Highlights of Spanish Astrophysics XII, Proceedings of the XVI Scientific Meeting of the Spanish Astronomical Society held on July 15 - 19, 2024, in Granada, Spain. M. Manteiga, F. González Galindo, A. Labiano Ortega, M. Martínez González, N. Rea, M. Romero Gómez, A. Ulla Miguel, G. Yepes, C. Rodríguez López, A. Gómez García and C. Dafonte (eds.), 2025

CATARSIS optimization: dithering strategies and source scheduling with deep reinforcement learning.

Montenegro-Montes, F.M.[1], Gil de Paz, A.[1], Sánchez-Blázquez, P.[1], Pérez-Calpena, A.[2] and the TARSIS Consortium, instrument and CATARSIS teams

[1] Departamento de Física de la Tierra y Astrofísica e Instituto de Física de Partículas y del Cosmos (IPARCOS). Universidad Complutense de Madrid (UCM)
[2] FRACTAL S.L.N.E. (Spain)

Abstract

CATARSIS is a deep unbiased 2D spectroscopic survey of galaxy clusters and constitutes the scientific project driving the construction of the TARSIS instrument, soon to be built and installed at the 3.5m telescope of the Calar Alto Observatory. It is a long-term program that requires about 700 pointings to be observed over the course of 6 years and a robust survey strategy is essential to maximize its scientific impact. Therefore, in parallel to the instrument design phase, we are working on the survey optimization strategy. We have developed a prototype TARSIS footprint simulator that includes the expected vignetting patterns and allows for field-of-view rotation. We have also proposed various dithering strategies to achieve a homogeneous coverage per cluster, suitable for CATARSIS but also for other open-time projects. Finally, we are prototyping an automated scheduling system based on deep reinforcement learning that aims at enhancing the survey execution efficiency.

1 The TARSIS instrument

The TARSIS (Tetra-Armed Super-Ifu Spectrograph) instrument has been selected in a competitive process to be the next generation instrument for the 3.5m telescope at the Calar Alto Observatory (CAHA) in Almería, Spain. TARSIS features in a compact configuration 4 integral field units (IFUs) based on image slicers, producing simultaneously moderate-resolution spectra (R $\sim$ 1000) of a large area on the sky of 2.8 x 2.8 arcmin^2. Three IFUs are optimized for the blue spectral range, and the remaining one to the red, with a global spectral coverage of the spectrographs in the range 320-810 nm. The spaxel size is about 2$\times$2 arcsec^2. All technical details about the TARSIS design and capabilities can be found in [3].

The right panel of Fig 1 shows the instrument footprint on sky, showing the 4 quadrants arranged in a 2x2 matrix, spaced with each other by a small gap of the size of 2 spaxels. There are 40 slices per quadrant leading to 2 pseudo-slits to feed each spectrograph and, as shown in the figure, the slices in each quadrant are orthogonal to those in the adjacent quadrant. The image slicer introduces a small amount of vignetting or shadowing between slices (of the order of 4%, on average), which leads to a slightly inhomogeneous sensitivity pattern, as can be seen in the left panel of Fig 1.

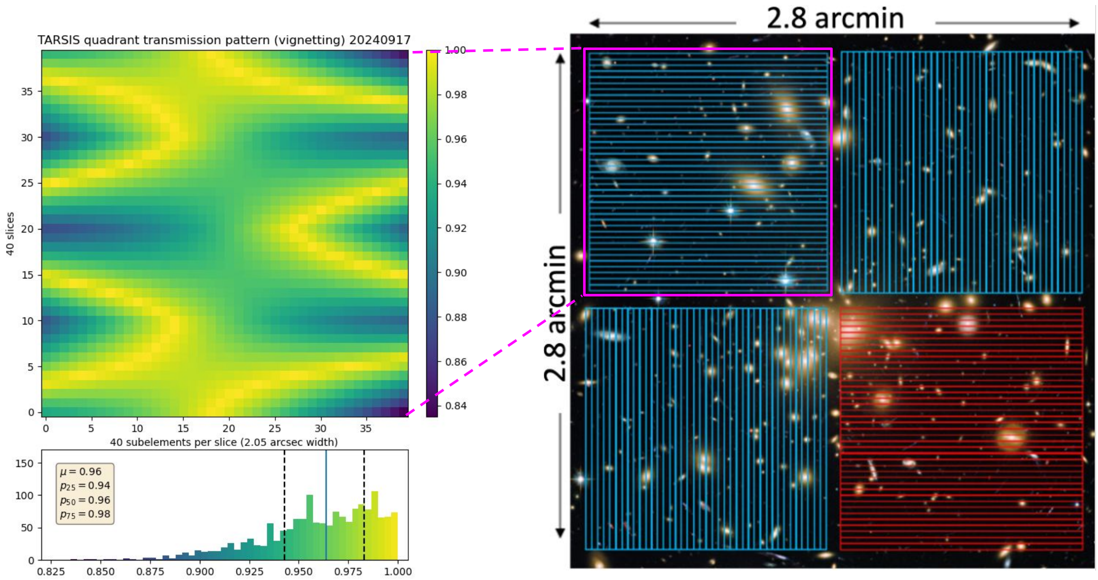

Figure 1: *Left*: Transmission pattern (per spaxel) due to vignetting in each of the quadrants of TARSIS. Below is shown the corresponding histogram highlighting the median and quartiles in the distribution. *Right:* On-sky footprint of TARSIS, with its 4 quadrants (3 blue-optimized and 1 red-optimized).

2 The CATARSIS survey

The Calar Alto Tetra-Armed Super-IFu spectrograph Survey, CATARSIS, is a self-contained survey aiming to map 16 galaxy clusters at $0.15 < z < 0.3$ from the cores to the outskirts, reaching the infall radius and the surrounding filaments. The main goal is to understand the formation of structures and the evolution of galaxies in a dynamical and growing environments as well as testing the standard cosmological models. CATARSIS will address questions that go from the nature of dark matter and energy, to how the galaxy evolution is affected by the surrounding environment. To achieve this, the star formation histories and chemical evolution connecting dark matter, stars, warm and hot gas will be studied, trying to get a coherent picture of how they evolve, together, and how are the interactions among them. Additional details about the goals and the structure of the survey can be found in [4] and [5].

3 Survey management

CATARSIS will be a legacy survey taking ∼6 years to complete, and a careful management becomes crucial to maximize its scientific impact. In this respect, the consortium is working on several areas since the feasibility study of the project, such as: the clusters sample selection and the collection of ancillary data, the design of an efficient observing strategy, the study of the calibration requirements, the strategy for data reduction and analysis, or the data archiving and delivery aspects. Some of these points will be discussed in more detail below.

3.1 Observing strategy

CATARSIS observations will cover a total area of 1.7 deg^2 distributed over the full LST range to ensure observability throughout the year. The survey requirements include reaching $m_{AB,r} < 22.0$ and limiting line fluxes around $1{-}2 \times 10^{-17}$ erg s^{-1} cm^{-2}. These deep observations will be carried out during clear dark nights (or with low Moon illumination) and down to a maximum airmass of 1.22, to limit the effect of extinction at the blue end of the spectra (see the NUV extinction curve for the CAHA site in [1]).
The Acquisition and Guiding (A&G) system will use stars catalogued in GAIA-DR3 [2] with 17<g<18 for pointing and guiding purposes. The patrol area covers around 2×2 $arcmin^2$, enough to match the expected stellar densities around the cluster positions.
As TARSIS observes in the blue (3 quadrants) and red (1 quadrant) parts of the spectrum simultaneously, a strategy has been developed to obtain full spectral coverage in all surveyed area. It consists of taking 4 consecutive scans, each of them rotating the instrument by 90° with respect to the previous one (around the FoV centre). This sequence of 4 pointings at 0°, 90°, 180° and 270° constitutes an observing block (OB).
Each cluster extends over a few hundreds $arcmin^2$ and needs to be covered with a mosaic of OBs. We consider a default integration time of 1920s per pointing in order to minimize the readout noise in the combined data cubes. Each position needs to be covered 15(5) times in the blue(red) to get to the required depth, needing 5 OBs in total.

3.2 Dithering strategies

As CATARSIS is a blind, flux-limited survey it would benefit from an homogeneous coverage over the full surveyed area. However, the presence of the small gap between quadrants and the slightly nonuniform response pattern over the focal plane due to vignetting (see Fig 1) introduces some difficulties. To mitigate these effects and try to maximize homogeneity as much as possible two dithering strategies have been proposed for CATARSIS. Both of them follow the scheme presented in Fig 2, with the 5 required OBs per position being done following a *quincunx* pattern (see Fig 2), with offsets of size **d** between positions. Then, a larger shift of angular distance **s** is introduced to continue mapping the adjacent area of the sky. Depending on the size of the offset **s** one can have:
Gap dithering: The shift **d** matches the gap size between quadrants (2 spaxels), so that every gap is covered almost completely by the other 4 OBs in the sequence.
Homogeneity dithering: A larger shift **d** tries to match the spatial frequency in the vignetting

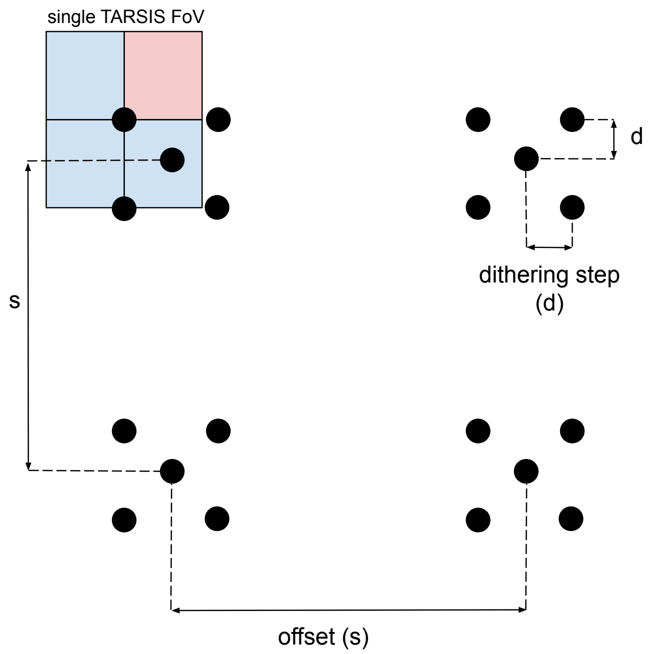

Figure 2: *Quincunx* dithering pattern proposed for CATARSIS with 5 points separated by a step **d**. The same pattern repeats at distances given by an offset **s** to map a larger area.

pattern so that the superposition of the 5 OBs becomes as uniform as possible, excluding the area covered in the narrow gaps.

3.3 Source scheduling with deep reinforcement learning

We are also exploring ways to efficiently schedule the observation of the CATARSIS fields. A good strategy should aim at optimizing the scientific quality of the collected data as well as making the best use of the observing time. Highly automated procedures are advantageous because they tend to minimize daily workload during the survey execution. They also minimize subjectivity in the decisions and unintended mistakes (such as repetitions) leading to unnecessary overheads. In addition, a certain level of flexibility is desirable to account for situations that cannot be planned in advance (instrumental failures, unexpected weather patterns, etc). We explored the use of Reinforcement Learning (RL) as a strategy fulfilling these requirements.
In the last decades, RL has been successfully applied to a variety of problems as a strategy of decision making in complex environments. This technique is based on having an *agent* exploring many different ways to solve a problem, while interacting with a certain environment. The agent finds rewards or penalties along the way, depending on the selected path. This methodology mimics the trial-and-error learning process that humans use to achieve their goals. The combination of RL and deep neural networks (DNN) has resulted in the

Deep Reinforcement Learning (DRL) algorithms, where the agent's decisions are better directed towards an optimal solution thanks to the multi-dimensional non-linear fitting power enabled by neural networks. There are notable recent successes of DRL applied, for instance, to gaming [6] or robotics [7].

DRL is an iterative algorithm and Figure 3 summarizes the different ingredients and pro-

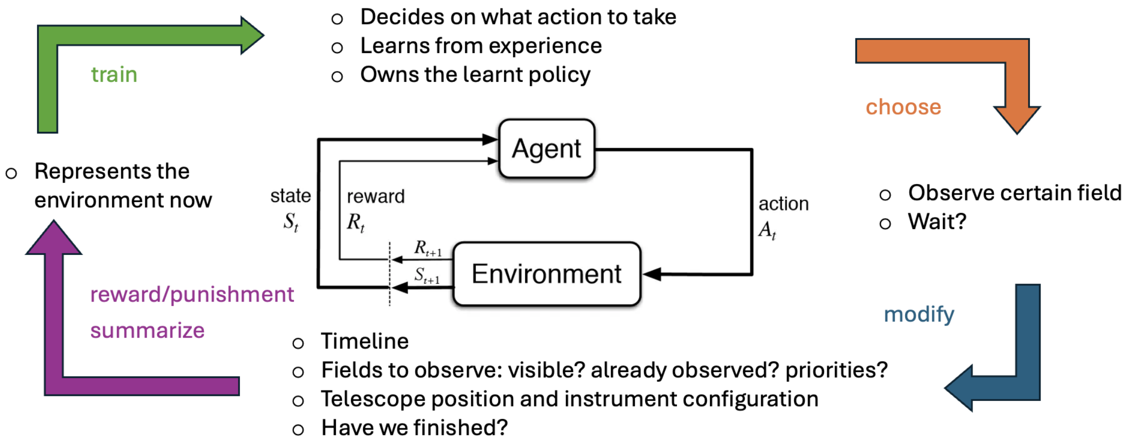

Figure 3: Scheme of the DQN algorithm applied to the CATARSIS scheduling problem.

cesses involved in a single iteration (episode) of DRL applied to the CATARSIS scheduling problem, which is based on the Deep Q-learning (DQN) algorithm. We run multiple episodes starting from the beginning of the survey, letting the agent decide what to observe at every moment while walking through the different states s, till it reaches the end of the survey. After every decision process we apply the *Bellmann equation* (Eq. 1) to the so-called action-value function $Q(s, a)$, which we approximate with a neural network.

$$Q(s, a) = r + \gamma \; max_{a'} Q(s', a') \tag{1}$$

In Eq. 1, r represents the (possibly) obtained reward after the agent has moved one step from the previous to the current state s. We update the action-value function by adding to this reward the maximum action-value function for the next step, after considering all possible actions available to the agent. The γ hyper-parameter is called the discount factor and can take values $\in$ [0-1], with γ closer to 1 meaning a higher weight given to long-term rewards compared to more immediate rewards.

We have implemented a DQN algorithm to tackle initially a simplified scheduling problem with a limited number of fields per cluster (up to 20) with equal priority and a single observing scan per pointing. This helps us to debug and improve the implementation while speeding up the model training. We have pre-computed the timeline corresponding to dark (or dark-grey) nights across 6 years so the agent can navigate through it and decide what to observe at every step (according to its learnt policy) among the observable fields. We then simulate these observations, as well as realistic telescope and instrument overheads (slew movements, FoV rotation, CCD readout, etc). Our reward policy consists of penalizing longer overheads and giving partial rewards after the full completion of a cluster in the sample. A big final reward is obtained when completing the survey. Throughout the episodes we collect triplets with states, decisions taken and obtained rewards to feed the neural network with a batch of experiences as input data. With the network output we update the agent's model driving its decision policy. We implemented an ϵ-greedy algorithm to let the agent occasionally explore randomly instead of following the learnt policy, to reduce the possibility of getting stuck in a local optima. For the learning process we have used a simple network architecture with input and output layers and two intermediate dense layers.

We run various experiments with different combinations of hyper-parameters to learn about their effect in the solution. From the experiments conducted so far, a reasonable setup consisted of 1000-episode training where the agent started to explore freely (ϵ=1) and we decreased ϵ by 0.5% after each episode, which resulted in a progressive improvement of the survey policy. After such a training the agent found a decision policy to decrease the slewing overhead by $\sim$20%, reaching 150h, compared to the $\sim$185h from a purely random selection of fields. This optimization also resulted in completing the survey 5 days earlier than the random selection.
Our next move will be simulating more realistically the survey execution, including some of the complexities skipped so far, like calibration procedures, the effect from weather and atmospheric extinction or the presence of satellite trails from mega-constellations.

Acknowledgments

This work is part of the project I+D+i PID2022-138621NB-I00, and also partially supported by project I+D+i PID2021-123417OB-I00, both funded by MCIN/AEI/10.13039/ 5011000110033/. This project received funds from the European Union - NextGenerationEU, through the TAU-UCM project with reference number PR47/21.

References

[1] Bellido-Hacar, I., Oñorbe-Bernis, J., Iglesias-Páramo, J., et al. 2024, *Study of NUV extinction in the Calar Alto sky and forecasts for the CATARSIS survey.* Poster presented at the XVI Reunión científica de la SEA, July 2024, Granada, Spain.

[2] Gaia collaboration et al. 2023j (DR3), A&A 674, pp. A1 DOI: 10.1051/0004-6361/202243940

[3] Gil de Paz, A., Iglesias-Páramo, J., Carrasco, E., et al. 2024, Proc. of SPIE, Vol 13096, Ground-based and Airborne Instrumentation for Astronomy X, 1309620. DOI: 10.1117/12.3016123

[4] Sánchez-Blázquez, P., Relaño, M., Iglesias, J., et al. 2020, XIV Reunión científica de la SEA, 84.

[5] Sánchez-Blázquez, P., et al. 2024, XVI Reunión científica de la SEA, July 2024, Granada, Spain

[6] Shao, K., Tang, Z., Zhu, Y. et al. 2019, eprint arXiv:1912.10944

[7] Tang, C., Abbatematteo, B., Hu, J. et al. 2024, eprint arXiv:2408.03539