

Study of the Galactic Centre Region of the Milky Way Using Slitless Spectroscopy

Soumya Shreeram¹

This project was supervised by *N. Neumayer*, *F. Nogueras-Lara*, and *H.-W. Rix*.
Specialization semester project with *Max-Planck Institut für Astronomie (MPIA)*,
Königstuhl 17, 69117 Heidelberg, Germany.

¹ École Polytechnique Fédérale de Lausanne (EPFL), Route Cantonale, 1015 Lausanne, Switzerland.
EPFL Supervisor: *Prof. Frédéric Courbin*.
January 31, 2021

Current gaps in our knowledge of the Galactic Center (GC) region motivates the proposal of mid-res observations $R = 4000$ of over $\sim 10^{5-6}$ GC stars in the near-infrared Ks-band $2.1 \mu\text{m} \leq \lambda \leq 2.4 \mu\text{m}$ using slitless spectroscopy. The aim is distinguish between early and late types stars. The data used for modelling was obtained from the GALACTICNUCLEUS survey. Since standard approaches fail in crowded field of views like those found in the GC, the Nearest Neighbour Approach (NNA) was formulated for observing in the slitless mode. The spectral types of the target stars in $\sim 60 - 75\%$ regions, with 5 – 12 neighbouring stars, could be accurately recovered using slitless spectroscopy. These results indicate that observations in the slitless mode are feasible with some further improvements.

Contents

I	Introduction	
II	Data Used for the Study	
II-A	The GALACTICNUCLEUS Survey . .	
II-B	Field of View and Foreground Stars	
III	Observing and Modelling the Slitless Mode	
III-A	Properties of Stellar Spectra	
III-B	Building the Slitless Image	
IV	Extracting Spectral Information in the Slitless Mode	
IV-A	The Naive Approach	
IV-B	The Nearest Neighbour Approach for Crowded Field of Views	
IV-C	Fitting the Templates with the Data Regions	
V	Analysing the Best-Fitting Templates	
VI	Optimizing the Accuracy of the Nearest Neighbour Approach	
VI-A	Effects on Accuracy due to Number of Templates	
VI-B	Effects on Accuracy due to Pixel Scale	
VII	Conclusions and Future Prospects	

I Introduction

1 The Galactic Center (GC) of the Milky Way is 8 kpc
away from the Sun (Abuter et al., 2019; Do et al., 2019). At
3 the very center lies the supermassive black hole, Sagittarius
3 A* (Sgr A*) with a mass of $4 \times 10^6 M_\odot$ (Ghez et al., 2008;
6 Gillessen et al., 2009). As we move radially outwards
by ~ 5 pc, the dense environment around Sgr A* hosts
7 a Nuclear Star Cluster (NSC), which has a mass of
7 $2 \times 10^7 M_\odot$, and luminosity of $\sim 4 \times 10^7 L_\odot$ (Launhardt
7 et al., 2002; Schödel et al., 2014). The NSC represents
a dense cluster of stars that show a wide range of
morphologies (Neumayer et al., 2020). The close proximity
9 the GC of the Milky Way allows us to resolve individual
9 stars down to milli-parsec scales. This enables the study of
the stellar structure and population of the NSC and Milky
9 Way in unprecedented detail. The presence of such star
clusters in the ‘nuclei’ of spiral galaxies has been detected
in nearly $\sim 75\%$ of the spiral galaxy population (Georgiev
& Böker, 2014). This further motivates the study of
physical processes in the GC of the Milky Way, making it
an excellent laboratory to probe phenomena and compare
the results with other distant galaxies.

14 There have been extensive studies that explore the central
parsec surrounding Sgr A* (Genzel et al., 2010), the
14 Arches (Najarro et al., 2004; Harfst et al., 2010), and
the Quintuplet star clusters (Figer et al., 1999; Liermann
15 et al., 2012; Rui et al., 2019). However, the properties
of these regions are not representative of larger areas in
the GC region (Ryde et al., 2015; Schultheis et al., 2019).
Continuing the radial journey outwards from Sgr A* and



Fig. 1: The Galactic Center (GC) of the Milky Way observed in the near-infrared by VLT’s HAWK-I instrument with an angular resolution of 0.2 arcseconds. It was observed as part of the GALACTICNUCLEUS survey (Nogueras-Lara et al., 2018a, 2019), covering a continuous rectangular area of $35' \times 16'$, which is centered on Sgr A*. The orientation of the image is with galactic north towards the top and the galactic east towards the left. This RGB image contains $\approx 2 \times 10^6$ stars and it was produced by combining three broad-band filters: J, H, and Ks, each centered at $1.25 \mu\text{m}$ (in blue), $1.62 \mu\text{m}$ (in green), and $2.2 \mu\text{m}$ (in red) respectively. Figure taken from ESO & Nogueras-Lara et. al (2018).

NSC, on the scale of a ~ 230 kpc, a distinct stellar structure called the Nuclear Bulge (NB) is encountered. The NB defines the boundaries of the GC region, with a total stellar mass $\sim 10^9 M_{\odot}$ and luminosity $\sim 10^9 L_{\odot}$ (Launhardt et al., 2002). It hosts a rich and fertile star-forming region due to harsh environmental conditions (Morris & Serabyn, 1996; Launhardt et al., 2002). There are many open questions about the metallicity, star formation rate, and stellar population in the NB. Observations of the GC suffer from high interstellar extinction from dust clouds, and dense stellar crowding (Nishiyama et al., 2008; Fritz et al., 2011; Nogueras-Lara et al., 2018a, 2019). Accurate data and wise observational strategies are key to studying the GC of the Milky Way.

The motive of this project is to bridge gaps in our current knowledge about the GC of the Milky Way. This can be done by observing the GC stars using slitless spectroscopy. The slitless mode allows us to know the spectral types of the stars in the NB, which paves the way to disentangle the star formation history of the NB. Theories on star formation (SF) predict the distribution of stellar masses, or the Initial Mass Function (IMF). Current observations of the Milky Way show no ambiguity in the IMF between different parts of the galaxy, with exceptions arising only in regions with star clusters (McKee & Ostriker, 2007). Spectroscopic wide-area observations in the GC provides the opportunity to probe SF theories in this region and complement previous work (Matsunaga

et al., 2011; Minniti et al., 2016; Nogueras-Lara et al., 2018b, 2020). Studies in this field are quintessential in understanding the formation and evolution of our Milky Way. In order to know the spectral types of the stars in the NB, the stars in the region must be modelled by three main parameters: the effective temperature of the star, T_{eff} , surface gravity, $\log g$, and the metallicity, $[\text{Fe}/\text{H}]$. In this project, the high-resolution PHOENIX stellar library (Husser et al., 2013) was used for modelling spectra.

Slitless spectroscopy is an economic technique that can cover large areas in the GC region, and these surveys can be conducted using ground-based telescopes like GranTeCan or Gemini South that are mounted with a grism-spectrograph (Garzón et al., 2006; Eikenberry et al., 2004). Observations in the slitless mode are a combination of three components: (1) the dispersed 2D spectroscopic image, which requires a spectrograph with a grism, (2) simultaneously obtained *standard* image of the same FoV, and (3) *prior* imaging of the same FoV that is used for source identification. The prior images must be preferably taken with the same broadband filter as the grism and it must contain the positions and magnitudes of the objects in the FoV. Only then can these prior images be used as a reference for the *standard* images (Ryan Jr et al., 2018). It was identified long ago that slitless or grism observations suffer from a high-level of sky background (Gai et al., 1997), deeming this technique as ‘difficult’. Indeed, decoding the spectra

of objects in the slitless mode is a challenge in itself, and over-crowded field of views, like those in the GC regions, supplement the difficulty of this problem. This is because over-crowded fields experience maximal spectral overlap when observed with a grism. However, these hurdles can be overcome by having accurate knowledge about the positions of the dispersed objects, careful calibration data, and accounting for unresolved background emissions appropriately. Additionally, conducting multiple scans of the same region with different dispersive orientations can minimize the overlap of stars. Thanks to the GALACTICNUCLEUS survey (Nogueras-Lara et al., 2018a, 2019), the information on positions and magnitudes of over 10^6 stars in the GC region is known. This survey can be used as *prior* imaging data required to do slitless spectroscopy.

In the past, slitless observations have been conducted using various HST instruments like the ACS, WFC3, and more recently, NICMOS (Walsh et al., 2010; Pasquali et al., 2006). Additionally, various software have been developed for extracting the spectral information from slitless images (Pirzkal et al., 2001; Bruzual & Charlot, 2003; Kümmel et al., 2009; Ryan Jr et al., 2018). However, these softwares are either specific to HST instruments or they are not equipped to deal with the scale of crowded fields that are encountered in the GC region of the Milky Way ($\sim 10^4$ stars in a $3.6' \times 6'$ patch). Therefore, this study devises a complete machinery for modelling and extracting spectral information from slitless images. The software developed is applicable for over-crowded fields like those found in the GC of the Milky Way (10^{5-7} pc^{-3} ; Launhardt et al., 2002).

This paper discusses the feasibility of proposing mid-res observations $R = 4000$ in the near-infrared Ks-band $2.1 \mu\text{m} \leq \lambda \leq 2.4 \mu\text{m}$ to observe over $\sim 10^{5-6}$ GC stars of the Milky Way using slitless spectroscopy. There are two absorption lines that aid in distinguishing stellar types: Br γ line ($2.16 \mu\text{m}$) present in spectra from hot, young stars (Beck et al., 2010) and carbon monoxide (CO) lines ($\sim 2.3 \mu\text{m}$) present in spectra from cooler stars (Feldmeier et al. (2014); Fig. 7). The proposed observation can draw distinction between hot and cold stars, thus, paving the way to infer properties of the stellar population in the GC region. Thereby, this technique provides us information about the probable regions with dense young stars, and if possible, also understand the proper motion and $\log g$ of objects in the Nuclear Bulge. The detection of young star regions, so-called Young Massive Clusters (YMC) is of particular interest. These young stellar clusters have shown to be the source of generating intermediate-mass black holes (Zwart et al., 2004; Gürkan et al., 2004). Matsunaga et al. (2011) have shown that $\sim 10^6$ stars must have

formed within the last 30 Myr, however, less than 10% of the predicted YMCs have been detected so far. The main reason for these YMCs escaping detections is due to high-interstellar extinction and high stellar density.

The structure of the paper is detailed here. Sec. II-A introduces GALACTICNUCLEUS survey, which encompasses the data used for modelling in this project. Comments of the chosen Field of View (FoV) and contributions from the foreground population along the line of sight (LoS) to the GC are addressed in Sec. II-B. A brief note on the instruments that can be used to observe in the slitless mode is mentioned in Sec. III. Sec. III-A explains how the spectra from the PHOENIX library (Husser et al., 2013) were used for the modelling procedure. Lastly, the procedure for building the slitless image using the stellar positions and magnitudes from the GALACTICNUCLEUS survey is explained in Sec. III-B. The crucial part of the report is devising a strategy to extract stellar information from the multiple-object image, which is modelled using the principles of slitless spectroscopy. The Nearest Neighbour Approach (NNA) is introduced and implemented for this purpose in Sec. IV-B, and details of the fitting procedure are provided in Sec. IV-C. The outputs from the fitting procedure are analyzed in Sec. V. Finally, the results of the analysis are optimized with respect to two parameters that are discussed in Sec. VI. The conclusions from the project are summarized in Sec. VII and a list of points to be executed in the future is provided.

II Data Used for the Study

The section aims to provide a general overview about the data that was used for modelling areas in the GC in this project. Sec. II-A introduces the GALACTICNUCLEUS survey. Sec. II-B explains the chosen FoV for the analysis. It also discusses the number distribution of foreground stars that contribute in the modelling process.

II-A The GALACTICNUCLEUS Survey

The data used for modelling in this project was obtained from the GALACTICNUCLEUS survey (Nogueras-Lara et al., 2018a, 2019). This survey was conducted using the Near-Infrared (NIR) camera mounted on the HAWK-I (High Acuity Wide field K-band Imager) instrument (Kissler-Patig et al., 2008). HAWK-I is installed on the unit telescope 4 (UT4), which is part of ESO's Very Large Telescope (VLT). The VLT facility is located at the Paranal observatory in Chile. Three broad-band filters were used to observe the GC: J , H , and K_s , each centered at $1.25 \mu\text{m}$, $1.62 \mu\text{m}$, and $2.2 \mu\text{m}$, respectively. The GALACTICNUCLEUS survey consisted of 49 pointings,

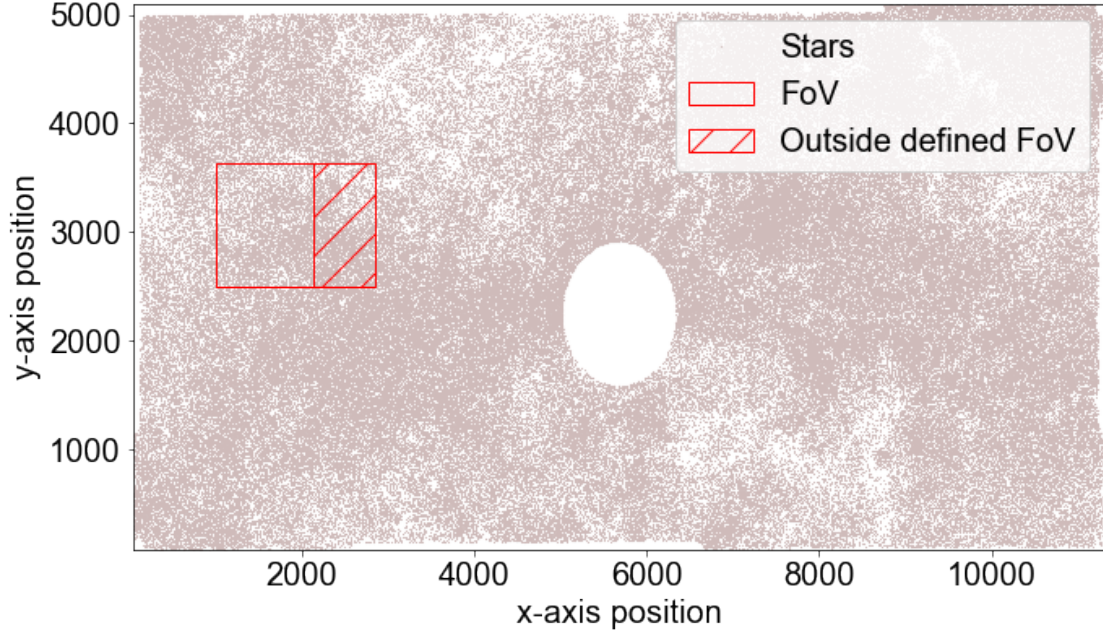


Fig. 2: All the stars (2×10^6) in the central catalogue covered by the GALACTICNUCLEUS survey (Nogueras-Lara et al., 2018a, 2019), as observed by the HAWK-I instrument (Kissler-Patig et al., 2008), are plotted. The extensively studied NSC has a half-light radius of ~ 5 pc (Schödel et al., 2014). Since the aim of the project is to study physics within the NB, the central 5 pc around Sgr A* was not considered for this analysis; hence, this central region has been cropped out. The highlighted red box (with the hatched region) displays a patch of $6.0' \times 3.6'$ of the GC region; this corresponds to 1855×1128 pixels² in an instrument like EMIR (see Sec. III for further details). The hatched region within the highlighted patch, shows the stars whose stellar information cannot be extracted (see discussion in Sec. III-B). However, these stars influence the defined FoV; therefore, they are considered in the analysis.

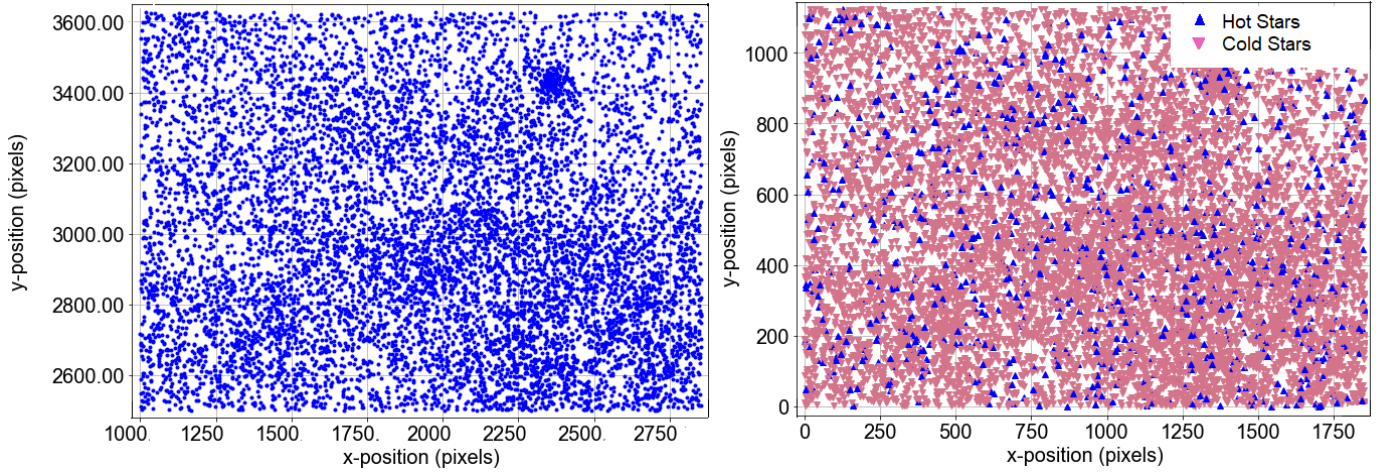


Fig. 3: **Left:** Zoomed in view of the highlighted-red box shown Fig. 2; this region shows the *spatial stellar* distribution of 9472 stars in the selected pointing. **Right:** Zoomed in view of the highlighted red box showing the *stellar type* distribution. All the stars were randomly assigned one of the two stellar types: hot or cold; 10% of the 9472 stars in the effective FoV were assigned to be hot stars. Only stars with magnitude in the $Ks < 15$ mag are shown (see Sec. II-B for discussion). These stars were corrected for extinction and reddening.

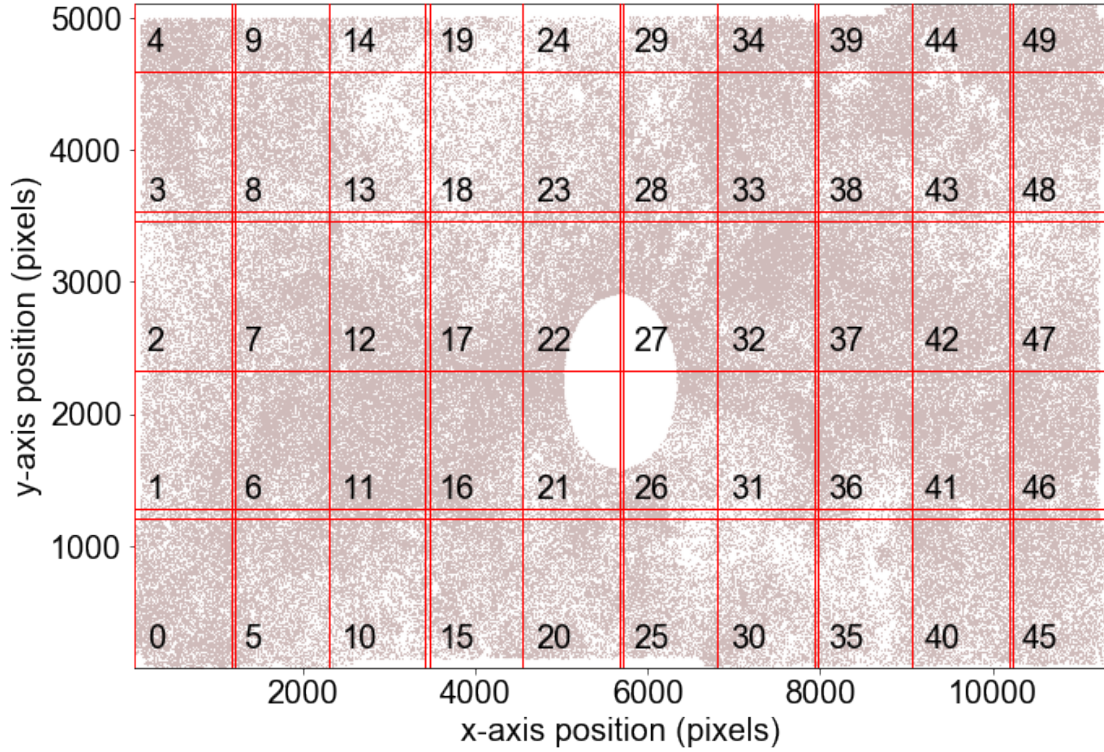


Fig. 4: 2×10^6 stars in the central catalogue covered by the GALACTICNUCLEUS survey. The catalogue has an area of $35' \times 20'$ with an angular resolution of 0.2 arcseconds. It is divided into grids of $\sim 3.6' \times 3.6'$, or equivalently into 1128×1128 pixels². These grids are labelled with numbers on the plot.

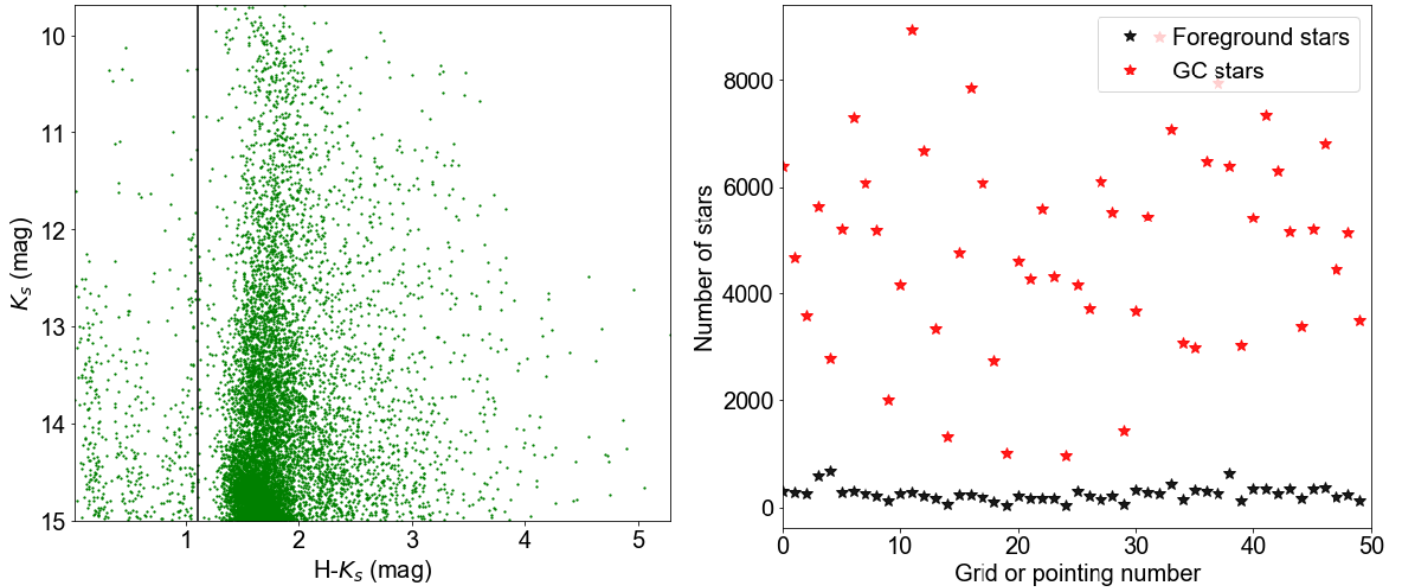


Fig. 5: Left: Color-magnitude diagram for the selected region of $3.6' \times 6.0'$, as shown in Fig. 2, from the central catalogue of the GALACTICNUCLEUS survey. The black line defines a cut at $(H - K_s) = 1.1$, which effectively removes the foreground population. This cut accounts for the NB contamination and enables better distinction between the foreground stars from the GC stars. In the case shown, there are 415 stars to the left of the cut that correspond to the foreground population. **Right:** The distribution of foreground (black) and GC stars (red) as a function of grids or pointings. The grid number is obtained by dividing the central catalogue, as shown at the top in Fig. 4. This corresponds to the information that can be extracted from approximately one pointing of the telescope. Although all the grids are symmetrically defined squares, they do not cover the same stellar surface area due to the central cropped region around Sgr A* and edge effects.

where each pointing defines a FoV of $\sim 7.2' \times 2.6'$ ¹. Of the 49 pointings, 30 pointings together define the central catalogue. Fig. 1 shows the central catalogue that mapped a continuous rectangular region of $35' \times 16'$ centered on Sgr A*. The remaining 19 pointings observed as part of the survey covered the following regions: east and west of the central catalogue, transition region between the GC and the inner bulge, and regions in the inner bulge of the Milky Way. The survey covered a total area of 0.3 deg^2 , which corresponds to $\sim 6000 \text{ pc}^2$. Additionally, it achieved 5σ detections for $J \sim 22 \text{ mag}$, $H \sim 21 \text{ mag}$, and $Ks \sim 22 \text{ mag}$. Of the 3.3×10^6 stars detected in the survey: 20% belong to J , 65% belong to H , and 90% belong to Ks . This survey supersedes previous surveys by several orders of magnitude. Hence, it was used as the prior imaging data for the analysis conducted in this paper. Only the central catalogue, which encompasses the nuclear bulge, will be used to study and model the GC region.

II-B Field of View and Foreground Stars

In order to study the feasibility of observing the GC in the slitless mode, a $3.6' \times 6.0'$ area was chosen from the central catalogue of the GALACTICNUCLEUS survey, as shown in Fig. 2 and 3. The dimensions of the area were chosen such that they are comparable to one pointing of ground-based telescopes like GranTeCan or Gemini South (see Sec. III). Within the chosen $3.6' \times 6.0'$ area, the stellar information of the stars in the hatched region in Fig. 2 cannot be extracted; this corresponds to an area of $2.4' \times 3.6'$. The reason why this information is cannot be extracted will be explained in Sec. III-B. Therefore, the term FoV is given to the $3.6' \times 3.6'$ area to the left. In principle, the stellar information of all the stars within the FoV can be extracted from an image that is observed in the slitless mode. For this analysis, the plate scale of $0.193''/\text{pixel}$ is used, as this corresponds to the plate scale of the EMIR instrument mounted on GranTeCan (Garzón et al., 2006). Therefore, $3.6' \times 6.0'$ area corresponds to $1128 \times 1855 \text{ pixel}^2$.

The observations of the GC are contaminated by foreground stars present along the LoS. These foreground stars could be useful in studies of the spiral arm structure of the Milky Way (Nogueras-Lara et al., 2018a; Vallée, 2005; Vallee, 2002). However, for the purposes of this project, these stars do not assist in studying the dynamics of the GC. One method to identify the foreground stars is by studying the Color-Magnitude Diagram (CMD), as shown in the left panel of Fig. 5. CMDs assist in identifying stellar populations (Morgan & Keenan, 1973, p. 400). The H band magnitude of a star is plotted as a function of

$H - Ks$. A cut at $H - Ks < 1.1 \text{ mag}$ draws distinction between the foreground stars and the GC stars. This is because it is known that these stars are likely to trace the spiral arm structure along the LoS from the Earth to the GC (Nogueras-Lara et al., 2019). In order to study the GC in the slitless mode, the number of foreground stars, which mostly belong to the Main Sequence, must be known. By applying the color-cut criterion, it was possible to study the number of foreground vs GC stars in all pointings, as shown in Fig. 5, of the central catalogue. To conduct this study, the central catalogue was divided into $3.6' \times 3.6'$ rectangles, so-called grids, as shown in Fig. 4. Overall, there were ≈ 49 grids. The number of foreground and GC stars is plotted as function of these grids in the right panel of Fig. 5. Thus, we can conclude that one pointing towards the GC could have $\approx 10^3 - 10^4$ stars, of which 2 – 5% belong to the foreground star population.

The foreground population can be dealt with using the following approaches:

- (A) Discarding the foreground stars completely. For the case shown in the left panels of Fig. 3 and Fig. 5, there are 9742 stars in the pointing, of which $\sim 4\%$ correspond to foreground stars.
- (B) Modelling the foreground star contributions by assuming fixed stellar types. These stars mostly belong to the Main-Sequence as indicated by the Nogueras-Lara et al. (2019). Therefore, their contributions can be subtracted from the slitless image, such that only stars from the GC are analyzed.
- (C) Keeping the foreground contributions by modelling all the stars in the observed image. This is a brute force approach that increases the complexity of the problem. However, if the techniques of extracting stellar information is successful, this procedure would not only give information about the GC stars, but also the stars along the LoS.

Of these three approaches, option (A) does not closely represent the image that is obtained from observations in the slitless mode. This is because the foreground contributions cannot be dismissed during the observations itself. Option (B) provides a smart solution since it takes advantage of the information from the central catalogue. However, there remains a level of ambiguity in assuming the exact stellar types of all the foreground stars. Additionally, the results become dependant of these assumptions. Therefore, option (C), although more difficult, appears to be the most conclusive way to approach the foreground star population.

Another important point to note is that only stars with magnitudes in the $Ks < 15 \text{ mag}$ were chosen for the analysis. The reason for this selection criterion is motivated by the fact that surveying the sky in the slitless mode results

¹The jittering technique used for the survey increased the effective FoV per pointing.

in high-levels of background contamination. The way to minimize this would be to observe the GC with short exposure times, which would result in the faint stars with magnitudes in the $Ks \geq 15$ mag being unresolved. It must be noted that, although the stellar properties of these faint stars will not be decodable, their presence contributes to the background noise that must be appropriately subtracted. Thanks to the GALACTICNUCLEUS survey, the positions and magnitudes of these faint stars are known. Therefore, the task of modelling them and appropriately subtracting their contributions remains feasible.

III Observing and Modelling the Slitless Mode

There are two possibilities to observe the GC region of the Milky Way in the slitless mode from ground-based observatories:

- 1) **EMIR** (Espectrógrafo Multiobjeto Infrarrojo) is a near-infrared camera-spectrograph of the Gran Telescopio Canarias (GranTeCan) that can observe in the K-band (Balcels, 1998; Garzón et al., 2006). It is a 10 m telescope located at La Palma, in Canarias, Spain, which has a FoV of $6.6' \times 4'$ and a plate scale of $0.193''/\text{pixel}$.
- 2) **Flamingos-2** is a near-infrared imaging spectrograph on the Gemini South telescope located at Chile (Eikenberry et al., 2004). It has a round FoV of $6'$ and a plate scale of $0.18''/\text{pixel}$.

The properties of stellar spectra used for modelling are detailed in Sec. III-A, and lastly, the steps implemented for modelling the final flux image in the slitless mode is explained in Sec. III-B.

III-A Properties of Stellar Spectra

Stellar spectra can be characterized by three parameters: the effective temperature of the star, T_{eff} , surface gravity, $\log g$, and the metallicity, $[\text{Fe}/\text{H}]$. Additional parameters like α -enhancement, which measures deviations from *standard* abundances, micro-turbulence velocity, and interstellar extinction assist in characterizing spectra, however, these can be fixed using prior information. There have been multiple studies in the literature to estimate the main stellar parameters to convert astrophysical observations into concrete physical interpretations (Decin et al., 2004; Bailer-Jones, 2000; Shkedy et al., 2007). Over the years, this motivated the need for a complete spectral library for analyzing spectra for large data surveys. The synthetic stellar spectra were taken from the PHOENIX library (Husser et al., 2013). The library covers a wide range of parameter space with $2300 \leq T_{\text{eff}} \text{ (K)} \leq 12,000$, $0.0 \leq \log g \leq +6.0$, and $-4.0 \leq [\text{Fe}/\text{H}] \leq +1.0$ with high-resolutions in the NIR spectra, $R = 10^5$. For the

purpose of this project, 10 arbitrary T_{eff} and $\log g$ were chosen, as shown in Tab. I, which spanned the entire range of temperatures covered by the PHOENIX library. The spectra are plotted in Fig. 6; they display two important absorption lines: the $\text{Br}\gamma$ line at $2.16 \mu\text{m}$ is distinct in spectra of hot, young stars (Beck et al., 2010) whereas the carbon monoxide (CO) lines at $\sim 2.3 \mu\text{m}$ are pronounced in cooler and older stars (Feldmeier-Krause et al., 2017).

As the aim is to observe at the resolution $R = 4000$, the synthetic spectra are binned such that only the $\text{Br}\gamma$ and the CO lines are recognizable. The value used for binning the spectra is explained in Sec. III-B. Furthermore, all the star spectra are normalized and damped with a Gaussian Line Spread Function (LSF), as shown in the right panel of Fig. 6. This implies that the reshift information of these stars is lost; therefore, the velocities of these stars cannot be determined. By damping the spectra, we also lose information on the $\log g$. This is because the spectral signature of $\log g$ is determined by the line widths, which is decipherable only at high-resolutions (Bailer-Jones, 2000). However, one of the future goals of this project would be to reconstruct the original spectra by appropriately subtracting the LSF. This would allow us to estimate the LoS velocity and $\log g$. The GC stars show solar or super-solar metallicities $[\text{Fe}/\text{H}]$ (Feldmeier-Krause et al., 2017; Schultheis et al., 2019; Nogueras-Lara et al., 2020) with no α -enhancement (Nogueras-Lara et al., 2020). Therefore, $[\text{Fe}/\text{H}]$ and α -enhancement was fixed to zero, i.e. the solar value. During the procedure of binning and damping the spectra, we are left with just one free parameter: the effective temperature of the star. From Fig. 6 it must be noted that the $\text{Br}\gamma$ line is present in hotter stars $T_{\text{eff}} > 5900$ K, whereas the CO lines are observed in the cooler stars, $T_{\text{eff}} < 5900$ K. This motivates a further simplification in the modelling procedure: all stars can be grouped into two spectral types, hot stars h_* with $T_{\text{eff}} > 5900$ K and cold stars c_* with $T_{\text{eff}} < 5900$ K. In essence, the rest of the project would aim to address the following question: *is it possible to extract the hot and cold stars observed in the slitless mode?*

III-B Building the Slitless Image

One of the most crucial steps in this project was to simulate 9742 stars shown in Fig. 3 in the slitless mode. The steps executed in order to model crowded fields in the slitless mode are listed below:

- **Binning the spectra and the dispersion range:** The right panel in Fig. 3 shows the spacial distribution of stellar types in the selected pointing. The aim is to associate each star shown in Fig. 3 with a spectrum that is dispersed along the spectral axis. In this case the spectral axis is defined to be parallel to the x -axis.

T_{eff}	12000	11800	10000	8400	7600	6900	5900	5100	4700	3900
$\log g$	6.00	2.50	3.00	6.00	3.50	3.00	3.00	2.00	1.50	6.00

TABLE I: The effective temperature, T_{eff} (Kelvin), and surface gravity, $\log g$, of 10 arbitrary stellar spectra chosen for analysis. The metallicity and α enhancement were set to zero for all the spectra. The highlighted T_{eff} and $\log g$ columns corresponds to the spectra that were chosen to represent hot (*blue*) and cold (*red*) stars.

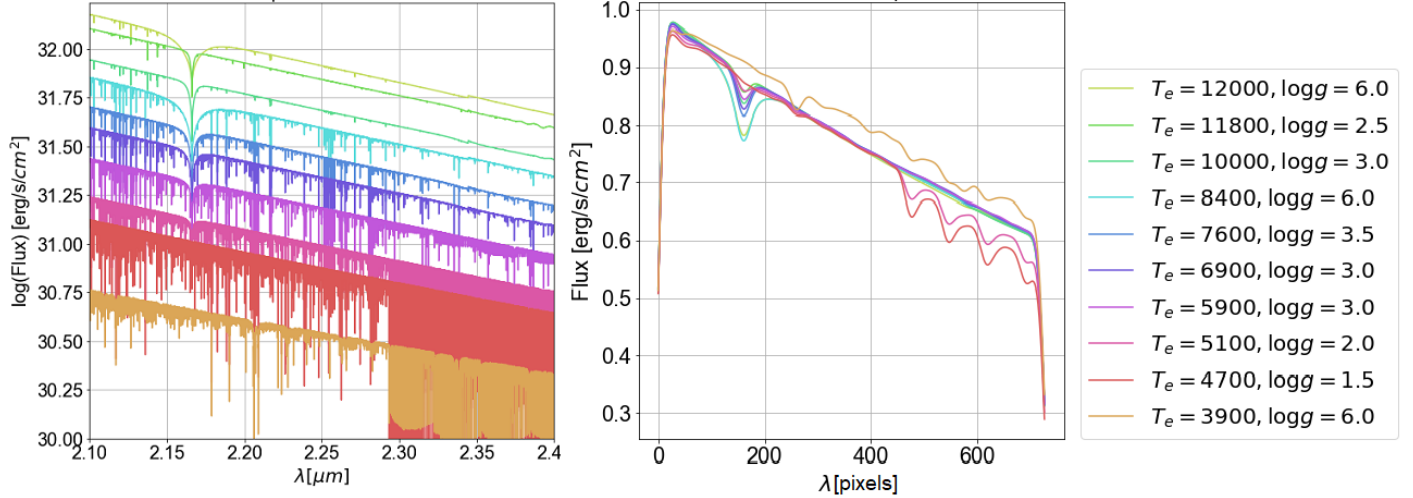


Fig. 6: **Left:** Synthetic stellar spectra of 10 stars chosen from the PHOENIX library (Husser et al., 2013) in the Ks-band with bandwidth $\Delta\lambda = 0.3 \mu\text{m}$. Each spectra has a different values of T_{eff} and $\log g$, as shown in the legend. The GC stars show solar or super-solar metallicities $[\text{Fe}/\text{H}]$ (Feldmeier-Krause et al., 2017; Schultheis et al., 2019; Nogueras-Lara et al., 2020) with no α -enhancement (Nogueras-Lara et al., 2020). Therefore, both these parameters were set to zero for all the spectra. **Right:** The binned, damped, and normalized synthetic spectra are shown to the left. The color codes of these spectra are consistent with the figure on the left. There two distinguishable absorption line features in the spectra: $\text{Br}\gamma$ line ($2.16 \mu\text{m}$) that correspond to hot stars and carbon monoxide (CO) lines ($\sim 2.3 \mu\text{m}$) from cooler stars.

However, in general, it could be tilted by an angle, θ , with respect to the x -axis; the modelling procedure accounts for this possibility. The *dispersion range*, d , of the spectra that are associated to stars is defined as

$$d = \frac{\Delta\lambda}{\lambda_0} R, \quad (1)$$

where λ_0 is the central wavelength in the observing broadband of width $\Delta\lambda$, and R is the resolution of the observing instrument. For an instrument like EMIR, $R = 4000$ pixels, and for the Ks -band in the NIR, $\Delta\lambda = 0.3 \mu\text{m}$ and $\lambda_0 = 2.2 \mu\text{m}$. Therefore, the high-res spectra from the PHOENIX library are binned into $d = 727$ pixels. Once these spectra are binned, they are damped by the Gaussian LSF.

- **Associating stars with their corresponding spectra:**

This step involves defining the spectral axis for dispersion for every object in the pointing. Before delving into a pointing with 9742 stars, the procedure of associating spectra was implemented in a simple 10 star FoV, shown in Fig. 7. Given a target star at position $\vec{x}_T = (x_T, y_T)$ that is dispersed along the direction $\vec{\theta} = (\cos \theta, \sin \theta)$, the star can be associated

with its corresponding spectra by the dispersion vector \vec{x}_D (pixels) and \vec{y}_D (pixels) along the x - and y -axis respectively as

$$\vec{x}_D = (x_T, x_T + 1, \dots, x_T + d - 1) \text{ and}, \quad (2)$$

$$\vec{y}_D = \tan \theta \vec{x}_D + \vec{c}, \quad (3)$$

where $\vec{c} = (\tan \theta - y) \vec{x}_D$ is the intercept vector. This dispersion vectors \vec{x}_D and \vec{y}_D are placed appropriately at the position of the star, resulting in the highlighted tail shown in the left panel of Fig. 7. The spectra are normalized; therefore, when the spectra are associated to a particular star, they are scaled according to the flux of that star. In the case of the selected pointing with 9472 stars, the value for the flux of each star was taken from the central catalogue of the GALACTICNUCLEUS survey. Additionally, a Gaussian PSF was added along the y -direction for every pixel along the dispersed spectral axis. The standard deviation for the PSF was set to be $\sigma_{\text{PSF}} = 3$ pixels. This value was motivated by the typical scale at which seeing influences the observations from an instrument.

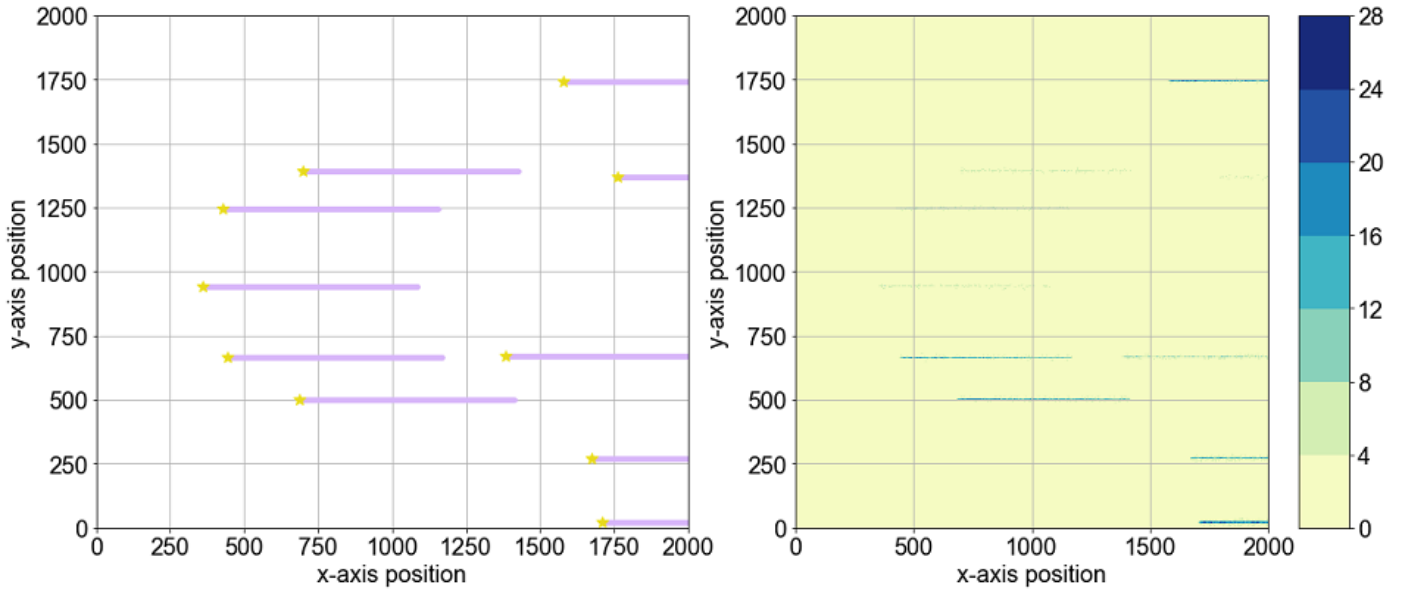


Fig. 7: Left: Ten stars (yellow) are dispersed along their spectral axis, which appears as a lagging tail behind the object. For the given case, the spectral axis is parallel to the x -axis. **Right:** A contour plot showing flux along the z -axis. The plot to the left is used as a template to add the flux for every (x, y) position. A Point Spread Function (PSF) is applied to all the dispersed stars. Lastly, Poisson Noise is added to the entire image.

- **Adding flux to the image:** It must be noted that the Fig. 7 (left) represents the collapse of the spectral dimension on the (x, y) spacial dimensions. The previous step laid the groundwork for adding flux along the z -dimension to produce the final image. For the 10 stars case, the result is shown in Fig. 7 (right), whereas for the selected pointing with 9472 stars, the result is shown in Fig. 8. The image corresponds to a two dimensional matrix, where every (x, y) point holds a flux value. The last step involved the addition of Poisson noise to all the pixels.

A final comment must be made on the edges affect slitless observations. Fig. 8 shows that some of the spectra are cut to the right due to the position of the star being close to the edge of the FoV. This implies that the spectral information can be extracted only for those stars whose complete spectra falls within the defined FoV. This corresponds to the $3.6' \times 3.6'$ area to the left of Fig. 8, or equivalently into 1128×1128 pixels² area. Hereafter, Fig. 8 is called the *data image* and it will be analyzed to extract stellar information.

IV Extracting Spectral Information in the Slitless Mode

The following section aims to discuss the techniques devised to extract stellar information in the slitless mode. The executions of these techniques are summarized in Sec. IV-A and Sec. IV-B.

IV-A The Naive Approach

A natural way to approach the problem would be to model the entire FoV. The pointing selected for analysis in this project contains 9472 stars, of which 5417 stars are within the FoV, and the remaining 4055 stars are outside the FoV. It must be noted that the stars *outside* the FoV also need to be modelled to find the stellar types of the stars *within* the FoV. This is because, although the stellar types of the stars outside the FoV cannot be known (since their spectra are cut due to the edges), they influence the stars within the FoV. Hence, in order to obtain accurate predictions of the stars within the FoV, we need to model all the 9472 stars in the chosen pointing.

Assuming that 10% of the stars are hot, the total number of distinct permutations that must be considered to extract the stellar types of all the stars in the FoV are $P_{947}^{9472} = 2.7 \times 10^{3744}$. Clearly, this method becomes unfeasible due to the huge number of templates that must be generated, resulting in overuse of computational resources and storage space. Additionally, all these templates are specific to the region of concern, so there is no scope for recycling the generated templates. Due to these pitfalls for over-dense regions with greater than 10^3 stars, a new and efficient method must be adopted.

IV-B The Nearest Neighbour Approach for Crowded Field of Views

Let's consider a target star, T , located in the FoV at position $\vec{x}_T = (x_T, y_T)$. Fig. 9 shows the nearest

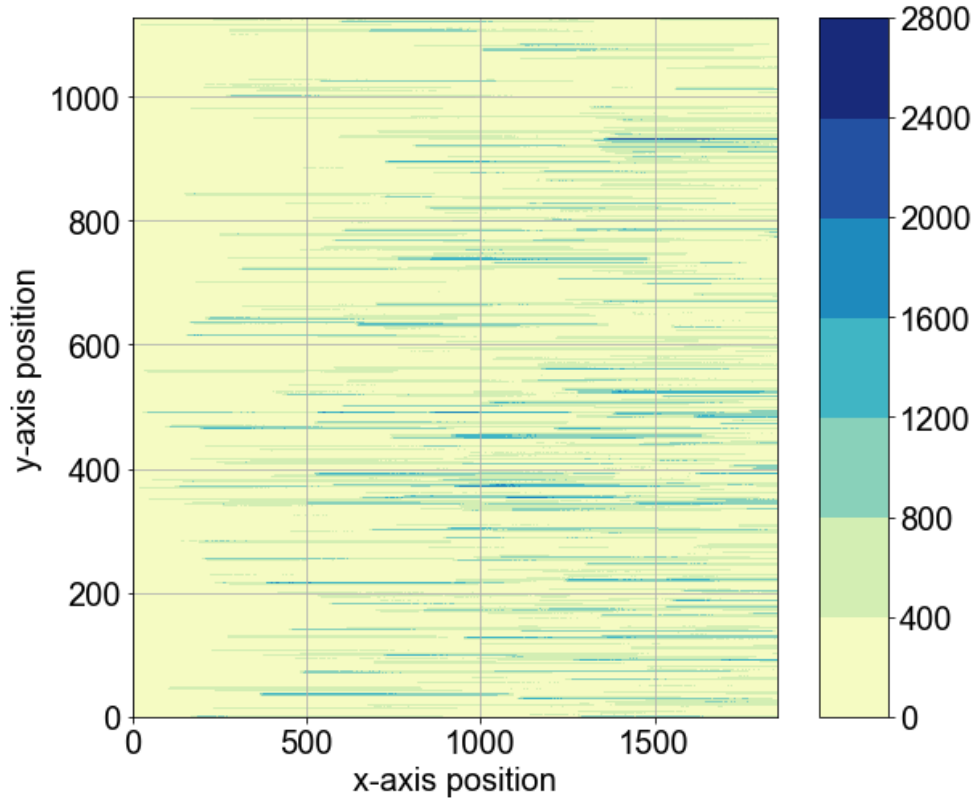


Fig. 8: Simulated image in the slitless mode containing 9472 stars. The stars belong to the selected pointing (see Fig. 2) that was chosen from the central catalogue of the GALACTICNUCLEUS survey. The spectra associated to the stars are damped with a Gaussian LSF. A Gaussian Point Spread Function (PSF) is applied along the y -direction for every point along the spectral axis with $\sigma_{PSF} = 3$ pixels; this was motivated by the typical scale at which seeing influences the observations. Additionally, Poisson noise is added to the image. Hereafter, this figure is called the *data image*; it will be analyzed to extract stellar information.

neighbours to T that influence it. The neighbouring stars, k_{\pm} , are chosen such that

$$\vec{x}_{k_+} \in [x_T + d, y_T \pm \sigma_{PSF}] \text{ and,} \quad (4)$$

$$\vec{x}_{k_-} \in [x_T - d, y_T \pm \sigma_{PSF}], \quad (5)$$

where d is the dispersion range of T along the x -direction, and σ_{PSF} is the spread in the PSF of the spectrum of T ; σ_{PSF} is defined by the seeing conditions of the telescope. The Nearest Neighbour Approach (NNA) relies on defining the Region of Influence (RoI) for T and modelling all the stars within this region, \mathcal{R} . The total number of stars, n , in $\mathcal{R} \equiv \mathcal{R}(n)$ is given by

$$n = T + k_+ + k_-. \quad (6)$$

The advantage of NNA is that over-dense regions can be modelled by discarding concerns regarding the total stars in a pointing. The original pointing chosen for analysis in this project contains 9472 stars, however, with NNA each star can be modelled by considering only its neighbours in \mathcal{R} . Fig. 10 illustrates the execution of the NNA.

The natural question arises when Fig. 10 is analysed: how many neighbours does a star have on average in the GC region? This question is addressed in Fig. 11, where the distribution of total number of neighbours is studied.

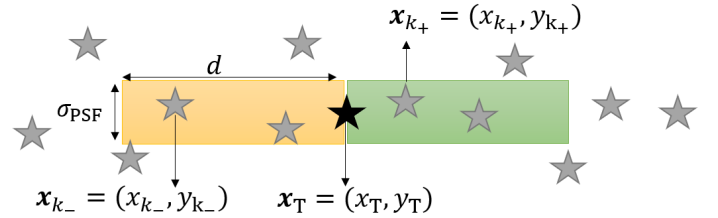


Fig. 9: Illustration of the Nearest Neighbour Approach (NNA). The target star, T , at position, \vec{x}_T , is influenced by the stars within the area $(2d \times \sigma_{PSF})$ pixel², so-called as the Region of Influence (RoI) or \mathcal{R} . The stars to the left (*yellow region*) of T in the RoI are labelled as k_- , and those to the right (*green region*) of T in the RoI are labelled as k_+ .

For the chosen pointing, containing 9472 stars, each star within the FoV has neighbours ranging from $\approx [5, 40]$ stars. The right panel of Fig. 11 shows that most stars in the defined pointing have ≈ 15 neighbours. Modelling regions with stars approximately of the $\mathcal{O}(10)$ is a feasible task. The methodology by which this task is executed is enumerated below:

- 1) **Grouping of the regions with same number of stars:** The data image is shown in Fig. 8. The first step consists of forming a group of all the regions,

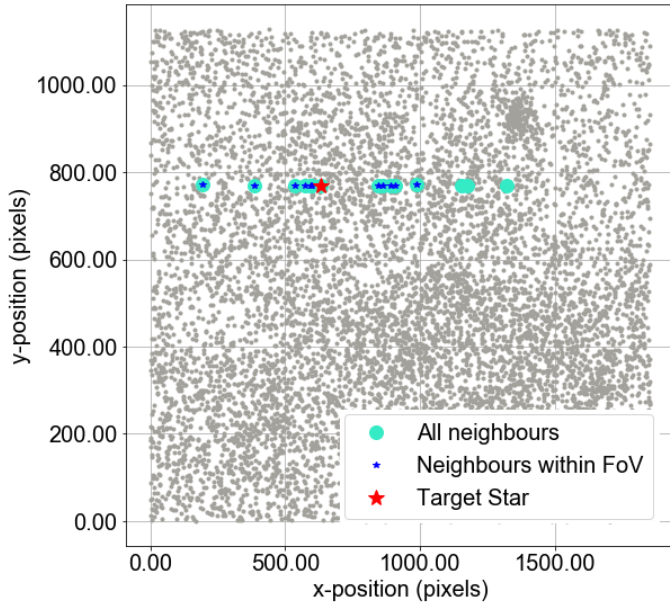


Fig. 10: Demonstration of the NNA approach. A random star is chosen as the target star, T , (in red) and its neighbours (in blue) are highlighted. For the given case, T has 14 neighbours: 11 within the FoV, and 3 outside the FoV (FoV is defined as 1128×1128 pixel²; see Sec. II-B for discussion of FoV). The neighbours and the target stars together define the Region of Influence $\mathcal{R}(n)$, where $n = 14 + 1$, where the $+1$ denotes that T is also counted.

\mathbf{R}_N , within the FoV that have the same number of stars n :

$$\mathbf{R}_N(n) = \{\mathcal{R}_1(n), \mathcal{R}_2(n), \dots, \mathcal{R}_N(n)\}, \quad (7)$$

where N denotes the size of the group and $\mathcal{R}_1(n)$ is a labelled region in the data image, containing n stars. In other words, all the regions containing the same number of neighbours form a group, and were stored in the same directory. As an example for $\mathcal{R}(n = 11)$: there exists 255 regions in the data image that contain 11 stars; therefore, $N = 255$. One arbitrary $\mathcal{R}(n = 11)$ is shown in Fig. 12, whereas Fig. 13 shows all the regions $\mathbf{R}_N(n = 11)$. The similar grouping procedure was applied for all $n \in [5, 44]$, where each target star is part of one \mathbf{R}_N .

- 2) **Creation of a template library:** Given $\mathcal{R}(n)$, there are possibilities for different number of hot stars, h_* , in the data region. For a fixed number of n and h_* , there are $P_{h_*}^n$ number of unique permutations of stars that can be constructed, where

$$P_{h_*}^n = \frac{n!}{(n - h_*)!h_*!}. \quad (8)$$

As an example, for $n = 5$ and $h_* = 3$, $P_{h_*}^n = 10$; this implies that there are 10 templates that can

be constructed to fit the region containing 5 stars, of which 3 are hot and the remaining are cold stars, $\mathcal{R}(n = 5, h_* = 3)$. The template library was constructed for $n \in [5, 20]$, where for each case of n , all possibilities of hot stars were accounted i.e. $h_* \in [0, n]$. The information on all these unique permutations were stored as template, which can be later called for generation of the flux image.

- 3) **Using the template library to build flux images:**

The templates generated in the previous step are completely recyclable to fit another region $\tilde{\mathcal{R}}(n = 5, h_* = 3)$, within the same pointing or a different one. This was made possible on account of the following: for n stars, and h_* hot stars, the spectra and order of stars were stored. For a particular $\mathcal{R}(n, h_*)$, the templates were used to associate the (normalized) spectra to the stars for the given spacial-distribution in the x - y plane. The normalized spectra were scaled appropriately according to the flux of the star that was obtained from the central catalogue of the GALACTICNUCLEUS survey. Therefore, this technique allows to use the template for any arbitrary $\tilde{\mathcal{R}}(n, h_*)$, even though the stars in this region could have different magnitudes (fluxes) and spacial distributions.

IV-C Fitting the Templates with the Data Regions

The final step to test the validity of the NNA technique involved fitting *all* the templates regions, $\mathcal{R}^{\text{temp}}(n)$, with their corresponding data regions, $\mathcal{R}^{\text{data}}(n)$. First, let's reiterate the physical meaning of $\mathcal{R}(n)$: a region \mathcal{R} corresponds to a 2-dimensional flux matrix. This flux matrix represents n stars being observed in the slitless mode. For the remainder of this section, it is assumed that $\mathcal{R} \equiv \mathcal{R}(n)$ is implicit, and hence, the notation will be simplified to just \mathcal{R} . In order to deduce the best template that depicts that data the following procedure was followed:

- 1) \mathcal{R} holds a flux value for every (x, y) pixel, and has a dimension of $(m \times k)$, where $m \in [d, 2d]^2$ and $k = \sigma_{\text{PSF}} \equiv 3$ pixels. The limits of m come from considering the extreme positions of the target star, T . The case where T is at the beginning of the data image results in \mathcal{R} to have the dimensions of $[d, \sigma_{\text{PSF}}]$. On the other hand, if T is positioned at $x_T > d$ pixels, for any y_T , \mathcal{R} has dimensions $[2d, \sigma_{\text{PSF}}]$.
- 2) The length of \mathcal{R} is reduced along the x -direction by binning together the fluxes. This can be done since

²Here, the dispersion range d is 727 pixels.

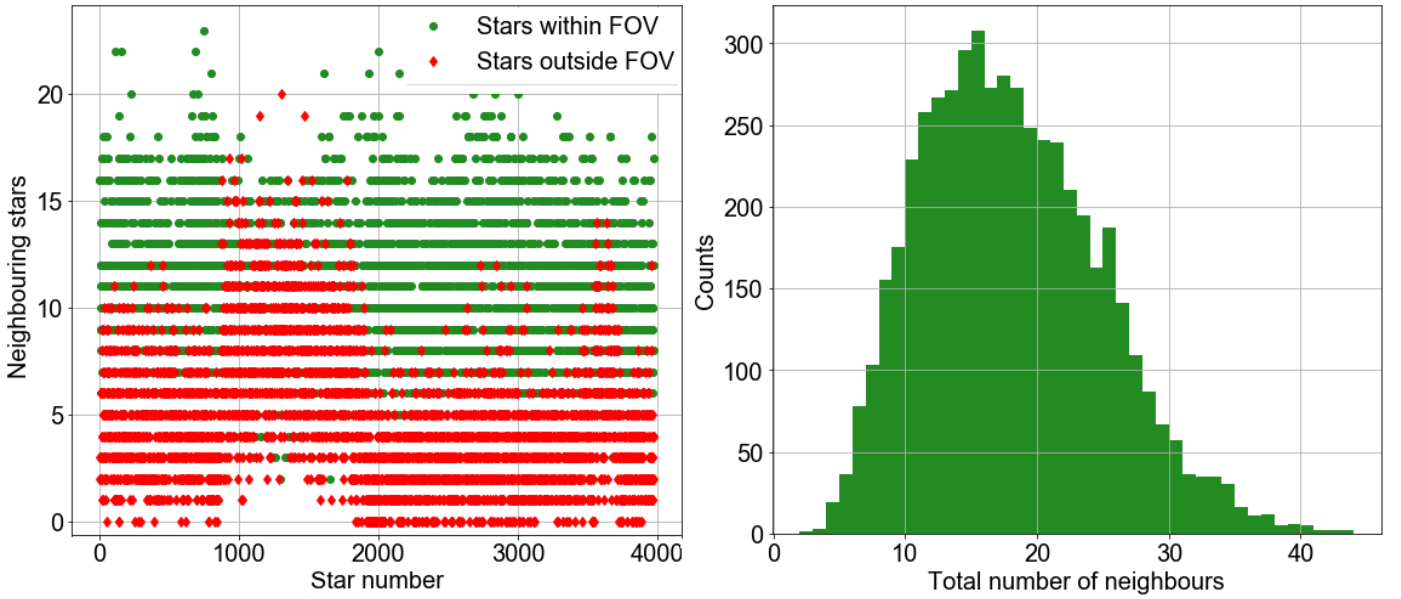


Fig. 11: Left: The number of neighbours for all the target stars T in the selected pointing (shown in Fig. 3; *left*) from the central catalogue. The neighbours within the FoV are shown in *green*, where the FoV is defined to be 1128×1128 pixel². However, the pointing is 1128×1855 pixel², resulting in additional contributions of stars outside the FoV (shown in *red*). **Right:** Histogram of the total number of neighbours of the stars in the selected pointing. A maximum at ~ 15 implies that most stars in the FoV have 15 neighbours, making NNA a feasible technique.

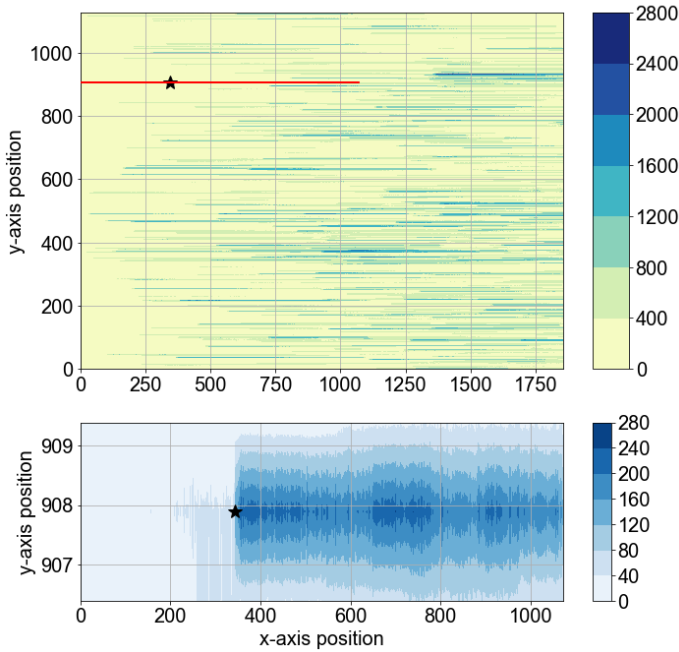


Fig. 12: Top: An eleven star region in the data image is highlighted. The target star, T , (in *black*) and its corresponding \mathcal{R} (in *red*) is zoomed in at the bottom. **Bottom:** Zoom in of an arbitrary $\mathcal{R}(n=11)$ case of the data image. The target star position is maintained for reference.

the spectra of the stars are damped sufficiently by the LSF along the x -direction. In order to bin the flux in a systematic way, we introduce the parameter called

pixel scale, Δ :

$$\Delta = m/\tilde{m}, \quad (9)$$

where m was the original length of \mathcal{R} and \tilde{m} is the reduced length. In this convention $\Delta \geq 1$, as the aim is to reduce the length of \mathcal{R} . Different values of Δ were considered during the fitting procedure, as detailed in Sec. VI. After reducing the length of \mathcal{R} , the new dimensions of \mathcal{R} becomes $(\tilde{m} \times k)$. The width of \mathcal{R} , which is given by k , remains unchanged as σ_{PSF} is typically of the $\mathcal{O}(1)$.

- 3) The last step of the fitting procedure involves defining a maximum likelihood estimator. Let's define $\mathcal{R} \rightarrow \mathcal{R}'$ after its length has been reduced by a factor $1/\Delta$. Correspondingly, the data and template regions are renamed as $\mathcal{R}'^{\text{temp}}$ and $\mathcal{R}'^{\text{data}}$, respectively. We can define the statistic, ω , where

$$\omega = \sum_{x=0}^{\tilde{m}} \sum_{y=0}^k [\mathcal{R}'^{\text{data}}(x \geq x_T, y) - \mathcal{R}'^{\text{temp}}(x \geq x_T, y)]^2. \quad (10)$$

In the convention used in this paper, \mathcal{R}' denotes a 2-dimensional matrix, however, $\mathcal{R}'(x, y)$ denotes a single-value, i.e. an element of the matrix \mathcal{R}' . An important point to note is that Eqn. 10 was minimized for all (x, y) to the right of the target star i.e. $x > x_T$. This corresponds to the *green* region in Fig. 9, where positions of the neighbours are given by \vec{x}_{k+} . This is because the aim of the NNA is to

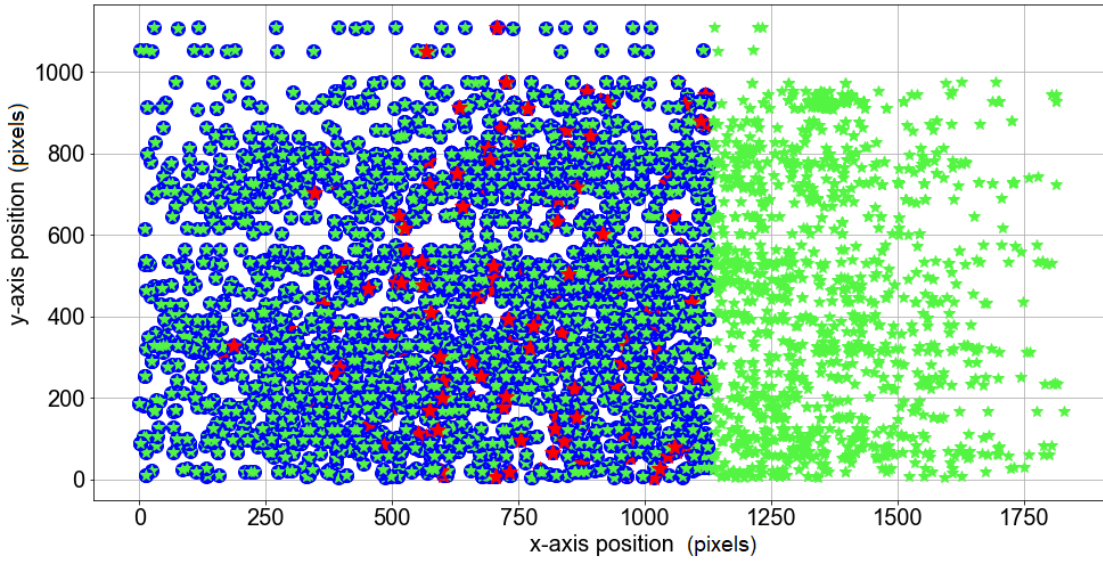


Fig. 13: All the 11 star regions, $\mathbf{R}_N(n=11)$; there are a total $N = 255$ regions in the chosen pointing. The targets stars T that each have 11 – 1 neighbours are highlighted in *red*, and their neighbours within the FoV (in *blue*) are distinguished from all the neighbours, which exist inside and outside the FoV (in *green*). All these regions are obtained by applying the NNA, shown in Fig. 9. This figure corresponds to the Step 1 of the methodology detailed in Sec. IV-B.

n	5	6	7	8	9	10	11*	12*
β	1	50	75	5	5	10	50	5
Δ	5	3	7	7	3	60	5	5
$A(n)$ (%)	70	73	69	65	73	60	54	48

TABLE II: Accuracy for regions containing different number stars, n , as a function of the optimization parameters: the number of templates, β , and the pixel scale, Δ . * The cases of $n = 11$ and 12 were varied for $\Delta = 3$ and 5 only, opposed to the other star cases where the following values were attempted: $\Delta = \{1, 3, 5, 7\}$.

decode the stellar information of T alone. In this case, the spectra of T extends to its right, and hence, only this region was minimized.

We obtain ω for every template that fitted to the data region. The total number of templates that are fit to $\mathcal{R}'^{\text{data}}$ is given by

$$\Gamma(n) = \sum_{h_*=0}^n P_{h_*}^n = \sum_{h_*=0}^n \frac{n!}{(n-h_*)!h_*!}, \quad (11)$$

where $P_{h_*}^n$ was introduced in Eqn. 8. Therefore, for fixed n , $\mathcal{R}'^{\text{data}}$ must be fitted with Γ number of template regions, $\mathcal{R}'^{\text{temp}}$. Each template region produces a value for ω , given by Eqn. 10. It is convenient to define the vector Ω that contains all the values of ω generated by the fitting procedure:

$$\Omega = [\omega_1, \omega_2, \dots, \omega_\Gamma]. \quad (12)$$

V Analysing the Best-Fitting Templates

This section aims to analyse the vector Ω (see Eqn. 10), which stores the information about the best-fitting templates. This fitting procedure was done for all the regions in the data image that contained stars in the range of $[5, 12]$. Fig. 14 shows one example of the output Ω obtained from fitting $\mathcal{R}'_0^{\text{data}}$ ³. In the case considered, $\mathcal{R}'_0^{\text{data}}$ contains 11 stars. The templates with the lowest value of Ω correspond to the templates that best represent the data region. By decoding the stellar type of the target star T in the best-fitting template(s), and comparing them to T in the data image, we can obtain a measure of *accuracy* of the NNA technique. There are two possible ways to proceed:

- 1) **Case I:** Choosing one template that corresponds to the lowest value in Ω , i.e. $\min\{\Omega\} = \omega_k^*$, where k is the template number. The corresponding template region is denoted as $\mathcal{R}'^{\text{temp}*}$.
- 2) **Case II:** Choosing a bunch of templates, say β , with the lowest values in Ω . This corresponds to

$$\min\{\Omega\}_\beta = \{\omega_{k_1}^*, \omega_{k_2}^*, \dots, \omega_{k_\beta}^*\}, \quad (13)$$

where k_1, k_2, \dots, k_β correspond to the template numbers that minimize $\mathcal{R}'^{\text{data}}$. Pictorially, this corresponds to defining a cut along the y -axis in Fig. 14.

The second case was preferred, because it minimizes the effect of the neighbouring stars in the analysis. However,

³Here the subscript of $\mathcal{R}'_0^{\text{data}}$ is 0. Here, 0 labels a region in the data image. Note that there are N total regions with n stars, so the $(N-1)^{\text{th}}$ region would be labelled as $\mathcal{R}'_{N-1}^{\text{data}}$

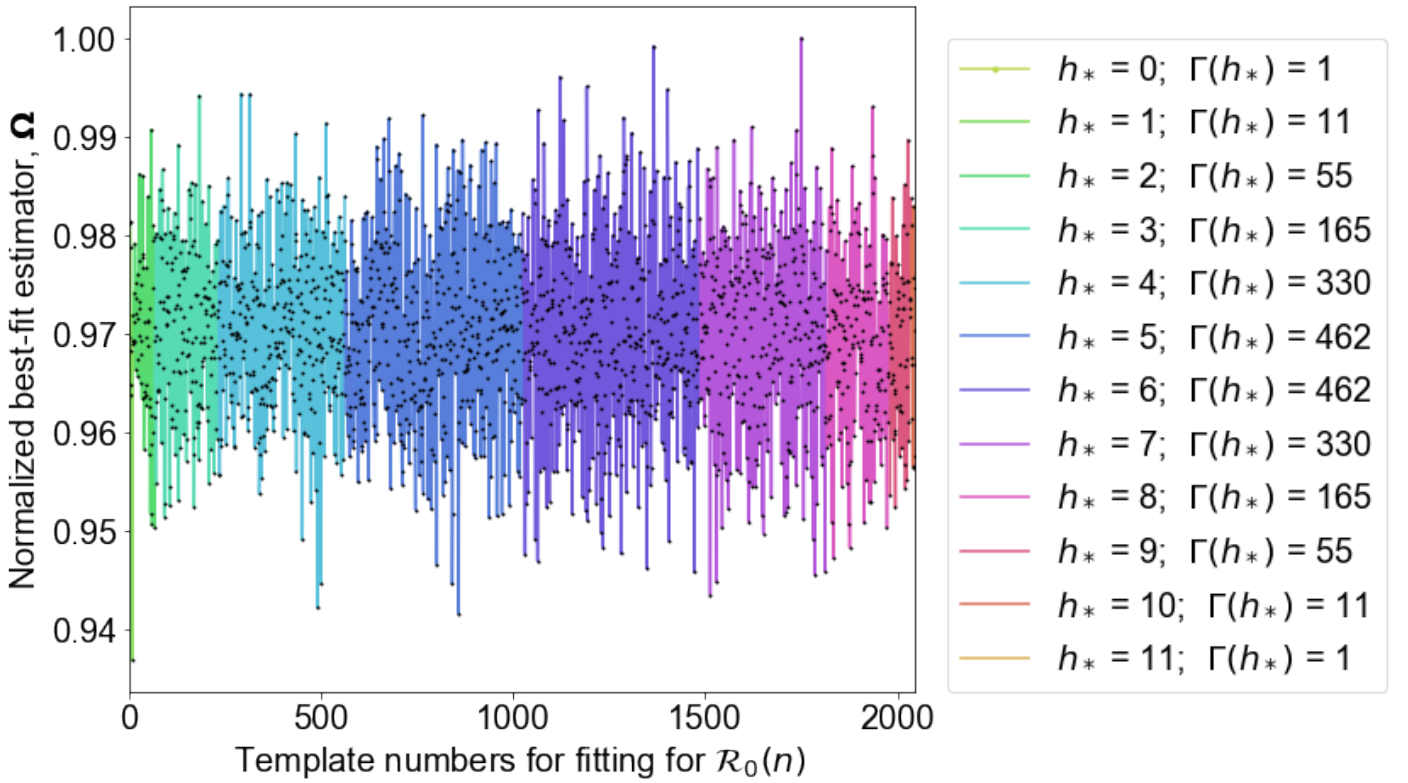


Fig. 14: The normalized best-fit estimator (defined in Eqn. 10), is plotted on the y -axis as a function of the template number. Each template number represents a templates region, $\mathcal{R}'^{\text{temp}}$, that was fitted to the data region, $\mathcal{R}_0'^{\text{data}}$ (refer to in Sec. IV-C for notation). The region \mathcal{R}' contains 11 stars and all possibilities for hot stars $h_* \in [0, n]$ were considered. For each specific case of h_* , there exist $\Gamma(h_*)$ templates; the different cases of h_* are represented with different colours, as shown in the legend. The sum of all these templates is $\sum_{h_*=0}^n \Gamma(h_*) = 2048$ for the considered case of 11 stars. Therefore, 2048 number of total templates that were fitted to $\mathcal{R}_0'^{\text{data}}$.

the results from the first case are also presented in Sec. VI-A. It must be noted that the ultimate goal is to decode the stellar type of the target star, and not the stellar types of all the stars in the region. So in principle, there could be multiple templates that closely depict $\mathcal{R}'^{\text{data}}$. In other words, the flux contributions from T and its neighbours could be added up in such a way that there are many $\mathcal{R}'^{\text{temp}*}$ that are a good fit to $\mathcal{R}'^{\text{data}}$. The only way to lift this degeneracy, or to minimize incorrect predictions, is to consider a sample of templates, β , that have the minimum values in Ω . If most templates in this sample predict that the target star is hot, it is less probable that the result is affected by the degeneracy in the templates. The next question to address is: *how many lowest templates β must be considered to estimate of the stellar types of the target star?* This answer is addressed in Fig. 15.

VI Optimizing the Accuracy of the Nearest Neighbour Approach

The technique for prediction and evaluation of the stellar types of the target stars was set up in the previous Sec. IV, and V. The final step is to optimize and recognize the effects of assumed parameters on the quality of the results.

The two main parameters of focus in this section will be: (a) The number of lowest ω templates, β , and (b) The pixel size, Δ . This effect of these two parameters are discussed in the following Sec. VI-A and Sec. VI-B respectively.

VI-A Effects on Accuracy due to Number of Templates

Fig. 15 shows the effect of the number of lowest ω templates, denoted as β , (Eqn. 13) on the accuracy of different cases of star regions \mathbf{R}_N . The following cases for β were considered, $\beta = \{1, 5, 10, 12, 25, 50, 75, 100\}$. Consider the case of $n = 11$, there are 255 \mathbf{R}_N in the data image. The accuracy, A , is obtained using the following formula:

$$A(n) = \frac{C \times 100}{\text{Total number of regions } (N)}, \quad (14)$$

where C is the number of regions with correct predictions for the target star. It is observed that for the low star cases, such as $n = 5$ and 6, low values of β give optimal results. This is understandable since for these cases, the total number of templates Γ that are being fit to a region

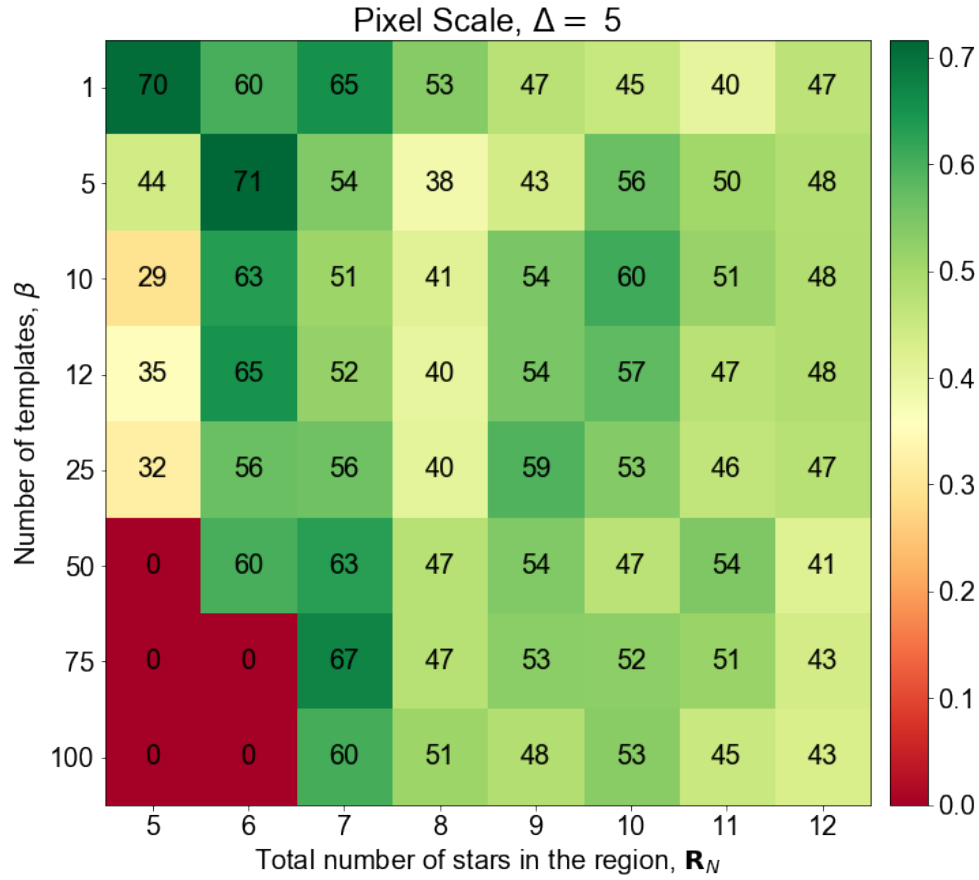


Fig. 15: Accuracy map showing the effect of number of templates, β (see Sec. V; Case II for further details) with maximum likelihood on the different star regions, \mathbf{R}_N . Here, N represents the total number of regions of n stars in the data image (Sec. IV; see Step 1). The number of lowest ω templates (see Eqn. 13) chosen to predict the stellar type of the target star affects the accuracy of the prediction. Cases using $n = 5$, and $n = 6$ stars show 0% accuracy for templates $\beta \geq 50$ and $\beta \geq 75$, respectively. This is because for such small n , these numbers of β imply an average over almost all the templates that are being fitted to the region; hence, they are not representative of the effectiveness of the fitting procedure.

are low. So having large sample sizes would result in considering majority of templates that are fit to the region. This explains the reason why for the case $\beta \geq 50$ for $n = 5$ and $\beta \geq 75$ for $n = 6$, the accuracy is 0%. As we go to higher star cases, there is no clear trend except for the $\beta = 1$ case, where the accuracy drops as n increases. Note that the $\beta = 1$ corresponds to Case I mentioned in Sec. V. An accuracy of $\sim 50\%$ is not optimal, especially as we approach the peak of Fig. 11 (right), because most of the regions in our data image contain ~ 15 neighbours. As will be discussed in the following Sec. VI-B, it was observed that the accuracy is also dependent on the pixel scale Δ that was chosen for the fitting procedure. Therefore, before making concluding remarks on the optimization procedure it is important to discuss the effects of Δ on the accuracy.

VI-B Effects on Accuracy due to Pixel Scale

The pixel scale, Δ , was introduced in Eqn. 9. This parameter was varied to see its effects on the accuracy of predicting the target stars. Fig. 16 summarizes the main

points that are discussed here. The following values of pixel scale were attempted, $\Delta = \{1, 3, 5, 7\}$, for the same range of β as mentioned in the previous Sec. VI-A. The effect of different values for Δ were studied for $n \in [5, 10]$. The accuracy for each case is calculated using the same formula given in Eqn. 14. It was observed that the accuracy was sensitive different values of Δ and the best values were obtained for $\Delta = 3, 5$.

VII Conclusions and Future Prospects

A pointing of $3.6' \times 6.0'$ of the Galactic Center region was chosen from the central catalogue covered by the GALACTICNUCLEUS survey. The 9472 stars in the pointing were randomly assigned one of the two stellar types: hot or cold; overall, 10% of the stars were assigned to be hot stars. The synthetic stellar spectra used to model images in the slitless mode were taken from the PHOENIX library. These spectra were binned together such that their resolution was equivalent to the dispersion range of the spectral axis in the slitless image. The instrumental details were chosen for

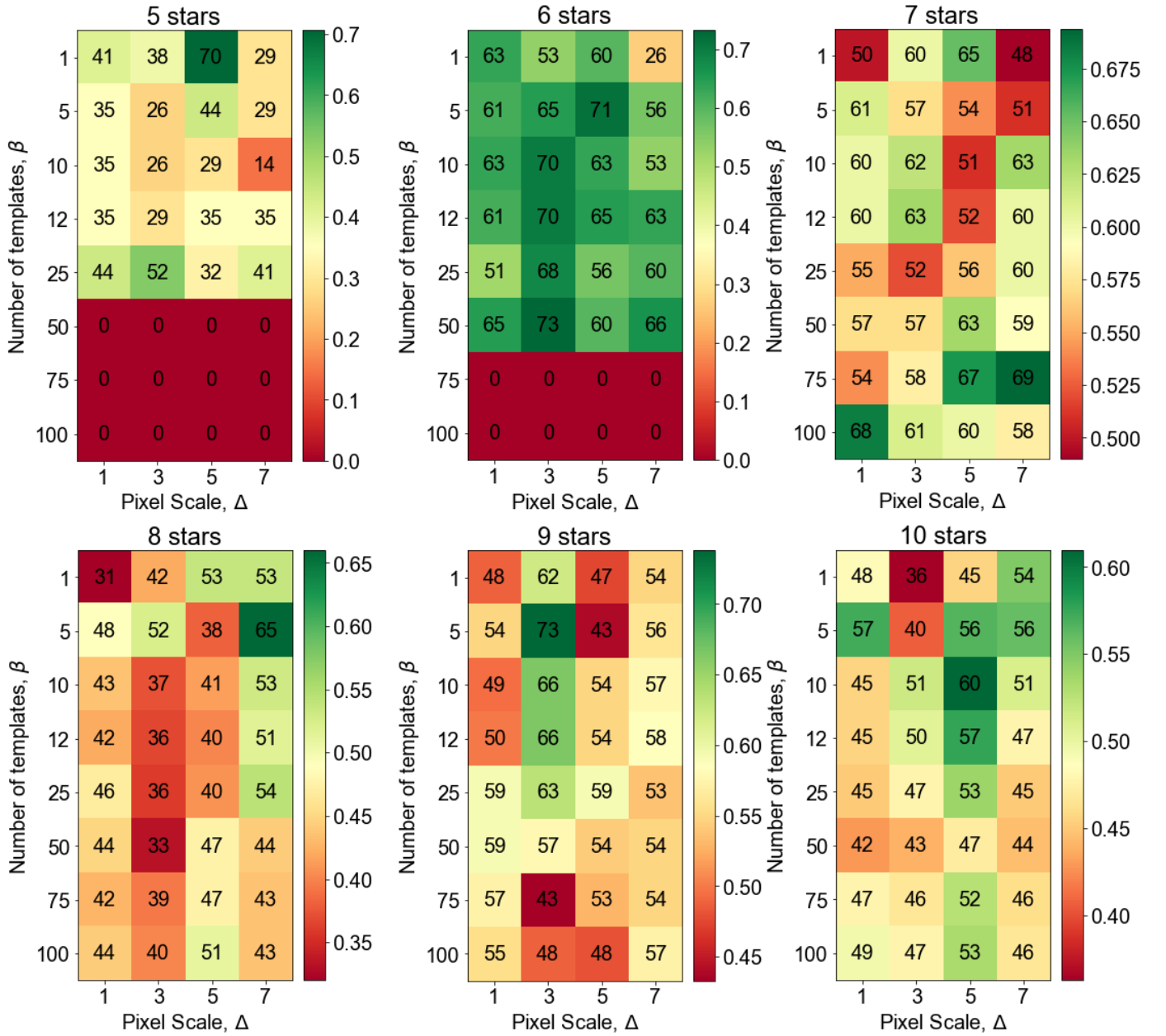


Fig. 16: Accuracy map showing the effect of number of templates and a function of the β pixel scale, Δ for different cases of star regions \mathbf{R}_N (refer to Eqn. 7, 9, and 13 for definitions). As in Fig. 15, the accuracy is plotted along the third-axis with its corresponding values for each case are shown in the colormap. Cases using $n = 5$, and $n = 6$ stars show 0% accuracy for templates $\beta \geq 50$ and $\beta \geq 75$, respectively. This is because for such small n , these numbers of β imply an average over almost all the templates that are being fit to the region; hence, they are not representative of the effectiveness of the fitting procedure. This plot is further discussed in Sec. VI-B.

observations with EMIR on GranTeCan. The modelling of the slitless image for the GC regions followed three steps: damping the spectra with a LSF, associating the spectra to the appropriate star, and adding flux to the image. The damping of the spectra implied the loss of velocity, and $\log g$ information in the stellar spectra; as a result, it was possible to extract just one free parameter from the slitless image, which is T_{eff} or the stellar type.

The nearest neighbour approach was formulated to tackle the problems posed by crowded FoV to observe in the slitless mode. This approach relied on defining a region of influence for every star, and analysing the stars within this region. The aim of such an approach was to extract stellar information star-by-star and span the entire FoV. For this purpose, the data image was divided into regions containing same number of neighbouring stars. The entire methodology for fitting regions, using this approach, was devised. The output from the fitting procedure were analysed.

The number of lowest ω templates, β , and the pixel scale, Δ , were recognized to be two parameters that affected the accuracy of the NNA approach. The fitting and analysis of the best-statistic templates were carried out for different values of β and Δ . By identifying the values that maximized the accuracy (to identify the stellar type of the target star), it was concluded that $\sim 60 - 75\%$ of all regions accurately predicted the stellar type of the target star, as shown in Tab. II. The accuracy of reproducing the target star stellar type varied based on the number of total stars in the region of influence. Additionally, the accuracy showed a strong dependence on the choice of β and Δ . The need to improve the NNA technique such that the accuracy is homogeneous over all cases of number of neighbours was identified. One additional parameters that could affect the accuracy is the level of Poisson noise added to the flux images. In order to test this hypothesis, the analysis must be performed again. The accuracy outputted from redoing the analysis must be compared with the current results. This will be one of the future goals of the project.

The future work on this project would involve the following:

- 1) Repeating the entire analysis with a different Poisson noise level for generation of the flux images. This would shed light on the effects that the noise levels have on the accuracy for extracting spectral information. Additionally, modifying the best-fit estimator such that the uncertainties in the pixels are taken into account.
- 2) Trying to model slitless images with different dispersive directions. This would aid in understanding the effects on the number of neighbours per star. The aim would be to see whether this improves the

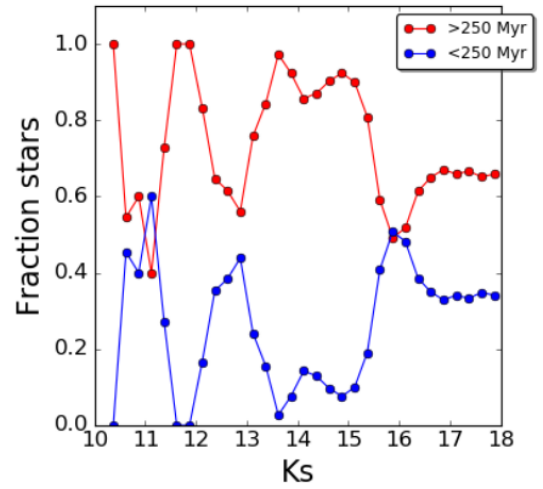


Fig. 17: Simulation of the age distribution of the stars for Ks magnitude bin according to the Star Formation History (SFH) derived in [Nogueras-Lara et al. \(2020\)](#) for the NB (excluding the NSC and Sgr A*). Old stars (< 250 Myr; in blue) show relatively few hot stars as opposed to young stars (> 250 Myr; in red), which show more number of hot stars. Figure credits to F. Nogueras-Lara.

results for stars that were not correctly identified in the analysis that was conducted in this project.

- 3) Recovering the original spectra by subtracting the LSF to determine velocity information about the observed stars. This would aid in the understanding of the proper motion of stars in the NB.
- 4) Model the foreground contribution and subtract it from the data image. This would simplify the modelling, as there would be few stars in the fitting procedure.
- 5) Modelling the background contribution. This corresponds to the faint stars that are unresolved, however, their dispersive contributions cannot be neglected.

The study aimed at extracting the stellar types of stars in the GC region of the Milky Way. This project was a pioneering study in setting up the basis for observing crowded FoVs in the slitless mode. The possibilities of such a study remain feasible; however, further improvements are required to this scheme to achieve higher accuracies in predicting stellar parameters of the GC stars. This technique will be useful to identify dissolved clusters by identifying regions that contain a higher-density of hot stars. Upon discovering such regions, follow-up observations with higher resolutions can be planned. The most important implication of this study would be in improving our current understanding on the Star Formation History (SFH) of the NB (excluding the NSC and Sgr A*). According to the SFH derived in [Nogueras-Lara et al. \(2020\)](#), a synthetic model was created

using 5 stellar populations (Parsec models, release v1.2S + COLIBRI S_35 + PR16; Bressan et al. (2012); Chen et al. (2014); Tang et al. (2014); Marigo et al. (2017); Pastorelli et al. (2019)), each accounting for 91%, 5%, 2%, 1%, and 1% of the stellar mass, respectively⁴. The results shown in Fig. 17 indicate that younger stars (< 250 Myr) must be hot-type stars opposed to older stars (> 250 Myr) that must be cold-type stars. The model predicts the fraction of hot stars in the NB, which is a measurable quantity in observations using the slitless mode. Therefore, this makes the study described in this paper a unique probe to test current predictions of SFH of the GC of the Milky Way. Deviations from current predictions would indicate the need to renew our current understanding on SF in the GC of the Milky Way.

All the code for this project is made available at the following GitHub repository:

https://github.com/SoumyaShreeram/Slitless_spectroscopy.

Acknowledgements: I would like to thank Nadine Neumayer for providing me this opportunity to work at MPIA. I have enjoyed all the enlightening discussions throughout the project with Paco and Nadine. I would like to thank both of them for guiding me through this project and creating a friendly and collaborative environment. I am grateful to have the opportunity to learn about the GCs of the Milky Way and other galaxies from the Galactic Nuclei group. I would also like to thank Hans-Walter Rix for his crucial ideas and for always bringing in the enthusiasm to the project. Last but not the least, I would like to thank Paco, Nils, Nadine, and my parents for their vital feedback on this paper.

This document contains 6664 words.

References

- Abuter, R., Amorim, A., Bauböck, M., et al. 2019, *Astronomy & Astrophysics*, 625, L10
- Bailer-Jones, C. A. 2000, arXiv preprint astro-ph/0003071
- Balcells, M. 1998, *Astrophysics and space science*, 263, 361
- Beck, T. L., Bary, J. S., & McGregor, P. J. 2010, *The Astrophysical Journal*, 722, 1360
- Bressan, A., Marigo, P., Girardi, L., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 427, 127
- Bruzual, G. & Charlot, S. 2003, *Monthly Notices of the Royal Astronomical Society*, 344, 1000
- Chen, Y., Girardi, L., Bressan, A., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 444, 2525
- Decin, L., Shkedy, Z., Molenberghs, G., Aerts, M., & Aerts, C. 2004, *Astronomy & Astrophysics*, 421, 281
- Do, T., Hees, A., Ghez, A., et al. 2019, *Science*, 365, 664
- Eikenberry, S. S., Elston, R., Raines, S. N., et al. 2004, in *Ground-based Instrumentation for Astronomy*, Vol. 5492, International Society for Optics and Photonics, 1196–1207
- ESO, c. & Nogueras-Lara et. al, F. 2018, ESO, <https://www.eso.org/public/images/eso1920a/>
- Feldmeier, A., Neumayer, N., Seth, A., et al. 2014, *Astronomy & Astrophysics*, 570, A2
- Feldmeier-Krause, A., Kerzendorf, W., Neumayer, N., et al. 2017, *Monthly Notices of the Royal Astronomical Society*, 464, 194
- Figer, D. F., McLean, I. S., & Morris, M. 1999, *The Astrophysical Journal*, 514, 202
- Fritz, T. K., Gillessen, S., Dodds-Eden, K., et al. 2011, *The Astrophysical Journal*, 737, 73
- Gai, M., Lattanzi, M., Guarnieri, M., Robberto, M., & Munari, U. 1997, in *Wide-field spectroscopy* (Springer), 121–122
- Garzón, F., Abreu, D., Barrera, S., et al. 2006, in *Ground-based and Airborne Instrumentation for Astronomy*, Vol. 6269, International Society for Optics and Photonics, 626918
- Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, *Reviews of Modern Physics*, 82, 3121
- Georgiev, I. Y. & Böker, T. 2014, *Monthly Notices of the Royal Astronomical Society*, 441, 3570
- Ghez, A. M., Salim, S., Weinberg, N., et al. 2008, *The Astrophysical Journal*, 689, 1044
- Gillessen, S., Eisenhauer, F., Fritz, T., et al. 2009, *The Astrophysical Journal Letters*, 707, L114
- Gürkan, M. A., Freitag, M., & Rasio, F. A. 2004, *The Astrophysical Journal*, 604, 632
- Harfst, S., Portegies Zwart, S., & Stolte, A. 2010, *Monthly Notices of the Royal Astronomical Society*, 409, 628
- Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, *Astronomy & Astrophysics*, 553, A6
- Kissler-Patig, M., Pirard, J.-F., Casali, M., et al. 2008, *Astronomy & Astrophysics*, 491, 941
- Kümmel, M., Walsh, J., Pirzkal, N., Kuntschner, H., & Pasquali, A. 2009, *Publications of the Astronomical Society of the Pacific*, 121, 59
- Launhardt, R., Zylka, R., & Mezger, P. 2002, *Astronomy & Astrophysics*, 384, 112
- Liermann, A., Hamann, W.-R., & Oskinova, L. M. 2012, *Astronomy & Astrophysics*, 540, A14
- Marigo, P., Girardi, L., Bressan, A., et al. 2017, *The Astrophysical Journal*, 835, 77
- Matsunaga, N., Kawadu, T., Nishiyama, S., et al. 2011, *Nature*, 477, 188
- McKee, C. F. & Ostriker, E. C. 2007, *Annu. Rev. Astron. Astrophys.*, 45, 565
- Minniti, D., Ramos, R. C., Zoccali, M., et al. 2016, *The Astrophysical Journal Letters*, 830, L14
- Morgan, W. W. & Keenan, P. 1973, *Annual Review of Astronomy and Astrophysics*, 11, 29
- Morris, M. & Serabyn, E. 1996, *Annual Review of Astronomy and Astrophysics*, 34, 645
- Najarro, F., Figer, D. F., Hillier, D. J., & Kudritzki, R. P. 2004, *The Astrophysical Journal Letters*, 611, L105

⁴Extinction curve from Nogueras-Lara et al. (2020) was assumed; the photometry was reddened accordingly. Lastly, Gaussian uncertainties in agreement with the results from the (Nogueras-Lara et al., 2019) were used.

- Neumayer, N., Seth, A., & Böker, T. 2020, *The Astronomy and Astrophysics Review*, 28, 1
- Nishiyama, S., Nagata, T., Tamura, M., et al. 2008, *The Astrophysical Journal*, 680, 1174
- Nogueras-Lara, F., Gallego-Calvente, A., Dong, H., et al. 2018a, *Astronomy & Astrophysics*, 610, A83
- Nogueras-Lara, F., Schödel, R., Dong, H., et al. 2018b, *Astronomy & Astrophysics*, 620, A83
- Nogueras-Lara, F., Schödel, R., Gallego-Calvente, A. T., et al. 2019, *Astronomy & Astrophysics*, 631, A20
- Nogueras-Lara, F., Schödel, R., Gallego-Calvente, A. T., et al. 2020, *Nature Astronomy*, 4, 377
- Pasquali, A., Pirzkal, N., Larsen, S., Walsh, J., & Kümmel, M. 2006, *Publications of the Astronomical Society of the Pacific*, 118, 270
- Pastorelli, G., Marigo, P., Girardi, L., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 485, 5666
- Pirzkal, N., Pasquali, A., Walsh, J., et al. 2001, *ST-ECF ISR ACS*, 1
- Rui, N. Z., Hosek Jr, M. W., Lu, J. R., et al. 2019, *The Astrophysical Journal*, 877, 37
- Ryan Jr, R., Casertano, S., & Pirzkal, N. 2018, *Publications of the Astronomical Society of the Pacific*, 130, 034501
- Ryde, N., Schultheis, M., Grieco, V., et al. 2015, *The Astronomical Journal*, 151, 1
- Schödel, R., Feldmeier, A., Kunneriath, D., et al. 2014, *Astronomy & Astrophysics*, 566, A47
- Schultheis, M., Rich, R., Origlia, L., et al. 2019, *Astronomy & Astrophysics*, 627, A152
- Shkedy, Z., Decin, L., Molenberghs, G., & Aerts, C. 2007, *Monthly Notices of the Royal Astronomical Society*, 377, 120
- Tang, J., Bressan, A., Rosenfield, P., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 445, 4287
- Vallee, J. P. 2002, *The Astrophysical Journal*, 566, 261
- Vallée, J. P. 2005, *The Astronomical Journal*, 130, 569
- Walsh, J., Kümmel, M., & Kuntschner, H. 2010, *Hubble after SM4. Preparing JWST*, 8
- Zwart, S. F. P., Baumgardt, H., Hut, P., Makino, J., & McMillan, S. L. 2004, *Nature*, 428, 724