Using Epsilon-Machines to understand Pulsar Variability and Nulling

Felix Weber[1,2], Sofia Sheikh[3], Clément Vidal[2], Stuart Bartlett[4]

[1] Oberlin College, Oberlin, OH
[2] Berkeley SETI Research Center, UC Berkeley, CA
[3] SETI Institue, Moutain View, CA
[4] Caltech Pasadena, CA

August 11, 2023

Abstract

We present a novel technique to study the inherent stochastic complexity and randomness of pulsar variability in a comparative case study of two non-nulling (J0332+5434, J0358+5413) and two nulling pulsars (J0953+0755, J1136+1551). Using epsilon-machines we can attempt to recreate the fundamental stochastic behavior of unknown processes at different resolutions [Brodu and Crutchfield, 2020], producing 'stochastic fingerprints' that can be used to compare and contrast different stochastic models [Feldman et al., 2008]. We also developed two suites of software to help analyze our data and extract pulse intensity time series. Using high-resolution data we collected with the Allen Telescope Array, we showed that J0332+5434 and J0358+5413 exhibit statistical complexity beyond Gaussian noise models, instead their stochastic characteristics are better modeled by simple stochastic nulling, suggesting greater similarity with nulling pulsars than first thought. J0953+0755 showed agreement with stochastic nulling, however J1136+1551 exhibited stochastic behavior not comparable to the benchmarks we used. We suggest further study with similar methods of more pulsars to help better understand how stochastic behavior may vary across different pulsars.

1 Introduction

Pulsars have long been a fascinating subject for many astronomers, allowing for a deeper understanding of interstellar space and time, yet they have presented us with many other confounding open puzzles. One such puzzle is the apparent "nulling" of pulsars, where we fail to observe one or more pulses from a pulsar in question. Long-term nulling — the loss of pulses over several tens or more of pulses — has been found to be explained by scintillation and caustic lensing: the active refraction of a pulse away from the earth due to the interstellar medium. However, short-term nulling — the loss of only one or a couple of pulses — has evaded any testable hypothesis. These short nulls have mostly been identified via statistical analyses, showing two (or more) resolvable, but overlapping distributions of pulse intensities.

Current thought states that fluctuations in pulsar intensities are uncorrelated and stochastic, including short-term nulling. The physical mechanism behind nulling is also not known, and as such, any complete theory would need to reproduce the stochastic behavior observed in pulsar nulling and pulsar variability. One school of thought argues that nulling is a special process not seen in 'standard' pulsars, with some studies having suggested that variations in pulsar intensities are the result of Poisson-like processes [Gajjar et al., 2012]. Since nulling is a result of pulsar intensity 'going to zero', we can test some of these claims by investigating the stochastic variability of both nulling and non-nulling pulsars. If 'nulling' pulsars are special types of pulsars, their stochastic variability should differ from 'standard' pulsars.

To better understand and build stochastic models, we implement epsilon-machines. Epsilon-machines are a type of Markovian automation where the states of the machines are defined by the 'causal states' of the stochastic system in question. These are groupings of system states with similar past and future system states, hence 'causal' [Brodu and Crutchfield, 2020]. These machines allow us to measure the 'statistical complexity' of a stochastic process — i.e. how complicated it is to describe a stochastic process. The main

advantage of this metric is that it gives 'pure' random processes low scores, while processes with rigid 'rules' very high scores [Shalizi and Shalizi, 2004]. These machines and metrics give us qualitative and quantitative tools to discriminate between fundamentally different stochastic processes, as will be shown later.

To analyze the variability and nulling of different pulsars, we directly observed with the Allen Telescope Array (ATA) with high resolution (16 μs). We also created a pipeline for automated RFI removal, de-dispersion, and folding of the observed power over time. We then also created an automated system to find pulses and remove the background such that we can find the integrated flux of observed pulses. Doing so and correlating across a large bandwidth of different observing nodes, allowed us to extract a final time series of pulse intensities independant of fine-scale scintillation effects.

Ultimately, we show a proof-of-concept of the application of epsilon-machines to pulsar intensity data to compare and contrast different stochastic models and theories. Through such argumentation we can reject models such as uncorrelated Poisson variability and Gaussian noise. Further, we find that there is not yet a clear stochastic difference between nulling and non-nulling pulsars with respect to the variation of their intensities over time, suggesting that nulling may not be an inherently different physical process from normal pulsar variability.

2 Pulse Intensity Extraction

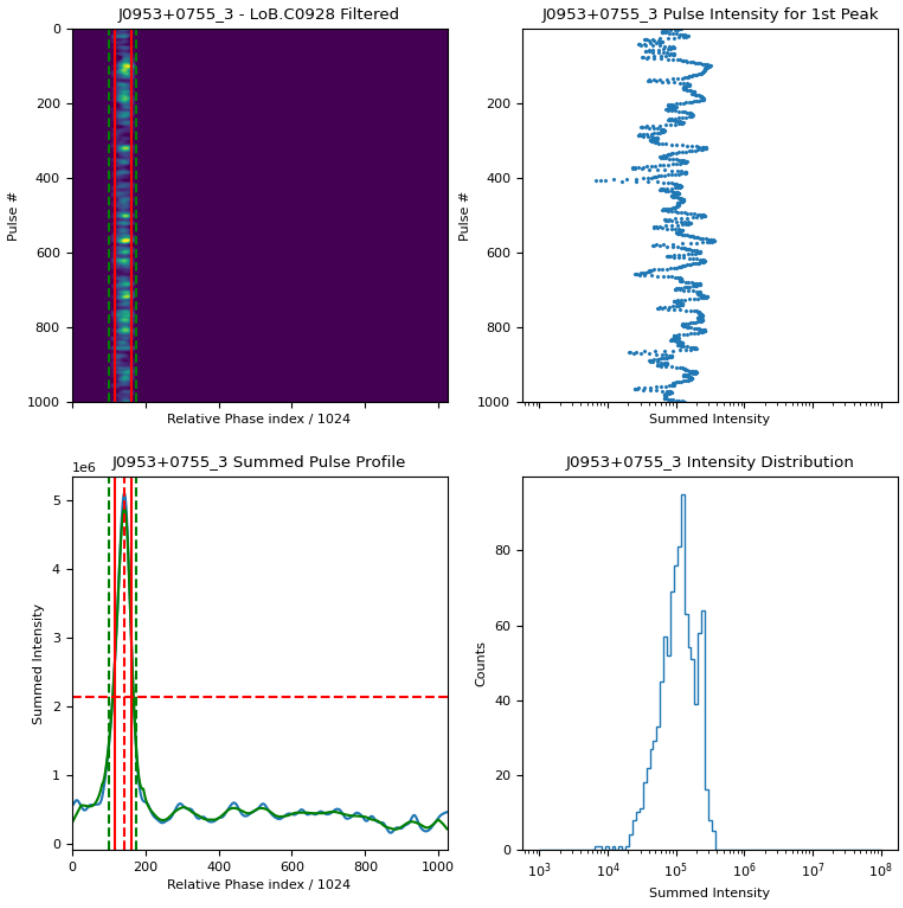


Figure 1: A diagnostic plot produced when extracting the pulse intensity of node LoB.C0928 for J0953+0755. The upper-left plot depicts our data after background filtering; the lower-left plot is the pulse profile folded over all observed pulses; the upper-right plot shows the integrated flux within the red bands of the left plots; the lower-right plot depicts a logarithmic histogram of extracted pulse intensities.

To extract accurate measurements of pulse intensity over time, we have developed our own suite of software to help in filterbank file processing (the native ATA file type). The first of the two softwares we use, P.I.P.E., has been developed as a wrapper for existing programs meant for RFI removal, dedispersion, TEMPO corrections, and accurate folding. In most cases we used SIGPROC command-line software [Lorimer, 2011] such as *rfifind*, *prepdata*, and *dedisperse*, alongside incorporating important corrections using TEMPO [Nice et al., 2015] before folding. All these programs have been wrapped together to calculate the observed power over time for each observation node (see section 4).

Our second piece of software takes the previously calculated power-series for each node and attempts to filter out pulses from the background. Figure 1 depicts a diagnostic plot of this process on data we collected from J0953+0755. One important step we include is temporal smoothing, as not all RFI is successfully removed. This is done using a Gaussian filter with a sigma of 8 relative phase bins (0.78% of a pulse period).

We also implement some periodic smoothing with a Gaussian filter with a sigma of 4 periods. We recognize there are some pros and cons to this approach. Firstly, this means that the time series we are recovering is not a direct measurement of the pulse intensity, but instead a local measurement of the expected intensity. This does help get rid of a lot of noise and allows us to visually inspect correlated changes in intensity over time.

To filter out the background, we use a simple peak finding algorithm to locate the pulse peak and widths at different pulse heights [Virtanen et al., 2020]. As a filter, we only investigate peaks in the summed pulse profile above one third of the way to the maximum value from the average intensity. If there are more than two peaks in that region, we flag and remove that observation due to potentially low SNR or interfering RFI. We then mask the pulse at 20% pulse height and interpolate the local background using a 2-degree spline. Using lower pulse heights for masking tends to over-mask our pulses, however this parameter can be changed on an observational basis to help better filter the background. Figure 2 depicts a result of this process.

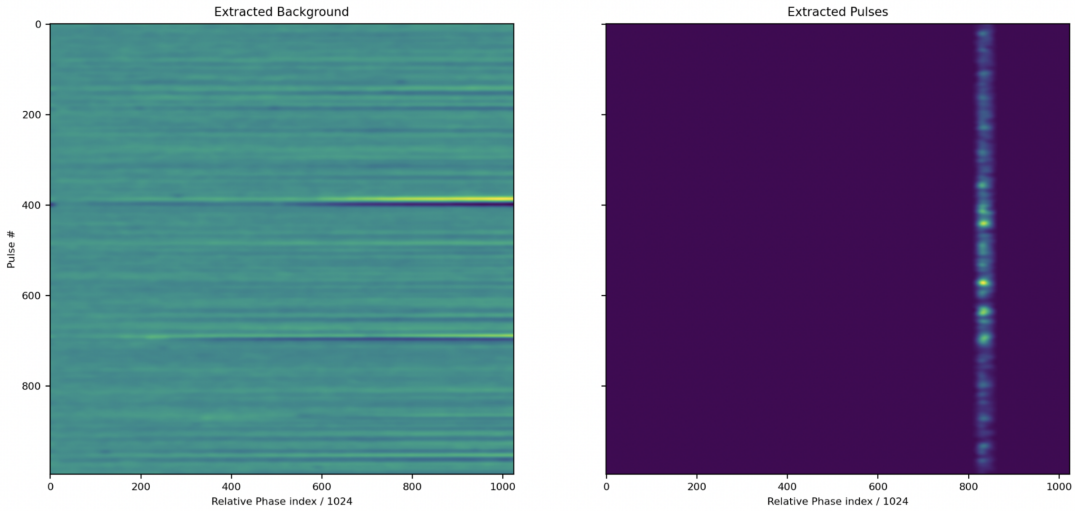


Figure 2: A typical result of background extraction. The left plot depicts the background, where the part under the pulses has been interpolated. The right plot depicts the result of taking our data and subtracting the background, clearly visualizing pulse intensity over time.

One effect that we want to avoid picking up in our extracted time series is scintillation drastically altering our observed power due to constructive and destructive interference in the ISM [Stinebring et al., 2022]. There are two ways we mitigate these effects: by limiting the time over which we observe, and by integrating observed power over a large bandwidth (0.5 to 1 GHz). The former allows us to effectively remove fine-scale scintillation, characterized by small scintillation bandwidths, while the first mitigates large-scale scintillation that effects brightness curves over large amounts of time. Most of our observations took place over 2 to at most 20 minutes, and as will be discussed later, we saw much more variability than could be explained by scintillation models.

Figure 3 shows a diagnostic plot of processed nodes from an observation of J0332+5434. The selected nodes are those that have been successfully filtered from background. Visually, this is a useful tool that

demonstrates clear correlation in pulsar intensity across observing nodes. The summed intensity time series shown in figure 3's right plot is what can then be used to study a pulsar's intrinsic variability.

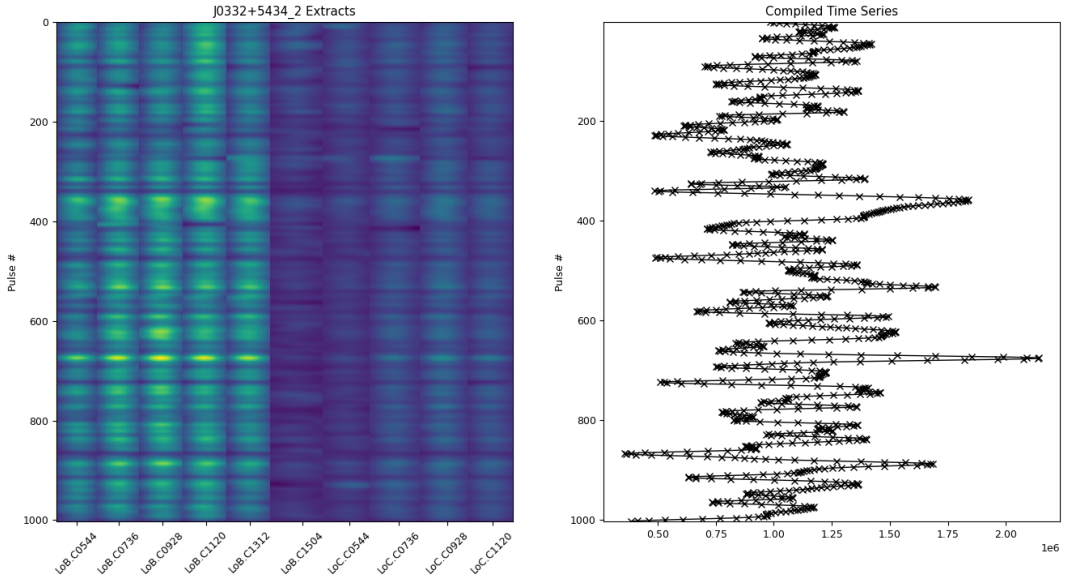


Figure 3: A diagnostic plot generated while compiling successfully filtered nodes from an observation of J0332+5434. Clear correlated behavior can be seen across nodes in terms of pulsar intensity.

3 Understanding Epsilon-Machines

Epsilon-machines were created as an answer to the question: "Given time series data produced by some stochastic process with certain rules, how can we begin to recover those rules from the given data?" Specifically, epsilon machines are a type of Markovian automation, whereby a stochastic process is modeled as a probabilistic walk between 'causal states'. Each 'causal state' is defined by the set of system states with similar future states - hence 'causal' Shalizi and Shalizi [2004]. Finding these causal states is done using the Causal State Splitting Reconstruction (CSSR) algorithm, which has been expanded and implemented by Brodu [2011] for symbolic and numerical data.

The power of this method is that it assumes the simplest process (pure randomness) and only increases complexity if statistically justified, alleviating issues with over-fitting. As an example, consider the symbolic even process: a random number of 0s followed by a random even number of 1s and so forth (e.g. 0011011110001100011111101...). In this case there are two causal states: one where the next symbol has equal chances of being 0 or 1, and one where we know there must be a 1 since there are an odd number of 1s. We can therefore create a simple epsilon-machine to model and recreate the even process:

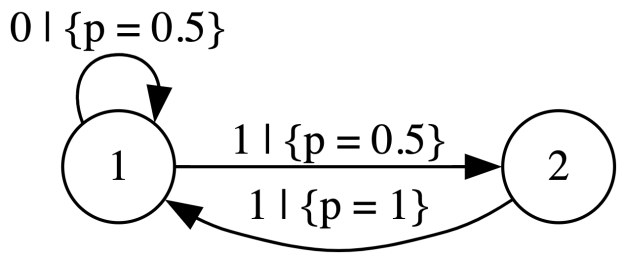


Figure 4: An Epsilon-Machine that accurately recreates the symbolic even process.

To 'read' figure 4, start in one state (one of the nodes), then with the given probabilities chose a transition (an arrow from the current node) to another state. Each arrow gives the transition symbol - i.e. the next symbol in the time series. Running this through your head a few times should convince you that figure 4 will always give an even process as an output. I'll note that figure 4 is an idealized output of CSSR. Often times there are issues with how much data is required to reconstruct an accurate machine, and other times starting states may interfere with accurate reconstruction [Shalizi et al., 2002]. Consider the even process starting with a '1': there is not enough information to know which state we actually are in, and therefore the starting symbol will be in it's own state in reconstructed models.

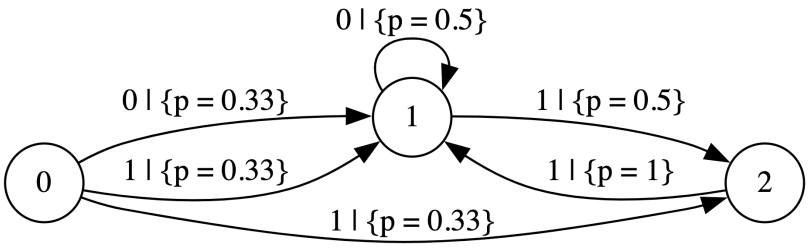


Figure 5: An Epsilon-Machine recovered with CSSR from even process data. State 0 represents starting with a '1', while the other two states represent the standard even process shown in figure 4

This issue of starting states reflects something rather powerful about CSSR: the ability to pick up on the smallest of statistical effects. This is what makes it a powerful tool in understanding stochastic and dynamic data. It can also be applied beyond pure symbolic series and applied to numerical data as well (since numbers are nothing but symbols as well), however there are some caveats. Numerical data is often too fine: 0.00001 and 0.00000 are effectively the same measurement in many contexts, yet are two different symbols. To get around this, we need to chose a 'kernel width' to bin close, but different numbers into the same symbolical categories. Beyond this binning, the application of CSSR is completely agnostic to any human biases as to how numbers 'should behave'. In fact, this extra parameter is useful in understanding how a process behaves at different 'resolutions'. Using large kernel widths is akin to listening to the beat in music, whereas smaller widths allow us to listen to changes in finer structures like harmonies and keys.

To characterize epsilon-machine behavior there are two important metrics: 'statistical complexity' ($\mathbb{C}_\mu$) and the 'entropy rate' (μ_r) of the recovered process. $\mathbb{C}_\mu$ categorizes the complexity of the epsilon-machine in question and is calculated by equation 1 (where α is a node in the machine, and p_α is the probability of that node). It is defined as the minimum amount of information, as measures in bits, required to describe the recovered epsilon-machine. μ_r, on the other hand, directly measures how random the recovered process is via the increase in uncertainty of future states for each future state predicted, and is calculated by equation 2 (where $H(L)$ is the expected entropy of sequence of length L).

$$\mathbb{C}_\mu = -\sum_\alpha p_\alpha \log_2 p_\alpha \tag{1}$$

$$\mu_r = \frac{1}{L} \lim_{L \to \text{inf}} H(L) \tag{2}$$

Philosophically, these metrics can be interpreted in different ways. For example, μ_r can be understood as the 'fundamental randomness' of a process that cannot be 'explained away' when accounting for all correlations present in the data. It also is analogous to thermodynamic entropy, giving a kind of 'temperature' to the process. $\mathbb{C}_\mu$ on the other hand tells us how hard a process is to describe. Taken together, these two metrics allow us to understand how well an epsilon machine is able to reconstruct an unknown process, as well as uncover some fundamental aspects of the given process.

Going back to numerical data, plotting $\mathbb{C}_\mu$ and μ_r for different kernel widths k, allows us to investigate the complexity and randomness of a process at given scales. Imagine some process where small oscillations directly influence the next state — this is something that would be not seen in large-k epsilon machines, but would then decrease μ_r for small-k epsilon machines as what initially appeared random is now accurately picked up by CSSR. These curves produce a type of 'stochastic fingerprint' that is directly related to the

stochastic mechanism in question. This allows us to compare and contrast different stochastic models on numerical data. Figure 6 shows an example of such a fingerprint that has been calculated from 2,000 points of a random, uncorrelated poissonian process.

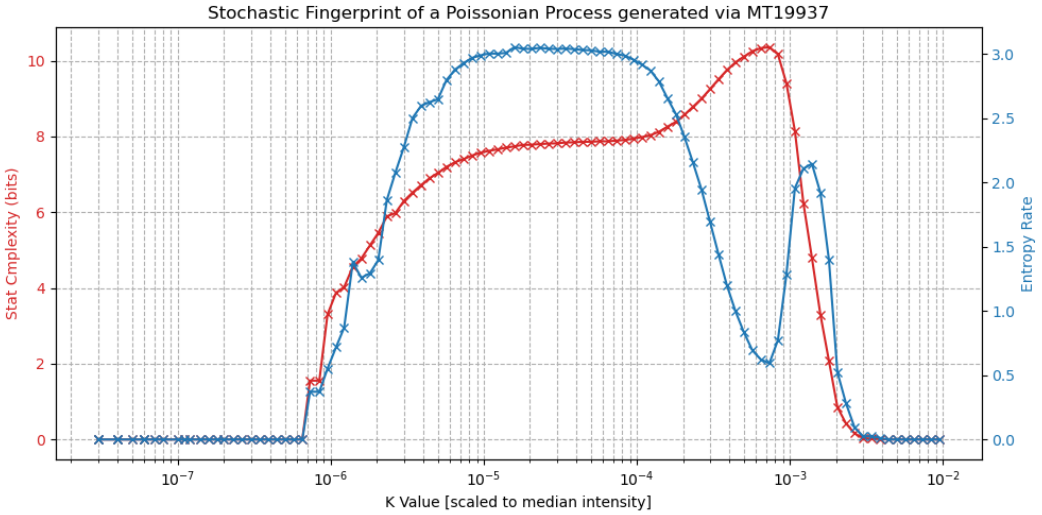


Figure 6: A 'stochastic fingerprint' calculated using CSSR on 2,000 points from a random, uncorrelated poissonian process.

One reason why these graphs are more useful to use than looking purely at an epsilon-machine for numerical data, is that numerical processes require magnitudes more information to accurately model due to higher degrees of freedom. In figure 5, we see areas for large- and small-k where both $\mathbb{C}_\mu$ and μ_r are 0 due to under- and over-resolving, respectively, our time series. Ironically, a random uncorrelated process has $\mathbb{C}_\mu = 0$ as there is only one causal state (see eq. 1). This gets to an important nuance of our stochastic fingerprints: they do not necessarily show us which epsilon-machine is most accurate, but it does show how a stochastic process inherently reacts to different resolution treatments.

It is with these new tools that we can begin to investigate the variability of pulsars and potentially many other phenomena.

4 Pulsar Observations

To test these new tools we selected pulsars on three criteria: nulling fraction, brightness at 1400 MHz, and availability at the Allen Telescope Array. Since we are trying to extract singular pulse fluxes, brighter pulsars allow for a higher SNR after pulse extraction. Annoyingly, many observable nulling pulsars at the ATA are relatively dim (<15 mJy at 1400 Mhz), so our selection came down to two individual pulsars: J0953+0755, J1136+1551. Secondly, as a benchmark to compare variability of our chosen nulling pulsars with, we observed J0332+5434 and J0358+5413. We recognize this is no where close to a comprehensive study of nulling and non-nulling pulsars, however we hope this selection will help highlight the potential of our proposed method. Table 1

Pulsar	Observations (Processed)	NF (%)	S_{1400} (mJy)	P (s)
J0332+5434	3 (2)	<0.5	203	0.714519699726
J0358+5413	3 (2)	0	23	0.156384121559
J0953+0755	5 (5)	5	100	0.253065165
J1136+1551	3 (2)	15	20	1.187913065936

Table 1: An overview of the pulsars we observed. Each observation was of roughly 1,000 pulses which were then extracted through P.I.P.E. and NullAnalysis.

We observed our targets using using the Allen Array Telescope, producing an effective diameter of 140 ft with 20 receiving antennae. All observations were recorded with 16 μs resolution across 14 nodes, each with 96 MHz bandwidth. Together, we observed from 899.5 to 2243.5 MHz: a total bandwidth of 1.344 GHz. Table 2 gives specific frequency information for each node. Before observing sessions we calibrated the ATA using standard quasar sources, such 3c48 and 3c84, over 10 minute observing sessions.

Node	f_0 (MHz)	f_1 (MHz)
LoB.C0352	899.5	995.5
LoB.C0544	995.5	1091.5
LoB.C0736	1091.5	1187.5
LoB.C0928	1187.5	1283.5
LoB.C1120	1283.5	1379.5
LoB.C1312	1379.5	1475.5
LoB.C1504	1475.5	1571.5
LoC.C0352	1571.5	1667.5
LoC.C0544	1667.5	1763.5
LoC.C0736	1763.5	1859.5
LoC.C0928	1859.5	1955.5
LoC.C1120	1955.5	2051.5
LoC.C1312	2051.5	2147.5
LoC.C1504	2147.5	2243.5

Table 2: Beginning (f_0) and final (f_1) frequencies of the nodes used at the ATA in L-Band observation settings. Each node has a bandwidth of 96 MHz, for a total observational bandwidth of 1.344 GHz.

In total we observed and successfully recovered over 14,000 individual pulses in an attempt to unearth the stochastic behavior of pulsars. We hope future studies will go beyond our own observational scope to observe several times more pulses than we have. Ultimately, the more data we can process, the more we will be able to extract about the processes that govern the unknown.

5 Stochastic Analysis with Epsilon-Machines

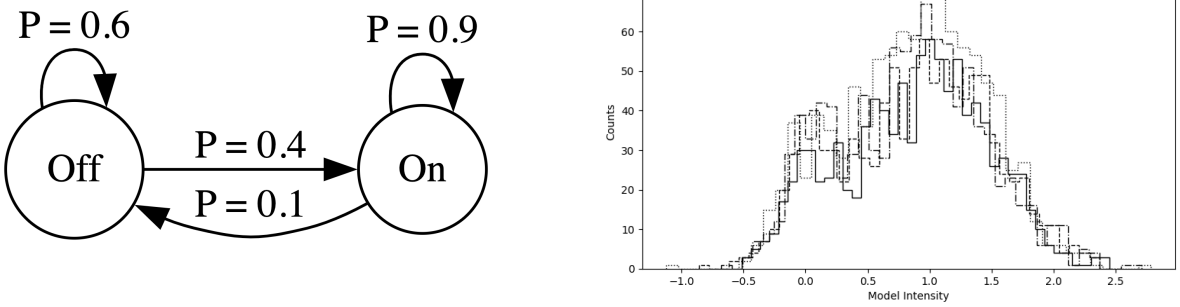


Figure 7: A simple stochastic nulling model used to generate benchmark data. NF is 13%. For an "On" state, an intensity is chosen from a normal distribution, $\mathcal{N}(1, \sigma_1)$, while for an "Off" state, we chose from $\mathcal{N}(0, \sigma_0)$, where σ_0 and σ_1 are set to reflect the noise of each state.

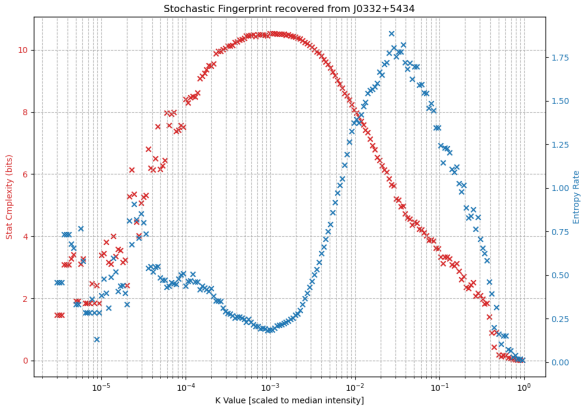
Taking the recovered intensity time series, we can group them together for each pulsar and run them through CSSR for different kernel widths to produce stochastic fingerprints for each pulsar. As a method to normalize the data, we divided by the median intensity of the recording. This is to help negate the effects of scintillation altering the baseline intensity observed from a pulsar over large periods of time. Further, this won't have any effects on our analysis as it is agnostic to geometric translations of our data.

To help interpret our data, we also prepared several benchmark observations assuming basic stochastic models. We selected three benchmark models: a randomized Poissonian process, a randomized Gaussian

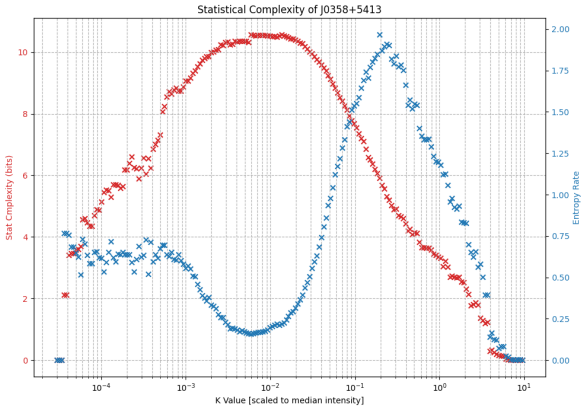
process, and a simple nulling model with Gaussian noise. The third model was generated using a simple Markov-chain model to choose an "On" or "Off" state, from which we choose from an associated normal distribution. Figure 7 shows an example of such a model, alongside the associated intensity distributions that one would observe in four modeled observations.

With each benchmark model, we made sure to treat each time-series the same way we have treated our actual observational data, primarily through pulse smoothing and median division. This is to ensure that differences between observed data and modeled data are the result of different stochastic processes, rather than subtle differences in processing.

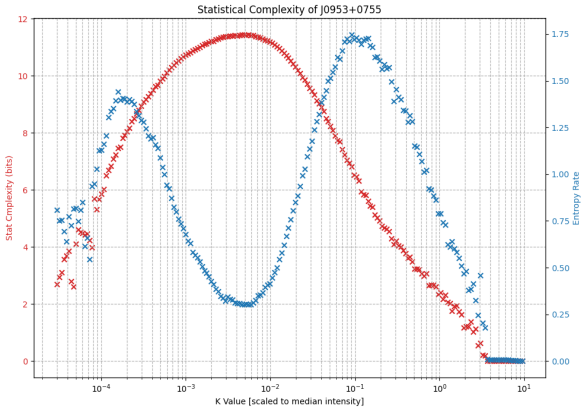
Figure 8 depicts the calculated stochastic fingerprints of the different pulsars we have observed. Visually, we see a clear coherence between the different pulsars, except for J0953+0755. This is due to having more observations on hand, allowing for a more complete fingerprint. On the whole we find maximum statistical complexities around 10.5 bits to 11.5 bits, depending on the amount of observations on hand. We also find minimum entropy rates between 0.2 to 0.3 bits, and maximum entropy rates of 1.7 to 2.0.



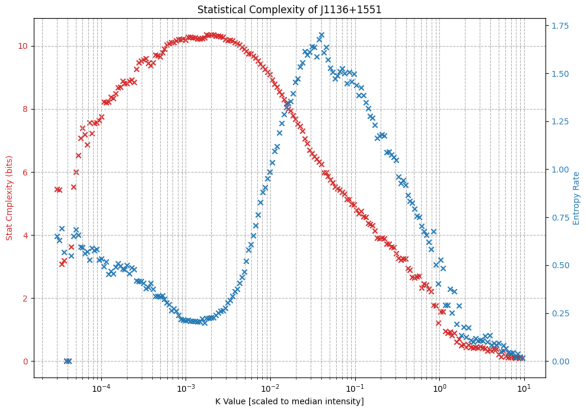
(a) J0332+5434: 2 Observations



(b) J0358+5413: 2 Observations



(c) J0953+0755: 5 Observations



(d) J1136+1551: 2 Observations

Figure 8: Stochastic Fingerprints derived from collected data. Statistical Complexity, $\mathbb{C}_\mu$ is plotted in red, and entropy rates, μ_r, are plotted in blue.

To help better compare these complexity and entropy rates across processes, we can use what are known as "complexity-entropy" diagrams. These help remove the spatial component of our analysis — the kernel width — and allow for a more agnostic approach (a similar but slightly different approach is used in Feldman et al. [2008]). To show how this approach is useful, figure 9 shows the stochastic fingerprint of J0953+0755 as compared to the fingerprint produced by our Gaussian process model for the same number of observations. One way to read such a graph is that each dot represents a single model derived from the data through CSSR. If two processes are stochastically equivalent, epsilon-machine models derived with CSSR will have

the same entropy rates and statistical complexities. We can clearly see some convergent behavior between a Gaussian process and our observed intensities, however the entropy-complexity curves exhibit important divergent behavior. For example, J0953+0755's most complex models (11.44 bits) are not only more complex than those of the Gaussian Process (11.19 bits), but also exhibit lower entropy rates (0.30 bits, 0.42 bits, respectively). Secondly, J0953+0755 exhibits much higher entropy rates in the complexity range between 4.0 and 10.0 bits. This usually indicates the presence of some non-random process that simpler models perceive as inherent randomness, also increasing the maximum statistical complexity recoverable by CSSR. It is safe to argue that the Gaussian process cannot be applied to J0953+0755.

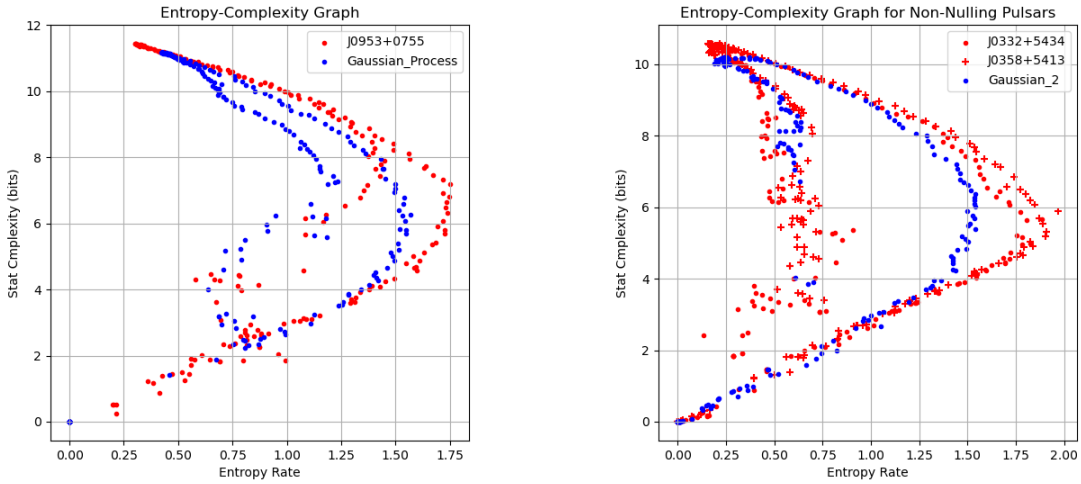


Figure 9: Left: a comparison of the entropy-complexity curves of our J0953+0755 data and modeled data produced from a simple Gaussian process for 5 observations. Some convergent behavior is visible, however it is clear that the Gaussian process does not fully encapsulate the process that generates J0953+0755's brightness. Right: a comparison of the EC curves of our non-nulling pulsars and modeled data produce from a simple Gaussian process for 2 observations. There is some greater coherence, however the pulsars still exhibit non-convergent behaviors in some regions.

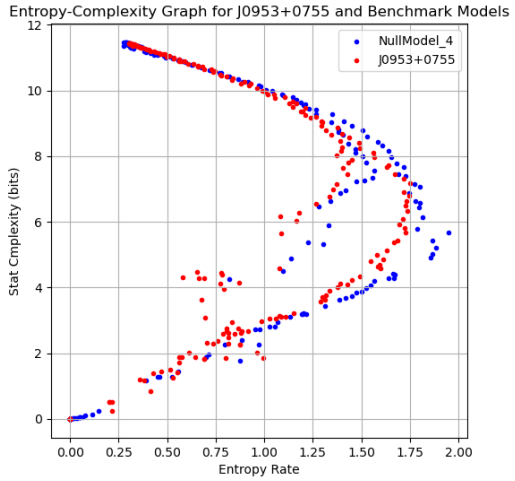
J0953+0755 is a known nulling pulsars, so what happens if we compare a simple Gaussian process to our two non-nulling pulsars: J0332+5434 and J0358+5413? Figure 9 also shows a comparison between the entropy-complexity curves of our non-nulling pulsars and a Gaussian process for the same number of observations. As in the last case, we see more coherence in some ares, but our pulsar data still exhibits greater statistical complexity (10.53 bits, 10.56 bits for J0332+5434, J0358+5413 respectively) than the Gaussian process (10.20). Secondly, for both pulsars, the recovered epsilon-machines exhibit higher entropy rates in the complexity range from 4 to 8 bits, hinting at some process not modelled by the Gaussian. However, how do these two pulsars compare to a simple nulling model?

Figure 10b shows a comparison between our non-nulling pulsars and the entropy-complexity curves of a simple nulling model. Surprisingly, the entropy-complexity curves overlap almost entirely! This suggests that the process of pulsar variability in the non-nulling pulsars that we have observed is more stochastically similar to current nulling models than a simple Gaussian noise model. If such a result is observed in further pulsar studies, it may suggest that the mechanism by which nulling pulsar brightness varies is universal across all pulsars. This raises the question as to why this mechanism manifests itself in such a way that some pulsars null, while others don't exhibit nulling or modal changes.

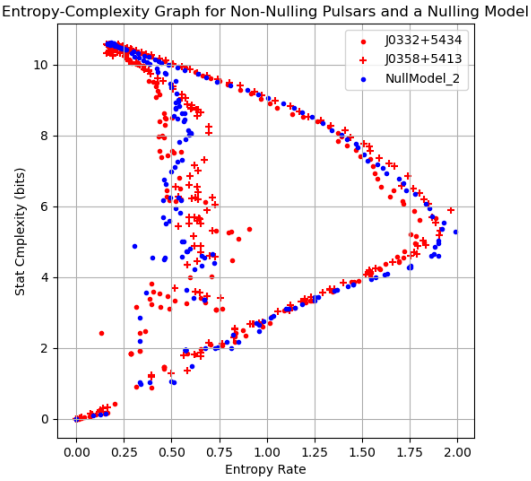
Likewise, figure 10c shows an improved coherence between the recovered EC curves for J0953+0755 and nulling model. There are some slight discrepancies: for example the nulling model displays some higher entropy rates in some regions. This may hint that the nulling model is slightly more random than actual nulling pulsar intensity curves, suggesting some underlying order to J0953+0755. What this may be will require further study and model comparison.

On the other hand, figure 10c shows that our second nulling pulsar's EC curves don't exactly match with either a stochastic nulling model or a Gaussian noise model. This could be due to a whole host of factors but it does warrant further investigation. It may be due to a more complex modal process beyond a simple "on"

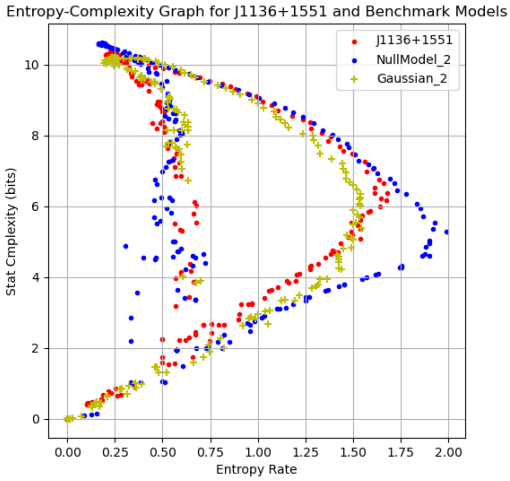
and "off" system. However the space of different models will take time to investigate. If anything, it appears that the variations observed from J1136+1551 are stochastically somewhere in-between a nulling model and a Gaussian noise model in terms of complexity, yet inherently less random at similar complexities.



(a) J0953+0755's EC curves compared to a nulling model over 5 observations.



(b) J0332+5434 and J0358+5413 EC curves compared to a nulling model over 2 observations.



(c) J1136+1551's EC curves compared to a nulling and Gaussian model over 2 observations.

Figure 10: An overview of our observed pulsar's EC curves compared against their closest benchmarks. Surprisingly, non-nulling pulsars agree the best with a nulling stochastic model. J0953+0755 is generally coherent with a nulling model but displays some lower entropy rates, suggesting lower levels of randomness. J1136+1551 does not agree with either and warrants further modeling and observation to help find a better approximate model.

Overall, it is curious that non-nulling pulsars are stochastically more similar to a simple nulling model that the standard assumption of Gaussian noise. This may hint that even standard pulsars are more complex and intricate than first thought. Also curious is that one of our two nulling pulsars did not stochastically agree with a nulling model, while the other does. All this raises questions as to how much entropy-complexity curves can vary across pulsars and their significance in understanding their variability processes. Investigating this will require not only more observations but a greater survey of pulsars to truly investigate the phenomena we are seeing.

6 Conclusion

We have shown the successful implementation of epsilon-machine modeling as a tool to help understand the stochastic variability of nulling and non-nulling pulsars. However, what we have shown here is a mere proof-of-concept to underscore the power of such an analysis. We have also created our own robust softwares to reliably extract pulsar intensity curves independant of scintillation effects. It's applications go beyond just our analysis and can be used to investigate other pulsar behaviors.

Using the Allen Telescope Array, we observed two non-nulling pulsars and two nulling pulsars, and successfully showed that a simple Gaussian noise model can be rejected in favor a standard stochastic nulling model. This suggests that standard pulsars may have more complex variability than previously thought. In fact, it may appear that they share a similar stochastic mechanism with nulling pulsars. However, of the two nulling pulsars we observed, J1136+1551 appeared to disagree with both our Gaussian and stochastic nulling benchmarks. Further benchmark testing is required to try and recreate the same EC curve that we observed.

We suggest a larger survey of pulsars, both nulling and non-nulling, to investigate how much these EC curves vary across different pulsars. We also suggest a survey of benchmark models to help create templates with which we can match pulsar EC curves. This may help uncover different stochastic behavior of pulsars that may have eluded us in the past, and help bridge a gap in current knowledge of how pulsars physically work.

References

Nicolas Brodu. Reconstruction of epsilon-machines in predictive frameworks and decisional states, 2011.

Nicolas Brodu and James P. Crutchfield. Discovering causal structure with reproducing-kernel hilbert space ϵ-machines. *CoRR*, abs/2011.14821, 2020. URL `https://arxiv.org/abs/2011.14821`.

David P. Feldman, Carl S. McTague, and James P. Crutchfield. The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(4), oct 2008. doi: 10.1063/1.2991106. URL `https://doi.org/10.1063%2F1.2991106`.

Vishal Gajjar, B. C. Joshi, and M. Kramer. A survey of nulling pulsars using the giant meterwave radio telescope. *Monthly Notices of the Royal Astronomical Society*, 424(2):1197–1205, jun 2012. doi: 10.1111/j.1365-2966.2012.21296.x. URL `https://doi.org/10.1111%2Fj.1365-2966.2012.21296.x`.

D. R. Lorimer. SIGPROC: Pulsar Signal Processing Programs. Astrophysics Source Code Library, record ascl:1107.016, July 2011.

D. Nice, P. Demorest, I. Stairs, R. Manchester, J. Taylor, W. Peters, J. Weisberg, A. Irwin, N. Wex, and Y. Huang. Tempo: Pulsar timing data analysis. Astrophysics Source Code Library, record ascl:1509.002, September 2015.

Cosma Rohilla Shalizi and Kristina Lisa Shalizi. Blind construction of optimal nonlinear recursive predictors for discrete sequences, 2004.

Cosma Rohilla Shalizi, Kristina Lisa Shalizi, and James P. Crutchfield. An algorithm for pattern discovery in time series, 2002.

Dan R. Stinebring, Barney J. Rickett, Anthony H. Minter, Alex S. Hill, Adam P. Jussila, Lele Mathis, Maura A. McLaughlin, Stella Koch Ocker, and Scott M. Ransom. A scintillation arc survey of 22 pulsars with low to moderate dispersion measures. *The Astrophysical Journal*, 941(1):34, dec 2022. doi: 10.3847/1538-4357/ac8ea8. URL `https://doi.org/10.3847%2F1538-4357%2Fac8ea8`.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert

Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.