NOISE REDUCTION OF SPEECH SIGNALS USING THE RANK-REVEALING ULLV DECOMPOSITION

Peter S. K. Hansen, Per Christian Hansen†, Steffen Duus Hansen and John Aasted Sørensen

Department of Mathematical Modelling, Section for Digital Signal Processing
Technical University of Denmark, DK-2800 Lyngby, Denmark
E-mail: pskh@imm.dtu.dk, sdh@imm.dtu.dk and jaas@imm.dtu.dk

†UNI•C, Technical University of Denmark, DK-2800 Lyngby, Denmark
E-mail: Per.Christian.Hansen@uni-c.dk

ABSTRACT

A recursive approach for nonparametric speech enhancement is developed. The underlying principle is to decompose the vector space of the noisy signal into a signal subspace and a noise subspace. Enhancement is performed by removing the noise subspace and estimating the clean signal from the remaining signal subspace. The decomposition is performed by applying the rank-revealing ULLV algorithm to the noisy signal. With this formulation, a prewhitening operation becomes an integral part of the algorithm. Linear estimation is performed using a proposed minimum variance estimator. Experiments indicate that the approximative method is able to achieve a satisfactory quality of the reconstructed speech signal comparable with eigenfilter based methods.

1 INTRODUCTION

Recently, a new approach for noise reduction of speech signals based on subspace decomposition has been proposed [1, 2, 4]. The idea is to organize the noisy speech signal in a Toeplitz structured data matrix, and to decompose the span into two mutually orthogonal components.

The noise reduction algorithm in [1] is based on the Singular Value Decomposition (SVD), which is a robust and widely used computational tool in noise suppression techniques. From the SVD of the data matrix, the Least Squares (LS) estimate of the signal-only matrix can be obtained by neglecting the smallest singular values and finally the Toeplitz structure of the estimate is restored to identify the time samples. The problem is that the method deals only with white noise and the LS estimate is sensitive to the number of retained singular values.

In [4] is a noise reduction method based on the Quotient Singular Value Decomposition (QSVD) presented, where a prewhitening is an integral part of the algorithm. Moreover, by using a Minimum Variance (MV) estimate [7] of the signal-only matrix, the algorithm is less sensitive to the choice of retained singular values [4].

Unfortunately, the SVD/QSVD is computationally expensive and resists updating. This paper uses the rank-revealing ULV/ULLV decomposition [8, 6, 5] to estimate the rank and the orthogonal subspaces in the noise reduction algorithm. A recursive ULLV algorithm for a sliding window has been developed and an approximate MV estimate is proposed.

2 SIGNAL AND NOISE MODEL

Let $\mathbf{x} = (x_1, x_2, \cdots, x_m)^T$ denote the noisy signal vector of m samples and assume that the noise is additive and uncorrelated with the speech signal, i.e.,

$$\mathbf{x} = \mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{s}$ contains the speech component and $\mathbf{n}$ represents the noise. A set of time shifted vectors can be organized in a data matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ with Toeplitz structure

$$\mathbf{X} = \begin{pmatrix} x_n & x_{n-1} & \cdots & x_1 \\ x_{n+1} & x_n & \cdots & x_2 \\ \vdots & \vdots & & \vdots \\ x_{m+n-1} & x_{m+n-2} & \cdots & x_m \end{pmatrix} = \mathbf{S} + \mathbf{N} \tag{2}$$

where $m \geq n$. Moreover, assume that the noise is broadbanded so $\mathrm{rank}(\mathbf{X}) = \mathrm{rank}(\mathbf{N}) = n$ and that the speech signal can be described by a low order model, giving a rank deficient matrix $\mathbf{S}$ with $\mathrm{rank}(\mathbf{S}) = p < n$. It includes, for example, the *damped complex sinusoid* model, which has often been attributed to speech signals.

Thus, the speech signal is known to lie in a subspace of order p, but the subspace is unknown. The noise reduction problem is to *estimate* the subspace, i.e., its dimension and a suitable basis, and use this information in a signal processing procedure. Note, that its not possible to find the *exact* subspace.

3 LINEAR SIGNAL ESTIMATORS

One approach for *nonparametric* speech enhancement is linear estimation of the clean signal from the noisy signal using signal subspace methods, which is based on the SVD of the data matrix $\mathbf{X}$ partitioned as follows

$$\mathbf{X} = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{pmatrix} \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2 \end{pmatrix} \begin{pmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{pmatrix} \tag{3}$$

where $\mathbf{U}_1 \in \mathbb{R}^{m \times p}$, $\mathbf{V}_1 \in \mathbb{R}^{n \times p}$ and $\mathbf{\Sigma}_1 \in \mathbb{R}^{p \times p}$.

A straightforward and simple solution to the estimation problem is obtained by use of the *Least Squares* (LS) criterion, which minimizes the squared fitting errors between the noisy measurements $\mathbf{X}$ and a low rank model $\mathbf{S}_p$, i.e.,

$$\min_{\mathrm{rank}(\mathbf{S}_p)=p} \mathrm{tr}\left((\mathbf{X} - \mathbf{S}_p)^T(\mathbf{X} - \mathbf{S}_p)\right) \;\; \Rightarrow \tag{4}$$

$$\hat{\mathbf{S}}_{LS} = \mathbf{S}_p = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T \tag{5}$$

The estimate $\hat{\mathbf{S}}_{LS}$ is easily obtained without any statistical knowledge about the signals.

Assume now that the estimator $\hat{\mathbf{s}} \in \mathbb{R}^m$ of the pure signal vector $\mathbf{s}$ is constrained to be a *linear function* of the measurement vector $\mathbf{x}$, i.e., $\hat{\mathbf{s}} = \mathbf{W}\mathbf{x}$ where $\mathbf{W} \in \mathbb{R}^{m\times m}$ is a filter matrix, then the *Linear Minimum Mean-Squared Error* (LMMSE) estimator problem is to find the matrix $\mathbf{W}$ that minimizes

$$\min_{\mathbf{W}} \mathrm{tr} E\{(\mathbf{W}\mathbf{x}-\mathbf{s})(\mathbf{W}\mathbf{x}-\mathbf{s})^T\} \ \Rightarrow \tag{6}$$

$$\mathbf{W}_{LMMSE} = \mathbf{R}_s\mathbf{R}_x^{-1} \tag{7}$$

This theory produces the *Wiener-Hopf equations* as the fundamental design equations, i.e., we require the covariance properties of the noisy signal and the noise process.

In practice, this information is not available and is estimated from the noisy data. Under stationary and ergodic conditions, the ensemble average operator $E\{\cdot\}$ can be implemented as the mean value of several time shifted vectors, i.e., by use of the data matrix $\mathbf{X}$ and the signal-only matrix $\mathbf{S}$ (2). This gives us the *Minimum Variance* (MV) estimator

$$\min_{\mathbf{W}} \mathrm{tr}\left((\mathbf{X}\mathbf{W}-\mathbf{S})^T(\mathbf{X}\mathbf{W}-\mathbf{S})\right) \ \Rightarrow \tag{8}$$

$$\mathbf{W}_{MV} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{S} \tag{9}$$

which converges asymptotically to the LMMSE estimator as the number of rows $m \to \infty$ [7]. Note that $\mathbf{W} \in \mathbb{R}^{n\times n}$ in this case.

Since speech signals are nonstationary, a time varying estimator must be used. Such an estimator provides nonstationary residual noise with annoying noticeable tonal characteristics referred to as *musical noise*. This can be reduced [2] by maintaining the residual noise below some threshold either global or local in each eigenfilter. An ULV/ULLV treatment of these estimators is outside the scope of this paper.

4 ULV BASED SIGNAL ESTIMATION

The ULV decomposition was first introduced by Stewart [8]. A basic feature is that the ULV decomposition of a full rank matrix $\mathbf{X}$ can be made rank-revealing, if there is a gap in the singular values, e.g., when $\mathbf{X}$ is the sum of a rank deficient signal matrix $\mathbf{S}$ and a full rank noise matrix $\mathbf{N}$.

Assume that $\mathbf{X} \in \mathbb{R}^{m\times n}$ has numerical rank $p < n \le m$ corresponding to a given tolerance τ, then its singular values satisfy

$$\sigma_1 \ge \cdots \ge \sigma_p \ge \tau \gg \sigma_{p+1} \ge \cdots \ge \sigma_n \tag{10}$$

and there exists a matrix $\mathbf{U} \in \mathbb{R}^{m\times n}$ with orthogonal columns and an orthogonal matrix $\mathbf{V} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{X} = \mathbf{U}\mathbf{L}\mathbf{V}^T = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{pmatrix} \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} \\ \mathbf{F} & \mathbf{G} \end{pmatrix} \begin{pmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{pmatrix} \tag{11}$$

where $\mathbf{L} \in \mathbb{R}^{n\times n}$, $\mathbf{L}_1 \in \mathbb{R}^{p\times p}$ and $\mathbf{G} \in \mathbb{R}^{(n-p)\times(n-p)}$ are lower triangular, and

$$\sigma_{min}(\mathbf{L}_1) \approx \sigma_p \tag{12}$$

$$\|\mathbf{F}\|_F^2 + \|\mathbf{G}\|_F^2 \approx \sigma_{p+1}^2 + \cdots + \sigma_n^2 \tag{13}$$

From the RRULVD we can estimate the signal- and noise subspaces defined by the gap in the singular values. The tolerance τ is defined based on a detection threshold in the underlying signal processing problem.

4.1 LS Estimate by RRULVD

An approximate least squares estimate $\hat{\mathbf{S}}_{ALS}$ of the signal matrix $\mathbf{S}$ can be computed by essentially substituting the ULV decomposition for the SVD based estimate [3], thus replacing one problem with a similar, nearby problem that can be solved more efficiently.

Based on (5) and (11), a useful rank-p matrix approximation to $\mathbf{X}$ is given by

$$\hat{\mathbf{S}}_{ALS} = \mathbf{U}_1\mathbf{L}_1\mathbf{V}_1^T = \mathbf{X}\mathbf{V}_1\mathbf{V}_1^T \tag{14}$$

where $\mathbf{U}_1$ and $\mathbf{V}_1$ approximate the numerical column space and row space as defined via the SVD of $\mathbf{X}$.

4.2 MV Estimate by RRULVD

The minimum variance estimate $\hat{\mathbf{S}}_{MV}$ of the signal matrix $\mathbf{S}$ can be obtained along the lines in [7] using an *idealized* rank-revealing ULV decomposition of $\mathbf{X} \in \mathbb{R}^{m\times n}$. With reference to (11), the necessary conditions are

1. The signal is orthogonal to the noise in the sense: $\mathbf{S}^T\mathbf{N} = \mathbf{0}$.
2. The matrix $\mathbf{N} = \sigma_{noise}\mathbf{Q}$, where $\mathbf{Q}$ has orthonormal columns: $\mathbf{N}^T\mathbf{N} = \sigma_{noise}^2\mathbf{I}_n$.
3. There is a distinct gap in the singular values of the matrix $\mathbf{X}$: $\sigma_p > \sigma_{p+1}$.
4. The off-diagonal matrix $\mathbf{F}$ is zero.
5. $\mathbf{G}$ is a diagonal matrix containing the noise-only singular values σ_{noise}.

Thus, we have

$$\mathbf{X} = \begin{pmatrix} \mathbf{U}_{X1} & \mathbf{U}_{X2} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{X1} & \mathbf{0} \\ \mathbf{0} & \sigma_{noise}\mathbf{I}_{n-p} \end{pmatrix} \begin{pmatrix} \mathbf{V}_{X1}^T \\ \mathbf{V}_{X2}^T \end{pmatrix} \tag{15}$$

Let the ULV decomposition of the matrix $\mathbf{S}$ be defined by

$$\mathbf{S} = \begin{pmatrix} \mathbf{U}_{S1} & \mathbf{U}_{S2} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{S1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V}_{S1}^T \\ \mathbf{V}_{S2}^T \end{pmatrix} \tag{16}$$

where $\mathbf{L}_{S1} \in \mathbb{R}^{p\times p}$, then we can write the idealized rank-revealing ULV decomposition of $\mathbf{X}$ in terms of the ULV decomposition of $\mathbf{S}$

$$\begin{aligned} \mathbf{X} &= \mathbf{U}_{S1}\mathbf{L}_{S1}\mathbf{V}_{S1}^T + \mathbf{N}\mathbf{V}_{S1}\mathbf{V}_{S1}^T + \mathbf{N}\mathbf{V}_{S2}\mathbf{V}_{S2}^T \\ &= \begin{pmatrix} (\mathbf{U}_{S1}\mathbf{L}_{S1} + \mathbf{N}\mathbf{V}_{S1})\mathbf{L}_{X1}^{-1} & \mathbf{N}\mathbf{V}_{S2}\sigma_{noise}^{-1} \end{pmatrix} \\ &\quad \times \begin{pmatrix} \mathbf{L}_{X1} & \mathbf{0} \\ \mathbf{0} & \sigma_{noise}\mathbf{I}_{n-p} \end{pmatrix} \begin{pmatrix} \mathbf{V}_{S1}^T \\ \mathbf{V}_{S2}^T \end{pmatrix} \end{aligned} \tag{17}$$

The matrix $\mathbf{L}_{X1}^T\mathbf{L}_{X1}$ can be obtained by comparing the matrix $\mathbf{X}^T\mathbf{X}$ using the definitions of $\mathbf{S}$ and $\mathbf{N}$ with the one based on the ULV decomposition of $\mathbf{X}$ (15), which gives

$$\mathbf{L}_{X1}^T\mathbf{L}_{X1} = \mathbf{L}_{S1}^T\mathbf{L}_{S1} + \sigma_{noise}^2\mathbf{I}_p \tag{18}$$

Using (16) and (17) in the MV definition (9) yields the desired MV estimate of $\mathbf{S}$

$$\begin{aligned} \hat{\mathbf{S}}_{MV} &= \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{S} \\ &= \mathbf{U}_X\mathbf{U}_X^T\mathbf{S} \\ &= \begin{pmatrix} \mathbf{U}_{X1} & \mathbf{U}_{X2} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{X1}^{-T}(\mathbf{L}_{S1}^T\mathbf{U}_{S1}^T + \mathbf{V}_{S1}^T\mathbf{N}^T) \\ \sigma_{noise}^{-1}\mathbf{V}_{S2}^T\mathbf{N}^T \end{pmatrix} \\ &\quad \times \begin{pmatrix} \mathbf{U}_{S1} & \mathbf{U}_{S2} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{S1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V}_{S1}^T \\ \mathbf{V}_{S2}^T \end{pmatrix} \\ &= \mathbf{U}_{X1}(\mathbf{L}_{X1} - \sigma_{noise}^2\mathbf{L}_{X1}^{-T})\mathbf{V}_{X1}^T \end{aligned} \tag{19}$$

where (18) has been used. This equation can be reformulated to avoid an explicit computation of $\mathbf{U}_X$

$$\hat{\mathbf{S}}_{MV} = \mathbf{X}\mathbf{V}_{X1}\mathbf{L}_{X1}^{-1}(\mathbf{L}_{X1} - \sigma_{noise}^2 \mathbf{L}_{X1}^{-T})\mathbf{V}_{X1}^T \quad (20)$$

The quantity σ_{noise}^2 can be obtained from (13)

$$\sigma_{noise}^2 = \frac{1}{n-p}\sum_{i=p+1}^{n}\sigma_i^2 \approx \frac{1}{n-p}\left(\|\mathbf{F}\|_F^2 + \|\mathbf{G}\|_F^2\right) \quad (21)$$

In practice, the above mentioned conditions are never satisfied exactly, but the rank-revealing ULV decomposition is robust with respect to mild violations of these conditions.

5 ULLV BASED SIGNAL ESTIMATION

If the additive noise $\mathbf{N}$ is colored, $\mathbf{N}^T\mathbf{N} \neq \sigma_n^2\mathbf{I}_n$, then a prewhitening transformation can be applied to the data matrix using the QR decomposition of $\mathbf{N} = \mathbf{Q}\mathbf{R}$

$$\mathbf{X}\mathbf{R}^{-1} = \mathbf{S}\mathbf{R}^{-1} + \mathbf{N}\mathbf{R}^{-1} = \mathbf{S}\mathbf{R}^{-1} + \mathbf{Q} \quad (22)$$

This transformation does not change the nature of the low order model of the speech signal while it diagonalizes the covariance matrix of the noise. In this application the noise matrix $\mathbf{N}$ can be estimated in periods without speech.

One problem concerning the prewhitening transformation is the complicated update of the matrix $\mathbf{X}\mathbf{R}^{-1}$ when $\mathbf{X}$ and $\mathbf{N}$ are updated, e.g., in a recursive application. This can be avoided by using the ULLV decomposition of the matrix pair $(\mathbf{X}, \mathbf{N})$, which allows each matrix to be updated individually and delivers the required factorizations without forming the quotients and products.

The definition given here for the rank-revealing ULLV decomposition (RRULLVD) of two matrices $\mathbf{X} \in \mathbb{R}^{m\times n}$ and $\mathbf{N} \in \mathbb{R}^{m\times n}$ is the one used by Luk and Qiao [6].

Assume that $\mathbf{X}\mathbf{N}^+$ ($\mathbf{N}^+$ is the pseudoinverse of $\mathbf{N}$) has numerical rank $p < n \leq m$ corresponding to a given tolerance τ, then its quotient singular values satisfy

$$\delta_1 \geq \cdots \geq \delta_p \geq \tau \gg \delta_{p+1} \geq \cdots \geq \delta_n \quad (23)$$

and there exist matrices $\mathbf{U}_X \in \mathbb{R}^{m\times n}$ and $\mathbf{U}_N \in \mathbb{R}^{m\times n}$ with orthogonal columns and a orthogonal matrix $\mathbf{V} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{X} = \begin{pmatrix} \mathbf{U}_{X1} & \mathbf{U}_{X2} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{X1} & \mathbf{0} \\ \mathbf{F} & \mathbf{G} \end{pmatrix} \mathbf{L} \begin{pmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{pmatrix} \quad (24)$$

$$\mathbf{N} = \mathbf{U}_N \mathbf{L} \mathbf{V}^T \quad (25)$$

where $\mathbf{L} \in \mathbb{R}^{n\times n}$, $\mathbf{L}_{X1} \in \mathbb{R}^{p\times p}$ and $\mathbf{G} \in \mathbb{R}^{(n-p)\times(n-p)}$ are lower triangular, and

$$\sigma_{min}(\mathbf{L}_{X1}) \approx \delta_p \quad (26)$$

$$\|\mathbf{F}\|_F^2 + \|\mathbf{G}\|_F^2 \approx \delta_{p+1}^2 + \cdots + \delta_n^2 \quad (27)$$

Thus, the ULLV decomposition reveals the rank of the matrix $\mathbf{X}\mathbf{N}^+$ assuming $\mathbf{N}$ has full rank

$$\mathbf{X}\mathbf{N}^+ = \begin{pmatrix} \mathbf{U}_{X1} & \mathbf{U}_{X2} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{X1} & \mathbf{0} \\ \mathbf{F} & \mathbf{G} \end{pmatrix} \begin{pmatrix} \mathbf{U}_{N1}^T \\ \mathbf{U}_{N2}^T \end{pmatrix} \quad (28)$$

Hence, working with the RRULLVD of $(\mathbf{X}, \mathbf{N})$ and the matrix $\mathbf{Q}$ is mathematically equivalent to working with the RRULVD of $\mathbf{X}\mathbf{R}^{-1}$.

5.1 LS Estimate by RRULLVD

An approximate LS estimate $\hat{\mathbf{S}}_{ALS}$ of the low-rank signal matrix $\mathbf{S}$ added colored noise can easily be obtained by first substituting the ULV decomposition of $\mathbf{X}\mathbf{N}^+$ for the SVD based estimate

$$(\mathbf{X}\mathbf{N}^+)_{ALS} = \mathbf{U}_{X1}\mathbf{L}_{X1}\mathbf{U}_{N1}^T \quad (29)$$

and then perform a denormalization of $(\mathbf{X}\mathbf{N}^+)_{ALS}$

$$\hat{\mathbf{S}}_{ALS} = (\mathbf{X}\mathbf{N}^+)_{ALS}\mathbf{N} = \mathbf{U}_{X1}\mathbf{L}_{X1}\mathbf{L}_1\mathbf{V}_1^T \quad (30)$$

which can be computed directly from the ULLV decomposition, i.e., the prewhitening is now an integral part of the algorithm. As before, equation (30) can be reformulated to avoid an explicit computation of $\mathbf{U}_X$

$$\hat{\mathbf{S}}_{ALS} = \mathbf{X}\mathbf{V}_1\mathbf{V}_1^T \quad (31)$$

5.2 MV Estimate by RRULLVD

The approximate minimum variance estimate $\hat{\mathbf{S}}_{AMV}$ of the low-rank signal matrix in the colored noise case follows from the least squares analysis.

Using (19) with $\sigma_{noise}^2 = 1$, the approximate MV estimate of the normalized data matrix $\mathbf{X}\mathbf{N}^+$ defined by (28) is

$$(\mathbf{X}\mathbf{N}^+)_{AMV} = \mathbf{U}_{X1}(\mathbf{L}_{X1} - \mathbf{L}_{X1}^{-T})\mathbf{U}_{N1}^T \quad (32)$$

To obtain the corresponding approximate minimum variance estimate of $\mathbf{S}$, we must denormalize $(\mathbf{X}\mathbf{N}^+)_{AMV}$

$$\hat{\mathbf{S}}_{AMV} = (\mathbf{X}\mathbf{N}^+)_{AMV}\mathbf{N} = \mathbf{U}_{X1}(\mathbf{L}_{X1} - \mathbf{L}_{X1}^{-T})\mathbf{L}_1\mathbf{V}_1^T \quad (33)$$

where we have used (25). Again, this equation can be reformulated to avoid an explicit computation of $\mathbf{U}_X$

$$\hat{\mathbf{S}}_{AMV} = \mathbf{X}\mathbf{V}_1\mathbf{L}_1^{-1}\mathbf{L}_{X1}^{-1}(\mathbf{L}_{X1} - \mathbf{L}_{X1}^{-T})\mathbf{L}_1\mathbf{V}_1^T \quad (34)$$

6 EXPERIMENTS

A recursive RRULLV algorithm has been developed based on the methods given in [8, 6, 5]. Starting with initial matrices, the decomposition is updated as $\mathbf{X}$ and $\mathbf{N}$ are taken into account one row at a time. A new row is processed in the following four steps. Updating: The current row of $\mathbf{X}$ or $\mathbf{N}$ is incorporated into the decomposition. Downdating: The oldest row of $\mathbf{X}$ or $\mathbf{N}$ is isolated and removed in the decomposition. Deflation: Establishes and maintains the rank-revealing nature of the decomposition. Refinement: The norm of $\mathbf{F}$ is reduced to improve the subspace quality. By using an exponential window, the downdating step can be omitted, but clearly, the sliding window method can track the change in the signal statistics more accurately when there is an abrupt change in data.

The recursive RRULLV algorithm was applied to speech signals contaminated by an AR(1,-0.7) noise process and the noise matrix $\mathbf{N}$ was only updated in periods without speech. All the signals were sampled at 8 kHz and the matrix dimension was $m = 141$ and $n = 20$.

The typical average SNR of a reconstructed speech segment (voiced) using 100 noise realizations and SNR = 5 dB is illustrated in Fig. 1 as a function of the signal subspace dimension p. Clearly, the MV estimate is less sensitive to the choice of p compared with the LS estimate. Thus, using a fixed value of $p = 12$ as in the following results, we are able to achieve a satisfactory quality of the reconstructed speech. The behavior of the reconstructed segment in the frequency domain was also analyzed using a tenth order LPC model spectra of noise-free, noisy and reconstructed (MV estimate) speech segments, respectively. As shown in Fig. 2, the MV estimate improves the spectrum in the regions near the dominant formants. These results closely match the QSVD based method [4].

The RRULLV algorithm using a sliding window was applied to the speech signal in Fig. 3 added broad-band noise (global SNR of 5 dB). Observe from Fig. 4 that the global SNR improvement using the MV estimate is about twice the LS based improvement due to the fixed p. Moreover, the variations among the local SNRs of the various segments are reduced.

In the RRULLV algorithm computations can be saved by using the exponential window, but as demonstrated in Fig. 5, the sliding window method gives up to 6 dB better SNR, when there is a change in the dynamics of the signal. The same is true by comparing the SNRs obtained from the RRULLV sliding window method with the QSVD segment based approach also illustrated in Fig. 5.

7 SUMMARY

A recursive signal subspace approach for noise reduction of speech signals is presented. The algorithm is formulated by means of the RRULLVD using a proposed MV estimator. The method was demonstrated to be comparable with eigenfilter based methods. Integration of the RRULLVD with perceptually more meaningful estimation criterias is a topic of current research.

References

[1] M. Dendrinos, S. Bakamidis, and G. Carayannis. Speech Enhancement from Noise: A Regenerative Approach. *Speech Communication*, 10(1):45–57, February 1991.

[2] Yariv Ephraim and Harry L. Van Trees. A Signal Subspace Approach for Speech Enhancement. *IEEE Trans. on Speech and Audio Processing*, 3(4):251–266, July 1995.

[3] Ricardo D. Fierro and Per Christian Hansen. Accuracy of TSVD Solutions Computed from Rank-Revealing Decompositions. *Numerische Mathematik*, 70:453–471, 1995.

[4] S. H. Jensen, P. C. Hansen, S. D. Hansen, and J. Aa. Sørensen. Reduction of Broad-Band Noise in Speech by Truncated QSVD. *IEEE Trans. on Speech and Audio Processing*, 3(6):439–448, November 1995.

[5] J. M. Lebak and A. W. Bojanczyk. Modifying a Rank-Revealing ULLV Decomposition. Technical report, Cornell Theory Center, June 1994.

[6] F. T. Luk and S. Qiao. A New Matrix Decomposition for Signal Processing. In M. S. Moonen et al., editor, *Linear Algebra for Large Scale and Real-Time Applications*, pages 241–247. Kluwer Academic Publishers, 1993.

[7] Bart De Moor. The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices. *IEEE Trans. on Signal Processing*, 41(9):2826–2838, September 1993.

[8] G. W. Stewart. Updating a Rank-Revealing ULV Decomposition. *SIAM Journal on Matrix Analysis and Applications*, 14(2):494–499, April 1993.

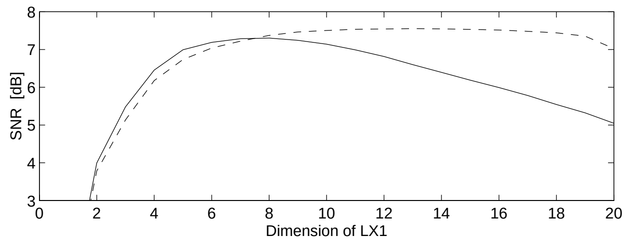

Figure 1 Average SNR of a reconstructed voiced speech segment, SNR=5dB, LS estimate (solid), MV estimate (dashed).

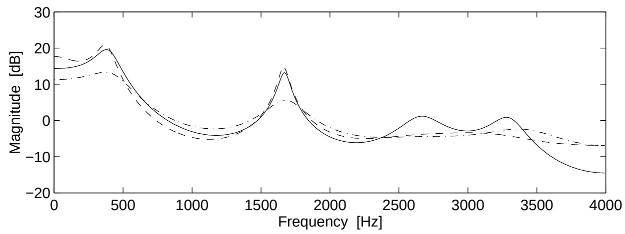

Figure 2 LPC model spectra of noise-free speech segment (solid), noisy speech segment, SNR=5dB (dash-dot) and MV estimate (dashed).

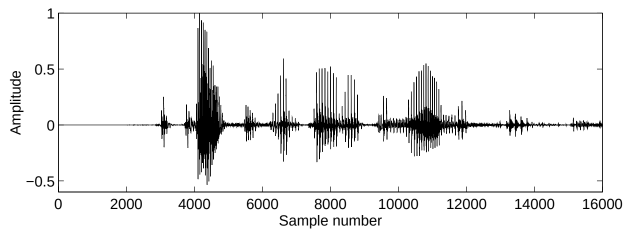

Figure 3 Noise-free speech signal.

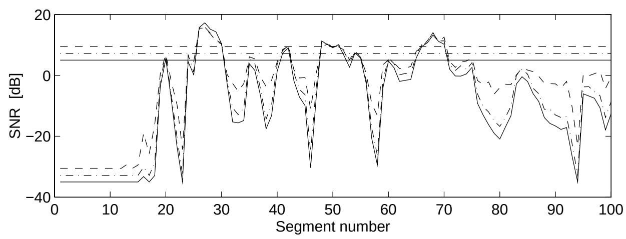

Figure 4 Local/global SNR of noisy speech signal (solid), LS estimate (dash-dot) and MV estimate (dashed).

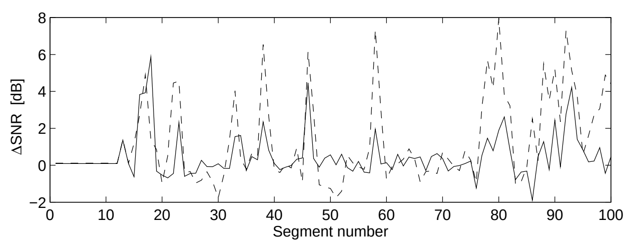

Figure 5 Difference in SNR between sliding and exponential window based MV estimate (solid) and between RRULLVD sliding window MV estimate and QSVD segment based MV estimate (dashed).