Neutrino physics with the PTOLEMY project: active neutrino properties and the light sterile case

M.G. Betti,$^{a,b}$ M. Biasotti,$^c$ A. Boscá,$^d$ F. Calle,$^d$ N. Canci,$^e$ G. Cavoto,$^{a,b}$ C. Chang,$^{f,g}$ A.G. Cocco,$^h$ A.P. Colijn,$^{i,j}$ J. Conrad,$^k$ N. D’Ambrosio,$^e$ N. De Groot,$^{i,l}$ P.F. de Salas,$^m$ M. Faverzani,$^n$ A. Ferella,$^k$ E. Ferri,$^n$ P. Garcia-Abia,$^o$ I. García-Cortés,$^p$ G. García Gomez-Tejedor,$^q$ S. Garfias,$^r$ F. Gatti,$^c$ C. Gentile,$^s$ A. Giachero,$^n$ J.E. Gudmundsson,$^k$ Y. Hochberg,$^l$ Y. Kahn,$^u$ A. Kievsky,$^v$ M. Lisanti,$^w$ C. Mancini-Terracciano,$^{a,b}$ G. Mangano,$^h$ L.E. Marcucci,$^{v,x}$ C. Mariani,$^{a,b}$ J. Martínez,$^d$ M. Messina,$^{e,y}$ A. Molinero-Vela,$^o$ E. Monticone,$^z$ A. Moroño,$^p$ A. Nuccioni,$^n$ F. Pandolfi,$^b$ S. Parlati,$^e$ S. Pastor,$^r$ J. Pedrós,$^d$ C. Pérez de los Heros,$^{aa}$ O. Pisanti,$^{h,ab}$ A.D. Polosa,$^{a,b}$ A. Puiu,$^n$ I. Rago,$^{a,b}$ Y. Raitses,$^s$ M. Rajteri,$^z$ N. Rossi,$^e$ I. Rucandio,$^o$ R. Santorelli,$^o$ K. Schaeffner,$^y$ C.G. Tully,$^w$ M. Viviani,$^v$ F. Zhao,$^w$ K.M. Zurek$^{f,ac}$

$^a$Università degli Studi di Roma La Sapienza, Roma, Italy
$^b$INFN Sezione di Roma, Roma, Italy
$^c$Università degli Studi di Genova e INFN Sezione di Genova, Genova, Italy
$^d$Universidad Politécnica de Madrid, Madrid, Spain
$^e$INFN Laboratori Nazionali del Gran Sasso, L’Aquila, Italy
$^f$Department of Physics, University of California, Berkeley, CA, USA
$^g$Argonne National Laboratory, Chicago, IL, USA
$^h$INFN Sezione di Napoli, Napoli, Italy
$^i$Nationale institut voor subatomaire fysica (NIKHEF), Amsterdam, Netherlands
$^j$University of Amsterdam, Amsterdam, Netherlands
$^k$Stockholm University, Stockholm, Sweden
$^l$Radboud University Nijmegen, Nijmegen, Netherlands
$^m$The Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, SE-10691 Stockholm, Sweden
$^n$Università degli Studi di Milano-Bicocca e INFN Sezione di Milano-Bicocca, Milano, Italy
Abstract. The PTOLEMY project aims to develop a scalable design for a Cosmic Neutrino Background (CNB) detector, the first of its kind and the only one conceived that can look directly at the image of the Universe encoded in neutrino background produced in the first second after the Big Bang. The scope of the work for the next three years is to complete the conceptual design of this detector and to validate with direct measurements that the non-neutrino backgrounds are below the expected cosmological signal. In this paper we discuss in details the theoretical aspects of the experiment and its physics goals. In particular, we mainly address three issues. First we discuss the sensitivity of PTOLEMY to the standard neutrino mass scale. We then study the perspectives of the experiment to detect the CNB via neutrino capture on tritium as a function of the neutrino mass scale and the energy resolution of the apparatus. Finally, we consider an extra sterile neutrino with mass in the eV range, coupled to the active states via oscillations, which has been advocated in view of neutrino oscillation anomalies. This extra state would contribute to the tritium decay spectrum, and its properties, mass and mixing angle, could be studied by analyzing the features in the beta decay electron spectrum.
1 Introduction

The Universe has expanded by a factor of over one billion between the early thermal epoch known as the neutrino decoupling stage and the present day. We observe this dynamics in many forms: the recession of galaxies (Hubble expansion), the dim afterglow of the hot plasma epoch, the Cosmic Microwave Background (CMB), and the abundances of light elements during Big Bang Nucleosynthesis (BBN). The epoch of neutrino decoupling produced another pillar of confirmation, the Cosmic Neutrino Background (CNB), perhaps one of the most important not yet directly probed predictions of the standard cosmological model. Because of the similarities shared with the CMB, its properties are theoretically expected to be very close to those of the photon background. In the so-called instantaneous decoupling limit, i.e. assuming that neutrino weak interactions become slower than the Hubble rate instantaneously, we have a simple expression for the average number density per neutrino state today

\[ n_0 = \frac{3\zeta(3)}{4\pi^2} T_{\nu,0}^3 = 56 \text{ cm}^{-3}, \]  

with a neutrino temperature \( T_{\nu,0} \approx 1.95 \text{ K} \), and where the subscript '0' refers to the present epoch. Actually, the CNB spectrum is expected to deviate from the one obtained considering a perfect Fermi-Dirac distribution at the percent level [1–3]. The reason is the partial overlap of the last stage of neutrino decoupling and the first instants of \( e^\pm \) annihilations in the primeval plasma. The same density is also shared by antineutrino species, assuming a vanishing neutrino-antineutrino asymmetry. This assumption is well justified since the ratio of neutrino asymmetry is strongly constrained by comparing the theoretical description of the BBN stage with astrophysical determinations of both primordial \(^4\)He and deuterium nuclei, 

\[ -0.71 \leq L_\nu \leq 0.054 \]  

[4, 5], where \( L_\nu \) is the total neutrino asymmetry, summed over the three flavors, normalized by the photon number density \( n_\gamma \). Actually, the neutrino asymmetry in the standard baryogenesis through leptogenesis model is expected to be much smaller than this bound, of the order of the baryon density today, \( n_B/n_\gamma \sim 6 \cdot 10^{-10} \).
The CNB acts as a source of gravity through its pressure and energy density, which are usually cast in terms of the effective number of relativistic neutrinos

\[ N_{\text{eff}} = \frac{8}{7} \left( \frac{11}{4} \right)^{4/3} \frac{\rho_\nu}{\rho_\gamma}, \tag{1.2} \]

where \( \rho_\nu \) (\( \rho_\gamma \)) is the neutrino (photon) energy density. The theoretical value for this parameter including the small non thermal distortion to the CNB is \( N_{\text{eff}} = 3.045 \) [2, 3], in good agreement with the bounds coming from BBN and Planck observations [6], which thus constrain any extra contribution to the radiation energy density due to exotic relativistic species and/or non standard features in the neutrino momentum distribution.

We know from neutrino flavor oscillation experiments, performed over more than thirty years [7–9], that at least two of the three standard neutrinos are massive particles. While the (squared) mass differences are very well measured today [10–12], we are still ignorant of the absolute mass scale, set by the lightest neutrino mass eigenstate, and the mass ordering, which is related to a possible hierarchy of the mass spectrum (see e.g. [12–14]). One of the best ways to measure the absolute neutrino mass scale is through the study of the endpoint of the \( \beta \) decay or the electron capture decay of some convenient atoms, such as tritium andholmium. The bound on the effective mass entering the tritium \( \beta \) decay, \( m_\beta \), obtained by studying the emitted electron energy spectrum near the endpoint, is presently \( m_\beta < 2 \) eV [15–17], but the KATRIN experiment, currently taking data, is expected to improve this bound by an order of magnitude in the near future, or obtain a measurement if \( m_\beta \) is at the level of fractions of eV [18, 19]. Other future experiments such as ECHo [20], HOLMES [21] or Project-8 [22] will also study the electron capture or \( \beta \) decay endpoint of holmium or atomic tritium, respectively, to obtain constraints on the neutrino mass scale. Interestingly, this range of neutrino masses can also be scrutinized using cosmological observations. We will present a short summary of this issue in the next Section, as well as the present status of oscillations, neutrino mass bounds and mass ordering. Moreover, since massive neutrinos became non-relativistic at a certain time, depending on their absolute masses, they can be gravitationally trapped under the effect of large enough gravitational potentials [23–25], and this enhances the local density at the Earth with respect to the homogeneous value \( n_0 \). This is quite important for their direct detection perspectives, since the signal rate is proportional to their density.

Experimental advances both in the understanding of massive neutrino physics and in techniques of high sensitivity instrumentation have opened up new opportunities to directly detect the CNB, an achievement which would profoundly confront and extend the sensitivity of precision cosmology data. The aim of the PTOLEMY project [26] is to develop a scalable design for a CNB detector, the first of its kind and the only one conceived that can look directly at the neutrino background. The scope of the work for the next three years is to complete the blueprint of the Cosmic Neutrino detector and to validate that the non-neutrino backgrounds are below the expected signal from the Big Bang with a direct measurement. An array of detectors of this design could reach discovery sensitivity for the CNB. The number and deployment of these detectors around the world will depend on the next phase of PTOLEMY developments, described in [26]. Yet, the physics case of the experiment is quite wide, including, as another major goal, the measurement of the standard neutrino absolute mass scale, in a way similar to that of the KATRIN experiment [18]. Moreover, some non standard scenarios could be tested, also in the preliminary phases of development of PTOLEMY. For example, an interesting issue concerns the physics of sterile neutrino states with masses.
in the eV range. Sterile eV neutrinos which mix with active states have been suggested since the LSND results [27] to solve some anomalies in neutrino oscillation experiments (short baseline data, reactor anomaly and gallium anomaly) [28–30] and their existence would have cosmological implications, see e.g. [31, 32].

In this paper we discuss in detail the theoretical aspects of the experiment and its neutrino physics goals. We study the sensitivity of PTOLEMY on neutrino mass detection, CNB detection and signals or bounds on sterile neutrinos with masses in the eV range, as a function of the expected energy resolution on the outgoing electrons and the employed mass of the tritium source. The scenario of keV neutrinos as warm dark matter candidates and their imprint on tritium spectrum, which appears well below the $Q$ value of the decay, will be considered elsewhere.

The paper is structured as follows. In Section 2 we review the to date information on neutrino masses, both from tritium decay and cosmological observables, mass ordering and oscillations, and on possible exotic neutrino states or properties, such as the possible existence of a light sterile neutrino. In Section 3 we introduce the formalism used in the analysis and describe how the finite electron energy resolution is taken into account. In Section 4 we illustrate our method, the Bayesian analysis we use to obtain forecasts of the experiment sensitivity. In Section 5 we then apply this analysis to the lightest neutrino mass parameter, showing the PTOLEMY discovery potential for the neutrino mass scale as a function of the tritium sample mass. We then study in Section 6 the perspectives of the experiment to detect the CNB via neutrino capture on tritium as a function of the neutrino mass scale and the energy resolution of the apparatus. The scenario of an extra sterile neutrino with mass in the eV range, coupled to the active states via oscillations, is considered in Section 7, where we describe how its properties, mass and mixing angle, could be constrained by analyzing the $\beta$ decay electron spectrum. Finally, we present our conclusions and outlooks in Section 8.

2 Theoretical context

As already introduced, the cleanest determination of the absolute scale of neutrino masses would proceed from a precise observation of the electron or positron spectrum close to the end point of $\beta$ decay. Current best limits on the effective electron antineutrino mass come from the observations of tritium decay in the Troitsk [15] and Mainz [16] experiments. As discussed in the next section, each neutrino mass eigenstate contributes to the suppression of the electron spectrum, but when the energy resolution is not sufficient to discriminate the different contributions it is safe to parameterize the effective neutrino $\beta$ decay mass as:

$$m_{\beta}^2 = \sum_i |U_{ei}|^2 m_i^2,$$

which is the experimentally determined quantity. It depends on the mixing matrix elements that describe the fraction of electron flavor for each mass eigenstate, $U_{ei}$, and on the mass of the $i$-th neutrino eigenstate, $m_i$. When the $m_i \simeq m_\nu$ are very similar and the neutrino masses are degenerate, the above definition becomes $m_{\beta} \simeq m_\nu$. The current best limit is $m_{\beta} < 2$ eV, at 95% CL [15–17]. In the incoming years, several experiments are expected to provide new and more stringent bounds. KATRIN [18, 19] started taking data in 2018 and the first results are expected soon, while the final sensitivity with 5 years of data will allow

---

1We assume neutrinos and antineutrinos share the same mass, although there is no experimental confirmation of this fact.
to reach $m_\beta \lesssim 0.2$ eV at 90% CL. Other experiments under development include Project-8 [22], which is expected to reach the sensitivity level of $m_\beta \lesssim 40$ meV at 90% CL using atomic tritium, and the two holmium experiments ECHo [20] and HOLMES [21], which aim at detecting the neutrino mass using the electron capture decay of $^{163}$Ho with a sensitivity around the eV. 

Another, indirect, probe of neutrino masses comes from cosmological observables. In fact, neutrinos have cooled during the expansion of the Universe and their presence and masses can be indirectly felt through the action of their diminishing thermal velocities on large-scale structure formation. In particular, Planck data constrain the sum of the three neutrino masses mainly via the lensing power spectrum and the lensing effects on CMB anisotropies. The bound on the sum of the neutrino masses is presently in the range $\sum_i m_i < (0.24 - 0.54)$ eV (95% CL) when considering CMB observations only [6], depending on the CMB data which are used in the analysis (see also [33, 34]). CMB probes are not the ideal way to constrain the neutrino masses, as neutrinos were relativistic at the time of photon decoupling and the effect of their masses is mainly imprinted through the late-time evolution of the CMB spectrum. For this reason, combinations of CMB data with low redshift probes such as determinations of the Baryon Acoustic Oscillations (BAO), of the matter power spectrum at late times or of other probes such as from the absorption spectrum measured from Lyman-\(\alpha\) forests provide a stronger constraint on the sum of the neutrino masses. When CMB observations are combined with BAO, for example, the limits tighten thanks to the combination of observations that are relevant at different epochs, and the limits become $\sum_i m_i < (0.12 - 0.16)$ eV (95% CL) [6]. From Lyman-\(\alpha\) data alone, on the other hand, a bound of $\sum_i m_i < 0.8$ eV (95% CL) is found [35], while Lyman-\(\alpha\) in combination with CMB observations leads to a limit on the sum of neutrino masses of $\sum_i m_i < (0.12 - 0.14)$ eV (95% CL) [35, 36]. Future observations of the matter power spectrum, such as those inferred by galaxy surveys like Euclid [37], will allow to obtain a determination of $\sum_i m_i$ to a precision of $\sim 0.02$ eV in combination with Planck [38]. One must always remember that cosmological constraints on the absolute neutrino mass are obtained indirectly under the assumption of a specific cosmological model. The results quoted above are for example derived assuming the simplest case, where the Universe evolution is described by the six parameters of the $\Lambda$CDM model plus the sum of the neutrino masses. When more parameters are varied, the limits can be relaxed up to a factor three (see e.g. [33, 39, 40]). Bayesian techniques allow to marginalize over the possible extensions of the minimal model, and a robust limit can be obtained relaxing the Planck ones in the $\Lambda$CDM+$\sum_i m_i$ model by approximately 50% [40].

Another unknown related to neutrino masses is their ordering, which can be normal if the lightest neutrino is the one with the largest mixing with the electron flavor, or inverted in the opposite case. The mass ordering is defined by the undetermined sign of the mass splitting that mostly affects atmospheric neutrino oscillations, $\Delta m^2_{31}$. The sign of such mass difference can be only determined by oscillation experiments which are sensitive to matter effects, such as atmospheric or long-baseline accelerator neutrino experiments. The neutrino mass ordering is very important to interpret the number of events which PTOLEMY can observe, as we will discuss in the next section. At present, the strongest preference for one of the two possible orderings by a single experiment comes from Super-Kamiokande, which favors normal ordering at $\sim 1.4\sigma$ [41]. Results from long-baseline oscillation experiments such as NO$\nu$A [42] or T2K [43], instead, provide a preference for normal ordering around the $2\sigma$ level (each) only when a prior on $\theta_{13}$ from reactor experiments is adopted. When all relevant neutrino oscillation experiments are combined in a global fit, the preference rises up
to $\sim 3\sigma$ \cite{10, 12, 13}. The addition of data from other probes (such as 0νββ, CMB, BAO or a prior on $H_0$) only improves slightly the preference towards a normal ordering \cite{14, 44}. A definite determination of the mass ordering is expected in the next few years when the currently ongoing experiments will improve the statistics, although is unlikely that any of them will be able to reject the wrong ordering alone. In the near future, however, the ORCA experiment from the KM3NeT collaboration will improve significantly the sensitivity to the neutrino mass ordering. ORCA, which is expected to deploy a working set of strings in the next years, can reach a 5σ preference in three years provided that the true ordering is normal \cite{45}. In the case of an inverted ordering, the reached sensitivity in three years will be of $\sim 3\sigma$, which will increase in combination with other experiments. The final mass ordering determination will be performed by DUNE, expected to start in 2026, which will reach the 5σ significance regardless of the true ordering after 7 years of data taking \cite{46}.

As we will detail in the next Section, the rate of CNB capture on a tritium nucleus, the process which PTOLEMY will exploit, can be written as

$$\Gamma_{\text{CNB}} = \sum_{i=1}^{N_\nu} |U_{ei}|^2 \bar{\sigma}_\nu f_{c,i} n_0,$$

(2.2)

where the sum is over neutrino mass eigenstates, $n_0$ is the average neutrino number density on large scales, see Eq. (1.1), $f_{c,i} \geq 1$ are the clustering factors, defined as the ratio between the local and the average number density $n_0$, that code the local overdensity of these particles due to the gravitational attraction of our galaxy, and $v_\nu$ the neutrino velocity in the Earth frame. Finally, the quantity $\bar{\sigma}$ is the average cross section for neutrino capture. As discussed in the following, the neutrino mass ordering enters the above equation through the various $U_{ei}$, which are different when considering a normal or inverted mass ordering.

Apart from the obvious dependence of the shape of emitted electron energy on neutrino masses, the latter also affects the $f_{c,i}$ parameters, which monotonically increase with $m_i$. Since at least two of the neutrino mass eigenstates are non relativistic in the recent times of structure formation, indeed, relic neutrinos tend to cluster in overdensity regions, such as the Milky Way or the Virgo cluster, to which our galaxy belongs. As shown in \cite{23–25} this leads to an increased capture rate at Earth, which can be larger by a factor 10-20% for neutrino masses around 60 meV or up to 200% for masses of 150 meV, also depending on the assumed matter profile of the halo where the clustering occurs. Focusing only on one single case for the Milky Way composition, one can write the clustering factor as a function of the neutrino mass using a power-law. Considering a generalized Navarro-Frenk-White \cite{47} profile and baryon content of the galaxy as in \cite{24}, one can find \cite{25}:

$$f_{c,i} = 76.5 \left(\frac{m_i}{\text{eV}}\right)^{2.21}.$$

(2.3)

The nature of neutrinos is still unknown, i.e. we ignore if they are Dirac or Majorana (self-conjugated) particles. As well known, neutrinoless double β decay is the most promising way to experimentally answer this issue. Interestingly, the expected event rate of CNB capture on tritium is in general larger for non-relativistic Majorana neutrinos with respect to the Dirac case \cite{48, 49}. The reason is that, when neutrinos become non-relativistic while free-streaming, helicity is conserved contrary to chirality. In the Dirac case, this leads to a population of half the original amount of left-handed neutrinos that are left-chiral, and therefore able to be captured in tritium, while in the Majorana case the original right-handed neutrinos also
contribute with a developed left-chiral component, which amounts to a twice larger local density of relic neutrinos that can be detected with respect to the Dirac case. The total effect is not always a factor two in the event rate of Majorana neutrinos with respect to Dirac ones. Depending on the mass hierarchy, if the lightest neutrino has an extremely small mass (below \(\sim 1\) meV) and is still relativistic today, its capture rate will be the same for both cases [49], and the increase in the total event rate is much smaller than two. A CNB detection with significant statistics would be a further way to understand the neutrino nature, though this goal seems quite demanding, in particular because the effect is only visible in the event rate, which also depend, for example, on the clustering factors.

The event rate would be also modified by the presence of neutrino interactions beyond the standard weak processes, as those predicted in most of the extended theoretical models where neutrinos acquire mass. The characteristics of these non-standard interactions (NSI) involving neutrinos depend on the specific model and the relevant operators [50, 51], but their existence could alter the event rate of relic cosmological neutrinos. For instance, in [52] it was shown that charged-current NSI involving Dirac neutrinos could change the capture rate on tritium in a PTOLEMY-like detector by a factor between 0.3 to 2.2, while in the Majorana case the possible variation would be restricted to a few percent. Thus, also this kind of hypothetical interactions would affect the information that can be inferred on the neutrino nature.

Let us now consider a more exotic scenario with an extra neutrino mass eigenstate \(\nu_4\), with mass around an eV, mainly mixed with a new sterile neutrino flavor \(^2\). The considered mass comes from the fact that a number of anomalies in neutrino oscillation experiments could be solved by the presence of a light sterile neutrino with a mass around this scale. The first anomaly was published by the LSND experiment [27], the result of which was immediately criticized for being incompatible with neutrino oscillations [53]. After the discovery of the gallium [54]\(^3\) and reactor [56] anomalies, more experimental efforts were planned to study more carefully the problem and obtain a final conclusion. Nowadays, the situation is unclear due to a tension between the disappearance measurements, including both the electron and muon neutrino channels, and the appearance observations. The tension arises from the fact that the reactor antineutrino experiments (discussed in details in the following paragraph) seem to prefer a non-zero mixing between active and sterile neutrinos through the mixing matrix element \(U_{e4}\), but in the muon neutrino disappearance channel, which is probed mainly by accelerator neutrino experiments as MINOS/MINOS+ [57, 58] or atmospheric neutrino detectors such as IceCube [59] and DeepCore [60], no oscillations have been observed. As a consequence, we have only upper limits on the mixing matrix element \(U_{\mu 4}\). On the other hand, neutrino appearance experiments as LSND [27] and MiniBooNE [61], which are sensitive to the product \(|U_{e4}|^2|U_{\mu 4}|^2\), observed an excess of neutrino events that cannot be explained using the values of \(U_{e4}\) and \(U_{\mu 4}\) inferred by disappearance probes. The tension is known since many years (see e.g. [29, 30, 32, 62]), but its significance has increase even more with the most recent MINOS+ and MiniBooNE results [63, 64].

Since in this paper we are only interested in the effect that a sterile neutrino may have on the \(\beta\) spectrum or on neutrino capture events, which can be described using only its mass

\(^2\) Since the active neutrino flavors are expected to have a small mixing with the fourth mass eigenstate in order to preserve the phenomenology of the three active neutrino mixing, the sterile neutrino flavor should mix almost only with the fourth mass eigenstate. For this reason it is approximately correct to say that the sterile neutrino has a mass \(m_s \simeq m_4 \simeq 1\) eV.

\(^3\) See also [55].
and the mixing with electron flavor ($U_{e4}$), we will only focus on the neutrino oscillation constraints that come from the electron (anti)neutrino disappearance channel, which is sensitive to the squared mass difference $\Delta m^2_{41}$ and the mixing matrix element $U_{e4}$. One among the best approaches to distinguish the effect of new neutrino oscillations from the presence of other systematic uncertainties, like for example a wrong theoretical spectrum of reactor antineutrinos, is to measure the flux at different distances from the source and to consider their ratios to do the analyses. Since this method decouples the neutrino oscillation effects from the theoretical description of the initial flux, it is referred to as “model-independent approach”. Nowadays, the strongest constraints come from the NEOS [65], DANSS [66], Neutrino-4 [67], PROSPECT [68] and STEREO [69] experiments, which all use a model-independent approach considering distances between 6 and 25 m from the reactor core. Among these experiments, Neutrino-4 is the one that claimed to have observed active-sterile oscillations with the highest significance [67], but the rather large values of the mixing parameters at the best-fit are excluded at more than 95% CL by PROSPECT [68]. The Neutrino-4 collaboration did not discuss their compatibility with the best-fit found in [66], for which the DANSS collaboration reports an improvement of the fit with respect to the standard three neutrino oscillation paradigm with a $\Delta \chi^2 \simeq 13$. This value corresponds to a $\simeq 2.8 \sigma$ preference for 3+1 neutrino oscillations according to the preliminary results presented at the Neutrino 2018 conference [70], and is in excellent agreement with the NEOS results [65] and the first PROSPECT results [68]. When the DANSS and NEOS results, which are not in tension with any other known observation, are analysed in a combined fit, the preferred value for the new mass splitting and mixing with the electron flavor are $\Delta m^2_{41} \simeq 1.29$ eV$^2$ and $|U_{e4}|^2 = s^2_{14} \simeq 0.012$ [63, 71]. Since more data are expected from the already mentioned experiments, a final result on the existence of short-baseline oscillations at reactors is expected soon.

From the cosmological point of view, data from both BBN and CMB are incompatible with a fully thermalized sterile neutrino with an eV mass and a mixing angle as required to solve the oscillation anomalies in the three neutrino picture (see e.g. [32, 72]). In fact, this would turn into a larger radiation content in the early universe, a faster expansion rate given by the Hubble factor, which would change both the amount of primordial deuterium (and to a less extent of $^4$He) produced during BBN [73], the relative height of the first acoustic peak and the damping tail in the CMB power spectrum [17, 29, 31]. In particular, a comparison of primordial deuterium observations with the theoretical analysis constrains the sterile state number density to a factor less than 0.8 (at 95% C.L.) with respect to the active neutrino one, $n_0 = 56$ cm$^{-3}$. We have obtained this result using the BBN public code PArthENoPE 2.0 [74] and the most recent determination of deuterium, $^2$H/H = (2.527 $\pm$ 0.030) $\cdot$ 10$^{-5}$ [75], exploiting the theoretical ab-initio calculation of the $d(p, \gamma)^3$He cross section from [76].

In the recent years, models have been considered to relax this tension between a relic density of sterile neutrinos and cosmological observables, using mechanisms which reduce the production of sterile states through oscillations. One possibility is to consider asymmetries in the active neutrino sector. In fact, while in a neutrino symmetric bath a thermal population of the sterile state would quickly grow, allowing for primordial neutrino asymmetries of order $L \nu \geq O(10^{-2})$ a self-suppression as well as a resonant sterile neutrino production can take place, depending on temperature and chosen parameters, see [77, 78]. This reduces the sterile neutrino contribution to the effective number of relativistic neutrinos $N_{\text{eff}}$, see Eq. (1.2). However, the active-sterile flavor conversions take place at later stages and this produces significant distortions in the electron (anti)neutrino spectra, which increase the $^4$He abundance in primordial nucleosynthesis [79].
Another possibility is to introduce sterile self-interaction processes, the so called secret interaction model. In this model the sterile states are coupled to a new $U(1)$ gauge boson [80, 81] or to a new pseudoscalar [82–84] with a mass much smaller than the $W$ boson. This new interaction induces a temperature dependent matter potential which suppresses the active-sterile mixing in the early universe and so their relic abundance. Yet, also in this scenario, significant distortions may be produced in the electron (anti)neutrino spectra, altering the abundance of light element during BBN, see [73, 85, 86].

3 Beta decay and neutrino capture

The PTOLEMY approach to detect the CNB exploits the neutrino capture processes on $\beta$-unstable nuclei [87–89], like the one with tritium

$$\nu_e + ^3\text{H} \rightarrow ^3\text{He} + e^-. \quad (3.1)$$

In fact, tritium has been chosen among other target candidates because of its availability, lifetime, high neutrino capture cross section and low $Q$ value [88]. The smoking gun signature of a relic neutrino capture is a peak in the electron spectrum above the $\beta$ decay endpoint.

Because flavor neutrino eigenstates are a composition of mass eigenstates with different masses, while propagating, relic neutrinos quickly decohere into those, in a time scale less than one Hubble time [90]. Therefore, the capture rate of relic neutrinos by tritium nuclei

$$\Gamma_{\text{CNB}} = \sum_{i=1}^{N_{\nu}} \Gamma_i, \quad (3.2)$$

must be computed from the capture rates, $\Gamma_i$, of the different neutrino mass eigenstates $\nu_i$:

$$\Gamma_i = N_T |U_{ei}|^2 \int \frac{d^3p_{\nu}}{(2\pi)^3} \sigma(p_{\nu}) v_\nu f_{\nu_i}(p_{\nu}) , \quad (3.3)$$

where $N_T = M_T/m_{^3\text{H}}$ represents the number of tritium nuclei in a sample of mass $M_T$ of this element, $U_{ei}$ are the mixing matrix elements, $p_{\nu}$ is the neutrino momentum, $v_\nu$ is the neutrino velocity as measured at Earth, $\sigma(p_{\nu})$ is the momentum-dependent cross section and $f_{\nu_i}(p_{\nu})$ is the momentum distribution function of the neutrino eigenstate $\nu_i$. Since the CNB distribution in the phase space is very narrow, the integral reduces to

$$\Gamma_i = N_T |U_{ei}|^2 \bar{\sigma} v_\nu f_{e,i} n_0 , \quad (3.4)$$

where $n_0$ is the average neutrino number density on large scales, see Eq. (1.1), and $f_{e,i}$ is the clustering factor, that is the local overdensity of these particles due to the gravitational attraction of our galaxy [23–25]. The quantity $\bar{\sigma}$ represents the average cross section for neutrino capture,

$$\bar{\sigma} = \frac{G_F^2}{2\pi \nu_\nu} F(Z, E_e) \frac{m_{^3\text{He}}}{m_{^3\text{H}}} E_e p_e (|F|^2 + g_A^2 |GT|^2) , \quad (3.5)$$

where $m_{^3\text{He}} \approx 2808.391$ MeV and $m_{^3\text{H}} \approx 2808.921$ MeV are the nuclear masses of the $^3\text{He}$ and $^3\text{H}$ nuclei, respectively, and $E_e$ ($p_e$) is the electron energy (momentum). The cross section

---

4 The nuclear masses $m_{^3\text{He}}$ and $m_{^3\text{H}}$ are related to the atomic masses $M_{^3\text{He}} \approx 2809.413$ MeV and $M_{^3\text{H}} \approx 2809.432$ MeV [91] according to $m_{^3\text{He}} = M_{^3\text{He}} - 2m_e + 24.58678$ eV and $m_{^3\text{H}} = M_{^3\text{H}} - m_e + 13.59811$ eV.
is written in terms of the “standard” Fermi ($F$) and Gamow-Teller ($GT$) matrix elements. The Fermi function $F(Z, E_e)$ describes the effect of the Coulomb attraction between a proton and the outgoing electron, which enhances the cross section. In order to account for this effect we use the approximation due to Primakoff and Rosen [93],

$$F(Z, E_e) = \frac{2\pi \eta}{1 - \exp(-2\pi \eta)},$$

(3.6)

where $\eta = Z\alpha E_e/p_e$, $Z = 2$ is the atomic number of $^3$He and $\alpha = 1/137.036$ [17] is the fine structure constant.

Notice the presence of the mixing matrix element $U_{ei}$ in the partial rate $\Gamma_i$. This is due to the fact that only electron neutrinos intervene in the process (3.1), while relic neutrinos are found in their mass eigenstates. In the usual 3 neutrino parameterization [17]

$$|U_{ei}|^2 = (c_{12}^2 s_{13}^2, s_{12}^2 c_{13}, s_{13}^2),$$

(3.7)

where $c_{jk} = \cos \theta_{jk}$ and $s_{jk} = \sin \theta_{jk}$, being $\theta_{jk}$ the corresponding mixing angle. In our case, we use the best fit values $s_{12}^2 = 0.32$, $s_{13}^2 = 2.16$ (2.22) $\times 10^{-2}$ for normal (inverted) ordering [10] (see also [11, 12]).

Because of the finite experimental energy resolution, the main background to the neutrino capture process comes from the most energetic electrons of the $\beta$ decay of tritium, since they can be measured with energies larger than the endpoint. To estimate the rate of such background, we need to account for the $\beta$ decay spectrum [94]

$$\frac{d\Gamma_\beta}{dE_e} = \frac{\bar{\sigma}}{\pi^2} N_T \sum_{i=1}^{N_\nu} |U_{ei}|^2 H(E_e, m_i).$$

(3.8)

Defining $y = E_{\text{end},0} - E_e - m_i$, with $E_{\text{end},0}$ the energy at the $\beta$ decay endpoint for massless neutrinos,

$$H(E_e, m_i) = \frac{1 - m_e^2/(E_e m_{3\text{He}})}{(1 - 2E_e/m_{3\text{He}} + m_e^2/m_{3\text{He}})^{3/2}} \sqrt{y + \frac{2m_i m_{\text{3He}}}{m_{3\text{He}}}} \times \frac{m_i}{m_{3\text{He}}} \left(y + m_i + m_{\text{3He}} + m_i\right).$$

(3.9)

To account for the experimental energy resolution $\Delta$, we introduce a smearing in the electron spectrum. This is done using a convolution of both the CNB signal and the $\beta$ decay spectrum with a Gaussian of full width at half maximum (FWHM) given by $\Delta$. The smeared neutrino capture event rate $\tilde{\Gamma}_{\text{CNB}}$ then reads

$$\frac{d\tilde{\Gamma}_{\text{CNB}}(E_e)}{dE_e} = \frac{1}{\sqrt{2\pi}(\Delta/\sqrt{8\ln 2})} \sum_{i=1}^{N_\nu} \Gamma_i \times \exp \left\{ -\frac{(E_e - (E_{\text{end}} + m_i + m_{\text{lightest}}))^2}{2(\Delta/\sqrt{8\ln 2})^2} \right\},$$

(3.10)

where $m_{\text{lightest}}$ is the mass of the lightest neutrino and $E_{\text{end}}$ is the energy at the $\beta$ decay endpoint, $E_{\text{end}} = E_{\text{end},0} - m_{\text{lightest}}$. In the same way, the smeared $\bar{\beta}$ decays reads

$$\frac{d\tilde{\Gamma}_\beta(E_e)}{dE_e} = \frac{1}{\sqrt{2\pi}(\Delta/\sqrt{8\ln 2})} \int_{-\infty}^{+\infty} dE' \frac{d\Gamma_\beta(E_e)}{dE_e} \exp \left[ -\frac{(E_e - E')^2}{2(\Delta/\sqrt{8\ln 2})^2} \right].$$

(3.11)
Figure 1. Expected event rates versus electron energy $E_e$ in a direct-detection experiment like PTOLEMY (assuming 100 g of tritium source) near the $\beta$ decay endpoint for different lightest neutrino masses and energy resolutions. Solid lines represent the total event rates convolved with a Gaussian envelope of FWHM equal to the assumed energy resolution, as computed from Eqs. (3.10) and (3.11). Dashed lines represent the signal event rates as it would be measured by the experiment without the background, while dotted lines show the background ($\beta$ decay) event rates without the convolution, i.e. for $\Delta = 0$. Red (blue) lines indicate normal (inverted) ordering. All lines are obtained considering Dirac neutrinos and neutrino overdensity according to the semi-analytic expression from [25].

The integrated number of signal events is expected to be, for Dirac neutrinos and ignoring the possible enhancement due to clustering, around 4 per year [48]. Depending on the neutrino masses and nature, this number can be enhanced. As already argued, for Majorana neutrinos the event rate can be a factor two larger, if neutrinos are massive enough to be all non-relativistic today or the correct mass ordering is the inverted one [49]. Concerning the enhancement due to the local relic neutrino density, it mostly depends on the mass of each mass eigenstate and on the mass ordering. The event rate may be unaltered if the mass ordering is inverted and the neutrinos are very light, it could be 10-20% larger for normal ordering and nearly minimal neutrino masses, or it may be significantly increased for masses above $\sim 100-150$ meV [24].

In Fig. 1 we show the expected event rates at energies close to the $\beta$ decay endpoint for different neutrino masses and energy resolutions, and comparing the two possible mass orderings, considering Dirac neutrinos and taking into account the neutrino overdensity ac-

\^5 For the form factors and the axial coupling we use $|F|^2 \approx 0.9987$, $|GT|^2 \approx 2.788$ and $g_A \approx 1.2695$ [92].
Neutrino capture events can be resolved from the $\beta$ decay background only for neutrino masses sufficiently larger than $\Delta$, as expected. In the inverted ordering case (blue) it is interesting to note a kink in the $\beta$ decay spectra due to the larger overlap of $\nu_e$ with the heaviest mass eigenstates. The feature can be hardly observed in the normal ordering (red) as $\nu_e$ has a much smaller mixing with $\nu_3$. This is also the reason why the CNB capture peaks have different heights, proportional to $|U_{ei}|^2$, see Eq. (3.10). For the smallest energy resolution we have considered, $\Delta = 10$ meV, the contributions due to the different mass eigenstates ($\nu_1$ plus $\nu_2$, and $\nu_3$) can be seen in the upper plots. In this case, it is also interesting to see that the peak due to the larger mass eigenstate, $\nu_3$ for normal and $\nu_1$ plus $\nu_2$ for inverted ordering, has a larger amplitude in the latter case, making the situation more favorable due to the fact that most of the interesting events have a larger separation from the $\beta$ decay background.

4 Data analysis method

To estimate the sensitivity of PTOLEMY to the neutrino mass scale we follow and adapt the procedure proposed in the KATRIN Design Report [18] and revisited from the Bayesian point of view in [95], see also [96]. We consider here in detail the standard active neutrino states, but the analysis can be easily extended to include an extra sterile state with mass in the eV range, see section 7. Following the notation adopted in the previous section, we define the number of $\beta$ decay and neutrino capture events within an energy bin centered at $E_i$ as

$$N^i_\beta = T \int_{E_i-\Delta/2}^{E_i+\Delta/2} \frac{d\tilde{\Gamma}_\beta}{dE_\nu} dE_\nu,$$

$$N^i_{\text{CNB}} = T \int_{E_i-\Delta/2}^{E_i+\Delta/2} \frac{d\tilde{\Gamma}_{\text{CNB}}}{dE_\nu} dE_\nu,$$

with $T$ the exposure time. In our Bayesian simulation we reconstruct the physical parameters given an initial fiducial model. We will indicate with hats the fiducial parameter values, while the quantities without hats refer to the varying parameters in the analyses. For the fiducial models we will select different values for lightest neutrino mass $\hat{m}_{\text{lightest}}$, while the other masses ($\hat{m}_i$) and mixing matrix ($\hat{U}$) parameters, as well as the true endpoint of the $\beta$ spectrum ($\hat{E}_{\text{end}}$), are fixed according to the currently known best fit values 6.

For the fiducial model, the number of expected events per energy bin is given by:

$$\hat{N}^i = N^i_\beta(\hat{E}_{\text{end}}, \hat{m}_i, \hat{U}) + N^i_{\text{CNB}}(\hat{E}_{\text{end}}, \hat{m}_i, \hat{U}).$$

The total number of events that will be measured in a bin is the sum of $\hat{N}^i$ and a constant background:

$$\hat{N}^i_t = \hat{N}^i + \hat{N}_b.$$
allow to distinguish the few signal events that are expected in the full-scale PTOLEMY configuration, but more detailed studies on the topic are left for future works where the detector characteristics will be considered in detail. We then estimate the experimental measurement in each energy bin using the Asimov dataset, i.e. with no statistical fluctuations around the number of events computed using the fiducial parameter values \[97\] :

\[ N_{\text{exp}}^i(\hat{E}_{\text{end}}, \hat{m}_i, \hat{U}) = \hat{N}_i^i \pm \sqrt{\hat{N}_i^i}, \quad (4.5) \]

assuming a statistical error of \(\sqrt{\hat{N}_i^i}\) in each bin. Systematic errors will be studied using dedicated Monte Carlo simulations once the detector design will be more defined.

The simulated measurement is fitted in order to reconstruct the values of the theoretical parameters that describe the physical model. We introduce a normalization uncertainty on the number of \(\beta\) events \((A_\beta)\), on the endpoint energy \((\Delta E_{\text{end}})\) and an unknown constant background \((N_b)\). For these parameters we use linear priors \(A_\beta \in [0, 2], \ln N_b \in [-1, 3]\) and \(\Delta E_{\text{end}} \in [-1, 1] \text{ eV}\) and their values will be determined by the fit. We additionally vary the mass of the lightest neutrino \((m_{\text{lightest}} \in [0, 1] \text{ eV})\), from which we compute the other mass eigenstates according to the mass splittings measured by current neutrino oscillation experiments, \(\Delta m^2_{31} = 7.55 \times 10^{-5} \text{ eV}^2\) and \(\Delta m^2_{31} = 2.50 \times 10^{-3} \text{ eV}^2\) for normal or \(\Delta m^2_{31} = -2.42 \times 10^{-3} \text{ eV}^2\) for inverted ordering \[10\].

In order to test the perspectives for CNB detection, we multiply the capture event number by an unknown normalization \(A_{\text{CNB}}\), whose fiducial value (i.e. the expected value in the standard theoretical scenario) is one, and for which we consider a linear prior \(A_{\text{CNB}} \in [0, 5]\). From the fitted value of \(A_{\text{CNB}}\) one can in principle extract information on the Dirac/Majorana nature of neutrinos, on the cross section dependence on NSI and on the neutrino clustering. As already mentioned, however, this task will be challenging due to the degeneracy of the various effects and will not be explored in this work, where we only assess the statistical reach of the PTOLEMY setup. A direct detection of the CNB (or a measure of the lightest neutrino mass) at a given C.L. can be claimed if the credible interval for \(A_{\text{CNB}}\) (or \(m_{\text{lightest}}\)) that at C.L. is found to be incompatible with zero. A more accurate test would require a comparison between the model with free \(A_{\text{CNB}}\) and the model with \(A_{\text{CNB}} = 0\), for example using the Bayes factor or a maximum likelihood ratio. We have checked that the results of the two methods are approximately equivalent, with the model comparison method based on the Savage-Dickey density ratio \[98\] giving slightly more pessimistic results. Since the sensitivity of the PTOLEMY experiment will be more precisely assessed only when we will know the systematic uncertainties related to the detector, we do not go in further details here.

For sake of brevity, in the following we will indicate the list of theoretical parameters with \(\theta = (A_\beta, N_b, \Delta E_{\text{end}}, A_{\text{CNB}}, m_i, U)\). The theoretical number of events in the bin \(i\) therefore reads

\[ N_{\text{th}}^i(\theta) = N_b + A_\beta N^\beta_b(\hat{E}_{\text{end}} + \Delta E_{\text{end}}, \hat{m}_i, U) \]

\[ + A_{\text{CNB}} N_{\text{CNB}}^i(\hat{E}_{\text{end}} + \Delta E_{\text{end}}, \hat{m}_i, U). \quad (4.6) \]

In order to perform the analysis and fit the desired parameters \(\theta\), we use a Gaussian \(\chi^2\) function:

\[ \chi^2(\theta) = \sum_i \left( \frac{N_{\text{exp}}^i(\hat{E}_{\text{end}}, \hat{m}_i, \hat{U}) - N_{\text{th}}^i(\theta)}{\sqrt{N_i^i}} \right)^2, \quad (4.7) \]
which will be converted into a likelihood function $L$ for the Bayesian analysis according to
\[ \chi^2 = -2\log L. \]
The Gaussian approximation is fully justified for the energy bins for which we have a large number of events, as expected for the $\beta$ spectrum. We checked that the presented results for the expected sensitivity on the CNB detection, which mostly come from bins with a small number of events, do not change when a Poissonian likelihood is considered instead:
\[
\ln \mathcal{L}(\theta) = \sum_i \left( N^i_{\text{exp}}(\hat{E}^{\text{end}}, \hat{m}_i, \hat{U}) \ln N^i_{\text{th}}(\theta) - N^i_{\text{th}}(\theta) - \ln \Gamma[N^i_{\text{exp}}(\hat{E}^{\text{end}}, \hat{m}_i, \hat{U}) + 1] \right). \quad (4.8)
\]
In the following series of simulations, we consider three possibilities for the detector mass. The full-scale PTOLEMY detector, aiming at the direct detection of the CNB, requires a tritium mass of at least 100 g, otherwise the signal event rate would be too small to be measurable. Such an amount of tritium is above the reach of current technology: the first phases of the experiment will therefore have a less ambitious goal, exploiting a smaller mass of tritium to test that the available techniques may allow to reach the final target. In order to demonstrate that the initial phases of the PTOLEMY project will offer interesting opportunities for studying neutrino properties, such as their mass and possibly the mass hierarchy, we will present some of the results considering both 1 g and 0.01 g of tritium mass. We adopt one year of data taking, an observed energy range between $\hat{E}^{\text{min}} = E_0 - 5$ eV and $\hat{E}^{\text{max}} = E_0 + 10$ eV and a constant background rate $\Gamma_b = 10^{-5}$ Hz over the whole energy range.

We have verified that increasing $\hat{E}^{\text{max}}$ has no impact on the results if the observed range is sufficient to cover the CNB events, while some effect may come from a different $\hat{E}^{\text{min}}$. If $\hat{E}^{\text{min}}$ is decreased, the precision in measuring the $\beta$ spectrum (its normalization and the endpoint) allows to slightly improve the sensitivity on the neutrino parameters, but this comes at the price of a larger number of events, which might be difficult to handle. On the other hand, an $\hat{E}^{\text{min}}$ closer to the endpoint allows to reduce the $\beta$ decay event rate at the expense of slightly worsening the precision on the neutrino mass determination. The best value for $\hat{E}^{\text{min}}$ will be determined once the technical properties of the apparatus are defined more precisely.

A final comment is about the constant background rate $\hat{\Gamma}_b$. For an amount of tritium of 100 g, the number of events expected from the $\beta$ decay is much larger than the background rate and the determination of the neutrino masses or the detection of a putative sterile neutrino will be possible even with a much larger $\hat{\Gamma}_b$. In other words, a much smaller tritium mass might be sufficient to measure the $\beta$ spectrum over the background and achieve sensitivity to neutrino mass scale, as we will discuss later. Yet, a 100 g target mass and $\hat{\Gamma}_b \lesssim 10^{-5}$ Hz are crucial to allow a detection of the relic neutrinos, due to the extremely small cross section.

To perform the analysis, we have adapted the generic Markov Chain Monte Carlo (MCMC) sampler used in CosmoMC [99]. The theoretical parameters that we will try to reconstruct are: the lightest neutrino mass $m_{\text{lightest}}$, the normalization of the signal spectrum, the mass ordering, and, for scenarios with an extra sterile neutrino state, the squared mass difference $\Delta m^2_{41} \in [0, 20]$ eV$^2$ and the mixing angle $s_{14}^2 = |U_{e4}|^2 \in [0, 1]$.

5 Neutrino mass sensitivity and mass ordering

The full-scale PTOLEMY experiment is expected to have an impressive performance in reconstructing the fiducial value for the lightest neutrino mass, thanks to the large amount of tritium and the correspondingly large statistics. The 1σ statistical error obtained from the simulations is of the order of $10^{-3}$ eV or below, with minimal dependences on the detector
configurations described in [26], energy range and background rate. Reasonably, the error slightly depends on the energy resolution of the experiment and on the value of the fiducial lightest neutrino mass, with smaller relative errors for larger masses. This can be seen in Fig. 2, where we show relative statistical errors obtained when reconstructing a given fiducial lightest neutrino mass \( \hat{m}_{\text{lightest}} \), for different energy resolutions \( \Delta \) and 100 g yr of PTOLEMY data (lower panel). As we can see, PTOLEMY may distinguish the neutrino mass almost independently of the experimental energy resolution, which has a very small impact on the statistical error on \( m_{\text{lightest}} \). The reason is that a larger lightest neutrino mass does not only induce a shift in the endpoint of the \( \beta \) decay spectrum, but also a change in the normalization of the spectrum at all energies, which can be measured very well thanks to the very large event rate.

It is worth mentioning that already in the possible initial configurations of the detector, with lower tritium masses, PTOLEMY may have the ability to measure the neutrino mass. Considering scenarios with only 10 mg (upper panel) or 1 g (central panel) of tritium, PTOLEMY has the statistical reach for a determination of the neutrino mass even for \( m_{\text{lightest}} = 10 \) meV, for which a 30% relative error on the true value of \( m_{\text{lightest}} \) could be obtained, even with 10 mg of tritium and in the pessimistic case of \( \Delta \sim 125 \) meV.

Another interesting result that PTOLEMY can obtain is the determination of the neutrino mass ordering. This is due to the fact that the shape of the \( \beta \) spectrum near the endpoint depends on the single mass eigenstates and on the mixing matrix elements as described in eq. (3.8) and already shown in Fig. 1. Currently, neutrino oscillation data prefer normal ordering (NO, \( \Delta m_{31}^2 > 0 \)) over the inverted one (IO, \( \Delta m_{31}^2 < 0 \)), with a preference of more than 3\( \sigma \) [14]. For this reason, we will mostly focus on the NO case.

To estimate the capabilities of PTOLEMY in determining the mass ordering, we assume as fiducial values the best-fit mixing parameters obtained within NO [10]. We then fit the simulated experimental data using both the NO and IO best-fit mixing parameters and we compute the Bayesian evidence \( Z \). We will then have two cases: fiducial NO fitted using NO (\( \hat{\text{NO/NO}} \) for sake of brevity) and fiducial NO fitted using IO (\( \hat{\text{NO/IO}} \)). The PTOLEMY sensitivity on the mass ordering is then obtained using the Bayes factor:

\[
\ln B_{ij} = \ln Z_i - \ln Z_j .
\]

The magnitude of the Bayes factor provides the strength of the preference for one of the two cases, while the sign of \( \ln B_{ij} \) indicates which of the two cases is preferred (case \( i \) if \( \ln B_{ij} > 0 \), case \( j \) if \( \ln B_{ij} < 0 \)). If PTOLEMY is able to distinguish the two orderings, we expect the fit performed using the same case as the fiducial choice (\( \hat{\text{NO/NO}} \)) to be better than the one that assumes a different ordering with respect to the fiducial one (\( \hat{\text{NO/IO}} \)):

\[
\ln B_{\text{NO,NO}, \text{IO}} \equiv \ln Z_{\text{NO/NO}} - \ln Z_{\text{NO/IO}} > 0 .
\]

The significance of the preference in favor of NO can be quantified using the absolute value of the Bayes factor. In terms of its logarithm, the preference for NO is equivalent to a 3\( \sigma \) probability against IO if \( \ln B_{\text{NO,IO}} \simeq 6 \), to 4\( \sigma \) if \( \ln B_{\text{NO,IO}} \simeq 10 \) or to \( \gtrsim 5\sigma \) if \( \ln B_{\text{NO,IO}} \gtrsim 15 \) [14]. In Fig. 3 we show \( \ln B_{\text{NO,IO}} \) as a function of the fiducial lightest neutrino mass \( \hat{m}_{\text{lightest}} \) and the energy resolution \( \Delta \). As we can see, PTOLEMY will be able to determine the

\footnote{For a review on Bayesian model comparison see e.g. [100], for its application in determining the neutrino mass ordering see [14, 44].}
mass ordering (if it is normal) in the non-degenerate region, while the distinction will not be feasible ($\ln B_{NO,IO}$ is inconclusive) for neutrino masses above $\sim 0.18$ eV. Even if the
excellent sensitivity of PTOLEMY for the mass ordering may be unexpected, it has a very simple explanation. The $\beta$ decay spectrum near the endpoint is significantly different for the normal and inverted ordering cases, due to the different role of the mixing matrix elements. The consequence is that when the lightest neutrino has a small mixing with the electron flavor (in the IO case), the number of events that one can observe close to the endpoint is significantly suppressed. Here we show only the estimates obtained for 100 g of tritium, but similar conclusions could be obtained when dealing with intermediate-scale experimental configurations. Considering only statistical errors, the count rate can change up to two orders of magnitude in the interesting region, as one can see in Fig. 4, where we compare the spectra obtained with different lightest neutrino masses and energy resolutions for normal (red) and inverted (blue) ordering. As expected, when the mass or the energy resolution are larger the difference between the two spectra diminishes, but not enough, if the neutrino mass is sufficiently small, to decrease below the statistical error and consequently completely lose the sensitivity to the mass ordering.

\section{CNB detection}

In this Section we investigate the possibility to detect the CNB capture events. As already mentioned, we fit the signal from CNB capture using a free normalization $A_{\text{CNB}}$, see eq. (4.6), and we can claim a detection if $A_{\text{CNB}}$ can be distinguished from zero. Figure 5 shows the C.L. which can be achieved as a function of the different fiducial lightest neutrino masses and energy resolutions. As we can see, it is crucial to achieve a very good energy resolution, but this may be not enough if the neutrino masses are very small and the ordering of the mass eigenstates is normal. While smaller amounts of tritium may be sufficient to study the

\textbf{Figure 3.} Statistical significance for the determination of the neutrino mass ordering, if the NO is assumed as true, as a function of the fiducial lightest neutrino mass $\hat{m}_{\text{lightest}}$ and the energy resolution $\Delta$, considering 100 g yr of PTOLEMY data. Positive values of $\ln B_{\text{NO,IO}}$ correspond to a preference for NO, which is statistically decisive ($\gtrsim 5\sigma$) if $\ln B_{\text{NO,IO}} \gtrsim 15$.\vspace{2cm}
Figure 4. Comparison of the electron spectra obtained assuming normal (red) or inverted (blue) ordering, for different values of the fiducial lightest neutrino mass $m_{\text{lightest}}$ (10 or 100 meV, left to right) and the energy resolution $\Delta$ (50 or 150 meV, top to bottom), considering 100 g yr of PTOLEMY data. Dashed lines show the signal energy spectrum from neutrino capture, given the considered energy resolution. Dotted lines indicate the true $\beta$ decay energy spectrum, as it would be measured with a perfect energy determination.

For the neutrino mass spectrum, experimental configurations with less than 100 g of tritium are not suitable for CNB searches, due to the too low event rate.

The situation does not change significantly if one takes into account the possible enhancement of the event rate due to the clustering of relic neutrinos in the local dark matter halo, or other effects that could increase the cross section of the process, such as a Majorana nature of neutrinos or the presence of NSI. The crucial point, in fact, is that these factors could help to increase the number of observed signal events only if the energy resolution allows to distinguish them from the $\beta$ decay background, which has a many orders of magnitude larger rate.

We already noticed that direct detection of relic neutrinos is generally easier for inverted than for normal mass ordering, when the neutrino masses are small and the energy resolution is sufficiently good. This is due to the fact that the primary CNB peaks are shifted at higher electron energies, because $m_1$ and $m_2$ are larger (see Fig. 1). As a result, the perspectives of CNB detection at small neutrino masses are slightly improved in IO with respect to NO, see the bottom panel of Fig. 5.
Figure 5. Statistical significance for the detection of the CNB as a function of the fiducial lightest neutrino mass $\hat{m}_{\text{lightest}}$ and the energy resolution $\Delta$, considering 100 g yr of PTOLEMY data. The top (bottom) panel represents normal (inverted) ordering of the neutrino mass eigenstates.

7 Sterile neutrinos

In this last section we consider the perspectives for detection of a putative light sterile neutrino at the eV scale as introduced in Section 2. It is interesting to scrutinize the information that can be obtained from the distortions they induce on tritium decay spectrum as well as the effect of their capture. For the latter case the feasibility of a direct measurement of cosmological sterile states is strongly related to the theoretical model under consideration for the thermalization, which determines the average number density of the fourth mass eigenstate
that appears in Eq. (3.4), and to the calculation of the clustering effect of the fourth neutrino, which might be quite large [24]. Since the thermalization according to standard oscillations is disfavored [72], we will not explore the possibility of achieving a direct detection of relic $\nu_4$ with PTOLEMY in this work.

On the other hand, measuring the $\beta$ spectrum would be extremely useful to constrain the new squared mass difference $\Delta m_{41}^2$ and mixing angle $s^2_{14}$ through the suppression of the spectrum at energies above $\sim E_0 - \sqrt{\Delta m_{41}^2}$, emerging from eq. (3.8). Considering the best-fit results from [63, 71] as fiducial values in our analysis, we find that PTOLEMY would be able to confirm the presence of the sterile neutrino or reject its existence in case of no observation. Considering a fiducial model with $\Delta m_{41}^2 = 0$ and $s^2_{14} = 0$, we can get a marginalized 3$\sigma$ limit $s^2_{14} \lesssim 10^{-4}$ on the relevant mixing angle. In this case it is very useful to be able to measure a larger fraction of the $\beta$ decay energy spectrum, since the suppression corresponding to the fourth neutrino starts to be relevant at energies $\sim E_0 - \sqrt{\Delta m_{41}^2}$, and what is crucial, thus, is to determine the normalization below and above this point. This helps to discriminate the contribution of the sterile neutrino from other effects.

In Figs. 6 (for the squared mass difference $\Delta m_{41}^2$) and 7 (for the mixing angle $\sin^2 \theta_{14}$) we show the prospects for detecting a light fourth neutrino mass eigenstate with a small mixing with the electron neutrino, considering a wide range of fiducial new squared mass differences and mixing angles. For very small mixing angles PTOLEMY will hardly distinguish the effect of a new kink in the $\beta$ spectrum, while a detection will be clearly possible given the current preferred values of the best-fit mixing parameters (see e.g. [63, 64, 71, 101]), shown in black in the figure, even in the case of a small detector with only 10 mg of tritium (top panel). We also show in red the expected KATRIN sensitivity [102], which is only marginally covering the present best-fit regions. The increase in statistics related to the larger tritium mass translates in a better sensitivity also for a light sterile neutrino search.

8 Conclusions and outlook

In this paper we have studied the discovery potential of an experiment like PTOLEMY on several aspects of neutrino physics. The PTOLEMY project has been first conceived as a unique possibility to detect the CNB, a robust prediction of the standard hot big bang model, which has been indirectly proven by several cosmological observables, like BBN and CMB anisotropies. Yet, since neutrinos are only weakly interacting and because of the very small kinetic energy of the CNB, their direct detection is still an extremely demanding challenge.

In fact, being based on the detection of neutrinos captured by tritium nuclei, a process with no energy threshold, PTOLEMY has a vast physics case, able to detect very low energy fluxes, and to provide constraints on neutrino properties, those of standard ones as well as of exotic sterile species. Furthermore, it was also realized that the PTOLEMY setup (but without tritium) could be a way to observe dark matter particles with masses in the MeV mass range, keeping track of their arrival directions [103, 104], while the complete detector setup (with tritium) can also serve as a “laboratory” to test “ab-initio” theoretical predictions of few-nucleon systems, and therefore ultimately of the models for nuclear interactions and weak currents [105].

Our analysis has been focused on three issues. What is the PTOLEMY sensitivity to the neutrino mass scale? What is its discovery potential of CNB? What are the prospects to constrain (or exclude) extra sterile neutrino states which are mixed with standard ones, and have a mass in the eV range? Exploiting a Bayesian approach, we have analyzed the
Figure 6. Statistical significance for a detection of $\Delta m^2_{41}$ from measurements of the $\beta$ spectrum, assuming various fiducial values for the new squared mass difference and mixing angle, considering 10 mg yr (top panel), 1 g yr or 100 g yr (bottom panel) of PTOLEMY data. Black contours denote the 3$\sigma$ constraints from NEOS and DANSS [71], while the red line shows the 90% CL sensitivity which is expected for KATRIN [102].
role of the two main parameters which affect the sensitivity of a PTOLEMY-like experiment, namely the emitted electron energy resolution and the tritium sample mass, forecasting an answer to each of the questions above.

The absolute neutrino mass scale is still an unknown parameter, though cosmological observations and present results from tritium decay constrain the sum of the neutrino masses

Figure 7. The same as Fig. 6, but for $\sin^2 \theta_{14}$. 

-21-
to be less than 0.12 – 0.6 [6, 33, 34, 40] or 6 eV [17], respectively. The KATRIN experiment is expected to provide a limit which enters into the region for the neutrino mass that is still marginally allowed by cosmology. In view of the larger tritium mass which would be used, this bound could be further improved by PTOLEMY, at the level of tens of meV or less with a 1 g tritium mass, independently of the electron energy resolution, provided \( m_{\text{lightest}} \) is in the range 10-150 meV. Moreover, we have also found that there are interesting opportunities to check the neutrino mass hierarchy, from a different and independent approach than oscillation experiments. This can be achieved not only with the largest tritium mass considered in this paper, 100 g, but also with intermediate smaller amounts of order of grams or fractions of grams.

Of course the most ambitious goal of PTOLEMY is the detection of the CNB. We have performed an analysis of the detection sensitivity as a function of the tritium mass and the energy resolution. A tritium mass of the order of 100 g, unless the neutrino local clustering is much larger than what is obtained by simulations, is required. We report the discovery potential as function of the lightest neutrino mass and energy resolution, as shown in Fig. 5. As expected, these two parameters are almost linearly correlated: a smaller electron energy resolution implies a better constraint on smaller neutrino mass scales, and, for example, with \( \Delta \sim 100 \) meV, it would be possible to detect the CNB at \( \sim 2 \sigma \) if \( m_{\text{lightest}} \gtrsim 120 \) meV. Despite the demanding technological issues in dealing with a large tritium mass, as well as in achieving high energy resolutions (which, however, are not so far from present values), we note that there are no other feasible approaches to directly unveil the CNB. The only other plausible one, the Stodolsky effect [106], is much more challenging to be detected, if not simply impossible, if cosmological neutrinos have zero (or exceedingly small) chemical potential.

We have also considered the case of an extra sterile neutrino state with a mass of order eV and quite large mixing with active neutrinos. This sterile particle may help in solving the anomalies found in oscillation experiments, though the picture is still puzzling, since appearance and disappearance observations in short baseline experiment are only marginally in agreement once the 3+1 neutrino scenario is assumed [63]. Cosmological data strongly disfavor a relic density of these sterile neutrinos of the order of the standard ones. Since their contribution to the neutrino capture rate is also suppressed by their mixing angle \( \sin^2 \theta_{14} \sim 0.012 \), it seems difficult that they could be measured by their capture on tritium nuclei. It is much more promising that the features they would induce on the standard \( \beta \) decay spectrum could give more stringent constraints on their mass and mixing angles. In particular, we found that a sterile neutrino “detection” could be achieved by PTOLEMY, given the current preferred values of the best-fit mixing parameters, even in the case of a small detector with only 10 mg of tritium. In case of no observation, strong limits on the mixing angles will be derived.

In the future we plan to address a similar analysis for the case of a sterile neutrino with mass in the keV range, which has been considered as a warm dark matter candidate. If we assume that the whole dark matter is made by such particles, their local energy density would be of the order of GeV cm\(^{-3}\), corresponding to a rather large local number density of \( 10^5 \) cm\(^{-3}\), while their average density on cosmological scales is five orders of magnitude smaller. This means that the sterile states cannot have been produced in equilibrium in the early universe. This, together with astrophysical constraints, bounds the sterile-active mixing angle to be very small, \( \sin^2 \theta_{14} \lesssim 10^{-8} \) [107]. We thus expect the sterile neutrino capture signal to be too small to be detected by PTOLEMY, while the analysis of the much larger event
number expected in the $\beta$ decay spectrum might provide bounds on both $m_4$ and mixing angles. For this analysis, however, a larger fraction of the $\beta$ spectrum needs to be measured, as an observation around the endpoint is not sufficient to single out the kink corresponding to a mass in the keV range and the sterile neutrino presence would be degenerate with the normalization of the $\beta$ spectrum.

Another interesting issue, which deserves further studies, is to use PTOLEMY-like experiments as a way to check the models for nuclear interactions and currents. The three-nucleon bound systems involved in the PTOLEMY experiment have been the object of intense theoretical studies since many years. They represent an ideal “laboratory” to test our understanding of how nucleons interact among themselves, as well as with external electroweak probes. In order to do so, the three-body Schrödinger equation has to be solved exactly. Nowadays, a variety of methods exists to this aim: among the most accurate ones we can list the so-called Hyperspherical Harmonics (HH) method (see Ref. [108] and references therein). Historically, the models for the nuclear interaction and currents have been derived within two different frameworks: a purely phenomenological one, and, more recently, the so-called chiral effective field theory approach ($\chi$EFT) (see Ref. [109] and references therein). The theoretical results for the cross section obtained within this “ab-initio” approach for the tritium decay and neutrino capture process will represent predictions which the PTOLEMY experiment results can check and validate for the unpolarized and, more interestingly, for the polarized rate.

Acknowledgments

Work supported by the Italian grant 2017W4HA7S “NAT-NET: Neutrino and Astroparticle Theory Network” (PRIN 2017) funded by the Italian Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR); by the Spanish grants ENE2016-76755-R, SEV-2014-0398 and FPA2017-85216-P (AEI/FEDER, UE), PROMETEO/2018/165 (Generalitat Valenciana), María de Maeztu Unit of Excellence CIEMAT - Particle Physics (MDM-2015-0509) and the Red Consolider MultiDark FPA2017-90566-REDC; by the Vetenskapsrådet (Swedish Research Council) through contract No. 638-2013-8993; by the Simons Foundation, USA (#377485) and John Templeton Foundation, USA (#58851); and by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie individual Grant Agreement No. 796941.

References


– 23 –


