AN EIGEN-BASED APPROACH FOR COMPLEX-VALUED FORECASTING

S. Enshaeifar, S. Sanei and C. Cheong Took

Department of Computing, University of Surrey, GU2 7XH, U.K.
Emails: {s.enshaeifar, s.sanei, c.cheongtook}@surrey.ac.uk

ABSTRACT

Forecasting one step ahead is generally straightforward. Forecasting two steps ahead a little more challenging. Forecasting further into the horizon may require prior forecasted samples, as the availability of historical data may not be adequate to do so. It is in this motivational context that we proposed an eigen-based approach for complex-valued multiple-step ahead forecasting. Here we establish an augmented complex-domain singular spectrum analysis framework to perform prediction beyond 50 step ahead. It is shown that other prediction algorithms such as the least mean square, though useful and adaptive, cannot use the predicted samples to predict further. In some cases, they may diverge from the trend. Simulations on real-world data support our approach.

***Index Terms*—** Singular Spectrum Analysis, Forecasting, Augmented Statistics

1. INTRODUCTION

There are many areas which require forecasting algorithms such as financial stock price prediction or wind forecasting in renewable energy [1, 2]. Although there are different prediction techniques, they are mostly based on restrictive assumptions such as Gaussianity or circularity of the data. Therefore, applying a method with less prior assumptions would be useful for modelling and forecasting purposes [3, 4]. Furthermore, prediction accuracy is likely to be affected by noise, and currently there are not many effective forecasting algorithms robust against noise [4]. Two general approaches for prediction of noisy signals can be found in the literature. In the first one, presence of noise is overlooked and a forecasting algorithm operates directly on the noisy data. In the second approach, noise is first mitigated by filtering the original data and new points are forecasted based on the filtered data [4]. It is crucial to select a suitable filtering algorithm which alleviates the effect of noise, yet maintains the structure of the signal. There are several nonlinear noise reduction algorithms, among which singular value decomposition (SVD) based method is accepted as an effective method for noise reduction [4].

Singular spectrum analysis (SSA) is a powerful model-free, SVD-based technique which does not need any prior statistical assumptions such as normality or Gaussianity of data [4, 5]. SSA has found many applications such as single-channel source separation, non-parametric signal decomposition, denoising and forecasting [1, 5, 6]. SSA consists of two major stages: decomposition and reconstruction [1]. Reconstruction stage is generally used to mitigate the effect of noise. The reconstructed component is then used to forecast the future samples [1, 4]. Note that SSA forecasting can only be employed for the data which approximately satisfy the linear recurrent formula [1](LRF) [1, 4]. SSA forecasting is a recursive procedure where at each time instant only one step ahead is predicted. Thus, the first step ahead is predicted from the reconstructed data and all other predictions are computed using the previously predicted values [1].

Despite all these advantages, one of the shortcomings of the SSA method is that this technique was mostly developed for real-valued signals. Many important signal processing areas such as in radar or telecommunications involve complex-valued data, thus several algorithms were developed in this direction [2, 7, 8]. For example, one of the most well known algorithms is the Least Mean Square (LMS) which was first introduced by Widrow and Hoff in 1959. Widrow then adapted LMS for complex-valued signals (CLMS) in 1960 [9]. Furthermore, recent advances in complex statistics brought to light the augmented statistics to cater for both circular and non-circular signals [7]. For instance, augmented CLMS (ACLMS) was proposed in [7]. In the same spirit, this work will provide a unified framework to cater for complex-valued signals by taking advantage of the complete second order information. To this end, the augmented complex SSA (A-CSSA) is proposed to address rigorously both circular and noncircular data. The usefulness and enhanced performance offered by our approach is illustrated on synthetic data as well as real-world signals such as financial and wind data in prediction applications.

2. COMPLEX SSA ALGORITHM

The objective of SSA is to represent the original signal as a sum of a small number of components which can be identi-

[1]It is assumed that a signal $\mathbf{f}_n$ satisfies linear recurrent formula of order d if $f_n = a_1 f_{n-1} + \cdots + a_d f_{n-d}$ where a_i are the coefficients.

978-1-4799-2893-4/14/$31.00

fied as a trend, periodic or quasi-periodic component or noise. This is achieved by mapping the given signal in terms of eigenvectors and eigenvalues of a matrix generated from the original signal. The stages in SSA are described as follows [1, 10].

2.1. Decomposition

The original signal is considered as one-dimensional vector $\mathbf{f}_S = (f_1, \dots, f_S)^T \in \mathbb{C}$ with length S. In the decomposition stage, this vector is mapped to the trajectory matrix $\mathbf{W} \in \mathbb{C}^{Le \times N}$. The parameter Le is defined as the window length where $1 < Le < S$ and $N = S - Le + 1$. Window length Le should be long enough to account for information about data variation [1, 4].

$$\mathbf{W} = (w_{ij})_{i,j=1}^{Le,N} = \begin{pmatrix} f_1 & f_2 & f_3 & \cdots & f_N \\ f_2 & f_3 & f_4 & \cdots & f_{N+1} \\ \vdots & \vdots & \ddots & \vdots & \\ f_{Le} & f_{Le+1} & f_{Le+2} & \cdots & f_S \end{pmatrix} \tag{1}$$

Observe that $\mathbf{W}$ is a Hankel matrix in which entries along the skew diagonals ($i + j =$ const) are equal. In the next step, SVD is applied to the generated trajectory matrix such that j-th component of SVD is defined by j-th eigenvalue (λ_j) and eigenvector ($\mathbf{q}_j$) of the covariance matrix $\mathbf{W}\mathbf{W}^H$. Since $\mathbf{W}\mathbf{W}^H$ is a Hermitian and positive definite matrix, it can be diagonalised as $\mathbf{Q}\mathbf{\Sigma}\mathbf{Q}^H$, in which $\mathbf{\Sigma}$ is a diagonal matrix of positive eigenvalues placed in the decreasing order ($\lambda_1 > \lambda_2 > \ldots > \lambda_{Le}$) and $\mathbf{Q}$ is an orthogonal matrix of the corresponding eigenvectors [1, 4]. Thus, SVD of the trajectory matrix can be rewritten as:

$$\mathbf{W} = \sum_{j=1}^{r} \mathbf{W}_j = \sum_{j=1}^{r} \sqrt{\lambda_j}\mathbf{q}_j\mathbf{v}_j^H \quad \text{where} \quad \mathbf{v}_j = \mathbf{W}^H\mathbf{q}_j/\sqrt{\lambda_j} \tag{2}$$

$\sqrt{\lambda_j}$ are the singular values of the matrix $\mathbf{W}$ and r is defined as $r = \max\{j : \lambda_j > 0\}$. The set $(\lambda_j, \mathbf{q}_j, \mathbf{v}_j)$ is called j-th eigentriple of the matrix $\mathbf{W}$ and $\mathbf{W}_j$ is defined as the elementary matrix.

2.2. Reconstruction

First part of the reconstruction stage consists of dividing the elementary matrices $\mathbf{W}_j$ into a number of groups and adding the matrices within each group [1]. Thus, each group is represented by corresponding matrix $\tilde{\mathbf{W}}_g \subset \mathbb{C}^{Le \times N}$ and SVD of the trajectory matrix can be rewritten as:

$$\mathbf{W} = \sum_{g=1}^{z} \tilde{\mathbf{W}}_g \tag{3}$$

where z determines the total number of groups. Index g defines g-th subgroup of eigentriples and $\tilde{\mathbf{W}}_g$ is the sum of elementary matrices within group g. After splitting the SVD components, a specific subgroup ($\tilde{\mathbf{W}}_g$) is selected and the subseries is reconstructed by using the Hankelization algorithm[2][1, 5]. In this case, if $\tilde{w}_{ij}$ refers to an entry of the matrix ($\tilde{\mathbf{W}}_g$), s-th term of the new reconstructed series is calculated by averaging all $\tilde{w}_{ij}$ along all i, j such that $(i + j = s + 1)$. Therefore, the following parameters can be initialized as $s = 1$, $\tilde{f}_1 = \tilde{w}_{11}$ and for $s = 2$, $\tilde{f}_2 = (\tilde{w}_{12} + \tilde{w}_{21})/2$ and so on [1].

$$\tilde{\mathbf{W}}_g = \begin{pmatrix} \tilde{w}_{11} & \tilde{w}_{12} & \cdots & \tilde{w}_{1,N} \\ \tilde{w}_{21} & \tilde{w}_{22} & \cdots & \tilde{w}_{2,N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{w}_{Le,1} & \tilde{w}_{Le,Le+1} & \cdots & \tilde{w}_{Le,S} \end{pmatrix}, \quad \tilde{\mathbf{f}} = [\tilde{f}_1, \tilde{f}_2, \ldots, \tilde{f}_S] \tag{4}$$

$\tilde{\mathbf{f}}$ is the reconstructed signal with length S. Observe that one of the main challenges in SSA is finding the group of the eigentriples for reconstructing the component of interest.

2.3. Forecasting

SSA forecasting algorithm was first introduced in [1] for the real-valued processes that satisfies LRF. If the rank of the trajectory matrix is smaller than window length ($r < Le$), the signal satisfies the linear recurrent formula. For such signals, SSA could be applied as the forecasting algorithm. In this case, the reconstruction stage of SSA aims to smooth the original data by removing the eigentriples corresponding to noise. Then, a recurrent forecasting procedure is applied to the reconstructed signal [1].

Consider $\check{\mathbf{q}} \in \mathbb{C}^{Le-1}$ consisting of the last $(Le - 1)$ elements of the eigenvector $\mathbf{q}$ and $\upsilon^2 = {\eta_1}^2 + \cdots + {\eta_r}^2$ in which η_i is the last element of the corresponding eigenvector. The vector $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_{Le-1})^T$ is then defined as $\boldsymbol{\delta} = \frac{1}{1-\upsilon^2}\sum_{i=1}^{r} \eta_i\check{\mathbf{q}}_i$ where i refers to index of the eigenvector. At this stage, the following equation is used to forecast h steps ahead using $\boldsymbol{\delta}$ and the reconstructed signal $\tilde{\mathbf{f}}$ [1, 4, 5].

$$\tilde{\mathbf{f}}_i = \begin{cases} \tilde{\mathbf{f}}_i & \forall i = 1, \ldots, S \\ \sum_{j=1}^{Le-1} \delta_j \tilde{\mathbf{f}}_{i-j} & \forall i = S+1, \ldots, S+h \end{cases} \tag{5}$$

Remark#1: Note that Equation (5) caters for real-valued signals. To adapt such procedure for complex data, we adopt the split complex approach whereby prediction is calculated from the real and imaginary part of the eigenvectors separately. This is because η_i represents the steering angle between a real-valued eigenvector and the real axis. On the

[2]Hankelization refers to averaging along cross-diagonals of matrix ($\tilde{\mathbf{W}}_g$), i.e. averaging along elements with indices ($i + j =$ const). Full description of Hankelization algorithm can be found in [1, 5].

other hand, for complex-valued data, the angle corresponds to the phase between real and imaginary components of the vector. This cross-information will be exploited in the form of augmented statistics (see next section), and therefore is not considered here explicitly in our prediction task.

3. AUGMENTED COMPLEX SSA (A-CSSA)

Typically, statistics of complex domain are considered as the direct extension of $\mathbb{C}$ statistics. For example, covariance matrix of a zero mean complex vector could simply be defined by replacing the standard transpose operator $(.)^T$ with the Hermitian transpose [2]. However, recent works showed that basic complex covariance matrix ignores the correlation between the real and imaginary part [7, 11]. Therefore, “augmented” statistics have been established to generalize the optimal second-order statistics for complex domain [7] in which both covariance and pseudo-covariance are considered. To incorporate the latest advances in complex-valued statistics into the CSSA framework and extend it as the A-CSSA, consider a univariate complex signal $\mathbf{f} \in \mathbb{C}^{1,S}$. Similarly to CSSA, the overall complex trajectory matrix $(\mathbf{W})$ is generated via (1).
Next step aims at calculating SVD of the trajectory matrix $\mathbf{W}$ using its covariance matrix. For basic CSSA, the covariance was generated as $\mathbf{W}\mathbf{W}^H$. In A-CSSA, the basic trajectory matrix is combined with its complex conjugate to generate the augmented trajectory matrix ($\mathbf{W}^a \subset \mathbb{C}^{2Le \times N}$) and then it is used to compute the new augmented covariance matrix ($\mathbf{C}^a \subset \mathbb{C}^{2Le \times N}$):

$$\begin{aligned} \mathbf{W}^a &= \begin{bmatrix} \mathbf{W} \\ \mathbf{W}^* \end{bmatrix} \\ \mathbf{C}^a &= \mathbf{W}^a {\mathbf{W}^a}^H = \begin{bmatrix} \mathbf{W}\mathbf{W}^H & \mathbf{W}\mathbf{W}^T \\ (\mathbf{W}\mathbf{W}^T)^* & (\mathbf{W}\mathbf{W}^H)^* \end{bmatrix} \end{aligned} \tag{6}$$

Remark#2: The new augmented covariance matrix in (6) for the A-CSSA, which includes both covariance and pseudo-covariance, takes into account the complete second order information in the complex-valued data. Thus, it is more likely that A-CSSA outperforms its basic complex-valued SSA for non-circular signals.
Remark#3: Note that the A-CSSA method generates $2Le$ eigenvalues which is double the size of its basic complex-valued counterpart. Hence, the higher computational complexity of A-CSSA vis-a-vis CSSA.

4. SIMULATION AND RESULTS

To evaluate the benefits of the proposed method, CSSA and A-CSSA were applied to three different data sets: synthetic data using ARMA model and real-world data such as financial stocks and wind data. Note that noise free signal is not available for training purpose or constructing predictions. Thus, the goal is to forecast the unavailable clean samples from the smoothed version of the noisy data computed by SSA filtering [12]. To evaluate the performance of our proposed algorithm, normalised root mean square (NRMS) error was calculated among the original data $(\mathbf{f})$ and final output of the SSA $(\tilde{\mathbf{f}})$ which is a combination of the reconstructed data and the predicted values [4]:

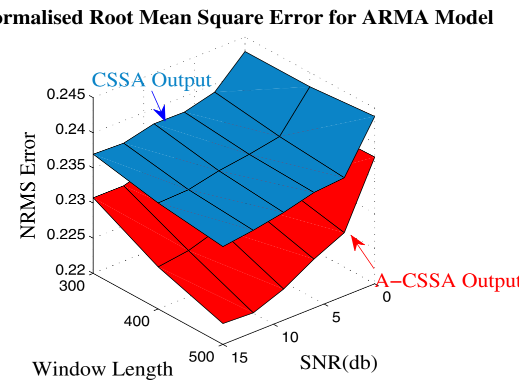

Fig. 1. NRMS error for the ARMA model simulation using C-SSA and A-CSSA. Note that the plotted NRMS error is averaged over 20 sets of simulation.

$$\text{NRMSE} = \frac{\sqrt{\dfrac{\sum_{i=1}^{S}(\mathbf{f}_i - \tilde{\mathbf{f}}_i)^2}{S}}}{\max(\mathbf{f}) - \min(\mathbf{f})} \tag{7}$$

.

4.1. Synthetic data

The first experiment evaluated the reconstruction and prediction accuracy for different levels of white noise. Simulations were performed on the widely linear ARMA signals in which $\mathbf{w}(s)$ is doubly white Gaussian noise [7]:

$$\begin{aligned} \mathbf{f}(s) &= 1.79\mathbf{f}(s-1) - 1.85\mathbf{f}(s-2) \\ &+ 1.27\mathbf{f}(s-3) - 0.41\mathbf{f}(s-4) \\ &+ 0.2\mathbf{f}(s-5) + 2\mathbf{w}(s) + 0.5\mathbf{w}^*(s) \\ &+ \mathbf{w}(s-1) + 0.9\mathbf{w}^*(s-1) \end{aligned} \tag{8}$$

In order to assess the effect of noise on the performance accuracy, white noise was added so as to vary the signal to noise ratio (SNR) from 0 to 15dB. For rigour, results were obtained for different window lengths. Note that the original data contains 1000 samples and 30 steps ahead were predicted. Although it is assumed that clean signal is not available for prediction purpose, in the case of synthetic data, the performance is still measured with regard to the noise free signal. Fig.1. represents that the A-CSSA provided higher accuracy compared to CSSA. Moreover, as the SNR increases, the benefits of exploiting augmented statistic becomes more prominent, since the rate of decrease in error of A-CSSA is much more significant than that of CSSA.

4.2. Real-world data

The data used in this part were obtained online from the Yahoo finance center[3] and the Iowa Environmental Mesonet (IEM) center recorded for five days at five-minutes intervals[4]. Recently, a novel model represented the wind as a complex-valued vector $\mathbf{f} = \boldsymbol{\vartheta} e^{j\boldsymbol{\theta}}$ where $\boldsymbol{\vartheta}$ is the speed and $\boldsymbol{\theta}$ is the wind direction. Thus, real and imaginary part of the generated $\mathbf{f}$ represent wind speed in east and north directions respectively [7]. Furthermore, two features of the *same* financial stock was used as a complex-valued vector in order to take advantage of the correlation between two features and analyse them simultaneously as one single complex-valued channel data such as $\mathbf{f} = \boldsymbol{\alpha} + j\boldsymbol{\beta}$ where $\boldsymbol{\alpha}$ is the closing price and $\boldsymbol{\beta}$ is the high price of the JPMorgan.
The mentioned datasets were analysed and forecasted using CSSA and A-CSSA. Additionally, forecasting was also performed using ACLMS method[5]. ACLMS is an adaptive algorithm which can be used for prediction of few steps ahead [2].
NRMS error was measured for the overall 1200 samples to evaluate both the reconstruction and forecasting accuracy. Furthermore, NRMS error was calculated for the last 90 predicted points to compare the accuracy of the forecasting stage. Note that NRMS error varies between 0 to 1, where 0 indicates perfect prediction.
Fig. 2. and Fig. 3. represent the original data and the forecasted samples for financial and wind data respectively. It is shown that for both reconstruction and forecasting stages, the proposed method A-CSSA followed the actual trend better than CSSA.
Remark#4: Note that LMS algorithm updates its filter coefficients (w) using $w(s+1) = w(s) + \mu\mathbf{e}(s)\mathbf{f}(s)$ where μ is the positive learning rate, $\mathbf{e}(s)$ is the error and $\mathbf{f}$ is the input [7]. Thus, for inputs with high mean values the influence of the accumulated error $\mathbf{e}(s)$ was magnified through the iterations. For such signals, the ACLMS algorithm diverged after first few prediction steps, while our approach showed robustness against the accumulated error. See Fig. 2. On the other hand, for the signals with smaller mean values, the error accumulation was less harmful. Thus the diverging effect was less pronounced in Fig.3. for the LMS algorithm. Numerical assessments also confirmed this observation; see Table.1. and Table.2.

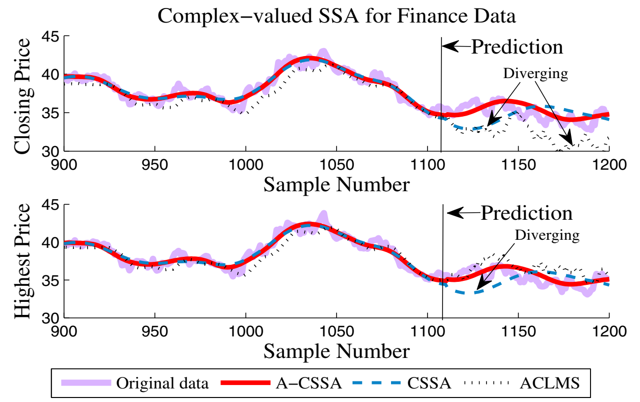

Fig. 2. SSA output for real-world financial data. Results are zoomed out for better visibility. Out performance of the A-CSSA is very significant for both prediction and reconstruction. ACLM forecasting diverged after first few prediction steps.

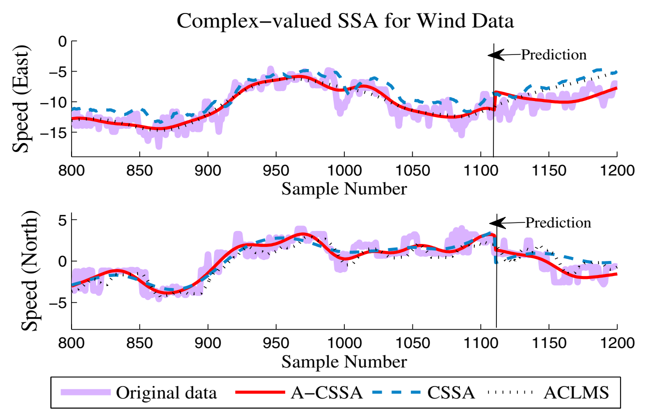

Fig. 3. SSA output for real-world wind data zoomed for the last 400 samples. A-CSSA shows significant advantage in following the trend for the predicted values.

Table 1. Performance Measurements for Financial Data.

	A-CSSA	CSSA	ACLMS
NRMSE Overall	0.03	0.04	0.31
NRMSE Forecast	0.21	0.45	0.61

Table 2. Performance Measurements for Wind Data.

	A-CSSA	CSSA	ACLMS
NRMSE Overall	0.05	0.06	0.15
NRMSE Forecast	0.27	0.43	0.49

5. CONCLUSION

An augmented complex-valued SSA algorithm (A-CSSA) has been introduced to address both second order circular and noncircular data. We have (i) reviewed the usefulness of singular spectrum analysis in the context of trend prediction for large number of steps (90 steps); (ii) shown the performance superiority of A-CSSA over normal CSSA for both reconstruction and prediction purposes. Future work goes on (i) reducing the computational complexity of A-CSSA by replacing the full SVD with so-called Truncated SVD [1], (ii) finding the optimal window length based on the individual datasets.

[3] Financial data available from: "http://finance.yahoo.com/q?s=VOD.L"
[4] Wind data available from: "http://mesonet.agron.iastate.edu/request/awos/1min.php"
[5] ACLMS refers to Augmented Complex Least Mean Square which is comprehensively described in [2]

6. REFERENCES

[1] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of time series structure: SSA and related techniques*, Chapman and Hall, Boca Raton, Florida, 2001.

[2] D. P. Mandic, S. Javidi, S. L. Goh, A. Kuh, and K. Aihara, “Complex-valued prediction of wind profile using augmented complex statistics,” *Renewable Energy*, vol. 34, no. 1, pp. 196 – 201, 2009.

[3] J. Navarro-Moreno, “ARMA prediction of widely linear systems by using the innovations algorithm,” *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3061–3068, 2008.

[4] H. Hassani and A. Zhigljavsky, “Singular spectrum analysis: methodology and application to economics data,” *Journal of Systems Science and Complexity*, vol. 22, no. 3, 2009.

[5] H. Hassani, S. Heravi, and A. Zhigljavsky, “Forecasting european industrial production with singular spectrum analysis,” *International Journal of Forecasting*, vol. 25, no. 1, pp. 103 – 118, 2009.

[6] S. Sanei, T. K. M. Lee, and V. Abolghasemi, “A new adaptive line enhancer based on singular spectrum analysis,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 428–434, 2012.

[7] D. P. Mandic and S. L. Goh, *Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models*, Wiley, 2009.

[8] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press, 2010.

[9] B. Widrow, J. McCool, and M. Ball, “The complex LMS algorithm,” *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719–720, 1975.

[10] C. L. Keppenne and U. Lall, “Complex singular spectrum analysis and multivariate adaptive regression splines applied to forecasting the southern oscillation,” *Exp. Long Lead Forecast Bull*, vol. 5, pp. 54–56, 1996.

[11] C. Cheong Took, S. C. Douglas, and D. P. Mandic, “On approximate diagonalization of correlation matrices in widely linear signal processing,” *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1469–1473, 2012.

[12] S. S. Kozat and A. C. Singer, “Competitive prediction under additive noise,” *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3698–3703, 2009.