

Removing Noise from Chaotic Dynamical Systems To Improve Accuracy of Takens Embedding and HAVOK

Igor Świerlikowski
Maastricht University
`i.swierlikowski@student.maastrichtuniversity.nl`

January 24, 2025

Abstract

This report explores the process of denoising chaotic dynamical systems and demonstrates the use of Takens Embedding and the HAVOK method for reconstructing and analyzing its dynamics. I propose a method for removing noise from chaotic signals and test the approach on the Lorenz Attractor system. My experiments show how some denoising techniques can be applied to minimize the impact of noisy measurements on our final reconstructed time delayed attractor.

1 Introduction

Chaotic dynamical systems are prevalent in many fields, including physics, biology, and engineering. Despite them being fully deterministic, chaos arises due to sensitive dependence on initial conditions, leading to behavior that appears random and is difficult to predict over long time horizons. One approach to characterizing such complexity is the Hankel Alternative View of Koopman (HAVOK) method [1], which tries to decompose complex nonlinear systems into a mostly linear framework with a separate nonlinear forcing terms. This decomposition can then be used for easier forecasting and analysis.

In practice, measurement noise is often unavoidable when studying chaotic systems. While such noise can't influence the underlying dynamics of a system, it is extremely difficult to describe using linear methods often "blinding" them. Consequently, efficient denoising techniques are critical for accurately reconstructing and predicting the true dynamics of these systems.

In this report, we focus on one main technique for analyzing chaotic systems:

- **HAVOK Method:** A data-driven approach for decomposing time series into dominant modes and dynamical features. HAVOK begins with a time-delay Hankel matrix (as part of Takens Embedding) and applies Singular Value Decomposition (SVD) to said matrix. Afterwards, it uses Sparse Identification of Nonlinear Dynamics (SINDY) [2] to represent the system with a main linear component and a separate nonlinear forcing term.

When this method is combined with a suitable denoising strategy, we can recover essential features of the underlying chaotic dynamics and potentially forecast intermittent events, such as lobe switching in the Lorenz system.

2 Methodology and Experiments

2.1 Data Description

In these experiments, I work exclusively with the Lorenz system, given by the following set of ordinary differential equations:

$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= x(\rho - z) - y, \\ \frac{dz}{dt} &= xy - \beta z,\end{aligned}\tag{1}$$

where the classical parameter set is

$$\sigma = 10, \quad \rho = 28, \quad \beta = \frac{8}{3}.$$

For the experiments, the Lorenz system is evolved from $t_0 = 0$ to $t_1 = 50$ with a time step $\Delta t = 0.001$. The initial conditions are set to $(x, y, z) = (-8, 8, 27)$. Numerical integration is carried out in MATLAB using the `ode45` solver with a strict error tolerance of 10^{-12} .

To simulate measurement imperfections, Gaussian noise is added to the time series *after* the system evolution is completed, thus modeling realistic sensor or measurement errors rather than influencing the underlying dynamics directly.

2.2 Noise Removal Techniques

Three primary methods are evaluated for noise removal:

- **Moving Average:** A simple method with only one parameter (*window size*) which takes the average of surrounding measurements to smooth the data.
- **Savitzky–Golay:** Instead of simply averaging, this method fits a low-degree polynomial to a set of neighboring points, thus preserving features like local maxima and minima better than a simple moving average.
- **Wavelet Denoising** [3]: This approach uses the Discrete Wavelet Transform, a technique similar to the Fourier Discrete Transform but instead of fitting sine and cosine waves, it uses wavelets instead. This separates the signal into components of different frequencies. Noise is usually concentrated in higher frequencies, so thresholding or cutting off the high-frequency components can effectively reduce noise while preserving the main structure of the signal.

These methods were selected to cover a range of filtering techniques and to evaluate their relative performance in preserving the chaotic structure while reducing noise. Each method’s parameters were optimized across multiple noise samples to ensure a fair comparison between them.

2.3 Takens Embedding

Takens Embedding creates a delay-embedded attractor from time series data. In the HAVOK framework, this method is used by constructing a Hankel matrix of time-delayed coordinates. Following [1], we choose:

- An embedding dimension (number of rows in the Hankel matrix) of 100.
- A delay time of 0.001, matching the numerical integration time step.

2.4 HAVOK Method

Once the Hankel matrix is formed, we perform SVD, yielding matrices U , Σ , and V^T . Let $\mathbf{v}(t)$ denote the relevant row components of V . In the HAVOK procedure:

- We select the first r rows of V (i.e., the most dominant modes). Typically, $r = 15$ is used.
- Then sparse regression (SINDY) is applied to the first $r - 1$ rows, while the r -th row is treated as an intermittent forcing term.

Hence, we arrive at a nearly linear system of the form:

$$\frac{d}{dt}\mathbf{v}(t) = A\mathbf{v}(t) + Bv_r(t), \quad (2)$$

where A and B are matrices identified through sparse regression, and $v_r(t)$ is the last mode representing nonlinear forcing.

2.5 Error Measure

To evaluate reconstruction performance, we compare the \mathbf{V} matrix of the true or target system to the \mathbf{v}_{sim} matrix (the reconstruction). A straightforward approach I used is:

$$\text{error}_x = \sqrt{\sum \left(V_x(t_{\text{span}}, 1:r) - v_{\text{sim}}(t_{\text{span}}, 1:r) \right)^2}, \quad (3)$$

where we sum over the relevant time points and dominant modes. However, as discussed later, this measure can be misleading because HAVOK does not treat all modes equally.

2.6 Experimental Procedure

Overall, the experiments follow this pipeline:

1. **Data Generation:** Numerically integrate the Lorenz system over $[0, 50]$ using $\Delta t = 0.001$.
2. **Noise Addition:** Inject Gaussian or uniform noise after integration.
3. **Denoising:** Apply one of the chosen noise removal techniques (Moving Average, Savitzky–Golay, or Wavelet).
4. **Takens Embedding & HAVOK:** Construct the Hankel matrix, perform SVD, and apply SINDY to form the HAVOK model.
5. **Analysis:** Compare the reconstructed attractors, measure predictive performance, and assess sensitivity to noise.

3 Results

3.1 Investigating the Cause of Poor Performance

My first experiment added a small amount of noise ($\text{sd} = 0.01$) to the Lorenz system. As we can see in Figure 1, the delay-embedded attractor (Takens Embedding) itself looks almost the same, and the main differences appear when we linearize with HAVOK. This indicates that the problem arises in the linearization step, not in Takens Embedding.

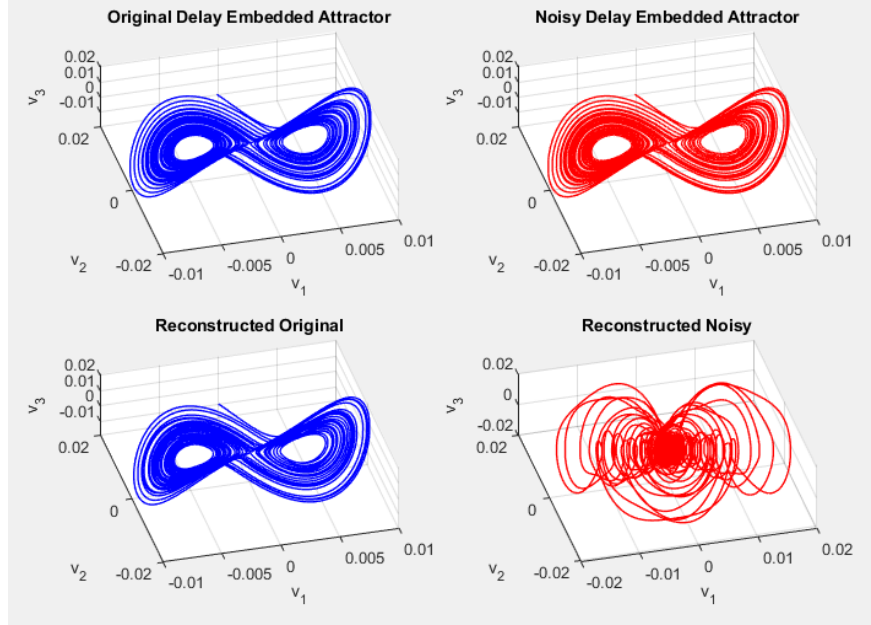


Figure 1: Comparing the original/noisy delay-embedded attractors and their reconstruction (small noise, $sd = 0.01$). The top panels show the embedded attractors (original vs. noisy), whereas the bottom panel show the HAVOK reconstructions.

3.2 Noise Removal Performance

Next, I applied Gaussian noise ($sd = 0.01$) to the Lorenz data and tested each denoising technique. Figure 2 shows the time-series (the x variable versus time), including the noisy and denoised signals, while Figure 3 shows a bar chart of the average RMSE across multiple trials for each method.

From Figure 3, Savitzky-Golay and Wavelet Denoising substantially lower error than the noisy dataset itself. Interestingly, the simple Moving Average filter performs worse on average than leaving the data noisy. This motivates focusing on Savitzky-Golay and Wavelet Denoising in future research.

3.3 Denoised HAVOK Reconstructions

With the same Gaussian noise settings, we now feed the denoised signals to the HAVOK pipeline. Figure 4 compares the delay-embedded attractors from noisy vs. denoised data, while Figure 5 shows the time evolution of an error metric (cf. Section 2.5) for each approach.

3.4 Error Measure Revisited

We observe that the error metric in Figure 5 often doesn't agree the actual quality of the reconstructed attractors. This occurs because our metric treats all rows in the \mathbf{V} matrix equally, while HAVOK itself focuses on the most dominant (first ones) modes for reconstruction. Figure 6 illustrates how the first few modes match closely, but noise takes over the higher modes that are less critical to the main dynamics. Future work will explore more nuanced error measures that weigh modes according to their importance in HAVOK.

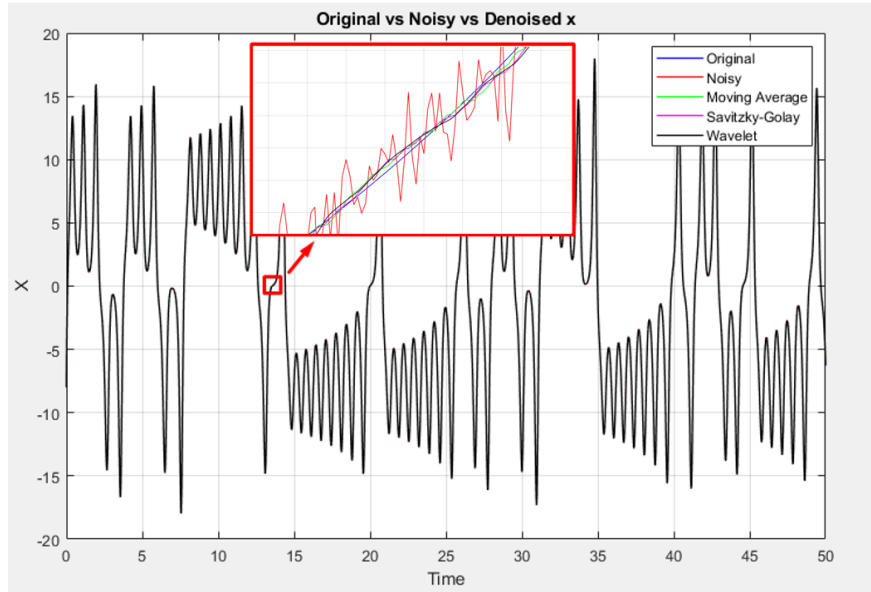


Figure 2: Graph of $x(t)$ for the Lorenz system with added noise and after denoising, including a zoomed-in portion to highlight subtle differences.

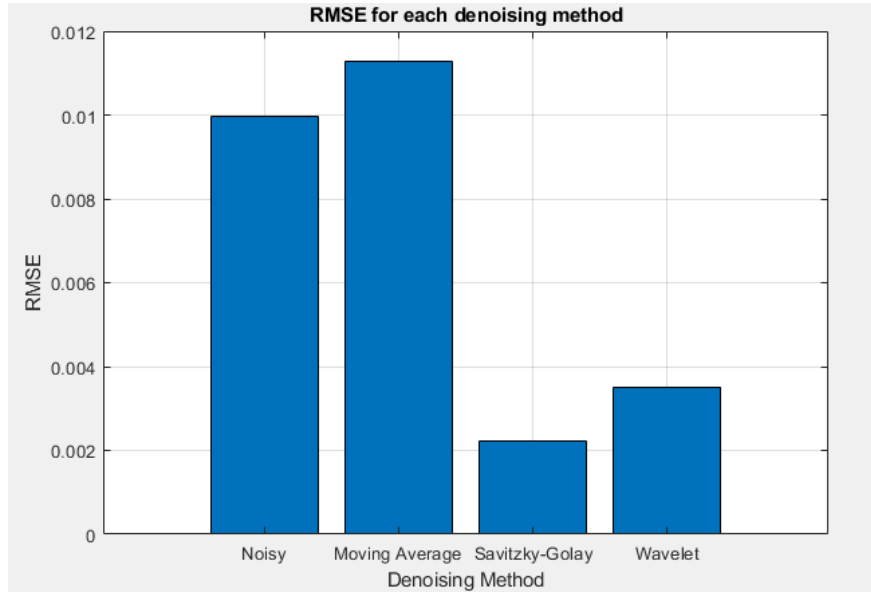


Figure 3: Bar chart of RMSE for different denoising methods. Lower RMSE indicates better alignment with the true (noise-free) signal.

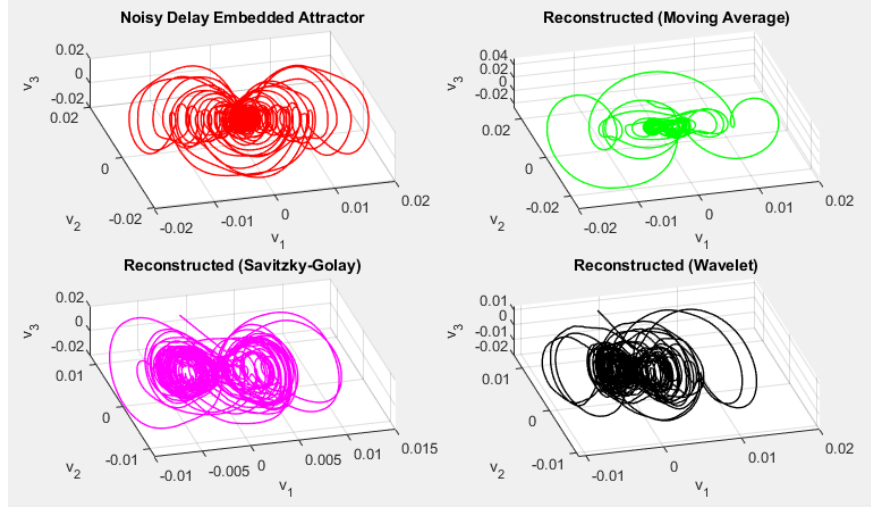


Figure 4: The noisy delay-embedded attractor compared to its denoised version (both using HA-VOK). Savitzky-Golay and Wavelet Denoising more closely preserve the two-lobe structure.

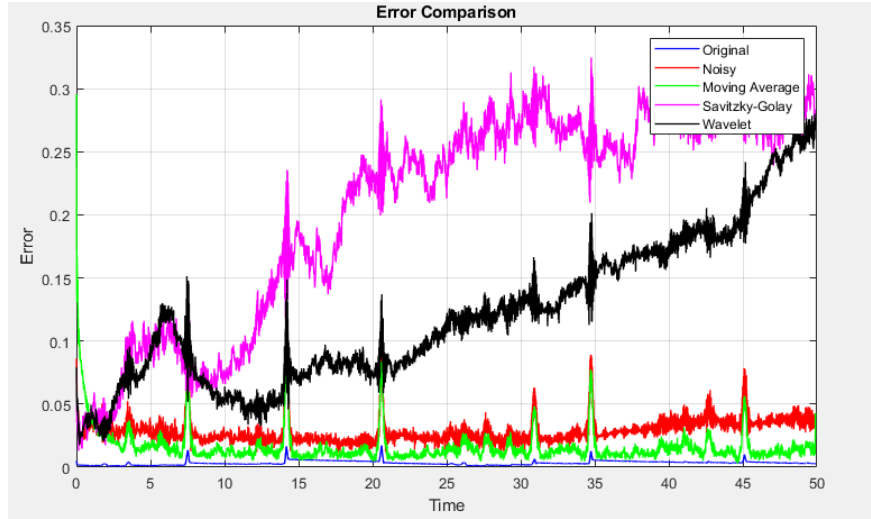


Figure 5: The time evolution of the reconstruction error for each denoising method. Although Savitzky-Golay and Wavelet appear worse by this metric, visual inspection reveals they preserve the Lorenz structure better (see text).

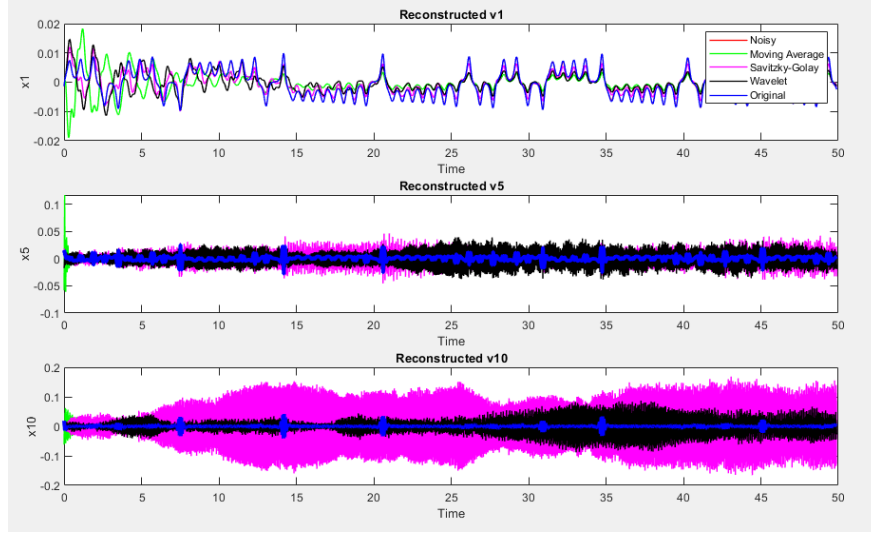


Figure 6: Comparison of the first, fifth, and tenth modes in \mathbf{V} for different denoising methods. The lower modes match closely, while higher modes exhibit very visible effects of noise.

4 Discussion

4.1 Investigating the Cause of Poor Performance

From Figure 1, we see that Takens Embedding is not the bottleneck but the linearization from HAVOK is what appears most sensitive to noise. One direction to mitigate this would be to increase r , the number of modes retained, so that small errors do not disproportionately affect the lower-dimensional model, but research is required to confirm it.

4.2 Noise Removal Performance

As shown in Figures 2 and 3, Savitzky-Golay and Wavelet Denoising are more effective than either leaving the signal noisy or applying a simple Moving Average. Because Wavelet Denoising offers many variable parameters (e.g., wavelet type and thresholds), it presents a rich area for further improvement so that will be my highest area of focus next semester.

4.3 Denoised Havok Reconstructions

Figure 4 demonstrates that Savitzky-Golay and Wavelet Denoising yield significantly clearer attractors with two distinct lobes. However, some noise remains, and improving these filters (especially wavelets) is an ongoing focus for the next semester. A more sophisticated wavelet thresholding strategy or a different wavelet choice might further enhance reconstruction.

4.4 Error Measure

Figure 5 highlights a key issue: a single RMSE measure across all modes of \mathbf{V} can be misleading. HAVOK prioritizes dominant modes, meaning that less important modes (which may contain mostly noise) can inflate the error without affecting overall attractor fidelity. Thus, a weighted error measure would be more appropriate and the first focus for next year to ensure good comparisons.

5 Research Questions

Looking ahead, the following questions guide the next phase of this study:

1. How can we effectively remove measurement noise from a chaotic time series while preserving the underlying structure?
2. To what extent does noise removal improve the accuracy of the HAVOK method for reconstructing the attractor?
3. How does the HAVOK method, combined with noise-reduced data, improve our ability to predict and model chaotic systems?
4. What are the limitations of the proposed denoising techniques when applied to different types and levels of noise?
5. What error measures can we use to better reflect the dominance of lower HAVOK modes?

6 Conclusion

In this study, I explored methods for removing noise from chaotic dynamical systems, specifically with of the Lorenz attractor, to improve the accuracy of HAVOK’s linear reconstructions. Through the application of three noise reduction techniques—Moving Average, Savitzky-Golay, and Wavelet Denoising—I evaluated their performance in preserving the real chaotic structure while removing as much of the noise (measurement error) as possible.

The results showed that Moving Average was least effective, usually performing even worse than leaving the data noisy. By contrast, Savitzky-Golay and Wavelet Denoising performed significantly better, retaining the attractor’s overall shape and reducing noise levels for the HAVOK procedure. Moreover, we made sure that the main sensitivity to noise lies in the linearization step of HAVOK rather than the Takens Embedding. This suggests possible improvements, such as increasing the number of modes r to capture the more of the important dynamics.

We also highlighted that the standard RMSE metric across all modes of the \mathbf{V} matrix is a poor measure of error. Because HAVOK focuses on the most dominant rows, noise in the higher modes can increase error measurements without drastically changing the reconstruction. Future work will focus on developing better error measures and on refining wavelet-based denoising by experimenting with different wavelet types and ways of thresholding.

Finally, these advances will help make data-driven modeling of chaotic systems more robust to noise which will make HAVOK a far more useful method, with potential applications in meteorology, finance, biological systems and many more.

References

- [1] Brunton, S. L., Brunton, B. W., Proctor, J. L., & Kutz, J. N. (2017). *Chaos as an intermittently forced linear system*. Nature Communications.
- [2] Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2017). *Discovering governing equations from data: Sparse identification of nonlinear dynamical systems*. PNAS.
- [3] Luo, G., & Zhang, D. (2012). *Wavelet Denoising*. In *Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology*. IntechOpen. <https://doi.org/10.5772/37424>