UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Chaos as an Intermittently Forced Linear System

Seminar Essay
by Siddhanth Janadri

Research group
Data Science for Engineering

Contents

Abstract

The state of a dynamical system often evolves with time over state space. Sometimes, these dynamics of a system give rise to complex behaviours that makes it difficult to predict its nature which is commonly known as Chaos. We make use of Lorenz system to present a data-driven approach to decompose chaotic dynamics as an intermittently forced linear system. Takens' delay embedding along with Koopman operator theory and spare regression are considered to represent non-linear dynamics. The process of decomposing chaotic dynamics into a linear model with forcing by low energy delay co-ordinates is known as Hankel Alternative View Of Koopman (HAVOK) analysis. This analysis is applied to a various real world chaotic systems such as Rossler system, double pendulum, earth's magnetic field, measles outbreak in the USA and so on. The resulting statistics that are obtained after applying forementioned analysis will have long tails that indicates rare or intermittent events that happens with a particular chaotic systems. This analysis paves way to go from data to dynamics because of abundance supply of data in the real world with the lack of methods to describe dynamical systems.

1 Introduction

Dynamical systems change their states over time and are represented over state space. Time and state at one time evolves to a state or set of states at later time. These states are ordered by time which can be thought of as a single quantity. These systems can be deterministic which means that there is a unique consequent to every state of the system or can be stochastic if there is a probability distribution of consequents[1]. Each dynamical system can be represented in the form:

$$\frac{d}{dt}\mathrm{x}(t) = \mathrm{f}(\mathrm{x}(t)) \tag{1}$$

Its discretized version can be of the form:

$$\mathrm{x}_{k+1} = \mathrm{F}\left(\mathrm{x}_k\right) \tag{2}$$

where x_k or x is the state variable of the system at time t that can take either scalar or vector value and F or f is the function that determines rules by which the system changes its state over time[BBP$^+$17]. Firstly, linear systems have single fixed point through which all the states evolve around. But non-linear systems have multiple fixed points and as we increase the non-linearity of the system we get more fixed points or periodic orbits. At some point, the non-linearity becomes more that the system becomes chaotic.

Chaotic systems are present in every field of domain such as physics, biology, engineering processes and financial markets. These systems make the scientists very difficult

[1] https://mathworld.wolfram.com/Chaos.html

to discern the pattern due to its unpredictable nature. A simple example is double pendulum which has two fixed points. The reason why double pendulum is considered as a chaotic system is because the tiny difference in the initial conditions causes the trajectories to diverge exponentially at least for some time. These trajectories are very sensitive to initial conditions. Chaotic systems are not random but they have the balance of order and structure. A phenomenon known as Butterfly effect is the defining property of a chaotic system. It states that a butterfly flapping its wings in one location can cause a tornado in a different location. This is because of its sensitive dependence on initial conditions. This concept has been leveraged in this approach to predict rare and intermittent events in chaotic systems. Though the governing equations of the dynamical systems have been very helpful in understanding and analyzing the systems, apparently, scientists are turning their sight towards limitless availability of data. The trend has been changed in recent years where the data driven analysis has been instrumental[BBP$^+$17].

Consequently, there are a couple of concepts and theories that have been leveraged to derive HAVOK analysis in order to predict the lobe switching activity in attractor and bursting phenomenon in some systems. In the following sections, we will brief about Koopman Operator theory to find the measurement subspaces, system identification techniques based on dynamic regression followed by Taken's Delay Embedding that generates delay co-ordinates in order to populate the Hankel matrix. Later, we will talk about HAVOK analysis and apply it on Lorenz system to explain with a real world example. Following this, we will discuss how HAVOK analysis serves its purpose on other real world examples that shows an irregular pattern of occurrences. In the end, we will discuss its scope of future developments and conclude it with a short summary of everything.

2 Background Knowledge

2.1 Koopman Operator Theory

This theory was derived from Koopman Spectral Analysis that delineates Hamiltonian systems and its evolution from measurements and data. This theory enables us to represent nonlinear dynamics linearly. In order to understand any dynamical system, governing equations and the dynamics of the state of the system play a major role. According to Koopman Operation Theory, instead of searching and propagating dynamics of the system like position, velocity etc we can propagate measurements of the state of the state[2].

$$\mathcal{K}g \triangleq g \circ \mathbf{F} \quad \Longrightarrow \quad \mathcal{K}g\left(\mathbf{x}_k\right) = g\left(\mathbf{x}_{k+1}\right) \tag{3}$$

where g is the measurement function of the state of the system and K is an infinite dimensional operator known as Koopman Operator. This operator takes the measurement function and propagates it forward in time. For an instance, if we have a measurement function at a particular time k i.e $g(x_k)$, then with koopman operator we will be able to

[2]https://www.youtube.com/watch?v=831Ell3QNck

advance the measurement forward in time to $g(x_{k+1})$. With the aim of representing nonlinear dynamics in a linear fashion, there is an agreement between nonlinear dynamics on a finite dimensional states and linear dynamics represented on an infinite dimensional Hilbert space of measurement functions[MSM20]. From the data driven perspective like weather prediction, financial market where there is immense amount of measurements and data, propagating these system measurements to Koopman operator gives us the linear representation in an infinite dimensional space if we choose the right measurements. Apparently, this depends on a measurement subspace that remains invariant to the Koopman operator[BBKK21].

2.1.1 Koopman invariant measurement subspaces

The goal is to find the finite dimensional approximation in data driven analysis. For any measurement function g in the measurement subspace ranged across a set of available measurement functions $g1, g2, g3\dots$ that is given by:

$$g = \alpha_1 g_1 + \alpha_2 g_2 + \cdots + \alpha_p g_p \tag{4}$$

koopman operator lets the measurement function be in the same subspace. This is basically what is known as Koopman invariance. If the function in the subspace that is hit by Koopman operator remains in the same subspace then such a subspace is known as Koopman invariant measurement subspace. Consequently, if we find such a subspace that remains invariant to the Koopman operator, then the operator that is restricted to that subspace becomes a finite dimensional linear operator on matrix. With the help of Koopman operator, p dimensional measurement subspace can be represented as $p * p$ matrix with a better choice of observable functions g. Matrix representations have always been very powerful as a means to achieve prediction and control of nonlinear systems represented in any dimension which, in this case, finite dimensional Koopman invariant subspaces[BBPK16].

Apparently, identifying such measurement subspaces with good measurement coordinates is one of greatest challenges because they carry a lot of information and they are related to eigen functions of the Koopman operator. Therefore, if we can manage to find such Koopman invariant measurement subspace there are a lot of methods and techniques that can be used for analysis except that finding such subspaces might be exceedingly difficult. In this paper, we propose an alternative perspective of getting Koopman invariant measurement subspaces for chaotic systems.

2.2 Data-driven Dynamic Regression

Due to immense availability of data, modelling of dynamical systems using regression techniques has earned the spotlight. In the following, we will discuss about two leading system identification techniques that are based on regression.

2.2.1 Dynamic Mode Decomposition

It is one of the data driven decomposition methods. It helps in reconstructing structure that come from nonlinear dynamical systems[3]. It extracts dynamic information based on sequence of snapshots of data. The two data matrices where one contains snapshots of system state in time and the other contains snapshots of system state with one step forward in time are utilized to provide a best fit linear model as shown in (6). The relationship between these two matrices can be defined by a best fit linear operator A as given in (7)[BBP+17].

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathrm{x}_1 & \mathrm{x}_2 & \cdots & \mathrm{x}_{m-1} \\ | & | & & | \end{bmatrix} \quad \mathbf{X}' = \begin{bmatrix} | & | & & | \\ \mathrm{x}_2 & \mathrm{x}_3 & \cdots & \mathrm{x}_m \\ | & | & & | \end{bmatrix} \tag{5}$$

$$\mathbf{X}' = \mathbf{A}\mathbf{X} \quad \implies \quad \mathbf{A} \approx \mathbf{X}'\mathbf{X}^{\dagger} \tag{6}$$

Then, we perform Singular Value Decomposition (SVD) on X because of large dimensionality of systems[BBP+17].

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{*} \tag{7}$$

After several steps which are not important in this discussion, we arrive at the below equation:

$$\Phi = \mathbf{X}'\tilde{\mathbf{V}}\tilde{\Sigma}^{-1}\mathbf{W} \tag{8}$$

The limitation of this method is its capability that is restricted only to linear measurements which clearly doesn't help representing nonlinear dynamical systems. The extended version of this method doesn't serve the purpose of choosing the right nonlinear measurements albeit it comes close to Koopman invariant measurement system. There seems to be two ways to determine measurement functions with one being the usage of information from the dynamical systems and another being the process of searching using brute-force technique in a particular Hilbert space[BBP+17].

2.2.2 Sparse Identification of Nonlinear Dynamics (SINDy)

This is another yet powerful technique to identify nonlinear dynamics from measurement subspace. A regression method which is similar to linear regression called sparse regression has been used in this technique. Furthermore, sparse regression is used when the number of samples are comparatively smaller than that of predictors that are used in the specific problem such as gene-expression measurement[4]. Every nonlinear function space has some active terms with respect to the system of interest. Finding these active terms in a brute force fashion is not an efficient way. This is when SINDy algorithm comes to use. It uses sparse regression method to identify active terms as non-zero terms. A matrix that contains time history of the state and the other one that contains its derivatives are arranged with data that is sampled at times $t_1, t_2, \ldots..t_m$ as shown below[BBP+17]:

[3] https://dc.uwm.edu/etd/1879/
[4] https://cims.nyu.edu/

$$\mathbf{X} = \begin{array}{c} \text{state} \\ \begin{bmatrix} x_1(t_1) & x_2(t_1) & \cdots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \cdots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_m) & x_2(t_m) & \cdots & x_n(t_m) \end{bmatrix} \end{array} \text{time} \qquad \dot{\mathbf{X}} = \begin{bmatrix} \dot{x}_1(t_1) & \dot{x}_2(t_1) & \cdots & \dot{x}_n(t_1) \\ \dot{x}_1(t_2) & \dot{x}_2(t_2) & \cdots & \dot{x}_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{x}_1(t_m) & \dot{x}_2(t_m) & \cdots & \dot{x}_n(t_m) \end{bmatrix}$$

Following the construction of the augmented library, theta(X) consists of candidate nonlinear values of the columns of the array. For instance, theta(X) may consist of constant, polynomial, and trigonometric values[BBP+17].

$$\Theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & & | & | & | & | & \\ 1 & \mathrm{X} & \mathrm{X}^{P_2} & \mathrm{X}^{P_3} & \cdots & \sin(\mathrm{X}) & \cos(\mathrm{X}) & \sin(2\mathrm{X}) & \cos(2\mathrm{X}) & \cdots \\ | & | & & | & & | & | & | & | & \end{bmatrix} \tag{9}$$

Like DMD technique, let's not go in depth about the next steps using sparse regression. The restricted isometry property may require that we normalize the columns of (X) first; this is particularly important when the entries in X are small, since the powers of X will be small. It can also be used to formulate discrete-time formulations, as the SINDy algorithm reduces to linear regression and uses transposed notation similar to the DMD algorithm above[BBP+17].

2.3 Taken's delay embedding

Takens' embedding theorem shows how a dynamical system's hidden states are captured in a series output. This technique has been used for a variety of tasks, such as time series prediction and attractor dimension estimation[5]. In one of the previous topics we understood that finding good measurement subspace is the most difficult task for nonlinear representation. Indeed, it is very important to model, predict and control dynamical systems. This embedding theorem allows for the robust analysis of measurement information in a nonlinear dynamical system. One can enrich a measurement $x(t)$ with copies of itself that are time-shifted $x(t-\rho)$, known as delay coordinates[BBP+17]. Dynamics of a system are mostly revealed from a time series of a single point measurement. Entire attractor of a dynamical system that is diffeomorphic to the original can be constructed which is based on singular value decomposition of a Hankel matrix. When the full state of a linear dynamical system can be determined from an analysis of the measurement history of the system, the concept of observability provides a basis for calculating the full state dynamically. This theorem justifies it's relation with nonlinear observability[Tak81].

[5]https://cnx.org/contents/k57$_M 8Tw$@2/Takens-Embedding-Theorem

3 Hankel Alternative View of Koopman (HAVOK)

Linearization of dynamics near fixed points or periodic orbits has long been applied for local linear representations of strongly nonlinear systems, and it stands to reason that the linearization of dynamics near fixed points or periodic orbits could revolutionize our ability to predict and control these systems[BBP+17]. In the following, we will generalize system identification using Hankel matrix to nonlinear dynamical system via the Koopman analysis. This analysis provides an alternative perspective to get Koopman invariant measurement subspaces for chaotic systems. It is based on time delay coordinates, embedding theory, Koopman operator and system identification of the Hankel operator. We apply HAVOK analysis on chaotic systems to get good intrinsic measurement coordinates where we can get finite dimensional Koopman linear approximation. These measurement coordinates are obtained from time history of a particular system which is better than considering just measurements of the state of the system. As DMD and augmented DMD don't approximate Koopman operator properly with nonlinear measurements, data driven method of approximation has been considered which seems to work better with nonlinear systems which predicts the future based on past time delay coordinates.

Taking a singular value decomposition (SVD) of the following Hankel matrix H, the eigen-time-delay coordinates can be obtained from a time series of single point measurements $x(t)$ by applying the Takens embedding theorem.

$$\mathrm{H} = \begin{bmatrix} x(t_1) & x(t_2) & \cdots & x(t_p) \\ x(t_2) & x(t_3) & \cdots & x(t_{p+1}) \\ \vdots & \vdots & \ddots & \vdots \\ x(t_q) & x(t_{q+1}) & \cdots & x(t_m) \end{bmatrix} = \mathrm{U}\Sigma\mathrm{V}^T \tag{10}$$

As we already know before, if we have the measurement function at a particular time, with the help of Koopman operator we can advance the measurement forward in time. Therefore, if we apply the Koopman operator to the above Hankel matrix, the resulting matrix looks as follows:

$$\mathrm{H} = \begin{bmatrix} x(t_1) & \mathcal{K}x(t_1) & \cdots & \mathcal{K}^{p-1}x(t_1) \\ \mathcal{K}x(t_1) & \mathcal{K}^2x(t_1) & \cdots & \mathcal{K}^p x(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{K}^{q-1}x(t_1) & \mathcal{K}^q x(t_1) & \cdots & \mathcal{K}^{m-1}x(t_1) \end{bmatrix} \tag{11}$$

This Hankel matrix is built by taking the time shifted copies of the measurement and stacking them up as rows of H. It consists of terms with different number of times Koopman operator is applied to the first point measurement. Hence, we call it a time series of a single point measurement. Upon performing Singular Value Decomposition on Hankel matrix H, three matrices are formed U, Σ and V^T.

A linear model cannot characterize multiple fixed points that describe chaos. There are still challenges to identifying a linear model for a chaotic system even with an approximately Koopman-invariant measurement system. The eigen-time delay coordinates in Hankel matrix motivate a linear regression model on the variables in V[CM79]. Therefore, an efficient workaround which turned out to be a better solution is employed where only first $r-1$ variables (rows) are treated for modelling while the last variable v_r is treated as a forcing term that correspond to transient events like lobe switching or bursting activity. This can be shown in the form of an equation as below:

$$\frac{d}{dt}\mathbf{v}(t) = \mathbf{A}\mathbf{v}(t) + \mathbf{B}v_r(t) \tag{12}$$

where $v = [v_1, v_2, ...v_{r-1}]^T$ is a vector of first $r-1$ eigen-time delay coordinates[BBP+17]. v_r can never be represented by a linear model because of which it is considered as a forcing term that approximate nonlinear dynamics. Therefore, the behavior of the system at a particular point in time can be judged based on the level of forcing. If the forcing is small, it tells that the system can be modelled linearly as opposed to situation when the forcing is large, events like bursting phenomenon or lobe switching are described.

In the following, we will go over the same explanation but on a simple chaotic system among well studied ones called Lorenz system.

4 HAVOK on Lorenz System

Let's assume that we have a single measurement of the system. For the sake of illustration, let's consider that we only measure time history of x variable. Once the time history of x variable is measured, we build a set of good delay coordinates by building a Hankel matrix. The first row of Hankel matrix is built by taking the time history $x(t_1), x(t_2), ...x(t_n)$. Now, the second row is populated by shifting one measurement forward in time $x(t_2), x(t_3)...x(t_{n+1})$. The third row will be another time shifted version of measurement. As a result, the Hankel matrix is built by taking the time shifted copies of measurement and stacking them up as rows.

In the next step, we take the Singular Value Decomposition of Hankel matrix to generate three matrices U, Σ, V^T. The columns of U and V are ordered hierarchically in terms of their magnitude in Σ. The first column of U and V are the most important eigen-time series that are most co-related with this measurement. The second column of U and V are the second most important or variance in H and so on. As a result, we get an ordered representation of these delay coordinates in terms of SVD of the Hankel matrix. Once we have these eigen-time delay coordinates we can plot the time history in the new coordinate system. We get a new embedded attractor by taking the columns of V i.e $v_1, v_2, ...$ and plotting the time history in those columns.

In this scenario, variable x is a good measurement and z variable would be a bad measurement. So, if we get the right measurement which is variable x and if we build

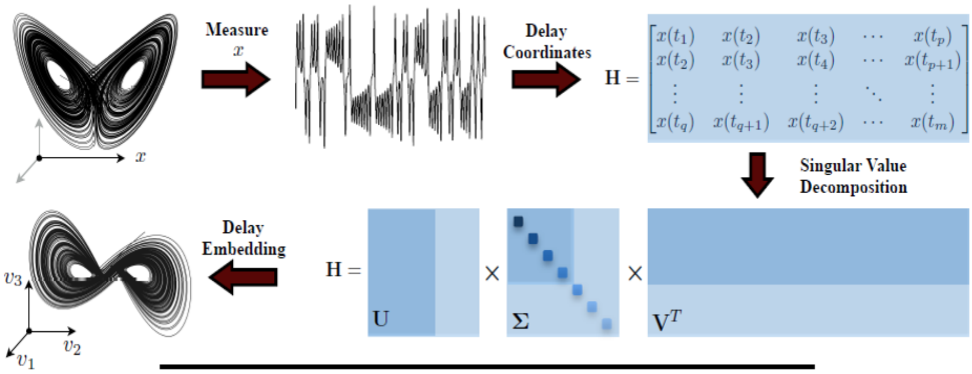

Figure 1: Decomposition of Chaos into a linear dynamical system with intermittent forcing[BBP+17]

sufficiently many delay coordinates on x, we will get an embedded attractor that is diffeomorphic to the original attractor. In a nutshell, all we need to do is take the time history of the right point measurement and we will be able to reconstruct an attractor that is diffeomorphic to the complete high dimensional system. The next step in this analysis is to find a measurement subspace that is invariant to the Koopman operator. Furthermore, if the variable is hit with the Koopman operator in the measurement subspace, the variable stays in the same subspace and as a result, data driven SVD of the Hankel matrix plays an important role.

The columns of U are the most important columns to express columns in H. In many systems, it has been found that very less number of columns in U are sufficient to express tons of columns of H. Therefore, the columns of H are approximately linearly dependent on the leading r columns of U. From prior experiments, the r value in Lorenz system is calculated as 15. Consequently, with 15 columns of U, every possible column in H can be expressed with machine precision. It means that the eigen-time delay coordinates in U is the measurement subspace where the time history will be measured and if these measurements are bombarded with Koopman operator, they stay in the same measurement subspace which is U because all the columns of H are spanned by the columns of U[6]. In summary, we get the data driven Koopman invariant measurement subspace in terms of eigen-time delay coordinates from the Hankel matrix. The connection with Koopman operator motivates to build a linear system on these time delay coordinates[BBP+17]. The more data we collect the better our linear dependence assumption gets and ultimately Koopman invariant measurement subspace becomes more approximate.

[6]https://www.youtube.com/watch?v=831Ell3QNck

4.1 Dynamic regression

The time history of V matrix components $v_1, v_2, ...v_{15}$ are considered to build a regression model. Algorithms like DMD and SINDy build regression models from normal data. Similarly, a regression model is built from eigen-time delay coordinates of Hankel matrix for fully nonlinear systems. The first r-1 delay cooridnates $v_1, ...v_{14}$ generates a very accurate regression fit. But the fit of the last term which is 15^{th} time series has a terrible fit. This whole framework was discovered from SINDy algorithm on time delay coordinates and it has been found that albeit including nonlinear terms (the last row) it was never well modelled with linear or nonlinear dynamics[6]. As a result, if the last row v_{15} is omitted and consider it as the forcing term, we will have a great linear regression model in the first $r-1$ eigen-time delay coordinates. It can also be interpreted as $\dot{V}$ term which is equal to matrix times V in addition to another matrix times the last v_r. In other words, it can be seen as the addition of a linear system and an input term where this input is the last eigen-time delay coordinates. Because of the connection to the Koopman operator, the HAVOK analysis guarantees that such models can be built for any chaotic system[BBP+17].

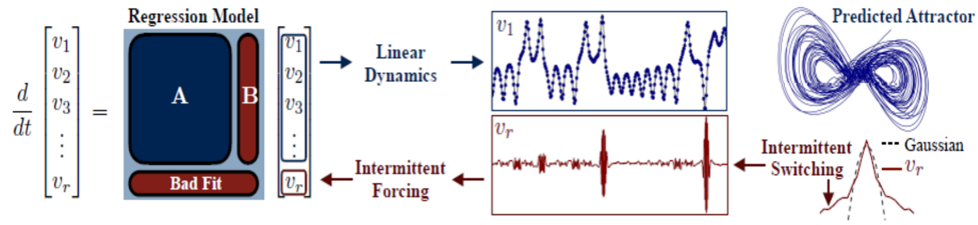

Figure 2: A best fit linear model from time-delay coordinates and lobe switching in chaotic dynamics[BBP+17]

The red curve is the rth eigen-time delay coordinates. It is the forcing term and it cannot be modelled properly. As a consequence, if the linear system is modelled with the red forcing, what we then get is blue. We get the attractor dynamics which aligns remarkably well with that of an embedded attractor and when the dynamics switch from one lobe to another, we get the bursting phenomenon in terms of an intermittent red forcing signature. If this model is run through the right forcing, we get a predicted attractor from a linear model that superbly matches the delay embedded nonlinear attractor. If we look at the statistics of the intermittent switching, the white dash curve is the gaussian curve and the intermittent forcing is abnormal as the long tails correspond to the extreme forcing events which in turn drives the lobe switching activity.

In conclusion, if we have sufficient data columns, a Hankel matrix can be built from which time delay coordinates can be found which helps in generating an embedded attractor. If the linear regression model is run on the coordinates of the generated embedded attractor, we get a linear model with a proper forcing which can be stochastic or intermittent that ultimately gives a heads up of when the system is going to switch.

4.2 HAVOK model of the Lorenz system

The HAVOK model is color coded by the magnitude and sign of the entries of the linear dynamics and the forcing term. This was discovered when then Sparse Identification of Nonlinear dynamics algorithm was applied to eigen-time delay coordinates. It has been found that the most parsimonious model even if the nonlinear terms in the dynamics are allowed is the beautiful linear structure as shown in the figure 3[BBP+17].

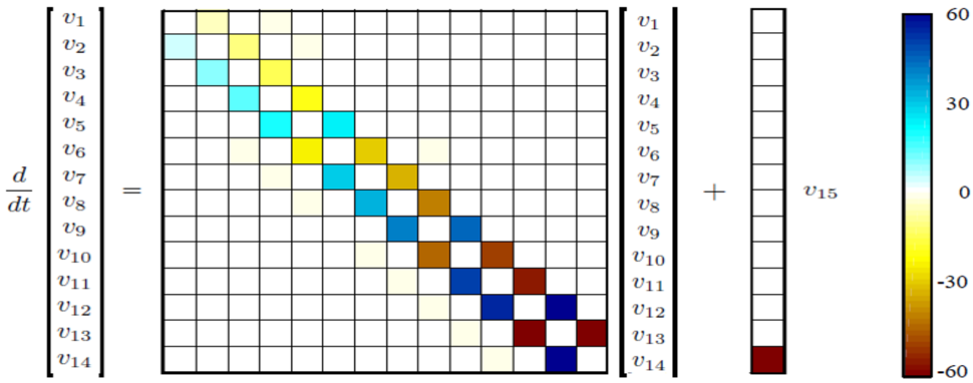

Figure 3: HAVOK model obtained from time-delay coordinates of the Lorenz system[BBP+17]

To eliminate terms with very small coefficients in the model, an additional sparsity penalizing factor can be added in addition to a straightforward linear regression procedure. Using this model, we produce a remarkable skew-symmetric structure[BBP+17]. The numbers directly above and below the diagonal are almost integer multiples of 5. We measure the time window in sliding window that evolves with the eigen modes in U matrix from SVD. This in turn gives us the initial conditions for V which also gives is the forcing term v_r. As we slide the window forward, we can predict the system. The eigen time histories that are used to get $v_1 \ldots . v_r$ tend to be polynomials and if v_r gets really active it means that we are about the switch the lobes.

5 Are HAVOK models predictive?

We are now aware of how this model helps us understand the behavior of chaotic systems and its nature of switching lobes. But we aren't sure if we get to know about such transient events in advance. For an instance, if we don't identify the possibility of an earthquake before it happens, it's of no use knowing its nature or the behavior. Therefore, these models can be utilized to the fullest if it can predict the behavior well before transient events like lobe switching or bursting phenomenon happen.

We don't have a future prediction of the last eigen-time delay coordinates v_r because it cannot be modelled very well with linear or nonlinear dynamics. Therefore, we don't get a satisfying model for v_r. As a solution, we collect a streaming data from the dynamical

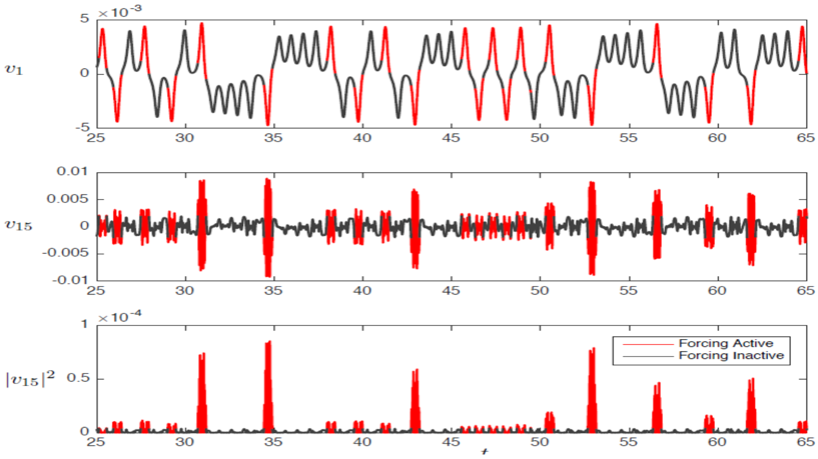

Figure 4: Delay coordinate v_1 of the Lorenz system, colored by the thresholded magnitude of the square of the forcing v_{15}[BBP+17]

system and the current value of v_r can be computed by taking the last bit of time history and dotting it with the columns of U. As a result, we will have a running window of v_r. Now, if we observe this running window as and when we collect the data, we can actually watch the forcing signal. During the observation, when the signal gets large, we can conclude that the system is about to do something nonlinear like lobe switching. In order to make the proper predictions, we set an upper threshold for the signal strength ${v_r}^2$. The plot of ${v_r}^2$ can tell us when the signal is larger than threshold as we can see from the figure 4. Therefore, we color the particular trajectory in red so that when we find the red trajectory, it says that the system is about to do something nonlinear. This way, we can actually see the signature of the system well before it actually switches to other lobe of Lorenz attractor so that we get almost full revolution to act upon it which makes the model predictive.

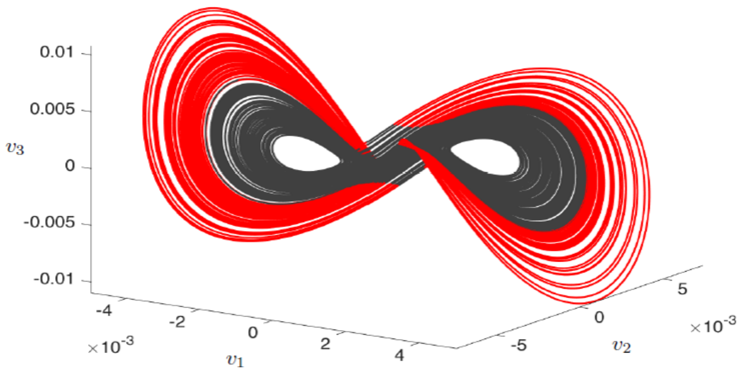

Figure 5: Color coded time delayed embedded attractor of the Lorenz system by the activity of v_{15}[BBP+17]

We can plot the color coded data on delay embedded attractor as shown in the figure 5. The red regions are where the external forcing is large and the black region is where the forcing is relatively small. When the forcing that the V_r is very small, the system can be well described by the linear part of dynamics. Therefore, the black region is like a Koopman linear model on time delay coordinates which is surrounded by three different fixed points: saddle point in the middle and two centers of each lobe in Lorenz attractor. So, the black region is where the forcing is not active and the system is well described by linear dynamics and the red region is where the forcing becomes large which is when the system does something nonlinear like lobe switching. When the trajectory turns red, we know that after one more revolution we are about to switch to another lobe. The black region is a Koopman linearized system and the red region on the outside contains trajectories with nonlinearity which can't be represented as a Koopman model.

The nonlinearities with trajectories are sometimes essential because multiple fixed points or multiple attractors cannot be represented by a finite dimensional representation of Koopman operator. This is because of the existence of only one fixed point at the origin for linear systems. Therefore, while modelling a chaotic or nonlinear system using Koopman analysis, there would be some regions of phase space where the linear model isn't very good and there will be in need of some nonlinearities like intermittent forcing.

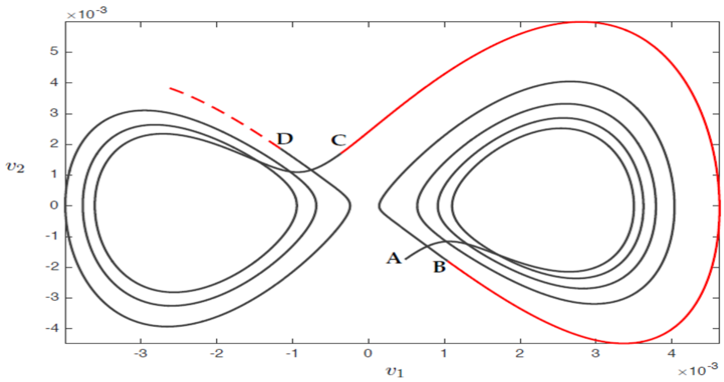

Figure 6: Illustration of one intermittent lobe switching activity[BBP+17]

For an illustration, figure 6 shows how lobe switching takes place in terms of revolutions. Firstly, we start at point A and do 4 inner revolutions until we reach point B. Due to red coloring of trajectory we can say that the nonlinear forcing become large from point B. This is the point where we get to know that the system is about to switch to the other lobe and exactly after one period or revolution later, it switches to the next lobe at point C. For the next 3-4 revolutions in the current lobe the forcing will be small and at point D we will come to know that the forcing becomes large which indicates that after one revolution we will go back to the right lobe indicating transient events.

6 HAVOK on other chaotic systems

HAVOK analysis can be done on any chaotic system which follows the same set of steps in order to create an embedded attractor along with the forcing statistics. The trajectories can be color coded to distinguish between normal and large forcing to represent intermittent events. For every system, a different measurement is considered to generate delay co-ordinates based on which the embedded attractors are created that are diffeomorphic to their original ones.

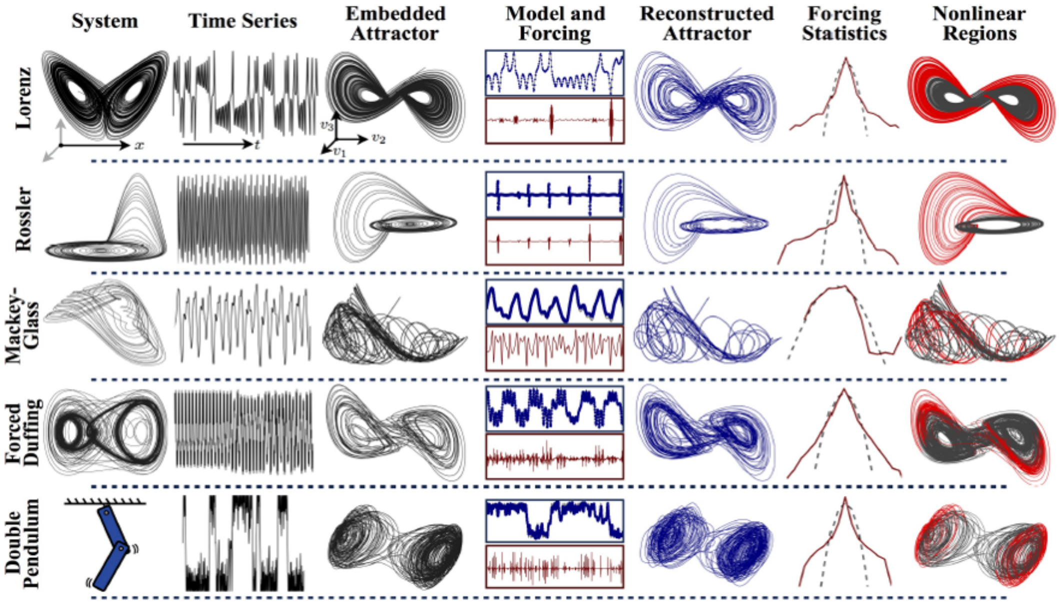

Figure 7: Results of HAVOK analysis on other systems[BBP+17]

7 Conclusion

This analysis has shown how data driven approach has been used to identify intermittent events in chaotic systems rather than going through the old approach of using governing equations. Oftentimes, we are not provided with dynamical systems but we can find tons of data. Therefore, we need to find the right data and right measurements which is why we make use of Koopman operator theory so that we can represent these measurements in a dynamical system either linear or nonlinear. With right measurements, we get delay co-ordinates using which we create an embedded attractor. Using attractor's coordinates, we create regression models which helps in constructing an attractor that distinguishes between linear and nonlinear terms using different colors. The nonlinear part signifies intermittent events and the linear part is basically what has been modelled properly. This is basically how we can go from data to dynamic representations that can be used to control and predict the so called chaotic systems.

8 References

[BBKK21] Steven L. Brunton, Marko Budišić, Eurika Kaiser, and J. Nathan Kutz. Modern koopman theory for dynamical systems, 2021.

[BBP+17] Steven L. Brunton, Bingni W. Brunton, Joshua L. Proctor, Eurika Kaiser, and J. Nathan Kutz. Chaos as an intermittently forced linear system. *Nature Communications*, 8(1), May 2017.

[BBPK16] Steven L. Brunton, Bingni W. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Koopman invariant subspaces and finite linear representations of non-linear dynamical systems for control. *PLOS ONE*, 11(2):e0150171, Feb 2016.

[CM79] A.J. Chorin and J.E. Marsden. *A Mathematical Introduction to Fluid Mechanics*. A Mathematical Introduction to Fluid Mechanics. Springer-Verlag, 1979.

[MSM20] Alexandre Mauroy, Yoshihiko Susuki, and Igor Mezić. *Introduction to the Koopman Operator in Dynamical Systems and Control Theory*, pages 3–33. Springer International Publishing, Cham, 2020.

[Tak81] Floris Takens. *Detecting strange attractors in turbulence*, volume 898, page 366. 1981.