

Denoising Time Series by way of a Flexible Model for Phase Space Reconstruction

Minhazul Islam Sk, Arunava Banerjee

University of Florida, Gainesville, Florida

Abstract. We present a denoising technique in the domain of time series data that presumes a model for the uncorrupted underlying signal rather than a model for noise. Specifically, we show how the non-linear reconstruction of the underlying dynamical system by way of time delay embedding yields a new solution for denoising where the underlying dynamics is assumed to be highly non-linear yet low-dimensional. The model for the underlying data is recovered using a non-parametric Bayesian approach and is therefore very flexible. The proposed technique first clusters the reconstructed phase space through a Dirichlet Process Mixture of Exponential density, an infinite mixture model. Phase Space Reconstruction is accomplished by time delay embedding in the framework of Taken's Embedding Theorem with the underlying dimension being determined by the False Neighborhood method. Next, an Infinite Mixtures of Linear Regression via Dirichlet Process is used to non-linearly map the phase space data points to their respective temporally subsequent points in the phase space. Finally, a convex optimization based approach is used to restructure the dynamics by perturbing the phase space points to create the new denoised time series. We find that this method yields significantly better performance in noise reduction, power spectrum analysis and prediction accuracy of the phase space.

1 Introduction

Noise is a high dimensional dynamical process which limits the extraction of quantitative information from experimental time series data. Successful removal of noise from time series data requires a model either for the noise or for the dynamics of the uncorrupted time series. For example, in wavelet based denoising methods for time series [14, 20], the model for the signal assumes that the expected output of a forward/inverse wavelet transform of the uncorrupted time series is sparse in the wavelet coefficients. In other words, it is presupposed that the signal energy is concentrated on a small number of wavelet basis elements; the remaining elements with negligible coefficients are considered noise. Hard-threshold wavelet [25] and Soft-threshold wavelet [4] are two widely known noise reduction methods that subscribe to this model. Principal Component Analysis, on the other hand, assumes a model for the noise: the variance captured by the least important principal components. Therefore, denoising is accomplished by dropping the bottom principal components and projecting the data onto the remaining components.

In many cases, the time series is produced by a low-dimensional dynamical system. In such cases, the contamination of noise in the time series can disable measurements

of the underlying embedding dimension [12], introduce extra Lyapunov Exponents [2], obscure the fractal structure [9] and limit prediction accuracy [5]. Therefore, reduction of noise while maintaining the underlying dynamics generated from the time series is of paramount importance.

A widely used method in time series denoising is Low-pass filtering. Here noise is assumed to constitute all high frequency components without reference to the characteristics of the underlying dynamics. Unfortunately, low pass filtering is not well suited to non-linear chaotic time series [23]. Since the power spectrum of low-dimensional chaos resembles a noisy time series, removal of the higher frequencies distorts the underlying dynamics, thereby, adding fractal dimensions [15].

In this article, we present a phase space reconstruction based approach to time series denoising. The method is founded on Taken’s Embedding Theorem [21], according to which a dynamical system can be reconstructed from a sequence of observations of the output of the system (considered, here, the time series). This respects all properties of the dynamical system that do not change under smooth coordinate transformations.

Informally stated, the proposed technique can be described as follows: Consider a time series, $x(1), x(2), x(3), \dots$ corrupted by noise. We first reconstruct the phase space by taking time delayed observations from the noisy time series (for example, $\langle x(i), x(i+1) \rangle$ forms a phase space trajectory in 2-dimensions). The minimum embedding dimension (i.e., number of lags) of the underlying phase space is determined via the False Neighborhood method [11], as detailed in Sec. 2.1. Next, we cluster the phase space non-parametrically without imposing any constraints on the number of clusters. Finally, we apply a non-linear regression to approximate the temporally subsequent phase space points for each point in each cluster via a Non-parametric Bayesian approach. Henceforth, we refer to our technique by the acronym NPB-NR, standing for non-parametric Bayesian approach to noise reduction in Time Series.

To elaborate, the second step clusters the reconstructed phase space of the time series through an Infinite Mixture of Gaussian distribution via Dirichlet Process [7]. We consider the entire phase space to be generated from a Dirichlet Process mixture (DP) of some underlying density [6]. DP allows the phase space to choose as many clusters as fits its dynamics. The clusters pick out small neighborhoods of the phase space where the subsequent non-linear approximation would be performed. As the latent underlying density of the phase space is unknown, modeling this with an Infinite mixture model allows NPB-NR to correctly find the phase space density. This is because of the guarantee of posterior consistency of the Dirichlet Process Mixtures under Gaussian base density [18]. Therefore, we choose the mixing density to be Gaussian. The posterior consistency acts as a frequentist justification of Bayesian methods—as more data arrives, the posterior density concentrates on the true underlying density of the data.

In the third step, our goal is to non-linearly approximate the dynamics in each cluster formed above. We use a DP mixture of Linear Regression to non-linearly map each point in a cluster to its image (the temporally subsequent point in the phase space). In this Infinite Mixtures of Regression, we model the data in a specific cluster via a mixtures of local densities (Normal density with Linear Transformation of the covariates (βX) as the Mean). Although the mean function is linear for each local density, marginalizing over the local distribution creates a non-linear mean function. In addition,

the variance of the responses vary among mixture components in the clusters, thereby varying among covariates. The non-parametric model ensures that the data determines the number of mixture components in specific clusters and the nature of the local regressions. Again, the basis for the infinite mixture model of linear regression is the guarantee of posterior consistency [22].

In the final step, we restructure the dynamics by minimizing the sum of the deviation between each point in the cluster and its pre-image (previous temporal point) and post-image (next temporal point) yielded by the non-linear regression described above. To create a noise removed time series out of the phase space, readjustment of the trajectory is done by maintaining the co-ordinates of the phase space points to be consistent with time delay embedding.

We demonstrate the accuracy of the NPB-NR model across several experimental settings such as, noise reduction percentage and power spectrum analysis on several dynamical systems like Lorenz, Van-der-poll, Buckling Column, GOPY, Rayleigh and Sinusoid attractors, as compared to low pass filtering. We also show the forecasting performance of the NPB-NR method in time series datasets from various domain like the ‘‘DOW 30’’ index stocks, LASER dataset, Computer Generated Series, Astrophysical dataset, Currency Exchange dataset, US Industrial Production Indices dataset, Darwin Sea Level Pressure dataset and Oxygen Isotope dataset against some of its competitors like GARCH, AR, ARMA, ARIMA, PCA, Kernel PCA and Gaussian Process Regression.

2 Mathematical Background

2.1 Time Delay Embedding and False Neighborhood Method

Time Delay Embedding has become a common approach to reconstruct the phase space from an experimental time series. The central idea is that the dynamics is considered to be governed by a solution traveling through a phase space and a smooth function maps points in the phase space to the measurement with some error. Given a time series of measurements, $x(1), x(2), \dots, x(N)$, the phase space is represented by vectors in D -dimensional Euclidean space.

$$y(n) = \langle x(n), x(n+T), \dots, x(n+(D-1)T) \rangle \quad (1)$$

Here, T is the time delay and D is the embedding dimension. The temporally subsequent point to $y(n)$ in the phase space is $y(n+1)$. The purpose of the embedding is to unfold the phase space to a multivariate space, which is representative of the original dynamics. [21] has shown that under suitable conditions, if the dynamical system has dimension d_A and if the embedding dimension is chosen as $D > 2d_A$, then all the self-crossings in the trajectory due to the projection can be eliminated. The False Neighborhood method [11] accomplishes this task, where it views the dynamics as a compact object in the phase space. If the embedding dimension is too low (the system is not correctly unfolded), many points that lie very close to each other (i.e., neighbors) are far apart in the higher dimensional correctly unfolded space. Identification of these false neighbors allows the technique to determine that the dynamical system has not been correctly unfolded.

For, the time series, $x(n)$, in d^{th} and $(d+1)^{th}$ dimensional embedding, the Euclidean distance between an arbitrary point, $y(n)$ and its closest neighbor $y'(n)$ is, $R_d^2(n) = \sum_{k=0}^{d-1} [x(n+kT) - x'(n+kT)]^2$ and $R_{d+1}^2(n) = \sum_{k=0}^d [x(n+kT) - x'(n+kT)]^2$ respectively. If the ratio of these two distances exceeds a threshold R_{tol} (we took this as 15 in this paper), the points are considered to be false neighbors in the d^{th} dimension. The method starts from $d = 1$ and increases it to D , until only $1 - 2\%$ of the total points appear as false neighbors. Then, we deem the phase space to be completely unfolded in \mathcal{R}^D , a D -dimensional Euclidean Space.

2.2 Dirichlet Process and its Stick-breaking Representation

A Dirichlet Process [7], $D(\alpha_0, G_0)$ is defined as a probability distribution over a sample space of probability distributions, $G \sim DP(\alpha_0, G_0)$. Here, α_0 is the concentration parameter and G_0 is the base distribution. According to the stick-breaking construction [19] of DP, G , which is a sample from DP, is an atomic distribution with countably infinite atoms drawn from G_0 .

$$v_i | \alpha_0, G_0 \sim \text{Beta}(1, \alpha_0), \quad \theta_i | \alpha_0, G_0 \sim G_0, \quad M_i = v_i \prod_{l=1}^{i-1} (1 - v_l), \quad G = \sum_{i=1}^{\infty} M_i \delta_{\theta_i} \quad (2)$$

In the DP mixture model [6], DP is used as a non-parametric prior over parameters of an Infinite Mixture model.

$$z_n | \{v_1, v_2, \dots\} \sim \text{Categorical}\{M_1, M_2, M_3, \dots\}, \quad X_n | z_n, (\theta_i)_{i=1}^{\infty} \sim F(\theta_{z_n}) \quad (3)$$

Here, F is a distribution parametrized by θ_{z_n} . $\{M_1, M_2, M_3, \dots\}$ is defined by Eq.3.

3 NPB-NR Model

3.1 Step One: Clustering of Phase Space

Given a time series $\{x(1), x(2), \dots, x(N)\}$, let the minimum embedding dimension be D (using the False Neighborhood). Hence, the reconstructed phase space is,

$$\begin{bmatrix} x(1) & x(2) & \dots & x(N - (D-1)T) \\ x(1+T) & x(2+T) & \dots & x(N - (D-2)T) \\ \vdots & \vdots & \ddots & \vdots \\ x(1 + (D-1)T) & x(2 + (D-1)T) & \dots & x(N) \end{bmatrix} \quad (4)$$

Here, each column represents a point in the phase space. The generative model of the points in the phase space is now assumed as,

$$v_i | \alpha_1, \alpha_2 \sim \text{Beta}(\alpha_1, \alpha_2), \quad \{\mu_{i,d}, \lambda_{i,d}\} \sim \mathcal{N}(\mu_{i,d} | m_d, (\beta_d, \lambda_{i,d})^{-1}) \text{Gamma}(\lambda_{i,d} | a_d, b_d) \\ z_n | \{v_1, v_2, \dots\} \sim \text{Categorical}\{M_1, M_2, M_3, \dots\}, \quad X_d(n) | z_n \sim \mathcal{N}(\mu_{z_n,d}, \lambda_{z_n,d}) \quad (5)$$

Here, $X_d(n)$ is the d^{th} co-ordinate of the n^{th} phase space point. $\{z, v, \mu_{i,d}, \lambda_{i,d}\}$ is the set of latent variables. The distribution, $\{\mu_{i,d}, \lambda_{i,d}\}$, is the base distribution of the DP. $\{M_1, M_2, M_3, \dots\}$ denotes the Categorical Distribution parameters determined by Eq.3. In this DP mixture, the sequence, $\{M_1, M_2, M_3, \dots\}$, creates an infinite vector of mixing proportions and $\{\mu_{z_n,d}, \lambda_{z_n,d}\}$ are the atoms representing the mixture components. This Infinite Mixtures of Gaussians picks clusters for each phase space point and lets the phase space data determine the number of clusters. From this perspective, we can interpret the DP mixture as a flexible mixture model in which the number of components (i.e., the number of cells in the partition) is random and grows as new data is observed.

3.2 Step Two: Non-linear Mapping of Phase Space Points

Due to the discretization of the original continuous phase space, our assumption is that a point in the phase space is constructed by a nonlinear map R whose form we wish to approximate. In this section, we approximate this non-linear map of the subsequent phase space points via the proposed non-linear regression. We assume that a specific cluster has N points. We reorder these points according to their occurrence in the time series. We then pick the corresponding image of these points (which are the temporally subsequent phase space points according to the original time delay embedding). We map each phase space points in the cluster through an Infinite Mixtures of Linear Regression to their respective images. The model is formally defined as:

$$y_1(n) = R_1(x(n)) \quad y_2(n) = R_2(x(n)) \quad \dots y_D(n) = R_D(x(n)) \quad (6)$$

Here, $R_{1:D}$ are non-linear Regressors which is described by the following set of equations. Here, $X_d(n)$ and $Y_1(n)$ represent the d^{th} co-ordinate of the n^{th} phase space point and the first co-ordinate of its post image respectively. $\{z, v, \mu_{i,d}, \lambda_{x,i,d}, \beta_{i,d}, \lambda_{y,i}\}$ is the set of latent variables and the distributions, $\{\mu_{i,d}, \lambda_{x,i,d}\}$ and $\{\beta_{i,d}, \lambda_{y,i}\}$ are the base distributions of the DP. $\{M_1, M_2, M_3, \dots\}$ is defined by Eq.3. Although these set of equations are for R_1 , the same model applies for $R_{2:D}$, representing $Y_{2:D}(n)$.

$$\begin{aligned} v_i | \alpha_1, \alpha_2 &\sim \text{Beta}(\alpha_1, \alpha_2), \{\beta_{i,d}, \lambda_{y,i}\} \sim \mathcal{N}(\beta_{i,d} | m_{y,d}, (\beta_{y,d}, \lambda_{y,i})^{-1}) \text{Gamma}(\lambda_{y,i} | a_y, b_y) \\ z_n | \{v_1, v_2, \dots\} &\sim \text{Categorical}\{M_1, M_2, M_3, \dots\}, \\ Y_1(n) | X(n), z_n &\sim \mathcal{N}\left(\beta_{z_n,0} + \sum_{d=1}^D \beta_{z_n,d} X_d(n), \lambda_{y,z_n}^{-1}\right) \end{aligned} \quad (7)$$

The Infinite Mixture model approach to the Linear Regression makes the covariate be associated with the model via a non-linear function, resulting from marginalizing over the other mixtures with respect to a specific mixture. Also, now the variance is different across different mixtures, thereby capturing Heteroscedasticity.

3.3 Step Three: Restructuring of the Dynamics

The idea here is to perturb the trajectory to make the modified phase space more consistent with the dynamics, which is equivalent to reducing the error by perturbing the phase space points from its original position and also the error between the perturbed position and the mapped position. We have to choose a new sequence of phase space points, $\widehat{x(n)}$, such that following objective is minimized.

$$\sum_{n=1}^N (\|\widehat{x(n)} - x(n)\|^2 + \|\widehat{x(n)} - R(\widehat{x(n)})\|^2 + \|R(\widehat{x(n)}) - (x_{post-image})\|^2) \quad (8)$$

R is the non-linear Regressors ($R_{1:D}$) that are used to temporally approximate the phase space (Described in the section above). N is the number of points in the specific cluster. This is done across all the clusters. In addition, to create the new noise removed time series, perturbations of $x_d(n)$'s are done consistently for all subsequent points, such that we can revert back from the phase space to a time series. For example, if the time delay is 1 and the embedding dimension is 2, then, the phase space points are perturbed in such a way that when $x(n) = (t(n), t(n+1))$ is moved to $\widehat{x(n)} = (\widehat{t(n)}, \widehat{t(n+1)})$, we make the first co-ordinate of $\widehat{x(n+1)}$ to be $\widehat{t(n+1)}$. These form a set of equality constraints. What results is a convex program, that is then solved to retrieve the denoised time series.

The entire algorithm is summarized in Table 1.

-
-
1. Form the phase space dynamics from the Noisy Time Series according to Eq. 4 with the embedding dimension determined by False Neighborhood method described in Section 3.1
 2. Cluster the points in the phase space via Infinite Mixture of Gaussian Densities formally defined in Eq. 5
 3. For each cluster, map each phase space point via an infinite mixtures of linear regression($R_{1:D}$) to its temporally subsequent point (post-image) which is formally defined in Eq. 6. and Eq. 7
 4. Infer the latent parameters for both Infinite Mixture of Gaussian Densities and Infinite Mixture of Linear Regression. $\{z, v, \mu_{i,d}, \lambda_{i,d}\}$ and $\{z, v, \mu_{i,d}, \lambda_{x,i,d}, \beta_{i,d}, \lambda_{y,i}\}$ were inferred through variational inference which we could not include due to lack of space problem. The inference gives us the form of the regressors, ($R_{1:D}$)
 5. Restructure the dynamics via optimizing the Convex function in Eq. 8. The restructuring is done consistently for all the subsequent points, which leads to the reconstruction of the noise removed time series.
-
-

Table 1. Algorithm: Step-wise Description of NPB-NR Process.

4 Experimental Results

4.1 An Illustrative Description Of The NPB-NR Process

First, we present an illustrative pictorial description of the complete NPB-NR process with a real world historical stock price dataset. Our model for the historical time series of the stock price is a low-dimensional dynamical system that was contaminated by noise and passed through a measurement function at the output. Our task was to denoise the stock price to not only recover the underlying original phase space dynamics and create the subsequent noise removed stock price via the NPB-NR process, but also to utilize it to make better future predictions of the stock price. We picked historical daily close out stock price data of IBM from March-1990 to Sept-2015 for this task. The original noisy time series is plotted in Fig.2. The various stages of NPB-NR are illustrated in the subsequent figures. The underlying dimension of the phase space turned out to be 3 from the False Neighborhood Method. The Reconstructed Phase Space with noise is shown in Fig.3. The completely clustered phase space and one specific cluster in the phase space by Dirichlet Process Mixture of Gaussian of NPB-NR (step one) is shown in Fig.4. For a 3-dimensional phase space, as is the case with the IBM stock price data, consider X and Y to be two temporally successive points in one cluster. Therefore, the non-linear regression model (Step Two) in NPB-NR is $Y(1) = R_1(X(1), X(2), X(3))$, $Y(2) = R_2(X(1), X(2), X(3))$ and $Y(3) = R_3(X(1), X(2), X(3))$. In Fig.5, we plot $Y(1)$ against $X(1)$, $X(2)$ and $X(3)$ (The first regression- $R(1)$) to depict the non-linearity of the regression model which we have modeled through the Dirichlet Process Mixtures of linear regression (step two). The trajectory adjusted (step three) and consequently the noise removed specific cluster and the complete noise removed phase space are shown in Fig. 6. Finally, the denoised time series is shown in Fig. 7. The error information for prediction for IBM stock data is reported in Table 2.

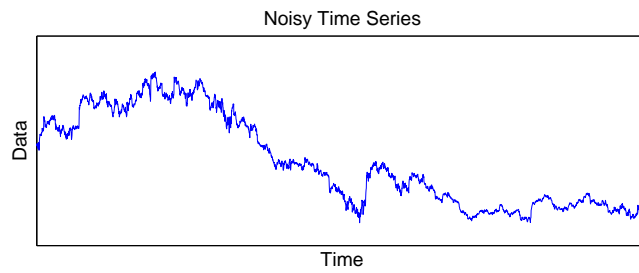


Fig. 1. Plot of the Noisy (15 db SNR) and Noiseless Time Series Data

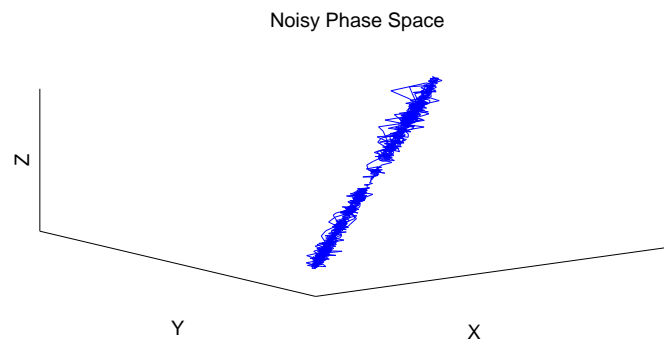


Fig. 2. Depiction of The Noiseless and Noisy Phase Space (Reconstructed).

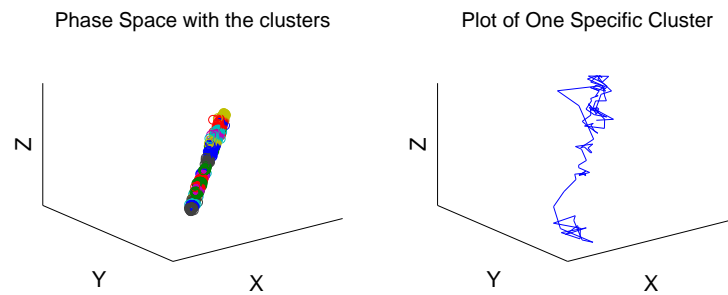


Fig. 3. Depiction of Whole Clustered Phase Space (Step One) and One Single Cluster

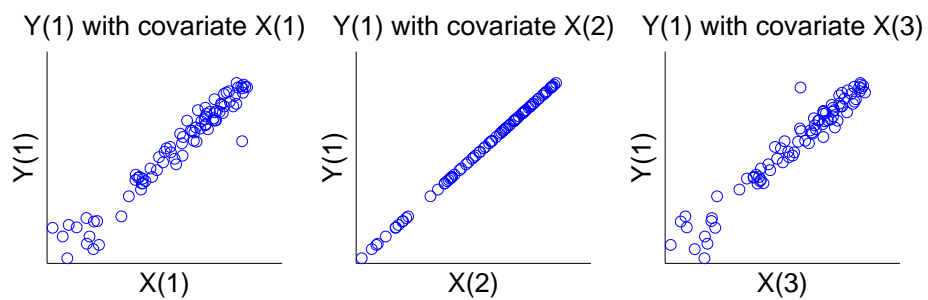


Fig. 4. The Regression Data: Y(1) Regressed with Covariate as X(1), X(2) and X(3)

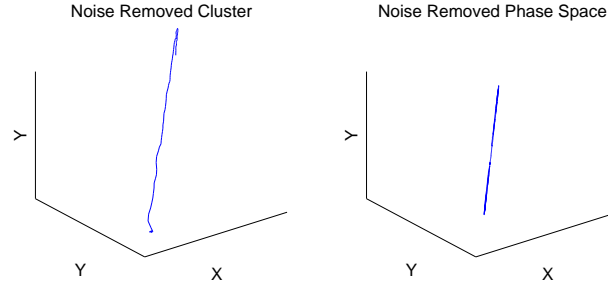


Fig. 5. The Single Noise Removed Cluster and Whole Noise Removed Phase Space

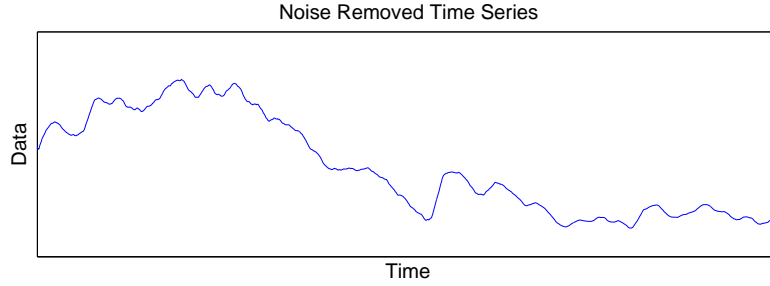


Fig. 6. Plot of the Noise Removed Time Series Data

4.2 Prediction Accuracy

NPB-NR was used for time series forecasting. The first dataset was drawn from the stock market. We choose 5 stocks (IBM, JPMorgan, MMM, Home-Depot and Walmart) from March, 2000 to Sept., 2015 with 3239 instances (time points) from “DOW30”. The next four datasets came from the Santa Fe competition compiled in [8]. The first is a Laser generated dataset which is a univariate time record of a single observed quantity, measured in a physics laboratory experiment. The next is a Currency Exchange Rate Dataset which is a collection of tickwise bids for the exchange rate from Swiss Francs to US Dollars, from August 1990 to April 1991. The next dataset is a synthetic computer generated series governed by a long sequence of known high dimensional dynamics. The fourth dataset is a set of astrophysical measurements of light curve of the variable white dwarf star PG 1159035 in March, 1989. The next set of datasets are the Darwin sea level pressure dataset from 1882 to 1998, Oxygen Isotope ratio dataset of 2.3 million years and US Industrial Production Indices dataset from Federal Reserve release. All of these are taken from [24]. NPB-NR was compared with the *GARCH*, *AR*(ρ), *ARMA*(p, q) and *ARIMA*(p, d, q) models, where ρ, p, d, q were taken by cross-validations ranging from 1 to 10 fold. We also compared NPB-NR to PCA and kernel PCA [3] with sigma set to 1, and Gaussian Process Based Auto-regression with ρ taken by cross-validations ranging from 1 to 5 fold. We also compared results from Hard Threshold Wavelet denoising using the “wden” Matlab function. All competitor algorithms were run with a 50-50 training-testing split. We report the Mean Square Error (MSE, L2) of the forecast for all the competitor algorithms in Table 2. Individual time series were reconstructed into a phase space with the dimension determined by

the False Neighborhood method, was passed through NPB-NR to find the most consistent dynamics by reducing noise, and subsequently fed into a simple auto-regressor with lag order taken as the embedding dimension of the reconstructed time series. In most datasets, NPB-NR not only yielded better forecasts, but also a smaller standard deviation among its competitors among the 10 runs.

4.3 Noise Reduction Experiment

We evaluated the NPB-NR technique for noise reduction across several well known dynamical systems, namely, Lorenz attractor (chaotic) [13], Van-der-poll attractor [16] and Rossler attractor [17] (periodic), Buckling Column attractor (non strange non chaotic, fixed point), Rayleigh attractor (non strange non chaotic, limit cycle) [1] and GOPY attractor (strange non-chaotic) [10].

Although noise was added to the time series such that the SNR ranged from 15 db to 100 db, it is impossible to calculate numerically or from the Power Spectrum how much noise was actually removed from the noisy time series. Therefore, for both the noise removed and the noisy time series we calculated the fluctuation error:

$$f_i = \|x_i - x_{i-1} - (dt) \cdot f(x_{i-1}, y_{i-1}, z_{i-1})\|$$

This measures the distance between the observed and the predicted point in the phase space. Here, measurement of the noise reduction percentage is given by,

$$R = 1 - \frac{E_{noise-removed}}{E_{noisy}}, \quad E = \left(\frac{\sum f_i^2}{N} \right)^{\frac{1}{2}}$$

We tabulated the noise reduction percentages of the NPB-NR, the low pass filter, and also wavelet denoising methods in Table 3. For the wavelet method, we used the matlab “wden” function in ‘soft’ and ‘hard’ threshold mode. The NPB-NR yielded the highest noise reduction percentage for 15-100 db SNR. Since the faithful reconstruction of the underlying dynamics intrinsically removes the noise, as the noise increases the noise reduction performance of NPB-NR got significantly better as opposed to the other techniques.

4.4 Power Spectrum Experiments

We ran a Power Spectrum experiment for a noise corrupted Van-der-poll attractor (periodic) [16] as well as a time series created by superimposing 6 Sinusoids and subsequently corrupting it with noise. The noise was additive white Gaussian noise with the SNR (Signal-to-Noise ratio) set at 15 db. Van-der-poll is a simple two dimensional attractor with $b = 0.4$; $x_0 = 1$; $y_0 = 1$ and the superimposition of Sinusoids is a simple limit cycle attractor with negative Lyapunov Exponents and no fractal structure. We plot the phase space and the Power Spectrum of the noisy time series generated from these attractors, the noise removed solution with a 6th-order Butterworth low-pass filter (cut-off freq. 30 Hz and 1000 Hz respectively) and the NPB-NR technique. The Power Spectrum and the phase space plot of the Van-der-poll and Sinusoid Attractors is shown in Fig.1. Note that NPB-NR successfully made the harmonics/peaks more prominent which was originally obscured by the noise. The filtering method was unable to restore the harmonics, although it removed some of the higher frequency components. We also observe that NPB-NR smoothened out the phase space dynamics better than the low pass filter.

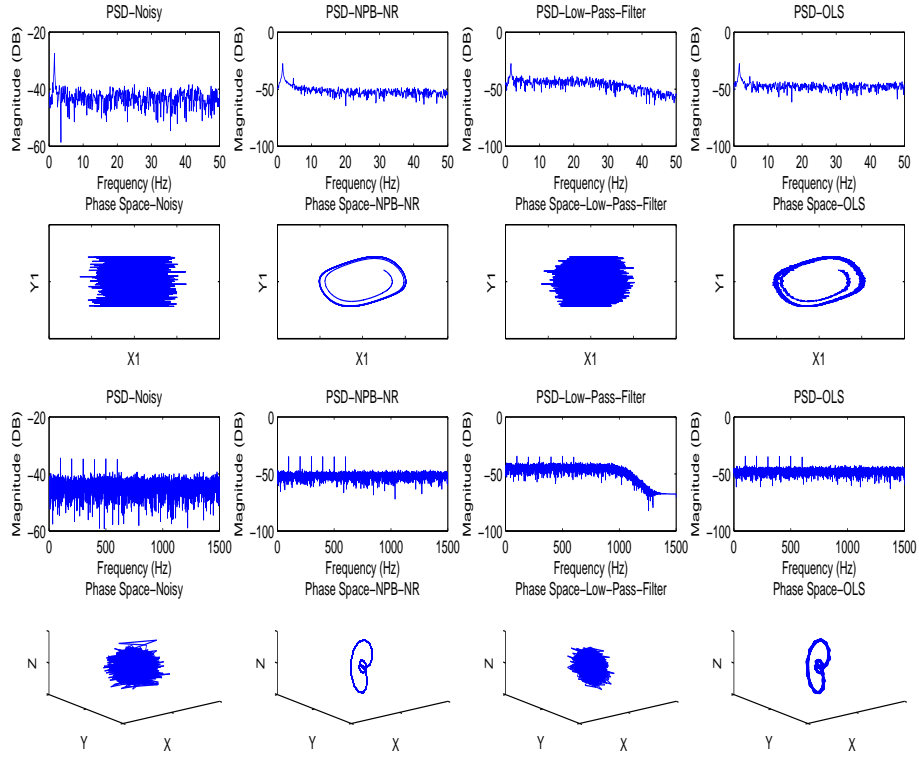


Fig. 7. The Power Spectrum and The Phase Space Plot of Van-der-poll and Sinusoid Attractor

5 Conclusion

In this article, we have formulated a Bayesian Non-Parametric Model for noise reduction in time series. The model captures the local non-linear dynamics in the time delay embedded phase space to fit the most appropriate dynamics consistent with the data. We have derived the mean field Variational Inference for the Dirichlet Process Mixture of Linear Regression by maximizing the evidence lower bound to obtain the variational parameters. Finally, we have evaluated the NPB-NR technique on various time series generated from several dynamical systems, stock market data, LASER data, Sea Level Pressure data, etc. The technique yields much better noise reduction percentage, power spectrum analysis, accurate dimension and prediction accuracy. In the experiments, we varied the scale factor from 1 to 5 in increments of .25. This scale factor modulates the number of clusters in the phase space. Developing theoretical insights into how the number of clusters affects the adjustment of the dynamics would be an interesting topic for future research. We also plan to explore which kind of physical systems can be analyzed using Non-parametric Bayesian based noise reduction methods. Finally, considerable effort should be given to analyzing time series generated from higher dimensional systems.

MSE	NPB-NR	GARCH	Wavelet	AR	ARMA	ARIMA	PCA	KERNEL-PCA	GPR
IBM	1.43	1.65	1.37	1.87	1.70	1.68	1.98	1.90	1.84
JPM	1.38	1.52	1.49	1.46	1.42	1.39	1.67	1.59	1.73
MMM	1.69	1.87	1.96	2.06	1.93	1.83	1.14	2.11	2.23
HD	1.74	1.58	1.46	1.73	1.69	1.62	1.79	1.72	1.86
WMT	1.24	1.47	1.58	1.39	1.35	1.29	1.49	1.57	1.38
LASER	.97	1.36	1.29	1.31	.86	1.15	1.42	1.35	1.34
CER	.82	.99	.93	.94	.88	.84	1.18	1.11	1.05
CGS	1.79	2.11	1.88	2.03	1.96	1.86	2.38	2.28	2.17
ASTRO	1.82	2.19	2.14	2.08	1.91	1.92	2.26	2.33	2.46
DSLPL	1.33	1.68	1.53	1.49	1.41	1.14	1.68	1.60	1.55
OxIso	1.19	1.38	1.87	1.32	1.26	1.53	1.48	1.41	1.45
USIPI	1.30	1.57	1.36	1.48	1.43	1.62	1.57	1.57	1.63
stan.dev.	NPB-NR	GARCH	Wavelet	AR	ARMA	ARIMA	PCA	KERNEL-PCA	GPR
IBM	1.34	1.89	1.86	1.67	1.78	1.39	3.26	2.37	1.42
JPM	1.63	1.98	1.97	1.78	1.94	2.01	2.35	1.69	2.21
MMM	1.82	1.48	1.29	1.42	1.36	1.82	1.73	1.66	1.59
HD	1.86	1.85	1.85	1.92	1.77	1.88	1.93	1.90	1.86
WMT	1.79	1.66	1.82	1.62	1.73	2.31	1.67	1.61	1.98
LASER	2.12	2.28	2.39	2.19	2.36	1.72	2.42	2.27	2.39
CER	1.69	1.78	1.93	1.84	1.74	1.71	1.80	1.91	1.72
CGS	2.34	1.92	1.88	1.97	2.05	1.87	1.95	2.17	2.11
ASTRO	1.13	1.82	1.41	1.37	1.69	1.79	1.55	1.29	1.62
DSLPL	1.58	1.49	1.19	1.27	1.35	1.45	1.26	1.42	1.25
OxIso	2.47	2.15	1.99	2.24	1.89	2.25	1.92	2.61	2.45
USIPI	2.23	2.33	2.49	2.42	1.89	2.29	2.37	2.72	2.25

Table 2. MSE and Standard Deviation of all the datasets for all the competitor algorithms in 50-50 random Training-Testing Split for 10 Runs.

	Lorenz	GOPY	Van-Der-Poll	Rosler	Rayleigh
Noise Level-15db SNR					
NPB-NR	40	45	54	29	34
Low Pass Filter	19	27	40	19	31
Wavelet_soft	15	13	29	21	25
Wavelet_hard	17	7	21	18	22
Noise Level-35db SNR					
NPB-NR	51	59	61	40	56
Low Pass Filter	26	31	40	28	39
Wavelet_soft	22	22	36	33	32
Wavelet_hard	23	14	29	24	28
Noise Level-60db SNR					
NPB-NR	63	71	75	79	82
Low Pass Filter	31	35	40	37	41
Wavelet_soft	32	29	41	40	42
Wavelet_hard	29	21	33	32	33
Noise Level-80db SNR					
NPB-NR	72	76	79	81	84
Low Pass Filter	34	39	43	43	44
Wavelet_soft	35	35	46	44	47
Wavelet_hard	34	27	39	36	38
Noise Level-100db SNR					
NPB-NR	80	79	85	85	89
Low Pass Filter	38	43	46	47	46
Wavelet_soft	41	39	50	50	60
Wavelet_hard	36	30	45	40	51

Table 3. Noise Reduction Percentage of the Attractors for the NPB-NR, the Low Pass Filtering Method and the Hard and Soft Threshold Wavelet method.

References

1. Abraham, R., Shaw, C.: Dynamics: The Geometry of Behavior. Ariel Press (1985)
2. Badii, R., Broggi, G., Derighetti, B., Ravani, M., Ciliberto, S., Politi, A., Rubio, M.: Dimension increase in filtered chaotic signals. *Phys. Rev. Lett.* 60, 979–982 (1988)
3. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer-Verlag (2006)
4. David, L., Donoho, J.: De-noising by soft-thresholding. *IEEE Trans. Inf. Theor.* 41(3), 613–627 (1995)
5. Elshorbagy, A., Panu, U.: Noise reduction in chaotic hydrologic time series: facts and doubts. *Journal of Hydrology* 256(34), 147–165 (2002)
6. Escobar, D.M., West, M.: Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association* 90(430), 577–588 (1995)
7. Ferguson, T.: A bayesian analysis of some nonparametric problems. *Annals of Statistics* 1, 209–230 (1973)
8. Gershenfeld, N., Weigend, A.: Time Series Prediction: Forecasting the Future and Understanding the Past. Addison-Wesley (1994)
9. Grassberger, P., Schreiber, T., Schaffrath, C.: Non-linear time sequence analysis. *International Journal of Bifurcation and Chaos* 1(3), 521–547 (1991)
10. Grebogi, C., Ott, E., Pelikan, S., Yorke, J.A.: Strange attractors that are not chaotic. *Physica D: Nonlinear Phenomena* 13(1), 261–268 (1984)
11. Kennel, M.B., Brown, R., Abarbanel, H.D.I.: Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* 45(6), 3403–3411 (1992)
12. Kostelich, E.J., Yorke, J.A.: Noise reduction: Finding the simplest dynamical system consistent with the data. *Phys. D* 41(2), 183–196 (1990)
13. Lorenz, E.N.: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences* 20(2), 130–141 (1963)
14. Mallat, S., Hwang, W.L.: Singularity detection and processing with wavelets. *Information Theory, IEEE Transactions on* 38(2), 617–643 (1992)
15. Mitschke, F., Moller, M., Lange, W.: Measuring filtered chaotic signals. *Phys. Rev. A* 37(11), 4518–4521 (1988)
16. Pol, B.V.D.: A theory of the amplitude of free and forced triode vibrations. *Radio Review* 1, 701–710 (1920)
17. Rossler, O.: An equation for continuous chaos. *Physics Letters A* 57(5), 397–398 (1976)
18. S. Ghosal, J.K.G., Ramamoorthi, R.V.: Posterior consistency of dirichlet mixtures in density estimation. *Annals of Statistics* 27, 143–158 (1999)
19. Sethuraman, J.: A constructive definition of dirichlet priors. *Statistica Sinica* 4, 639–650 (1994)
20. Site, G., Ramakrishnan, A.G.: Wavelet domain nonlinear filtering for evoked potential signal enhancement. *Computer and Biomedical Research* 33(3), 431–446 (2000)
21. Takens, F.: Dynamical systems and turbulence, warwick 1980. Detecting strange attractors in turbulence 898, 366–381 (1981)
22. Tokdar, S.T.: Posterior consistency of dirichlet location-scale mixture of normals in density estimation and regression. *Sankhya: The Indian Journal of Statistics* 68(1), 90–110 (2006)
23. Wang, Z., Lam, J., Liu, X.: Filtering for a class of nonlinear discrete-time stochastic systems with state delays. *Journal of Computational and Applied Mathematics* 201(1), 153–163 (2007)
24. West, M.: <http://aiweb.techfak.uni-bielefeld.de/content/bworld-robot-control-software/>
25. Zhang, L., Bao, P., Pan, Q.: Threshold analysis in wavelet-based denoising. *Electronics Letters* 37(24), 1485–1486 (2001)