

# New insights into the population of young, massive stars near Sagittarius A\*

Gallego-Cano, E.<sup>1</sup> and Fritz, T.<sup>2</sup>

<sup>1</sup> Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía s/n, 18008 Granada, Spain

<sup>2</sup> School of Data Science, University of Virginia, 1919 Ivy Road, Charlottesville, VA 22903, USA

## Abstract

Several studies have revealed the presence of young massive stars in the central 0.5 pc of the Galactic Centre, prompting questions about their properties and formation in the immediate vicinity of the massive black hole Sagittarius A\*. Intriguingly, the shape of the initial mass function (IMF) in this region deviates from the standard Salpeter/Kroupa law. To enhance our understanding of the number and distribution of young, massive stars, we employ intermediate band (IB) near-infrared imaging. In agreement with previous studies, but going significantly beyond them, we find a core-like distribution of late-type stars. In contrast, the density of the early-type stars increases steeply towards the black hole. The analysis reveals a top-heavy IMF for young stars within about 0.4 pc of Sagittarius A\*. A standard IMF may be adequate at greater distances, suggesting that various star formation mechanisms are involved. Additionally, we demonstrate how IB photometry can be used to determine metallicities and temperatures, with around 6% of the late-type stars in our sample exhibiting metal-poor characteristics.

## 1 Introduction

The Galactic Centre (GC), located about 8 kpc from Earth, hosts the supermassive black hole Sagittarius A\* (Sgr A\*) with a mass of  $4.04 \pm 0.06 \times 10^6 M_{\odot}$  [1], surrounded by a nuclear star cluster (NSC) of  $2.5 \times 10^7 M_{\odot}$  and a half-light radius of  $\sim 4.2 - 7$  pc [16, 6]. The proximity of the GC allows for the resolution of stars on milliparsec scales, offering a unique opportunity to study star formation and stellar interactions with a massive black hole.

Recent studies have revealed young, massive stars (2 – 8 Myr) within 0.5 pc of Sgr A\*, including a population of B stars in the S-cluster [9] and two rotating disks: a clockwise (CW)

disk and a debated counterclockwise (CCW) disk [15, 2]. This discovery is unexpected, as star formation is typically thought to be suppressed in the vicinity of a massive black hole due to its strong tidal field. Additionally, discrepancies exist regarding the initial mass function (IMF) of these stars, with reported slopes varying significantly, from a top-heavy value of  $-0.45$  reported by [3] to a steeper value of  $-1.7$  reported by [12]. Recent work by [17] confirms the CW disk and identifies additional kinematic features, providing further insights into the star formation history of the region. Nevertheless, the IMF remains a subject of ongoing debate, which is essential for understanding the formation mechanisms of young stars near Sgr A\*.

Despite progress, studying the stellar population near Sgr A\* is challenging due to extreme extinction and crowding. Spectroscopic classifications using instruments like SINFONI/ERIS and OSIRIS are limited to bright stars ( $K_S < 15$  mag). Intermediate-band (IB) photometry, as shown by previous studies [14], offers a promising alternative for identifying fainter stars over larger fields, though follow-up spectroscopy remains crucial.

We have recently undertaken a new study using seven intermediate bands in the near-infrared (NIR) obtained with NACO at the ESO-VLT, reanalyzing data from [4] with improved techniques. Our study expands the analysis region, improves image reduction to obtain deeper data, and enhances photometry. We also apply advanced classification methods, including Bayesian inference, a basic neural network, and a gradient-boosted tree. By leveraging hundreds of classified stars, our approach enhances precision and accuracy in stellar classification near Sgr A\*. This methodology allows us to identify new candidates for early stars in this region. With the new data, we investigate the surface-density profile, offering insights into the dynamical state of the cluster and examining the luminosity function, a fundamental parameter for determining properties such as age and star formation history (Sect. 3). Furthermore, in Sect. 4, we constrain the metallicities of late-type stars, estimating the metallicities for over 600 of these stars. We also compute the IMF of the cluster, as detailed in Sect. 5.

## 2 Data and methodology

The data were obtained using seven IB filters, each with a wavelength width of  $0.06 \mu\text{m}$  in the  $K_S$ -band. The observations were conducted with the S27 camera on NACO at the VLT, with a pixel scale of  $0.027''$ . The reduction data and analysis are detailed in [7]. As explained in the study, we applied spatially variable point spread function (PSF), which led to deeper data and improved photometry compared to previous work. Additionally, we determined robust photometric uncertainties through a bootstrapping procedure and by accounting for PSF uncertainties. Following a basic calibration, a critical aspect of our work was the implementation of local photometric calibration across the field of view (FoV) to account for spatial variations in Adaptive Optics (AO) performance, number of exposures, and differential extinction using Red Clump (RC) stars as references. This approach homogenized the calibration quality across the field, ensuring consistency in the final star list for all bands.

With the seven filters, we obtained low-resolution spectral energy distributions (SEDs)

for the stars. To classify these stars, we take advantage of the characteristic CO bandhead absorption that late-type stars exhibit at wavelengths  $\lambda > 2.27 \mu\text{m}$  (see the right panel in Fig. 1). By analyzing the observed and calibrated SEDs for this feature, we can categorize stars as late-type candidates (spectral types  $\sim\text{GKM}$ ) when the feature is present, or as early-type candidates (spectral types  $\sim\text{OB}$ ) when it is absent. The first four IB filters ( $\lambda = 2.00 - 2.27 \mu\text{m}$ ) are used to estimate the continuum. To measure the CO band depth (CBD), we employed SED fitting methods: a linear fit to the first four data points and an exponential fit to all seven data points, as illustrated in Fig. 1.

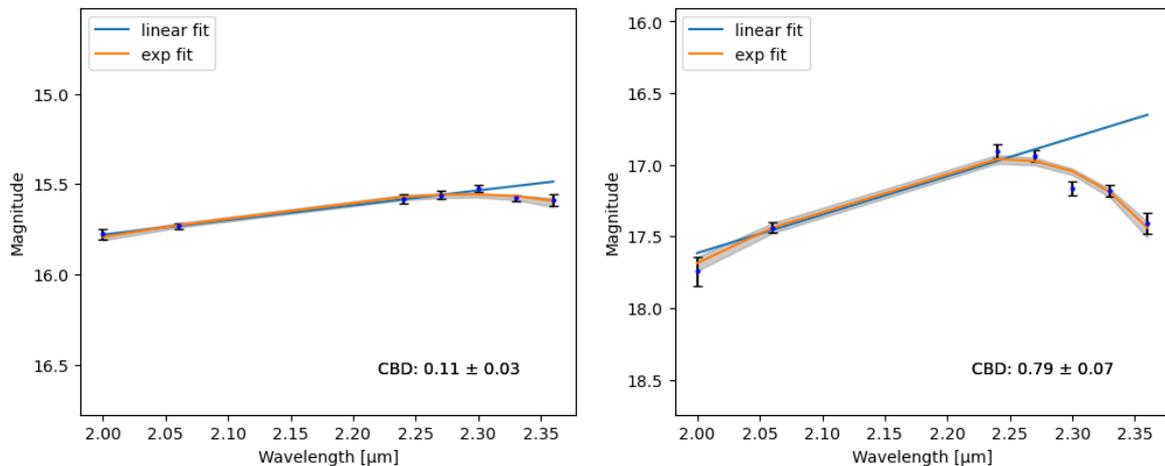


Figure 1: SEDs of two spectroscopically classified stars. The left panel displays an early-type star, and the right panel shows a late-type star. The blue lines denote the linear fit to the first four data points, while the orange lines represent the exponential fit to all seven data points. The grey shadows indicate the uncertainties of the exponential fits, as obtained from Monte Carlo simulations. The panels include notations for the CBDs for the stars.

We obtained low-resolution SEDs for the 6590 detected stars across all IB filters, 1181 of which have been previously classified spectroscopically (982 late-type stars and 212 early-type stars). The left panel in Fig. 2 shows the computed CBD values for the entire sample, with the y-axis representing the magnitude in the IB filter centred on  $\lambda = 2.24 \mu\text{m}$  (IB224). Typical errors for CBD values are displayed in bins of 0.5 magnitudes.

To ensure robust classification, we employed three complementary methods to distinguish between early- and late-type stars: Bayesian inference, a Multi-Layer Perceptron (MLP) neural network for supervised learning, and Gradient Boosted Trees (GBT) using the XGBoost algorithm.

Bayesian analysis offers interpretability and a probabilistic framework for handling uncertainty, which helps make more informed decisions based on the available data. However, it heavily relies on a well-informed prior, which can introduce bias if it does not accurately represent the true distribution of star types. On the other hand, machine learning methods, such as the neural network and gradient-boosted trees, can learn complex patterns in the data, though they can be prone to overfitting. Further details of the procedure can be found

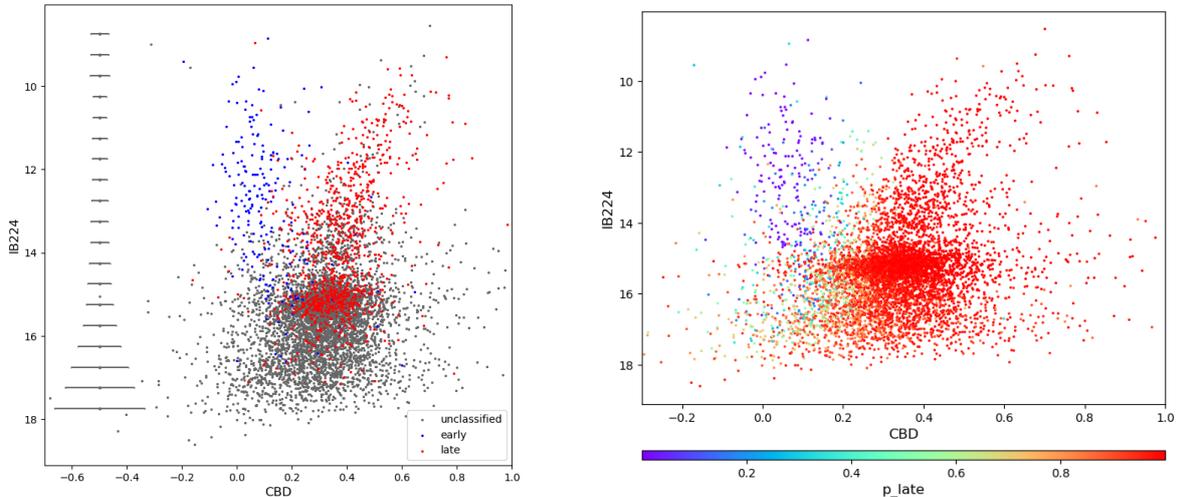


Figure 2: The left panel displays the CBD diagram for the entire sample, highlighting spectroscopically classified stars, with early-type stars marked in blue and late-type stars in red. The right panel shows the final CBD diagram, including the posterior distribution from our Bayesian model. In both diagrams, the Red Giant Branch is distinguishable for stars brighter than  $\sim 14$  mag and CBD values greater than 0.2. Additionally, the RC appears as a concentration of stars with magnitudes between  $\sim 15$  and 16 mag.

in [7]. The right panel in Fig. 2 displays the posterior (the product of the prior and the likelihood) of our Bayesian model.

The number of new early-type candidates detected by the three methods is shown in Table 1. We defined a candidate early-type star as one with an unknown spectral type and a probability of being early larger than 0.5. The *Common* row includes the early-type candidates identified by all three methods. The magnitudes  $m$  are extinction-corrected IB224 magnitudes.  $R$  is the projected distance from Sgr A\*. We identified 155 new early-type candidates and 4806 late-type candidates, all consistently identified by the three methods and previously lacking spectroscopic identification.

Table 1: Number of new early-type candidates.

Method	all	$m < 16$	$R < 14''$	$m < 16 \ \& \ R < 14''$
MLP	498	229	351	148
XGBoost	311	141	294	124
Bayesian	204	134	185	116
Common	155	98	145	88

### 3 Structure of the cluster

For our study on the cluster structure, we use probabilities directly instead of relying on candidates identified by a probability cut. For the analysis of the radial profile, we excluded all stars fainter than 16 magnitudes and those located in regions where only two observed frames are available for a given epoch (specifically, the four corners). Consistent with previous studies, we find that broken power laws provide adequate fits to the radial density distribution of both late-type and early-type stars (see the left panel in Fig. 3). Our investigation reveals a break radius of  $7.4 \pm 1.2''$  for the late-type distribution and  $9.2 \pm 0.6''$  for the early-type distribution. The slopes for the late-type distribution are determined to be  $\beta = 0.46 \pm 0.07$  (outside the break radius) and  $\gamma = 0.00 \pm 0.08$ , while the early-type distribution exhibits slopes of  $\beta = 3.65 \pm 0.47$  and  $\gamma = 0.81 \pm 0.08$ . The results are obtained using the Bayesian approach, although the other methods yield very similar results, except for the outer slope of early-type stars, where the methods differ.

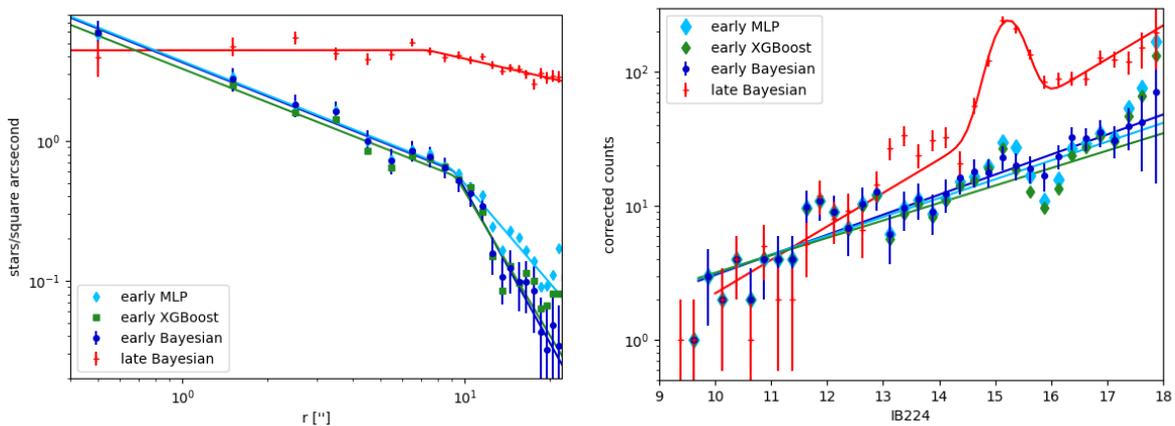


Figure 3: The left panel displays the radial distributions obtained using different methods, summing up probabilities. The right panel presents the completeness-corrected luminosity function and the corresponding fits for the inner region ( $R \leq 9''$ ).

In order to compute the  $K_S$  luminosity function (KLF), we sum the probabilities in magnitude bins (bin size = 0.25 mag). The data become increasingly incomplete at  $m \geq 16.5$  mag. To estimate and correct for completeness, we utilize the KLF from [8] (see the right panel in Fig. 3). We fit star counts down to the limit  $m = 18.5$  mag, omitting the brightest bins, with a power-law function for early-type stars and a power-law plus Gaussian function (to account for the RC) for late-type stars. We determine a slope parameter for the KLF of late-type stars, obtaining a value of 0.25. This parameter exhibits only minor variations across different classification methods and radial ranges. For early-type stars, we observe a shallower slope compared to late-type stars, with a calculated value of 0.18 using the Bayesian method. The obtained values for the inner region ( $R < 9''$ ) and the entire range are very consistent across the three methods. However, variations become more apparent at larger distances, where the MLP method yields a steeper slope.

## 4 Temperature and metallicity of late-type stars

Both temperature and metallicity significantly influence the strength of CO bandhead absorption, suggesting that IB photometry could be a valuable tool for measuring these stellar properties. We initiated our analysis by focusing on the effective temperature of late-type stars, which is more directly correlated with the CO bandhead than metallicity. Our comparison with spectroscopic temperatures indicates that the CBD primarily exhibits a linear relationship with temperature. However, metallicity also plays a notable role, particularly for stars with  $[M/H] < 0$  (see more details in Section 7.4 of [7]). When predicting metallicities for over 600 stars, we find that approximately 6% show metal-poor characteristics ( $[M/H] < -0.5$ ), a result that aligns well with spectroscopic metallicity measurements.

## 5 IMF of early-type stars

To infer the IMF of the young cluster formed by early-type candidates brighter than 16 mag and located within a distance smaller than the computed break radius of  $9''$ , we employed an optimization algorithm to compare observational data with theoretical simulations. Using SPISEA [10], we constructed a grid of synthetic single-age star clusters varying in age, cluster mass, extinction, and IMF slope index. The mass limits for stars in the simulated clusters were set to  $5 M_{\odot} - 120 M_{\odot}$  due to the absence of lower-mass stars in our observed magnitude range. We utilized the MIST (MESA Isochrones & Stellar Tracks; [5]) evolution model and adopted the reddening law from [13]. Isochrones spanning cluster ages from 2 to 8 Myr were examined, and synthetic KLFs were constructed for each parameter combination. We then compared the theoretical KLFs to the observed KLF by computing the chi-squared statistic and identified the parameter set that minimizes this value. The optimal fit corresponds to a cluster age of 4 Myr. To incorporate uncertainties arising from the sampling of the IMF in each cluster realization, we used a Monte Carlo approach. The best power-law IMF index value is  $1.6 \pm 0.1$ , obtained for the three sets of early-type candidates using different methods (see Fig.4).

## 6 Conclusions

In this work, we demonstrate that IB photometry is a powerful tool for classifying stars, particularly within the challenging environment of the central parsec around Sgr A\*. Three complementary methods—Bayesian inference, a neural network, and a gradient-boosted trees algorithm—were applied to distinguish between early and late-type stars. This approach led to the identification of 155 new early-type star candidates, underscoring the effectiveness of IB photometry in selecting promising targets for future spectroscopic follow-up.

With the new information we provide a new insight into the structure of the cluster, and we find that the IMF of early-type stars near Sgr A\* is top-heavy, with a power-law slope of 1.6. At larger distances, the mass function aligns more closely with the standard Salpeter/Kroupa IMF. This suggests distinct star formation mechanisms between the inner and outer regions.

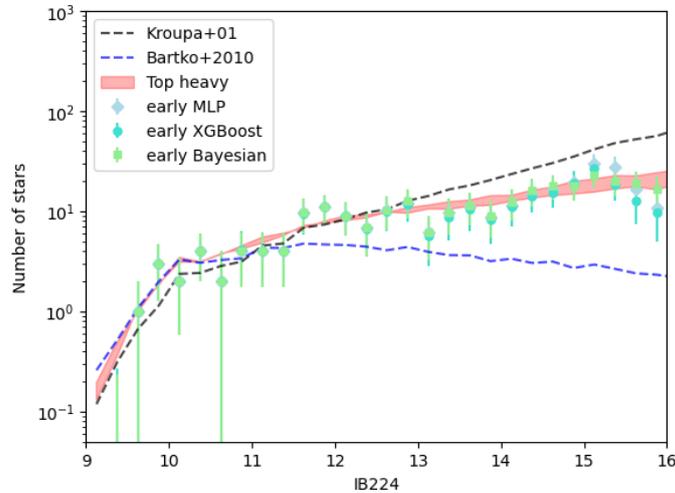


Figure 4: Comparison of different models simulated with SPISEA and the observed KLF, corrected for completeness. The red line represents the KLF for a simulated cluster with the optimal power-law IMF (slope = -1.6), while the shaded region indicates model uncertainty. For comparison, KLFs using the Kroupa IMF ([11]) and a top-heavy IMF ([3]) are also shown.

Beyond classification, IB photometry enables the exploration of stellar properties such as metallicity and temperature, providing valuable insights into the formation history of the NSC. This work highlights the importance of IB photometry not only for current studies but also for the design of future telescope filters.

## Acknowledgments

This work is based on observations made with ESO Telescopes at the Paranal Observatory under programs 073.B-0084(A), 073.B-0745(A), and 077.B-0014(A). EGC acknowledges financial support from the Severo Ochoa grant CEX2021-001131-S funded by MCIN/AEI/10.13039/501100011033, the grant EUR2022-134031 funded by MCIN/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRT, and the grant PID2022-136640NB-C21 funded by MCIN/AEI/10.13039/501100011033 and the European Union.

## References

- [1] Abuter, R., Amorim, A., Bauböck, M., et al. 2021, *A&A*, 647, A59
- [2] Bartko, H., Martins, F., Fritz, T. K., et al. 2009, *ApJ*, 697, 1741
- [3] Bartko, H., Martins, F., Trippe, S., et al. 2010, *ApJ*, 708, 834
- [4] Buchholz, R. M., Schödel, R., & Eckart, A. 2009, *A&A*, 499, 483
- [5] Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823, 102
- [6] Gallego-Cano, E., Schödel, R., Nogueras-Lara, F., et al. 2020, *A&A*, 634, A71

- [7] Gallego-Cano E., Fritz T., Schödel R., et al. 2024, A&A 689 A190
- [8] Gallego-Cano E., Schödel R., Dong, H., et al. 2018, A&A, 609, A26
- [9] Ghez, A., Duchêne, G., Matthews, K., et al. 2003, ApJ, 586, L127
- [10] Hosek, M. W., Lu, J. R., Lam, C. Y., et al. 2020, AJ, 160, 143
- [11] Kroupa, P. 2001, MNRAS, 322, 231
- [12] Lu, J. R., Do, T., Ghez, A. M., et al. 2013, ApJ, 764, 155
- [13] Nishiyama, S., Tamura, M., Hatano, H., et al. 2009, ApJ, 696, 1407
- [14] Nishiyama, S., Funamoto, N., & Schödel, R. 2023, ApJ, 951, 148
- [15] Paumard, T., Genzel, R., Martins, F., et al. 2006, ApJ, 643, 1011
- [16] Schödel, R., Feldmeier, A., Kunneriath, D., et al. 2014, A&A, 566, A47
- [17] von Fellenberg, S. D., Gillessen, S., Stadler, J., et al. 2022, ApJ, 932, L6