International Journal on Recent and Innovation Trends in Computing and Communication
Volume: 5 Issue: 10
ISSN: 2321-8169
41 – 48

Moment Preserving Approximation of Independent Components for the Reconstruction of Multivariate Time Series

Amr Goneid
Department of Computer Science & Engineering,
The American University in Cairo,
Cairo, Egypt
e-mail: goneid@aucegypt.edu

Abstract—The application of Independent Component Analysis (ICA) has found considerable success in problems where sets of observed time series may be considered as results of linearly mixed instantaneous source signals. The Independent Components (IC's) or features can be used in the reconstruction of observed multivariate time seriesfollowing an optimal ordering process. For trend discovery and forecasting, the generated IC's can be approximated for the purpose of noise removal and for the lossy compression of the signals.We propose a moment-preserving (MP) methodology for approximating IC's for the reconstruction of multivariate time series.The methodologyis based on deriving the approximation in the signal domain while preserving a finite number of geometric moments in its Fourier domain.Experimental results are presented onthe approximation of both artificial time series and actual time series of currency exchange rates. Our results show that the moment-preserving (MP) approximations of time series are superior to other usual interpolation approximation methods, particularly when the signals contain significant noise components. The results also indicate that the present MP approximations have significantly higher reconstruction accuracy and can be used successfully for signal denoising while achieving in the same time high packing ratios. Moreover, we find that quite acceptable reconstructions of observed multivariate time series can be obtained with only the first few MP approximated IC's.

Keywords:*Signal Approximation**, Moment-Preserving Approximations,Independent Component Analysis, Time Series Reconstruction*

*****

I. INTRODUCTION

The application of Independent Component Analysis (ICA) has proved to be quite successful in problems where observed multivariate time series may be considered as results of linearly mixed instantaneous source signals [1,2]. This is due to the fact that ICA can generate*independentcomponents*(IC's) that represent the true sources of the observed time series (see survey in [3]). ICA has also shown great promise in trend discovery and forecasting [4,5] of time series through the use of IC's in the reconstruction of observed data. In this reconstruction process, optimal ordering of IC's is requiredand several algorithms have been proposed for such ordering [4,6, 7].

For applications to trend discovery and forecasting, there is a need for approximating the generated IC's for the purpose of noise removal and for the compression of the signals to reduce redundancy.For this purpose, usual approximations in the signal space can be used where the values of the approximating function at selected nodal points are taken to be the same as the original function at those points. Such values can then be joined using some interpolation technique subject to a certain error minimization criterion.However, a more precise approximation method approaches the problem by deriving the approximation in the signal domain while preserving a finite number of geometric moments that are related to its Fourier domain.This moment-preserving (MP) method has been applied to piecewise linear approximations[8] and has also been extended to higher order polynomials for approximating 1-D and 2-D signals [9, 10].

In the present paper, we propose a moment-preserving (MP) method for approximating IC's generated by ICA of multivariate time series. We also investigate the efficiency of applying such approximated IC's for the reconstruction of the observed time series. Experimental results are presented on the MP approximation of artificial 1-D signals and IC's of multivariate time series as well as the reconstruction of actual financial multivariate time series using approximated IC's.

The paper is organized as follows: section II introduces the MP approximation methodology; section IIIgives results on the application of MP approximation to simulated 1-D signals; section IV presents results on applying the MP method to multivariate financial time series; section V presents the method for reconstruction of observed seriesusing MP approximated IC'sand finally section VI presents the summary and conclusion of our work.

II. MOMENT-PRESEVING APPROXIMATION METHODOLOGY

A. General

Consider a function $f(x)$ specified at a finite set of discrete points $\{x_i\ ,\ i = 0\ ,1\ ,\ ..\ N\}$. The objective of an approximation method is to seek an approximating function $g(x)$ defined at a set of distinct nodal points $\{z_j\ ,\ j = 0\ ,1\ ,....,\ M\}$, where generally $M < N$. For a nodal points sampling rate R, the nodal points will be located at $z_j = x_0 + j\ R\ d$, where d is the x sampling interval. In the usual approximation in the signal space, the values of the approximating function at the nodal points are taken to be the same as the original function at those points, i,e., $g(z_j) = f(z_j)$. Such values can then be joined using some interpolation technique subject to a certain error minimization criterion.

A more precise approximation method approaches the problem by deriving the approximation in the signal domain while preserving a finite number of geometric moments that are related to its Fourier domain [9,10]. Such approach can be realized by considering the function *f(x)* to be characterized by a PDF *p(x)* so that its k^{th} geometric moment S_k isgiven by:

$$S_k = E_x[x^k] = \int x^k p(x)\,dx = (-j)^k [d^k \varphi(j\nu)/d\nu^k]_{\nu=0} \qquad (1)$$

The characteristic function $\varphi(j\nu)$ represents the Fourier transform of *p(x):*

$$\varphi(j\nu) = E_x[\exp(j\nu x)] = \int \exp(j\nu x)\, p(x)\,dx \qquad (2)$$

If such characteristic function has a Taylor-series expansion valid in some region about the origin, it is uniquely determined in this interval by the geometric moments since,

$$\varphi(j\nu) = \sum_k [d^k\varphi(j\nu)/d\nu^k]_{\nu=0}\, \nu^k / k! = \sum_k S_k (j\nu)^k / k! \qquad (3)$$

Therefore, the moments S_k do uniquely determine the characteristic function as well as the PDF *p(x)*. It follows that a moment-preserving approximation to the function *p(x)* in the *x*-domain will also serve as an approximation constraint in the ν-domain.

B. Moment-Preserving Approximation

Let$S_k(i,j) = E_x[x^k]_{i,j}bek^{th}$ moment of the variable *x* over the finite interval *(i,j)* of the function *f(x)*. Also, let α be a scale reduction factor so that $x = \alpha y$ and hence we may define a scaled moment as:

$$\sigma_k(i,j) = \alpha^{-(k+1)} S_k(i,j) = \int_{i,j} y^k f(\alpha y)\,dy = E_{\alpha y}[y^k]_{i,j} \qquad (4)$$

With the function *f(x)* specified by a finite set of discrete points $\{x_i\ ,\ i = 0, 1, .. N\}$, the scaled moment σ_k is the sum over all (*N*) segments (*i, i+1*) covering the above domain:

$$\sigma_k = \sum_i \sigma_k(i,i+1) \quad , \; i = 0,1,\ldots,N-1 \qquad (5)$$

On the other hand, if we seek an approximating function *g(x)* defined at a set of distinct nodal points $\{z_j\ ,\ j = 0, 1, \ldots, M\}$, then over the interval between two nodal points *(p,q)* we obtain scaled moments $\mu_k(p,q)$ whose sum over the nodal intervals gives the scaled moments μ_k.For the moment preserving approximation, we require that:

$$\sigma_k = \mu_k \qquad for\ \ k = 0,1,..,M \qquad (6)$$

The above moment-preserving constraint leads to a system of *M+1* equations:

$$\boldsymbol{\sigma} = \boldsymbol{E}\,.\,\boldsymbol{G} \qquad (7)$$

where ***E***is an *M+1* by *M+1* square matrix of coefficients depending on the approximating polynomial, and ***G***is a column vector representing the approximations $g(z_j)$ to the function *f(x)* at the nodal points $\{z_j\ ,\ j = 0, 1, \ldots, M\}$.

C. Moment-Preserving using Quadratic Approximation

Following our previous work [10], we present here the method for estimating the moment-preserving values of the approximations $g(z_j)$ to the function *f(x)* at the *M* +1 nodal points using a quadratic approximation. For this purpose, we assume that between the nodal points $z_p$and z_q the function is piecewise quadratic, and we use equally spaced nodal points with an internal point z_r between the points $z_p$and z_q. In this case we may use Lagrange's classical formula to obtain the k^{th} scaled moment over that region:

$$\mu_k(p,q) = \int_{p,q} y^k g_{pq}(\alpha y)\,dy \qquad (8)$$

$$= g(\alpha y_p)\,B_{p,q}(r,q,k,y) - 2g(\alpha y_r)\,B_{p,q}(p,q,k,y) + g(\alpha y_q)\,B_{p,q}(p,r,k,y)$$

where,

$$B_{p,q}(i,j,k,t) = [2/D^2_{pq}(t,0)]\{\,D_{pq}(t,k+2) - (t_i + t_j)\,D_{pq}(t,k+1) + t_i\, t_j\, D_{pq}(t,k)\,\}$$

and,

$$D_{pq}(t,n) = \int_{p,q} t^n dt = (t_q^{\,n+1} - t_p^{\,n+1})/(n+1)$$

Following the work in [10], the total scaled moment can be expressed as a vector $\boldsymbol{\mu}$with elements:

$$\mu_k = \sum_j C_j(k,y)\, g(\alpha\, y_j) \quad , \; j, k = 0,1,\ldots,M \;, M\ even \qquad (9)$$

where

$$Cj(k,y) = \begin{cases} B_{0,2}(1,2,k,y) & for\ \ j = 0, \\ -2\,B_{j-1,j+1}(j-1, j+1,k,y) & for\ \ j\ odd, \\ B_{j-2,j}(j-2, j-1,k,y) + B_{j,j+2}(j+1, j+2,k,y) & for\ \ j\ \ even, \\ B_{m-2,m}(M-2, M-1,k,y) & for\ \ j = M \end{cases}$$

Notice that in the above equations, the values of y_j represent the scaled coordinates of the nodal points. When the scaled coordinates of the actual function points are used, then we obtain the actual scaled moments vector $\boldsymbol{\sigma}$. Accordingly, moment preservation ($\boldsymbol{\mu} = \boldsymbol{\sigma}$) leads to the system:

$$\boldsymbol{G} = \boldsymbol{E}^{-1}.\,\boldsymbol{\sigma} \qquad (10)$$

where the elements of the square matrix ***E*** are given by $e(k,j) = C_j(k,y)$.

III. EXPERIMENTS ON 1-D SIGNAL APPROXIMATION

A. The Simulated 1-D Signals

For verification of the above moment-preserving (MP) method, we have applied it to obtain piecewise approximations for 1-D signals *f(x)*for which the nodal points $\{z_j,\ j = 0, 1, \ldots, M\}$were chosen to be evenly spaced across the *x*-space. The vector of approximants ***G***at those points was computed using the quadratic method outlined above to obtain an approximation *g(x)* to the function. For comparison, an approximation *h(x)* was also obtained using the original function values $f(z_j)$at the selected nodal points.We have used the mean-squared error (MSE) as a measure of the error norm between *f(x)* and each of the approximations *g(x)* and *h(x).*

For the piecewise approximation, a signal sequence *f(x)* of length N_T is divided into $n_b$blocks of equal length. In a given

 ISSN: 2321-8169
41 – 48

block, $n_k = M+1$ nodal points are selected such that their sampling rate is $R = N_T/(n_b M)$. A block will therefore contain $n_f=N+1$ function points located at x_i , $i = 0 ,1 , .. N$, with $x_i = x_0 + i\, d$, where d is the x sampling interval .and $N = R\, M – 1$. Of these function points, the M +1 nodal points will be located at z_j , $j = 0 ,1 ,...., M$, with $z_j = x_0 + j\, R\, d$.

As an example, we have used a simulated series $f(x)$ given by:

$$f(x) = 2sin(0.2\, x) + 5cos(0.3\, x) + w\, r \qquad (11)$$

where $r$represents an additive random noise and w is an amplitude factor.

B. The MP Approximation Algorithm

The algorithm for obtaining the MP approximations at the nodal points is summarized in the following steps:

- Input sequence $f(x)$ of length N_T
- Select number of blocks $n_b$and number of nodal points per block $n_k = M+1$
- Compute nodal points sampling rate R and the number $n_f=N+1= R\, M$of function points per block
- Set the starting point x_0 and the x sampling interval (d)
- Repeat for each block:
  1. Obtain locations of nodal points z_j , $j = 0,1,...., M$ in the block
  2. Obtain $f(x_i)$ at x_i for $i = 0 ,1 , .. N$
  3. Set a scale reduction factor α (e.g. $\alpha = x_N$)
  4. Compute scaled nodal point locations y_j relative to start of block
  5. Compute scaled moments vector $\boldsymbol{\sigma}$ at the n_k nodal points and the $\boldsymbol{E}$ matrix ($n_k x n_k$)
  6. Compute vector of MP values $G = E^{-1} * \sigma$ at the nodal points
  7. Obtain vector of MP approximated $f(x)$ in block by quadratic interpolation
  8. Set start of next block
- Join blocks to obtain final MP approximated $g(x)$

C. Experimental Results

The above algorithm has been applied to the 1-D signal example given by the function (11). For that example, we have used an x-domain with 1152 function points at a sampling interval d = 0.03. For more accuracy and to reduce the need for reconditioning the matrices in the inversion process, we have used a scale factor $\alpha = x_N$ over the block.

To illustrate how MP approximations differ from original function values at the nodal points, Fig. 1 shows examples of the results obtained for n_b = 12 blocks and n_k = 5 nodal points per block, yielding a nodal points sampling rate R = 24 (i.e. one nodal point every 24 function points). Fig. 1a and Fig. 1b show the results for blocks 5 and 12, respectively, using a noise factor w = 0.8.

In these figures, the function $f(x)$ is shown for the given block together with the MP approximated values (circles) and the original function values (*) at the selected nodal points.

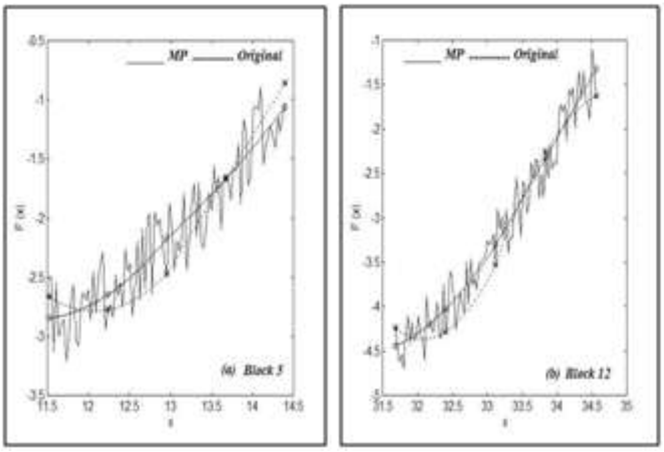

Figure 1. Function approximation for (a) block 5 and (b) block 12 (o: MP function values, *: Original function values)

It is clear from the examples shown in Fig. 1 that the MPapproximated values differ from those values of the original function at the same nodal points. The figure also shows the results of quadratic interpolation of the MP values ($g(x)$ approximation, solid curve) and the original function values ($h(x)$ approximation, dotted curve), again highlighting the difference between these two approximations.

To examine the effect of the number of nodal points on the piecewise approximations, we have computed the Mean Square Error (MSE) over the whole series between the original signals $f(x)$ and the corresponding MP approximation $g(x)$ using different numbers of nodal points per block $n_k = M+1$. Fig. 2 shows a plot of the MSE over all blocks against n_k. For comparison, the figure also shows the MSE for the corresponding approximation $h(x)$ that uses the original function values at the nodal points.

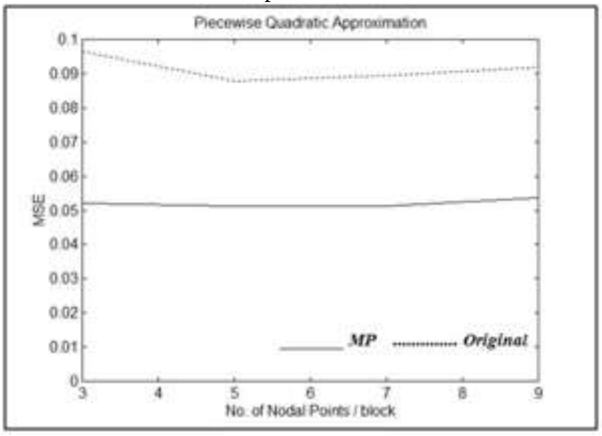

Figure 2. Mean Square Error Vs No. of nodal points / block.

The figure shows that the MSE is almost constant for the MP approximation $g(x)$over n_k = 3 – 9, and that the MSE for $g(x)$ is about 57% only of the corresponding MSE for the approximation $h(x)$obtained using the original function values.

We have also examined the effect of noise on the approximations $g(x)$ and $h(x)$ with the results shown in Fig. 3.

ISSN: 2321-8169
41 – 48

In the calculations shown in this figure, we have used n_k = 7 nodal points / block corresponding to a nodal point sampling rate of R = 16.

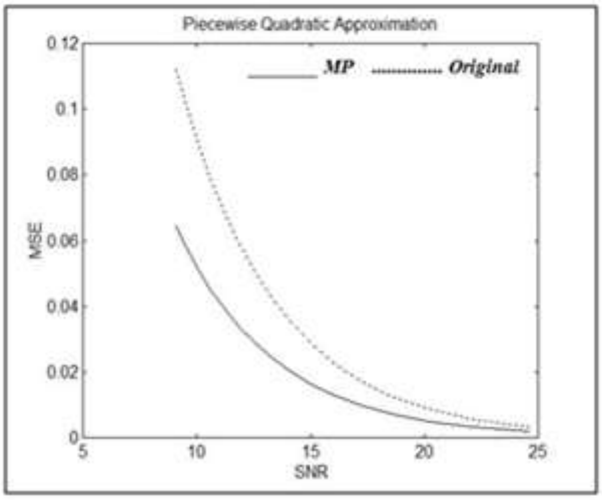

Figure 3. Mean Square Error Vs Signal-to-Noise Ratio.

As we might expect, the MSE drops with increasing SNR. Moreover, the MSE for the MP approximation *g(x)* is significantly lower than that for the original function values approximation *h(x)*, particularly in the presence of significant noise (low SNR). Only in the case of noise-free signals that we might expect the two approximations to give low and close values for the MSE. Similar patterns are obtained for n_k = 3 or 5 or 9.

From the results shown in Fig, 1 to Fig. 3, we might conclude that the moment-preserving (MP) approximations of 1-D signals are superior to other interpolation approximation methods, particularly when the signals contain significant noise components. It follows that MP approximations have significantly higher reconstruction accuracy and can be used successfully for signal denoising while achieving in the same time high packing ratios *R*.

IV. EXPERIMENTS ON MULTIVARIATEFINANCIAL TIME SERIES APPROXIMATION

A. *The Observed Financial Time Series*

We have conducted a set of experiments on the MP approximation of independent components derived from actual multivariate time series representing financial data. For this purpose, 6 foreign exchange rate series were selected representing USD versus Brazilian Real (BRL), Canadian Dollar (CAD), Danish Krone (DKK), Japanese Yen (JPY), Swedish Krona (SEK), and Swiss Franc (CHF) in the period from January 4, 2010 till December 31, 2015. The dataset size was 6 time series over 1504 days collected from different historical exchange rates data sources such as [11, 12, 13]. Fig. 4 shows these time series.

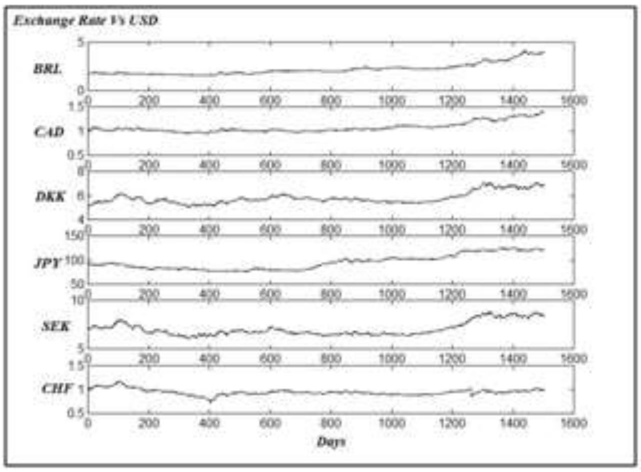

Figure 4. Exchange rate time series (January 4, 2010 - December 31, 2015)

B. *Computing the Independent Components*

We consider the above observed series $\boldsymbol{X}$ to be a linear mixture of independent sources $\boldsymbol{S}$ related to $\boldsymbol{X}$ by an instantaneous linear noiseless mixing model represented by:

$$\boldsymbol{X} = \boldsymbol{A}\,\boldsymbol{S} \qquad (12)$$

where $\boldsymbol{S}$is a random matrix of hidden sources with mutually independent components (IC's), and $\boldsymbol{A}$ is a non-singular mixing matrix. Given $\boldsymbol{X}$, the basic problem in Independent Component Analysis (ICA) is to find an estimate $\boldsymbol{Y}$ of the IC's $\boldsymbol{S}$ and the mixing matrix $\boldsymbol{A}$ such that:

$$\boldsymbol{Y} = \boldsymbol{W}\,\boldsymbol{X} = \boldsymbol{W}\,\boldsymbol{A}\,\boldsymbol{S} = \boldsymbol{G}\,\boldsymbol{S} \approx \boldsymbol{S} \qquad (13)$$

where $\boldsymbol{W} = \boldsymbol{A}^{-1}$ is the *unmixing* matrix, and $\boldsymbol{G} = \boldsymbol{W}\,\boldsymbol{A}$ is usually called the Global Transfer Function or Global Separating-Mixing (GSM) Matrix. The linear mapping $\boldsymbol{W}$ is such that the unmixed signals $\boldsymbol{Y}$ are statistically independent. However, the sources are recovered only up to scaling and permutation. In practice, the estimate of the unmixing matrix $\boldsymbol{W}$ is not exactly the inverse of the mixing matrix $\boldsymbol{A}$. Hence, the departure of $\boldsymbol{G}$ from the identity matrix $\boldsymbol{I}$ can be a measure of the error in achieving complete separation of sources.

For computing the independent components $\boldsymbol{Y}$ from the observed mixtures $\boldsymbol{X}$, we adopt the modified algorithm given by [14] which is based on the Fast ICA algorithm originally given by [15]. Basically, the algorithm uses a fixed-point iteration method to maximize the negentropy using a Newton iteration method. Details of applying this algorithm to financial series are given in a previous paper [7]. For the present financial series, it was found necessary to choose the appropriate non-linearity for the fast ICA algorithm. Analysis of the normalized kurtosis for the given series $\boldsymbol{X}$ showed that they represent a mixture between super-gaussian and sub-gaussian signals. We found that the algorithm would converge in fewer number of iterations if we use a mixture non-linearity derived from a bimodal Exponential Power Distribution (EPD) symmetric mixture density [7, 16].

C. *Application of MP Approximation to Noisy IC's*

In the present application of MP approximations to the IC;s obtaied by the above mentioned Fast ICA algorithm, we have examined the effect of the number of nodal points on the piecewise approximations. We have computed the Mean Square Error (MSE) between the original independent components *Y* and their corresponding MP approximations *g(x)* using different numbers of nodal points per block, as well as the corresponding approximations *h(x)* that use the original function values at the nodal points. In these computations, we have used a total of1440 points/IC divided into 15 blocks with 96 function points/block.

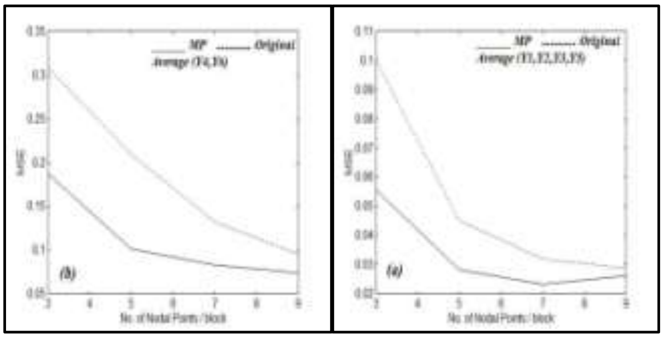

Fig 5. Mean Square Error Vs No. of nodal points / block. (a) Average MSE for Y1,Y2,Y3,Y5 and (b) Average MSE for Y4 and Y6

Fig. 5 shows a plot of the MSE over all blocks against n_k for noise free IC's. The results shown in Fig. 5a relate to the average MSE over the IC's *Y1,Y2,Y3* and *Y5* while those shown in Fig. 5b are for the averages over IC's *Y4* and *Y6*. As in the case of 1-D simulated signals, these results show that the MSE is almost constant for the MP approximation *g(x)*over n_k = 5 – 9, and that the MSE for *g(x)* is significantly lower than the corresponding MSE for the approximation *h(x)*obtained using the original function values.

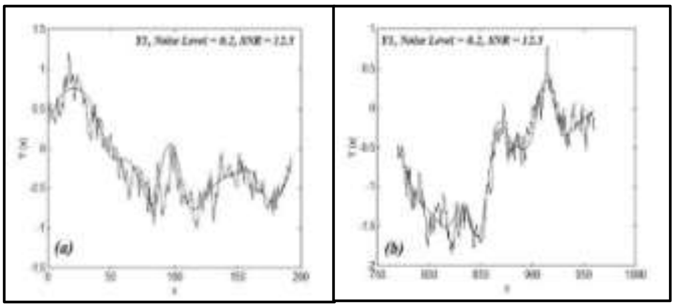

Fig 6. IC *Y1* MP approximation using random noise w = 0.2 (SNR = 12.5)

To investigate the effect of random noise on the MP approximations of IC's, we have simulated the presence of noise in the resulting dataset ***Y*** by adding a noise component w r, where r is a uniform random number {-1,+1} and w is the noise contribution weight. Fig. 6 shows examples of the results obtained for n_b = 15 block and n_k = 7 nodal points per block, yielding a nodal points sampling rate *R* = 16 (i.e. one nodal point every 16 function points).

Fig. 6a and Fig. 6b show the results for the MP approximations for IC *Y1* for two different regions of the time series using a noise factor *w* = 0.2 (SNR = 12.5). The results shown in these figures indicate that MP approximations can effectively denoise the IC's for more effective reconstruction processes. Fig. 7a and Fig. 7b show similar examples of denoising using a higher noise content (w = 0.4, SNR = 6.57).

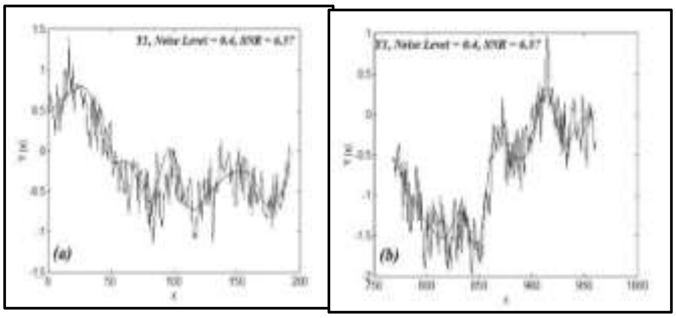

Fig 7. IC *Y1* MP approximation using random noise w = 0.4 (SNR = 6.57)

In the presence of noise, the MP approximation is expected to denoise the signals since it preserves the moments in the frequency domain. Fig. 8 shows examples of comparison of the ability of denoising IC's between MP approximations *g(x)* and the corresponding approximations *h(x)* that use the original function values at the nodal points using n_k = 7 nodal points per block. The figure shows that the MSE drops with increasing SNR and that the MSE for the MP approximation *g(x)* is significantly lower than that for the original function values approximation *h(x)*, particularly in the presence of significant noise (low SNR). Only in the case of noise-free signals that we might expect the two approximations to give low and close values for the MSE. Similar patterns are obtained for n_k = 3 or 5 or 9.

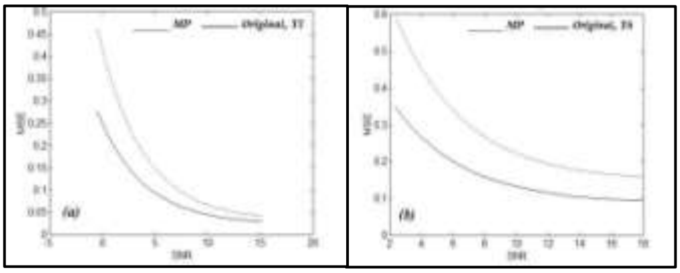

Fig 8. Mean Square Error Vs Signal-to-Noise Ratio using MP approximation *g(x)* (solid curve) and original values *h(x)*: (a) Approximations for IC *Y1*, (b) Approximation for IC *Y6*

V. RECONSTRUCTION OF FINANCIAL TIME SERIES USING MP-APPROXIMATED IC'S

A. *Methodology*

There is a significant interest in the process of reconstruction of observed multivariate time series for trend discovery and forecasting [4,5]. Using IC's derived from ICA of the series has proved to be quite efficient in the reconstruction process [7]. In the present work, we add the advantage of using Moment-Preserving approximations (MP) of the IC's instead of the original ones. The basic advantages are that MP approximations act effectively in denoising the IC series and they also allow for the compression of the IC series by the factor R expressing the sampling rate of the nodal points.

Consider the observed k time series $\boldsymbol{X} = x(t) = [x_1(t),\ldots x_k(t)]^T$, $1 \le t \le N$ and let Y_j, $j = 1\ldots k$ be their MP approximated independent components (IC's). To reconstruct the observed time series $\boldsymbol{X}$ from these IC's, it is necessary to determine for each series $x_i(t)$ a list L_i of independent components indices in descending order of the dominance of the given IC's in the corresponding reconstructed series. The process of reconstructing time series from the estimated independent components can then be done by summing their contributions in the order given by the list L_i. Following [4], [7], the contribution may be expressed as:

$$u(i,j,t) = W^{-1}(i,j)Y_j(t)\,, \quad 1 \le j \le k \qquad (14)$$

In the above equation, W represents the demixing matrix derived from ICA of the observed series and $W^{-1}(i,j)$ is the *(i,j)th* element in the inverse of the W matrix.

Investigation of the different methods to obtain optimal ordering lists from the contributions (14) has been done in detail in [7]. In such work, it is found that three methods give almost the same ordering:

1. ES: This is basically an exhaustive search method in which the optimal list is determined by performing$k!$ reconstruction steps for each series and selecting the list with the least error profile.
2. EL: This method uses a strategy of excluding the least contributing IC first. This method first selects from the set of k IC's the component that when excluded from the list will minimize the reconstruction error. This component is then removed from the set of IC's and its index becomes the last in the order list L_i. The process is repeated on what remains in the component set to select the second-last in the order list, and so on. This algorithm will involve $k(k+1)/2 - 1$reconstruction steps for each time series.
3. ME: This method involves minimizing the reconstruction error for Individual IC's Contribution. Given a certain error measure, the list is obtained by sorting in ascending order the error between individual contributions $u(i,j,t)$ and the observed series x_i. This method involves only k reconstruction steps for each time series.

In the present work, we adopt the third method for obtaining optimal ordering lists due to its linear complexity. We compute the reconstructed version of the time series $x_i(t)$ using the first m independent components in the optimal list L_i by summing the contributions of the individual components. Such sum is given by:

$$\hat{x}_i^{\,m}(t) = \sum_{s=1}^{m} W^{-1}(i,s)Y_s(t) \qquad (15)$$

where (s) denotes the s[th] element of L_i..Also, we use the Mean Square Error (MSE) for the overall reconstruction error for a given series x_i using the first m independent components in the optimal list L_i as:

$$Q_i(m) = \frac{1}{N}\sum_{t=1}^{N}[\hat{x}_i^{\,m}(t) - x_i(t)]^2 \qquad (16)$$

B. Reconstruction Results

The ME method was usedto obtain optimal ordering lists which involves only $k = 6$ reconstruction steps for each time series. For such method, we follow the work of [7] in using the MSE as an error measure since it has lowest error profile. Using this method, we obtain the set of ordered lists for the 6 exchange rate time series $\boldsymbol{X}$ as given in Table (1).

TABLE 1. Obtained ordered lists

Series	Label	Ordered List					
BRL	X_1	3	1	5	4	2	6
CAD	X_2	3	1	5	6	4	2
DKK	X_3	3	1	4	6	5	2
JPY	X_4	3	5	1	4	6	2
SEK	X_5	3	1	2	5	4	6
CHF	X_6	2	6	5	1	4	3

Calculations have also been made for the percentage cumulative contribution of the first m MP approximated IC's from the lists to the reconstruction of the exchange rate time series. These contributions are calculated as 1 $-MSE(x_i, \hat{x}_{L_i}^m)$ / $variance(x_i)$ = 1- $Q(m)$, since the series have zero mean and unit variance. Fig. 9 shows the average of these contributions over the 6 observed time series.

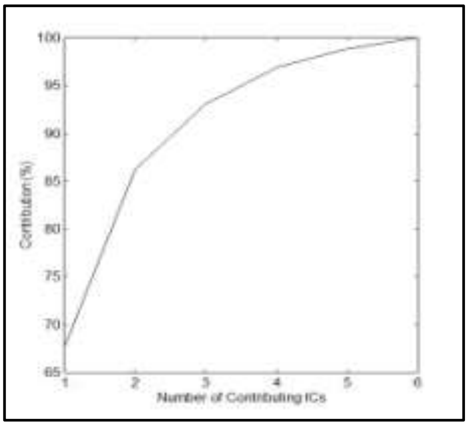

Fig 9. Average percentage cumulative contribution of the first mMP approximated IC's from the lists to the reconstruction of the exchange rate time series.

The results shown in Fig. 9 indicate that it is possible to reconstruct the general trends in most of the observed exchange rate series considered here using only the first 2 MP approximated IC' (contribution ≈86 %). This is evident from the results shown in Fig. 10 where such reconstructed series are compared with the corresponding observed ones for three different exchange rate series.

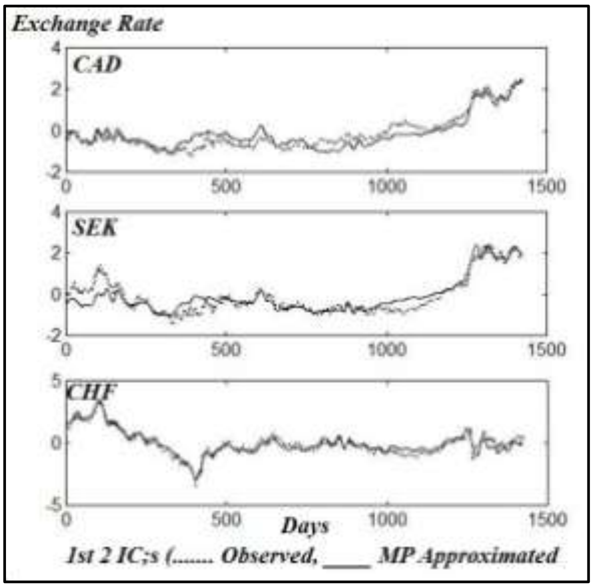

Fig. 10. Comparison between Observed Series (dotted curves) and Reconstructed Series (solid curves) using first 2 MP approximated IC's

Fig. 9 also indicates that the observed series can be reconstructed to an excellent degree using the first 4 MP approximated IC's in their respective ordered lists (contributions ≈ 97 %). For the majority of series, quite acceptable reconstructions (≈ 93 %) can also be obtained with only the first 3 IC's in the lists.

For a comparison between the observed exchange rate time series and those reconstructed using MP approximated IC's, weshow in Fig. 11a the results of such comparison using the first three IC's in the corresponding lists given in Table (1).

Similar results are shown in Fig. 11b for reconstructions using the first four MP approximated IC's. Notice that the series compared in these figures have zero mean and unit variance as obtained from the preprocessing of the data for the ICA algorithm.

It can be seen from these figures that the reconstruction of observed series using MP approximated IC's successfully preserves the general trends of the series. Moreover, quite acceptable matching can be realized with only the dominant 3 or 4 IC's in the lists.

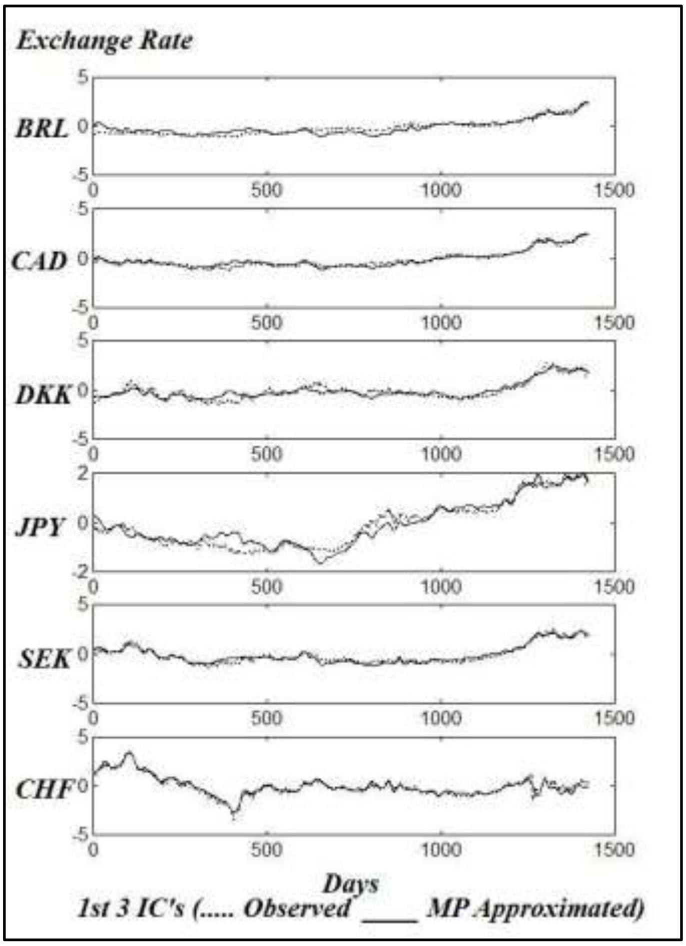

Fig 11a. Comparison between Observed Exchange Rate Series (dotted curves) and Reconstructed Series using the first 3 MP approximated IC's (solid curves)

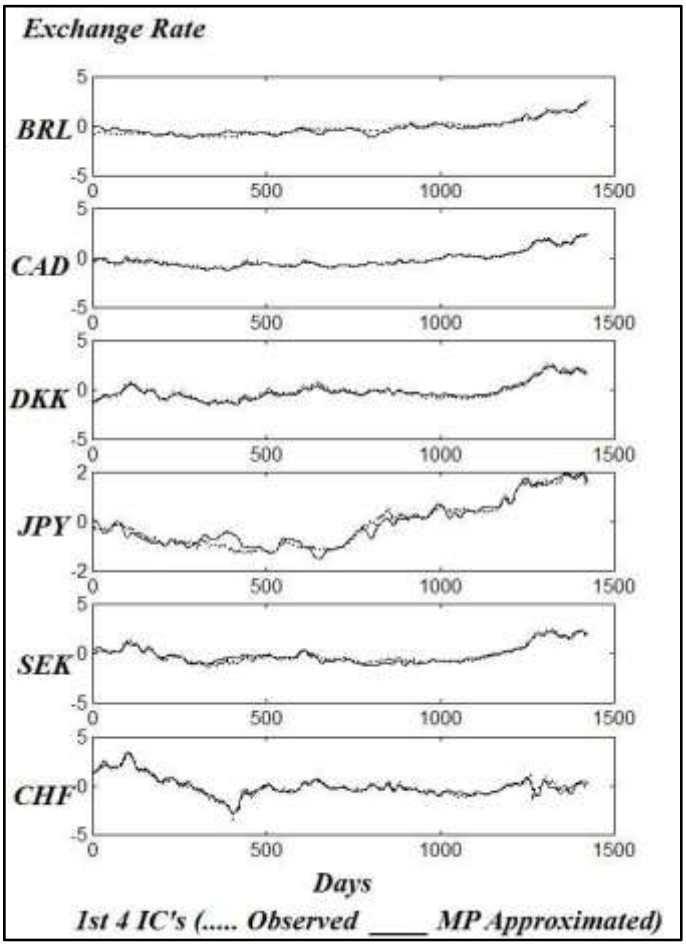

Fig 11b. Comparison between Observed Exchange Rate Series (dotted curves) and Reconstructed Series using the first 4 MP approximated IC's (Solid Curves)

VI. SUMMARY AND CONCLUSIONS

The present paper proposes a moment-preserving (MP) method for approximating time series for trend discovery and forecasting applications. The method is based onderiving the approximation in the signal domain while preserving a finite number of geometric moments that are related to its Fourier domain. MP values of the time series are derived at selected nodal points using a quadratic approximation. Piece-wise approximation is conducted on blocks of the series having equal numbers of nodal points in a block.

The proposed MP method has been applied to a 1-D simulated series and was compared with the usual approximation method in which the original values of the function at the nodal points were used in the approximation process. In general, the MP values were found to differ significantly from the original function values at the same nodal points leading to a better approximation of the time series.

We have also investigated the effect of the number of nodal points in a block and the presence of noise componentin the simulated series using the Mean Square Error (MSE) between the observed and approximated series. The results show that the MSE is almost constant for the MP approximation when using 3 – 9 nodal points per block, corresponding to packing ratios 48 – 12 and that such MSE is only about 57% of the corresponding MSE for the approximation obtained using the original function values.

Moreover, investigation of the effect of random noise on the approximation efficiency has shown that the MSE for the MP approximation is significantly lower than that for the original function values approximation, particularly in the presence of significant noise (low SNR). Only in the case of noise-free signals that we might expect the two approximations to give closer values for the MSE.

From the above results, we may conclude that the moment-preserving (MP) approximations of 1-D signals are superior to other interpolation approximation methods, particularly when the signals contain significant noise components. It follows that MP approximations have significantly higher reconstruction accuracy and can be used successfully for signal denoising while achieving in the same time high packing ratios.

The present MP approximation method has also been applied to IC's derived from ICA of actual financial multivariate time series of currency exchange rates. As in the case of 1-D simulated signals, the results for the financial series show that the MSE for the MPapproximated IC's is significantly lower than the corresponding error for the approximation obtained using the original function values. Also, results obtained with noisy IC's indicate that, for more effective reconstruction processes, the MP approximations can more effectively denoise the IC's compared to the usual approximation methods that use the original function values. This is because the MP method also preserves the geometric moments inthe frequency domain.

For trend discovery and forecasting applications, we have used the MP approximated IC's in reconstructing the observed financial series considered here. Using an efficient method for obtaining optimal ordered lists of dominant IC's, we were able to compute the percentage contribution of cumulative IC's from such lists. Comparison of the reconstructed series using the MP approximated IC's with the observed financial series shows that it is possible to reconstruct the general trends in most of the observed series using only the first 2 MP approximated IC's (contribution ≈ 86 %). Moreover, the observed series can be reconstructed in this way to quite acceptable (~93 %) and excellent (~ 97 %) degrees using the first 3 and 4 IC's, respectively.

REFERENCES

[1] A. Hyvärinen, J. Karhunen, and E. Oja, "Independent Component Analysis", John Wiley and Sons , New York , NT, 2001
[2] A. Hyvärinen, "Independent component analysis: recent advances",Phil. Trans.R. Soc. A 371: 20110534 http://dx.doi.org/10.1098/rsta.2011.0534, 2013
[3] A. Mansour, and M. Kawamoto, "ICA papers classified according to their applications and performances", IEICE Trans. Fundamentals, Vol. E86-A, No. 3, pp. 620-633, 2003
[4] Cheung, Yiu-ming, and Lei Xu. "Independent component ordering in ICA time series analysis", Neurocomputing, Vol. 41.1, pp. 145-152, 2001
[5] Chi-Jie Lu, Tian-Shyug Lee and Chih-Chou Chiu, "Time series forecasting using independent component analysis and support vector regression", Decision Support Systems, Vol. 47-2, pp. 115-125, 2009
[6] A. Kamel, A. Goneid, and D. Mokhtar, "Ordering of dominant independent components in time series analysis using fast ICA algorithm", Egyptian Computer Science Journal (ISSN-1110-2586), Vol. 41–2, pp. 1-10, 2017 http://ecsjournal.org/Archive/Volume41_Issue2.aspx
[7] A. Goneid, and A. Kamel, "Reconstruction of time series using optimal ordering of ICA components", International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), ISSN: 2321-8169, Volume 5 Issue 7, PP: 297 – 305, July 2017
[8] T.B. Nguyen and B.J. Oommen, "Moment-Preserving piecewise linear approximations of signals and images", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19(1), pp. 84-91, 1997
[9] A. Goneid, and S. AbuSeif, "Moment-Preserving piecewise approximations for 1-D and 2-D signals", Proc. Int. Conf. on Computer, Communication and Control Technologies (CCCT'03), Orlando, Florida, USA, July 31 – Aug 2, 2003, Vol. 4
[10] A. Goneid, "Moment-Preserving methods for signal piecewise approximation and denoising", Egyptian Computer Science Journal, (ISSN-1110-2586), Vol. 36 (2), pp. 18 – 32, 2012
[11] www.bankofengland.co.uk/boeapps/iadb/
[12] www.oanda.com/solutions-for-business/historical-rates
[13] www.xe.com/currencytables/
[14] A. Kamel, A. Goneid, and D. Mokhtar, "Ordering of dominant independent components in time series Analysis using fast ICA algorithm", Egyptian Computer Science Journal (ISSN-1110-2586), vol. 41–2, pp. 1-10, 2017 http://ecsjournal.org/Archive/Volume41_Issue2.aspx
[15] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis", *IEEE Trans. on Neural Networks*, vol. 10(3), pp. 626-634, 1999
[16] A. Goneid, A. Kamel, and I. Farag, "Generalized mixture models for blind source separation", Egyptian Computer Science Journal, (ISSN-1110-2586), vol. 34-1, pp. 1-14, 2010