MAJOR REPORT • OPEN ACCESS

KATRIN: status and prospects for the neutrino mass and beyond

To cite this article: M Aker *et al* 2022 *J. Phys. G: Nucl. Part. Phys.* **49** 100501

View the article online for updates and enhancements.

Major Report

KATRIN: status and prospects for the neutrino mass and beyond

M Aker[1], M Balzer[2], D Batzler[1], A Beglarian[2], J Behrens[1],
A Berlev[3], U Besserer[1], M Biassoni[4], B Bieringer[5], F Block[6],
S Bobien[7], L Bombelli[8], D Bormann[2], B Bornschein[1],
L Bornschein[1], M Böttcher[5], C Brofferio[4,9], C Bruch[10,11],
T Brunst[10,11], T S Caldwell[12,13], M Carminati[14,15],
R M D Carney[16], S Chilingaryan[2], W Choi[6], O Cremonesi[4],
K Debowski[17], M Descher[6], D Díaz Barrero[18], P J Doe[19],
O Dragoun[20], G Drexlin[6], F Edzards[10,11], K Eitel[1],
E Ellinger[17], R Engel[1], S Enomoto[19], A Felden[1], D Fink[11],
C Fiorini[14,15], J A Formaggio[21], C Forstner[10,11],
F M Fränkle[1], G B Franklin[22], F Friedel[1], A Fulst[5], K Gauda[5],
A S Gavin[12,13], W Gil[1], F Glück[1], A Grande[14,15], R Grössle[1],
M Gugiatti[14,15], R Gumbsheimer[1], V Hannen[5], J Hartmann[2],
N Haußmann[17], K Helbing[17], S Hickford[1], R Hiller[1],
D Hillesheimer[1], D Hinz[1], T Höhn[1], T Houdy[10,11], A Huber[1],
A Jansen[1], C Karl[10,11], J Kellerer[6], P King[14,15], M Kleifges[2],
M Klein[1], C Köhler[10,11], L Köllenberger[1], A Kopmann[2],
M Korzeczek[6], A Kovalík[20], B Krasch[1], H Krause[1],
T Lasserre[23], L La Cascio[6], O Lebeda[20], P Lechner[24],
B Lehnert[16], T L Le[1], A Lokhov[3,5], M Machatschek[1],
E Malcherek[1], D Manfrin[14,15], M Mark[1], A Marsteller[1],
E L Martin[12,13], E Mazzola[4,9], C Melzer[1], S Mertens[10,11],
J Mostafa[2], K Müller[1], A Nava[4,9], H Neumann[7], S Niemes[1],
P Oelpmann[5], A Onillon[10,11], D S Parno[22,*], M Pavan[4,9],
A Pigliafreddo[14,15], A W P Poon[16], J M L Poyato[18],
S Pozzi[4,9], F Priester[1], M Puritscher[6], D C Radford[25],
J Rális[20], S Ramachandran[17], R G H Robertson[19],
W Rodejohann[26], C Rodenbeck[5], M Röllig[1], C Röttele[1],
M Ryšavý[20], R Sack[1,5], A Saenz[27], R W J Salomon[5],
P Schäfer[1], L Schimpf[5,6], K Schlösser[1], M Schlösser[1,*],
L Schlüter[10,11], S Schneidewind[5], M Schrank[1], A-K Schütz[16],
A Schwemmer[10,11], A Sedlak[11], M Šefčík[20], V Sibille[21],

* Author to whom any correspondence should be addressed.

D Siegmann[10,11], **M Slezák**[10,11], **F Spanier**[28], **D Spreng**[10,11], **M Steidl**[1], **M Sturm**[1], **H H Telle**[18], **L A Thorne**[29], **T Thümmler**[1], **N Titov**[3], **I Tkachev**[3], **P Trigilio**[8], **K Urban**[10,11], **K Valerius**[1], **D Vénos**[20], **A P Vizcaya Hernández**[22], **P Voigt**[10,11], **C Weinheimer**[5], **S Welte**[1], **J Wendel**[1], **C Wiesinger**[10,11], **J F Wilkerson**[12,13], **J Wolf**[6], **L Wunderl**[10,11], **S Wüstling**[2], **J Wydra**[1], **W Xu**[21], **S Zadoroghny**[3] and **G Zeller**[1]

[1] Institute for Astroparticle Physics (IAP), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

[2] Institute for Data Processing and Electronics (IPE), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

[3] Institute for Nuclear Research of Russian Academy of Sciences, 60th October Anniversary Prospect 7a, 117312 Moscow, Russia

[4] INFN-Sezione di Milano-Bicocca, 20126 Milano, Italy

[5] Institute for Nuclear Physics, University of Münster, Wilhelm-Klemm-Str. 9, 48149 Münster, Germany

[6] Institute of Experimental Particle Physics (ETP), Karlsruhe Institute of Technology (KIT), Wolfgang-Gaede-Str. 1, 76131 Karlsruhe, Germany

[7] Institute for Technical Physics (ITEP), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

[8] XGLab srl, Bruker Nano Analytics, Via Conte Rosso 23, 20134 Milano, Italy

[9] Università di Milano-Bicocca, Dipartimento di Fisica, 20126 Milano, Italy

[10] Technische Universität München, James-Franck-Str. 1, 85748 Garching, Germany

[11] Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München, Germany

[12] Department of Physics and Astronomy, University of North Carolina, Chapel Hill, NC 27599, United States of America

[13] Triangle Universities Nuclear Laboratory, Durham, NC 27708, United States of America

[14] DEIB, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

[15] INFN-Sezione di Milano, 20133, Milano, Italy

[16] Institute for Nuclear and Particle Astrophysics and Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States of America

[17] Department of Physics, Faculty of Mathematics and Natural Sciences, University of Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany

[18] Departamento de Química Física Aplicada, Universidad Autonoma de Madrid, Campus de Cantoblanco, 28049 Madrid, Spain

[19] Center for Experimental Nuclear Physics and Astrophysics, and Dept. of Physics, University of Washington, Seattle, WA 98195, United States of America

[20] Nuclear Physics Institute, Czech Academy of Sciences, 25068 Řež, Czech Republic

[21] Laboratory for Nuclear Science, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, United States of America

[22] Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, United States of America

[23] IRFU (DPhP & APC), CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France

[24] Halbleiterlabor of the Max Planck Society, Otto-Hahn-Ring 6, 81739 München, German

[25] Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States of America
[26] Max-Planck-Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany
[27] Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, 12489 Berlin, Germany
[28] Institute for Theoretical Astrophysics, University of Heidelberg, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
[29] Institut für Physik, Johannes-Gutenberg-Universität Mainz, 55099 Mainz, Germany

E-mail: dparno@cmu.edu and magnus.schloesser@kit.edu

Abstract

The Karlsruhe Tritium Neutrino (KATRIN) experiment is designed to measure a high-precision integral spectrum of the endpoint region of T_2 β decay, with the primary goal of probing the absolute mass scale of the neutrino. After a first tritium commissioning campaign in 2018, the experiment has been regularly running since 2019, and in its first two measurement campaigns has already achieved a sub-eV sensitivity. After 1000 days of data-taking, KATRIN's design sensitivity is 0.2 eV at the 90% confidence level. In this white paper we describe the current status of KATRIN; explore prospects for measuring the neutrino mass and other physics observables, including sterile neutrinos and other beyond-Standard-Model hypotheses; and discuss research-and-development projects that may further improve the KATRIN sensitivity.

(Some figures may appear in colour only in the online journal)

Contents

Executive summary

The Karlsruhe Tritium Neutrino (KATRIN) experiment, sited at the Karlsruhe Institute of Technology (KIT) in Germany, is designed to make a high-precision, high-statistics measurement of the endpoint region of the integral spectrum of T_2 β decay. KATRIN's primary physics goal is the measurement of the absolute neutrino-mass scale via tritium-decay kinematics; recently, the experiment has set the world-leading limit of $m_\nu < 0.8$ eV (90% C.L.), based on its first calendar year of data taking [1]. β decays are provided by a high-luminosity, windowless, gaseous, cryogenic, molecular tritium source, while the β energies are analyzed by a magnetic adiabatic collimation with electrostatic filtering (MAC-E filters) spectrometer with eV-scale filter width. In this white paper, in addition to explaining the basics of the experiment, its analysis, and its recent results, we address KATRIN's status and prospects.

- Two additional calendar years of regular neutrino-mass data have been acquired, and data-taking continues. Our sensitivity target on m_ν is 0.2 eV at the 90% C.L.
- KATRIN's background is higher than anticipated. We enumerate the residual backgrounds affecting the measurement (section 2.3), and explore several complementary R & D programs into further background mitigation (section 7.1).
- The KATRIN collaboration is making continual progress on understanding and mitigating sources of systematic uncertainty, via theory, computational improvements, and dedicated measurements of systematics (section 3.4).
- In addition to the neutrino-mass scale, KATRIN's exquisite spectral measurement enables a broad range of new-physics searches (section 5) ranging from sterile neutrinos to Lorentz-invariance violation to exotic weak currents. Additional physics observables, related to T_2 molecular physics and to the conversion-electron decays of our ^{83m}Kr calibration standard, are accessible through KATRIN spectroscopy.
- By extending the KATRIN β-spectrum measurement to lower β energies, we obtain sensitivity to sterile neutrinos at keV mass scales in a complementary measurement to astrophysical probes. We are actively developing new detector technology that will accommodate the higher rates at these lower energies (section 7.3.2).

1. Introduction

The discovery of neutrino oscillations [2, 3] implies a non-vanishing rest mass of the neutrino, which is physics beyond the Standard Model (BSM). Unlike other BSM physics, such as dark matter or the baryon asymmetry of the Universe, neutrinos can be probed directly in the lab. In the standard picture, three massive Dirac (Majorana) neutrinos mix among each other, and seven (nine) physical parameters exist. Neutrino oscillations can measure six of these, but the neutrino mass-scale (and two Majorana CP phases) must be determined by other means. Apart from being a fundamental parameter of nature, the neutrino-mass scale has deep implications. For instance, in many models for neutrino-mass generation the neutrino-mass scale is inversely proportional to the scale of its origin. Together with neutrinoless double-beta decay experiments [4], knowledge of (or constraints on) the neutrino-mass scale allows determination of the Majorana or Dirac character of neutrinos. The neutrino-mass scale is one of a very limited number of cosmological observables that is directly measurable in the laboratory. For instance, a signal in KATRIN accompanied by no signal in current and upcoming cosmology surveys [5] would imply that the ΛCDM model would need to be significantly modified, or that neutrinos possess exotic new features or interactions. Similarly, a signal in KATRIN but no signal in neutrinoless double-beta decay experiments would strongly imply that neutrinos

are Dirac particles. No signal in KATRIN but a signal in neutrinoless double beta-decay experiments could mean that an alternative diagram leads to that decay, which would have interesting consequences for collider physics and for baryogenesis. As the only model-independent way to probe neutrino mass, direct searches such as KATRIN are crucial.

Building on some 70 years of kinematic searches [6], KATRIN probes the neutrino mass via a precise measurement of the electron energy spectrum resulting from the β decay of molecular tritium:

$$T_2 \rightarrow {}^3HeT^+ + e^- + \overline{\nu}_e. \tag{1}$$

The nonzero rest mass of the neutrino results in a distortion of the β spectrum that is statistically most significant at the endpoint of the spectrum. This shape distortion grants sensitivity to the square of the effective neutrino mass

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

an incoherent sum of the distinct neutrino-mass values m_i weighted by their contributions to the electron-flavor state, given by the elements U_{ei} of the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix. Here, we denote the observable m_ν, but m_β is often used in the literature for the same observable.

Tritium is the isotope of choice for a number of reasons. First, it has the second lowest known β-decay endpoint energy, increasing the significance of the contribution of the neutrino mass. Second, its short half-life of 12.3 years enables high β-decay rates. Third, the relatively simple electron structure of both T and ${}^3He^+$ facilitates calculation both for the atom and the molecule. Fourth, the low Z of the nucleus reduces energy loss by inelastic scattering of the β as it escapes the source. Finally, tritium β decay is a super-allowed transition, and no shape correction to the spectrum is needed.

Only about one part in 10^{13} of the β decays appear in the last eV of the spectrum; a statistically significant measurement at the endpoint of the spectrum therefore requires an extremely bright source and excellent resolution for measuring the energy of the decay electron. The Tritium Laboratory Karlsruhe (TLK) provides an intense 10^{11} Bq gaseous tritium source while the electromagnetic spectrometer techniques pioneered by the earlier Mainz [7] and Troitsk [8] experiments provide sub-eV energy discrimination. The ability to precisely measure the β-decay spectrum grants sensitivity to additional physics topics such as sterile neutrinos, the number density of relic neutrinos and Lorentz-invariance violation. Such an experimental setup is not without its challenges, however. The intense gaseous source requires an understanding of the systematic uncertainties, e.g. plasma properties and instabilities. The low event rate at the endpoint of the spectrum demands identification and mitigation of sources of background. Significant progress has been made on these challenges as the KATRIN collaboration works to reach its 0.2 eV design goal. A high-rate detector, denoted TRISTAN, is planned to expand the search for sterile neutrinos.

Section 2 describes the KATRIN apparatus, including recent technological achievements (section 2.1), operational history (section 2.2), and sources of background (section 2.3). General analysis strategies are summarized in section 3, including a discussion of systematic uncertainties in section 3.4. Section 4 gives details about the inference of the neutrino mass, including KATRIN's recent results. Prospects for KATRIN tests of BSM physics are described in section 5, with additional physics observables in section 6. Section 7 gives an overview of
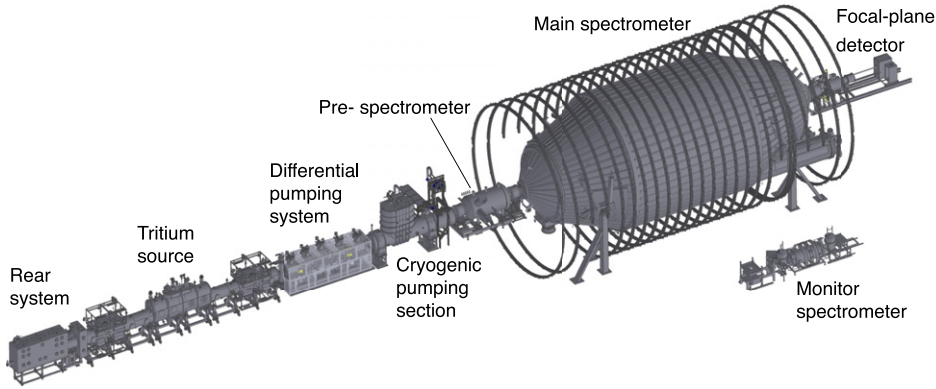
Figure 1. Engineering drawing of the KATRIN beamline.

current research and development efforts, ranging from background mitigation (section 7.1) to acceptance improvements (section 7.2) to plans for a future keV-scale sterile-neutrino search (section 7.3).

2. The KATRIN apparatus

The 70 m KATRIN beamline is designed to perform high-precision energy analysis of β electrons from a high-luminosity, gaseous T_2 source. Here, we give a high-level overview of the apparatus (figure 1); a complete technical description may be found in reference [9], and details of selected systems, chosen based on recent technical progress, are given in section 2.1. Briefly, T_2 gas is purified in a tritium loop system (section 2.1.1), which continuously delivers cold T_2 gas to the center of the source system. The gas diffuses to both ends of the source-system cryostat, where the first pumping stages are located. A fraction of the T_2 molecules experience β decay during their flight within the source beam tube. Of the resulting βs, those that are emitted in the downstream direction are guided by magnetic field lines through the chicanes of two consecutive pumping stages that together reduce the flow of neutral tritium by some 14 orders of magnitude. The first, differential pumping stage relies on turbomolecular pumps (TMPs), while the second is a cryogenic pumping stage in which tritium is adsorbed onto an Ar frost layer that lines the beam tube.

Past the pumping systems, the βs reach a pair of tandem spectrometers designed according to the principle of MAC-E filters. This operating principle will be described in detail in section 2.1.2. Briefly, a MAC-E filter uses magnetic-field gradients to collimate the β-electron flux, allowing a longitudinal retarding potential U to set a threshold on the total kinetic energy of the βs. A β with energy $E > qU$, where q is the electron charge, will pass through the spectrometer to the downstream exit; if $E < qU$, the electron cannot pass and is instead reflected in the upstream direction. To a good approximation, a MAC-E filter thus acts as an integrating high-pass filter, with a characteristic relative filter width set by the ratio of the minimum to the maximum magnetic fields. By scanning through a set of retarding energies $\{qU_i\}$, KATRIN is able to reconstruct an integral spectrum.

In the original KATRIN design, the first MAC-E filter—the pre-spectrometer—was envisioned as a pre-filter, excluding the low-energy bulk of the spectrum to avoid backgrounds arising from collisions of βs with residual gas in the larger, high-resolution main spectrometer with a readily adjustable qU. However, following the discovery that tandem spectrometer

operation led to an effectively energy-dependent background (section 2.3), KATRIN now runs without energizing the pre-spectrometer. β electrons with sufficient energy pass through the main spectrometer and are counted in the focal-plane detector, a monolithic silicon p–i–n diode segmented into 148 equal-area pixels. A 10 kV post-acceleration electrode, immediately upstream of the detector, elevates the signal energies above local backgrounds.

Calibration and monitoring systems are located in several positions along the beamline. In the tritium circulation loops, a laser-Raman spectroscopy system measures the tritium purity of the gas being injected into the source. At its upstream end, the source section terminates in a gold-plated rear wall which can be held at a small voltage to better control the source plasma conditions. Behind the rear wall, silicon drift detectors monitor the source activity by detecting x-rays from β interactions in the rear wall. An electron gun provides calibration electrons with controlled angle and energy; these electrons enter the source via a small aperture in the center of the rear wall, and travel the entire length of the beamline. Gaseous ^{83m}Kr can be introduced into the source to circulate either by itself or along with other gases (section 2.1.3), providing monoenergetic electron lines through internal-conversion decays.

Further downstream, in the cryogenic pumping section, the forward beam monitor (FBM) uses silicon detectors to sample the β rate at the edge of the flux tube. A condensed ^{83m}Kr source is available, providing calibration electrons from a specific location and controlled starting potential. Parallel to the main spectrometer, a monitor spectrometer (MoS)—the refurbished MAC-E filter once used by the Mainz collaboration [7]—allows real-time comparison of the main-spectrometer retarding potential to a ^{83m}Kr standard. The retarding potential is also monitored by high-precision voltage dividers. Finally, the detector section is fitted with electron and γ sources for energy calibration and efficiency measurements.

2.1. Selected technical details

Here, we highlight technical advances since the preparation of reference [9]. These advances include the system for circulating the source gas (section 2.1.1), a new operational mode for the primary MAC-E filter (section 2.1.2), and a very intense, gaseous ^{83m}Kr calibration source (section 2.1.3).

2.1.1. Source-gas circulation. The tritium loop system supplies the windowless gaseous tritium source (WGTS) with a stable flow of tritium. Gas is fed into an injection chamber from two capillaries (the tritium and the tritium + krypton injection capillary) which are thermally coupled to the beam tube over a length of 5 m. It is then injected into the central part of the WGTS beam tube via circumferential orifices [10]. Two differential pumping ducts are attached to both ends of the source tube, each equipped with two or four TMPs. A stationary gas-density profile inside the source tube is formed by continuous injection of gas in the middle of the beam tube and pumping at both ends [10], as can be seen in figure 2. The integral of this density profile along the length of the WGTS, the column density, is an experimentally accessible parameter for measurements and monitoring [11].

The main requirements for the WGTS are 24/7 operation over long periods of time with tritium at a high purity of $>95\%$, and a stability of the column density better than $<0.1\%$ [9, 12]. A stable column density is achieved when the injection pressure and temperature of the beam tube are kept stable. Temperature stability is provided by the WGTS cryogenic system, while pressure stability and composition are the task of the tritium loop system. In order to achieve the necessary source activity, a stable cumulative throughput of up to $\sim$40 g tritium d^{-1} is required. Such high flow rates of tritium can only be achieved in a closed-loop operation [13, 14] as described below.
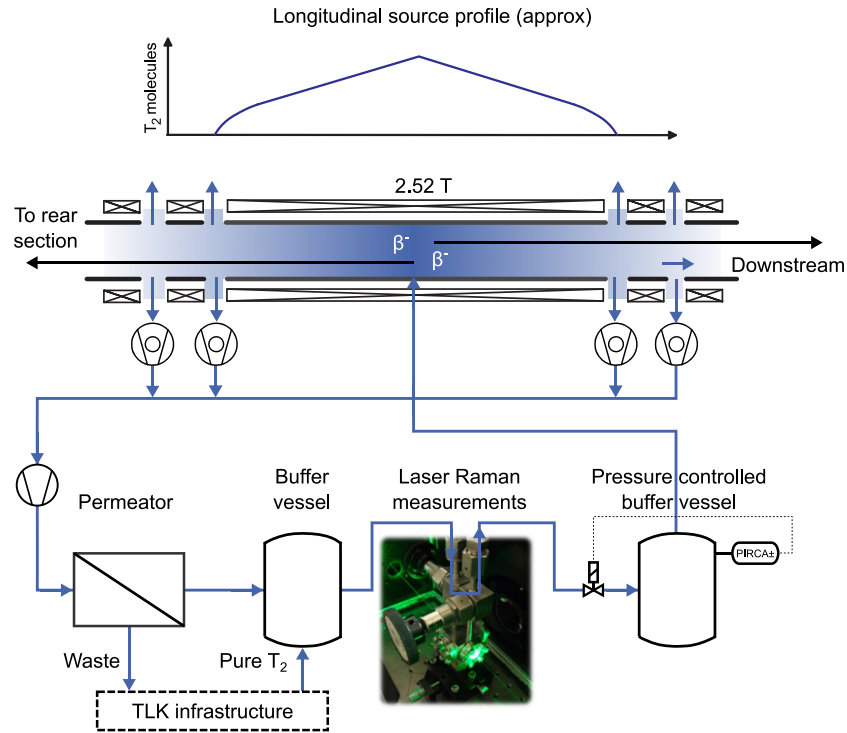
Figure 2. Simplified schematic of the WGTS and tritium circulation loop. The temperature of the WGTS beam tube was 30 K in the first two campaigns [1, 15] and has been changed to 80 K for subsequent campaigns.

Tritium is injected from a pressure-controlled buffer vessel via the tritium-injection capillary and injection chamber into the WGTS beam tube. The majority of the gas ($>99.9\%$) is pumped out of the beam path with the first-line TMPs at both ends. The remaining gas is pumped off further along the beamline [16]. The fore vacuum of the first-line TMPs, which directly attach to the pumping ducts, is provided by 4 s-stage TMPs, a scroll and a metal bellows pump in series. The tritium + krypton injection capillary connects the volume behind the second-stage TMPs with the injection chamber via a valve. The metal bellows pump pushes the gas through a palladium membrane filter ('permeator'). Only hydrogen isotopologs may pass through this permeator, and all non-hydrogen impurities are thus removed [17, 18]. In order to avoid impurity accumulation and subsequent blocking of this filter, a small fraction ($\sim 1\%$) of the gas is continuously extracted in front of the permeator. The gas that passes through the permeator enters into an intermediate buffer vessel, supplemented with high-purity tritium to replace the gas extracted at the permeator. This high-purity tritium is drawn from a system of buffer vessels that serve as the interface to the TLK infrastructure. From there, the gas is led over a laser Raman sampling cell [19] and a regulation valve back into the pressure-controlled buffer vessel, completing the closed loop.

In addition to this standard configuration, the tritium loops system can be operated in further distinct modes [20], depending on the usage of the tritium- and the tritium + krypton injection capillary. When doing calibration measurements using ^{83m}Kr, the closed loop described above would filter out the krypton immediately, rendering the measurement impossible. If the tritium + krypton injection capillary is included in addition to the tritium-injection capillary, a fraction

of the gas pumped out of the WGTS is directly re-injected. This so-called double-injection mode allows co-circulation of krypton through the source. However, a large part of the ^{83m}Kr is still filtered out of the gas stream. In order to achieve the maximum possible ^{83m}Kr activity in the source, the WGTS loop-only mode is used. By feeding all of the gas pumped out of the WGTS beam tube directly back into it, no ^{83m}Kr is lost. The small fraction of gas lost to pumps outside the WGTS is replenished using the tritium-injection capillary.

The design source temperature is 30 K, achieved by a two-phase neon cooling at 10^{-3} stability [21]. In order to perform ^{83m}Kr calibrations without condensation in the cold source, one needs to operate the source tube at a higher temperature (e.g. at 120 K with two-phase argon cooling, or at 80 K with two-phase nitrogen cooling). The operation at higher temperature not only changes the Doppler broadening of the β energy, but it also limits the maximum achievable column density for a given tritium throughput. Thus, krypton-calibration and neutrino-mass operation can only be performed under rather different states of the source with regard to column density and temperature, which both impact the secondary ionisation in the source plasma. Experimental studies indicated that the plasma effects could be more relevant than initial simulations indicated. It was therefore decided in 2020 (after the campaigns for which neutrino-mass results have been published [1, 15, 22]), that neutrino-mass and calibration measurements should subsequently be performed at a common temperature of 80 K. In this way, one can switch from one mode to the other to perform calibration measurements at any time, and the calibration is performed under more similar conditions to the primary measurement. The drawback of this choice, is however, that the maximum possible column density is reduced by about 10%.

2.1.2. MAC-E filter.   A sensitivity to the neutrino mass of less than 1 eV requires energy discrimination at a similar level. Such resolution is obtained in KATRIN using the MAC-E filter technique.

The basic working principle is shown in figure 3. Electrons are born in a high magnetic field (2.5 T) in the tritium source and are guided magnetically through the transport section toward the spectrometers, which analyze their kinetic energy. The electrons in the source have an isotropic starting-angle distribution. If an electron's motion is adiabatic and the magnetic-field strength B is reduced by several orders of magnitude, the conserved orbital magnetic moment ($\mu = \frac{E_\perp}{B}$ in the non-relativistic limit) requires a corresponding reduction of the electron's kinetic energy due to motion perpendicular to the magnetic field lines, $E_\perp$. The total kinetic energy $E_{\text{tot}} = E_\perp + E_\parallel$ is conserved and for large pitch angles (the angle θ between the magnetic field $\vec{B}$ and the electron momentum $\vec{p}$ at the electron's starting location) is mainly converted into $E_\parallel = E_{\text{tot}} \cdot \cos\theta$. Without an electric field, $E_\perp$ reaches its minimum at the position where the magnetic field is minimal (typically, $\mathcal{O}(1\text{ mT})$). A high voltage U (typically -18.6 kV) applied to the spectrometer vessel results in an energy threshold qU for an electron with charge q. The potential reaches its maximum in the central part of the spectrometer, the so-called analyzing plane (AP, see figure 3), creating a barrier for electrons with insufficient kinetic energy $E_\parallel < qU$, while the other electrons are transmitted. The filter width of the spectrometer is given by $\Delta E = \frac{B_{\min}}{B_{\max}} \cdot E$—the maximal residual transverse energy $E_\perp$ for which an electron with the energy E will be transmitted by the spectrometer.

The KATRIN main spectrometer has a design filter width of $\Delta E = E \cdot$ 0.3 mT/6 T ≈ 1 eV for energies of 18.6 keV. The β-spectrum is measured by scanning the retarding potential U near the endpoint.

The magnetic flux of 134 T cm^2 is conserved along KATRIN's beamline. The high statistics required by the precise measurement of the shape of the tritium endpoint spectrum can be obtained by increasing the activity of the tritium source. The luminosity of the source, however,
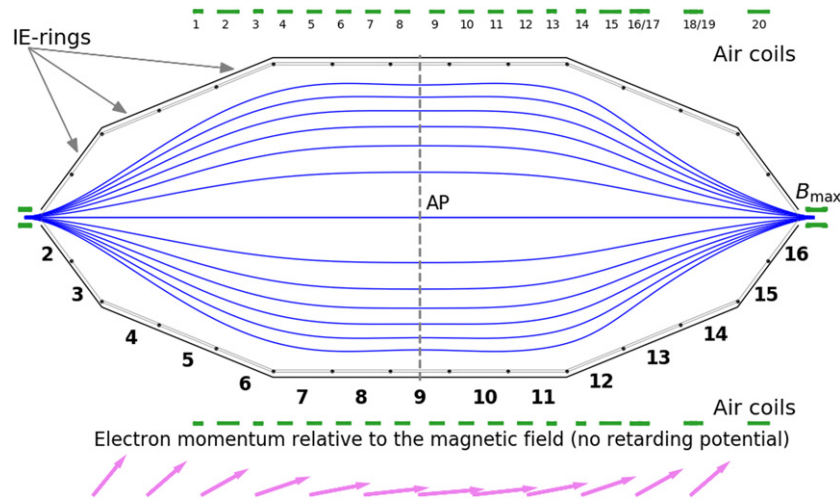
Figure 3. Schematic view of the KATRIN main spectrometer, a MAC-E filter. Electrons from the tritium source are born in a high magnetic field and enter the spectrometer from the left, guided adiabatically along magnetic field lines (solid blue) to a region with low magnetic field. The lower part of the figure shows the evolution of the electron momentum due to the conservation of the orbital magnetic moment, in the absence of a retarding potential: the momentum is almost aligned with the magnetic field lines at the analyzing plane (dashed gray, AP). The magnetic field is defined by the stray fields of the superconducting solenoids at the entrance and exit of the spectrometer and by the system of 20 air coils (thick, green bars above and below the vessel).

is limited in practice by the electron scattering in the source: a higher source length and column density would make it effectively opaque to electrons at relevant energies. Therefore the diameter of the source tube is the primary lever for increasing the luminosity. The KATRIN source has a diameter of 9 cm. The reduction of the magnetic field in the main spectrometer increases the area of the flux, whose diameter reaches ∼9 m for the design magnetic-field configuration. The inner diameter of the main spectrometer is 9.8 m.

Adiabatic motion is another key requirement for the MAC-E filter; that is, the gradients of the magnetic fields must be small. The adiabatic transformation of transverse into longitudinal motion occurs over a distance of several meters and sets the 23.23 m length scale of the KATRIN main spectrometer.

To avoid collisions of electrons with the residual gas in the main spectrometer and to reduce backgrounds, the spectrometer vessel volume of 1240 m^3 is kept at 10^{-11} mbar. Reference [23] describes the vacuum systems that maintain this pressure.

The magnetic field inside the main spectrometer is defined by the superconducting solenoids at the entrance and the exit of the main spectrometer. The solenoid at the exit of the spectrometer, the so-called pinch magnet, provides the highest magnetic field in the beam line. The magnetic field is fine-tuned by a system of air-cooled magnet coils ('air coils') surrounding the main spectrometer (see figure 3). A high-sensitivity magnetometer system monitors the fields at several external positions during measurements.

The retarding potential U is generated by a high-precision, high-voltage supply with a distribution and monitoring system. The stability and precision of the voltage applied to the spectrometer vessel and inner electrode (IE) reach the ppm (10^{-6}) level [24]. The IE

system also allows fine shaping of the retarding-potential profile along the symmetry axis of
the spectrometer to ensure the required transmission properties.

In the design configuration of the main spectrometer, the maximum of the absolute retarding
potential and the minimum of the magnetic field are placed in the center of the spectrometer.
The cylindrical part of the vessel and IE (rings 7–11 in figure 3) provides a homogeneous
potential, and the magnetic field is kept minimal and constant along the symmetry axis of
the vessel. With the analyzing plane (AP) in the center of the spectrometer, the variation of
the electric potential and magnetic field across the AP is minimized. The symmetrical field
configuration for the neutrino-mass measurements was defined using simulations [25].

The observation of an elevated background level (see section 2.3 for details) motivated a
modification of the nominal magnetic flux in the spectrometer to improve the statistical sen-
sitivity to the neutrino mass. The most significant background contributions scale with the
imaged volume of the main spectrometer—that is, the number of counted background elec-
trons is proportional to the volume of the flux tube within the spectrometer vessel. The first
modification was made during the first measurements of the tritium spectrum (section 2.2). The
magnetic field in the AP was increased by a factor of 3, reducing the corresponding flux-tube
volume at the cost of a broader filter width.

The next background-mitigation measure was a new field configuration: the shifted analyz-
ing plane (SAP) [26]. This operating mode further reduces the effective volume from which
any background electrons can reach the detector, by shifting the maximum of the absolute
retarding potential and the minimum of the magnetic field toward the detector. The low-energy
background electrons cannot overcome the potential barrier and are reflected back to the source
side. Only those that are emitted between the AP and the detector are counted. With this mode
the background of the main spectrometer is reduced by a factor of two. The challenge of find-
ing the optimal configuration is to minimize variations of the electromagnetic field in the AP
while preserving the transmission conditions and narrow filter width of the spectrometer. For
that purpose we use the air coils and IE system to place the magnetic field minima and the elec-
tric potential maxima at the same positions for all the radii in the spectrometer. Nonetheless,
the increased field variation requires more fine-grained analysis of spectra recorded in different
regions of the detector. After calibration measurements and assessment of the related systemat-
ic uncertainties, the SAP configuration was implemented as the default operating mode of
the KATRIN main spectrometer.

2.1.3. Intense ^{83m}Kr calibration source. The monoenergetic conversion electrons from the
radioactive decay of the ^{83m}Kr isomeric state provide a unique calibration tool, since ^{83m}Kr
gas can be injected into the source where the conversion electrons are produced in a simi-
lar environment to the β electrons. The continuous β spectrum is distorted due to the energy
losses of the β electrons in the source gas as well as due to the time instabilities and spatial
inhomogeneity of the source plasma. Measuring the spectra of the individual electron conver-
sion lines provides information on the distortions. At the same time, due to its short half-life
($T_{1/2} = 1.8620$ h), ^{83m}Kr does not remain in the source long-term.

At the Nuclear Physics Institute, Czech Academy of Sciences, Řež, KATRIN's ^{83m}Kr source
[27] was developed, based on the deposition of the parent radionuclide ^{83}Rb into a few zeolite
spherules (15–30 pieces). The krypton generator for injection of ^{83m}Kr into the KATRIN source
was then constructed using Swagelok components [28]. The long-lived ^{83}Rb ($T_{1/2} = 86.2$ d)
remains firmly attached to the aluminosilicate skeleton of zeolite while about 80% of the ^{83m}Kr
atoms born in the ^{83}Rb decay emanate from zeolite and are thus available for injection. The
parent ^{83}Rb was produced in the proton-induced reactions ^{nat}Kr(p, xn)^{83}Rb at the U-120 M
and TR-24 cyclotrons of the NPI within the center of accelerators and nuclear analytical

methods (CANAM), where ^{nat}Kr refers to Kr gas with natural isotope abundances. Increasing activity demands resulted in gradual development of several generations of dedicated, pressurized, ^{nat}Kr gas targets. Continuous optimization of the target allowed us to enhance the ^{83}Rb production rate from the initial value of 14 MBq h^{-1} to the current value of 132 MBq h^{-1}. Several ^{83}Rb/^{83m}Kr emanation sources with a typical ^{83}Rb activity of $\sim$1.2 GBq were provided for krypton measurements at KATRIN.

The first krypton measurement was accomplished in 2017, when ^{83m}Kr was allowed to propagate freely from the krypton generator output into the source [29, 30]. Then the gaseous krypton source (GKrS) was assembled [9], consisting of the krypton generator (in secondary containment to prevent potential tritium contamination), WGTS tube and necessary loops. The GKrS system allows three principal application modes: ^{83m}Kr alone, T$_2$ + ^{83m}Kr and D$_2$ + ^{83m}Kr. The decision to neglect plasma effects when analyzing the first KATRIN neutrino-mass data, when only limited initial β-spectrum statistics were available, was based on the L$_3$-32 krypton line spectrum [22]. In 2020, extensive krypton measurements were performed to quantitatively study plasma effects (section 3.4). To achieve the KATRIN design goal, an ultra-high-intensity ^{83m}Kr source with $\sim$10 GBq of the parent ^{83}Rb was required, see also [1]. The amount of ^{83}Rb corresponds to the activity limit defined in the handling license of the TLK for this radionuclide.

The fabrication of this source presented several technical challenges related to the large-scale production of ^{83}Rb: long-lasting bombardment of the single target with a high-current proton beam, semi-automated and remote processing of the target, and deposition of the ^{83}Rb into the zeolite. The new technology had to provide a high-quality ^{83}Rb/^{83m}Kr emanation source, while minimizing the contamination risk and the radiation burden on personnel.

The final target system, in particular the beam entrance window, demonstrated high stability during the longest continuous bombardment performed on the TR-24 cyclotron: 4.5 days with a 24 MeV, 50 μA beam. Remote elution of ^{83}Rb from the target chamber was performed in a dedicated hot cell equipped with telemanipulators. Roughly 94% of the ^{83}Rb present in the target was recovered into the aqueous solution. Subsequent deposition of ^{83}Rb into the zeolite spherules was a smooth process resulting in overall deposition efficiency of $\sim$92%. In order to reduce unnecessary radiation exposure, some routine test measurements of the source were minimized, a measure made possible by the experience and data obtained during previous investigations of the sources and their behavior. The calibration source was successfully used at KATRIN in summer 2021.

2.2. Operational history

The KATRIN collaboration was founded in 2001, and a comprehensive baseline design of the experiment was completed in 2004 [12]. Initial measurements were conducted in 2012 to study transmission properties and background processes of the main spectrometer and focal-plane detector, which were the first major parts of the experimental apparatus to become operational [31].

'First light' was achieved in 2016, when photoelectrons and inactive ions from an ultraviolet source were transported for the first time through the complete 70 m-long setup, including the source and transport systems inside the TLK. The following year, 2017, saw a first campaign of precision spectroscopy, using monoenergetic conversion electrons from gaseous and condensed ^{83m}Kr sources to demonstrate the high-resolution and high-stability performance of KATRIN's MAC-E type energy filter [29, 30, 32].

In 2018 the KATRIN collaboration celebrated the official inauguration of the beamline with the first tritium operation of the source (albeit at low tritium content). Initial β-decay spectra showed an excellent match to the model description [33]. Subsequently, the column density of

the source was increased in successive steps; at about 20% of the nominal column density, the first neutrino-mass data were taken in Spring 2019 [15, 22]. After a final tritium burn-in phase of the source section in June 2019, stable conditions at high tritium throughput and isotopic purity were achieved in advance of the second neutrino-mass measurement campaign [1].

Since entering regular operations, the annual measurement program aims for at least 210 calendar days of scientific data taking, which has been continuously achieved from 2018 onwards. Typically, the beamline and tritium laboratory are shut down for maintenance work from December until February. The measurement program, starting in March, is split into three campaigns per year to allow a two- to three-week break in between each pair of campaigns to regenerate the tritium-loaded cryogenic pumping section. At least once per year, there is a measurement campaign with krypton (pure or mixed with tritium; see sections 2.1.1 and 2.1.3) for calibration and systematics studies.

2.3. Backgrounds

The sensitivity of the KATRIN experiment is currently limited by a higher background than anticipated, underlining the importance of understanding all background processes. During KATRIN's early operations, many different background sources and processes were investigated, understood, and mitigated. Detector background is comparatively small and can easily be measured by decoupling the detector from the rest of the system. Since it is nearly homogeneous and constant in the energy region of interest, consideration in the final analysis is not an issue; muon-induced events in the detector system are mitigated with a muon-veto system [34, 35]. The overall background in KATRIN is dominated by low-energy electrons that originate in the main spectrometer. This type of background is especially pernicious because it is not subject to energy discrimination: since signal βs are decelerated by the retarding potential until they are almost at rest in the central part of the main spectrometer (section 2.1.2), before being re-accelerated by the retarding potential as they travel toward the detector, low-energy secondary electrons created near the AP will arrive at the detector with close to the same energy that the surviving signal βs do.

External sources—cosmic muons and gammas from environmental radioactivity—can stimulate the emission of low-energy secondary electrons from the spectrometer walls. By design, these charged secondaries are excluded from the flux tube via electrostatic reflection (from the IEs, which are held at a more negative potential than the walls) and via magnetic reflection (from the fields that contain the flux tube). Dedicated studies have confirmed that no significant spectrometer background is associated with tagged muons [36] or with enhanced or suppressed external radioactivity [37].

Charged particles, traveling down the beamline, can create secondary electrons in the main spectrometer via ionization interactions with the residual gas. Tritium ions, created in the source via β decay or scattering with β electrons, are guided along the beamline by the same magnetic fields that guide β electrons, and are not simply pumped away like neutral tritium gas. Dedicated electrodes in the beamline of the differential pumping system have been demonstrated to block and remove these ions before they can generate background in the spectrometers [38, 39]. Meanwhile, during the fourth neutrino-mass campaign it was realized that tandem spectrometer operation resulted in a background that increased with the amount of time spent at a given qU setting, due to the trapping of βs in the potential minimum between the pre-spectrometer exit and the main-spectrometer entrance. This trap is cleared by a wire electron catcher inserted during each qU setpoint change [40], so that a longer measurement time at a given setpoint results in a longer time between emptyings of the trap [1]. To avoid the trap and remove this effectively energy-dependent background source, KATRIN now runs
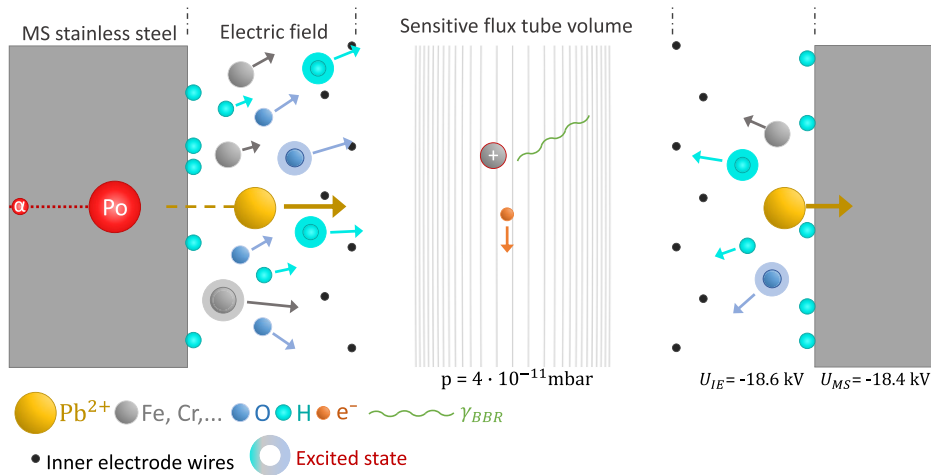
Figure 4. Illustration of Rydberg-atom generation (not to scale). Various atomic species are sputtered by the energetic recoil ion. Additional sputtering occurs if the ion leaves the vessel surface and strikes the opposite surface. BBR can ionize such excited atoms, releasing low-energy electrons that are either guided to the source or to the focal plane detector by the magnetic field lines, creating the sensitive flux tube volume.

with the pre-spectrometer vessel grounded and its downstream IE held at a moderate voltage of -100 V.

The most significant background contributions are those induced by radioactivity inside the spectrometers, by two major generation mechanisms. First, we discuss background electrons due to magnetically trapped particles inside the sensitive flux-tube volume [41, 42]. Depending on their initial energy and pitch angle, electrons can be trapped within the magnetic bottle of the spectrometer; since only scattering can reduce the electron energy, trapping is very efficient in the 4×10^{-11} mbar ultra-high vacuum of the main-spectrometer vessel. Trapped electrons can scatter with residual gas and ionize atoms to generate low-energy secondary electrons, which may then propagate to the detector if they are created with low enough energy. A single initial trapped electron may produce a cluster of secondary electrons at the detector, so that the resulting background is non-Poissonian.

In KATRIN, this mechanism is fed by radon decays in the volume of the main spectrometer. There are three naturally occurring isotopes of radon, all of which emanate from the extensive non-evaporable getter material that (along with TMPs) achieves the spectrometer vacuum: ^{219}Rn, ^{220}Rn, and ^{222}Rn. With half-lives of 56 s and 3.8 d, respectively, the latter two are efficiently pumped out of the volume by TMPs before decaying. With a short half-life of about 4 s, ^{219}Rn in the main-spectrometer volume will typically decay there, releasing shakeoff electrons with energies up to several keV. However, this background source can be efficiently mitigated with a baffle system in the pump ports, consisting of copper plates cooled with liquid nitrogen, on which ^{219}Rn atoms adsorb and decay [43, 44]. The SAP configuration of the MAC-E filter (section 2.1.2) further reduces this background.

The remaining background is generated by emission from the spectrometer walls of neutral particles, which can penetrate the electromagnetic shielding of the flux tube and be ionized within the main-spectrometer volume. These neutral messengers are highly excited Rydberg atoms, sputtered from the walls due to naturally implanted radioactivity within the steel; because of their high initial quantum numbers n, they can be ionized by blackbody radiation (BBR).

During installation of the IEs, ambient air circulated within the main spectrometer, and naturally occurring ^{222}Rn adsorbed onto the surface, decayed and implanted radioactive impurities into the vessel walls. Eventually, the long-lived daughter ^{210}Pb ($t_{1/2} = 22$ yr) accumulated, feeding a decay chain leading to the alpha emitter ^{210}Po that decays to stable ^{206}Pb with 5.4 MeV excess energy. The recoil ion can sputter atoms from the inner vessel surface, as shown in figure 4. These sputtered atoms could be iron, chromium, or oxygen from the steel, or hydrogen adsorbed on the surface; the total sputtering yield (released atoms per alpha decay) depends on the surface conditions. Depending on the excitation state of these Rydberg atoms, the ionization electrons will have energies in the meV range. This hypothesis is supported by detailed investigations, including short-term contamination with shorter-lived decay chains from ^{223}Ra and ^{228}Th sources [45]. Since the rate of ionization electrons detected depends on the volume imaged by the detector, this background can also be mitigated by the SAP (section 2.1.2).

At about 150 mcps, the current background level is about 15 times larger than initially projected. An increased background level reduces the sensitivity to the spectral imprint of the neutrino mass. Further reducing the background level is of high importance. KATRIN's background-related research-and-development efforts are presented in section 7.1.

3. KATRIN analysis tools and strategies

The squared neutrino mass, m_ν^2, is obtained from a fit of the calculated analytical spectrum to the measured data. An overview of the KATRIN model is given in section 3.1; the full details of the source and spectrum calculation may be found in reference [46]. The spectrum fitting and limit setting procedures are briefly discussed in section 3.2, and the data combination methods are described in section 3.3. Section 3.4 summarizes major sources of systematic uncertainty.

3.1. The analytic model

The analytical spectrum is a convolution of the theoretical differential spectrum with the experimental response function. The response function describes the number of β electrons that are transported through the flux tube and counted at the detector, including both the energy loss experienced as the β electrons propagate through the source and the transmission properties of the MAC-E filter (section 2.1.2).

The differential decay rate of a tritium nucleus is described by Fermi's golden rule,

$$\frac{d\Gamma}{dE} = \frac{G_F^2 \cos^2(\theta_C)}{2\pi^3} |M_{nuc}|^2 F(Z, E) p(E + m_e) \cdot \sum_f P_f \epsilon_f \sqrt{\epsilon_f^2 - m_\nu^2} \Theta(\epsilon_f - m_\nu), \qquad (3)$$

where G_F is the Fermi constant, $\cos^2(\theta_C)$ is the Cabibbo angle, $|M_{nuc}|^2$ is the nuclear matrix element, and $F(Z, E)$ is the Fermi function with the atomic charge of the helium daughter nucleus $Z = 2$. $\epsilon_f = E_0 - V_f - E$ is the neutrino energy with an energy correction from the molecular final state f of T_2, where E_0 is the effective tritium endpoint and V_f is the molecular excitation energy populated with probability P_f, as calculated in quantum chemical theory [15, 47, 48]. The Heaviside function $\Theta(\epsilon_f - m_\nu)$ ensures conservation of energy. Doppler broadening and broadening arising from plasma effects within the source are emulated in the differential spectrum. The Fermi function is treated relativistically and radiative theoretical corrections on the atomic and nuclear level are included in the analytical model.

Electrons experience energy loss as they scatter from molecules within the gaseous tritium source. The energy-loss function, $f(\epsilon)$, describes the probability of a certain energy loss ϵ in a

scattering process, illustrated by the red dotted line in the bottom panel of figure 5. The scattering probability functions, $P_s(\theta)$, describe the probability of an electron with pitch angle θ to scatter s times before leaving the source. Electrons leaving the source without scattering experience no energy loss, $f_0(\epsilon) = \delta(\epsilon)$. For s-fold scattering the energy-loss function is convoluted s times with itself leading to a total energy loss function of

$$\sum_s P_s(\theta)f_s(\epsilon) = P_0(\theta)\delta(\epsilon) + P_1(\theta)f(\epsilon) + P_2(\theta)(f \otimes f)(\epsilon) + \cdots. \tag{4}$$

illustrated by the blue solid line in the bottom panel of figure 5.

The transmission condition, $\mathcal{T}$, is either 1 or 0 (i.e. electrons are either transmitted or not through the MAC-E filter) depending on the starting energy E and the starting pitch angle θ of the electron, and on the retarding energy qU. For an isotropic electron source the angular distribution of electrons is $\omega(\theta)\mathrm{d}\theta = \sin(\theta)\mathrm{d}\theta$ and $\mathcal{T}$ is integrated over the electron pitch angle θ,

$$
T(E, qU) = \int_0^{\theta_{\max}} \mathcal{T}(E, \theta, qU) \cdot \sin(\theta)\mathrm{d}\theta
$$

$$
= \begin{cases}
0 & (E - qU) < 0 \\[2mm]
1 - \dfrac{\sqrt{1 - \frac{(E-qU)}{E}\left(\frac{B_{\mathrm{src}}}{B_{\mathrm{ana}}}\right)}}{1 - \sqrt{1 - \left(\frac{B_{\mathrm{src}}}{B_{\mathrm{max}}}\right)}} & 0 \leqslant (E - qU) \leqslant \Delta E, \\[4mm]
1 & (E - qU) > \Delta E
\end{cases} \tag{5}
$$

where $(E - qU)$ is the surplus energy. The transmission function is therefore mainly determined by the magnetic fields of the source, B_{src}, the AP, B_{ana}, and the beamline maximum, B_{max}. The magnetic reflection imposes an upper limit on the pitch angle given by $\theta_{\max} = \arcsin\sqrt{(B_{\mathrm{src}}/B_{\mathrm{max}})}$ which reduces the effective width, or energy resolution, of the transmission function. These MAC-E-filter properties are described in section 2.1.2 and illustrated by the grey dash-dotted line in the top panel of figure 5. Synchrotron energy losses due to the electron motion through magnetic fields is also included in the transmission-function calculation.

The final experimental response is the convolution of the energy-loss function with the transmission function, given by

$$R(E, qU) = \int_0^E \int_0^{\theta_{\max}} \mathcal{T}(E, \theta, qU) \cdot \sin(\theta) \cdot \sum_s P_s(\theta)f_s(\epsilon)\mathrm{d}\theta\,\mathrm{d}\epsilon. \tag{6}$$

The response function is illustrated by the blue solid line in the top panel of figure 5. This experimental response function, together with the theoretical differential spectrum, constitutes the analytical model which is fit to the measured data. It can be computed for any region of the detector under analysis.

3.2. Spectral fitting and limit setting

The number of observed counts measured at the detector is compared to the expected number of counts predicted from the analytical model (section 3.1). The probability to have an
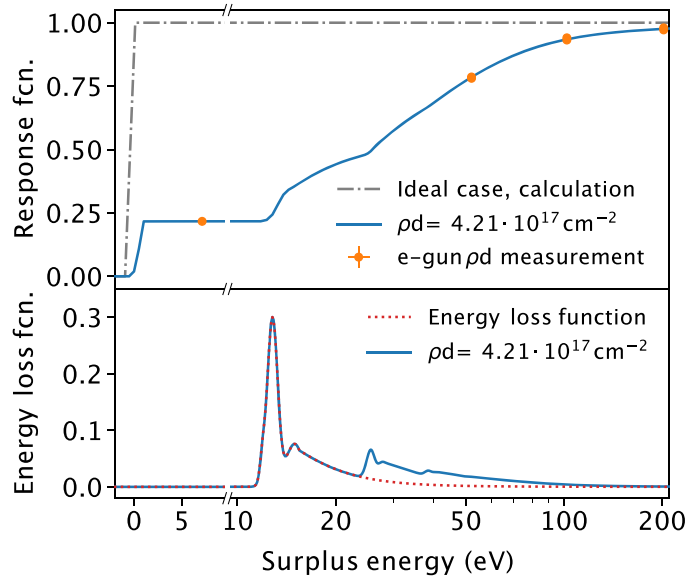
Figure 5. Top: transmission function of the KATRIN spectrometer (gray dash-dotted line). When convolved with the energy-loss function for a given column density, the full response function for β-decay electrons is described (solid blue line). This agrees well with measurements from the electron gun with column density $\rho d = 4.21 \times 10^{17}$ cm^{-2} (orange dots). Bottom: energy-loss function measured *in situ* with the KATRIN electron gun (red dotted line). It is convolved with itself s times to describe the energy loss of electrons passing through the source with s-fold scatterings (blue solid line). Electron-gun measurements are described in detail in reference [49].

observed outcome given the predicted number of counts, that is, the likelihood function, can be maximized to obtain the parameters of interest. In a 'statistics-only' fit there are four fit parameters:

- Squared neutrino mass, m_ν^2
- Effective tritium endpoint, E_0
- Signal amplitude, Sig
- Background rate, Bg

In practice the χ^2, or negative log-likelihood,

$$\chi^2 = -2 \log \mathcal{L}\left(m_\nu^2, E_0, \text{Sig}, \text{Bg}\right) \tag{7}$$

is minimized to obtain the best-fit spectral parameters. If systematic uncertainties are treated via the 'pull term' method (section 4.1) the number of fit parameters can be higher.

If the sensitivity of a spectral measurement is not good enough to resolve the neutrino mass, then statistically we expect that half the time the measurement will show too many counts in the endpoint region, and half the time it will show too few. However, equation (3) is valid only for $m_\nu^2 \geqslant 0$. In order to perform a well-behaved fit at this sensitivity boundary, the spectral function must be extended to negative m_ν^2. The KATRIN analysis performs this extension by allowing m_ν^2 to be negative in equation (3); the resulting χ^2 function is asymmetric, but the analysis of early data sets is robust against a different choice of spectral extension [6].

When a fit of the analytical model to the data has been completed, an upper limit on the neutrino mass is set from the Neyman construction [50] of the confidence belt. Statistically unfluctuated, 'Asimov' data sets are generated for a set of true squared neutrino mass values, $m_{\nu,\mathrm{true}}^2$, using the fit results as input parameters. Each Asimov data set is fit over a range of hypothetical measured squared neutrino masses, $m_{\nu,\mathrm{meas}}^2$, and the confidence belt is then constructed by integrating under the resultant likelihoods to the desired confidence level. The ordering principle can be taken from the Feldman–Cousins method [51] or the Lokhov–Tkachov method [52]. These methods yield consistent results for positive best-fit neutrino mass values; when the best-fit neutrino-mass value is negative, the Lokhov–Tkachov method gives the sensitivity as the upper limit.

3.3. Data combination

Neutrino-mass measurements are performed in 'scans' that step through the retarding-potential set points over a few hours under stable source conditions. The tritium β-decay electrons are counted at the focal-plane detector on 148 pixels (see section 2) which provide spatial resolution over the cross-sectional area of the flux tube. In order to perform the neutrino mass analysis, the spectra from scans and pixels are combined.

- **Scan combination:** scans performed under the same operating conditions (during one measurement phase for example) are combined via 'stacking'. This entails summing the counts from each detector pixel, and averaging parameters such as the retarding potential, tritium purity, and column density.
- **Pixel combination:** pixels that have similar transmission properties are combined into 'patches'. The simplest combination is all pixels to a uniform detector, but pixels can alternatively be combined into subsets taking into account the field configuration and the alignment of the beamline. Pixel combination entails summing the counts from each detector pixel within the defined patch, and averaging parameters such as the AP magnetic field and maximum magnetic field.
- **Measurement-period combination:** neutrino-mass scanning periods with the same measurement conditions are combined in the same ways as the individual scans. However, if the parameters of the experimental setup cannot be averaged, the periods can be treated via a simultaneous fit of several periods using the squared neutrino mass as a common parameter, while other fit parameters and systematic contributions can be defined independently for different periods. Section 4.2 describes alternative schemes for combining data from different measurement periods.

When the data combination has been defined, the neutrino-mass analysis is performed with a fit of the analytic model to the measured data. There are several models with differing input parameters, corresponding to each of the spectral combinations. The minimization of the parameters of interest is therefore over several spectra. The common parameter for all spectra is the squared neutrino mass, and the log-likelihood profile is the summation of the log-likelihoods from each spectrum i. For the 'statistics-only' case, equation (7) is extended to

$$-\log \mathcal{L} = \sum_i -\log \mathcal{L}_i\left(m_\nu^2, E_{0i}, \mathrm{Sig}_i, \mathrm{Bg}_i\right). \tag{8}$$

The endpoint, E_0, is not a common parameter, allowing it to absorb systematic shifts arising from, e.g., plasma effects. In practice, additional parameters are included when systematic uncertainties are considered. These parameters can be either common or independent for each scan and pixel combination.

3.4. Sources of systematic uncertainty

Inference of the squared neutrino mass, m_ν^2, from the measured β-decay spectrum requires very accurate knowledge of the theoretical as well as experimental ingredients entering the analytic expression of the spectral shape. This analytic model, which incorporates both the theoretical β spectrum and the experimental response, forms the fit function, as detailed in equations (3) and (6) in section 3.1. Uncertainties of the parameters used to calculate the spectrum model translate into uncertainties of the parameters obtained from the fit. For the two neutrino-mass results released by KATRIN thus far (section 4.3), the budget of systematic uncertainties is presented in table 2.

It is instructive to group the sources of systematic uncertainties into signal-related and background-related effects. In the following we will inspect these two classes separately.

3.4.1. Signal-related systematic uncertainties.

The environment of the magnetized tritium plasma inside the windowless, gaseous tritium source determines the electric and magnetic fields in which the β electrons are produced. These boundary conditions impact the β-decay spectrum directly or indirectly—for instance through the acceptance angle of signal electrons, which is determined by the ratio of source magnetic field to maximum magnetic field strength encountered along the trajectory, or through the energy scale determined by the difference in electric potential between the point of origin in the source and the analyzing point in the spectrometer.

The strong magnetic fields of the source ($B_{src} = 2.52$ T) and of the pinch magnet ($B_{max} = 4.74$ T) were measured with precision probes before final system integration, when the sensitive field region was still accessible. Their stability is monitored during operations. We estimate the field uncertainties at $\sigma(B_{src}) = 1.7\%$ and $\sigma(B_{max}) = 0.1\%$, respectively. The much weaker magnetic field inside the main spectrometer of $\mathcal{O}(10^{-4}$ T) is inaccessible to direct measurement during operations, but can instead be inferred from the data delivered by an extensive network of precision magnetic sensors outside of the vacuum vessel, combined with simulations using the Kassiopeia field-modeling and particle-tracking software [53]. The resulting uncertainty is estimated at the 1% level.

The molecular tritium gas forms a self-ionizing plasma which is bounded on the sides by the stainless steel walls of the beam tube and at the upstream end by a gold-plated rear wall which can be supplied with a variable bias voltage. While a stable, non-zero absolute source potential is no concern for the neutrino-mass analysis since it would be absorbed in the effective, fitted endpoint of the spectrum, temporal drifts or spatial nonuniformities can cause a broadening of the response function. To assess such possible spatial variations, KATRIN makes use of monoenergetic conversion electrons from ^{83m}Kr co-circulated with the tritium gas (section 2.1.3). Plasma parameters can be inferred from the narrow lines in the spectrum of conversion electrons from this nuclear standard. In order to ensure that calibration conditions are directly applicable to neutrino-mass measurements, KATRIN has run both at an elevated source temperature of 80 K since 2019 (section 2.1.1).

A key component of the model applied to fit the data of the integrated β-decay spectrum is the instrumental response function. The response is composed of two main ingredients (section 3.1): the transmission function of the main spectrometer as a high-pass energy filter, and the energy-loss description of signal electrons traversing the gaseous tritium column in the source. The working point of the tritium source (temperature, gas density or flow rate) is optimized to yield a high luminosity of β electrons at an acceptable opacity from scattering. Depending on their point of origin inside the extended source beam tube, β electrons have a sizable probability to undergo scattering and thus lose a certain amount of energy in inelastic

encounters. This energy loss has to be taken into account with high accuracy in the description of the response. The two relevant factors are the precise value of the column density (ρd) and the inelastic scattering cross section (σ_{inel}). KATRIN employs a custom-made precision photoelectron source for monoenergetic electrons emitted at selected pitch angle in order to directly measure the product $\rho d \cdot \sigma_{\text{inel}}$ as well as the full energy-loss function [49]. The monitoring of $\rho d \cdot \sigma_{\text{inel}}$ is done periodically during a neutrino-mass measurement campaign and is accompanied by a combined online monitoring of the source activity (through beta-induced x-ray spectroscopy at the upstream end of the beamline and through a FBM at the exit of the source and transport section) and of the tritium purity (through Raman spectroscopy), whereas the energy-dependent loss function was determined in dedicated measurements (figure 5). As a result of these efforts, the energy-loss function is now well characterized and understood based on a physics model, and the uncertainty on m_ν^2 due to $\rho d \cdot \sigma_{\text{inel}}$ is now at the level of 0.01 eV^2.

In addition to the systematics related to operating conditions or experimental effects, another uncertainty is introduced at the level of the differential β-decay spectrum (analytic model, section 3.1) by the energy distribution allocated to electronic and rotational-vibrational excited states of the $(^3\text{HeT})^+$ molecule in the final state of the decay. This probability distribution over the various molecular states of the daughter $^3\text{HeT}^+$ ion has been computed in a series of very precise calculations adopting sophisticated molecular-structure theory [15, 47, 48]. The resulting uncertainty on the neutrino mass has recently been re-assessed and found to be sufficiently small that it does not limit the neutrino-mass determination by the KATRIN experiment.

3.4.2. Background-related systematic uncertainties.

Understanding the background characteristics, such as a possible spatial dependence and variation with time or energy (section 2.3), is key to a thorough assessment of background-related uncertainties. These effects can best be mitigated by further reducing the overall background level; efforts in this direction are described in section 7.1.

Apart from the increase of the statistical uncertainty due to a high background level, several effects related to the background could influence the shape and uncertainties of the measured spectrum. The first effect is related to the decay of radon in the volume of the main spectrometer; as described in section 2.3, each radon decay can result in several secondary electrons that can eventually reach the detector. Since they originate from the same primary electron, these counts are no longer independent and do not follow the Poisson distribution, but rather a wider distribution, approximated by a Gaussian. This over-dispersion effectively increases the spread of the measured count rates and therefore the statistical uncertainty of the spectrum. This was a dominant contribution to the total neutrino-mass uncertainty in the first two measurement campaigns (table 2).

The next effect is related to the possible dependence of the background rate on the absolute retarding potential applied to the spectrometer. The dependence is described by a linear function (slope) in the range of retarding energies near the tritium endpoint. An estimate of the slope is obtained from dedicated measurements of the background rate over a wider range of retarding potentials, with the valve to the tritium source closed. The typical size of the effect is $\mathcal{O}(1 \text{ mcps keV}^{-1})$ and the contribution to the neutrino-mass uncertainty is shown in table 2.

Finally, the Penning trap in the region between the pre- and main spectrometer causes an increase of the background rate during each scan step (section 2.3), so that the measured rate depends on the measurement time spent at each setting. The magnitude of the effect is estimated by fitting a linear increase of the rate within each set point. The rate increase is of the order of $\mathcal{O}(1 \text{ }\mu\text{cps s}^{-1})$, which leads to an $\mathcal{O}(0.1 \text{ mcps})$ rate increase for a 100 s measurement step. The time spent at each scan step varies from 20 s to 700 s and therefore the additional

rate is different for individual retarding-potential set points, leading to a shape distortion of the measured spectrum. To mitigate this significant source of systematic uncertainty, the pre-spectrometer vessel was grounded beginning in 2021, thereby effectively lifting the trapping conditions (section 2.3).

Here and in the results reported in section 4.3, we take into account the systematics as assessed for KATRIN's initial two neutrino-mass phases. Significant progress has since been made in controlling both the signal- and background-related uncertainties, and will be reflected in future neutrino-mass results.

4. Measuring the neutrino-mass scale with KATRIN

As described in section 3.2, the neutrino-mass analysis requires fitting model spectra—incorporating the square of the neutrino mass among other fit parameters—to the measured, integral spectrum. The KATRIN analysis strategy includes several complementary methods for handling systematic uncertainties, as detailed below. The fitting frameworks, developed by independent teams, rely on a shared implementation of the underlying spectral model and experimental response function. Additional physics searches, as described in sections 5 and 6, can typically be performed by incorporating new terms into the spectral model.

The KATRIN analysis is model-blinded: a designated collaborator generates an arbitrary, secret Gaussian broadening of the electronic ground state of the daughter molecule in T_2 β decay, which alters the spectral model in a manner correlated much more strongly to the squared neutrino mass than to the other fit parameters. This broadening is implemented independently for each measurement campaign, in a software module that is automatically synchronized with the analysis packages. The analysis procedures are first developed and tested on Monte-Carlo data generated using slow-control data recorded during the measurement. Then, the model-blinded fits are performed on the actual data. The broadening is revealed, and the analysis performed with the correct model, only after all input parameters and systematic uncertainties have been finalized.

4.1. Parameter inference

KATRIN's free parameters $\vec{\Theta}$ are inferred by minimizing the χ^2 function,

$$\chi^2 = \left(\vec{R}_{\text{data}}(q\vec{U},\vec{r}) - \vec{R}(q\vec{U},\vec{r}|\vec{\Theta},\vec{\eta})\right)^{\text{T}} \cdot C^{-1} \left(\vec{R}_{\text{data}}(q\vec{U},\vec{r}) - \vec{R}(q\vec{U},\vec{r}|\vec{\Theta},\vec{\eta})\right), \qquad (9)$$

where $\vec{R}_{\text{data}}(q\vec{U},\vec{r})$ gives the measured count rates at a retarding energy qU_i for the detector region r_j, $\vec{R}(q\vec{U},\vec{r})$ gives the predictions of these rates, and C is the covariance matrix that includes the statistical uncertainties and, in some cases, systematic uncertainties; systematic parameters are collected in $\vec{\eta}$. As described in section 3.3, by combining data from multiple scans and multiple pixels, we ensure that the counts can be described by a Gaussian distribution, so that this approach is valid.

To propagate uncertainties on the systematic parameters $\vec{\eta}$, KATRIN pursues four parallel strategies as summarized below. Additional details can be found in references [1, 15]. This analysis is very computationally intensive, but the recent development of a neural-network method for computing the KATRIN physics model promises to reduce the computational load of KATRIN analysis by up to three orders of magnitude [54].

4.1.1. Covariance matrix. This approach explicitly incorporates the systematic uncertainties in the covariance matrix C that appears in equation (9). The diagonal entries describe

uncorrelated uncertainties for each $R(qU_i, r_j)$, while the off-diagonal terms describe the correlated uncertainties between different data points $R(qU_i, r_j)$. The covariance matrix is pre-computed from $\mathcal{O}(10^4)$ decay spectra that are simulated while varying the systematic parameters η_i within their probability density functions.

4.1.2. Pull terms. In this method, a systematic parameter η_i with an estimated value $\hat{\eta}_i$ and an uncertainty σ_{η_i}, as determined from separate measurements, is treated as a free parameter in the fit. A corresponding pull term η_i is added to the χ^2 function from equation (9),

$$\chi^2_{\text{tot}}(\vec{\Theta}, \vec{\eta}) = \chi^2(\vec{\Theta}, \vec{\eta}) + \sum_i \left(\frac{\hat{\eta}_i - \eta_i}{\sigma_{\eta_i}} \right)^2 . \tag{10}$$

Each η_i is therefore free to vary, but an increasing χ^2 penalty is imposed the further it moves from its best estimation, depending on our uncertainty on this parameter.

4.1.3. Monte Carlo propagation. This technique involves propagation of uncertainties via $\mathcal{O}(10^5)$ fits to simulated spectra. To assess the statistical uncertainty, we generate 10^5 statistically randomized Monte Carlo spectra, which are fit with a constant model. To assess the systematic uncertainties, a statistically unfluctuated spectrum is simulated and fit 10^5 times with the systematic parameters $\vec{\eta}$ varied each time according to their probability density functions. To obtain the total uncertainty, these steps are combined, producing a histogram of best-fit parameters. Each entry in this histogram is weighted by the likelihood of its corresponding fit, and the best-fit value and uncertainty are taken from the mode and width of the resulting distribution.

4.1.4. Bayesian. In contrast to the frequentist methods described above, in Bayesian inference the posterior probabilities of $\vec{\Theta}$ are computed from a prior probability and the data, according to Bayes' theorem. The KATRIN analysis uses a flat, positive prior on m_ν^2, restricting the posterior distribution to physically allowed values. In principle, the systematic effects $\vec{\eta}$ are included as free parameters in the Bayesian analysis, with each η_i constrained by prior knowledge. Due to computational constraints, this is presently done only for the uncertainty on a qU-dependent background, with all other effects implemented either with a covariance matrix or with model variation in the inputs of a large number of Markov Chain Monte Carlo simulations.

4.2. Combining data from different measurement campaigns

KATRIN measurement campaigns are separated by planned breaks during which maintenance activities are conducted—such as pumping away the T_2 gas, cleaning the rear wall, regenerating the Ar frost in the cryogenic pumping system, baking out the main spectrometer, or replacing the wafer of the focal-plane detector. Often, commissioning or calibration measurements are conducted under conditions that differ significantly from neutrino-mass running. Different measurement campaigns may see different background levels, different energy scales (and hence effective endpoints), and different column densities (and hence energy-loss profiles), among other significant differences; see table 1, further below, for a comparison of the first two neutrino-mass measurement campaigns. In general, it is not possible to describe and fit multiple campaigns with a single model spectrum.

As summarized in reference [1], the KATRIN collaboration has explored several approaches for combining multiple data sets, building on the analysis approaches described above in section 4.1. For example, when the data sets are statistically dominated, a straightforward frequentist approach is to combine the results of multiple measurement campaigns by summing

Table 1. Key operational parameters for the first (KNM1 [22]) and second (KNM2 [1]) KATRIN neutrino-mass campaigns. The total β-electron count is given for the analysis interval extending from $E_0 - 40$ eV to E_0. The β-electron-to-background ratio is given by the ratio of this number to the extrapolated background counts in the same energy range. Table reproduced from reference [1].

	KNM1	KNM2
Number of scans	274	361
Total scan time	521.7 h	743.7 h
Background rate	290 mcps	220 mcps
T_2 column density	1.11×10^{17} cm^{-2}	4.23×10^{17} cm^{-2}
Source activity	2.5×10^{10} Bq	9.5×10^{10} Bq
Total number of β-electrons	1.48×10^6	3.68×10^6
β-electron-to-background ratio	3.7	9.9

their respective χ^2 curves. Alternatively, one may perform a joint frequentist fit of the two data sets, in which m_ν^2 is a shared fit parameter for both campaigns, whereas the parameters A_s, R_{bg}, and E_0 are allowed to differ between the two campaigns (section 3.3). Each systematic uncertainty may be treated jointly between the data sets, or separately for each data set, as appropriate. In a Bayesian framework, the posterior distribution from analyzing one data set can simply be used as prior information for the analysis of the next, neglecting correlations between the two spectra. As additional data sets are added and systematic uncertainties become more important, these first approaches will be further refined.

4.3. Recent results

The KATRIN experiment has been operated with tritium since 2018 [33], and has been performing neutrino-mass measurement campaigns from 2019 on. The multi-month campaigns usually take place twice a year and are designated KNM, for 'KATRIN neutrino mass campaign', followed by a sequential number. In spring 2022 KATRIN conducted its KNM7 campaign. At present, results from the KNM1 and KNM2 campaigns (both from 2019) have been published. We review these results here. Table 1 summarizes the operational parameters and performance of these campaigns.

4.3.1. KNM1—spring 2019. In the first neutrino-mass campaign, the KATRIN source was operated at a reduced column density of 1.11×10^{17} molecules cm^{-2}, roughly one fifth of the design value. The lower tritium throughput of about 4.9 g d^{-1} was necessary because materials previously unexposed to tritium have a high rate of impurity formation from radiochemical reactions [14]. These impurities can then accumulate and freeze in the tritium-injection capillary, gradually reducing the throughput. The reduced column-density setting was chosen to stabilize the tritium column density. With this source setting, the KNM1 measurement campaign had a total duration of four weeks, yielding 272 spectral scans that passed data-quality selection. The 148 detector pixels were also subjected to quality cuts, which led to an exclusion of 31 pixels mainly in the outer detector rings. The spectra recorded from these selected pixels were combined into a single spectrum (section 3.3) on which the final analysis was performed.

Three different methods, two frequentist and one Bayesian, were applied for the parameter inference (see section 4.1), as described in detail in reference [15]. The choice of experimental input parameters for the KATRIN model directly affects the neutrino-mass observable, so a model-blind analysis was performed (see beginning of section 4). The best fit from this

Table 2. Breakdown of the most significant uncertainties for the first (KNM1 [15]) and second (KNM2 Reproduced from [1]. CC BY 4.0.) KATRIN neutrino-mass campaigns. Comprehensive tabulations are found in the respective publications. The statistical uncertainty, even though much smaller in KNM2 compared to KNM1, was the dominant uncertainty on the neutrino-mass square derived in both campaigns. Several systematics, such as the ones related to the background characteristics, are reduced substantially from KNM3 onwards.

Effect (1σ uncertainty on m_ν^2)	KNM1 (eV^2)	KNM2 (eV^2)
Statistical	0.97	0.29
Non-Poissonian background	0.30	0.11
Source-potential variations (plasma effects)	Neglected	0.09
Scan-step-duration-dependent background	Neglected	0.07
qU-dependent background	0.07	0.06
Magnetic fields	0.05	0.04
Molecular FSD	0.02	0.02
Column density $\times$ inelastic scat. cross-section ($\rho d\sigma$)	0.05	0.01
Total uncertainty	1.02	0.34

analysis was $m_\nu^2 = -1.0^{+0.9}_{-1.1}$ eV^2 [22]. This result is consistent with a vanishing neutrino mass. The Lokhov–Tkachov prescription [52] was applied to derive the resulting limit of $m_\nu < 1.1$ eV (90% C.L.). Table 2 gives the uncertainty budget. Compared to the statistical uncertainty (0.97 eV^2), the total systematic uncertainty (0.32 eV^2) was of secondary importance. The largest systematic effect was attributed to the non-Poissonian nature of the radon-induced background (section 2.3).

The best-fit value for the effective endpoint, $E_0 = 18573.7(1)$ eV, allows comparison of the KATRIN result to external measurements of the ^{3}He–T atomic-mass difference [55]. Combined with our knowledge of the KATRIN energy scale, these values are in agreement [15].

4.3.2. KNM2—fall 2019. The issue of radiochemical impurity generation was resolved for the second campaign starting in September of the same year. After extensive tritium exposure in the loop system during and following KNM1, impurity formation ceased due to depletion of available surface reactants. For example, carbon present on the stainless-steel wall surface had been mostly converted to tritiated hydrocarbons, which were then removed by the PdAg-membrane filter in the tritium loop [14]. This allowed an increase of source activity by almost a factor of four, to 9.5×10^{10} Bq. Furthermore, KNM2 saw a 30% increase in total scan time recorded, in comparison to KNM1. All of these factors led to significantly higher statistics. At the same time, the background conditions were improved substantially: since the copper baffles of the main spectrometer were regenerated during a spectrometer bakeout prior to the KNM2 campaign, the radon-induced non-Poissonian background contribution was reduced significantly, as reflected by the three-times-smaller corresponding systematic uncertainty (table 2). The overall background rate was approximately 25% lower in KNM2 compared to KNM1. Scan-to-scan fluctuations were also reduced, due to enhanced high-voltage reproducibility.

The analysis again uses 117 detector pixels after quality selection. By virtue of the improved statistics for KNM2, these pixels can now be grouped into 12 concentric rings or into four annuli, apart from being combined into a uniform detector. An analysis according to rings or annuli can probe for radial patterns in the resulting fit parameters, notably the effective endpoint E_0, which might be a sign of plasma effects impacting the energy scale at different regions of
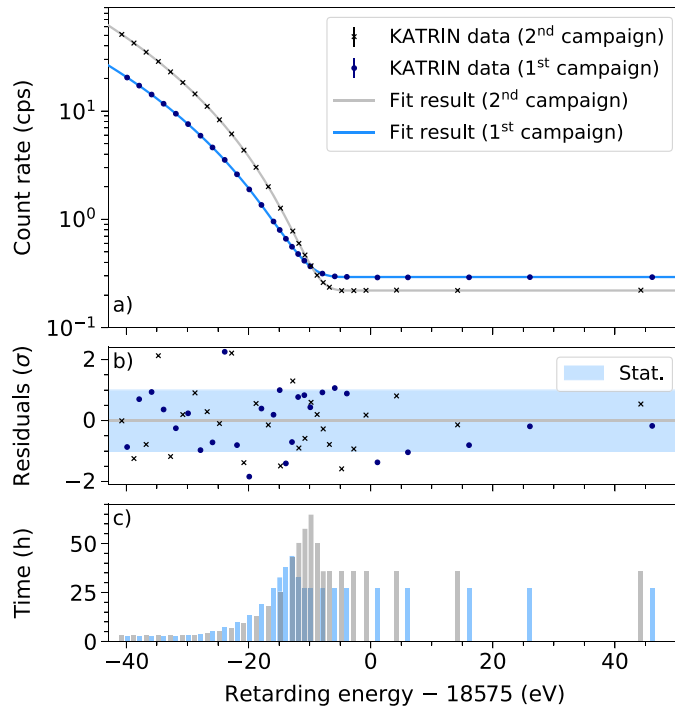
Figure 6. Measured tritium β spectra and fit model for KATRIN's KNM1 (blue; [22]) and KNM2 (gray; [1]) campaigns. (a) Simultaneous, uniform fit to both data sets. (b) Residuals for the uniform fit. (c) Distribution of measurement time near the endpoint, covering the signal and background regions. The share of measurement time allocated to individual scan steps is different for the two campaigns because it is optimized based on signal-to-background ratio, which was larger for the KNM2 data. Reproduced from [1]. CC BY 4.0.

the beam cross section. Eventually, no such radial dependence was found at a significant level in the analysis of KNM2 data; figure 6 therefore displays the result of a uniform fit to the data. A detailed account of the different analysis methods applied to the data is presented in reference [1].

The best-fit values of squared neutrino mass and effective endpoint are $m_\nu^2 = 0.26^{+0.34}_{-0.34}$ eV^2 and $E_0 = 18573.69(3)$ eV, respectively. Again, the results obtained with the different analysis methods are in agreement. Table 2 contains the breakdown of systematic uncertainties, highlighting numerous improvements made through dedicated calibration measurements, more precise determination of various model inputs, and background reduction leading up to KNM2. The total uncertainty on the squared neutrino mass is cut by a factor of three compared to KNM1.

From the best-fit value of m_ν^2, which is compatible with zero within uncertainties, we deduce an upper limit of $m_\nu < 0.9$ eV (90% C.L.). Given that the central value is slightly positive, both the Feldman–Cousins and Lokhov–Tkachov methods result in the same numerical value of the frequentist limit. The Bayesian approach yields an upper bound of $m_\nu < 0.85$ eV (90% credible interval).

4.3.3. Combined result of KNM1 and KNM2 campaigns   Several approaches are pursued to combine results from the two presently unblinded data sets, as detailed in reference [1] and
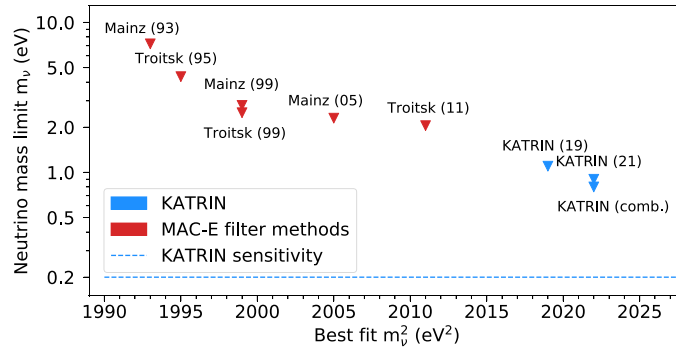
Figure 7. Overview of direct neutrino-mass measurements from the past three decades. Previous upper limits, all obtained using the MAC-E-filter technique, are marked in red; those from KATRIN are highlighted in blue. The projected sensitivity of KATRIN after 1000 full measurement days is indicated by the dashed line. References for data points: Mainz (93) [56], Troitsk (95) [57], Mainz (99) [58], Troitsk (99) [59], Mainz (05) [7], Troitsk (11) [8], KATRIN (19) [15, 22], KATRIN (comb.), KATRIN (21) [1].

summarized in section 4.2. One such method is a simultaneous fit of both data sets, as presented in figure 6. Here, the data from all pixels are merged into a uniform spectrum for each campaign. The joint fit uses m_ν^2 as a shared parameter, while A_s, R_{bg}, and E_0 are implemented as individual parameters for KNM1 and KNM2. In this analysis, systematic uncertainties are treated via pull terms; systematic inputs that differ from one measurement campaign to the next are allowed to vary for KNM1 and KNM2, whereas other pull terms are fit as common parameters. The result of the combined fit yields $m_\nu^2 = 0.07(32)$ eV^2, which translates to an upper limit of $m_\nu < 0.75$ eV (90% C.L.). Alternatively, adding the χ^2 curves from the fits to the two individual data sets gives a consistent upper limit of $m_\nu < 0.81$ eV (90% C.L.). Our reported upper limit, $m_\nu < 0.8$ eV (90% C.L.), is rounded to the digit where both approaches agree.

A Bayesian approach, using the posterior distribution from the analysis of the first data set as an informative prior for the analysis of the second one, finds $m_\nu < 0.73$ eV (90% credible interval).

With the release of its second data set, KATRIN became the first kinematics-based neutrino-mass experiment to reach a sub-eV sensitivity. The relevance of this achievement is illustrated in figure 7, which places these initial results obtained by KATRIN in the historical context of tritium-based neutrino-mass experiments from the past three decades.

5. KATRIN tests of new physics

Testing physics beyond the light-neutrino standard picture adds exciting and potentially groundbreaking physics potential to the experiment. In the upcoming sections we will discuss the sensitivity of KATRIN to a variety of possible scenarios. The motivation of many of these scenarios originates in the facts that neutrino mass is generated in a different way from the other masses of Standard-Model (SM) particles, and that neutrinos are in general perfect candidates to be connected to BSM physics. Frequent features of such frameworks are additional sterile-neutrino states, new neutrino interactions, new particles coupling to neutrinos, etc. The scale at which new physics appears is in general not clear, in particular for sterile

neutrinos. KATRIN is sensitive to two interesting mass scales, the eV-scale, which matches the design analysis interval of the experiment, and the keV scale, which falls within the Q-value of tritium decay. There are long-standing experimental hints for eV-scale sterile neutrinos [60], and keV-scale neutrinos are excellent candidates to be warm dark matter [61].

We separate our discussion into eV-scale sterile neutrinos (section 5.1), keV-scale sterile neutrinos (section 5.2), general neutrino interactions and charged currents BSM (section 5.3), new light bosons (section 5.4), Lorentz-invariance violation (section 5.5), and other new physics (section 5.7). Moreover, we discuss the sensitivity to relic neutrinos (section 5.6), i.e. the ubiquitous background of very low-energy neutrinos that froze out in the early Universe at temperatures of MeV, equivalent to an age of 1 s after the Big Bang. While this is standard physics, it is clear that if observed, there must be a mechanism that leads to large overdensities in our local neighborhood.

The separation between topics is unavoidably blurry. For instance, eV- or keV-scale sterile neutrinos can have new interactions, or Lorentz-invariance violation may be present only for sterile neutrinos. Some of the scenarios we discuss here are better motivated than others, but more often than not neutrino physics has provided unexpected results, and one should remain open to all possibilities.

We stress that the characteristics of the new physics that may show up in β decay can be distinguished from the now-standard physics of active neutrino mass. The new physics can manifest in the measured spectrum in diverse ways, e.g. a kink-like spectral feature (heavy sterile neutrinos), shape distortion (exotic weak interactions, light sterile neutrinos, or new light bosons), appearance of a line feature (relic-neutrino capture) or a sidereal modulation of fit parameters (Lorentz-invariance violation).

5.1. Sterile neutrinos at the eV scale

Right-handed neutrinos are a well-motivated SM extension. The introduction of a right-handed partner to the left-handed neutrinos provides a natural way to create neutrino masses [60]. Right-handed neutrinos, as opposed to the known left-handed neutrinos, would not take part in the weak interaction and would therefore be sterile. No gauge symmetry of the SM forbids the introduction of a Majorana mass term of arbitrary scale for the right-handed neutrino. As a consequence, new neutrino-mass eigenstates arise that can have an admixture with the active SM neutrinos [62]. In the following, these new mass eigenstates are referred to as 'sterile' neutrinos.

Light sterile neutrinos on the eV-mass range are motivated by long-standing and accumulating anomalies in short-baseline neutrino oscillation experiments, such as the reactor antineutrino anomaly (RAA) [63] and the recently reaffirmed gallium anomaly [64–66]. The postulated sterile-neutrino solutions of both anomalies cover a wide range of fourth mass eigenvalues. While the small-mass solutions, $\Delta m_{41}^2 < 10$ eV, are largely excluded by short-baseline reactor experiments [67–71], KATRIN is the only running laboratory experiment that is sensitive to the remaining large Δm_{41}^2 solutions.

The sterile-neutrino search in KATRIN considers a minimal extension of the SM by one additional sterile neutrino ($3\nu + 1$), associated with a mass eigenvalue m_4 or m_s. Here, the standard β-decay spectrum $R_\beta(E, m_\nu^2)$ is complemented by a sterile branch $R_\beta(E, m_4^2)$:

$$R_\beta(E, m_\nu^2, m_4^2, |U_{e4}|^2) = (1 - |U_{e4}|^2) \cdot R_\beta(E, m_\nu^2) + |U_{e4}|^2 \cdot R_\beta(E, m_4^2) \qquad (11)$$

with U being the 4×4 unitary PMNS matrix. The sterile branch is weighted with the mixing of the electron flavor with the fourth mass eigenstate $|U_{e4}|^2$, simply referred to below as
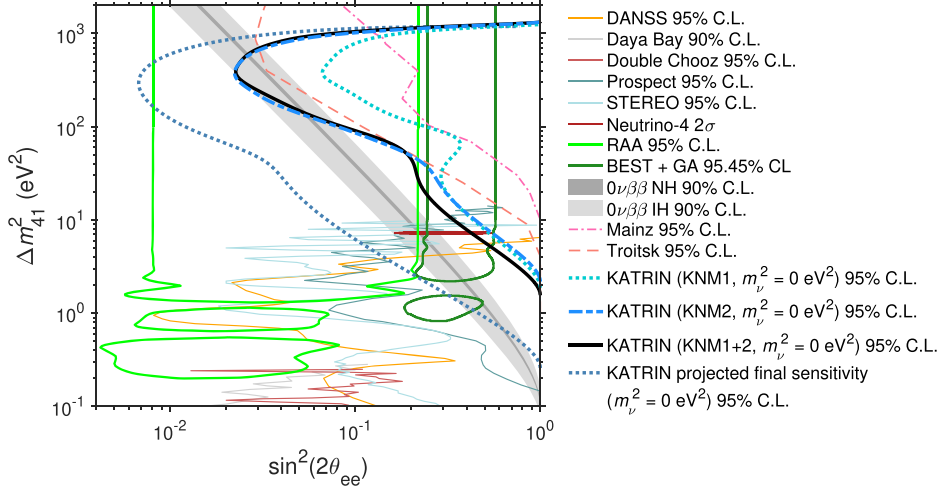
Figure 8. 95% C.L. exclusion contours of the first and second KATRIN science run with $m_\nu^2 = 0$ eV^2, as well as the combined exclusion from both campaigns. The final sensitivity is given by the projection described in reference [72]. The second measurement is able to exclude a wider range than both the Mainz [74] and Troitsk [75] experiments for $m_4^2 \lesssim 300$ eV^2. Large Δm_{41}^2 can be excluded to a great extent for the solutions of the RAA and BEST + gallium (GA) anomalies [63, 66]. The combined analysis is in tension with the positive results claimed by neutrino-4 [76] for $\sin^2(2\theta_{\rm ee}) \gtrsim 0.4$. KATRIN data improve the exclusion bounds set by short-baseline oscillation experiments for $\Delta m_{41}^2 \gtrsim 10$ eV^2 [67–71]. Constraints from $0\nu\beta\beta$ with $m_{\beta\beta} < 0.16$ eV are shown as gray bands [77–79]. Reproduced from [73]. CC BY 4.0.

active-to-sterile mixing. The signature of a sterile neutrino in KATRIN is a kink-like distortion of the β spectrum, most prominent at $E \approx E_0 - m_4$. Figure 9 is an illustration of the shape distortion for a keV-scale m_4.

The sterile-neutrino analyses of the first two measurement campaigns have been published in references [72, 73]. The analyses are sensitive to $m_4^2 \lesssim 1600$ eV^2 and $|U_{e4}|^2 \gtrsim 6 \times 10^{-3}$. As no sterile neutrino signal was found at 95% C.L., exclusion bounds were set.

Figure 8 displays the latest KATRIN results at 95% C.L. and compares them to other experiments. These exclusion contours assume $m_\nu^2 = 0$ eV^2, resulting in $\Delta m_{41}^2 \approx m_4^2$; reference [73] includes an extensive discussion of the relationship between active and sterile neutrino masses in such a search. Analysis of the KNM2 data improves on constraints set by the Mainz [74] and Troitsk [75] experiments for the mass range $m_4 \lesssim 300$ eV^2. As short-baseline experiments are sensitive to different observables from kinematic searches, namely Δm_{41}^2 and $\sin^2(2\theta_{\rm ee}) = 4|U_{e4}|^2 (1 - |U_{e4}|^2)$, the KATRIN, Mainz and Troitsk exclusion contours have been transformed accordingly. A considerable fraction, 20 eV$^2 \lesssim \Delta m_{41}^2 \lesssim 1000$ eV^2, of the gallium-anomaly space can be excluded by KATRIN. Moreover, the RAA is excluded for 50 eV$^2 \lesssim \Delta m_{41}^2 \lesssim 1000$ eV^2. The hint of a signal from the neutrino-4 collaboration is disfavored for $\sin^2(2\theta_{\rm ee}) \gtrsim 0.4$ at 95% C.L.

The present exclusion limits are dominated by statistical uncertainties. Therefore, additional measurement time and statistics from future KATRIN campaigns will improve its sensitivity considerably. The projected final sensitivity demonstrates that KATRIN will be able to examine the entire large-Δm_{41}^2 parameter space of the gallium anomaly and a significant fraction of the RAA.
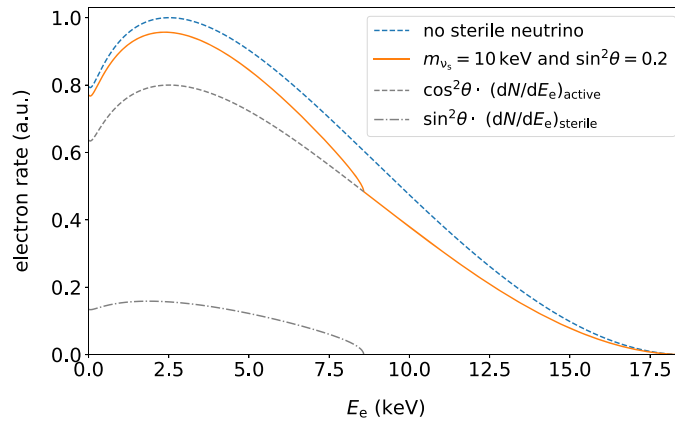
Figure 9. Signature of a 10 keV sterile neutrino in a differential tritium β-decay spectrum: the blue dashed line shows the spectrum without sterile neutrinos. The solid orange line shows the spectrum with a sterile neutrino with $m_s = 10$ keV and an exaggerated mixing amplitude of $\sin^2 \theta = 0.2$. The latter is composed of the active branch (gray dashed line) and the sterile branch (gray dot-dashed line).

5.2. Sterile neutrinos at the keV scale

Sterile neutrinos with a mass in the kilo-electronvolt (keV) regime are promising dark-matter candidates [80–82]. Due to their mass scale and production mechanism, they can act as warm dark matter, which could mitigate tensions between cosmological observations and predictions of purely cold-dark-matter scenarios [80].

As for eV-scale sterile neutrinos (section 5.1), the presence of a sterile neutrino in the keV mass range would result in a characteristic kink-like signature (figure 9) at energy $E_0 - m_s$, with a magnitude governed by the mixing amplitude $\sin^2 \theta$ of the sterile neutrino with the active SM electron neutrino [83, 84].

The KATRIN experiment has several advantageous features for a keV-scale sterile neutrino search. With an endpoint of $E_0 = 18.6$ keV, tritium β-decay permits a search for sterile neutrinos on a mass range of multiple keV. Due to the short-half life of 12.3 years, high decay rates can be achieved with reasonable amounts of tritium which in turn leads to rapid statistical significance on the sterile-to-active mixing amplitude.

The use of the KATRIN experiment for a keV-scale sterile neutrino search has been investigated in several studies (for example in [84–86]). It was shown that the experiment can be used with only minor modifications for a first, low-statistics measurement. As a proof of principle, KATRIN data taken in 2018 has been used to search for keV-scale sterile neutrinos in a mass range of up to 1.6 keV; a publication is in preparation.

For higher-statistics measurements, a new detector system for KATRIN is under development. For a technical description of the TRISTAN detector project, an overview of performance tests, and the current project status, see section 7.3.2.

With 1 yr of KATRIN operation at full source strength, a statistical sensitivity of the order of $\sin^2 \theta = 10^{-8}$ can be reached. It is, however, extremely challenging to control systematic uncertainties at this level. Therefore, the targeted design sensitivity is $\sin^2 \theta = 10^{-6}$, which would surpass the sensitivity of previous laboratory-based searches and reach a region of cosmological interest. Note that while astrophysical constraints on the mixing are quite strong, there are ways to evade them and to bring the mixing up to scales reachable by the TRISTAN
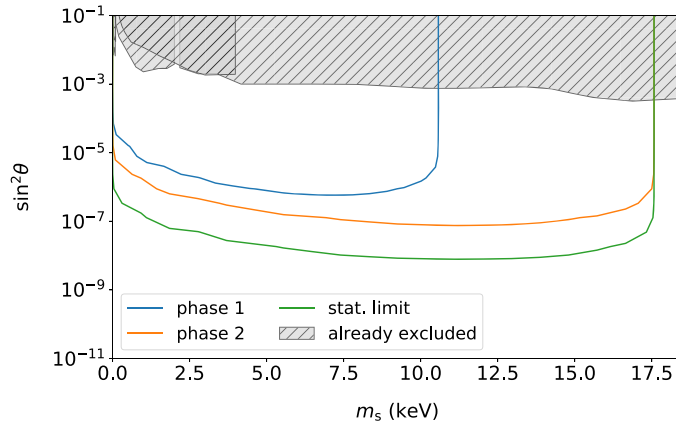
Figure 10. KATRIN sensitivity to keV-scale sterile neutrinos in different scenarios. The statistical limit assumes 10^{18} electrons over the full energy range, corresponding to a measurement time of 1 yr at the full KATRIN source strength. Phase 1 denotes the first stage of TRISTAN operation with a reduced amount of TRISTAN modules (approx. 1000 pixels) and a lowered KATRIN source strength (0.3% of the nominal column density). The spectrum is scanned on an energy interval of 8 keV to 18 keV with a total measurement time of $t_{meas} = 0.5$ yrs. In the follow up stage (phase 2), the TRISTAN detector is operated with all 3500 pixels. The tritium spectrum is scanned on the full energy range at a reduced source activity of 2% of the nominal column density for $t_{meas} = 1.0$ yrs. The already excluded parameter space is based on references [74, 89–91].

detector upgrade [87, 88]. Figure 10 shows the sensitivity for different measurement scenarios, using the KATRIN experiment and a TRISTAN detector to search for sterile neutrinos.

5.3. GNIs and exotic charged currents

Generalized neutrino interactions (GNI) can be derived by embedding a basis of gauge-invariant dimension-six operators into the SM effective field theory, respecting the full $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry. This generalized theory includes a variety of interactions such as scalar, pseudoscalar, vector, axial vector, and tensor interactions with SM fermions [92]. The wide range of interactions could arise from heavy new physics and would cause energy-dependent distortions of the measured β-decay spectrum. As this generalized parameter space is rather complex, it is sensible to look at different scenarios separately.

Precision kinematics of weak decays in particular allow searches for exotic forms of interactions, such as additional non-standard right-handed currents departing from the established $V - A$ structure of SM weak interactions as a special case of GNI. Left-right symmetric models are hypothetical extensions of the SM in which right-handed weak currents occur. The imprint of right-handed currents on the tritium β-decay spectrum with active neutrinos has been the subject of earlier studies (e.g., references [93, 94]). Dedicated experiments (most notably, neutron decay, reviewed in reference [95]) have yielded very stringent limits. The sensitivity of the standard KATRIN analysis is limited by the fact that the endpoint E_0 is one of the free fit parameters (section 3.2). Sensitivity to right-handed currents with active neutrinos can potentially be improved by fixing the endpoint in the fit, which would require a very accurate measurement of the Q-value (through the ^{3}H–^{3}He mass difference) and knowledge of the absolute KATRIN energy scale at the level of 30 meV–100 meV or better.
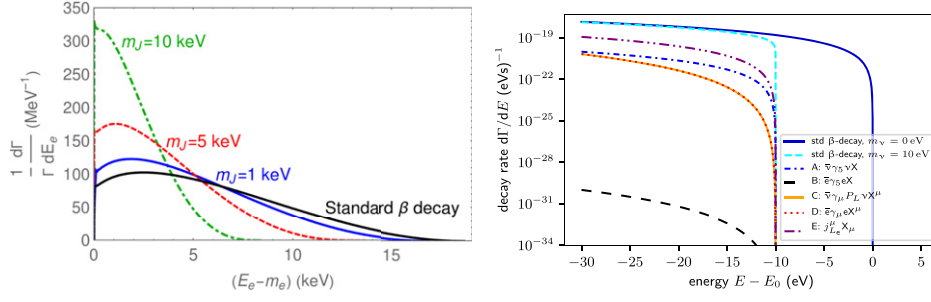
Figure 11. Left: normalized electron-energy spectrum of the decay $^3\text{H} \to {}^3\text{He}^+ + e^- + \bar{\nu}_e + J$ (with the pseudoscalar J emitted from the neutrino) and the standard β-decay spectrum. The coupling $g_{\nu J}$ has been set to 1; the spectrum scales with $g_{\nu J}^2$. Right: differential spectrum for standard β decay without (solid blue) and with 10 eV (dashed cyan) neutrino mass, as well as five spectra expected for the additional emission of a light boson X with $g_X = 1$, $m_X = 10$ eV and various coupling types. Reproduced from [99]. CC BY 4.0.

Recently, the interplay of right-handed currents with potential sterile neutrinos has become a focus of interest (see, e.g., references [96–98]). It has been shown in these works that KATRIN will be able to set bounds on the strength of the left–right interference terms as a function of m_s, a low-energy search complementary to LHC-based efforts. New limits could be obtained for a Fierz-like interference term that does not originate in the left–right symmetric model [98].

5.4. New light bosons

New particles coupling to known particles are a generic prediction of BSM theories. As no heavy new particles have been found at collider experiments, focus on light new physics has increased in recent years. Light scalar or vector bosons coupling to electrons, neutrinos, u or d quarks may be emitted in β decay, if their mass is below the Q value of tritium. KATRIN is particularly sensitive to this interesting low-mass regime. Reference [99] has analyzed this in detail for emission from the leptons.

It is common to find massive CP-even (scalar) and CP-odd (pseudoscalar) fields in new physics theories. For instance, axions and axion-like particles fall into this category. In some cases, these pseudoscalar particles play the role of mediators between SM particles and the dark-matter sector. In some other cases, they are connected to neutrino masses and lepton-number violation, most notably in so-called Majoron models. Agnostic to its possible origin, we can assume that the pseudoscalar J is coupled to neutrinos via $ig_\nu \bar{\nu}\gamma_5\nu J$. A scalar coupling and a lepton-number-violating coupling have the same effect on β decay. Figure 11 (left) shows the effect on the electron spectrum.

Vector bosons Z' are a common feature in many BSM frameworks with additional gauge symmetries. Their masses are typically determined by a (combination of) gauge coupling(s) and by the energy scale at which the additional gauge symmetry is spontaneously broken. There is a subtlety depending on whether the Z' couples only to electrons or neutrinos, or to both. In the former case the current is not conserved and the theory is non-renormalizable, leading to a $1/m_{Z'}^2$ enhancement of the emission process. In contrast, if the Z' couples to both the neutrino and electron, the theory is free of conceptual issues. Focusing on leptonic couplings, we thus have five relevant cases: a pseudoscalar coupling to neutrino, to electron, and a Z' coupling to neutrino, electron and both. Figure 11 (right) shows the differential spectrum for these cases.
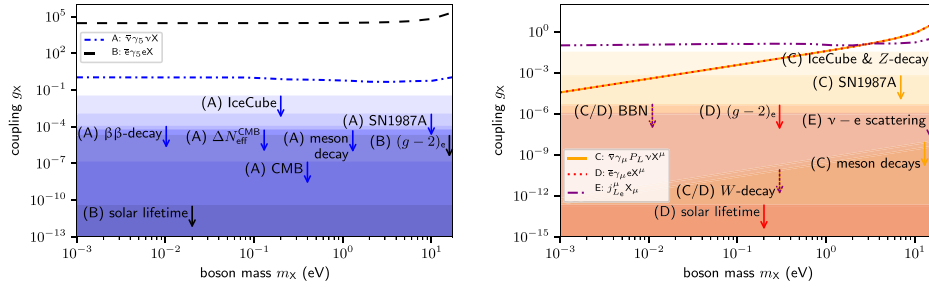
Figure 12. The lines show the statistical sensitivity (90% C.L.) of KATRIN for the detection of eV-scale light pseudoscalars emitted from neutrinos (A) or from electrons (B) (left) and vector bosons emitted from neutrinos (C), from electrons (D), or from both (E) (right). The shaded areas mark the parameter regions allowed from other constraints. Reproduced from [99] CC BY 4.0.

The light-boson spectrum superimposed on the standard β-decay spectrum has six fit parameters:

(a) Neutrino mass squared m_ν^2
(b) Endpoint E_0
(c) The amplitude ensuring correct normalization of the superposition of light-boson and standard β-decay spectra
(d) Background rate
(e) Light-boson mass m_X
(f) Light-boson coupling g_X

Using the standard settings defined in the design report [12], including an analysis interval $[E_0 - 30\,\text{eV}, E_0 + 5\,\text{eV}]$, a true neutrino mass of zero, and a measuring time of three full years, gives expected limits shown in figure 12.

As figure 12 shows, KATRIN's sensitivity region is apparently already excluded by previous laboratory, astrophysical and cosmological searches. However, the laboratory constraints are extrapolated from searches for heavier states, conducted at a far larger energy scale; meanwhile, the astrophysical and cosmological limits are highly sensitive to specific flux and source hypotheses. KATRIN's data, acquired directly at the relevant energy scale, will thus offer an important complement to these earlier searches. Furthermore, KATRIN is an excellent probe for exotic scenarios with strongly energy-dependent couplings [100].

5.5. Lorentz-invariance violation

A variety of BSM theories predict deviations from Lorentz invariance, the possibilities for which are parameterized in the SM extension [101–104]. In the neutrino sector, many of the possible Lorentz-invariance-violating parameters are strongly constrained by time-of-flight or neutrino-oscillation experiments. However, some 'countershaded' parameters, so named because they are virtually invisible to a wide range of experiments, are still constrained only weakly, if at all [104–106].

The complex oscillation-free parameter $\left(a_\text{of}^{(3)}\right)_{jm}$, where $\{j, m\}$ are angular-momentum quantum numbers with $j = \{0, 1\}$, would introduce a preferred direction in space. In a precise measurement of the β-decay spectrum, this could manifest as a variation of the spectral
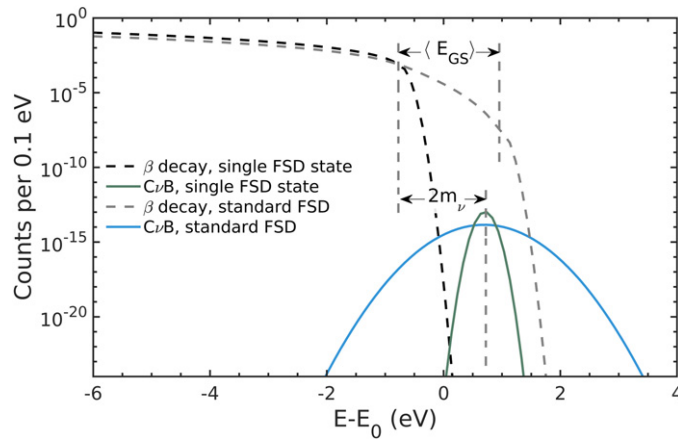
Figure 13. Simulation of the effects of the molecular FSD on the differential spectra from β decay and from relic-neutrino capture, assuming $m_\nu = 0.7\,\mathrm{eV}$ and $\eta = 1$; Reproduced with permission from [111]., from which this figure is adapted. Doppler broadening is included for all spectra. $\langle E_{\mathrm{GS}} \rangle \approx 1.7\,\mathrm{eV}$ is the mean energy of the electronic ground state.

endpoint depending on the latitude and orientation of the β-electron beam, or as a sidereal oscillation of the spectral endpoint E_0 [104, 107]. Because KATRIN analyzes only those β-electrons that are emitted within $50.4°$ relative to the direction of the beam axis, it is in principle sensitive to this parameter.

KATRIN has performed a search for Lorentz-invariance violation based on its first science run. This campaign featured about 300 β-scans over a time period of two weeks. For the best sensitivity, we defined an effective time point for each scan, based on the ordering of scan steps and the region of the scanned spectrum that most dominates determination of the spectral endpoint. Based on these timepoints and the best-fit endpoint of each scan, KATRIN searched for an oscillatory signature with the required period of 23.7 h. This analysis finds no sign of Lorentz-invariance violation, thus setting the first limit on a_{of} [108]. With additional statistics acquired over more sidereal cycles, this limit will be improved.

5.6. Relic neutrinos

Modern cosmology predicts the existence of a cosmic (or relic) neutrino background (CνB) that fills the entire Universe. Cosmic neutrinos decoupled from the rest of the Universe about 1 s after the Big Bang. Nowadays relic neutrinos cluster around galaxies, and the local relic-neutrino overdensity η is predicted to be 1.2 to 20 for neutrino masses below 0.6 eV [109, 110]. The direct detection of these ancient neutrinos would provide direct insight into the early history of the Universe. However, this measurement remains one of the greatest challenges in neutrino physics.

Relic neutrinos can interact with the tritium located in the KATRIN source via the neutrino capture reaction $\nu_e + {}^3\mathrm{H} \rightarrow {}^3\mathrm{He}^+ + e^-$ [112, 113], leading to a monoenergetic electron signal just beyond the endpoint. Figure 13 shows the resulting spectral contributions as well as the broadening imposed by the molecular final-state distribution (FSD) of T_2 decay. KATRIN's high-precision spectral measurement can thus provide new constraints on the local overdensity of relic neutrinos.
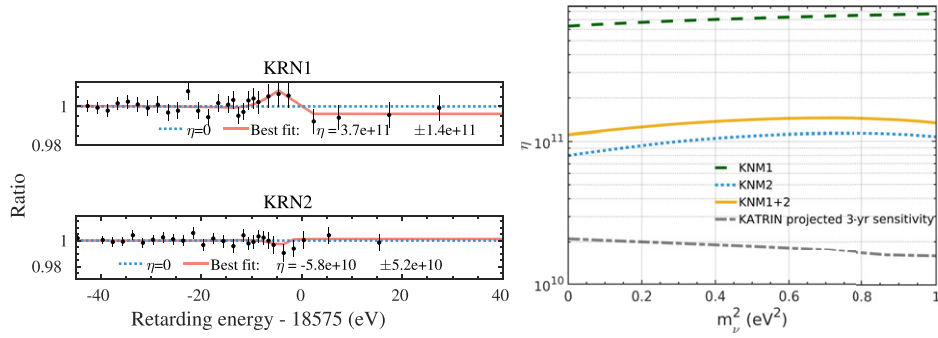
Figure 14. Left: spectral ratio of the best fits for relic neutrinos with respect to the null hypothesis $\eta = 0$ for the first two measurement campaigns. The non-zero values of the best fit of η are both consistent with background fluctuations. Right: exclusion contours for the relic-neutrino overdensity η at 99% C.L. for both measurement campaigns and their combination, and the projected final KATRIN sensitivity. Reproduced with permission from [111].

Using the data from the first two science runs conducted in 2019 [1, 15, 22], KATRIN has established new constraints on the local overdensity of relic neutrinos [111]. The data set includes 5.16×10^6 β decay electrons and 0.72×10^6 background events below the endpoint. No significant relic-neutrino signal is observed, and the η parameter was constrained to be smaller than $9.70 \times 10^{10} (1.10 \times 10^{11}, 1.30 \times 10^{11})$ at 90% (95%, 99%) C.L. for neutrino masses less than 1 eV. This new result, shown in figure 14, improves on previous upper limits set by direct kinematic experiments at Los Alamos [114] and Troitsk [59].

KATRIN is continuing to operate toward the goal of collecting 1000 days of data by 2024. The current measurements at increased background triggered a reassessment of the final sensitivity for the relic neutrino overdensity, now at $\eta < 1.00 \times 10^{10} (1.40 \times 10^{10}, 1.80 \times 10^{10})$ at 90% (95%, 99%) for a background rate of 130 mcps summed over all detector pixels. We note that the irreducible background from β decay, as broadened by the FSD, makes it infeasible to distinguish a relic-neutrino signal for $m_\nu < 0.85$ eV unless $\eta > 10^6$; simply increasing the tritium mass does not improve sensitivity, since the β-decay rate scales in the same way as the neutrino-capture rate [111].

5.7. Other new physics

In this section we briefly describe additional new physics scenarios that could be probed by KATRIN, but that the collaboration has not yet studied in detail.

If a dark-matter particle is light, and decays into neutrinos, tritium may capture them and generate via inverse β decay an electron with energy beyond the endpoint. This interesting application of a neutrino-mass experiment as an indirect dark-matter experiment has been proposed in references [115, 116], and applied to the proposed PTOLEMY experiment, see figure 15 (left). Of course, KATRIN could already set limits on such scenarios.

Another example is β-decay in a medium, i.e. a background field, which could be formed by dark matter, dark energy or a fifth-force potential. Neutrinos couple to this background field, and this modifies their dispersion relation, which has an impact on the measurable electron spectrum (figure 15, right). The interaction can be of scalar, pseudoscalar, vector or axial-vector type. Scalar and pseudoscalar interactions mimic the effect of neutrino mass, and therefore could explain a disagreement between cosmological and KATRIN results on neutrino mass.
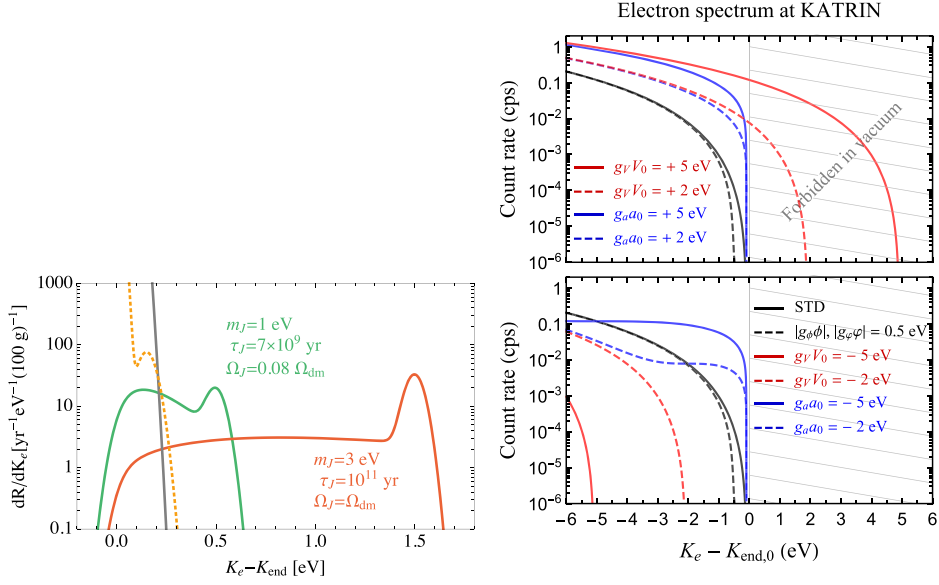
Figure 15. Left: event rates for capture on tritium for neutrinos from Majoron decay $J \rightarrow \nu\nu$ for $m_J = 1$ eV, $\tau_J = 7 \times 10^9$ yr, $\Omega_J/\Omega_{dm} = 0.08$ (green, solid) and $m_J = 3$ eV, $\tau_J = 10^{11}$ yr, $\Omega_J/\Omega_{dm} = 1$ (red, solid) and natural β decay (gray, solid) with an electron energy resolution of 0.1 eV. For comparison, we plot the total rate for signal plus background for a Majorana relic-neutrino signal with $m_\nu = 0.15$ eV (yellow, dashed). Reproduced from [115]. CC BY 4.0. Right: the integrated β-decay spectra at KATRIN in various types of media: scalar, pseudoscalar and two different vector and axial-vector potentials, with neutrino mass fixed to 0.1 eV. Reproduced from [117]. CC BY 4.0.

Even more interesting are the axial-vector and vector forms. The former can lead to a characteristic spectrum distortion; the latter can feature the spectacular effect of the β spectrum extending beyond the normal kinematic limit, because processes which are not kinematically allowed in vacuum can take place if the medium modifies the dispersion relations. Taking publicly available information, reference [117] set limits on the interaction potentials that are expected to dominate, depending on the explicit model realization.

6. Additional physics observables

6.1. Molecular final states from T_2 decay

With a molecular tritium (T_2) source, KATRIN measures a spectrum that is modified by the molecular FSD: the spectrum of rotational, vibrational and electronic modes excited in the daughter $^3HeT^+$ molecule by the beta decay of one tritium nucleus in the parent molecule. As a result, the neutrino-mass analysis requires knowledge of the energy levels of these distributions, as well as the probability that each is excited in the beta decay. KATRIN's spectral analysis relies on precise calculations of this distribution, computed using the geminal-basis method over the last 20-odd years [15, 47, 48]. It is impractical to measure the entire FSD, but observables predicted by this theoretical framework have good agreement with a wide range of experiments, summarized in reference [118]. Recently, the TRIMS experiment followed up on the one area of discrepancy with existing experiments, the branching ratio to the bound

molecular ion $^{3}\mathrm{HeT}^{+}$. The TRIMS results agree closely with the theoretical calculations for both HT and TT parent molecules [119].

The KATRIN spectrum provides a unique sensitivity to an observable that has never yet been measured: the onset to the electronic excited states. This observable is of interest beyond the neutrino-mass community. Given the prevailing abundance of H and He in the Universe, $^{4}\mathrm{HeH}^{+}$—which is described using the same theoretical framework as its isotopolog $^{3}\mathrm{HeT}^{+}$—played an important role in the chemistry of the early Universe, and has recently been detected for the first time in an astrophysical source, in the planetary nebula NGC 7027 viewed by SOFIA [120]. Although ample spectroscopic data are available for the electronic ground states of this molecule (reviewed in references [118, 121]), the electronic excited states have a purely repulsive character in the Franck–Condon window and are very difficult to probe through traditional spectroscopic methods. The available experimental data on these states are very limited [122–124].

KATRIN's sensitivity to the electronic excited states arises from the theoretical prediction of a separation of about 20 eV between the highest-lying states of the electronic ground-state manifold, and the first electronic excited state. This relatively sudden onset should cause a distortion in the spectral shape, through which KATRIN can measure the first electronic excited state's location and amplitude.

6.2. Precision spectroscopy of ^{83m}Kr conversion-electron lines

KATRIN is not limited to tritium spectroscopy, but can in principle be used to measure other electron spectra with comparable energies. In fact, the first measurement of a gaseous source at KATRIN was in 2017 with the ^{83m}Kr calibration source described in section 2.1.3. These repeated measurements are performed for their crucial contributions to the understanding of important systematic effects, but they also allow spectroscopy of the source; some results of the 2017 measurement campaign have already been published [30, 125]. This section discusses the possibility of further spectroscopic measurements, based on the experience of several ^{83m}Kr campaigns since 2017; however, some findings should be transferable to the spectroscopy of other nuclides.

Isomeric ^{83m}Kr decays via two consecutive γ transitions with energies E_γ of roughly 32.2 and 9.4 keV. The direct transition with an energy of 41.6 keV is suppressed but possible, and KATRIN has observed clear indications of this transition. In addition to γ emission the transition can also occur by internal conversion, where the transition energy is transferred to an electron in the atomic shell s, which is then emitted. The ratio of emitted electrons to γ photons is called the conversion coefficient $\alpha^{(s)}$. To calculate the energy $E_{\mathrm{ce}}^{(s)}$ of the emitted electron, the binding energy $E_{\mathrm{ce}}^{(s)}(\mathrm{bind})$ of the shell and the recoil energy of the nucleus ($E_\gamma(\mathrm{recoil})$ and $E_{\mathrm{ce}}(\mathrm{recoil})$) need to be subtracted from the transition energy, giving [126]

$$E_{\mathrm{ce}}^{(s)} = E_\gamma + E_\gamma(\mathrm{recoil}) - E_{\mathrm{ce}}^{(s)}(\mathrm{bind}) - E_{\mathrm{ce}}(\mathrm{recoil}). \tag{12}$$

In x-ray notation the electron shell is given as $s \in \{\mathrm{K}, \mathrm{L}_{l_1}, \mathrm{M}_{l_2}, \mathrm{N}_{l_3}, \ldots\}$, where the capital letters $\{\mathrm{K}, \mathrm{L}, \mathrm{M}, \mathrm{N}, \ldots\}$ denote the principal quantum number and the index l_i orbitals of different angular momentum.

The binding energies vary strongly between the shells, ranging from around 14 keV for the innermost K-shell electrons to 14 eV for the outermost N-shell electrons. The recoil energies are very small in comparison. Consequently, the electron spectrum is divided into several line groups. An overview of the 32.2 keV ^{83m}Kr conversion spectrum is shown in table 3.

Table 3. Mean line position E_{ce}, width Γ and conversion coefficient α of the 32.2 keV transition of the ^{83m}Kr conversion spectrum (Reproduced from [126]. © 2018 IOP Publishing Ltd and Sissa Medialab. All rights reserved.).

Line	E_{ce} (eV)	Γ (eV)	α
K	17824.2(5)	2.71(2)	478.0(50)
L_1	30226.8(9)	3.75(93)	31.7(3)
L_2	30419.5(5)	1.25(25)	492.0(50)
L_3	30472.2(5)	1.19(24)	766.0(77)
M_1	31858.7(6)	3.5(4)	5.19(5)
M_2	31929.3(5)	1.6(2)	83.7(8)
M_3	31936.9(5)	1.1(1)	130.0(13)
M_4	32056.4(5)	0.07(2)	1.31(1)
M_5	32057.6(5)	0.07(2)	1.84(2)
N_1	32123.9(5)	0.40(4)	0.643(6)
N_2	32136.7(5)	0.03	7.54(8)
N_3	32137.4(5)	0.03	11.5(1)

All electron conversion lines of the 9.4 keV transition, and the K conversion line of the 32 keV transition, are below the tritium β-spectrum endpoint of 18.6 keV and are therefore hidden in the β spectrum, when the source operates at the nominal tritium activity. This puts a first constraint on the spectroscopy of those lines, which would require a tritium-free source, possibly including a cleaning of residual tritium from the rear-wall surface. Consequently, only the conversion electrons of the 32.2 keV transition have been used in most of the ^{83m}Kr campaigns since the first exposure with tritium.

The analysis of the krypton spectrum needs no less attention than that of the tritium spectrum, especially since all measurements of the gaseous source are exposed to the same source systematic effects. This is exactly what one makes use of in calibration measurements: if it is given that the experimental conditions are the same and that one knows the krypton spectrum very well, one can use the krypton data to determine systematic effects and take them into account in the modeling of the tritium measurement. One of the systematic effects which the krypton measurement aims to quantify is the electric starting potential of the source, whose mean value shifts the position of the selected krypton line and whose inhomogeneity broadens the line. Conversely, one can only determine the absolute positions and widths of the krypton lines well if one knows the source potential well.

Thus, for spectroscopic measurements the electric potential of the source is a systematic effect. When deciding which lines to use for calibration or for purely spectroscopic measurements, two additional systematic effects must be considered in addition to typical indicators such as the signal-to-background ratio:

- **Inelastic scattering:** it is usually advisable to use a carrier gas in addition to the krypton to achieve high rates and rate stability. Moreover, the usage of tritium also leads to a smaller inhomogeneity of the electric potential of the source, since the β decay creates a plasma which shields inhomogeneities of the source-tube work function. On the other hand, inelastic scattering of the krypton conversion electrons on the carrier gas produces a shape distortion of the krypton spectrum starting typically around 10 eV below the line, such that only the electron conversion lines of outer shells can be measured without precise modeling of the scattering.
- **Background slope:** due to the integral measurement principle of KATRIN, the conversion spectra of higher-shell electrons contribute to the background in any spectrum

measurement of the inner-shell conversion electrons of ^{83m}Kr. Since the transmission conditions are non-adiabatic for high surplus energies, leading to a loss of rate, this background can have an energy slope, which needs to be modeled precisely.

For the regular calibration measurements the $N_{2,3}$-32 doublet is typically chosen for several reasons: since the $N_{2,3}$ electrons belong to the outermost atomic shell of krypton, there is no background contribution from other conversion electrons from the 32.2 keV transition. Also, since there is no filling up of the $N_{2,3}$ vacancies after emission of the conversion electrons from outer shells, they have a long lifetime and the lines therefore have a quasi-vanishing width. The ratio of the line width and the standard deviation of the source potential dictates which uncertainties dominate; for the $N_{2,3}$-32 doublet the $\mathcal{O}(10\,mV)$ measured value of the standard deviation of the source potential is much larger than the line width, such that uncertainties of the width have a minor effect on the determination of the standard deviation. For all other lines of the krypton spectrum this ratio is reversed, which means that they are not optimally suited for determining the source potential. However, KATRIN has the capability to determine their line widths even given uncertainties in the modeling of the source potential.

The determination of the absolute line positions requires knowledge of the absolute electric potential of the source, which in turn is calibrated using ^{83m}Kr measurements. Conversely, the determination of the absolute electric potential is limited by the literature uncertainties of the line positions. These uncertainties arise from uncertainties in the binding energies of krypton and the uncertainty on the γ transition energy (equation (12)), with the latter strongly dominating. KATRIN plans to decrease this uncertainty using the excellent linearity of the KATRIN spectrometer [127]: by determining the relative positions of lines of the conversion electrons emitted from the same atomic subshells of all three γ transitions (9.4, 32.2, 41.6 keV) KATRIN can measure the energy with high precision [128].

The determination of the absolute activity or half-life requires precise modeling of both the gas dynamics and emanation of the ^{83m}Kr from the rubidium generator. Due to the continuous supply of ^{83m}Kr from the generator, it is in quasi-equilibrium and only the half-life of rubidium is actually observed. The expected and measured krypton activity agree within uncertainties [129], but the modeling is not tailored for precise determination of the absolute activity.

Using a ^{83m}Kr source in the GBq range, KATRIN is capable of achieving a statistical sensitivity to line positions of a few meV, to line width in the 10 meV range and to intensities in the per mil range in a few hours, depending on the chosen lines and mode of ^{83m}Kr operation. In the literature, the uncertainty of the absolute line positions is in the 500 meV range (dominated by the uncertainty of the γ energy), and that of the relative line positions within one γ transition in the 40 meV range (dominated by the uncertainties of the binding energies), while the uncertainties of the line widths are in the 100 meV range [126]. Thus, KATRIN could in principle significantly improve these values.

In summary, using KATRIN for spectroscopy of ^{83m}Kr (or other electron spectra beyond tritium) is possible but it requires an assessment of systematic effects which is specific for each chosen line. For the $N_{2,3}$-32 doublet, measurements are performed regularly and the relative intensity and position of N_2 and N_3 can be obtained by KATRIN with unprecedented sensitivity. In addition, a measurement of the N_1-32 line region was performed in 2021 using helium as carrier gas. The latter requires at least 21.2 eV energy for incident electrons to scatter inelastically, so that scattered electrons from the $N_{2,3}$-32 doublet appear only below the N_1-32 line. Thus, while the current application of ^{83m}Kr measurements is focused on reaching KATRIN's physics goal, some spectroscopic byproducts have already been obtained and will be part of a future publication.
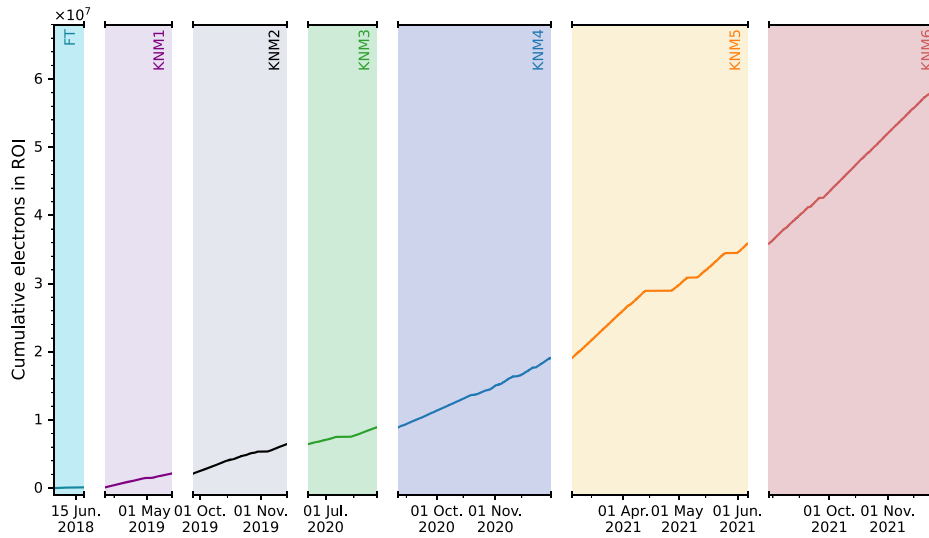
Figure 16. Integrated luminosity in the analysis interval since the beginning of tritium data-taking with KATRIN (section 2.2). FT denotes the first tritium campaign; each neutrino-mass measurement campaign is numbered sequentially in the format KNMx. White gaps correspond to scheduled maintenance periods.

7. R & D for the future of KATRIN

In 2005, the KATRIN design report set a sensitivity goal of 0.2 eV at 90% confidence [12]. This goal was based on an assumption of three full years of running, and a set of projections about achievable backgrounds and systematic uncertainties. KATRIN has now successfully acquired neutrino-mass data during three calendar years (figure 16). Several systematics once regarded as especially difficult are now exceptionally well controlled, e.g. the description of β energy loss in the source; others, not originally anticipated, are now recognized as particularly important (section 3.4). Meanwhile, KATRIN's designed countermeasures against the dominant backgrounds of prior generations, including cosmic-ray muons [36] and environmental gammas [37], have been largely successful—but previously unforeseen backgrounds (section 2.3) far exceed the original design budget.

In this section, we discuss R&D efforts—some quite early, and some advanced—with the goal of improving KATRIN's sensitivity to m_ν^2 and other physics observables. One of the foremost efforts is toward background reduction (section 7.1). Several countermeasures have already been successfully applied (section 2.3), including more frequent regeneration of liquid-nitrogen-cooled baffles on the main-spectrometer pump ports; reduction of the pre-spectrometer operating voltage; and operation of the main spectrometer with a SAP [26] that more than halves the spectrometer volume imaged by the detector (section 2.1.2). In addition, KATRIN is investigating both active (section 7.1.1) and passive (section 7.1.2) transverse filters, to distinguish background Rydberg electrons from signal βs. THz radiation is a possible means of actively de-exciting Rydberg atoms, reducing the number of background electrons produced (section 7.1.3). Electron taggers (section 7.1.4) could flag electrons entering the main spectrometer, allowing straightforward rejection of backgrounds originating inside the main spectrometer.

An increase in the KATRIN acceptance angle would improve statistics for any given amount of running time; investigations are underway to establish that this would not result in an unacceptable worsening of systematics (section 7.2). A further statistics gain could be realized by increasing the size of the analysis interval by reducing the lower limit of energies considered to, for example, $E_0 - 60$ eV from the present $E_0 - 40$ eV; this requires establishing high confidence in the models for electron energy loss and the molecular FSD (section 3.4).

Dedicated searches for keV-scale sterile neutrinos (section 5.2) require upgrades to the KATRIN apparatus. To scan deeper into the spectrum, the rear wall must be redesigned to reduce backscattering (section 7.3.1). Significantly higher rates will require a new detector system, which we call TRISTAN (section 7.3.2).

7.1. Background mitigation

After the successful implementation of initial measures to reduce the Rydberg background component [26] (section 2.1.2), KATRIN's background is still an order of magnitude higher than anticipated. Therefore we are investigating several methods to reduce the background further.

7.1.1. Active transverse-energy filter.
As described in section 2.3, our evidence suggests that a significant share of the background electrons in the KATRIN spectrometer originate from highly excited Rydberg atoms, which are ionized by the BBR emitted from the spectrometer walls at room temperature [45]. These secondary electrons possess very low starting energies [130], but are accelerated by the electric-potential gradient of the spectrometer and arrive with essentially the same kinetic energy at the detector as the signal electrons. However, since the initial transverse energy of the electrons from this process is only $\mathcal{O}(k_{\mathrm{B}} T \approx 25$ meV$)$ and they do not acquire transverse energy in the electric-field gradient towards the exit of the spectrometer due to the strong magnetic field there, they reach the detector with incidence angles of typically less than $10°$ in contrast to the signal electrons which fill the phase space up to the maximum angle of $51°$ allowed by the pinch magnet (figure 17).

The idea of the active transverse energy filter (aTEF) is to utilize the different angular distributions of signal and background electrons at the detector to further reduce the background rate. Standard methods to differentiate different angles of incidence, like low-density tracking detectors (e.g. gas-filled drift chambers) or a ΔE–E arrangement, are not applicable because the electrons in the strong magnetic field at the KATRIN detector possess sub-mm cyclotron radii and ranges of a few μm. In addition, a gas-filled detector with a thin entrance window is not ultra-high-vacuum-compatible. Therefore, we propose to geometrically distinguish the impact angles via the cyclotron radius of the spiraling motion around the magnetic field lines. At first glance a suitable device for this purpose could be a micro-channel plate (MCP) which we will thus use to explain the idea [131].

Typically, when a charged particle such as an electron or ion hits the side walls of a channel within the MCP, it triggers the emission of several secondary electrons from the wall, which is coated with a material with a particularly high secondary-electron yield. These secondary electrons are accelerated by the electric field applied across the channel, hit the wall again and, thus, trigger a secondary-electron cascade [132]. In commercial MCPs, the channels with diameter d are tilted by $8°$ to $15°$ [132] with respect to the surface normal. We deliberately neglect this tilt angle of the channels in figure 18 to illustrate the principle of our idea.

If an electron with a small cyclotron radius $r_{\mathrm{c}} \ll d/2$ (orange spiral track) enters a channel of the MCP, it usually passes through unaffected and causes a small signal in the detector (here in light blue). If, on the other hand, an electron with a large cyclotron radius $r_{\mathrm{c}} > d/2$ (purple spiral track) enters a channel, the particle will most likely hit the channel walls and trigger
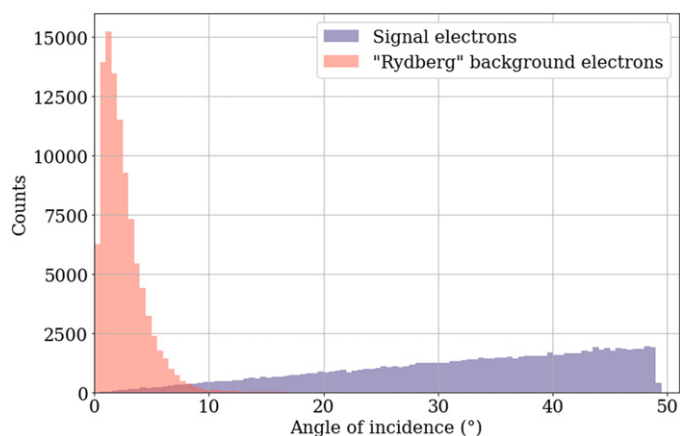
Figure 17. Simulated angular distributions of signal (purple) and 'Rydberg' background electrons (orange) at the detector of the KATRIN experiment simulated for 1×10^5 electrons using input from [130]. The simulation considers the energy spectrum of background electrons that are liberated by room-temperature BBR from hydrogen (36%) and oxygen atoms (64%) in highly excited Rydberg states. These Rydberg atoms are sputtered from the spectrometer walls due to α-decays of implanted ^{210}Pb progenies from the ^{222}Rn decay chain [45] (section 2.3). Reproduced with permission from [131].
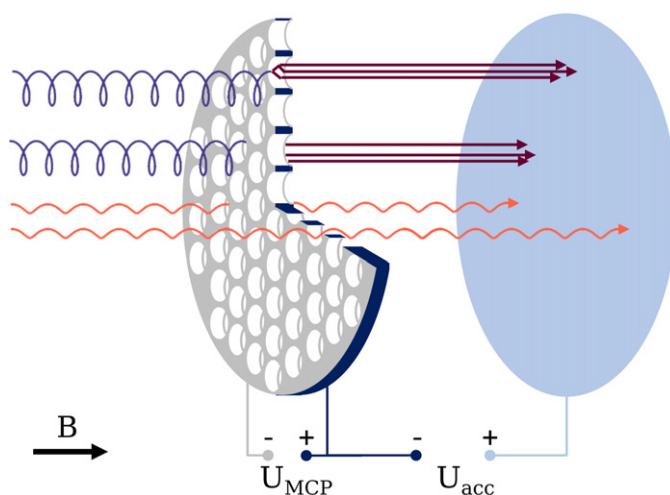
Figure 18. Principle of an MCP-based aTEF in front of the KATRIN detector (light blue) in a horizontal magnetic field *B*. Electrons with a small cyclotron radius (orange) pass the MCP without interaction and hit the detector. Electrons with a large cyclotron radius (purple) hit the MCP walls, giving rise to a cascade of secondary electrons (dark red) that is accelerated via U_{acc} toward the detector. Since the secondary electrons are produced with practically no transverse momentum, they follow the magnetic-field lines without a sizable cyclotron radius. Reproduced with permission from [131].

a secondary-electron avalanche (dark red arrows). An applied voltage U_{acc} will accelerate the electron avalanche toward the detector where it will cause a comparably larger signal. When we
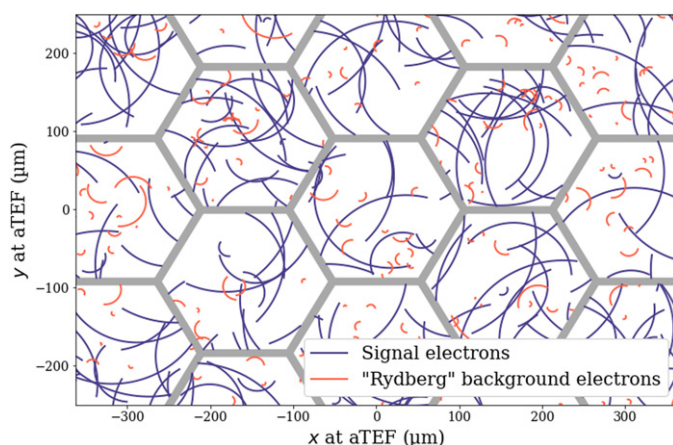
Figure 19. Simulation of signal (purple) and background (orange) electron tracks in a 2.4 T axial magnetic field with the angular distributions shown in figure 17 for a honeycomb-like aTEF structure with a side length of 100 μm, a wall thickness of 10 μm, and a depth of 400 μm. The tracking of an electron is stopped when it either hits the aTEF surface or when it reaches the full depth without any hit. Reproduced with permission from [131].
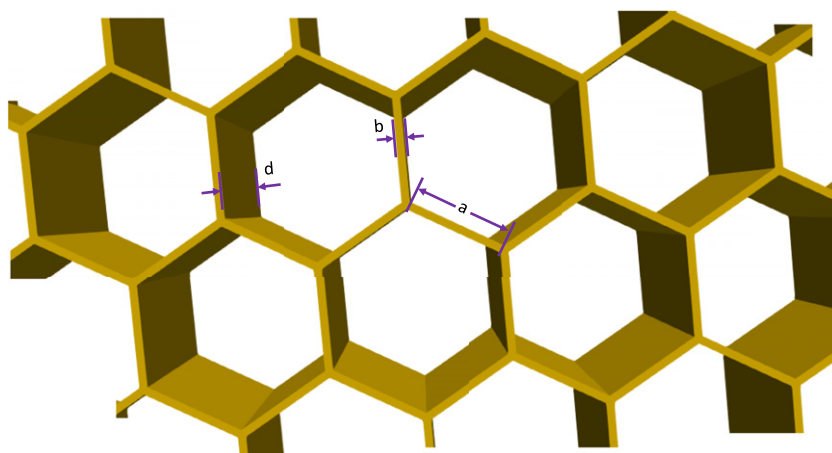
Figure 20. Cutout of the geometric pTEF model as simulated in Kassiopeia. The hexagonal channels are described with the parameters side length a, wall thickness b, and depth d.

discriminate on the measured signal height, the channels thus act as a 'transverse energy filter'. Given that the filter is not a passive element but aids in the detection of the signal electrons, we name this principle 'aTEF' [131].

Of course a standard MCP does not meet the KATRIN requirements. First, we have to adapt the diameter-to-length ratio of the channels to allow background electrons with small transverse energy to pass the filter unhindered while most of the signal electrons are detected. A channel bias angle is disfavored, unlike in typical commercial MCPs where it is mainly used to
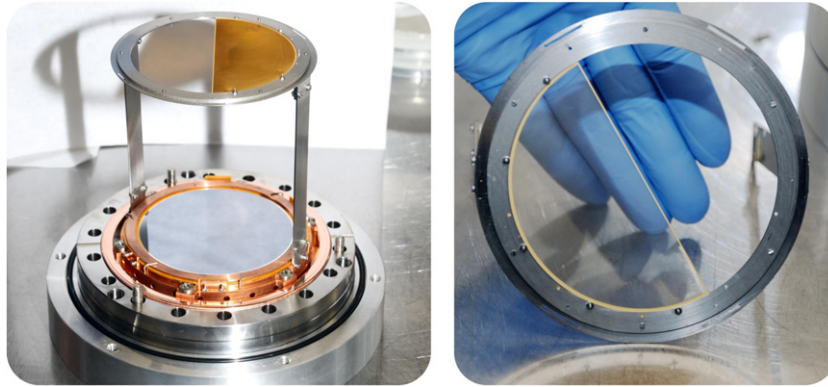
Figure 21. Left: pTEF (gold plate) mounted with a dedicated holding structure on the focal-plane-detector wafer flange before installation into the KATRIN beamline. Right: pTEF installed in holding structure.

reduce ion feedback [133]. In contrast, the channels for the proposed aTEF have to be parallel to the magnetic-field lines. Secondly, the open-area ratio (OAR) needs to be maximized to reduce losses in non-background events due to signal electrons being blocked by the material in between the channels. Using a hexagonal channel shape instead of the circular shape of regular MCPs allows much larger OARs. To overcome the inherent dark-count rate of the glass materials usually used for MCP production, a radiopure material like silicon with an intrinsic low background rate could alternatively be used. The targeted geometry can be manufactured in silicon via a highly anisotropic cryo-etching procedure for deep silicon structures.

There are obvious disadvantages of an MCP-based aTEF for our application, like the low secondary-electron yield of most materials for 18.6 keV electrons or the sensitivity to the very strong magnetic field of 2.4 T. We do not discuss these here, since the aTEF idea is more general. This principle can be realized on any kind of geometric aperture that is instrumented with a suitable method of particle detection. A generally suitable aTEF-aperture geometry is a honeycomb structure due to its ideal OAR. The instrumentation of such a honeycomb structure could, for instance, also be realized using scintillator or semiconductor materials to generate photons or electron–hole pairs and thereby detectable signals. The advantage of the latter is that the energy of the signal electrons could still be determined, an interesting feature for suppression of other kinds of backgrounds.

A Monte Carlo simulation of electron tracks through a filter design with hexagonal channels, a side length of 100 μm, a channel wall thickness of 10 μm and a depth of 400 μm is illustrated in figure 19 (see figure 20 for definitions of these parameters). Electrons are generated with angles drawn from the two angular distributions shown in figure 17, relative to the 2.4 T magnetic field at the entrance of the three-dimensional hexagonal structure. Corresponding to the angular distribution used, the simulated electrons are either of Rydberg-background type (orange tracks) or of signal type (purple tracks). The magnetic-field lines and, therefore, the guiding centers of the electron's motion are directed perpendicular to the viewing plane. The tracking stops when the electron either hits the wall or reaches the full depth of the channel of 400 μm. Electrons are counted as 'detected' in the former and as 'not detected' in the latter case. The percentage of detected electrons differs noticeably between the two types of electrons: 90% of the signal electrons and 11% of Rydberg background electrons interact with the wall. Taking into account an OAR of 90%, we therefore retain 81% of the signal events and 10% of the background events. Research and development on this concept is continuing.

7.1.2. Passive transverse-energy filter. To inform the development of the aTEF (section 7.1.1), we have built a passive transverse-energy filter (pTEF) to measure the angular distribution of the background electrons, testing the prediction in figure 17. The pTEF is a gold semicircle with a lithographed, hexagonal honeycomb structure, designed to block electrons within a certain range of transverse kinetic energy. This semicircle is installed to shadow one half of the detector while the other half is left clear (figure 21). In this way, different pitch-angle distributions are distinguishable based on the rate on the respective detector sections. The geometry (figure 20) is described by three parameters: the side length a, wall thickness b and depth d. These parameters were optimized in terms of signal-to-background ratio, based on simulations of the KATRIN beamline performed with Kassiopeia [53] code for tracking charged particles in varying electromagnetic fields. Additional parameters, such as diameter or OAR, are derived from these. The values of the pTEF design are:

$$a = 100 \ \mu\text{m}, \quad b = 8 \ \mu\text{m}, \quad d = 250 \ \mu\text{m}, \quad \text{and} \quad \text{OAR} = 91.4\%. \tag{13}$$

The electron motion can be simplified to a two-dimensional case. Electrons with non-zero pitch angles circulate around magnetic-field lines due to the Lorentz force. If we let the motion and magnetic-field lines lie along the beam axis (z), the circular motion leads to a circle in the xy plane. Therefore, the electron direction has components parallel and perpendicular to $\vec{B}$. The size of the circular motion is described by the cyclotron radius

$$r_\text{c} = \frac{\gamma \cdot m \cdot v_\perp}{|q| \cdot B}, \tag{14}$$

and its frequency $\omega = q \cdot B/m$, where B is the magnetic field strength, q the charge and m the mass of the particle, γ the Lorentz factor, and $v_\perp$ the velocity perpendicular to the magnetic field. The pitch angle θ, between the electron velocity and the magnetic-field line, gives the relative size of the $v_\perp$ component. The energy due to transverse motion, along with the cyclotron radius, is higher for larger θ. Figure 22 shows the dependence of various parameters on θ for selected electron energies. The cyclotron radius as a function of the polar angle θ is proportional to $\sin \theta$ (top left). The top right shows the $\cos \theta$ dependence of the projection of the electron path along the z-axis, during one period of circular motion. For energies relevant to KATRIN, these path lengths always exceed the 250 μm pTEF depth. By comparing the electrons' area of motion in the xy-plane to the area of the circle inscribed in a hexagonal cell (bottom left), we find that for 18.6 keV the area of motion already exceeds the open area for $\theta = 27°$. However, the depth of the filter is insufficient for a full period of circular motion, so electrons with even higher pitch angles can be transmitted. Figure 22 (bottom right) shows a simulation of electron transmission through a hexagonal filter. For each 1° angular bin, 50 000 electrons were simulated and counted as transmitted if they passed the filter without striking its surface. Electrons with 0° pitch are not 100% transmitted since the boundaries between the hexagons fill about 8.6% of the area. Small variations of the transmission occur between different electron kinetic energies, resulting from a combination of the different cyclotron radii and paths along z.

A multifaceted measurement campaign was performed in winter 2021, including detailed investigations into transmission through the pTEF and therefore into the θ distribution of the background electrons generated in the spectrometers. The results are under analysis.

7.1.3. De-excitation of Rydberg atoms with THz radiation. An alternate approach to mitigate the Rydberg background is to manipulate the atomic states of the Rydberg atoms so as to reduce the ionization process by thermal radiation (section 2.3). This method was pioneered
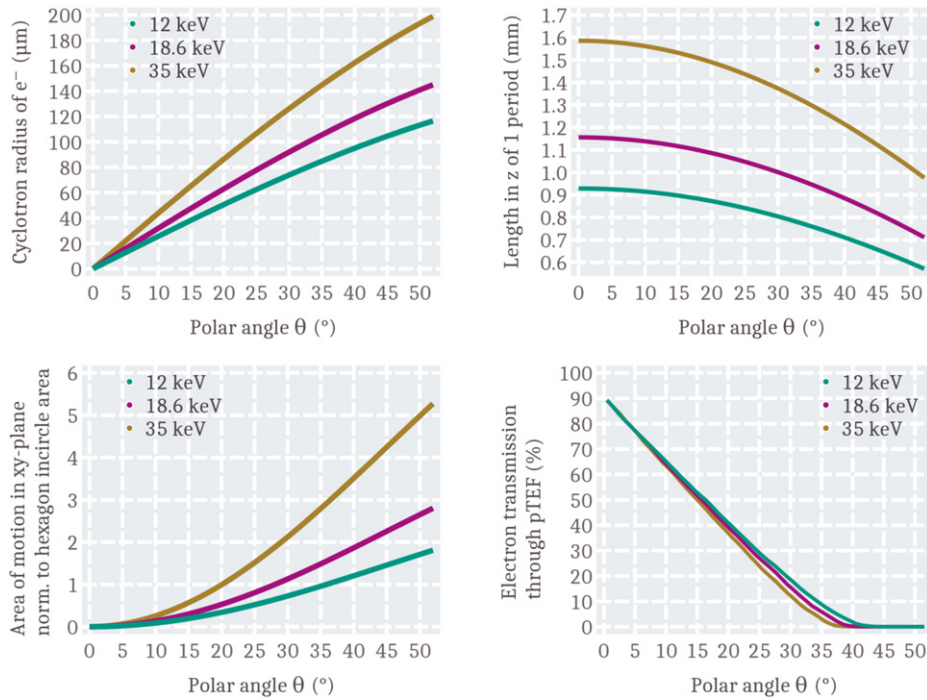
Figure 22. Dependence of relevant quantities on pitch angle θ for electron energies of 12, 18.6, and 35 keV at 2.5 T magnetic field. The minimum and maximum energies give the largest spread that can be measured in a practical setup. Top left: cyclotron radius. Top right: axial path length of one full circular motion. Bottom left: area of circular motion in xy plane, normalized to the area of a circle inscribed in a hexagonal pTEF channel. Bottom right: simulation of electron transmission through a hexagonal microstructured plate. Refer to figure 17 for the predicted angular distributions of signal and background electrons.

by the CERN antihydrogen group [134, 135]. KATRIN is currently investigating this idea via numerical simulations.

Here, we assume that the Rydberg atoms are produced by charge-exchange mechanisms during the sputtering processes in the main-spectrometer walls [130] (see figure 4 above). Such mechanisms tend to produce a Rydberg-state distribution favoring low principal quantum numbers n and high angular quantum numbers l, with probabilities P according to $P(n, l) = (2l + 1) \cdot n^{-5}$. During their transits though the main spectrometer, the Rydberg atoms undergo one or more of the following processes: spontaneous decay with transition rate Γ_{sp}, BBR-induced excitation or de-excitation ($\Gamma_{\text{trans}}^{\text{BBR}}$) and BBR-induced ionization ($\Gamma_{\text{ion}}^{\text{BBR}}$). $\Gamma_{\text{ion}}^{\text{BBR}}$ peaks for all l at about $n = 25$ (figure 23, left).

Therefore, we aim to mitigate the background with stimulated de-excitation after creation of the Rydberg atoms, quickly depopulating these levels. The goal is to drive multiple $\Delta n = -1$ transitions that de-excite the Rydberg atoms to states from which spontaneous decay is fast ($\tau \approx n^3 l(l + 1) \cdot 10^{-10}$ s). The frequencies of these transitions range from 2.2 THz for $n = 15 \rightarrow 14$ to 0.26 THz for $n = 30 \rightarrow 29$. However, a light source (LS) will not only cause LS-induced excitation/de-excitation ($\Gamma_{\text{trans}}^{\text{LS}}$) but also LS-induced ionization ($\Gamma_{\text{ion}}^{\text{LS}}$). To reduce additional ionization, narrow-band LSs can be chosen to target specific transitions. The LS
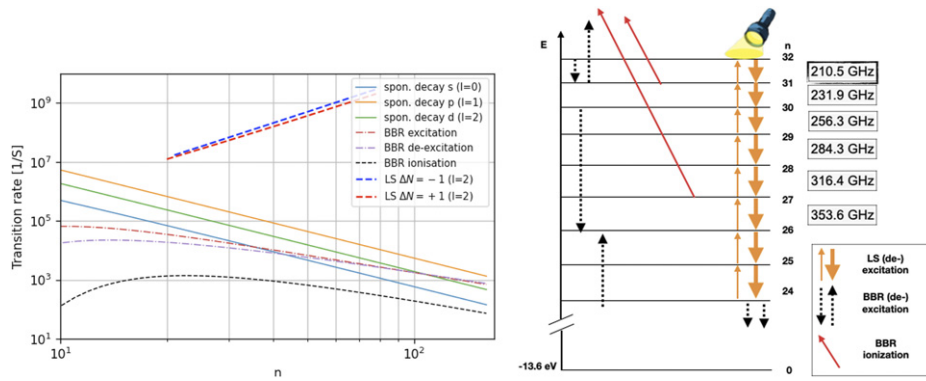
Figure 23. Left: transition rates for the different processes the Rydberg atoms undergo in the main spectrometer for low angular momenta l. The BBR-induced ionization peaks at about $n = 25$. Even at a radiance of only 0.005 W m^{-2}, the transition rate for LS-induced de-excitation is about five orders of magnitude higher than for ionization induced by BBR. Right: simplified transitions and binding-energy diagram of hydrogen. The thicker arrow for LS-induced de-excitation indicates the higher ionization rates for this process compared to the excitation channel (compare with left panel).

is simulated with a Lorentzian spectral distribution described by the center frequency, the radiance and the spectral width.

The Monte Carlo code calculates the Rydberg-state development according to the full transition rate

$$\Gamma_{\text{full}} = \Gamma_{\text{sp}} + \Gamma_{\text{trans}}^{\text{BBR}} + \Gamma_{\text{ion}}^{\text{BBR}} + \Gamma_{\text{trans}}^{\text{LS}} + \Gamma_{\text{ion}}^{\text{LS}}. \tag{15}$$

Figure 23 (left) shows the transition rates for low angular momenta.

Combined with a dedicated tracking code, which includes the initial kinetic energies and angular distribution of the different species of Rydberg atoms from SRIM [136], the ionization location and hence the background contribution can be calculated. We investigated a large parameter space, varying the number of targeted transitions and the intensities of the corresponding LSs. As a first, promising result, we found that a >50% reduction of the Rydberg background is possible with a set of eight equal-radiance, 0.005 W m^{-2} LSs driving $\Delta n = -1$ transitions from $n = 32$ to $n = 24$ (compare with figure 23 right).

So far the Stark and Zeeman effects are not considered in our model. Full inter-l-mixing with microwave radiation might be beneficial to our de-excitation goal, but will also induce potential losses through new excitation and ionization channels. However, these effects are expected to be very small as both the magnetic and the electric fields are weak in the bulk of the main spectrometer.

The practical challenge of this approach is that powerful sources with the necessary frequencies for the targeted transition are not easily available. Although THz technology is a fast-moving field, mW power is roughly the maximum achievable.

7.1.4. Electron taggers. As described in section 2.3, most of the KATRIN background electrons are produced within the volume of the main spectrometer. These backgrounds could, in principle, be vetoed based on the fact that they do not pass through any of the beamline components upstream of the main spectrometer. Detecting the passage of β electrons upstream of the main spectrometer allows electron times of flight to be measured for electrons incident on the

detector. The majority of tritium β electrons have times of flight of less than 100 μs through the main spectrometer. Detector events from background electrons produced in the main spectrometer could then be vetoed if an electron passage has not been detected within a suitable time window.

An electron tagger to perform this coincidence measurement must provide timing resolution better than ∼10 μs. The rate of electron tagging must be less than the inverse of the required timing resolution in order to avoid random coincidence rates dominating the time-of-flight measurement. This requires an electron tagger to be implemented in the limited space between an active pre-spectrometer and the main spectrometer. Therefore, the tagger would be sensitive to electrons in the pre-spectrometer Penning trap, but this would produce a distinctive oscillating signal in the tagger that could be vetoed. The electron energy loss to the tagger must also be known to better than 1 eV so as not to degrade the MAC-E filter width excessively. Previous studies have shown that there are no fundamental physical impediments to these requirements [137]. Further, with a timing resolution of <100 ns, a tagger would enable a differential, time-of-flight spectral measurement that would improve the KATRIN statistical sensitivity at a given run time [138].

The signal-to-noise ratio and required timing resolution are challenges for measuring the induced signal of the electron by an electron tagger. The development of an electron tagger involves two distinct sections: the pickup structure and the amplifier system. Loop pickups, plate capacitors, and resonators are being investigated as possible pickup structures used to couple with the passing electrons. Fundamental limits of a number of different amplifier systems, including SQUIDs, amplifiers operating at and below the standard quantum limit, and qubit state measurements, are also being investigated. The final goal is to find a combination of pickup and amplifier that provides a sufficient signal-to-noise ratio.

7.2. Acceptance improvements

KATRIN's design limits the acceptance angle of tritium electrons to $\theta_{max} =$ arc since the source magnetic field B_{src} is smaller than the pinch magnetic field B_{max}. The magnetic-field ratio is chosen to reject electrons with large starting angles and therefore larger systematic uncertainties. The goal is to balance out the statistical and systematic contributions to the neutrino-mass uncertainty.

One approach to improve the statistical uncertainty of the experiment is to increase the acceptance angle, so that more electrons reach the detector during a neutrino-mass scan. This can be achieved by changing the ratio of the magnetic fields, B_{src}/B_{max}. In our particular case, it is desirable to keep the source at its nominal field strength of ≈2.5 T, while reducing the pinch magnetic field together with all other beamline magnetic fields between source and detector. Hence, the entire beamline downstream of the source is operated at lower magnetic field compared to the nominal configuration. This has several advantages. First, by using the same ratio of analyzing and pinch magnetic fields, B_{min}/B_{max}, the active spectrometer volume that is viewed by the detector remains unchanged. Thus, the volume-dependent background from the spectrometer is the same as under nominal conditions. The main-spectrometer filter width is also unchanged. Furthermore, the systematic uncertainty on B_{min} depends partially on the total magnetic field strength in the AP. Lowering the analyzing magnetic field reduces this uncertainty accordingly. Finally, reducing the magnetic field in the transport section between source and spectrometers reduces the amount of synchrotron energy loss and its impact on the total systematic uncertainty.

On the other hand, the magnetic field in the source and in the rear section is at nominal strength. Contributions to the total systematic uncertainty from these sections have been
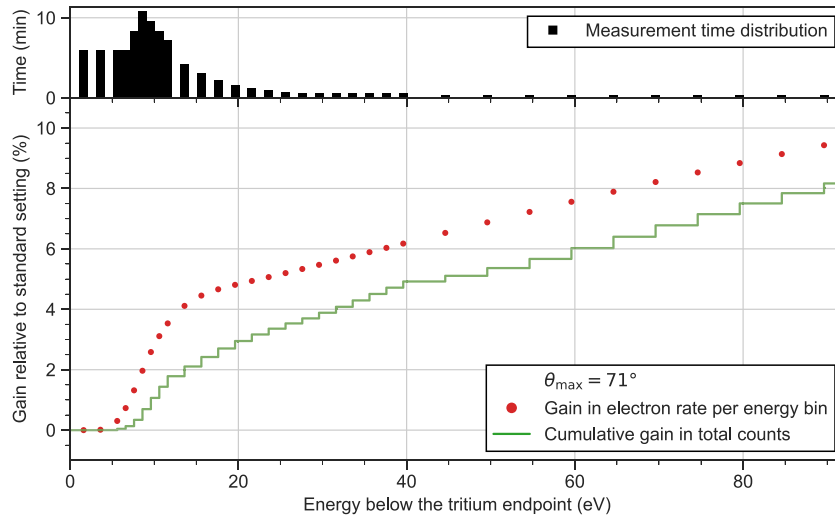
Figure 24. Expected statistics gain from an increased acceptance angle over an energy range of 90 eV below the tritium endpoint, at $\theta_{\max} = 71°$. The red data points show the rate gain at each filter energy, compared to nominal conditions from KATRIN's third measurement campaign. The green stepped line shows the cumulative gain in total statistics (counts) for the corresponding energy window [E; 0], with the distribution of measurement time shown in the top panel.

thoroughly investigated over past measurement campaigns. If these magnetic fields remain unchanged, these time-consuming commissioning measurements remain applicable.

By reducing the pinch magnetic field from its nominal strength of 4.2 T to 2.8 T, the acceptance angle increases from 50.4° to 70.9° at a source magnetic field of 2.5 T. The accepted forward solid angle is then $\Delta\Omega/2\pi = 1 - \cos\theta_{\max} = 0.67$, about 1.9 times larger than in the nominal configuration [15].

In addition to the accepted forward solid angle, the maximum electron rate at the detector is defined by the magnetic flux Φ. According to the equation $\Phi = \int B(r)\mathrm{d}r \approx B \cdot \pi r^2$ the flux is conserved in the beam line. The flux visible by the detector is given by the detector magnetic field, $B_{\det}$, and the detector radius of 4.5 cm. Scaling down the detector magnetic field by the same ratio as the pinch magnetic field ($B'_{\det} = 1.6$ T) reduces the transmitted flux to 102 T cm². This setting has the advantage of avoiding collisions of the flux tube with the beam-tube walls in the source due to the lower magnetic field in the beamline.

Figure 24 shows the expected gain in statistics relative to the nominal magnetic-field configuration, at different energies below the tritium endpoint E_0. To estimate the statistics gain, we determine the total electron counts (including background electrons) at the detector for each retarding potential, using the scenario from KATRIN's third measurement campaign with 125 active pixels at $B_{\max} = 2.8$ T and $B_{\min} = 0.4$ mT. The rate gain increases steeply over the first 15 eV due to the increasing contribution of single-scattered electrons. The cumulative gain over the full energy window depends on the measurement-time distribution and reaches an improvement in statistics by 8% over the investigated 90 eV energy window.

A higher acceptance angle allows electrons that have a high probability of scattering to reach the detector. Electrons lose a fraction of their kinetic energy in inelastic scattering processes on source-gas molecules. Hence, electrons with higher angles θ contribute mainly to the energy region that is several tens of eV below the tritium endpoint. A maximum gain in

statistics therefore requires increasing the energy window of the neutrino-mass analysis to $\gtrsim 100$ eV as shown in figure 24. This in turn requires an excellent understanding of the systematic uncertainties associated with source scattering processes and the tritium molecular FSD (section 3.4). Fortunately, recent developments in these areas indicate a much smaller contribution to the systematic uncertainty budget than originally anticipated. Hence, the gain in statistics from an increased acceptance angle becomes favorable in terms of balancing the statistical and systematic uncertainties.

7.3. R & D for a keV-scale sterile-neutrino search

This section describes research and development efforts in support of a high-statistics search for keV-scale sterile neutrinos (section 5.2), anticipated after the conclusion of neutrino-mass running.

7.3.1. Modified rear wall. A keV-scale sterile-neutrino search requires scanning over a much wider energy range than for the neutrino-mass determination [84]. Simulation studies show that backscattering of signal electrons on the rear wall of the source leads to a significant systematic uncertainty. By replacing the gold-plated rear wall with a beryllium disk, the backscattering coefficient can be suppressed, which leads to a reduction of the absolute size of this systematic effect by up to one order of magnitude [86].

7.3.2. TRISTAN detector. Since the mass associated with a possible sterile neutrino is unconstrained, the kink-like distortion of the tritium β-decay spectrum, caused by the emission of a sterile neutrino, could be located several keV away from the endpoint, see figure 9. As a consequence, the electron count rate is increased up to levels of $\mathcal{O}(10^8)$ cps), yet the current focal-plane detector is limited to a total rate of 10^5 cps integrated over all 148 pixels. To handle the exceedingly high rate and maintain excellent spectroscopic properties at the same time, i.e. an energy resolution of 300 eV FWHM at 20 keV, a multi-pixel SDD and readout system is currently being developed. The ultimate goal of the TRISTAN detector is to enable KATRIN to reach a sensitivity to the active-to-sterile mixing amplitude at the ppm level. With this, current laboratory limits could be improved by three orders of magnitude and the parameter space of cosmological interest could possibly be reached.

The TRISTAN detector follows the general design idea of SDDs used for x-ray spectroscopy [140]. The detectors are being produced at the Semiconductor Laboratory of the Max Planck Society (HLL). SDDs use the basic principle of sideways depletion, see figure 25. A volume of a high-resistivity semiconductor material, *n*-type silicon, is covered by rectifying p-doped junctions on both surfaces. A small substrate contact in reverse bias to the p-regions depletes the silicon bulk. The *p*-junctions are segmented strip-like and biased such that they generate an electric field with a strong component parallel to the surface. Signal electrons released within the depleted volume by the absorption of ionizing radiation drift toward the readout contact, i.e. the collecting anode. Due to the small physical dimensions of the anode, the detector has a small capacitance which is almost independent of the detector area. Compared to a conventional diode of equal area, this feature translates into larger amplitudes of the output signals. The anode of every pixel is read out by a charge-sensitive amplifier (CSA) with a JFET integrated into the anode structure of the chip, followed by a low-noise application-specific integrated circuit (ASIC) specifically developed for the TRISTAN detector [141]. The integrated JFet allows the ASIC chip to be placed at several cm distance to the detector chip, while keeping the total anode capacitance at only 180 fF. This provides an excellent signal-to-noise ratio. The read-out components are developed by the company XGLab, Politechnico di Milano
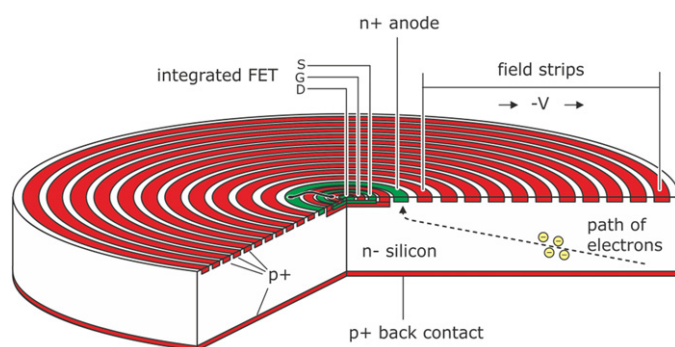
Figure 25. Schematic of a single pixel of the TRISTAN SDD, Reprinted from [139], Copyright (1996), with permission from Elsevier. Electrons are guided by an electric field to the collecting n+ anode in the center of the detector, which is surrounded by several drift rings. The entrance window for radiation is on the opposite side (p+ back contact). The detector features an integrated FET close to the anode which forms the first stage of the readout electronics.
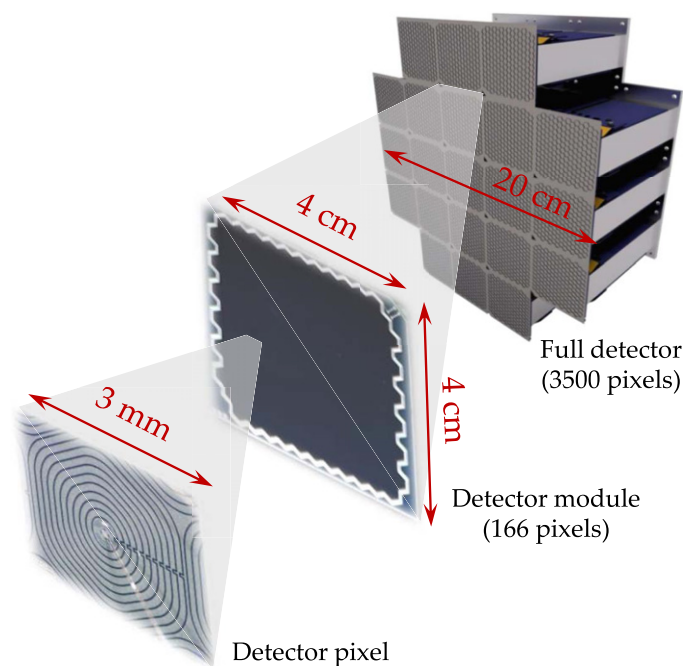

Figure 26. Design of the TRISTAN detector. The full detector array consists of 21 identical modules each consisting of 166 pixels. One module has a size of 4×4 cm^2 with the individual pixels having a diameter of 3 mm. The pixel anodes are connected to the CSA via wire bonds.

[142], the Institute of Data Processing and Electronics at KIT, and the Max Planck Institute of Physics.
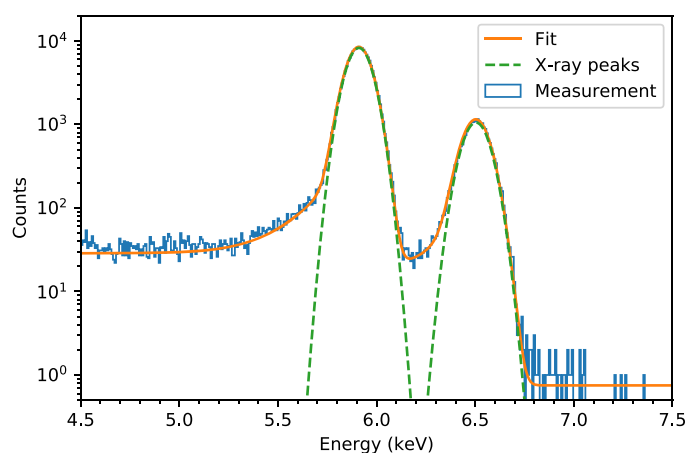
Figure 27. Energy spectrum of ^{55}Fe recorded with a 2 mm pixel-diameter SDD array at a temperature of $-30\,^{\circ}$C. The MnKα line at 5.9 keV has an energy resolution of 139 eV FWHM and is clearly separated from the MnKβ line at 6.5 keV. Two phenomenological functions similar to a hypermet [148] were used to fit the spectrum. Reproduced from [145]. CC BY 3.0.
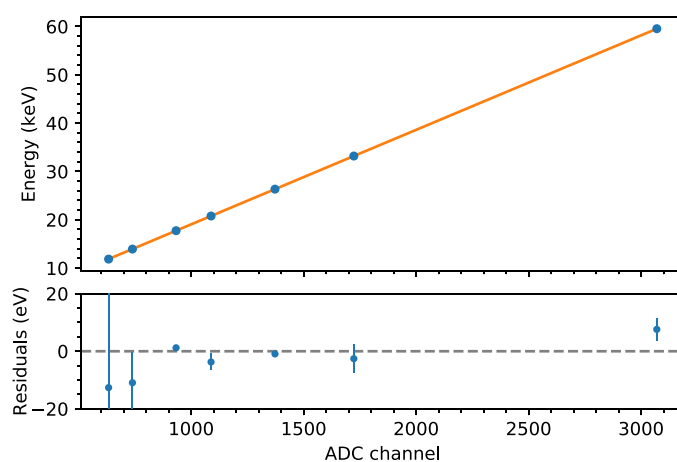
Figure 28. Linear calibration curve based on a measurement with an ^{241}Am source. The blue data points correspond to several photon lines. The error bars include the uncertainties of the line positions and of the fit. The maximal deviation from linearity is 0.1% at the lowest-energy peak of 11.9 keV. Reproduced from [145]. CC BY 3.0.

The SDD concept is very flexible in shape and size. However, a large single-pixel detector would have inherent limitations in terms of drift time and count rate. Therefore, the TRISTAN detector system will be based on multi-cell SDDs combining a large sensitive area with the energy resolution and the count rate capability of a single SDD (figure 26). A multi-cell SDD is a continuous, gapless arrangement of a number of SDDs with individual readout, but with common voltage supply, entrance window, and guard ring structure. The novel detector system is optimized to minimize effects which can alter the shape of the detector response.
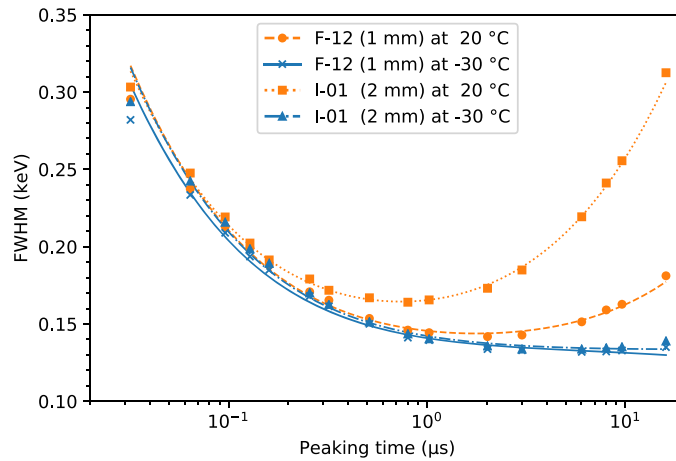
Figure 29. Energy resolution (FWHM) of the 5.9 keV line of ^{55}Fe as a function of the trapezoidal-filter peaking time, comparing the performance of the 1 mm and 2 mm pixel diameter detectors at room temperature (orange) and -30 °C (blue). Reproduced from [145]. CC BY 3.0.

The pixel size is chosen to be 3 mm in order to minimize charge-sharing and pixel changes after backscattering and back-reflection [85]. The entrance window has a minimal thickness of about 50 nm in order to minimize energy loss. Detector and readout electronics are optimized for low noise, and thus for good energy resolution. Finally, a full waveform digitization is chosen to minimize the effect of ADC non-linearities.

The development of the TRISTAN detector system follows a staged approach. Starting with seven-pixel detector prototypes with a simple mechanical design (first without an integrated JFET into the anode structure of the field-effect transistor (FET)), the detector chip was scaled to a more complex module consisting of 47 pixels, which has already been tested successfully [143]. The final focal-plane array will consist of nine (phase 1) and 21 (phase 2) detector modules of 166 pixels each.

7.3.2.1. Detector characterization. To test the general performance of the novel detector system, several measurements were performed with 1-, 7- and 47-pixel prototype detectors in dedicated setups [144–147]. In particular, we investigated the energy resolution, linearity, and electronic noise of the system. For all measurements, suitable electron, x-ray, and gamma calibration sources were used. For the seven-pixel detectors, we used devices without a JFET integrated into the anode structure, as this feature was only added at a later development stage. Instead, the anode of each pixel was directly connected to the input of an external ASIC via a wire bond.

A typical ^{55}Fe energy spectrum recorded with a 2 mm pixel-diameter SDD array at -30 °C is shown in figure 27. The MnKα and MnKβ lines at 5.9 keV and 6.5 keV, respectively, can be well approximated by Gaussian functions, with the entire structure best described by hypermet-type functions [148]. At the 5.9 keV x-ray line, an excellent energy resolution of 139 eV FWHM was obtained. To test the linearity of the system, an ^{241}Am source was used. As can be seen in figure 28, the system shows excellent linearity over the entire energy range from 10 keV to 60 keV. All measured line positions deviate from linearity by <0.1%. The noise performance of the detector system was investigated by measuring the energy resolution of the
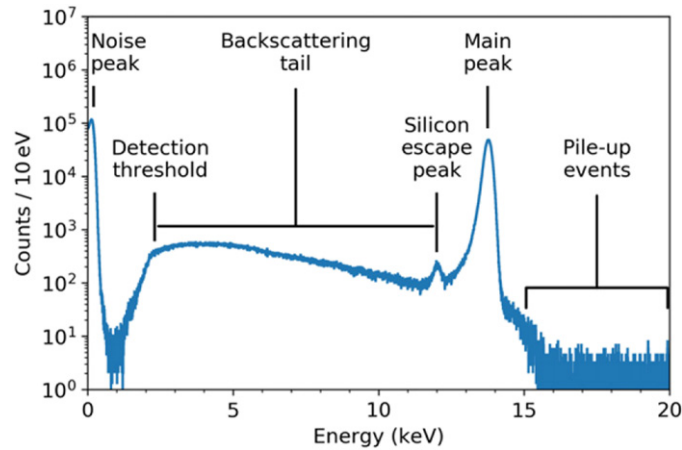
Figure 30. Spectrum of a 14 keV monoenergetic electron beam from an electron microscope measured with a TRISTAN prototype detector. Reproduced from [146]. CC BY 4.0.

5.9 keV MnKα line of ^{55}Fe as a function of the peaking time, which is a parameter of the trapezoidal filter applied to determine the energy of an event from the waveform. Figure 29 shows the typical dependence of the resolution on the peaking time for 1 mm- and 2 mm-pixel-diameter detectors, at room temperature and at −30 °C. The increased resolution at higher peaking times at room temperature is due to leakage current. As expected, it is more pronounced for the 2 mm detector since the leakage current grows with the area of the pixel. In contrast, at −30 °C, the leakage current is suppressed to a negligible level for both detectors. This shows that the performance is not degraded for larger pixel areas. An energy resolution of ⩽140 eV FWHM is reached for peaking times of about 1 μs. More details on the x-ray characterization measurements can be found in reference [145].

The application of SDDs to high-precision electron spectroscopy is novel to the TRIS-TAN detector. The energy-deposition profile is different for massive charged particles and for photons. This fact has several consequences: first, electrons deposit a significant fraction of their energy close to the entrance-window surface, where the electric fields are too weak to transport charge carriers to the readout electrode. Therefore, the energy of the electron is not fully detected. Second, low-energy electrons have a certain probability to scatter back from the detector surface, again resulting in a partial energy measurement. As a consequence, a detailed characterization of the detection system with electrons is of major importance.

Here, we present first characterization measurements of a seven-pixel prototype SDD with monoenergetic electrons at room temperature [144, 146, 147]. A scanning electron microscope was used as a calibration source. A typical electron energy spectrum is shown in figure 30. An empirical model was developed to describe the detector's response to electrons [146]. Each physical effect is modeled by a separate term including a Gaussian function, a low-energy tail, a silicon escape peak, and a backscattering tail. An example of the fit to the obtained energy spectrum is shown in figure 31, indicating very good agreement of the model with the data. More details on the electron-characterization measurements can be found in references [144, 146, 147], including an estimation of the entrance-window thickness as well as the impact of the detector response on the sterile-neutrino sensitivity.
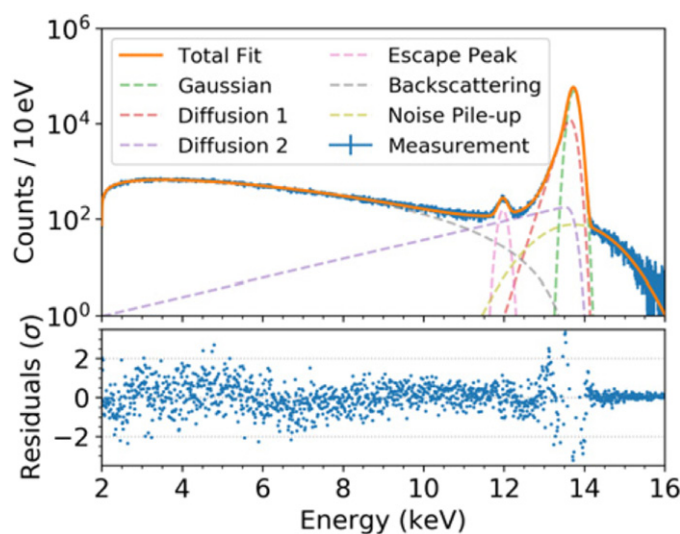
Figure 31. Fit of the electron response model to the measured energy spectrum. The energy of the electron microscope was set to 14 keV. The silicon escape peak and noise pile-up are approximated by two additional functions. Reproduced from [146]. CC BY 4.0.
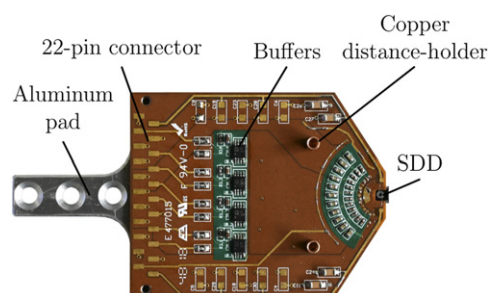

Figure 32. Photograph of the board hosting the TRISTAN FBM detector. Reproduced with permission from [149]. (a) 47 ch planar TRISTAN detector. (b) 47 ch 3D TRISTAN detector.

7.3.2.2. TRISTAN detector in the KATRIN beamline.   A seven-pixel TRISTAN prototype detector with a pixel diameter of 250 μm has been integrated as the FBM in the KATRIN beam line. A photograph of the printed circuit board hosting the SDD is shown in figure 32. Located upstream of the main spectrometer, the FBM continuously monitors the flux of β-electrons. In addition, the excellent energy resolution of the SDD permits measurement of the spectral shape of the β-decay spectrum. In reference [149] it was demonstrated that the activity of the tritium source can be monitored on the <0.1% level on time scales of minutes. Moreover, measurements with the FBM can be used for initial searches for keV-scale sterile neutrinos with limited statistics and systematic uncertainties. These measurements can help to identify challenges and prepare for operations with the full TRISTAN detector.

A 47-channel TRISTAN prototype detector was successfully integrated into the KATRIN MoS in November 2020. The main objective was to gain a better understanding of the response
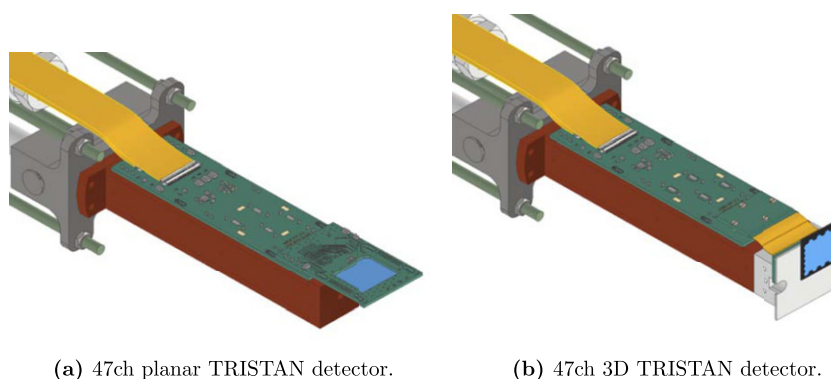
(a) 47ch planar TRISTAN detector.   (b) 47ch 3D TRISTAN detector.

Figure 33. CAD drawings of the 47-pixel TRISTAN detector. The left figure shows the planar detector design, whereas the right one shows the three-dimensional design. The green printed circuit board hosts the signal readout electronics and is mounted on a copper cooling block. (a) 47 ch planar TRISTAN detector. (b) 47 ch 3D TRISTAN detector.
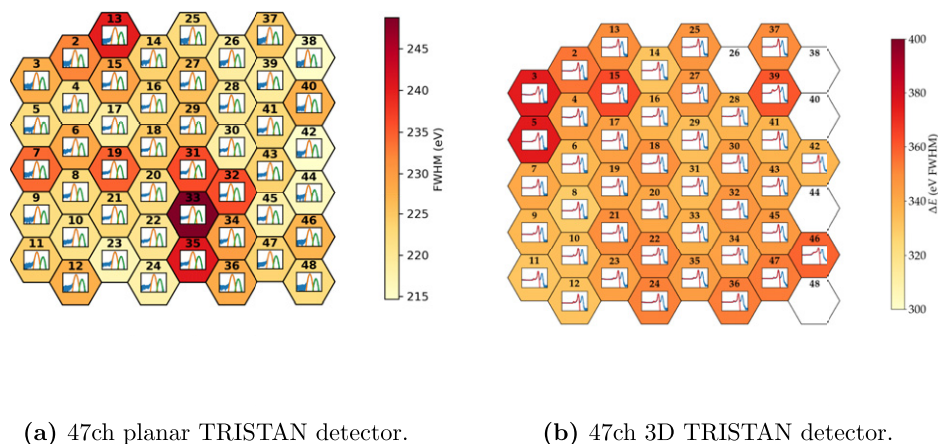
(a) 47ch planar TRISTAN detector.   (b) 47ch 3D TRISTAN detector.

Figure 34. Pixel maps of the 47-pixel TRISTAN detector (planar and 3D design) in the KATRIN MoS. The color bar indicates the energy resolution of the individual pixels. The plot insets show typical energy spectra (^{55}Fe calibration spectra for the planar detector and ^{83m}Kr spectra for the 3D detector). The spectra were fit with suitable functions. For the 3D design, five pixels did not work properly and were excluded from the analysis; they are marked here in white.

of the SDD in a realistic MAC-E filter environment, i.e. operation in a strong magnetic field (up to 0.4 T), at high temperatures (40 °C), and in vacuum (10^{-9} mbar). In a first step for initial tests, a planar detector was installed (figure 33(a)). It was demonstrated that all pixels are functional with an energy resolution of about $\lesssim$300 eV FWHM at the 5.9 keV line of ^{55}Fe, (figure 34(a)). In the next step, a three-dimensional version (SDD and readout electronics mounted orthogonally to each other) was installed in the MoS (figure 33(b)). Eventually, this device was used to detect electrons from the spectrometer walls, from an electron gun, and from a ^{83m}Kr source. Figure 35 shows the recorded energy spectra of the wall electrons with the spectrometer set to different retarding potentials. For 30 keV wall electrons, an energy
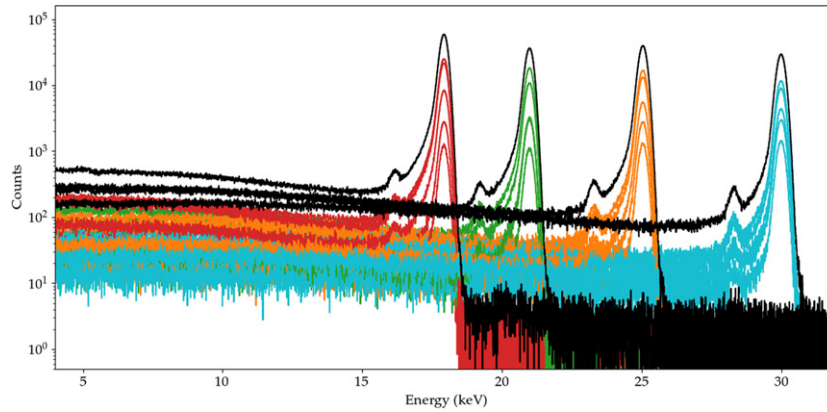
Figure 35. Energy spectra (wall electrons) measured with a 47-pixel TRISTAN detector in the KATRIN MoS at different high voltages. The black spectra represent stacked spectra of several pixels.

resolution of about $\lesssim$440 eV FWHM was obtained (figure 34(b)). In 2022, a TRISTAN 3D detector with 166 pixels will be deployed for the first time in the MoS.

8. Discussion and conclusions

First envisioned in 2001, with its conceptual design formalized in 2005 [12], the KATRIN experiment is the world-leading effort to measure the absolute mass scale of the neutrino. Via an exquisite measurement of the endpoint region of the tritium β-decay spectrum, KATRIN probes for the telltale shape distortion induced by the tiny neutrino rest mass. After the first two neutrino-mass data-taking campaigns, both in 2019, KATRIN has set an upper limit of $m_\nu < 0.8$ eV (90% C.L.) (section 4.3). The experiment continues to acquire data and is expected to run for at least two more years; the collaboration is continuing to develop increasingly sophisticated analysis methods (sections 3, 4.1 and 4.2). Several promising upgrades, including active background-mitigation measures (section 7.1) and improvements to the acceptance (section 7.2), are currently in development with the aim of reducing or compensating for the observed KATRIN background, and thereby further improving the neutrino-mass sensitivity toward the 0.2 eV sensitivity goal.

In addition to the effective neutrino mass, KATRIN's measurement of the tritium endpoint β spectrum provides sensitivity to a variety of hypothesized BSM phenomena, including sterile neutrinos at eV (section 5.1) and keV scales (section 5.2); GNI and exotic weak currents (section 5.3); new light bosons (section 5.4); Lorentz-invariance violation (section 5.5); local relic-neutrino overdensities (section 5.6); and other new physics ideas (section 5.7). In several cases, initial searches have already been performed with the first data sets. The KATRIN data sets also hold promise for spectroscopy of tritium molecular final states (section 6.1) and ^{83m}Kr conversion electrons (section 6.2), used for calibration.

Once KATRIN neutrino-mass running is complete, the next phase of the experiment will be a dedicated search for keV-scale sterile neutrinos (section 5.2). Optimizing KATRIN's sensitivity in this mass range requires running with a lower-luminosity source, a modified rear wall, and a new detector with improved energy resolution at high electron rates. The effort to realize such a detector is well underway (section 7.3.2). With these improvements, KATRIN will

achieve a world-leading laboratory sensitivity to the presence of a sterile neutrino with m_4 in the keV mass range, setting complementary limits to those from astrophysics measurements. The future of KATRIN is bright.

Acknowledgments

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

T Brunst ⬤ https://orcid.org/0000-0001-5810-4345
M Schlösser ⬤ https://orcid.org/0000-0002-3091-8088

References

[1] Aker M *et al* (KATRIN) 2022 *Nat. Phys.* **18** 160–6
[2] Fukuda Y *et al* (Super-Kamiokande) 1998 *Phys. Rev. Lett.* **81** 1562–7
[3] Ahmad Q R *et al* (SNO) 2002 *Phys. Rev. Lett.* **89** 011301
[4] Dolinski M J, Poon A W P and Rodejohann W 2019 *Annu. Rev. Nucl. Part. Sci.* **69** 219
[5] Abbott T M C *et al* 2022 *Phys. Rev.* D **105** 023520
[6] Formaggio J A, de Gouvêa A L C and Robertson R G H 2021 *Phys. Rep.* **914** 1
[7] Kraus C *et al* 2005 *Eur. Phys. J.* C **40** 447–68
[8] Aseev V N *et al* 2011 *Phys. Rev.* D **84** 112003
[9] Aker M *et al* (KATRIN) 2021 *J. Instrum.* **16** T08015
[10] Kuckert L, Heizmann F, Drexlin G, Glück F, Hötzel M, Kleesiek M, Sharipov F and Valerius K 2018 *Vacuum* **158** 195–205
[11] Babutzka M *et al* 2012 *New J. Phys.* **14** 103046
[12] Angrik J *et al* (KATRIN) 2005 *KATRIN Design Report Tech. Rep. 7090 FZKA Scientific Reports* http://bibliothek.fzk.de/zb/berichte/FZKA7090.pdf

[13] Priester F, Sturm M and Bornschein B 2015 *Vacuum* **116** 42–7
[14] Sturm M *et al* 2021 *Fusion Eng. Des.* **170** 112507
[15] Aker M *et al* (KATRIN) 2021 *Phys. Rev.* D **104** 012005
[16] Marsteller A *et al* 2021 *Vacuum* **184** 109979
[17] Bornschein B, Glugla M, Günther K, Le T L and Simon K H 2005 *Fusion Eng. Des.* **75–79** 645
[18] Goto R 1970 *Chem. Econ. Eng. Rev.* **2** 11
[19] Aker M *et al* (KATRIN) 2020 *Sensors* **2020** 4827
[20] Marsteller A 2020 Characterization and optimization of the KATRIN tritium source *PhD Thesis* Karlsruher Institut für Technologie (KIT) https://publikationen.bibliothek.kit.edu/1000127553
[21] Grohmann S, Bode T, Hötzel M, Schön H, Süßer M and Wahl T 2013 *Cryogenics* **55–56** 5–11
[22] Aker M *et al* (KATRIN) 2019 *Phys. Rev. Lett.* **123** 221802
[23] Arenz M *et al* 2016 *J. Instrum.* **11** P04011
[24] Rodenbeck C, Wüstling S, Enomoto S, Hartmann J, Rest O, Thümmler T and Weinheimer C 2022 *J. Instrum.* **17** P06003
[25] Glück F, Drexlin G, Leiber B, Mertens S, Osipowicz A, Reich J and Wandkowsky N 2013 *New J. Phys.* **15** 083025
[26] Lokhov A *et al* 2022 *Eur. Phys. J.* C **82** 258
[27] Vénos D, Slezák M, Dragoun O, Inoyatov A, Lebeda O, Pulec Z, Sentkerestiová J and Špalek A 2014 *J. Instrum.* **9** P12010
[28] Sentkerestiová J, Dragoun O, Lebeda O, Ryšavý M, Sturm M and Vénos D 2018 *J. Instrum.* **13** P04018
[29] Arenz M *et al* (KATRIN) 2018 *J. Instrum.* **13** P04020
[30] Altenmüller K *et al* (KATRIN) 2020 *J. Phys.* G **47** 065002
[31] Fraenkle F M (KATRIN) 2017 *J. Phys.: Conf. Ser.* **888** 012070
[32] Arenz M *et al* (KATRIN) 2018 *Eur. Phys. J.* C **78** 368
[33] Aker M *et al* (KATRIN) 2020 *Eur. Phys. J.* C **80** 264
[34] Amsbaugh J F *et al* 2015 *Nucl. Instrum. Methods Phys. Res.* A **778** 40
[35] Schwarz J S 2014 The detector system of the KATRIN experiment—implementation and first measurements with the spectrometer *PhD Thesis* Karlsruher Institut für Technologie (KIT) http://nbn-resolving.org/urn:nbn:de:swb:90-427724
[36] Altenmüller M *et al* (KATRIN) 2019 *Astropart. Phys.* **108** 40
[37] Altenmüller K *et al* (KATRIN) 2019 *Eur. Phys. J.* C **79** 807
[38] Friedel F R 2020 Ion and plasma systematics during the first KATRIN neutrino mass measurements *PhD Thesis* Karlsruher Institut für Technologie (KIT) https://publikationen.bibliothek.kit.edu/1000126837/94741900
[39] Vizcaya Hernández A P 2021 Toward a measurement of the neutrino mass with tritium: ion studies for the KATRIN and TRIMS experiments *PhD Thesis* Carnegie Mellon University https://kilthub.cmu.edu/articles/thesis/Toward_a_Measurement_of_the_Neutrino_Mass_with_Tritium_Ion_Studies_for_the_KATRIN_and_TRIMS_Experiments/18658334/1
[40] Aker M *et al* (KATRIN) 2020 *Eur. Phys. J.* C **80** 821
[41] Fränkle F M, Bornschein L, Drexlin G, Glück F, Görhardt S, Käfer W, Mertens S, Wandkowsky N and Wolf J 2011 *Astropart. Phys.* **35** 128–34
[42] Mertens S *et al* 2013 *Astropart. Phys.* **41** 52–62
[43] Drexlin G, Harms F, Jansen A, Krause M, Müller F, Schlösser K and Wolf J 2017 *Vacuum* **138** 165–72
[44] Görhardt S *et al* 2018 *J. Instrum.* **13** T10004
[45] Fränkle F M *et al* 2022 *Astropart. Phys.* **138** 102686
[46] Kleesiek M *et al* 2019 *Eur. Phys. J.* C **79** 204
[47] Saenz A, Jonsell S and Froelich P 2000 *Phys. Rev. Lett.* **84** 242–5
[48] Doss N, Tennyson J, Saenz A and Jonsell S 2006 *Phys. Rev.* C **73** 025502
[49] Aker M *et al* (KATRIN) 2021 *Eur. Phys. J.* C **81** 579
[50] Neyman J 1937 *Trans. R. Soc.* A **236** 333 http://jstor.2012.org/stable/91337
[51] Feldman G J and Cousins R D 1998 *Phys. Rev.* D **57** 3873
[52] Lokhov A V and Tkachov F V 2015 *Phys. Part. Nuclei* **46** 347–65
[53] Furse D *et al* 2017 *New J. Phys.* **19** 053012
[54] Karl C, Eller P and Mertens S 2022 *Eur. Phys. J.* C **82** 439
[55] Myers E G, Wagner A, Kracke H and Wesson B A 2015 *Phys. Rev. Lett.* **114** 013003
[56] Weinheimer C *et al* 1993 *Phys. Lett.* B **300** 210–6

[57] Belesev A I *et al* 1995 *Phys. Lett.* B **350** 263–72
[58] Weinheimer C, Degenddag B, Bleile A, Bonn J, Bornschein L, Kazachenko O, Kovalik A and Otten E W 1999 *Phys. Lett.* B **460** 219–26
 Weinheimer C, Degenddag B, Bleile A, Bonn J, Bornschein L, Kazachenko O, Kovalik A and Otten E W 1999 *Phys. Lett.* B **464** 352 (erratum)
[59] Lobashev V M *et al* 1999 *Nucl. Phys.* B **77** 327–32
[60] Abazajian K N *et al* 2012 Light sterile neutrinos: a white paper (arXiv:1204.5379 [hep-ph])
[61] Adhikari R *et al* 2017 *J. Cosmol. Astropart. Phys.* JACP01(2017)025
[62] Volkas R R 2002 *Prog. Part. Nucl. Phys.* **48** 161–74
[63] Mention G, Fechner M, Lasserre T, Mueller T A, Lhuillier D, Cribier M and Letourneau A 2011 *Phys. Rev.* D **83** 073006
[64] Hampel W *et al* (GALLEX) 1998 *Phys. Lett.* B **420** 114–26
[65] Abdurashitov J N *et al* (SAGE) 2009 *Phys. Rev.* C **80** 015807
[66] Barinov V V *et al* (BEST) 2021 Results from the Baksan experiment on sterile transitions (BEST) *Phys. Rev. Lett.* **128** 232501
[67] Andriamirado M *et al* (PROSPECT) 2021 *Phys. Rev.* D **103** 032001
[68] Danilov M (DANSS) 2020 *PoS EPS-HEP2019* p 401
[69] Adamson P *et al* (MINOS+) (Daya Bay) 2020 *Phys. Rev. Lett.* **125** 071801
[70] Abrahão T *et al* (Double Chooz) 2021 *Eur. Phys. J.* C **81** 775
[71] Almazán H *et al* (STEREO) 2020 *Phys. Rev.* D **102** 052002
[72] Aker M *et al* (KATRIN) 2021 *Phys. Rev. Lett.* **126** 091803
[73] Aker M *et al* (KATRIN) 2022 *Phys. Rev.* D **105** 072004
[74] Kraus C, Singer A, Valerius K and Weinheimer C 2013 *Eur. Phys. J.* C **73** 2323
[75] Belesev A I, Berlev A I, Geraskin E V, Golubev A A, Likhovid N A, Nozik A A, Pantuev V S, Parfenov V I and Skasyrskaya A K 2013 *JETP Lett.* **97** 67–9
[76] Serebrov A P *et al* (NEUTRINO-4) 2019 *Pisma Zh. Eksp. Teor. Fiz.* **109** 209–18
[77] Zyla P A *et al* (Particle Data Group) 2020 *Prog. Theor. Exp. Phys.* **2020** 083C01
[78] Gando A *et al* (KamLAND-Zen) 2016 *Phys. Rev. Lett.* **117** 082503
[79] Agostini M *et al* (GERDA) 2020 *Phys. Rev. Lett.* **125** 252502
[80] Dodelson S and Widrow L M 1994 *Phys. Rev. Lett.* **72** 17–20
[81] Canetti L, Drewes M and Shaposhnikov M 2013 *Phys. Rev. Lett.* **110** 061801
[82] Shi X and Fuller G M 1999 *Phys. Rev. Lett.* **82** 2832–5
[83] Fujikawa K and Shrock R E 1980 *Phys. Rev. Lett.* **45** 963–6
[84] Mertens S *et al* 2015 *J. Cosmol. Astropart. Phys.* JCAP02(2015)020
[85] Korzeczek M 2020 Sterile neutrino search with KATRIN—modeling and design-criteria of a novel detector system *PhD Thesis* Karlsruher Institut für Technologie (KIT) https://publikationen. bibliothek.kit.edu/1000120634
[86] Huber A 2020 Analysis of first KATRIN data and searches for keV-scale sterile neutrinos *PhD Thesis* Karlsruher Institut für Technologie (KIT) https://primo.bibliothek.kit.edu/permalink/f/ 4jne3t/KITSRCE1000128344
[87] Bezrukov F, Chudaykin A and Gorbunov D 2017 *J. Cosmol. Astropart. Phys.* JCAP06(2017)051
[88] Benso C, Brdar V, Lindner M and Rodejohann W 2019 *Phys. Rev.* D **100** 115035
[89] Hiddemann K-H, Daniel H and Schwentker O 1995 *J. Phys. G: Nucl. Part. Phys.* **21** 639–50
[90] Holzschuh E, Kündig W, Palermo L, Stüssi H and Wenk P 1999 *Phys. Lett.* B **451** 247–55
[91] Abdurashitov J N *et al* 2017 *Pisma Zh. Eksp. Teor. Fiz.* **105** 723–4
[92] Bischer I and Rodejohann W 2019 *Nucl. Phys.* B **947** 114746
[93] Bonn J, Eitel K, Glück F, Sevilla-Sanchez D, Titov N and Blaum K 2011 *Phys. Lett.* B **703** 310–2
[94] Stephenson G J Jr, Goldman T J and McKellar B H J 2000 *Phys. Rev.* D **62** 093013
[95] Severijns N, Beck M and Naviliat-Cuncic O 2006 *Rev. Mod. Phys.* **78** 991–1040
[96] Barry J, Heeck J and Rodejohann W 2014 *J. High Energy Phys.* JHEP07(2014)081
[97] Ludl P O and Rodejohann W 2016 *J. High Energy Phys.* JHEP06(2016)040
[98] Steinbrink N M N, Glück F, Heizmann F, Kleesiek M, Valerius K, Weinheimer C and Hannestad S 2017 *J. Cosmol. Astropart. Phys.* JCAP06(2017)015
[99] Arcadi G, Heeck J, Heizmann F, Mertens S, Queiroz F S, Rodejohann W, Slezák M and Valerius K 2019 *J. High Energy Phys.* JHEP01(2019)206
[100] Dvali G and Funcke L 2016 *Phys. Rev.* D **93** 113002
[101] Colladay D and Kostelecký V A 1997 *Phys. Rev.* D **55** 6760–74
[102] Colladay D and Kostelecký V A 1998 *Phys. Rev.* D **58** 116002

[103] Kostelecký V A 2004 *Phys. Rev.* D **69** 105009
[104] Kostelecký V A and Mewes M 2012 *Phys. Rev.* D **85** 096005
[105] Díaz J S, Kostelecký V A and Lehnert R 2013 *Phys. Rev.* D **88** 071902
[106] Lehnert R 2009 *Hyperfine Interact.* **193** 275–81
[107] Lehnert R 2022 *Phys. Lett.* B **828** 137017
[108] Aker M *et al* (KATRIN) 2022 Search for Lorentz violation with the first KATRIN data (arXiv 2207.06326)
[109] Ringwald A 2005 arXiv:hep-ph/0505024
[110] de Salas P F, Gariazzo S, Lesgourgues J and Pastor S 2017 *J. Cosmol. Astropart. Phys.* JCAP09(2017)034
[111] Aker M *et al* (KATRIN) 2022 New constraint on the local relic neutrino background overdensity with the first KATRIN data runs *Phys. Rev. Lett.* **129** 011806
[112] Weinberg S 1962 *Phys. Rev.* **128** 1457–73
[113] Faessler A, Hodák R, Kovalenko S and Šimkovic F 2017 *Int. J. Mod. Phys.* E **26** 1740008
[114] Robertson R G H, Bowles T J, Stephenson G J, Wark D L, Wilkerson J F and Knapp D A 1991 *Phys. Rev. Lett.* **67** 957–60
[115] McKeen D 2019 *Phys. Rev.* D **100** 015028
[116] Chacko Z, Du P and Geller M 2019 *Phys. Rev.* D **100** 015050
[117] Huang G y and Rodejohann W 2021 arXiv:2110.03718 [hep-ph]
[118] Bodine L I, Parno D S and Robertson R G H 2015 *Phys. Rev.* C **91** 035505
[119] Lin Y-T *et al* (TRIMS) 2020 *Phys. Rev. Lett.* **124** 222502
[120] Güsten R *et al* 2019 *Nature* **568** 357–9
[121] Jonsell S, Saenz A and Froelich P 1999 *Phys. Rev.* C **60** 034601
[122] Pedersen H B *et al* 2007 *Phys. Rev. Lett.* **98** 223202
[123] Pedersen H B, Lammich L, Domesle C, Jordon-Thaden B, Heber O, Ullrich J, Treusch R, Guerassimova N and Wolf A 2010 *Phys. Rev.* A **82** 023415
[124] Wustelt P, Oppermann F, Yue L, Möller M, Stöhlker T, Lein M, Gräfe S, Paulus G G and Sayler A M 2018 *Phys. Rev. Lett.* **121** 073203
[125] Seitz-Moskaliuk H 2019 Characterisation of the KATRIN tritium source and evaluation of systematic effects *PhD Thesis* Karlsruher Institut für Technologie (KIT) https://publikationen.bibliothek.kit.edu/1000090748
[126] Vénos D, Sentkerestiová J, Dragoun O, Slezák M, Ryšavý M and Špalek A 2018 *J. Instrum.* **13** T02012
[127] Rest O, Winzen D, Bauer S, Berendes R, Meisner J, Thümmler T, Wüstling S and Weinheimer C 2019 *Metrologia* **56** 045007
[128] Rodenbeck C 2022 A method for determining the transition energies of ^{83m}Kr at the KATRIN experiment *Eur. Phys. J.* C **82** 700
[129] Marsteller A 2022 Operation modes of the KATRIN experiment tritium loop system using ^{83m}Kr (in preparation)
[130] Trost N R M 2019 Modeling and measurement of Rydberg-state mediated background at the KATRIN main spectrometer *PhD Thesis* Karlsruher Institut für Technologie (KIT) https://publikationen.bibliothek.kit.edu/1000090450
[131] Gauda K *et al* 2022 An active transverse energy filter to differentiate low energy particles with large pitch angles in a strong magnetic field (arXiv:2203.06085 [physics.ins-det])
[132] Ladislas Wiza J 1979 *Nucl. Instrum. Methods* **162** 587–601
[133] Va'vra J and Sumiyoshi T 2004 *IEEE Symp. Conf. Record Nucl. Sci.* vol 2 p 1142
[134] Wolz T, Malbrunot C, Vieille-Grosjean M and Comparat D 2020 *Phys. Rev.* A **101** 043412
[135] Vieille-Grosjean M, Dimova E, Mazzotta Z, Comparat D, Wolz T and Malbrunot C 2021 *Eur. Phys. J.* D **75** 27
[136] Ziegler J F 2004 *Nucl. Instrum. Methods Phys. Res.* B **219–220** 1027–36
[137] Steinbrink N, Hannen V, Martin E L, Robertson R G H, Zacher M and Weinheimer C 2013 *New J. Phys.* **15** 113020
[138] Martin E 2017 Electron detection systems for KATRIN detector and spectrometer section *PhD Thesis* University of Washington https://digital.lib.washington.edu/researchworks/handle/1773/40285
[139] Lechner P *et al* 1996 *Nucl. Instrum. Methods Phys. Res.* A **377** 346
[140] Lechner P *et al* 2001 *Nucl. Instrum. Methods Phys. Res.* A **458** 281

[141] Trigilio P *et al* 2018 ETTORE: a 12-channel front-end ASIC for SDDs with integrated JFET *IEEE Nucl. Sci. Symp. Med. Imaging Conf. Proc. NSS/MIC (IEEE)* p 1
[142] King P *et al* 2021 *J. Instrum.* **16** T07007
[143] Gugiatti M *et al* 2022 *Nucl. Instrum. Methods Phys. Res.* A **1025** 166102
[144] Gugiatti M *et al* 2020 *Nucl. Instrum. Methods Phys. Res.* A **979** 164474
[145] Mertens S *et al* 2019 *J. Phys. G: Nucl. Part. Phys.* **46** 065203
[146] Mertens S *et al* 2020 *J. Phys. G: Nucl. Part. Phys.* **48** 015008
[147] Biassoni M *et al* 2021 *Eur. Phys. J. Plus* **136** 125
[148] Eggert T 2004 Die spektrale Antwort von Silizium-Röntgendetektoren *PhD Thesis* Technische Universität München http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss2004101114551
[149] Urban K 2018 Application of a TRISTAN silicon drift detector as forward beam monitor in KATRIN *Master Thesis* Technical University of Munich (TUM) https://publications.mppmu.mpg.de/2019/MPP-2019-356/FullText.pdf