Original Research Article

# Blue Straggler Star Formation in the Field Regions of M31

Yiyang Zheng

*Basis Independent Silicon Valley, 2321 Ridgecliff Ct., San Jose, CA 95131, United States*

## ABSTRACT

During stellar evolution, some stars evolve into blue straggler stars (BSS), stars that are more blue and luminous compared to the main sequence turnoff point (MSTO) of the parent generation. Two main theories for BSS formation exist, stellar collision and binary mass transfer. One way to determine which theory led to BSS formation is by determining if there is a correlation between local stellar density and BSS frequency. A strong positive correlation suggests that collision is the main method, while a very weak or negative correlation suggests that binary mass transfer dominates BSS formation. While most previous research have focused on BSS formation within stellar clusters, this paper intends to find a trend between local stellar density and binary mass transfer in the field regions of a large spiral galaxy, Andromeda (M31). Data from the Panchromatic Hubble Andromeda Treasury (PHAT) was used, and BSS candidates were identified. Local stellar density was then plotted against BSS frequency for each star in a logistic regression, and the coefficient was recorded. Out of 11 fields, 8 fields showed a weak trend with a confidence interval containing zero for the coefficient, with coefficients ranging from -2.818 to 2.222, while 3 fields showed a negative correlation between BSS frequency and local stellar density with coefficients from -4.231 to -3.587. This supports the theory that binary mass transfer is the primary mechanism for BSS formation in field stars.

**Keywords:** Blue Straggler Stars; Field Stars; Binary Mass Transfer; Stellar Evolution; Andromeda Galaxy (M31)

## INTRODUCTION

Stellar evolution is composed of several key stages including the main sequence and giant phases. This can be plotted on a Hertzsprung-Russel diagram (H-R diagram). On the H-R diagram, stars in the main sequence lie on a diagonal line to the left and upwards in the H-R diagram depending on its mass. In the giant phase, stars move off the main sequence and occupy a region above and to the right of it on the H-R diagram becoming cooler and more luminous. While most stars evolve off the main sequence and become a giant by cooling and expanding, blue straggler stars (BSS) instead appear hotter and more luminous than the parent stars at the main sequence turnoff point and are located beyond the main sequence band on the H-R diagram. They are also more massive than main sequence stars.

Two main theories are proposed for the formation of blue stragglers. The first is the creation of blue stragglers through stellar collisions (1, 2). When two stars collide, they merge together, forming a more massive star. As the star is more massive, the gravitational force on the core increases, increasing nuclear fusion rates and the temperature and luminosity of the star. The second is creation of blue stragglers through mass transfer between binary stars. In a binary system, when one star

**Corresponding author:** Yiyang Zheng, E-mail: benjaminzyy@gmail.com.

**Accepted** November 3, 2025
https://doi.org/10.70251/HYJR2348.36364369

fills its Roche lobe, mass is transferred via gravity to the other star, creating one star of increased mass, and thus increased temperature and luminosity. One way of determining the validity of these theories is by finding if there is a relation between local stellar density and blue straggler frequency. If there is a strong positive relation between the two, that supports the stellar formation collision theory, as stellar collisions occur far more frequently in areas of high local stellar density. If there is a very weak or no relation between the two, that supports the binary transfer theory. Binary transfer does not require high local stellar density environments such as in globular clusters and should be independent of its environment. Therefore, there shouldn't be a relation between local stellar density and blue straggler frequency in field environments. Thus, stellar collisions are thought to dominate blue straggler formation in globular clusters, while binary mass transfer is thought to dominate blue straggler formation in field regions. Many studies have been done to highlight this. For example, in Santana 2013, the authors tracked blue stragglers in low density systems in a dwarf galaxy, eventually finding no significant trend between local stellar density and blue straggler frequency in these areas, thus supporting the binary mass transfer theory (3). In Sollima, Lanzoni, and Beccari 2008, a similar trend is found in these areas (4).

While many studies have studied the relation between local stellar density and blue straggler frequency in globular clusters and dwarf galaxies, none have been done examining this relation on the field regions of a large spiral galaxy. This study focuses on determining if there is a relation between local stellar density and blue straggler frequency within the field regions of the Andromeda Galaxy (M31), a large spiral galaxy. The aim of the paper is to show evidence supporting or refuting the theories on the formation of field blue stragglers. The hypothesis is that there is no or a very weak correlation between local stellar density and blue straggler frequency in the field regions of M31. This would be consistent with previous research theorizing that blue straggler formation is dominated by mass transfer in areas of less local stellar density like the field regions of a galaxy. If little to no trend between local stellar density and blue straggler frequency is discovered, then further support will be given to the theory that binary mass transfer is the main source of field blue straggler formation. However, if there is a clear increasing trend between the two, then it would provide support for the theory that stellar collisions is instead the main source of field blue straggler formation. In low-density environments, binary mass transfer is expected to dominate blue straggler star formation, resulting in little or no correlation between local stellar density and BSS frequency.

The data used in this study come from Panchromatic Hubble Andromeda Treasury (PHAT) (5, 6, 7). Only field stars in the disk regions of M31 were used to identify blue straggler candidates, and no stars from the bulge or halo were used in the study. Thus, the results can only be generalized to field stars outside of clusters in the disk regions of the galaxy. In addition, only certain "fields," an individual HST pointing inside a "brick", a rectangular survey unit in M31, captured by PHAT, was used in the analysis, so results cannot be generalized to regions of local stellar densities not represented by the bricks. In addition, celestial objects other than stars could contaminate the data, causing information for some stars to be different than in reality. The brightness of some stars are also too low, and thus cannot be used for data analysis. Additionally, some young stars may be bluer and luminous compared to the main sequence turn-off point but are not blue stragglers (7). While PHAT UV photometry using F336 filters could separate them from BSS candidates, as the BSS selection method used is through the 475W and 814W filters, they cannot be differentiated.

## METHODS AND MATERIALS

### Methodology Overview

Initially, good star catalogs from the Panchromatic Hubble Andromeda Treasury (PHAT) of selected bricks and fields of the galaxy were selected and downloaded (6. 7). The data was cleaned; non-star objects, stars with unreasonable magnitude measurements, or stars with poor signal quality were removed, following procedures described in the PHAT photometric pipeline (6, 7). Each field was then plotted onto a HR Diagram, where blue stragglers were then identified. The local stellar densities of the stars were then found, and it was plotted against blue straggler frequency in a logistic regression test. This is an observational and cross-sectional study, as data was taken by observing stars at one point in time.

### Data Download and Cuts

Initially, good star catalogs which contain data on a star's properties from PHAT were downloaded. Only certain fields, smaller subsections, within the bricks were used. They include Brick 7 Field 8, Brick 7 Field 11, Brick 9 Field 13, Brick 9 Field 18, Brick 11 Field 5, Brick 11 Field 6, Brick 13 Field 12, Brick 17 Field 6, Brick 17 Field 7, Brick 20 Field 7, and Brick 20 Field 17. These fields were picked as they gave a wide range of local stellar density values.

Then, the good star catalogs were cleaned to contain only candidates with high quality data and no contamination. Stars with magnitude values 99+ or 0 were cut, as it symbolizes an invalid brightness measurement. These aren't real brightness measurements and rather placeholders, so they were cut. Stars with a signal-to-noise ratio below 4 were cut to ensure the star's location was captured accurately; any lower would find an inaccurate location of the star. Stars with a SHARP value above 0.3 were cut to cut objects that do not have a shape of the star, and stars with a difference in SHARP values of more than 0.5 in the blue 475W filter and the red 814W filter were also cut, as the star should appear similar in both filters. Additionally, any star whose square of the sum of SHARP values in both the 475W and 814W filters above 0.075 was cut. Stars whose sum in crowding values in both the blue and red filters was above 1 was cut to remove stars whose light profiles were too contaminated. Stars whose light profile did not match a typical star were also cut, as they may not be a star shouldn't be included in the data.

**Blue Straggler Candidate Selection**

BSS candidates lie beyond the main sequence in a H-R diagram or a color-magnitude diagram, therefore first the main sequence is identified by finding the main sequence turn-off point. First, a filter is implemented to cut off all stars from the cleaned data that are clearly not part of the main sequence such as giants. Then, the ridge line of the main sequence is determined. The main sequence is split into multiple bins based on its magnitude, and the median color of each bin is found. These medians form the main sequence ridge line. A 4th-order polynomial is then fitted with the ridge line to find the color of the center of the main sequence at each magnitude, forming the model of a typical star. Then, an estimation of the main sequence turn-off point, the point where a star leaves the main sequence and becomes a giant, is found by finding the point on the main sequence with the lowest magnitude using the ridge line. Finally, BSS candidates are selected if they are significantly brighter than the MSTO estimation and much bluer of a star of the same brightness on the ridge line, and Figure 1 shows an example for Brick 7 Field 11.

**Density Measurement**

The next step is to find the local stellar density value for each star. A BallTree data structure was used to find the nearest neighbors to each star as it is much faster and more efficient compared to what was used in Sollima *et al*. 2008 (4). The distance to the 5th nearest neighbor for each star is found, and the formula $\Sigma = (k-1) / (\pi * r_k^2)$ with k = 5 for local surface density was used to find the local stellar density for each star[8]. This formula follows the established formula used in previous studies, with a slight change to the value of k

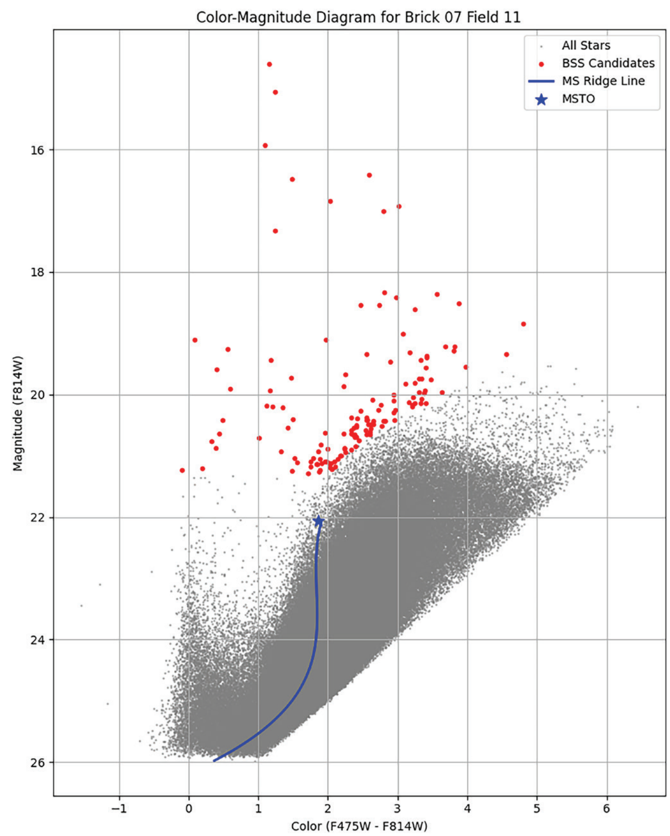


**Figure 1. Color-Magnitude Diagram for Brick 7 Field 11.** All stars of a field were plotted out on a color-magnitude diagram. The median of multiple bins is found to form the MS ridgeline. A 4th-order polynomial is then fitted with the ridgeline to form the model for a typical star, then an estimation of the main sequence turn-off point is found by finding the point on the main sequence with the lowest magnitude using the ridge line. BSS candidates are selected if they are far brighter and bluer of a main sequence star of the same brightness. This diagram shows the BSS selection process for Brick 7 Field 11.

to increase spatial resolution (4). The logarithm of this value was then taken due to vast differences in orders of magnitudes of stellar densities between different stars.

### Logistic Regression

BSS are bright (mag is low), and bright stars can affect density measurements. A multiple logistic regression model is used: d the probability that a given star is a BSS candidate versus not using a logistic regression was modelled and controlled for the effect of magnitude, yielding a coefficient that only shows the strength and direction of the relationship between local stellar density and BSS frequency.

The model used is:

$$\text{Logit}(P(BSS=1)) = \alpha + \beta_1 \log \Sigma + \beta_2 \text{ mag},$$

where $\Sigma$ is local stellar density and mag is the star's magnitude. Bootstrapping provided confidence intervals for $\beta_1$.

## RESULTS AND DISCUSSION

The $\beta_1$ coefficient represents how likely a star is to be a BSS as local stellar density changes. If it is positive, stars in regions of high local stellar density are more likely to be a BSS, if negative it is less likely, and if the value of the coefficient is near zero there is no significant relationship between BSS likelihood and local stellar density. As local stellar density increases by 10 times, the odds of the star being a BSS changes by $e^{\beta 1}$.

For most fields, there is no relation between local stellar density and BSS frequency, as the confidence interval for the regression coefficient contains zero (Figure 2). Eight fields exhibited this property as seen in Table 1, indicating no significant correlation between local stellar density and the likelihood of a star being a blue straggler. The coefficients range from -2.818 to + 2.222, meaning that as local stellar density increases by 10 times, the chance of the star being a BSS changes by $e^{-2.818}$ or $e^{2.222}$, or 0.06x to 9.22x. This suggests

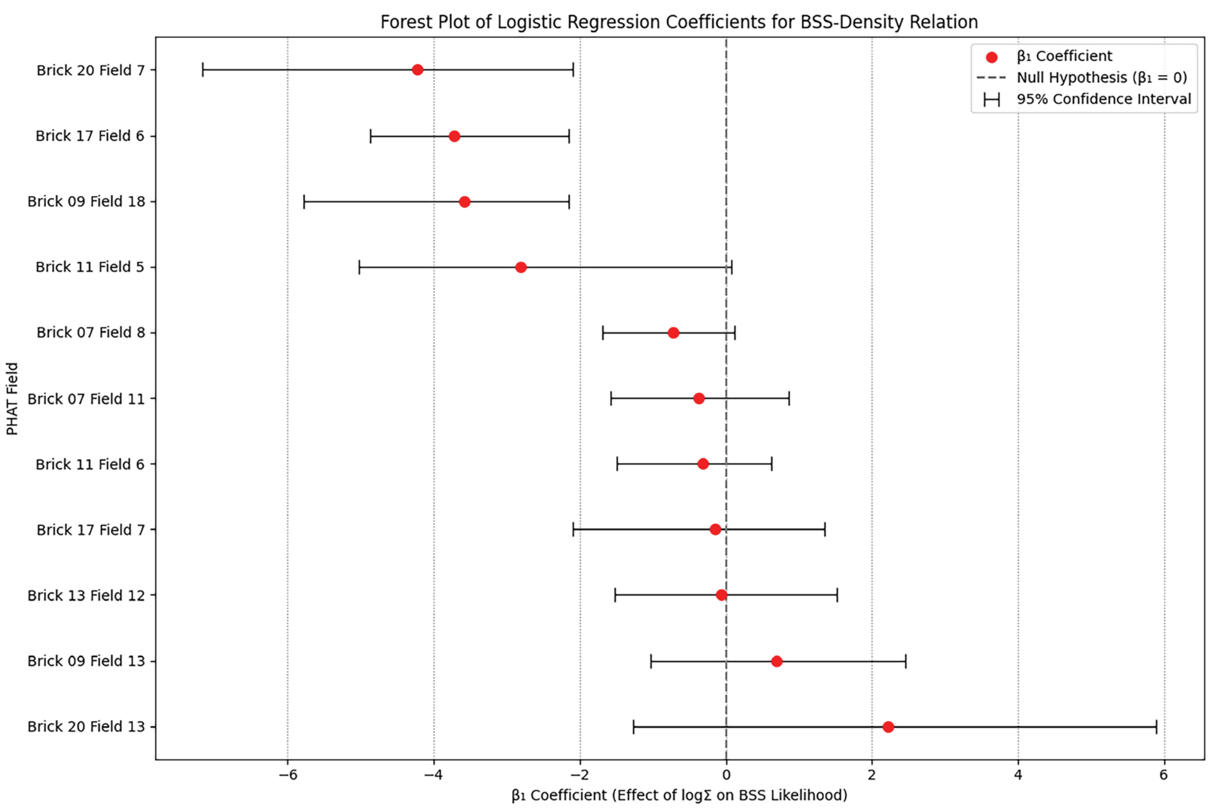


**Figure 2. Forest Plot of Logistic Regression Coefficients for BSS-Density Relation.** This forest plot shows the 95% confidence intervals of logistic regression coefficients for each field. Brick 20 Field 7, Brick 17 Field 6, Brick 9 Field 18 have negative coefficients, while the confidence intervals for all other fields include zero.

**Table 1.** Fields with No Correlation with $\beta_1$ Coefficient

| Field (PHAT GST) | Total Stars (count) | BSS Candidates (count) | $\beta_1$ Coefficient (dimensionless) | 95 % CI (Low, High) (dimensionless) |
|---|---|---|---|---|
| B20 F13 | 63,318 | 22 | +2.222 | (–1.266, +5.895) |
| B9 F13 | 320,609 | 139 | +0.697 | (–1.027, +2.453) |
| B11 F5 | 289,668 | 55 | –2.818 | (–5.018, +0.072) |
| B7 F8 | 348,188 | 303 | –0.720 | (–1.689, +0.112) |
| B7 F11 | 343,885 | 140 | –0.373 | (–1.583, +0.855) |
| B11 F6 | 260,223 | 34 | –0.318 | (–1.488, +0.620) |
| B17 F7 | 194,592 | 27 | –0.149 | (–2.101, +1.352) |
| B13 F12 | 231,344 | 46 | –0.063 | (–1.523, +1.517) |

Shown are the amount of BSS candidates, and the regression coefficient and confidence interval for fields with no correlation.

**Table 2. Fields with Negative Correlation with $\beta_1$ Coefficient**

| Field (PHAT GST) | Total Stars (count) | BSS Candidates (count) | $\beta_1$ Coefficient (dimensionless) | 95 % CI (Low, High) (dimensionless) |
|---|---|---|---|---|
| B9 F18 | 303,748 | 124 | –3.587 | (–5.782, –2.159) |
| B17 F4 | 91,140 | 36 | –3.717 | (–4.871, –2.147) |
| B20 F7 | 64,090 | 26 | –4.231 | (–7.161, –2.102) |

Shown are the amount of BSS candidates, and the regression coefficient and confidence interval for fields with a negative correlation.

that binary mass transfer is the main source of BSS formation in these regions, consistent with the previous theory that binary mass transfer dominates in areas of low local stellar density.

However for some fields such as Brick 9 Field 18, Brick 17 field 6, and Brick 20 field 7, a strong negative coefficient was found as seen in Table 2. The coefficients range from -2.818 to + 2.222, meaning that as local stellar density increases by 10 times, the chance of the star being a BSS changes by $e^{-4.231}$ or $e^{-3.587}$, or 0.015x to 0.028x, clearly showing that it is less likely for a star to be a BSS as local stellar density increases. This could be due to other stellar phenomena, such as age and metallicity gradients or photometric incompleteness. However, it is consistent that stellar collisions played little to no role in the formation of these stars. As the confidence intervals of the majority of coefficients have values including zero, it shows that binary mass transfer dominates BSS formation while stellar collisions play far less if any role. The fact that there are no significant correlations across fields is consistent with the theory that BSS formation is caused by binary mass transfer in field regions of large spiral galaxies supports the hypothesis.

## CONCLUSION

The purpose of this study is to identify which method causes BSS formation in field areas of large spiral galaxies: stellar collision or binary mass transfer. This is found through finding the $\beta_1$ coefficient for each field, with a positive coefficient supporting stellar collision and a negative or zero coefficient supporting the binary mass transfer theory. As all fields used in the study had a negative or zero coefficient, the binary mass transfer theory is supported, consistent with previous papers stating that binary mass transfer dominates BSS creation in areas of lower local stellar density.

Further research could include more fields representing different densities. It could also include fields within the bulge or halo, though that may be difficult due to the presence of stellar clusters. Using

UV photometry to differentiate young stars from BSS would also be necessary to get rid of false positives. It can also be expanded to other galaxies.

Currently, this result can only be extended to field regions within large galaxies. PHAT data may also be inaccurate, and some BSS candidates may not have been captured. The main-sequence turnoff point is also an estimate. There is also a very small amount of BSS, which could affect the data.

## CONFLICT OF INTEREST

There are no conflicts of interest regarding the publication of this article.

## REFERENCES


1. Ferraro FR, Lanzoni B, Dalessandro E, Beccari G and Pasquato M. Blue straggler stars: A direct comparison between globular clusters and dwarf galaxies. *Nature*. 2012; 492 (7429): 393-395. https://doi.org/10.1038/nature11686
2. Leigh NWC, Sills A and Knigge C. Blue straggler stars beyond the Milky Way: evidence for multiple formation mechanisms. *Monthly Notices of the Royal Astronomical Society (MNRAS)*. 2011; 416 (2): 1410-1421. https://doi.org/10.1111/j.1365-2966.2011.19136.x
3. Santana FA, Muñoz RR, Geha M, *et al*. Blue stragglers in the lowest stellar density systems. *Astrophysical Journal (ApJ)*. 2013; 774 (1): 106. https://doi.org/10.1088/0004-637X/774/2/106
4. Sollima A, Lanzoni B and Beccari G. The formation of blue straggler stars in dwarf spheroidal galaxies: A photometric study of the Sagittarius dwarf spheroidal galaxy. *Astronomy & Astrophysics (A&A)*. 2008; 481 (3): 701-706. https://doi.org/10.1051/0004-6361:20079082
5. Dalcanton JJ, Williams BF, Lang D, *et al*. The Panchromatic Hubble Andromeda Treasury (PHAT). *Astrophysical Journal Supplement Series (ApJS)*. 2012; 200 (2): 18. Available from: https://archive.stsci.edu/hlsp/phat (accessed on 2025-6-15).
6. Williams BF, Dalcanton JJ, Bell EF, *et al*. The Panchromatic Hubble Andromeda Treasury. IV. A Photometric Catalog of Over 100 Million Stars. *Astrophysical Journal Supplement Series (ApJS)*. 2014; 215 (1): 9. Available from: https://archive.stsci.edu/hlsp/phat (accessed on 2025-6-15). https://doi.org/10.1088/0067-0049/215/1/9
7. Johnson LC, Seth AC, Dalcanton JJ, *et al*. The Panchromatic Hubble Andromeda Treasury. X. Ultraviolet to Infrared Photometry of 117 Million Stars. *Astrophysical Journal (ApJ)*. 2016; 827 (1): 33. Available from: https://archive.stsci.edu/hlsp/phat (accessed on 2025-6-15).
8. Casertano S and Hut P. Core radius and density measurements in N-body experiments. *Astrophysical Journal (ApJ)*. 1985; 298: 80-94. https://doi.org/10.1086/163589