Draft version August 24, 2025
Typeset using LaTeX **default** style in AASTeX7

# Understanding Clustering Algorithm Applied to Planetary Nebulae

Lizette Rodriguez[1] and Rodolfo Montez Jr.[2]

[1] *University of Texas at San Antonio*
[2] *Center for Astrophysics | Harvard & Smithsonian*

## ABSTRACT

At the end of a sun-like star's life, the mass lost by the star will be shaped into shells of ionized gas that are illuminated by the evolving white dwarf at the core. Planetary nebulae PNe are left with complex formations that also emit X-rays. What makes some of these PNe formations complex is a shocked region known as a hot bubble, which emits soft X-ray photons. Occasionally, point-like sources of hard X-ray emission are also found at the core and are associated with the central star or close binary companions. Generally, hot bubble emission is detected early in the development of the PN and seems to disappear as the PN evolves, while the compact sources can persist into later stages. Identifying the presence of X-ray emission in PNe is typically performed by visual inspection. In this project, the primary objective of this research was to perform a parameter study that uses machine learning algorithms to further quantify X-ray emission in NGC6543. In particular, we considered a range of parameter values for the parameters eps and min_samples to further analyze hot bubbles and some with compact point source emission embedded within the hot bubbles. Our analysis shows that clustering algorithms can identify the best set of parameters to identify hot bubbles and compact point sources. As a bonus, some clustering algorithms can also spatially isolate source and background photons. Ultimately the results of these clustering algorithm experiments suggest future directions to apply these algorithms across similar sources with extended and compact emissions.

## 1. INTRODUCTION

Towards the end of a main sequence of star lives a star will expand into the red giant phase and enter the AGB phase (asymptotic giant branch phase). From the AGB phase, the star loses material that ends up in space. After the star's core collapses the star dies and leaves ionized gas in space. What is left from a main sequence star death is a Planetary nebula. Planetary Nebulae (PNe) are often characterized by their shells of ionized gas with white dwarfs at the center. These formations after a star's death often more than not leave behind complex morphologies or evolutionary patterns. So many things play a role in PNe formation, so much so, that no PNe is the same. Many planetary nebulae come in all different shapes sizes and even symmetry. Common PNe shapes include but aren't limited to: Elliptical, Spherical, Bipolar, and Irregular. Furthermore, the two symmetries that exist to classify PNe include asymmetric and symmetric. The particular PNe that was studied in this parameter study was NGC 6543 also known as Cat's Eye Planetary Nebula. NGC 6543 was chosen because of its well-sourced and extensive data sets by Chandra.

The Chandra X-ray telescope observes planetary nebulae in this X-ray range (0.5 keV). The X-ray emission typically observed by PNe are seen in two flavors: Diffuse and point-like X-rays. These two different flavors of X-ray emission have different characteristic X-ray emissions. Typically, the differences seen in these two X-ray emissions are primarily observed by the location and energy level. Point-like X-rays (Or central point X-rays) are defined by the X-ray emission from the center of the planetary nebulae or compact X-ray emission near the white dwarf(s). Vise-versa diffuse X-ray sources are defined by their characteristic emission from the hot bubbles. Hot bubbles are formed through shocked winds and gas that collide with each other. Typically, these hot bubbles emit soft X-rays from 0.3 to 1 keV.

The Cat's Eye Nebula archival Chandra data classifies the cat's eye as an Emps(h)/Mcspa (J. H. Kastner et al. 2012). As seen in (J. H. Kastner et al. 2012) this kind of morphology is known as: elongated, multipole shells present, have point symmetry, a distinct outer halo, has multipolar closed outers. This classification is particularly important to

Email: lizette.rodriguez@my.utsa.edu

understanding planetary nebulae because it gives potential insight into how this PNe might have been formed millions of years ago. Additionally, it's the notion of physical characteristics that leaves behind very important questions that are still quite unclear in the PNe community. With this how do PNe form 'such complex elongated elliptical morphologies such as the one listed above? Do these complex formations happen when in theorized binary star systems through stellar shocks? What do stellar shocks tell us about x-ray emission and what are their connections to hot bubbles? These questions are all valid questions that the beginnings of this research aim to tackle. Primarily so the fact that the physical characteristics of PNe are categorized by eye. What if we had the best sets of parameters to run an algorithm to characterize the different categories of X-ray emission? What would that then say about the evolution of the PNe? Would this give us results that aren't already provided through X-ray spectroscopy?

Throughout the past decade, X-ray spectroscopy has been used to identify certain types of elemental abundances found within PNe. The elemental abundances within PNe suggest just as much the physical composition just as much as the chemical composition. Primarily the interactions of the central(s) are the biggest source of information and often more than not their association with complex morphological structures is categorized as PNe that are results of binary star systems. Particularly in (M. Freeman et al. (2014)) where the researchers utilized Chandra and XMM-Newton to further research the extended diffuse X-ray shocks, known as hot bubbles. Like previous methodologies (J. H. Kastner et al. (2012)), this study identified 24 new subjects for X-ray morphological analysis. This investigation ultimately establishes a theoretical timeline for shocked winds in planetary nebulae, characterized by the presence of hot bubbles in the earlier stages of their life. Similarly, (R. M. Jr et al. (2015)) conducted another survey of CSPNe, focusing on compact emission PNe. In this survey, binaries were examined through pure X-ray analysis and were specifically chosen for their extremely bright point sources at the center of each PN. All 20 X-ray sources demonstrated a high correlation between the X-ray energy levels and the mechanical processes involved in the evolution of planetary nebulae over time (R. M. Jr et al. (2010)). These observed behaviors are as follows:

- Bipolar and asymmetric PNe → More likely to have X-ray-luminous CSPNe (suggesting binarity and strong magnetic fields).
- Round PNe → Less likely to emit X-rays (suggesting single-star evolution with weaker wind interactions).
- Diffuse X-ray emission → Traces shocked gas from fast stellar winds interacting with older AGB winds.

These behaviors in previous studies are important to note within the context of the research because they are foundational to the understanding of PNe physical composition. Particularly in our case study for the cat's eye nebulae, our target behaves in a manner that is conducive to the sets of behaviors in CSPNe with large amounts of diffuse X-ray emission. What this investigation seeks to computationally accomplish is whether a parameter study of the algorithm DBSCAN can physically help find a better method of categorizing different PNe morphologies more methodologically. Instead, rather than relying on human categorization usually done to analyze specific physical properties of PNe, an attempt to automate user parameters with the algorithm to see if automation in clustering a PNe will categorically produce a more accurate understanding between the photon counts within the hot bubble and the central(s) that correlate with relationships between the photon count and X-ray emission of PNe.

## 2. METHODS

### 2.1. *Data from the Chandra X-ray Telescope*

The data used for this paper was obtained from the Chandra X-ray telescope using the Chandra archive[3]. The archive was used to retrieve primary data on the target NGC 6543. The observation ID is noted in the Chandra archive as 630 where the relevant logged parameters such as photon count, energy level, and distance in x and y pixels were collected by Chandra in the year 2000. After data retrieval was complete data analysis was run in a Jupyter Notebook in which the researcher read the fits file that originally contained over 462401 events within the event table. To better combat the oversaturation of noise within the data set the data set was reduced significantly to only 2446 entries in the new chancleantable. This reduction was done using a series of filters in Python where specific energy distribution values can be attributed to noise detection and not the actual nebula itself.

[3] Chandra Archive:Chandra X-ray Observatory

**Figure 1.** An image by Chandra/HST showing the Cat's Eye Nebulae (NGC 6543) optical image background in the visible wavelength with an x-ray overlay at the center of the PNe by Chandra x-ray telescope.

Once data reduction was complete on the original event table the new event table (chancleantable) underwent a series of calculations to better later cluster different sets of photon population densities within NGC 6543. First, the offsets of x(1) and y(2) were calculated from the central position of the PNe ($x_0, y_0 = 4054.9101, 4033.6436$) and subtracted from all of the x and y distance entries within the chancleantable. These offsets are crucial for determining a uniform spatial scale between all photons. Besides the special scale, the chancleantable also underwent more np.where filters where the chancleantable not only went another energy filter that further reduced any more energies that whereas within the range 300 and 8000 eV but chancleantable also underwent an offset resolution were this paper study did not consider offset positions above the value 50. With more calculated parameters to better apply the clustering algorithm, the researcher created a new table that kept parameters of interest that include x and y pixel distance, energy, x, and y offsets, and the Chandra energy filter. This new table as a result was called evt_cln. evt_cln is an important table because this table gives the base parameters that were used to calculate the radius of every photon within evt_cln.

Another important thing to mention is the mean nearest neighbor distance. This means nearest neighbor distance is defined in this scope of research as the average distance between each feature and its nearest neighbor. This particularly is useful because the distribution of distances between photons is random. Understanding this, a loop for the average nearest neighbor distance column was created. First, a set parameter of nearest neighbor counts was set to 8. This number in previous test studies was shown to be the best for the scope of this parameter study. After setting a base parameter of nearest neighbor counts evt_cln the Euclidean distance to all other events was calculated based on the distances of $x_{\text{pix}}$ and $y_{\text{pix}}$(4). These distances were sorted and the average of the 8 closest non-zero distances was taken(6). Then nnd8 was created when we stored the nearest neighbor distances in a new column of our evt_cln table (now nnd).

#### 2.1.1. *Radius Calculations*

$$X_{\text{offset}} = x_{\text{pix}} - x_0 \tag{1}$$

$$Y_{\text{offset}} = y_{\text{pix}} - y_0 \tag{2}$$

$$Radius = \sqrt{(X_{\text{offset}})^2 + (Y_{\text{offset}})^2} \tag{3}$$

#### 2.1.2. *Mean Nearest Neighbor Distance Calculations*

$$x_{\text{current}}, y_{\text{current}} = (x_{\text{pix}})(100) + (y_{\text{pix}})(100) \tag{4}$$

$$Radii = \sqrt{(x_{\text{pix}} - x_{\text{current}})^2 + (y_{\text{pix}} - y_{\text{current}})^2} \tag{5}$$

$$\text{NND}_8 = \frac{1}{n_{\text{nn}}} \sum_{i=1}^{n_{\text{nn}}} Radii_i \tag{6}$$

### 2.2. *Understanding Spatial Clustering Algorithms*

#### 2.2.1. *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)*

DBSCAN, or Density-Based Spatial Clustering of Applications with Noise, is a clustering algorithm that works to cluster data distribution sets of high and low-density populations. DBSCAN is an algorithm that primarily functions on two set user parameters eps and min_samples. User parameters such as epsilon (or eps) and minimum sample value (or min_samples) are user parameters that define the distance away from a chosen data point in correlation to the minimum number of points it takes to make a cluster. The primary purpose of DBSCAN in this parameter case study was to determine whether DBCSAN could cluster populations of different photon densities (concerning characteristic X-ray emission types mentioned above) based on optimized parameters eps and min_samples.

#### 2.2.2. *Core, Non-core, and Noise Points*

After defining your set of user parameters, the DBSCAN clustering algorithm will first draw the set epsilon (radius) you previously selected when using DBSCAN and then elevate based on the user parameter min_samples to determine if it makes the requirements to join a cluster. If a certain point within a high population density consistently fulfills both user parameters, then the point that DBSCAN is analyzing will characterize this point as a core point. These core points are associated with high population densities and in our case will be associated with high populations of photons with put observational target NGC 6543. Say if a point is within a certain radius of a group of points but does not meet the requirement to have enough min_samples values to join a cluster this point is designated as a non-core point. These non-core points are criticized for being in lower population densities and often are the boundary points of a core cluster, hence the name is also known as: Boundary points. Found in the lowest population densities are the noise points also known as outliers. These noise points are classified by not fulfilling any of the requirements to join a cluster. This means that points are so far away from a general population of points that there are no proximity points to even the radius of the noise point.

### 2.3. *DBSCAN Cluster Identification*

For our particular observation modifications to the DBSCAN and twinx tutorials by scikit[4] and Matplotlib[5] respectively were modified to best fit the goals of this parameter study.

Before modifications were ever made to matplotlib demos for parameter analysis, modifications to the DBSCAN algorithm were made primarily to provide more information on population density samples. A function named run_dbscan accepts DBSCAN parameters eps and min_samples that fit the DBSCAN model to the data and returns several diagnostic metrics. Run_dbscan desired returns are made up of the following: the total number of clusters (excluding noise), the number of noise points, and the number of core versus non-core points (boundary points) for each cluster. This is particularly important because these metrics provide insight into the internal structure of the cluster(s) and help optimize the choice of parameters for analyzing photon density variations in NGC 6543.

## 3. RESULTS

### 3.1. *The Parameter Study*

Understanding how the behavior of eps and min_samples is primary to identifying how the cluster results of DBSCAN will look. In the previous section, the discussion focused on how if you varied these parameters you would receive

[4] scikit Clustering Demo:scikit

[5] Matplotlib Twinx Demo:Matplotlib

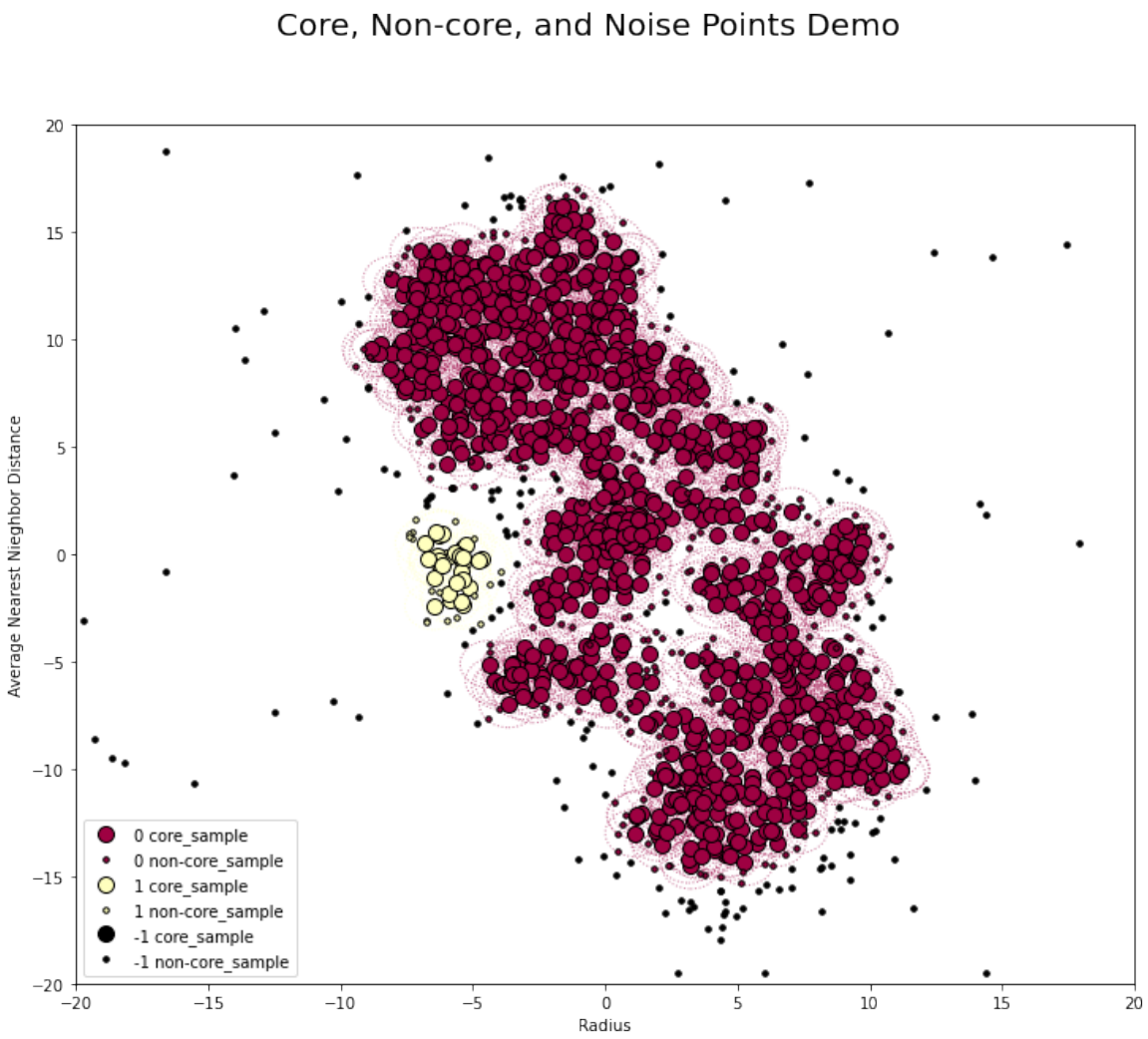


**Figure 2.** A Graph Demo Created by the author with the user parameters of eps =1.1 and min_samples = 10 to illustrate the differences between core, non-core points, and noise.

different clustering results based on the density population of a data set. This section will discuss how the algorithm's results (run_dbscan function) will be used to cross-analyze between ranges of parameter values compared to the number of core and noncore points.

Primary for the scope of this research modification to the Twinax demo[6] was made to better compare the axes of different numerical scales against each other, as the core and noncore points of any variation of eps (range = 5,40) would be sure to produce a different scale for 10 min_samples. The parameter values were run through a range loop where the (run_dbscan function was used to run through a series of eps values for 10 min_samples what exactly would yield what kind of core counts.

Once a full distribution (see Figure 6)) of core and non-core counts was complete over a range of eps values. The next step was to further deduce how the full distribution of core and non-core points would be visualized by DBSCAN. Figures 5 and 6 accurately show how specific ranges of eps values were limited in studying how different cross-sections of core and non-core counts cluster photons. The lines throughout Figures 5 and 6 are vertical lines attached to the plots to accurately signify eps points of interest that later get placed into the DBSCAN demo for Clustering analysis.

The DBSCAN page demo has little modification with only the parameter values of interest and of course, the file name is changed. Additionally, the subplot below is a subsequent modification of the DBSCAN code where the

[6] Matplotlib Twinx Demo:Matplotlib

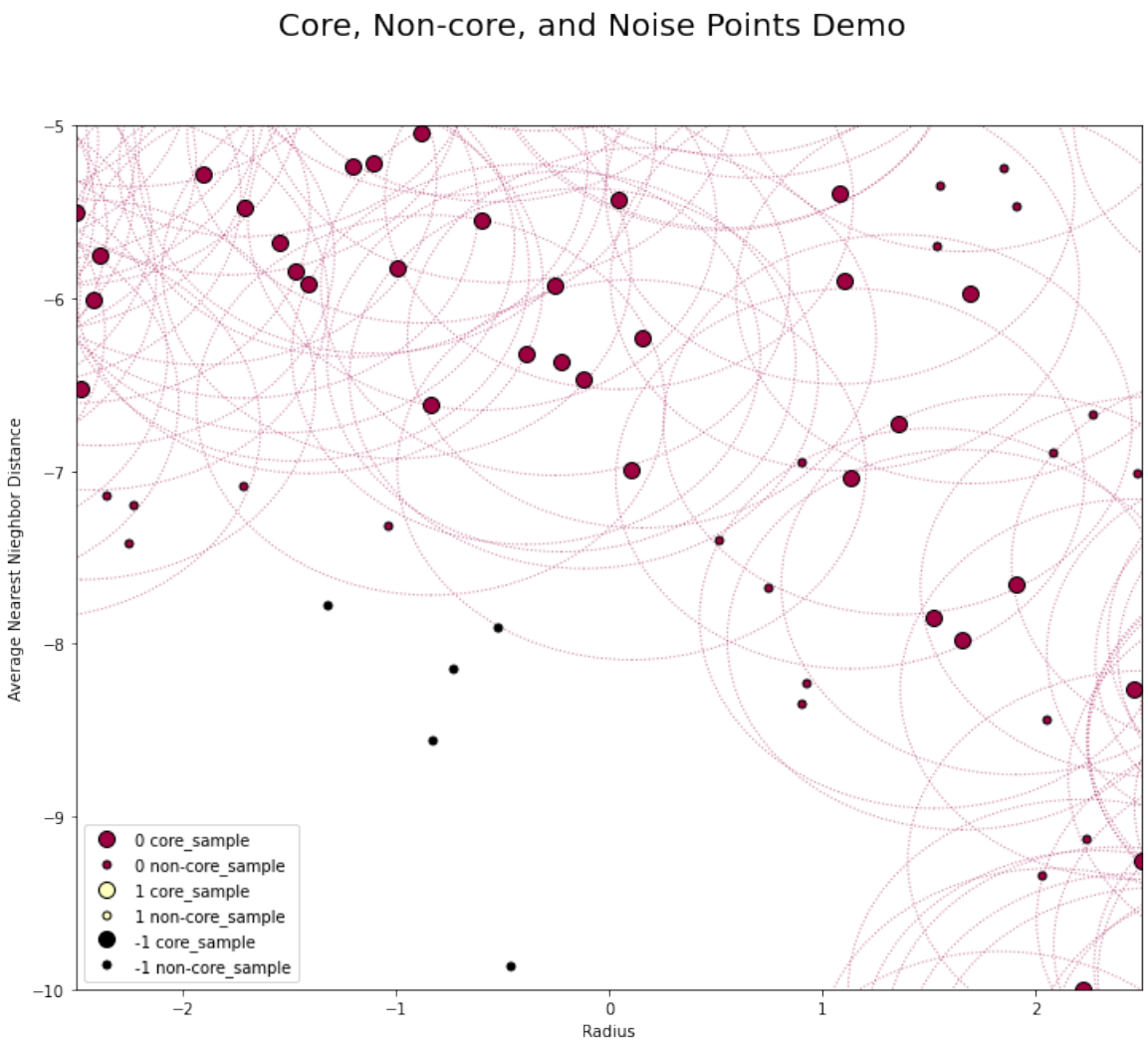


**Figure 3.** A Zoomed version of the core and non-core and noise points demo to further illustrate how the radius of DBSCAN is drawn by DBSCAN itself.

researcher used a fig, axes command to index every resultant graph of all eps vertical lines of interest all across the same min_samples value.

## 3.2. *Paramter Study Results*

### 3.2.1. *To Many Core Counts?*

To understand the output information going in chronological order (first, see Figure 5)) there must be an understanding of the relationship between core and non-core counts. This is crucial as the number of core counts increases the number of boundary points decreases. In the scope of this research, it is more probable that the density of points is in a high-density point population cluster as opposed to multiple-point population clusters. This is accurate to the clustering nebulae (as seen in the clustering figures below) and to what is represented in the parameter results for a certain range of eps. Below is a full distribution of eps values 5 through 40 where the core and non-core distribution of points proves the assertion that the more the core points increase the more the non-core points decrease (please see the ranges of eps= 15 to 40). Here the over-saturation of core counts virtually yields no boundary points making every point (excluding noise) a member within a major cluster. This is particularly not only seen in the full arbitrary distribution of eps values generated but also in Figure 6) and Figure 7) where the same behaviors of core points become over-saturated for our data set at (eps = 1 to 3.0, eps = 1.5 to 3.0 and eps = 2.0 to 3.0). Additionally, to further investigate the behavior of over-saturation or core to non-core points the number of clusters for a given arbitrary eps

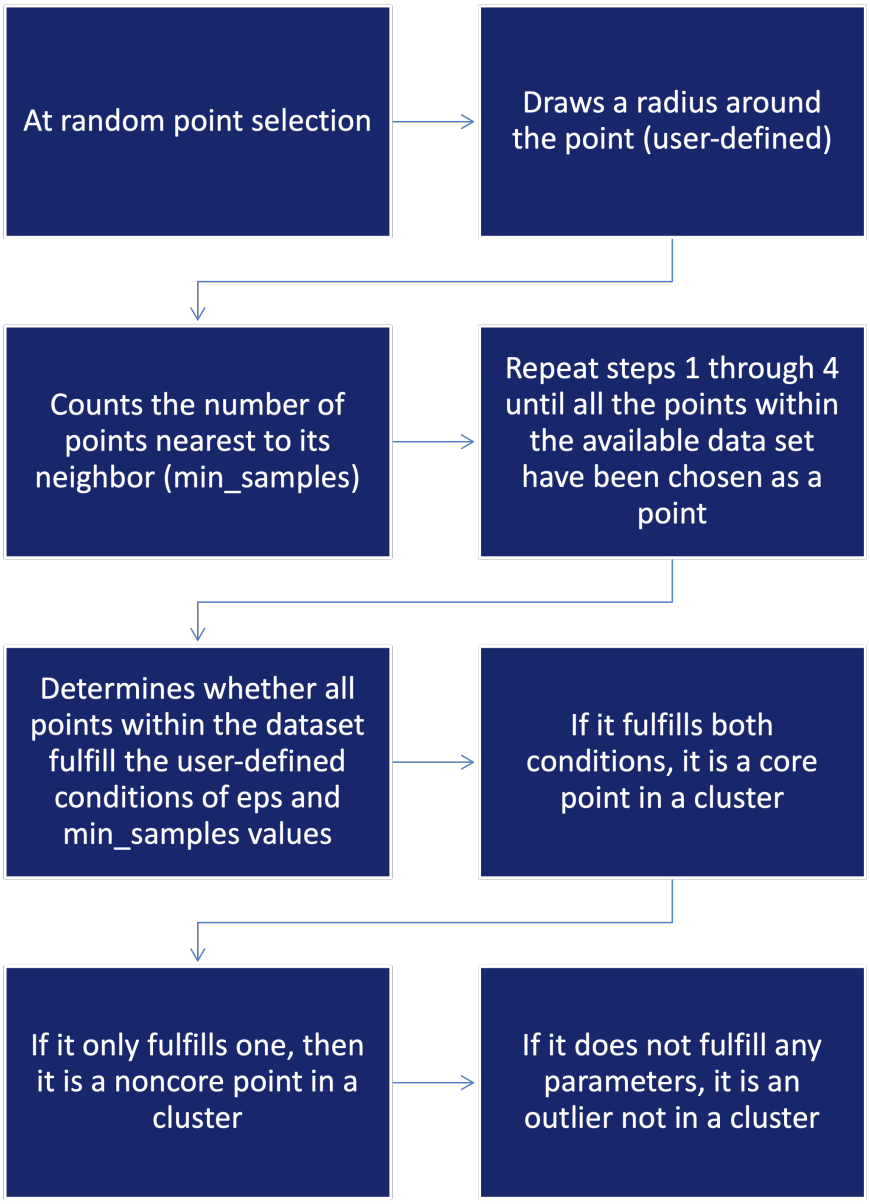


**Figure 4.** A flow chart of the process of how DBSCAN works

value was generated to provide further information on the relation the number of clusters had on the given eps value. As previously stated through the generalization of the clustering behavior of core and noncore clusters you will notice how generally (in all figures) as an abundance of core points begins to expand into a place the core count becomes 1. This is generally seen in all the behaviors no matter the min_samples value. Furthermore, to complete this set of generalizations of all the figures notice how not many changes between the behaviors of all figures (figures 5 through 8). Every parameter plot no matter the min_samples value starts the same with steady increases in the number of clusters and non-core counts just as the eps radius value starts at the beginning of the plot. With an increasing eps value an eventual decrease in the number of clusters and non-core points begins to happen just as the core points become prominent within the population density of photons. It's because of the similarity in behaviors in the parameter this study chose to focus on the min_samples value of 10.

#### 3.2.2. *Intersections, Transitions, and Vertical Line Designations*

As the title of this section suggests the more interesting part of this project relied on the visual inspection of the intersectional and transitional behavior of joining curve lines between core, non-core, and number of points in the

parameter study plots. Particularly as in chronological order if we start with Figure 5), you can see a strong line intersection at eps = 15. But notice that in the next figures from the earlier study, our data range is constrained for further behavior analysis in the ranges of (roughly eps = 0 to 3), why is that? Why in this parameter study did we not account for the parameter we don't account for the ranges of say eps = 10 to 20? The straight answer is because of the core point situation. Recall in the previous method section the total amount of photon points within this data set is around 2000 the photon point distribution values for eps = 10 to 20 would not only effectively cluster all photons into one major cluster but also cluster noise into the major cluster. It's why within the context of this parameter study the ranges of eps = 1 to 3 were meticulously studied as the spacial radius between each photon has to be equivocally small enough to not be saturated with core counts to provide meaningful results between the different physical characteristics of our target NGC 6543.

Something particularly meaningful in the scope of this project is the smaller eps values (ie: the determination of Figure 9). Firstly, notice how Figure 9) has magenta vertical line designations. What exactly do these line designations mean? How did this study effectively determine what eps patterns to study and what do these designations mean to the scope of research? These are fundamental questions that outline the transitional behaviors between all of the behaviors of our partners eps, core, non-core, and number of clusters. Firstly, because the goal of this research was to investigate small eps values with valid singular cluster boundary points the transitions of the dip in non-core courts and the uptick in core points were marked as potential candidates of review for the DBSCAN examination. Furthermore, to further investigate the dips and peaks of the data vertical line ticks were also placed in between the other candidates to examine how the number of clusters and the ratio of core to non-core points changed across selected eps candidates below.

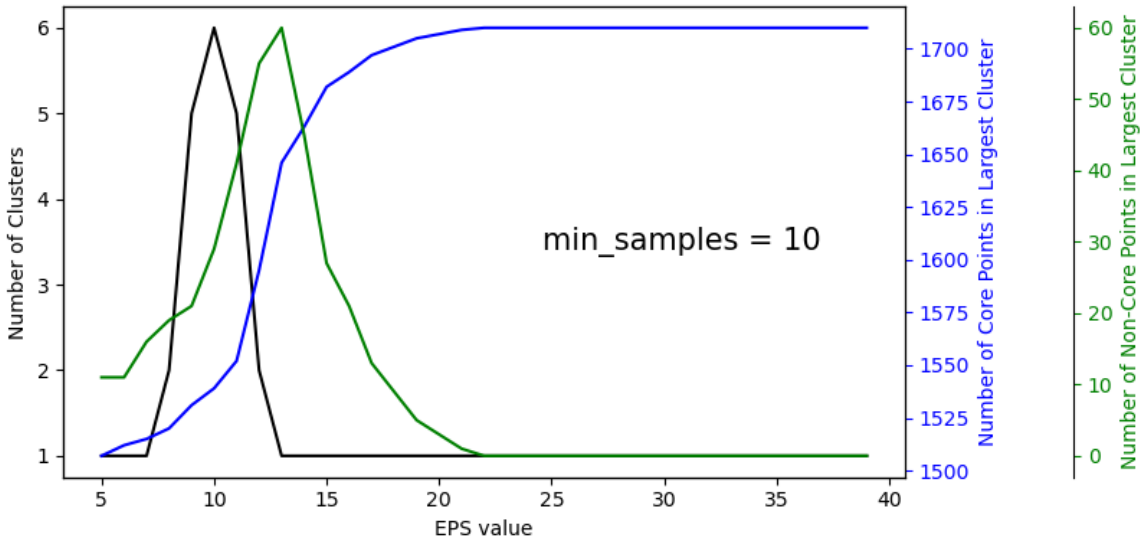


**Figure 5.** A full range distribution of all considered eps values (5 through 40) for the min_samples value of 10.

### 3.3. *Examining DBSCAN Outputs*

Once the eps candidates were selected the algorithm was utilized to perform clustering visualization on all 7 vertical line designations to further determine which were the best set of parameters using DBSCAN. How might one know whether or not the visualizations are accurate to the main objective of this study? Primarily a PNe cluster in question should have different clustering properties in the center of the PNe in comparison to say the extended X-ray sources farther away from the central star. This means that the PNe in question should have the most core counts in the center and boundary points that connect across the ridges of the PNe itself. This part of the parameter study relied on visual inspection of the PNe to further determine which cluster(s) best described our target NGC 6543.

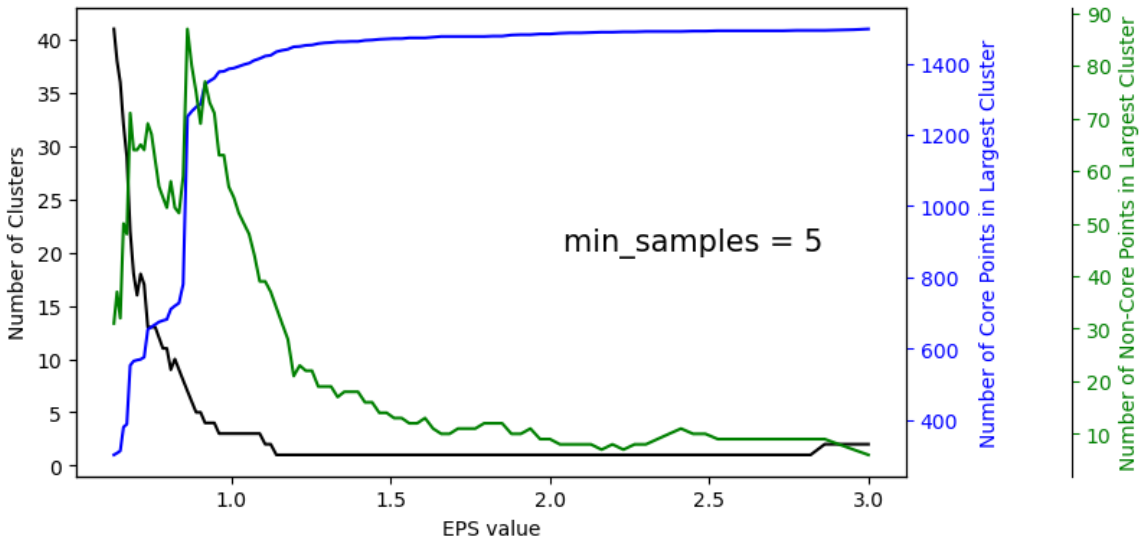


**Figure 6.** A eps Value graph illustrating behaviors between the number of clusters, core, and non-core points ranging from 1 to 3 representing the change in min_samples value of 5.

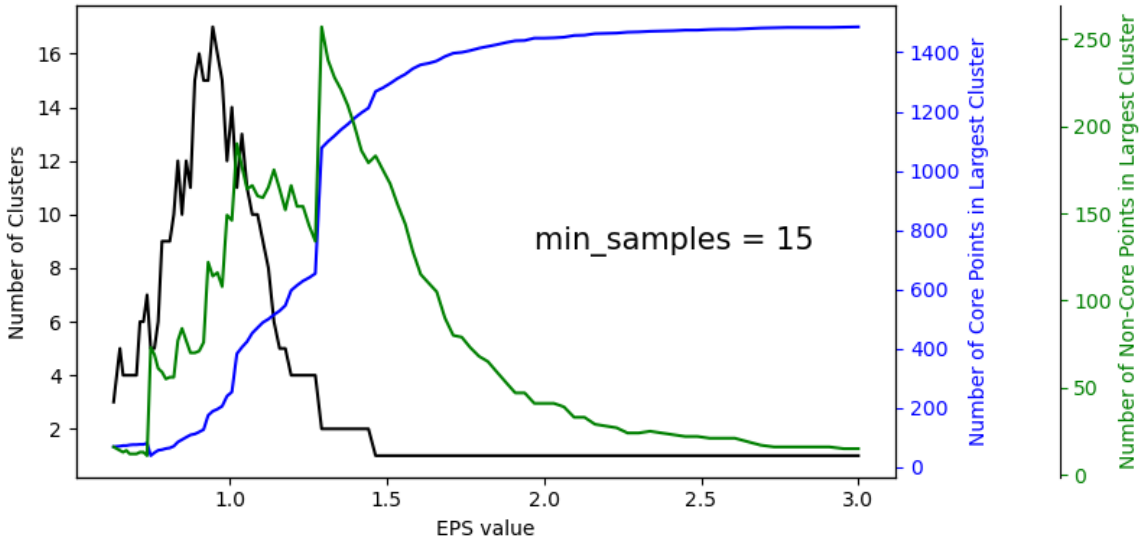


**Figure 7.** A eps Value graph illustrating behaviors between the number of clusters, core, and non-core points ranging from 1 to 3 representing the change in min_samples value of 15.

### 3.3.1. *2 Cluster Designations*

As the title of this section suggests this section will be mining the DBSCAN outputs that have 2 cluster formations (Figure 10) through 11).

- In Figure 10) the clustering algorithm examines an eps of 1.09 and a min_samples value of 10. In this figure, there is a clear separation between the two major clusters with slight gaps in the formation of the boundary points at an offset of -5 and y offer of 2. This is a clear indication of potential boundary points that are being separated into noise labels.

- In Figure 11) the clustering algorithm examines an eps value of 1.28 and the same min_samples value. Here in the determination of this graph, the results from the last slightly vary. Notice how the boundary points in this plot appear to be more complete. furthermore, there is not a huge level of separation between the first and

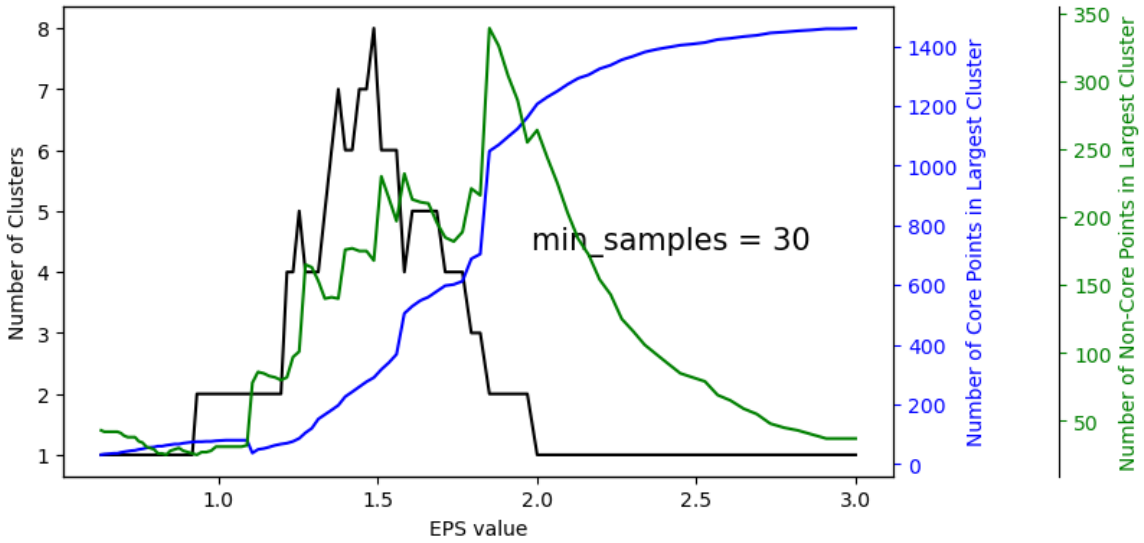


**Figure 8.** A eps Value graph illustrating behaviors between the number of clusters, core, and non-core points ranging from 1 to 3 representing the change in min_samples value of 30.

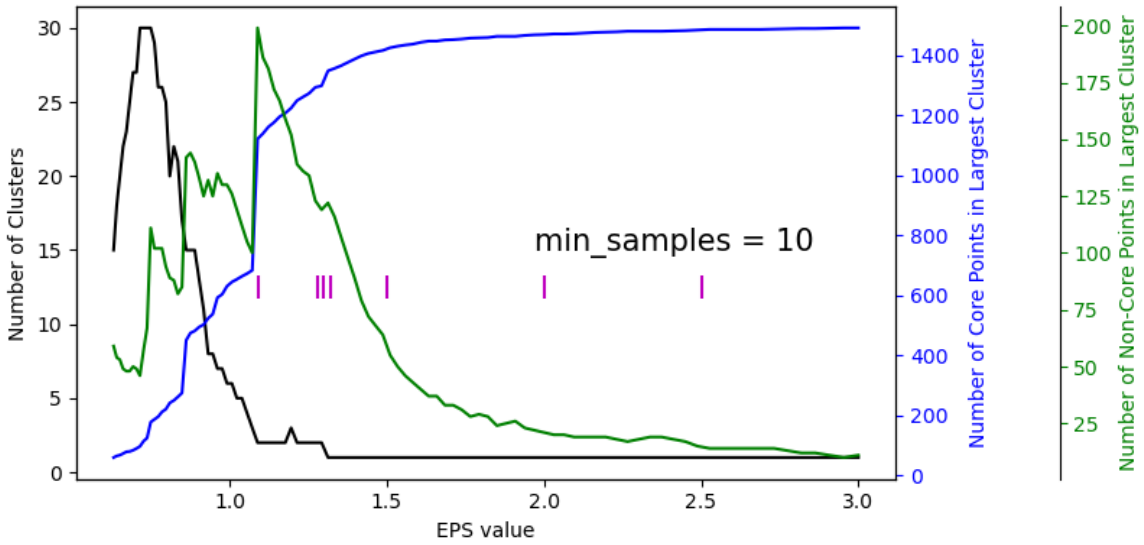


**Figure 9.** A eps Value graph with vertical line designations representing the selected eps values for that illustrate the most interesting behaviors between the number of clusters, core, and non-core points ranging from 1 to 3 with a min_samples value of 10.

the second cluster. Recall Figure 9 where you will notice that the transition from eps = 1.09 to eps = 1.28 is significant because it shows a huge reduction in non-core points and the first abundance of core points. This allows the algorithm to include more points within the major cluster closing the preliminary gap between the first and the second cluster as seen in Figure 11)

- (Figure 11)) increases the eps incrementally to 1.3 to further analyze the transition between 1.28 and 1.32. the reason why this eps value in particular is important to the parameter study is that it is the transition value for the number of clusters to two going to one. This means as seen in (Figure 11)) you can see a distribution of photons in a major cluster slightly overtaking cluster two that remains slightly reduced in size from the previous figure (Figure 11)). The significance of this partially makes sense as in Figure 9 this intersection shows a dip in non-core and a peak in core points.

#### 3.3.2. *1 Cluster Designations*

As the title of this section suggests this section will be mining the DBSCAN outputs that have 1 major cluster formation (Figure 11) through 13).

- In Figure 11) the eps value is 1.32 this is the transitional phase in which the prior two clusters prior to this cluster now become one major cluster. This correlated to be presented in Figure 9 where the further expansion of non-core points becomes greater because the 2nd cluster got absorbed into the major cluster model as seen in Figure 13). It also makes sense as a result that primarily core points have a slight increase as the photon distribution within this cluster now has a bigger population of points than it did at eps = 1.3. Notice how the boundary points are slightly disconnected and form the core points within the PNe? There is some formations of noise within some of the constructed core and non-core points of the PNe (see x offset of 3 and y offset of -12).

- in figure 14 at eps = 1.5 there is a slight variation of boundary points present within the major cluster. This variety of points collected some of the noise from the previous visualization of DBSACAN and provides more connected boundary points for our target which effectively shows optimized regions of boundary points for NGC 6543.this makes sense going back and looking at our plot in Figure 12) as the non-core distribution researchers a general decline and the core counts are about to reach its maximum plate it makes sense that at these boundary regions the visualization of DBSCAN is picking up points to make a fully elliptical morphology of our PNe NGC 6543.

- In Figure 12) the eps value is 2.0. notice from the transmission between going from eps = 1.5 to eps = 2.0. In both visualization plots there is a consistent connection between all the boundary points forming more of that lexical morphology we are familiar with based on previous studies of NGC 6543. One difference to note between both figures is in this formation of the PNe there are fewer boundary points towards the center offsets of the PNe. This formation makes sense in correlation to Figure 9 (whereas indicated via the vertical line designation) the noncore points decrease as the core points have a slight increase.

- In Figure 13) the eps partner value is eps = 2.5. In this figure, the boundary points become core points effectively barely leaving any trace of previous boundary lines that correlate with the formation of the PNe. What does this mean? This means that the algorithm is effectively pulling in points from the noise formations around the nebula that are not consistent with actual X-ray formations that are important in the scope of this research. This is our present in Figure 9 where based on the vertical designation you can notice for eps = 2.5 the core count formation is an overabundance leaving very few core counts descanting the entire pantry nebula all into one cluster (including some noise points in Figure 15.)

## 4. DISCUSSION

To understand how meaningful the results above we must first understand what exactly the scope of this research is first looking for. Our target NGC6543 is generally an elliptical planetary nebula and as previously stated in the Chandra section of this paper the physical characteristics of NGC6543 have not only been studied by Chandra before this but ultimately present the idea that PNe like NGCS6543 has complex planetary nebulae formations because of a theoretical presence of a companion star in tandem to a white dwarf host star. This dynamic theoretically in the literature is said to correlate complex morphologies such as irregular shape patterns and the composition of hot bubbles to the presence of the binary star system that is gravitationally bound to one another. This theory stems from general observations of single-star planetary nebulae and their more uniform morphological distributions such as very common circular morphology formations.

In connection to these theoretical formations, a bigger overarching question that this parameter aims to tackle is what exactly could be understood from these complex formations and how can we use computation to further study x-ray subjects that are not as well sourced as NGC6543.

This ultimately brings us to the formation of DBSCAN and the meaning of the results in tandem with the meaningful results already provided by literature and the Chandra X-ray telescope. DBSCAN itself has interesting sets of behaviors that ultimately lead to the next most important question for this project. How effective is DBSCAN in clustering photons and what are the limitations of using DBSCAN? Referring back to the eps ranges of 1.5 to 2.0 the eps ranges are pretty comparable to the formation of a hot bubble and provide pretty meaningful results in the context of this research. It's because of these results that. This range is the best set of parameters to describe the extended sources of

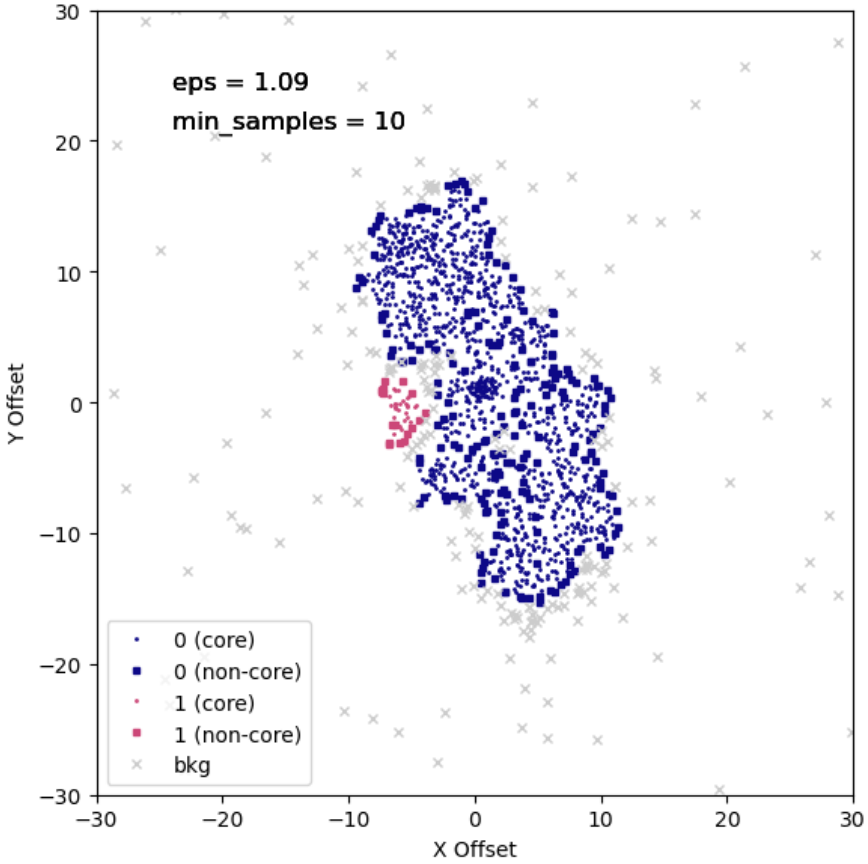


**Figure 10.** A DBSCAN cluster plot of NGC 6543 with an eps value of 1.09 and a min_samples value of 10.

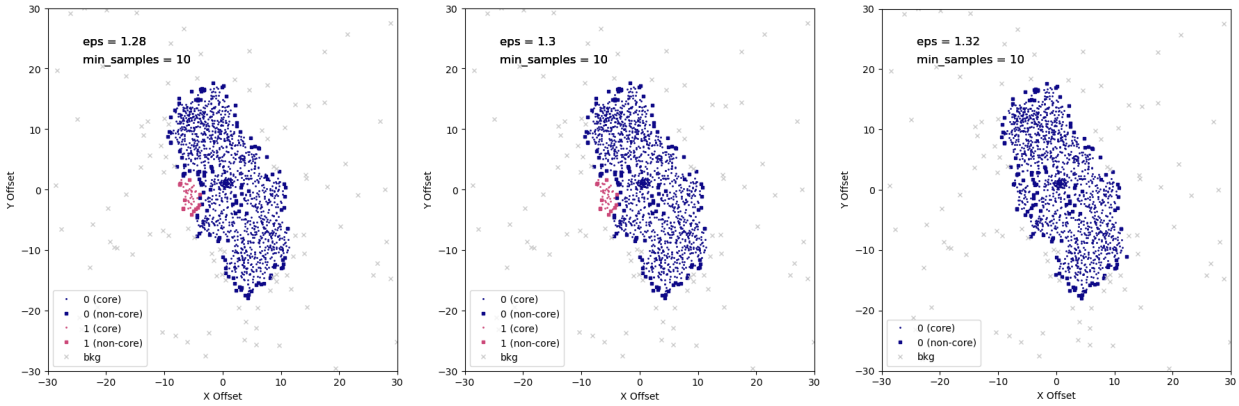


**Figure 11.** A DBSCAN cluster plot of NGC 6543 with an eps value of 1.28,1.3, and 1.32 with a min_samples value of 10.

NGC 6543. Overall this method did have comparable limitations mostly because the results still require lots of human intervention. Thus because it's still so controlled it takes human intervention to realize what exactly is significant in this scope of the research and what eps ranges aren't particularly useful for identifying extended vs point sources.

## 5. CONCLUSION

The primary object of this research is to conduct a parameter study to find the best sets of eps values based on our best selected min_samples value of 10 in sight of further understanding the correlations between cluster formations and physical characterizations of our target PNe NGC 6543. Typically the physical characteristics of our target source

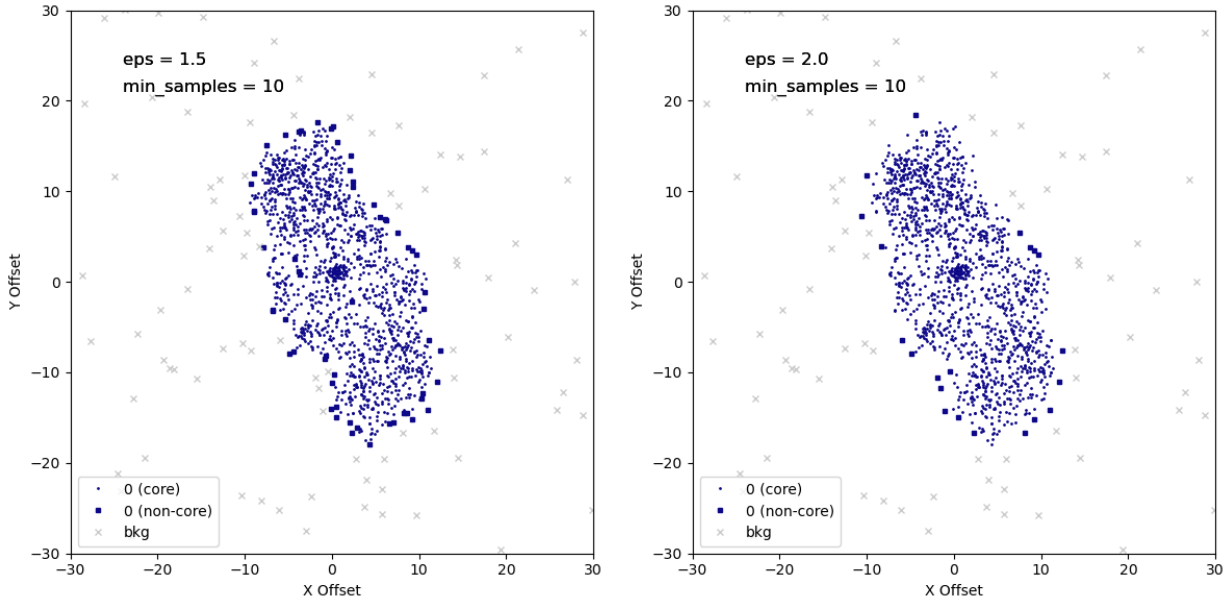


**Figure 12.** A DBSCAN cluster plot of NGC 6543 with an eps value of 1.5 and 2.0 with a min_samples value of 10.

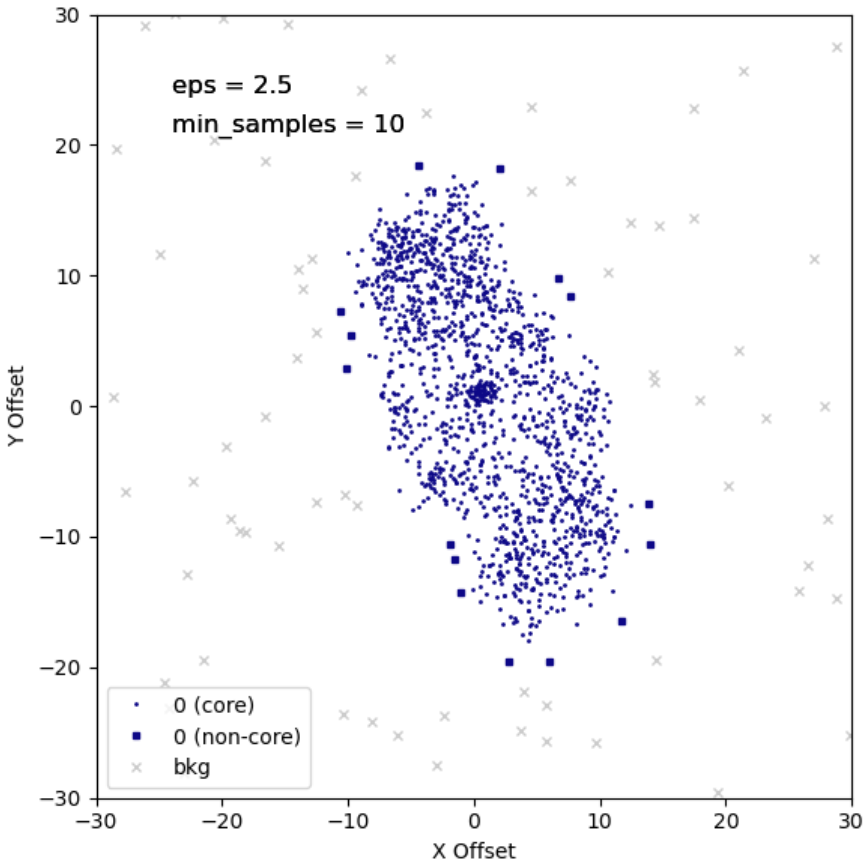


**Figure 13.** A DBSCAN cluster plot of NGC 6543 with an eps value of 2.5 with amin_samples value of 10.

have plenty to do with the number of photons within the PNe clusters meaning that for the context of this parameter study the larger the photon population the harder the x-rays and the less the population size the fewer x-ray counts

are detected by Chandra. This generalization allowed for the context of this research to further use the algorithm DBSCAN to find the best sets of parameters to separate physical characteristics such as the hot bubble from the central star's huge photon count. Furthermore, this was done by understanding the identification of core and non-core counts of NGC 6543. Fundamentally the non-core counts in correlation to the hot bubble and the core counts in correlation to the central star photon source. This is because of the spatial distances between photons where the special distance between boundary points is larger making it categorically easy to label the photon emission from the hot bubble as non-core points. Oppositely, the core counts of the central state are so packed together in a huge population density that it was far simpler to label them as core counts using DBSCAN.

DBSCAN was able to identify a range that best showed the separation between diffuse X-ray counts and hard X-ray counts by using the number of photons within a certain region and spatial the distance between each photon. The range identified was between eps = 1.5 to 2.0 for a meaning full min_samples value of 10. This range ultimately gave the best results by having the most prominent boundary points within the constraints of all the parameters listed in Figure 9. These prominent boundary points are particularly a point of interest because our interest is in understanding the physical characteristics of NGC 6543.

Among all of these vast generalizations and meaningful results, DBSCAN was able to provide meaningful results to the morphology of NGC 6543. This ultimately can be applied to fainter sources that the human eye can't categorize due to instrumentation limitations. The basis of this research can serve as a case study for what can the shapes of objects or clusters tell us about what is currently going on in the formation of objects we are uncertain about and what may have happened to them a thousand light years ago.

## REFERENCES


Freeman, M., Jr, R. M., Kastner, J. H., et al. 2014, Astrophysical Journal

Jr, R. M., Marco, O. D., Kastner, J. H., & Chu, Y.-H. 2010, Astrophysical Journal

Jr, R. M., Kastner, J. H., Balick, B., et al. 2015, Astrophysical Journal

Kastner, J. H., Jr, R. M., Balick, B., et al. 2012, Astrophysical Journal