. The considerable problem is whether the CMA based on stochastic gradient descent (SGD) minimization of a specific mean cost function will always converge to its global rather than local minima [18, 19, 20]. It is very crucial because the undesirable local minima [12, 21] can cause insufficient removal of inter-symbol interference (ISI), which can produce errors in the reconstructed data stream at the receiver output.

Much work [21, 22] has been reported to indicate that these CMA families might ill-converge, i.e., converge to local minima, if they are not properly initialized, due to the local minima of the corresponding cost function. To improve the situation, the normalized constant modulus algorithm (NCMA) has been proposed [23, 24]. This type of algorithm has a stable operation for values of the step size in a larger range, in contrast to unnormalized algorithms, for which the step size range is often hard to determine. A larger

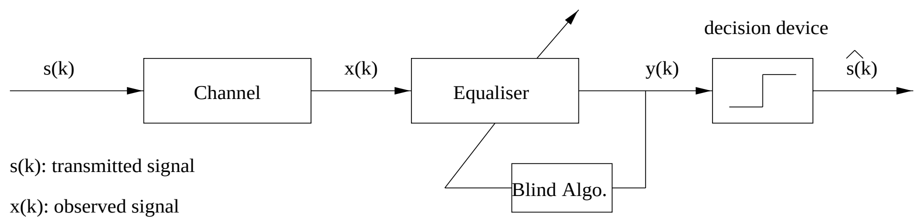

FIG. 1.1. *Blind Equalisation*

step size makes the convergence faster and often helps to avoid the undesirable local minima. However, despite the beneficial effects of the flexible step size range, the initialization problem remains. If the initial weights of the equalizer are chosen far away from the optimal point, equalization may not be achievable. In order to overcome this problem, an alternative CMA algorithm is proposed by Tanrikulu et al. in [25]. It is observed that the algorithm very often converges globally independent of the initialization. In the normalized family of algorithms, normalization affects the denominator of the gradient part of the update equations. This is in terms of a 2-norm of the input vector, and hence when the signal level is small or zero (e.g., during the silent period in speech), it can force the algorithm to diverge. An alternative algorithm through the Lagrange Neural Network is proposed to avoid such divergence. Complete work for the blind algorithm family has been collectively reexamined [26].

1.2. Outline of the report. The report consists of four main sections. The current section has introduced the goals of this report, namely the motivation fowards the development of a new predictive modelling based on blind equalization principles and its relevance to digital signal processing and control-related problems. Section 2 summarizes the new approach to adaptive prediction [26, 27, 28] along with the significant deviations between this new method and the conventional one. The problem is explored from the definition of blind equalization, from which adaptive prediction is derived and so are the updating algorithms. It appears speech coding can be done in a straightfoward manner by this method. In section 3, linear algebra used in advanced system identification is applied to tackle the same problem. Three algorithms are derived from equality and inequality constraints. The details of derivation are given and simulations on speech coding illustrate their performance. The final section, section 4, concludes with an assessment of the work from constrained optimization and analytical algebra. Some future directions currently being examined are also involved.

2. Blind adaptive prediction. The basic idea of blind equalization for binary input data, CMA, imposes the constraint of constant modulus at the output to recover the transmitted signal while the desired signal is unknown and so is the channel [13, 29]. In this section, this principle is applied to formulate the problem of signal prediction in the context of a linear model. In this novel prediction approach the error of representation at each estimate is kept to within a predefined set of bounds. This is in direct contrast to the procedures so far that minimize a functional on the assembly of a set of errors. This new approach offers significant advantages in comparison to the conventional methods. The error is kept under control at each sample, and hence no training sequence is needed in advance and no statistical assumptions and long term attributes of signals need be made.

The predictor coefficients are updated between each consecutive sample until an objective constraint is optimized. After that, the coefficients represent an interesting result, in which they resemble piecewise constants. The variation of the coefficients in a single decomposition is different for different update algorithms.

Moreover, it dominates the prediction performance.

The significant result from this modelling is the extraction of piecewise constant parameters, i.e., the tap weights of the equalizer. With piecewise constant parameters, the reconstruction of the signal is easily achievable. In this section, the experiment is carried out on a speech signal. The performance of the model indicates that this novel prediction method is reliable and robust to the practical problems.

2.1. Principle. For the general blind equalization problem, the problem has been expressed as a constrained optimization in the form [24]

$$\begin{aligned} & min\{\|\mathbf{W}(k+1)-\mathbf{W}(k)\|_2^2\} \\ & st. \mid \mathbf{X}^T(k)\mathbf{W}(k+1) \mid = 1 \end{aligned} \tag{2.1}$$

where the vector $\mathbf{W}(k)$ contains the coefficients of the equalizer $\mathbf{W}(k) = [\theta_0(k)\ \theta_1(k)\ \theta_2(k) \ldots \theta_n(k)]^T$, and $\mathbf{X}(k)$ is an observed sequence, $\mathbf{X}(k) = [x(k)\ x(k-1)\ x(k-2)\ \ldots.\ x(k-n)]^T$.

The constraint for our current purpose can be changed to

$$\mid \mathbf{X}^T(k)\mathbf{W}(k+1) \mid = \epsilon \tag{2.2}$$

where ϵ is the error signal, which is an arbitrary constant [27].
Equation(2.2) can be rewritten by the form

$$\mid x(k)\theta_0(k+1) + \mathbf{X}(k-1)^T\mathbf{\Theta}(k+1) \mid = \epsilon \tag{2.3}$$

where $\mathbf{\Theta}(k) = [\theta_1(k)\ \theta_2(k)\ \ldots.\ \theta_n(k)]^T$ and $\mathbf{X}(k-1) = [x(k-1)\ x(k-2)\ \ldots.\ x(k-n)]^T$. The output of the equalizer is then

$$y(k) = \mathbf{X}(k)^T\mathbf{W}(k) = \theta_0(k)x(k) + \mathbf{X}(k-1)^T\mathbf{\Theta}(k) \tag{2.4}$$

The first coefficient of the equalizer, θ_0, is fixed to be unity, and the ϵ is chosen to be much smaller than the average of the observed sequence for example. When equalization is achieved, equation(2.4) becomes

$$\pm\epsilon = x(k) + \mathbf{X}(k-1)^T\mathbf{\Theta}(k) \tag{2.5}$$

and under the assumption that ϵ is kept small, the estimated output value at the time instant k is obtained from the linear combination of the past samples as

$$\hat{x}(k) = -\mathbf{X}(k-1)^T\mathbf{\Theta}(k) \tag{2.6}$$

Assuming constant predictor weights between instants k and $k+1$, the adaptive prediction is then formulated as

$$\hat{x}(k+1) = -\mathbf{X}(k)^T\mathbf{\Theta}_{opt} \tag{2.7}$$

$\mathbf{\Theta}_{opt}$ is the coefficients of the adaptive predictor.

2.2. Blind Algorithms. Imposing the principal approach in the previous section to blind filtering algorithm, the prediction is carried out sample by sample. Because the variation in the constancy of the coefficients differs from the adapting algorithms and furthermore, the constancy is a measure of the accuracy achieved in prediction, different blind algorithms are derived and compared by simulating speech signal in this section. The two algorithms being discussed are as follows:

(1) Normalized Constant Modulus Algorithm (NCMA);

(2) Soft Constraint Satisfaction Algorithm (SCS).

2.2.1. Modified Normalized Constant Modulus Algorithm (MNCMA). In [23], 1992, Hilal and Duhamel have derived a normalized constant modulus algorithm (NCMA) by nulling the a posteriori error of constant modulus algorithm (CMA) at each iteration. These normalized algorithms can be also developed in an alternative fashion as suggested and derived by Papadias and Slock [15, 24, 30, 31]. We apply this principle to derive the modified normalized constant modulus algorithm to implement our prediction. The problem now can be reformulated as

$$\begin{aligned} &min\{\|\mathbf{\Theta}(k+1) - \mathbf{\Theta}(k)\|_2^2\} \\ &st. \mid x(k) + \mathbf{X}(k-1)^T\mathbf{\Theta}(k+1) \mid = \epsilon \end{aligned} \tag{2.8}$$

The solution yields the update algorithm to be

$$\mathbf{\Theta}(k+1) = \mathbf{\Theta}(k) + \frac{\mu}{\|\mathbf{X}(k-1)\|_2^2}(\epsilon \times sgn(y(k)) - y(k))\mathbf{X}(k-1) \tag{2.9}$$

wherer μ is the step size for the weights, $\|.\|$ denotes 2-norm, and k is time index. This normalized algorithm has a stable adaptation for a step size in a larger range, and hence it helps avoiding the undesirable local minima and, i.e., ill-convergence [24].

In this simulation, we use the speech signal shown in figure 2.1. The signal is normalized with a maximum amplitude of 1. The constant ϵ is set to be 0.1, the length of the equalizer is chosen as 11, and the step size, μ, is 0.001. The unit-center tap anchoring strategy [32] used in general blind equalization is not suitable for this prediction case. Indeed, the first coefficient of the predictor is actually fixed to 1, as indicated in equation(2.5), and hence the taps are initialized to $\mathbf{\Theta}(0) = [-1\ 0\ ....\ 0]^T$. Figure 2.2 shows the error signal obtained by this procedure, which is bounded as expected. Prediction is carried out in a sample by sample fashion according to equation(2.7) and the result is shown in figure 2.3. The corresponding coefficients are found in figure 2.4. It is seen from this figure that the coefficients exhibit an approximate piecewise constant behavior. This may be a useful means of carrying out speech coding. With the piecewise constant parameters, the reconstruction of speech signal is easily achievable.

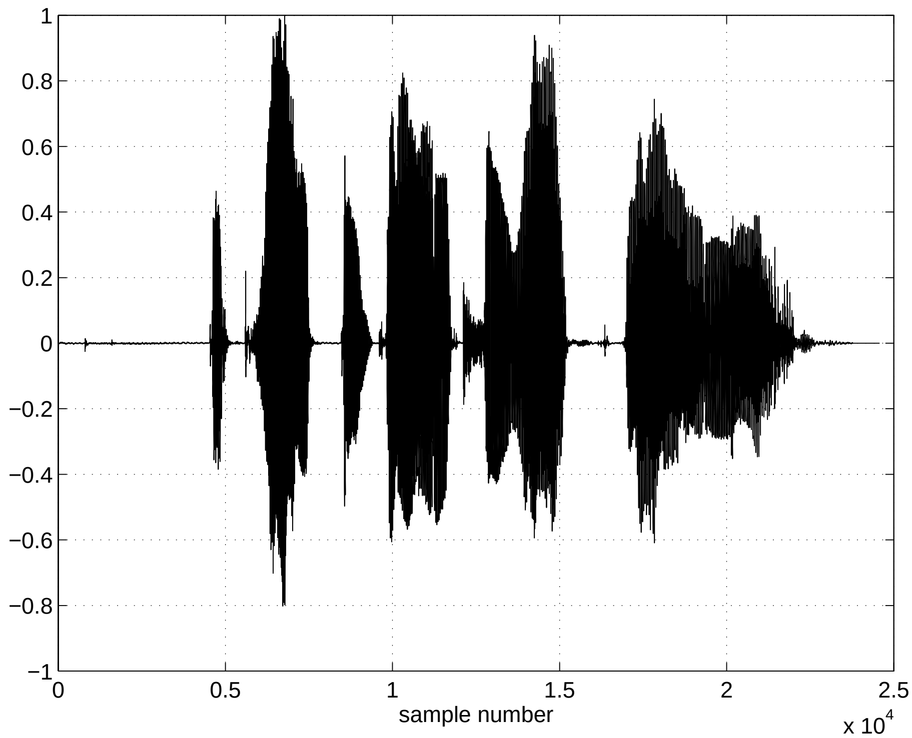

FIG. 2.1. *The original speech signal "the pipe began to rust while new"*

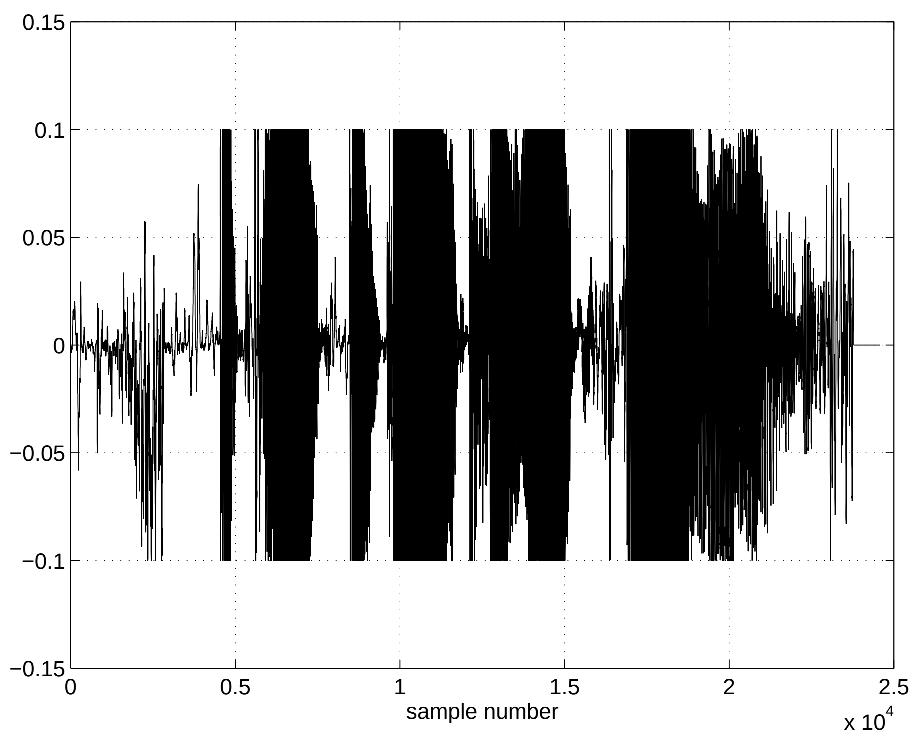

FIG. 2.2. *The error signal of prediction from MNCMA*

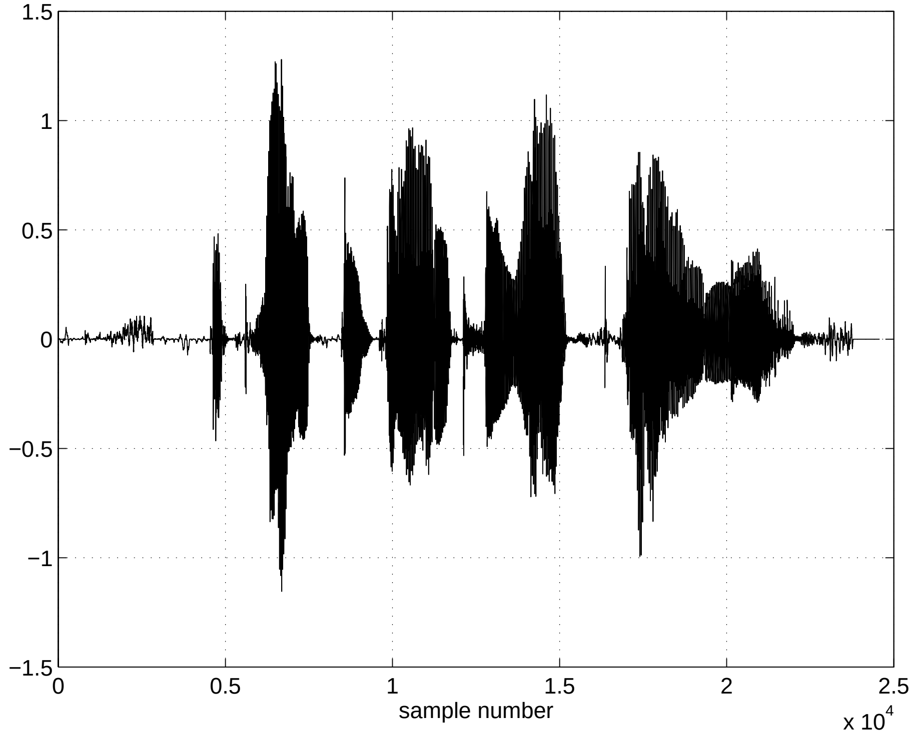

FIG. 2.3. *The predicted signal from MNCMA*

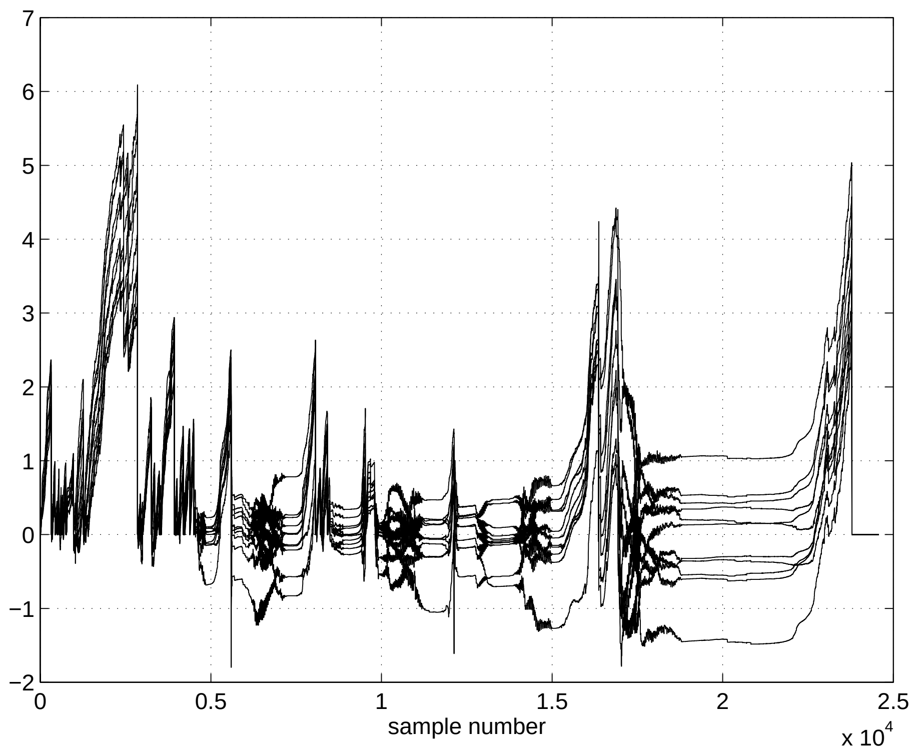

FIG. 2.4. *The coefficients of the predictor from MNCMA*

2.2.2. Soft Constraint Satisfaction Algorithm (SCSA). In [28], the optimization problem in section 2.2.1 is modified by imposing a quadratic constraint

$$\begin{aligned} min\ \{\|\boldsymbol{\Theta}(k+1) - \boldsymbol{\Theta}(k)\|_2^2\} \\ st.\ (x(k) + \mathbf{X}(k-1)^T \boldsymbol{\Theta}(k+1))^2 = \epsilon^2 \end{aligned} \tag{2.10}$$

and hence the Lagrange equivalent can be formed as

$$L = \|\boldsymbol{\Theta}(k+1) - \boldsymbol{\Theta}(k)\|_2^2 + \lambda((x(k) + \mathbf{X}(k-1)^T \boldsymbol{\Theta}(k+1))^2 - \epsilon^2) \tag{2.11}$$

By the same derivation in [25], the SCS algorithm takes the update form [33]

$$\begin{aligned} \boldsymbol{\Theta}(k+1) = \boldsymbol{\Theta}(k) + \frac{1}{\|\mathbf{X}(k-1)\|_2^2}(\varphi[y(k)] - y(k))\mathbf{X}(k-1) \\ \varphi[y(k)] = \frac{y(k)}{1 - \mu(1 - \frac{|y(k)|}{\epsilon})} \end{aligned} \tag{2.12}$$

For our simulation work we used the speech signal in figure 2.1. The constant ϵ is also set to be 0.1. The length of the equalizer is set to 11, and the step size is 0.001. We initialize the adaptation with $\boldsymbol{\Theta}(0) = [-1\ 0\ ....\ 0]^T$. Figure 2.5 shows the error signal bounded as the constant modulus constraint satisfied, and the predicted result is in figure 2.6. However, the prediction coefficients in figure 2.7 show a larger deviation than the corresponding result from MNCMA. Nevertheless, the speed of convergence is faster in that in about 100 iterations we reach constraint optimization.

2.3. Summary. From these simulations, the coefficients of the equalizer are approximately piecewise constant. It is evident that different algorithms produce different behavior with respect to the expected constancy of the coefficients. From equations(2.6) and (2.7), the expected constancy in the coefficients is a measure of the effectiveness achieved in prediction. The predicted results in figure 2.3 and figure 2.6 illustrate this clearly. If computational cost is to be considered, then the SCS algorithm performs better. However, the SCS exhibits the larger variation in the coefficients.

Another interesting result is the relation between the satisfaction of constraints and the constancy of the predictor coefficients. It appears that the predictor coefficients can respond fast enough to satisfy the constraint absolutely, which can be referred to as the equality constraint. However, meanwhile, the filter tracks noise to cause the large deviation of the coefficients. Therefore, the coefficients in figure 2.4 behave more constant corresponding to a bounded constraint, the inequality constraint, in figure 2.2.

3. Novel algorithms. A brief review of blind adaptive prediction has been given and the results produced by constrained optimization have been discussed in section 2. In this section, we will tackle the same problem by employing the fundamental algebra used in advanced system identification instead of the proposed solution in [27, 28]. Moreover, according to the summary in section 2.3, the requirements, the satisfaction of constraints and the piecewise constancy of the coefficients, relate to equality and inequality constraints, so we put both of them into effect and derive different algorithms. These novel algorithms need no iterations between each estimate and therefore, they are feasible in real-time computation. The fundamental linear algebra is outlined and the derivation of algorithms are also detailed. Experiments are simulated with applications to speech signal.

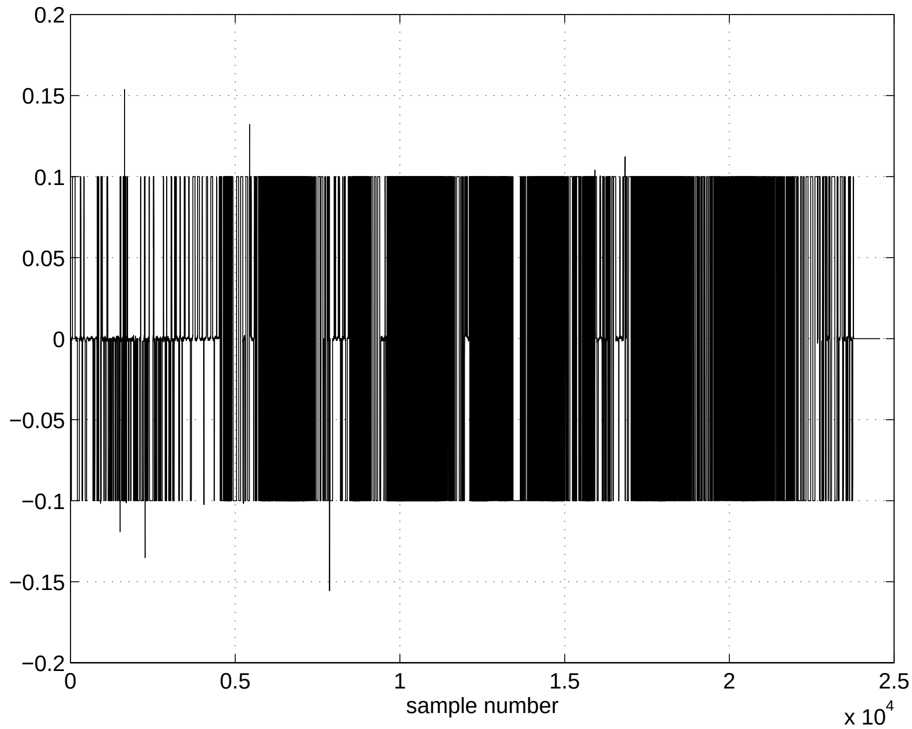

FIG. 2.5. *The error signal of prediction from SCS*

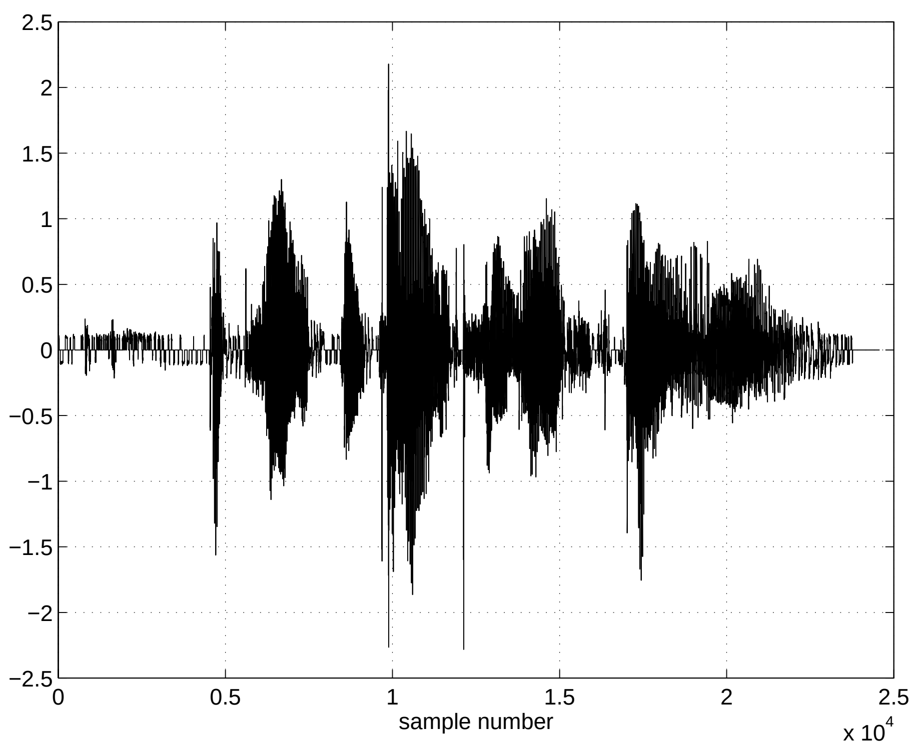

FIG. 2.6. *The predicted signal from SCS*

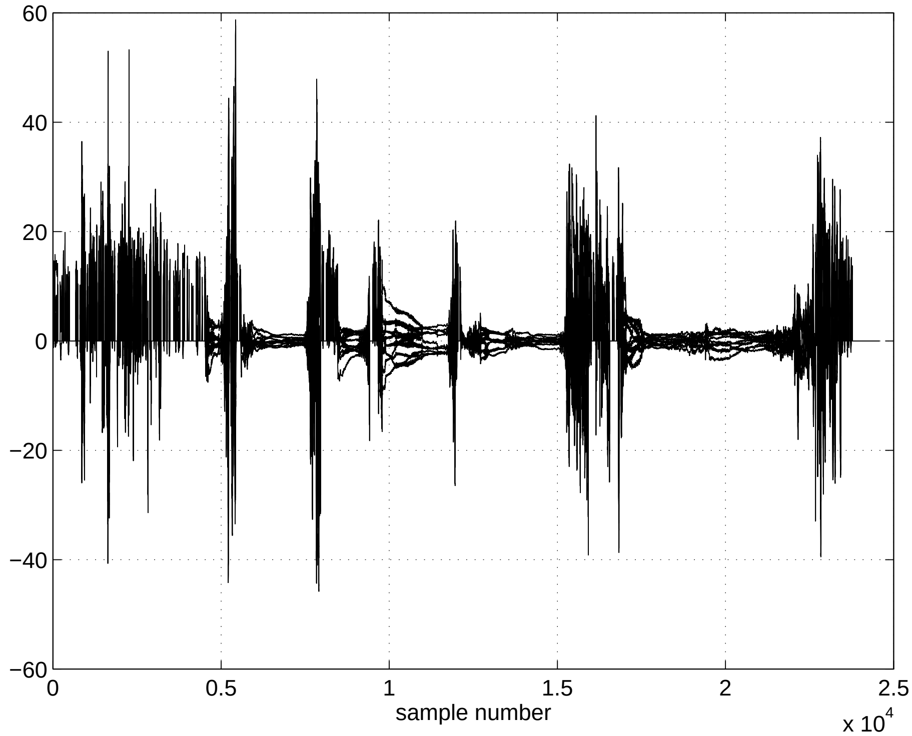

FIG. 2.7. *The coefficients of the predictor from SCS*

3.1. Equality constraint. In this section, an equality constraint is imposed to carry out the blind adaptive prediction. By means of equality constraint, the error signal is fully satisfied to the predefined value. Therefore, the ignored error in prediction, equation(2.7), is well controlled. However, from the results shown in the previous section, the fully satisfied constraint causes the variance of the coefficients to be larger. To conquer this trade-off problem, a preprocess for coefficient smoothing is employed to enhance the constancy of the coefficients.

3.1.1. Non-iterative Method. The first novel algorithm is conducted by the fundamental algebra and the method of least squares. In this derivation, we simply solve the problem to obtain the analytical solution, so there is no need to approach a desired objective iteratively, which the blind algorithms in previous section do.

Problem definition and algorithm derivation. The problem of blind adaptive prediction is defined by minimizing the disturbance of parameters of the predictor subject to a constant modulus constraint. The formulation of the output of the equalizer is now rewritten

$$y(k) = x(k) + \mathbf{X}^T(k-1)\mathbf{\Theta}(k) \tag{3.1}$$

Instead of utilizing constrained optimization methods, from equation(2.8), we may solve equation(3.1) using linear algebra [12] to obtain a *general solution* for the coefficients of the predictor

$$\mathbf{\Theta}(k+1) = [\mathbf{X}^T(k-1)]^{\dagger}(\epsilon - x(k)) + \mathbf{X}_0(k-1)\mathbf{C} \tag{3.2}$$

where $\mathbf{X}_0(k-1)$ contains n column basis vectors of the null space of the column vector $\mathbf{X}(k-1)$, and $[\mathbf{X}^T(k-1)]^\dagger$ is the *pseudoinverse* of the row vector $\mathbf{X}^T(k-1)$, which can be computed by

$$[\mathbf{X}^T(k-1)]^\dagger = [\mathbf{X}^T(k-1)\mathbf{X}(k-1)]^{-1}\mathbf{X}(k-1) \tag{3.3}$$

if $\mathbf{X}^T(k-1)$ is not a null row vector. The null-space basis vectors in $\mathbf{X}_0(k-1)$ may be generated by using the singular value decomposition [12]. Note that the matrix $\tilde{\mathbf{X}}_0(k-1)$ computed by

$$\tilde{\mathbf{X}}_0(k-1) = \mathbf{I} - [\mathbf{X}^T(k-1)]^\dagger \mathbf{X}^T(k-1) \tag{3.4}$$

is in the same space as $\mathbf{X}_0(k-1)$ but has a different size of dimension. The matrix $\tilde{\mathbf{X}}_0(k-1)$ may be used in equation(3.2) to replace $\mathbf{X}_0(k-1)$. However, special caution must be taken into account that $\tilde{\mathbf{X}}_0(k-1)$ is ill-conditioned with rank short at least by one.

After the constraint, equation(3.1), is fully satisfied by the solution in equation(3.2), the cost function in equation(2.8) is taken into consideration. The minimization of the disturbance of the coefficients has been defined as

$$\Delta\Theta = \|\mathbf{\Theta}(k+1) - \mathbf{\Theta}(k)\|_2^2 \tag{3.5}$$

Substituting the solution from equation(3.2) into equation(3.5) yields the disturbance equation

$$\Delta\Theta = \|\mathbf{X}_0(k-1)\mathbf{C} + \{[\mathbf{X}^T(k-1)]^\dagger(\epsilon - x(k)) - \mathbf{\Theta}(k)\}\|_2^2 \tag{3.6}$$

Thus, the disturbance can be minimized by solving for the *least-squares solution* [12, 34] of the arbitrary vector $\mathbf{C}$ from equation(3.6) as

$$\mathbf{C} = -\mathbf{X}_0^\dagger(k-1)\{[\mathbf{X}^T(k-1)]^\dagger(\epsilon - x(k)) - \mathbf{\Theta}(k)\} = \mathbf{X}_0^\dagger(k-1)\mathbf{\Theta}(k) \tag{3.7}$$

The second equality holds because $\mathbf{X}(k-1)$ is orthogonal to $\mathbf{X}_0(k-1)$, i.e.,

$$\mathbf{X}^T(k-1)\mathbf{X}_0(k-1) = 0 \Longrightarrow \mathbf{X}_0^\dagger(k-1)[\mathbf{X}^T(k-1)]^\dagger = 0$$

The coefficients obtained from equation(3.2) satisfy the constraint equation(3.1) exactly. No iteration is necessary between each estimate. As the arbitrary vector $\mathbf{C}$ is the least-squares solution, the cost function in equation(3.5) is globally minimized.

Simulation. The speech signal shown in figure 3.1 is part of the whole sequence in figure 2.1 from time index $k = 16851, ..., 23500$. The length of the equalizer is taken to be 11, and the error signal is set to 0. According to equation(2.5), the first coefficient of the equalizer, θ_0, is fixed to be unity to carry out the prediction, equation(2.7). From figure 3.2, it shows the constraint is fully satisfied as the error signal is predefined by 0. Because the adaptation tracks each estimate to fulfill the constraint such that the coefficients of the predictor vary in a large deviation, which can be found in figure 3.3. Hence, if the median of these coefficients is taken to reconstruct the original speech signal, the result, figure 3.4, turns out in a slightly different scale from the original one. From these results, once again, it illustrates that the

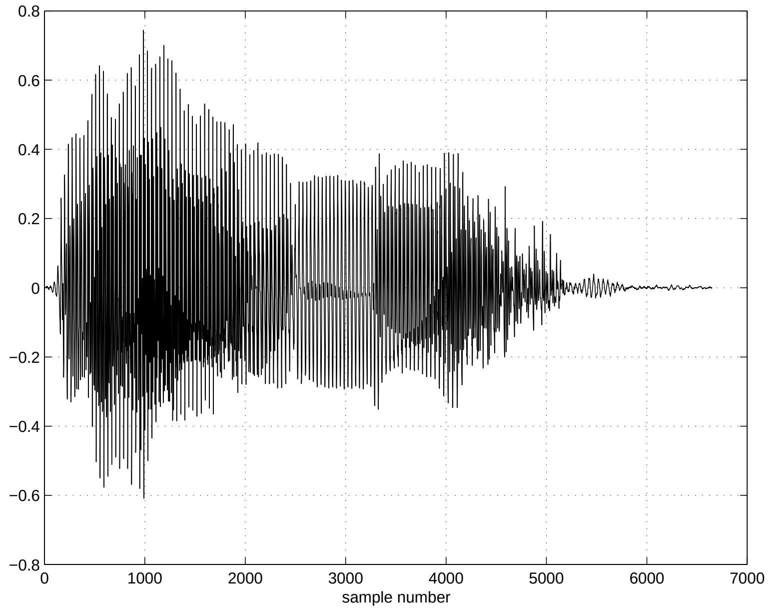

FIG. 3.1. *Original speech signal*

satisfaction of the constraint and the demand of piecewise constant coefficients are mutually influenced. If the requirement of satisfaction of the constraint comes first, the adaptation tracks all the signals including noises at the same time. On the contrary, if the piecewise constant coefficients are required by the prediction in equation(2.7), the adjustment of the constraint should be reconsidered. From the practical point of view in coding, compression, and synthesis, the piecewise constancy of the predictor coefficients are more crucial than the satisfaction of the constraint.

3.1.2. 2-step Method (Enhancement). The enhanced algorithm is aimed at smoothing the coefficients to improve piecewise constancy of the coefficients. A preprocess is added using a conventional recursive least-squares technique to compute a set of smoothing coefficients. The same update procedure from the non-iterative method is then used to fine tune the coefficients to satisfy the constraint equation.

Problem definition and algorithm derivation. In coefficient smoothing, the method is exploited directly from what has been widely used in system identification [34]. Indeed, a set of smoothing coefficients can be determined by the minimization defined as

$$min \parallel \sum_{i=1}^{k} \lambda^{k-i}[\mathbf{X}^T(i-1)\mathbf{\Theta}_s(k) - x(i)]^2 \parallel \tag{3.8}$$

where $\mathbf{\Theta}_s$ is the smoothing coefficients and λ is a parameter such that $0 < \lambda \leq 1$. The parameter λ is called the forgetting factor that is a time-varying weighting of the data. The most recent data is given unit weight, but data that is n time steps old is weighted by λ^n. The method is commonly called exponential forgetting.

The parameter $\mathbf{\Theta}_s(k)$ which minimizes the equation(3.8) is given recursively by

$$\mathbf{G}(k) = \frac{\mathbf{X}^T(k-1)\mathbf{P}(k-1)}{\lambda + \mathbf{X}^T(k-1)\mathbf{P}(k-1)\mathbf{X}(k-1)}$$
$$\hat{x}(k) = \mathbf{X}^T(k-1)\mathbf{\Theta}_s(k-1)$$

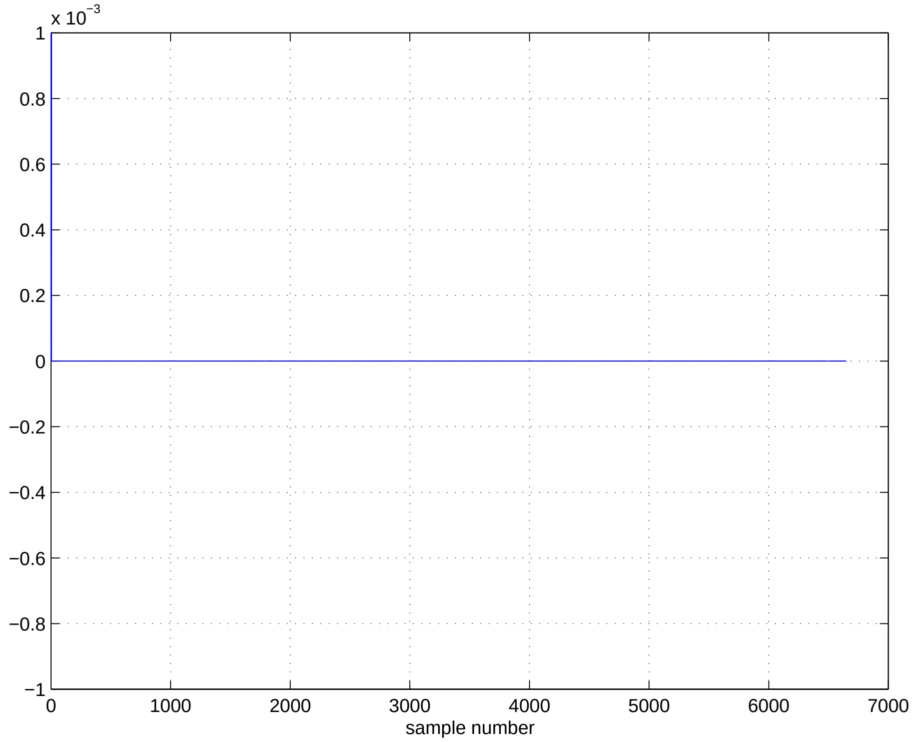

Fig. 3.2. *Error signal (Constraint)*

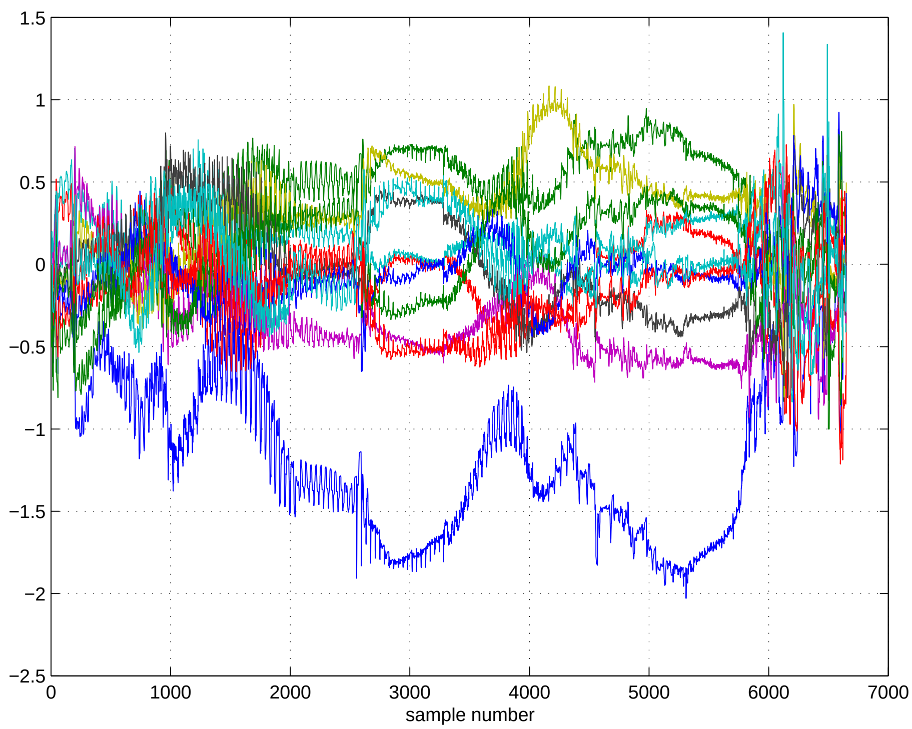

Fig. 3.3. *Coefficients of the predictor*

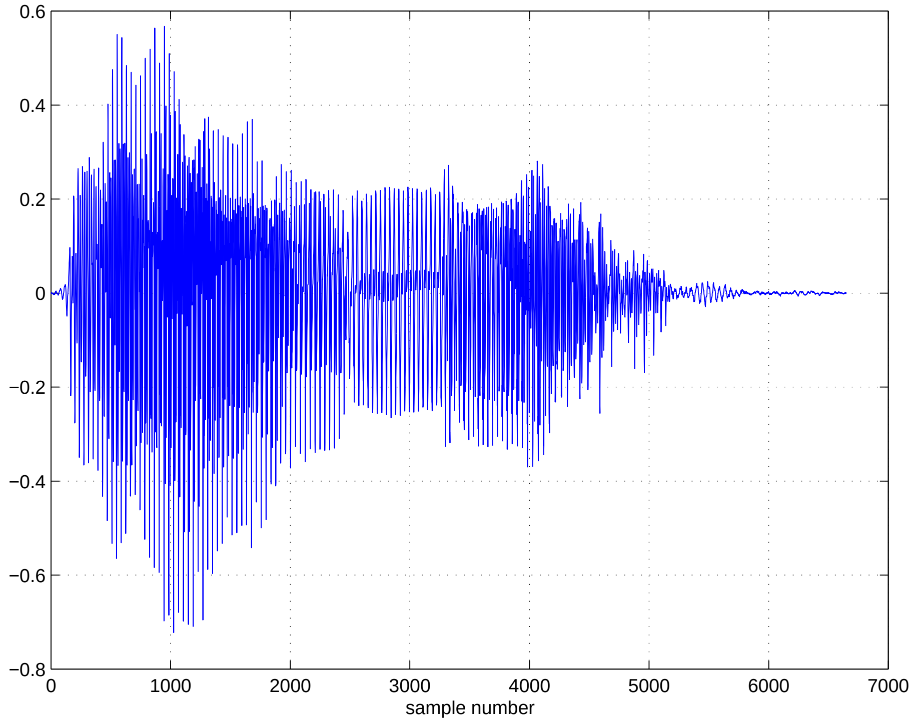

FIG. 3.4. *Open-loop prediction by median coefficients*

$$\mathbf{P}(k) = \frac{\mathbf{P}(k-1)\left[\mathbf{I} - \mathbf{X}^T(k)\mathbf{G}(k)\right]}{\lambda}$$
$$\mathbf{\Theta}_s(k) = \mathbf{\Theta}_s(k-1) + \left[x(k) - \hat{x}(k)\right]\mathbf{G}(k)$$

where $\mathbf{G}(k)$ is the update gain determined by the matrix $\mathbf{P}(k-1)$, the vector $\mathbf{X}(k-1)$, and the scalar λ. The initial values of $\mathbf{P}(0)$ and $\mathbf{\Theta}_s(0)$ can be arbitrarily assigned. Conventionally, $\mathbf{P}(0)$ and $\mathbf{\Theta}_s(0)$ are assigned as $d\mathbf{I}_n$ and $\mathbf{0}_{n\times 1}$, respectively where d is a large positive number, $\mathbf{I}_n$ is an identity matrix of dimension $n \times n$, and $\mathbf{0}_{n\times 1}$ is a zero matrix of dimension $n \times 1$.

Rather than minimizing the disturbance of two consecutive coefficients, the disturbance is taken with respect to the smoothing coefficients.

$$\Delta\Theta = \|\mathbf{\Theta}(k+1) - \mathbf{\Theta}_s(k)\|_2^2 \tag{3.9}$$

The minimization steps are the same as that presented in section 3.1.1. The predictor coefficients can then be obtained after the complete 2-step adaptation.

Simulation. In order to compare the enhanced algorithm to the previous non-iterative one, the same experiment in section 3.1.1 is used again. Let the forgetting factor set to $\lambda = 0.999$ and the intial value of $\mathbf{P}(0) = 1000\mathbf{I}_{11}$. The constraints are examined with respect to the smoothing coefficients and the predictor coefficients, as shown in figure 3.5 and figure 3.6, respectively. The constraint from the smoothing coefficients does not satisfy the predefined value while the predictor coefficients does. However, from figure 3.7 and figure 3.8, it can be seen that the predictor coefficients vibrate more rapidly than the smoothing coefficients. This is because predictor coefficients have to track noise to fulfill the constraint at each estimate. By taking the medians of these two sets of coefficients, the speech signals are reconstructed in figure 3.9 and figure 3.10, respectively.

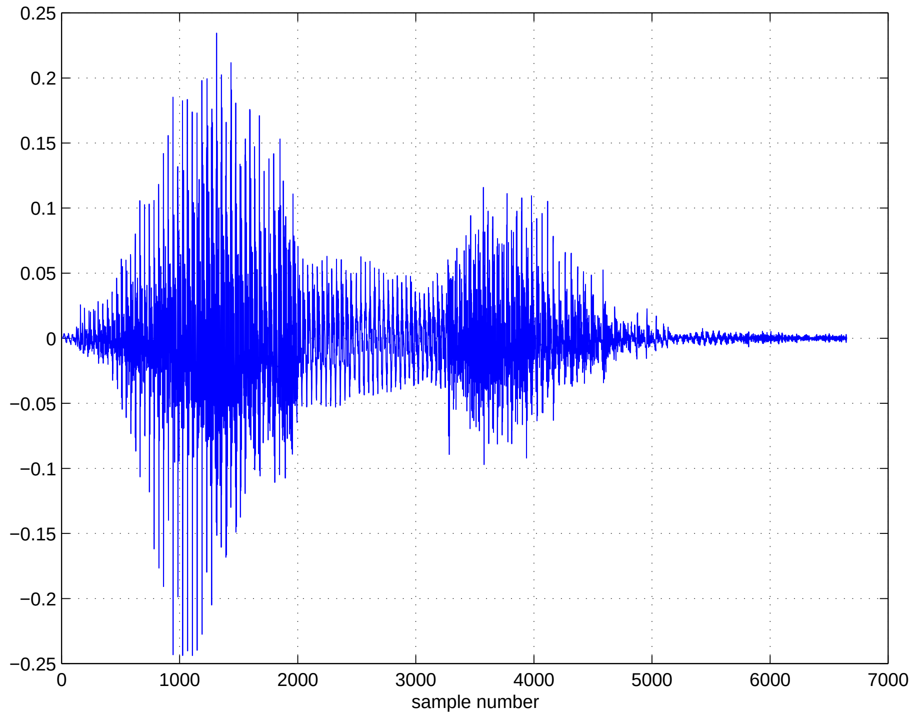

FIG. 3.5. *Error signal (Constraint) from smoothing-only step*

Another simulation is done by increasing the length of the equalizer to be 16. The interesting results shown in figure 3.11 and figure 3.12 indicate that the coefficients can approximate to constants overall without the scale jump due to the segment of words. This may reduce the memory for signal storage.

3.1.3. Summary. From the above results, it is clear that the algorithms derived in this section can achieve blind adaptive prediction without iterations. This makes real-time computation feasible. The enhanced algorithm makes a progress in the piecewise constancy of the predictor coefficients. This improvement enhances the reconstructed signal more reliable. From the point of view of signal compression, the low rate storage can be achieved when the filter length is taken long enough, which can be referred to the case, $L = 16$.

3.2. Inequality constraint. Different from the equality constrained methods, this method is based on an inequality constraint, which makes the error being bounded within the predefined value rather than exactly equivalent to it. When the disturbance below the predefined bound, which is mostly due to noise, is measured, the coefficients stop tracking the signal, and retain the same as the previous ones. It will be seen that the coefficients behave the expected constancy as the 2-step algorithm does. However, it requires no preprocess for coefficient smoothing.

3.2.1. Problem definition and algorithm derivation. The problem is now in the form of the following equation after imposing an inequality constraint

$$\begin{aligned} & min\{\|\mathbf{\Theta}(k+1) - \mathbf{\Theta}(k)\|_2^2\} \\ & st. \mid x(k) + \mathbf{X}^T(k-1)\mathbf{\Theta}(k+1) \mid \leq \epsilon \end{aligned} \tag{3.10}$$

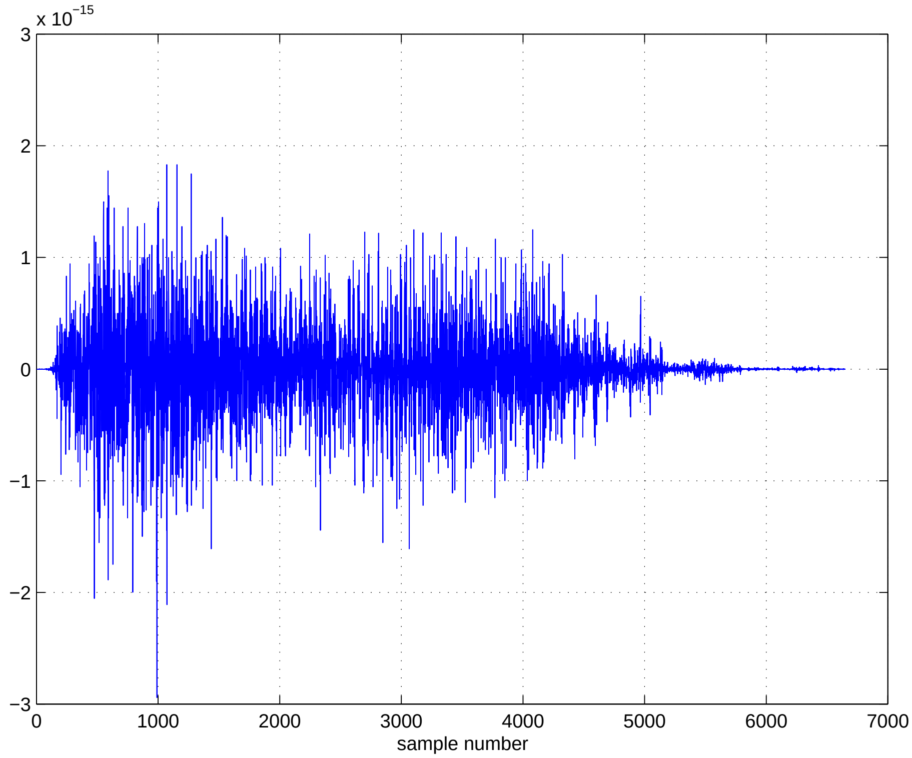

FIG. 3.6. *Error signal (Constraint) from smoothing+non-iterative*

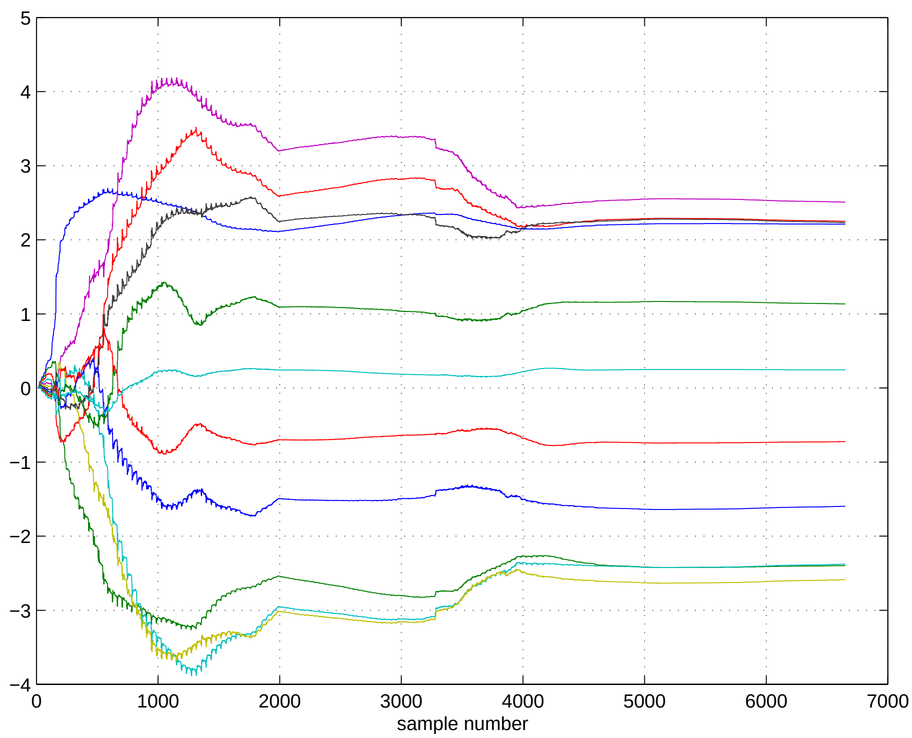

FIG. 3.7. *Coefficients of the predictor from smoothing-only step*

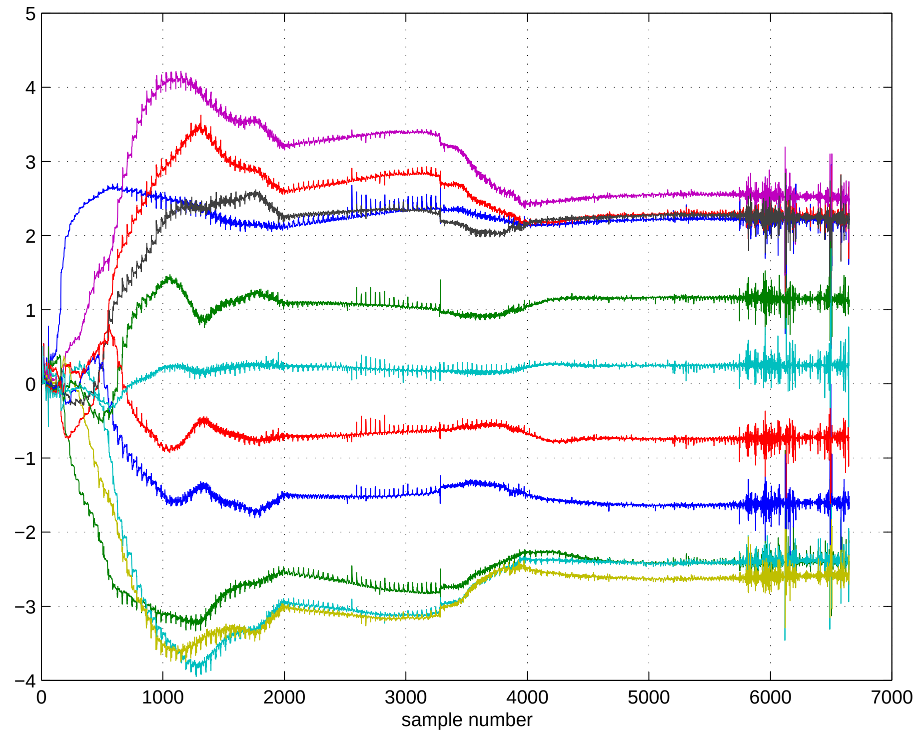

FIG. 3.8. *Coefficients of the predictor from smoothing+non-iterative*

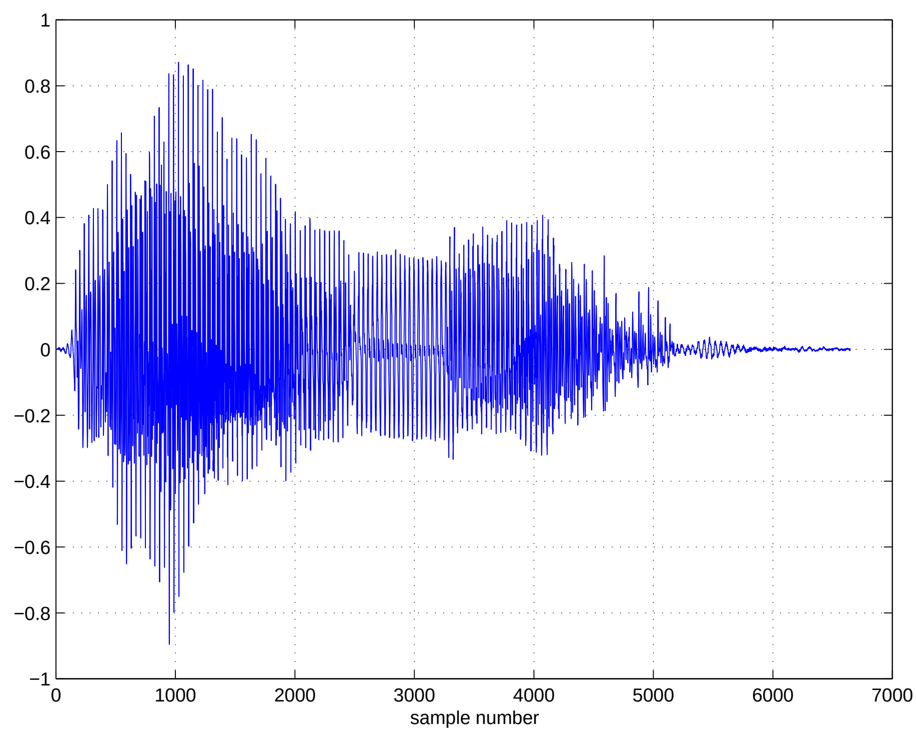

FIG. 3.9. *Open-loop prediction by median coefficients from smoothing-only step*

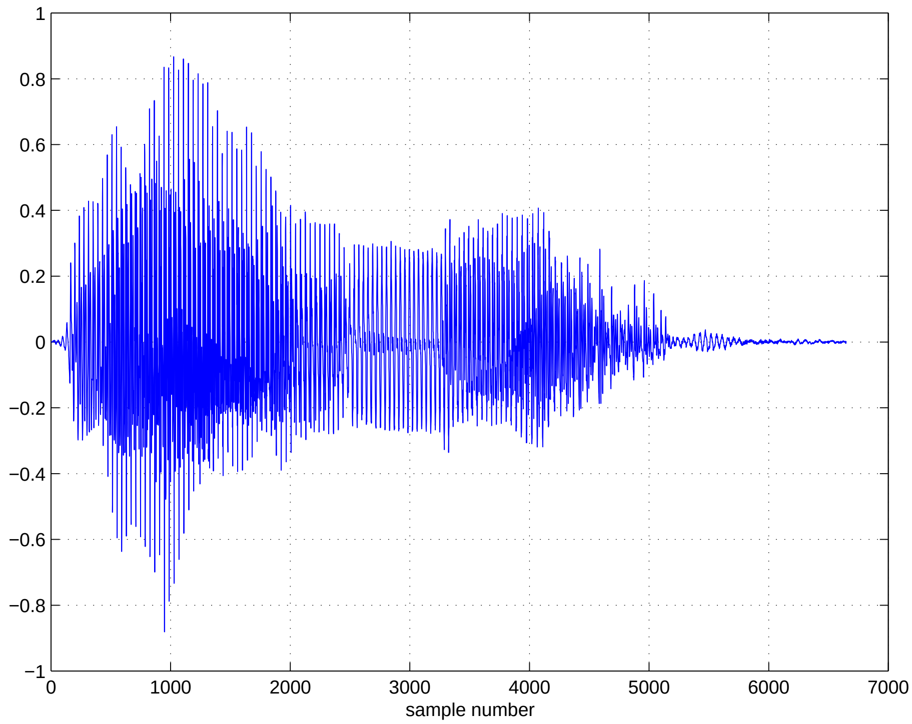

FIG. 3.10. *Open-loop prediction by median coefficients from smoothing+non-iterative*

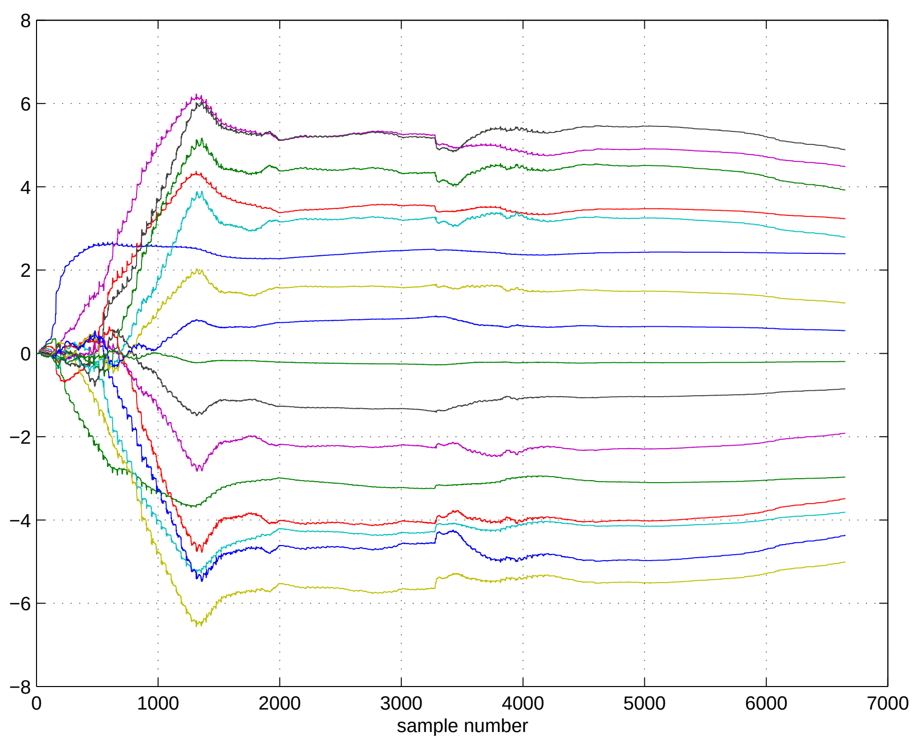

FIG. 3.11. *Coefficients of the predictor from smoothing-only step while L=16*

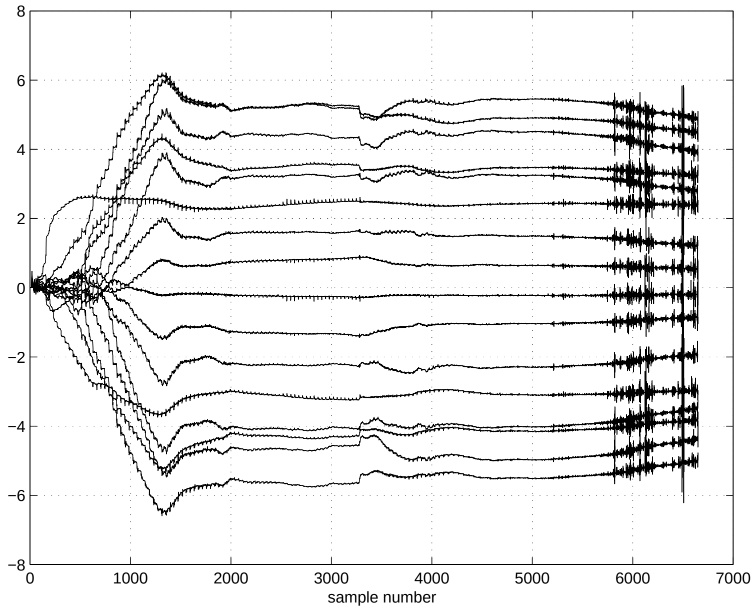

FIG. 3.12. *Coefficients of the predictor from smoothing+non-iterative while L=16*

The computation can be briefly summarized

$$If \mid x(k) + \mathbf{X}^T(k-1)\mathbf{\Theta}(k+1) \mid \leq \epsilon$$

$$\mathbf{\Theta}(k+1) = \mathbf{\Theta}(k).$$

$$else$$

$$\mathbf{\Theta}(k+1) = \mathbf{\Theta}(k) + \Delta\Theta(k)$$

When the constraint is not bounded within the expected value, the error signal can be presented as

$$\mid x(k) + \mathbf{X}^T(k-1)\mathbf{\Theta}(k+1) \mid > \epsilon \tag{3.11}$$

Assume that

$$x(k) + \mathbf{X}^T(k-1)\mathbf{\Theta}(k+1) = \epsilon_r \tag{3.12}$$

where ϵ_r is the real value of the error signal. The difference between the error signal and the desired constraint bound can be computed by

$$\mid \epsilon_r \mid - \epsilon = \Delta\epsilon \tag{3.13}$$

and thus

$$\mid x(k) + \mathbf{X}^T(k-1)\mathbf{\Theta}(k+1) \mid - \Delta\epsilon = \epsilon \tag{3.14}$$

The update algorithm for the coefficients can then be obtained

$$\mathbf{\Theta}(k+1) = [\mathbf{X}^T(k-1)]^{\dagger}[sgn(x(k)) \times \epsilon - x(k)] + \mathbf{X}_0(k-1)\mathbf{C} \tag{3.15}$$

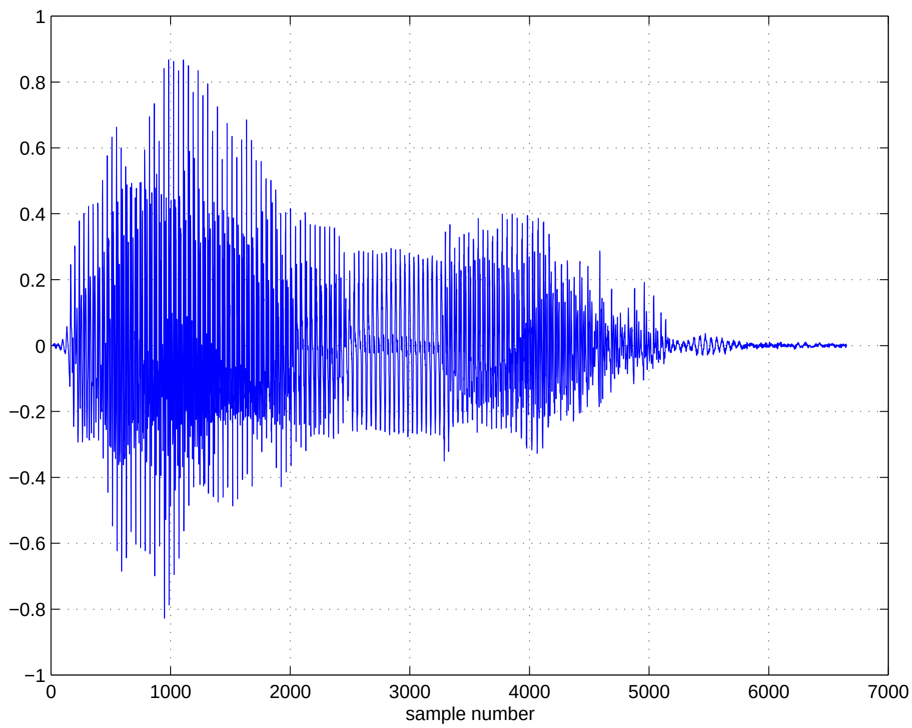

FIG. 3.13. *Open-loop prediction by median coefficients from smoothing-only step while L=16*

where $\mathbf{X}_0(k-1)$ is the null space of the vector $\mathbf{X}(k-1)$. The coefficient vector $\mathbf{C}$ is obtained using the same way as presented in equation(3.7) for equality constraint.

3.2.2. Simulation. The same speech signal shown in figure 3.1 is simulated again. The length of the equalizer is set to be 11, and the bound of the constraint is predefined as 0.1. In figure 3.15, the result indicates the error is bounded within 0.1 rather than exactly equivalent to 0.1. The coefficients shown in figure 3.16 behave piecewise constant clearly except the transient part. In this case, the median values of the coefficients can represent precisely the overall coefficients. The reconstructed signal by the median values of the coefficients, figure 3.17, keeps the features of the original signal quite well.

3.2.3. Summary. In the inequality constrained method, the constraint condition and the constancy of the desired coefficients can be treated synchronously. As long as the error signal is bounded within the constraint, the update of the coefficients is not necessary. In this case, we can simply leave the coefficients unchanged to satisfy the demand for piecewise constant coefficients. As a result, the preprocess of coefficient smoothing is no longer required in the inequality constrained method, but the constancy of the coefficients is better than that from the equality constrained algorithm. This is because the coefficients only track the signal beyond the predefined bounds. The disturbance below some bounds usually results from the existence of irrelevant noise. With a reasonable bound for the inequality constraint, the variation of the coefficients can be considerably reduced. This can be applied to noise reduction, or signal enhancement.

4. Conclusions and Future work. Conventional linear predictive coding has limitations because their theoretical development is based on the assumption of stationarity and thus the corresonding adaptive prediction is non-blind. In practice, prediction on a sample-by-sample basis is desirable. The sample-by-sample concept is especially useful for short data records, since the conventional linear predictive coding

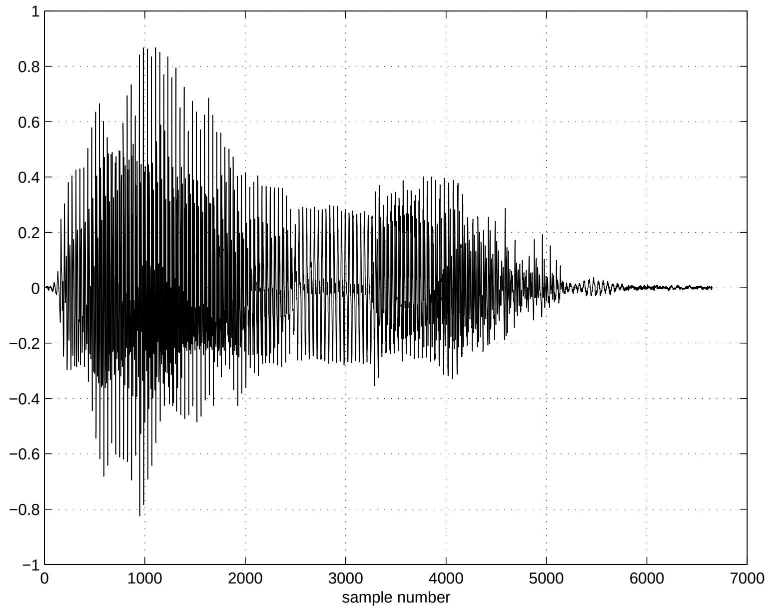

FIG. 3.14. *Open-loop prediction by median coefficients from smoothing+non-iterative while L=16*

exhibits a large variability due to a lack of averaging for a short data record.

In section 2, the blind adaptive prediction is reviewed and some results are also given to illustrate its performance. The implementation is completed using constrained optimization mechanisms. The prediction processes are driven by error signal and the corresponding predictor coefficients. The extraction of ultimately piecewise constant parameters is a novel means of modelling signals.

In comparison to the algorithms based on constrained optimization mechanisms, the novel algorithms proposed in section 3 can not only achieve the same performance but also speed the computation because of no need of iteration. This makes real-time computation feasible. These algorithms are derived from fundamental linear algebra, and some concepts of advanced system identification. The results of the work can be extended to a number of control-related problems, such as disturbance identification.

Other work of interest being examined includes stability study for closed-loop prediction. Also, the mechanism of blind adaptive multi-stage prediction [28] is developing.

5. Acknowledgments. This report was written as a result of a five-week visit of Ms. Ya-Chin Chen at ICASE under the supervision and guidance of Dr. Jer-Nan Juang. The first draft of this report was completed on September 12, 1997. Ms. Ya-Chin Chen would like to take this opportunity to express her appreciation of the ICASE support.

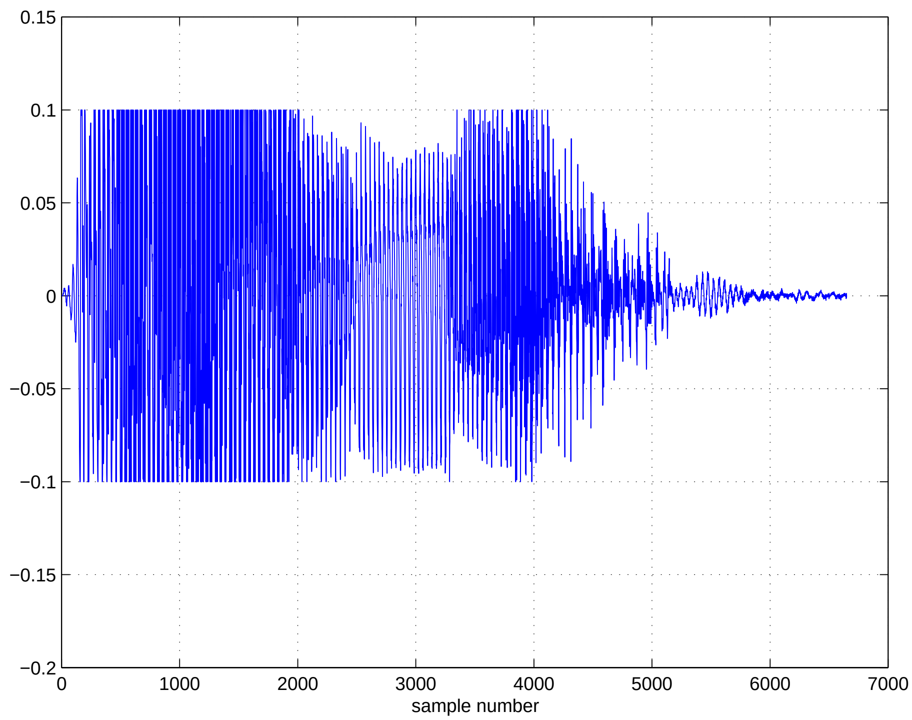

FIG. 3.15. *Error signal (Constraint)*

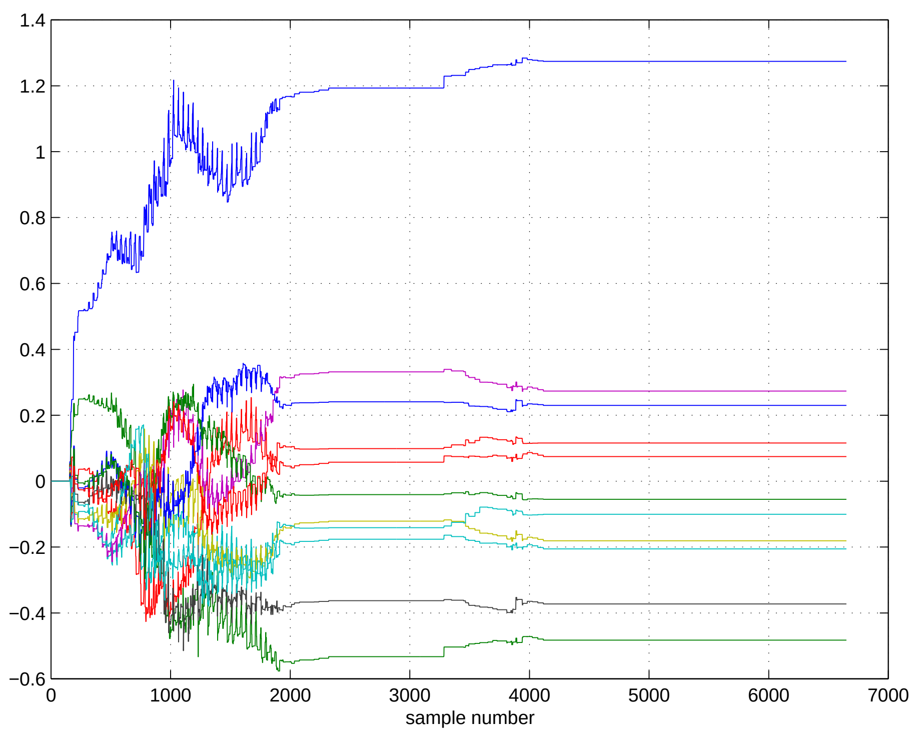

FIG. 3.16. *Coefficients of the predictor*

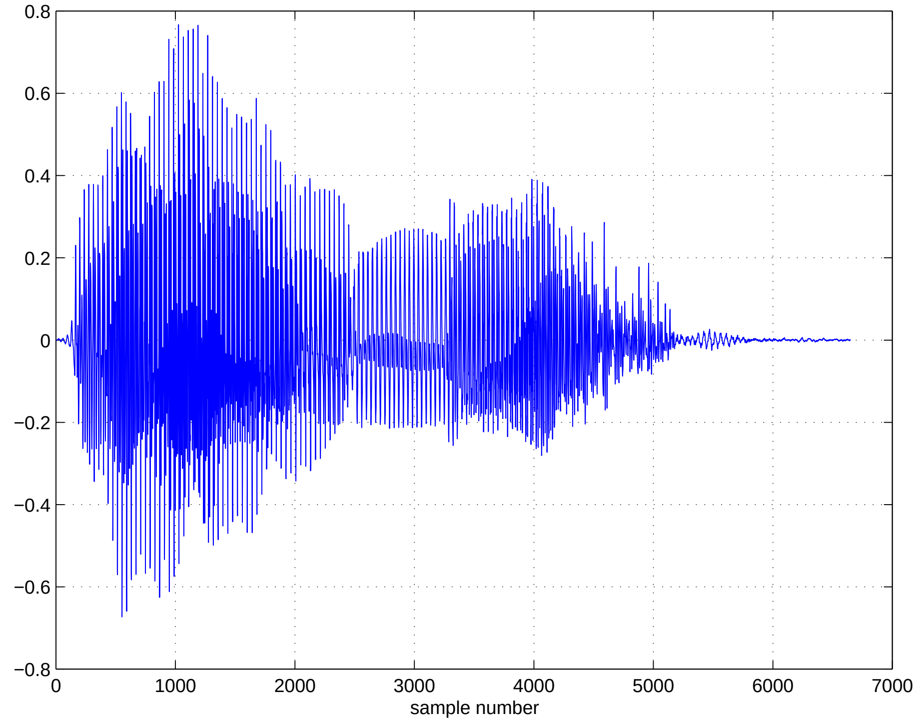

FIG. 3.17. *Open-loop prediction by median coefficients*

REFERENCES

[1] A. V. Oppenheim and R. W. Schafer, *Discrete-time Signal Processing*, Englewood Cliffs, NJ, Prentice-Hall, 1989.

[2] J. G. Proakis, *Digital communications*, 3rd ed., New York, McGraw-Hill, 1995.

[3] J. A. Chambers, *Digital Signal Processing Algorithms and Structures for Adaptive Line Enhancing*, PhD Thesis, 1990.

[4] F. Itakura and S. Saito, "Analysis Synthesis Telephony Based on the Maximum Likelihood Method," *Proc. 6th Int. Cong. Acoust.*, C-5-5, 1968.

[5] B. S. Atal and M. R. Schroeder, "Predictive Coding of Speech Signals," *Proc. 6th Int. Cong. Acoust.*, C-5-4, 1968.

[6] F. Sadaoki, *Speech Processing Systems*, 2nd ed., Marcel Dekker, 1989.

[7] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Englewood Cliffs, NJ, Prentice-Hall, 1978.

[8] W. B. Kleijn and K. K. Paliwal, *Speech coding and synthesis*, Elsevier, Amsterdam, 1995.

[9] J. R. Deller, Jr., J. G. Proakis and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*, Prentice-Hall, Inc., NJ., 1987.

[10] T. E. Landers and R. T. Lacoss, "Some Geophysical Applications of Autoregressive Spectral Estimates," *IEEE Trans. Geosci. Electron.*, Vol. GE15, pp. 26–32, Jan. 1977.

[11] J. Makhoul, R. Viswanathan, L. Cosell, and W. Russell, "Natural Communication with Computers: Speech Compression Research at BBN," *BBN Rep. 2976II*, Bolt, Beranek, and Newman, Inc., Cambridge, Mass., 1974.

[12] S. Haykin, *Adaptive filter theory*, 3rd ed. Englewood Cliffs, NJ, Prentice-Hall, 1996.

[13] D. Godard, "Self-recovering Equalization and Carrier Tracking in Two-dimensional Data Communication Systems," *IEEE Trans. Commun.*, Vol. COM-28, pp. 1867–1875, Nov. 1980.

[14] L. Tong, G. Xu and T. Kailath, "A new approach to blind identification and equalization of multipath channels," *Proc. 25st Asilomar Conf. Signals, Syst., Comput.*, pp. 856–860, Nov. 1991.

[15] D. T. M. Slock and C. B. Papadias, "Further Results on Blind Identification and Equalization of Multiple FIR channels," *IEEE ICASSP-95*, Vol. 3, pp. 1964–1967, 1995.

[16] J. Treichler and B. Agee, " A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-31, pp. 459–471, 1983.

[17] S. Verdú, B. Anderson, and R. Kennedy, "Blind equalization without gain identification," *IEEE Trans. Inform. Theory*, pp. 292–297, Jan. 1993.

[18] Z. Ding, R. Kennedy, B. Anderson, and C. Johnson, "Existence and avoidance of ill-convergence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, Vol. 39, pp. 1313–1327, Sep. 1991.

[19] Z. Ding, C. Johnson, and R. Kennedy, "On the (non)existence of undesirable equilibria of Godard blind equalizers," *IEEE Trans. Signal Processing*, Vol. 40, No. 10, pp. 2425–2432, Oct. 1992.

[20] L. Ye and Z. Ding, "Convergence Analysis of Finite Length Blind Adaptive Equalizers," *IEEE Trans. on Signal Proc.*, Vol. 43, No. 9, Sep. 1995.

[21] S. Haykin, *Blind deconvolution,* Englewood Cliffs, NJ:Prentice-Hall, 1994.

[22] C. R. Johnson, Jr., S. Dasgupta, and W. A. Sethares, "Averaging Analysis of Local Stability of a Real Constant Modulus Algorithm Adaptive Filter," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 36, No. 6, Jun. 1988.

[23] K. Hilal and P. Duhamel, "A Convergence Study of the Constant Modulus Algorithm Leading to a Normalized-CMA and a Block-Normalized-CMA," *Proc. EUSIPCO 92, VIth European Signal Processing Conference*, pp. 135–138, Aug. 1992.

[24] C. B. Papadias and D. T. M. Slock, "On the Convergence of Normalised Constant Modulus Algorithms for Blind Equalisation," *Proc. DSP-93*, Nicosia, Cyprus, Jul. 1993.

[25] O. Tanrikulu, A. G. Constantinides, and J. A. Chambers, "A normalised constant modulus blind channel equalisation algorithm via constrained optimisation," *ICT-96*, Vol. 2, pp. 836–839, April 1996.

[26] Y. C. Chen, *Adaptive Signal Modelling Based on Constant Modulus Algorithms*, Imperial College of Science, Technology, and Medicine, London, Mar. 1997.

[27] Y. C. Chen, T. Stathaki and A. G. Constantinides, "Adaptive Prediction Based on Blind Equalisation Principles," *ISMIP-96*, Taiwan, pp. 301–307, Dec. 1996 (invited paper).

[28] ——————-, "Nonlinear Adaptive Prediction of Nonstationary Signals with Application to Speech Coding," *IEEE Signal Processing, MMSP-97*, Princeton, New Jersey, Jun. 1997.

[29] A. Benveniste and M. Goursat, "Blind Equalizers," *IEEE Trans. Commun.*, Vol. COM-32, pp. 871–882, Aug. 1982.

[30] C. B. Papadias and D. T. M. Slock, "Normalised Sliding Window Constant Modulus Algorithms for Blind Equalisation," *14-th GRETSI Symposium on Signal and Image Processing*, Juan Les Pins, France, Sep. 1993.

[31] ——————-, "New Adaptive Blind Equalization Algorithms for Constant Modulus Constellations," *IEEE ICASSP-94*, Vol. 3, pp. 321–324, 1994.

[32] G. Foschini, “Equalization without altering or detecting data,” *Bell Syst. Tech. J.*, pp. 1885–1911, Oct. 1985.

[33] A. G. Constantinides, “Relaxation of the Normalised Constant Modulus Algorithm through Lagrangean Optimisation,” Signal Processing and Digital Systems Section, Department of Electrical and Electronic Engineering, Imperial College, 1996.

[34] Jer-Nan Juang, *Applied System Identification*, Englewood Cliffs, New Jersey 07632, 1994.