UNIVERSITÀ DI TRENTO

Department of Physics
Doctoral School in Physics - XXXVI Cycle

Towards the formation of a forward-boosted beam of neutral antihydrogen for gravitational studies

Dissertation presented for the degree of
Philosophiæ Doctor

Candidate:

Marco Volponi

Supervisors: Prof. Roberto Sennen Brusa
Dr. Michael Doser
Co-supervisors: Dr. Ruggero Caravita
Prof. Lucio Pancheri

In collaboration with
the European Organisation for Nuclear Research - CERN

Defended on 17 May 2024

A Claudia, che mi ha fatto scoprire i colori della vita.

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are.
If it doesn't agree with experiments, it's wrong.

R. P. Feynman

Declaration

Parts of this thesis are based on the work presented in the articles [1, 2, 3], of which I am first and main author, and the article [4] and technical reports [5, 6, 7], of which I am co-author. Excerpts and figures of these publications have been reused, both identical or modified.

Abstract

The matter/antimatter asymmetry is one of the greatest mysteries in modern physics. In fact, the observable universe's lack of antimatter poses a considerable problem for physicists, especially because the Standard Model predicts that the Big Bang should have produced identical amounts of matter and antimatter. The deviations observed experimentally in the production of matter with respect to antimatter, both in the leptonic and the hadronic sectors, are still insufficient to account for the matter's domination of the universe. Therefore, one possible explanation for the discrepancy is that antimatter could interact differently than matter with gravity, resulting in a violation of the (Weak) Equivalence Principle.

The Equivalence Principle, one of the pillars of general relativity, postulates that the gravitational charge of antimatter should be equal to its inertial mass, as it is for matter. However, some physics beyond the Standard Model (BSM) theories admit a difference in the gravitational properties of matter and antimatter [8].

Measuring antimatter's gravitational force is a challenging task, both because of its scarcity and the complexities involved in trapping and handling antimatter particles. Indirect limits based on astronomical data have been posed [9], and lately, indirect measurements have progressed significantly [10]. Nonetheless, these indirect methods depend on theoretical models involving additional hypotheses: therefore they cannot be as conclusive as direct measurements. To date, the only direct measurement of the matter-antimatter gravitational interaction is from ALPHA [11]: while eliminating the possibility of "pure-repulsion", its precision still lacks towards what is considered theoretically intriguing.

This PhD thesis has been developed in the context of the AEgIS experiment (Antimatter Experiment: Gravity, Interferometry, Spectroscopy), which aims at measuring directly the acceleration exerted on antimatter by the gravitational field of the Earth and potentially by other gravity-like interactions. Specifically, antihydrogen has been selected as the test particle, as it is the simplest system of a neutral antimatter that can be synthesized[1]. To perform such a measurement, the methodology chosen is to create a pulsed beam of antihydrogen atoms, well defined in time, and let it pass through the grids of a moiré deflectometer [13] while accelerating because of the influence of the Earth's gravitational field. By the pattern created by the particles passing through the grids, the displacement due to gravity can be determined. For the formation of antihydrogen, AEgIS relies on the charge-exchange reaction between a positronium atom (Ps) (which is the bound state of an electron and a positron) and an antiproton: in the reaction, the electron is swapped with the antiproton, and the antihydrogen atom is thus formed.

The AEgIS experiment is located at the European Organisation for Nuclear Research (CERN), in the Antimatter Factory (AD, from Antiproton Decelerator). The AD is the sole source of bunched trappable antiprotons in the world: they are a key ingredient for the formation of antihydrogen.

The AEgIS experimental apparatus consists of multiple subsystems linked together, used in different combinations depending on the experiment's needs (e.g. antihydrogen formation instead

[1]There have been efforts to observe the free fall of charged antiparticles within Earth's gravitational field [12], which have evidenced the need to use neutral antimatter systems for such experiments.

of positronium cooling). The principal piece of hardware is the main vacuum chamber, hosting the cryostat, the magnets system and the traps. The magnet system divides the traps into two zones, one at 5 T and one at 1 T: in the first region, the *capture trap* is present, with electrodes tunable in the ±200 V range and three dedicated electrodes going up to 15 kV, and it is used to capture the antiprotons coming from the decelerators, after being moderated by a material degrader. The high magnetic field is used to provide stronger radial confinement to the more energetic particles and to ensure a more efficient $\bar{p}$ cooling via electron sympathetic cooling (which is proportional to B^2). In the 1 T region, instead, is present the *formation trap*: here the cold antiprotons are transferred from the *capture trap*, and the antihydrogen production takes place. A lower magnetic field ensures a higher rate of formation: in fact, the antihydrogen charge-exchange cross-section is proportional to the fourth power of the Rydberg level of the positronium involved ($\sigma_{\bar{H}} \propto n_{Ps}^4$). But the maximum Rydberg level achievable by Ps is limited by the magnetic field, since it is bounded by the motional Stark effect induced field ionisation ($n_{max} \propto B^{-1/4}$) [14]. Therefore, a lower magnetic field in the formation region can enhance greatly the production rate of antihydrogen.

To the main vacuum chamber, the positrons line is connected, which is used to inject positrons in the formation trap, to form the positronium atoms needed for the $\bar{H}$ formation. It consists of a ^{22}Na source emitting e^+, moderated by a solid Neon moderator and stored in a Surko trap. To form positronium, a bunch of positrons are extracted and collided into the $e^+ \rightarrow$Ps converter, consisting of a nanochanneled porous silica plate. A series of two lasers, then, is used to excite the so-formed Ps from the ground state to a high Rydberg state ($^1S \rightarrow^3 P \rightarrow 17 \div 32$). The cloud of Ps then expands towards the $\bar{p}$ plasma, and $\bar{H}$ formation can take place. The number of antihydrogens thus created is determined by observing, using the scintillators posed around the cryostat, the difference in annihilation rates arising when both positrons, antiprotons, and lasers are present, in opposition to the lack of (at least) one of them.

AE$\bar{g}$IS successfully produced cold antihydrogen in pulsed mode in 2018 [15], marking the end of its Phase 1. The formation rate was determined to be approximately 0.05 $\bar{H}$ per decelerator cycle, which lasts $\sim$ 110 s.

This result showed that to arrive at a gravity measurement with a precision in the order of 1 %, antihydrogen formation needs to be improved significantly: in order to gather enough statistical data, a rate of approximately $1 \div 10$ $\bar{H}$ per decelerator cycle is necessary, which is more than two orders of magnitude greater than what was obtained in Phase 1. A substantial reduction in temperature is also required, from the previously reached level of around 400 K to a few tenths of kelvin. Therefore, AE$\bar{g}$IS entered in 2019 (together with CERN LS2), its Phase 2 (which is going to last until CERN LS3, in 2025), with four goals: the two just aforementioned aims (2-3 order of magnitude higher rate and one order of magnitude colder $\bar{H}$, with respect to Phase 1), together with the formation of a forward-boosted beam of antihydrogen, and the development of a moiré deflectometer prototype for inertial measurement.

To achieve these objectives, major upgrades of the apparatus have been deemed necessary, starting from the formation scheme. The Ps target has been positioned on the axis of the trap, to illuminate the $\bar{p}$ plasma collinearly, and not perpendicularly as before: this has raised the maximum Ps Rydberg level from $\sim$ 19 to above 32 by reducing the Ps ionisation due to the motional Stark effect ($n_{Ps}^{max} \propto \theta_{Ps|B}^{-1/4}$). The updated formation scheme necessitated a redesign of the formation trap, which was built and installed in 2022. Additionally, the $e^+ \rightarrow$Ps target has been optimised by fine-tuning the morphology of the nano-channel, so to increase up to five times the efficiency of positronium generation.

In the context of the upgrades to the AE$\bar{g}$IS apparatus, my main contribution has been the development of the new control system controlling the entire experiment. In fact, AE$\bar{g}$IS has been operated using multiple independent control systems for each subsystem, coordinating them by a grown-up program and by the manual labour of multiple experts simultaneously. This already resulted demanding from the scientists at the end of Phase 1, when antihydrogen production was successfully attempted. But with the introduction, in 2021, of ELENA (Extra Low Energy

Antiproton ring) [16], the modality of antiprotons delivery changed from 8 h daily shifts to continuous cycles: keeping the experiment up day and night would have become completely unfeasible with the previous system, ultimately leading to beam time loss and team overwork. Therefore, it was decided to completely rewrite the control system, creating a distributed system with stability and automation as the main targets.

The new control system is called CIRCUS (Computer Interface for Reliably Controlling, in an Unsupervised manner, Scientific experiments) [1], whose main component is the TALOS framework (Total Automation of Labview Operations for Science) [2]. TALOS creates an environment where all individual control programs are subdivided into atomic modules, called *MicroServices*, which are integrated into a single, coordinated, distributed system: these features are the base that enables the complete automation of experimental procedures since high-level decisions often rely on parameters residing on multiple computers. TALOS is built upon the distributed system architecture provided by the Actor Model [17], which facilitates the modular structure. In particular, it is realised by dividing the code into standalone units, called *MicroServices*, each with a precise scope and function, which operate in parallel, communicating through non-blocking messages, to ensure complete asynchronous functionality. The unification of the distributed system is realised by a common process, called *Guardian*, an instance of which is executing on each machine. Every *Guardian* monitors the status of the other *Guardian* in the experiment's computer network and oversees the local MicroServices, thanks to a series of three distributed watchdog systems. This system ensures that no component becoming unresponsive can pass unnoticed: it significantly strengthens the reliability and safety of the system, ultimately leading to unsupervised operation for extended periods. Furthermore, the automation of the system was augmented by the introduction of an optimiser, that leverages the feedback given by the online analysis of the data acquired, to autonomously find the best parameters that optimise a series of predefined observable

CIRCUS, per se, is experiment-agnostic, being general enough to be usable by other experiments than AE$\bar{g}$IS, in particular nuclear, atomic and quantum ones. Therefore, to be able to adopt it in AE$\bar{g}$IS, I also coded the majority of the MicroServices necessary to manage most of the detectors and actuators, plus the interface with the decelerators and the data acquisition interface, among others.

The implementation of CIRCUS would not have been possible without upgrading the control electronics from the previous custom-made hardware (with a very limiting software interface) to a new one based on ARTIQ/Sinara [18], an open hardware & software ecosystem expressly created for quantum physics experiment. The hardware modularity has enabled progressive migration and guarantees future-proofness; its ns synchronisation (internal and w.r.t. an external clock) capability has further improved the timing accuracy of the apparatus. Conversely, the programming interface done in ARTIQ, a Python-based real-time language, has greatly simplified the generation of experimental procedures, enabling both a library-based approach (which minimises debugging time and code duplication) and the integration with TALOS, to ensure the automation of the full system. In this respect, during this thesis, I contributed to the installation and integration of the Sinara hardware, and I wrote part of the libraries used to operate it and, in particular, to interface the FPGA seamlessly with the CIRCUS control system.

During the three antiprotons campaigns (2021, 2022, 2023) that occurred during the development of this thesis, the system was tested with the particles, and the various operations needed to form antihydrogen were re-developed using the new control system. In this thesis, the physics motivations that brought these procedures to the form that we found most performant, which is also presented, are explained. The main operations tested and optimised are: antiprotons capture, trap closure time optimisation, antiproton beam alignment, traps voltage reshape, electrons loading, $\bar{p}$ sympathetic cooling, $\bar{p}$ plasma compression with rotating-wall technique, electrons removal before antiprotons transfer, antiprotons transfer and recapture in the formation trap, $\bar{p}$ ballistic transfer to the formation trap, electrons recycling, antiprotons partial recycling, positronium formation optimisation, laser synchronisation for Ps excitation, antihydrogen formation.

These operations were implemented sequentially, by consolidating each one before passing to the subsequent one. A key ingredient was the extensive use of custom libraries for the Python code of the various experiments: every time an atomic operation was tested and optimised, it was defined as a function in the AE$\bar{g}$IS libraries, so to be able to recall it in all the subsequent operations scripts. This greatly facilitated the development, since the new code was progressively added to the already consolidated one, minimising code duplication, script proliferation, and debugging time.

To guide the success and performance of the operations, two detectors were mainly used: the scintillating slabs surrounding the main apparatus, acquired via photomultipliers (PMTs) and digitised, and a combination of a multichannel plate (MCP), a phosphor screen and a camera, placed downstream with respect to the formation trap. The firsts are fundamental to monitoring and understanding the lifetime and the quality of the operations with the antiprotons, by understanding the time and position of the various annihilations. The second is mainly used to understand radial plasma profiles and to perform time-of-flight (ToF) analyses.

In the end, to determine the formation rate of antihydrogen with the new apparatus, I took part in the data analysis, by developing one of the three parallel analyses that have been performed. One analysis, analogue to the one already performed in 2018, used the scintillators counts to discriminate the difference in $\bar{p}$ annihilation profiles in case that both $\bar{p}$, positrons and laser were employed together, or that one of them was missing: by the difference, $\bar{H}$ production can be inferred. My data analysis, instead, used different techniques to look at the MCP data to understand not only the formation rate but also, by comparison with the scintillators' data, the possible forward-boosted antihydrogen formation. The images were algorithmically selected, background corrected, and binarised; then, clusters were extracted. A Bayesian test was finally performed between runs where $\bar{H}$ production was attempted, versus runs where it was suppressed by omitting one of the lasers. The analyses hint toward a successful production of antihydrogen with the upgraded apparatus, at a higher rate than the one seen at the end of Phase 1: nevertheless, it has been difficult to determine the direction of production with the MCP analyses, and deeper analyses and/or more data is needed to fully conclude.

Overall, the work accomplished during this PhD thesis has been fundamental to achieving the formation of antihydrogen using the ballistic antiprotons transfer, which leads to the creation of a forward-boosted beam of neutral antihydrogen. This is an important milestone towards measuring directly the gravitational interaction between matter and antimatter, which could lead to a confirmation of the WEP on antimatter or could find a violation and, so, hint at new physics.

Contents

Chapter 1

Where is all the antimatter?

One of the greatest still unsolved mysteries of modern physics is the imbalance between the amount of matter and antimatter present in our universe. The Standard Model (SM) of particle physics predicts an almost perfect symmetrical abundance of the two classes of particles, but in practice, the visible part of our universe (i.e., all the particles of the observable universe, excluding the Dark Matter sector) is constituted of nearly only matter. In fact, if regions of space where antimatter is predominant exist, we would be able to detect the gamma rays emitted by the annihilations at the boundary of the regions [19, 20]. Moreover, a universe made of patches respectively dominated by matter and antimatter would generate anisotropies (mainly in the form of large-scale isocurvature fluctuations) in the Cosmic Microwave Background (CMB), which are not present [21].

Even if the presence of small amounts of antimatter organised in dense objects cannot be completely excluded – the existence of antistars, even in our galaxy, cannot be ruled out by CMB measurements [22] –, their number has to be very limited: a recent survey [23] using data from the Fermi Large-Area Telescope has put a upper bound of $2.5 \cdot 10^{-6}$ to the maximum fraction of antistars allowed in our galaxy by gamma-ray sources[1]. Therefore, the imbalance remains.

Despite being the universe basically devoid of any antimatter, the matter/antimatter asymmetry needed to explain the imbalance is very small: in fact, in the early moments of the universe after the Big Bang, when the temperature of the universe went approximately below 10^{12} K, quarks and antiquarks combined into baryons and mesons, mostly annihilating with each other, but leaving an excess of $1.67 \cdot 10^9$ baryons over the number of photons [24, 25]. A possible explanation for the excess is that the Big Bang originally started with a baryonic number in the order of 10^{80}([2]), but this hypothesis is a huge fine-tuning. A preferred solution is that the universe started perfectly symmetric, and then a combination of processes took place, unbalancing the content towards matter.

1.1 Sakharov conditions

In 1967, Andrei Sakharov suggested a set of three conditions that are necessary in order to have a mechanism that can produce matter in greater amounts than antimatter [27]. They are:

- **Baryon number violation:** the existence of a process that violates baryon number (B) conservation is mandatory, otherwise baryogenesis is impossible;
- **C&CP violation**: at least one process violating both Charge (C) and Charge-Parity (CP) symmetries is necessary, otherwise the ΔB-violating process would be compensated by its CP-symmetric counterpart, and no asymmetry could arise;

[1] Actually, the fraction could go up to ~ 10 % if the distribution of antistar masses would be very different from the stars one (i.e. only made up of Sun-size antistars): but again, this would be an unexplained fine-tuning.

[2] Which is the estimated number of nucleons in the visible universe [26].

- **Out of thermal equilibrium:** the reactions violating the conservation principles just discussed must happen at a rate out of thermal equilibrium, otherwise CPT symmetry would ensure that the increasing ΔB process would compensate with its own symmetric decreasing ΔB process. The rate of the process(es) must be smaller in relation to the Universe expansion rate at the time of the baryon generation.

No definitive explanation has been found to address these points: in the following, the most common hypotheses are presented.

1.1.1 Baryon number violation

All the interactions of the SM conserve both baryon B and lepton L numbers independently, therefore a new physical process is required to introduce the violation. Several mechanisms have been theorised to violate the baryon number, either directly (true *baryogenesis*), or first violating the lepton number and then transforming the generated ΔL into ΔB through a $B - L$ conservation mechanism (*leptogenesis*).

A number of theories suppose the violation to happen at very high energy, where possibly Grand Unification Theories (GUT) may be dominant. One of the first of this class was the one exposed in [28], followed by many others (some notable examples are [29, 30, 31]). Typically these models involve the postulation of very massive particles (so that they can only exist in the very early instants of the universe), which have decay channels slightly favouring matter over antimatter.

Another set of theories, instead, foresees that the baryon number violation took place close to, or during, the electroweak (EW) phase transition [32, 33, 34, 35]. In particular, it has been hypothesised that, if the EW transition is a first-order phase transition, the non-perturbative processes forming (going with the name of *sphalerons*) could in principle break baryon number conservation without the need for additional physics Beyond the Standard Model (BSM) [36, 37, 38].

Conversely, other theories explore leptogenesis, where a ΔL-violating process can produce more leptons than anti-leptons, ultimately leading to a matter-dominated universe [39, 40, 41]. The asymmetry can be, for example, generated by introducing three heavy Majorana right-handed neutrinos, coupled to the SM neutrino; the baryon asymmetry is then formed by a ΔL-ΔB transfer mechanism happening before, or during, the electroweak phase transition [42].

Further theories exists, involving Dark Matter (DM) [43, 44, 45], gravity modifications [46, 47, 48], extra-dimensions [49, 50], and many more. At present, no experimental hint towards any of them has been established.

1.1.2 C&CP violation

Contrarily for the baryon number violation, the violation of the CP symmetry in the SM is possible [42], and can happen in multiple manners.

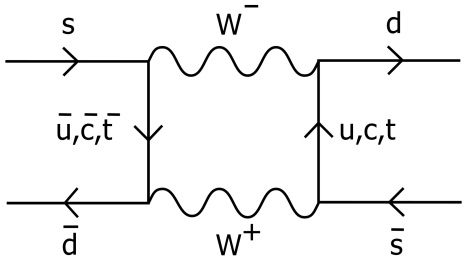
In the hadronic electroweak sector, it is linked to the complex phase inside the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which mixes quark families: the transformation between mass-eigenstates (right) and flavour eigenstates (left) is given by equation 1.1.1.

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \tag{1.1.1}$$

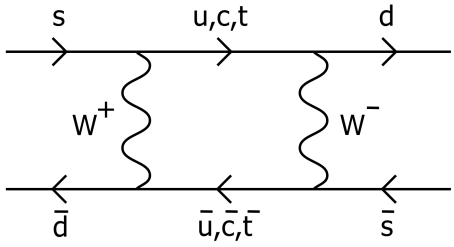
With three generations of quarks, this matrix can be reduced to three mixing angles (θ_{12}, θ_{23} and θ_{13}) and a complex phase (δ_{13}), which cannot be re-absorbed into the redefinition of the fields: an example of parametrisation is (here with c_{jk} and s_{jk} are indicated the cosines and sines of the respective angles):

$$V_{\mathrm{CKM}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{13}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{13}} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.1.2}$$

The mixing leads to particle-antiparticle oscillations (like $\mathrm{K}^0 \rightleftarrows \overline{\mathrm{K}}^0$ and $\mathrm{B}^0 \rightleftarrows \overline{\mathrm{B}}^0$), through 1-loop suppressed processes as the ones depicted in figure 1.1, a mechanism that has been demonstrated to violate CP symmetry [51]. Nevertheless, the amount of violation is several orders of magnitude too low to explain the baryon imbalance[3] [52].



(a) Kaon oscillation *box* diagram 1.



(b) Kaon oscillation *box* diagram 2.

Figure 1.1 – The two Feynman diagrams giving the leading contribution for the $\mathrm{K}^0 \rightleftarrows \bar{\mathrm{K}}^0$ oscillation: their combination can violate CP symmetry (from [53]).

Another place where CP violation is admitted in the SM is in the leptonic sector. Assuming neutrinos to have masses as suggested by oscillation experiments in the framework of some suitable, minimal extension of the SM, then also neutrinos can mix between flavour and mass eigenstates similar to what happens to quarks, through the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix:

$$\begin{bmatrix} \nu_{\mathrm{e}} \\ \nu_\mu \\ \nu_\tau \end{bmatrix} = \begin{bmatrix} U_{\mathrm{e}1} & U_{\mathrm{e}2} & U_{\mathrm{e}3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} \tag{1.1.3}$$

where 1, 2 and 3 are the mass eigenstates, while e, μ and τ are the flavour eigenstates. The mixing leads to neutrino oscillations, where a neutrino generated by a weak interaction in a precise flavour state has an oscillating probability to mutate to the other flavours (and back) during the propagation [54]. In the simple case of only two neutrino families (a valid approximation in some real cases, e.g. $\nu_\mu \leftrightarrow \nu_\tau$ in the atmosphere, or $\nu_e \leftrightarrow \nu_x$ for solar neutrinos, where ν_x is a superposition of the other two), the oscillation takes the form (using natural units, and indicating with θ the mixing angle between the two families):

$$P_{\alpha\to\beta} = \sin^2(2\theta)\,\sin^2\left(\frac{\Delta m^2 L}{4E}\right). \tag{1.1.4}$$

Evidences in difference in neutrinos and antineutrinos oscillation have been reported [55], and the non-zero complex phase δ_c of the PMNS matrix would lead to a CP violation in the leptonic sector. If confirmed by future experiments (like "Tokay 2 Hyper-Kamiokande" (T2HK) [56] and "Deep Underground Neutrino Experiment" (DUNE) [57]), this source of CP violation could propel us forward in the understanding of the matter/antimatter asymmetry of our universe.

The last source of CP violation inside the SM is in QCD, by the terms of the Lagrangian with θ and θ'[4]:

[3] The ineffectiveness is not only a matter of the smallness of the phase δ_{13} but it is further enhanced by the difference in mass between quarks and the weak interaction scale: all quarks except the top t are basically massless with respect to the W mass [19].

[4] For simplicity, only one species of quark is used in the discussion here.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{\theta g^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu} + \psi(i\gamma^\mu D_\mu - me^{i\theta'\gamma^5})\psi. \tag{1.1.5}$$

A chiral transformation can eliminate one of the two elements, but not both (so typically the second term is removed, to only deal with real masses). Therefore QCD implies a violation of CP: this would lead to a non-null neutron electric dipole moment [58] (here $\bar{\theta}$ is the real SM CP violating angle)

$$d_N = (5.2 \cdot 10^{-16} e \cdot m)(\bar{\theta}) \tag{1.1.6}$$

Recent measurements of the neutron electric dipole moment give an upper bound of $|d_n| < 2.9 \cdot 10^{-26}$e cm [59], which leads to a violation term of $\bar{\theta} \lesssim 10^{-10}$. In conclusion, from QCD there is no real CP violation in the SM[5].

In synthesis, if the CP violation arising from the leptonic sector will also be measured to be too small to explain the unbalance, other solutions involving BSM physics will be necessary.

1.1.3 Gravity-like interactions

A different possibility proposed to explain the asymmetry in the matter/antimatter content of the universe lies in a different gravitational effective interaction between them [61]. Several ways for this difference to happen have been theorised [8, 62]. A possible manner to accommodate such a difference is to introduce additional gravity-like interactions, often in the form of scalar and vector-bosons [8], leveraging the fact that also between same-charge particles, the sign of the force can be modulated changing the spin of the mediator [63]:

$$U_{12} = -\frac{Gm_1m_2}{r}\left(1 \mp ae^{-\frac{r}{v}} + be^{-\frac{r}{s}}\right) \tag{1.1.7}$$

where the minus sign in front of the second addend holds for 1 and 2 being of the same species (matter-matter or antimatter-antimatter), while the plus sign is for the matter-antimatter case. It can be immediately seen that if $a = b$ and $v = s$, the vector and scalar interactions cancel out in the same-species case, but a non-null force arises in the matter-antimatter case. The magnitude of the allowed deviation with this simple model depends strongly on the force range: for $v = s < 1000\,\mathrm{m}$, the limits are in the order of $10^{-5}g$, while for longer ranges the limit relaxes, and a force up to $10^{-2}g$ can arise [8, 64].

To probe the existence of a difference in the matter/antimatter gravitational effective coupling (i.e. the force exerted by the normal gravitational interaction plus the one generated by a gravity-like interaction), a good technique is to perform a Weak Equivalence Principle (WEP) test using antimatter.

1.2 Einstein Equivalence Principle

We too often take for granted that

$$m_{inertial} = m_{gravitational}\,. \tag{1.2.1}$$

In practice, there is no first-principles argument from which to derive the equivalence between the inertial constant of a body and its gravitational charge. Nevertheless, since this was first experimentally established by Galileo, and exposed in his famous "*Discorsi e dimostrazioni matematiche intorno a due nuove scienze*" [65], and then raised to fundamental axiom by Isaac Newton in his

[5] Conversely, the smallness of $\bar{\theta}$ is considered "unnatural" from the theoretical point of view – it would be expected to be of order 1 –: the vanishing of the axial coupling is called *strong CP problem*, and it is one of the greatest unsolved puzzles in physics. Possible solutions to the strong CP problem have been theorised, the most famous being the Peccei-Quinn mechanism [60], but no experiment to date has confirmed any of them.

"*Philosophiæ Naturalis Principia Mathematica*" [66] with the name of *Universality of Free-Fall* (UFF), it has never been disproved, becoming a pillar of modern science. In fact, on this equivalence, Albert Einstein based his work on the derivation of the theories of Special and General Relativity [67, 68], extending it to the principle that "a falling object follows exactly the same trajectory inside a gravitational potential and in an accelerating frame of reference" (the other pillar of Relativity being "the speed of light c is independent of the frame of reference").

Customarily, the Einstein Equivalence Principle is divided into three parts[69]: the universality of free fall, also known as the weak equivalence principle (WEP), the local Lorentz invariance (LLI) and local position invariance (LPI). Furthermore, an additional distinction can be made, between the Weak Equivalence Principle, where the UFF is tested in an experiment among bodies whose gravitational self-binding energy can be neglected (e.g. a stone), and the Strong Equivalence Principle (SEP), where the gravitational energy is allowed to be of the same order of the masses and energy in play (e.g. stars, black holes, etc.). A more precise classification can be made with respect to the dimensionless quantity ϵ [70]:

$$\epsilon = \frac{GM}{Rc^2} \tag{1.2.2}$$

where M and R are respectively the mass and length scale of the phenomenon. For the entire observable universe and on the events horizon of a non-rotating black hole, $\epsilon \sim 1$; on the surface of a neutron star, $\epsilon \sim 0.2$; in the Solar System, $\epsilon \sim 10^{-5}$. The first two cases are of SEP, and the latter is of WEP.

The Universality of Free-Fall and the Einstein Equivalence Principles have been tested multiple times, both with laboratory experiments (on Earth and beyond) and with astronomical observation. In the following, the most relevant ones performed both on matter and antimatter bodies are listed.

1.2.1 WEP tests on matter

Starting from the first experiments performed by Galileo with tilted planes and pendulums(and the more famous, dropping two different metal spheres from the leaning tower of Pisa[6]) [71], the Universality of Free-Fall/WEP has been subject of several tests of increasing stringent precision. A good evolving summary is presented in [70], which is pictorially condensed in figure 1.2, showing the most relevant constraints posed up to 2014.

To determine if the WEP is violated, a dimensionless variable is defined as the normalised difference between the accelerations of the two test bodies:

$$\eta = 2\frac{|a_1 - a_2|}{|a_1 + a_2|} \tag{1.2.3}$$

1.2.1.1 Torsion balances

Historically, Eötvös' torsion balance experiment (performed in 1909) was the first measuring the UFF with a very stringent precision[7], going to $\eta < 10^{-8}$. The original torsion balance experiments were constituted of two test masses posed at the two extremities of a rod, suspended by its centre by a very thin wire. A mirror in the centre of the rod was used to reflect light to determine the rotation of the rod with great accuracy. If $m_i = m_g$ for both masses, the centrifugal force caused by the rotation of the Earth and the pull from gravity would exactly cancel out, leaving the rod at rest; conversely, if the equivalence does not hold even only for one of the two test masses, the rod would start to rotate. This type of experiments has been refined multiple times (Braginsky [73], Eöt-Wash group [74]), arriving to precision below $\eta = 3 \cdot 10^{-14}$.

[6] Although most historians believe that it has been only a thought experiment.

[7] Before him, the record was established by Bessel in 1827 with a speciallly devised pendulum [72], arriving to a precision of one part in 60000.

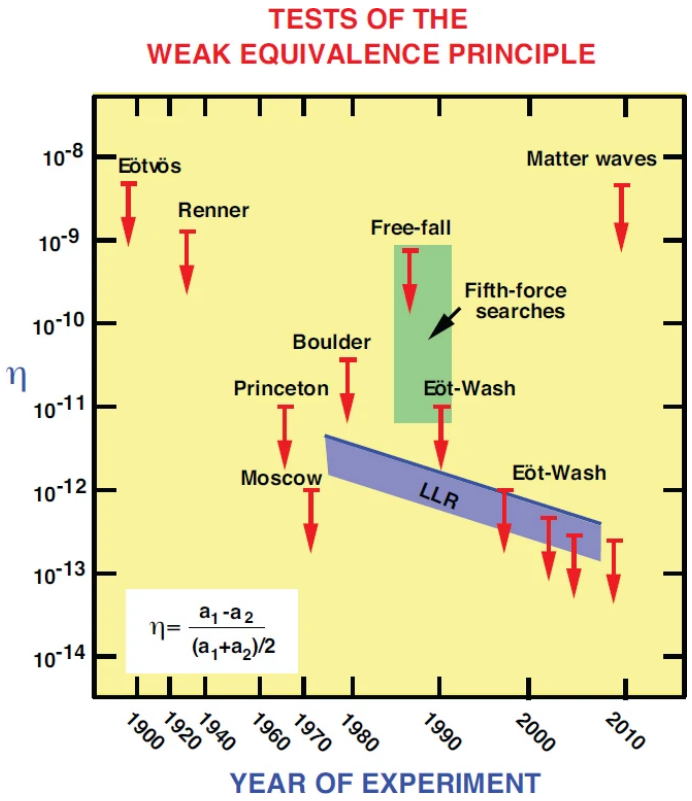


Figure 1.2 – Most relevant tests of the Weak Equivalence Principle, showing constraints placed on η, which measures the relative normalised difference in acceleration of different bodies (definition given in the inset). The blue band shows the evolving bounds on η from Lunar Laser Ranging (LLR) (from [70], data from 2014).

1.2.1.2 Lunar Laser Ranging

Another class of experiments is the Lunar Laser Ranging (LRR) [75]. The distance between the Earth and the Moon is constantly monitored via the time of flight of laser beams reflected thanks to mirrors placed on the surface of the Moon by the Apollo missions (11, 14 and 15), the Lunokhod rovers (1 and 2) and the recent Chandrayaan-3 lander. Knowing the trajectory of the Moon makes possible to test the UFF of the Moon on the Earth, with increasing precision (best value is $\eta < 7 \cdot 10^{-14}$ [76])[8].

1.2.1.3 MICROSCOPE

The best WEP limit to date has been set by MICROSCOPE (MICRO-Satellite with Compensated drag for Observing the Principle of Equivalence, or MICRO-Satellite à traînée Compensée pour l'Observation du Principe d'Equivalence) [78]. It is a satellite orbiting the Earth in a Sun-synchronous orbit in Low-Earth Orbit (LEO) (at approximately 712 km), with two test masses that are kept on a perfect trajectory by the onboard attitude control system, compensating for external factors, like residual atmospheric drag, Sun radiation pressure, etc. By the difference of the corrections needed to be exerted on the two different test masses, the UFF can be tested. At the end of the campaign, MICROSCOPE reported a limit of $\eta < 3 \cdot 10^{-15}$ [79].

[8]The LLR can also test the SEP, being performed using two planetary-size objects: there the limit is $\eta_{\text{strong}} < 10^{-4}$ [77].

1.2.2 WEP on antimatter

The Weak Equivalence Principle has been investigated also using antimatter systems. The interest is twofold: together with performing yet another test of the UFF, there is also the scientific interest in searching for gravity-like interactions in the antimatter sector, as discussed above.

In spite of that, measurements of WEP on antimatter are scarce and not of great precision, mainly because of the enormous experimental difficulties in producing and cooling antiparticles to temperatures in the range of $\sim$ K, necessary to perform gravity experiments. Conversely, indirect bounds have often better sensitivity, but are limited by the employment of a model to generate the constraint. Here I present the most relevant measurements for historical and precision reasons.

1.2.2.1 Attempts with charged antimatter

Already in 1965, Witteborn and Fairbanks tried to measure the UFF using electrons and positrons inside a vertical tube. Unfortunately, their results [80, 12] mainly pointed out the impossibility of performing such a test, since the systematic effects linked to the charged nature of the particles used were overwhelming (like the electric field patch effect [81], and the residual charge on materials [82] – a single electron at 5 m from the experiment cancel out gravity on electrons/positrons! –).

1.2.2.2 Neutrinos and antineutrinos from supernova SN1987A

On February 24th 1987, the supernova SN1987A exploded, sending a burst of neutrinos and antineutrinos in all directions, which were detected by three neutrino observatories on Earth (12 by Kamiokande II [83], 8 by the Irvine-Michigan-Brookhaven (IMB) detector [84], and 5 by the Baksan Neutrino Observatory [85]), over a span of $\sim$12.5 s. Since both neutrinos and antineutrinos arrived at the same time, this measurement was used to pose a limit on the WEP for antimatter: admitting a maximum difference of 12.5 s over a total travel time of $\sim$168 000 years, the diversity in gravitational coupling was calculated to be smaller than $\eta = 10^{-6}$ [86]. This limit was disproved in [87], making it completely null, mainly with the argument that the (anti)neutrinos were emitted by the supernova with an energy of tens of MeV, and therefore gravity has coupled with their energy-momentum tensor only via their kinetic part, and very marginally with their rest mass (which is smaller than 1 eV).

1.2.2.3 BASE redshift of p and $\bar{\text{p}}$ cyclotron frequency

The Baryon Antibaryon Symmetry Experiment (BASE) [88] is an experiment in the Antimatter Factory at CERN (same location of AE$\bar{\text{g}}$IS, see *Brief description of AE$\bar{g}$IS* (1.3.1)) which uses a Penning trap to compare the cyclotron frequencies of an antiproton and of a H^- ion, and from them infer with extreme precision matter/antimatter symmetry tests. In 2022 they published a paper [10] where they constrained the ratio of the masses of proton and antiproton to be identical to 16 parts per trillion; in the same article, they also used the data acquired along the year to perform an indirect WEP test on antiprotons.

The measurement method is based on the original proposal from the TRAP collaboration [89] of comparing the cyclotron frequency of particles and antiparticles in the trap. They argued that the cyclotron frequency is influenced by the absolute value of the gravitational potential at the time of measurement: if matter and antimatter would interact differently with gravity, their cyclotron frequencies would be modified as

$$\frac{\nu_{\bar{p}} - \nu_p}{\nu_{\text{avg}}} = \frac{3\Phi}{c^2}(\alpha_g - 1) \tag{1.2.4}$$

where $\nu_{\bar{p}/p}$ is the cyclotron frequency of the antiproton and proton respectively, Φ is the absolute value of the gravitational potential, and $(\alpha_g - 1)$ is a factor accounting for the deviation of the interaction ($\alpha_g = \bar{g}/g$). They determined the limit of the potential violation to be $|\alpha_g - 1| < 1.8 \cdot 10^{-7}$,

assuming the gravitational potential to be null at infinity and the CPT symmetry to hold. The measurement is extremely stringent, but the problem lies in the determination of the absolute potential inside the local supergalactic cluster [90, 91][9].

Nevertheless, they have made use of the fact that they have taken data in different periods of the year, and therefore on different points of the Earth's orbit around the Sun. In this manner, a differential measurement is possible, which depends only on the gravitational potential of the Sun:

$$\frac{\Delta R(t)}{R_{\text{avg}}} = \frac{3GM_\odot}{c^2}(\alpha_{g|\text{diff}} - 1)\left(\frac{1}{O(t)} - \frac{1}{O(t_0)}\right) \tag{1.2.5}$$

where $\Delta R(t)$ is the cyclotron frequency variation at time t, and $O(t) = d_p\left(1-\epsilon^2\right)\left[1+\epsilon\cos\left(\frac{2\pi t}{t_{\text{sid}}}\right)\right]$ indicate the position of the Earth along its orbit at time t (being d_p the Sun-Earth distance at the perihelion, t_{sid} the time of the sidereal year and ϵ the eccentricity of the orbit). In this manner, the potential WEP violation for antimatter is constrained as $\left|\alpha_{g|\text{diff}} - 1\right| < 3$ %.

To compare these values with the ones for the WEP test on matter, we can notice that

$$\begin{aligned}
\eta = 2\,\frac{|a_1 - a_2|}{|a_1 + a_2|} &= \\
= 2\,\frac{\left|\dfrac{F}{m_{i,1}} - \dfrac{F}{m_{i,2}}\right|}{\dfrac{F}{m_{i,1}} + \dfrac{F}{m_{i,2}}} &= \\
= 2\,\frac{\left|\dfrac{m_{g,1}}{m_{i,1}} - \dfrac{m_{g,2}}{m_{i,2}}\right|}{\dfrac{m_{g,1}}{m_{i,1}} + \dfrac{m_{g,2}}{m_{i,2}}} &= \\
= 2\,\frac{|\alpha_g - 1|}{1 + \alpha_g} &\approx \\
\approx 2\,|\alpha_g - 1| \qquad\qquad \text{for} \quad \alpha_g \ll 1\,. &
\end{aligned} \tag{1.2.6}$$

With this, we have that the absolute constraint is approximately $\eta < 4 \cdot 10^{-7}$, and the Sun-relative differential WEP test (i.e. the one considering only the effect of the gravitational potential of the Sun) is $\sim \eta_{\text{diff}} < 6$ %.

1.2.2.4 ALPHA sign determination

The Antihydrogen Laser PHysics Apparatus (ALPHA) [92] is another experiment of the Antimatter Factory at CERN (again, see *Brief description of AEgIS* (1.3.1)), mainly focusing on studying the properties of antihydrogen, both spectroscopically (like the $1S-2S$ transition) and gravitationally, with their ALPHA-g apparatus. It consists of a Penning trap positioned vertically, equipped with a radial time-projection chamber (rTPC) detector and veto scintillators all along the sides, capable of trapping neutral antihydrogen exploiting its magnetic moment. By opening the top and bottom end-caps of the trap (i.e., bringing in ~ 20 s to zero the confining voltage) and detecting the vertical distribution of the annihilations of the antihydrogen atoms on the side walls with the rTPC, the force exerted by gravity on it can be estimated. To come to a more precise measurement, a slightly more complicated technique has been employed in 2022: the confining magnetic field gradient was modified to change the distribution of the annihilation along the walls, and the value of $\bar{g}$ was extrapolated from the annihilations variation. With this technique, they arrived at a value of $\bar{g} = (0.75 \pm 0.29)g$, therefore managing to constrain the sign of the gravitational acceleration of antihydrogen in the Earth's field (it falls down) [11]. Although in terms of WEP test this measurement is very limited ($\eta \sim O(1)$), it is still the most accurate direct measure of the gravitational influence over antimatter to date.

[9] And, technically, to where the zero of the potential is posed since it is a direct use of an absolute value.

1.3 WEP test with antihydrogen at AEgIS

One experiment designed to shine a light on the matter/antimatter asymmetry is indeed AEgIS (Antimatter Experiment: Gravity, Interferometry, Spectroscopy) [93], whose main aim is to measure directly the fall of antimatter in the Earth's gravitational field: specifically, the aim is to measure the vertical displacement, caused by gravity and possible gravity-like interactions, of a pulsed beam of antihydrogen ($\bar{H}$).

The AEgIS collaboration has developed a unique pulsed scheme which is able to provide precise knowledge of the $\bar{H}$ formation time, control the final antihydrogen temperature, and manipulate its excitation state, among others. The formation of antihydrogen is based on the charge-exchange reaction between Rydberg-excited positronium atoms (Ps) and trapped, cold antiprotons from the CERN decelerators [94, 95] (as shown in figure 1.3). The detection scheme is based on measuring the vertical displacement of the $\bar{H}$ beam with a gravimeter consisting of a combination of a moiré deflectometer and a Talbot-Lau interferometer.

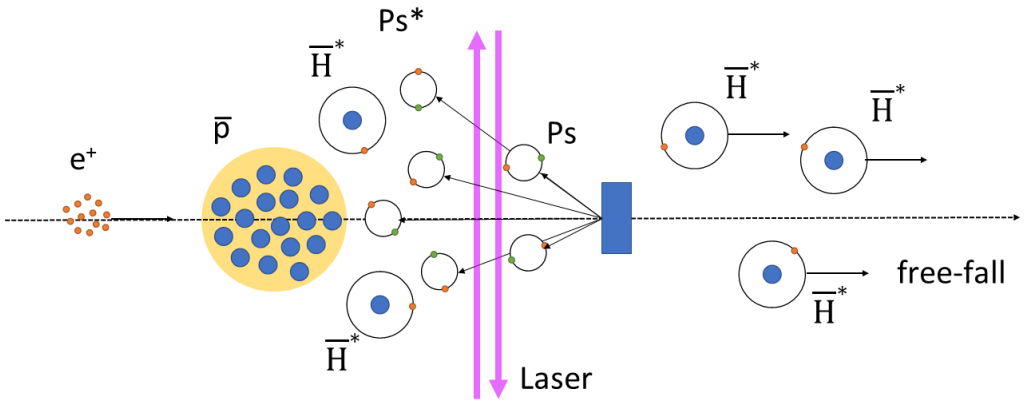


Figure 1.3 – Scheme of the antihydrogen production technique used in AEgIS. It leverages a charge-exchange reaction, where a trapped cold plasma of antiprotons is invested by a cloud of Rydberg-excited positronium atoms, created by impinging positrons on a nanochanneled silica converter and subsequently excited by a series of two lasers.

In this section, first, a brief description of the experiment and its apparatus is given, followed by the gravimeter functioning principle. Finally, the first gravimeter prototype is exposed, with an estimation of its expected sensitivity.

1.3.1 Brief description of AEgIS

The AEgIS Experiment is located at the Antimatter Factory, hosted at CERN (the European Organization for Nuclear Research) (see figures 1.4 and 1.5). The Antimatter Factory (shortened to AD, from *Antimatter Decelerator*[10]) is a very special place at CERN because its focus is the investigation of the properties of antimatter at low energies. In addition, the AD is also a very special place in the world, since it is the only existing source of low energy antiprotons ($\bar{p}$) from a synchrotron: this opens possibilities for antimatter studies that are otherwise unthinkable. In the specific, symmetry test on cold baryonic antimatter [88], on anti-atoms (antihydrogen) [14, 92, 96],

[10] AD is the original decelerator that was providing antiprotons to the experiments with a kinetic energy of 5.3 MeV. From 2021, a new decelerator is present in the complex, called ELENA (see figure 1.4), which has lowered the energy of the supplied antiprotons to 100 keV.

creation of matter-antimatter exotic system [97, 98], and many more. Being able to trap antimatter to study it opens the doors to precision physics, to stress the Standard Model in search of new physics.

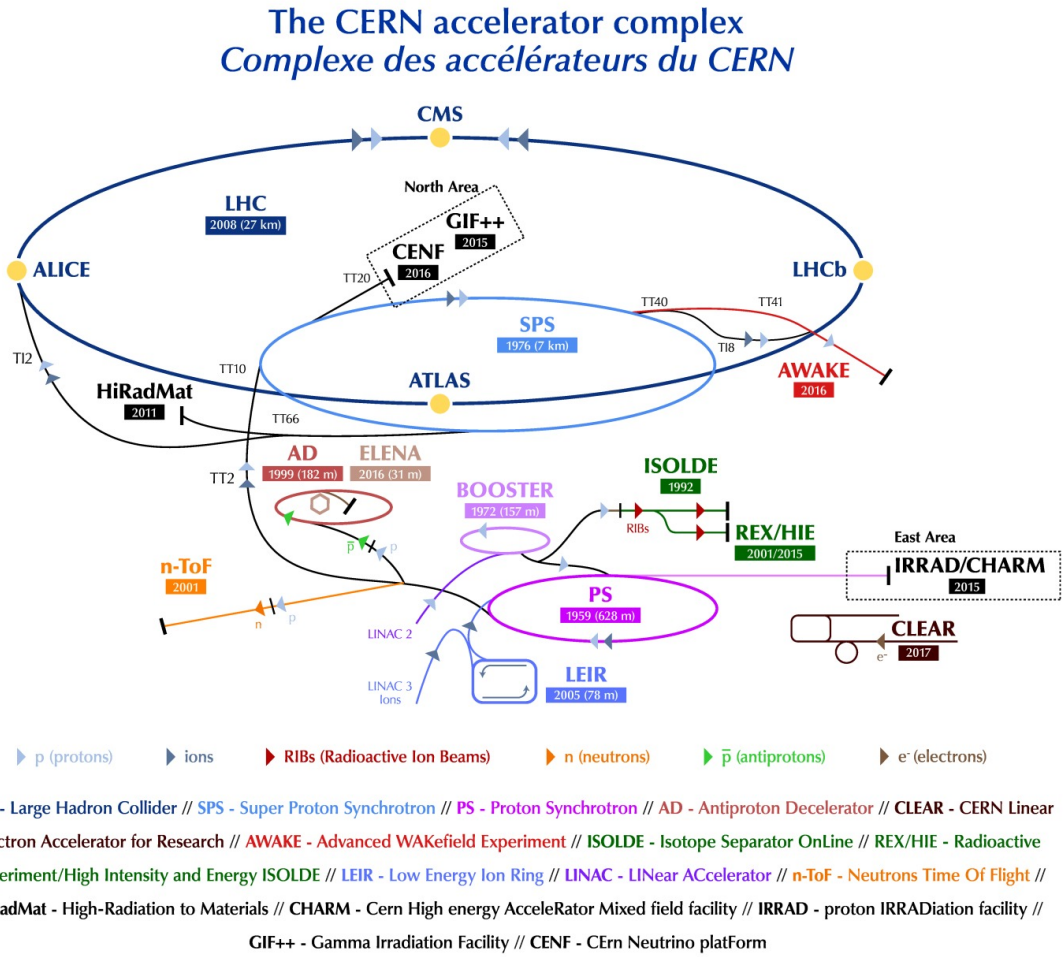


Figure 1.4 – Scheme of the CERN accelerators complex (from [99]).

As mentioned before, AE$\bar{g}$IS relies on the charge-exchange reaction between antiproton and positronium to form the antihydrogen to be used for the gravity test: to achieve this complex task, a state-of-the-art apparatus has been constructed [100]. It comprises a cylindrical cryostat containing two superconducting magnets of 5 T and 1 T respectively, both housing a series of collinear Penning-Malmberg traps. The traps in the 5 T region are optimised for trapping and cooling antiprotons, while the trap in the 1 T region is used to form antihydrogen. The axial confinement of charged particles is achieved by the more than 60 electrodes forming the traps (some of which can go as high as 15 kV) and, to minimise the losses of trapped antiprotons, an ultra-high vacuum of 10^{-13} mbar or better is maintained. Additionally to the electrodes, the manipulation of the anti-atoms is done with a set of q-switched pulsed lasers, relevant for the excitation of positronium to high Rydberg states ($n = 17 \div 32$), essential in order to efficiently produce antihydrogen. The apparatus is also equipped with a Micro-Channel Plate (MCP) detector at the end of the traps system, a two-layer scintillator fibre tracker for detecting the annihilation [101, 102], and plastic scintillators for detecting the annihilations [103]. The entry region of the antiproton beam from AD/ELENA also serves to bring positrons from the positron line inside the main apparatus, which are then turned to positronium atoms in a dedicated silicon nano-channel $e^+ \rightarrow$Ps converter [104, 105, 106]. The antihydrogen produced can be detected via its decay products using the scintillators and the

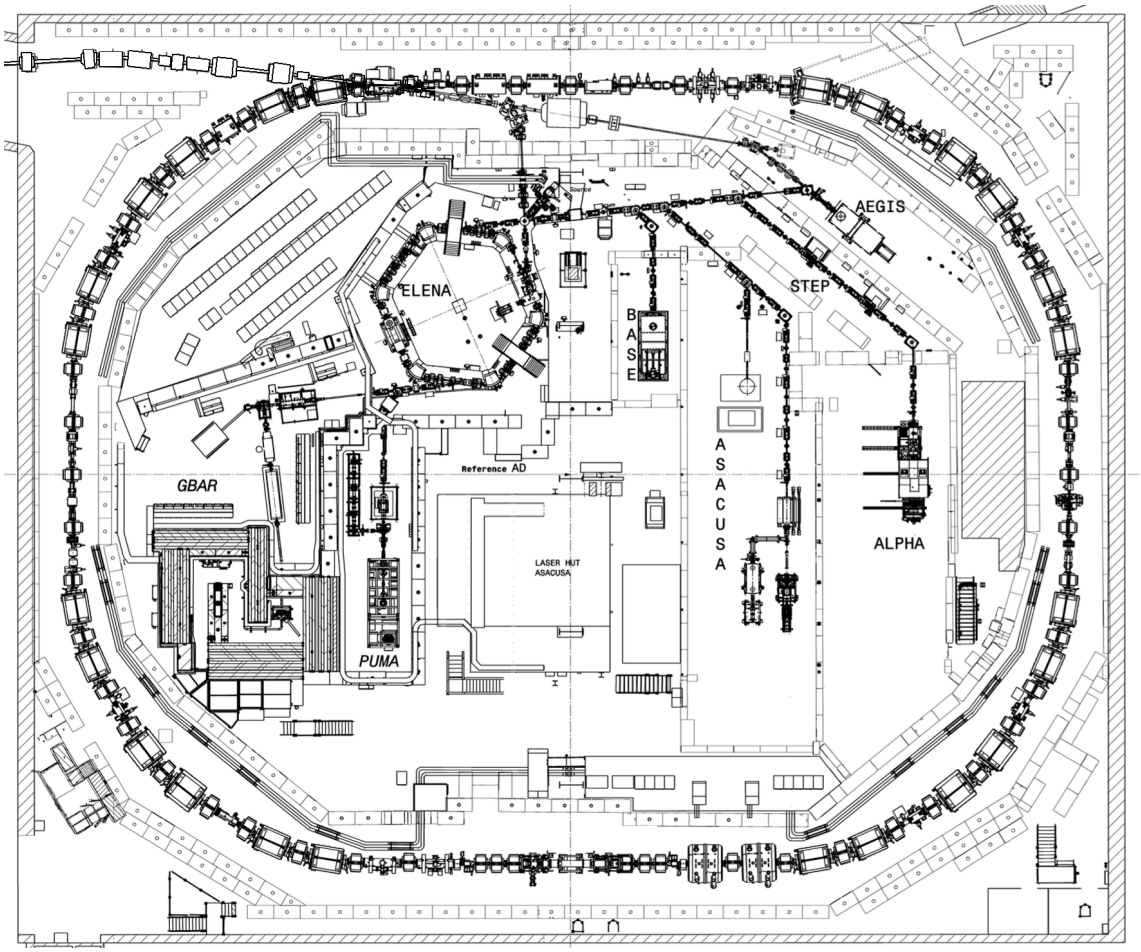


Figure 1.5 – Bird's-eye view of the AD hall (courtesy of CERN). On the perimeter is visible the Antimatter Decelerator, which pre-cools the $\bar{p}$ and feeds them to ELENA (center-left). Inside the decelerator, all the experiments of the Antimatter Factory are present: AE$\bar{g}$IS is on the top-right.

MCP, confronting the number of annihilations when the formation is favoured or suppressed by the procedure utilised.

The complexity of the apparatus gives the possibility to investigate multiple different phenomena: for example, the laser-cooling of positronium atoms, recently performed using the experience of positronium generation and the recently upgraded laser system (see *Positronium laser cooling* (A.1)). The installation of an additional trap for heavy ions generation is also ongoing, which will enable AE$\bar{g}$IS to perform studies on the formation processes of highly-charged antiprotonic heavy ions (see *Formation of trapped cold Highly Charged Ions* (A.2) for preliminary results).

A detailed description of the apparatus and the techniques used to generate the beam of neutral antihydrogen are presented in chapters *The AE$\bar{g}$IS Experiment* (2) and *Antihydrogen production methods* (4). In the following, the method to measure the vertical displacement of the $\bar{H}$ beam is explained, together with the design of the first gravimeter prototype.

1.3.2 AE$\bar{g}$IS $\bar{g}$ measurement technique

To determine the acceleration exerted on the antihydrogen atoms by the Earth's gravitational field, the AE$\bar{g}$IS collaboration intends to exploit the most simple and direct method available: measure the vertical drift caused by the parabolic motion of free-falling $\bar{H}$ (as schematised in figure 1.6).

Afterwards, the simple relation ($\tau = L/v_z$)

$$\Delta y = -\bar{g}\Delta\tau^2 \tag{1.3.1}$$

permits to extract the value of $\bar{g}$.

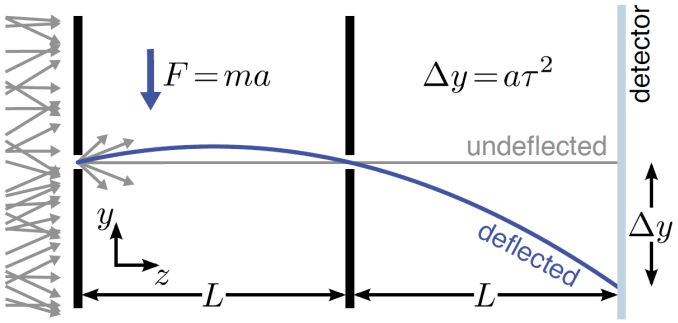


Figure 1.6 – The technique that AE$\bar{g}$IS will employ to measure gravity on antihydrogen, by leveraging the simplicity of the parabolic free-fall trajectory (from [107]).

To determine with precision the trajectories followed by the antiatoms, in the AE$\bar{g}$IS experiment a moiré deflectometer [108] will be used. The same device will be also employed as a Talbot-Lau interferometer for light, to have a reference for the displacement evaluation. In the subsequent sections, I will first present the detector functioning scheme and the first prototype design for AE$\bar{g}$IS, and afterwards, an estimation of the sensitivity reachable will be given.

1.3.2.1 The moiré deflectometer

The moiré deflectometer is a fully classical device, constituted of three identical material gratings equally spaced and a time-sensitive detector, where the first two are fixed, and the third is movable in the direction transversal to the particles beam (see figure 1.7). The first two gratings serve to select the particle trajectories out of an uncollimated source, creating, at distances multiple of their distance L, a fringe pattern similar to the one generated by interferometers. The application of a deviating force on the particles during their travel results in a transversal shift of the fringes: from the amount of the shift, the magnitude of the force can be deduced, using equation 1.3.1 (in case of a constant force). The third, movable gratings is therefore employed to scan over the fringe pattern, to resolve precisely their position by detecting the periodic modulation of the transmitted intensity.

In this manner, there is the need for a detector which has to be only precise in time resolution: the spatial resolution is compensated by the third grid, acting as a filter. The time resolution is critical since the determination of the time-of-flight (ToF) depends directly on it: for this precise reason, this measurement is possible in AE$\bar{g}$IS, given its unique pulsed production scheme.

It is also possible, in principle, not to use the third gratings, and employ a spatial and time-resolved detector in its place to image directly the fringe pattern and the ToF (as shown in figure 1.8): such a device has been developed in the context of the AE$\bar{g}$IS experiment [110], and it might be used in future refined version of the gravimeter.

1.3.2.2 The Talbot-Lau interferometer

To precisely establish the displacement of the fringes of particles with respect to the theoretical un-deflected trajectories, a reference is mandatory. A manner to create such a reference is to use the moiré deflectometer just described as a Talbot-Lau interferometer for light [107]. The Talbot-Lau interferometer has the characteristic of reproducing the pattern of the grids as light and shadow strips on the detector, provided that between the geometry of the interferometer and the wave light the following relation holds:

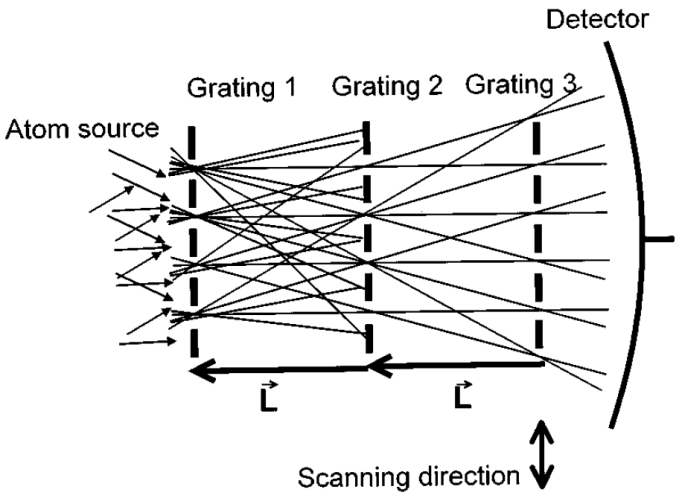


Figure 1.7 – Functioning principle of a moiré deflectometer. A series of three identical material gratings are placed at an equal distance L among them. The first two are used to select the trajectories out of an uncollimated particle beam, creating a fringe pattern at distances multiple of L; the third one is moved transversally, to scan over the position of the fringes, in order to detect their precise positioning leveraging the modulation of the transmitted intensity (from [109]).

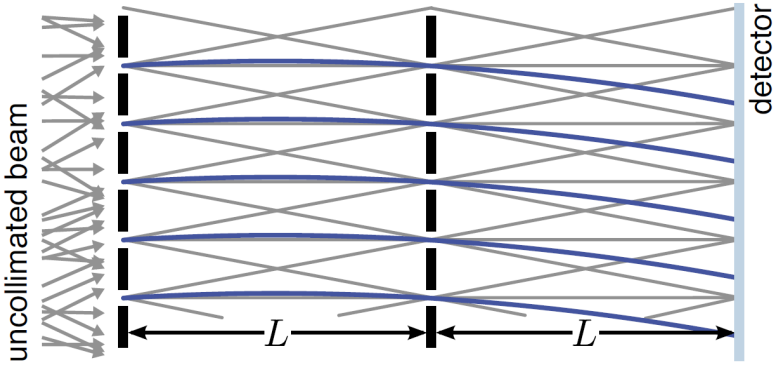


Figure 1.8 – Example of using a moiré deflectometer as a gravimeter, substituting the third grating with a position-and-time sensitive detector. The grey lines represent the trajectories that the particles would follow in the absence of the external force, while the blue lines show the parabolic trajectories of the particles subject to a constant force (from [107]).

$$L = kL_T = k\frac{d^2}{\lambda} \tag{1.3.2}$$

where d is the period of the gratings, λ is the light wavelength, L_T is the Talbot length of the interferometer, and k is a positive integer.

The Talbot-Lau interferometer exploits the Talbot effect [111], where the interference of coherent light passing through a grid forms an image of the grids itself every multiple of the Talbot length L_T: the effect is not present if the illumination is done with incoherent light (see figure 1.9).

The Talbot-Lau interferometer [112] builds on this effect, noticing that, even with incoherent light, the image of the grid is generated if a second grid (identical to the first) is positioned in the middle between the detector and the first grid, at a mutual distance equal of an integer multiple of the Talbot length (see figure 1.10). The fringes are recovered because the first grating creates spatial coherence in the passing light, which in turn generates an interference pattern after traversing the second grating.

This solution not only gives a reference for the displacement measurement, but it also eases the

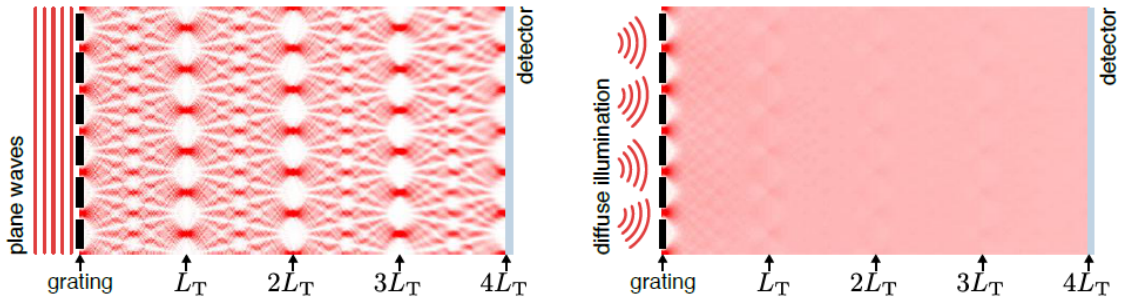


Figure 1.9 – The Talbot effect, where the coherent illumination of gratings produces an interference pattern mimicking the grid itself after a distance multiple of the Talbot length L_T (left). Conversely, incoherent illumination does not create the same effect (from [107]).

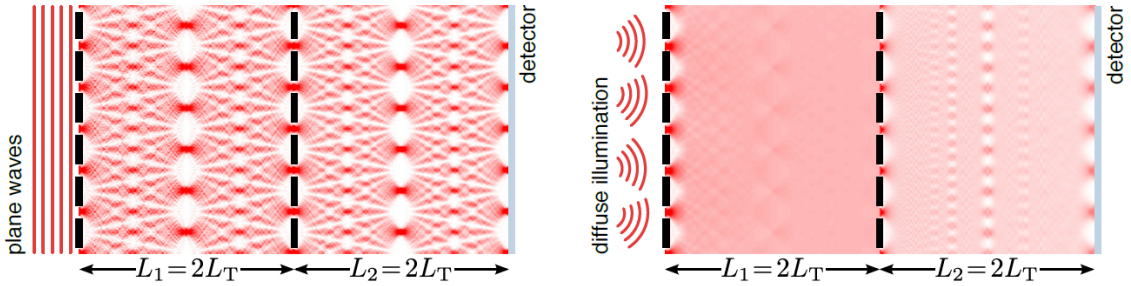


Figure 1.10 – The Talbot-Lau interferometer, where a second grating (identical to the first) is placed between the first gratings and the detector. In the case of coherent light (left), the behaviour is identical to the single-grid Talbot effect; but if incoherent light is used (right), the fringe pattern is recovered, because the first grid creates spatial coherence in the passing light that in turn generate the interference pattern traversing the second grid (from [107]).

positioning of the three gratings to be at the same distance L: in fact, by posing the middle grating on rails so as to be moved by a fine actuator, and by varying the frequency of the light used to illuminate the system, each of the two gratings inter-distance can be set to the same integer multiple of the Talbot length by maximising the visibility $\mathcal{V}$ of the Talbot fringes, defined as ($I_{\max/\min}$ is the intensity respectively in the centre of a fringe and in the centre of a shadow)

$$\mathcal{V} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \,. \tag{1.3.3}$$

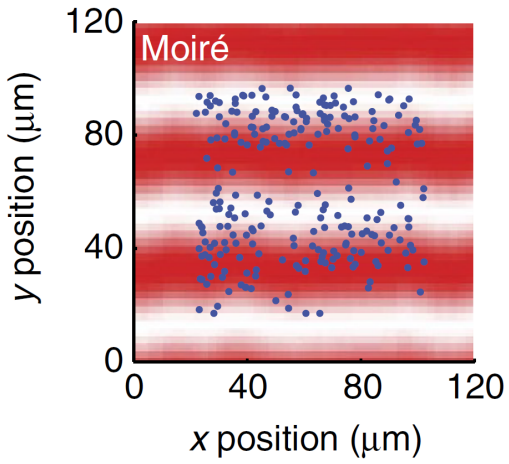
The grids longitudinal alignment precision that can be obtained in this manner accounts for approximately $L_T/10$: with a red laser diode of $\lambda \approx 650\,\text{nm}$ and a grid periodicity of 100 µm, it is circa 1.5 mm. If the grid periodicity is lowered to 20 µm, the alignment becomes possible up to $\sim 60\,\mu\text{m}$.

An example of the combination of the signals from both the antiprotons and the light traversing the gravimeter is given in figure 1.11a. From it, the two intensity curves can be derived, out of which the displacement can be extracted, as exemplified in figure 1.11b (incidentally, this was the first "interferometric" measurement ever done with antimatter).

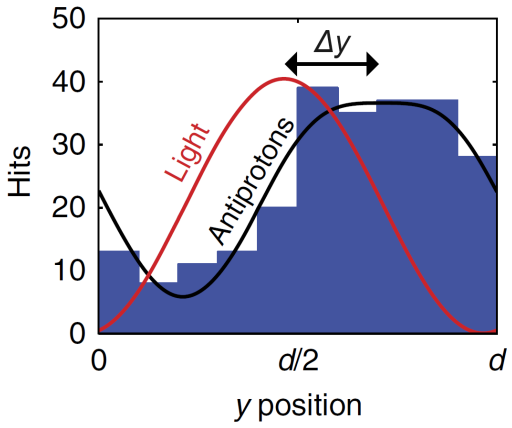
1.3.2.3 Gravimeter prototype design

In 2023, the design of the first AE$\bar{\text{g}}$IS gravimeter prototype has been produced, and it will be finalised in 2024, so to build, install and test the gravimeter in the antiprotons campaign of 2025.

The internal design is shown in figure 1.12. The three gratings (40 x 40 mm, 150 µm thick) have a periodicity of 100 µm and an opening 40 µm, for an open fraction of 40 %. They are mounted on a series of rails so that the mutual distance can be adjusted by cryo-actuators (while in cryogenic



(a) Example of a gravimeter signal.



(b) Example of displacement determination.

Figure 1.11 – On the left: example of a signal caused by the combination of light and antiprotons particles traversing the AEgIS gravimeter. On the right: an example of the two signal intensity curves extracted from the gravimeter signal. The distance between the light and the particle peaks corresponds to the displacement (both from [113]). This was the first "interferometric" measurement ever done with antimatter.

conditions) around a value of 500 mm. The entire assembly is mounted on two rotating frames, to be able to rotate the entire deflectometer of 90°, in order to perform the measurement also on the horizontal plane, to better verify the presence of eventual systematics.

These internal components will be housed in a vacuum tube that will be attached to the AEgIS main apparatus (explained in detail in *The Apparatus* (2.1)) as shown in figure 1.13. The light and the particles traversing the grids will be imaged with an MCP or a TimePix3 detector [110]; three scintillators will be placed outside the gravimeter housing, in correspondence with the three gratings, to monitor the annihilations of the antihydrogen atoms on the grids lines. A gate valve will separate the body of the main apparatus from the gravimeter, to guarantee the ability of quick interventions on the latter without the need to disrupt the main vacuum. Last, an *anti-gradient* coil will be placed at the end of the apparatus, to eliminate to the maximum extent the magnetic field gradient present in the external region, where the gravimeter will be installed. This minimises the acceleration given by the magnetic field on the flying antihydrogen atoms.

1.3.2.4 Expected sensitivity

To estimate the minimum acceleration that this first prototype can detect, we can use the following formula (taken from [114, 113]):

$$a_{min} = \frac{d}{2\pi \mathcal{V} \tau^2 \sqrt{N_{\text{det}}}} = \\ = \frac{d v_{\parallel}^2}{2\pi \mathcal{V} L^2 \sqrt{N_{\text{det}}}} \,. \tag{1.3.4}$$

In table 1.1 the meaning of the various parameters used in equation 1.3.4 is given, together with their values.

The grating visibility is estimated based on [114]: with the periodicity and the open fraction given in *Gravimeter prototype design* (1.3.2.3), the moiré interferometer is fully into the classical regime, and therefore approach the limit value of 80 %. The $\bar{\text{H}}$ axial velocity has been assessed at the end of chapter *Antihydrogen production methods* (4): being tunable by the launching potential

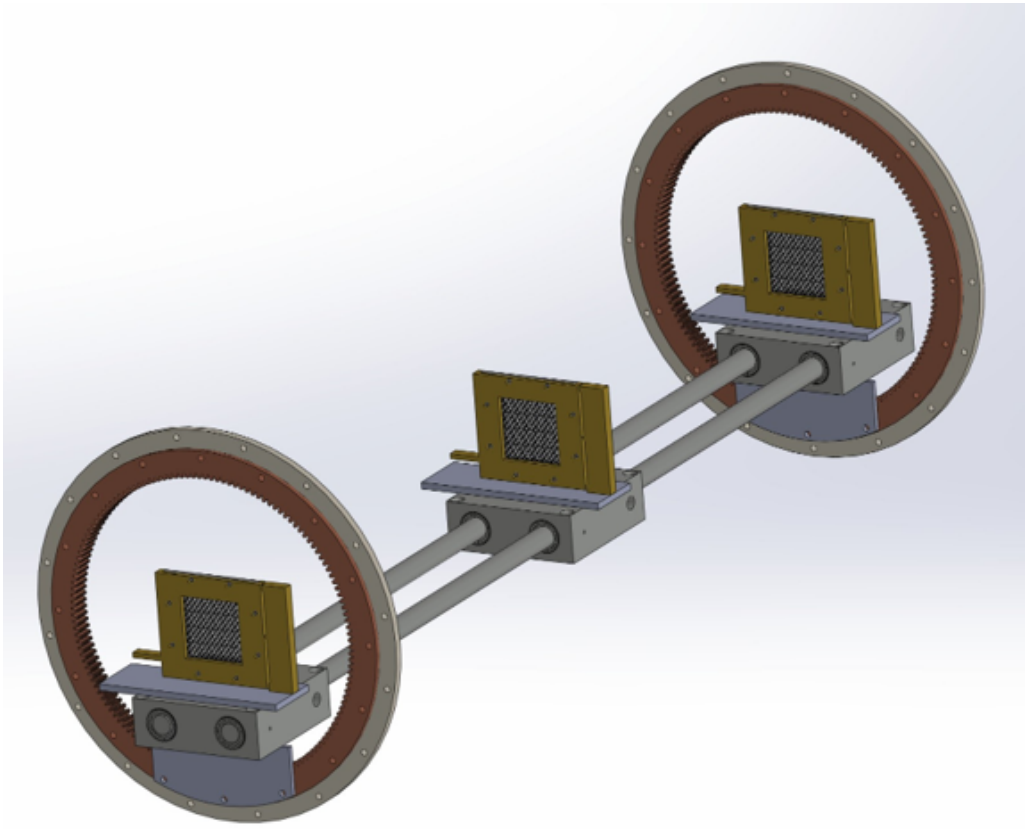


Figure 1.12 – Preliminary design of the internal main components of the AEḡIS gravimeter prototype. The three gratings (40 x 40 mm, 150 µm thick, periodicity d 100 µm, opening 40 µm) of the moiré deflectometer/Talbot-Lau interferometer are clearly visible, mounted on rails so that their distance may be varied, in cryogenic conditions, by cryo-actuators (L, the mutual distance between grids, is around 500 mm). The entire assembly is mounted on two rotating frames, so to be rotated of 90°, in order to perform the measurement also on the horizontal plane, for control (courtesy of Stefan Haider).

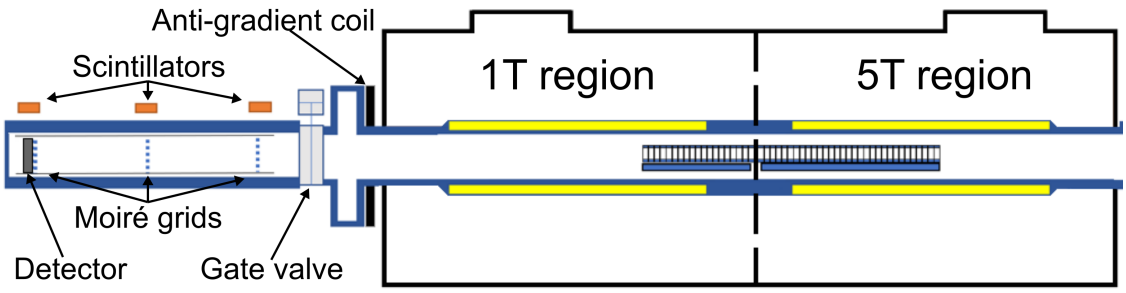


Figure 1.13 – Sketch showing the positioning of the gravimeter with respect to the AEḡIS apparatus.

Symbol	Meaning	Value	Uncertainty
d	gratings periodicity	100 µm	±1 µm
$\mathcal{V}$	pattern visibility	80 %	± 1%
L	distance between the gratings	500 mm	±1.5 µm
$v_{\parallel}$	$\bar{\mathrm{H}}$ axial velocity	1700 m/s	±100 m/s
N_{det}	number of $\bar{\mathrm{H}}$ detected	1000	32

Table 1.1 – Meaning and values of the parameters used to estimate the AEgIS gravimeter sensitivity, with their uncertainty.

configuration, lower values are achievable, the limit being dictated by the antihydrogen temperature. The total number of antihydrogen detected by the MCP at the end of the gravimeter are achievable in one antiproton campaign with a production rate of 10 $\bar{\mathrm{H}}$/minute (see *Expected new $\bar{H}$ production rate* (2.2.9)), given the grid open fraction of 40 %([11]).

Plugging the values given in table 1.1 in equation 1.3.4, the expected minimal acceleration detectable is

$$a_{\mathrm{min}} \approx 7 \ \frac{\mathrm{m}}{\mathrm{s}^2} \tag{1.3.5}$$

with an uncertainty of $1\,\mathrm{m/s^2}$.

In the future, the AEgIS target is to build such a gravimeter with 20(1) µm periodicity, 30 % open fraction, and therefore 70 ± 5 % visibility, with colder and slower antihydrogen, moving at 300(30) m/s, and a total number of $\bar{\mathrm{H}}$ detected equal to 50000, the sensitivity will increase up to

$$\delta a_{\mathrm{min}} \approx 7 \pm 2 \cdot 10^{-3} \ \frac{\mathrm{m}}{\mathrm{s}^2} \tag{1.3.6}$$

finally enabling AEgIS goal of a gravity measurement with a precision higher than 1 %.

[11] A simple $\frac{1}{r^2}$ law does not hold, for the $\bar{\mathrm{H}}$ flux, both because of its boosted nature, and because of the focusing effect given by the magnetic field in the second section of the 1 T section of the apparatus.

Chapter 2

The AEgIS Experiment

As briefly presented in *Where is all the antimatter?* (1), the AEgIS experiment aims at verifying the WEP on antimatter using a beam of antihydrogen atoms passing through the gratings of a moiré deflectometer. This task is far from trivial, hence it necessitates an extremely complex and varied setup, coupled to heterogeneous and precise operations. In the following chapter, the AEgIS experimental apparatus is described, giving an overview of its most important components and their functioning. Following, a bit of history of AEgIS is presented, where the transition from Phase 1 to Phase 2 is explained, together with the hardware and software upgrades that took place during CERN Long Shutdown 2 (LS2).

2.1 The Apparatus

The AEgIS apparatus (see figure 2.1) consists of a plethora of different subsystems linked together, which are used in different combinations depending on the specific experiment's needs (e.g. antihydrogen formation instead of positronium cooling). The principal piece of hardware is the main vacuum chamber, hosting the cryostat, the magnets system and both the 5 T and 1 T traps. Directly connected to it, but external, there is the positrons line, the two lasers systems, a multi-port deflection chamber for ions injection and extraction (that will host, from 2024, the heavy ions source, and multiple future experimental devices), and of course, all the detectors and electronics which were not bound to be inside the main chamber.

Before continuing, it is better to introduce the AEgIS coordinate system: it is a classical cartesian right-handed system, with the origin located at the geometrical centre of the experiment cryostat, which connects the 5 T and 1 T region of the main chamber. The z axis is the axis of the axis of the main chamber itself, pointing from 5 T to 1 T, and the y axis points vertically upwards.

In the following, a brief description of the main subsystems is given, to give the reader a better understanding of the functioning of the experiment.

2.1.1 Magnets

The magnet system of AEgIS is located inside the main vacuum chamber. It is divided into two different regions of magnetic field (which give the name of the two sections of the experiment): towards ELENA the magnetic field strength is 5 T, while in the other half is 1 T. Each magnetic field is generated by a solenoid made of superconducting Niobium-Titanium alloy (Nb:Ti), kept at 4.2 K by liquid helium. To ensure a field homogeneity of $1/10^5$, a sharper field termination at the end caps, and to ease the mechanical stresses induced by the magnetic field gradients, two sets of ten correction coils are present (see 2.2).

The double magnetic field configuration was designed in order to optimise the performances of two opposite necessities: a strong magnetic field in the $\bar{p}$ capture and cooling region, to maximise both

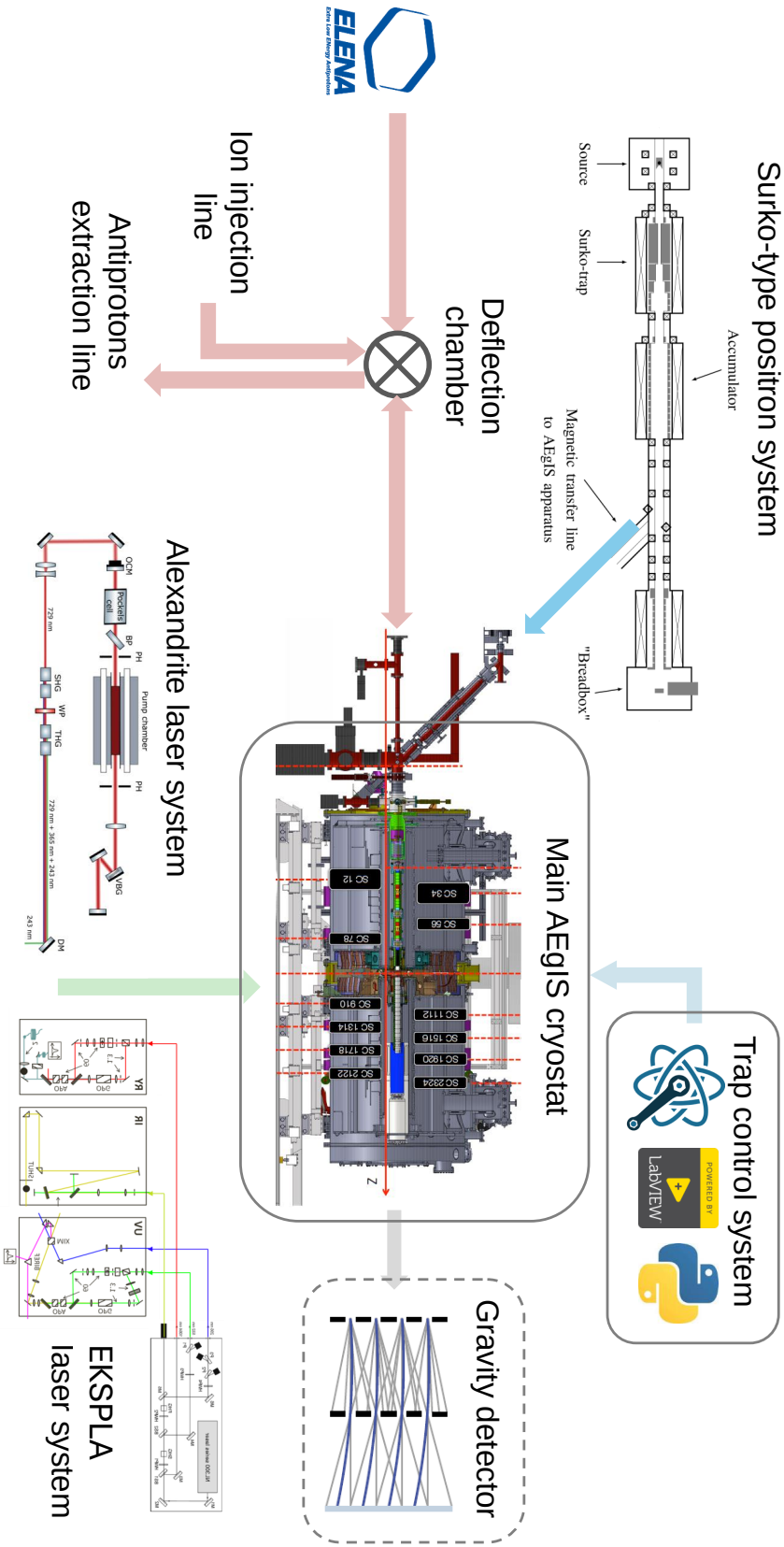


Figure 2.1 – Schematic drawing of the entire AEgIS experimental apparatus.

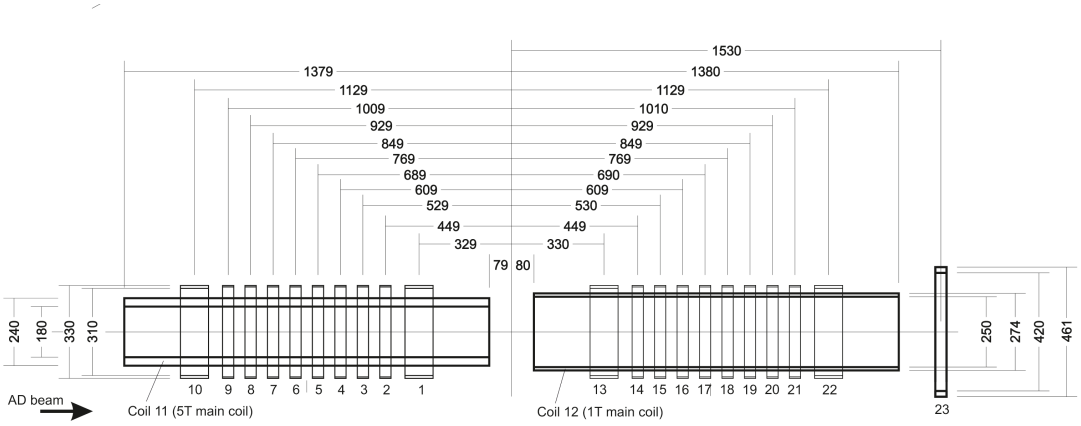


Figure 2.2 – Schematic drawing of the configuration of the magnets. On the bottom, the coil identification number is present; on the top and on the sides, measures and distances are indicated (in mm) (from [115]).

particle radial confinement and $\bar{p}$ sympathetic cooling with electrons, which is bound by the efficiency of electrons cooling via cyclotron radiation ($\propto B^2$: see *Sympathetic $\bar{p}$ cooling with e^-* (4.1.1)); and a smaller magnetic field in the $\bar{H}$ formation region, to limit the least possible the maximum Ps Rydberg level obtainable (bound by motional Stark effect induced field ionisation: $N \propto B^{-1/4}$: see *The new formation scheme* (2.2.1)) [14].

Moreover, in order to maximise the plasma lifetimes inside the various traps, the magnetic field needs to be as much as possible aligned with the traps' axis: this has been guaranteed by aligning, during the construction, the various magnets to the central region flange, with an accuracy of 300 µrad, and by aligning the traps with the geometrical centre of the magnet (see *1 T section* (2.1.5.2)).

In figure 2.3 a plot of the magnetic field strength on the axis of the magnet is given, obtained with a COMSOL simulation.

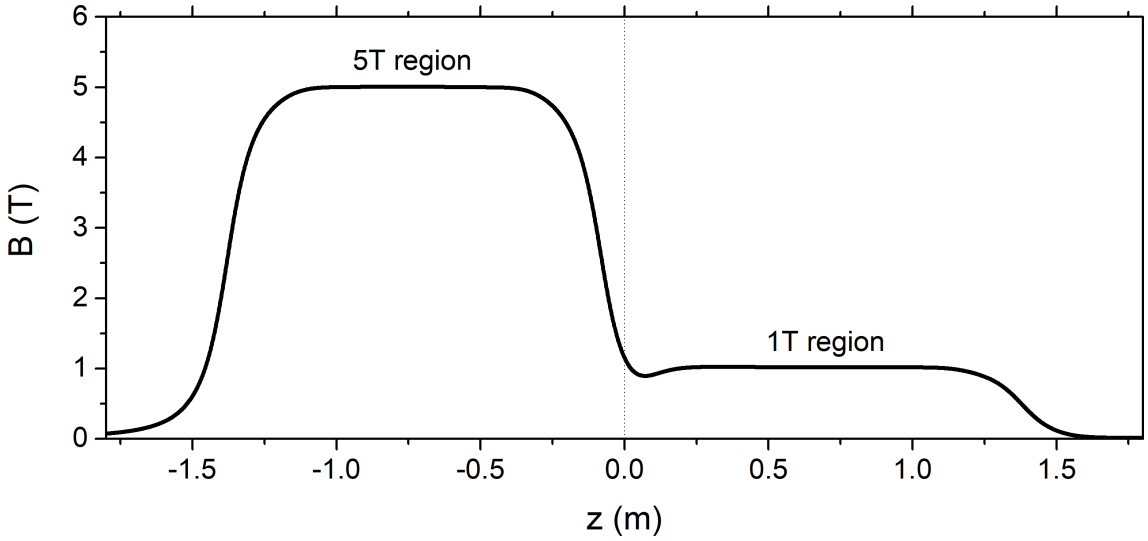


Figure 2.3 – Plot showing the magnetic field strength on the axis of the trap, obtained with a COMSOL simulation (courtesy of Ruggero Caravita).

2.1.2 Cryogenics

The Nb:Ti magnets just described are superconducting only below 9 K, so it is necessary to keep them at the temperature of liquid helium (4.2 K). To accomplish this, while minimising helium consumption, a complex cryogenics structure of nested coaxial cylindrical vessels was put in place: the Ultra-High Vacuum (UHV) region, hosting the traps, and the surrounding magnets are enclosed in the vessel containing the liquid helium, which in turn is enveloped by a part of the Outer Vacuum Chamber (OVC). This is further encircled by a vessel of liquid nitrogen (boiling point 77.4 K), which has both the use of shielding the interior from the heat radiation coming from outside, limiting the consumption of liquid helium, and accelerates the cool-down of the experiment after opening up; finally, the second OVC region is present[1], completed by a super-insulation material (multi-layer "space blanket") (see figure 2.4).

2.1.3 Vacuum

In order to store antiprotons for extensive periods of time (minutes or hours), a state-of-the-art ultra-high vacuum is needed, to minimise the annihilation of $\bar{p}$ with the residual gas, which is the ultimate limiting factor. To achieve these lifetimes, a vacuum of 10^{-13} mbar or better is necessary.

This vacuum level is maintained in AEgIS in the UHV region described in *Cryogenics* (2.1.2) by a combination of pumps: the internal surfaces of the UHV region, being kept at cryogenic temperature, acts as cryo-pump for all the residual gases except helium and hydrogen[2]; four big getter pumps (mainly for hydrogen); and an ionic pump for helium, installed outside the main vacuum chamber. The primary UHV pumping station is situated in the upstream section of the apparatus (referred to as the *Sun*, due to its distinctive shape) (see figure 2.4).

A second pumping station, connected to the 1 T part of the cryostat, is responsible for keeping the OVC vacuum levels (which have more relaxed requirements: typically the vacuum is around 10^{-7} mbar). All electrical, optical and thermal connections pass through two series of vacuum feed-throughs (external-to-OVC and OVC-to-UHV), with wires that also minimise the heat flux brought from outside.

The entire process of closing the main chamber, leak testing, pumping, thermal treatment ("baking"), and cooling, is extremely laborious and demands approximately one month to complete.

2.1.4 Antiprotons production in AD

Before introducing the AEgIS traps system, it is better to give an overview of how the antiprotons are generated in the Antimatter Factory, decelerated and finally supplied to the various experiments.

A bunch of $2 \cdot 10^{13}$ protons, with an energy of 26 GeV, is extracted from the Proton Synchrotron and shot through a 55 cm-long iridium target, where the highly energetic particles lose abruptly their energy, and proton-antiproton pair-production can occur. On average, every proton bunch produces approximately $5 \cdot 10^{7}$ antiprotons, which are separated via an electromagnetic selector, and injected at ~ 3.5 GeV in the AD. In the AD, the $\bar{p}$ are decelerated by a RF system down to 5.3 MeV. To compensate for the bunch explosion during the deceleration (especially in the longitudinal direction), two cooling mechanisms are employed: stochastic cooling, and electron cooling.

The first one employs a detector to measure the momentum fluctuations of the particles around the bunch average momentum and applies a corrective pulse to reduce, on average, the momentum spread. When applied multiple times, this technique can effectively reduce both the radial and the horizontal momentum spread of the antiproton bunch.

[1] Which is connected to the inner OVC region, so they are practically a unique zone to be evacuated.

[2] Which, conversely, is the most important to be removed, since it has the highest antiproton annihilation cross-section.

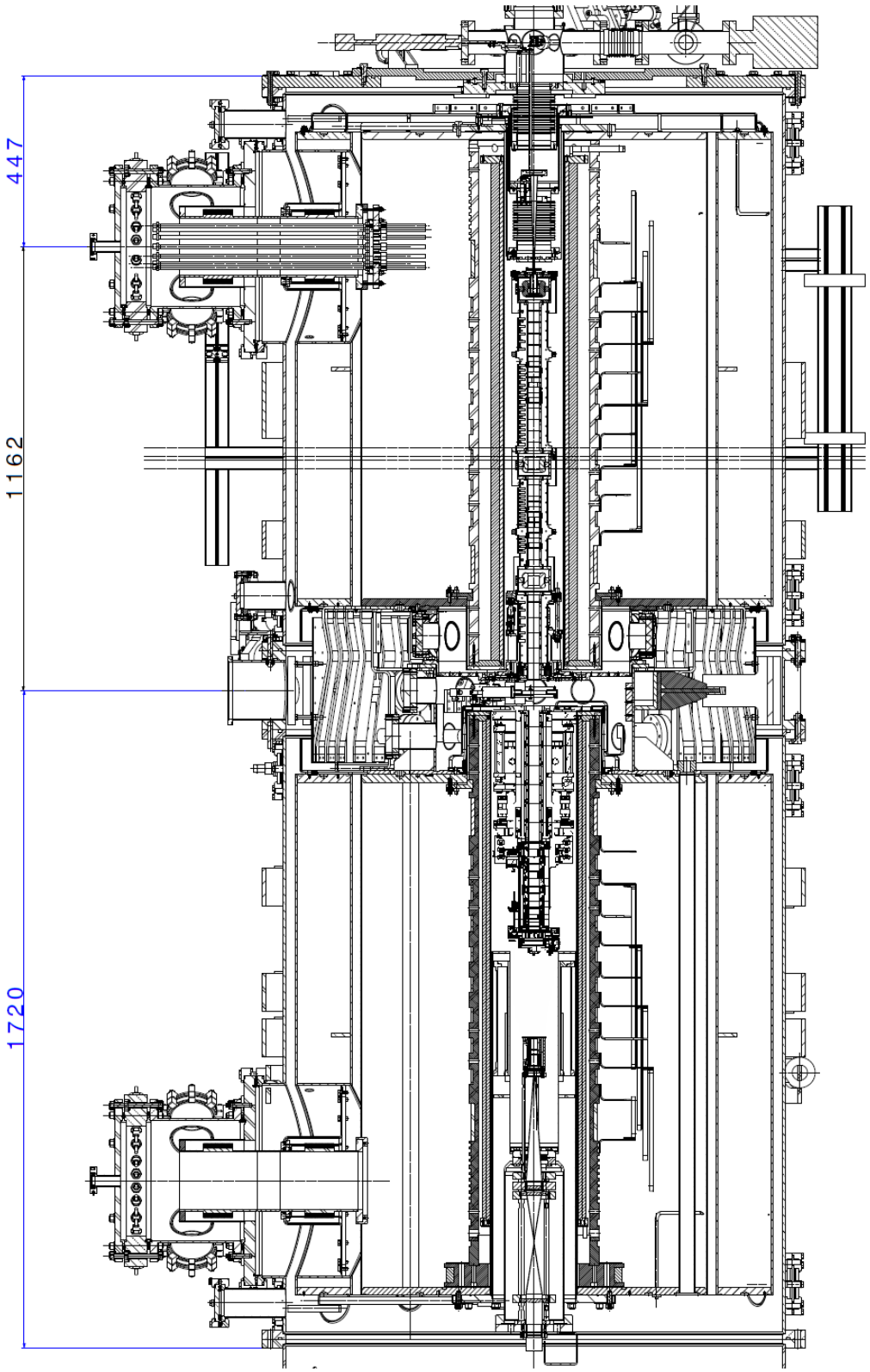


Figure 2.4 – A cut view of the AEgIS main cryostat. On the right there is the 5 T side, and on the left the 1 T one.

The second one consists of injecting a packet of electrons with the same average speed of the $\bar{p}$ bunch (and therefore, less kinetic energy), and letting them travel together for a while (in the AD, for 2.2 m). The antiprotons experience Coulomb scattering within the electron cloud, leading to momentum exchange with the electrons until thermodynamic equilibrium is achieved when both $\bar{p}$ and electrons share the same momentum. Given the significantly lower mass of electrons compared to antiprotons, the electrons must reach much higher velocities: consequently, (thermal) energy is transferred from the $\bar{p}$ to the electrons, effectively cooling the former down. The electrons are then easily extracted from the beam, and the operation is repeated. The functioning of AD is well detailed in [116, 117].

The entire process takes approximately 100 s: then, the antiprotons are injected into ELENA, a similar machine that further decelerates and cools the particles to 100 keV of kinetic energy, by making use of an RF drive and another electron cooler. It also splits the beam into four identical packets, to be able to serve each of them to a different experiment. After circa 110 s from the impingement of the proton on the target, the $\bar{p}$ are ready and are given simultaneously to the experiments in bunches of approximately 150 ns duration.

To synchronise experiments' behaviour with the AD and ELENA cycles, the decelerator complex gives the experiments a series of hardware triggers, listed in table 2.1. In AEgIS, the first trigger ("Injection in AD") is used to start the preparation of long actions, like loading in the FPGA's (Field Programmable Gate Array) kernel the operations to be performed on the traps, or configuring detectors for acquisition. The second trigger ("Injection in ELENA") is used to moderately long operations, e.g. power up the high-voltage power supply for the high-voltage electrodes, or trigger detectors that take seconds to go in acquisition mode. The third trigger ("ELENA Ejection −20 µs"), happening just 20 µs before the arrival of the antiprotons, is used to trigger fast detectors and to start the ns-precise series of operations that are subsequently fully synchronised with the $\bar{p}$ bunch arrival thanks to the last trigger ("ELENA Ejection").

Trigger	Time before $\bar{p}$ arrival
Injection in AD	~ 110 s
Injection in ELENA	~ 20 s
ELENA "Ejection −20 µs"	20 µs
ELENA Ejection	$0 \div 500$ ns (tunable)

Table 2.1 – Summary of the triggers given by the AD decelerator complex to the various experiments. On the right, the corresponding time before the actual arrival of antiprotons is displayed. The *ELENA Ejection* trigger (last line in the table) is tunable to the experiment demand in the range given in the table (e.g., for AEgIS ~ 350 ns are typically used).

The performances of the AD/ELENA system are very high: more than 80 % of the antiprotons created in the target are delivered to the experiments. As an example, in figure 2.5, the statistics of the $\bar{p}$ given to AEgIS in 2023 is shown. The average bunch size is $7.2 \cdot 10^6$ $\bar{p}$, with a standard deviation of $1.7 \cdot 10^6$ $\bar{p}$.

2.1.5 Traps' system

The AEgIS experiment traps system is very complex, due to the multitude of different (and sometimes concurrent) operations needed to be performed on electrons, antiprotons and positrons. Overall, it consists of a long (~ 162 cm) series of electrodes, all with a diameter of 30 mm (but different lengths), whose configuration divides the entire trap system in a series of smaller traps. In the following, a more detailed description is given.

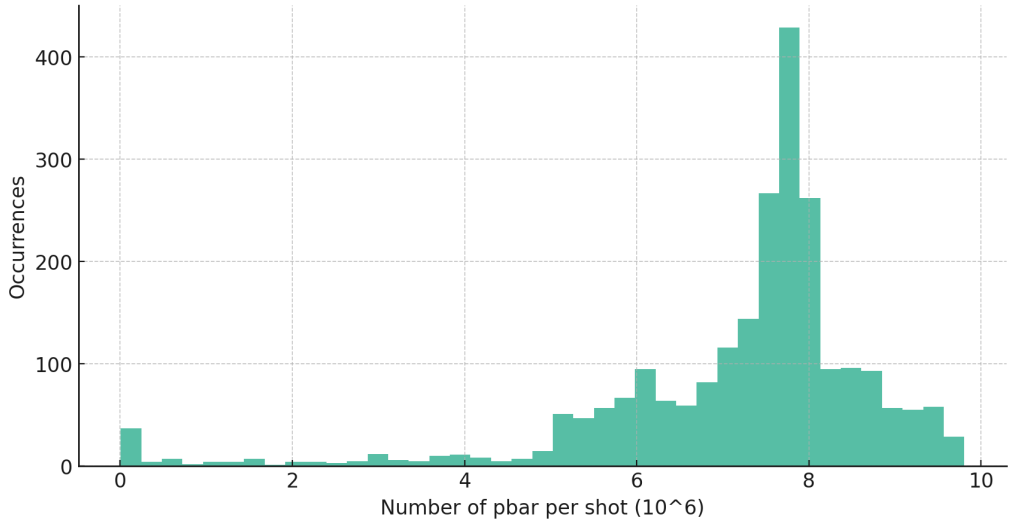


Figure 2.5 – The statistics of the number of antiprotons per bunch given to AEḡIS in 2023 from the AD/ELENA decelerator complex. The mean value is $7.2 \pm 1.7 \cdot 10^6$ (data from CERN Beam Performance Tracking).

2.1.5.1 5 T section

The 5 T section (see figure 2.6) is subdivided into two distinct regions, referred to as C and P traps[3], which are characterized by having, at their core, shorter electrodes (long 15 mm), designed to shape harmonic potentials. At their centre, electrodes sectorised in four 90°-parts are present, which are used to drive the rotating wall technique (RW: see *Plasma compression via Rotating Wall technique* (4.1.2)). Both the C and P traps are terminated by (in total) three high-voltage electrodes, which can reach potentials of up to 15 kV (HV1 to HV3, in the z direction): they are used to catch the antiprotons moderated by the degrader. Downstream of the C trap, a series of Transfer (T) electrodes establish connections between the 5 T and the 1 T section of the traps' system, guiding the particles as they navigate through the magnetic field gradient. At the centre of the transfer section (located at $z = 0$)[4], another fourfold sectorised electrode (B0, measuring 50 mm in length) is present, enabling particle steering by biasing differently the various sectors, so to create an electric field orthogonal to the magnetic field.

2.1.5.2 1 T section

The 1 T section of the traps' system is way more complex than its 5 T counterpart. First, the entire electrodes' complex is mounted on a ballast, that is fixed on the central region of the experiment: this has enabled the alignment of all the electrodes with respect to the axis of the trap at the CERN Metrology Lab with a deviation smaller than 11 µm. Afterwards, the entire trap ballast was aligned, on the cryostat, to the geometrical centre of the magnets, with a precision of 500 µm. Furthermore, the trap can be further aligned, in cryogenic conditions, thanks to two linear piezoelectric actuators positioned at the end of the electrodes' complex (one for the x direction, and one for the y one). Moreover, the ballast acts as a thermal sink for the trap system, ensuring cryogenic conditions are maintained.

The electrodes' complex (see figure 2.7) is characterised by a series of long transfer electrodes

[3]This nomenclature comes from "Catch" and "Positron", which were the original designation of these traps' functions: since, actually, they are general purpose traps, nowadays the original names have lost their meaning, but the letters have stuck.

[4]Technically, this electrode is located in the central region, so, strictly speaking, it is neither in the 5 T nor in the 1 T section of the experiment: for simplicity, sometimes it is included in the 1 T part of the CAD drawings, but it is not connected to the ballast holding the 1 T trap (see 2.1.5.2).

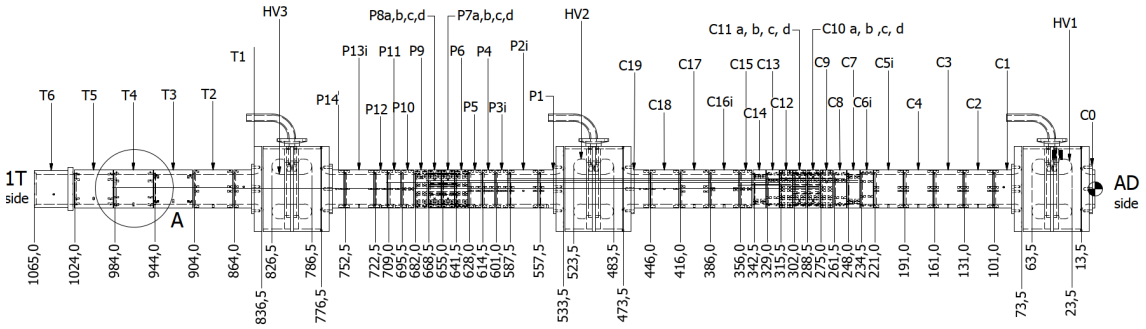


Figure 2.6 – Drawing of the electrodes assembly of the 5 T trap. The big electrodes (HV1/2/3) are the ones used for the capture; on the right, the C trap is present, while in the centre is located the P trap. On the left, the first part of the T transfer line takes place.

(similar to the ones terminating the 5 T trap, but 42 mm long), followed by two shorter transfer electrodes (B8 and B9, 15 mm long), and by B10, a fourfold sectorised electrode for rotating wall. After, three high-voltage electrodes are located, capable of sustaining up to 5 kV: they are used for the field-ionisation of Rydberg-Ps (in order to detect its correct formation), and they will serve for ions trapping, for the future antiprotonic atoms campaigns. Further downstream, the A trap (also called *formation trap*, since it is where the charge-exchange reaction takes place) is placed, made up of four 15 mm-long electrodes (the first one sectorised) followed by four 7.5 mm-long electrodes.

The progressive reduction of the electrode length, from right to left, is to have a more precise shaping of the big parabola (ranging from P10, in the 5 T section, all through A1) used for the ballistic production of antihydrogen (see *Antiprotons transfer* (4.2.4)): the potential gradient is greater at the extremities of the parabola, hence the necessity of shorter electrodes.

At the end of the trap, mounted on a circular actuator – so to be inserted or not, depending on the needs –, a special electrode is placed: it is the *target holder* (also called A0, for continuity with the rest of the A trap), which hosts the $e^+ \rightarrow$Ps target at its centre (see *$e^+ \rightarrow$Ps converter* (2.1.12)). A miniaturised heating system, coupled to a thermocouple, can bring the target up to 500 K: this is used to periodically regenerate the target by performing regular bake-outs.

After the target holder, mounted on another analogous actuator, the ionisation grids can be found: they are made up of two metal grids, each 0.2 mm thick, with a grid pitch of 2 mm and a wire thickness of 0.2 mm (so that the section of the wires is a square); the two grids are 2 mm apart. The grids can be mutually biased up to 2.5 kV. Their use is to ionise the antihydrogen passing through them while repelling the antiprotons: the $\bar{p}$ produced as an effect of the $\bar{H}$ ionisation can be imaged by the Micro-Channel Plate present at the end of 5 T region (see *MCP* (2.1.9.1)). This can be used to estimate the number of $\bar{H}$ produced forward-boosted (useful for the gravity measurement) with respect to the ones emitted in the other directions (see *Antihydrogen formation* (5.3)).

The ballast of the 1 T traps also holds the electronic board housing the custom-made radio-frequency (RF) filters that isolate the trap electrodes (except for the RW ones) from the noise carried from outside, minimising the heat transferred on the plasmas by the noise energy.

Last, but not least, this section houses also the series of prisms used for guiding the excitation lasers in and out without depositing any heat inside the cryostat, a manoeuvre doubling their passage onto the particles (so to increase the excitation efficiency); the little un-reflected portion of the light is carried outside the vacuum vessel by a fibre bundle (see *Laser positioning with fibres bundle* (2.1.9.6)), where is imaged by a series of camera, enabling the steering of the lasers with live feedback.

All these items are illustrated in figure 2.8, and a photo of the entire 1 T trap system is portrayed in figure 2.9.

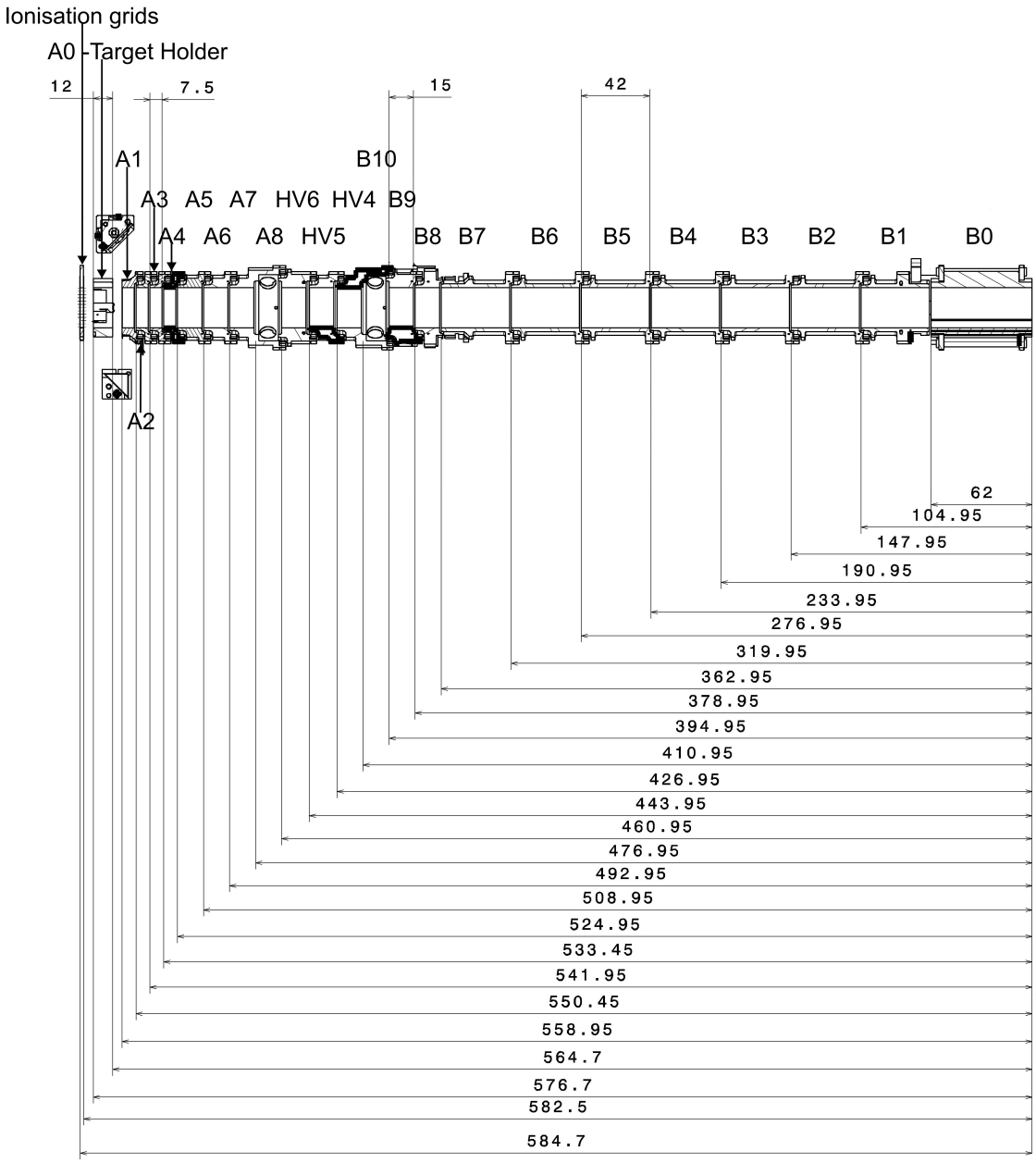


Figure 2.7 – Drawing of the electrodes assembly of the 1 T trap.

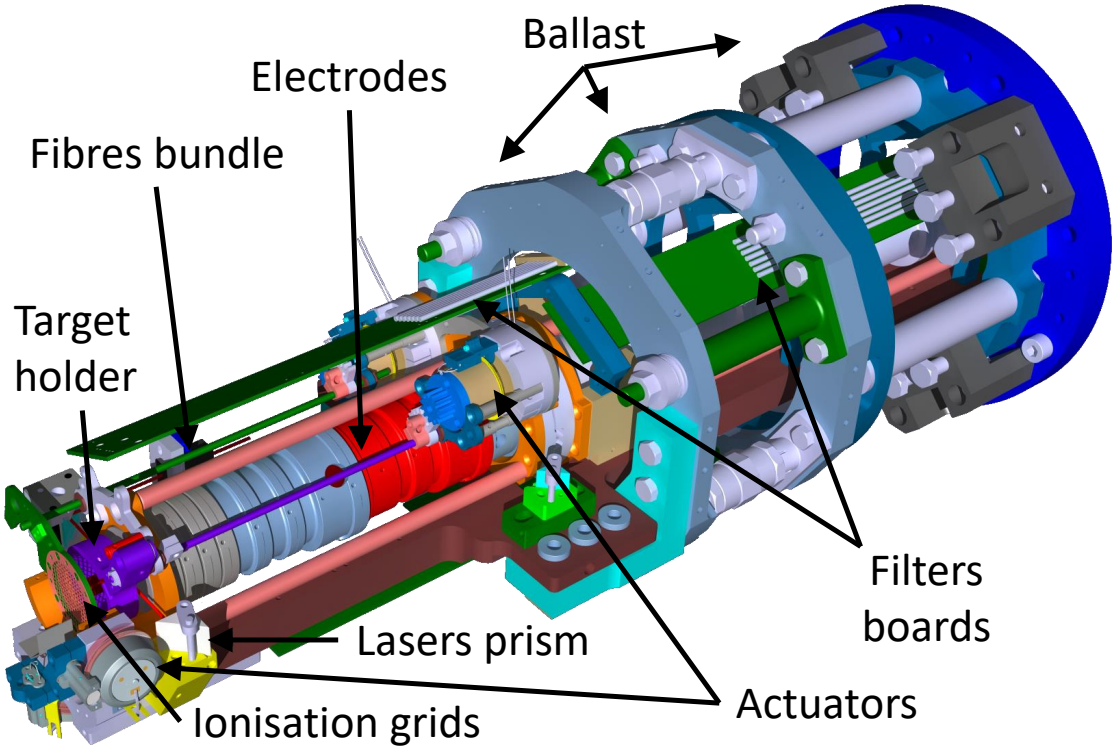


Figure 2.8 – CAD drawing of the new trap (courtesy of Stefan Haider). All the components presented in the text (target holder, ionisation grids, actuators, laser prisms, fibres bundle, electrodes, and ballast) are visible here and labelled.

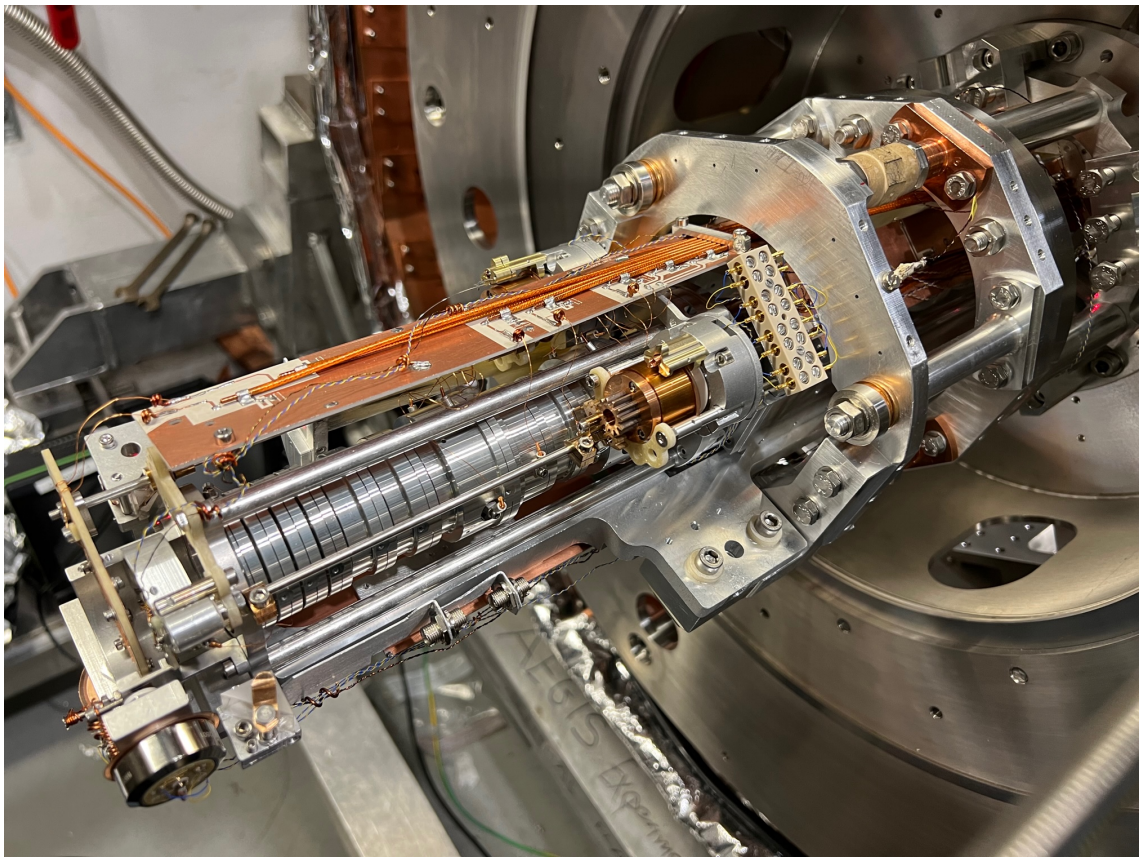


Figure 2.9 – A photo of the 1 T trap (2022, courtesy of Stefan Haider).

2.1.6 Lasers

In AEgIS there are two laser systems: the "EKSPLA", used to excite the Positronium to a high Rydberg state and/or photo-ionise it, and the "ALEX", specifically designed to perform Ps cooling.

2.1.6.1 EKSPLA

The formation of antihydrogen through the charge-exchange reaction is heavily dependent on the excitation level of the Ps ($\sigma \propto n_{Ps}^4$): therefore, the ability to reliably bring Ps to high Rydberg numbers is vital for the experiment. In AEgIS, this is obtained by a two-photon resonant excitation (see *Scheme of the Ps excitation and/or ionisation processes used in AEgIS.* (2.10)), from $1^3S \rightarrow 3^3P$ and from $3^3P \rightarrow$ Rydberg/continuum: for this, the "EKSPLA" laser system (called like this from the pump company's name) has been set up, as comprehensively detailed in references [118, 119].

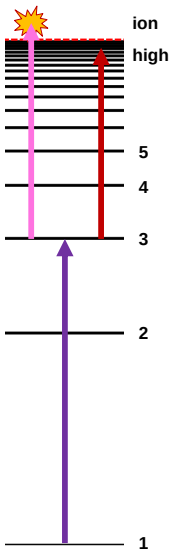


Figure 2.10 – Scheme of the Ps excitation and/or ionisation processes used in AEgIS.

In short, a single Nd:YAG (neodymium-doped yttrium aluminium garnet) pump, operated in Q-switched mode (i.e. a pulsed mode which can produce short, but very powerful, laser pulses), generates infrared (IR) pulses at 1064 nm with a rate of 10 Hz; the pump further produces a laser at 532 nm (second harmonic) and another at 266 nm (fourth harmonic). These three harmonics are combined in different stages of Optical Parametric Generation (OPG), Optical Parametric Amplification (OPA) and Sum Frequency Generation (SFG), to create ultraviolet pulses (205 nm) and mid-infrared pulses (1650 nm to 1715 nm). The first laser pulse is utilised for the $1^3S \rightarrow 3^3P$ Ps transition, while the second one is designed for the transitions to Rydberg states, the specific level being selected by the tunable wavelength. Additionally, the first harmonic is directly employed for the selective photo-ionisation of Positronium atoms from the 3^3P state. A schematic of the EKSPLA laser system can be seen in figure 2.11.

Under ideal conditions, the first Ps transition should take place with an efficiency close to 100 %: in practice, because of a non-complete spectral overlap of the lasers with the Ps cloud (both geometrical and in the velocity phase space, because of limited bandwidth), an excitation rate of approximately 15 % was observed with a single laser passage [120], increasing to about 28-30 % with a double laser passage.

2.1.6.2 Alexandrite Laser ("Alex")

During CERN LS2, a new laser was installed into the newly built *Lighthouse* [5] (a dedicated laser hut with accurate temperature and environmental stabilization), based on an alexandrite crystal rod (hence its nickname, "Alex"). This laser is built to provide, with high accuracy, a 729 nm wavelength, 100 ns-long laser pulse with a broad bandwidth: a Volume Bragg Grating (VBG) is used to select the wavelength while ensuring a broad spectral bandwidth of about 100 GHz. Moreover, thanks to the VBG, it is possible to fine-tune the wavelength to a precision of 10 pm at 729 nm, matching the requirements for distinguishing the Doppler profile of the positronium cloud via the $1^3S \rightarrow 2^3P$ transition. This transition, crucial for positronium laser cooling, is achieved by converting Alex's fundamental wavelength to its third harmonic ($\lambda = 243$ nm), using non-linear tripling crystals. A photo of the setup is visible in figure 2.12a, and the graphs showing the two harmonics generated are displayed in figure 2.12b.

Both the non-linear crystals and the VBG are mounted on motorized stages: once the system is initially manually calibrated, the target wavelength and corresponding phase matching angles of the crystals can be set in complete automation, thanks to TALOS (see *TALOS, the framework* (3.4)).

The Alex laser produces pulses around 100 ns in duration, which are ideal for positronium laser

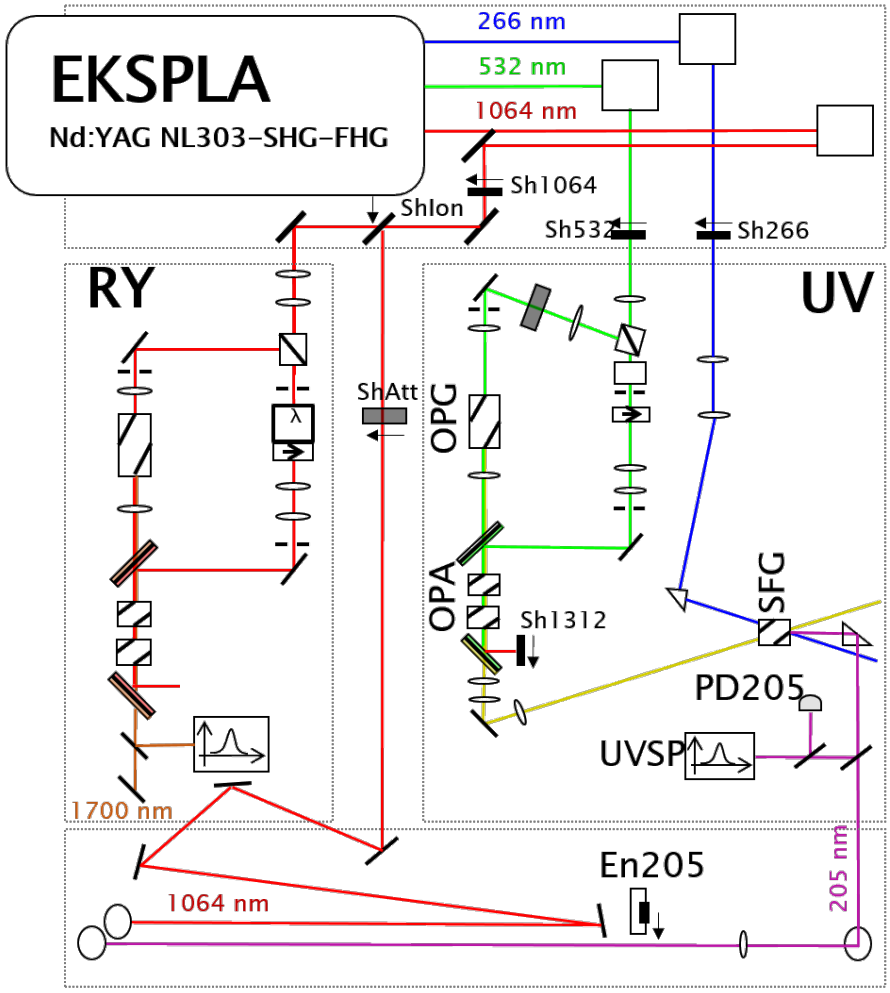
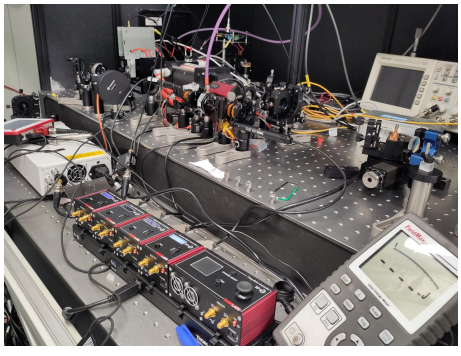
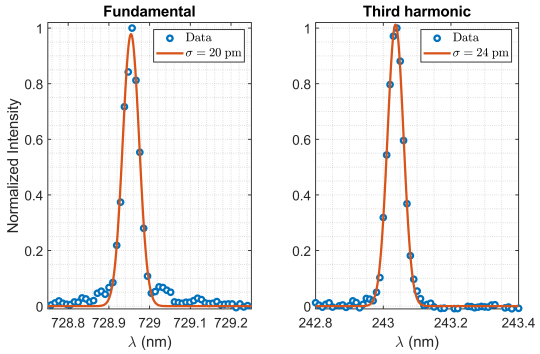


Figure 2.11 – Scheme of the EKSPLA laser system (courtesy of Ruggero Caravita).



(a) The Alexandrite "Alex" laser setup.



(b) "Alex" fundamental and third harmonic spectra.

Figure 2.12 – On the left, a picture of the alexandrite "Alex" laser setup. On the right, typical fundamental and third harmonic spectra, together with the Gaussian fits estimating the spectral bandwidths.

cooling given the 142 ns annihilation lifetime of positronium in its triplet ground state. The system can generate up to 3 mJ at 243 nm, leading to peak power of approximately 10 kW.

Furthermore, being the laser extremely sensitive to temperature and humidity variations, to achieve the aforementioned wavelength precision and stability, and especially time determination of the pulses, it has been paramount to develop an active feedback loop for exactly calibrating the firing time before each shot. This endeavour, rendered possible by the new control system, is well described in *In-Run autonomous parameter stabilisation* (5.1.4).

Moreover, a high-voltage electronic switch was installed to generate a sharp falling edge and precisely control the termination of the interaction with positronium, by abruptly lowering the voltage on the Q-switch of the oscillator to end the laser emission. With this addition, the Alex holds exciting prospects not only for manipulation and spectroscopy of positronium (in particular, Doppler cooling) but also as a source for applications in hydrogen, antihydrogen and muonium spectroscopy.

2.1.7 Degrader(s)

As mentioned before (see *Antiprotons production in AD* (2.1.4)), ELENA gives the antiprotons in bunches with a kinetic energy of 100 keV: since this energy is too high for the experiments to stop $\bar{p}$ directly, a common escamotage is to use an energy degrader[5]. It consists of a thin foil that the particles need to traverse to enter the experiment: by doing so, they lose some energy by multiple scattering with the atoms of the degrader, so the particles exit it with a lower axial energy. The energy reduction comes at a price: a part of the antiprotons annihilates on the atoms of the degrader, and the beam spreads out at the exit since the collisions redistribute part of the axial momentum to radial momentum.

In AE$\bar{g}$IS, there is a main degrader installed at the entrance of the 5 T region (so to minimise the impact of the beam spreading: the magnetic field will tend to contain it). It is constituted by two Mylar foils, one with a thickness of 900 nm, and one with a thickness of 500 nm: the foil towards the traps is completely aluminised (with 10 nm of aluminium) one the side looking the traps, so to defy possible charging up of the insulator material and to be usable as a Faraday Cup (see figure 2.13a), while the foil facing the ELENA pipe is aluminised in four 90°-sectors, separated by a cross of 1 mm thickness, so to use it as a beam position monitor (see figure 2.13b). A dedicated GEANT4 simulation showed that such a degrader should produce a population of $\bar{p}$ with energies ranging from $\sim$ 3 keV to 20 keV, with a transmission above 96 % (see figure 2.14). Moreover, the simulation also analysed the effect of adding a supplementary layer of degrading material, in the form of a thin Parylene N foil, for thicknesses ranging from 100 nm to 500 nm. A foil of 200 nm was then used, during the antiprotons campaign of 2022, to perform the record of efficiency of catching antiprotons in the trap (see *Efficient antiprotons capture* (5.2)).

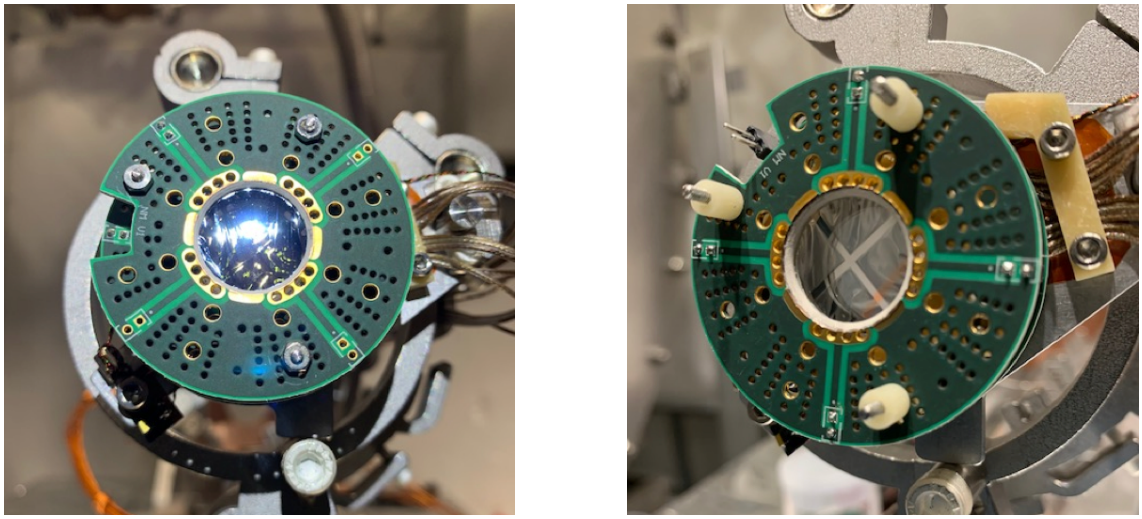
2.1.8 Electron Gun

The entire cooling and compression of the $\bar{p}$ plasma in AE$\bar{g}$IS relies on the presence of an electron plasma. In fact, the energy is mainly lost by bremsstrahlung, which power, for a circularly moving charged particle, is:

$$P_{a\perp v} = \frac{q^2 a^2 \gamma^4}{6\pi\varepsilon_0 c^3} \tag{2.1.1}$$

where $\gamma = \frac{1}{\sqrt{1-(v/c)^2}} = \frac{E}{mc^2}$. The electron therefore emits $(m_p/m_e)^4 \approx 10^{13}$ times the power of an antiproton: the speed of the cooling of the $\bar{p}$ is limited, then, by the power transferred by the collision of the $\bar{p}$ with the electrons (a better treatment is given in 4.1.1).

[5]Another technique worth mentioning is the one used by GBAR[96]: by using an inverted pulsed tube, they manage to electromagnetically decelerate antiprotons down to $\sim$ 200 eV, with negligible losses, at the price of increasing apparatus complexity.



(a) Degrader - Faraday Cup

(b) Degrader - fourfold beam position monitor

Figure 2.13 – Two photos showing the two foils constituting the degrader for the incoming $\bar{p}$ beam. On the left: the foil facing the trap, completely aluminised, to work as a Faraday Cup. On the right: the foil facing the ELENA tube, aluminised in four sectors, to be used as a beam position monitor (from [5]).

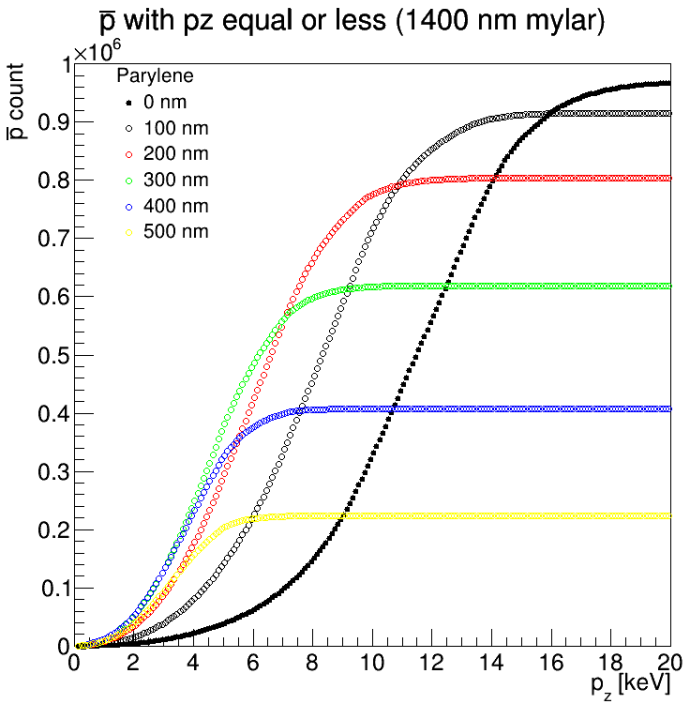


Figure 2.14 – GEANT4 simulation showing the number of antiprotons transmitted by a 1400 nm Mylar foil as a function of their horizontal momentum. Moreover, a study was performed to see the effect of adding an additional Parylene foil of thickness ranging from 100 nm to 500 nm (from [5]).

Therefore, a fundamental part of the experiment is the source of electrons, simply called *electron gun* (or e-gun for short). It involves a thin metal filament which emits electrons in virtue of the thermionic effect: the filament is heated up via an electric current up to a temperature where the average electron's thermal energy surpasses the extraction work function from the metal, and they are therefore extracted and accelerated thanks to an electric field.

In practice, this is realised by encasing in an insulated aluminium bar a thin barium oxide filament[6], which is powered up by a custom-made power supply (current-regulated), supplying 1130(1) mA to reach the correct temperature, and biasing the filament with respect to the electron gun case up to −200 V (normally operated around −40 V). The entire electron gun is fixed to a retractable high-vacuum linear actuator in the Sun, to be inserted for electron loading and removed to let $\bar{p}$ (and positrons) pass.

In regular operation, the electron gun generates an e^- current around 10 µA: the magnitude of the current produced is normally assessed measuring it directly on the front face of the 1T MCP, operated as a Faraday Cup (i.e., without any HV present on the plates to prevent damaging the detector, see *Faraday Caps* (2.1.9.5)), in a given time.

2.1.9 Detectors

The AEgIS experiment comprises a great variety of different detectors, each with a specific scope and function, in order to fulfil its scientific goals. Not all of them are used in every measure, but each is crucial in its specific function. Here the most important ones are presented.

2.1.9.1 MCP

A Micro-Channel Plate (MCP) is a particle detector composed of a resistive material (typically glass), featuring an array of closely spaced micrometre-scale channels, slightly tilted (∼ 8 %) with respect to the normal of the detector surface [121]. When low-energy particles impinge against its surface, they interact with its channel walls, inducing the emission of secondary electrons, which in turn are accelerated by an electric field applied between the two faces of the plate. This leads to a cascade effect since each electron generates more secondary electrons through collisions with the channel walls. The electron multiplication process results in substantial signal amplification, often exceeding thousands; the gain can be further multiplied (to go into the million range) by stacking two plates, rotated 180°, in what is called the "chevron" configuration. The electrons so produced are then converted into photons by means of a phosphor screen, which in turn are imaged by a digital camera. MCPs are very versatile detectors, where low-intensity signals, rapid response times, and high spatial resolution are needed.

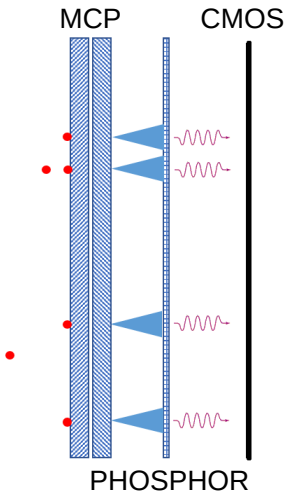


Figure 2.15 – Scheme of functioning of a MCP sensor, in double chevron configuration, coupled with a phosphor screen and acquired with a CMOS camera.

In AEgIS, after about 30 cm from the end of the formation trap, an MCP from Hamamatsu (F2223), with a fast phosphor screen (P67), is placed, which is imaged by an ORCA-Flash4.0 V2 CMOS camera with very high quantum efficiency (87 %); moreover, the phosphor screen is also connected to the NI 5152 oscilloscope (see *NI 5152 - Fast digitiser* (2.1.9.4)), enabling it to work as a Faraday Cup (FC), detecting the charge deposition in time. This sensor is crucial for a variety of applications, from the steering of the ELENA beam on the axis of the trap, to the characterisation of electrons, positrons and antiprotons plasma (shape, position, total charge), to

[6] For redundancy reasons, two independent filaments are present, so that in case of a failure (e.g., melting), operations can be continued without the need to open the cryostat.

time-of-flight analyses, to the detection of Ps and $\bar{H}$ formation (in combination with the ionisation grid (see *1 T section* (2.1.5.2))).

2.1.9.2 Scintillators

Around the exterior of the cryostat 12 external scintillator slabs are installed, semi-encircling the apparatus' body, which serve as detectors for antiproton and positron annihilations (see figure 2.17). Each scintillator is curved, to perfectly fit around the cryostat, with a section measuring approximately 100 mm × 20 mm (with the exception of the two slabs closer to the degrader, which are 200 mm × 20 mm), and made of polystyrene doped with POPOP wavelength shifter. Each individual slab is optically linked to two shielded photomultiplier tubes (PMT), one per each side, so to eliminate the background by counting only the event which generates a signal in each PMT in a windows smaller then 50 ns (see figure 2.16). Each scintillator possesses a coverage area encompassing approximately 3 % of the solid angle concerning an annihilation event occurring at their centre (∼ 6 % of the bigger ones).

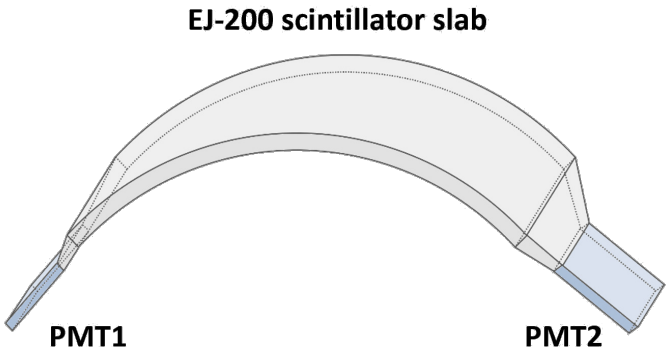


Figure 2.16 – The figure shows one of the scintillator slabs of AEgIS, coupled to two PMTs (from [103]).

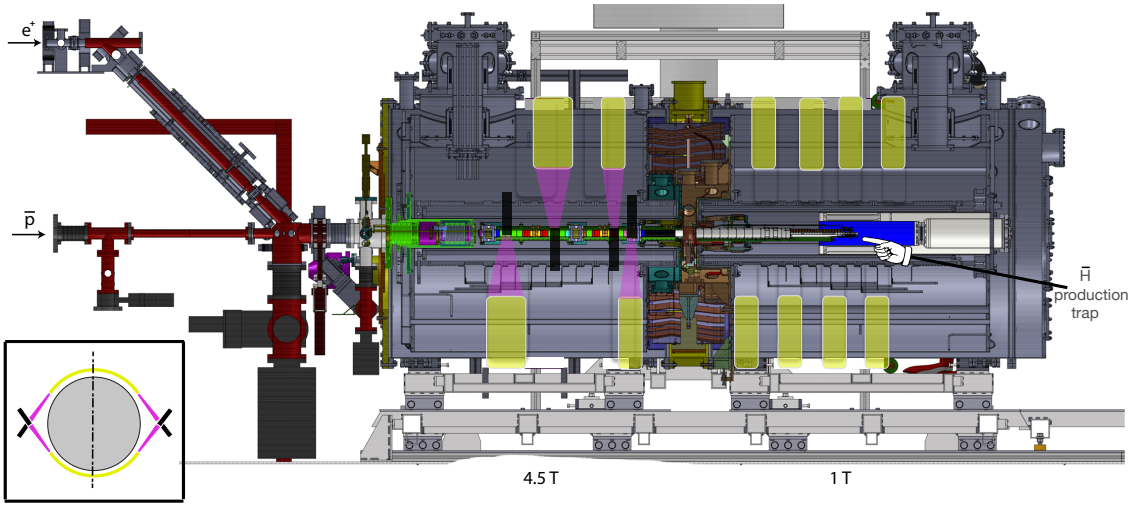


Figure 2.17 – A drawing of the position of the various scintillator detectors present on the main body of the experiment (lateral view). The scintillating slabs are shown in yellow; the light guides (in purple) and the PMTs (in black) are shown only for the first four slabs for the sake of clarity. In the small inset on the bottom-left, a schematic front view is presented, with the same colour coding (from [103]).

Their principle of operation is the following: the material with which the scintillators are made

has the peculiarity of emitting photons when traversed by ionising radiation. A charged particle with enough energy, in fact, excites or ionises atoms along its track inside the scintillator, while a gamma ray releases energy via the photoelectric effect or Compton scattering. The excited atoms then relax emitting photons, which travel inside the plastic until they reach one of the two ends, where they are detected by a PMT.

The latter is constituted of an evacuated tube, with a photocathode at the detecting end, a collecting anode at the opposite end, and multiple dynodes. The photon impinging the photocathode creates an electron thanks to the photoelectric effect; this electron travels toward the first dynode, because of the potential bias between the two: upon collision with the dynode, multiple secondary electrons are emitted. The dynodes are geometrically configured, and electrically progressively more biased so that an exponential electrons cascade is generated starting with the first electron: at the end, on the collection anode, several millions of electrons are deposited, resulting in a very clear signal (see figure 2.18 for a schematic representation).

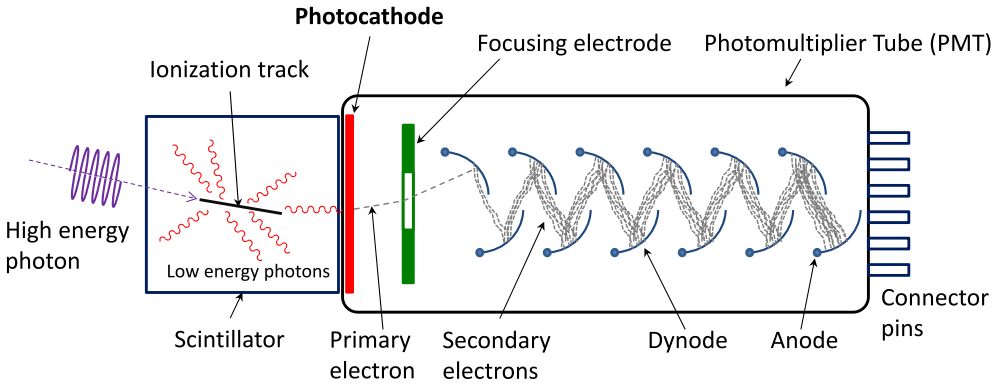


Figure 2.18 – Schematic view of a photomultiplier coupled to a scintillator, illustrating detection of gamma rays (from [122]).

2.1.9.3 FACT

Another detector dedicated to the observation of antihydrogen production is the Fast Annihilation Cryogenic Tracker (FACT) [101, 123]. It consists of two concentrical *superlayers* covered with 794 tiny scintillating fibres (1 mm in diameter), each superlayer housing two layers of fibres. The two superlayers are in turn concentrical to the formation trap and in thermal contact with the liquid helium bath (see figure 2.19).

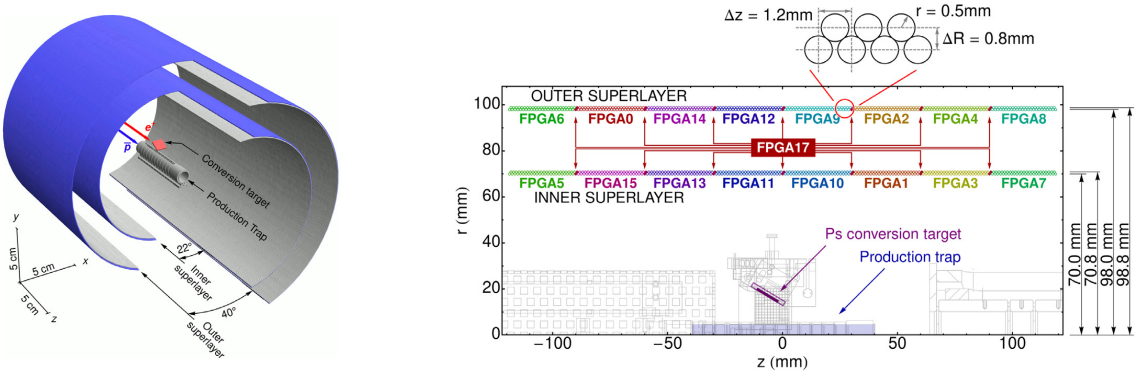
The fibres are read by multiple silicon photomultipliers from Hamamatsu, located inside the outer vacuum layer of the cryostat. The readout electronic, instead, is constituted by a total of 17 FPGAs that are fixed on the outer flange of the 1 T region, and they are connected to the various PMTs via vacuum feed-throughs.

FACT is not very sensitive to the typical 511 keV produced by a positron annihilation[7], but the pions produced from an antiproton annihilation can be clearly detected, each creating on average 30-50 photons inside the fibres.

The double (double) layer structure has been specifically designed to function as a vertex locator for $\bar{p}$ annihilation events. Additionally to the geometrical sensitivity to annihilation position, the longitudinal one can be inferred from the temporal coincidence of events in the two superlayers.

FACT has been the subject of a software rework during CERN LS2, in order to integrate it with the new control system (see *CIRCUS, the new AEgIS autonomous control system* (2.1.15)): despite

[7] FACT sensitivity to low energy gamma-ray is in the order of 10^{-3}: while it can see a positron bunch implantation event, it cannot distinguish the single annihilation arising from antihydrogen annihilation.



(a) Cutaway layout of FACT.

(b) Longitudinal projection of FACT.

Figure 2.19 – On the left: cutaway layout of FACT, encircling the (old: see*A bit of history: from Phase 1 to Phase 2* (2.2)) antihydrogen production trap, visible in the centre. The four layers of scintillating fibres are depicted in blue and grey. On the right: longitudianl projection of FACT, displaying the assembly of the fibres in the two superlayers, and their division among the 17 FPGAs readout boards (from [123]).

the efforts placed, the integration has not been finalised in time for the antihydrogen campaign of 2023, so its capabilities have not been leveraged in the determination of $\bar{H}$ formation.

2.1.9.4 PXI

One core component of the electronics of AEgIS is the PXI[8] system from NI[9]: it is a high-performance modular instrumentation platform developed for test, measurement, and control systems. Strictly speaking, the AEgIS PXI is not a detector: its main controller acts also as a PC (with Microsoft Windows 10), and among its modules, there are also signal-generating ones. For the purpose of this thesis, though, it fits better into the *Detector* section, since some acquisition modules are crucial in the running of the experiment.

Thanks to the high-speed performance and compact size of the CompactPCI standard, the PXI system is specifically engineered to excel in demanding environments where efficient, rapid, and precise data acquisition and analysis are paramount. Its architecture is based on a main controller connected via a high-bandwidth backplane to multiple modules, with functions ranging from multifunction data acquisition and dynamic signal acquisition to RF signal generation and analysis, thus offering great flexibility and scalability.

Furthermore, its integration with LabVIEW™ rendered the interface of the various modules with the new AEgIS control system (see *The CIRCUS* (3)) straightforward, stable and reliable.

In the following, the most salient characteristics of the most used modules are given. A picture of the AEgIS PXI crate is visible in figure 2.20.

NI 5152 - Fast digitiser The NI 5152 is a high-performance digitiser/oscilloscope module that embodies versatility in high-frequency signal acquisition and analysis. This module has an analogue bandwidth of 300 MHz, with a maximum sampling rate of 2 GS/s (giga-samples per second); coupled with its 8-bit Analog-to-Digital Converter (ADC) converter, it can acquire rapid transient events with high fidelity. The amplitude range can go up to ±10 V.

Coupled with the NI 2546, which is a four-channel multiplexer, up to four different sources can be sampled (serially) during the same measurement.

[8] *PXI* is the acronym for the communication standard used by the modules of the system: Peripheral Component Interconnect (PCI) eXtensions for Instrumentation

[9] Formerly *National Instruments Corp.*.

Figure 2.20 – A picture of the PXI system present in the AEgIS experiment.

NI 6133 - Slow multichannel digitiser The NI 6133 is another digitiser module, characterised by a more precise 14-bits ADC, at the price of a slower maximum sampling rate of 2.5 MS/s. Nevertheless, its standout feature is the capability of sampling simultaneously from its 8 input channels, ensuring synchronised data acquisition, a critical factor in many advanced analysis and monitoring tasks.

NI 6682 - Timing Clock The NI 6682 is used in AEgIS to provide the 10 MHz clock to all devices, synchronising everything to its time base. This module has an integrated GPS receiver, enabling it to maintain time accuracy to within ±50 ns, with a stability of $\pm 5 \cdot 10^{-11}$, ensuring exceptionally reliable frequency reference signals.

2.1.9.5 Faraday Caps

A Faraday Cap (FC) detector, for low energy particles, is simply a metallic foil used as a charge collector. It is an excellent method to destructively measure the total population of electrons in a cloud. The number of electrons deposited onto its surface can simply be calculated by integrating the current flowing over time[10]: considering the gain of the amplifier G, and the resistance R of the sampling circuit, we have:

$$N_e = \frac{1}{e}\int_{t_0}^{t_1} I(t)dt = \frac{1}{eRG}\int_{t_0}^{t_1} V(t)dt \tag{2.1.2}$$

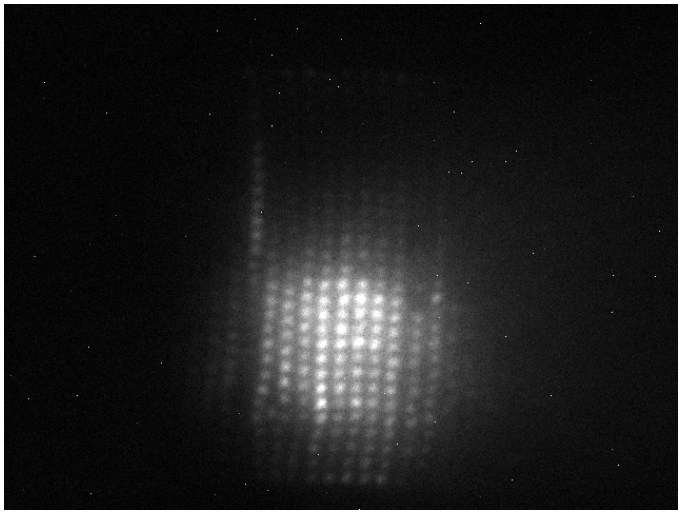
In AEgIS, multiple Faraday Caps are present, as already mentioned in *Degrader(s)* (2.1.7) and in *MCP* (2.1.9.1), so to be able to reconstruct the total population of electronic plasmas by dumping it in either z direction.

2.1.9.6 Laser positioning with fibres bundle

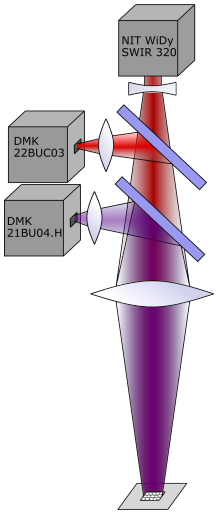
To align the laser beams with the positronium cloud inside the AEgIS apparatus, a fibre bundle was constructed using 1300 Thorlabs FG200AEA multi-mode silica fibres, each with a diameter of 200 µm and 1 m long. One end of the bundle is positioned within the UHV section of the experiment, and it transports part of the light emitted by the lasers (a series of prisms inject into the fibres only a small percentage of the total laser light) outside the cryostat. There, a complex imaging system (see figure 2.21b) separates the three laser frequencies and images them separately: in fact, the light is first focused by a 2" fused silica lens, and then a custom dichroic mirror diverts

[10]This method works perfectly below 10^8 e^-: above, the digitiser starts to saturate and the estimate loses its validity. For plasma bigger than 10^8 e^-, it is better to fit the curve of the potential of the discharge of the FC with $\tau = RC$, and from the peak one can have $Q = CV_{peak}$.

only the 205 nm beam to the first camera (ImagingSource DMK 21BU04.H, naphthalene coated). Subsequently, a Thorlabs DMLP1180 dichroic mirror distinguishes the 1064 nm and 1650 nm beams. The former is captured by an ImagingSource DMK 22BUC03 camera, while the latter is imaged by a NIT WiDy SWIR 320 camera. Furthermore, a light-tight box encases the system to reduce light noise. The dichroic mirrors effectively separate the wavelengths, eliminating cross-talk between the cameras. In figure 2.21a is shown an example of the image generated acquiring the 205 nm laser: to be noted the array of fibres placed inverted, so to be used as a reference during the horizontal alignment of the beam.



(a) Example of a laser bundle image.



(b) Laser bundle imaging system

Figure 2.21 – On the left: an example image of 205 nm beam imaged by the fibre bundle. To be noted the the flipped array of fibres, used as a reference for a precise laser alignment. On the right: pictorial scheme of the setup for the fibre bundle imaging (both from [7]).

2.1.10 Positron system

For the formation of antihydrogen, positrons are as essential as antiprotons: therefore, in AEgIS a reliable source of these particles is present. It consists of a positron source made of ^{22}Na, which had an initial activity of 1.02 GBq on March 8, 2018, translating approximately to 220 MBq during the antihydrogen campaign of autumn 2023. This sodium is deposited on a tantalum plate, acting as a positron reflector, and sealed into a steel-titanium capsule, so as to be exposed safely to ultra-high vacuum. This capsule is then positioned inside a copper block with a cone-shaped aperture, maintained at 7 K by a helium cryocooler. In this way, filling the area with neon gas, it is possible to grow a positron moderator made of neon ice, which naturally forms on the cold surface (an annealing at 9.3 K for several minutes is required for moderator optimal performances). The emitted positrons are further accelerated to 8 eV, and filtered by a magnetic field selector: the most energetic positrons cannot properly follow the field lines and annihilate against a tungsten barrier. The so-formed positron flux is continuously trapped in the *Surko Trap*. The overall system has an efficiency of $\sim 2.5 \cdot 10^{-3}$, translating into approximately $5 \cdot 10^5$ e^+/s flux into the *Surko Trap*.

The *Surko Trap* consists of a buffer-gas Surko trap [124]. It is composed of a series of electrodes (25.4 mm in diameter) placed inside a vacuum tube, where a 0.07 mT magnetic field ensures radial confinement. The electrodes are used to perform particle manipulations and assure axial confinement; a special electrode in the centre enables perform rotating-wall on the e^+ cloud.

The positrons flying inside the trap are cooled by inelastic collisions with a nitrogen and SF_6 gas mixture (in accordance with [125]). A pressure of $8 \cdot 10^{-4}$ mbar was found to be sufficient to

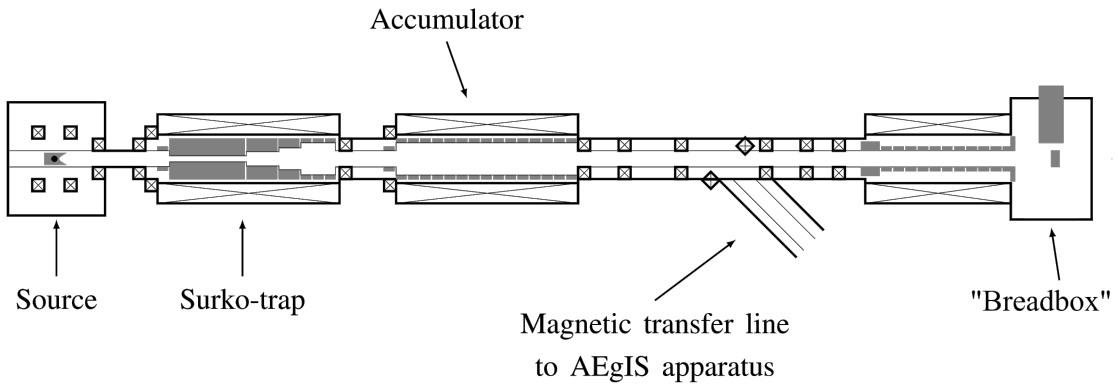


Figure 2.22 – Schematic drawing of the positron system of AEgIS. On the left the ^{22}Na source emits positrons, which are moderated by the solid neon and caught in the *Surko Trap* (in the middle). The trap periodically discharges its contents into the *Accumulator* (on the right), which can then transfer the positrons on demand, either to the "Breadbox" or towards the main apparatus, thanks to the *Transfer Line*.

slow the positrons enough so as not to be able to escape the trap when coming back after being reflected by the last electrode: in this way, a continuous incoming flux can be sustained during the entire filling phase, lasting 150 ms. Afterwards, the trap is closed by raising the initial electrode, the floor is raised, and after waiting 2 ms for further cooling, the end electrode is opened to transfer the e^+ cloud to the *Accumulator*. The cycle is then repeated indefinitely, steadily amassing cold positrons in the *Accumulator*.

The *Accumulator* is a Penning trap, constituted by multiple electrodes (so to be able to shape a smoother electric harmonic potential) and a lower gas pressure ($\sim 10^{-8}$ mbar of N_2), in order to minimise positrons annihilation. Also here, one electrode is segmented to perform RW. When requested, the electrodes in the trap can be quickly pulsed to form a potential gradient going from 466 V to 0 V, so to bunch and eject the stored particles: this technique minimises the axial spread of the outgoing cloud.

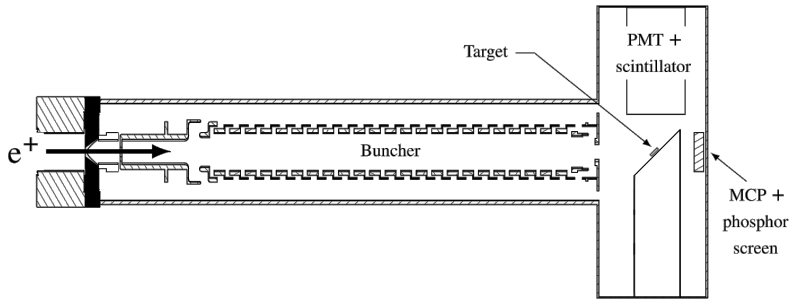
Thanks to an electromagnetic diverter, the so formed e^+ cloud can then be either sent to the "Breadbox" (see following section) for positronium experiment or to the AEgIS main apparatus. In the second case, the cloud passes inside the *Transfer Line*, which guides the positrons toward an accelerator, bringing the energy of the particles up to 2.5 keV: in this way, the high velocity impedes axial spread during the travel of the positrons towards their destination.

2.1.11 The "Breadbox"

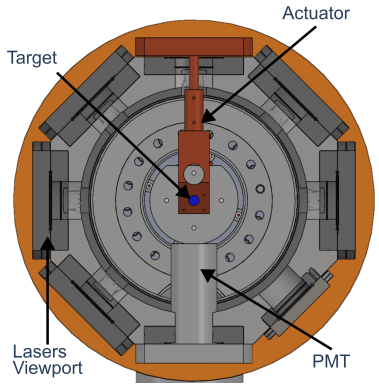
The AEgIS experiment has a dedicated environment to perform positronium physics studies, called the "Breadbox", which is situated on the same line of the positron system (see figure 2.22).

It is constituted by an electrostatic buncher, which extracts the positrons from the *Accumulator* (see *Positron system* (2.1.10)), accelerates them to the required energy (3.3 keV) to be implanted into the Ps target, and focuses the e^+ cloud into it. Being fully electrostatic, it maintains the magnetic-field-free environment necessary to prevent the level splitting of the positronium and perform precise spectroscopic measurements. A test chamber is placed right after this accelerating line, where the Ps target is placed, in the centre, plus a MCP detector and one (or more) scintillator, to observe the annihilation products. A viewport permits the lasers necessary for excitation and ionisation to enter the chamber. A schematic drawing of the entire assembly and more detailed images of the test chamber are shown in figure 2.23.

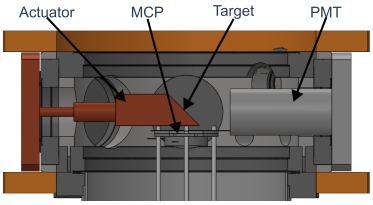
This test setup has been instrumental to multiple experiments in AEgIS, and in particular for the laser cooling of positronium atoms (see *Positronium laser cooling* (A.1)).



(a) The entire "Breadbox" setup.



(b) Detail of the "Breadbox" test chamber, top view (courtesy of Malgorzata Grosbart).



(c) Detail of the "Breadbox" test chamber, side view (courtesy of Malgorzata Grosbart).

Figure 2.23 – On the top: a schematic drawing of the whole "Breadbox" setup for positronium physics. The buncher is used to extract, accelerate and focus the e^+ onto the target positioned inside the test chamber, where an MCP and/or a PMT detect the annihilation products. On the bottom: detailed views of the "Breadbox" test chamber (courtesy of Malgorzata Grosbart)

2.1.12 $e^+ \rightarrow$Ps converter

As seen in the *Brief description of AEgIS* (1.3.1), AEgIS employs positronium atoms to form antihydrogen, therefore an efficient $e^+ \rightarrow$Ps converter is fundamental for the experiment.

The most common and effective manner of producing Ps is by implanting positrons into a solid medium: the advantage is mainly due to the naturally high electron density in materials, very difficult to achieve in a trap. Furthermore, the momentum exchange occurring inside the medium promotes positronium formation. Nevertheless, it is fundamental to ensure that the positronium thus formed can exit the material: otherwise, annihilation will occur. Once emitted in vacuum, the positronium can be excited to higher Rydberg states via laser pulses, both extending its lifetime and enhancing the probability of interacting with other particles.

The multitude of interactions and processes that positrons undergo within a solid are well-detailed in [126]. In short, when entering the bulk of the material at keV, positrons are scattered by the electric field generated by nuclei and electrons, and they rapidly lose their energy mainly via bremsstrahlung and core excitations. Going down in energy, other processes become dominant, namely the excitation of plasmons, phonons and electron-hole pairs formation, until, below the eV, only phonons excitation remains available as a cooling mechanism.

Subsequently, the thermalised e^+ undergo random diffusion, the duration of which is influenced by factors such as the core and electron density of the material, the presence and nature of defects, and the temperature. The positrons may eventually undergo direct annihilation, delayed annihilation (by getting trapped in a defect with low electron density), or they might be re-emitted in vacuum, either directly or in the form of positronium. But since the space travelled when being

implanted (O(µm)) is way higher than the diffusion process (hundredth of nm), the chances of Ps being freed outside the material are very low.

A good solution to this problem is to use silicon, where channels orthogonal to the surface, the size of several nm, are produced by chemical etching, and subsequently, the surface is oxidised by exposure to air. The positron, once inside the material, will likely diffuse into one of these nanochannels, possibly picking up an electron either from the silicon bulk or if not, from the internal surface of the pore. The last process is favoured from the fact that SiO_2 has a negative work-function for positrons: the picking up of an electron is very likely, and then the so formed Ps is re-emitted inside the channel with an energy of $\sim 1\,\text{eV}$. When inside the channel, the Ps undergoes scattering along the walls, progressively cooling by the collisions with the walls of the nanochannel (a process known by *collisional cooling*): if the pore is connected to the surface of the silica target (*open pore*), the positronium atom can escape into vacuum, being effectively "emitted" by the $e^+ \rightarrow$Ps converter[11]. A pictorial schematisation of the $e^+ \rightarrow$Ps process is shown in figure 2.24.

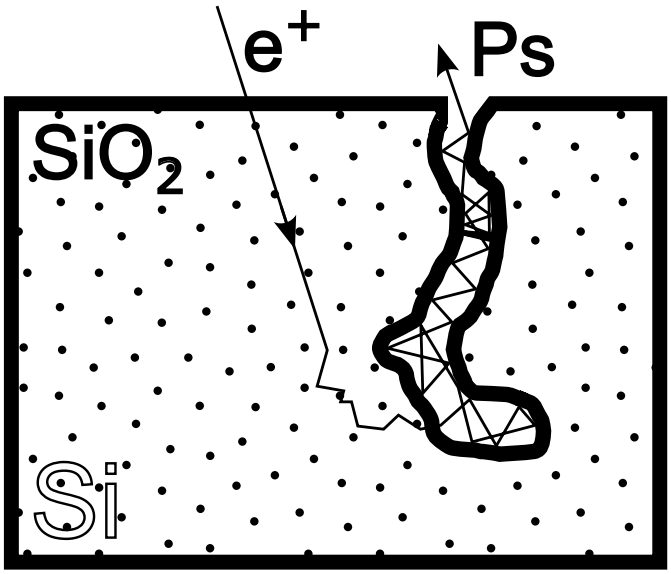


Figure 2.24 – A drawing of the process taking place inside the $e^+ \rightarrow$Ps converter, where the impinging positron gets implanted into the bulk of the material, loses energy, tears an electron and binds to it, and the newly formed Ps escapes through the nanochannel, collisional-cooling while doing so.

Afterwards, the positronium atoms can be either excited to various Rydberg states (2/3 or higher) or photo-ionised by a series of two lasers, as presented in *Lasers* (2.1.6).

The AEgIS Ps target, at room temperature ($\sim 300\,\text{K}$), has an efficiency of conversion of ∼30 %, with a positron implantation energy of 3.3 keV [127].

2.1.13 Data Acquisition System

The AEgIS experiment operates an integrated measurement and monitoring data acquisition (DAQ) system, to capture and preserve all the data generated in the various measurements, to have them available for analyses when needed. All data is subdivided into *data atoms*, i.e. units of the smallest data container used, which have a standard format, composed by:

[11] It has to be noted that the positronium atoms so emitted are almost all ortho-Ps (i.e. Ps where e^- and e^+ have parallel spins), since the process takes several nanosecond, thereby causing the annihilation of all the Ps in the para-state (i.e. with e^- and e^+ with antiparallel spins).

- **Name:** alphanumeric string containing a (possibly hierarchical) unique name for the data atom. The format of the data associated with a given name should not be changed.
- **Timestamp:** instant when the data was acquired, in three formats:
  - character string, parsable by `strptime(3)`
  - `struct timespec` containing time since the UNIX epoch
  - 64-bit unsigned integer with RF clock count, if applicable
- **Data:** instance of a scalar, vector, or structured (cluster) data type, compatible with LabVIEW™ types, and their conversion into either json-formatted files or ROOT `TTree`s

Data atoms are generated at various locations in the experiment, transferred over the local-area network, saved to local storage, and then saved to long-term disk and tape storage systems at CERN. Data sources and sinks, along with the data transfer paths over the Local Area Network (LAN), are identified in figure 2.25. This system is designed for the vital parameter monitoring needs for experiment commissioning and the long-term data logging for experimental measurements, and has been running for over a decade.

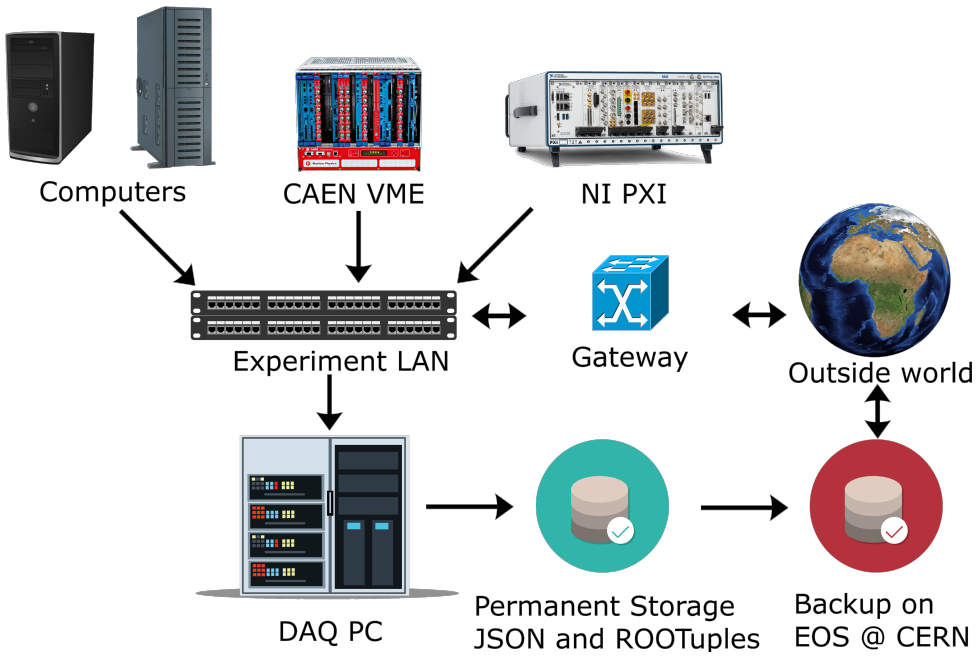


Figure 2.25 – Scheme of the data flow in AEgIS. All devices (computers, VME and real-time) are connected to a common LAN subnet and send data to the DAQ PC as GXML Data Objects over TCP or SCP (Secure Copy Protocol). The DAQ computer permanently stores the data on hard drives as JSON files and ROOTuples. A further backup copy of the data is generated on EOS [128] at CERN. The data can be accessed from outside CERN from EOS, or directly from the DAQ computer via a dedicated gateway.

The data are saved in JSON-formatted files, which provide a compact, clearly structured standard for efficient generation and transfer and are compatible with the GXML reference library (for serialisation) of the LabVIEW™ architecture used in many experiments.

For online access to monitoring data, CERN's ROOT data format system is currently still used preferentially thanks to its high data compression functionalities.

A side-by-side comparison of text representations of the general-purpose AEgIS data atom in the GXML and JSON formats is shown in figure 2.26.

```
<gxml_Root>
    <Name type='String'>
        test_cluster
    </Name>
    <Timestamp mems='4'>
        <str type='String'>
            16:17:18.020212 10/20/2021</str>
        <tv_sec type='U64'>1634739438</tv_sec>
        <tv_nsec type='U32'>20212223</tv_nsec>
        <Clock type='U64'>7856432</Clock>
    </Timestamp>
    <Data mems='3'>
        <double_val type='DBL'>
            1.2345</double_val>
        <int_val type='I32'>12345</int_val>
        <float_array dim='[3]' type='SGL'>
            <v>1.1</v><v>2.2</v><v>3.3</v>
        </float_array>
    </Data>
</gxml_Root>
```

```
[ { "test_cluster": {
    "Timestamp": {
        "clock": 7856432,
        "str": "16:17:18.020212 10/20/2021",
        "tv_nsec": 20212223,
        "tv_sec": 1634739438
    },
    "Type": "",
    "double_val": {
        "Type": "DBL",
        "__value": 1.2345
    },
    "float_array": {
        "MemberDims": "[3]",
        "Type": "Array",
        "v": [1.1, 2.2, 3.3]
    },
    "int_val": {
        "Type": "I32",
        "__value": 12345
}}}]
```

Figure 2.26 – Left: Example of GXML serialisation of an AEgIS data atom containing a cluster of two numeric scalar values and one numeric array. Right: The corresponding JSON equivalent representation.

2.1.14 Control system electronics

The timing precision needed in modern physics experiments like AEgIS (typically, at least, on the order of the ns) cannot be handled by PC Operating Systems (OS) without real-time hardware and OS extensions whose complexity often makes it preferable to delegate time-critical operation to one or more *ad-hoc* Field Programmable Gate Arrays (FPGAs). In AEgIS, the Sinara [18] ecosystem has been chosen as the base of the (new) control system electronics.

Sinara comprises a varied collection of open-source hardware components that were initially designed for quantum information experiments. This FPGA is programmed with a custom programming language called ARTIQ (Advanced Real-Time Infrastructure for Quantum physics) [129] (described in depth in *ARTIQ* (3.2)). The Sinara hardware provides compact, modular, reproducible and reliable electronics capable of controlling intricate, time-critical experiments. It is particularly optimised for experimental setups which are limited in space, as is the case inside the AD, and, thanks to its standardised and modular nature, assures the long-term maintainability of the control system.

As shown in figure 2.27, the hardware of the AEgIS trap control system is organised in rack-standard Eurocard 84 HP electronics crates, each hosting a variety of modules. The main controller is called Kasli (see figure 2.27). It comprises an Artix-7 FPGA and can be used as a stand-alone core device or in combination with additional carriers as a repeater or satellite of Distributed Real-Time Input/Output (DRTIO) communication through optical fibre links, facilitating a stable, high-speed Gbps transfer of time and data information between the devices. This second option allows for fast propagation of both a clock signal (internally generated or externally connected) and the control communication between controllers, thus offering straightforward adaptations and extensions of the experiment. Software communication with Sinara electronics is facilitated via Kasli's high-speed Gigabit Ethernet port. Each Kasli is capable of controlling up to twelve extension modules with various purposes.

Each Sinara crate used in AEgIS contains a Kasli carrier combined with digital I/O units and fast DAC modules, called Fastino, from the Sinara repertoire, as well as several voltage amplifier modules, which have been custom-designed for the requirements of the AEgIS experiment.

The digital I/O cards are used for the reception and provision of high-speed ns TTL trigger signals

Figure 2.27 – Photograph of one of three fully equipped Sinara electronics crates of the AEgIS trap control system, including (from left to right) power module, Kasli carrier, digital I/O units, Fastino DAC, and four high-voltage amplifier boards.

between the sub-systems of the entire experimental setup. On each single, thin module 16 MCX connectors are compactly arranged, and their direction of input or output can be individually configured in batches of four.

Each Fastino provides simultaneous $3\,\mathrm{Mbit\,s^{-1}}$ digital-to-analog conversion for 32 channels, yielding stable output voltages in the range of $\pm 10\,\mathrm{V}$ with a 16 bit resolution. The Fastino DAC channels can either be used directly to supply low voltages in this range or be connected in batches of eight to the voltage amplifier modules.

One such amplifier unit comprises eight channels, each of which is capable of a 20-fold amplification of the output voltage of one Fastino channel respectively, i.e. yielding voltages of up to $\pm 200\,\mathrm{V}$. The voltage amplifiers are equipped with individual OptoMOS® relays, allowing the isolation of the outputs and preventing the noise from the amplifiers from propagating to the connected load.

The main control electronics of the AEgIS setup are formed by three of the described Sinara crates: two (one Kasli acting as DRTIO repeater, the other as satellite) provide the voltage output channels for the synchronous potential steering of the low-voltage electrodes of the 5 T trap region of the experiment, required for electrons and cold antiprotons manipulations, while the third crate is used for the control of the 1 T antihydrogen production trap electrodes. The first two are physically referred to as *5TC1* and *5TC2*, while the third *1TC1*. Their *logical* names, i.e. the names used to distinguish them from a programming point of view, are simply 5 T Kasli and 1 T Kasli, being the *5TC1* and *5TC2* daisy-chained and controlled as a single device.

During the ELENA/AD antiproton run campaigns, the fast digital I/O units have demonstrated reliable acquisition and processing of the incoming trigger signals, essentially enabling the steering of the trap potentials with the required timing for the capture of antiprotons.

In figure 2.28, examples of output signals of three voltage amplifier channels are shown. They are produced by sending an external trigger pulse to the digital I/O unit and subsequently setting a voltage on three of the Fastino channels. The voltage is amplified by the connected amplifier units. The final output is recorded using an oscilloscope and read out via LabVIEW™. The Sinara system is thus found to be able to satisfy the timing requirements of the AEgIS experiment: reactions to triggers on the microsecond scale and synchronous control of the output channel voltages are reliably provided.

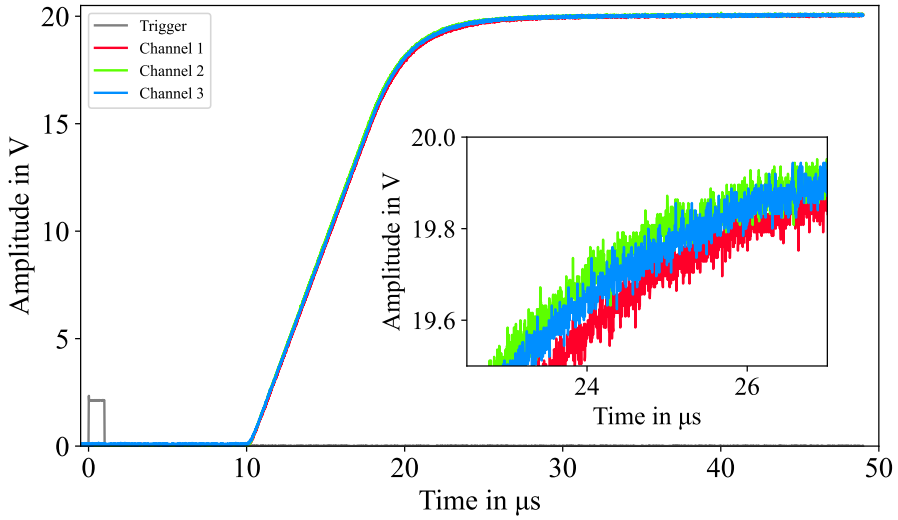


Figure 2.28 – Synchronous voltage ramp-up to 20 V on three voltage amplifier channels 10 µs subsequent to the arrival of a common trigger pulse at zero time in the figure. The inset shows a zoom to the shoulder region for a better visualisation of the synchronicity.

All amplifier channels have been calibrated individually together with their corresponding Fastino DAC channels to reliably provide the required voltage despite their different offset and voltage precision step values. With this calibration, each channel voltage can be set with an accuracy of few mV, which is comparable to the precision of the 6 mV steps generated by the 16 bit nature of the Fastino.

In addition to the electronics controlling the trap system and providing inter-system triggers, two additional Sinara crates have been successfully commissioned to run the laser system and provide synchronisation between the two involved lasers despite their difference in frequency (see *In-Run autonomous parameter stabilisation* (5.1.4)). In order to be able to do this, the 1 T Kasli, in addition to controlling the respective trap system, is used as the master core for two satellite Kasli devices, both of which control a digital I/O card with BNC output for triggering the sequences needed for laser operation.[12] Furthermore, the new control electronics have been successfully integrated as part of the AE$\bar{g}$IS positron system to provide triggers for the positron preparation sequence and synchronize it to the rest of the experiment. Further extensions of the control infrastructure, e.g. dedicated Sinara crates for the positron system and to future-proof the use of the rotating wall technique for plasma compression, are ongoing.

There are two relevant additional electronics components, which have been integrated into the new control system setup and are fully steerable programmatically. The first is a pulser device which provides ns-synchronised pulses of variable length to the electrodes, with tunable amplitude in the voltage range provided by the Sinara Fastino plus amplifier channels. The trigger signals for this pulser are given by the Sinara digital I/O units, while the amplitude is determined by internal DAC units. The inclusion of this functionality is vital for the efficient and fast transport of particles between the different trapping regions inside the experiment. The second component is a waveform synthesizer with multiple channels, which can be used to add phase-shifted sinusoidal signals of up to 5 V in a frequency range of 0 MHz to 30 MHz to the so-called sectorised electrodes. These electrodes are separated into four sectors around their centre, i.e. around the central axis of the trap. By applying the sinusoidal signals with a phase shift between the four sectors, it is possible to employ the rotating wall technique for manipulation of the radial dimension of the contained particle plasma.

[12]The BNC digital I/O units work in the same way as the MCX units except for comprising only eight channels instead of 16.

2.1.15 CIRCUS, the new AEgIS autonomous control system

Control systems are, generally speaking, combinations of hardware and software with the ability to modify the operation and/or configuration of other elements of a system and are in charge of the management of that system. Autonomous control systems are such that can operate with little to no human supervision. They are applied in any imaginable field, from satellites to dishwashers. Control systems for antimatter, nuclear, atomic, and quantum physics experiments, like AEgIS, are a special category because they need to deal with systems that are continuously upgraded, fixed, and reshaped. For this reason, they need to maintain stability, reliability and reproducibility while allowing for the flexibility necessary for the experiment to mutate: this is different from the demands of the control systems of the big observational experiments (e.g., neutrino telescopes), which are less prone to change. The nature of these experiments puts a range of constraints on the control system: nanosecond-precise execution, multiple computer synchronisation, interfacing with different hardware using multiple interfaces, and easy extendability, among others.

A common characteristic of this class of experiments is that the synchronised coordinated behaviour of multiple components is required to ensure the success of the complex measurements involved. More often than not, the various subsystems are constructed independently: while this enables parallelisation during the development, it frequently leads to difficulties during the integration and operation phases. In fact, each part is built according to the likes and experience of a single group of scientists: in some cases, e.g. for nuclear, atomic, and quantum physics experiments, the fields of expertise can be very different, ranging from lasers to electromagnetic traps, to ultra-cold and ultra-high vacuum, to various detection techniques.

Particularly suffering from this phenomenon are, indeed, experiments' control systems' software. As a matter of fact, albeit interfaces for system integration might have been decided a priori, the various parts are coded with different paradigms, with different styles, and often even in different programming languages, leading to difficult subsequent iterations of unification. In addition, this limits the interplay among different scientists, and creates a great barrier to knowledge transfer: every time a new way of reasoning has to be learned.

The AEgIS experiment is a perfect example of this situation: as mentioned before, AEgIS has already effectively generated antihydrogen atoms in pulsed mode (see *A bit of history: from Phase 1 to Phase 2* (2.2)).

Utilising the knowledge gathered from previous experience, the AEgIS collaboration has designed a generalised experiment control system that is customisable to individual experiments' specific requirements, the CIRCUS (Computer Interface for Reliably Controlling, in an Unsupervised manner, Scientific experiments) [1]. This flexibility benefits the AEgIS experiment (as it allows it to evolve smoothly to changing requirements), but equally importantly, the system was constructed with the needs of the much wider atomic and quantum physics community in mind. This control system incorporates a programmable end-user interface, providing advanced synchronisation, watchdogs, error management, and online decision-making features, reinforced by an active feedback loop from the acquired data. This re-design specifically targeted complexity reduction of experimental procedures by standardising established sub-procedures into libraries, and by increasing stability, reliability, and autonomy. With this as the baseline, the subsequent implementation of increasing layers of automation and autonomy becomes feasible, strengthening the trust in the system through cycles of campaigns of implementation and debugging.

The implemented solution merges the capabilities of a real-time code with a distributed slow-control system that unifies the computer in a single entity and brings together all the features described above, so as to partially remove the operators' need to control the running procedures. The control system itself is completely experiment-agnostic (technically, it could be used to control experiments outside the realm of physics as well), and it is released open-source so that other experiments can profit from the effort. CIRCUS is presented in greater detail in the dedicated chapter *The CIRCUS* (3).

The high level of automation is a choice motivated further by the introduction of ELENA (see *Antiprotons production in AD* (2.1.4)): with the new continuous delivery modality of antiprotons to the experiment, the previous control system would have demanded an operators' workload unable to sustain, ultimately leading to loss of beam time and people overwork.

Furthermore, recognising the broad applicability of this approach, it was decided to develop a novel framework for control systems in atomic and quantum physics experiments, and then construct the AEgIS control system atop it. This framework, TALOS (Total Automation of Labview Operations for Science) [2], is extensively described in *TALOS, the framework* (3.4).

The development of CIRCUS and TALOS have been one of the greatest contributions I have given during the upgrade of the AEgIS experiment from Phase 1 to Phase 2: I have participated in the conceptualisation, performed the architecture design in full autonomy, and written and tested most of the code and documentation. In 2022, I formed and supervised another doctoral student, Jakub Zielinski, who then contributed to the development of the control system.

2.2 A bit of history: from Phase 1 to Phase 2

During AEgIS Phase 1, the experiment goal was to demonstrate the feasibility of using the charge-exchange reaction between Rydberg-excited positronium atoms and antiprotons to form antihydrogen. This result was achieved in 2018, when antihydrogen production was demonstrated [15], as shown in figure 2.29. The achieved rate was calculated to be approximately 0.05 $\bar{H}$ per experimental cycle, which lasted $\sim 110\,s$.

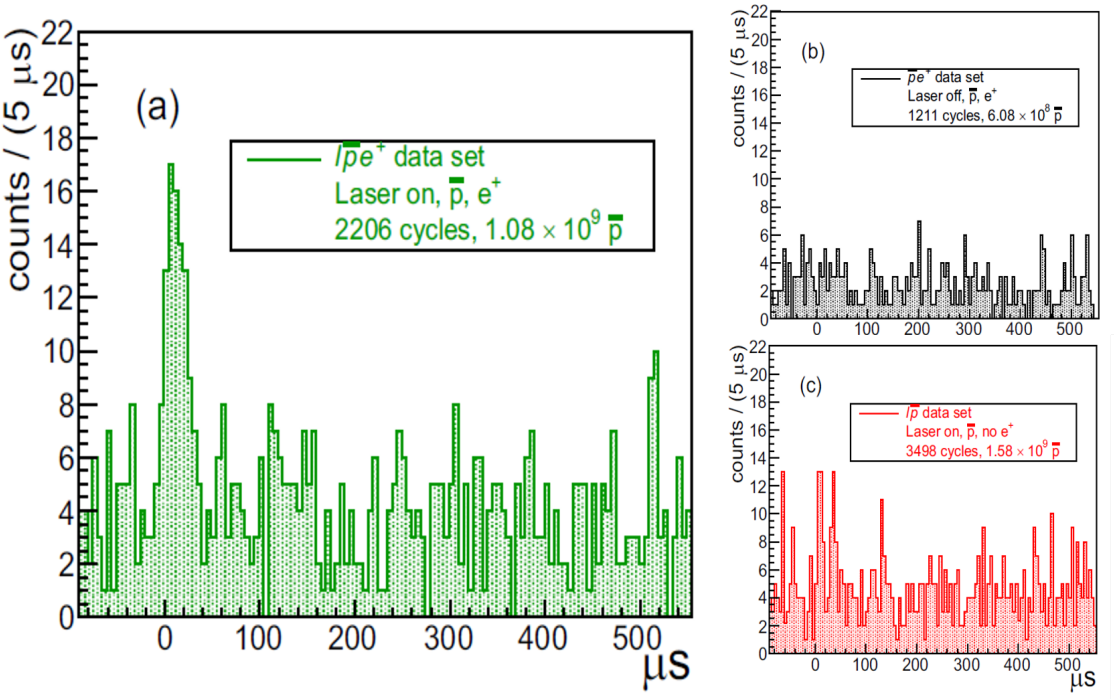


Figure 2.29 – Evidence of $\bar{H}$ formation in the AEgIS experiment, using the charge-exchange reaction (from [15]). The three graphs show the annihilation counts in case of (a) lasers, $\bar{p}$ and e^+ present; (b) only $\bar{p}$ and e^+; (c) only lasers and $\bar{p}$. The peak on the left of figure (a) represents the excess of annihilations caused by the $\bar{H}$ formed.

The second phase of AEgIS has started in 2019, together with CERN LS2, and it is going to last until the beginning of CERN LS3, in 2025. It is set to enhance and solidify the antihydrogen formation procedure, and to perform a proof-of-concept inertial measurement using antimatter. A critical insight from Phase 1 was that, in order to accumulate sufficient data for a gravity measurement with

a target precision of around 1 %, the need for a substantial increase in antihydrogen production was necessary, to have ca. $1 \div 10$ $\bar{\mathrm{H}}$ per ELENA cycle, i.e. at least two orders of magnitude higher than Phase 1.

Moreover, another key development to achieve such a measurement is to lower the temperature of the produced antihydrogen (by further cooling antiprotons) to tens of kelvin, a significant reduction from the previously achieved $\sim 400\,\mathrm{K}$. Together with these, the additional two objectives for Phase 2 include the creation of a beam of antihydrogen forward-boosted and the development of a moiré deflectometer prototype for inertial measurement.

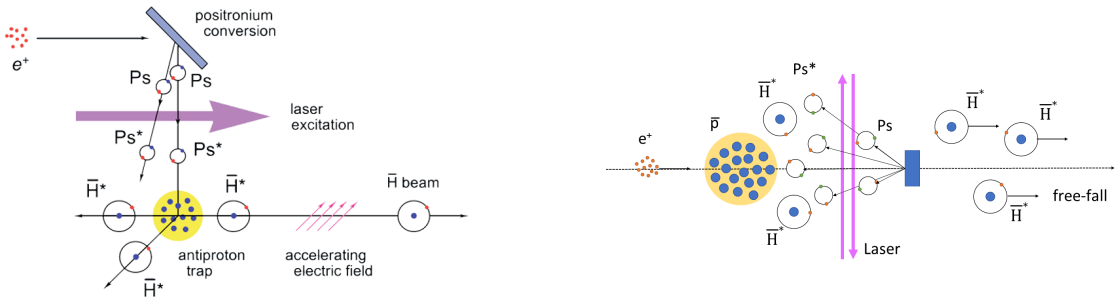
To achieve these results, a series of major upgrades have been performed on the experimental apparatus [6]. Among these, there have been changing how the Ps illuminates the $\bar{\mathrm{p}}$ plasma (from orthogonal to collinear), a completely redesigned formation trap, efficiency and stability improvements to the positronium line, a higher-yield $\mathrm{e}^+ \rightarrow$Ps converter and a more powerful laser system, and the migration of most of the electronic and software control system to the ARTIQ/Sinara and TALOS.

In the following, the upgrades are briefly presented, together with the antihydrogen production enhancements they are estimated to bring; finally, based on these numbers, the expected new $\bar{\mathrm{H}}$ rate is calculated.

2.2.1 The new formation scheme

The biggest innovation is the antihydrogen improved production scheme. In fact, during AEgIS Phase 1, a double channel formation trap was used: in the inner, coaxial trap, the antiprotons were kept, while the e^+ were sent into the parallel channel to hit a 30°-tilted Ps target (toward the beam axis), which was illuminating the $\bar{\mathrm{p}}$ plasma orthogonally. This scheme, albeit functional, was of higher complexity.

Now, in Phase 2, the charge-exchange reaction is still exploited, but the positronium atoms move along the axis to the trap, and hence to the magnetic field. This is obtained by placing the Ps target on the axis of the trap, at its end (see *1 T section* (2.1.5.2)), and forming the Ps cloud by launching the e^+ through the trapped $\bar{\mathrm{p}}$. A pictorial schematisation of the two formation schemes is depicted in figure 2.30.



(a) Perpendicular $\bar{\mathrm{H}}$ formation scheme, used in Phase 1.

(b) New collinear formation scheme, used in Phase 2.

Figure 2.30 – Left: orthogonal formation scheme, where the Ps illuminate the $\bar{\mathrm{p}}$ perpendicularly (from [5]). Right: collinear formation scheme, minimising Ps ionisation caused by the motional Stark effect.

The new formation process has many advantages: the first one is to raise the maximum Ps Rydberg level achievable from $n \approx 19$ to 32 (and possibly more). This is because the Ps, travelling collinear with the magnetic field, experience a smaller relative electric field:

$$E = \left|\vec{v} \times \vec{B}\right| = vB \sin\theta \tag{2.2.1}$$

where θ is the angle between the magnetic field lines (i.e. the axis of the trap, to a good extent) and the velocity vector of the positronium atom. In turn, this raises the threshold for the maximum n allowed by the ionisation limit caused by the motional Stark effect [130]:

$$E_{ion} = 1.3 \cdot 10^8 \frac{1}{9n^4} \frac{\text{kV}}{\text{m}} \tag{2.2.2}$$

Bringing together these two equations, we find[13]:

$$n_{max} = \sqrt[4]{\frac{1.4 \cdot 10^{10}}{vB \sin\theta}} \tag{2.2.3}$$

To find the maximum Ps Rydberg level allowed, let us make a simple model: the positronium atoms need to illuminate a $\bar{\text{p}}$ plasma of 5 mm radius, distant 35 mm from the target, and it is moving at a speed of $\sim 10^5$ m/s. Under these assumptions, the plasma is seen under an angle of $\sim 8°$: at this angle, the $n_{max} = 32$, which means already a tenfold higher cross-section than with $n = 17$. Going within the inner 2 mm of the plasma, it could get as high as $n = 40$, meaning 30 times higher cross-section.

The second benefit is the possibility of creating the antihydrogen "ballistically", i.e. by launching the antiprotons from the capture trap to the formation trap, and while they are in motion, proceeding with the positrons injection as before: this modality would enable the creation of a forward-boosted antihydrogen beam without requiring additional hardware. In addition, by keeping the A0 electrode – which is the electrode holding the Ps target (see figure 2.31 and *1 T section* (2.1.5.2)) – and/or the first ionisation grid biased to a negative potential, the antiprotons so launched can be filtered out, letting only the antihydrogen formed passing over and reaching the future gravity detector. Even more, the antiprotons can be made "swing" back, eventually recapturing them and recycling them for the subsequent shot.

In figure 2.31 is shown how the new $\bar{\text{H}}$ formation process actually takes place inside the new trap that needed to be built for the purpose (see next section).

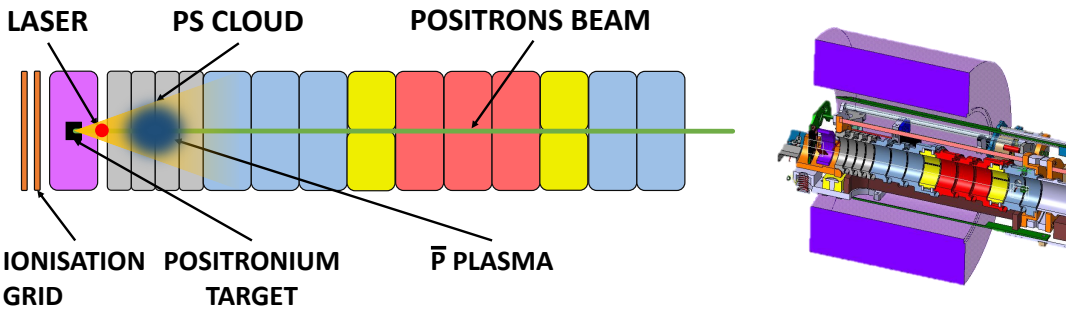


Figure 2.31 – On the left: a pictorial schematic on how the charge-exchange reaction takes place in the AEgIS formation trap. On the right: 3D technical drawing of the corresponding trap section (courtesy of Stefan Haider).

2.2.2 New 1 T trap

The old 1 T trap was made up of an initial section, of 45 mm diameter electrodes, followed by two sections, stacked one on top of the other, of 10 mm electrodes. The one coaxial with the axis of the experiment was terminating with the old formation trap section, where one of the electrodes had an opening to let the positronium atoms fly inside from the Ps target placed above.

[13] Here, v is expressed in m/s and B in T.

The new trap, exposed in *1 T section* (2.1.5.2), simplifies drastically the design, eliminating all the variations in diameter and openings that were causes of field inhomogeneities, leading to instabilities into the antiproton plasma. Furthermore, the bigger diameter electrodes towards the end generate an electric field in the centre with a small radial gradient, improving the conditions for the $\bar{p}$ plasma.

2.2.3 ELENA instead of AD

The upgrades to the apparatus were also deemed necessary by the introduction, in the Antimatter Factory, of ELENA (see *Antiprotons production in AD* (2.1.4)), which became active in 2021. In fact, before then, the antiprotons were provided to the experiments directly from AD, in shifts of 8 h per experiment, with a kinetic energy of 5.3 MeV, while now ELENA supplies $\bar{p}$ at an energy of 100 keV, simultaneously to up to four experiments every 2 min, 24 h/day, for the total duration of each antiproton yearly campaign.

While the lower energy of the antiprotons has boosted the trapping efficiency of AEgIS significantly (although it required the development of new degraders, see *Degrader(s)* (2.1.7)), the modality of receiving them would have not been sustainable with the previous control system, which was needing a high degree of supervision: it would have demanded a huge effort to the operators, which could have ultimately resulted to not fully profit of the extended period of beam time. Moving to a control system with a high degree of automation was, therefore, necessary, both to maximise the amount of data taken and to lower the time that the users would have spent overseeing it, which could be allocated to physics and development.

2.2.4 The new control system software and hardware

As already mentioned in *CIRCUS, the new AEgIS autonomous control system* (2.1.15), in Phase 1 AEgIS was operated by multiple control system software operating a series of custom-made control electronics.

In particular, the old software architecture was composed as outlined in table 2.2. When performing coordinated experiments, the trap control system acts as the overall coordinator, sending commands to the other control systems. The series of operations of each Run[14] were specified as different lines of a Microsoft Excel file: this seriously impeded the scalability of the system, making clustering operations into higher level functions impossible. Moreover, the readability of the instructions was extremely low, making the system very difficult to be operated by persons different from the main developers. Furthermore, the code of the trap system was a monolithic LabVIEW™ system: the little scalability and maintainability had led it to "grow" in an uncontrolled manner, rendering it extremely complex and time-consuming for each debug and/or addition needed[15].

Of this entire multitude of software systems, the *Trap System*, the *Vacuum and cryogenics* and the *Laser system* have been ported and superseded by the new CIRCUS control system. The *Positron system* was quickly interfaced to it, and it is foreseen to be fully ported inside CIRCUS in 2024. The *Data acquisition* has been heavily interfaced with CIRCUS, so to be operable as if it were a native component (see *DAQ Manager* (3.4.5.8) and *DAQ Sender* (3.4.5.9)). The *Magnet system* is operating standalone under the control of the CERN Magnet Team, and it has remained untouched.

The electronics, instead, were constituted by several DAC boards, grouped in multiple crates, each equipped with an FPGA: everything was custom-made for AEgIS. The NI PXI (see *PXI* (2.1.9.4)) was acting as a supervisor, and it was used not as a Windows machine, but as an Embedded

[14] AEgIS is inherently a pulsed experiment, and as such is organised in Runs: each Run is an atomic experiment, identified by the execution of the instructions defined in a single file (Microsoft Excel in Phase 1, ARTIQ Python script in Phase 2).

[15] It was a good example of what, in LabVIEW™ jargon, is called "spaghetti monster".

Subsystem	Software architecture
Trap system	LabView/Windows + LabView/RTOS
Magnet system	LabView/Windows + PLC
Data acquisition	Custom/Linux
Vacuum and cryogenics	LabView/Windows
Positron system	LabView/Windows
Laser system	Python/Windows + FPGA

Table 2.2 – List of the old AEgIS software control systems (from [115]).

Real-Time controller (RT). The sequence of operations specified in the Excel file fed to the *Trap system* was downloaded to the PXI at the beginning of the Run, and then the PXI was taking care of periodically sending at the correct time the specific programming for each FPGA.

Albeit the system was expressing good performances, being completely custom-made was making it increasingly difficult to repair; more to that, its software interfaces, completely non-standard, were impeding its integrability in a different control system than the original one. Therefore, moving to an open-source hardware and software ecosystem like the ARTIQ/Sinara one was necessary both in terms of system simplification, extendability, and futureproofness.

2.2.5 Improved positron line

In 2021, a huge work to improve the positron line took place. It mainly consisted of the realignment of the entire beamline with respect to the central axis of the solenoid producing the confining magnetic field, into the servicing of the cryogenics and vacuum system, and the rationalisation of the cabling. This thorough revision of the positron line has increased the efficiency of the positrons entering the accumulator by approximately a factor of three: as can be seen in figure 2.32, the intensity of a pulse of positrons extracted from the *Accumulator* as a function of the accumulation time is roughly the same in 2021 than in 2017, despite the activity of the ^{22}Na source (half-life of 2.6 years) is three times lower.

Up until 2022, positrons in the AEgIS setup could be directed either to the main apparatus for antihydrogen experiments or to the *Breadbox* chamber (described in *The "Breadbox"* (2.1.11)), initially intended for magnetic field-free Ps experiments, even if in the first experiments a pair of Helmholtz coils were employed to enhance the number of positrons correctly hitting the target.

In a significant development that happened in 2022, the AEgIS team achieved truly magnetic field-free positron transport: this breakthrough involved installing a mu-metal shield around the test chamber to block all magnetic fields (reducing residual fields to less than 2.5 G when AD is functional, and less than 0.5 G when AD is off), adding new correction coils for better positron transport, and optimising the rotating wall technique for positron plasma compression. These advancements now permit precision Ps experiments using lasers, free from magnetic quenching effects (see [131]) or perturbations in the Ps energy levels.

Moreover, the system includes a fast switch that precisely turns off electrostatic potentials in the test chamber with ns accuracy: this feature allows for the implantation of positrons at an energy level of 3.3 keV onto the Ps target, ensuring that Ps emission takes place in an environment completely free of both magnetic and electric fields.

This improvement was foremost towards the success of laser cooling positronium atoms (see *Positronium laser cooling* (A.1)).

2.2.6 New ^{22}Na positron source

The positrons source, made of ^{22}Na radioisotope (half-life 2.6 years), naturally decays with time and consequently needs to be periodically replaced. A new 50 mCi ^{22}Na source was ordered in 2022

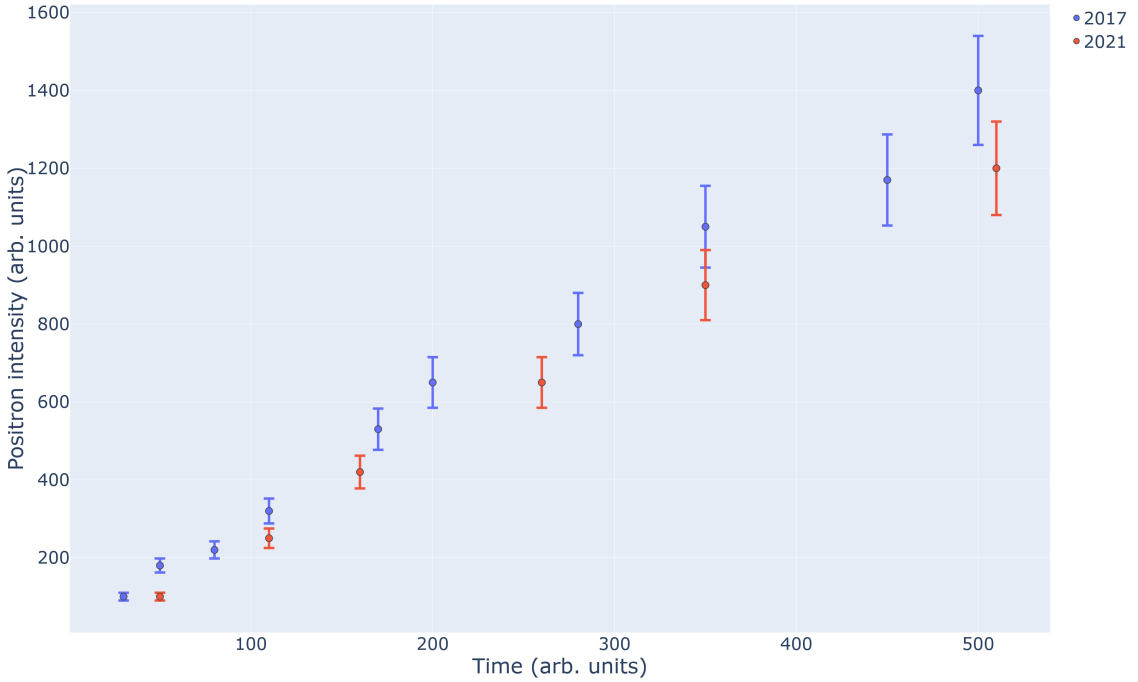


Figure 2.32 – Positron intensity as a function of the accumulation time, both for 2017 and 2021. It can be seen how the two results are very similar, despite a reduction of the source activity by approximately a factor of three: it is compensated by the improved performance of the positron line.

to be installed during the 2023 antiproton run, so to substitute the old one, installed during Phase 1, that in 2023 had an activity of 6 mCi ($\sim$ 220 MBq): this would lead to more than 8 times more positronium generation, which in turn means a similar improvement in antihydrogen production rate.

Unfortunately, the sole supplier, iThemba LABS, faced significant delays, due to the impact of the pandemic and the political situation in South Africa, resulting in a backlog that affected the delivery of the new positron source both for AEgIS and other experiments (like ALPHA and ASACUSA). As a consequence, all the 2023 data collection was conducted using the older 6 mCi source.

2.2.7 Improved $e^+ \rightarrow$Ps converter

Since the number of $\bar{H}$ produced is proportional to the number of positronium atoms present in the reaction, a huge work has been put in place to enhance the efficiency of the $e^+ \rightarrow$Ps converter.

A study was performed [104] to tune the morphology of the nanochannels such that the emission of Ps from the surface was maximised. It was found that the optimal setting consisted of nanochannels with diameter of 7-10 nm and depth of $\sim$ 3.89 µm, converting approximately 28 % of the e^+ implanted at 3.3 keV. Moreover, another key aspect taken into consideration is the Ps thermalisation inside the target: in fact, it was found that shortening too much the length of the pores (e.g. to 1.13 µm) was detrimental to the cooling of the formed Ps.

This development served as a basis to realize the Ps target that was installed inside AEgIS main apparatus.

2.2.8 New crystal for the EKSPLA to higher Ps Rydberg level excitation

Having lifted the limitation on the maximum Ps Rydberg level achieved, imposed by the motional-induced Stark effect ionisation, an upgrade to the excitation laser has been performed, to be able

to increase the reachable level range. The laser modification consisted of substituting one of the crystals that generated the infrared wavelength used for the $3^3\mathrm{P} \rightarrow n$. The new crystal should have ensured to be able to go as high as $n = 40$, but for an error in the manufacturing process, the crystal received can only produce wavelengths up to $\sim 1665\,\mathrm{nm}$, which corresponds to a Ps level of about $n = 24$. It has given an improvement to the $\bar{\mathrm{H}}$ formation rate, but relative: the increase has been of a factor of 4, instead of the planned 12.6 (for $n = 32$) or more.

Nevertheless, the laser has been also fully robotised, by adding several actuators that can pilot remotely the various tuning lenses and crystals that before had to be regulated manually. This has led to the capability of aligning the laser from remote, easing the pressure from the operators on-site. Moreover, it also made the laser capable of being tuned autonomously by the control system: a simple test had shown that the optimiser (see *Automatic parameter optimisation* (3.4.7.3)) can perform a full alignment, starting from a random configuration, in less than two hours.

2.2.9 Expected new $\bar{\mathrm{H}}$ production rate

All the aforementioned modifications have been carried out with the purpose of producing antihydrogen at a rate higher than 1 $\bar{\mathrm{H}}$/min, necessary to perform a direct WEP test on antihydrogen.

To estimate the increase in antihydrogen formation, the approximate analytical rate of formation is sufficient [95]. At first order

$$\frac{dN_{\bar{H}}}{dt} = N_{\bar{p}} N_{Ps} \sigma_{\bar{H}} \Delta v \tag{2.2.4}$$

Defining k_v as the ratio of the mutual velocity v and the classical orbital speed $v_{e^\pm}$ of $e^\pm$ in a positronium atom:

$$k_v = \frac{v}{v_{e^\pm}} = \frac{2nv}{\alpha c} \tag{2.2.5}$$

an analitical approximation for the antihydrogen formation cross-section (with the charge-exchange reaction) is:

$$\sigma_{\bar{H}} = n_{Ps}^4 (10^{-19}\mathrm{m}^2) \frac{1.2 + 0.1 k_v^{-2}}{1 + (k_v/1.84)^{18}} \tag{2.2.6}$$

In table 2.3 the expected $\bar{\mathrm{H}}$ production rate for Phase 2 is estimated, combining the calculated or measured improvement values for each upgrade. As it can be seen, antihydrogen should be formed at a rate higher than 1 $\bar{\mathrm{H}}$/cycle, going up to 100 $\bar{\mathrm{H}}$/cycle.

Unfortunately, in 2023 these numbers could not be achieved, due to several concurrent factors: some upgrades could not be carried out (like the new source, see *New ^{22}Na positron source* (2.2.6)), or we were only able to implement them partially (as in the case of the n_{Ps}, see *New crystal for the EKSPLA to higher Ps Rydberg level excitation* (2.2.8)); some were not solid enough, and we could not benefit of the fully optimal conditions (e.g. the degrader: only the 1400 nm Mylar foil was used); lastly, a problem with the Ps target (see *Ps formation* (4.2.5)) seriously impaired the production of positronium.

Nevertheless, even in those drastically low-performance conditions, antihydrogen production was indeed achieved already in 2023, as better described in *Results* (5).

Item	Phase 1 (meas.)	Phase 2 (est.)	Improvement	Phase 2 (2023)
$\bar{p}$ amount	$2-3\ \% \times 3 \cdot 10^7$	$70\ \% \times 7 \cdot 10^6$	$\sim 6\times$	$70\ \% \times 7 \cdot 10^6$
$\sigma_{\overline{H}} \propto n_{Ps}^4$	$n_{Ps} = 17$	$n_{Ps} \geq 32$	$\sim 12.6\times$	$n_{Ps} = 21$
$e^+ \rightarrow$Ps efficiency	7 %	28 %	$\sim 4\times$	~ 5 %
New ^{22}Na source	24 mCi	50 mCi	$\sim 2\times$	6 mCi
Laser coverage	15 %	30 %	$\sim 2\times$	30 %
Total production	~ 0.05 $\bar{H}$/cycle	10÷100 $\bar{H}$/cycle	$\sim 2000\times$	~ 0.5 $\bar{H}$/cycle

Table 2.3 – This table summarises the various upgrades the AEgIS apparatus has undergone at the beginning of Phase 2, each with the calculated, or measured, improvement that should have on the $\bar{H}$ production rate. In the end, the total expected $\bar{H}$ rate increase is derived. The last coloumn evidences the expected formation for 2023, taking into account the suboptimal performance of some components.

Chapter 3

The CIRCUS

As introduced in *CIRCUS, the new AEḡIS autonomous control system* (2.1.15), the CIRCUS consists mainly of two parts: ARTIQ and TALOS (both presented in greater detail in the following sections), with the integration of ALPACA. ARTIQ is the high-level programming language for scripting the ns-precise routines to be executed by Kasli, which we empowered with libraries to streamline the coding of experimental routines and to integrate its operations with TALOS. In contrast, TALOS is the framework that constitutes the slow control: it provides the interface between the operators and the apparatus, and its flexibility makes it compatible with any range of hardware and control software units independent of their precise characteristics.

It is in the interplay of this ns-precise hardware control on the one hand and its full integration and automation of the surrounding experiment on the other hand that the CIRCUS control system manifests its strength in such a way that it can be applied to any experiment with similar requirements.

This interplay is evident especially when it comes to executing a sequence of measurements. In fact, the schedule of scripts (with parameters) is defined using the apposite TALOS interface (see *The Scheduler* (3.4.5.2)), and when the schedule is launched, it is TALOS that assesses if the conditions for running the experiment are met. If positive, it passes the command to Kasli, which executes the script, and TALOS remains available to forward calls from the running script to any part of the experimental apparatus. When the script terminates, the command passes back to TALOS, which, based on the outcome of the script, decides what action is to be taken – most of the time, running the same or the subsequent script in the schedule.

The integration of ALPACA has further empowered the CIRCUS to find in autonomy the best settings to optimise pre-defined observables. ALPACA is the AEḡIS framework for data analysis, that provides to the CIRCUS the capabilities of automation based on feedback on data acquired via a Bayesian optimisation algorithm.

In figure 3.1 the schematic of the CIRCUS control system is given, outlining the relationship of its constituent parts and their functionality, together with the connection with the other software and hardware components of AEḡIS.

The CIRCUS control system, and the TALOS framework, are available open-source in two *git* repositories (CIRCUS: 10.5281/zenodo.10371799; TALOS: 10.5281/zenodo.10371404).

3.1 Requirements

The design of the control system is driven by the requirements that this class of experiments has. A review of the literature was performed, to take some examples of atomic and quantum experiments [132, 133, 134, 135, 136], and relate their requirements to the ones derived from the experience of realising the first pulsed source of antihydrogen in AEḡIS [15]. The comparison

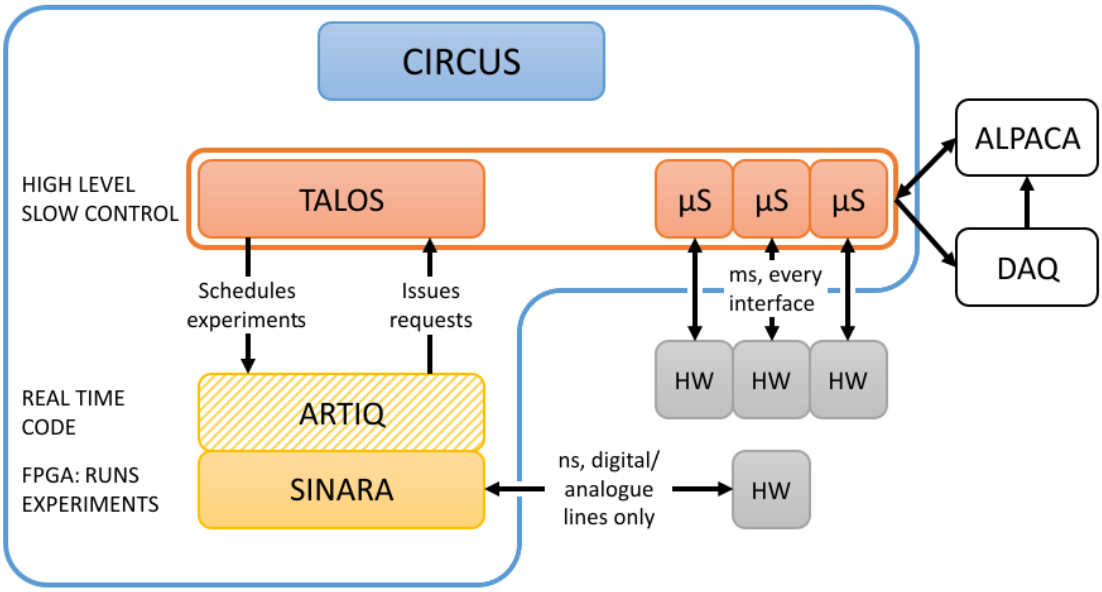


Figure 3.1 – Schematic of the CIRCUS control system and its constituent parts (ARTIQ/Sinara and TALOS, with various MicroServices (µService)), together with its relationship with other software and hardware subsystems (e.g. ALPACA).

showed that this class of experiments share similar requirements, which can be subdivided into four categories: interface requirements with the particle source; trap operations; particle and laser synchronisation; general slow control, data acquisition and networking.

Therefore, we decided to use the AEgIS requirements as a base for the design of the control system: the generality of these requirements renders a system satisfying them applicable to a broad range of tasks. In the following, their rationale is exposed, and the requirements are then summarised in table 3.1.

Requirements of the particle source interface: AEgIS obtains the antiprotons in bunches from the AD-ELENA complex (as explained in *Antiprotons production in AD* (2.1.4)). Consequently, the experiment is synchronised to the decelerator stack via a set of hardware triggers occurring at different times during each $\approx$ 110 s antiproton cycle: the AD injection trigger (occurring at the beginning of the cycle), the AD extraction/ELENA injection trigger (occurring $\approx$ 20 s before antiproton delivery), a bunch pre-arrival trigger (occurring $\approx$ 20 µs before antiproton extraction from ELENA) and a bunch arrival trigger (synchronous with the extraction from ELENA). The bunch is approximately 150 ns FWHM long. Antiprotons are delivered to the experiment at 100 keV energy, which is further reduced by a thin foil degrader (ca. 1400 nm of Mylar, see *Degrader(s)* (2.1.7)) to about 10 keV. Antiprotons are subsequently caught by means of a pulsed high-voltage Malmberg-Penning trap operated up to 15 kV in a 5 T magnetic field (see *5 T section* (2.1.5.1)). The timing of the trap has to be fine-tuned in $\approx$ 10 ns steps.

Requirements for trapped particle manipulations: a typical antihydrogen production sequence involves several manipulation steps of trapped particles (in the form of non-neutral plasmas, see *Antihydrogen production methods* (4)), performed with low-voltage electrodes of the Malmberg-Penning trap in the 1 T region (see *1 T section* (2.1.5.2)). These have to be controlled in the $\pm$ 200 V range, by arbitrary function generators. An accuracy of 10 mV or better is needed to allow for accurate potential ramps and thus enable measurements of the plasma space charge and temperature [137] as well as evaporative [138] and adiabatic cooling [139]. Standard manipulations in traps include both slow (several seconds) and fast (less than a millisecond) ramps, fast extraction of particles with $\approx$ 100 ns ($\approx$ 100 µs) pulses for electron (antiproton) extraction respectively, as

well as application of RF in the 1 kHz-100 MHz range for plasma heating or cooling and density control with the Rotating Wall technique [140]. Often, these procedures are combined, and the ability to synchronise events with the accuracy of 1 ns over several hours is required.

Requirements of particle and laser synchronisation: antihydrogen formation via charge-exchange reaction with trapped antiprotons requires the control of the times of positronium formation and laser excitation to its Rydberg levels at the ns accuracy level, as well as triggering the diagnostic scintillation and MCP detectors, as detailed in [120, 130]. This is due to the fact that the excitation laser has to be carefully synchronised according to its beam shape and position to obtain efficient positronium excitation. Hardware trigger lines allowing time adjustment of 1 ns or better and jitters of < 0.5 ns are required.

Slow control, DAQ and networking requirements: these include all the procedural sequences of trap initialisation, synchronisation on slow scales, computer responsiveness, data upload to the DAQ, etc., which admit a considerable jitter between the moment the command is issued and its execution and must not interfere with the experimental sequence (typically in the order of 100 ms). Network communication has to guarantee a smooth control flow: the communication speed among the various machines needs to be at least an order of magnitude faster than the timescale of PC operations.

3.2 ARTIQ

As explained in *Control system electronics* (2.1.14), the Sinara hardware relies upon the ARTIQ (Advanced Real-Time Infrastructure for Quantum physics) [129] language for straightforward, reliable software control. ARTIQ is a Python-based, high-level programming language which supplies specialised pre-generated functions for communicating with the hardware. The resulting control routines are formed by clear and short-run scripts, preventing long familiarisation phases of semi-experienced programmers and allowing for quick adaptations during data taking.

ARTIQ is designed to script experiments with nanosecond resolution and microsecond latency. To meet the requirements of real-time programming, ARTIQ code consists of two parts which can interact with one another: the first one, composed of regular Python commands, is executed on the host, while the *ARTIQ kernel* is executed on the CPU of the core device. This CPU can directly access a part of the *gateware*, i.e. the code specifying the configuration of the digital logic gates of an FPGA, responsible for specialised programmable I/O timing logic. A timeline, constituted by all programmed input and output events, keeps the synchronisation of the experimental routines: output events with a given timestamp are executed in a first-in-first-out (FIFO) mode when matching an internal, high-resolution clock, and input events are recorded with a stamp for the current clock value.

The ARTIQ environment includes a dedicated function to observe a given I/O TTL channel and register rising or falling edge events for a specified duration. A sequence of actions can then be performed within a deterministically programmed time window after receiving a trigger, one example of this being another ARTIQ function designed to set a specified voltage on a given Fastino channel. To control multiple different trap electrodes synchronously, the use of the provided Direct Memory Access (DMA) is essential, as it allows for pre-define RTIO sequences in the Kaslis' SDRAM, which can then be run directly by the FPGA core.

3.2.1 ARTIQ custom libraries: the AERIALIST

ARTIQ allows for a library-based approach to programming run routines of an experiment. Every script consists of the definition of a class, inheriting from the `EnvExperiment` class (which is the ARTIQ default parent class for every experiment), and in the override of the `build()` and `run()` functions, which respectively contain the code to be executed before the FPGA is started, and the actual experimental routine. Therefore, it has been natural to create a library structure to both

Particle source interfacing	
High-voltage catching potential range	0–15 kV
High-voltage potential accuracy	< 10 V
High-voltage gate raising edge duration	< 100 ns
High-voltage gate temporal tuning accuracy	< 10 ns
AD injection trigger synchronization	< 5 s
ELENA injection trigger synchronization	< 1 s
Bunch pre-arrival trigger synchronization	< 1 µs
Bunch arrival trigger synchronization	< 10 ns

Trapped particles manipulation	
Low-voltage potential range	$\pm$ 200 V
Low-voltage potential ramps duration	100 µs - 10 s
Maximum absolute calibration error	< 5 mV
Maximum noise band amplitude	< 1 $\mathrm{mV_{rms}}$
Fast pulse settling time	< 30 ns
Fast pulse duration range	100 ns - 100 µs
Fast pulse timing adjustment	< 10 ns
Fast pulse timing jitter	< 1 ns
RF frequency range	1 kHz - 100 MHz
RF amplitude range	10 mV - 5 V

Particles and laser synchronisation	
Positron/laser triggers time adjustment	< 1 ns
Positron/laser trigger jitter	< 0.5 ns
Detector arming timing accuracy	< 100 ms
Detector triggering timing accuracy	< 10 ns

Slow control, DAQ and networking	
PC–PC message delay	< 100 ms
Real-time–PC message delay	< 10 ms
DAQ run start/stop time	< 10 s
Data availability for online analysis	< 5 s

Table 3.1 – Summary of the different technical requirements set on the control system from experiment needs.

simplify AEgIS experimental procedures coding and to be able to release open-source as well the components necessary for the TALOS/FPGA integration.

Figure 3.2 shows the AERIALIST (Antimatter Experiment Realtime Integration of Artiq LIbraries and Sinara Technology), the library structure developed in ARTIQ/Python code that is used as the basis of the hardware communication of the presented control system. Each library contains all relevant functionality and methods to control one part of the entire experimental procedure: there are base classes for the communication with TALOS (*TCP Library*), the configuration of the hardware (*Build & Init Library*), error handling (*Error Library*), and overall often used functions such as the triggering of detectors or timeouts (*Utility Library*). Additionally, specialised classes govern trap electrode potentials (*Trap Library*), the feedback of data analysed online (*Analysis Library*), and specific parts of the experimental physics procedures and the different involved particles (*Physics Libraries*). Furthermore, a set of JSON configuration files has been created, which is used to set up the software to match the hardware, including calibration data, channel

mapping, and location information. The `_AegisExpOfficial` class is then created as a child of all of these parent classes such that each experimental script inheriting from this main class has all functionality readily available for use and can control all needed aspects of the experiment.

Of course, several of the functions, particularly in the lower experimental libraries, are specific to the AEgIS experiment and would need to be replaced by corresponding functionalities in other environments. On the other hand, the base functions in the *TCP Library*, used to interface with the TALOS part of the control system infrastructure, as well as the standard routines to configure and initialise the used hardware (with adapted configurations) and those general functions related to timing synchronisation, information logging, and data retrieval in the *Utility Library* and *Analysis Library* are re-usable as general functionalities of CIRCUS.

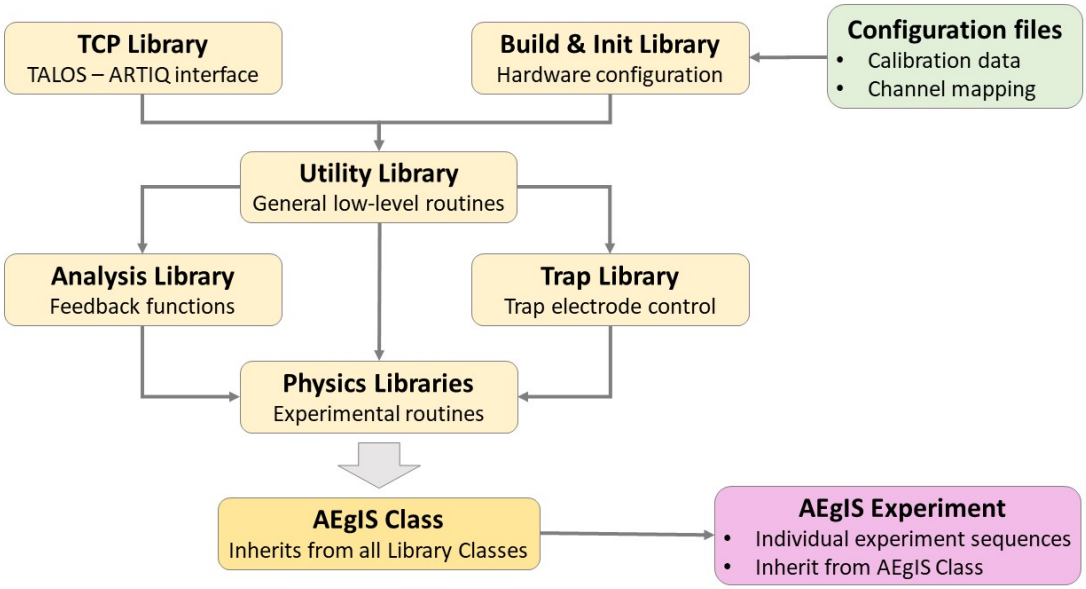


Figure 3.2 – Schematic of the AERIALIST, the ARTIQ/Python library structure of CIRCUS, as used in AEgIS. Each library defines a class, which all the experimental scripts of AEgIS inherit from. Most of the functions defined in the top-level libraries (*TCP*, *Build & Init*, *Utility* and *Analysis* libraries) are generic and could be utilised by other experiments as well.

The effect of the outlined library-based approach can be observed in figure 3.3, which shows a very simple experimental routine. In both cases, the resulting sequence is the same: the system waits for an incoming trigger signal on one of the digital I/O lines and subsequently produces a voltage ramp to 1 V on three of the Fastino channels (which is amplified to 20 V by the corresponding amplifier units). The application of the calibration constants for the amplifier boards is included in the function used in the routine on the right. The functionality to set up and initialise the used hardware, which is part of the first two function definitions on the left, is included in the standardised `Build()` and `Init()` functions on the right. All other functions defined explicitly in the script on the left are included in the library structure and available without re-definition to all experiment scripts. This means that only one additional function call is needed in the actual run routine shown on the right side to achieve the same result as the code on the left.

The use of the AERIALIST reduces the ARTIQ script to a few lines of code when importing the parent classes and yields an immediate, simple overview of the routine. This effect is rendered more and more distinct the more complex (and closer to realistic run sequences) the experimental routines become.

The simplification generated by this library approach also renders it possible for people external to the main developing team to easily and quickly use the system: as an example, by using the

high-level functions, two researchers who wanted to develop a new sensor for antiprotons managed to write a script to successfully manipulate the captured antiprotons in a novel way in only two days.

```
from artiq.experiment import *
from artiq.coredevice.kasli_i2c import port_mapping

class HVAamp_Trigger(EnvExperiment):

    def build(self):
        self.setattr_device("core")
        self.setattr_device("fastino0")
        self.setattr_device("ttl0")
        self.setattr_device("ttl_hvamp0_sw1")
        self.setattr_device("ttl_hvamp0_sw2")
        self.setattr_device("ttl_hvamp0_sw3")
        self.setattr_device("dio_mcx_dir_switch")
        self.setattr_device("i2c_switch0")
        self.dio_mmcx_i2c_port = port_mapping["EEM0"]

    @kernel
    def set_dio_outputs(self):
        self.i2c_switch0.set(self.dio_mmcx_i2c_port)
        self.dio_mcx_dir_switch.set(0b00000001)
        self.core.break_realtime()
        self.ttl0.input()
        self.core.break_realtime()

    @kernel
    def SignalAtTrigger(self):
        t_gate = self.ttl0.gate_rising(120*s)
        t_trig = self.ttl0.timestamp_mu(t_gate)
        at_mu(t_trig)
        delay(10*us)
        self.fastino0.update(1<<3|1<<2|1<<1)

    @kernel
    def SetVoltages(self):
        self.fastino0.set_dac(1, 1.0)
        self.core.break_realtime()
        self.fastino0.set_dac(2, 1.0)
        self.core.break_realtime()
        self.fastino0.set_dac(3, 1.0)
        self.core.break_realtime()
        self.SignalAtTrigger()

    @kernel
    def run(self):
        self.core.reset()
        self.fastino0.init()
        self.core.break_realtime()
        self.fastino0.set_hold(1<<3|1<<2|1<<1)
        self.SetVoltages()
```

```
import sys
sys.path.insert(1, 'C:\kasli-code\Libraries')
from AEgIS_imports import *
from AegisExperiment import _AegisExpOfficial

class HVAamp_Trigger(_AegisExpOfficial):

    def build(self):
        self.Build()

    def run(self):
        self.Init()
        self.SetVoltagesAtTrigger("Trigger", 10*us,
            "Channel1", 20.0, "Channel2", 20.0,
            "Channel3", 20.0)
```

Figure 3.3 – Left: Experimental routine to set a specified output voltage on three amplifier channels of the Sinara hardware system after an incoming trigger pulse, programmed in the ARTIQ environment. Right: The same experimental routine as on the left, reduced to a few lines of code when implementing library-based programming.

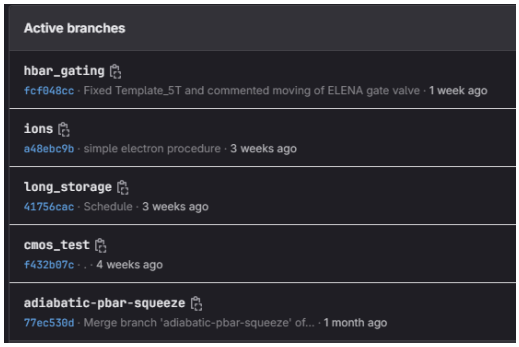
3.2.2 *git* branches

Another great improvement has come from the widespread adoption and use of *git* for the entire control system and for the totality of the experimental scripts. In fact, in AEgIS, already during the Phase 1 some pieces of software were handled using *git* (for example, the GEANT4 MC analysis software), but it was a solution employed by only a part of the collaboration and for a limited scope. My push has brought the use of *git* to be widespread, with the advantages mentioned in this section.

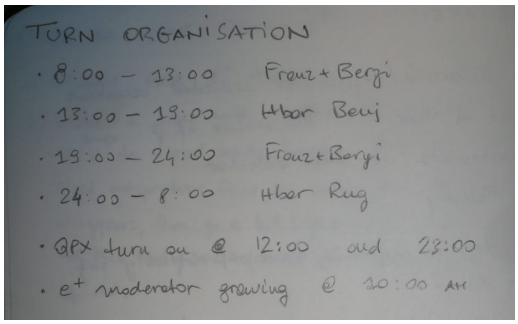
The code on the machine from which the scripts were usually run was mirrored to a repository held on CERN *GitLab* servers. The adoption of this working modality led to multiple advantages: first and foremost, the commits history keeps track of the code implemented, and in case of issues arising that were not easily solved, reverting to a previous state of the system was straightforward, also easing the pinpointing of problems. Then, different colleagues could develop their piece of

the system on their own branch, on their computer, and only when it was tested and finalised, it was merged into the main experiment branch by the repository maintainers, who overlooked the operation and ensured that no inconsistencies would arise. But most importantly, the adoption of *git* brought the possibility of developing and running in parallel different experiments, each on its own branch, without any interference among them: a simple branch switch was necessary to pass from one to the other. Again, after development was terminated, the repository maintainer would take care of the unification of the branches, porting into the AERIALIST the new features developed.

This working modality was leveraged heavily, especially during the final part of the antiproton campaign of 2023, where several different lines of research were advanced together (developing a new CMOS sensor, manipulation of antiprotons for $\bar{H}$ formation, long storage of high amount of antiprotons, antiprotons recycling, positronium formation optimisation, HCI formation, among others), by diverse research teams. In figure 3.4a the most active branches during that period are shown, and in figure 3.4b the turns of experiments switch during a single day, which exemplifies perfectly the flexibility given by the approach.



(a) Main Git branches alternated during November 2023 data taking.



(b) Data taking turns.

Figure 3.4 – On the left: the most active branches in November 2023, each of them represents a different line of research developed in parallel. On the right: the turns for alternating among different experimental endeavours, during the same day.

3.2.3 Error handling inside the experiment scripts

Normally, ARTIQ scripts are meant to be executed by a user from a command shell. However, in the CIRCUS system, the handling of the FPGA is entirely operated by the TALOS system (see *ARTIQ integration* (3.4.4) for the details).

One crucial aspect of the stability of the control system is the resistance to errors. In fact, a consistent, centralised, and dependable error management system guarantees that no exception in the system can pass unnoticed, taking the appropriate action to assure repeatable data acquisition and apparatus safety (see *The Error Manager* (3.4.5.1)).

Clearly, the errors encountered during the execution of the experimental scripts are no exception to this rule. A custom exception, called `Banana`, was defined to ensure a system-wide consistent response. It can carry, together with the values inherited from the standard Python `Exception`, a numerical value from 0 to 4 representing the *Criticality* of the exception raised (the concept of *Criticality* in the CIRCUS is explained in section *The Error Manager* (3.4.5.1)). The CIRCUS *Error Manager* actively looks for these particular exceptions during the execution of an ARTIQ script and acts accordingly.

Nevertheless, the code in the script positioned after the exception would not execute. If the error is severe enough to stop the execution of the schedule, some activated devices could remain

on for extended periods, potentially causing hardware damage (an example is the high-voltage power supplies). Therefore, the structure of the `run()` routine was coded explicitly directly in the AERIALIST by defining two new methods, `experiment()` and `closure()`, to be overridden by each experimental script, and encapsulating them into a `try...except...finally` clause. This structure guarantees the execution of the code present in `closure()`, independently from errors potentially thrown by the rest of the experimental routine. Together with it, it also ensures the completion of `StopRun()`, containing the code necessary for TALOS to understand the termination of the script correctly and its return code. In turn, this enables the correct handling of exceptions, differentiating between a user-defined `Banana` (or `KernelBanana`, see below), carrying *Criticality Code* details, and an unhandled native `Exception` [1].

A problem was still present, though: the ARTIQ framework is coded such that whatever error is generated by code running on the kernel, the latter halts, even if the exception is raised inside a `try` clause. This behaviour was causing the execution of the code in `StopRun()` to be impossible, since it contains a kernel call. Therefore, the `Banana` exception cannot be used inside code to be executed on the kernel, hindering the capability of error-controlled programming. A solution was found by raising a specific ARTIQ exception, `RTIOOverflow` [2] , which does not halt the kernel execution. By embedding in its message a string carrying the *Criticality Code* and creating a dedicated `except` clause in the main `run()` method of the `_AegisExpOfficial` class, the ability to generate user-defined errors on the kernel was recovered.

The code of the `_AegisExpOfficial` class is shown in figure 3.5.

As a final addition, in the extreme case where an actual kernel error gets raised, therefore halting the execution of the kernel, and impeding the correct completion of the `StopRun()` routine (which would disrupt the saving of the data acquired in the subsequent Run), the TALOS system is instructed to run a dedicated script, containing only the code present in the `StopRun()` method, before proceeding into the evaluation of the errors encountered (and at least skipping the script, otherwise the behaviour would likely repeat). This way, everything goes as the script would have correctly terminated, and the schedule can continue normally.

I carried out this endeavour to guarantee that even new scripts, poorly tested, could be safely scheduled during the night time without hindering the execution of the remaining scripts or jeopardising the safety of the apparatus.

3.2.4 Kaslis ns synchronisation

To control the totality of the electrodes of all the traps inside the main vacuum chamber, a total of three Sinara crates have been necessary, to host all the needed Fastinos and amplifiers. To perform antihydrogen formation, the three crates must operate synchronously, with a precision of ns between them, in order to not disrupt the particles' manipulation.

A method could have been to use the Kasli in one crate as master and configure the other two as satellite devices, by daisy-chaining them with optical fibre links. This technique would have been straightforward to implement, but it would not have been scalable in the future when other Kaslis will be integrated into the AEḡIS apparatus (for example, to manage the *Starship*, or to migrate the current positron system electronics). For this reason, the Kaslis were configured with one master for the 1 T, and one master and one slave for the 5 T region, respectively.

Consequently, to perform antihydrogen production, it was necessary to synchronise the two masters with ns precision: and not only at the very beginning of the script, since the necessary operations have different and not fixed lengths (e.g. the waiting for antiproton injection). TALOS cannot provide such a functionality, since it is bound to typical computer time scales (O(ms)).

[1] Which, by default, is interpreted as carrying *Criticality Code*= 2 (i.e. "Skip"): in fact, it means that the programmer has incorrectly handled an exception in the code, so the script is not safely executable. Hence it has to be not try again.

[2] Together with `RTIOUnderflow`, is the only exception exhibiting this behaviour.

```
class _AegisExpOfficial(*Libraries):

    def experiment(self):
        pass

    def closure(self):
        pass

    def build(self):
        pass

    def run(self):
        try:
            # Start procedure
            self.Simplified_Init()
            self.StartRun()
            # Experiment routine
            self.experiment()
        # Exception handling and script closure
        except Banana as error:
            self.ExceptBanana(error)
        except RTIOOverflow as error:
            self.ExceptKernelBanana(error)
        except Exception as error:
            self.ExceptException(error)
        finally:
            # Post-measurement cleanup
            print("Starting closure procedure...")
            self.closure()
            # Closing
            self.StopRun(self.RetCode)
            return
```

Figure 3.5 – The `AegisExpOfficial` class, parent of all AEgIS experimental scripts. A `try... except... finally` clause is used to ensure the execution of the `closure()` method, independently from any error arising from the `experiment()` function. This structure also correctly enables user-defined errors (`Banana` and `BananaKernel`), which can carry *Criticality Code*, so that TALOS can correctly interpret it and take the appropriate action.

The solution was found in the creation of a barrier function, `Kasli_Sync()`, which has to be called by both Kaslis' scripts when a (re-)synchronisation is wanted. The code after this function is executed only when both Kaslis have finished executing it. Hardware-wise, a digital input and a digital output have to be dedicated to the functionality on each Kasli, and connected mutually, so to ensure a direct bi-directional low-level communication.

To briefly explain how the scheme works, let us imagine having a setup with two Kaslis called *A* and *B*. When the function is executed on Kasli A, it sends a pulse to Kasli B on the output line and then waits, listening on the input line for a reply from Kasli B. If Kasli B replies with a single pulse, it denotes that the first pulse was not received, and Kasli B has just entered the `Kasli_Sync()` function. Nevertheless, Kasli A acknowledges the reception by sending two pulses back to Kasli B, waits a pre-defined amount of time (empirically determined), and exits the function, knowing that Kasli B will exit simultaneously. On the contrary, if Kasli A receives a series of two pulses immediately after sending the initial pulse, it deduces that Kasli B was already waiting for the handshake to occur, and, as a result, it has already started exiting the routine. Therefore, after another empirically determined delay,Kasli A exits the `Kasli_Sync()` function and continues executing the subsequent code.

In this way, only one digital line-in input and one in output are necessary to ensure ns synchronisation between two Kaslis. Furthermore, it guarantees the scalability and flexibility necessary

for the AEgIS experiment to advance in its scientific goals. This solution can be extended to synchronise more than two Kaslis by defining one particular Kasli as server and all the others as clients and modifying slightly the aforementioned function.

In addition, if coupled with the possibility of sending messages among Kaslis enabled by TALOS (see *Kaslis' Server* (3.4.5.10)), arbitrarily complex operations could be performed synchronously, which ultimately led to the creation of the *Particle Server* (see following) and enabled the parabolic potential shape for antiprotons ballistic transfer (see *Antiprotons transfer* (4.2.4)).

3.2.5 Particle Server

Building on what was done for the Kaslis synchronisation, the coding and execution of the 5 T Kasli was also streamlined. A new experimental script was created, called *Particle Server*, which waits indefinitely for messages to arrive from other Kaslis (the 1 T, in this specific case), each corresponding to a specific task to be performed (e.g., "Capture Antiprotons", or "Dump $\bar{p}$ on degrader"). Every time a recognised message is received by the *Particle Server*, it executes the corresponding task, giving back an acknowledgement when completed, and then it goes back to listen.

This approach allowed disentangling logically the operations to be performed on the 5 T side of the traps, which were consolidated by two years of tests and improvements, from the operations for the formation of antihydrogen, taking part on the 1 T part of the trap and still in full development. As an example, in figure 3.6 a script for the 1 T Kasli is presented, assuming the *Particle Server* running on the 5 T one. The script loads electrons in the *capture trap*, catches antiprotons, lets them electron-cool for 30 s, reshapes the trap to dump the cold $\bar{p}$ on the MCP, and performs the dump. Each message executes a lengthy procedure, masking its complexity and making the script written almost in natural human language: these benefits have proved fundamental during the development of the procedures for $\bar{H}$ formation.

```
class dump_pbar_mcp_with_particle_server(_AegisExpOfficial):

    def closure(self):
        self.Send_Msg_Blocking("5TC1", "reset_trap", timeout = 300)

    def experiment(self):
        # Catch pbar, electron-cool for 30 s and dump on degrader
        self.Send_Msg_Blocking("5TC1", "wait_injection", timeout = 300)
        self.Send_Msg_Blocking("5TC1", "load_electrons", timeout = 300)
        self.Send_Msg_Blocking("5TC1", "load_antiprotons", timeout = 300)
        self.DelayHost(30*s)
        self.Send_Msg_Blocking("5TC1", "reshape_cooling2dump", timeout = 60)
        self.Send_Msg_Blocking("5TC1", "dump_mcp", timeout = 60)
```

Figure 3.6 – An example of a script for the 1 T Kasli using the *Particle Server* on the 5 T one. The script loads electrons in the *capture trap*, catches antiprotons, lets them electron-cool for 30 s, reshapes the trap to dump the cold $\bar{p}$ on the MCP, and performs the dump. Each message executes a lengthy procedure, masking its complexity and making the script written almost in natural human language.

This also goes in the direction of managing the two Kaslis more asynchronously in the future. Essentially, the 5 T Kasli can operate the *Particle Server* in a continuous catch-and-accumulate-antiprotons mode, while the 1 T one can execute a list of different experiments, requesting $\bar{p}$ from the 5 T Kasli when needed. Furthermore, this scheme can be easily scaled by creating other "Servers", e.g., for the positrons, and the ions.

3.3 ALPACA

Every experiment has the need for a series of algorithms to analyse the obtained data. Often, part of the data analysis is used to tune and improve the subsequent data acquisition: the capability of a control system to perform this task autonomously is of great advantage to the scientists.

All Python Analysis Code of AEgIS (ALPACA) is a Python data analysis framework written specifically for the AEgIS experiment's different physics tracks. It leverages the functionality of the *NumPy* [141], *SciPy* [142] and *Plotly* [143] libraries to collect and transform the raw data acquired by all the detectors into observables, which can then be utilised by the scientists to perform dedicated studies. Figure 3.7 depicts the framework's linear architecture, where pipelines transform the data into different processing states.

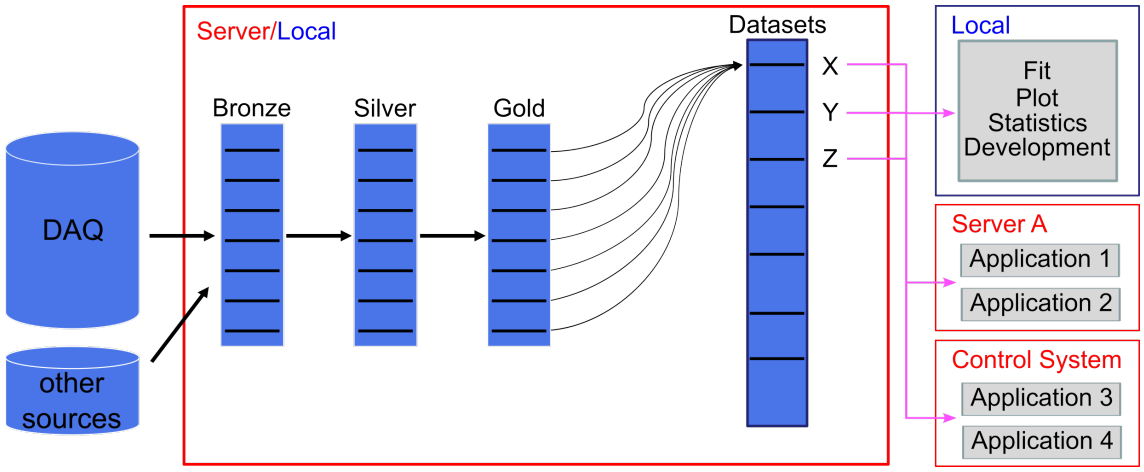


Figure 3.7 – Representation of the architecture of the ALPACA analysis framework, including the stepwise processing of the data as well as the local or server-based deployment.

First, all raw sources of an experiment's data, stored on different servers and in different formats (e.g. ROOT, JSON, png, txt, etc., either originally plain or zipped) are concatenated into a unique Python dictionary cointaing all the raw data, therefore called to be in a *bronze state*. Raw sources include the data of each detector triggered, the settings of the detectors and the environmental data (for example, temperature and vacuum readings) during the experiment. At this stage, the originally stored files are just saved in a Python native format but no data manipulation is applied.

The data dictionary is then refined, passing from the bronze to the silver state: the data is restructured depending on how each detector stores the acquired data according to its own configurations. For example, the JSON files for the acquired voltage readout of the MCP detector always contain, as the first and second entries, the start time of the acquisition and the time increment, while the remaining entries hold the actual voltage readings after each time increment. In the bronze $\rightarrow$ silver pipeline, these data are parsed such that the start time, the time increment and the voltage readout become variables accessible on their own. Moreover, a three-layer nested data structure is established with the detector on the top, the acquisition number in the middle and the acquired data and run-specific configurations of the detectors at the deepest level.

Subsequently, the silver $\rightarrow$ gold pipeline computes and appends observables for each detector and acquisition, still preserving all the original data in the gold state as well. For example, in this step, the image taken from the MCP camera is first normalised for the set gain of the MCP, then the background is evaluated and subtracted, before the sum, mean and standard deviation of all pixels are calculated and made available as three different observables.

In the last step, user-specified datasets of observables over many experiments are concatenated and made available for the user's personal analysis as well as for applications. Additionally, a

dedicated package for the generation of statistics fits and plots as well as for the training, evaluation, and use of machine learning models using the generated datasets has been developed.

Thanks to the single end-point for querying datasets from ALPACA as well as the independence of the pipelines from each other, ALPACA is easily scalable in the number of applications as well as in the data sources and processing pipelines. Special emphasis is put on the scalability and reusability of the source code, which allows the seamless integration of new detectors installed at the AEgIS apparatus as well as new analysis pipelines.

Table 3.2 includes samples of the current runtime performance on a set of 177 experiments (constituting the entire dataset of a parametric scan taken during the antiproton beam time of 2022), which produced an average of 21.4 MB of raw data. As can be seen, a significant speed-up in development and analyses is achieved by reloading the data from the different processing states. Loading the data from "Raw" takes exceptionally long due to the necessary download from the AEgIS servers, while the locally stored datasets are available almost instantaneously. Processing the data of a single experiment usually takes a few seconds, which is feasible for feedback loops with the control system.

Number of experiments	Loading times, in seconds, from:			
	Raw	Bronze	Gold	Datasets
1	3.96	1.42	0.22	0.009
10	36.20	15.86	2.33	0.015
100	380.27	176.76	24.47	0.06
177	658.81	310.85	46.04	0.09

Table 3.2 – Runtime performance of the analyses framework generating observables using the experimental data from a parameter scan during the antiproton beam time 2022. These times are characteristics of the AEgIS system.

The integration of ALPACA extends the capabilities of the CIRCUS control system beyond simply scheduling experiments, giving the possibility to perform, in full autonomy, optimisations of observables over a given parameter space. The linear and scalable architecture allows for easy integration of new observables depending on the envisioned experiments, and the automated deployment makes these observables accessible to the control system without any user interaction. ALPACA has been heavily leveraged in CIRCUS to render the control system capable also of taking decisions driven by online data feedback, as shown in *Automatic parameter optimisation* (3.4.7.3) and *Quality of Run assessment* (3.4.7.4).

In the framework of the presented control system, ALPACA is a powerful tool to aid automation and enable self-optimisation, and it is used as the main analysis framework in AEgIS. In principle, its design serves as a foundation. However, different software architectures that fulfil this purpose can also be used in its place. In particular, the capability of CIRCUS to autonomously modify the experiment parameters based on the feedback loop from the data taken relies on a simple interface with the analysis framework. It consists of two shell commands: one for retrieving the last measured value of a specified observable, and another one for receiving new parameters to use, given a list of parameters used and results obtained. Any analysis framework capable of producing such a simple interface would be straightforwardly integrable in CIRCUS.

3.4 TALOS, the framework

As stated in section *CIRCUS, the new AEgIS autonomous control system* (2.1.15), to address the common problems of nuclear, atomic and quantum experiments' control systems, it has been decided to devise an experiment-agnostic framework which would serve as the common denominator to all the various components of the entire control system. It was required to be general enough so

that every conceivably needed application could be written in it, inherently safe to operate under all circumstances, scalable, future-proof and reliable. Moreover, to minimise the required time for operators to run it, it should be capable of a great amount of automation.

From all these necessities, the framework should be modular, in order to be easily expandable, self-checking, to be reliable, and able to operate concurrently on different machines in a harmonious manner. These requirements paved the way for the foundations of the TALOS framework. We refer to TALOS as a framework because it does not only come with the functionalities described in this section, but it also creates a specific way of coding, in the form of guidelines to write µServices.

Fundamental to its development has been the adoption of *git*[3], a distributed version control system, that enables agile development. Every time a new feature is to be added, a new branch is created, and the development is merged into the main branch only after rigorous testing. This modality of working has sped greatly the construction of the system, guaranteeing the functionality of the code in the main branch, used only for the real running system. Furthermore, it eases collaborative work, by enabling multiple developers to work on different branches, which are then merged in the main one by the maintainers, consistently assuring code integrity.

3.4.1 The two main pillars

To meet the aforementioned necessities, the TALOS framework is founded upon two main pillars, which are the "everything is a MicroService" (or, simply, µService) approach and the distributed system architecture.

The first concept is to divide the code into independent and autonomous parts, called µServices, each with a clear scope and task. The µServices are meant to be separate applications running asynchronously side-by-side, interacting with each other via a built-in messaging system. They all inherit from a common class, called *Father Of All µServices* (FOAM), which ensures, at the same time, code uniformity and maintainability: in fact, updates to the framework can be pushed "from the back", dealing at the same time with the back-end and the FOAM, but leaving the code, written by any user and specific to each µService, untouched.

The second concept comprises the idea of having an identical process, named Guardian, running on every machine, which monitors both the status of the other Guardians present on the network and the µServices running locally. Moreover, the Guardian supplies a common infrastructure to share messages and data between various µServices and among different computers. On the one hand, it provides a way to unify all the computers as a single, distributed, entity, and on the other hand, it enhances the stability, reliability, and safety of the system by having a distributed watchdog system, so that no single computer becoming unresponsive can pass unnoticed. The unification of all the computers into a single entity is the key feature that enables the possibility to automatise the entire system: in fact, if every PC runs independently from the others, a single computer cannot react to a problem occurring to another PC on the network, because it is unaware of it. Furthermore, the possibility of making high-level decisions often depends on parameters generated by different computers, therefore the automation of such processes again requires the sharing of information.

To fulfil these requirements, it has been decided to base the TALOS framework on the NI (*National Instruments Corp.*) LabVIEW™ Actor Framework, a LabVIEW™ implementation of the *Actor Model*, both of which are briefly introduced in the following subsection.

3.4.2 The Actor Model and NI LabVIEW™ Actor Framework

The Actor Model [17] is a computational model that offers a flexible framework for designing and implementing concurrent and distributed systems. It provides a way to organize and structure software components in a manner that promotes scalability, fault tolerance, and responsiveness. In

[3]Similarly to git *branches* (3.2.2).

the Actor Model, computation is centred around autonomous entities called actors. Each actor is a separate process with its own state, behaviour, and message inbox, and it works by executing serially the messages present in its own queue, typically in a First-In-First-Out (FIFO) manner. Actors communicate by passing messages to each other, and these messages can contain data or instructions. Importantly, actors can create new actors, allowing for dynamic, hierarchical systems. Actors operate asynchronously, meaning they can process messages independently without being blocked by other actors: this characteristic makes the Actor Model particularly suitable for systems with high concurrency and responsiveness requirements. As actors can work concurrently, they can achieve efficient utilization of resources, making it possible to handle large workloads and scale systems easily.

One of the key benefits of the Actor Model is its decentralized nature. Actors interact solely through message passing, eliminating the need for shared memory and global synchronisation mechanisms. This design choice simplifies the development of distributed systems and enhances fault tolerance: in fact, if an actor encounters an error or fails, it does not stop the functioning of the overall system, as other actors can continue to operate independently. The Actor Model also provides a natural way to handle concurrency and synchronisation. As actors process messages one at a time, they ensure that their internal state remains consistent, avoiding data races and other common concurrency issues. Moreover, the message-based nature of communication allows for easy isolation of actors, making it easier to reason about and test individual components.

In the LabVIEW™ programming language [144] a native implementation of the Actor Model is offered, called the *Actor Framework*. Actors in LabVIEW™ are represented by individual Virtual Instruments (VIs) that encapsulate both data and behaviour. The Actor Framework provides a structured way to create, manage, and coordinate these actors, making it easier to develop complex systems. The framework manages the execution of actors and ensures that messages are processed in a controlled and synchronised manner: this enables developers to take advantage of multi-core processors and distributed computing architectures, making their applications more efficient and scalable. The Actor Framework also provides features for handling fault-tolerance and error recovery: actors can be designed to handle errors locally, allowing the system to recover from failures gracefully; otherwise, the framework includes mechanisms for supervision and error communication, enabling the detection and handling of errors across multiple actors. Additionally, the Actor Framework promotes modularity and code reusability, by leveraging the power of object-oriented programming (each actor is a class). In fact, by encapsulating functionality within actors, developers can create independent and reusable components: this modular approach simplifies application design, maintenance, and testing, leading to more robust and maintainable software.

The characteristics of the Actor Model (and the Actor Framework) made it the ideal starting point for the development of the control system, following the two aforementioned pillars. As a matter of fact, it is, by nature, suitable for building distributed systems, and the actor features (modularity, independence, asynchronicity) are exactly what was envisaged for the µServices.

3.4.3 TALOS structure

As mentioned before, the TALOS framework is based on the LabVIEW™ Actor Framework, since it is designed for the implementation of multiple asynchronously interacting processes. The actors hierarchy is very simple: on each PC, a single *Root Actor* exists, called the Guardian, and all the µServices are *Nested Actors* of it (see figure 3.8). This flat organisation simplifies the management of the various components, and makes the system more resistant to failures, since the stopping of one µService does not affect its siblings.

The power of class inheritance is used for the µServices, since *Father Of All µServices* (see *The Father Of All Microservices (FOAM)* (3.4.3.2)) is the parent actor of all the µServices in the system. It both masks the Actor Framework complexity to the end-developers and manages the interaction of the µService with the Guardian. Moreover, it guarantees the maintainability of the system: in

fact, all the µService-specific code inserted into each child of FOAM remains untouched during updates to TALOS, since only the FOAM specific VIs are modified.

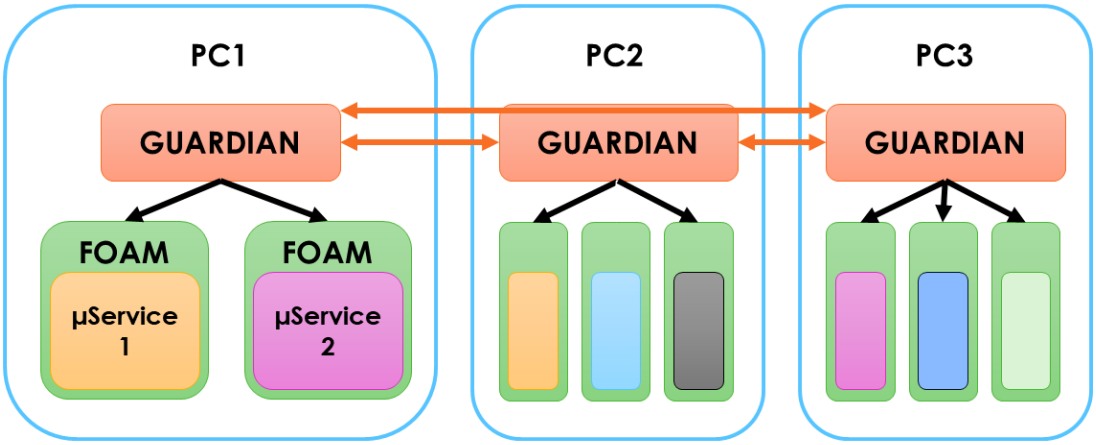


Figure 3.8 – The TALOS structure: on each PC, an instance of Guardian is launched, which in turn launches all the µServices that should run on that machine. The Guardians maintain the communication and watchdog layers that unify the distributed system into a single entity (enabling, for example, each µService to message any other µService, regardless of the computers they are running on, possibly different), and they monitor the local µServices for unresponsiveness. The interaction between each Guardian and its own µServices is coded into FOAM, the parent class of every µService.

3.4.3.1 The Guardian

The Guardian main feature is to act as the *Root Actor* of all the µServices on each machine: the actors' hierarchy has been chosen to be completely flat because, while it does not limit in any way the development and the growth of the system, it simplifies the interface between each µService and the Guardian. In practice, the Guardians are the real core of TALOS: in fact, each Guardian establishes a network connection with all the other Guardians present in the experiment, providing the messaging system among them and µServices, so that they can interact as if they were on the same computer (for more detail, see 3.4.3.3).

Furthermore, the Guardian is also responsible for running the three different watchdogs that are the base of the reliability of TALOS. They consist of:

- Guardians Watchdog: it consists of each Guardian sending periodically a special message to every other Guardians. Every Guardian collects all the messages and checks the times of arrival of the last received messages: if too old, the corresponding Guardian is marked as unresponsive and the system is put into *Safe Mode*. The *Safe Mode* is a special state of the system where everything is put to idle, so to minimise eventual damage caused by e.g. a hardware fault. Manual intervention is needed to bring the system out of this state;

- µServices Watchdog: similar to the Guardians Watchdog, all the µServices sends to their root Guardian a periodic message, whose timestamp oldness gets checked. If the check is not passed, the Guardian tries to restart the µService in question, by terminating it and relaunching it. If the µService is successfully relaunched, the measurements are restarted with no need for the user's input.

- `ABORT` Watchdog: to ensure the safety of the distributed system, a boolean *Shared Vari-*

able (SV)[4], called `ABORT`, is present in the system, and kept as a singleton in the network. It can be set to true any time, by whichever Guardian encounters a critical error that requires the entire system to halt and go to safe mode (e.g. finding a Guardian not online). The change of this SV immediately generates an event on every Guardian, so it is a very effective method to propagate urgently that a major issue has arisen. But keeping it as a single element in the entire network is far from trivial: to ensure this property, a third periodical watchdog has been coded. This watchdog scans periodically the entire network for the presence of the `ABORT` SV, in order to assess its accessibility and presence. In case the variable is not found, a new instance is deployed locally, and raised; instead, in case more than one instance is found (which may happen if, for example, a computer gets briefly disconnected from the network), the Guardian who is later in the alphabet (by computer's name) will un-deploy its copy. Although it is complex to keep the variable as a singleton across the network, with this solution at least one copy is always present and its marking as true is univocal[5].

Each Guardian does not assume that every PCs present in the network shall be interpreted as making part of the experiment: the actual experiment's network boundary is defined once in the CIRCUS configuration file, common to all the PCs.

3.4.3.2 The Father Of All Microservices (FOAM)

All the µServices are children of the *Father Of All µService*, or simply FOAM, and therefore they are all *Nested Actors* of the Guardian, like FOAM, completely separated from each other, running asynchronously side-by-side. This inheritance structure has several advantages: it masks completely the complexity of the Actor Framework, and of the rest of the internal mechanics of TALOS, to the end-user/developer; it gives the possibility of updates to the framework without affecting the per-µService specific code; and it automates and/or enables several functions via dedicated methods.

By inheriting from FOAM, all the µServices become structured as a *Queued Message Handler* (QMH), i.e. the combination of a part reacting to internal or external events (like a button pressed on its Graphical User Interface (GUI), or a message from another µService), generally referred to as "event handler", and a "consumer", which executes the queue of tasks populated both by preceding tasks or by the event handler part.

To host the µService-specific code, each µService has to override a certain number of VIs, each serving a specific function, as follows:

- <µService Name>.ctl: This is the µService actor/class private data container.
- *Init*: This VI is executed before the µService is started: it is useful to perform initial checks and to populate the values of the private data of the µService.
- *Close*: This is used to safely close everything that was opened in *Init* and during the µService execution. This VI is guaranteed to be executed during the µService shutdown procedure.
- *Consumer*: This VI represent the Consumer loop of a QMH. It executes the messages in the internal queue of the µService, dequeuing them in the order they are sent and according to their priority. In general, no case should take more time than the value of the *µService Watchdog Timeout* variable, otherwise, the µService will be flagged as unresponsive and force-restarted.

[4]It is a LabVIEW™ tool to share a variable across the network, making it available to all the PCs with the right to access it.

[5]Another solution for the uniqueness of `ABORT` could have been "the absence of abort is abort itself". The idea is to define a priori a machine which should deploy the variable: if it is not found by another Guardian, it means that the chosen PC is down or unreachable, and therefore interpreted as if `ABORT` would be true. Albeit of simpler implementation, this solution would break the symmetry of the distributed system and therefore was discarded.

- *Event Handler*: Every time an event related to the µService is triggered, this VI is called. The event types available are:
  - Button Pressed: it gets called every time a button is pressed on the µService GUI.
  - Shared Variable Updated: it is executed when a SV the µService subscribed to is updated.
  - Message: this case defines which messages the µService will react to. The messages can be sent internally using *Send Message to Event Handler*, or externally with *Send Message to µService*.
  - Ready to launch?: it is called at the beginning of each Schedule to assess each µService health. Each µService is responsible for performing a self-check and sending the results to the *Monkey*, which will decide if the Schedule should start (see *The Monkey* (3.4.5.3) for more information on Run execution).
  - Start Run: this case is executed every time a new Run is started.
  - Stop Run: it is run every time a Run is terminated (successfully or not).
  - Safe Mode: this case is called from the Guardian when a critical error has been detected, and all the subsystems need to be put in a safe state. The case should put the µService in a safe state, dormant but from which the normal operation can be resumed.
- *µService GUI*: The Front Panel of this VI is the actual GUI of the µService. It is displayed inside the main GUI when the corresponding µService is selected.
- *Process Error*: This VI gets called every time an error is generated in whatever part of the µService. This VI should serve as a first filter of the errors, where a tentative solution can be provided. If an error is left going out from this VI (which is the default case), it is forwarded to the distributed system *Error Manager*.

In figure 3.9 is schematised how an equivalent QMH VI would look like to mimic the behaviour of a µService. However, the actual internal mechanics is completely different, because it leverages the functionality of the Actor Framework: so, for example, the *Consumer* is a dedicated actor message, the *Event Handler* is inserted into a dedicated separated loop inside an override of *actor-core.vi*, and the *Consumer* messages queue is the internal actor one.

The execution of multiple tasks is eased by the presence of several methods, which every µService inherits from FOAM. They are available as a dedicated palette that can be installed in LabVIEW™ (see figure 3.10), to be quickly accessible during the coding of any µService. In the following, a list of the most important methods is given.

- Message to Event Handler: This VI generates a user event to be handled by the *Message* case of the Event Handler.
- Get GUI Item Value: Returns the value currently present in the µService GUI Control. The correct datatype to output can be selected (Variant, Bool, Str, Dbl, Int).
- Send to DAQ: To send data to the Data Acquisition System.
- Update GUI: Use this VI to update the value of an Indicator of the µService GUI Front Panel.
- System Log: This VI logs the text given in the TALOS System Log. µService ID and a timestamp are automatically added.
- Send to Consumer: This VI enqueues a new message to be processed by the Consumer.
- Launch New µService: This VI is used to programmatically launch a µService.
- Message to µService: Send a message to another µService, using its ID.

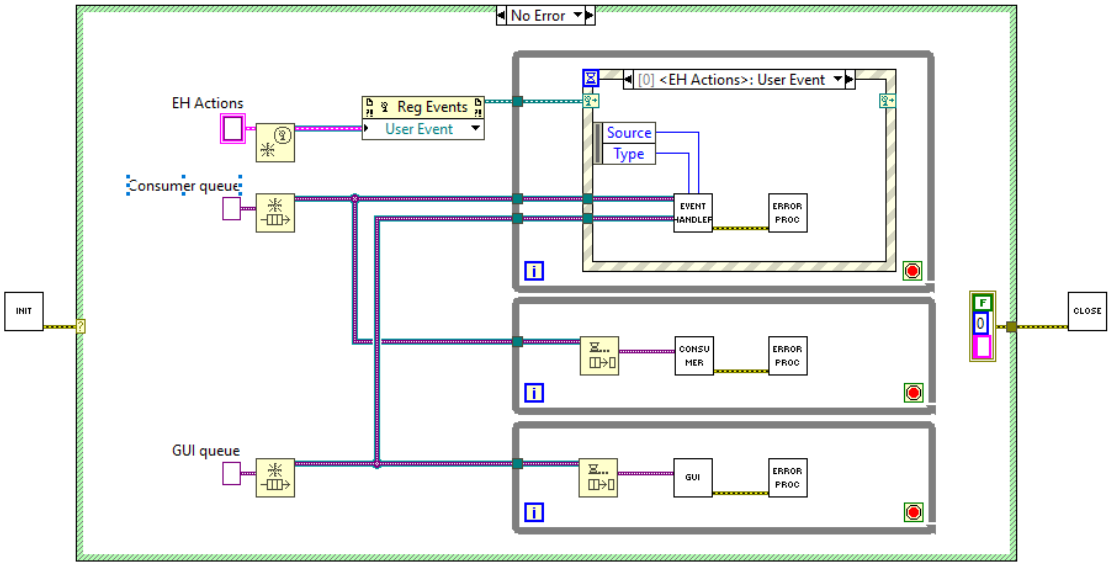


Figure 3.9 – Schematics of an equivalent QMH VI mimicking the behaviour of a µService.

- Config Param: To access the parameters present into the CIRCUS configuration file.
- Generate Error: This VI generates a custom error, given its number.
- Substitute Error: This VI is useful for translating LabVIEW™ errors, which are meaningless if extrapolated from the context, to project-specific errors, easier to handle (see *The Error Manager* (3.4.5.1)). The error(s) that gets substituted are logged in the TALOS Error Log.
- Check ABORT: Use this VI to check the status of the `ABORT` SV. It is optimised with respect to just opening the SV on the network.
- Raise ABORT: Use this VI to raise `ABORT` (also optimised).
- Infinite Message Manager: This VI is useful to manage (start, check, stop and flush) Consumer cases to be enqueued periodically. This is very useful in case a polling action needs to be performed, to avoid triggering the µServices Watchdog for apparent unresponsiveness.
- Filter Future Errors: With this VI is possible to memorise an error which will be filtered in the future, for a specified amount of time or occurrences.
- Stop: Used to cleanly stop the execution of the µService.

3.4.3.3 Communication

In the first days of the project, the communication between various Guardians was passing through a set of SV specially created for this purpose. But this proved immediately to be highly unscalable since the LabVIEW™ *Shared Variable Engine* (which is the LabVIEW™ process handling all the read and write requests of all the accessible SV) already displayed very long access times (order of the second) when three Guardians where simultaneously trying to access the same SV. Therefore the architecture was completely changed, moving the communication between Guardians and µServices on a pure *Transmission Control Protocol* (TCP) [145] system. It consists of using two separate pairs of µServices, the *TCP Listener* and *TCP Writer*, to instantiate two tunnels between each two Guardians where data flows in a single direction, allowing parallel and asynchronous

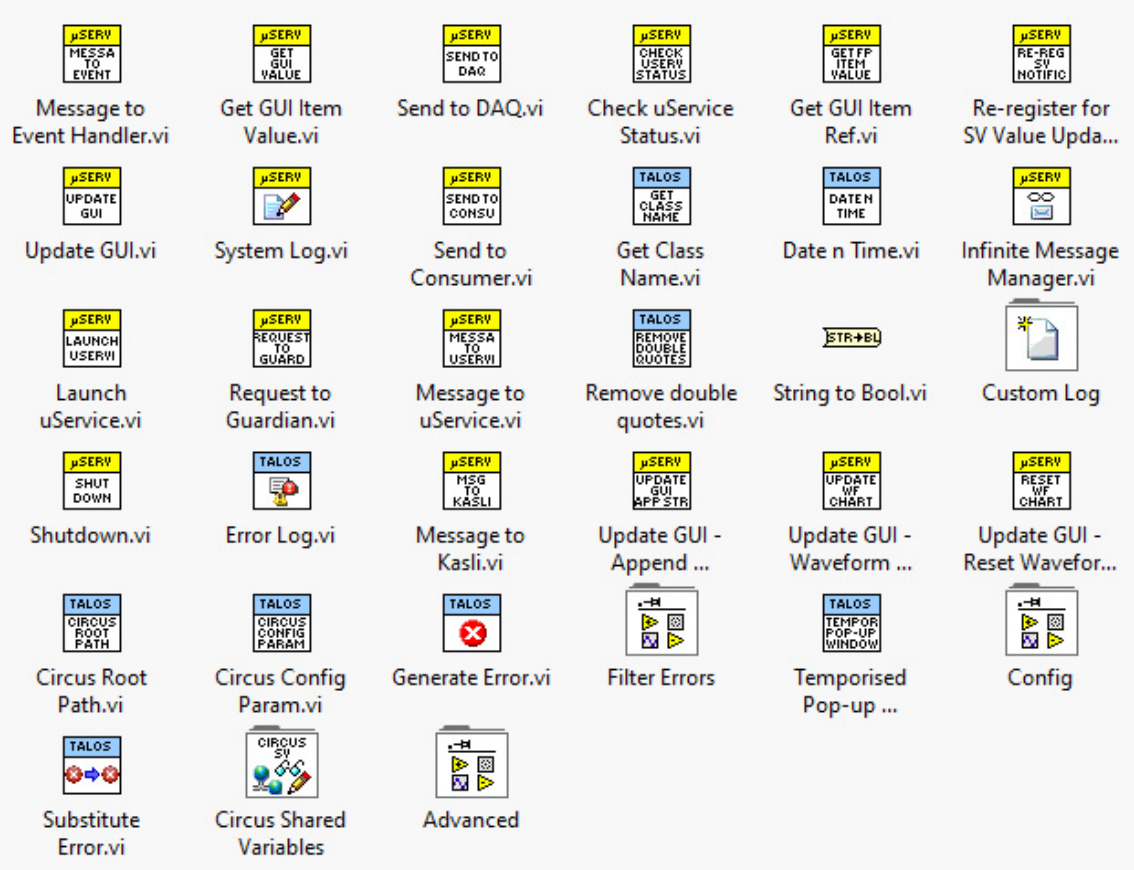


Figure 3.10 – The TALOS LabVIEW™ palette containing the methods for coding µServices.

message handling; in this way, the dead times between sending and receiving a packet is completely eliminated. With this method, the average round-trip time for an `echo` message (a message sent in one direction, which gets re-sent back immediately after arrival) is 3(1) ms.

This solution is also used for the communication between TALOS and the FPGAs. In fact, the Kasli normal user interface is a command-line terminal, where a user controls the outputs and provides eventual input. In the AEgIS Experiment, though, Kasli is the actual orchestrator of the operations at the ns level, so it necessitates bi-directionally exchange with computers, to send commands and receive inputs. Re-routing the terminal communications (i.e. *StdIn*, *StdOut* and *StdErr*) proved unfeasible as it can interfere with the real-time operation of the FPGA. The double, asynchronous TCP tunnel between TALOS and the Kasli[6] allows the Kasli to communicate without losing the ns-precise schedule of the sequence of operations.

3.4.4 ARTIQ integration

As previously mentioned, while TALOS combines all the experiment's computers into a symmetric distributed system and manages the slow control of the experiment, the ns-precise sequences of operations are offloaded to the Kasli. Normally, the Kasli is foreseen to be operated manually from a computer shell: most simply, the script containing the sequence of instructions is sent to the Kasli using `artiq_run <script.py>`; all the outputs – or necessary input – are then exchanged through the shell. Albeit very useful for quick debugging, this modality has several drawbacks, like needing an operator to monitor the FPGA operations and the necessity of submitting each script manually.

[6]Actually, the TCP tunnels are instantiated between TALOS and the PC hosting the Kasli, i.e. the computer where the *Kasli Wrapper* runs (see **??** (**??**)), and made available on the FPGA kernel via a set of custom-designed remote procedure calls (RPCs).

A more advanced mode of operation can be achieved via the *master-client* method. It consists of opening the *Master Shell* process in a shell, acting as a server for scripts, by executing, and leaving running, the `artiq-master` command. It maintains a queue of scripts, which can be populated from other shells with the command `artiq-client`, and automatically dequeues and sends scripts to the Kasli when found idle. This is the modality offered by the ARTIQ environment to handle a sequence of scripts[7] (at the price of not being able to pass inputs to the FPGA during the execution, but being an automated mode of operations, this is actually a design choice). For the automation goals of the AEgIS experiment, this solution was still not optimal because the control of the FPGA would have been outside of the overall control system, rendering the manipulation of the schedule of scripts difficult and the full handling of the errors generated by Kasli impossible.

Therefore, a specific µService has been created to manage the execution of the scripts of Kasli, the *Kasli Wrapper*. This µService encapsulates the ARTIQ Master Shell and makes the schedule manipulation easier by defining commands to operate it; moreover, it intercepts *StdIn*, *StdOut* and *StdErr*, giving the ability to send inputs to the running FPGA, and parsing both the output and the errors generated. The first point internalises the handling of the schedule inside TALOS, giving full control over its flow (see *The Scheduler* (3.4.5.2) and *The Monkey* (3.4.5.3)); the last point, instead, is crucial to building a reliable autonomous control system: in this way, also the low-level exceptions generated by Kasli are forwarded to the distributed error management system, enabling a more conscious execution of the experiments.

The integration does not stop at the encapsulation of the executing shell. As already outlined in *Communication* (3.4.3.3), an alternative communication protocol has been developed which uses the double TCP tunnel. Aside from the ability to treat, from the TALOS standpoint, the Kasli messages as any other µService message, the greater advantage is that the Kasli can send messages to TALOS without interrupting or delaying the ns-precise schedule of operations, contrarily to what happens when the normal *StdOut* is used.

This is further simplified by the creation of a custom library for ARTIQ, called *TCP Library* (see *ARTIQ custom libraries: the AERIALIST* (3.2.1)), where the implementation of the TCP communication is hidden inside a few functions to be called from the importing script: for example, the control of a µService via Kasli operations becomes straightforward. Here follows an example of the syntax of a Python function to trigger a specific message event of a specific µService, carrying also data.

```
def NewFunctionName(parameters):
    TCP_Send("<MicroService-Name>;<Message-Case-Name>;<Data1>;<Data2>;<etc.>")
```

It has to be noted that Sinara Kasli is the FPGA integrated by default into TALOS. However, by overriding the *Kasli Wrapper*, any other FPGA capable of TCP communication can be accommodated in the same manner.

3.4.5 Main System µServices

µServices are coded in the system as needed by the experiment. Nevertheless, some special µServices are already integrated in the TALOS library because they perform functions which are generally useful to every experiment: an example are the aforementioned µServices responsible for communications. In the following part, other notable µServices present in the system are described (the list is not exhaustive).

3.4.5.1 The Error Manager

A key part of the automation provided by TALOS lies in the distributed error management system. In particular, the fact that the system will respond to every error generated in any of its parts is the

[7] The ARTIQ environment also provides `artiq-dashboard`, a graphical interface for managing the queue of scripts of the Master Shell, but it is beyond the scope of this thesis.

core foundation of the safety of TALOS. The three key concepts upon which the error management system is based are the following:

- **Error substitution** relies on converting general error codes into user-defined ones, which carry also context-specific information: in fact, the same error can have a completely different meaning (and so require a different action to be taken) if generated in different contexts. As an example, the error *Connection lost* given by the *TCP Listener/Writer* of the Kasli is perfectly normal at the end of each script (since the connection is closed from the Kasli side), but the same error during a Guardian messaging another sibling is critical because it means a computer is unreachable.
- **Error concentration**, instead, is the process of gathering all the errors generated by the various parts of the distributed system to a single µService, the *Error Manager*. This solution ensures the generation of a consistent response to every error, regardless of its source (both process and location).
- **Error criticality**, is defined by assigning a numerical value to every error, ranging from 0 to 4, indicating the appropriate action (*Continue*, *Retry*, *Skip*, *Stop*, *Abort*) to be taken in response to that error. The system keeps track of the criticality codes generated during the experiment execution and, upon completion, the action corresponding to the highest criticality level encountered is performed (see section *The Monkey* (3.4.5.3)).

The strength of this error management system lies in its simplicity: in fact, the three key concepts have proven to guarantee its reliability, stability, and scalability to very high numbers of user-defined errors (more than 500, in AEḡIS, at the moment of writing). Conversely, it is not "too simple": the five levels of error criticality give the system the ability to respond flexibly in every situation.

3.4.5.2 The Scheduler

An indispensable functionality of any autonomous system is the ability to define a sequence of tasks the system needs to perform. For this reason, the *Scheduler* µService was created to allow users to define a schedule of scripts the system will execute. A picture of the *Scheduler* GUI is presented in figure 3.11.

A schedule is defined in Schedule Blocks (SBlocks), each defining a series of runs that execute the same script with possibly varying parameters. Each SBlock creates a sub-schedule from the provided parameters. After pressing the *SCIENCE* button, the schedules generated from the SBlocks information are concatenated and propagated further into the system, to be executed. There are two types of SBlocks, *Scan* and *Optimisation*, differing in the operation modality.

Scan SBlocks perform a pre-defined scan over some script parameter lists, while *Optimisation* SBlocks use the integration of ALPACA (see *ALPACA* (3.3)) to autonomously explore the parameter space to find optimal values of some parameters to bring some pre-defined observable as closes as possible to a target value. To add new SBlocks, the user can use the *Add Run* window as shown in figure 3.12. The script to be executed can be chosen via a browser, and the *Common Parameters*, fixed throughout the entire SBlock, can be defined. Furthermore, the *Quality of Run* can be specified, to retry the script if this quality criterion is not met (see *Quality of Run assessment* (3.4.7.4)).

If the Scan mode is selected (see figure 3.12a), the sequence of parameters to scan over has to be defined: either the full list can be given, or (if numeric) *start value* (x_0), *stop value* (x_N), and the *step size* (dx) or *number of steps* (N) can be provided, and the *Scheduler* will generate the corresponding list in linear, exponential or linked progression. When two (or more) parameters are linked together, the scan is performed varying them together, using the *number of steps* value. The parameter p_i's value is calculated as:

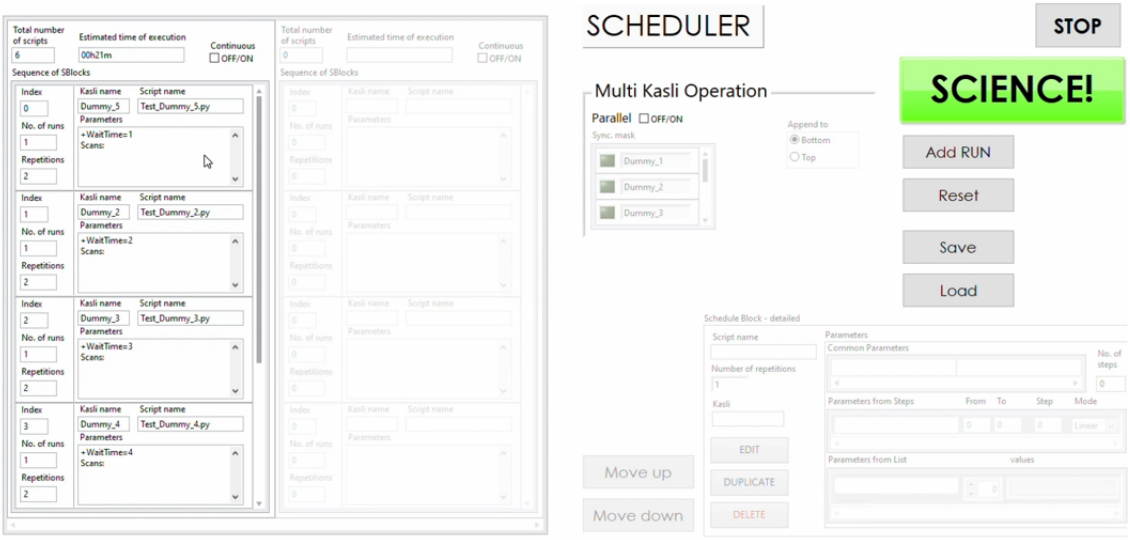


Figure 3.11 – The *Scheduler* GUI. On the left, the schedule being created is present; in the centre, the selector for multi-kasli operation is visible, together with the *synchronisation mask*. On the right, there are buttons to create a new SBlock (by opening the windows shown in figure 3.12), the "SCIENCE" button, to send the schedule to the *Tamer* to be executed, a button to clear the current schedule, and two buttons to save and load the current schedule on a file. On the bottom right, the little panel gives a preview of the selected SBlock, which can be modified at will by clicking the corresponding control, and buttons to re-organise the schedule, by moving, duplicating or deleting the selected SBlock are present.

- linear: $p_i = x_0 + i * dx$
- exponential: $p_i = x_0 * dx^i$
- linked: $p_i = x_0 + (x_N - x_0)/N$

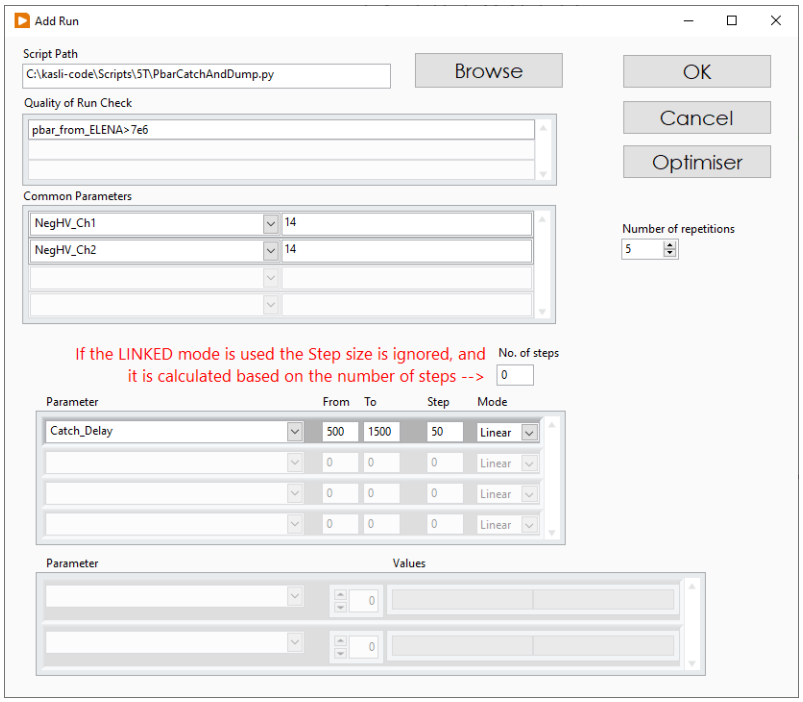
Both the first and the last values are included. Parameters can also be specified with a list of values: this allows for the use of strings or specific values not defined by a simple equation. Then the scan is performed over all combinations of parameters $(a_i, b_i, c_j, d_k, ...)$: in this example, parameters a and b are linked in the scan. The total number of Runs is calculated with $N = N_{linked} * \prod N_i$, where N_{linked} is the number of runs for the linked scans and N_i is the number of scans for the i-th parameter.

In the Optimisation mode (see figure 3.12b), the user defines the parameters (numbers and/or strings) that are used to perform the optimisation over, together with their boundaries, and the list of observables that the optimiser uses to evaluate the run results, together with their optimisation strategy, i.e. maximum, minimum or specific value.

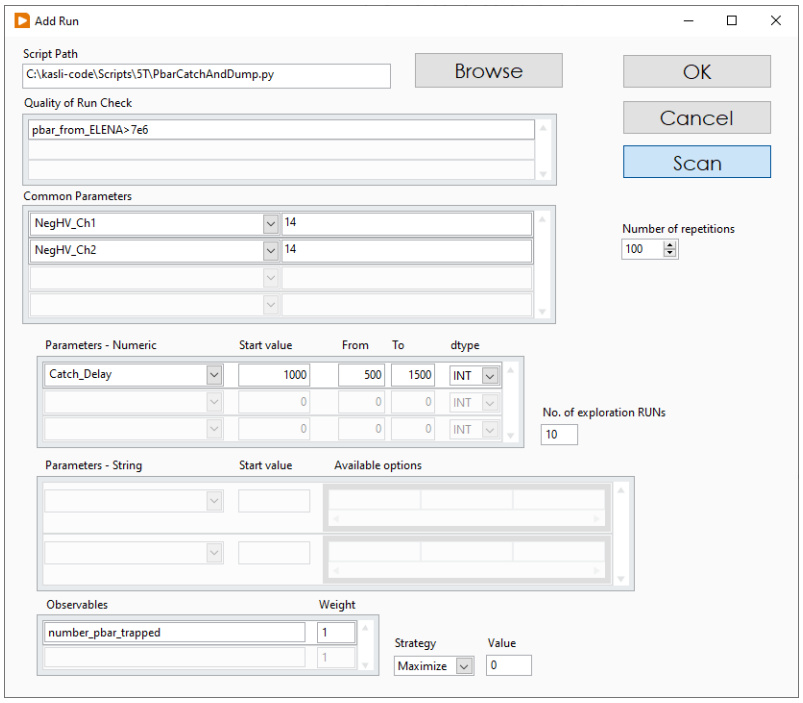
All created SBlocks can be modified, and their order can be changed at any time when creating schedules. The *Scheduler* supports the creation of multiple schedules when working in parallel multi-kasli operation mode. If this modality is selected, the various schedules become automatically populated so that each one contains scripts for one specific Kasli one. Moreover, a *synchronisation mask* becomes available, to select how the various Kaslis are synchronised, or not, among them (see section *Automation flow* (3.4.7.2) for more details).

3.4.5.3 The Monkey

The *Monkey* plays a central role in the automation of the entire control system, taking most of the high-level decisions that normally would be entitled to the user operating the experiment. This µService executes the sequence of scripts defined in the schedule generated by pressing *SCIENCE!*



(a) Scan SBlock



(b) Optimisation SBlock

Figure 3.12 – *Add Run* panels for different Schedule Blocks. The selector for the script choice and the first table for the fixed parameters of the execution are always present. In the Scan case (top), multiple tables are available to define the parameter sequences to scan over. In the Optimisation case (bottom) instead, the tables defining the parameters and the ranges over which they can be optimised by evaluating the observable(s) are edited at the bottom of the window.

on the *Scheduler* (see *The Scheduler* (3.4.5.2)), and it decides the best action to be taken upon each script completion based on its outputs and the errors that occurred. In fact, during the execution of each script, the *Error Manager* (3.4.5.1) sends to the *Monkey* the criticality level of any error raised, which is duly memorised by the *Monkey*. Then, each script always ends with a *Basic quAlity Notification After the eNd of an Action*, or `BANANA` for short, with a return value depending on its execution. The final assessment of the action to be taken is then evaluated as:

$$\text{action} = max(\text{error criticalities}, \texttt{BANANA}) \tag{3.4.1}$$

The numbers correspond to the following possible actions:

- 0: *Continue* - Everything finished well. The system will execute the next script.
- 1: *Retry* - There was a minor problem in the execution (e.g., some data was not saved): the script will be re-executed to ensure that the given configuration of parameters is correctly measured.
- 2: *Skip* - There was a problem that prevented the completion of the script, e.g. an incorrect configuration was provided. Since retrying the execution will give the same result, the script will be skipped, and the system will move on to the subsequent one.
- 3: *Stop* - There was an error during the operation, like, e.g., the Kasli controller not responding. In this case, the entire schedule will be skipped and the system will remain idle, but the *Monkey* won't try to stop the execution of the current Run.
- 4: *Abort* - A critical error happened: the *Monkey* skips the entire schedule and stops the execution of the current script. Meanwhile, the Guardians will put the entire control system in *Safe Mode*.

The full decision process is shown in figure 3.13.

All the skipped scripts are saved in a new schedule called *Skipped RUNs*. They can later be rescheduled (i.e. sent back to the *Scheduler* for editing) or run again. This can be seen in figure 3.14.

Upon the correct ending of a script from an Optimisation SBlock, the *Monkey* contacts ALPACA to retrieve the new parameter values to be used on the subsequent re-execution of the script (or end the optimisation if the target is reached).

3.4.5.4 The Tamer

One important feature of TALOS is the possibility of using multiple Kasli controllers at once. This is obtained through the simultaneous operation of multiple instances of *Monkeys* and *Kasli Wrappers*, a pair for each Kasli controller used. To ensure and control the proper flow of data, the *Tamer* µService is used as a distributor: it receives the error criticalities and `BANANA` messages from the *Kasli Wrappers* and re-routes them to the addressed *Monkeys*. This scheme allows for synchronous and asynchronous execution of scripts in parallel mode (see *Automation flow* (3.4.7.2)). Moreover, it simplifies the communication between the main system µServices – shown in this section – during the experiment by limiting *Scheduler* and *Kasli Wrappers* to send messages to the *Tamer* exclusively. Additionally, the *Tamer* is the second level of protection from errors occurring during the execution: it is the µService that decides to Stop the run in case of high error criticality levels; otherwise, it lets *Monkeys* retry and skip scripts, based on the evaluated value.

In figure 3.15, a screenshot of the Tamer subpanel managing more Kaslis is presented. The decision tree of *Tamer* after receiving error criticality is shown in figure 3.16.

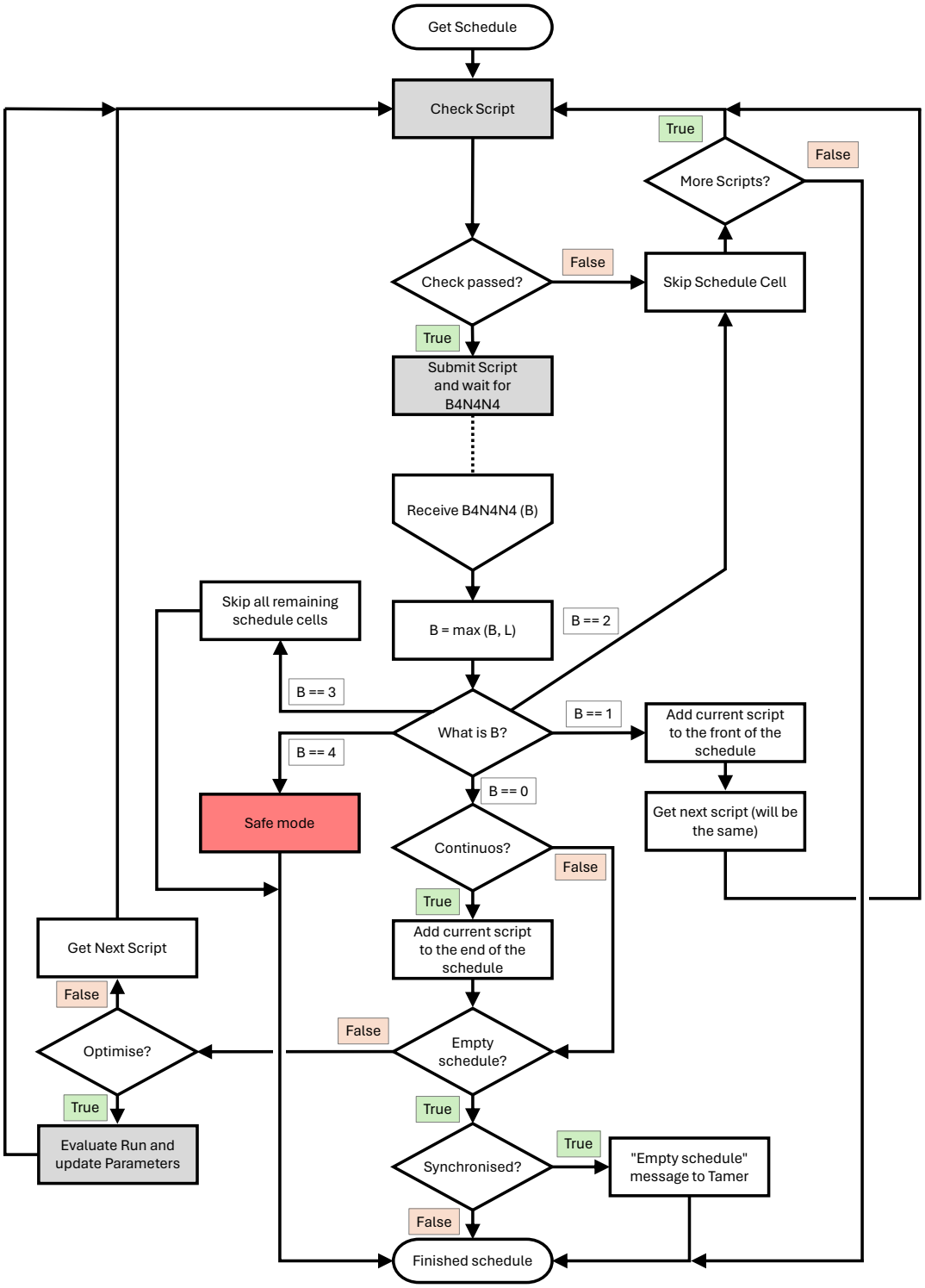


Figure 3.13 – The decision tree of *Monkey* during operations.

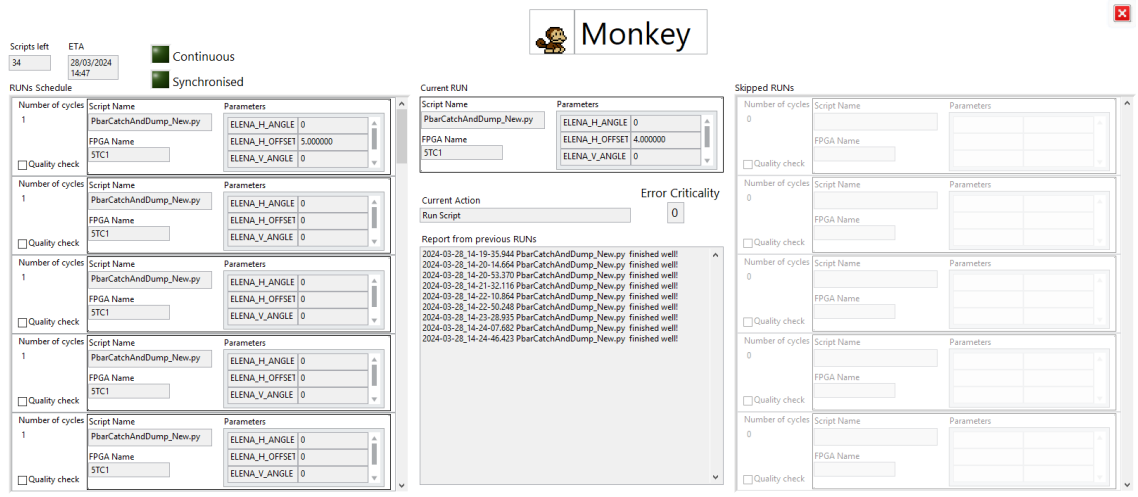


Figure 3.14 – The *Monkey* subpanel: on the left, the list of SBlocks to be still executed is present; in the centre, the running script is shown, together with the inherent parameters, and the status of the previous scripts execution; on the right, the list of the skipped scripts is saved, to be rescheduled or re-executed when the running schedule is terminated.

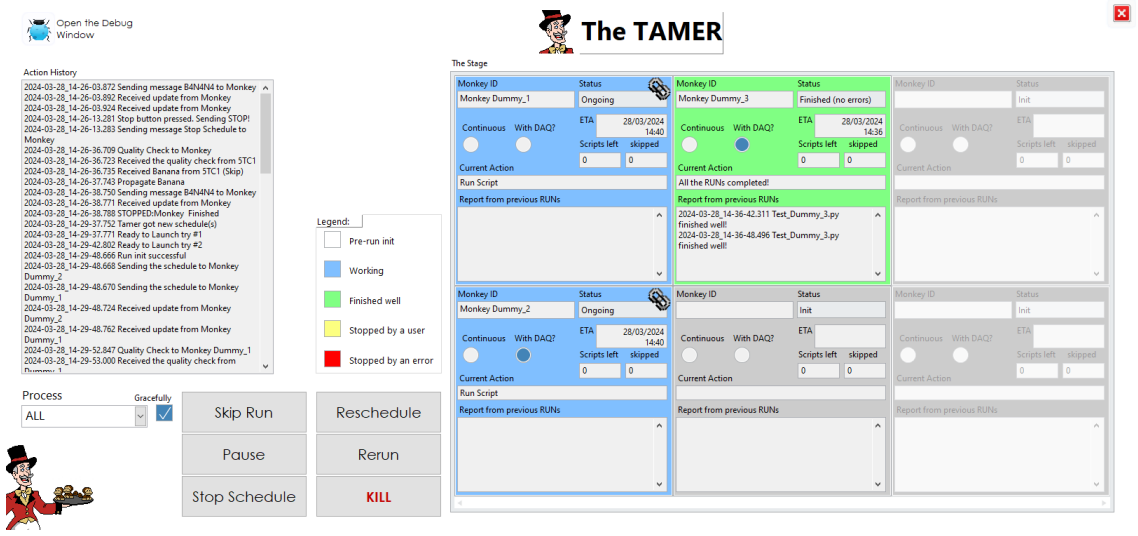


Figure 3.15 – The *Tamer* subpanel. On *The Stage* it can be seen that the *Tamer* is managing three *Monkeys*, two of them running synchronised (the ones with the blue background, hence still running), and one asynchronous (green background, i.e. terminated without errors). The background colour-code legend is visible on the left, above the various buttons; the buttons, in turn, are used to manage the execution of the current schedule. On the left, from top to bottom, there is the button to open the *Debug window*, the *Action history*, displaying the last actions performed by the *Tamer*, and the picture of the *Tamer* itself (holding as many *Monkeys* on the tray as the one on *The Stage*).

3.4.5.5 TCP Server

This µService, which runs headless (i.e. without a GUI to be displayed) in the background, listens for incoming connections from *TCP Writers*, and it spawns the corresponding *TCP Listener* to create the communication tunnel between Guardians. In the Kasli case, both sides of the connections are initiated by the FPGA, so the *TCP Server* creates a new instance of both *TCP Writers* and *TCP Listener* to establish the two tunnels.

Moreover, it also functions as a server for incoming errors that can be sent via TCP as strings on a dedicated port: in this way, it is possible to interface with TALOS whatever other program

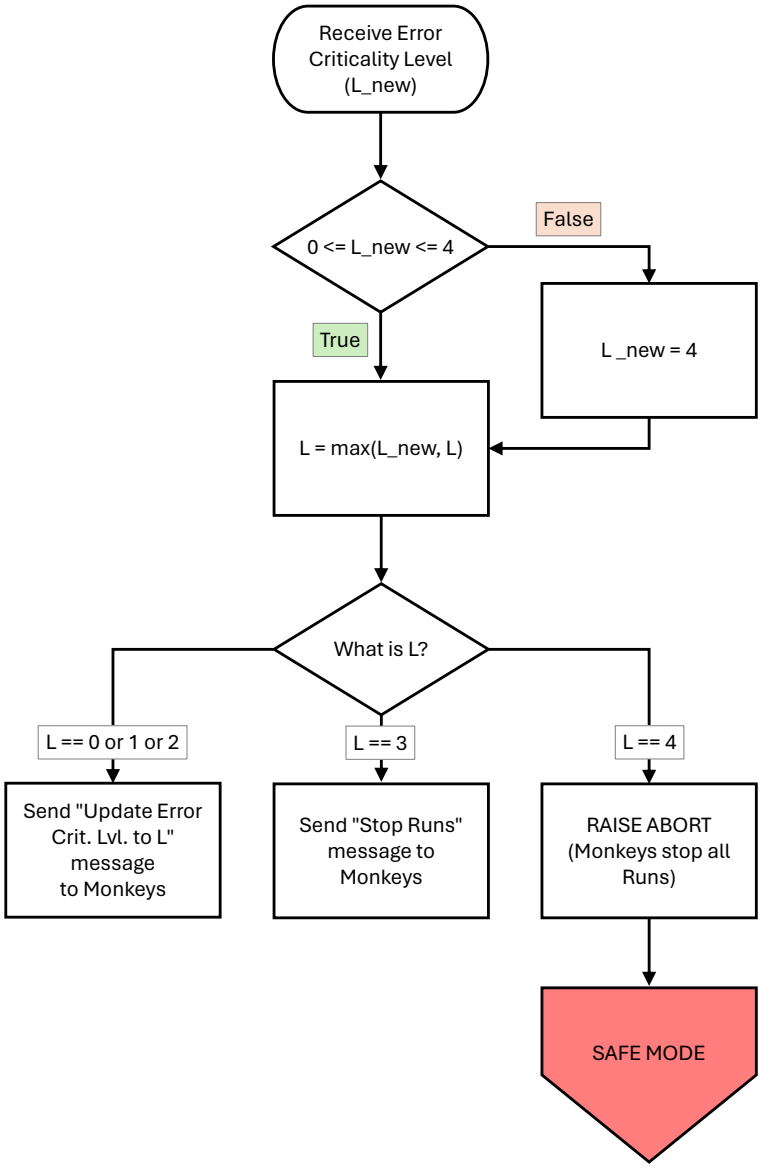


Figure 3.16 – The Error Criticality update process of the *Tamer*.

able to communicate over the network, so as to render the CIRCUS control system aware of their functioning state, and able to react to their errors natively, as if they were µServices.

3.4.5.6 Common GUI

This µService generates the overall GUI of the control system (and, funnily enough, runs headless, i.e. without a µService GUI to be displayed). In figure 3.17 a screenshot of the entire GUI is shown.

It is a special µService because it uses a direct override of the *actor-core.vi* to create a Front Panel where the other µService GUI are embedded dynamically, depending on the chosen value of the selector. In the upper left corner are present the Guardians and µServices watchdogs, showing respectively the Guardians connected and the µServices running, with an LED to display their health status and the timestamp of the last *I'm alive!* message received. Moving to the right, the CIRCUS logo is present, followed by the error list: this mirrors on every machine the list of errors received by the *Error Manager*, very useful when working on a µService present on a machine that does not run it. On the top right, various LED and indicators are embedded, to let the user understand what the status of the current measurement and the value of `ABORT` is, completed by the big, red STOP button for halting the system (*Ctrl* needs to be pressed to arm it, in order to avoid to inadvertently click it and, possibly, stop an ongoing measurement). In the right-hand column, the top section provides specific details regarding the error selected from the above error list and, underneath it, a real-time log displays Kaslis's operational activities, independently across all experiment machines.

3.4.5.7 Guardian Subpanel

The user interaction with the Guardian is managed by the *Guardian Subpanel*. On its GUI various buttons are present, some of which can trigger actions performed simultaneously by all the Guardians.

The best example of it is the *Update from Git* action. Every time a modification on the system is done, TALOS needs to be updated on every machine, which can easily become a long and tedious job when it has to be performed on multiple computers. Pressing this button broadcasts to all Guardians[8] the need to initiate an update sequence, which consists of shutting TALOS down, issuing a *git pull* command to synchronise the local copy of the data with the updated one on the cloud, and then restarting the control system.

Figure 3.18 shows the GUI of this µService, with the process of updating the distributed system initiated (pressing the button actually opens a confirmation panel: again, tiredness can bring a user to mis-click, sometimes...).

3.4.5.8 DAQ Manager

This µService manages the interaction with the DAQ system of AEḡIS. Every time a new Run is started, it has the task of opening a new Run folder on the system, checking the availability of the system throughout the entire data acquisition, and safely closing the connection when the measurement is finished. The actual sending of the data to the DAQ is handled by the *DAQ Sender* (3.4.5.9).

Has to be noted that this specific implementation of the *DAQ Manager* presented here is characteristic of the AEḡIS CIRCUS: but other data acquisition systems, based on different hardware and software setups, can of course be easily integrated into the overall control system structure analogously. Provided that the data acquisition system supports an interface with the commands *Start*, *Stop*, *Check Status* and *Send data*, its integration in CIRCUS would simply consist of creating a child of the *DAQ Manager* and of the *DAQ Sender* (see next section), and

[8]The computer flagged as "critical" on the CIRCUS configuration file is not updated and rebooted, by default: their inclusion needs to be manually triggered by clicking the little box on its left.

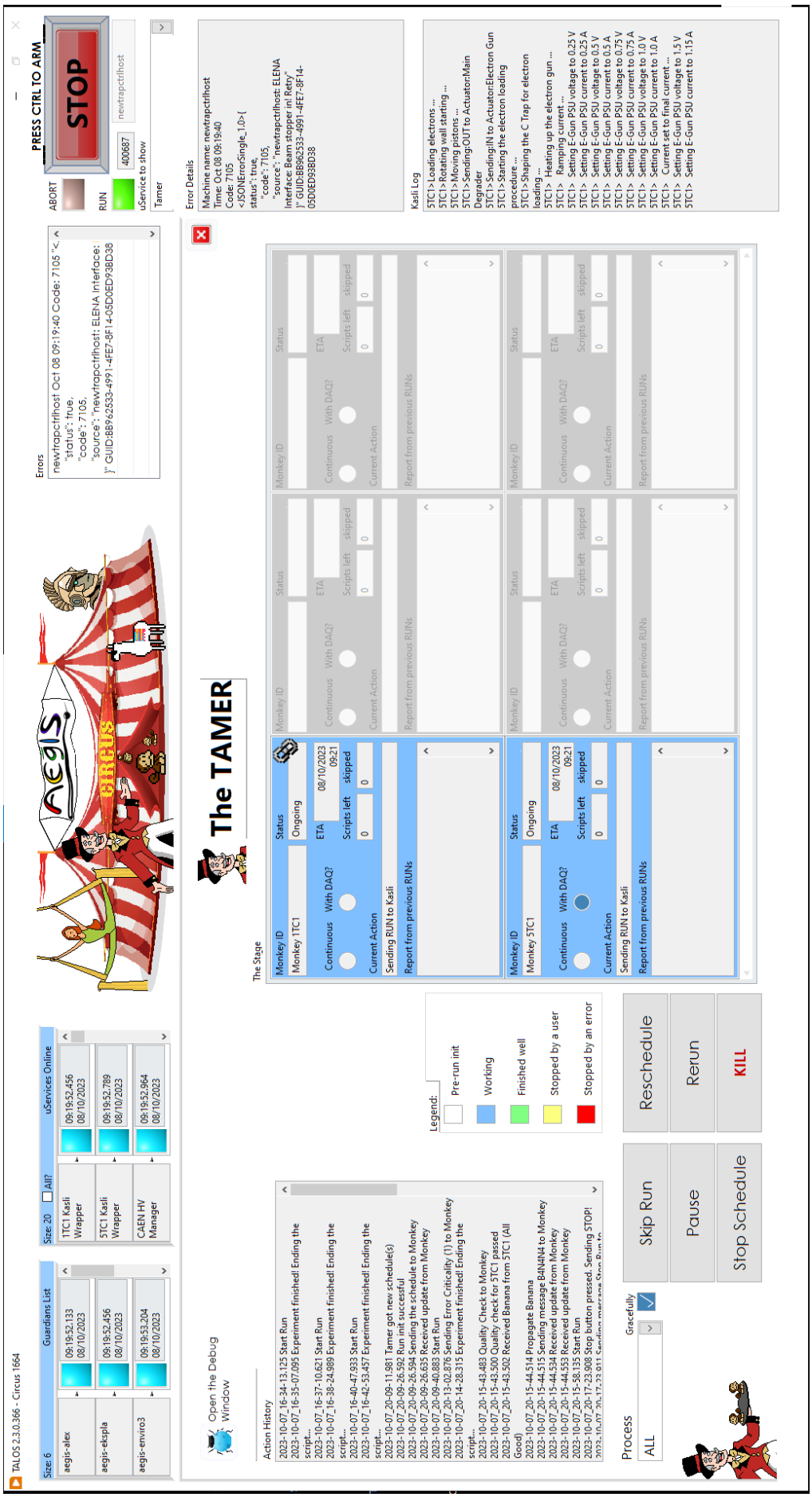


Figure 3.17 – A screenshot of the CIRCUS control system, showing the *Common GUI*. Located in the upper left corner are the Guardians and µServices watchdogs, while the error list is positioned in the upper right corner. In the right-hand column, the top section provides specific details regarding the selected error and, underneath it, a real-time log displays Kasli's operational activities. This interface is uniform across all experiment machines. Within the main window, the *Tamer* µService is displayed, presently overseeing two *Monkeys* performing an active measurement schedule with two Kaslis.

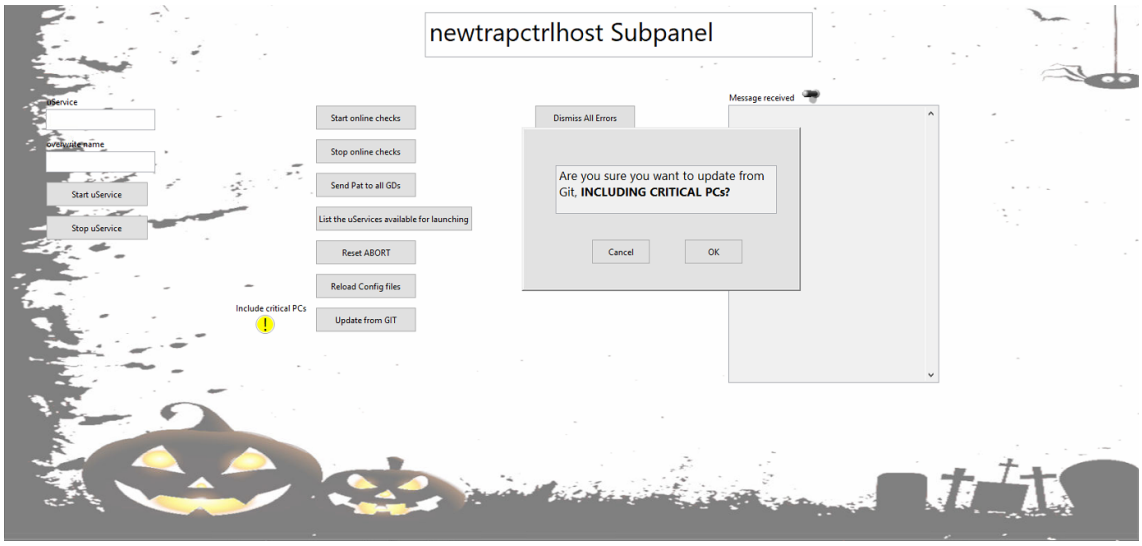


Figure 3.18 – The *Guardian-subpanel* of the *NewTrapCtrlHost* PC is about to start the process of updating from Git the entire distributed system, pending the confirmation on the appropriate window provided (when the screenshot was taken, it was Halloween period: TALOS can change skins and figure to be adapted to themes, and this is one of the cases. See *Artworks* (B.3) for more details and more artworks).

implementing inside it the proper interface with these commands. After that, TALOS and all the other µServices will immediately use the new data acquisition system for data saving, without any further change in the code.

3.4.5.9 DAQ Sender

The sending of the data to the DAQ has been externalised to a separate µService, for various reasons. First, this µService provides a data buffer to all the µService (and Guardian): in fact, when data is sent to the DAQ using the *Send data to DAQ* µService method (see *The Father Of All Microservices (FOAM)* (3.4.3.2)), the packet is stored in a queue inside this µService, and sent to the DAQ serially in a FIFO manner. This solution makes the data sending essentially instantaneous for every µService, minimising the time lost between two subsequent acquisitions.

Moreover, in this way, the data sending can go on in the background also after the Run has terminated, if the DAQ supports such feature (like the case of the AEgIS one): this results very useful in case a big amount of data gets acquired all together at the very end of a Run, minimising the time lost between the end of a script and the start of the subsequent one.

3.4.5.10 Kaslis' Server

This simple µService enables the sending of messages among Kaslis. In fact, there is no native way of sending a string of text between two Kaslis, but the TCP connection each Kasli establishes with TALOS opened for this possibility. The *Kaslis' Server* acts as a relay: every message addressed to a Kasli is actually sent to this µService, which properly reformats the message (into the usual semicolon-separated syntax used for the outgoing messages, as shown in *ARTIQ integration* (3.4.4)) and then it delivers it to the correct Kasli. Albeit simple in functionality, the *Kaslis' Server* enables coordination among multiple Kaslis before impossible, leading to applications like the *Particle Server* (3.2.5), and it is fundamental for the complex procedures needed to produce antihydrogen (see *Antihydrogen production methods* (4)).

3.4.6 Other notable µServices

The µServices described in this section are not part of the core of TALOS (some of them are specific to AE$\bar{g}$IS), but they show the implementation capability that the framework gives.

3.4.6.1 Detector Manager & Father Of All Detector

The AE$\bar{g}$IS experiment, like most physics experiments, is characterised by a high number of detectors. Albeit very different in nature and function, their pattern of operation essentially boils down to the configuration for the acquisition of the signal, the acquisition itself, and the saving of captured data.

In the light of this general schematisation, to fully capture the power of code reuse via class inheritance, the combination of a µService, *Detector Manager*, and a hardware class, *Father Of All Detectors* (or, simply, FOAD), was created, to speed up µServices coding and minimising debug time. In fact, every time a new µService needs to be created to manage a new detector, instead of generating a completely new µService, it is way more convenient to only create two children, one per each of the classes just mentioned (i.e. FOAD and *Detector Manager*), and fill their components.

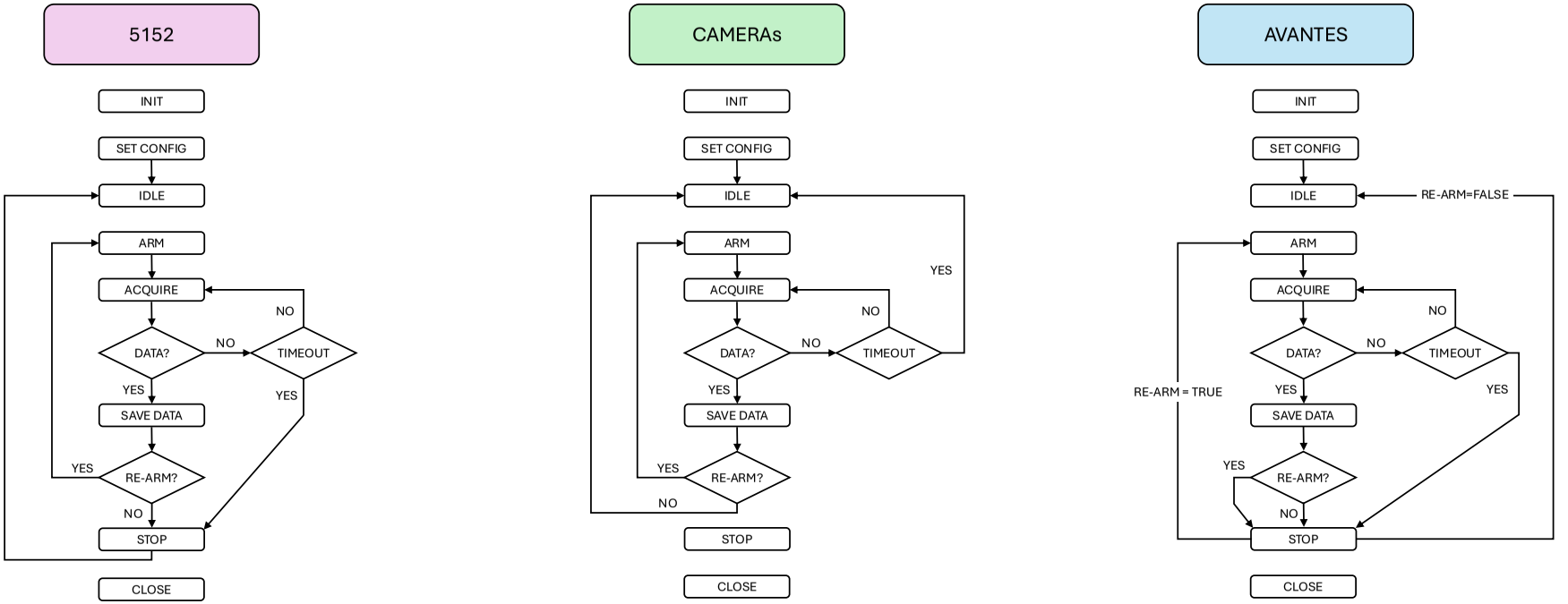
The class FOAD represents the generalisation of the detector's hardware functionalities: each of its VIs (see following list) is a place-holder that is called in the implementation of the *Detector Manager* µService, and it represents a specific action for the detector. Each child, which implements the software interface with a real detector, needs to override them.

- Init: executed during the corresponding phase of the µService, used to perform static configurations and basic hardware checks.
- Set Config: it needs to be called at least once before arming the device; it is used to configure the detector for a specific acquisition mode.
- Arm: to ready the device for acquisition.
- Acquiring: it polls the status of the acquisition, to see if data is present. If positive, it moves to Save Data; if timeout, it moves to Stop (or goes idle).
- Save Data: in this VI the data acquired can be formatted in the most suitable manner to be saved in the chosen destination place(s) (locally and/or over the network).
- Stop: to stop the current measurement.
- Close: guaranteed to be performed at µService shutdown, it ensures both the proper shutdown of the detector and the termination of the communication with the hardware (which often avoids problems in subsequently re-establishing it if incorrectly closed).

Detector Manager contains all the instructions needed for the correct functioning of the µService, from the interaction, via messages, with the rest of the system, to the GUI. In each of its children, it is only needed to specify the corresponding child of FOAD to manage (so to substitute the FOAD VIs with the corresponding override VIs specified in the child detector hardware class), and together with some flags, namely

- Auto-ReArm: to go again to Arm after saving data, if the timeout has not yet expired
- Stop after Save Data: to Stop, or not, between Save Data and going idle
- Stop before ReArm: if it is needed to pass through Stop between a Save Data and a following (Re)Arm

the patterns shown in figure 3.19 can then be obtained, which covers most cases of detectors to implement. Two examples of the *Detector Manager* controlling, respectively, the Avantes spectrometers and the MCP, are given in figure 3.20.

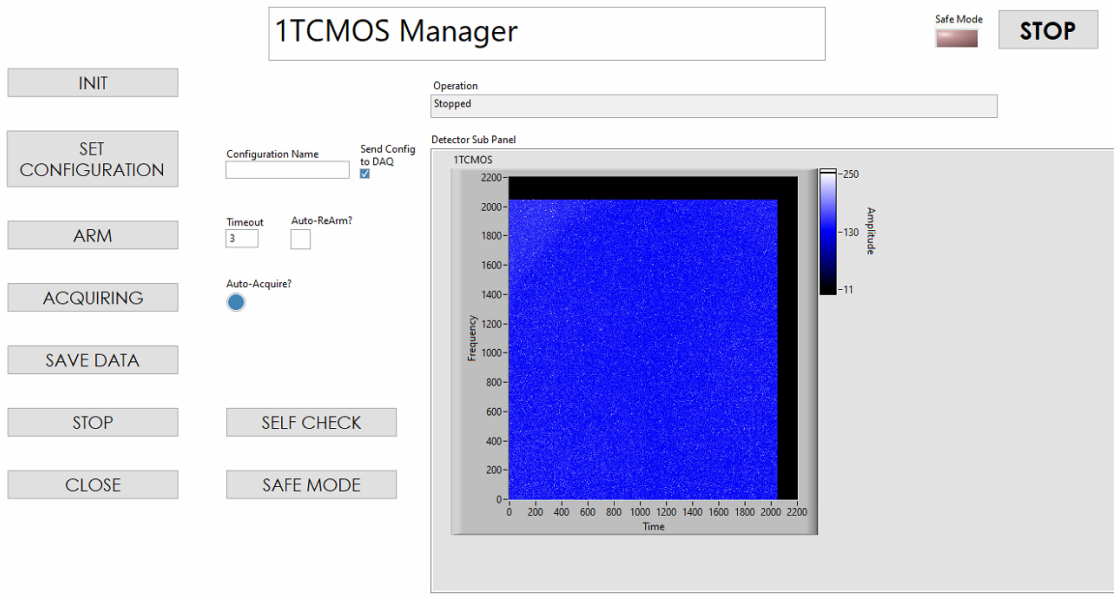


(a) Stop after Save Data: True; Stop before ReArm: False

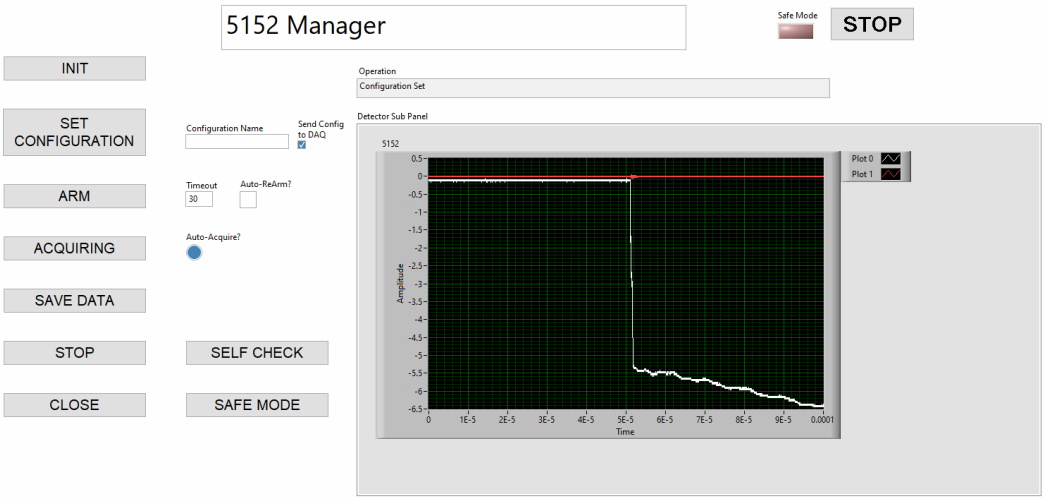
(b) Stop after Save Data: False; Stop before ReArm: False

(c) Stop after Save Data: True; Stop before ReArm: True

Figure 3.19 – Flowcharts indicating the three different logics obtainable with the *Detector Manager* parent µService, varying the values of the `Stop after Save Data` and `Stop before ReArm` flags.



(a) MCP Manager



(b) 5152 Manager

Figure 3.20 – The *Detector Manager* controlling the *1TCMOS* (top) and the *5152* (bottom). It can be seen how the common interface, and particularly the underlying µService, can support two completely different pieces of hardware.

3.4.6.2 ELENA Interface

This µService centralises our interface with the decelerators, automating checks and actions that before were demanded to the operator in charge of the experiment. In fact, this µService has a two-fold use: on one side, it controls and records a series of parameters (e.g. the actual beam position) and can generate errors in case any of them is not nominal, automatically flagging a Run as not good and asking the system to *Retry* it. It is the case, for example, when the valve letting the antiprotons in the experiment is closed, because a quick intervention has to be performed in the experimental zone; or when ELENA delivers an empty shot.

On the other hand, the *ELENA Interface* also enables the settings of specific parameters programmatically, allowing the automation of tasks that before were only possible to be performed tediously manually. A clear example of this is beam steering optimisation: ELENA has four

parameters to tune the horizontal and vertical displacement and angle of entrance of the $\bar{p}$ bunch inside the experiment. These four values need to be periodically corrected in order to maximise the number of antiprotons entering the trap system (instead of colliding and annihilating on the pipes and trap walls): before the introduction of this µService, the only way was to perform a series of measurements, at the beginning of each the operator was changing manually the configuration of the steering, to scan the parameter space with the intended grid resolution. At the end of the series of acquisitions, an analysis was performed to find the best point, and the full operation was repeated with a smaller grid spacing, and so on, until the best spot was found with reasonable precision. With the *ELENA Interface*, the scanning over the grid can be completely automated by simply creating a schedule of measurements where the parameters are changing, and letting it run overnight, so to analyse the data taken the following day. With the further introduction of the *Automatic parameter optimisation* (3.4.7.3), the entire process of finding the best point is done in total automation, enabling the periodic optimisation of the beam steering without any burden on the scientists (see *Autonomous parameters optimisation* (5.1.5) for a detailed evaluation of the performance of the system).

3.4.6.3 Telegram Bot

To ease the life of the operators in charge of the remote monitor of the experiment[9], a µService acting as server for a Telegram [146] bot was coded: it enables the fruition of a simple interface to survey the running of the experiment by chatting with the corresponding bot on the Telegram app. The actual status of the system can be queried, and the bot can keep the user informed about it, by sending a periodic status message in the user's chat (the user defines the period, in minutes), until the user cancels the operation, or until the schedule of measurement is terminated. An example of a chat with the TALOS Telegram Bot can be seen in figure 3.21.

Clearly, multiple people can chat with the bot at the same time, and the various requests are handled separately (also the periodic updates); the bot can even be included in groups, for a collective interaction.

3.4.6.4 Environmental Manager

At the end of 2022, after the antiproton campaign, having TALOS shown already its reliability, it was decided to port the code managing the environmental control of the AE$\bar{g}$IS into a µService. This was no trivial task since it is one of the most critical pieces of code of the entire experiment: in fact, a failure can also cause major hardware damage to the apparatus (for example, opening the wrong valve, or colliding two movable actuators in the Sun).

This µService controls the status of the gate valves present in the main vacuum chamber and in the lines connected to it, and the position of the various actuators in the Sun; in addition, it continuously monitors the values of the different pressures present in the system. To perform these tasks, it manages a total of three devices: it was not possible to do otherwise (i.e. splitting the µService in three smaller µServices, one per piece of hardware), because of their logical interconnection. For example, when it is asked to open a valve, the *Environmental Manager* before controls the values of the pressures at both sides and only performs the action if the values are normal: otherwise, an error is generated. Coding this in more than one µService would have increased the complexity of the system, ultimately raising the probability of inserting a malfunction in the logic, jeopardising the safety of the experiment.

Instead, the µService has been running smoothly since the beginning of 2023, letting the user control the AE$\bar{g}$IS environment easily and reliably, also enabling the programmatic movement of actuators: it is the best proof of the stability and dependability that TALOS provides.

[9] Even when the experiment is let to run in full autonomy, a person designated to periodically check the behaviour of the system is always present, the only exception being the night hours, completely unsupervised.

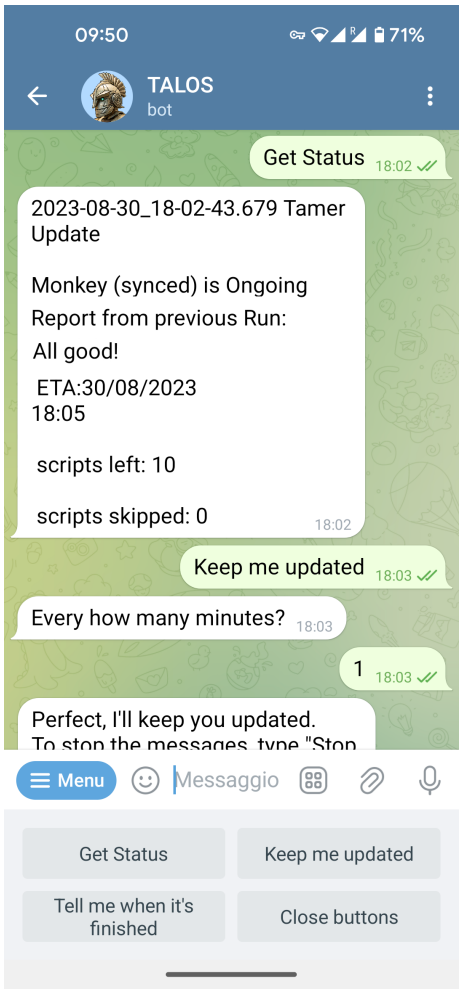


Figure 3.21 – A chat with TALOS Telegram Bot: on the bottom, the buttons for calling the available actions are displayed.

3.4.6.5 Thorlabs KCubes Managers

To programmatically control the various lasers present in AEgIS, numerous devices from Thorlabs, Inc. are employed. Since they share a common API framework, it was thought to heavily leverage the power of class inheritance enabled by TALOS to minimise the coding of all the related µServices necessary to manage all the devices.

To do this, a placeholder hardware class, called *Father of all KCubes* is created, containing the following methods: *Create*, *Close*, *Set Parameters*, *Get Position* and *Set Position*. All these VIs are left blank, the real code to be inserted in the corresponding override in children's classes. A µService, called *KCube Manager*, is created, using this hardware class VIs to program every action the µService is needed to perform.

Afterwards, four children of *Father of all KCubes* are generated: *KCubePiezo*, *KCubeServoDC*, *KCubeStepper* and *FilterFlipper* [10]. Each of them implements in the overrides of the parent VIs the interface with a particular class of devices, leaving out only the serial number, to be specified externally (i.e. from the calling µService), so as to be able to be employed to control multiple hardware pieces of the same kind.

Finally, every time a new Thorlabs component is added in the AEgIS experiment, to generate the µService to control it, it is only needed to create a child of *KCube Manager*, specifying which

[10]The *FilterFlipper* child was implemented at a much later stage, so the name of the parent class does not reflect this inclusion.

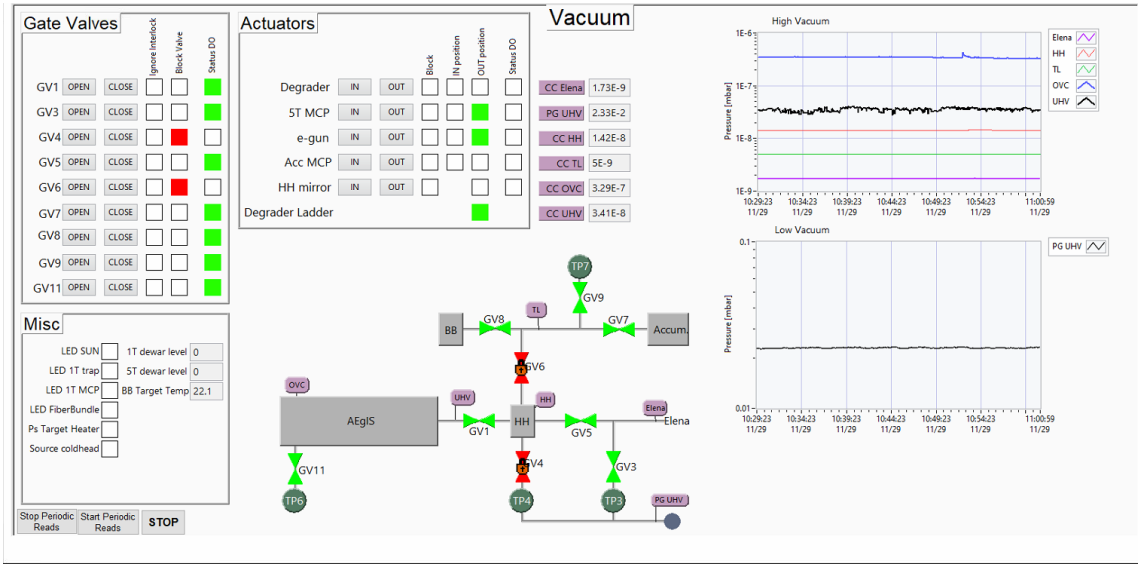


Figure 3.22 – The *Environmental Manager* µService. On the top right, the controls for the valves and the actuators are present: veto to the movements can be manually inserted, and they are automatically reloaded in case of restart of the µService (or the entire system). In the centre, a pictorial description of the valves is present, where the valves' status is updated in real time. On the left, the history of the various pressures in the system is displayed. On the bottom left, other miscellaneous readings and controls are placed.

hardware class has to use, and the actual serial of the device: in this way, 13 devices have been interfaced in the experiment, and more are foreseen, since the simplicity of the system.

The integration of the various Thorlabs devices in the AEgIS CIRCUS is proof of the power of inheritance upon which it is built, which guarantees the straightforward extendability and maintainability of the control system.

3.4.7 Autonomous operation

The most notable characteristic of TALOS is its ability to handle the execution of entire schedules of experiments without the need for human supervision. This was one of the main goals from the start, so the entire system was specifically designed for this task. This capability is crucial in order to maximise the amount of data taken while minimising the operators' time devoted to caring for the machine; moreover, it increases the repeatability of experiments by minimising random events and human errors.

3.4.7.1 A boat with two captains

To safely and reliably execute a schedule of particle physics experiments, both real-time system status awareness and nanosecond precision timing are essential. These two properties cannot be satisfied by either TALOS or Kasli alone: in fact, TALOS has the system overview, but it is limited to the ms-level scheduling precision of regular, non-real-time PCs; in contrast, Kasli offers ns-precision, but its scope is limited to its electrode values and internal status. Therefore, the control system resembles a boat with two captains, periodically switching the control of the helm depending on the needs.

As a matter of fact, at the beginning, when the schedule of Runs commences, TALOS assumes control. It verifies the correct functioning of each µService and, if positive, sends the first script to Kasli. Here, the helm control shifts to Kasli, which executes the script, while TALOS remains available to redirect to the correct µService every message request originated by Kasli (TALOS only intervenes in the event of an `ABORT`, halting the execution and entering *Safe Mode*). At script

completion, Kasli returns control to TALOS that, based on the return code of the script and any eventual error that occurred, decides the action to take, which typically involves submitting another script to Kasli. The process repeats until the schedule is completed or the execution halts due to an error or user intervention.

3.4.7.2 Automation flow

With the implementation of the *Tamer* and *Monkey*, the system allows for three modes of operation (depending on how and how many Kaslis are controlled): sequential (standard), asynchronous (parallel), and synchronous (parallel). These modes are achieved by specific handshake communication between *Tamer* and *Monkey*. As described in *Main System µServices* (3.4.5), each *Monkey* is capable of assessing the outcome of its Run, while the *Tamer* is responsible for the propagation of schedules and messages to the correct *Monkey*.

The automation flow begins with the schedule being defined by the user in the *Scheduler* and, by pressing the *SCIENCE* button, it gets sent to the *Tamer*. Upon reception of the new schedule, the *Tamer* starts with performing the *pre-run* procedure: it parses the schedule and decides how many *Monkeys* and *Kasli Wrappers*[11] needs to be running for the execution of all schedules, and it spawns them if necessary. Subsequently, after having verified that, the *Tamer* asks all Guardians and µServices if they are ready to start runs (see *The Father Of All Microservices (FOAM)* (3.4.3.2)): if everything responds positively, the *Tamer* propagates to each *Monkey* the corresponding part of the schedule, and the first Run starts. This procedure is a common start that is independent of the operation mode. Each *Monkey* starts executing its schedule and waits for the Run's outcome.

In the sequential operation mode (a scheme is given in figure 3.23), only one *Monkey* is running; nevertheless, the (single) schedule can have scripts designated for different Kaslis. Before running each script, the *Monkey* performs a formal check on it (like verifying that the file exists and that is correctly coded for the specified Kasli): if the check fails, the script is skipped, and the subsequent one is evaluated; otherwise, the script is executed, and the *Monkey* waits for its termination. The `BANANA` message from the Kasli is received by the *Tamer*, which redirects it to the running *Monkey*, signalling the end of the script. The *Monkey* runs the scripts one by one, waiting for a `BANANA` message (i.e. script termination) before starting a new script, even if they are designated for different Kaslis.

The asynchronous parallel operation mode (schematised in figure 3.24) is analogous to the sequential mode but with multiple *Monkeys* running multiple schedules simultaneously. Each schedule is a set of scripts for a specific Kasli, which is assigned to a dedicated *Monkey*, to have a 1-1 correspondence. Each *Monkey* executes its own schedule in total independence. After the end of each script, the Kasli sends the `BANANA` message to the *Tamer*, which re-routes it directly to the *Monkey* responsible for that specific Kasli: in this manner, each *Monkey* performs the check and evaluate the Run independently from the status of the other *Monkeys*, allowing for multiple scripts to run in parallel, asynchronously, on different Kaslis.

In the synchronous parallel operation mode (visible in figure 3.25), the communication between the *Tamer* and the *Monkeys* is more intense. The main goal of this mode is to start all the scripts belonging to different Kaslis at the same time: therefore, each *Monkey* needs to know the status and outcome of all its siblings. This functionality is achieved by using the *Tamer* as an execution barrier and response collector, both at the start and at the end of each script. It gathers all the outcomes of the various *Monkeys* script checks, and the `BANANA` messages, and redistributes the summary to every *Monkey* simultaneously, to ensure their synchronicity. The workflow develops as follows: before starting a script, each synchronized *Monkey* sends the outcome of the script check to the *Tamer*, which collects all the responses and evaluates the overall check outcome. Next, it propagates the overall check outcome to all synchronized *Monkeys*, so to ensure that the action after the check is the same by all *Monkeys* (and the synchronisation is not broken by a wrong

[11] It is customary to refer to multiple instances of the same µService simply using the plural, e.g. *Monkeys*.

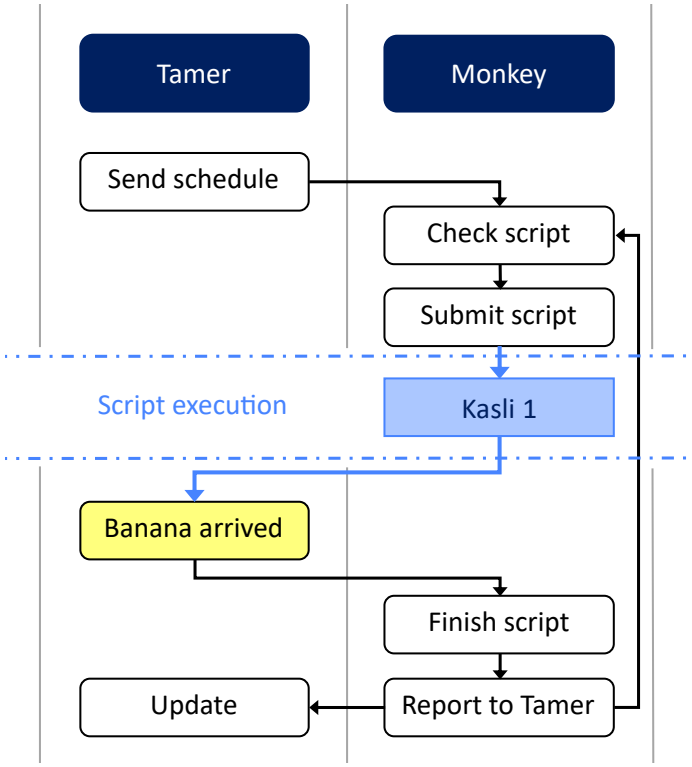


Figure 3.23 – CIRCUS autonomous operation flow in case of sequential operation mode.

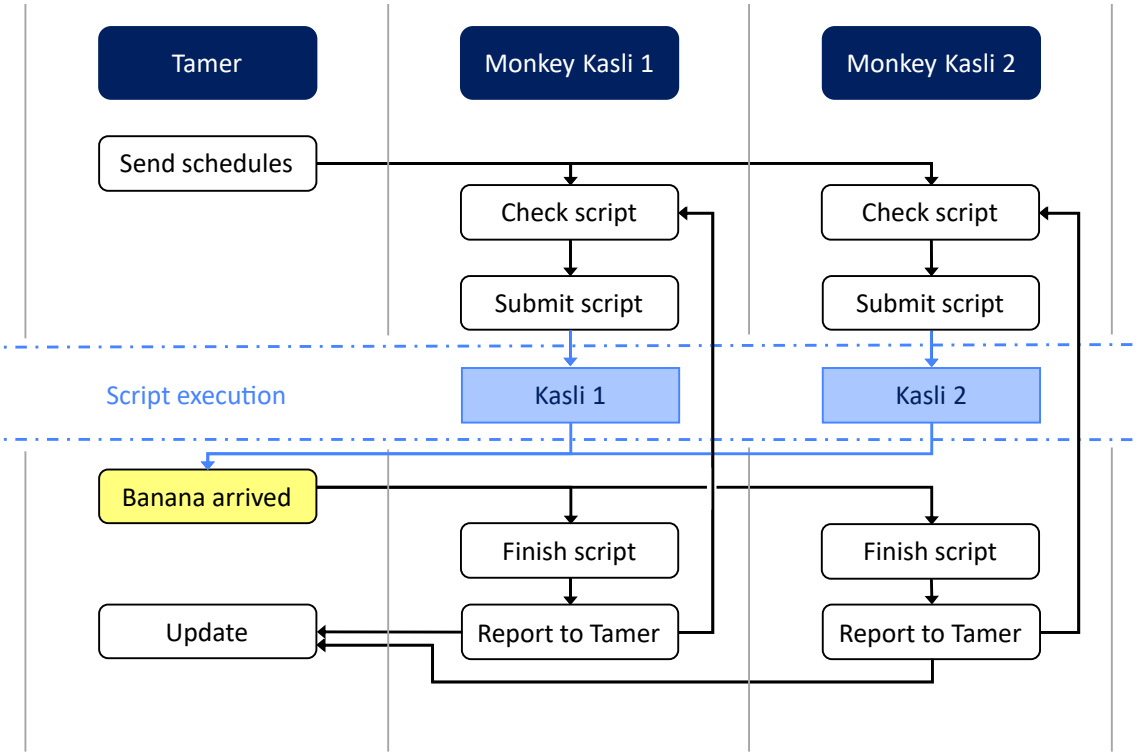


Figure 3.24 – CIRCUS autonomous operation flow in case of parallel asynchronous mode.

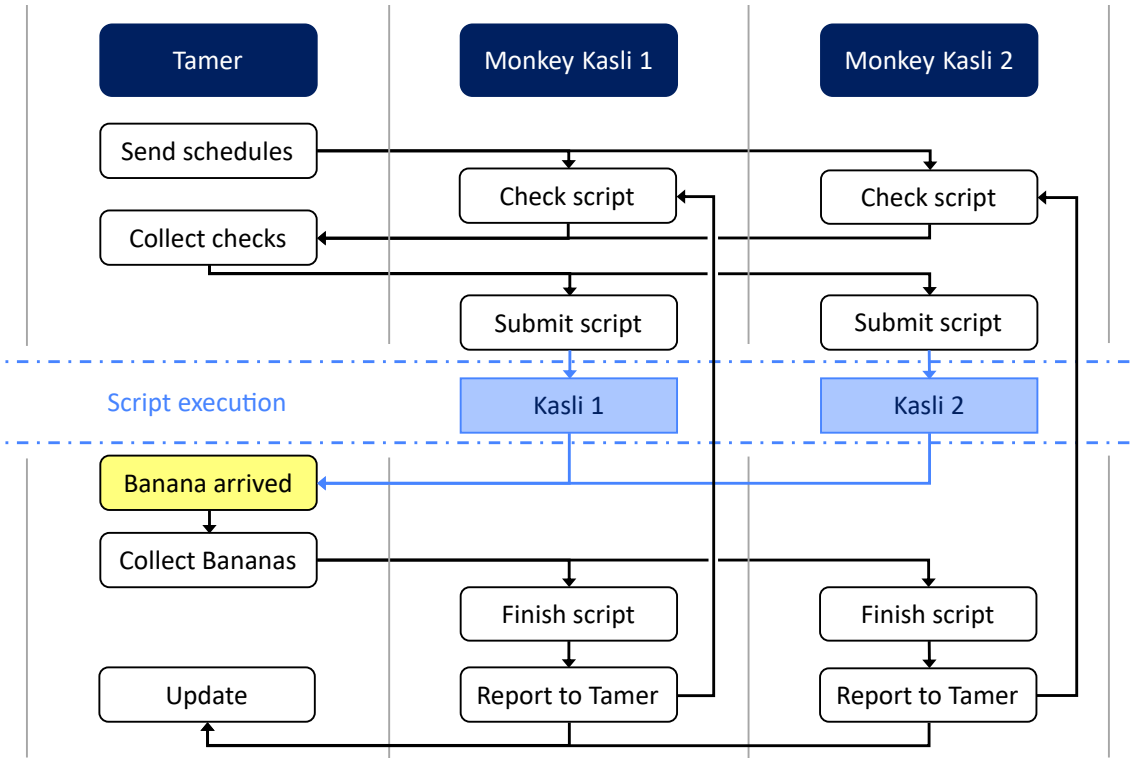


Figure 3.25 – CIRCUS autonomous operation flow in case of parallel asynchronous mode.

script in a single schedule), and that all the *Monkeys* start the execution simultaneously. A similar operation is performed at the script completion: the *Tamer* collects all the `BANANA` messages, and sends back the highest value to all the synchronised *Monkeys*, ensuring the end of the script action is the same for all of them.

The mode of operation is defined by the user in the *Scheduler* thanks to the synchronisation mask (see *The Scheduler* (3.4.5.2)), which allows the system to synchronize multiple Kaslis while running others asynchronously. The results of the evaluation of synchronicity are shown in section *Kaslis synchronisation* (5.1.3).

3.4.7.3 Automatic parameter optimisation

Although the level of user independence explained until here is already high, the automation, as intended in the sense of taking high-level decisions autonomously, is limited to reacting to errors and external events. This already had a huge positive impact on running the experiments at AEgIS, but it did not change conceptually how the experiments were conceived (and executed).

Thus it was decided to go a step further in this direction, by integrating into TALOS ALPACA, the AEgIS analysis framework, in order to empower the former with the ability to change the parameters of the scripts based on the results of the experiments performed previously. This feature can be leveraged by selecting the *Optimisation Mode*, as explained in *The Scheduler* (3.4.5.2). The optimisation script is executed normally (using the starting parameters provided by the user) but, upon completion, ALPACA suggests the next point to explore in the parameter space, thanks to its Bayesian optimiser, based on the results of the previous execution(s). TALOS then re-executes the script, using the ALPACA feedback. This procedure is iterated until ALPACA declares that the optimisation has converged, or the maximum number of iterations allowed by the user is reached. The rest of the schedule is then executed normally (but other Optimisation SBlocks can be present in the schedule).

This new mode of operation is truly disrupting, because it changes how the experiments are thought of and then defined. In fact, they pass from the spanning of wide parameter space to be

manually analysed afterwards in order to understand where the desired point of operation lays, to code directly the optimisation problem in the script, and let the machine reach it autonomously. One added benefit of this is the speed of convergence: because of the nature of the optimisation code in ALPACA, the number of points necessary to arrive at a desired observable value is smaller than the one using the usual grid scan; this is particularly true for multi-parametric optimisation, since with the Bayesian optimiser implemented in ALPACA the point increase seems to scale linearly with the number of dimensions of the parameters' space, instead of the exponential in the grid scan (see *Autonomous parameters optimisation* (5.1.5)). This means that complex, multi-parameters optimisation tasks, that before were to be performed fractioning the phase space into mono-dimensional problems to be optimised separately (implying that the parameters are independent, which often is true only to an extent), now can be fully explored, possibly finding better operational settings. An example of this is the calibration of one AEgIS laser: before it required the active manipulation of an expert for several hours, while it has been brought to the same setpoint by TALOS in a couple of hours. Once the calibration procedures are automated, they can be also thought of as being run periodically (even placed at the beginning of a long schedule of experiments), to ensure the system performs the measurements at its best every time.

The new mode of operation requires, clearly, a mentality change, since a lot of measurements can – and have to – be re-thought in terms of optimisation problems. As an example, the entire trapping and preparation of antiprotons in the AEgIS experiment can be broken down as a series of optimisation tasks, like: the best delay between receiving the trigger from ELENA and raising the trap electrode, to maximise the captured fraction; the best settings of the rotating wall technique, to maximise $\bar{p}$ plasma compression; the best "electromagnetic shape" of the trap (i.e. voltages of the various electrodes) to maximise $\bar{p}$ lifetime.

3.4.7.4 Quality of Run assessment

Another use of the TALOS-ALPACA integration is to assess the *Quality* of a Run. As explained in *The Scheduler* (3.4.5.2), checks can be specified during SBlocks definition (in the form of *Observable* $< / \leq / = / \geq / >$ *value*) to determine the data quality upon script completion. When the script is finished, in case the overall return code is *Continue* (i.e. 0), ALPACA is contacted by the *Monkey* to return the values of the observables specified to be tested. The value returned is then evaluated as per user instruction, and in case at least one of the checks is not passed, the Run is marked as unsatisfactory and is retried.

This addition allows the system to pre-filter data, preventing the need for manual exclusion during the analysis stage and subsequent re-taking of points. This is particularly advantageous when combined with the optimiser, as pre-filtering avoids biasing the system during its autonomous search for optimal parameter values.

3.5 TALOS and the CIRCUS

TALOS is intended to be the underlying framework upon which the control system of an experiment can be based[12]. It is thought to be used as the core engine of the CIRCUS control system, where TALOS is inserted as a LabVIEW™ *Packed Project Library*[13] , pre-loaded into a LabVIEW™ example project, also containing a template for µServices. The CIRCUS configuration file, the custom errors file, and *Startup.bat* come bundled with it.

To start using it, the CIRCUS configuration file has to be filled with the list of all the PCs where a Guardian should run, together with their IPs. Then, for every PC, the list of the µServices that should be automatically started at the Guardian boot should also be filled.

[12]Actually, its generality would make it suitable to also handle distributed systems which are not strictly experiments.

[13]A pre-compiled LabVIEW™ binary library that enhances the speed of execution and prevents unwanted file modification. It is the analogue of Windows DLL.

Subsequently, a copy of the example LabVIEW™ project shall be made for each PC of the experiment, each project name reflecting the name of the machine it represents. Afterwards, all the µServices need to be created, by copying the template provided, renaming it and modifying it to suit the purpose. Newly defined errors need to be inserted into the corresponding file.

The entire folder structure should be present on all used PCs (for this, *Git* is suggested: there are also utilities inside TALOS that leverage its capabilities, like the *Update from Git* of the *Guardian Subpanel* (3.4.5.7)). To start operating the CIRCUS, it is simply necessary to launch *Startup.bat* on all the machines and wait for the Guardian boot to complete. When all the lights of all the Guardians connections are on, the system is ready to run experiments.

Together with the normal, schedule-driven mode of operation, a *Debug Mode* is also available (see figure 3.26). To quickly prove if a new script is working while developing it, in fact, the script can be executed manually from a shell (`artiq_run script_name.py debug='True'`), and TALOS will accept the connection and redirect messages to the corresponding µServices, exactly as the script would be run by the *Monkey*[14]. The advantage of this modality lies in the control that the user can exert on the script during its execution: for example, it can easily be terminated (by `Ctrl-C`) at any point, if unresponsive, or if a particular error arises, or just to test the functioning up to a certain part.

3.6 The AEgIS CIRCUS

Throughout 2021 and 2022, the existing AEgIS control system underwent a gradual migration to use the CIRCUS. While this transition represented a substantial effort due to the complexity and diversity of the control system[15], the framework has demonstrated immediately its effectiveness and its power. The architecture of the system was reorganized into µServices, facilitating incremental development. Each of these was thoroughly tested and debugged in isolation, ensuring smooth operation in every foreseeable scenario before its integration into the operational control system: this approach minimized periods during which the entire control system was inoperable due to bugs. Additionally, the independence of each µService from the rest of the system meant that, in the event of issues arising post-integration, the system could be promptly reverted to its former stable state by simply removing the problematic µService. This strategy made identifying the source of problems straightforward – invariably located inside the most recently added µService –, significantly quickening and simplifying the debugging process.

Another salient characteristic of the µServices structure is that the exact same response can be triggered by a button on its GUI or by an external message (from the Kasli or another µService): this guarantees that, after testing the µService manually from its GUI, the expected behaviour will be maintained during its programmatic use.

At the time of writing, the new control system encompasses a network of 6 computers, running a collection of 120 µServices, some of which are different instances of the 42 unique µServices coded in the project. These manage several pieces of hardware, including 3 cameras, 3 different spectrometer types, 2 laser crystal heaters, 7 actuators for laser-optic components, 2 oscilloscopes, the electron gun, the high voltage generator, the pulser, the rotating wall generator, etc. An overview of the AEgIS CIRCUS topology is given in table 3.3. The stability of the system has led to the integration, at the end of 2022, of the environmental control system as a µService (see *Environmental Manager* (3.4.6.4)): it is one of the most critical pieces of software of AEgIS, since it checks and maintains the status of the vacuum and the cryogenic temperature of the entire experiment (preserving the life of the superconducting magnets, not immune to quench damage). More than 500 custom errors have been defined, and the system has been online since August 2021.

[14] To work in this modality, though, no running schedule has to be present, even if using different Kaslis, in order to avoid conflicts.

[15] To get an idea of the size of the system, see *The most important metric* (B.2).

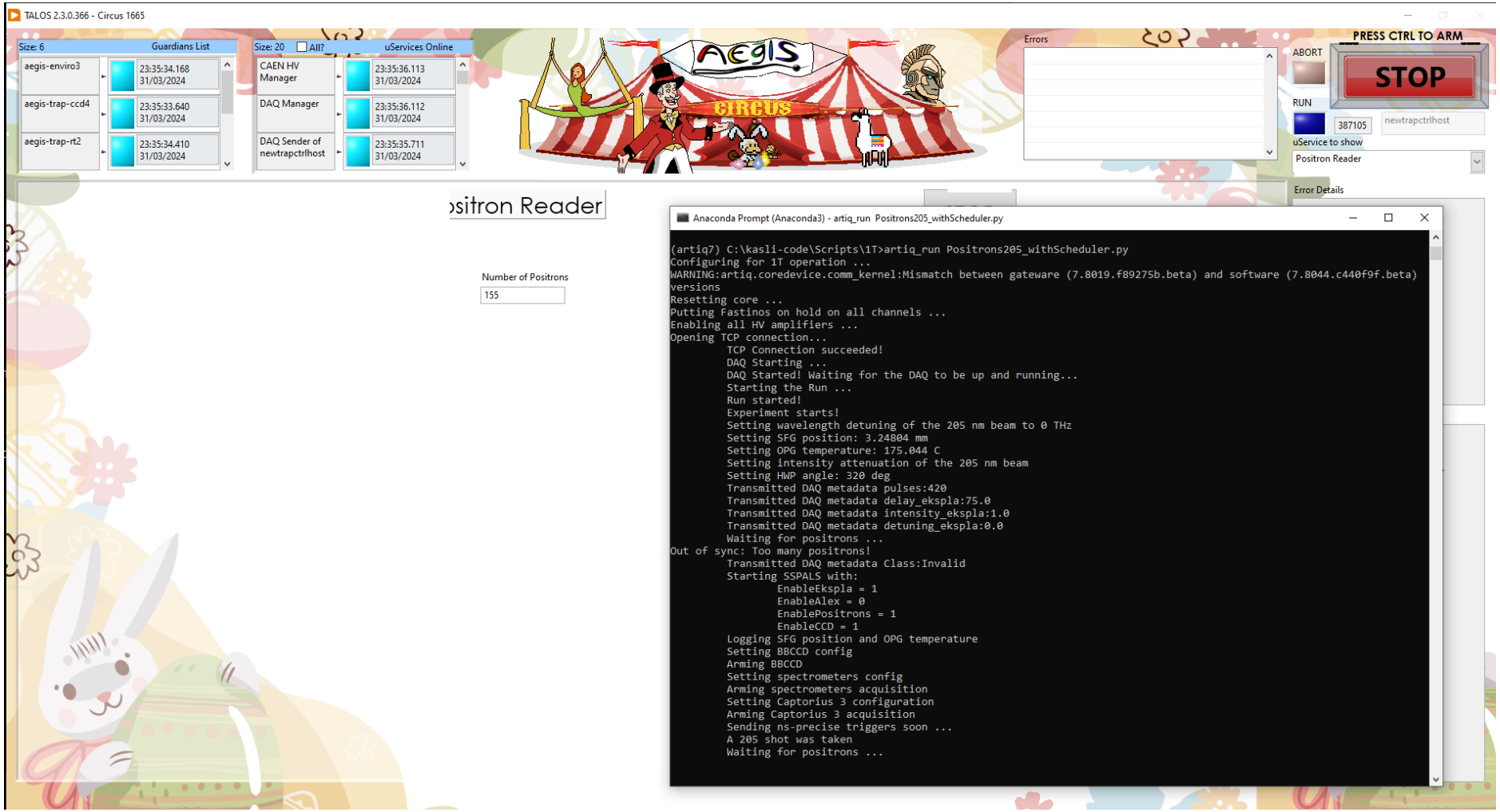


Figure 3.26 – A screenshot of the CIRCUS control system running in debug mode. On the front of the main CIRCUS window can be seen the terminal executing the script (the screenshot was taken during the Easter period).

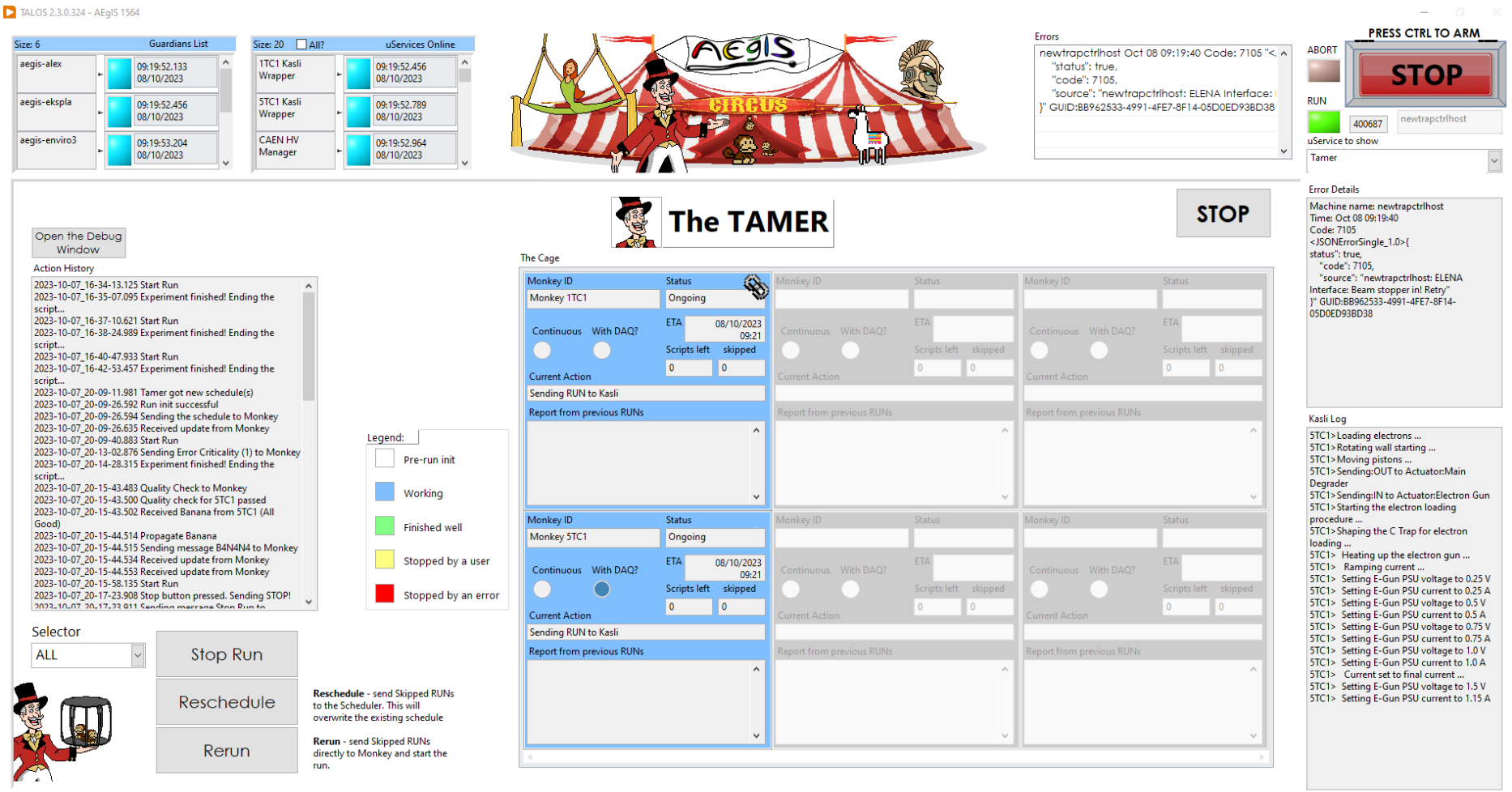


Figure 3.27 – A screenshot of the CIRCUS control system executing a schedule of experiments involving antiprotons. The primary interface is provided by the *Common GUI* (see *Common GUI* (3.4.5.6)). Located in the upper left corner are the Guardians and µServices watchdogs, while the error list is positioned in the upper right corner. In the right-hand column, the top section provides specific details regarding the selected error and, underneath it, a real-time log displays Kasli's operational activities. This interface is uniform across all experiment machines. Within the main window, the *Tamer* µService is displayed, presently overseeing two *Monkeys* performing an active measurement schedule with two Kaslis.

NewTrapCtrlHost	AEgIS-Trap-RT2	AEgIS-Trap-CCD4	AEgIS-EKSPLA	AEgIS-Alex	Enviro-3
TCP Server*					
TCP Listener (5x)*					
TCP Writer (5x)*					
DAQ Sender*					
Common GUI*					
Error Manager	TimingClock	1TCMOS	Avantes Spm	Alex SHG1	Environment Mngr
DAQ Manager	HV	5TCCD	TC200 UV	Alex SHG2	Temperature Mngr
ELENA Interface	5152	Captorius (2x)	TC200 IR	Alex THG1	CryoMotors
Kasli Wrapper (2x)	6133	Kasli UART Logger (3x)	SFG205	Alex THG2	
Telegram Bot	Rotating Wall	Degrader Ladder	HWP532	Alex VBG Rotator	
Monkey	Pulser Manager	Electron Gun	HWP1064	Alex Prism Mngr	
Scheduler			Energy Meter 2	Energy Meter 1	
Positron Reader				Moglabs MWM	
Tamer					
Kasli Writer (2x)					
Kasli Listener (2x)					
PMT Manager					
CAEN HV Mngr (2x)					
Magnet Updater					

Table 3.3 – Topology of all the µServices running on all the computers of AEgIS, at the time of writing. The µServices marked with an asterisk are present, by default, on all machines, and they run headless.

The AEgIS Experiment has not only benefited from the new control system in terms of purely enhanced performances but especially from new capabilities that were unthinkable before. A good example of this is the *ELENA Interface* (3.4.6.2), our interface with the accelerator, which enables the automation of tasks that were previously only possible manually, like the beam steering (see *Autonomous parameters optimisation* (5.1.5)).

But by far the biggest addition in terms of capabilities is the smart automation. It reduces enormously the pressure on the operators, while also enhancing the repeatability of the experiments by lowering the possibility for human errors. This feature was fundamental in achieving all the results presented in the following section.

Chapter 4

Antihydrogen production methods

To success in the difficult task to create (anti-)atoms of antihydrogen, the quality of the hardware and software used is, of course, foremost, but so they are the procedures describing the complex of operations to perform. During the antiproton campaigns of 2021, 2022 and 2023, the knowledge acquired in AEgIS Phase 1 has been utilised to develop the operations needed for $\bar{H}$ production with the renewed apparatus, in the context of the new control system.

In this section, I will first outline some of the physics principles underlying the operations that are described in detail immediately afterwards. These procedures range from loading electrons in the trap, capturing antiprotons, cooling and compressing the $\bar{p}$ plasma, transferring the antiprotons in the formation trap, generating excited Ps, and finally forming antihydrogen.

The main outcomes of these efforts are presented in *Results* (5).

4.1 Physics principles

4.1.1 Sympathetic $\bar{p}$ cooling with e^-

The idea of sympathetic antiproton cooling through electron radiative cooling revolves around harnessing the high cyclotron radiation generated by electron plasmas (which is several orders of magnitude greater than the one emitted by the $\bar{p}$ themselves, see 2.1.8). This power is utilised to induce friction and dissipate the kinetic energy of antiprotons, shortly after their capture, through Coulomb collisions with the electron plasmas. The fundamental principle behind electron radiative cooling is based on the fact that a charged particle, confined within a trap, will lose energy over time via cyclotron radiation since the trap potential is exerting a force – and therefore, an acceleration – on the particle to not let it escape. The formulas for the power emitted both from the axial and from the radial energy components, derived from the power emitted by an accelerated charge (as calculated in [147, 148]), are:

$$\begin{aligned} \frac{dE_z}{dt} &= -\gamma_z E_z(t) \\ \frac{dE_r}{dt} &= -\gamma_r E_r(t) \end{aligned} \tag{4.1.1}$$

where γ_z and γ_r are, respectively:[1]

[1] Here ω_z is the (angular) frequency of axial oscillation of the particles in the trap.

$$\begin{aligned}\gamma_z &= \frac{e^2\omega_z^2}{6\pi\varepsilon_0 mc^3} = 6.25\cdot 10^{-10}\text{Hz}\cdot\left(\frac{m}{m_e}\right)^{-1}\cdot\left(\frac{\omega_z}{10\text{MHz}}\right)^2 \\ \gamma_r &= \frac{e^4B^2}{3\pi\varepsilon_0 m^3c^3} = 0.39\text{Hz}\cdot\left(\frac{m}{m_e}\right)^{-3}\cdot\left(\frac{B}{1\text{T}}\right)^2\end{aligned} \tag{4.1.2}$$

Thanks to its efficiency and simplicity, this technique is prominently employed in AD experiments (as detailed in [149]). In the case of AE$\bar{\text{g}}$IS, the cooling is different if taking place in the 5 T or the 1 T region of the experiment: the characteristic cooling times for $\bar{\text{p}}$ are given in table 4.1.

	5 T	1 T
$\gamma_{r\|e^-}^{-1}$	0.11 s	2.6 s
$\gamma_{z\|e^-}^{-1}$	44 d	44 d
$\gamma_{r\|\overline{p}}^{-1}$	20 years	500 years
$\gamma_{z\|\overline{p}}^{-1}$	212 years	212 years

Table 4.1 – Characteristic radial and axial cyclotron cooling times for electrons and antiprotons in the two regions of AE$\bar{\text{g}}$IS.

When dealing with individual particles, the cooling process predominantly impacts the radial motion of electrons through radiation emission. However, in a highly collisional plasma, the frequent Coulomb collisions between particles serve as an effective mechanism for transferring heat from the axial to the radial degrees of freedom: this results in efficient dissipation of axial energy as well, as explained in [150].

In the context of a mixed plasma consisting of both electrons and antiprotons, as elaborated in [149, 151], the cooling process can be represented, assuming two radially overlapping clouds with differing axial lengths denoted as $\alpha = L_e/L_{\overline{\text{p}}}$, in the case where $\alpha < 1$, by two rate equations that establish the relationship between the two instantaneous plasma temperatures, T_e and $T_{\overline{\text{p}}}$:

$$\begin{aligned}\frac{dT_{\overline{\text{p}}}}{dt} &= -\alpha\gamma_c\left(T_{\overline{\text{p}}} - T_e\right) \\ \frac{dT_e}{dt} &= \alpha\gamma\frac{N_{\overline{\text{p}}}}{N_e}\left(T_{\overline{\text{p}}} - T_e\right) - \gamma_r\left(e^-\right)\left(T_e - T_0\right)\end{aligned} \tag{4.1.3}$$

The first equation models the transfer of energy from the antiprotons to the electrons only through collisions, while the second one links the variation of the electrons' temperature to collision with $\bar{\text{p}}$ (using the same term above, but rescaled with the populations' numbers) and to the dissipation via cyclotron radiation, where T_0 is the limit where only electrons are present in the trap.

Furthermore, in [151] the collision rate is evaluated analytically as:

$$\gamma_c^{-1} = 4\pi\varepsilon_0^2\cdot\frac{3m_e m_{\overline{\text{p}}}c^3}{8\sqrt{2\pi}e^4}\cdot\frac{1}{n_e\log\left(\Lambda\right)}\cdot\left(\frac{k_B T_{\overline{\text{p}}}}{m_{\overline{\text{p}}}c^2} + \frac{k_B T_e}{m_e c^2}\right) \tag{4.1.4}$$

where with $\log\left(\Lambda\right)$ we have indicated the *Coulomb logarithm*, which evaluates the particles' collision impact factor from their temperatures. An analytical expression can be found, for example, in [152]:

$$\Lambda = 4\pi\left(\frac{\varepsilon_0 k_B}{e^2}\right)^{\frac{3}{2}}\sqrt{\frac{T_e}{n_e}}\left(T_e + \frac{m_e}{m_{\overline{\text{p}}}}T_{\overline{\text{p}}} + 2\sqrt{\frac{m_e}{m_{\overline{\text{p}}}}T_e T_{\overline{\text{p}}}}\right) \tag{4.1.5}$$

In [151], a computational analysis was performed by numerically integrating the previous set of equations, considering realistic initial plasma conditions. The research found that the time scale required to cool a hybrid plasma could extend to tens of seconds, under conditions where the

electron density is at least $5 \times 10^7\,\mathrm{cm}^{-3}$ and the proportion $n_e/n_{\bar{\mathrm{p}}}$ does not surpass 10^3. Though this calculation offers a rough estimation for the sympathetic cooling duration in practical scenarios, its application to real-world data, especially at energy approximations around or below the eV, demands careful consideration: in fact, one of the model's underlying assumptions is that the temperatures have no influence over the overlapping of particles' clouds. This assertion often does not correspond with reality, both in the axial degree of freedom, where antiprotons typically form a longer plasma, and in the radial one, where temperatures above $\sim$ eV start centrifugal separation, diminishing the overlap between the two species of particles.

4.1.2 Plasma compression via Rotating Wall technique

Confining a non-neutral plasma in a trap for a long time is not an easy task, since the trap's imperfections and electrical noise will necessarily increase the fluctuations inside the plasma, resulting in a heating process that expands the plasma, ultimately leading to its loss. An efficient technique to counteract the radial expansion is the rotating wall technique (often shortened to RW) [153, 154], which consists of applying a sinusoidal electric field to an azimuthally segmented electrode (divided into N sectors), shifting the wave by an increasing phase $2j\pi/N$ for each electrode (where j is the sector number):

$$\phi\,(j,t) = A_{RW}\cos\left[2\pi m\left(\frac{j}{N} + f_{RW}t\right)\right] \tag{4.1.6}$$

Here A_{RW} and f_{RW} represents the amplitude and the frequency of the wave, while m selects the multipolarity ($m = 1$ for dipole, $m = 2$ for quadrupole, etc.).

From the point of view of the plasma, everything goes as if the entire trap would rotate[2] with a frequency f_{RW}: this couples to the plasma and it can increase or decrease its rotational frequency. The torque exerted will, in turn, modify the angular momentum of the rotating plasma, and hence change its radial distribution.

In fact, in general, an ideal uniformly dense cylindrical plasma (density ρ, particles' charge q) produces a radial electric field:

$$E_r = \frac{\rho q r}{2\varepsilon_0} \tag{4.1.7}$$

Each particle in the plasma therefore starts to rotate around the axis, so that the magnetic centripetal drift compensates for the outward electric force:

$$qE_r = qv_\theta B \quad \rightarrow \quad v_\theta = \frac{E_r}{B} \tag{4.1.8}$$

As a result, the plasma behaves as a rigidly rotating cylinder with frequency f_E:

$$f_E = \frac{v_\theta}{2\pi r} = \frac{E_r}{2\pi r B} = \frac{\rho q}{4\pi\varepsilon_0 B} \tag{4.1.9}$$

Various RW working regimes exist, depending on how the RF driving force couples to the plasma [154, 155, 156], but in general, a RW wave will force the plasma to follow it: so if $f_{RW} > f_E$ the plasma will spin up, while if $f_{RW} < f_E$ the force will act as a drag, slowing the rotation down. In turn, this torque modifies the angular momentum of the plasma cylinder, causing the plasma to radially compress or dilate, as can be seen from equation 4.1.9.

The rotating wall technique inherently increases the energy of the plasma, so it is always needed to couple it to a cooling technique (like *Sympathetic $\bar{p}$ cooling with e^-* (4.1.1)) to have the desired containing effect.

In the case of AEḡIS, being the number of electrons trapped together with the antiprotons order of magnitude greater than the number of the latter, the rotating wall is applied assuming only the

[2] Even if, by physically rotating the trap, it would be impossible to attain angular velocities of MHz and above, due to finite structural strength of materials.

electrons being present in the trap: by the momenta redistribution between the two species by Coulomb collision (as already treated in *Sympathetic $\bar{p}$ cooling with e^-* (4.1.1)), the antiprotons follows naturally, varying their orbiting radius together with the electrons.

In AEgIS, to drive the electric field necessary, a custom-made controller (based on AD9959 DDS DACs) generates pure sine waves from 0 to 50 MHz, with a maximum 2.5 V amplitude with 10 bits resolution. In the near future, it is foreseen to move towards the *Urukul* modules, from the Sinara ecosystem.

4.2 $\bar{H}$ formation procedures

The production of antihydrogen necessitates a series of complex and precise particle manipulations, each needing a substantial amount of consolidation and fine-tuning before developing the subsequent one. In fact, each operation relies heavily on the successful performance of the chain of manipulations preceding it, therefore stabilising and optimising each procedure is paramount before moving to test the following one.

In this section, the first operation described is antiproton capture: although in the chain comes second – after electron loading –, it is consolidated first being the most crucial operation. Then, it is followed by electron loading, paving the way for antiproton cooling and compression. Subsequently, it is described how the electron skimming and antiproton launch towards the 1 T region is performed. Afterwards, the parallel work to bring positrons onto the $e^+ \rightarrow$Ps target and to create high Rydberg Ps is presented. Finally, all these efforts are summed up to arrive at the formation of antihydrogen, where the ballistic formation was tested to produce a forward (and backward) boosted beam of antihydrogen.

The introduction of the *Particle Server* (3.2.5) has been instrumental in the development of the procedures to form antihydrogen. As a matter of fact, it rendered possible to develop, quite in an "agile" manner, the scripts to run on the 1 T Kasli, by iterating quickly over possibilities and failures, while leaving untouched the single sub-procedures to be performed by the 5 T Kasli, which were developed and consolidated separately, mainly in the previous antiproton campaigns. These procedures are implemented as actions triggered by messages on the *Particle Server*: examples are *Antiprotons capture* (4.2.1), *Electron loading* (4.2.2) and *$\bar{p}$ cooling and compression* (4.2.3).

In the following, the main steps outlined before are explained, together with the plots showing the corresponding shape of the potentials given to the traps' electrodes.

4.2.1 Antiprotons capture

The first and most important procedure is, of course, the trapping of the antiprotons coming from ELENA (see *Antiprotons production in AD* (2.1.4)). As mentioned in *Degrader(s)* (2.1.7), first a material degrader is used to slow the incoming antiprotons from the initial 100 keV to below 20 keV. Then, the trapping of the degraded antiprotons is performed using the high-voltage electrodes HV1 and HV3[3] (see figure 2.6): HV3 is raised to 15 kV well beforehand the $\bar{p}$ bunch arrives, while HV1 is connected to a fast high-voltage switch, so to be able to raise it to 15 kV in few ns.

The trapping is then performed by tuning the raising of (the voltage on) HV1 so that the $\bar{p}$ bunch is all inside the trap when the operation starts but should arrive to its maximum value before the $\bar{p}$ reflected by HV3 would escape again from the rear end of the trap. For antiprotons of 20 keV energy, given a total length of the trap of the order of 1 m, the average residency time is ($\bar{p}$ are not relativistic at this energy):

$$\delta t = \frac{2L}{v} = \frac{2L}{\sqrt{\frac{2E}{M}}} \approx 1\,\mu s \tag{4.2.1}$$

[3]It can also be performed between either HV1 and HV2, or HV2 and HV3, but clearly, the combination HV1 and HV3 maximises the length between the electrodes, easing the trapping of the entire bunch.

The trapping operation is optimised to maximise the number of $\bar{p}$ captured[4] at the beginning of every yearly campaign, and the results for the best closing time, available in *Autonomous parameters optimisation* (5.1.5), are in good accordance with the estimation done above.

After the electrode closing time is optimised, the number of antiprotons captured is then increased by properly steering the particle beam coming from ELENA into the centre of the trap. The electromagnets at the end of the ELENA delivery line can be adjusted by setting four parameters, specifically *horizontal* and *vertical offset*, and *horizontal* and *vertical angle*. This operation is paramount to achieve high $\bar{p}$ trapping efficiency, and it has to be repeated regularly to ensure optimal efficacy: therefore, it was fully automated, and results can be seen in *Autonomous parameters optimisation* (5.1.5).

4.2.2 Electron loading

Antiprotons captured in the trap are bound to be lost by radial losses if not properly cooled[5]: therefore, before performing the actual trapping, electrons need to be loaded into the 5 T region, so to enable electron sympathetic cooling of antiprotons (see next section).

To do so, first, a trap for the electrons is created by biasing some electrodes (from P4 to P8) to a positive potential of 150 V, to make a potential well where the electrons can fall inside. The electrodes downstream (P9, P10, P11) are then biased to −180 V, in order to create a barrier to reflect the e^- coming from the e-gun (see *Electron Gun* (2.1.8)). In this way, the reflected electrons collide with the incoming ones, lose energy and ultimately fall inside the potential well. The current is kept on for 14 s, so to be sure to fill the well up to the space-charge limit (order of 10^7 e^-). The electrons' number can be destructively measured by opening the trap towards a FC.

The potential configuration of the trap for the electron loading is shown in figure 4.1. On the x-axis, the experimental z-coordinate is used, with the zero placed at the beginning of the 5 T trap (as in figure 2.6), and on the y axis the real potential at the centre of the electrodes is plotted. For further clarity, the drawing of the trap is placed on top of the image, aligned and in scale with the values on the x-axis, to immediately visualise the position of the particles. To calculate the real potential on the axis, the influence of all the electrodes is taken into account: using a finite-elements model (FEM), the potential on the axis for every mm z-position was determined when setting a specific electrode to 1 V, and saved as a matrix. Afterwards, the potential is calculated by convolving this matrix with the actual voltages set on every electrode. This method is particularly useful when dealing with traps with end-caps of not equal length (that can produce asymmetric potential wells if biased identically), or when potential-precise operations are needed (like for skimming electrons before launching antiprotons, see *Antiprotons transfer* (4.2.4)).

Moreover, electrons were also employed to find the geometrical axis of the trap on the MCP, and to centre the e-gun flux to it. Initially, electrons were loaded into the trap until reaching the space-charge limit, as outlined just before; then they were released and imaged with the MCP, and the plasma centre on the image is used as a reference point (marked by a yellow cross in figure 4.2) to align the main current. Afterwards, electrons are pulse-extracted from the e-gun and were visualized on the same MCP. The electron gun position and angle were then adjusted manually using alignment screws between each pulse, aligning the current's centre (indicated by a red cross in figure 4.2) with the plasma's centre. This new procedure enabled us to precisely align the e-gun to the axis of the traps and opened the possibility for aligning other particle sources with the MCP (for example, determining the beam steering with the un-degraded direct $\bar{p}$ beam, by centring the beam and minimise impinging area – i.e. particle spiralling –, ultimately aligning the particle with

[4]The number of $\bar{p}$ captured is obtained by slowly opening the containing electrode HV1, so to let the particles spill over towards the degrader (a procedure called "slow hot dump"), and then integrate the annihilations recorded by the PMTs near the degrader over time.

[5]For high energy antiprotons, the losses for annihilations with the residual gas are actually negligible: in fact, the cross-section of annihilation is heavily dependent on the average temperature[157], so an antiparticle will virtually never annihilate if not stopped before.

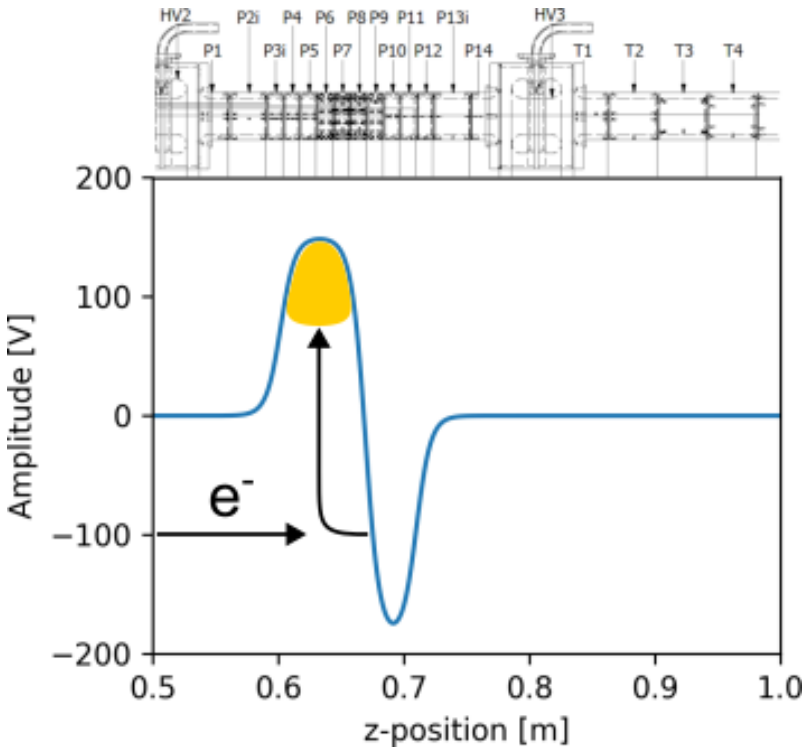


Figure 4.1 – Potential on the axis of part of the 5 T section of the traps used to perform electrons loading. The drawing of the trap is put on top of the graph for clarity: it is in scale and aligned with respect to the x-axis of the graph, to immediately visualise the position of the particles. In yellow the electron plasma is stylised.

respect to the magnetic field lines).

4.2.3 $\bar{p}$ cooling and compression

As explained in detail in *Sympathetic $\bar{p}$ cooling with e^-* (4.1.1), to obtain a reduction of temperature of the antiprotons in the trap (on a human timescale), electrons play a fundamental role: therefore, the trapping procedure actually starts with loading e^- in the trap. For this purpose, we proceed as described in the previous section, ramping down the negative barrier once the well is full of electrons, to arrive at the potential configuration shown in figure 4.3. Then the procedure follows what is explained in *Antiprotons capture* (4.2.1): the antiprotons trapped between the high-voltage electrodes slowly lose energy by collision with the cold electrons, slow down and eventually fall also inside the positive potential well (typically antiprotons are left cooling for ~ 30 s). Subsequentially, the high-voltage electrodes are ramped down, so to discard the small part of $\bar{p}$ that did not cool in the trap[6].

After the cooling, the rotating wall procedure typically takes place, to compress the antiproton plasma to a smaller radial size. The operation was successfully conducted, compressing the plasma of about 50 % of its original size; ultimately, it was decided not to perform it by default, since it was not strictly necessary with the new, larger electrodes in the formation trap (the initial plasma radius was small enough not to cause too many losses during the transfer, see *Antiprotons transfer* (4.2.4)). Even more, it could have been detrimental to antihydrogen beam formation: in fact, having a $\bar{p}$ plasma too concentrated on the axis would produce $\bar{H}$ that would necessarily collide with the Ps target, now on the axis and masking a good portion of the trap section. Most of the antihydrogen

[6]The cooling is performed only on the the part of the $\bar{p}$ plasma that shares the same volume with the electrons. If, for example, the $\bar{p}$ plasma is radially more extended than the e^- one (as is often the case), a circular corona of hot antiproton will remain uncooled.

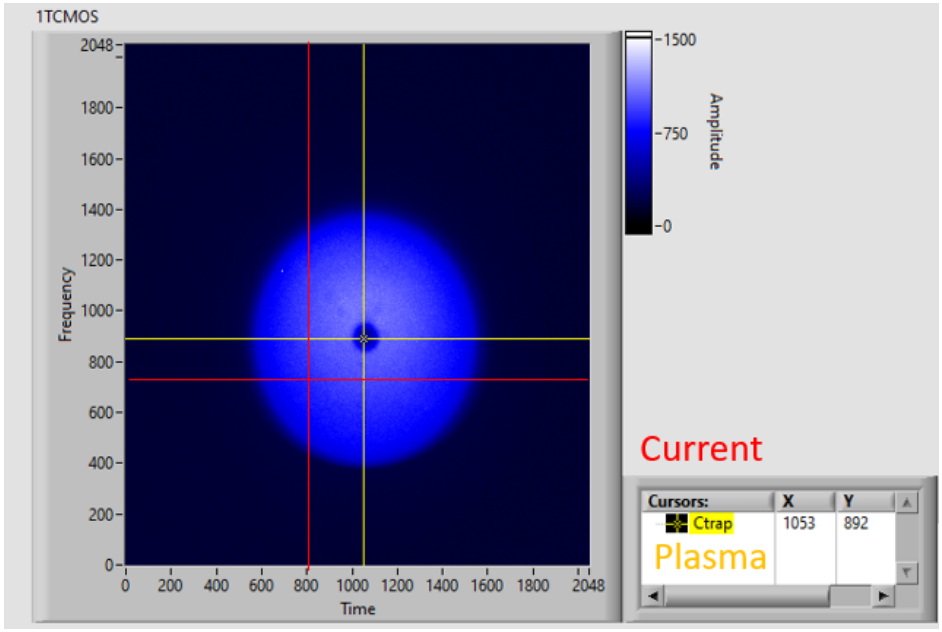


Figure 4.2 – Example of the steering procedure used to centre the electron current on the experiment axis with high accuracy. The centring is obtained by aligning the e-gun current (whose position corresponds to the red cross in the figure) with respect to the centre of the electron plasma dumped from the trap (yellow cross) (from [6]).

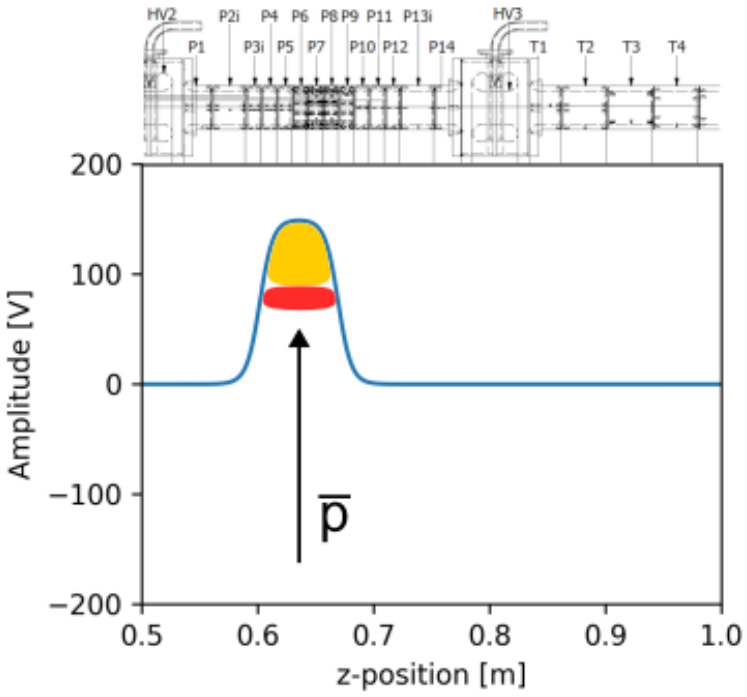


Figure 4.3 – Potential on the axis of part of the 5 T section of the traps to perform $\bar{p}$ sympathetic cooling. The drawing of the trap is put on top of the graph for clarity: it is in scale and aligned with respect to the x-axis of the graph, to immediately visualise the position of the particles. In yellow the electron plasma is stylised, in red the antiproton one.

forming a beam would in reality come from a ring of radius bigger than $\sim 5\,\text{mm}$. By not performing the RW, the complexity of the procedure is reduced and the total operation time is shortened.

Afterwards, the electrodes P3 and P9 are raised to a negative *high*[7] potential (around $-180\,\text{V}$), and all the electrodes in the middle (the *floor* of the trap) to a slightly negative potential ($\sim -10\,\text{V}$): this is what we called "ready to dump" configuration, and it is the standard configuration from where procedures involving the use of cold antiprotons begins (the potential configuration is displayed in figure 4.4). The name of the configuration originates from the fact that, by simply opening one of the two *wall*[8] electrodes (i.e. either P3 or P9), the antiprotons would escape toward the desired end, since the trap floor is biased to a slightly negative potential (while outside the trap, the potential is kept to ground). This procedure can be used, for example, to calibrate annihilations seen by the PMTs with the charge measured on one of the FC (MCP or degrader).

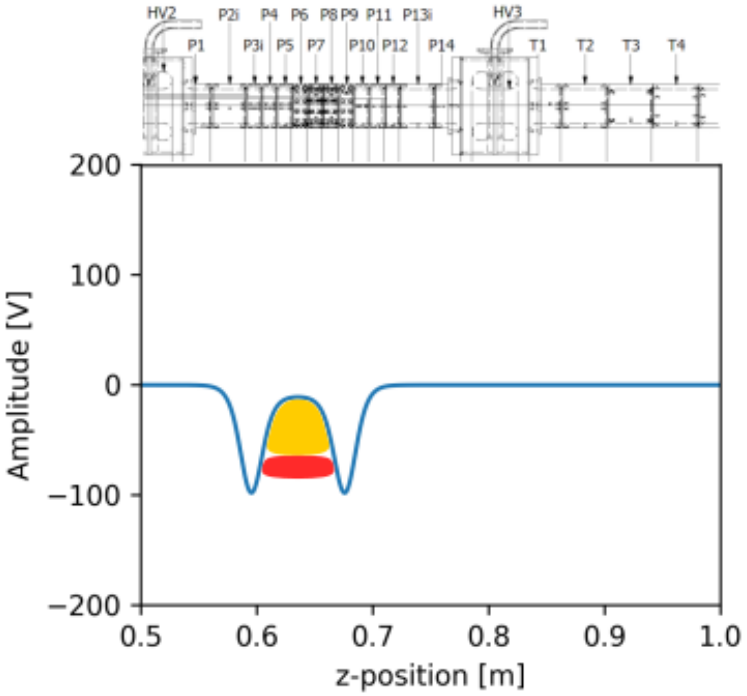


Figure 4.4 – Potential on the axis of part of the traps used as a standard starting point for operations with antiprotons. The drawing of the trap is put on top of the graph for clarity: it is in scale and aligned with respect to the x-axis of the graph, to immediately visualise the position of the particles. In yellow the electron plasma is stylised, in red the antiproton one.

4.2.4 Antiprotons transfer

To create antihydrogen, the antiprotons need to be transferred to the 1 T side of the traps' system. To do this, the potential well where they are kept needs to undergo a series of potential manipulations, to arrive at a shorter trap with a higher floor. The first condition is necessary in order to compress the $\bar{\text{p}}$ plasma longitudinally, to minimise the horizontal spread and better concentrate the $\bar{\text{p}}$ when they arrive close to the Ps target; the second condition, instead, guarantees a faster flight of the particles throughout the central region of the experiment, where the magnetic field abrupt gradient causes plasma modifications[9].

[7] When we refer to normal electrodes (not the "high-voltage", i.e. kV ones), for *high voltage* we mean a potential between 150 and 200 V absolute.

[8] Also called *end-caps*.

[9] Actually, when going from the 5 T to the 1 T magnetic field, the particles undergo an "inverse magnetic bottle" effect, with the result of collinearisation of the momentum of the particles, a welcomed effect. But when antiprotons

The trap profile mutation is materialised in four steps, shown in figure 4.5. First, the electrodes before the upstream end-cap and after the downstream end-cap (P2, P10, P11, P12 and P13) are raised to match the same high-voltage (figure 4.5a), so to later elongate the floor of the trap to effectively create a longer trap (figure 4.5b). The trap is then squeezed from upstream, by raising each electrode a little (typically $\frac{1}{100}$ of the total excursion to be done, i.e. around 1.7 V), starting from the upstream one, and progressively moving into raising each one on the right, so to form a reclined potential slope moving to the right (figures 4.5c and 4.5d illustrate the manoeuvre: the slope is accentuated to better visualise it), ending with a very short trap (figure 4.5e). The floor of the trap is then raised to a higher potential (~ -70 V) slowly, to minimise $\bar{p}$ spilling (figure 4.5f).

Starting from the latter configuration, further reshapes are needed. In fact, it is mandatory to separate and remove the electrons from the antiprotons' plasma: if launched together, being ~ 1836 times lighter, the electrons would bounce back and forth from the trap to the end of the 1 T trap several times while the antiprotons are flying, ultimately perturbing the $\bar{p}$ plasma.

Moreover, if launched directly on a constant grounded potential, the $\bar{p}$ plasma would undergo an immediate abrupt expansion, from which it would be impossible to recover by simply letting the particle climb an uphill potential toward the end. To overcome this problem, a solution was found to create a parabolic potential going from the bottom of the *launch trap* (i.e. electrode P12) all the way to the end of the 1 T trap (i.e. electrode A0 to ~ -70 V), with the lowest point[10] around ~ 0 V.

In this manner, the $\bar{p}$ cloud slowly accelerates, passing fast through the magnetic field gradient, and then it slows down and halts just in front of the Ps target, without spreading substantially. In addition, this method guarantees that the $\bar{p}$ will do the same flight pattern when going back (except for the $\vec{B}$ gradient, now reversed), lowering annihilation rate – which would ultimately lead to an increased background during the antihydrogen annihilation – and enabling the possibility for "antiprotons recycling".

The idea of antiprotons recycling lies in re-trapping them after one full swing (back and forth), so to stack the antiprotons captured from the following shot coming from ELENA on top of them, eventually increase their number at the subsequent launch: if correctly performed, this operation could greatly increase the amount of $\bar{H}$ formed (by a factor equal to the number of shots stacked, minus the losses).

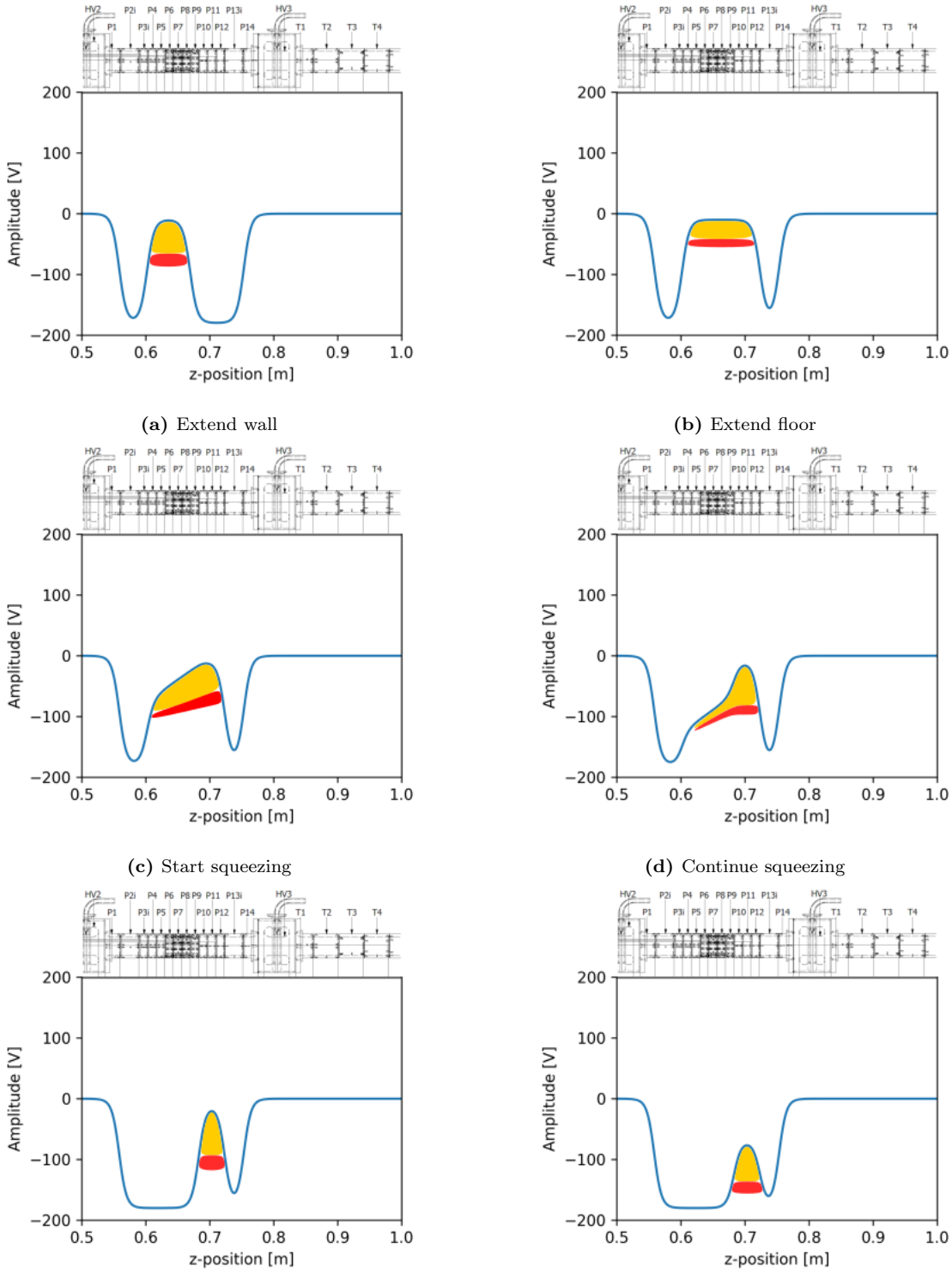
To remove the electrons, and later attempt to re-catch the antiprotons, a second trap is formed on the left of the the *launch* one, called *recycling* trap (see figure 4.6a). The electrode separating the two traps, P9, is kept to a potential that is high enough not to let $\bar{p}$ spill over, but low enough so that, when pulsed, it is possible to create a plateau between the two traps. This minimises the acceleration given to the particles such that, by fine-tuning the pulse length given to P9, it is possible to separate the two species, leveraging their mass difference (see figure 4.6b). Multiple trials were conducted, by first performing the manoeuvre, subsequently opening in turn the two traps towards the MCP, and finally looking at the signal on the MCP itself and on the scintillators: they confirmed that indeed it was possible to remove more than 90 % of the electrons present, while retaining more than 90 % of the antiprotons in the *launch* trap.

The removal of the electrons is an operation that requires fine-tuning of the potentials and the timing of the pulses, enabled by the precision of Kasli and by the resolution and extremely fast potential raise of AE$\bar{g}$IS pulser. Forming the parabola, instead, is less time-critical, but it requires the coordinated behaviour of both Kaslis, made possible only by the *Kaslis ns synchronisation* (3.2.4) and the *Kaslis' Server* (3.4.5.10), which culminated into the *Particle Server* (3.2.5). In fact, the operation is initiated by the 1 T Kasli sending a message to the 5 T one to reshape all the electrodes after P14 to form that part of the parabola, specifying the height of the potential of P12 (which is one of the two ends of the parabola, the other one being A1)[11] and the potential in the middle of

come back (which happens if the transfer is done ballistically, and not adiabatically), the magnetic field gradient opens up the particle cloud and unwelcomed annihilations occur.

[10] Which is actually the highest point of the potential, being the trap negative!

[11] The potential of P13 does not get modified at this moment, since it still acts as a wall of the *launch* trap: but



(a) Extend wall

(b) Extend floor

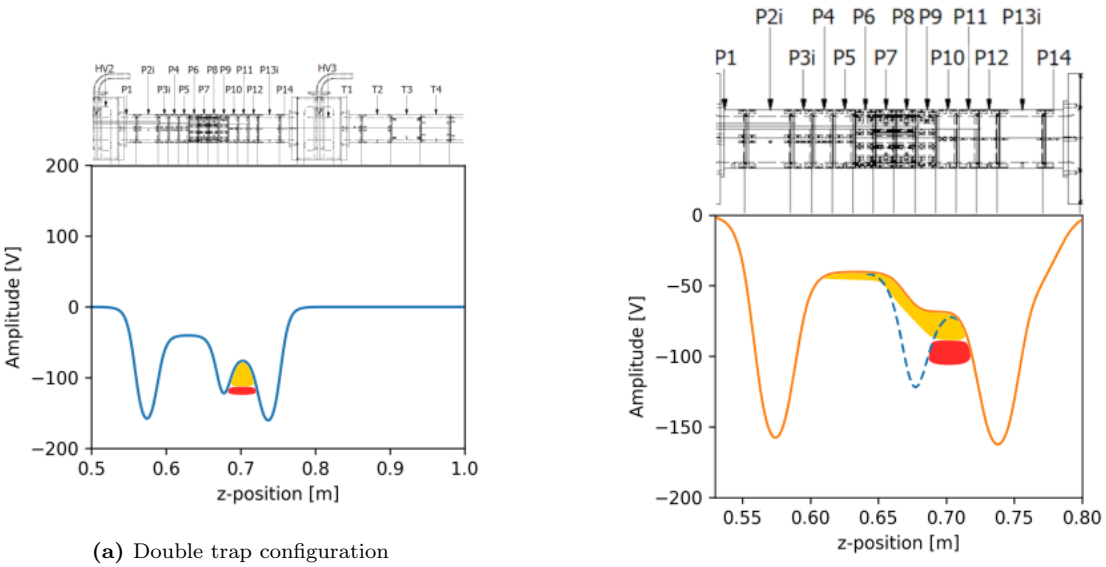
(c) Start squeezing

(d) Continue squeezing

(e) Squeezed trap

(f) Raise floor

Figure 4.5 – Potential on the axis of part of the traps showing the reshaping of the trap profile to pass from the standard "ready to dump" configuration to the raised and squeezed trap to launch $\bar{p}$. The drawing of the trap is put on top of the graph for clarity: it is in scale and aligned with respect to the x-axis of the graph, to immediately visualise the position of the particles. In yellow the electron plasma is stylised, in red the antiproton one.



(a) Double trap configuration

(b) Electron removal

Figure 4.6 – On the left: potential on the axis of the traps propaedeutic to the removal of the electrons, and to attempt $\bar{p}$ recapture. On the right: zoom on the potential profile when P9 is pulsed (full line versus dashed addition). It is seen the plateau formed between the two traps, letting the electrons quickly escape: if P9 pulse length is properly timed, the $\bar{p}$ do not move enough and the two species are separated. In both figures the drawing of the trap is put on top of the graph for clarity: it is in scale and aligned with respect to the x-axis of the graph, to immediately visualise the position of the particles. In yellow the electron plasma is stylised, in red the antiproton one.

the parabola, so that the shape of the latter can be dynamically determined (the resulting potential is shown in figure 4.7).

Afterwards, also the 1 T Kasli changes the potential of the electrodes of its part of the traps' system. Furthermore, it also sets the voltage of the electrode of the target holder (A0) and the two ionisation grids (see *1 T section* (2.1.5.2)), so to both let the antiprotons being transmitted or reflected back, and the eventually formed antihydrogen being ionised. The full potential shape of the entire traps' system ready to perform the antiprotons ballistic transfer is plotted in figure 4.8.

Putting everything together, the full launch procedure, starting from the "ready to dump" potential (the final point of *$\bar{p}$ cooling and compression* (4.2.3), shown in figure 4.4), goes as follows. First, the potentials are reshaped as described above, so to arrive at the configuration displayed in figure 4.8. Afterwards, a pulse ~ 40 ns-long is given to P9, so to skim away the majority of the electrons (sending them inside the *recycling* trap, or losing them on the degrader if also P2 is pulsed), as explained above. A waiting time of $\sim 10\,\mu$s is reserved, to let the $\bar{p}$ inside the launch trap re-equilibrate from the electron removal. Then, P13 is pulsed, and the antiprotons start to swing towards the end of the 1 T trap; the length of this pulse depends on the operation that is wanted for the $\bar{p}$. In case they are sent against the MCP for a destructive measurement, or if they are let bounce inside the parabola (for example, to precise determine the oscillation period: see figure 5.19), a pulse length of $10\,\mu$s is more than sufficient to guarantee total drainage of the launch trap, while ensuring re-entering to be impossible, since the oscillation period inside the parabola is greater than that (see equation 4.2.2). If, instead, they are meant to be re-trapped inside the recycling trap, the pulse has to be carefully tuned so to let them in when coming back, and to impede their re-exit shortly after.

To estimate the antiprotons oscillation time inside the parabola, we can calculate the oscillation

the amplitude of the pulse to be given later is calculated so that also this electrode will nicely smooth the curve.

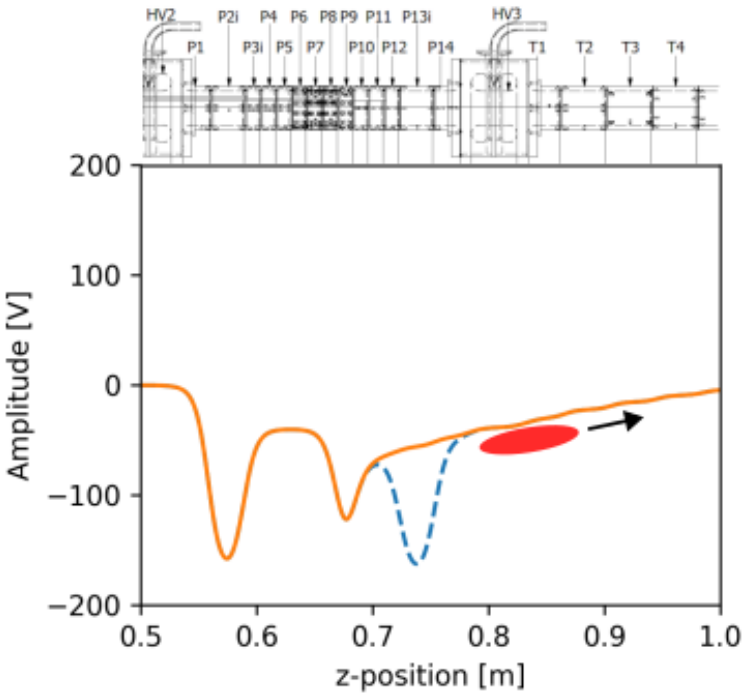


Figure 4.7 – Detail of the potential on the axis of the traps used to perform the ballistic transfer towards the 1 T section of the trap: here only the 5 T section is shown. The dashed line shows the shape difference between the parabola preparation (with P13 raised), and the transfer occurring (continuous line, with P13 pulsed). The drawing of the trap is put on top of the graph for clarity: it is in scale and aligned with respect to the x-axis of the graph, to immediately visualise the position of the particles. The red bubble represents the antiproton plasma moving towards the formation trap.

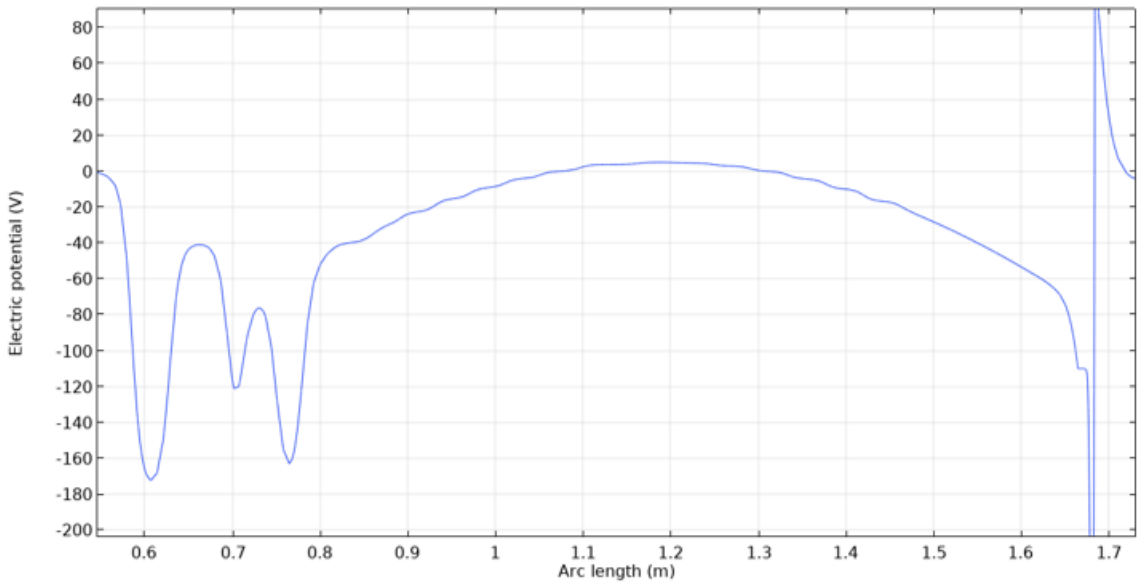


Figure 4.8 – Potential on the axis of the entire traps' system used to perform the antiprotons ballistic transfer (from [7]).

period of an (anti)proton inside a parabolic potential well of $L = 0.93\,\mathrm{m}$ and $V = 75\,\mathrm{V}$:

$$T = 2\pi\sqrt{\frac{mL^2}{8V}} \approx 24\,\mu\mathrm{s} \tag{4.2.2}$$

This value is in good agreement with what was found empirically.

For test purposes, the recycling was tempted, and a pulse length of $\sim 23\,\mu\mathrm{s}$ indeed demonstrated the capability of re-trapping the antiprotons into the *recycling* trap. Nevertheless, this operation was deemed a refinement to be developed only after the consolidation of the full procedure for antihydrogen formation and therefore left for future exploration.

In addition, tests of $\bar{\mathrm{p}}$ adiabatic transfer were also performed: with a series of steps of elongation and subsequent reduction of one electrode at a time towards the Ps target, the *launch* trap was slowly moved downstream. Unfortunately, this transport method yielded unsatisfactory results, mainly because of the long permanence inside the magnetic gradient section, causing huge particle losses.

4.2.5 Ps formation

Until now the procedures dealt with the preparation of the antiprotons, but positronium is the other fundamental pillar for the formation of antihydrogen. The procedures to form and excite Ps has been therefore carefully re-developed and tested, in the context of the new control system and the new collinear geometry, separately from the antiprotons manipulations, being independent from them.

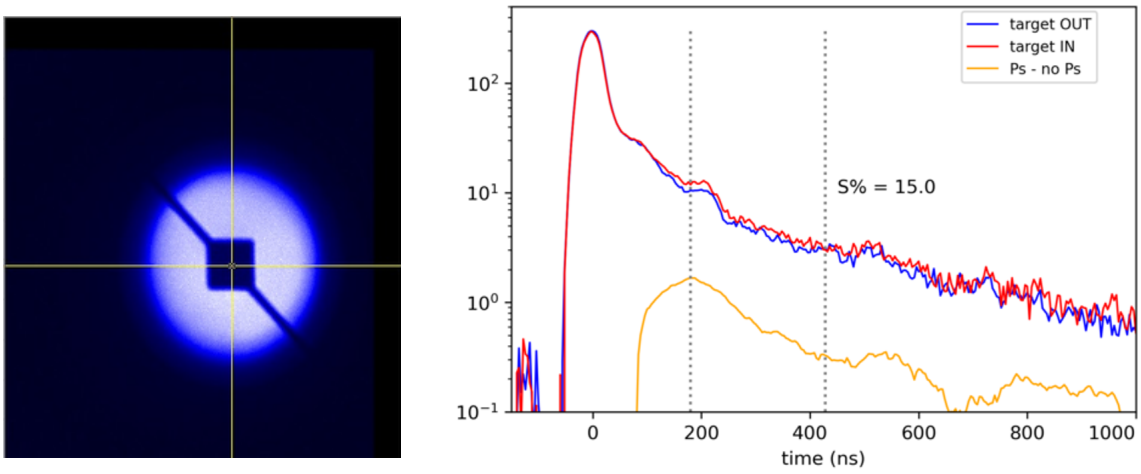
Everything starts, as explained in *Positron system* (2.1.10), with the accumulation of positrons in the *Accumulator*: the accumulation time was set to be 110 s, in order to match the ELENA cycle to which it needs to be synchronised when used for $\bar{\mathrm{H}}$ production. In this period, the system is able to store approximately $2 \cdot 10^6$ e^+ with the 6 mCi ^{22}Na activity source of autumn 2023, which are then bunched and accelerated to 2.5 keV towards the Ps target in the main apparatus, in order to have the same implantation conditions (3.3 keV) as in the Ps test chamber where the target is mounted under 45° angle (hence $3.3/\sqrt{2}$). The energy of the particles is tuned such that they can easily enter into the 5 T magnetic field at the beginning of the traps, and be implanted at the correct depth into the silica target[12].

For the precise alignment of the bunch (measured to have a temporal spread of about 35 ns FWHM and a transverse diameter of $\sim 2\,\mathrm{mm}$ FWHM) with the target centre, a novel method has been invented, leveraging once more the MCP at the end of the 1 T trap. First, a "picture" of the target was taken by launching electrons (previously loaded into the P trap) against the MCP, with the Ps target electrode inserted: the shadow of the target is projected clearly onto the MCP (see figure 4.9a). Afterwards, with the target electrode removed from the beamline, the e^+ bunch can be steered at the centre of the $\mathrm{e}^+ \rightarrow$Ps target shadow taken before, guaranteeing a perfect alignment.

Notwithstanding the careful regulation, positronium production evidence was lacking, when compared to implantation into the aluminium border of the target electrode, where no Ps formation is expected. After a long trial and error, positronium formation was finally established, but only when the silica target was heated (and kept) at 300 K: the amount was, in any case, lower than expected, by almost one order of magnitude (see figure 4.9b).

Several hypotheses have been formulated for the cause of such poor efficiency: the most probable is the long exposure to water in air, at room temperature, which might have caused clogging of the nanochannels, and even damage to them when cooled down. Other suppositions lay into: exposure to oxygen from the air (at room temperature, as before), causing a thickening of the SiO_2 layer;

[12] Moreover, when the implantation is performed for $\bar{\mathrm{H}}$ production, this energy also helps to avoid deviations from inelastic collisions with the $\bar{\mathrm{p}}$ cloud they need to traverse to get to the Ps target (in fact, for Coulomb scattering, $\theta_{deflection} \approx \frac{Ze^2}{4\pi\epsilon_0 b}\frac{1}{v^2}$).



(a) Shadow image of the Ps target.

(b) Ps production evidence.

Figure 4.9 – On the left: image of the $e^+ \rightarrow$Ps target, taken by launching electrons towards the MCP with the target electrode inserted. Its shadow is cast onto the MCP, and it is later used to centre the positron bunch on the target (by fixing the yellow cross). On the right: annihilations measured with the PMTs by launching positrons with (red) and without (blue) the target inserted. The curves difference proves that Ps was formed (both from [7]).

bad etching due to contaminants into the acidic solution; adsorption of pollutants like silicone, acetone, isopropanol, etc., used in the target surroundings during the cryostat closure process.

Albeit more improbable, it could also have been a combination of these causes: further investigations will be carried out during 2024 in order to establish the source of the problem to avoid and/or mitigate it in future Runs.

4.2.6 Ps excitation

The positronium atoms formed, as described in the previous section, necessitate being excited to a high Rydberg level in order both to augment the $\bar{H}$ cross-section and to prolong the Ps lifetime enough to reach and have sufficient time for the interaction with the $\bar{p}$.

To attain a good excitation efficiency is fundamental to have a good spatial and temporal superposition of the laser beams with the particles. The spatial alignment has been performed by manually steering each of the three lasers of the EKSPLA (see *EKSPLA* (2.1.6.1)) separately, while imaging the light with the system described in *Laser positioning with fibres bundle* (2.1.9.6): this method has proven to be very efficient, both in terms of operators' time and precision.

The temporal coordination has been achieved by synchronising the light of each laser, imaged this time with a fast photo-diode (few ns resolution), placed inside the laser cabin, before the laser enters the viewport that brings the light inside the experiment, and connected to a Teledyne LeCroy HDO6104A-MS oscilloscope, to the annihilations occurring when the e^+ are implanted into the Ps target. Taking into account particle dynamics and the various delays (cables, laser travel time, and PMTs response (41 ns)), the synchronisation was performed, as shown in figure 4.10.

Subsequentially, a measurement to estimate the excitation efficiency was carried out. To better align to the needs of the $\bar{H}$ campaign, it was decided to confront the annihilations occurring when firing both the 205 nm UV and the 1675 nm IR laser (leading to Ps in the $n = 21$ state), and when firing only the ultraviolet one, which would bring Ps to 3^3P, but only for 12 ns, before spontaneously decaying again into the ground state. These two classes of measurements were chosen because antihydrogen formation can only occur when both lasers are active: in fact, in the second case, the production is both heavily suppressed by the cross-section ($\propto n^4$), and especially by the interaction

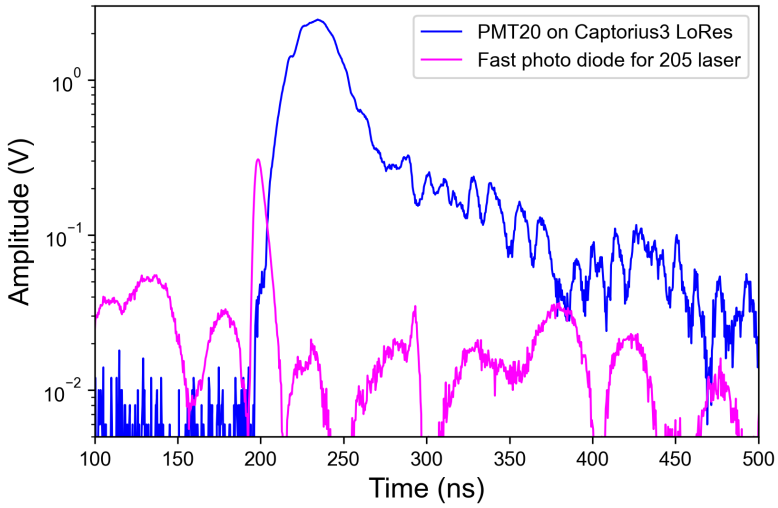


Figure 4.10 – Temporal synchronisation of the 205 nm laser beam with the annihilation deriving from positrons implantation into the Ps target, both occurring around 190 ns in the plot (from [7]).

time[13], since Ps ground state lifetime is much shorter. Moreover, the choice of using as control measurement the case with the UV laser active, with respect to just not firing any laser at all, was motivated by the need to change as minimum as possible the background on the the MCP between the measurement and the control class (see *Antihydrogen formation* (5.3): for example, the UV laser can cause photo-electron emission from the ionisation grid. Not firing it in the control class would cause an excess of signal in the measurement class that could be mistaken for antihydrogen).

In figure 4.11a the annihilations occurring in the two measurement classes just described are displayed. The two lines are almost overlapping due to the very little positronium formation achieved (due to the problems described before), and therefore the background, generated by the e^+ annihilation in the target, is dominant. To obviate this problem, in figure 4.11b the relative difference between the two curves is given, using a moving window with a width of 200 ns: between 300 ns and 800 ns a negative signal is clearly present, pointing towards a reduced annihilation rate, evidence of high Rydberg excitation (with its long lifetime and thus reduces in-flight annihilations).

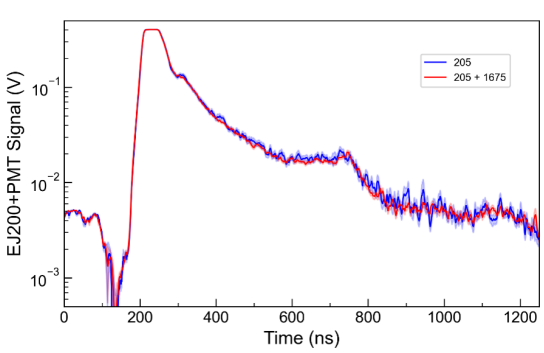
4.2.7 $\bar{H}$ formation

The two parallel developments for manipulating the antiprotons and the positronium atoms were finalised to the formation of antihydrogen, with the additional tentative aim of forming it as a forward-boosted beam.

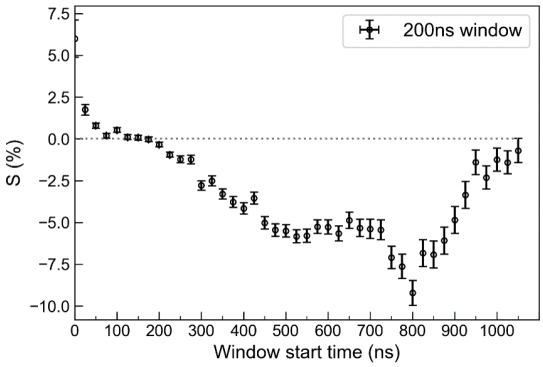
Practically, this consisted of combining all the aforementioned operations into a single procedure, by carefully merging the various parts of the experimental scripts. When a functional script was assembled and tested, it manifested the need for a precise synchronisation between the launching of the $\bar{p}$ and the Ps formation. To achieve it, we measured the time difference between the annihilation of positrons on the target and the one of $\bar{p}$: for a careful determination, the two timings were evaluated separately, and in particular for the antiprotons a voltage scan was performed, by varying the potential of the target holder from -200 V to progressively higher voltages, so to find the voltage when the $\bar{p}$ were just grazing the target surface (and, hence, the annihilation peak appearance).

In figure 4.12 the peak of annihilations of the positrons impinging on the target is visible: from it, we determined that the delay between the launch trigger and their arrival was 70.73(1) µs. In figure 4.13, instead, the variation of the annihilations of the antiprotons on the target varying the potential set on the target holder electrode is plotted. In this manner, the annihilation peak was

[13]Actually, it does not even reach the $\bar{p}$ cloud.



(a) Ps annihilations with UV+IR laser vs. UV laser only.



(b) Ps annihilation difference in the two measurement classes.

Figure 4.11 – On the left: comparison between the annihilation of Ps when only the ultraviolet laser is shot (blue), versus when both UV and IR are present (red). On the right: the relative difference between the two curves on the left plot, using a moving window 200 ns large. The negative values between 300 ns and 800 ns is an indication of high-Rydberg Ps excitation (both from [7]).

established to occur 67.4(5) µs after the launch. The time spread of the positron bunch is 35 ns FWHM (see *Ps formation* (4.2.5)), therefore comparable with the uncertainties of the measurements, and it has been neglected.

From these two measurements combined, the timing for having the antiprotons cloud stationary in front of the $e^+ \rightarrow$Ps converter was found: it corresponds to delaying the launch of the antiprotons 3.3 µs after triggering the e^+ bunching.

Afterwards, to further enhance the spatial superposition between $\bar{p}$ and Ps, a technique similar to what explained in *Ps formation* (4.2.5) was used to align the antiproton cloud with the target. The antiprotons were launched ballistically from the P trap, and the sectorised central electrode (B0, see *5 T section* (2.1.5.1)) was used to steer the beam both in the x and in the y direction: finally, the $\bar{p}$ were imaged on the MCP, leaving the target inserted so that its shadow was cast on the image itself. The alignment so achieved can be seen in figure 4.14.

To determine the formation of antihydrogen, a binomial test was set up, so to establish if there is a statistical difference between runs where antihydrogen formation is enhanced versus the ones where it is suppressed (as explained in *Ps excitation* (4.2.6)). Furthermore, to test the possibility of forming $\bar{H}$ with a precisely directed momentum, two further measurement classes were added: one where the antiprotons are launched 500 ns earlier than the timing determined to have them stationary when Ps emission occurs, and one where they are 500 ns after. In fact, in the first case (indicated with "-500 ns", or "Backwards"), the antiprotons have already started to go back when they are illuminated with Ps: therefore the antihydrogen atoms thus formed, free from the trap electromagnetic confinement, retain the momentum the antiprotons had during the formation, which is directed backwards. A similar argument applies to the second class ("+500 ns", or "Forward"), which generates a forward-directed beam of antihydrogen. An estimation of the velocity the $\bar{H}$ formed in such manner can be calculated by modelling the movement of the antiprotons inside the parabolic potential as a perfect harmonic oscillator (as done in *Antiprotons transfer* (4.2.4)), and evaluating the value at $\tau \pm 500$ ns. The momentum transferred by the positronium atom, and the one caused by the electron emission during the charge-exchange reaction are disregarded, thanks to the fact that $\frac{m_p}{m_e} \approx 1836$. The antiprotons bunch length at the end of the parabola is comparable to the one at the start, which is in the order of centimetres, corresponding to the size of the launch trap, and has been disregarded as weel, in first approximation. The equation of motion of the antiprotons inside the parabolic well is, then:

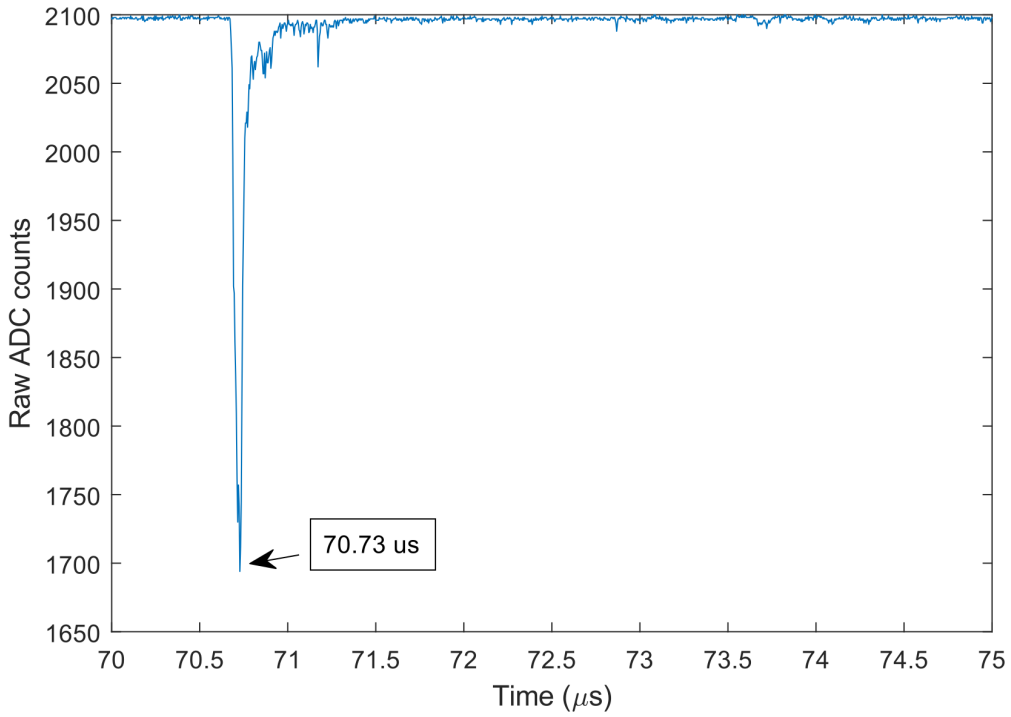


Figure 4.12 – Plot showing the peak of annihilations of the positrons launched against the target: the time of arrival was determined to be 70.73(1) µs after the launch trigger was issued (courtesy of Ruggero Caravita).

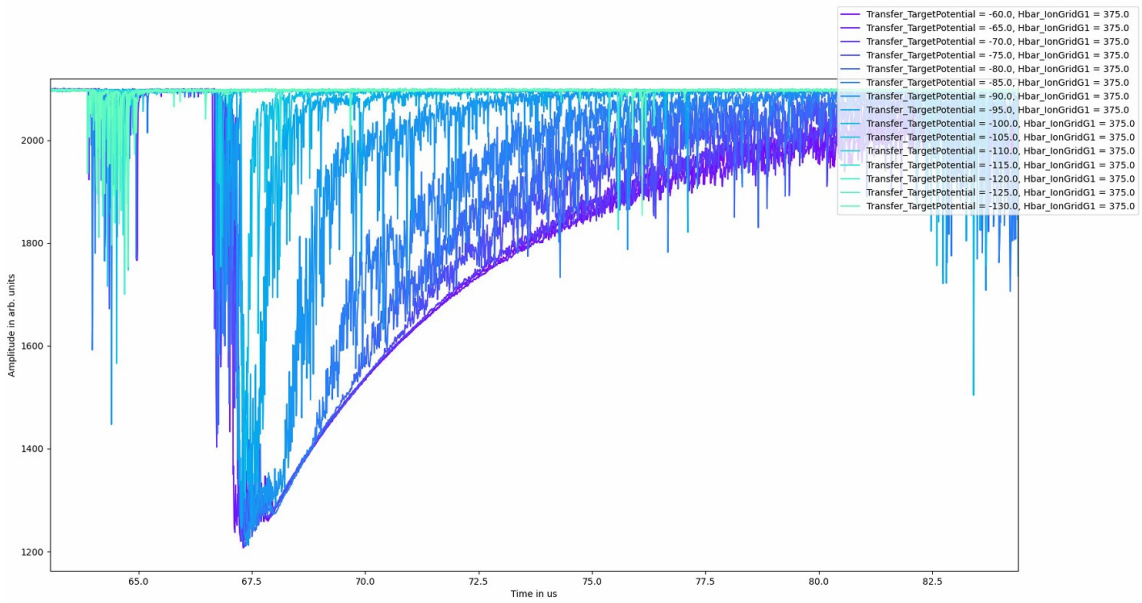


Figure 4.13 – Plot showing the variation of the annihilations of the antiprotons on the target, varying the potential of the target holder electrode. The peak of the annihilations was determined to occur at 67.4(5) µs (courtesy of Saiva Huck).

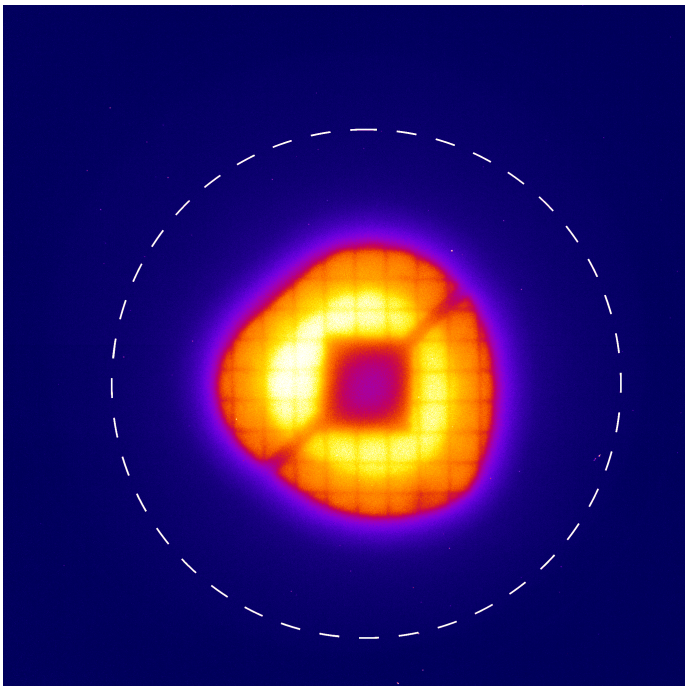


Figure 4.14 – The shadow cast by the target when imaged with antiprotons, after the careful alignment performed by biasing the B0 sectorised electrode. The dashed line represents the trap's inner diameter.

$$v(t) = -\frac{l\pi}{2\tau}\sin\left(\pi\frac{t}{\tau}\right) \tag{4.2.3}$$

where $l \approx 0.93\,\mathrm{m}$ is the distance between the centre of the launch trap (see *Antiprotons transfer* (4.2.4)) and the target and $\tau \approx 11\,\mu\mathrm{s}$ is the travel time. With this, we have:

$$v(\tau \pm 500\text{ ns}) \approx \pm 2 \cdot 10^4 \frac{\mathrm{m}}{\mathrm{s}} \tag{4.2.4}$$

This velocity is negligible with respect to the one of the Ps atoms emitted by the target at room temperature ($O(10^5)$), and should therefore influence negligibly the formation of antihydrogen.

This velocity is one order of magnitude greater than the thermal average velocity of an antiproton at ∼100-300 K:

$$v_{th|100\,\mathrm{K}-300\,\mathrm{K}} = \sqrt{\frac{3k_BT}{m}} \approx 1570 - 2700\mathrm{m/s} \tag{4.2.5}$$

In this manner, a beam of antihydrogen with an aperture smaller than 8° (4° at 100 K) can be formed.

In the case of colder antihydrogen, a boosted beam can be produced also at slower velocities. In fact, $v_{th|10\,\mathrm{K}-30\,\mathrm{K}} \approx 500 - 870\mathrm{m/s}$: therefore, a three-times boosted beam can be formed leaving to the antiprotons a velocity of 1500-2000 m/s (at the moment of $\bar{\mathrm{H}}$ formation), obtainable with an initial launch delay of 40-53 ns.

To gain evidence of the antihydrogen formation, both the scintillators and the MCP have been leveraged: the results of the preliminary analyses carried out are presented in detail in *Antihydrogen formation* (5.3).

Chapter 5

Results

Although the last years were marked by the transition, for AEgIS, from Phase 1 to Phase 2, and therefore much efforts had been put into the upgrade and consolidation of the apparatus and the related operations (as extensively presented in the previous chapters), a great focus was always held towards the preliminary and enabling technical and physics results that we wanted to obtain during this endeavour.

In this chapter, the most significant technical and physics results originating from the antihydrogen research line are described (other two very significant physics results, i.e. the world-first laser cooling of positronium atoms, and the formation of cold trapped highly-charged ions, are not presented here, but they can be found in *Additional results* (A)). The first section shows the outcomes of the performance investigations of CIRCUS, which span from the control system stability and safety to its synchronisation capabilities, arriving at tests and demonstrations of its automation features. Afterwards, the measurements performed to estimate the efficiency of the capture of antiprotons are outlined, and utilised to determine the record of trapped antiprotons per ELENA shot in the entire Antimatter Factory. Finally, in the last section, the most important physics result is presented, the formation of neutral antihydrogen in the new apparatus. Three different analyses are shown, to assess both the formation of $\bar{H}$ and the formation of a boosted beam (forward and backwards): one is performed using the scintillator data, while the other two look at the images captured with the MCP at the end of the trap system.

5.1 CIRCUS performances

The flexibility and modularity allowed the system to be operative in a very short time. The first version of CIRCUS was deployed in 2021 and, during its first $\bar{p}$ campaign, we managed to successfully trap antiprotons in the experiment with it in less than a week, several times faster than with the previous control system. The possibility of running automated scans overnight, while debugging and development took place in the daytime, was exploited immediately, to explore the effect of the trap closure time on the trapping efficiency, and to characterise the energy of the antiprotons thus captured. A scan in steps of 50 ns assessed that the optimal trap closure time was between 800 ns and 1150 ns, in accordance with what estimated in equation 4.2.1. The result of the scan of the number of trapped $\bar{p}$ versus the capture electrode voltage is visible in figure 5.1: the ratio is in good accordance with our GEANT4 simulation (see figure 2.14), considering that for this test, a 1500 nm Parylene N total thickness was employed, which has the equivalent degrading capabilities of the 1400 nm of Mylar installed afterwards.

The following year, CIRCUS demonstrated the reproducibility of its procedures: in fact, the capture of the antiprotons was achieved on the first day of beam taken in the 2022 antiproton campaign, by simply running the `PbarCatchNDump.py` procedure developed the year before. After that, the trap-closing-time scan was re-executed with a much smaller step size (7 ns), and it improved

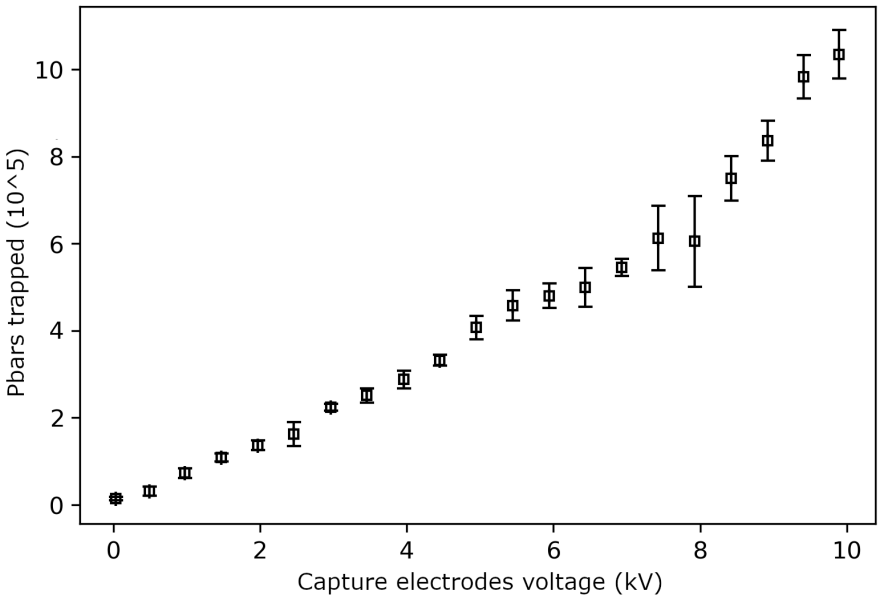


Figure 5.1 – The graph shows the number of antiprotons captured varying the potential of the catching electrodes. Each point corresponds to a different ELENA shot.

the results of 2021, narrowing down the best closure timing to few tens of ns (see figure 5.2). The implementation of the *ELENA Interface* (see 3.4.6.2) enabled both further increases in the trapping efficiency, by automatically scanning over the beam parameters to find the best ones (see figure 5.3), and a higher stability and uptime, by making TALOS react to external events like *no beam*, *valve closed*, *empty shot* and *beam stopper in* (see the following section *Control system stability and error handling* (5.1.1)). All these improvements have contributed toward achieving the record of $\bar{p}$ trapping efficiency (see *Efficient antiprotons capture* (5.2)).

In 2023, CIRCUS was extended with the focus shifted towards a great level of automation and multiple Kaslis support. The first resulted in the possibility of live parameters stabilisation (see *In-Run autonomous parameter stabilisation* (5.1.4)) and, thanks to the integration of ALPACA, it enabled autonomous parameters optimisation based on data feedback driven by a machine-learning algorithm (see *Automatic parameter optimisation* (3.4.7.3)). The multi-Kaslis support was instrumental in managing the series of operations necessary to achieve antihydrogen production, culminating in the coding of the *Particle Server* (see *Particle Server* (3.2.5)). Some stress-tests of the multi-Kaslis information flow and synchronisation were performed, and their results are available in *Kaslis synchronisation* (5.1.3).

In the following sections, the main CIRCUS performance tests are presented, each accompanied by the corresponding results.

5.1.1 Control system stability and error handling

The TALOS framework has demonstrated very high stability, running since its first deployment in 2021 almost continuously (at least idle), the only moments offline during system upgrades. A few reboots were needed during the first year due to the control system entering unrecoverable error states. With the consolidation of the code, such reboots have become less and less necessary: only two were needed in 2023.

The control system uptime was constantly monitored, as a direct benchmark for stability. During the AE$\bar{g}$IS 2022 antiprotons campaign (spanning 35 days, from October 12th to November 16th) the control system conducted measurements for 552.3 h (almost 22 days), equivalent to $\sim$ 62 % of the total time, which corresponds to most of the nights and weekends of the period: in fact,

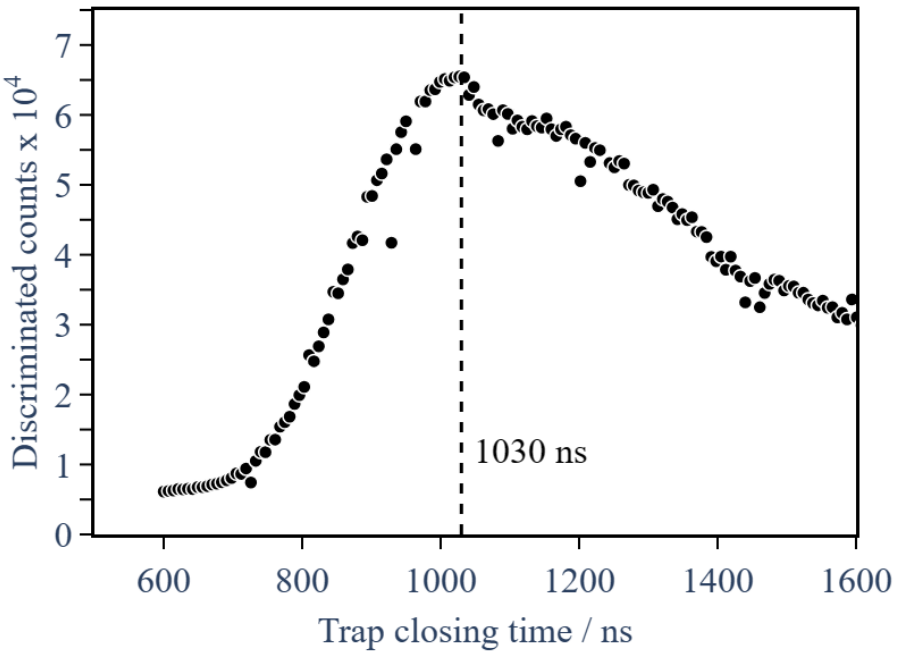


Figure 5.2 – A trap-closing-time scan, showing the number of antiprotons captured versus the closure timing of the trap: it clearly shows the presence of the best working point. Closing too fast lets some $\bar{p}$ out, and, conversely, closing too slow lets some $\bar{p}$ escape after the bounce on the second electrode.

daytime was mainly devoted to development. Throughout this period, the system faced various situations that prevented measurements due to external factors, like conditions identified by the *ELENA Interface* (e.g. "Beam Stopper In", "Valve Closed" and "No triggers"), differences in run conditions and data rates causing congestion and variations in the time needed by the DAQ to sync data to disk within a fixed timeout, or other minor hardware or software-related exceptions, with *Retry* as the associated action.

In table 5.1, a summary of these exceptions is displayed, totalling approximately 66 h 27 min, which accounts for 12.7 % of the measurements' total time and to 7.9 % of the entire antiproton campaign. It is important to note that the total time calculation does not simply add up individual exception times, since multiple errors could occur simultaneously during a single script execution (e.g., during an intervention in the zone, both "Beam Stopper In", "Valve Closed" and "No triggers" are likely thrown).

This capability streamlines both data collection and data analysis. Without the ability to react to these exceptions, identifying and manually re-acquiring affected Runs would have imposed a significant overhead on the scientists. Therefore, the CIRCUS de-facto "saved" the experiments this amount of time, by avoiding them the manual labour (the time to re-take the points has to be invested in any case).

In the summer of 2023, AE$\bar{g}$IS performed its first experimental campaign on Highly-Charged Ions (HCI). Antiprotons were trapped as usual, and the interaction with nitrogen injected into the chamber was probed. During $\bar{p}$ collisional cooling with the gas, HCI are formed either by collisional ionisation or by antiproton's capture, which leads to a cascade of electrons emitted while it falls on the nucleus. In this experimental campaign (lasted 32 days, from July 22nd to August 23rd), CIRCUS acquired data for 516.7 h (almost 22 days), equivalent to $\sim$ 67 % of the total time; by correctly handling the exceptions presented in the paragraph above, it saved approximately 161 h 21 min (i.e. 31.2 % of the measurements' total time and to $\sim$ 21 % of the entire campaign). A review of the exception handling is presented in table 5.2. Preliminary results from

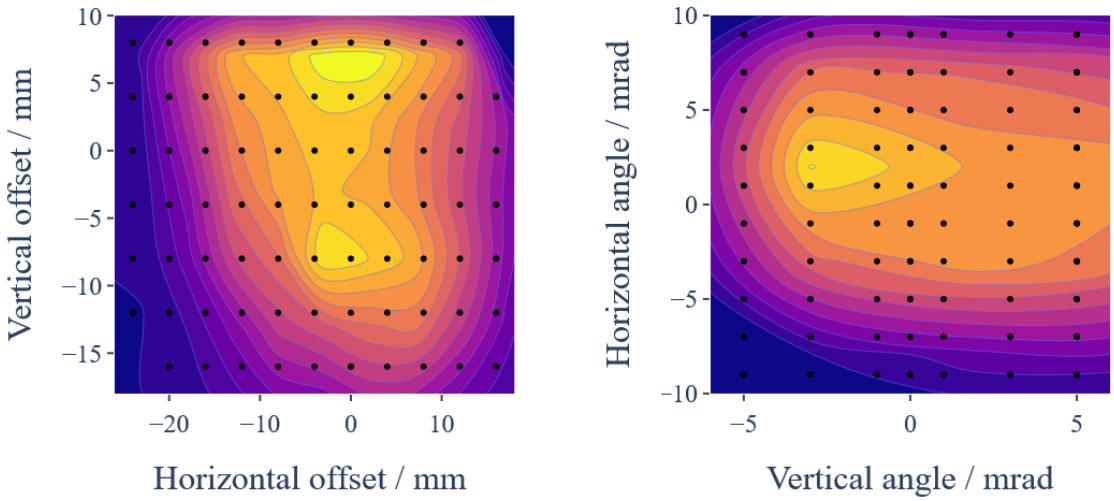


Figure 5.3 – The two figures show the results of the scan over the horizontal and vertical displacement of the antiproton beam (on the left) and over the horizontal and vertical angle (on the right). The colour represents the number of trapped antiprotons (in arbitrary units, warmer is higher). The parameter space has been organised in this way, assuming that displacements and angles have independent effects, not for physics reasons, but because scanning over the full parameter space would have been impossible time-wise (10 steps per dimension ˆ 4 dimensions × 5 minutes of duration of the script ≈ 35 days!).

the HCI campaign are given in *Formation of trapped cold Highly Charged Ions* (A.2).

In the autumn of 2023 (from September 4th to November 11th), instead, the formation of antihydrogen was tempted, the first time since the end of Phase 1. During this measurement campaign (69 days), data was acquired for a total of 894.2 h (more than 37 days), corresponding to ∼ 54 % of the entire duration. Also this time, CIRCUS saved us ∼ 283 h 30 min (equivalent to 31.7 % of the measurements' time and to 17.1 % of the whole campaign). A summary of the errors encountered is given in table 5.3.

Furthermore, during the 2023 HCI campaign, CIRCUS has been put to a new level of stress. The novel and exploratory nature of these experiments resulted in a continued modification of the ARTIQ script(s) managing the measurements, but due to hardware development and ELENA unavailabilities, little to no script testing was possible during the daytime. Therefore, during the night time (which was dedicated to data taking), often the control system stopped to an idle state because of exceptions thrown from Kasli, which were treated equally from TALOS at the same

	Beam Stopper In	Valve closed	Empty shot	No triggers	DAQ	Total
Events	330	172	27	251	864	1644
Blocks	66	28	19	48	148	244
Total duration	20 h 31'	13 h 3'	2 h 56'	27 h 39'	24 h 40'	66 h 27'

Table 5.1 – Details of the handling of the most frequent exceptions in the AEḡIS experiment during the antiproton campaign of 2022. With "Events" we mean the number of occurrences, while "Blocks" is the number of groups of contiguous Runs where the error keeps on appearing. The "Total duration" is the total script time that was invalidated by the corresponding exception. The column "Total" is not simply the sum of all the previous columns: in fact, exceptions that were thrown during the same Run are cumulated and counted as one. This is especially necessary to correctly evaluate the total time.

	Beam Stopper In	Valve closed	Empty shot	No triggers	DAQ	Total
Events	105	189	421	299	1601	2615
Blocks	42	39	83	48	129	242
Total duration	14 h 5’	27 h 26’	62 h 11’	50 h 37’	30 h 40’	161 h 21’

Table 5.2 – Summary of the most thrown exceptions during the Highly-Charged Ions (HCI) campaign of 2023. The meaning of the various terms is the same as in table 5.1.

	Beam Stopper In	Valve closed	Empty shot	No triggers	DAQ	Total
Events	134	508	492	146	16477	17757
Blocks	48	46	96	25	159	252
Total duration	18 h 17’	66 h 28’	46 h 32’	35 h 37’	117 h 10’	283 h 30’

Table 5.3 – A review of the errors encountered during AEḡIS antihydrogen campaign of 2023. The meaning of the various terms is the same as in table 5.1.

extent as unhandled exceptions. Therefore, it was decided to implement also in the ARTIQ libraries the same error conversion structure ideated for TALOS (see *Error handling inside the experiment scripts* (3.2.3)): dedicated exceptions where defined, with a specifiable *Criticality Code*, and the entire experimental routine was wrapped into a *try...except...finally...* clause. The *finally...* section guarantees the execution of the script termination routine, which could therefore always give to TALOS the proper *Criticality Code* for the exception raised. With this structure in place, the number of stopping exceptions rapidly declined to zero, guaranteeing full nights of acquisition even in place of bugs in the procedures, caused by the little testing time available.

A good indicator of the stability of the system is also given by the total number of antiproton bunches taken by AEḡIS over the course of the 2023 antiproton campaign (comprising both the HCI and the antihydrogen experiments). In figure 5.4 the integrated number of antiproton shots delivered to the various experiments in the Antimatter Factory during 2023 is shown: it can be seen that AEḡIS was the experiment utilising the highest number of them (totalling to approximately 64 000 shots).

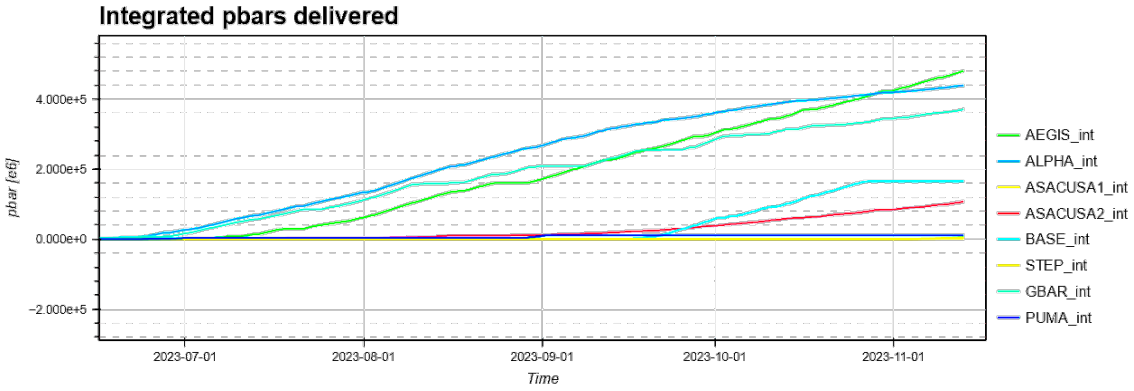


Figure 5.4 – Integrated number of shots delivered by ELENA to the various experiments in the Antimatter Factory during the 2023 antiproton campaign. It can be seen that AEḡIS was the experiment utilising the highest number of bunches (courtesy of CERN).

5.1.2 Safety

It is important to highlight that the ABORT system, running quietly in the background and typically unnoticed (except when a Guardian is restarted, at which point the ABORT LED on all other Guardians lights up), once prevented potential hardware damage during an unsupervised nightly data taking. On that occasion, the high-voltage power supply connected to the MCP began to fail while biasing the MCP's front face at 2800 V. The corresponding µService reported a lost connection error, and TALOS raised ABORT (being a potentially hardware-damaging error, it was given *Criticality Code* 4, the highest). During the transition to *Safe Mode*, the system shut off all the high-voltage power supply lines. On the same night, a minor vacuum incident occurred, which could have caused damage to the MCP if it had remained continuously powered on.

This event demonstrated the importance of having a distributed system capable of reacting to hardware exceptions in full autonomy, enhancing and ensuring the safety (and self-preservation) of the experiment.

5.1.3 Kaslis synchronisation

A dedicated series of measurements have been carried out to test the performance of the multi-Kaslis parallel operations described in *Automation flow* (3.4.7.2). To execute the test, a simple script was defined, consisting of a waiting routine, and a total of 5 Kaslis[1] have been used, so to put the system under stress. For each parallel mode (asynchronous and synchronous), two series of 50 Runs of the aforementioned script were executed. The first series was done with the same wait time of 2 s for all the 5 Kaslis used; in the second series, the wait time increased by 2 s between Kaslis, so from 0 s for the first Kasli to 8 s for the fifth Kasli. The order of the Kaslis is defined by the one in the synchronisation mask (see *The Scheduler* (3.4.5.2)). To make considerations about this analysis easier, we define *neighbouring Kaslis* two Kaslis adjacent in the synchronisation mask (e.g. the first and the second, or the third and the fourth).

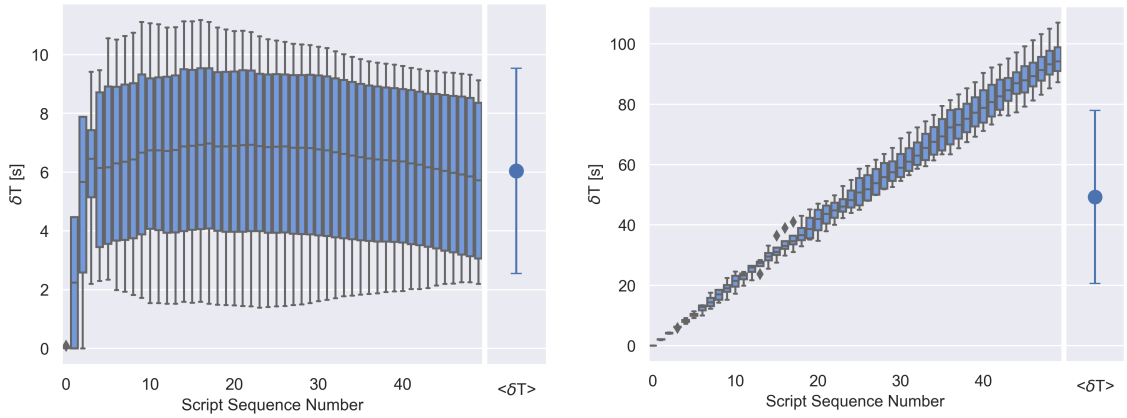
To quantify the results, a measure, denoted as δT, was defined: it is the difference, in seconds, between start times of the n-th script of a schedule of two neighbouring Kaslis. This is a good observable to establish the (a)synchronicity of the parallel operations since the synchronisation mask determines the order used by the *Tamer* to send collective messages to the Kaslis during the synchronous operation (see *Automation flow* (3.4.7.2)). Because of this, in the synchronous operation mode, the difference between the starting times of two Kaslis in consequent positions has to be smaller than the time between one of them and whatever other Kaslis.

In figures 5.5 and 5.6 the calculated values of δT at different series iterations are presented as a box plot. Each point shows the range of all values, with the top and bottom of the box corresponding to the 75 and 25 percentile of measured values. Each plot is accompanied by the calculated average of the δT, with error bars calculated as standard deviations.

Noticeable trends in asynchronous operation are evident, as values of δT tend to increase for the same-duration test (figure 5.5a) until reaching a plateau. This behaviour can be attributed to the definition of the measure: in fact, in the same-duration asynchronous case, well-defined neighbours among Kaslis only exist at the start of the schedule. Subsequently, each *Monkey* independently manages its respective Kasli, so the concept of neighbours loses significance. These effects, however, are perfectly normal and demonstrate the asynchronicity of the operations.

In the asynchronous operation, with different durations of the scripts (figure 5.5b), the trend is a linear increase with a slope of around 2 s per iteration. This is exactly what would be expected, since in this case the neighbour of a given Kasli never changes. In fact, the BANANA message always

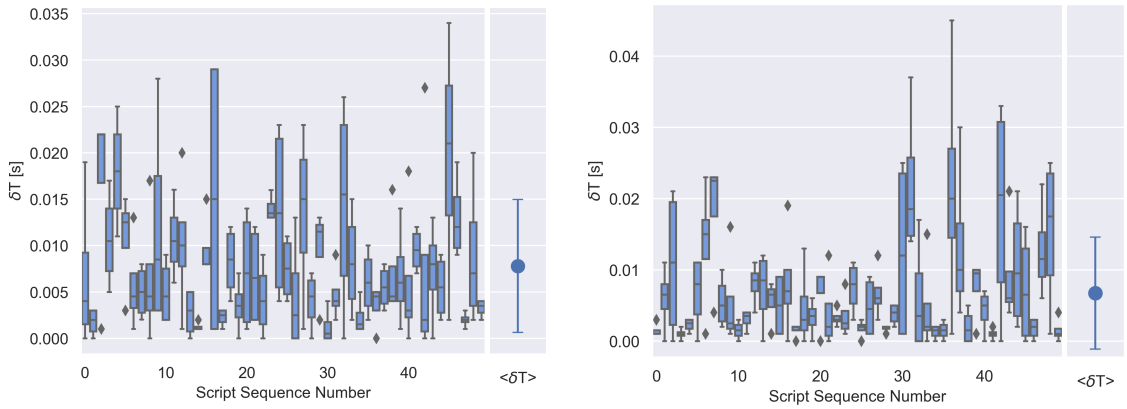
[1] Actually, lacking the necessary hardware, we coded a little program to simulate the behaviour of a simple Kasli module: with this, we were also able to test the control system – together with the simulated hardware – on computers different from the experiment ones, even during data taking campaigns.



(a) Same duration, asynchronous

(b) Different duration, asynchronous

Figure 5.5 – Synchronisation results in the asynchronous parallel operation mode. In general, it can be seen that the time difference of the start of the n-th script between two neighbouring Kaslis is generally not null. In the same-duration case (left), the randomisation of the Kaslis order makes δT lose its original meaning, so the value averages to a plateau. In the different-duration case (right), δT increases linearly by 2 s per iteration, corresponding to the accumulated delay due to the parallel execution.



(a) Same duration, synchronous

(b) Different duration, synchronous

Figure 5.6 – Synchronisation results in the synchronous parallel operation mode. Both in the same-duration case (left) and in the different-duration case (right), the time difference between the start of different parallel scripts is stable around 10 ms, independently of the duration of the scripts.

arrives later for Kaslis further in the synchronisation mask:

$$t_{\texttt{BANANA}}(n) = 2s \cdot n \cdot position_{Kasli}. \tag{5.1.1}$$

So, for the first Kasli the end of the scripts (and therefore, the start of the subsequent one) happens approximately at 2 s, 4 s, 6 s, etc., while for the second one occurs circa at 4 s, 8 s, 12 s, etc. Hence the neighbour stability and the linear δT increase.

On the contrary, when looking at the synchronous operation results, it is clear that there is no trend whatsoever. In both the cases of same-duration script (figure 5.6a) and different-duration script (figure 5.6b), the average δT is around 10 ms. This measurement clearly shows the synchronisation of the multiple scripts running in parallel and gives a precise value of the jitter to be expected between the start of different scripts on parallel Kaslis.

5.1.4 In-Run autonomous parameter stabilisation

The combination of the new CIRCUS control system and the Sinara programming yields another desirable feature: decision-making based on feedback loops. In fact, complex systems typically depend on a multitude of parameters, of which not all are directly controllable. A good study case to demonstrate the capabilities introduced by CIRCUS to improve the control of such systems is the stabilisation of the pulse timing of the "Alex" laser (see *Alexandrite Laser ("Alex")* (2.1.6.2)).

The "Alex" laser system, used for positronium excitation to $n = 2$ state, displays a strong correlation between ambient humidity and the resulting generation instant of the light pulse. The humidity in the environment, on the other hand, is coupled to the temperature, which in turn affects the output laser energy. Since the current "climate control" system of the *Lighthouse* can either stabilise the humidity or the temperature, the other needs to be allowed to run freely. The nanosecond-precise control system opens up the opportunity to tune the timing of the laser pulse by means of triggering a Pockels cell at the right moment, whereas the energy of the laser cannot be adjusted that easily. Thus, the temperature (and consequentially the energy of the laser pulse) is chosen to be controlled by the climate system, while the humidity is left to run freely. In turn, the time drift caused by the humidity variation is compensated by the control system via a feedback loop, which is detailed below.

A few seconds before the actual positronium production occurs, a test laser pulse is produced by triggering the Pockels cell and the data acquisition chain. The generation instant of this pulse, depending on the environmental conditions, may vary with respect to the moment the Pockels cell is triggered, for example, because of humidity drift over time. The acquired spectrum of a photodiode is immediately stored by the DAQ system and analysed by a dedicated function in the experimental script, which extracts the arrival time of the test laser pulse: its value is then compared to the user-defined one, and a correction term is calculated. Imminent to positron implantation into the converter target, the Pockels cell is triggered again for the actually used pulse, applying the correction term obtained from the test pulse to account for the temporal offset. As a result, the synchronisation is now sufficiently precise to guarantee an overlap of the laser pulse and the positronium cloud, independently of the origin of the drift: this can be seen in figure 5.7, where the timings of the test laser pulses (red squares) and the desired laser pulses (blue circles) are plotted for a series of experimental trials executed over the course of one hour (with some interruptions). The user-defined value is given as the horizontal line. The statistical errors in the determination of the timings are of the order of a few hundred picoseconds and thus are not visible in the plot.

The capability of correctly firing the "Alex" laser at the exact chosen moment has been fundamental to achieving the world-first laser cooling of positronium atoms, which necessitated a perfect time overlap between laser emission and positrons implantation on the target (see *Positronium laser cooling* (A.1)).

This active feedback loop, exemplified by the timing of a laser pulse, is versatile and can be applied to any parameter of any part of the system, given that there is enough time to obtain the test data and analyse it before the real experiment occurs. In this manner, CIRCUS is de-facto self-stabilising, being able to autonomously tune parameters during a Run to obtain consistently optimal results.

5.1.5 Autonomous parameters optimisation

As explained in *Automatic parameter optimisation* (3.4.7.3), the integration of ALPACA (see *ALPACA* (3.3)) in TALOS has rendered the control system able to autonomously search the best parameter setpoint to optimise a given observable. A good example is the optimization of the beam steering to maximise the number of antiprotons trapped, a task which has to be performed repeatedly during every antiproton beam time (which automatisation was enabled by the introduction of the *ELENA Interface* (3.4.6.2)).

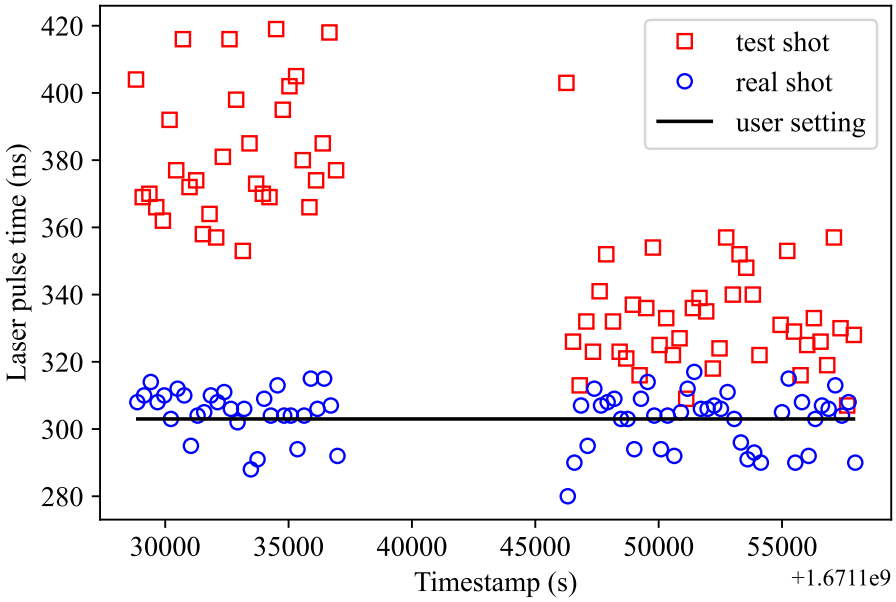


Figure 5.7 – A feedback loop uses the uncorrected laser pulse timings (red squares) to calculate the deviation from the user setting (solid black line) over an hour and corrects the timing of the subsequent desired laser pulse that is used for the actual experiment (blue circles). Independent of short-term to long-term drifts or even sudden jumps, the resulting timing is always close to the desired value.

In this example, the optimisation variables' parameter space is constituted by the vertical and horizontal offsets and angles of the incoming antiproton bunch relative to the axis of the injection line, expressed as real numbers. The number of trapped antiprotons is proportional to the number of annihilation events detected by the scintillator detectors upon a "dump" of the content of the trap towards either the MCP or the main degrader.

As the ELENA power supplies responsible for setting the voltages of the beam steering electrodes only allow for certain combinations of ranges, the parameter space is bound to unknown correlations among the steering parameters. In fact, the setting of a precise value of a parameter may constrain the ranges of the other parameters still to be set. When given an invalid combination of parameters, the power supplies set the closest combination of admitted values: thus, in order to perform a scan over a valid parameter space, one has to explore the boundaries of the parameter space beforehand. The Bayesian optimiser is agnostic of these correlations, and only processes valid evaluation of the parameter space; therefore, we chose to set the parameter space with the full range for all the parameters, thus allowing for valid and invalid combinations. In case an invalid combination is requested by the Bayesian optimiser, the corresponding result is anyway saved together with the actual parameter values set by the power supplies, so no false point is taken.

Figure 5.8a shows the largest number of observed annihilation events over the course of 101 consecutive measurements with the parameters suggested by the Bayesian optimizer provided by ALPACA. Highlighted in yellow are the initial 30 runs used to randomly explore the parameter space. For benchmark purposes, we defined a reference convergence criterion, which is evaluated as soon as the number of conducted experiments exceeds the initial exploration experiments:

$$\left| \frac{\sigma_{best10}}{\mu_{best10}} \right| < \delta. \tag{5.1.2}$$

Here, μ and σ reflect the mean and standard deviation of $X_{best10} = \{x_0, x_1, ...x_9\}$ (with

$x_k \in X = \{x_1, x_2, ..., x_n\}$, and n being the number of conducted experiments), while δ has to be chosen by the experimenters.

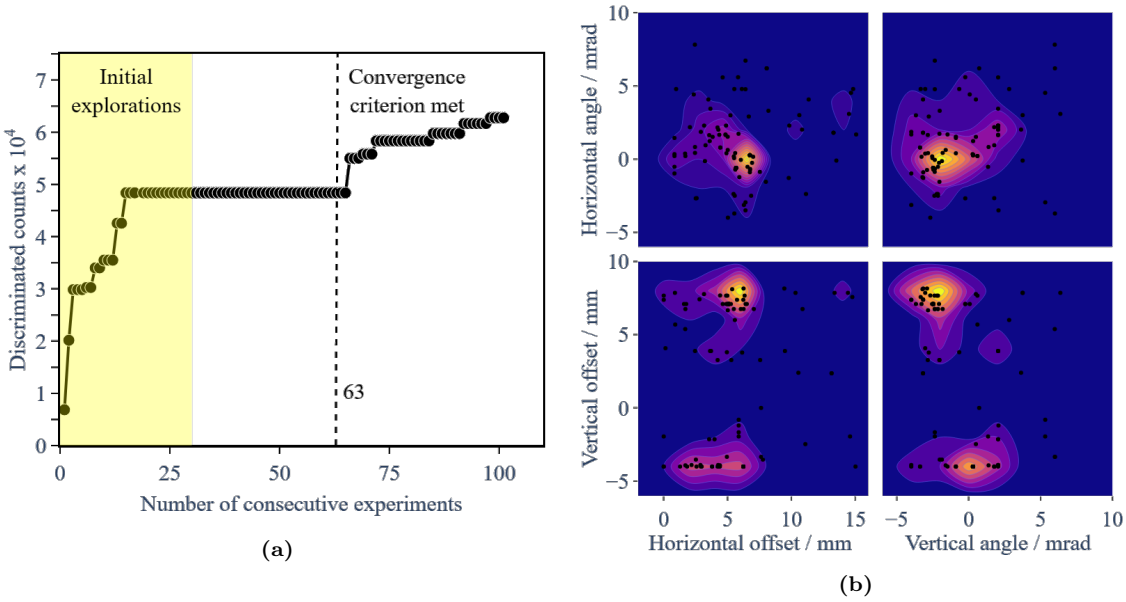


Figure 5.8 – Results of the beam steering optimisation: on the left, the convergence plot of the largest observed count of annihilation events; on the right, the density plots of the evaluations in the four-parameter space dimensions. Each black dot corresponds to a measurement point.

The convergence criterion was pre-defined with $\delta = 0.05$ and is met after 63 Runs. Given the relative improvement of additional ~ 23 % over the course of next 38 Runs, we find that either the convergence criterion (5.1.2) or the value chosen for δ could have been stricter. On the other hand, setting a precision threshold below 5 % could have caused instabilities, since the number of antiprotons provided by ELENA can vary by up to 10 % among different cycles.

The highest value over the course of the 101 Runs was observed for the parameters shown in table 5.4, which matches with the density of evaluations in the parameter space shown in figure 5.8b. This result is in excellent agreement with the results from the previous year (see figure 5.3). The number of Runs that were required to reach a result comparable to the manual scan performed in 2022 shows a speed increase of about 146 %. The performance increase is much higher, since the Bayesian optimiser was operated without restrictions on the parameter space, eliminating the previous need to conduct scans in multiple sub-spaces.

Parameter	Best value
Horizontal offset	6.23 mm
Vertical offset	7.38 mm
Horizontal angle	−0.90 mrad
Vertical angle	−3.17 mrad

Table 5.4 – Values of the beam steering parameters that maximise the number of antiprotons trapped in the *catching* trap. These values were autonomously found by the CIRCUS control system thanks to the Bayesian optimiser provided by ALPACA.

As explained in *Antiprotons capture* (4.2.1), the capture of the $\bar{p}$ in AE$\bar{g}$IS is performed thanks to two high-voltage electrodes (HV1 and HV3), the former being connected to a fast switch, to be able to raise its potential to 15 kV in few nanoseconds. The timing of raising the voltage on HV1 is crucial: in fact, if performed too early, it can block the entrance to (some) antiprotons, and if performed too late, the antiprotons escape the trap after being reflected by HV3.

In figure 5.9b the relationship between the trap closing time and the number of distinct annihilation events after the dump of the trapped antiprotons is depicted. The evolution of the best-observed setting shown in figure 5.9a indicates that the initial random exploration of the one-dimensional parameter space already got almost the best setting. The convergence criterion (5.1.2), using – this time – a more rigorous threshold of $\delta = 0.02$, was met after 31 experiments, which is in agreement with the scan performed in 2022 with a total of 140 Runs: it corresponds to a speed improvement of ~ 450 %.

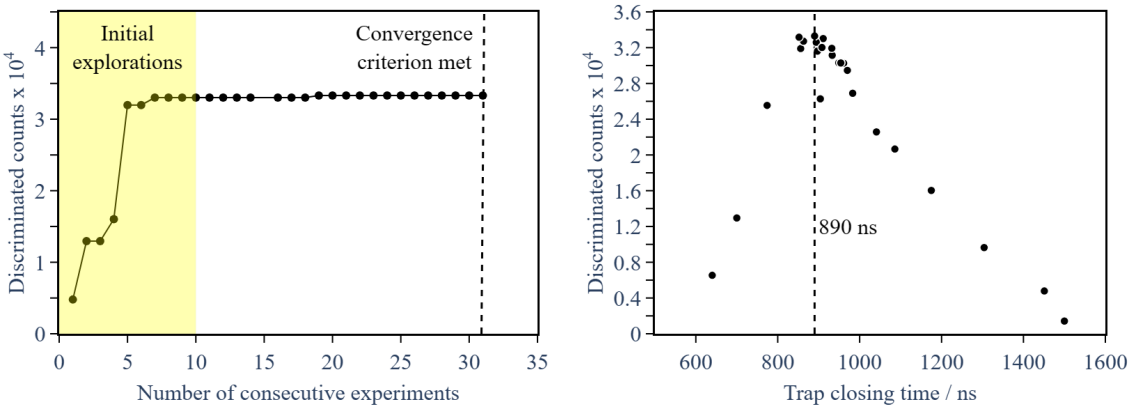


Figure 5.9 – Results of the trap closing time optimization: on the left, the convergence plot of the best-observed count of annihilation events; on the right, the number of observed annihilation events as a function of the trap closing time. The latter result is different from the one in 2022 (figure 5.2) because the capture electrodes voltage was raised from 10 kV to 15 kV.

5.2 Efficient antiprotons capture

AEgIS requires a high antihydrogen flux to obtain the necessary statistic to precisely determine the behaviour of antimatter in Earth's gravitational field: it is then obvious that maximising the capability of capturing the antiprotons furnished by ELENA is extremely important.

To go in this direction, as previously mentioned, a careful study has been carried out to design the new degrading system (see *Degrader(s)* (2.1.7)). The simulation further motivated an upgrade to the high-voltage system controlling the HV electrodes in the 5 T trap: as it can be seen in figure 2.14, by going from the 9 kV available in Phase 1 to 14 kV, the $\bar{p}$ stopping capability would be boosted from ~ 30 % to ~ 80 %, reaching almost 100 % at 20 kV. This work was accomplished in 2022, bringing the maximum voltage deliverable to the HV electrodes to 15 kV.

Afterwards, to test the combination of the new degrader coupled to the improved HV electrodes, a "barrier scan" was performed: the voltage on the electrode HV3[2] was varied, and the corresponding fraction of antiprotons not stopped – and hence, hitting the MCP – was measured by the charge deposited on the MCP itself and by the scintillators. In figure 5.10 the result of this scan is displayed: it can be seen that the new system guarantees a stopping capability of ~ 80 %.

Subsequently, during the antiproton campaign of 2022, a detailed study of the trapping capabilities of the upgraded system was performed [6], in order to evaluate the captured fraction of $\bar{p}$. Two separate techniques were employed: the first entailed the measurement of all the losses relative to the trapped amount; the second, instead, consisted of determining the absolute number of antiprotons by accurately counting the annihilations on the scintillators and comparing it to the amount of $\bar{p}$ assessed by the ELENA team.

[2]Similar scans have been performed also with HV1 and HV2, both to test their correct functioning, and to cross-verify the measurements: all the three electrodes gave identical results.

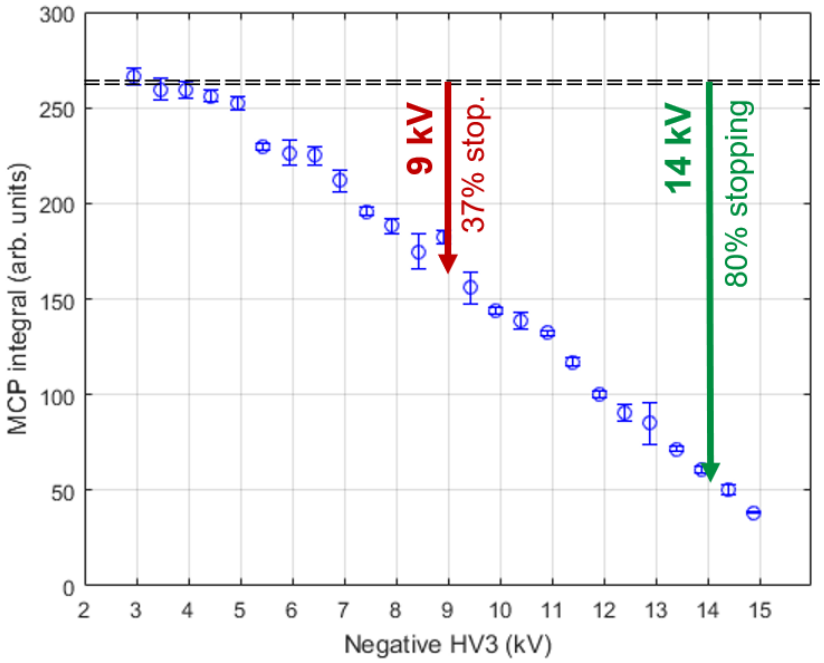


Figure 5.10 – MCP signal intergral versus the voltage on HV3. The characterisation was performed by varying the voltage of the stopping electrode HV3 and measuring the antiprotons, emerging from AEgIS main degrader, that were hitting the MCP by counting them with the MCP itself. The enhancement given by the improvement of the HV capabilities is highlighted (from [6]).

5.2.1 Determination via losses estimation

For the first evaluation, some considerations on the various loss sources need to be carried out. In the beginning, the $\bar{p}$ bunch form ELENA passes through the degrader, where a fraction of the particles annihilate (f_{deg}), while most of them traverse it moderated[3] (f_{mod}). Most of the moderated particles get reflected back from the high-voltage electrode HV3 ($f_{R|HV3}$), but again a small part cannot be stopped and annihilates at the end of the traps system, on the MCP (f_{MCP}). The optimised, fast raising of HV1 (see *Autonomous parameters optimisation* (5.1.5)) maximises the percentage of $f_{R|HV3}$ being reflected again towards the centre of the trap – and thus trapped ($f^{(1)}_{trap}$) –, but cannot avoid that a share passes through and ultimately annihilates on the degrader (f_{deg2}). Furthermore, likely due to collisions inside the trap, causing radial annihilations, a non-negligible portion of the trapped antiprotons is lost in the very few seconds of storage (f_{store}).

A pictorial schematisation of the process outlined is presented in figure 5.11, and in formulas 5.2.1 the relations among the various fractions just presented are summarised (the incoming bunch is put as "1", since all the numbers are relative to it).

$$\begin{aligned} 1 &= f_{deg1} + f_{mod} \\ f_{mod} &= f_{MCP} + f_{refl} \\ f_{refl} &= f_{deg2} + f_{trap} \end{aligned} \tag{5.2.1}$$

A sequence of measurements was therefore designed to isolate and quantify each loss mechanism. The loss fractions f_{deg1} and f_{deg2} were determined by observing $\bar{p}$ annihilations on the degrader, as detected by the first scintillator detector. It was possible to distinguish the two different contributions thanks to their time separation, given by the travel time of antiprotons (back and

[3]The amount of $\bar{p}$ back-scattered from the degrader is perfectly negligible.

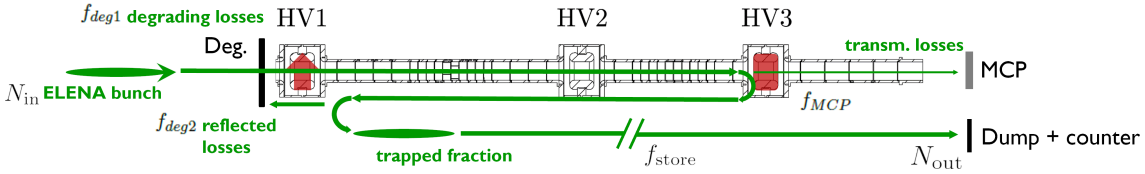


Figure 5.11 – A scheme of the various fractions of antiprotons (losses and trapped) used to estimate the total number of capture $\bar{p}$. f_{mod} is the fraction lost during the moderation in the degrader; $f_{R|HV3}$ is the part reflected back by HV3, while f_{MCP} are the $\bar{p}$ annihilating on the MCP; on the first collision with HV1, f_{deg2} traverse it and dies on the degrader, while $f^{(1)}_{trap}$ is finally securely captured and stored in the trap. Nevertheless, in the first seconds of storage, due to internal collisions, f_{st} is lost in annihilations.

forth) inside the trap ($\sim 1.1\,\mu s$). The normalization of the signals with respect to the total beam intensity was critical, and it was performed by directing the entire beam to annihilate on the MCP and measuring on PMT2 the annihilations occurring from the impact on the degrader and on the MCP itself. Knowing the charge deposition ratio of 1:71.4 (calculated as the ratio of the solid angles[4]), the fraction annihilated on the degrader was derived. Conversely, f_{MCP} was derived from the ratio of $\bar{p}$ annihilations on the MCP, with and without a 14 kV stopping barrier (from the measurement depicted in figure 5.10).

The values found are listed in table 5.5. Notably, the fraction of $\bar{p}$ annihilations in the main degrader aligned well with GEANT4 Monte-Carlo predictions (shown in *Degrader(s)* (2.1.7)).

Fraction	Amount
f_{deg1}	$\approx 0.5 \div 1.5\%$
f_{MCP}	$\approx 16 \div 21\%$
f_{deg2}	$\approx 4 \div 6\%$

Table 5.5 – Estimated values of the various $\bar{p}$ loss fractions occurring during the capture operation.

With these numbers, the total fraction of antiprotons captured was calculated to be:

$$f^{(1)}_{trap} = 1 - f_{\mathrm{deg1}} - f_{\mathrm{MCP}} - f_{\mathrm{deg2}} \approx 71 \div 79\% \tag{5.2.2}$$

5.2.2 Determination via absolute number estimation

The second method involved counting directly the number of antiprotons stored in the trap after the trapping was terminated. This consisted of slowly releasing the $\bar{p}$ towards the degrader by progressively reducing the potential on HV1, to let them annihilate on the degrader (a procedure that we normally call *hot dump*, because performed without – or simply before – cooling the antiprotons with the electrons): the annihilations are then computed from the readings of the various PMTs. The release is performed slowly, spreading the annihilation over the course of several seconds, so to avoid saturation of the PMTs. In figure 5.12 is shown one of such signals, taken from the scintillator SC56.

The transformation from the number of scintillating events to the number of annihilations is based upon a precise calibration [103], which has been verified in 2022 specifically for this purpose. First, all the PMTs gains are tuned to uniform their reading, confronting the cosmic rays spectrum detected by means of a charge digitiser (QDC). Then, a GEANT4 simulation is used to calibrate the effective relative detection acceptance of each scintillator with respect to an annihilation taking place in the same point of the apparatus (e.g., the degrader), taking into account the geometry

[4] For a better estimation, a Montecarlo simulation should be performed, to keep into account the different materials thicknesses to be traversed by the particles.

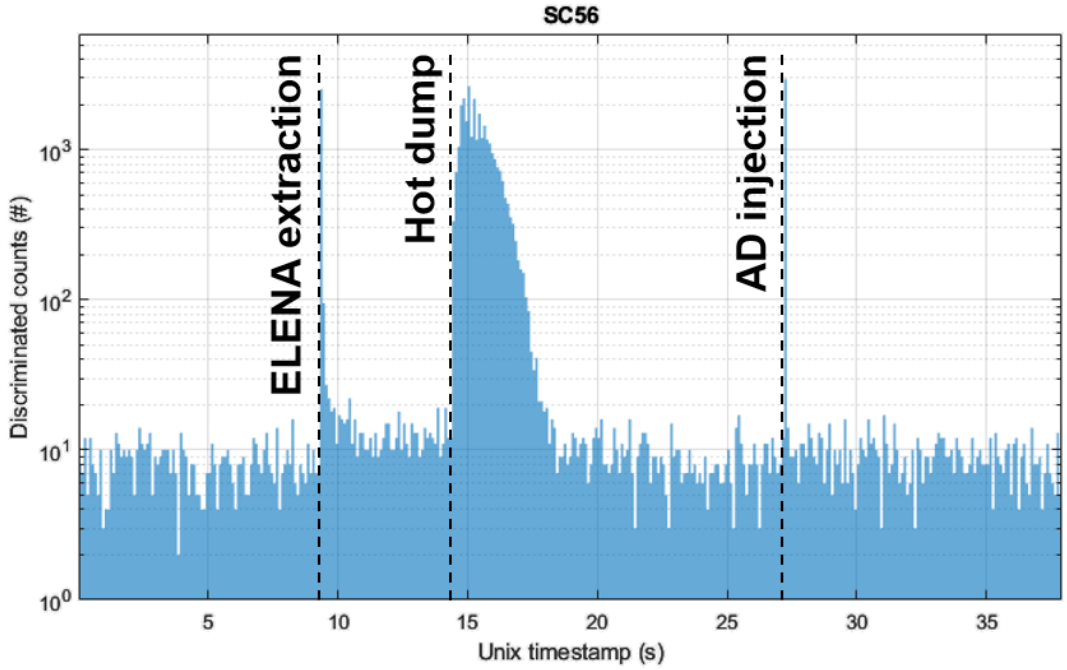


Figure 5.12 – Example of signal produced by a *hot dump* on one of the scintillators (in this case, SC56). The procedure consists of slowly releasing $\bar{p}$ towards one end of the trap (in this case, the degrader), to count the annihilations produced, without saturating the PMTs. This signal is clearly visible when the antiprotons arrive from ELENA (first dashed line, some annihilations traversing the degrader); afterwards, after 5 s of hot storage in the trap, the *hot dump* starts (second dashed line), and the events from the annihilations on the degrader are detected and counted. The graph also presents the pions shower generated by the injection in AD of the following $\bar{p}$ bunch (third dashed line) (from [6]).

and the various materials (metals, cryogenic liquids, etc.). The re-assessment of the calibration has confirmed a mutual agreement of their reading with less than a 30 % discrepancy, which remains our biggest source of uncertainty: further studies to understand better the scintillators' systematics are foreseen.

With these considerations in mind, in 2022 we determined the maximum number of antiprotons trapped and stored (from a single ELENA shot) to be $\sim 3.7 \cdot 10^6$, as shown in figure 5.13. The record number was achieved by bringing the high-voltage electrodes up to 14 kV, by using an additional degrader layer, consisting of a 200 nm-thick Parylene N foil, to further moderate the $\bar{p}$ to increase the stoppable fraction, and by optimising beam steering and capture time as explained in *Autonomous parameters optimisation* (5.1.5). This number corresponds to approximately 62 % trapping efficiency, given that the ELENA team estimated the real amount of antiprotons in each bunch to be circa $6 \cdot 10^6$. This number aligns reasonably well with the efficiencies calculated in *Determination via losses estimation* (5.2.1) (i.e. 71 % ÷ 79 %), considering that typically 10-20 % of trapped antiprotons are lost within the first second of being stored, likely due to collisions within the trap (f_{store}, see figure 5.11).

5.2.3 Record accumulation of antiprotons

Strong of the results obtained by the record trapping efficiency, a study was performed to evaluate the capability of storing the highest possible number of cold antiprotons in the 5 T trap. To do so, a technique called *stacking* was employed: a $\bar{p}$ bunch was captured and cooled (as described in

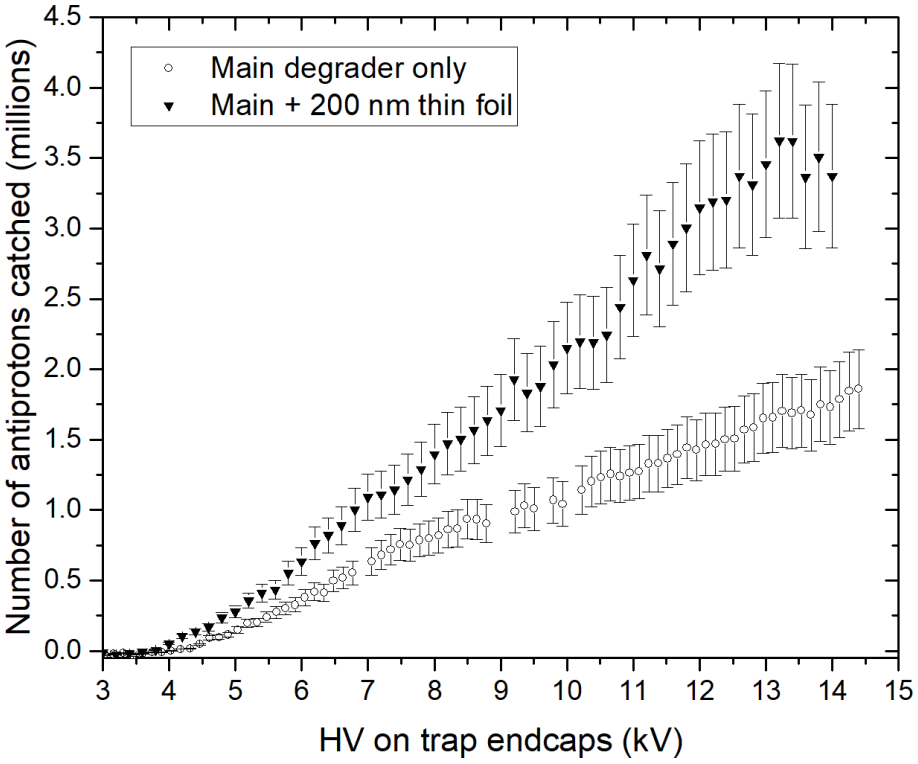


Figure 5.13 – Plot of the various measurements of caught $\bar{p}$ as a function of the high-voltage electrode potential. The series of white dots was performed with the main degrader (1400 nm of Mylar) only, and non-optimal steering. The series of black dots are the measurements performed with an additional degrader foil (200 nm of Parylene N), and both catching time and beam steering optimised. The highest amount of trapped antiproton is $\sim 3.7 \cdot 10^6$, corresponding to approximately 62 % of the number of $\bar{p}$ present in the bunch given from ELENA (from [6]).

Antihydrogen production methods (4)), and kept into a trap similar to what shown in figure 4.3, but longer, spanning from P2 to P13, to minimise the effect of space charge saturation. Subsequently, a new bunch was captured, and these new antiprotons, while cooling, fell down in the aforementioned trap, together with the previously stored antiprotons. In this manner, the amount of antiprotons over multiple shots can be accumulated, potentially leading to a very high number of antiprotons stored cold in the trap.

Nevertheless, radial losses became quickly significant, and as a result, it became necessary to utilise the rotating wall to maintain the plasma radially confined. The accumulation was tried both using a rotating wall frequency of 2 MHz and 4 MHz, and a series of accumulation trials was carried out, progressively increasing the number of shots stacked. The number of trapped $\bar{p}$ was then calculated by performing an extremely slow *cold dump*, of the order of tens of minutes, to avoid saturation of the scintillators counter.

The result of this endeavour is visible in figure 5.14: the higher frequency rotating wall has shown a better compression efficiency at higher antiproton numbers. A preliminary estimation gave a maximum amount of trapped antiprotons around 60 million: this is the highest amount of trapped cold antiprotons ever stored in a Malmberg-Penning trap.

This technique will be further studied and refined in the next years, in order to arrive both at higher numbers and at a more compressed and smaller plasma. The capability of stacking multiple ELENA shots would become immensely beneficial once the ability of "recycling" launched antiprotons (as explained briefly in *Antiprotons transfer* (4.2.4)) will be mastered: in fact, in that case, the effort of stacking and recycling the $\bar{p}$ would be largely repaid by an increase in antihydrogen production proportional to the number of antiprotons accumulated.

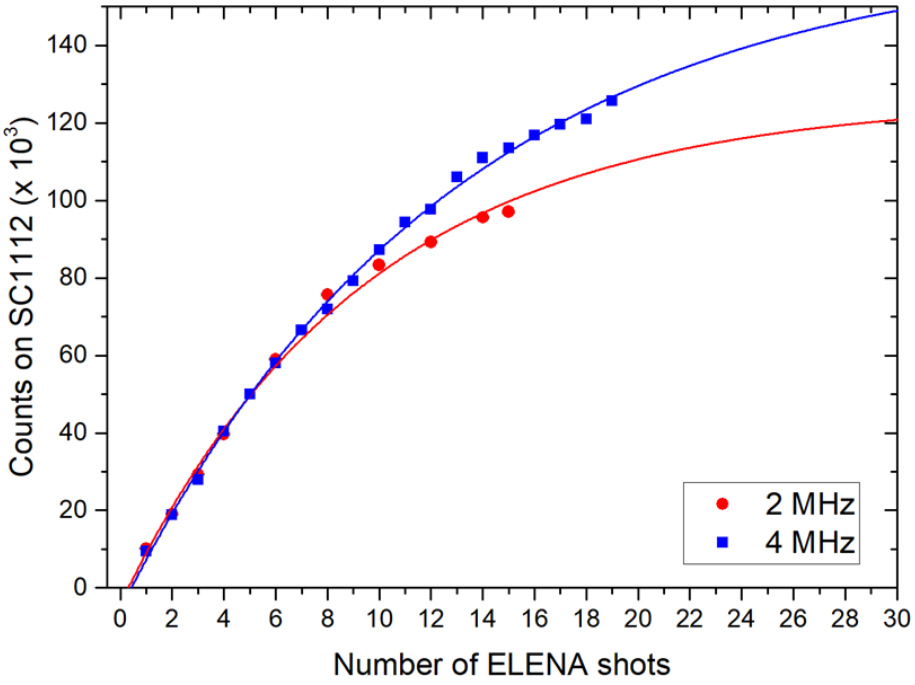


Figure 5.14 – Test of antiprotons accumulation in the 5 T trap by stacking multiple ELENA shots. Each trapped shot gets cooled by electrons and falls inside the low-voltage positive trap, where it accumulates. The amount of antiprotons is then estimated by the annihilations on the MCP, counted via the scintillators, generated via *cold dumping* the $\bar{p}$ on it extremely slowly (order of tens of minutes), not to saturate the counter. The rotating wall technique was employed to maintain radially confined antiproton plasma. In the figure the total number of annihilation events is displayed, versus the number of ELENA shots accumulated, using two different rotating wall frequencies. The maximum number of antiprotons accumulated was estimated to be above 60 million (courtesy of Ruggero Caravita).

5.3 Antihydrogen formation

Finally, at the end of the antiproton campaign of 2023, antihydrogen production was tested, operating as described in *Antihydrogen production methods* (4). In total, 916 Runs over 12 Schedules (see *The Scheduler* (3.4.5.2)) were successfully acquired, evenly spread across four measurement classes: "Forward", or "500 ns"; "Backward", or "−500 ns"; and "Still", or "0 ns"; and the control group "No IR" (see *$\bar{H}$ formation* (4.2.7) for classes details). Actually, to minimize the contribution from eventual drifts, the schedules were constructed as repetitions of four measurements, one per each class.

Unfortunately, the low-efficiency problem of the Ps target (see *Ps formation* (4.2.5)) impaired our capability of forming $\bar{H}$ in absolute numeric terms, potentially jeopardising our ability to distinguish the signal over the measurement backgrounds. A lower production rate, compatible with the reduced Ps target yield, was expected.

Three independent analyses have been carried out in parallel, to assess both the antihydrogen formation rate, and to determine the capability of forming a forward-boosted beam of neutral antimatter. In this respect, the analyses were carried in parallel over the three measurement classes used (Forward, Backward, and Still), comparing them with the same control group "No IR", to assess both the antihydrogen formation rate, and to determine the capability of forming a pulsed forward-boosted beam of neutral antimatter.

The first analysis leveraged the difference in the annihilation counts seen on the scintillators (see *Scintillators* (2.1.9.2)) in the few µs after the formation to determine the rate of formation,

and its logic and preliminary results are briefly shown in *$\bar{H}$ formation analysis using scintillator signals* (5.3.1). The other two, instead, examined the images acquired with the MCP (see *MCP* (2.1.9.1)) at the end of the traps system, with different methods. One of the latter was developed by me, and it is presented here below in great detail (see *$\bar{H}$ formation analyses using MCP images* (5.3.2)).

5.3.1 $\bar{H}$ formation analysis using scintillator signals

The first method to assess the formation of antihydrogen consists of evaluating the difference in annihilation counts seen on the scintillators between measurement and control classes, and focusing the search on the first µs after the reaction took place where most of the events are expected.

This analysis builds on the similar one performed at the end of AE$\bar{g}$IS Phase 1 to demonstrate the successful production of antihydrogen in the perpendicular geometry [15]. It is divided into four main steps:

- **Bad Runs filtering:** in this step the bad Runs are filtered out using the UV laser and the positrons as quality estimators;
- **Events discrimination:** it builds up a database of events for each PMT, starting from the raw data obtained by the scintillators, digitised;
- **Coincident events filtering:** this routine processes the events that occurred in each couple of PMTs connected to the same scintillating slab, keeping only coincident events;
- **Charge deposited filtering:** it processes the global coincident events database by filtering the events with a threshold on the charge deposited, so to maximise the signal-to-noise ratio (SNR);
- **Significance Test:** this final step assesses the significance of an excess of events in each measurement class with respect to the control one.

In the subsequent sections, the aforementioned steps are outlined in detail, and the results derived from the data acquired in 2023 are presented.

5.3.1.1 Bad Runs filtering

At the beginning, the data in each Run is briefly analysed to determine its quality. For this purpose, the spectra of the UV laser (acquired with the same photodiode and oscilloscope used in *Ps excitation* (4.2.6)) and the annihilations due to the positrons implantation on the target (determined by looking at one scintillator signal (PMT20), acquiring it analogically with a Teledyne LeCroy oscilloscope) are evaluated, both with the same algorithm. A typical signal from the lasers and the positrons is shown in figure 5.15, after being smoothed with a low-pass filter.

The algorithm starts by searching, inside the acceptance window, the maximum of the waveform, and then the FWHM of the corresponding peak is estimated. In addition, the integral inside the FWHM of the curve is calculated: it is a more robust estimator for the laser/positron intensity than the maximum amplitude. These three numbers are subsequently filtered according to pre-defined acceptance windows: the time of the maximum serves to verify the correct time alignment of the laser/e^+ shot; the integral is used to understand if the amplitude of the signal is nominal, and the FWHM is used mainly for the positrons, to assess the correct bunching of the particle cloud. An example of these analyses is given in figure 5.16.

The Runs where at least one of these checks is not passed are discarded, and the $\bar{H}$ formation analysis continues on the remaining ones. Typically, about 4 % of Runs are discarded in this manner; however, this filtering step is a necessity as those Runs showing a poor positron bunching can lead to a strong background component from delayed positrons in the signal window.

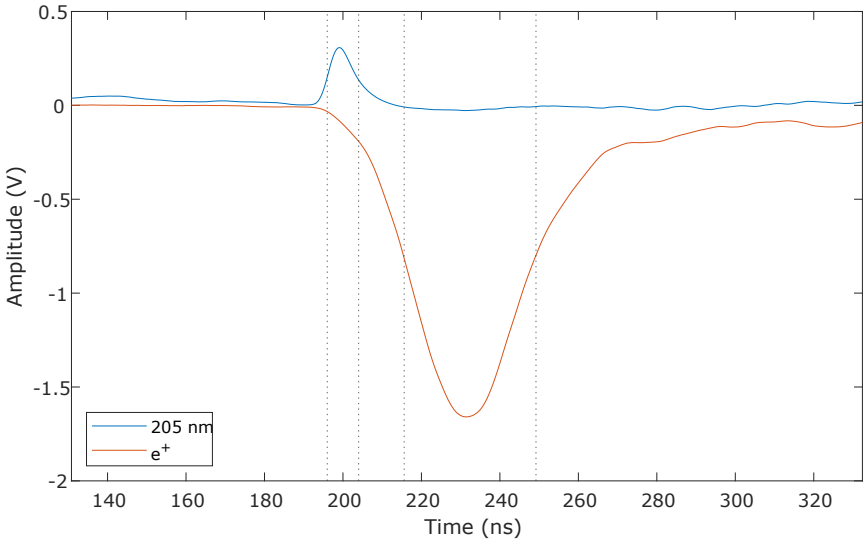


Figure 5.15 – Example of a typical signal obtained with a fast oscilloscope of both the UV laser (positive excess) and the positron annihilations from the implantation on the target (negative curve), smoothed. The vertical dashed lines depict the FWHM boundaries of the two curves (courtesy of Ruggero Caravita).

5.3.1.2 Events discrimination

The scintillators' data from the filtered Runs are then processed to extract all scintillations events above the threshold, which can be originated by several sources: dark counts of the PMTs, cosmic ray muons, gamma rays from positron annihilations, after-pulses of the PMT after the strong e^+ gamma-ray burst, antiproton and antihydrogen annihilations.

For this analysis, the raw signals from all the scintillators are used, digitised by a CAEN V1720 digitiser (12 bits, 250 MS/s). Each trace is analysed in independence, and all the events are cumulated[5]. An example of such signal is visible in figure 5.17 (red line).

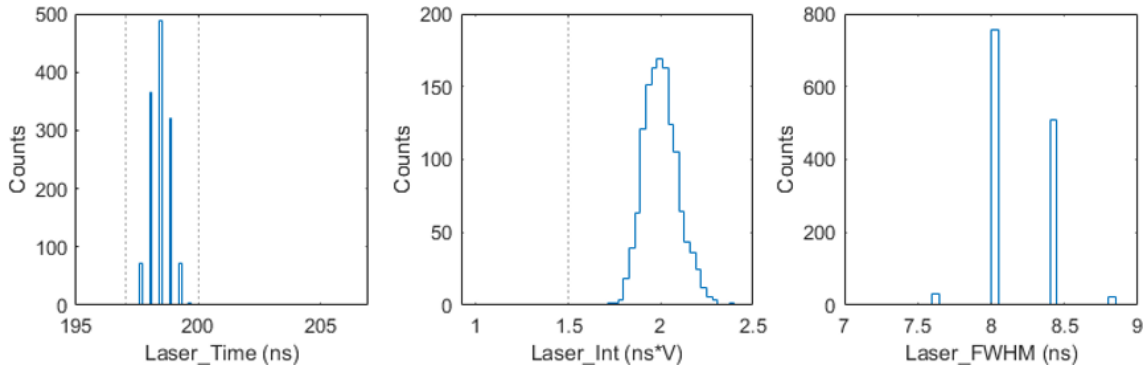
For each trace, a time-wise scanning is implemented, and two actions are carried out: the following of the baseline, and the discrimination of events. As a matter of fact, the baseline sometimes experiences drifts in either the positive or negative direction over the course of many tens of microseconds (probably due to minimal fluctuations of the potential given from the high-voltage power supply of the PMTs), and therefore the thresholds used to discriminate events from the noise have to be adjusted relatively to it. This operation is possible because the baseline variation characteristic times are orders of magnitude slower than the characteristic event duration (~ 10 ns). Such baseline fluctuations can be seen in figure 5.17.

To follow the baseline $\mathcal{B}$, a simple software low-pass follower is implemented through the (iterative) evaluation of

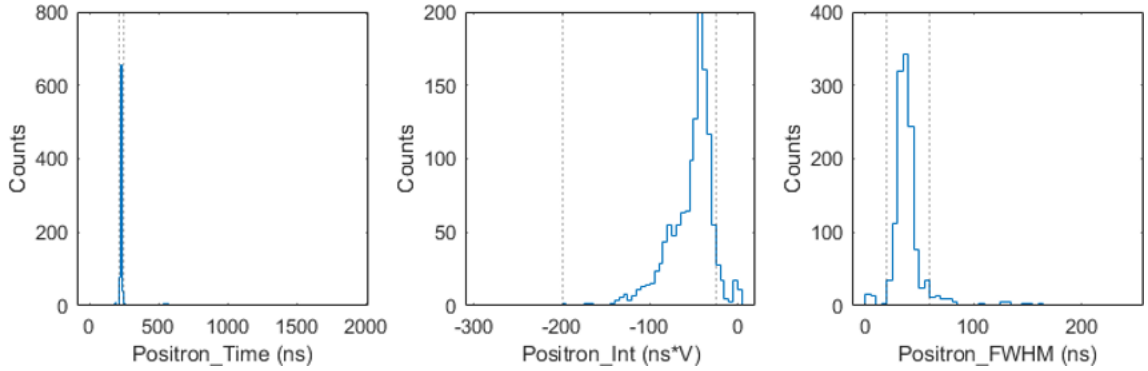
$$\mathcal{B}(t_{i+1}) = (1-w)\mathcal{B}(t_i) + wA(t_i)\,, \quad \text{with} \quad w = 0.005 \tag{5.3.1}$$

where $A(t_i)$ is the amplitude of the waveform at the time t_i, and w is the weight of the new point, which implicitly defines the filter time constant. The value of w used in the analysis ($w = 0.005$) has been empirically determined, so to follow the baseline with acceptable precision, without introducing an excessive delay in following the variation and suppressing the high-frequency behaviour that could be caused by the noise. The baseline value is updated with equation 5.3.1 during all the time steps that are outside what is determined to be an event window: inside it, the baseline is kept as a

[5] A past simulation [15] had estimated an overall probability of detecting an antihydrogen annihilation event to be approximately 40 %.



(a) Laser analysis with thresholds.



(b) Positrons analysis with thresholds.

Figure 5.16 – An example of analyses performed on the laser signal (on top) and on the positrons annihilations signal (bottom) to filter out bad Runs: the vertical dashed lines show the various acceptance thresholds (courtesy of Ruggero Caravita).

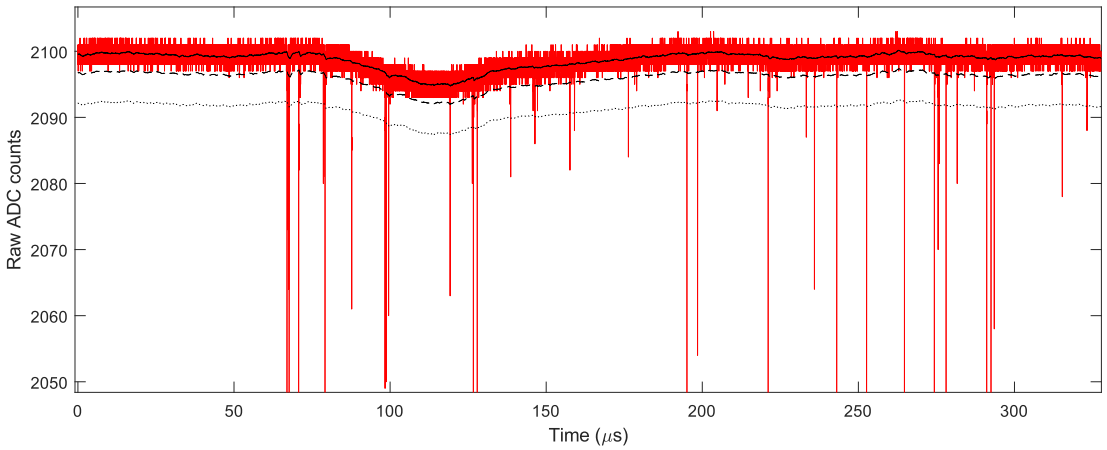


Figure 5.17 – Example of a typical digitised scintillator signal. The slow change of the baseline is visible, and the continuous black line is the following value $\mathcal{B}$. The dashed and pointed lines respectively correspond to the two thresholds used to discriminate the start (T_{up}) and the end (T_{down}) of an event (courtesy of Ruggero Caravita).

constant.

The other action is indeed to identify events by comparing the difference between the amplitude and the baseline with the event start threshold T_{up}: an event is considered to be taking place if $\mathcal{B}(t_i) - A(t_i) > T_{\text{up}}$, where $B(t_i)$ is the baseline (black curve in figure 5.17), $A(t_i)$ is the digitized PMT signal (red samples in figure 5.17), and (T_{up}) is the event start threshold (dotted curve in figure 5.17)[6]. The value of t_i is memorised as t_{start}, and in the subsequent steps the baseline is kept constant and the relative amplitude is checked to be greater than the event end threshold T_{down} (dashed line in figure 5.17). When $\mathcal{B}(t_j) - A(t_j) < T_{\text{down}}$, the event is considered finished: t_j is saved as t_{end}, and the maximum amplitude inside the event window is searched. After the closure of the event, the baseline update restarts, and the check for another event starting is resumed.

The motivation for using two different thresholds, a lower T_{up} for the falling edge (i.e. event start) and a higher T_{down} for the raising edge (i.e. event end) is dictated from the intrinsic behaviour of the scintillators and PMTs, and presence of electronic noise on the digitized trace. Indeed, typical pulse shapes from these detectors feature fast-raising edges and slower tails, sometimes (as in the cases of positron implantation peaks) extending up to several hundreds of nanoseconds. If a single threshold were employed for both the event start and stop conditions, the risk of immediately starting another event during the noise oscillations happening during the signal descent would be considerable. Instead, by keeping higher the end event threshold (the signals are negative, so in modulus is a lower threshold), we ensure that the event has fully finished, and the baseline has recovered well past the point of electronic noise oscillations.

Furthermore, being electronic noise on every digitiser channel slightly different, the two thresholds are not set as unique fixed numbers, but rather as multiples of the Root Mean Square (RMS) amplitude of the signal evaluated in the first 50 µs of the baseline, before any antiproton manipulation occurred. In the specific, the following values

$$\begin{aligned} T_{\text{up}} &= 10 A_{\text{RMS}} \\ T_{\text{down}} &= A_{\text{RMS}} \end{aligned} \tag{5.3.2}$$

have proven to robustly guarantee a correct event beginning and termination identification.

Each event found in this manner is saved in the events database as a new entry, containing:

- start time t_{start}
- end time t_{end}
- maximum amplitude A_{max}
- IR status (On/Off)
- antiprotons-to-positrons delay, in ns (-500, 0, 500)

The first three values summarise the information carried by the event shape, while the last two are used to identify the control or measurement class to which the event belongs.

When all the traces from all the scintillators for every Run are parsed, the database building is complete, and the events are further processed to form coincidences at the slab level.

5.3.1.3 Coincident events filtering

In order to eliminate background counts caused by dark counts and PMT after-pulses from positron annihilations after their implantation on the $e^+ \rightarrow$Ps target, an event coincidence analysis is performed. In fact, both dark counts and PMT after-pulses events are time-wise uncorrelated: by looking for coincident events taking place on the PMTs connected to the same scintillating slab, it

[6]The inequality is written in this manner because the events values produce a negative signal.

is possible to reject them efficiently, effectively suppressing these backgrounds from 500 ns after the positrons implantation onwards.

The analysis develops as follows. For each slab in the scintillators' array, the databases of the events from the two PMTs connected to it are parsed, and all the possible couples of events (each from one of the two databases) are generated. Afterwards, the modulus of the difference between the two t_{start} of the events of each couple is calculated, and compared against half of the coincidence window (set to be 50 ns, see *Scintillators* (2.1.9.2)): if the events fall outside the acceptance window, they are discarded; otherwise, a new event is saved in a new global events database, inheriting the parameters from the original couple of events.

All the events from all the Runs are filtered in this manner, and the new global events database is filled. Typically, about 75 % of the overall events are identified to be part of coincidences, with a $\sim$ 10 % fluctuation slab-by-slab. These resulting events are further filtered by discriminating their amplitude.

5.3.1.4 Charge deposited filtering

In comparison, the routine for analysing the coincident events saved in the global database is simpler than the ones building the database: the events are solely discriminated based on an amplitude threshold, and separated in the classes they belong according to IR laser status (used to activate or deactivate antihydrogen production) and the antiproton-to-positron delay. Afterwards, significance tests are evaluated to determine the production of antihydrogen.

The event selection via an amplitude threshold cut is mainly motivated by the necessity of eliminating all the counts generated by the gamma rays emitted by the positronium atoms annihilations. An analysis done in [103] has estimated as 220 mV the minimum threshold to effectively discriminate away gamma ray events while retaining most of the events generated by an antiproton annihilation (the results from the analysis are visible in figure 5.18). Being now the transmission line from the PMT to the digitiser attenuated by a further 20 dB filter (corresponding to an overall factor of 20 smaller signal amplitude as if the digitizer would be directly connected to the PMT), done to avoid "ringing" on the transmission line caused by the splitting of the signal (to the digitiser and a separated event counter), the effective threshold used for this analysis and applied to the raw data is in fact of 25 mV, corresponding to an equivalent 500 mV at the PMT output: a more conservative choice than cutting at 11 mV (corresponding to 220 mV), motivated by the higher ADC quantization error due to the introduction of the anti-ringing attenuators. From now on, we only refer to threshold and amplitudes at the PMT output, implicitly assuming this factor $\times 20$ multiplication to account for the transmission line attenuation.

This threshold cut manages also to eliminate other two sources of spurious events: HV power supply spikes, which have zero average, but they can sometimes surpass briefly the T_{up} threshold, and the dark counts, caused by electrons thermally emitted by the PMT. Both these false signals have amplitudes below 100 mV, more than a factor of five lower than the signal threshold, and hence easily discarded.

From the dataset refined in this manner, where all the events coming from neither antiproton/antihydrogen annihilation pions nor cosmic muons (essentially negligible, as they contribute only as a small constant background, see figure 2.29) are excluded, a preliminary study of the annihilations caused by the $\bar{\text{p}}$ travelling back and forth inside the parabolic potential was performed by looking at the number of events in the time window between 80 µs and 200 µs (corresponding to $\sim$ 9.3 µs after the positrons implantation: no $\bar{\text{H}}$ should survive long enough to annihilate inside this window). An example of this signal is given in figure 5.19.

The periodic annihilations are evident, giving a total "swinging" period of $\sim$ 20 µs. Two things can be noticed from this graph. First, the delay impressed during launch to the $\bar{\text{p}}$ (which defines the measurement class, see *$\bar{H}$ formation* (4.2.7)) is maintained throughout the various oscillations: this is an indication of the time-stability of the parabolic potential used for the transfer. Second,

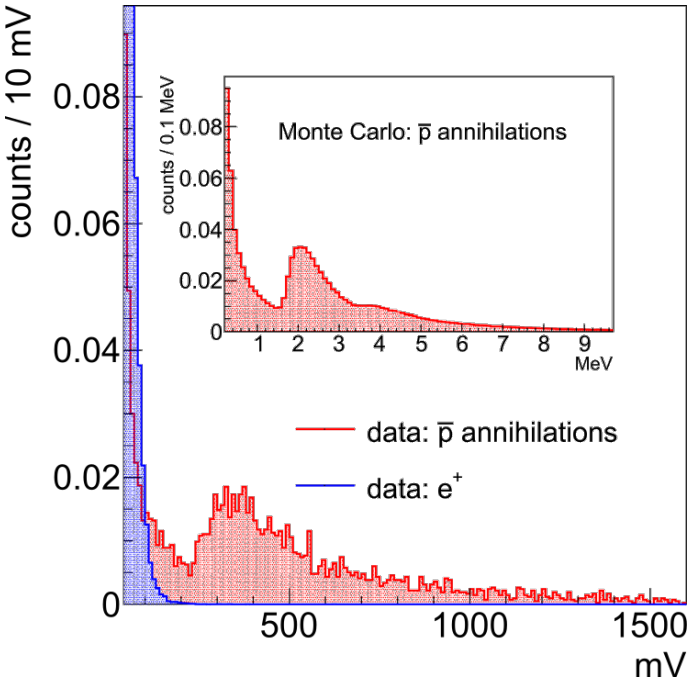


Figure 5.18 – Measured scintillator signal caused by a superposition of sources, mainly positronium annihilations gamma from target implantation (blu curve) and pions from antiprotons annihilations (red curve) - corresponding to the same signal expected from antihydrogen annihilations. The plot is remarkably similar to what was obtained in the inset via a Monte Carlo simulation of antiprotons annihilating on the trap walls. The threshold to confidently reject the signals coming from the positrons has been derived to be 220 mV(from [15]).

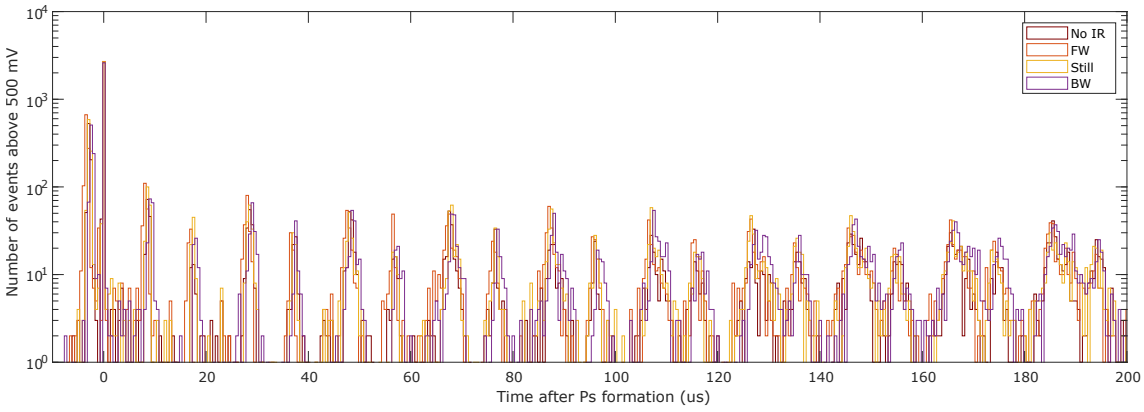


Figure 5.19 – Example of antiprotons annihilations during the oscillations inside the parabolic transfer potential, captured by the scintillators, in the three measurement and control configurations (FW - Forward, Still, BW - Backward, and No IR). It can be noticed that the delay impressed to the antiprotons at launch time – which defines the measurement class – is maintained during all the oscillations, an indicator of the stability of the parabolic potential. Moreover, the periodic annihilations represent the passage of the antiprotons in the central region of the experiment: the azimuthal component of the magnetic field on the axis (see figure 2.3) has a minimum point in the middle of the apparatus, causing the $\bar{p}$ to traverse a positive gradient in both directions, leading to an expansion of the plasma that generates the annihilation of the tail of the radial distribution (courtesy of Ruggero Caravita).

the periodic annihilations have two alternating amplitudes. A possible explanation can be derived from the antiproton plasma oscillation dynamics inside the real magnetic field of AEgIS. Indeed, the magnetic field inside the cryostat has a global minimum, in the shape of a little bump, in the field-changing region in the middle of the apparatus (see figure 2.3). Therefore the particles see a positive field gradient both passing from the 5 T to the 1 T region, and vice versa, causing a sudden expansion of the plasma. This causes the tail of the radial distribution of $\bar{p}$ to annihilate on the trap wall at every passage. But the magnitude of the negative magnetic field gradient is stronger when going towards the 1 T section than going in the other direction (because of the small "bump", there is a little negative magnetic field gradient also when going from the 1 T towards the 5 T section: see figure 2.3): hence the difference in annihilation counts. Another possible explanation takes into account the fact that the travel time inside the 1 T section of the traps system is longer than in the 5 T section, and the magnetic confinement is less strong: therefore, the $\bar{p}$ plasma expands more during the permanence in the 1 T section, and when traversing the $\vec{B}$ gradient the bigger tail causes more annihilations. Clearly, the solution can be a combination of these factors and will require further investigation.

In addition, the fact that the annihilation peaks are well distinct even after many oscillations is a sign of the smoothness of the shaped parabolic potential (roughness would likely cause plasma spreading).

By plotting the total number of events captured in the window defined above (from 9.3 µs to 129.3 µs after positrons implantation), we noticed that an abrupt change in the number of events per Run occurred on November 6th, after Run 404638: from a maximum number of events well below 100, it passed to more then 200 (see figure 5.20).

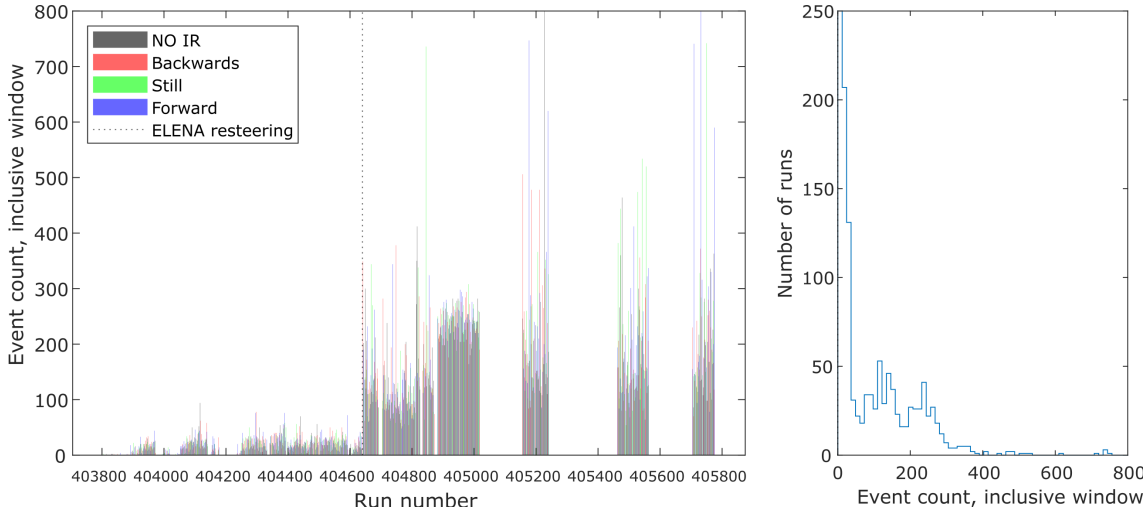


Figure 5.20 – On the left: plot showing the events count per Run (on the x-axis, the Run number is displayed), captured in a window from 9.3 µs to 129.3 µs after positrons implantation. The four colours distinguish the different measurement and control classes. A change of trend can be seen after Run 404638: a possible explanation is a change in the configuration of ELENA. Three different datasets were defined: one only taking Runs with before Run 404638, called "Before" dataset, one taking only the Runs after, called "After", and one accepting all the Runs taken, called "All". On the right: graph depicting the distribution of events per Run (courtesy of Ruggero Caravita).

This occurrence was synchronous to a change in magnetic steering configuration from ELENA. In fact, in the last days of the beam, the ELENA operators tried (with success) to increase the amount of antiprotons delivered in each single bunch per experiment to 10^7 by changing the configuration of ELENA. During the optimisation, one of the particle steerers along the ELENA-AEgIS line was found to be in a faulty state and fixed, introducing yet another discontinuity with the previous conditions. This led to a higher number of available antiprotons, and possibly also to a change

in their radial distribution (to be further investigated) at the instant of the ballistic transfer, and during the subsequent swings in the parabolic potential (since a burst of annihilations happens every time the magnetic field gradient is traversed). However, being the starting antiproton number greater, a higher amount of antihydrogen could potentially be produced. Sic stantibus rebus, we chose to organise the analysis in three different datasets: one comprising only the Runs before the change (called *"Before"*), one comprising only the Runs after the change (called "*After*"), and one including all the Runs collected in the 2023 $\bar{\mathrm{H}}$ campaign (*"All"*). The time window to look for $\bar{\mathrm{H}}$ has been restricted to between 0.5 µs and 6 µs after the positrons implantation inside the target. The first 500 ns are excluded to avoid the counting generated by the positron annihilations due to the implantation in the target, and the data is cut before 6 µs to avoid the annihilations caused by the antiprotons traversing the magnetic field gradient (and therefore the radial enlargement of the bunch make the outermost $\bar{\mathrm{p}}$ collide with the trap walls).

To check the validity of the amplitude threshold described at the beginning of the section, a scan over it has been performed, both in the "Before" (figure 5.21a), in the "After" (figure 5.21b) and in the "All" cases (figure 5.21c). The events count has been divided into the three measurement classes (Backward, Still, and Forward) and the control.

5.3.1.5 Significance tests

With the data processed in this manner, it is already possible to perform some significance tests to prove that antihydrogen has indeed formed. The method chosen consists of calculating the deviation from the hypothesis that no antihydrogen is formed, which would result in an identical number of events between the measurement classes and the control class. This evaluation is carried out with the use of a binomial test, which tests the probability that the number of events in the two classes is the same. To obtain the number of standard deviations from the null hypothesis we can use the following formula (from [158])

$$S = \sqrt{2\left(N_{\mathrm{on}} \ln\left[\frac{1+\alpha}{\alpha}\left(\frac{N_{\mathrm{on}}}{N_{\mathrm{on}}+N_{\mathrm{off}}}\right)\right] + N_{\mathrm{off}} \ln\left[(1+\alpha)\left(\frac{N_{\mathrm{off}}}{N_{\mathrm{on}}+N_{\mathrm{off}}}\right)\right]\right)} \tag{5.3.3}$$

where N_{on} is the number of events in the measurement case, N_{off} is the number of counts in the control case, and $\alpha = t_{\mathrm{on}}/t_{\mathrm{off}}$ correspond to the ratio of the number of acquisitions been done in the measurement case versus the control case (in our case, $\alpha \approx 1$ when confronting directly one measurement class with the control class, being the statistics approximately the same [316,317,314,310], and $\alpha \approx 3$ when summing up all the measurement classes and confronting their totality with the same control class). The number of antiprotons given by the ELENA team has not been taken into account for normalisation purposes, because of its high uncertainty ($\sim 10 \div 20\%$).

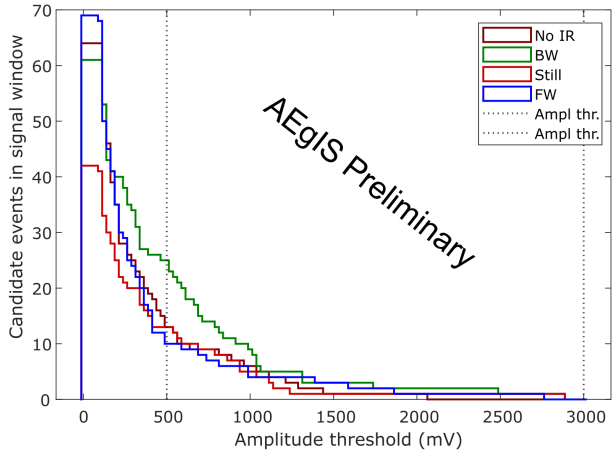
In the "Before" cases, though, the numbers are too small to be treated with equation 5.3.3 (which requires $N_{\mathrm{on}}, N_{\mathrm{off}} \gg 10$). Therefore, a pure binomial test has been performed, transforming the so-found probability to standard deviations using the inverse cumulative distribution function Φ:

$$S = \Phi^{-1}(1 - P(N_{\mathrm{on}}, N_{\mathrm{on}} + N_{\mathrm{off}}, p = 0.5)) = \Phi^{-1}\left(1 - \binom{N_{\mathrm{on}} + N_{\mathrm{off}}}{N_{\mathrm{on}}} 0.5^{N_{\mathrm{on}}+N_{\mathrm{off}}}\right) \tag{5.3.4}$$

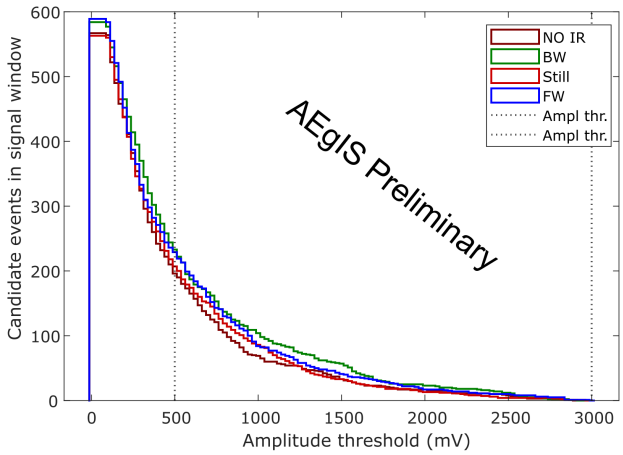
with the notable exception of the "Before" overall case, since the heavy difference of the acquisition times ($\alpha = 3$) makes a direct binomial test meaningless. In this case, equation 5.3.3 has still been used, taking the result with the necessary care[7]. The results of all these calculations are shown in table 5.6.

In the Backward "Before" case, we have clear evidence of antihydrogen formation already by bare counting the events in the entire signal window; in the overall "Before" case, the evidence

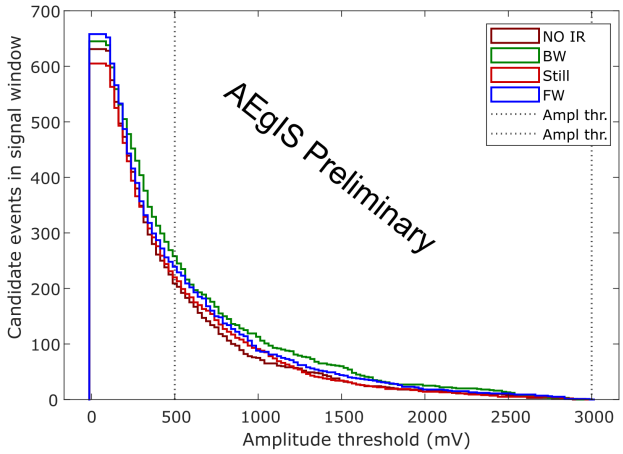
[7] Conversely, it cannot simply be done a binomial test with $N_{\mathrm{on}} = 41$ and $N_{\mathrm{off}|equiv} = \alpha N_{\mathrm{off}}$, because of the non-linearity of the binomial distribution.



(a) Amplitude distribution of events in the "Before" case.



(b) Amplitude distribution in the "After" case.



(c) Amplitude distribution in the "All" case.

Figure 5.21 – Amplitude distribution of events, both in the "Before" (above) and in the "All" (bottom) cases. The events count is divided into the three measurement classes ("BW" - Backward, Still, and "FW" - Forward) and the control ("NoIR"). The dashed vertical line corresponds to the 500 mV threshold (as explained in the text) (courtesy of Ruggero Caravita).

Dataset cut	Case	N_{on}	N_{off}	σ
"Before"	Backward	25	4	4.3
	Still	13	4	2.5
	Forward	13	4	2.5
	Overall	41	4	2.7
"After"	Backward	233	205	1.7
	Still	226	205	1.3
	Forward	207	205	0.1
	Overall	676	205	1.2
"All"	Backward	258	209	2.3
	Still	239	209	1.4
	Forward	220	209	0.5
	Overall	717	209	1.7

Table 5.6 – Results of antihydrogen formation in the three datasets ("Before", "After" and "All"), both overall and subdivided into the three different measurement classes (Forward, Still and Backward). It can be seen that there is clear evidence of formation in the "Before" Backward case. In the overall cases, a clear excess over the background is not present.

is weaker, but it is a still good indication. Unfortunately, in the "All" case, no real evidence is present, because the higher background caused by the antiproton annihilations completely covers any potential $\bar{H}$ signal.

A stronger observable of antihydrogen formation is obtained by slicing the events count with a rolling time window of 1.5 µs. In this manner, it is possible to better isolate the antihydrogen produced from the background. The results of this analysis are plotted in figure 5.22.

In the "Before" dataset, only in the Backward case, an excess can be identified; in the "All" dataset, there is strong evidence of $\bar{H}$ formation in the Backward case, good evidence in the Forward case and less evidence in the Still case. In this case, it is remarkable, though, that the maximum significance, for the three classes, lays approximately at 3 µs, 3.5 µs and 4 µs: the time separation of the maxima is consistent with the difference in delays impressed to the antiprotons when launched.

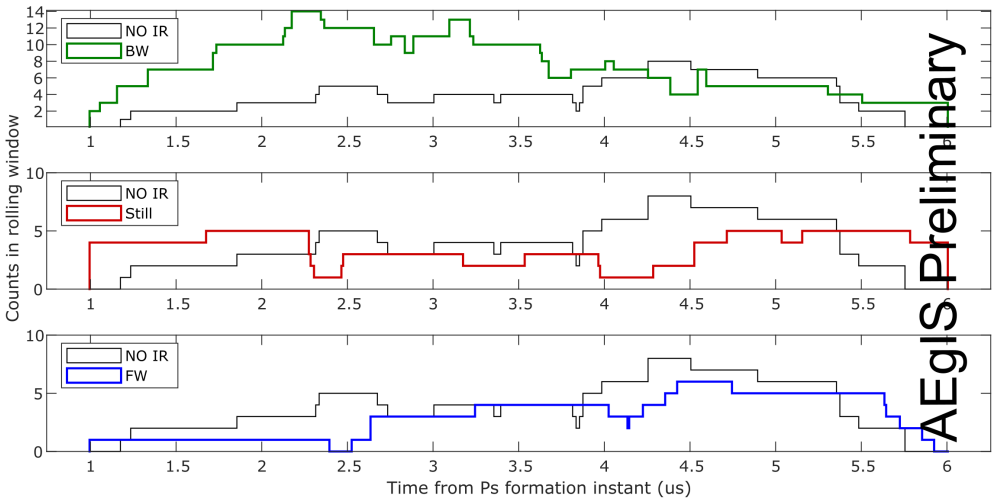
The corresponding deviations from the null hypothesis are presented in table 5.7.

Dataset cut	Case	N_{on}	N_{off}	σ
"Before"	Backward	14	3	3.0
"All"	Backward	115	60	4.2
	Still	97	67	2.3
	Forward	111	70	3.1

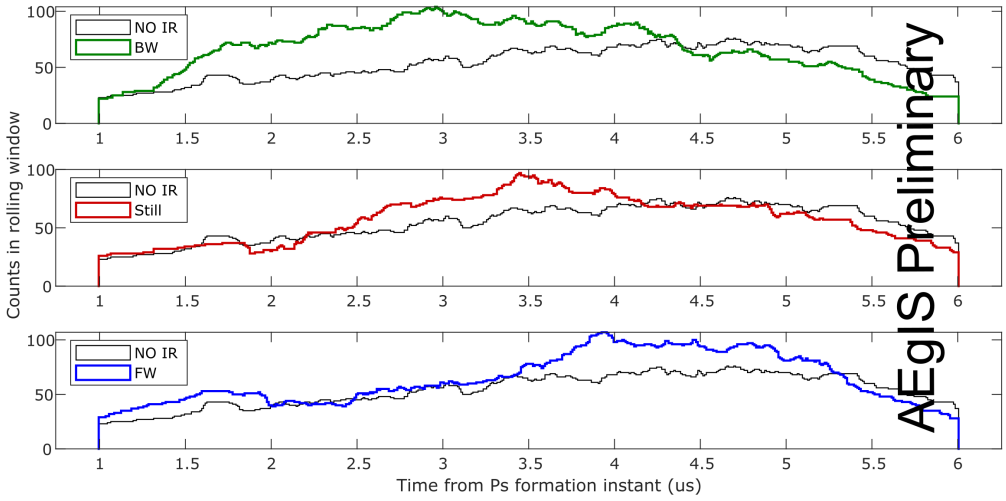
Table 5.7 – Results of antihydrogen formation in the "Before" Backward, "All" Backward, "All" Still and "All" Forward cases, in the best bin of a 1.5 µs rolling window. Evidence of antihydrogen formation is present in the first and last case, and strong evidence in the "All" Backward. This analysis confirms the pulsed nature of the production.

In the results of both tables 5.6 and 5.7 we can see an increased efficiency of production in the Backward cases with respect to the Still and Forward cases. We can advance three hypotheses: the first is that the formation of antihydrogen has been favoured in the Backward cases because the antiprotons have a similar velocity to the emitted positronium atoms (same direction and same order of magnitude of modulus), which enhances the formation rate, being $R_{\overline{H}} \propto |v_{\overline{p}} - v_{Ps}|^{-1}$.

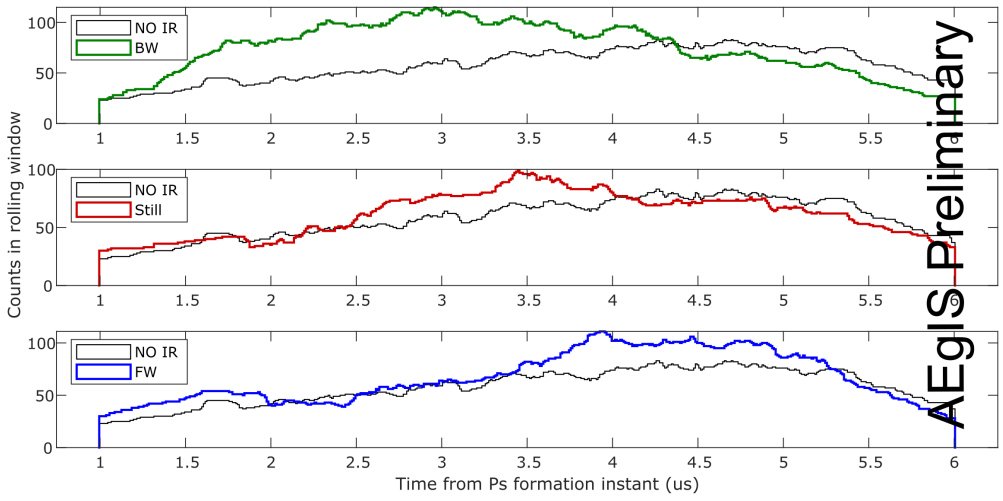
The second hypothesis lies in a not perfect match in time of the two particles clouds: in fact,



(a) Time distribution of events in the "Before" case.



(b) Time distribution in the "After" case.



(c) Time distribution in the "All" case.

Figure 5.22 – Time distribution of events, both in the "Before" (above, left), "After" (above, right) and in the "All" (below) cases. The events count is divided into the three measurement classes ("BW" - Backward, Still, and "FW" - Forward) and the control ("NO IR"). (courtesy of Ruggero Caravita).

if the time alignment performed as explained in *$\bar{H}$ formation* (4.2.7) has not brought the clouds to be aligned in the "Still" case, it is possible to see a variation of production rate depending on the real cloud time alignment. It is also possible that the formation always peaks in a determined instant (for example, in the moment that the $\bar{p}$ has the same velocity of the travelling Ps): this would explain why, in the three plots seen in figure 5.22, the maxima of the significance are spaced of $\sim 500\,\text{ns}$, consistent with the different delays of the three classes, as discussed before.

To efficiently form a forward-boosted beam of antihydrogen, then, a higher overall production rate will be needed, coupled to a better $\bar{p}$-Ps velocity and time-overlap fine-tuning. This analysis hints toward the possible great benefits that the development of a $e^+ \rightarrow$Ps transmission target (as outlined in [159]) would bring for an efficient $\bar{H}$ beam formation.

5.3.2 $\bar{H}$ formation analyses using MCP images

This analysis, ideated and carried out by me, centred on searching a signal excess between each of the three $\bar{H}$-production-favoured classes and the control class, using the images captured by the MCP at the end of the trap system. The logic of the analysis consists, for each image, to equalise it with respect to the gain of the camera, and subtract the camera background; then, only the circle part of the image inside the ionisation grid is kept, and the image binarised with the application of threshold. The analysis then proceeded with the application of a median filter to remove the salt-and-pepper noise and the counting of the number of spots with the use of a clustering algorithm.

At the end, a binomial test is performed, testing the probability that the same amount of annihilations occurs in all classes considered.

Being the threshold for binarisation the parameter that the analysis pipeline is most sensitive to, and the most arbitrary, a scan over its value has been performed, in order to find the best value for SNR and output significance. Moreover, also a scan over the parameter for median filtering has been executed, to evaluate the impact of this passage on the result.

In the following, the analysis algorithm is presented in detail, and the results obtained are commented on.

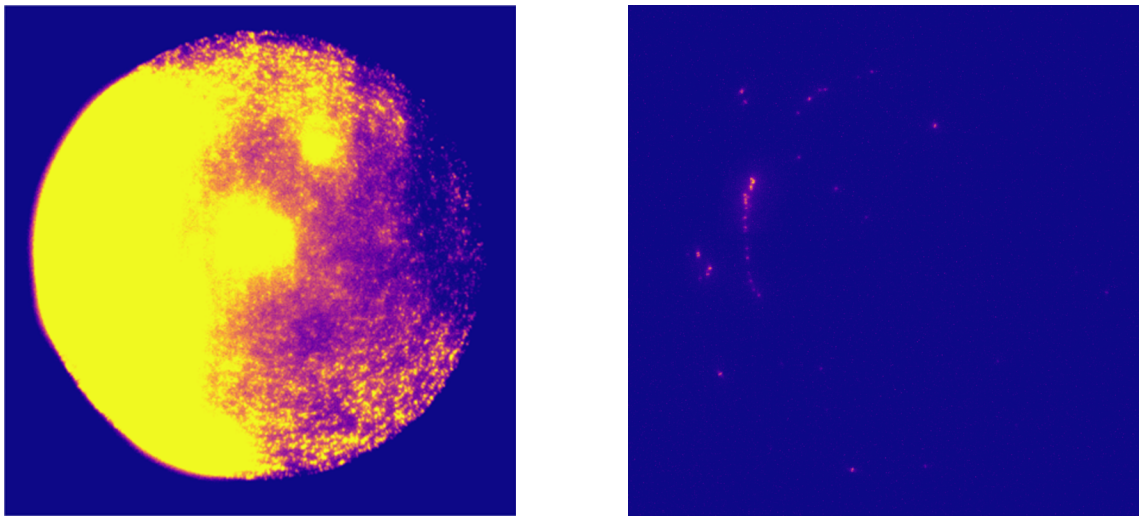
5.3.2.1 Physical background removal

The images were initially heavily polluted by a huge background, rendering them unusable. This was unexpected since the ionisation grid facing the trap was kept at $-400\,\text{V}$ to fully filter incoming antiprotons. A possible explanation for such a high background, present both in the measurement and the control classes, is the emission of electrons from the grid by the photoelectric effect with the UV laser.

To mitigate this problem, we restricted the total acquisition time of the MCP below its minimum setting (1 ms) by gating the voltage given to the phosphor screen with the same Pockels cell used for the precise time control of the "Alex" laser (see *In-Run autonomous parameter stabilisation* (5.1.4)). The Pockels cell was inserted in the middle of the line between the phosphor screen power supply and the phosphor screen itself, to be able to interrupt the power supplied to the sensor: this was performed by a series of two pulses, so to isolate the portion of time around antihydrogen production with great precision. In this way, we were able to reduce the effective acquisition time of the MCP down to 100 µs, starting 500 ns after firing the lasers, reducing the background substantially, and de-facto enabling the analysis of the images. The difference is shown in figure 5.23, where the left image is taken before the introduction of the Pockels cell and the right one after.

5.3.2.2 Camera gain equalisation and background subtraction

The first step in the image analysis pipeline consists of equalising the figure with respect to the camera gain and removing the dark counts noise background.In fact, the Hamamatsu camera used



(a) MCP image before introducing the Pockels cell.

(b) MCP image after introducing the Pockels cell.

Figure 5.23 – MCP images showing the difference of background noise before (left) and after (right) introducing the Pockels cell to gate the power given to the phosphor screen, so to effectively diminish the exposure down to 100 µs.

to acquire the images has a non-settable varying gain which multiplies the values of each pixel of the image by a power of 2: therefore, is necessary to equalise all the images with respect to this gain, to be able to compare them. To do that, the average of the pixel values inside the four squares contained in the four corners of the sensor unexposed to light is computed (because outside the trap electrodes area, see figure 5.24), and the base-2 logarithm of this average is calculated: if close than 0.2 to an integer, that integer number is interpreted as the gain, and the image corrected correspondingly; otherwise, the image is discarded[8]. Last, 1 is subtracted from every pixel value, corresponding to the *dark counts* value after normalisation.

In formulas:

$$\text{gain} = \text{round}\left(\log_2\left(\frac{\sum_{corner=1}^{4}\sum_{j=1}^{N_{pixels}} A_{j,corner}}{4N_{pixels}}\right)\right) \tag{5.3.5}$$

$$A_{i|corr} = \frac{A_i}{gain} - 1 \tag{5.3.6}$$

where A_i is the value of the pixel i.

5.3.2.3 Images region-of-interest selection

To improve the speed of the analysis pipeline, and to remove possible fake signals coming from regions of the MCP where is geometrically impossible for antihydrogen atoms to fly, the images are cropped so to keep only the part inside the inner contour of the ionisation grids. For this, an image of the grid was taken precisely with electrons, and the equation of the circle was determined as the best fit of eight points manually selected (see figure 5.25). In this manner, the images are reduced to a circle of 500 pixels of radius, diminishing the number of pixels to be processed to 19 %.

The region-of-interest (ROI) is further refined, having noted that on most images one or two very bright spots are always present (both in measurement and in control classes), exactly in the same position, on the top centre of the image[9] (see figure 5.26). These bright spots have been

[8]Typically, less than 1‰ of the images is discarded in this way.

[9]Which, in reality, corresponds to the centre-left of the image: the acquisition chain flips the x coordinate with the y-coordinate.

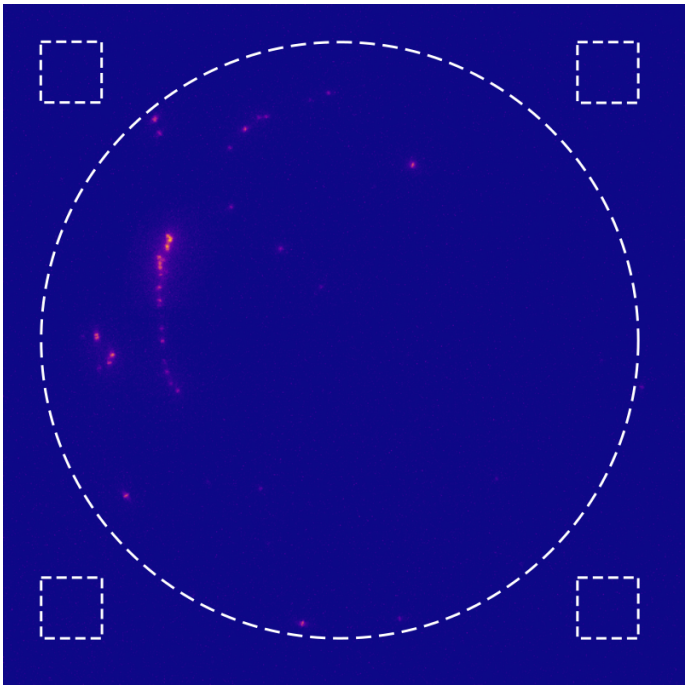


Figure 5.24 – The dashed four squares depict the areas used to correct the non-settable varying gain of the Hamamatsu camera used to acquire the images and to remove the *dark counts* background from every image. They are well outside the image of the perimeter of the trap (dashed circle).

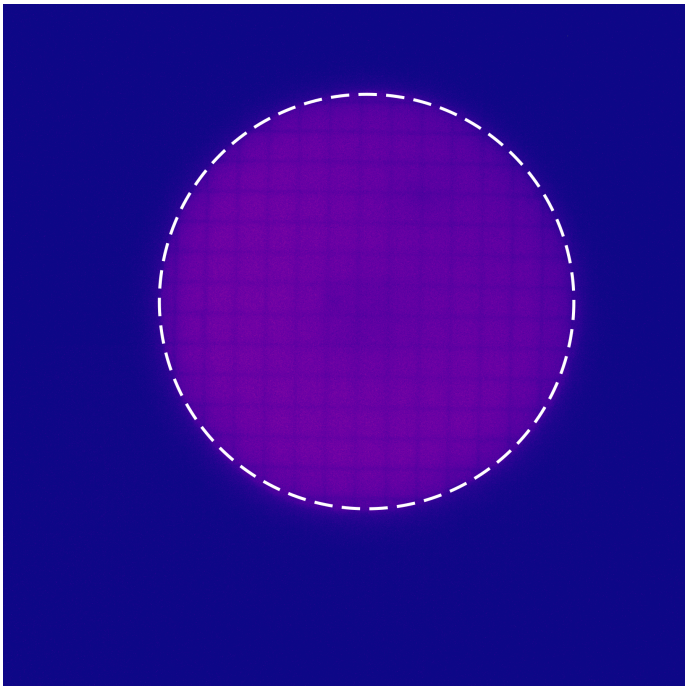


Figure 5.25 – The ionisation grid imaged launching electrons towards the MCP. The dashed circle is used to define the contour of the region of interest.

interpreted as electron emission from the grid, probably due to a particular reflection point for the UV laser, or to a minimal imperfection of the grid, causing electron emission due to the high gradient electric field. Consequently, a rectangle of 40x50 pixels is always removed ($\sim 2.5‰$ of the useful surface).

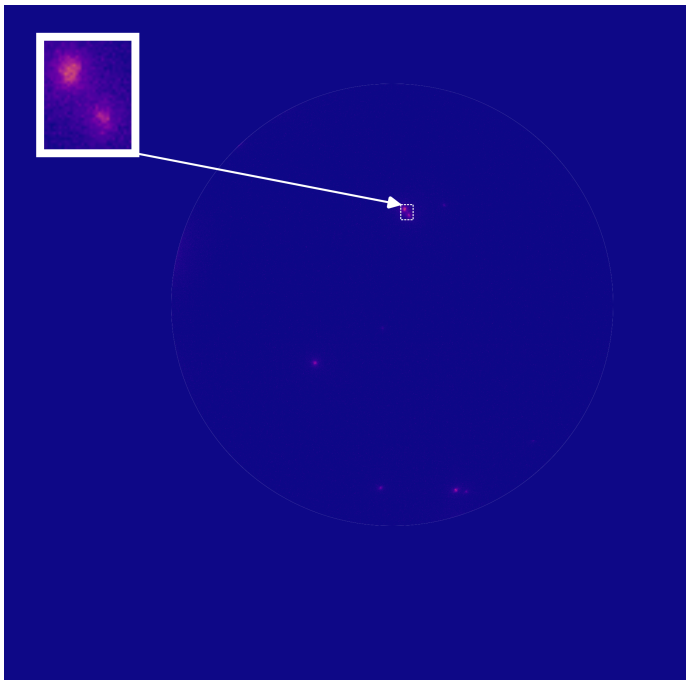


Figure 5.26 – MCP image where the double recurring bright spots are highlighted. The dashed rectangle is the area excluded from the analysis, and a zoomed version is given in the top-left inset.

5.3.2.4 Exclusion of bad images

After the isolation of the ROI, the images are filtered so to remove the problematic ones.

First, all the images corresponding to Runs marked by CIRCUS as "bad" are, of course, excluded (40 Runs). Afterwards, an analysis based on the sum of the values of all the pixels is performed: in fact, some images displayed some very large, extremely bright spots, which caused the sum to jump from the average value below 10^4 to above 10^5. The images with a sum bigger than 10^4 were isolated and visually inspected, and 33 of them were finally discarded. An example of these images is given in figure 5.27.

With this selection in place, the rest of the analysis was executed on 843 images, subdivided into the four classes almost equally (less than 1 % difference).

5.3.2.5 Image binarisation

The images so treated are further simplified, to better distinguish and isolate the events. To do so, the images are "binarised", i.e. all the pixels are transformed from integers into boolean values by confronting them with a predetermined threshold. The pixels whose values fall below such threshold are set to `0`, and the ones equal or above are set to `1`. Being a determination a priori of the correct threshold impossible[10], a scan binarising over multiple threshold values is performed, from 0.3 to 2.5 in step of 0.1. Smaller threshold values would have picked up too much noise, and

[10] All the images were visually inspected for the presence of traces or "stars", which would have been an unmistakable sign of an antiproton annihilating on the phosphor surface: the central region of such a formation could have given an indication for the pixels' value to expect for an antihydrogen signal, and consequently for the threshold to adopt. Unfortunately, no such structure was found, and therefore the a priori determination of the threshold was deemed impossible.

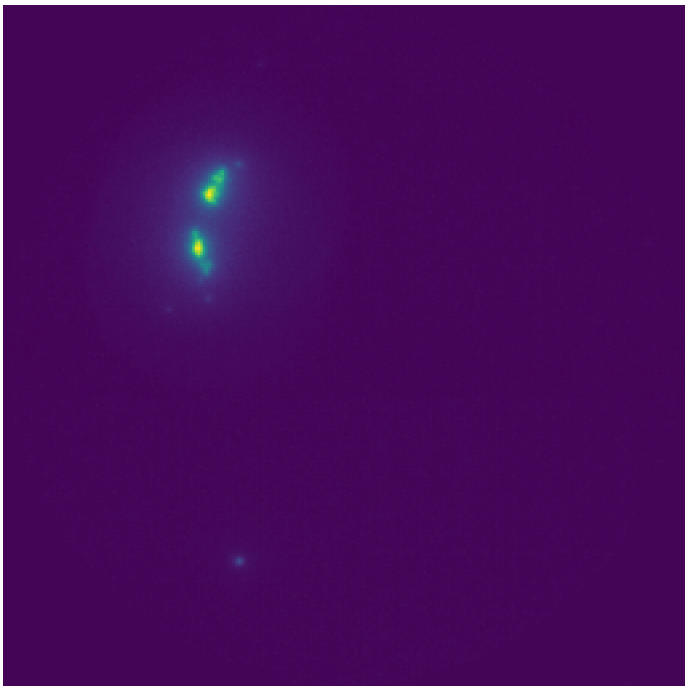


Figure 5.27 – An example of a visually bad image with a high sum value, that was discarded from the analysis.

above 2.5 basically no points survived on the images after binarisation. An example of the effect of the binarisation of the images with different thresholds is given in figure 5.28.

5.3.2.6 Median filtering

Subsequently, to remove the salt-and-pepper noise, a median filtering technique is adopted [160]. It is especially valuable in this context because it removes some noise while preserving the contours of objects. The median filter operates by substituting to each pixel value the median of the values of the surrounding pixels: the amount of pixels to be considered is determined by the `kernel-size` parameters. A kernel value of 5 was found to effectively remove all single and double pixels noise clusters, without significantly affecting the bigger clusters: but being still an arbitrary choice of the parameter, a scan was performed over `kernel-size` values of 0, 3, 5, 7, and 9, where in 0 no median filtering was done[11].

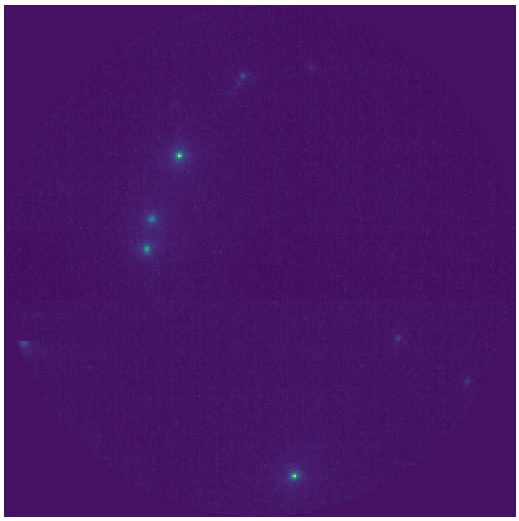
The effect of the entire pipeline (from original, to binarised, to median filtered) is shown in figure 5.29.

5.3.2.7 Clustering

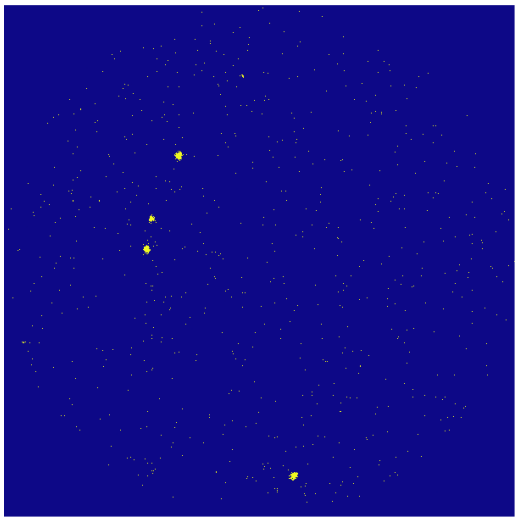
The goal of this analysis is to confront separately the number of events on the three measurement classes with the control class: therefore, a fast and reliable algorithm to determine and count the events is fundamental.

Since the various clusters have very different sizes, point densities, and shapes, I have chosen HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) [161] as clustering algorithm available in the *scikit-learn* Python package [162]. HDBSCAN is an extension of DBSCAN, which interprets as individual cluster areas with a high density of positive pixels surrounded by zones of low pixel density. The generality of the definition of a cluster in DBSCAN makes it suitable for the classification of clusters with any shape. In short, everything revolves around the concept of *core samples*, defined as pixels that have at least `min-samples` pixels at a

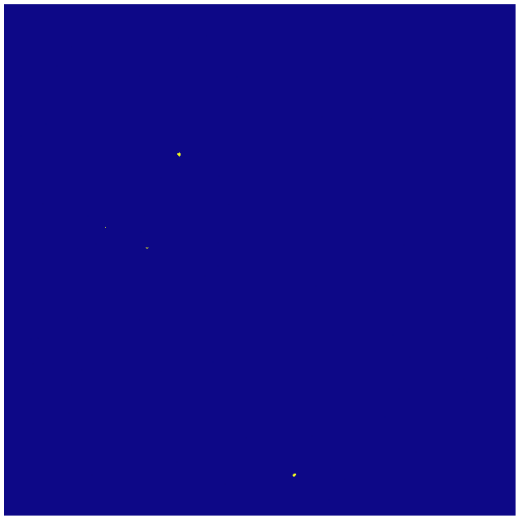
[11] The median filter `kernel-size` takes only odd values (3, 5, 7, etc.)



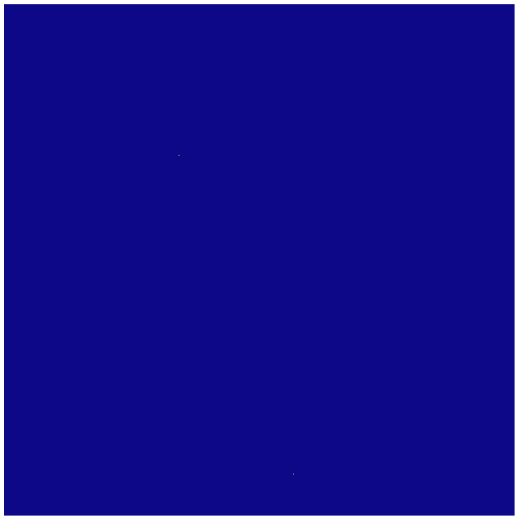
(a) Pre-filtered image.



(b) After binarisation with threshold = 0.3.

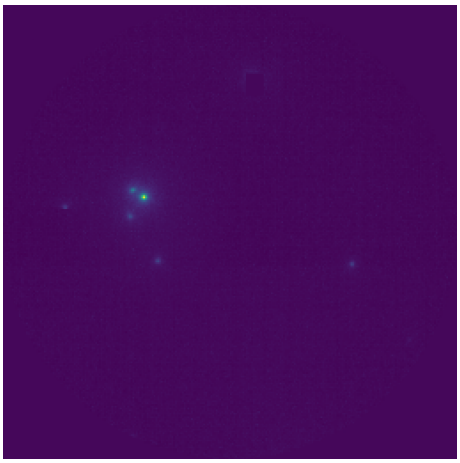


(c) After binarisation with threshold = 0.8.

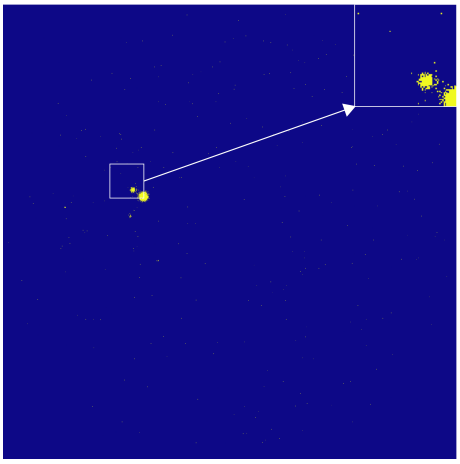


(d) After binarisation with threshold = 1.3.

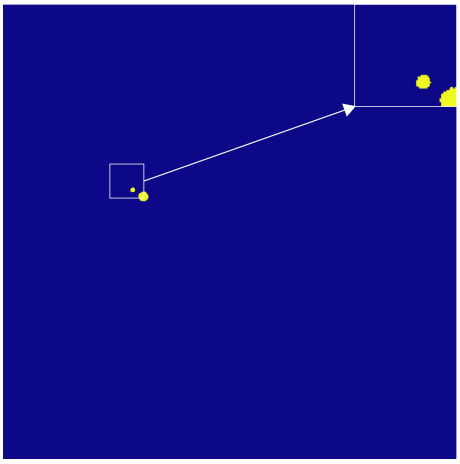
Figure 5.28 – Example of the effect of binarising the MCP image (original in top-left) using different thresholds (0.3, 0.8 and 1.3).



(a) Pre-filtered image.



(b) After binarisation with threshold = 0.5. In the upper-right inset a zoom over the central region is displayed, showing the presence of little noise dots.



(c) After binarisation and median filtering with kernel = 5. In the upper-right inset, a zoom over the same central region as above is displayed, now devoid of noise dots.

Figure 5.29 – Example of the effect of the filtering pipeline: on top of the original image, in the middle of the image after the binarisation with a threshold = 0.5, and at the bottom after the subsequent application of median filter with kernel = 5. In the upper-right insets of the second and third image the zoom over the same central region is displayed, showing the efficacy of the median filtering in removing the noise dots.

distance of `eps` or less (which are called *neighbours*). A cluster is then recursively built by including all the *core samples* neighbouring each other, plus all the other pixels which are neighbours of at least one of the *core samples* belonging to the cluster. The weak spot of this algorithm is that the density of the clusters (represented by `eps`) needs to be homogenous across all the clusters to be identified.

HDBSCAN solves this issue by using an algorithm equivalent to performing DBSCAN over multiple values of `eps`. In this way, it is more robust over identifying clusters with different densities, at the price of increasing the computation time required.

In figure 5.30 some examples of the functioning of this clustering are shown: it can be seen that it correctly isolates and identifies the various clusters present in the pictures, regardless of their size or density.

The number of clusters per image is then counted, and all the results are summed up, so to arrive at the total number of events reconstructed per each measurement class and the control class. These values are used to test, per each measurement class, the probability that the events are in the same number with respect to the control class: a deviation from this equality would point towards an excess of antiproton annihilations on the detector, corresponding to a production of antihydrogen.

5.3.2.8 Binomial test

To evaluate the probability of having formed antihydrogen, a binomial test is performed, in a similar manner to what is done in the scintillators analysis (see *Significance tests* (5.3.1.5)).

As mentioned before in *Image binarisation* (5.3.2.5), the threshold used for the binarisation of the images is a priori an arbitrary parameter; the same holds for the `kernel-size` value used for the median filtering[12].

Therefore a double scan over their value has been carried out: from 0.3 to 2.5 for the binarisation threshold, and over (0,3,5,7,9) for the `kernel-size` of the median filtering. In figure 5.31 the results of the threshold scan are visible, with the median filtering executed with `kernel-size` of 5: it was the value that visually gave the most convincing filtering, and indeed this was confirmed by the significance tests, bearing the highest results (see the following). The scans over threshold values from data filtered with the other values of `kernel-size` are visible in $\bar{H}$ *formation analysis using MCP images - Additional graphs* (B.4).

The values of the number of clusters individuated are subsequently used to perform the binary test of antihydrogen production, performing a one-tail binomial test, i.e. calculating the probability that there are more events in each measurement class than in the control class. The probability is then transformed into standard deviations by the use of an inverse cumulative distribution function Φ^{-1} (similarly to what is done in *Significance tests* (5.3.1.5)). In formulas

$$S = \Phi^{-1}\left(1 - \sum_{k=N_{\text{on}}}^{N_{\text{tot}}} \binom{N_{\text{tot}}}{k} p^k (1-p)^{N_{\text{tot}}-k}\right) \tag{5.3.7}$$

where $N_{\text{tot}} = N_{\text{on}} + N_{\text{off}}$ and $p = 0.5$.

An exact value of 0 was given to the significance in case $N_{\text{tot}} = 0$ or if it was giving a negative result (meaning that there were more counts in the control class than in the measurement one).

In figure 5.32 the results for the `kernel-size` $= 5$ is displayed (the other plots are visible in $\bar{H}$ *formation analysis using MCP images - Additional graphs* (B.4)).

It can be seen that an indication of the presence of an $\bar{\text{H}}$ signal is present in the Forward case, but it is very low.

[12] In principle, also the `min-cluster-size` and `min-samples` of HDBSCAN are arbitrary parameters, but they have been fixed with a logic: they have been set equal to `kernel-size`, being that the minimum size for a spot to survive the median filtering. Still, scans over small ranges of these parameters have been performed, and no substantial difference in the results has been found.

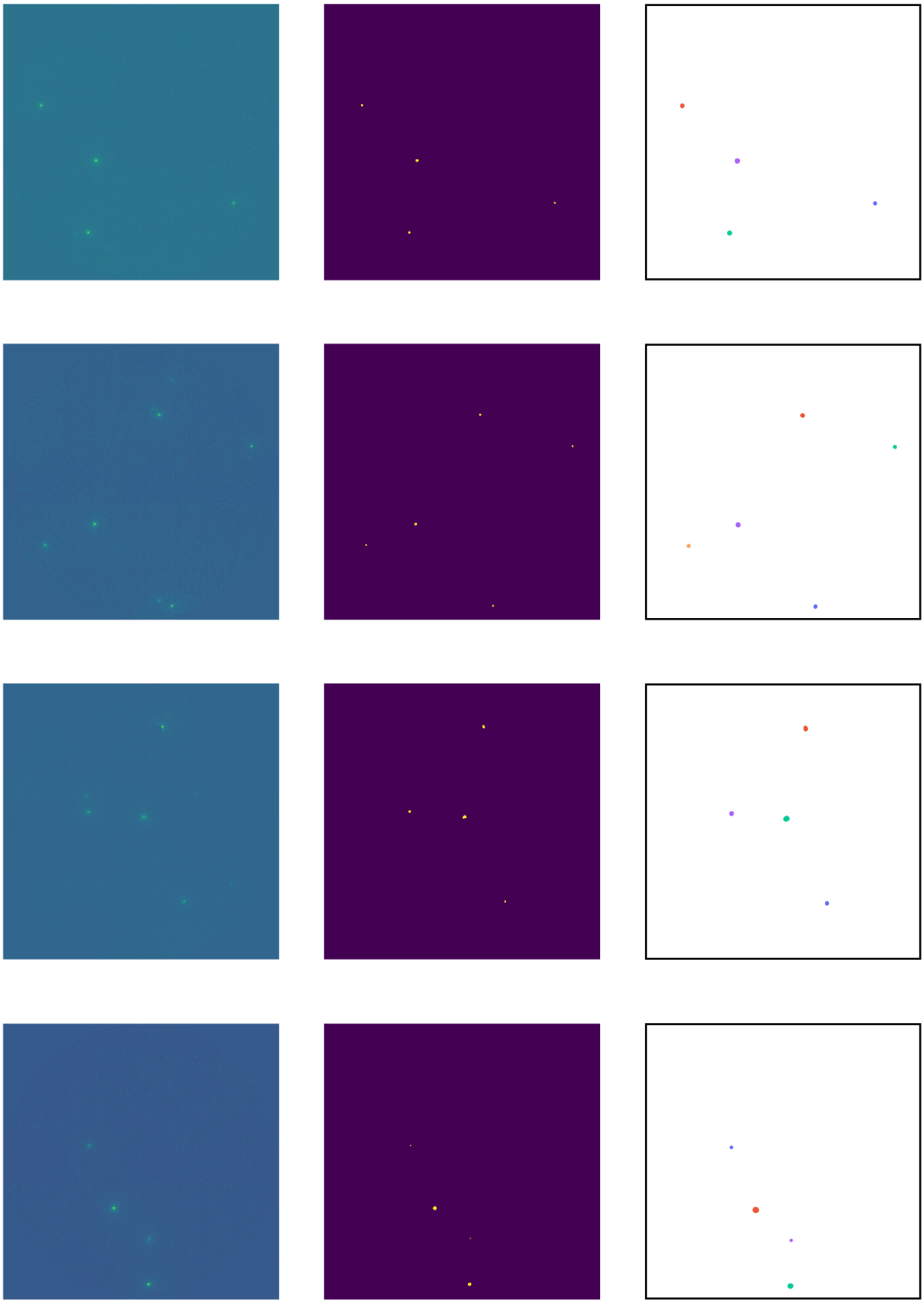


Figure 5.30 – Series of MCP pictures to show the results of the HDBSCAN clustering algorithm: on the left the original images are shown, in the centre the result of the filtering is displayed (binarisation threshold = 0.4, `kernel-size` = 5), and on the right the corresponding clusters identified is presented.

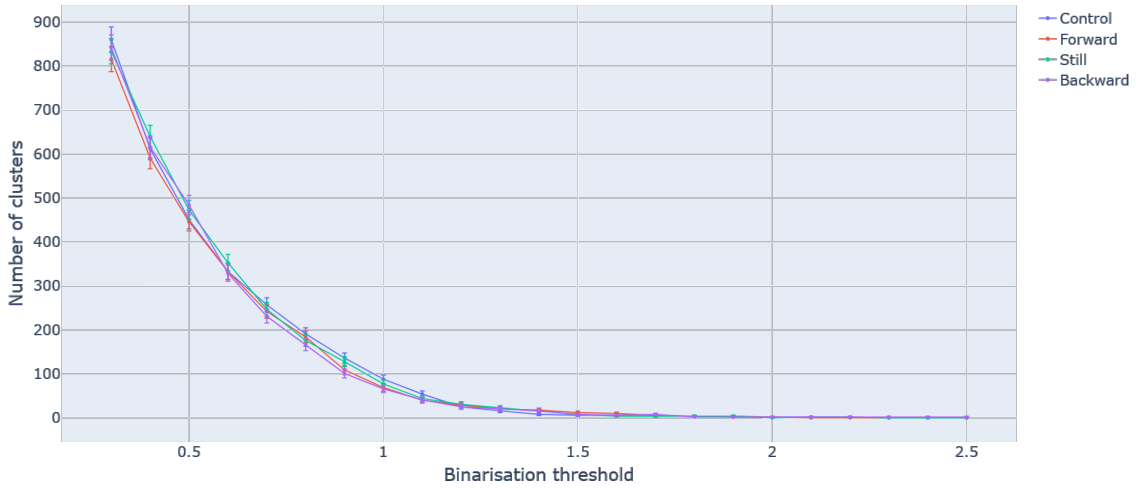


Figure 5.31 – The total number of clusters derived combining the results of all the images of each measurement class ("Backward", "Still", and "Forward") and the control class, varying the threshold used for binarisation. The median filtering was performed with `kernel-size` = 5.

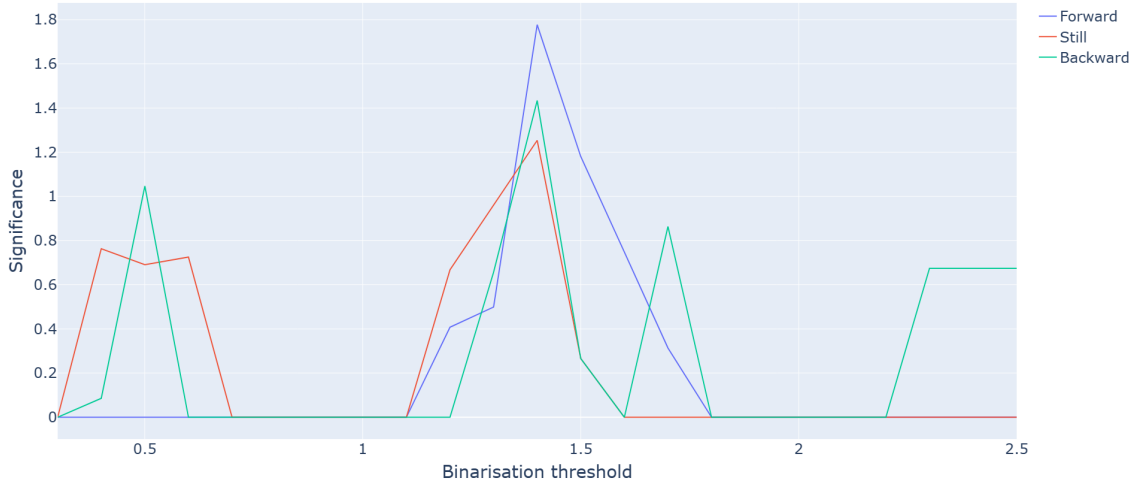


Figure 5.32 – The significance of having successfully produced antihydrogen resulting from the binary test for the three measurement classes ("Backward", "Still", and "Forward") against the control class, varying the threshold used for binarisation. The median filtering was performed with `kernel-size` = 5.

Following what was found in the antiprotons background study performed with the scintillators (see *Charge deposited filtering* (5.3.1.4)), also in this analysis I chose to re-evaluate the entire pipelines restricting the Runs to the one before November 6th, before the possible change in ELENA configuration (the "Before" dataset selection).

The graphs displaying the results of the number of clusters versus the binarisation threshold (with `kernel-size` = 5), and the corresponding significance test derived, restricting the Runs to the "Before" case, are displayed in figures 5.33 and 5.34 (again, the analysis have been carried out also for different values of `kernel-size`: these plots are available in $\bar{H}$ *formation analysis using MCP images - Additional graphs* (B.4)). A discussion of these results is given in *Analyses results and discussion* (5.3.2.9).

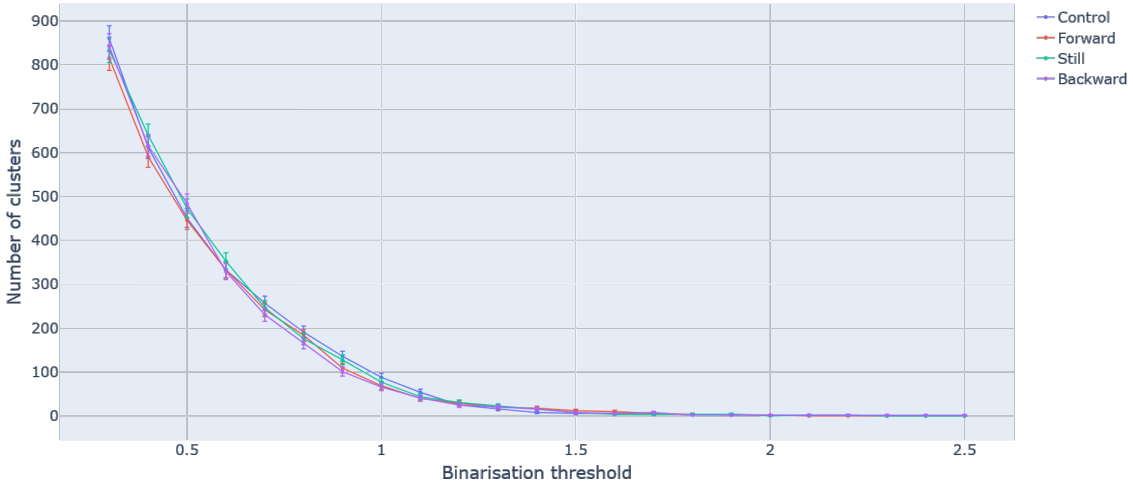


Figure 5.33 – The total number of clusters derived combining the results of all the images of each measurement class ("Backward", "Still", and "Forward") and the control class, varying the threshold used for binarisation, restricting the dataset to the Runs before Nov. 6th. The median filtering was performed with `kernel-size` = 5.

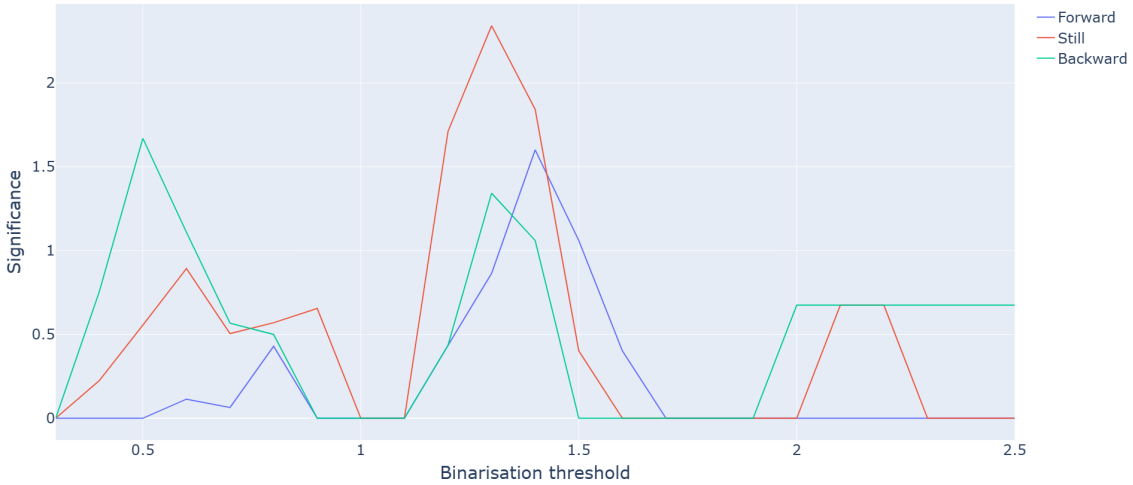


Figure 5.34 – The significance of having successfully produced antihydrogen resulting from the binary test for the three measurement classes ("Backward", "Still", and "Forward") against the control class, varying the threshold used for binarisation, restricting the dataset to the Runs before Nov. 6th. The median filtering was performed with `kernel-size` = 5.

5.3.2.9 Analyses results and discussion

The analysis I have performed for determining the formation of antihydrogen using the images from the MCP has been a completely new development, and therefore a thorough discussion of the method and the results obtained is mandatory.

The starting point is the gating of the phosphor voltage with the Pockels cell, which has been the key enabler of this analysis: without it, the images produced by the MCP would have been useless.

The subsequent camera gain equalisation and camera background subtraction are standard procedures (aside from the algorithm for determining the non-settable gain of the camera).

The definition of a region of interest inside the ionisation grid is fundamental to both speed up the analysis time – by reducing by a factor of 5 the number of pixels to be processed –, and to remove possible noise coming from a region of the sensor that is poorly reachable by the Hbar particles, for geometric reasons. Nevertheless, a simulation should be carried out in the future to verify with precision this last assumption.

The removal of the two almost omnipresent dots described in *Images region-of-interest selection* (5.3.2.3) has been executed to purge a clear source of noise (its cause is still to be determined with precision): the loss of detector surface is negligible, but the choice is still questionable, albeit conservative. The source of this noise has to be investigated, to resolve and/or mitigate it.

The exclusion of the fault images has been demonstrated to be prime towards reaching meaningful results: the inclusion of a single image with an elevated number of bright spots inside a measurement class can mistakenly boost the significance up to 7 standard deviations. Therefore, a conservative approach is mandatory in this respect. The process that I developed is already satisfactory because it lowered to ~ 10 % the number of images to be afterwards visually inspected for rejection, but it needs to be better automated in the future, in order to diminish further the need for user input, both for reasons of users workload (in 2024, the antihydrogen data taking campaign will likely produce 10 times more data) and to lower the potential bias introduced by human decisions.

The binarisation of the images is the key step around which the entire analysis revolves, but also the more delicate, given the arbitrariness of the threshold. From both results (5.32 and 5.34) we can see that in the $1.3 \div 1.5$ range of the binarisation threshold value there is a moderate excess, but it is not found to be significant.

The median filtering has helped remove the salt-and-pepper noise arising from the binarisation of the images: a `kernel-size` value of 5 has demonstrated to be most effective in this type of noise removal. Even if, in the "Before" dataset, a slightly higher significance value has arisen in the case of no median filtering with respect to where it was executed (see *$\bar{H}$ formation analysis using MCP images - Additional graphs* (B.4)), I consider the application of such a filter a better approach, since the resulting numbers of clusters are less prone to mistakenly count noise for signal.

The choice of HDBSCAN as a clustering algorithm has been motivated by its capability of discerning clusters of points independently from the cluster shape (elongated, curved, concave) and density; also the number of clusters expected per image does not have to be specified a priori and can vary. It has indeed shown remarkably good capabilities of correctly discerning clusters (see figure 5.30). The main drawback of this clustering algorithm lies in its time consumption: in case of low binarization thresholds, with a consequent high number of clusters to be distinguished, the clustering of a single image can take many seconds. Mitigation in this respect can be utilised (e.g. parallelisation of the clustering with a multi-threaded code, utilisation of a more powerful workstation or cloud farm), but also other algorithms can be explored (e.g. connectivity-based [163], OPTICS [164]).

Nevertheless, the results of this analysis on the formation of antihydrogen during the 2023 campaign are not conclusive: in the "All" dataset, basically, no real signal evidence is present; in the "Before" dataset, an indication is present only in the Still configuration. This is not in contradiction with the results obtained with the scintillators' data (see *Significance tests* (5.3.1.5)),

where also the "Before" dataset yielded higher significances with respect to the "All" one. It is also not surprising not to see the evidence, present in the scintillators analysis, of a formation in the "Before" Backward case: the MCP detector is unfavoured in revealing $\bar{\mathrm{H}}$ formed with momentum in the opposite direction.

In general, one of the main issues – for all the different analyses used – toward precisely determining the formation of antihydrogen has been the limited statistics acquired. One culprit, in this respect, has been a vacuum leak extremely difficult to find, which has been localised and temporarily fixed only in the summer of 2023, enabling less than three months of data taking. The faulty component was substituted last winter, so a nominal length of data taking (i.e., at least 6 months) is foreseen for 2024. In addition, most of 2023 antiproton beam time has been devoted to the development of the antihydrogen formation procedures, which obviously will not be repeated. Therefore, for the antihydrogen campaign of 2024, at least an order of magnitude increase in data-taking time is foreseen.

To precisely determine the formation of a forward-boosted beam of antihydrogen, further development is needed in this respect, both to eliminate noise at the source and to render this analysis pipeline more robust (exploring the various possibilities outlined before).

In this respect, the elimination of the noise given by gating the voltage of the phosphor of the MCP with the Pockels cell has been the key enabler of the image analysis: with a better fine-tuning, the signal acceptance time-window can be further shrunk to 10 µs or less, improving the signal-to-noise ratio of more than one order of magnitude.

Even better, the substitution of the MCP with an imaging sensor capable of distinguishing the signal of an antiproton from the one of an electron would virtually eliminate the noise issue. In fact, the main pollution of the images was caused by different sources of electrons, and the impossibility of separating the two signals has led to vague analysis results. In this respect, such a detector is currently in development inside the AEgIS collaboration [7].

Regardless, the analysis pipeline described in this section is the first fundamental development toward a reliable indicator for the formation of a forward-boosted beam of antihydrogen.

Chapter 6

Conclusions

The matter/antimatter unbalance in the visible universe is one of the greatest open puzzles in modern physics. The consensus of the majority of physicists places the solution in the existence of a process that, in the first instants following the Big Bang, has progressively generated the asymmetry, which has been maintained since then.

In the last decades, experiments have progressively constrained the possibility for such a process to arise from the physics of the Standard Model: therefore, the necessity for the existence of new physics to solve this problem is increasingly likely.

Out of the multiple theoretical frameworks envisaged to answer this question, a possibility put forward lays the source of the imbalance in a different gravitational effective interaction between matter and antimatter. To verify such a statement, tests of the Einstein Equivalence Principle on antimatter are being pursued by several experiments. Among the most promising experiments investigating this possibility is AE$\bar{g}$IS, whose main goal is the measurement of the vertical fall of a beam of antihydrogen in the Earth's gravitational field.

My doctoral studies have taken place indeed working for the AE$\bar{g}$IS experiment, during extremely unusual and "interesting" times. In fact, I began my PhD in 2020, when the COVID-19 pandemic was still a significant issue, and this rendered the start not easy (I literally worked the first six months remotely). Subsequently, in February 2022, the Russian invasion of Ukraine and the following resurgence of a new cold war completely upset the world's geopolitical balance, with the practical effect of having impaired the procurement of several hardware components (the most impacted have been the new positron source). Thankfully, I was granted – both from the University of Trento and from CERN – an extension of 6 months to my doctoral contract, with which I managed to properly conduct and terminate my studies.

During the years of my PhD studies, the AE$\bar{g}$IS experiment went through a period of great renovation. As a matter of fact, 2020 marked the beginning of the transition from Phase 1 of the experiment, where the main result was the demonstration of the technique chosen to form antihydrogen, to Phase 2, where the first gravitational study with an antihydrogen beam will be attempted. This shift has been characterised by several hardware and software upgrades to the apparatus, to improve the components that revealed being underperforming. These modifications encompassed multiple sectors: first of all the formation scheme itself, passing from the perpendicular to the collinear illumination of the antiproton plasma with positronium atoms; in turn, this caused the complete redesign and assembly of a new formation trap. The electronics were exchanged, passing from custom-made hardware to a modular system from the Sinara ecosystem; the positron line was reworked, and the lasers improved.

The most salient and drastic change has been the new control system software, which has been my biggest technical contribution to the apparatus upgrade. The new control system, CIRCUS, has been written anew, from scratch, basing it on the asynchronous modular formalism of the Actor Model. It has the capability of unifying all the computers of the experiment in a unique entity,

enabling the possibility of centralised control of the apparatus, and ensuring a consistent response to errors and unforeseen events that are the base for its reliability and a high degree of automation. The last feature has completely changed the approach to the control of the experiment, from a continuously manned control room with shifters to extended periods of unsupervised operations. Moreover, it has enabled the ability of autonomous parameters optimisation via live feedback from the data acquired, which has further lowered the scientists' workload, and rendered possible experiments unreachable before (like the world-first laser cooling of positronium atoms).

In addition, CIRCUS has been developed to be experiment agnostic (mainly targeting nuclear, atomic, and quantum experiments), and it has been released open-source, so to empower experiments other than AEgIS.

Furthermore, my push towards the embrace of the *agile* development paradigms, coupled with the modular formalism of the actors and precise version control with *Git*, has enormously increased the software development speed in AEgIS, also guaranteeing a better interaction among developers, fostered by the standardisation brought by CIRCUS.

My physics contribution to the experiment started with the discussions and the developments of the multiple procedures necessary to arrive at the formation of antihydrogen in the context of the renewed apparatus. This has both strengthened my knowledge of low-energy particle physics and plasma manipulations in traps (complementing my theoretical background in high-energy particle physics), and it has challenged me to propose solutions for experimental problems that I (and often, anyone in the AEgIS team) never faced.

The work on the development of procedures, and my knowledge of the novel control system, have put me at the forefront of the data-taking team. The data-taking campaigns have been particularly stressful, mainly due to their shortness caused by the delay accumulated because of the aforementioned world conditions (COVID-19 pandemic and war in Ukraine), and by some major hardware faults. Nevertheless, each year we managed to conclude the campaigns successfully achieving the predetermined goals: in 2021 we managed to capture antiprotons from the new decelerator (ELENA), validating the design of the new degrader and the novel control system software and electronics; in 2022 we set the record for the efficiency of trapping the incoming antiprotons; in 2023, we have good indications of having produced again antihydrogen in the new formation trap, using the ballistic antiprotons transfer method, the key step towards the formation of a forward-boosted beam of antihydrogen atoms.

My other main physics contribution to the experiment has been in the data analysis. Notwithstanding the little experience and the limited time, I managed to construct and refine an analysis pipeline that utilises the images taken with the MCP positioned at the end of the traps system to evaluate the probability of having formed antihydrogen. It was the first time that, in AEgIS, such an analysis was tempted: before, the detection of antihydrogen formation had relied upon the scintillators' signals only. The pipeline wandered from gain equalisation to image selection, binarisation, and filtering, to spot clusterisation and, finally, significance tests varying the binarisation threshold, the most arbitrary parameter of the analysis. The resulting evidence was not strong enough to claim a successful formation: still, the indication of the susceptibility of such a detector to noise was made clear, and ameliorations for the future were put forward. In particular, this analysis will be the starting point of the one used for the determination of the gravitational acceleration of the beam of antihydrogen, being the gravimeter signal based on a novel imaging detector currently in development inside the AEgIS collaboration [7].

In future prospects, to enhance the yield of the antihydrogen formation process, multiple improvements are already foreseen. The first is a more performant EKSPLA laser crystal, to bring the maximum positronium Rydberg level excitable from less than 25 to more than 32, impacting the formation by at least a factor of 3. Then, the new positrons source that, due to the current geopolitical situation, suffered more than a year of delay, not arriving on time to be installed in 2023: its impact is a factor of 8 increase in $\bar{H}$ formation cross-section. Most importantly, a new $e^+ \rightarrow Ps$ converter, whose partial fault (whose causes are still under investigation) almost

jeopardised the formation results, having more than an order of magnitude decreased efficiency. All these improvements will be carried out in 2024, and therefore the nominal production for Phase 2 (i.e., more than 1 $\bar{\text{H}}$/min) could be achieved.

To further improve the formation of a forward-boosted beam of antihydrogen, the development of a $e^+ \rightarrow$Ps transmission target would be an ideal evolution of the actual hardware. In fact, in the current ballistic transfer scheme, the resulting momentum of the antihydrogen atoms is defined by the shape of the potential used for transferring the antiprotons: to form a forward-boosted beam, then, the velocity of the antiprotons, when it gets illuminated by the positronium atoms, has to be directed towards the gravimeter. As a result, in the current production scheme, the two velocities have opposite signs. If a transmission target will be manufactured, the two reactants would have both velocities in the same direction: being the antihydrogen formation cross-section $\sigma_{\overline{H}} \propto |v_{\overline{p}} - v_{Ps}|^{-2}$, and us capable of tuning the antiproton velocity at will – by changing the shape of the parabolic transfer potential –, the formation rate in the forward-boosted mode could be raised by orders of magnitude. In AE$\bar{\text{g}}$IS, studies have already been performed to produce such a target [159].

For what concerns the analyses, while the determination of the antihydrogen formation can be successfully concluded by studying the signal of the scintillators, the measurement of the deviation of the antihydrogen beam has to pass through an analysis of images: consequently, an improvement of the analysis that I have developed during my doctoral studies is paramount towards an efficient determination of the result.

In conclusion, the technical and scientific advancements to the AE$\bar{\text{g}}$IS experiment performed in the last years, to which I have contributed substantially during my doctoral studies, have led to the formation of antihydrogen atoms with the novel ballistic antiprotons transfer, a key enabler for producing a forward-boosted beam of neutral antihydrogen. These developments pave the way towards the direct measurement of the gravitational interaction of antimatter with a precision smaller than 1 %, which ultimately may shed some light on the great mystery of the universe matter-antimatter asymmetry.

Appendix A

Additional results

In this chapter I briefly present the two main results obtained by the other two AEgIS lines of research, parallel to the antihydrogen line: positronium studies, and highly-charged ions formation.
I have given a substantial contribution to both results: in particular, for the positronium result I have coded most of the software interfaces with the hardware used for the e^+/Ps manipulations, and set up the feedback loop for increasing the laser timing accuracy; for the highly-charged ions formation, I have developed most of the routines used for the particles manipulations, performed the major part of the data taking, and contributed to the data analysis. Moreover, in both studies, I have taken active participation in the physics discussions.

Most importantly, the CIRCUS control system I developed during my PhD has been one of the key enablers of both results, thanks to its innovative features of automation and data feedback.

In the following, the two results are explained: for the Ps studies, the relative article published is reported here *verbatim*, and for the highly-charged ions, an extract from [7] is given.

A.1 Positronium laser cooling

In 2023 AEgIS managed to perform, for the first time in the world, the laser cooling of positronium atoms. This achievement, long-awaited in the positronium community, will open up many doors, from more precise spectroscopy to WEP tests with a purely leptonic system (by studying the redshift caused by the different gravitational potential exerted by the Sun in different points along Earth's orbit), arriving to the creation of positronium atoms Bose-Einstein condensate, potentially leading to a coherent emission similar to a gamma-ray laser.

To achieve these results, multiple fundamental components have been necessary, the most important being:

- the $e^+ \rightarrow$Ps reflection target able to produce positronium atom at room temperature (see *$e^+ \rightarrow$Ps converter* (2.1.12))
- the alexandrite laser (see *Alexandrite Laser ("Alex")* (2.1.6.2))
- the coordination and automatic laser temporal stabilisation given by CIRCUS (see *The CIRCUS* (3) and *In-Run autonomous parameter stabilisation* (5.1.4))

In the following, the article describing the result is reported, as published on *Physical Review Letters*.

Positronium laser cooling via the 1^3S–2^3P transition with a broadband laser pulse

(PRL Editor's Suggestion)

L. T. Glöggler, N. Gusakova, B. Rienäcker, A. Camper, R. Caravita, S. Huck, M. Volponi, T. Wolz, L. Penasa, V. Krumins, F. P. Gustafsson, D. Comparat, M. Auzins, B. Bergmann, P. Burian, R. S. Brusa, F. Castelli, G. Cerchiari, R. Ciuryło, G. Consolati, M. Doser, Ł. Graczykowski, M. Grosbart, F. Guatieri, S. Haider, M. A. Janik, G. Kasprowicz, G. Khatri, Ł. Kłosowski, G. Kornakov, L. Lappo, A. Linek, J. Malamant, S. Mariazzi, V. Petracek, M. Piwiński, S. Pospisil, L. Povolo, F. Prelz, S. A. Rangwala, T. Rauschendorfer, B. S. Rawat, V. Rodin, O. M. Røhne, H. Sandaker, P. Smolyanskiy, T. Sowiński, D. Tefelski, T. Vafeiadis, C. P. Welsch, M. Zawada, J. Zielinski, N. Zurlo (the AEgIS Collaboration)

Abstract:
We report on laser cooling of a large fraction of positronium (Ps) in free-flight by strongly saturating the 1^3S–2^3P transition with a broadband, long-pulsed 243 nm alexandrite laser. The ground state Ps cloud is produced in a magnetic and electric field-free environment. We observe two different laser-induced effects. The first effect is an increase in the number of atoms in the ground state after the time Ps has spent in the long-lived 2^3P states. The second effect is one-dimensional Doppler cooling of Ps, reducing the cloud's temperature from 380(20) K to 170(20) K. We demonstrate a 58(9) % increase in the fraction of Ps atoms with $v < 3.7\times10^4\,\mathrm{m\,s^{-1}}$.

Positronium (Ps), discovered in 1951, is the lightest known atomic system, consisting only of an electron (e^-) and a positron (e^+) [165]. Ps has been extensively studied for its exotic properties as a purely leptonic matter-antimatter system. So far, experiments researching Ps have relied on formation processes that result in clouds with a large velocity distribution, in the order of several $10^4\,\mathrm{m\,s^{-1}}$ [166, 130, 167]. This, for instance, has been limiting the precision of spectroscopy studies due to the large Doppler broadening of the transition lines [168, 169]. The idea of using laser cooling to narrow the Ps velocity distribution dates back to 1988 [170], following the first demonstration of laser cooling on neutral atoms by just a few years [171]. Despite significant efforts [172], Ps laser cooling has not been experimentally achieved yet. A whole range of new fundamental experiments would become feasible with a sufficient amount of cold Ps [173, 174]. These include 1^3S–2^3S precision spectroscopy at the 100 kHz level, which will enable testing bound state QED at the $\alpha^7\,m_{e^+}$ order [175], measuring the m_{e^+}/m_{e^-} mass ratio with unprecedented accuracy [176], and testing the Equivalence Principle (EP) with a purely leptonic system by looking at the transition red-shift around the Sun's orbit [177]. Testing the EP with atomic systems consisting of antimatter is the primary goal of AEgIS, building on the availability of cold Ps for efficient Rydberg antihydrogen ($\overline{\mathrm{H}}^*$) production through the charge exchange reaction $\mathrm{Ps}^* + \overline{\mathrm{p}} \rightarrow \overline{\mathrm{H}}^* + \mathrm{e}^-$ between cold antiprotons ($\overline{\mathrm{p}}$) and Rydberg-excited Ps (Ps*) [178, 15]. This reaction's efficiency can be significantly increased by reducing the temperature of the Ps* cloud [95]. Moreover, forming a Ps Bose-Einstein condensate (BEC) [179, 173] will allow studying stimulated annihilation, producing coherent light in the gamma radiation range [180, 181]. This objective (plus precision spectroscopy) is being pursued by the UTokyo group [182], which is actively developing Ps laser cooling with a chirped laser pulse [183].

Here, we report on the first experimental demonstration of Ps laser cooling by strongly saturating the 1^3S–2^3P transition (Fig. A.1a) for 70 ns, employing an alexandrite-based laser system developed

specifically to meet these requirements: high intensity, large bandwidth and long pulse duration. The velocity distributions with and without laser cooling were obtained by Doppler-sensitive two-photon resonant ionization along the 1^3S–3^3P transition [130] (Fig. A.1a).

Ps is produced by implanting e^+ bunches with 3.3 keV energy into a nanochanneled silicon converter [127, 167]. This e^+/Ps converter has a Ps yield of about 30 %, corresponding to a few 10^5 Ps per bunch. It is mounted at a 45° angle with respect to the e^+ beam axis, as shown in Fig. A.1b. Typically, both e^+ beam and emitted Ps cloud are about 5 mm in diameter. All experiments are conducted in a magnetic and electric field-free environment. An electrostatic buncher [184] with fast potential switch-off [185] and a mu-metal shield are used to minimize residual fields. A low residual magnetic field, measured to be below 1 mT in the Ps production area, is important towards Ps laser cooling, as in intermediate magnetic field ranges the saturation of the 1^3S–2^3P transition leads to fast annihilation due to singlet-triplet state mixing in the excited state manifold [186]. This effect, called "magnetic quenching", prevents efficient cooling [187] and strongly encourages to work either in the Paschen-Back regime [188] or, as we do, in a magnetic field-free environment.

The Ps 1^3S–2^3P transition is driven by the third harmonic of a Q-switched alexandrite laser [189], as proposed in [173], whose main features are briefly summarized hereafter. The optical length of the cavity is 1 m. The pulse length of 70 ns, much longer than the spontaneous emission lifetime of the 1^3S–2^3P transition (3.19 ns), allows several cooling cycles per pulse [173]. The central wavelength is set by an intra-cavity Volume Bragg Grating (VBG) [190]. Rotating the VBG finely tunes the fundamental wavelength with an absolute accuracy of 10 pm. Two lithium triborate (LBO) crystals and two beta barium borate (BBO) crystals are used to generate up to 2.3 mJ at the third harmonic. At 243 nm, the measured root-mean-square (rms) spectral bandwidth is $\sigma_{243} = 101(3)$ GHz. The laser [189] was specifically designed to deliver an irradiance of 100 kW cm^{-2} when focusing 0.7 mJ on an area of 10 mm^2. As a result, the power on the 20 MHz rms resonance transition linewidth amounts to 100 kW cm^{-2} × 20 MHz/101 GHz = 20 W cm^{-2}, much higher than the saturation intensity of the 1^3S–2^3P transition of 0.45 W cm^{-2} [187]. The laser fluence fills in the spectral gaps in the laser bandwidth [191] and the population in the excited state is saturated within a 360(15) GHz large spectral bandwidth.

It should be noted that in these conditions less than 1 % of the atoms are photo-ionized [187]. The transverse Doppler profile was probed by fine-tuning the wavelength of a 1.5 ns-long 205 nm pulse with a RMS spectral bandwidth of $\sigma_{205} = 179(9)$ GHz or 25(1) pm, populating the 3^3P states. A 4 ns-long 1064 nm pulse synchronized with the 205 nm pulse induced photo-ionization of the excited states [120].

The Pockels cell of the alexandrite laser cavity is connected to a high-voltage electronic switch, which opens and closes the cavity with nanosecond precision to generate a Q-switched pulse featuring a controllable sharp falling edge. Consequently, the laser emission can be suppressed imminent to the arrival of the 205 nm pulse probing the velocity profile, avoiding a temporal overlap of the cooling and probing laser pulses. The 205 nm pulse interacts with Ps about 12 ns after the 243 nm pulse has subsided. This ensures that the transiently excited Ps have spontaneously decayed to the ground state before probing the velocity distribution of the cloud. Nanosecond synchronizations between laser pulses, HV switch and e^+ were realized by ARTIQ/Sinara control electronics [192] coordinated by a LabVIEW™-based distributed control system [193]. The 243 nm laser beam is co-propagating with the 205 nm (Fig. A.1b) and retro-reflected by a dichroic mirror transmitting the 205 nm light. All laser beams are linearly polarized.

The time distribution of the gamma radiation resulting from Ps annihilations, the so-called Single-Shot Positron Annihilation Lifetime Spectroscopy (SSPALS) [194] spectrum, was acquired in different laser configurations. As illustrated in Fig. A.2, the configurations "no lasers", "205 nm+1064 nm", "243 nm only", and "243 nm+205 nm+1064 nm" were used. A 25 × 25 × 25 mm $PbWO_4$ scintillator placed 40 mm above the e^+/Ps converter, coupled to a Hamamatsu R11265-100 photomultiplier tube and a Teledyne LeCroy HDO4104A oscilloscope, was used to acquire SSPALS spectra [120].

The long tail in the SSPALS spectrum measured without lasers (black dotted curve), extending

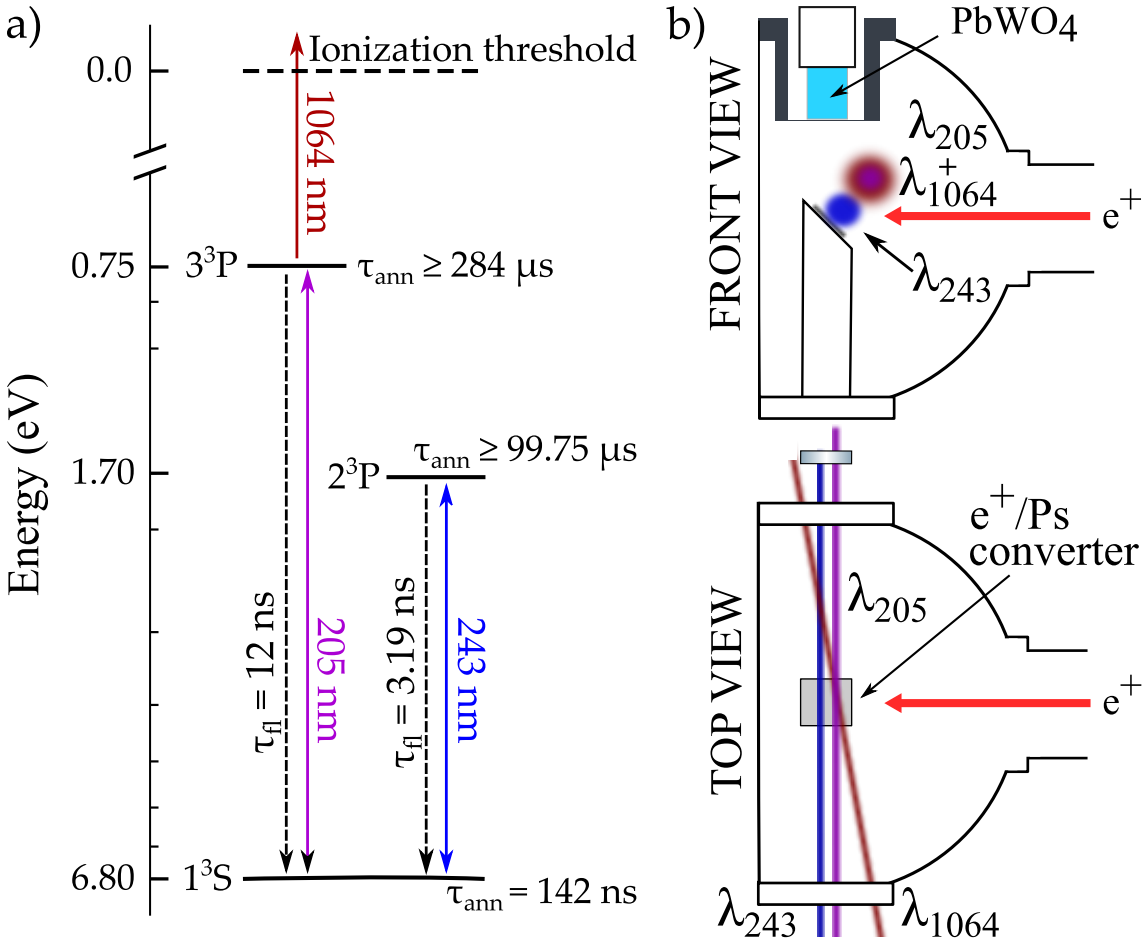


Figure A.1 – Experimental layout for Ps laser cooling: a) diagram of relevant Ps energy levels and transitions, highlighting transition wavelengths and annihilation/fluorescence lifetimes; b) front and top view of the vacuum setup, featuring the cooling (λ_{243}) and the two probing lasers (λ_{205} and λ_{1064}), the $PbWO_4$ detector and the dichroic mirror used to selectively retro-reflect the cooling beam.

from 150 ns to 400 ns in Fig. A.2, reflects the 142 ns lifetime of 1^3S Ps in vacuum. Firing the 243 nm laser only, a large fraction of Ps is excited to the 2^3P level, where the annihilation lifetime is much longer than in the ground state. Consequently, the annihilation rate at later times increases (green dash-dotted line in Fig. A.2) due to an increase in the number of annihilating atoms in the ground state. By sending the 205 nm+1064 nm pulses only, a fraction of the atoms is selectively photo-ionized. This leads to an immediate small increase in the gamma emission as a fraction of the isotropically emitted photo-dissociated e^+ hits the conversion target and annihilates (small bump at 90 ns in Fig. A.2), followed by a reduction of the number of ground state Ps annihilating with 142 ns lifetime (red dashed curve). The interaction of the Ps cloud with all three lasers induces a combined effect (blue solid curve in Fig. A.2).

In order to study the effects caused by the different laser configurations, S-parameters are constructed as:

$$S = \frac{f_{\mathrm{ON}} - f_{\mathrm{OFF}}}{f_{\mathrm{OFF}}}\,, \tag{A.1.1}$$

where f_{ON} and f_{OFF} denote the integrated SSPALS spectra in the time window between 150 ns and 400 ns (gray band in Fig. A.2). "ON" refers to laser(s) interacting with the Ps cloud, and "OFF" to no laser interaction. In the following, we will refer to $S_{205+1064}$ when only the probing lasers are present, and S_{243} when only the cooling laser is present. We further define S_{cool} as the

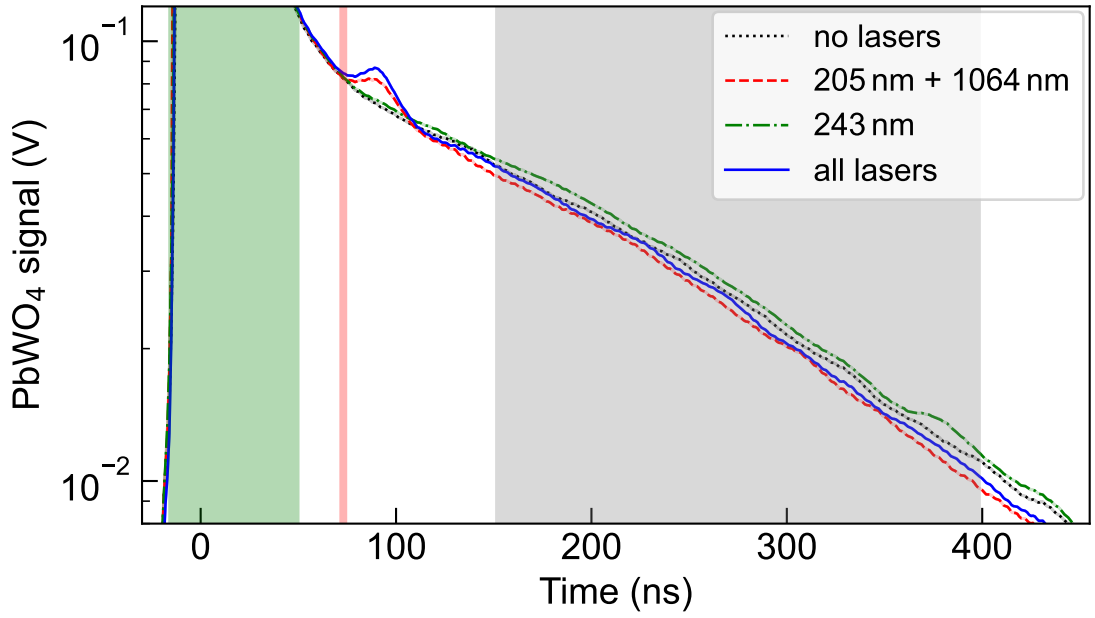


Figure A.2 – SSPALS spectra of Ps in vacuum without lasers (black dotted curve), with the 205 nm + 1064 nm lasers (red dashed curve), with the 243 nm laser only (green dash-dotted curve), and with all three lasers 243 nm + 205 nm + 1064 nm (blue solid curve). The 243 nm laser is firing during the time window from −20 ns to 50 ns (green band), while the 205 nm+1064 nm (red vertical line) are injected 75 ns after e^+ implantation time ($t = 0$ ns). Each curve is an average of 90 individual spectra. The statistical error is smaller than the linewidths. For analysis, the spectra were integrated between 150 ns and 400 ns (light grey area).

difference between $S_{243+205+1064}$ (when all three lasers are present) and S_{243}:

$$S_{\text{cool}} = S_{243+205+1064} - S_{243} \tag{A.1.2}$$

S_{cool} reflects the number of photo-ionized Ps atoms by the probing laser after the interaction with the cooling laser, normalized by the number of Ps atoms annihilating in the absence of any laser. The S-parameter values are calculated by averaging over many sets of spectra acquired consecutively in all the above-mentioned laser configurations. A de-trending procedure is applied [195] to correct for slow changes in the amount of Ps produced over time, caused by moderator ageing during the long measurement period. A trend function is built by applying Gaussian radial basis regression [162] to the f_{OFF} data set. Subsequently, S-parameters are calculated by evaluating the trend function at exactly the time at which the SSPALS spectra with laser(s) are acquired.

First, the Ps velocity distribution without laser cooling was measured by scanning the λ_{205} detuning with the 205 nm and 1064 nm laser beams grazing the surface of the e^+/Ps converter. In this experiment, the laser pulses were fired 50 ns after the e^+ implantation time. The transverse Doppler profile (Fig. A.3a) is fitted with a Gaussian function, yielding a rms-width of 44(1) pm. This width corresponds to a Ps rms-velocity of $5.3(2) \times 10^4\,\text{m}\,\text{s}^{-1}$ after de-convoluting the σ_{205} probing laser bandwidth. This rms-velocity is associated with a transverse temperature of 370(30) K. The resulting line is centred at $\lambda^c_{205} = 205.044(3)$ nm, which is compatible with the theoretical value [120].

Secondly, we performed Saturated Absorption Spectroscopy on the 1^3S–2^3P transition [196] to determine the center of the line and to characterize the effect of the cooling pulse on the SSPALS spectrum. Fig. A.3b) displays the S_{243} parameter calculated from SSPALS spectra, recorded as a function of the λ_{243} detuning. The cooling laser pulse is synchronized with the e^+ implantation time (see Fig. A.2). It is worth noting that the resulting S-values are now positive, in contrast to

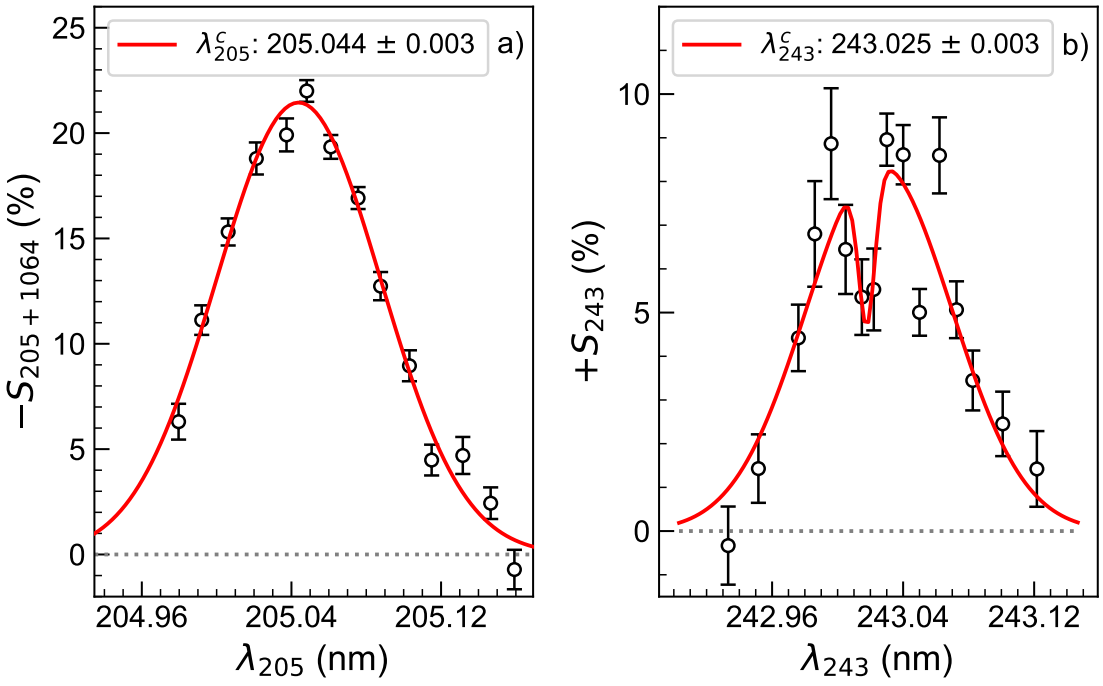


Figure A.3 – Ps velocity distribution measured by SSPALS.
a) Transverse Doppler profile measured by two-photon resonant ionization. A Gaussian fit yields a rms-width of 44(1) pm, which translates to a Ps rms-velocity of $5.3(2)\times10^4\,\mathrm{m\,s^{-1}}$ after de-convoluting the σ_{205} laser bandwidth.
b) Velocity-resolved increase in the number of ground state Ps atoms, induced by the 243 nm transitory excitation to the 2^3P level. At resonance, the expected Lamb dip is observed. A 2-Gaussian fit yields a rms-width of the enveloping Gaussian of 44(3) pm, which corresponds to a Ps rms-velocity of $4.9(4)\times10^4\,\mathrm{m\,s^{-1}}$.

what was observed in the two-photon resonant ionization experiment. To the best of our knowledge, such an increase in the number of ground state Ps atoms caused by a transitory laser excitation to the 2^3P level has never been observed and can be classified as a laser-induced, spectrally tunable preservation of Ps. This effect has the same physical origin as the one observed in the lifetime enhancement of Ps atoms excited to Rydberg-states [31].
The observed line shape shows a Lamb dip, demonstrating the saturation of the 1^3S–2^3P transition by the laser [196]. A 2-Gaussian model is fitted to the data, following the approach taken in Ref. [196]. The transition line is centred at $\lambda^c_{243} = 243.025(3)$ nm, which is in agreement with previous measurements [196]. The enveloping Gaussian features a rms-width of 44(3) pm, which corresponds to a Ps rms-velocity of $4.9(4)\times10^4\,\mathrm{m\,s^{-1}}$ (320(50) K) after de-convoluting the σ_{243} cooling laser bandwidth.

With this understanding of the individual laser interactions with the Ps cloud, we then performed experiments combining the 243 nm cooling laser and the 205 nm+1064 nm probing lasers. The cooling laser remains synchronized with the e^+ implantation time and in the same spatial position (Fig. A.1b, blue spot). The probing laser pulse is delayed by 75 ns with respect to the positron implantation time [185, 130] (red vertical line in Fig. A.2) and moved to a position at a distance of 7 mm from the converter surface (Fig. A.1b, violet spot) corresponding to the distance covered by the atoms in the peak-velocity component of the axial velocity distribution during 70 ns. These parameters are chosen to reach a $S_{205+1064}$ of $\sim 10\%$ with the probing laser tuned at resonance. To characterize the change in the Ps velocity distribution induced by the cooling laser, the detuning of the 243 nm laser is set to −200 GHz (corresponding to $\lambda_{243} = 243.061$ nm) and a photo-ionization Doppler scan is performed. The S_{cool} parameter measured as a function of the detuning of the

probing laser is shown in Fig. A.4. The curve is compared to the $S_{205+1064}$ distribution measured in the same configuration (75 ns delay and 7 mm away from the e^+/Ps converter), but without prior interaction with the cooling laser. Both of the one-dimensional transverse Doppler profiles were obtained by applying a moving average to the ~ 350 single S-values with a square window (350 GHz in width).

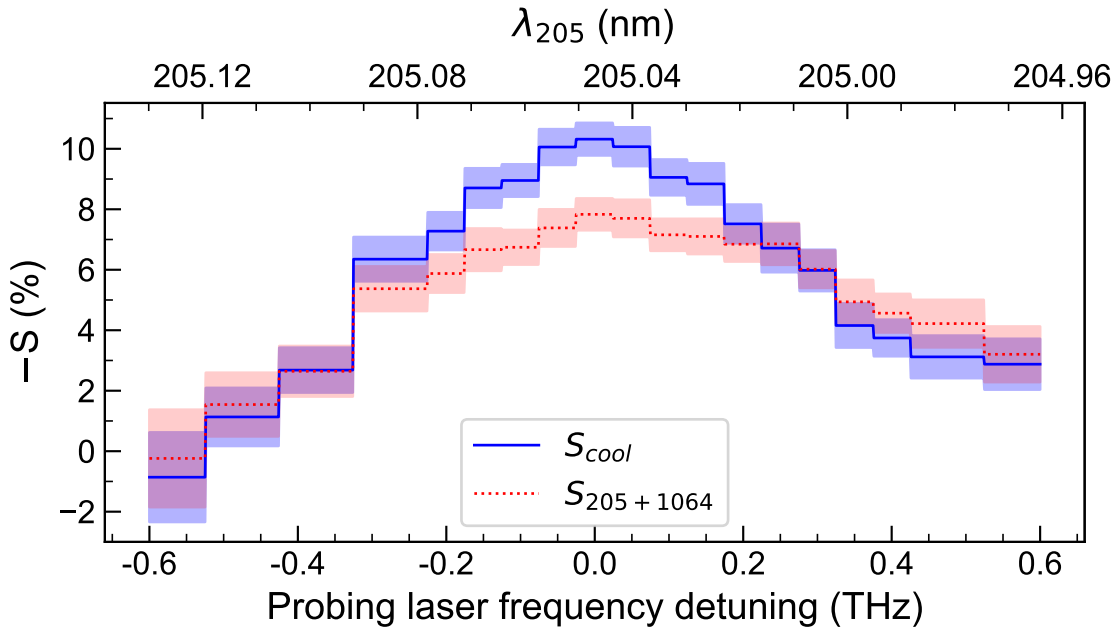


Figure A.4 – One-dimensional transverse Doppler profiles of the Ps cloud with (solid curve), and without (dotted curve) interaction with the 243 nm cooling laser beam at a fixed frequency detuning of −200 GHz. The semi-transparent bands represent the statistical measurement error (one standard deviation of the average).

The one-dimensional transverse Doppler profile obtained in the presence of the 243 nm cooling laser is narrower than the one measured without it. The asymmetry of the two profiles is caused by a slight increase in the pulse energy of the 205 nm probing laser toward blue-detuned wavelengths. A simple Gaussian fit on each of the two distributions was used to quantify the change in the velocity profile. With cooling, we find a rms-width of 269(1) GHz, in contrast to 330(2) GHz without cooling. After de-convoluting the standard deviation of the moving average window (350 GHz/$\sqrt{12}$) and the σ_{205} laser bandwidth, the Ps rms-velocities corresponding to these widths are $5.4(2) \times 10^4\,\mathrm{m\,s^{-1}}$, associated with a temperature of 380(20) K, and $3.7(2) \times 10^4\,\mathrm{m\,s^{-1}}$ associated with 170(20) K, respectively. The obtained rms-velocity in the absence of the cooling laser is in agreement with the results reported in Fig. A.3. The interaction with the 70 ns-long 243 nm laser pulse reduces the Ps rms-velocity by $1.7(3) \times 10^4\,\mathrm{m\,s^{-1}}$, corresponding to a temperature reduction of $\Delta T = 210(30)$ K. The systematic error associated with the arbitrary choice of a Gaussian fitting model is estimated to be ±10 K.

Given the high optical intensity of the 243 nm laser, the average time for all addressed Ps atoms to undergo a single cooling cycle is 6.38 ns [173]. Consequently, a maximum of 11 cooling cycles can be reached within the 70 ns-long laser–Ps interaction. Since the recoil velocity for a single 1^3S–2^3P transition of Ps is $v_{\mathrm{recoil}} = 1.5 \times 10^3\,\mathrm{m\,s^{-1}}$ [187], the velocity reduction can reach $11 \times v_{\mathrm{recoil}} = 1.65 \times 10^4\,\mathrm{m\,s^{-1}}$, corresponding to a temperature reduction of about 200 K, in agreement with our measurements.

To evaluate the maximum fraction of fast Ps that can be pushed toward null velocity via recoil effect, the cooling laser detuning was scanned from −0.65 THz to 0.25 THz while the 205 nm laser remained at resonance. The result of this scan is shown in Fig. A.5. The horizontal dashed line is

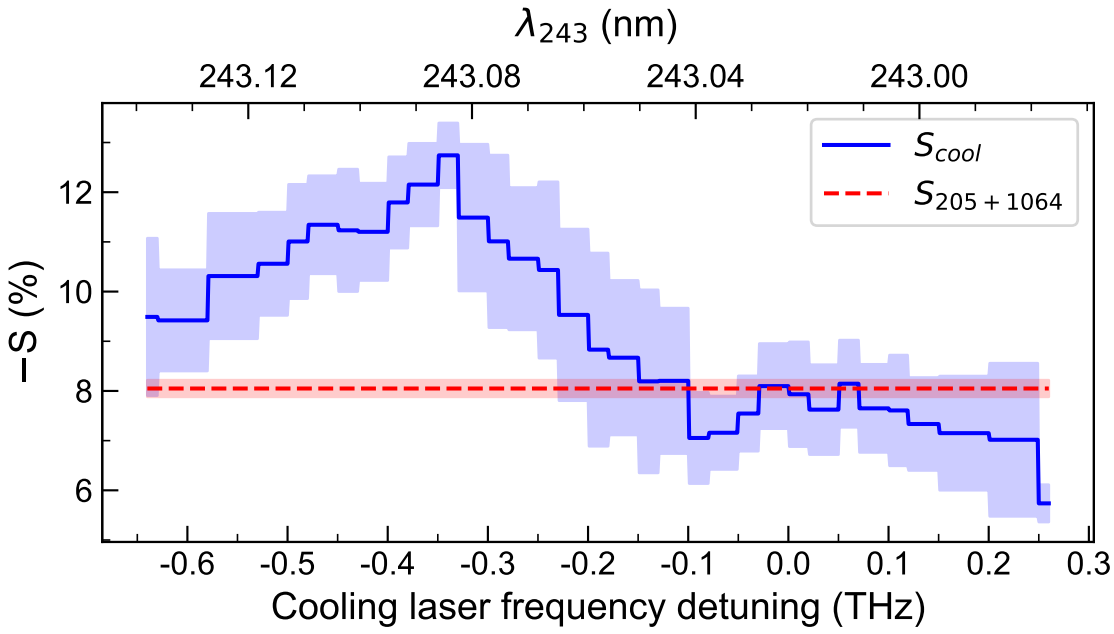


Figure A.5 – Number of Ps atoms with $v < 3.7 \times 10^4\,\mathrm{m\,s^{-1}}$, as a function of the cooling laser frequency detuning, normalized to the number of Ps atoms in the absence of all lasers. The dashed horizontal line represents the reference population of Ps in this velocity range with the cooling laser off. The highest observed relative increase is 58(9) % at a cooling laser frequency detuning of −350 GHz. The semi-transparent bands represent the statistical uncertainties (one standard deviation of the average).

the signal measured when only the probing laser interacts with the Ps cloud ($S_{205+1064}$), yielding $-S = 8.0(2)\,\%$ as the reference for the population near resonance. The blue curve is the S_{cool} parameter as defined in Eq. A.1.2. The curve was obtained by applying a moving average with a window size of 200 GHz. For a given λ_{243} detuning, the difference between S_{cool} and $S_{205+1064}$ corresponds to the fraction of Ps atoms cooled within the bandwidth of the 205 nm laser i.e., having velocities smaller than $3.7 \times 10^4\,\mathrm{m\,s^{-1}}$. The $-S$ value at -200 GHz detuning in Fig. A.5 is compatible with that of Fig. A.4 at resonance ($9.2 \pm 1.8\,\%$ and $10.4 \pm 0.5\,\%$, respectively). We find a maximum relative increase of 58(9) % at a detuning of −350 GHz.

In conclusion, we have experimentally demonstrated laser cooling of a large fraction of a thermal Ps cloud in a magnetic and electric field-free environment and reduced the velocity rms by $11 \times v_{\mathrm{recoil}} = 1.65 \times 10^4\,\mathrm{m\,s^{-1}}$ in 70 ns, which is the limit of the efficiency allowed by standard Doppler cooling. A temperature decrease from 380(20) K to 170(20) K was observed. Our study also gives an in-depth understanding of the different laser–Ps interactions and their manifestation in the SSPALS spectra. In particular, we observed an increase in the number of ground state atoms after Ps has been transiently excited to the 2^3P states. Consequently, this cooling method has the unique feature of delaying annihilation, which allows us to preserve a larger number of Ps atoms while cooling the ensemble. Furthermore, the estimate of the cooling laser intensity suggests that cooling is driven in the strongly saturated regime. Starting from a colder source at 150 K [197] and adding a second cooling stage with a narrower spectral bandwidth set closer to resonance, will allow to reach the recoil velocity in 22 cooling cycles (∼140 ns). Alternatively, coherent laser cooling [198, 199] may be adapted to the positronium case. Ps laser cooling opens the door to an entirely new range of important fundamental studies, including precision spectroscopy, Bose-Einstein condensation of antimatter and tests of the Equivalence Principle with a purely leptonic matter-antimatter system.

A.2 Formation of trapped cold Highly Charged Ions

The formation of antiprotonic atoms has so far only been studied in beam-on-bulk target experiments. At AEḡIS we are developing methods for the controlled formation of antiprotonic atoms in vacuum, by co-trapping negative ions with cold $\bar{p}$ and triggering the $\bar{p}$ capture through laser photo-detachment followed by Rydberg excitation of the atom. This controlled process enables the triggered charge-exchange of Rydberg atoms with the antiproton into a specific Rydberg orbit, facilitating precision studies of antiprotonic atoms within the trap. Furthermore, the antiprotonic atom cascade and subsequent annihilation of the antiproton on the nucleus will result in exotic nuclear fragments stripped of electrons, inside the trap.

Simulations performed using GEANT4 [200] showed that a significant fraction of these highly charged fragments can be captured and trapped for further nuclear structure studies. The novel formation mechanism and capture of these highly charged fragments could pave the way for new studies complementary to those at existing radioactive beam facilities. In order to facilitate these measurements multiple developments have taken place as discussed in *The AEḡIS Experiment* (2).

The initial suboptimal vacuum conditions at the beginning of 2023 beamtime constrained activities that needed extended $\bar{p}$ storage times to mere tens of seconds. However, this limitation inadvertently gave a unique opportunity for developing new ion techniques using the positive ions generated from the annihilation of $\bar{p}$ on the rest gas inside the trap (as schematised in figure A.6).

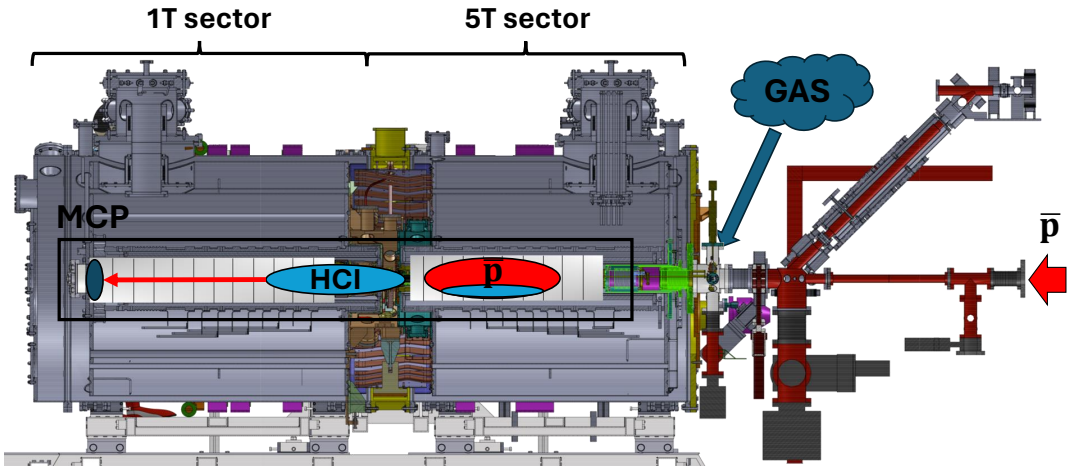


Figure A.6 – Scheme depicting the formation of Highly Charged Ions in the AEḡIS apparatus in the antiproton campaign of 2024.

A new nested trap procedure was developed for capturing and manipulating the resulting positive ions formed from the interaction with trapped $\bar{p}$. A set of new procedures for storing, manipulating, and identifying ions formed from the interaction with low energy $\bar{p}$ was developed (see figure A.7), before the vacuum problem was fixed, removing the air leak into the system. To identify the charge state and species of the resulting ions a time-of-flight (TOF) technique was used by triggering the release of ions in the 5 T trapping region towards the downstream 1 T MCP: an example signal obtained is given in figure A.8. The charge-to-mass ratio can then be extracted as

$$\frac{M}{Q} = \frac{2Vt^2}{L^2} \tag{A.2.1}$$

where V is the potential of the floor of the trapped ions, t is the travel time and L is the distance between the centre of the launch trap and the MCP.

The TOF technique was later calibrated using trapped $\bar{p}$, followed by a controlled measurement injecting pure nitrogen gas into the AEḡIS via a feedthrough installed in the vacuum jacket. These

opportunistic studies offered new insights into $\bar{p}$ interactions with atoms at low energies within the trapping region, facilitating the establishment of protocols at AE$\bar{g}$IS for co-trapping highly charged ions with $\bar{p}$. This advancement was crucial for forthcoming research on nuclear structure and antiprotonic atoms at AE$\bar{g}$IS.

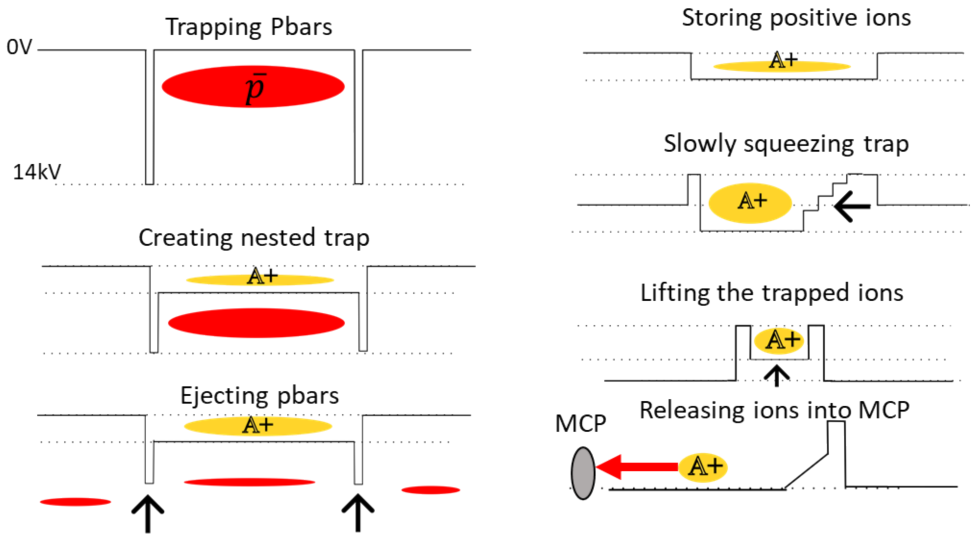


Figure A.7 – Steps used to form the Highly Charged Ions in the AE$\bar{g}$IS apparatus. First, antiprotons from ELENA were trapped; subsequently, a nested trap was created, to capture the positive ions formed. The antiprotons were then ejected, and the captured ions squeezed into a shorter trap. Afterwards, the floor of the trap was lifted, and then one of the two end-caps pulse-opened, to release the ions towards the MCP and perform the time-of-flight measurement.

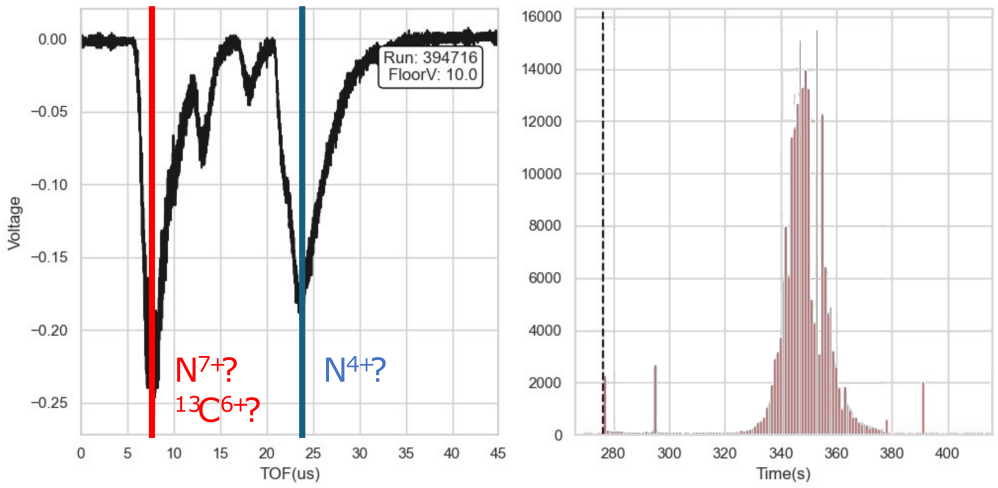


Figure A.8 – On the left: example of a time-of-flight measurement, where the peaks are tentatively identified thanks to their charge-to-mass ratio. On the right: antiproton annihilations monitored during the measurement.

Appendix B

Additional materials

B.1 Abbreviations and acronyms

AD	Antiproton Decelerator
ADC	Analog-to-Digital Converter
AEGIS	Antimatter Experiment: Gravity, Interferometry, Spectroscopy
AERIALIST	Antimatter Experiment Realtime Integration of Artiq LIbraries and Sinara Technology
ALPHA	Antihydrogen Laser PHysics Apparatus
ARTIQ	Advanced Real-Time Infrastructure for Quantum physics
ASACUSA	Atomic Spectroscopy And Collisions Using Slow Antiprotons
C	Charge (symmetry)
CAD	Computer Assisted Design
CI/CD	Continuous Integration/Continuous Development
CERN	European Organization for Nuclear Research
CIRCUS	Computer Interface for Reliably Controlling, in an Unsupervised manner, Scientific experiments
CKM	Cabibbo-Kobayashi-Maskawa (matrix)
CMB	Cosmic Microwave Background
CP	Charge-Parity (symmetry)
CPT	Charge-Parity-Time (symmetry)
DAC	Digital-to-Analog Converter
DAQ	Data AcQuisition system
DM	Dark Matter
DMA	Direct Memory Access
DRTIO	Distributed Real Time Input/Output
ELENA	Extra Low ENergy Antiproton (ring)
EW	ElectroWeak
FACT	Fast Annihilation Cryogenic Tracker
FC	Faraday Cap
FEM	Finite-Elements Model
FIFO	First-In-First-Out
FPGA	Field-Programmable Gate Array
FWHM	Full Width at Half Maximum
GUT	Grand Unification Theory
GXML	G eXtensible Markup Language
HV	High Voltage
IMB	Irvine-Michigan-Brookhaven (detector)
I/O	Input/Output

IR	InfraRed
LAN	Local Area Network
LabVIEW™	Laboratory Virtual Instrument Engineering Workbench
LEO	Loew Earth Orbit
LLR	Lunar Laser Ranging
MCP	Micro-Channel Plate
OS	Operating System
OVC	Outer Vacuum Chamber
PMNS	Pontecorvo-Maki-Nakagawa-Sakata (matrix)
PMT	Photo-Multiplier Tube
PPI	Pixels Per Inch
PXI	Peripheral Component Interconnect (PCI) eXtensions for Instrumentation
QDC	Charge (Q) Digital Converter
RF	Radio-Frequency
ROI	Region-Of-Interest
RMS	Root Mean Square
RT	Real-Time
rTPC	Radial Time-Projection Chamber
RW	Rotating Wall
SBlocks	Schedule Blocks
SCP	Secure Copy Protocol
SM	Standard Model (of particle physics)
SNR	Signal-to-Noise Ratio
TALOS	Total Automation of Labview Operations for Science
ToF/TOF	Time-Of-Flight
TPC	Time-Projection Chamber
T.R.	Tassilo Rauschendorfer
UFF	Universality of Free-Fall
UHV	Ultra-High Vacuum
UV	UltraViolet
VBG	Volume Bragg Grating
VME	Versa-Module Eurocard

B.2 The most important metric

CIRCUS and TALOS have been very great endeavours, in which both Jakub Zielinski and I have put an immense amount of hours coding and testing the software in LabVIEW™. Transmitting the effort to our colleagues, though, was not trivial, since we could not express the progress into the usual "number of lines of code" written, as is customary for textual programming languages.

Nevertheless, we come across a solution: the *WireLength Calculator* [201]. This program opens all VIs of a project (or inside a specific folder) and measures the length of all the wires traced inside the *Block Diagrams*, finally returning the overall sum. In this way, we could both better get a sense of the amount of work done, and easily communicate the progress during periodic Collaboration Meeting (and, most importantly, have a laugh!).

The original code of the *WireLength Calculator* has been slightly modified, to fix a known bug (inserting a factor of 10 error in the calculation) and to insert the possibility to specify the pixel-per-inch (PPI) of the screen used for the measurement[1]. The new version has been published in the original source thread [201].

In table B.2, the latest measurements for both the AEgIS CIRCUS and the TALOS projects are given. The metres of wires have been calculated using a 24" screen with Full HD resolution (1920 x 1080).

	Number of VIs	Pixels of wires	Metres of wires
TALOS	939	5.5 M	1518 m
AEgIS CIRCUS	801	5.3 M	1478 m
Total	1740	10.8 M	3.00 km

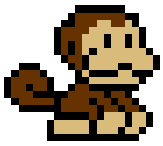
Table B.2 – Table summurising the most important metric, i.e. the total kilometres of wires traced, for both the AEgIS CIRCUS and the TALOS projects, calculated with an updated version of the *WireLength Calculator* [201].

Thanks Piotr Demski for the coding the LabVIEW™ WireLength Calculator!

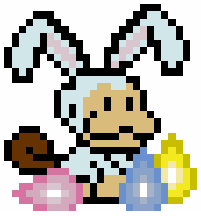
[1]The code measures the total number of pixels traced: to get the metres of wires, a conversion is performed using the PPI of the screen.

B.3 Artworks

During my three and a half years of PhD in AEgIS, together with my partner-in-crime Jakub Zieliński[2], we produced many pixelated artworks to make the CIRCUS more beautiful, and to bring out some laughs. Some of them are thematic and were used only on specific occasions (e.g. Christmas). Here I gathered a collection of the most beautiful pieces we produced.



(a) Standard Monkey

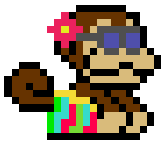


(b) Easter Monkey

(c) Christmas Monkey

(d) Halloween Monkey

(e) Darth Monkey (May 4th)

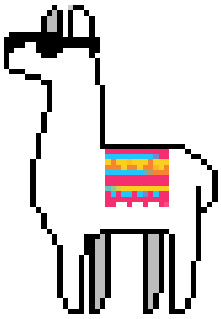


(f) Summer Monkey

Figure B.1 – The Monkey, with multiple skins.

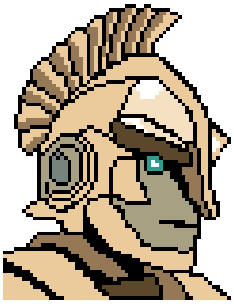
Figure B.2 – The Tamer.

[2]Who actually introduced me to this type of art, and made most of the drawings. Thanks, Kuba!



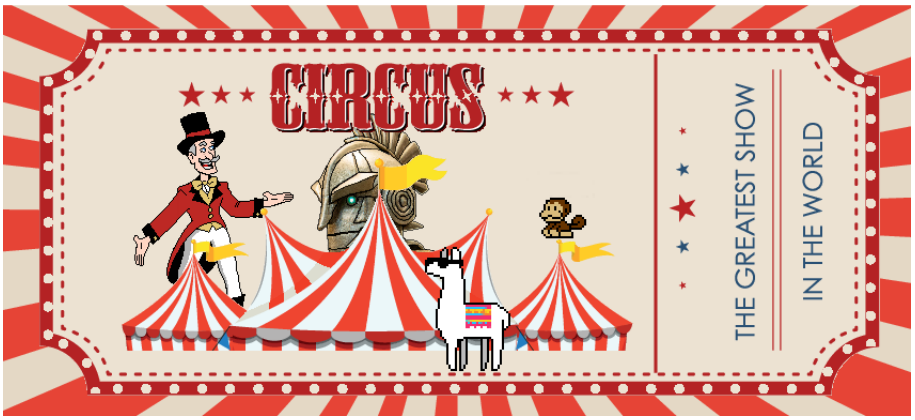
(a) Alpaca.

(b) The Aerialist.



(c) TALOS.

Figure B.3 – Other characters.



(a) A ticket for the CIRCUS.

(b) "Ladies and gentlemen, please come in!"

Figure B.4 – TALOS loading screens.

(a) Standard CIRCUS logo.

(b) AEḡIS CIRCUS logo.

(c) AEḡIS CIRCUS logo, Halloween version.

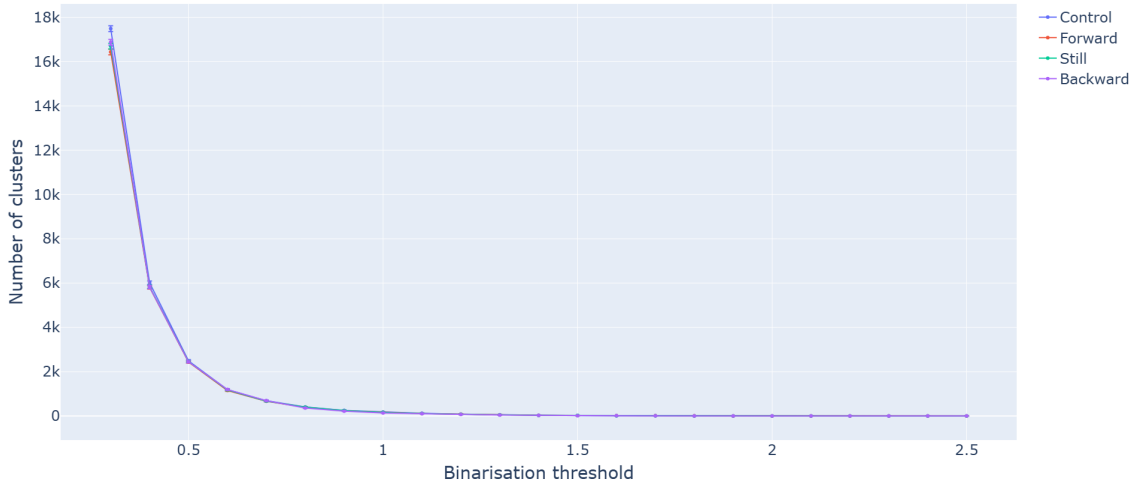
(d) AEḡIS CIRCUS logo, Christmas version.

Figure B.5 – Various CIRCUS logi.

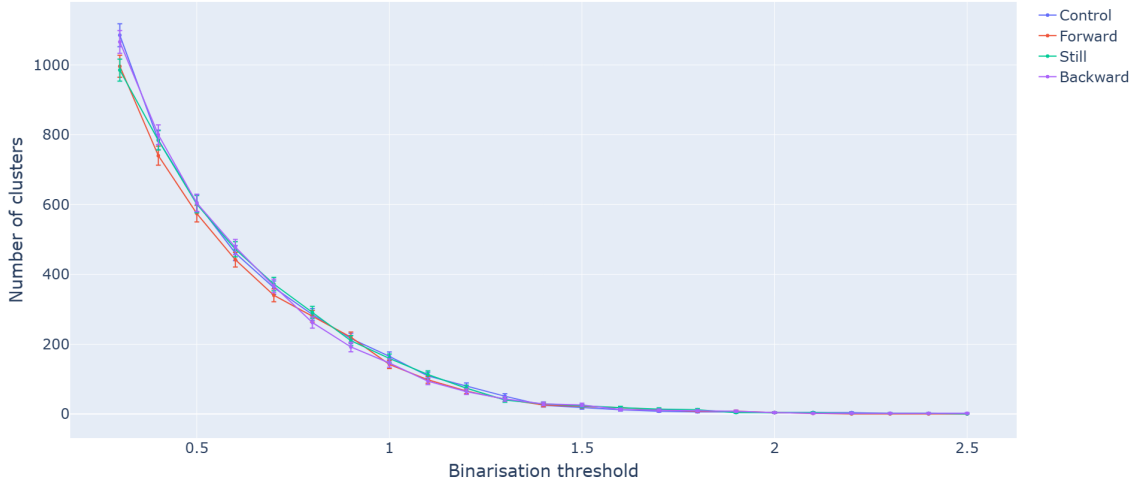
B.4 $\bar{H}$ formation analysis using MCP images - Additional graphs

As presented in *$\bar{H}$ formation analyses using MCP images* (5.3.2), the variation of the number of clusters determined by the analysis pipeline has been assessed over the difference in binarization threshold utilised. This is the parameter giving the most arbitrariness in the entire analysis.

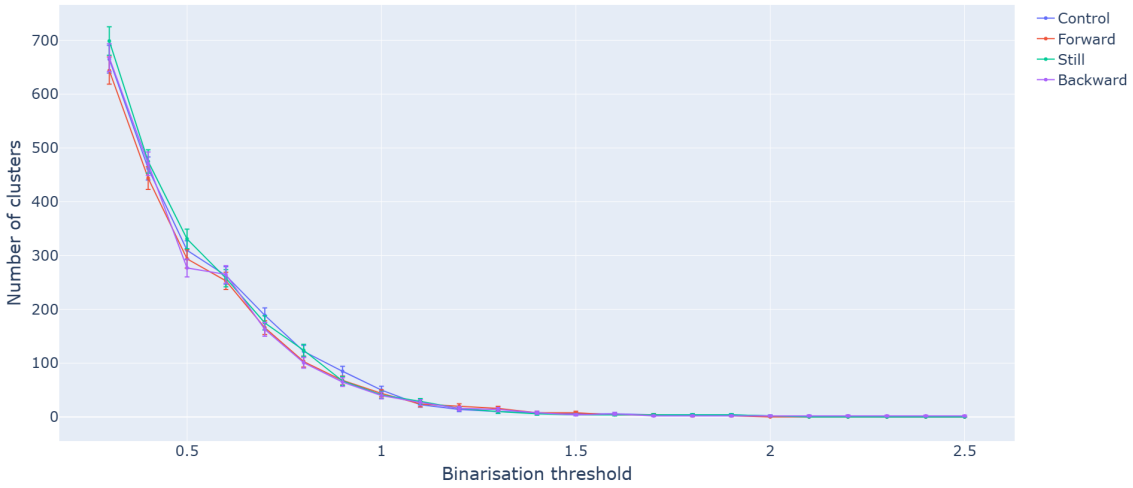
In addition, scans over different `kernel-size` values for the median filtering have also been performed, being another source of possible arbitrariness. The results are presented in this appendix (except for `kernel-size` $= 5$, given in *$\bar{H}$ formation analyses using MCP images* (5.3.2)). The best visual results are given by the value of 5, effectively removing all the noise spots smaller than 3 pixels: this is confirmed by the significance plots. The scan included the avoidance of median filtering, which actually gave the highest significance value in the "Before" dataset: despite the seemingly good result, I preferred the application of the median filtering since the very high amount of clusters found when not using the filter is an indication of the insertion of a lot of noise in the data, leading to smaller robustness of the algorithm.



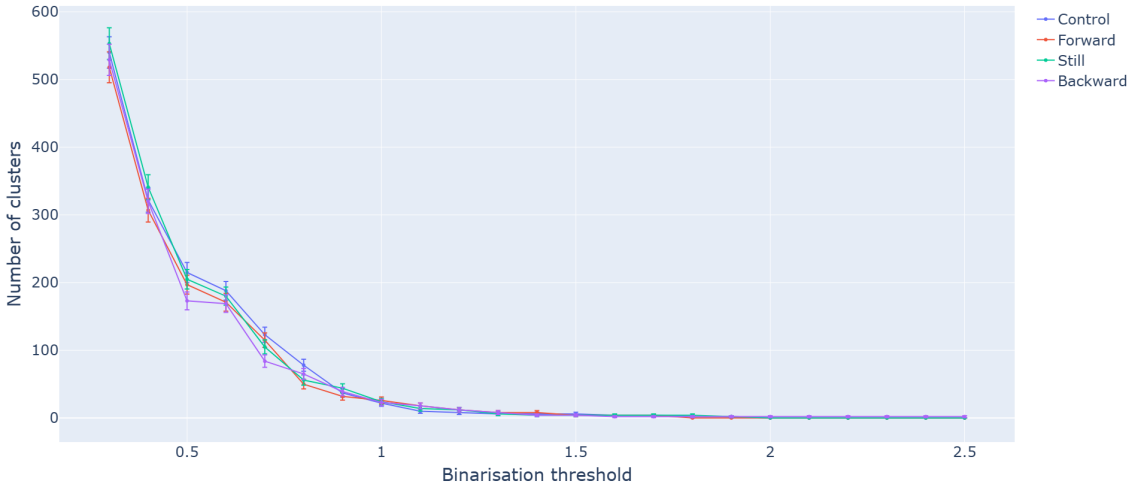
(a) Clusters vs. Threshold scan for no median filtering.



(b) Clusters vs. Threshold scan for `kernel-size` $= 3$.

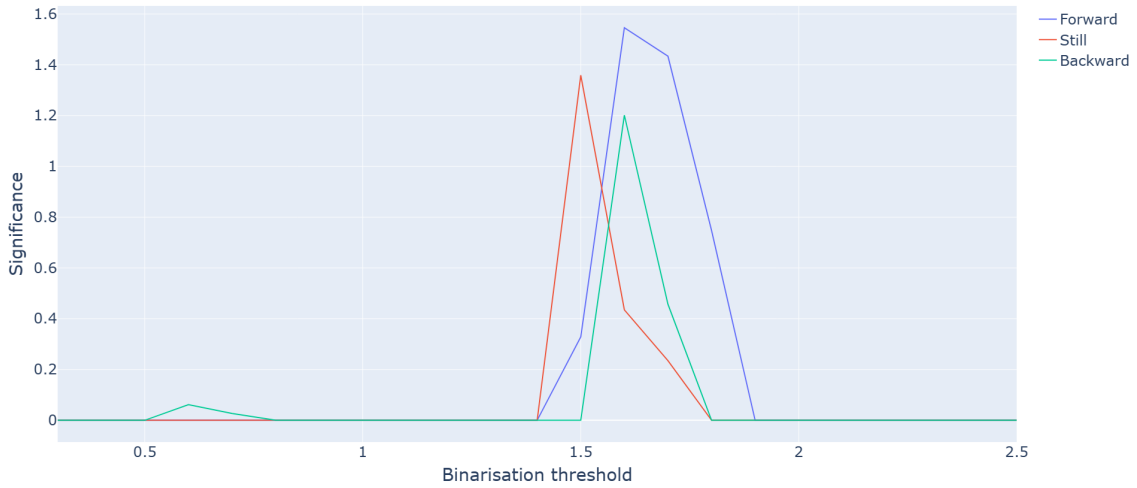


(c) Clusters vs. Threshold scan for `kernel-size` = 7.

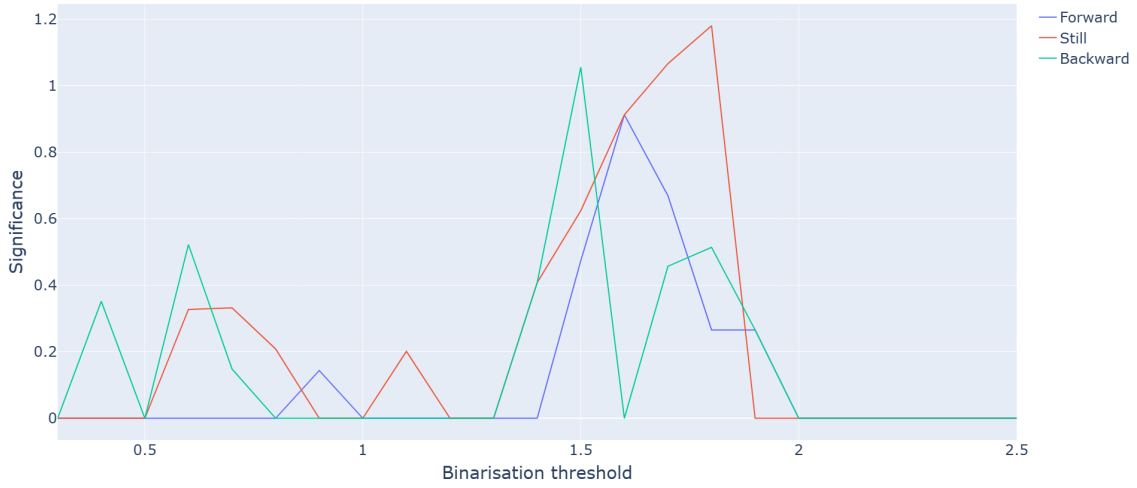


(d) Clusters vs. Threshold scan for `kernel-size` = 9.

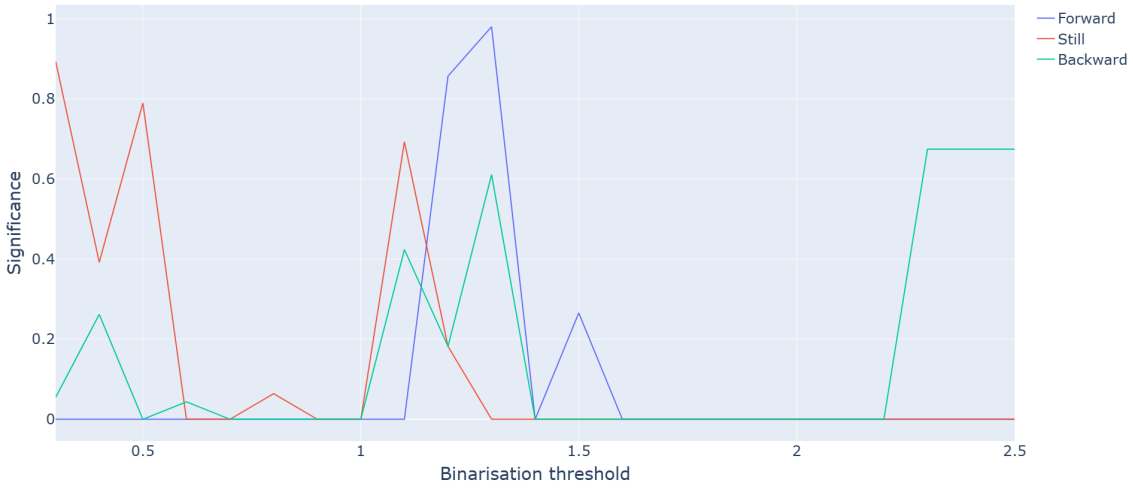
Figure B.6 – Plots showing the results of scanning the number of clusters found on all the images of every measurement and control class (Forward, Still, Normal, and Control) over the binarisation threshold. In the pipeline of the upper one, no median filtering was performed, while in the subsequent ones, a median filtering with `kernel-size` $= 3, 7, 9$ is present.



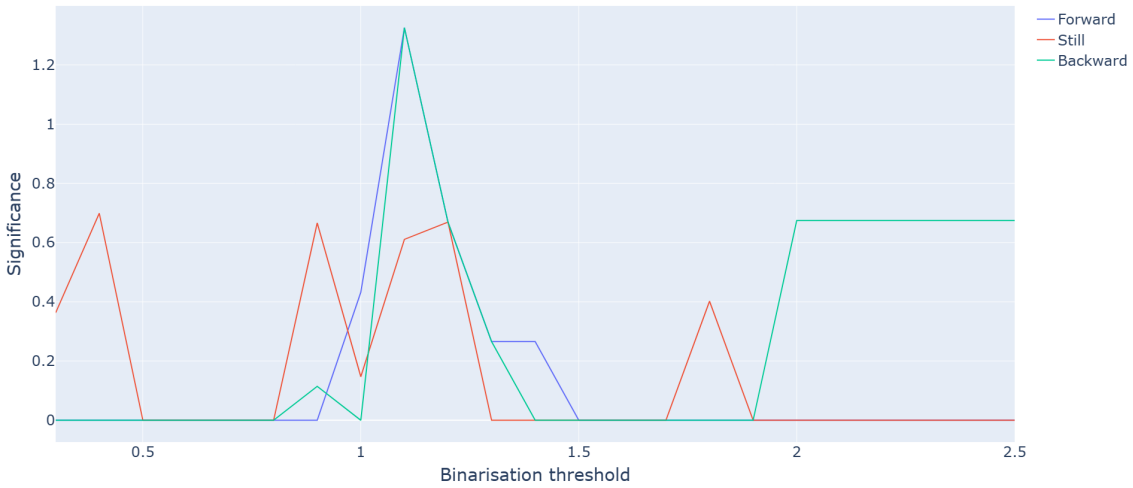
(a) Significance vs. Threshold scan for no median filtering.



(b) Significance vs. Threshold scan for `kernel-size` = 3.

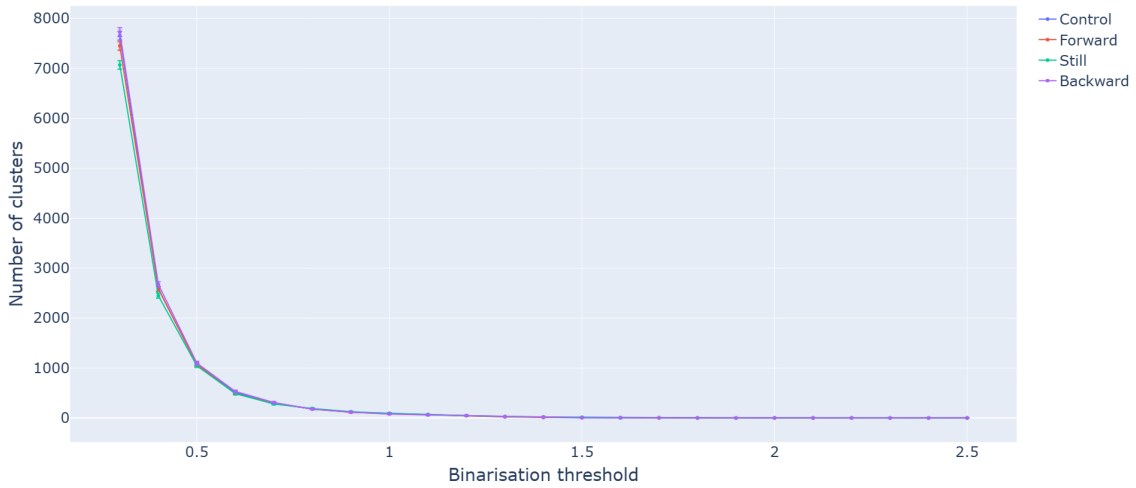


(c) Significance vs. Threshold scan for `kernel-size` = 7.

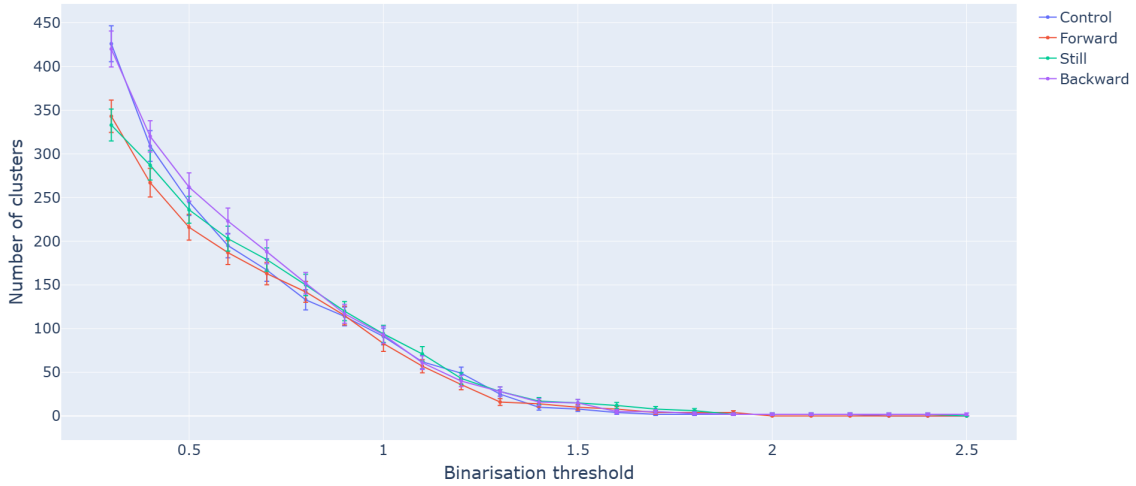


(d) Significance vs. Threshold scan for `kernel-size` = 9.

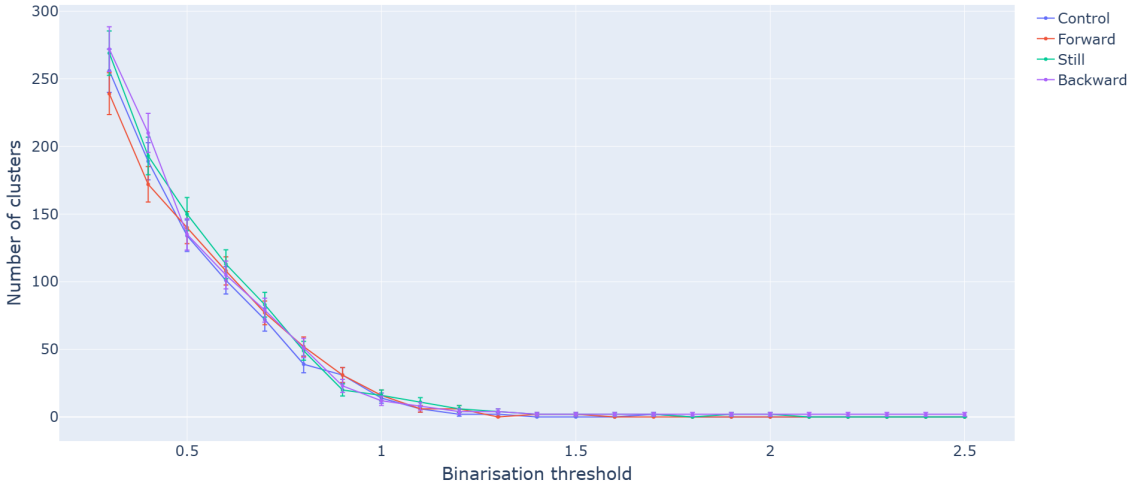
Figure B.7 – Plots showing the significance resulting from the binary test for the three measurement classes ("Backward", "Still", and "Forward") against the control class, varying the threshold used for binarisation. In the pipeline of the upper one, no median filtering was performed, while in the subsequent ones, a median filtering with `kernel-size` $= 3, 7, 9$ is present.



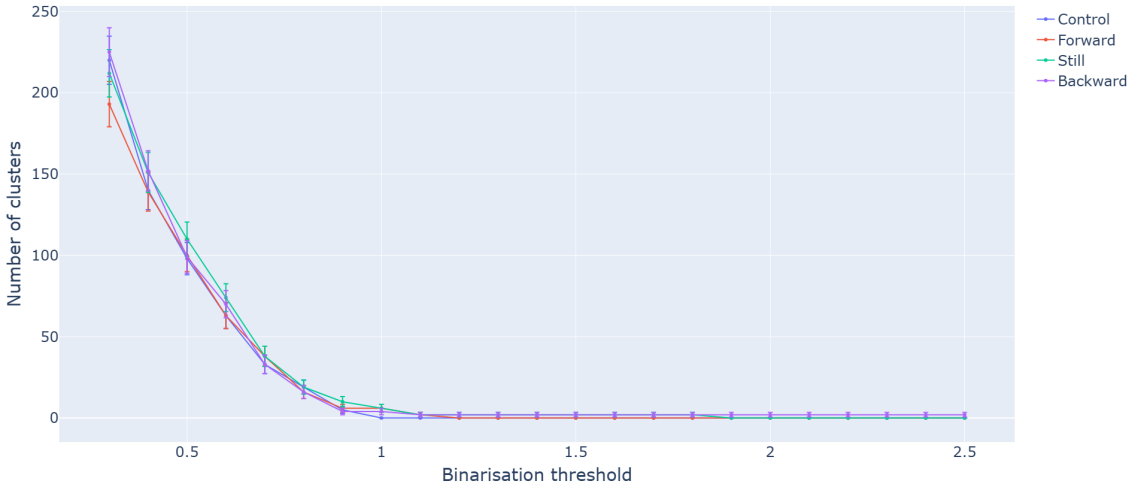
(a) Clusters vs. Threshold scan for no median filtering, in the "Before" dataset cut.



(b) Clusters vs. Threshold scan for `kernel-size` = 3, in the "Before" dataset cut.

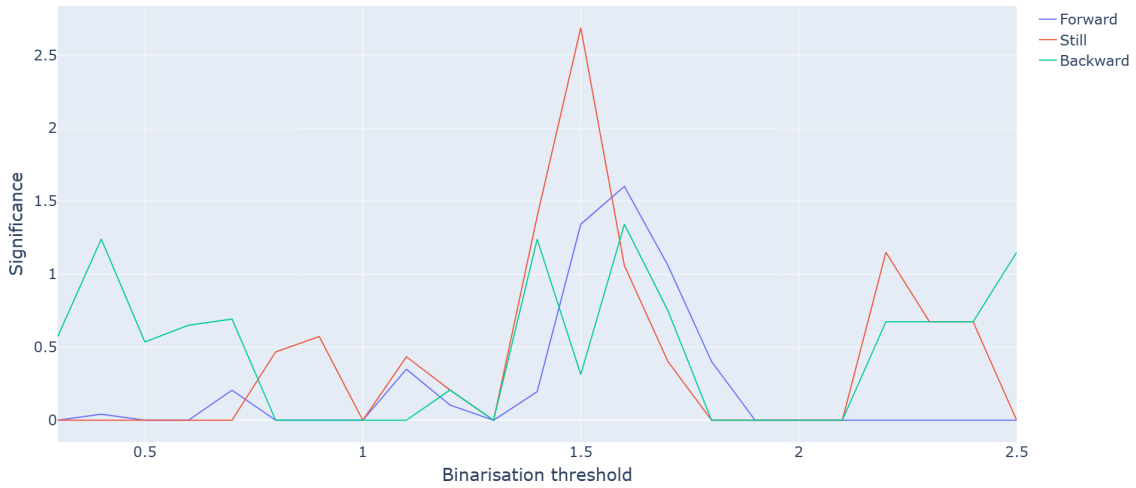


(c) Clusters vs. Threshold scan for `kernel-size` = 7, in the "Before" dataset cut.

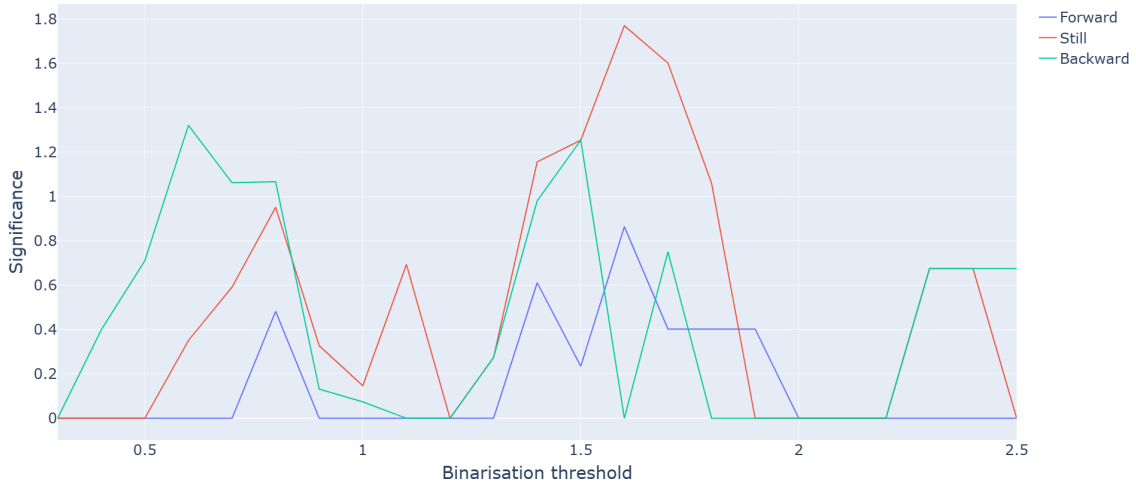


(d) Clusters vs. Threshold scan for `kernel-size` = 9, in the "Before" dataset cut.

Figure B.8 – Plots showing the results of scanning the number of clusters found on all the images of every measurement and control class (Forward, Still, Normal, and Control) over the binarisation threshold, in the "Before" dataset cut. In the pipeline of the upper one, no median filtering was performed, while in the subsequent ones, a median filtering with `kernel-size` = 3, 7, 9 is present.

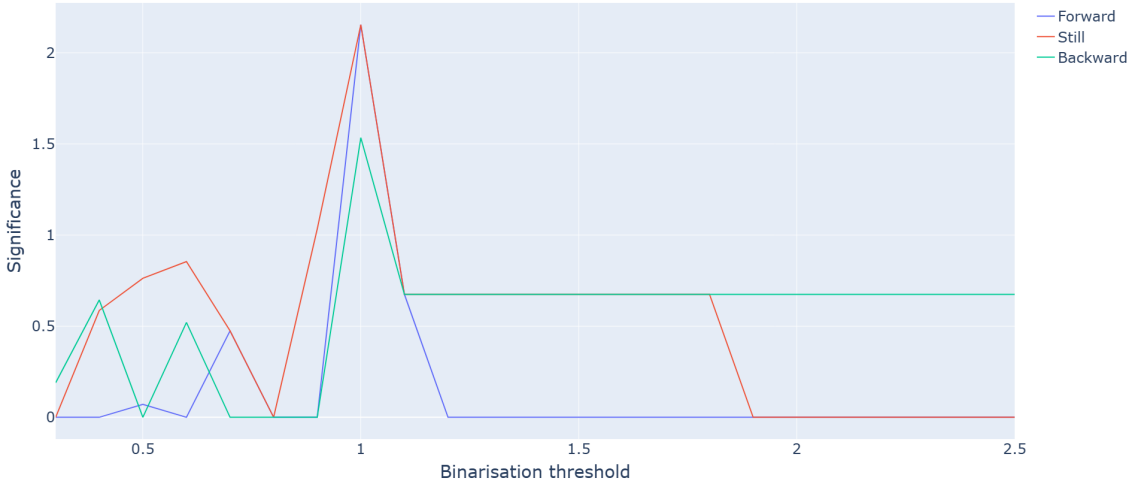


(a) Significance vs. Threshold scan for no median filtering, in the "Before" dataset cut.



(b) Significance vs. Threshold scan for `kernel-size` = 3, in the "Before" dataset cut..

(c) Significance vs. Threshold scan for `kernel-size` $= 7$, in the "Before" dataset cut..



(d) Significance vs. Threshold scan for `kernel-size` $= 9$, in the "Before" dataset cut..

Figure B.9 – Plots showing the significance resulting from the binary test for the three measurement classes ("Backward", "Still", and "Forward") against the control class, varying the threshold used for binarisation, in the "Before" dataset cut. In the pipeline of the upper one, no median filtering was performed, while in the subsequent ones, a median filtering with `kernel-size` $= 3, 7, 9$ is present.

Ringraziamenti

Ho sempre pensato che, in una tesi di laurea o di dottorato, la pagina dei ringraziamenti fosse una delle parti più importanti, perché nello scriverla l'autore si deve necessariamente confrontare con la sua vita passata e le persone che, nel bene o nel male, hanno lasciato una traccia nella sua esistenza. Ora che è venuto il mio di turno di scrivere questa sezione, il bilancio che ne esce è positivo, e la lista delle persone da ringraziare perché mi hanno portato qualcosa di buono nell'esistenza è bella lunga.

Il primo ringraziamento non può che non andare a Rob e a Micheal[3], per aver creduto di prendere me come dottorando, pur senza conoscermi, e per avermi affidato un compito così delicato e sostanziale come riorganizzare l'intero sistema di controllo dell'esperimento. Per carità, mi sembra che lo scommettere nel prendere un dottorando trentenne con esperienze multiple all'infuori dell'accademia abbia pagato ;) Ma è stato tutt'altro che scontato, quindi grazie per l'opportunità!

Similmente, un grazie va anche a Rugge, prima amico e poi anche superiore, per avermi proposto questo dottorato senza mai aver veramente lavorato insieme (beh, di belinate assieme ne avevamo fatte: vedi la *Mussa Neigra*), e per aver cercato di mantenere il difficile rapporto colleghi/amici durante tutti questi tre anni.

Ringrazio anche Lucio Pancheri per la fiducia ripostami nelle mie competenze "quantum", anche se purtroppo abbiamo avuto ben poche opportunità per scambi fruttiferi. Spero che avremo più occasioni, in futuro!

Clearly, how could I not say thanks to all the folks of the AEgIS collaboration, and in particular to all the people who have shared some of their time with me in that swirling vortex of entropy that is the Control Room? So, thanks, Benji (especially for your time-delayed jokes – it always takes a second of two to catch them! –), Franz (& Bergi!), Fredrik, Georgy, Gianni (sempre un sorriso sincero ed una buona parola), Gosja (I'll never bring you something from Sicily), Lisa (the Stef Bingo and mug will always remain in my heart, and on my desk), Matthias, Nicola, the Prelz (fonte di infinita saggezza informatica e storie incredibili), Saiva (remember: keep calm and buy another Kasli), Stef (how will we do without you? Ah, there's Theo! =P), Theo, Tim, Valts (how can you eat so much?).

A special thanks goes to Jakub, that from being a PhD student to train has become my partner-in-crime of the CIRCUS development, ending up writing a good part of it. He introduced me to the pixel art, which has been one of the most funny thing I've done during this PhD (see *Artworks* (B.3)). He's an endless source of nerdy jokes: despite we established the Rule of the Two, he still hasn't killed his Master, so apparently he still needs to learn something on the Dark Arts of LabVIEW™ ;)

Come non rigraziare quei belinoni dei fisici (Ale Calzona, Ale Caminata, Ale Casale, Bianca, Chiaretta, Coppe (appari sempre doppio te), Denise, il Dodo, FedeFerra, Laura, Lollo, Nello, il Pres, Ruben, Rugge (pure tu!), etc.), con i quali l'amore per la fisica è iniziato, e più in generale, ai "genovesi" (Sivo, Cate, la Fra, Caro, Chiara (anche se hai la casa a Chiavari =P), Marika, Circe e il Narcos (per me siete sempre "genovesi" :D), Viola (anche se vivi in Germania), e i nuovi arrivati Daniele e Gabriele), tutti indistintamente compagni di mille avventure, passate... e future.

[3] Che parla benissimo l'italiano, quindi non devo farglieli separati in inglese =P

Un grazie va anche agli amici forgiati durante il mio predecente tentativo di dottorato: Andrea, Lodewijk, Mario, Maarten, Sander, and all the others. E come dimenticare quegli scemi del SEEDS: Lo (si sposa!), Ja, Ma, Albi, Silvio. Ci si vede presto!

Grazie prof. Davide per raccontarmi le tue vicissitudini accademiche degne dell'Università Invisibile della Magia, e per il continuo (e continuamente in ritardo) scambio di indovinelli laterali.

A peculiar thanks then goes to Marieke, with whom the first doctorate attempt did not work out (I wasn't ready, sorry). But she granted me the benefit of the second attempt, and sustained me in the application to this position: dank je wel!

Merci à Samuel de m'avoir aidé à apprendre le français avec sympathie !

Poi ci sono i miei fratelli acquisiti chiavaresi, Angelo, Coppe, Dep, Paolo, Matte e Vitto, che anche se la vita ci sta allontanando sempre di più, basta fare una nottata a Villacella per ritornare immediatamente a quando avevamo vent'anni.

Voglio pensare di dare un saluto speciale a Simo, che non ha potuto vedere sta belin di tesi, ma i cui consigli e ricordi mi hanno aiutato nel percorso. Felice e onorato di averti conosciuto, ti porterò sempre nel cuore, belin1 di un brucia.

Un grazie molto speciale va anche a Elisa, senza la quale non avrei iniziato questa tesi, e a Donata, senza la quale non sono sicuro l'avrei finita. Ma grazie anche, e soprattutto, per tutto il resto.

Chiaramente, non posso non ringraziare i miei genitori, per avermi cresciuto e supportato con amore. Siete un po' due testine ogni tanto, ma d'altronde, la mela non cade troppo distante dall'albero ;) In particolare, questa tesi non sarebbe sicuramente uscita in tempo senza il vostro sostegno datoci a Marzo: grazie, mamma e papà. Vi voglio un mondo di bene.

Un grazie va anche ad Antonella e Tommaso, che mi hanno supportato (e sopportato!) durante questo mio percorso di dottorato, e mi hanno accolto con affetto come un membro della famiglia. Lo stesso vale per Giorgia e Stephan: non vedo l'ora di fare un'altra cena alla Giuggiola o, magari, a Berlino!

Questa tesi l'ho iniziata dedicandola a Claudia, e la chiudo ringraziando lei, per essere sempre al mio fianco, anche nei momenti peggiori, per essere abbastanza matta da aver mollato tutto per seguirmi in quest'avventura che è la mia vita (e che ora è diventata la nostra avventura!), e per aver sempre creduto ciecamente nella nostra relazione, anche quando ha vacillato al punto che molti – inclusi noi stessi – pensavano fosse finita. Ma soprattutto, Cla, ti voglio ringraziare per avermi trasformato la vita da un piatto film in bianco e nero ad una montagna russa di emozioni, piena di colori.

List of Figures

List of Tables

Bibliography

[1] M. Volponi et al. "CIRCUS: An Autonomous Control System for Antimatter, Atomic and Quantum Physics Experiments". In: *EPJ Quantum Technology* 11.1 (Feb. 15, 2024), p. 10. ISSN: 2196-0763. DOI: 10.1140/epjqt/s40507-024-00220-6. URL: https://rdcu.be/dyK5f.

[2] M. Volponi et al. "TALOS: a framework for autonomous control systems for complex experiments". In: *Submitted* (2023).

[3] Marco Volponi. "AEgIS: upgrades and first data". In: *International Journal of Modern Physics A* (2024, in publication).

[4] L. T. Gloggler et al. "Positronium laser cooling via the 1^3S-2^3P transition with a broadband laser pulse". In: (Oct. 2023). arXiv:2310.08760 [hep-ex, physics:physics]. DOI: 10.48550/arXiv.2310.08760. URL: http://arxiv.org/abs/2310.08760.

[5] Michael Doser. *AEgIS / AD-6 annual report 2022*. Tech. rep. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2799304.

[6] Ruggero Caravita. *AEgIS/AD-6 annual report 2023*. Tech. rep. Geneva: CERN, 2023. URL: https://cds.cern.ch/record/2846698.

[7] Ruggero Caravita. *AEgIS/AD-6 annual report 2024*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2887577.

[8] M. M. Nieto and T. Goldman. "The arguments against "antigravity"' and the gravitational acceleration of antimatter". In: *Phys. Rep.* 205 (1991), pp. 221–281.

[9] Mark Fischler, Joe Lykken, and Tom Roberts. *Direct Observation Limits on Antimatter Gravitation*. en. arXiv:0808.3929 [hep-th]. Aug. 2008. URL: http://arxiv.org/abs/0808.3929.

[10] M. J. Borchert et al. "A 16-Parts-per-Trillion Measurement of the Antiproton-to-Proton Charge-Mass Ratio". In: *Nature* 601.7891 (Jan. 6, 2022), pp. 53–57. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-021-04203-w. URL: https://www.nature.com/articles/s41586-021-04203-w.

[11] E. K. Anderson et al. "Observation of the Effect of Gravity on the Motion of Antimatter". In: *Nature* 621.7980 (7980 Sept. 2023), pp. 716–722. ISSN: 1476-4687. DOI: 10.1038/s41586-023-06527-1. URL: https://www.nature.com/articles/s41586-023-06527-1.

[12] F. C. Witteborn and W. M. Fairbank. "Experiments to determine the Force of Gravity on Single Electrons and Positrons". In: *Nature* 220 (1968), p. 436.

[13] S. Aghion et al. (AEgIS collaboration), AEGIS Collaboration. "A moiré deflectometer for antimatter". In: *Nat. Commun.* 5 (2014), p. 4538.

[14] G Drobychev et al. *Proposal for the AEgIS experiment at the CERN antiproton decelerator (Antimatter Experiment: Gravity, Interferometry, Spectroscopy)*. Tech. rep. SPSC-P-334. CERN-SPSC-2007-017. Geneva: CERN, June 2007. URL: http://cds.cern.ch/record/1037532.

[15] C. Amsler et al. (AEgIS Collaboration). "Pulsed production of antihydrogen". In: *Communications Physics* 4 (2021), p. 19.

[16] W Bartmann et al. "Progress in ELENA Design". en. In: *Proceeding of IPAC 2013, Shanghai (China)* (2013), pp. 2651–2653.

[17] Carl Hewitt, Peter Bishop, and Richard Steiger. "A universal modular ACTOR formalism for artificial intelligence". In: *Proc. International Joint Conference on Artificial Intelligence*. 1973, pp. 235–245.

[18] Grzegorz Kasprowicz et al. *ARTIQ and Sinara: Open Software and Hardware Stacks for Quantum Physics*. 2020. DOI: 10.1364/QUANTUM.2020.QTu8B.14. URL: http://www.osapublishing.org/abstract.cfm?URI=QUANTUM-2020-QTu8B.14.

[19] E. Sather. "The mystery of the matter asymmetry". In: *Beam Line* 1.26 (1999), pp. 31–37.

[20] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. "Matter and Antimatter in the Universe". In: *New Journal of Physics* 14.9 (Sept. 2012), p. 095012. ISSN: 1367-2630. DOI: 10.1088/1367-2630/14/9/095012. URL: https://dx.doi.org/10.1088/1367-2630/14/9/095012.

[21] A. G. Cohen, A. De Rújula, and S. L. Glashow. "A Matter-Antimatter Universe?" In: *The Astrophysical Journal* 495.2 (Mar. 1998), pp. 539–549. ISSN: 1538-4357. DOI: 10.1086/305328. URL: http://dx.doi.org/10.1086/305328.

[22] Maxim Yu Khlopov, Sergei G. Rubin, and Alexander S. Sakharov. "Antimatter Regions in the Baryon-Dominated Universe". In: XIVth Rencontres de Blois on Matter-Antimatter Asymmetry. Blois, France, Oct. 1, 2002. DOI: 10.48550/arXiv.hep-ph/0210012. arXiv: hep-ph/0210012. URL: http://arxiv.org/abs/hep-ph/0210012.

[23] Simon Dupourqué, Luigi Tibaldo, and Peter von Ballmoos. "Constraints on the antistar fraction in the Solar System neighborhood from the 10-year Fermi Large Area Telescope gamma-ray source catalog". In: *Phys. Rev. D* 103 (8 Apr. 2021), p. 083016. DOI: 10.1103/PhysRevD.103.083016. URL: https://link.aps.org/doi/10.1103/PhysRevD.103.083016.

[24] Andrew Liddle. *An Introduction to Modern Cosmology*. 3rd ed. Hoboken: Wiley, 2015. 201 pp. ISBN: 978-1-118-69027-7. URL: http://swbplus.bsz-bw.de/bsz449487482cov.htm.

[25] Pavel Fileviez Pérez, Clara Murgui, and Alexis D. Plascencia. "Baryogenesis via Leptogenesis: Spontaneous B and L Violation". In: *Physical Review D* 104.5 (Sept. 8, 2021), p. 055007. DOI: 10.1103/PhysRevD.104.055007. URL: https://link.aps.org/doi/10.1103/PhysRevD.104.055007.

[26] Antonio Padilla. *Fantastic Numbers and Where to Find Them: A Journey to the Edge of Physics*. Allen Lane, 2022. 352 pp. ISBN: 978-0-241-44537-2.

[27] A. D. Sakharov. "Violation of CP invariance, C asymmetry, and baryon asymmetry in the universe". In: *Journal of Experimental and Theoretical Physics Letters* 5 (1967), pp. 24–27.

[28] I. Affleck and M. Dine. "A new mechanism for baryogenesis". In: *Nuclear Physics* 249 (1985), pp. 361–380. DOI: 10.1016/0550-3213(85)90021-5.

[29] T. Matsuda. "Baryon number violation, baryogenesis and defects with extra dimensions". In: *Physical Review D* 66 (2002), p. 023508. DOI: 10.1103/PhysRevD.66.023508.

[30] Michael Dine, Lisa Randall, and Scott Thomas. "Baryogenesis from flat directions of the supersymmetric standard model". In: *Nuclear Physics* (1996). DOI: 10.1016/0550-3213(95)00538-2.

[31] Kolb, Linde, and Riotto. "Grand-Unified-Theory Baryogenesis after Preheating." In: *Physical review letters* 77 21 (1996), pp. 4290–4293. DOI: 10.1103/PHYSREVLETT.77.4290.

[32] V. Rubakov and M. Shaposhnikov. "Electroweak baryon number non-conservation in the early Universe and in high-energy collisions". In: *Physics-Uspekhi* 39 (1996), pp. 461–502. DOI: 10.1070/PU1996v039n05ABEH000145.

[33] G. Anderson and L. Hall. "Electroweak phase transition and baryogenesis." In: *Physical review. D, Particles and fields* 45 8 (1992), pp. 2685–2698. DOI: 10.1103/PHYSREVD.45.2685.

[34] M. Dine et al. "Creating the baryon asymmetry at the electroweak phase transition". In: *Physics Letters B* 257 (1991), pp. 351–356. DOI: 10.1016/0370-2693(91)91905-B.

[35] Lawrence M. Krauss and Soo-Jong Rey. "Baryon number violation at the electroweak scale : first Yale-Texas Workshop, Yale University, 19-21 March 1992". In: (1992). DOI: 10.1142/9789814525183.

[36] Glennys R. Farrar and M. E. Shaposhnikov. "Baryon asymmetry of the Universe in the minimal standard model". In: *Phys. Rev. Lett.* 70 (19 May 1993), pp. 2833–2836. DOI: 10.1103/PhysRevLett.70.2833. URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.2833.

[37] G. 't Hooft. "Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle". In: *Physical Review D* 14.12 (Dec. 15, 1976), pp. 3432–3450. DOI: 10.1103/PhysRevD.14.3432. URL: https://link.aps.org/doi/10.1103/PhysRevD.14.3432.

[38] J. Kripfganz. "Sphalerons and the electroweak phase transition". In: *International Journal of Modern Physics C* 03 (1992), pp. 783–797. DOI: 10.1142/S0129183192000476.

[39] M. Fukugita and T. Yanagida. "Barygenesis without Grand Unification". In: *Physics Letters B* 174.1 (June 26, 1986), pp. 45–47. ISSN: 0370-2693. DOI: 10.1016/0370-2693(86)91126-3. URL: https://www.sciencedirect.com/science/article/pii/0370269386911263.

[40] H. Nielsen and Y. Takanishi. "Baryogenesis via lepton number violation in Anti-GUT model". In: *Physics Letters B* 507 (2001), pp. 241–251. DOI: 10.1016/S0370-2693(01)00357-4.

[41] Riccardo Barbieri et al. "Baryogenesis through leptogenesis". In: *Nuclear Physics* (2000). DOI: 10.1016/S0550-3213(00)00011-0.

[42] Glennys R. Farrar and M. E. Shaposhnikov. "Baryon Asymmetry of the Universe in the Minimal Standard Model". In: *Physical Review Letters* 70.19 (May 10, 1993), pp. 2833–2836. DOI: 10.1103/PhysRevLett.70.2833. URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.2833.

[43] Wan-Zhe Feng, A. Mazumdar, and P. Nath. "Baryogenesis from dark matter". In: *Physical Review D* 88 (2013), p. 036014. DOI: 10.1103/PhysRevD.88.036014.

[44] Gilly Elor, M. Escudero, and A. Nelson. "Baryogenesis and dark matter from B mesons". In: *Physical Review D* (2018). DOI: 10.1103/PhysRevD.99.035031.

[45] T. Banks, S. Echols, and J. L. Jones. "Baryogenesis, dark matter and the pentagon". In: *Journal of High Energy Physics* 2006 (2006), pp. 046–046. DOI: 10.1088/1126-6708/2006/11/046.

[46] Jeremy Sakstein and A. Solomon. "Baryogenesis in Lorentz-violating gravity theories". In: *Physics Letters B* 773 (2017), pp. 186–190. DOI: 10.1016/j.physletb.2017.08.039.

[47] G. Lambiase and G. Scarpetta. "Baryogenesis in f(R) theories of gravity". In: *Physical Review D* 74 (2006), p. 087504. DOI: 10.1103/PhysRevD.74.087504.

[48] K. Hamada, Azusa Minamizaki, and A. Sugamoto. "Baryogenesis with Quantum Gravity". In: *Modern Physics Letters A* 23 (2007), pp. 237–244. DOI: 10.1142/S021773230802639X.

[49] T. Matsuda. "Enhanced baryon number violation due to cosmological defects with localized fermions along extra dimensions". In: *Physical Review D* 65 (2002), p. 107302. DOI: 10.1103/PhysRevD.65.107302.

[50] A. Masiero et al. "Baryogenesis versus proton stability in theories with extra dimensions". In: *Physical Review D* 62 (2000), pp. 063515–063520. DOI: 10.1103/PhysRevD.62.063515.

[51] J. et al. (PDG) Beringer. "CP violation in meson decays". In: *PR* D86 (2012), p. 010001. URL: http://pdg.lbl.gov.

[52] M. B. Gavela et al. "Standard Model Cp-Violation and Baryon Asymmetry". In: *Modern Physics Letters A* 09.09 (Mar. 21, 1994), pp. 795–809. ISSN: 0217-7323. DOI: 10.1142/S0217732394000629. URL: https://www.worldscientific.com/doi/abs/10.1142/S0217732394000629.

[53] Linas Vepstas. *Feynman diagrams showing the neutral kaon* K^0 *turning into* $\bar{K}^0$ *by means of the exchange of a pair of W bosons*. CC-BY-SA-3.0. 2006.

[54] B. Pontecorvo. "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". In: *Soviet Journal of Experimental and Theoretical Physics* 26 (May 1, 1968), p. 984. ISSN: 1063-7761. URL: https://ui.adsabs.harvard.edu/abs/1968JETP...26..984P.

[55] K. Abe et al. "Constraint on the Matter-Antimatter Symmetry-Violating Phase in Neutrino Oscillations". In: *Nature* 580.7803 (7803 Apr. 2020), pp. 339–344. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2177-0. URL: https://www.nature.com/articles/s41586-020-2177-0.

[56] T. Ishida and Hyper-Kamiokande. "T2HK: J-PARC Upgrade Plan for Future and beyond T2K". In: XVth International Workshop on Neutrino Factories, Super Beams and Beta Beams (NUFACT2013). Beijing, China, Aug. 19–24, 2013. DOI: 10.48550/arXiv.1311.5287.

[57] R. Acciarri et al. *Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 1: The LBNF and DUNE Projects*. Jan. 20, 2016. DOI: 10.48550/arXiv.1601.05471. arXiv: 1601.05471 [hep-ex, physics:physics]. URL: http://arxiv.org/abs/1601.05471. preprint.

[58] *Quantum Field Theory and Standard Model*. URL: https://schwartzqft.fas.harvard.edu/home.

[59] C. A. Baker et al. "Improved Experimental Limit on the Electric Dipole Moment of the Neutron". In: *Physical Review Letters* 97.13 (Sept. 27, 2006), p. 131801. DOI: 10.1103/PhysRevLett.97.131801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.131801.

[60] R. D. Peccei and Helen R. Quinn. "CP Conservation in the Presence of Pseudoparticles". In: *Phys. Rev. Lett.* 38 (25 June 1977), pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440. URL: https://link.aps.org/doi/10.1103/PhysRevLett.38.1440.

[61] J. P. Hsu. "Cosmic Matter-Antimatter Asymmetry and Gravitational Force". In: *Lettere al Nuovo Cimento (1971-1985)* 28.4 (May 1, 1980), pp. 128–132. ISSN: 1827-613X. DOI: 10.1007/BF02772916. URL: https://doi.org/10.1007/BF02772916.

[62] Alfred Scharff Goldhaber and Michael Martin Nieto. "Photon and Graviton Mass Limits". In: *Reviews of Modern Physics* 82.1 (Mar. 23, 2010), pp. 939–979. DOI: 10.1103/RevModPhys.82.939. URL: https://link.aps.org/doi/10.1103/RevModPhys.82.939.

[63] K. Jagannathan and L. P. S. Singh. "Attraction/Repulsion between like Charges and the Spin of the Classical Mediating Field". In: *Physical Review D* 33.8 (Apr. 15, 1986), pp. 2475–2477. DOI: 10.1103/PhysRevD.33.2475. URL: https://link.aps.org/doi/10.1103/PhysRevD.33.2475.

[64] M. M. Nieto and B. E. Bonner. *Looking for New Gravitational Forces with Antiprotons*. 2. review, planning workshop on antiproton science, and technology, Santa Monica, CA, USA, 6 Oct 1987. Jan. 1, 1987. URL: https://digital.library.unt.edu/ark:/67531/metadc1088131/.

[65] Galileo Galilei. *Discorsi e dimostrazioni matematiche intorno a due nuove scienze*. Lodewijk Elzevir, 1638.

[66] Isaac Newton. *Philosophiae Naturalis Principia Mathematica*. Pepys Press, 1687.

[67] A. Einstein. "Zur Elektrodynamik Bewegter Körper". In: *Annalen der Physik* 322.10 (1905), pp. 891–921. ISSN: 1521-3889. DOI: 10.1002/andp.19053221004. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19053221004.

[68] Albert Einstein. "The Field Equations of Gravitation". In: *Koniglich Preußische Akademie der Wissenschaften (Berlin). Sitzungsberichte* (1915), pp. 844–847.

[69] R. H. Dicke. "Republication of: The Theoretical Significance of Experimental Relativity". In: *General Relativity and Gravitation* 51.5 (Apr. 30, 2019), p. 57. ISSN: 1572-9532. DOI: 10.1007/s10714-019-2509-2. URL: https://doi.org/10.1007/s10714-019-2509-2.

[70] Clifford M. Will. "The Confrontation between General Relativity and Experiment". In: *Living Reviews in Relativity* 17.1 (Dec. 1, 2014), p. 4. ISSN: 1433-8351. DOI: 10.12942/lrr-2014-4. URL: https://doi.org/10.12942/lrr-2014-4.

[71] Anna M. Nobili and Alberto Anselmi. "Relevance of the Weak Equivalence Principle and Experiments to Test It: Lessons from the Past and Improvements Expected in Space". In: *Physics Letters A*. Special Issue in Memory of Professor V.B. Braginsky 382.33 (Aug. 25, 2018), pp. 2205–2218. ISSN: 0375-9601. DOI: 10.1016/j.physleta.2017.09.027. URL: https://www.sciencedirect.com/science/article/pii/S0375960117304899.

[72] Harold H. Potter and Owen Willans Richardson. "Some Experiments on the Proportionality of Mass and Weight". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 104.728 (Jan. 1997), pp. 588–610. DOI: 10.1098/rspa.1923.0130. URL: https://royalsocietypublishing.org/doi/10.1098/rspa.1923.0130.

[73] V. B. Braginsky and V. I. Panov. "The Equivalence of Inertial and Passive Gravitational Mass". In: *General Relativity and Gravitation* 3.4 (Dec. 1, 1972), pp. 403–404. ISSN: 1572-9532. DOI: 10.1007/BF00759178. URL: https://doi.org/10.1007/BF00759178.

[74] T. A. Wagner et al. "Torsion-Balance Tests of the Weak Equivalence Principle". In: *Classical and Quantum Gravity* 29.18 (Aug. 2012), p. 184002. ISSN: 0264-9381. DOI: 10.1088/0264-9381/29/18/184002. URL: https://dx.doi.org/10.1088/0264-9381/29/18/184002.

[75] J. O. Dickey et al. "Lunar Laser Ranging: A Continuing Legacy of the Apollo Program". In: *Science* 265.5171 (July 22, 1994), pp. 482–490. DOI: 10.1126/science.265.5171.482. URL: https://www.science.org/doi/10.1126/science.265.5171.482.

[76] V Viswanathan et al. "The New Lunar Ephemeris INPOP17a and Its Application to Fundamental Physics". In: *Monthly Notices of the Royal Astronomical Society* 476.2 (May 11, 2018), pp. 1877–1888. ISSN: 0035-8711. DOI: 10.1093/mnras/sty096. URL: https://doi.org/10.1093/mnras/sty096.

[77] Stephen M. Merkowitz. "Tests of Gravity Using Lunar Laser Ranging". In: *Living Reviews in Relativity* 13.1 (Nov. 2, 2010), p. 7. ISSN: 1433-8351. DOI: 10.12942/lrr-2010-7. URL: https://doi.org/10.12942/lrr-2010-7.

[78] Pierre Touboul. "The Microscope Mission and Its Uncertainty Analysis". In: *Space Science Reviews* 148.1 (Dec. 1, 2009), pp. 455–474. ISSN: 1572-9672. DOI: 10.1007/s11214-009-9565-y. URL: https://doi.org/10.1007/s11214-009-9565-y.

[79] MICROSCOPE Collaboration et al. "MICROSCOPE Mission: Final Results of the Test of the Equivalence Principle". In: *Physical Review Letters* 129.12 (Sept. 14, 2022), p. 121102. DOI: 10.1103/PhysRevLett.129.121102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.129.121102.

[80] F. C. Witteborn and W. M. Fairbank. "Experimental comparison of the gravitational force on freely falling electrons and metallic electrons". In: *Phys. Rev. Lett.* 19 (1967), p. 1049.

[81] C. Herring and M. H. Nichols. "Thermoionic emission". In: *Rev. Mod. Phys.* 21 (1949), p. 185.

[82] Michael Holzcheiter. *Gravitational Measurements on Charged Particles: can it be done?* talk at the 3rd Workshop on Antimatter and Gravity, London (UK). 2015.

[83] T. Kajita, M. Koshiba, and A. Suzuki. "On the origin of the Kamiokande experiment and neutrino astrophysics". In: *The European Physical Journal H* 37 (2012), pp. 33–73. DOI: 10.1140/EPJH/E2012-30007-Y.

[84] R. C. Svobod. "The IMB Proton Decay Detector". In: 107 (1983), pp. 363–366. DOI: 10.1007/978-94-009-7166-0_28.

[85] V. Kuzminov. "The Baksan Neutrino Observatory". In: *The European Physical Journal Plus* 127 (2012), pp. 1–13. DOI: 10.1140/EPJP/I2012-12113-0.

[86] Sandip Pakvasa, Walter A. Simmons, and Thomas J. Weiler. "Test of equivalence principle for neutrinos and antineutrinos". In: *Phys. Rev. D* 39 (6 Mar. 1989), pp. 1761–1763. DOI: 10.1103/PhysRevD.39.1761. URL: https://link.aps.org/doi/10.1103/PhysRevD.39.1761.

[87] C. S. Unnikrishnan and George T. Gillies. "Falling right while moving slow: true tests of the Weak Equivalence Principle for antiparticles". In: *International Journal of Modern Physics D* 21 (2012), p. 1242016. URL: https://api.semanticscholar.org/CorpusID:123440459.

[88] S Ulmer et al. *Direct High-Precision Measurement of the g-Factor of a Single Antiproton Stored in a Cryogenic Penning Trap*. Tech. rep. Geneva: CERN, 2012. URL: https://cds.cern.ch/record/1455847.

[89] G. Gabrielse et al. "Precision Mass Spectroscopy of the Antiproton and Proton Using Simultaneously Trapped Particles". In: *Physical Review Letters* 82.16 (Apr. 19, 1999), pp. 3198–3201. DOI: 10.1103/PhysRevLett.82.3198. URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.3198.

[90] I. R. Kenyon. "A recalculation on the gravitational mass difference between the K0 and mesons". In: *Physics Letters B* 237 (1990), pp. 274–277. URL: https://api.semanticscholar.org/CorpusID:121514000.

[91] C. Tchernin et al. "Characterizing Galaxy Clusters by Their Gravitational Potential: Systematics of Cluster Potential Reconstruction". In: *Astronomy and Astrophysics* 644 (Dec. 1, 2020), A126. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201937028. URL: https://ui.adsabs.harvard.edu/abs/2020A&A...644A.126T.

[92] Jeffrey S Hangst and Paul Bowe. *ALPHA Proposal*. Tech. rep. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/814351.

[93] M. Doser et al. (AEgIS collaboration), AEGIS Collaboration. "AEgIS at ELENA: outlook for physics with a pulsed cold antihydrogen beam". In: *Phil. Trans. Roy. Soc. Lond. A* 376 (2018), p. 20170274.

[94] M. Charlton. "Antihydrogen production in collisions of antiprotons with excited states of positronium". In: *Physics Letters A* 143.3 (1990), pp. 143–146. ISSN: 0375-9601. DOI: https://doi.org/10.1016/0375-9601(90)90665-B. URL: https://www.sciencedirect.com/science/article/pii/037596019090665B.

[95] D. Krasnický et al. "Cross-section for Rydberg Antihydrogen Production via Charge Exchange Between Rydberg Positronium and Antiprotons in Magnetic Field". In: *Phys. Rev. A* 94 (2016), p. 022714.

[96] G Chardin et al. *Proposal to measure the Gravitational Behaviour of Antihydrogen at Rest.* Tech. rep. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1386684.

[97] T Azuma, J S Bakos, and H Bluhme. *Atomic spectroscopy and collisions using slow antiprotons.* Tech. rep. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/622250.

[98] G Kornakov et al. "Experiments with Mid-Heavy Antiprotonic Atoms in AEgIS". In: *PoS* PANIC2021 (2022), p. 446. DOI: 10.22323/1.380.0446. URL: https://cds.cern.ch/record/2836714.

[99] Esma Mobs. *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018.* Tech. rep. 2018. URL: https://cds.cern.ch/record/2636343.

[100] G. Testera et al. (AEgIS collaboration), AEGIS Collaboration. "The AEgIS experiment". In: *Hyperfine Interact.* 233 (2015), pp. 1–3, 13–20.

[101] J. Storey et al., AEGIS Collaboration. "Particle tracking at 4K: the Fast Annihilation Cryogenic Tracking (FACT) detector for the AEgIS antimatter gravity experiment". In: *Nucl. Instrum. Methods Phys. Res., Sect. A* 732 (2013), pp. 437–441.

[102] C. Amsler et al. "A cryogenic tracking detector for antihydrogen detection in the AEgIS experiment". en. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 960 (Apr. 2020), p. 163637. ISSN: 01689002. DOI: 10.1016/j.nima.2020.163637. URL: https://linkinghub.elsevier.com/retrieve/pii/S0168900220302151.

[103] Nicola Zurlo et al. "Calibration and Equalisation of Plastic Scintillator Detectors for Antiproton Annihilation Identification Over Positron/Positronium Background". In: *Acta Phys. Polon. B* 51 (2020), pp. 213–223. DOI: 10.5506/APhysPolB.51.213.

[104] S. Mariazzi et al. "High-yield thermalized positronium at room temperature emitted by morphologically tuned nanochanneled silicon targets". In: *J. Phys. B* 54.8 (2021), p. 085004. DOI: 10.1088/1361-6455/abf6b6.

[105] Sebastiano Mariazzi, Paolo Bettotti, and Roberto S. Brusa. "Positronium Cooling and Emission in Vacuum from Nanochannels at Cryogenic Temperature". en. In: *Physical Review Letters* 104.24 (June 2010), p. 243401. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.104.243401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.104.243401.

[106] S. Mariazzi et al. "High positronium yield and emission into the vacuum from oxidized tunable nanochannels in silicon". en. In: *Physical Review B* 81.23 (June 2010), p. 235418. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.81.235418. URL: https://link.aps.org/doi/10.1103/PhysRevB.81.235418.

[107] Philippe H. M. Bräunig. "Atom Optical Tools for Antimatter Experiments". PhD thesis. Ruperto-Carola University of Heidelberg, Heidelberg, Dec. 2014.

[108] Markus K. Oberthaler et al. "Inertial sensing with classical atomic beams". In: *Phys. Rev. A* 54 (4 Oct. 1996), pp. 3165–3176. DOI: 10.1103/PhysRevA.54.3165. URL: https://link.aps.org/doi/10.1103/PhysRevA.54.3165.

[109] M. K. Oberthaler et al. "Inertial sensing with classical atomic beams". In: *Phys. Rev. A.* 54 (1996), p. 3165.

[110] L Glöggler et al. "High-resolution MCP-TimePix3 imaging/timing detector for antimatter physics". In: *Measurement Science and Technology* 33.11 (Aug. 2022), p. 115105. DOI: 10.1088/1361-6501/ac8221. URL: https://dx.doi.org/10.1088/1361-6501/ac8221.

[111] H.F. Talbot. "LXXVI. Facts Relating to Optical Science. No. IV". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 9.56 (Dec. 1, 1836), pp. 401–407. ISSN: 1941-5966. DOI: 10.1080/14786443608649032. URL: https://doi.org/10.1080/14786443608649032.

[112] E. Lau. "Beugungserscheinungen an Doppelrastern". In: *Annalen der Physik* 437.7-8 (1948), pp. 417–423. ISSN: 1521-3889. DOI: 10.1002/andp.19484370709. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19484370709.

[113] S. Aghion et al. "A Moiré Deflectometer for Antimatter". In: *Nature Communications* 5.1 (July 28, 2014), p. 4538. ISSN: 2041-1723. DOI: 10.1038/ncomms5538. URL: https://www.nature.com/articles/ncomms5538.

[114] P. Bräunig et al. *Limits on a gravity measurement with a non-collimated antihydrogen source.* Tech. rep. Kirchoff-Institute for Physics, Heidelberg, 2015.

[115] Ruggero Caravita. "Towards measuring gravity on neutral antimatter". PhD thesis. Università degli Studi di Genova, Apr. 2017.

[116] S A Baird et al. *Design Study of the Antiproton Decelerator: AD.* 1996. DOI: 10.17181/CERN.KNXM.AYKR. URL: https://cds.cern.ch/record/317704. preprint.

[117] J. Y. Hémery and S. Maury. "The Antiproton Decelerator: Overview". In: *Nuclear Physics A*. Proceedings of the Fifth Biennial Conference on Low-Energy Antiproton Physics 655.1 (Aug. 9, 1999), pp. c345–c352. ISSN: 0375-9474. DOI: 10.1016/S0375-9474(99)00223-7. URL: https://www.sciencedirect.com/science/article/pii/S0375947499002237.

[118] S. Cialdi et al. "Efficient two-step Positronium laser excitation to Rydberg levels". In: *Nucl. Instrum. Methods Phys. Res. B* 269.13 (2011), pp. 1527–1533.

[119] Ruggero Caravita. "Laser apparatus for exciting positronium in AEgIS positronium spectroscopy experiment". MA thesis. Università degli Studi di Milano, Apr. 2013.

[120] S. Aghion et al. (AEgIS collaboration), AEGIS Collaboration. "Laser excitation of the n = 3 level of positronium for antihydrogen production". In: *Phys. Rev. A* 94 (2016), p. 012507.

[121] Joseph Ladislas Wiza. "Microchannel plate detectors". en. In: *Nuclear Instruments and Methods* 162.1-3 (June 1979), pp. 587–601. ISSN: 0029554X. DOI: 10.1016/0029-554X(79)90734-1. URL: https://linkinghub.elsevier.com/retrieve/pii/0029554X79907341.

[122] Qwerty123uiop. *Schematic view of a photomultiplier coupled to a scintillator, illustrating detection of gamma rays.* CC-BY-SA-3.0. 2022.

[123] C. Amsler et al. "A Cryogenic Tracking Detector for Antihydrogen Detection in the AEgIS Experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 960 (Apr. 2020), p. 163637. ISSN: 01689002. DOI: 10.1016/j.nima.2020.163637. URL: https://linkinghub.elsevier.com/retrieve/pii/S0168900220302151.

[124] T. J. Murphy and C. M. Surko. "Positron trapping in an electrostatic well by inelastic collisions with nitrogen molecules". In: *Phys. Rev. A* 46 (1992), p. 5696.

[125] C. M. Surko, M. Leventhal, and A. Passner. "Positron Plasma in the Laboratory". In: *Physical Review Letters* 62 (1989), pp. 901–904. DOI: 10.1103/PhysRevLett.62.901.

[126] P. J. Schultz and K. G. Lynn. "Interaction of positron beams with surfaces, thin films, and interfaces". In: *Reviews of Modern Physics* 60.3 (1988), pp. 701–779. DOI: 10.1103/RevModPhys.60.701.

[127] S. Mariazzi et al. "High positronium yield and emission into the vacuum from oxidized tunable nanochannels in silicon". In: *Phys. Rev. B* 81 (2010), p. 235418.

[128] Andreas J Peters and Lukasz Janyst. "Exabyte Scale Storage at CERN". In: *Journal of Physics: Conference Series* 331.5 (Dec. 2011), p. 052015. DOI: 10.1088/1742-6596/331/5/052015. URL: https://dx.doi.org/10.1088/1742-6596/331/5/052015.

[129] Sébastien Bourdeauducq et al. *ARTIQ 1.0*. May 2016. DOI: 10.5281/zenodo.51303. URL: https://doi.org/10.5281/zenodo.51303.

[130] M. Antonello et al. "Rydberg-positronium velocity and self-ionization studies in a 1T magnetic field and cryogenic environment". In: *Phys. Rev. A* 102 (1 July 2020), p. 013101. DOI: 10.1103/PhysRevA.102.013101. URL: https://link.aps.org/doi/10.1103/PhysRevA.102.013101.

[131] Christian Zimmer et al. "Positronium Laser Cooling in a Magnetic Field". In: *Physical Review A* 104.2 (Aug. 18, 2021), p. 023106. DOI: 10.1103/PhysRevA.104.023106. URL: https://link.aps.org/doi/10.1103/PhysRevA.104.023106.

[132] C.-W. Chou et al. "Frequency Comparison of Two High-Accuracy Al+ Optical Clocks". In: *Physical Review Letters* 104.7 (Feb. 2010), p. 070802. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.104.070802. arXiv: 0911.4527 [quant-ph]. URL: http://arxiv.org/abs/0911.4527.

[133] F. M. Gonzalez et al. "Improved Neutron Lifetime Measurement with UCNτ". In: *Physical Review Letters* 127.16 (Oct. 2021), p. 162501. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.127.162501. arXiv: 2106.10375 [hep-ex, physics:nucl-ex, physics:nucl-th]. URL: http://arxiv.org/abs/2106.10375.

[134] N. Hinkley et al. "An Atomic Clock with $10\hat{}\{-18\}$ Instability". In: *Science* 341.6151 (Sept. 2013), pp. 1215–1218. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1240420. arXiv: 1305.5869 [physics, physics:quant-ph]. URL: http://arxiv.org/abs/1305.5869.

[135] James Millen and Benjamin A. Stickler. "Quantum Experiments with Microscale Particles". In: *Contemporary Physics* 61.3 (July 2020), pp. 155–168. ISSN: 0010-7514. DOI: 10.1080/00107514.2020.1854497. URL: https://doi.org/10.1080/00107514.2020.1854497.

[136] Rodrigo A. Thomas et al. "Entanglement between Distant Macroscopic Mechanical and Spin Systems". In: *Nature Physics* 17.2 (Feb. 2021), pp. 228–233. ISSN: 1745-2481. DOI: 10.1038/s41567-020-1031-5. URL: https://www.nature.com/articles/s41567-020-1031-5.

[137] B. R. Beck, J. Fajans, and J. H. Malmberg. "Measurement of collisional anisotropic temperature relaxation in a strongly magnetized pure electron plasma". In: *Phys. Rev. Lett.* 68 (3 Jan. 1992), pp. 317–320. DOI: 10.1103/PhysRevLett.68.317. URL: http://link.aps.org/doi/10.1103/PhysRevLett.68.317.

[138] G. B. Andresen et al. "Evaporative Cooling of Antiprotons to Cryogenic Temperatures". In: *Phys. Rev. Lett.* 105 (1 July 2010), p. 013003. DOI: 10.1103/PhysRevLett.105.013003. URL: http://link.aps.org/doi/10.1103/PhysRevLett.105.013003.

[139] G. Gabrielse et al. "Adiabatic Cooling of Antiprotons". In: *Phys. Rev. Lett.* 106 (7 Feb. 2011), p. 073002. DOI: 10.1103/PhysRevLett.106.073002. URL: http://link.aps.org/doi/10.1103/PhysRevLett.106.073002.

[140] S. Aghion et al. (AEgIS collaboration), AEGIS Collaboration. "Compression of a mixed antiproton and electron non-neutral plasma to high densities". In: *Eur. Phys. J. D* 72.4 (2018), p. 76. DOI: 10.1140/epjd/e2018-80617-x. URL: https://doi.org/10.1140/epjd/e2018-80617-x.

[141] C.R. Harris, K.J. Millman, S.J. van der Walt, et al. "Array programming with NumPy". In: *Nature* 585 (2020), pp. 357–362.

[142] P. Virtanen, R. Gommers, T. E. Oliphant, et al., SciPy 1.0 Contributors. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 3.17 (2020), pp. 261–272.

[143] Plotly Technologies Inc. *Collaborative data science*. 2015. URL: https://plot.ly.

[144] Rick Bitter, Taqi Mohiuddin, and Matt Nawrocki. "LabVIEW: Advanced programming techniques. 2nd Ed." In: (2017). DOI: https://doi.org/10.1201/9780849333255.

[145] V. Cerf and R. Kahn. "A Protocol for Packet Network Intercommunication". In: *IEEE Transactions on Communications* 22.5 (1974), pp. 637–648. DOI: 10.1109/TCOM.1974.1092259.

[146] Telegram FZ LLC and Telegram Messenger Inc. *Telegram*. URL: https://telegram.org.

[147] Daniel Krasnický. "Antiproton capture and cooling for production of cold antihydrogen". PhD thesis. Universitá degli Studi di Genova, Mar. 2013.

[148] John David Jackson. *Classical Electrodynamics, 3rd edition*. John Wiley and sons, 2015.

[149] X. Feng et al. "Capture and cooling of antiprotons". In: *Hyperfine Interactions* 109 (1997), p. 145.

[150] F. Anderegg et al. "Measurement of correlated-enhanced collision rates using pure ion plasmas". In: *Phys. Plasmas* 17 (2010), p. 055702.

[151] Daniela Manuzio. "Ottimizzazione dell'efficienza di cattura e di cooling di antiprotoni nell'esperimento ATHENA". MA thesis. via Dodecaneso, 33, Genova: Facoltá di Scienze M. F. N, Universitá di Genova, 2000.

[152] L. J. Spitzer. *Physics of Fully Ionized Gases*. Wiley and Sons, 1962.

[153] R. G. Greaves and C. M. Surko. "Radial compression and inward transport of positron plasmas using a rotating electric field". In: *Phys. Plasmas* 8.5 (2001), pp. 1879–1885.

[154] X. P. Huang et al. "Steady-state confinement of non-neutral plasma by rotating electric fields". In: *Phys. Rev. Lett.* 78 (1997), p. 875.

[155] F. Anderegg, E. M. Hollmann, and C. F. Driscoll. "Rotating field confinement of pure electron plasmas using Tivelpiece-Gould modes". In: *Phys. Rev. Lett.* 78 (1998), p. 4875.

[156] J. R. Danielson and C. M. Surko. "Torque-balanced high-density steady states of single-component plasmas". In: *Phys. Rev. Lett.* 94 (2005), p. 035001.

[157] David L. Morgan and Vernon W. Hughes. "Atomic Processes Involved in Matter-Antimatter Annihilation". In: *Physical Review D* 2.8 (Oct. 15, 1970), pp. 1389–1399. DOI: 10.1103/PhysRevD.2.1389. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1389.

[158] T. -P. Li and Y. -Q. Ma. "Analysis methods for results in gamma-ray astronomy." In: *Astrophysical Journal* 272 (Sept. 1983), pp. 317–324. DOI: 10.1086/161295.

[159] Benjamin Rienäcker. "Creation and Manipulation of Positronium for Efficient Antihydrogen Production at AEGIS". PhD thesis. Munich, Tech. U., 2021.

[160] B. R. Frieden. "A new restoring algorithm for the preferential enhancement of edge gradients". In: *J. Opt. Soc. Amer.* 66 (1976), pp. 280–283.

[161] Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. "Density-Based Clustering Based on Hierarchical Density Estimates". In: *Advances in Knowledge Discovery and Data Mining*. Ed. by Jian Pei et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 160–172. ISBN: 978-3-642-37456-2.

[162] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[163] Poornachandra Sarang. "Connectivity-Based Clustering". In: *Thinking Data Science: A Data Science Practitioner's Guide*. Ed. by Poornachandra Sarang. Cham: Springer International Publishing, 2023, pp. 185–195. ISBN: 978-3-031-02363-7. DOI: 10.1007/978-3-031-02363-7_10. URL: https://doi.org/10.1007/978-3-031-02363-7_10.

[164] Mihael Ankerst et al. "OPTICS: ordering points to identify the clustering structure". In: *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*. SIGMOD '99. Philadelphia, Pennsylvania, USA: Association for Computing Machinery, 1999, pp. 49–60. ISBN: 1581130848. DOI: 10.1145/304182.304187. URL: https://doi.org/10.1145/304182.304187.

[165] M. Deutsch. In: *Phys. Rev.* 82 (1951), p. 455.

[166] D. B. Cassidy et al. "Positronium cooling in porous silica measured via Doppler spectroscopy". In: *Phys. Rev. A* 81 (1 Jan. 2010), p. 012715. DOI: 10.1103/PhysRevA.81.012715. URL: https://link.aps.org/doi/10.1103/PhysRevA.81.012715.

[167] S. Mariazzi et al. (AEgIS Collaboration). "High-yield thermalized positronium at room temperature emitted by morphologically tuned nanochanneled silicon targets". In: *J. of Phys. B* 54 (2021), p. 085004. DOI: 10.1088/1361-6455/abf6b6.

[168] M. S. Fee et al. "Measurement of the positronium 1s-2s interval by continuous-wave two-photon excitation". In: *Phys. Rev. Lett.* 70.10 (1993), p. 1397.

[169] M. S. Fee et al. "Measurement of the positronium $1\ ^3S_1$–$2\ ^3S_1$ interval by continuous-wave two-photon excitation". In: *Phys. Rev. A* 48 (1 July 1993), pp. 192–219. DOI: 10.1103/PhysRevA.48.192. URL: https://link.aps.org/doi/10.1103/PhysRevA.48.192.

[170] Edison P. Liang and Charles D. Dermer. "Laser cooling of positronium". In: *Optics Communications* 65.6 (1988), pp. 419–424.

[171] William D. Phillips. "Nobel Lecture: Laser cooling and trapping of neutral atoms". In: *Rev. Mod. Phys.* 70 (1998), p. 721.

[172] T. Kumita et al. "Study on laser cooling of ortho-positronium". In: *Nucl. Instrum. Methods Phys. Res. B* 171.192 (2002), pp. 1527–1533.

[173] D. B. Cassidy, H. W. K. Tom, and A. P. Mills. "Fundamental Physics with Cold Positronium". In: *AIP Conference Proceedings* 1037 (2008), p. 66.

[174] D. B. Cassidy. "Experimental progress in positronium laser physics". In: *Eur. Phys. J. D* 72 (2018), p. 53.

[175] G. S. Adkins, D. B. Cassidy, and J. Pérez-Rios. "Precision spectroscopy of positronium: Testing bound-state QED theory and the search for physics beyond the Standard Model". In: *Physics Reports* 975 (2022), pp. 1–61.

[176] C. Amsler et al. (Particle Data Group). "Reviews, Tables, and Plots - Contents". In: *Physics Letters B* 667.1 (2008), pp. 101–102. DOI: https://doi.org/10.1016/j.physletb.2008.07.025.

[177] Savely G Karshenboim. "Positronium, antihydrogen, light, and the equivalence principle". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 49.14 (June 2016), p. 144001.

[178] M. Charlton. In: *Phys. Lett. A* 143 (1990), p. 143.

[179] P. M. Platzman and A. P. Mills. "Possibilities for Bose condensation of positronium". In: *Phys. Rev. B* 49 (1994), p. 454.

[180] A. P. Mills Jr. *Proceedings of the International School of Physics Enrico Fermi, course 174*. Ed. by A. P. Dupasquier Mills Jr and R. S. Brusa. Italian Physical Society, 2010.

[181] H. K. Avetissian, A. K. Avetissian, and G. F. Mkrtchian. "Self-Amplified Gamma-Ray Laser on Positronium Atoms from a Bose-Einstein Condensate". In: *Phys. Rev. Lett* 113 (2014), p. 023904.

[182] K. Yamada et al. "Theoretical Analysis and Experimental Demonstration of a Chirped Pulse-Train Generator and its Potential for Efficient Cooling of Positronium". In: *Phys. Rev. Appl.* 16 (1 July 2021), p. 014009. DOI: 10.1103/PhysRevApplied.16.014009.

[183] K Shu et al. "Study on cooling of positronium for Bose-Einstein condensation". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 49.10 (Apr. 2016), p. 104001. DOI: 10.1088/0953-4075/49/10/104001.

[184] S. Aghion et al. (AEgIS collaboration), AEGIS Collaboration. "Positron bunching and electrostatic transport system for the production and emission of dense positronium clouds into vacuum". In: *Nucl. Instrum. Methods Phys. Res., Sect. B* 362 (2015), pp. 86–92.

[185] C. Amsler et al. "Velocity-selected production of $2\,^3$S metastable positronium". In: *Phys. Rev. A* 99 (3 Mar. 2019), p. 033405. DOI: 10.1103/PhysRevA.99.033405.

[186] K P Ziock et al. "Optical saturation of the 13S-23P transition in positronium". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 23.2 (Jan. 1990), p. 329. DOI: 10.1088/0953-4075/23/2/015.

[187] C. Zimmer et al. "Positronium laser cooling in a magnetic field". In: *Phys. Rev. A* 104 (2021), p. 023106.

[188] D. B. Cassidy et al. "Laser excitation of positronium in the Paschen-Back regime". In: *Phys. Rev. Lett.* 106 (2011), p. 173401.

[189] N. Gusakova et al. "An alexandrite laser system for positronium laser cooling". In: (submitted).

[190] Michaël Hemmer et al. "Volume Bragg Grating assisted broadband tunability and spectral narrowing of Ti:Sapphire oscillators". In: *Opt. Express* 17.10 (May 2009), pp. 8212–8219. DOI: 10.1364/OE.17.008212.

[191] D. B. Cassidy et al. "Efficient production of Rydberg positronium". In: *Phys. Rev. Lett.* 108 (2012), p. 043401.

[192] D. Nowicka et al. (AEgIS Collaboration). "Control system for ion Penning traps at the AEgIS experiment at CERN". In: *J. of Phys.: Conf. Ser.* 2374 (2022), p. 012038. DOI: 10.1088/1742-6596/2374/1/012038.

[193] M. Volponi, S. Huck et al. (AEgIS Collaboration). "CIRCUS: an autonomous control system for antimatter, atomic and quantum physics experiments". In: *EPJ quantum technology* (2023). submitted.

[194] D. B. Cassidy et al. "Single shot positron annihilation lifetime spectroscopy". In: *Appl. Phys. Lett.* 88 (2006), p. 194105.

[195] S. Aghion et al. (AEgIS collaboration), AEGIS Collaboration. "Producing long-lived 23S Ps via 33P laser excitation in magnetic and electric fields". In: *Phys. Rev. A* 98 (2018), p. 013402.

[196] D. B. Cassidy et al. "Positronium hyperfine interval measured via saturated absorption spectroscopy". In: *Phys. Rev. Lett.* 109 (2012), p. 073401.

[197] S. Mariazzi, P. Bettotti, and R. S. Brusa. "Positronium cooling and emission in vacuum from nanochannels at cryogenic temperature". In: *Phys. Rev. Lett.* 104 (2010), p. 243401.

[198] Christopher Corder, Brian Arnold, and Harold Metcalf. "Laser Cooling without Spontaneous Emission". In: *Phys. Rev. Lett.* 114 (4 Jan. 2015), p. 043002.

[199] John P. Bartolotta et al. "Laser cooling by sawtooth-wave adiabatic passage". In: *Phys. Rev. A* 98 (2 Aug. 2018), p. 023404.

[200] G. Kornakov et al. "Synthesis of Cold and Trappable Fully Stripped Highly Charged Ions via Antiproton-Induced Nuclear Fragmentation in Traps". In: *Physical Review C* 107.3 (Mar. 23, 2023), p. 034314. ISSN: 2469-9985, 2469-9993. DOI: 10.1103/PhysRevC.107.034314. URL: https://link.aps.org/doi/10.1103/PhysRevC.107.034314.

[201] Piotr Demski and Marco Volponi. *LabVIEW WireLength Calculator*. https://forums.ni.com/t5/Community-Documents/Useless-Code-Corner-Wires-Length/ta-p/3502746. 2024.