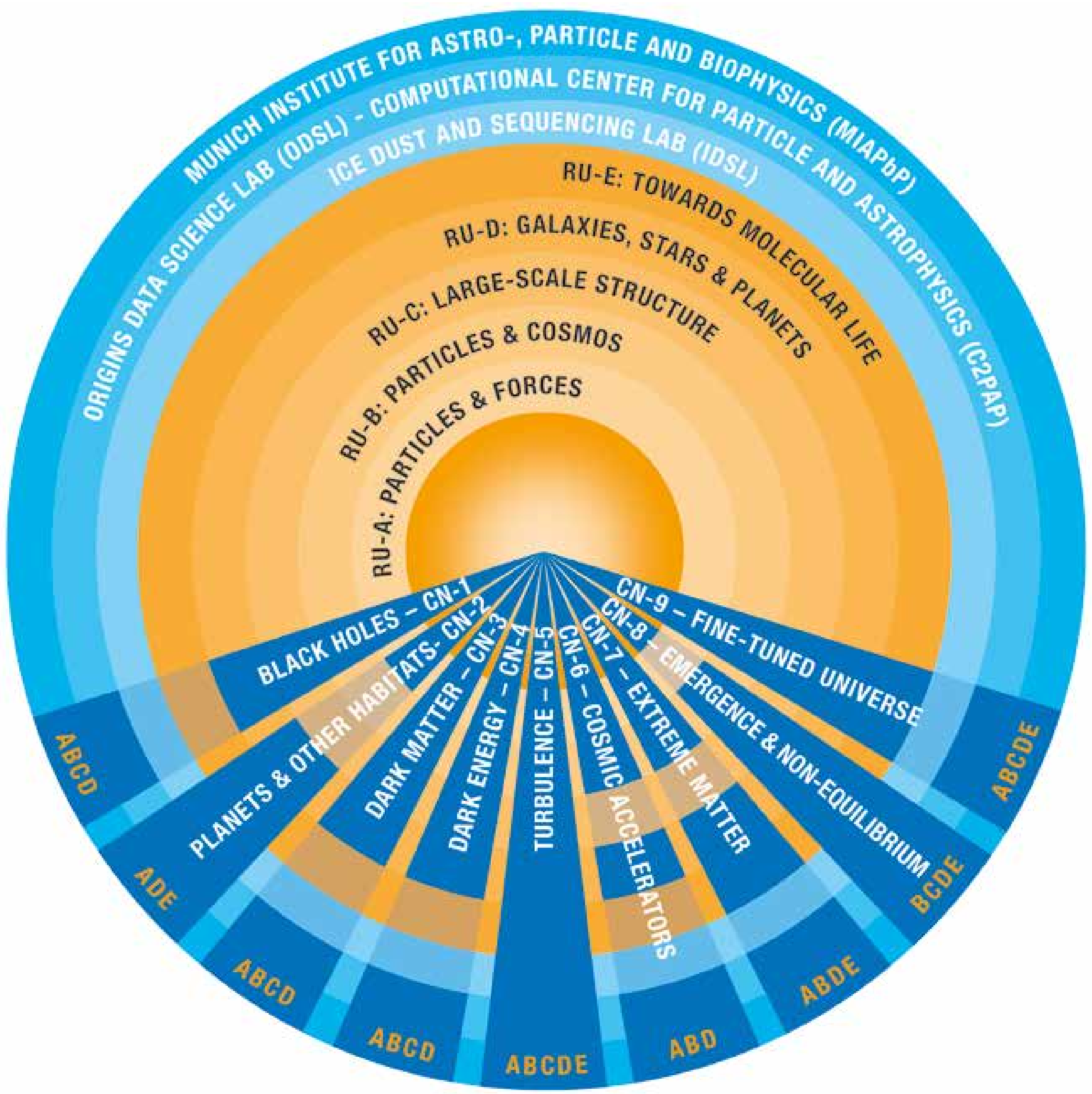
Volume 2

ORIGINS EXCELLENCE CLUSTER RESEARCH HIGHLIGHTS

From the origin of the Universe to the first building blocks of life

CONTENT



The interconnected Universe of the ORIGINS Excellence Cluster.
The ORIGINS infrastructures MIAPbP, ODSL, C2PAP and IDSL (blue rings) are the global structures that facilitate workshops, interaction and support corresponding projects in almost all of the different research units (RU-A, B, C, D, and E) and connectors (CN-1-9). © ORIGINS

"The question is," said Alice, "whether you can make words mean so many different things."

(L. CAROLL, THROUGH THE LOOKING-GLASS AND WHAT ALICE FOUND THERE, 1871)

FOREWORD

Language is a method of communication, a dynamic system of units and rules for combining them. Spoken language evolves, and words can take on different meanings depending on how they are used in different communities. This includes scientific communities. Take, for example, the term "building blocks." Are they quarks and leptons? Or are they lipids, proteins, nucleic acids and carbohydrates? Stardust?

The ORIGINS cluster of excellence brings together scientists from very different disciplines. Understanding each other is, therefore, a significant challenge. To meet this challenge, the cluster provides structures that enable and promote communication and translation as well as places and opportunities to meet and interact, discuss and visualize. Our print series "ORIGINS research highlights" serves as a communication tool to better understand fascinating science done in other areas where one is not an expert, disseminate significant developments, provide motivations and opportunities for successful and fruitful exchanges between disciplines, and encourage new collaborations.

This second collection of research highlights focuses on ORIGINS infrastructures. They are vital hubs in the daily life of the Cluster, where scientists from different fields come together and talk to each other. We present selected projects that have been initiated, developed, or completed thanks to the ORIGINS-funded infrastructure environment and thus are unique. These are the Munich Institute for Astro-, Particle- and Biophysics (MIAPbP), which hosts local and international guests and stimulates scientific exchange; the Computational Center for Particle and Astrophysics (C2PAP), where the next generation of numerical codes is developed, and most of the simulations and computationally intensive calculations take place; the ORIGINS Data Science Lab (ODSL), which specializes in the development of advanced techniques for pattern recognition, statistical analysis of complex, multi-connected data, and machine learning to deal with noisy and large data sets; the Dark Matter Data Centre (DMDC), which provides an open science repository for dark matter related experimental data, numerical simulations and associated software to improve accessibility within the dark matter community; the Ice, Dust and Sequencing Lab (IDSL), which provides the setup for biophysical experiments to search for the initial conditions of life; the Lab for Rapid Space Missions (LRSM), which has just launched a nanosatellite payload to the ISS; and last but not least, the Virtual Reality Lab (VRLab), where ORIGINS scientists can visualize complex multidimensional data to better understand the objects that they are studying. Highlights from our research units and connectors that did not appear in volume 1, from CN-1 "Black Holes" and RU-B "Particles and the Cosmos," are now included in this volume.

We have compiled this second set of research highlights with the kind help of Christian Alig (VRLab scientist), Heerak Banerjee (DMDC postdoc/ODSL fellow), Ludwig Böss (CN-5 manager), Dieter Braun (IDSL coordinator), Francesca Capel (ODLS fellow, former ODSL postdoc), Philipp Eller (ODSL fellow, former ODSL postdoc), Barbara Ercolano (CN-2 coordinator), Lorenz Gärtner (LMU Ph.D. student), Fabian Grübel (LMU Ph.D. student), Antony Hartin (C2PAP staff scientist), Lukas Heinrich (ODSL coordinator), Andreas Jäger (MIAPbP co-program manager), Florian Kaspar (ODSL fellow for RU-A), Jakob Knollmüller (ODSL postdoc), Rolf Kudritzki (MIAPbP director), Thomas Kuhr (RU-A coordinator), Martin Losekamm (LRSM coordinator), Christof Mast (RU-E scientist), Susanne Mertens (RU-B PI), Alexander Ruf (ODSL fellow for RU-E), Stefan Schönert (RU-B coordinator), Valentin Thoss (CN-1manager) and Michael Weber (CN-2 manager).

Thank you for taking the time to summarize the scientific highlights made possible by the ORIGINS infrastructures.

November 2023

Andreas Burkert and Stephan Paul
(Spokespersons)
Alice Smith-Gicklhorn
(Science Manager)
Odele Straub
(Science Communication)

Introduction to ORIGINS infrastructures

MIAPBP · MUNICH INSTITUTE FOR ASTRO-, PARTICLE AND BIOPHYSICS

Our MIAPbP "think-tank" workshop center will celebrate its 10th anniversary in 2024. Founded as part of the second funding period of the Excellence Cluster Universe, MIAPP smoothly transitioned to MIAPbP in ORIGINS, adding biophysics and data science to their program. The ORIGINS Cluster and TUM jointly fund MIAPbP. Its founding director, Prof. Rolf Kudritzki (LMU and Univ. of Hawaii) and co-director Prof. Andreas Weiler (TUM), together with a dedicated administrative staff comprised of the MIAPbP program manager and two full-time administrative coordinators, run the workshop center, hosting about 500 visitors annually. Over 700 publications acknowledge MIAPbP support. https://www.munich-iapbp.de/

C2PAP · COMPUTATIONAL CENTER FOR PARTICLE AND ASTROPHYSICS

Our C2PAP computational center has been a "work-horse" for high-performance computing for the Cluster scientists since 2013. Dedicated cluster hardware is hosted at Leibniz Supercomputing Center's (LRZ's) HPC system SuperMUC-NG and GPU and CPU servers in the LRZ compute cloud. PD Klaus Dolag (LMU) and Dr. Guenter Duckeck (LMU) run C2PAP together with a team of two scientific computing experts. Since the current C2PAP central cluster will be phased out in early 2024, Prof. Lukas Heinrich (TUM, ODSL coordinator) has joined forces to propose the successor system. The new C2PAP Cluster, with a manifold increase in CPU and GPU power, will have 500 TB of shared storage accessible on all LRZ systems - SuperMUC, Linux cluster and LRZ cloud.

ODSL · ORIGINS DATA SCIENCE LAB

As the name suggests, ODSL was founded with the ORIGINS Cluster as a new pillar for novel statistical data analysis and artificial intelligence in all ORIGINS physics and chemistry disciplines. Its founder, Allen Caldwell (MPP Director) and a team of postdocs and fellows conducted productive research and provided consulting and educational opportunities to ORIGINS members. ODSL fellows are researchers from any of the participating institutes, attracted through particular interest in statistical data analysis issues. They contribute through special lecture series and seminars and add domain knowledge. In recognition, ODSL fellows are entitled "Fellows of the ORIGINS Data Science Lab" with special access to ODSL resources. Since 2022, ODSL has been run by Lukas Heinrich, tenure-track Professor for Data Science at TUM, who, together with his rapidly growing team, is expanding opportunities for ORIGINS scientists and forging ties to Munich Data Science centers and C2PAP. With the funding of ORIGINS, ODSL supports our scientists with bleeding-edge hardware and the latest machine-learning techniques in modern physics.

DMDC · DARK MATTER DATA CENTER

The ORIGINS DMDC is integrated into ODSL. The experimentalists Prof. Stefan Schönert (TUM) and Dr. Federica Petricca (MPP) founded the center. The MPP postdoc Nahuel Ferreiro built up the center until TUM postdoc and ODSL fellow Heerak Banerjee took over. Since dark matter experiments are based on widely varying technologies and the data from these experiments are also very different in the way they need to be understood and analyzed, it isn't easy to combine and compare data. The DMDC supports open science by offering a repository for experimental data, models and code. The focus is on sharing knowledge in all its relevant forms: data, methodologies and software. Their goal is to facilitate data comparison, combination and interpretation using transparent and reproducible methods and easing the usability of this data so that, ultimately, the dark matter community can develop a consistent and unified view of the field in all its facets.

IDSL · ICE, DUST AND SEQUENCING Lab

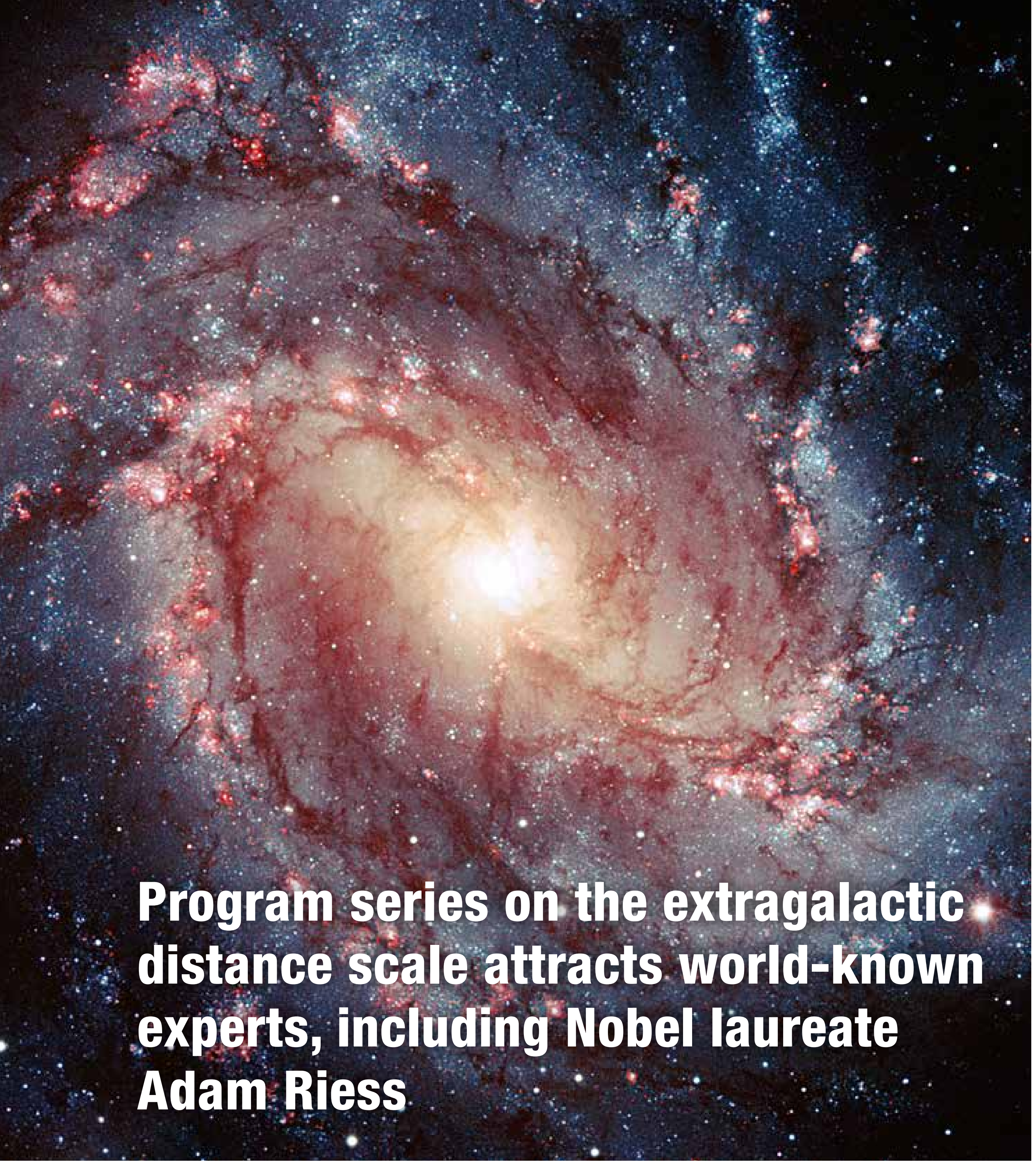
IDSL is a central infrastructure facility for all ORIGINS groups performing emergence of life experiments. The state-of-the-art instrumentation provides the best possible sensitivity to follow the first molecules in minute amounts and complex mixtures. IDSL provides methods, such as sequencing and various mass spectrometric techniques, which are not only very costly but require cumbersome maintenance, making the lab particularly attractive for young investigators of the Cluster. Measureables include the mass profile, binding thermodynamics, replication results, and RNA and DNA sequence in micro-fluid and vacuum low-temperature settings with optical access for irradiation with multiple sources and fluorescence detection. Recent additions include a MALDI (matrix-assisted laser desorption ionization) source for mass spectrometry to analyze rocky or icy surfaces and a Fourier-transform infrared spectrometer for molecular "fingerprints" complementing current space telescopes such as JWST. Conveniently located next to the LMU Physics and Biophysics building in Munich, IDSL is run by Prof. Dieter Braun (LMU) and his group along with a dedicated lab technician and aims to give broad support for all biophysical experiments in the search for the initial conditions of life.

LRSM · LABORATORY FOR RAPID SPACE MISSIONS

At the LRSM, Prof. Stephan Paul of the TUM Department of Physics and Prof. Ulrich Walter of the TUM Department of Aerospace and Geodesy have joined forces to use small satellites and other spacecraft as platforms for scientific investigations. With Prof. Susanne Mertens and Prof. Philipp Reiss (both at TUM), and under Martin Losekamm's (TUM) coordination, they are developing innovative detectors for astrophysics experiments to demonstrate the underlying technologies and collect scientific data. Besides early scientific successes, the LRSM has allowed PhD students to delve into (nearly) all aspects of a space mission, including launch and on-orbit operations.

VR Lab · VIRTUAL REALITY LAB

Prof. Andreas Burkert, University Observatory of the LMU and MPE, and Dr. Christian Alig (LRZ and LMU) founded the VR Lab to visualize astrophysical simulation data. The aim is to create a universal tool for ORIGINS scientists to visualize complex multi-dimensional data. In addition to this, virtual reality allows a cooperative working space and access to small details by flying through or viewing from any angle. Scientific, educational, and outreach projects benefit immensely from this facility. The VR Lab collaborates with the Leibniz Rechenzentrum (LRZ) Virtual Reality and Visualization Team (V2C) led by Dr. Thomas Odaker (LRZ) as well as C2PAP, which provides computational resources.



Program series on the extragalactic distance scale attracts world-known experts, including Nobel laureate Adam Riess

The galaxy Messier 83, captured by the Wide Field Imager at ESO's La Silla Observatory, located on the outskirts of the Chilean Atacama desert. The distance to the spiral galaxy – which is 4.81 Megaparsec or 15.68 million lightyears away from the Earth – has been determined recently by a small team of astronomers, including MIAPbP director Rolf Kudritzki. Extragalactic distance determination methods and the Hubble constant describing the expansion of the Universe were intensely discussed during a four-week MIAPbP program this summer. © ESO

SUMMARY

Today, progress in science does not rely on the ingenuity of individual researchers but on teamwork. The idea of MIAPbP, the Munich Institute for Astro-, Particle and BioPhysics, is at the heart of this: MIAPbP, which is an integral part of the ORIGINS cluster, is an international guest researcher center that hosts numerous several workshops a year. Opposed to standard conferences, MIAPbP programs typically last four weeks, and with only a few talks per day, they give participants – senior as well as young scientists – plenty of time to discuss their ideas. As one of this year's highlights, a program on the extragalactic distance scale attracted not only world-leading experts in cosmology, such as Nobel laureate Adam Riess, but also led to an improved determination of the Hubble constant.

THE STORY

Since its foundation in 2014, the Munich Institute for Astro-, Particle and BioPhysics (MIAPbP) has hosted many workshops. One of the recurring highlights are the workshops on the extragalactic distance scale – methods for measuring distances across the Universe, many of which, put simply, use the luminosity of certain types of stars. The first program of this kind was held in 2014. In fact, it was the very first ever MIAPbP program. A second one followed in 2018, which focused on the implications of the Gaia space telescope on the field. And now, in 2023, with the James Webb Space Telescope (JWST) providing new insights, there was a third program. In this regard, the "Extragalactic Distance Scale" workshops have become a program series. Not coincidentally, 2014, 2018, and 2023 were the years Nobel laureate Adam Riess traveled to Garching to participate in each of the "Extragalactic Distance Scale" programs. In fact, he was one of many leading experts in cosmology who came. What attracted them is both the structure and the infrastructure: MIAPbP programs differ from a standard conference in such a way that there are only a few talks per day. The focus is on discussions in small, mixing groups embedded in loose schedules. In terms of infrastructure, every guest gets his own office space, there is an auditorium for lectures and a kitchen for coffee breaks. Located in the middle of the IPP campus, close to TUM and LMU physics, MPIs, and ESO, the MIAPbP building can be easily reached by public transport. The proximity to Munich and the Alps is another plus, offering many activities after workdays, like hiking or going to a beer garden.

In an interview, we asked Adam Riess what makes him keep coming back. He said that there were very few places where you could get together with colleagues from different parts of the world and get to intensely focus on one subject. "Particularly since the pandemic, I think we all were reminded how important it is to interact with people directly rather than over Zoom, texting, and whatnot. People are built to understand each other in many forms of communication. Some of them are verbal, some of them are, you know, over a beer. It's hard to have a beer over Zoom." Apart from the social events, MIAPbP programs have always initiated scientific progress, which can be traced back to many papers that were published in the aftermath of workshops. However, some scientific questions are not agreed upon readily. On the contrary, a typical feature of MIAPbP programs is vivid, contentious discussions. This year's "Extragalactic Distance Scale" workshop is a good example: The participants debated on the Hubble constant, one of the most critical proportionality factors in cosmology, whose exact value some of the participants questioned. In the end, a group of astronomers, Adam Riess one of them, published a paper in The Astrophysical Journal Letters with an improved determination of the Hubble constant.

Many things in the Universe are still poorly understood, however. "We are like babies. We have to learn our way around the Universe by finding things that we can recognize," Riess said. By using novel telescopes like the JWST, which enables imaging of individual stars far away, we could get a perspective, a sense of proportion. Because according to Riess, "the things we know and understand the most are stars. They are the best tools we have."

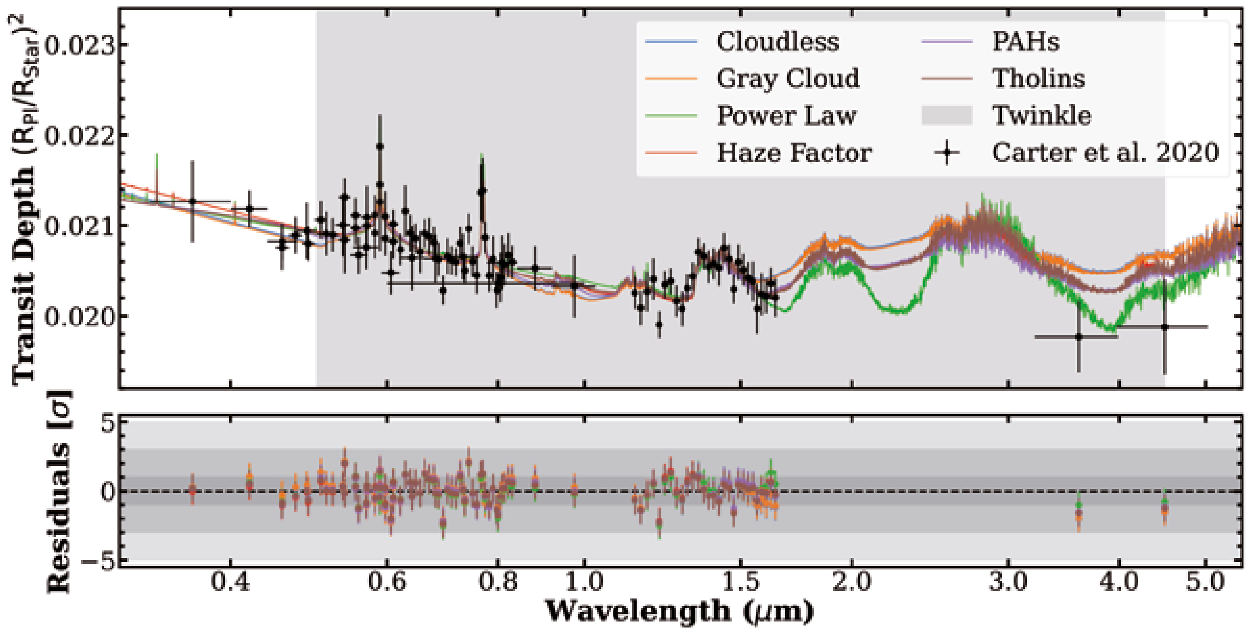
PUBLICATIONS

* **A. Bhardwaj, A.G. Riess, Catanzaro et al., "High-resolution Spectroscopic Metallicities of Milky Way Cepheid Standards and Their Impact on the Leavitt Law and the Hubble Constant," ApJL 955 L13 (2023)**
* **F. Bresolin, R.P. Kudritzki, M.A. Urbaneja, "The Metallicity and Distance of NGC 2403 from Blue Supergiants," ApJ 940 32 (2022)**

FUNDING INFORMATION

* **MIAPbP workshops, staff and guests are funded by the ORIGINS Excellence Cluster and the Technical University of Munich.**

Detectability of polycyclic aromatic hydrocarbons in the atmosphere of WASP-6b



Top: Transmission Spectrum of WASP-6b. The data (HST, Spitzer, VLT, TESS) including uncertainties are shown in black. Overlayed are six best-fit models that describe different atmospheric properties. The purple line represents the model with PAHs. The gray-shaded area marks the Twinkle wavelength coverage. Bottom: The residuals for each model spectrum. These are defined as the deviation of every data point from the best-fits in units of its uncertainty [σ]. © Grübel at al.

SUMMARY

Aiming to unlock the mysteries of distant worlds, researchers of the ORIGINS Excellence Cluster from the diverse fields of astrophysics, astrochemistry, and chemistry are investigating exoplanetary atmospheres to detect polycyclic aromatic hydrocarbons (PAHs) for their potential to shape these environments and even spark the beginnings of life. This recent case study focuses on the distant exoplanet WASP-6b and assesses the observability of PAH molecules using the powerful processing resources of the Computational Center for Particle and Astrophysics (C2PAP).

THE STORY

Even though probably every human has encountered PAHs at least once in their lives, as they are a common byproduct of the combustion of organic compounds and thus ubiquitous in the modern world, this molecular group remains relatively unknown to the public. However, PAHs are also found all over the Universe. For example, they appear in significant numbers in the hazy atmosphere of Saturn's moon Titan, play a crucial role in heating the interstellar medium, and influence the evolution of protoplanetary discs and exoplanetary atmospheres. They are even considered essential for the emergence of life as they represent a vital step toward creating amino acids and nucleotides.

Despite their significance, PAHs remain undetected on exoplanets until now. Yet, in a previous study, researchers from the ORIGINS Excellence Cluster determined the necessary environmental properties to maximize the production of PAHs in planetary atmospheres. They found that temperatures around 1000 degrees Celsius highly enhance the corresponding reaction rates. Since such extreme conditions occur only on so-called hot Jupiters, planets similar to Jupiter in size but around 100 times closer to the stellar host, the scientists used a grid search to find the most promising targets. This way, they came across the exoplanet WASP-6b.

Conveniently, the planet has already been observed several times by previous ground and space-based facilities such as the Hubble and Spitzer telescopes during the specific orbital phase where the object passes in front of its host star, commonly named transit. This technique allows to probe the atmosphere and learn about the molecular contents and properties of the planet by creating a transmission spectrum. The observations revealed a possibly ideal environment for the presence and first detection of PAHs.

In the subsequent analysis, Ph.D. student Fabian Grübel (LMU) in collaboration with fellow cluster scientists Barbara Ercolano (LMU), Karan Molaverdikhani (LMU), Dwaipayan Dubey (LMU), Rosa Arenales-Lope (LMU), Christian Rab (MPE), and Oliver Trapp (LMU) studied the available data in search of an unambiguous spectral signal of these molecules using the computational resources of the ORIGINS Excellence Cluster. Fabian Grübel applied the commonly employed method of atmospheric retrievals. In these algorithms, a large set of model spectra is sampled from a prior parameter space and compared with the observations in a Bayesian statistics framework to eventually reveal the most likely properties of the atmosphere. The results indicate that PAHs might be present on WASP-6b, but the data quality did not yet allow for a clear signal. The team thus additionally assessed if this situation would change using more advanced instruments with higher resolution and sensitivity such as the James Webb space telescope (JWST). They concluded that an observation with JWST can reveal currently unresolved molecules and thus enable an unambiguous detection.

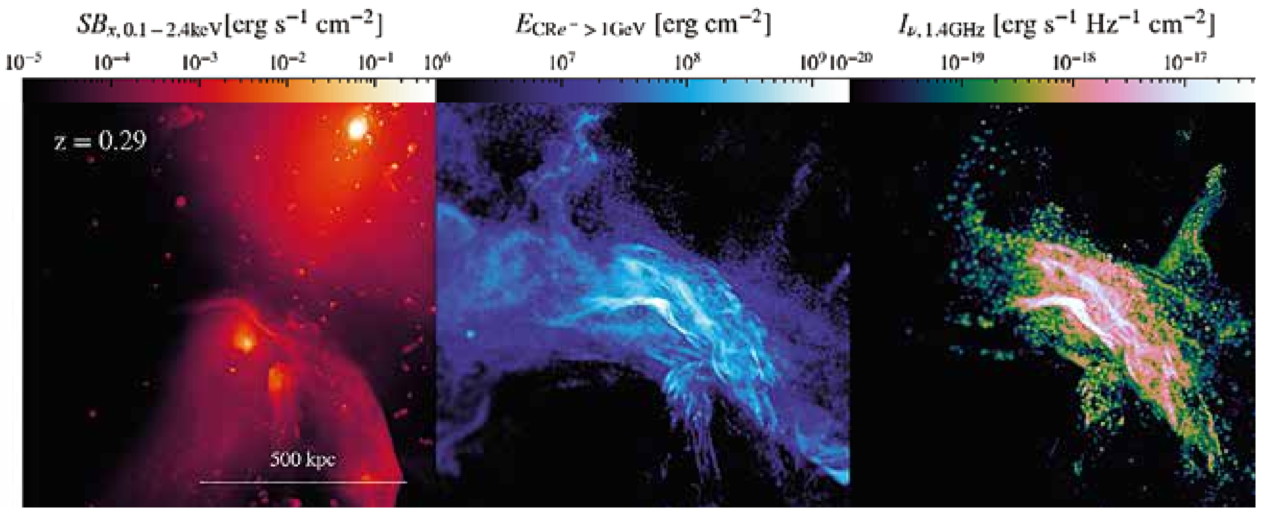
Furthermore, the Excellence Cluster ORIGINS is among the few founding members of the upcoming Twinkle private satellite mission to be launched in 2024. With exclusive access to this new powerful telescope, researchers can observe countless exoplanetary atmospheres and will be able to characterize the role of PAHs in these environments.

PUBLICATIONS

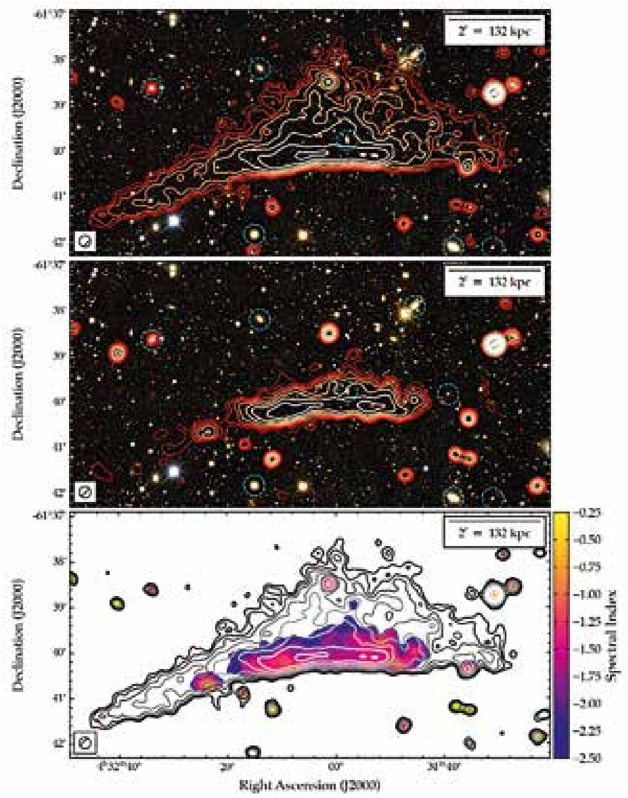
* F. Grübel et al., in preparation
* D. Dubey et al., "Polycyclic aromatic hydrocarbons in exoplanet atmospheres," A&A, 678, A53 (2023)

FUNDING INFORMATION

* Ph.D. student Dwaipayan Dubey is funded by ORIGINS
* Ph.D. student Rosa Arenales-Lope is funded by ORIGINS
* ORIGINS is a founding member of the Twinkle satellite mission



Simulated "wrong way relic" as it transitions through various morphologies found in observations. The three panels show, from left to right, the X-ray surface brightness, the cosmic ray electron energy, and the synchrotron surface brightness of the in-falling group in the xz-plane of the simulation. © Böss et al. 2023



Observed "wrong way relic" to the SE of galaxy cluster Abell 3266, shown as a radio-on-optical overlay. Contours indicate the surface brightness. © Riseley et al. 2022

A formation mechanism for "wrong way" radio relics

SUMMARY

Radio relics are typically found to be arc-like regions of synchrotron emission in the outskirts of merging galaxy clusters, bowing out from the cluster center. Many radio relics break with this ideal picture and show morphologies bent the opposite way. ORIGINS scientists recognized these "wrong way relics" in simulations of the intracluster medium. Using the C2PAP computational resources to analyze these, they found that "wrong way relics" can form when an outwards traveling shock wave is bent inwards by an in-falling galaxy cluster or group.

THE STORY

Synchrotron emission from relativistic electrons provides one of the most powerful tools to gain insight into the dynamics and magnetic field of the intracluster medium (ICM). So-called "radio relics" are some of the most prominent examples of synchrotron emission in the ICM. These radio relics are giant radio arcs with roughly one Mpc in diameter, typically found in the outskirts of colliding galaxy clusters. These collisions between galaxy clusters drive Mpc-sized shock waves through the ICM and dissipate more energy than any other single process since the Big Bang.

Radio relics are observed to have strongly polarized synchrotron emission and show a synchrotron spectrum that follows a powerlaw in frequency and steepens towards the cluster center. This powerlaw indicates that the source of the synchrotron emission is a population of cosmic ray (CR) electrons that have been accelerated at a shock in the ICM in a process called "Diffusive Shock Acceleration." On the other hand, the steepening can be explained as the electrons starting to lose their energy as they are swept away from the front of the shock and interact with the magnetic and photon fields in the ICM.

With the increasing capabilities of radio telescopes, more and more radio relics are found that break with this ideal picture. These relics show complex morphologies such as S-shapes, flat relics with varying thickness and filamentary structures. A newly discovered kind of relic shows even more puzzling features, by being bent "the wrong way." These relics do not bow out from the cluster center but instead appear to be bent inwards and show complex synchrotron spectral indices, indicating that more than one shock is responsible for accelerating the synchrotron bright electrons.

In a recent publication ORIGINS scientists Ludwig Böss (LMU) and Klaus Dolag (LMU) as well as Ulrich Steinwandel (Flatiron Institute) proposed that these "wrong way relics" can be formed through the collision of shock waves with in-falling substructure. They performed the first cosmological magnetohydrodynamic simulation to study these systems, including a spectral CR model. Using a spectral CR model allowed them to trace the time evolution of CR electrons accelerated at shocks and study their synchrotron emission as they move through the ICM and lose their energy due to synchrotron emission and inverse Compton scattering. This allowed them to explore the complex interaction of shocks with the complex and turbulent ICM at unprecedented fidelity.

While these simulations were performed at a supercomputer in New York, most of the analysis was performed at the ORIGINS-owned computing infrastructure C2PAP. After the relevant part of the simulation volume was found, the data was transferred to C2PAP to make use of its flexible and high-performance hardware.

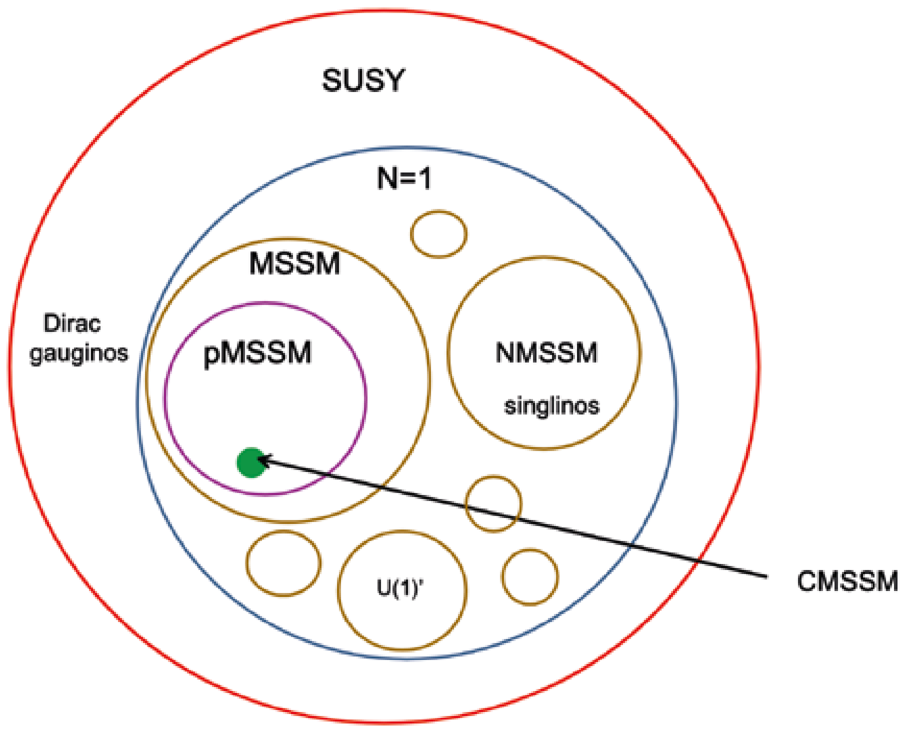
Having this level of high-performance computing readily available "in-house" allows exploring the data interactively. Once the analysis pipeline is complete, the resources available at C2PAP make it possible to analyze the terabytes of data produced in this simulation and scale the analysis up to various performance requirements.

PUBLICATIONS

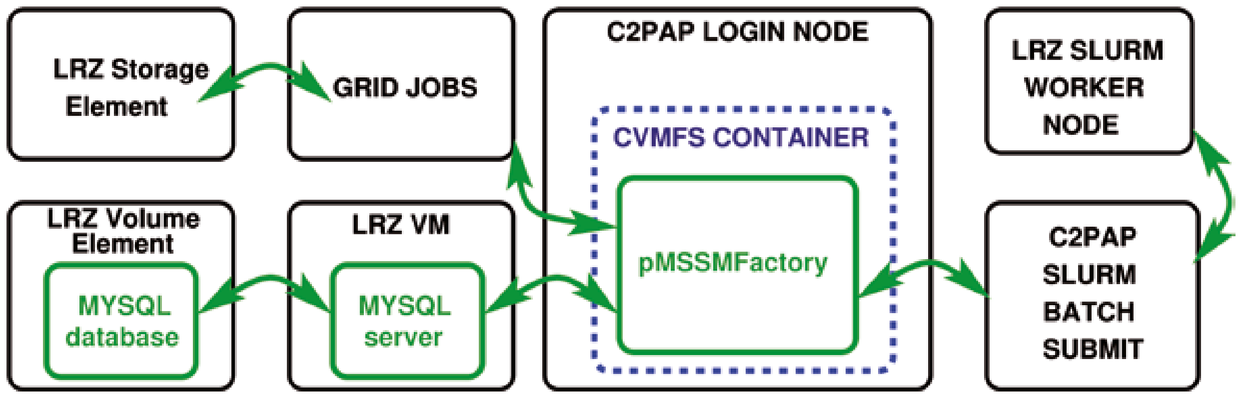
* L. Böss, U. Steinwandel, K. Dolag, "A formation mechanism for "Wrong Way" Radio Relics," arXiv:230900046 accepted for publication in ApJL, 957, 2, L16 (2023)

FUNDING INFORMATION

* Ph.D. student Ludwig Böss is funded by ORIGINS
* C2PAP is an ORIGINS infrastructure



Supersymmetry models. © A. Hartin



pMSSMFactory@C2PAP block setup. © A. Hartin

Efficient supersymmetry searches

SUMMARY

Supersymmetry (SUSY) is a theory of physics beyond the Standard Model, which posits a plethora of new fundamental particles. SUSY is an attractive theory because it provides a consistent way to explain physics that must still lie beyond the Standard Model. For example, SUSY provides a unifying formalism that enables high-energy scale unification of the electromagnetic, weak and strong forces. SUSY also provides candidate particles for dark matter. Searches for SUSY particles are ongoing at the Large Hadron Collider (LHC) at CERN. The C2PAP staff and ORIGINS scientists strongly contribute to the software development and optimization to efficiently perform SUSY searches.

THE STORY

Minimal supersymmetry (MSSM) is a subset model of supersymmetry (SUSY), which requires 105 additional free parameters to the standard model parameters. It is difficult to analyze such a large parameter dimensionality experimentally, so well-motivated theoretical conditions are applied, leading to SUSY subsets with fewer free parameters. The constrained MSSM (CMSSM) has only four additional parameters, which makes detailed analysis possible but is considered overly restrictive. The phenomenological MSSM (pMSSM) is a larger subset of the MSSM with 20 additional, low-energy scale, free parameters (upper figure).

pMSSM searches are enabled at the LHC by running large-scale analyses over the possible parameter space many times. For a parameter dimensionality of 20, this still requires enormous computing effort. However, using sufficient computing power, pMSSM searches can be designed that are both efficient and technically achievable. A dedicated and automated computing stream for pMSSM searches helps speed up the process. Such a stream must bundle together a number of software packages used in Monte Carlo production, detector simulation, fitting, and analysis (lower figure). This software stream must also interface with grid computing resources, global data transfer protocols, database servers and batch computing resources. And pMSSMFactory is just such a software stream that performs the required work. However it is constrained to the CERN computing environment.

This is where C2PAP enters the game: C2PAP staff scientist Anthony Hartin (LMU) has reconfigured the software to run in house at the C2PAP/LRZ computing environment. The pMSSMFactory port runs in a CERN virtual machine file system container served by a singularity environment on the C2PAP login node. Relevant CERN software packages run via the GRID or the local slurm batch system. A MySQL server running on an LRZ cloud virtual machine coordinates the pMSSMFactory. There is a software bridge from the CVMFS container to the local slurm batch system using the singularity bind function.

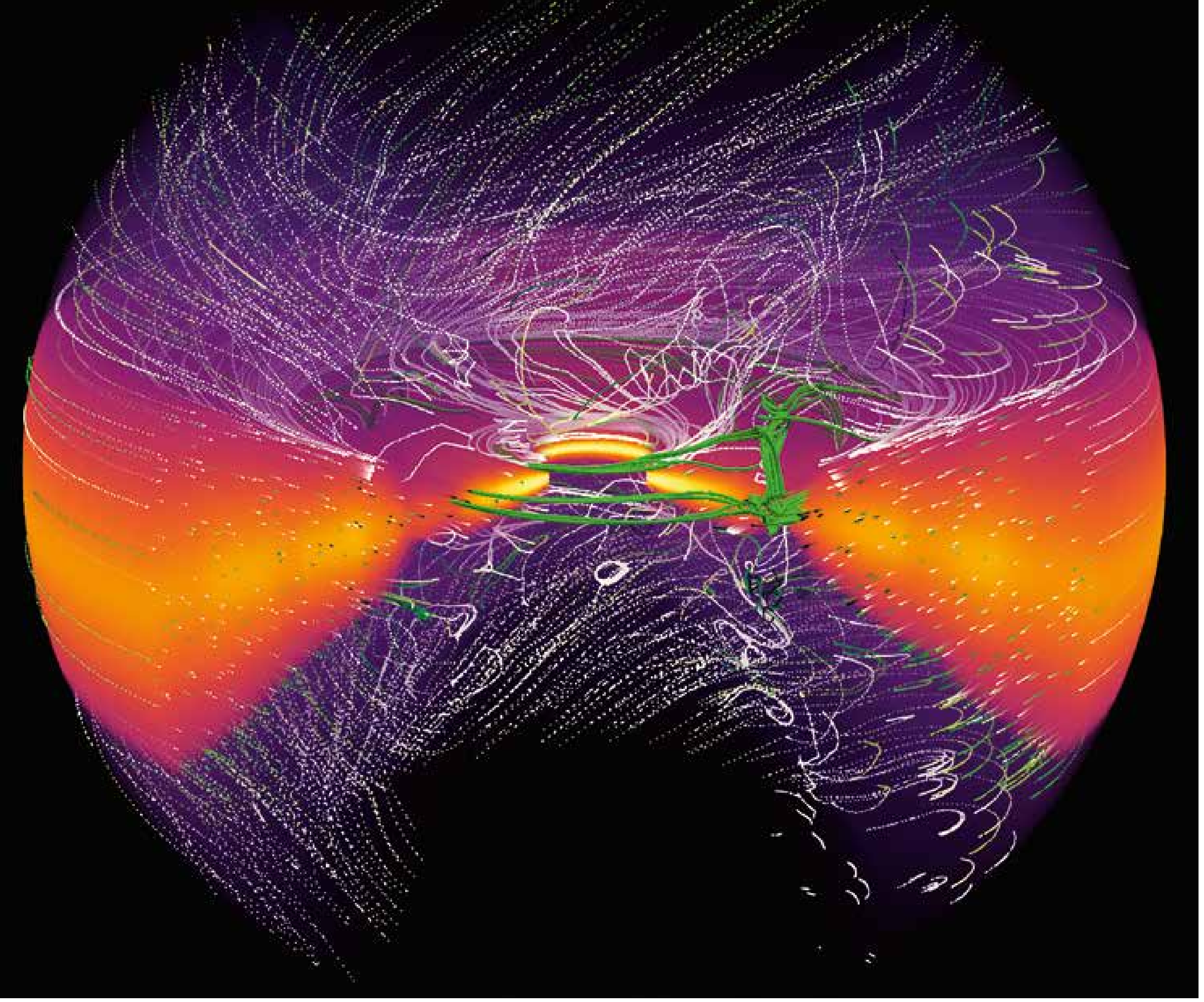
Ph.D. student Eric Schanet (LMU) and master student Markus Eck (LMU), supervised by ORIGINS scientist Jeanette Lorenz (LMU, now Frauenhofer), heavily used pMSSMFactory to perform extensive analysis. By enabling pMSSMFactory to run at C2PAP, ORIGINS contributes to the broader effort to analyze supersymmetry searches at the LHC and constrain the remaining parameter space in which supersymmetric particles may be found. If such particles are found, it opens the way to a more comprehensive understanding of the as-yet unresolved physics beyond the Standard Model.

PUBLICATIONS

* ATLAS SUSY Working Group, pMSSMFactory @ CERN, https://gitlab.cern.ch:8443/atlas-phys-susy-wg/summaries/pmssm/pMSSMFactory
* A. Hartin, pMMSSMFactory @ C2PAP, https://gitlab.cern.ch/hartin/pMSSMFactory_c2pap
* E. Schanet, "Search for charginos and neutralinos in a signature with a Higgs boson and an isolated lepton with the ATLAS detector and its reinterpretation in the pMSSM," Dissertation (2021)

FUNDING INFORMATION

* Anthony Hartin holds an ORIGINS funded C2PAP staff position



3D visualization of the gas density in a simulation with a Jupiter-like planet. The green tubes show different possible accretion pathways gas can take before the planet accretes it. The white dotted lines show gas streamlines in the disk and wind. © Weber et al.

Simulating the interplay between protoplanetary disk winds and accreting planets

PUBLICATIONS

- L. Wölfer et al., “A highly non-Keplerian protoplanetary disc,” A&A 648, A19 (2021)
- M. Weber, B. Ercolano, G. Picogna, C. Rab, “The interplay between forming planets and photoevaporating discs I: forbidden line diagnostics,” MNRAS 517, 3, 3598-3612 (2022)
- Weber et al. (2023), submitted

FUNDING INFORMATION

- ORIGINS funded Ph.D. student and CN-2 “Planets and other Habitats” manager: Michael Weber
- C2PAP is an ORIGINS infrastructure

SUMMARY

Protoplanetary disks are the birthplaces and nurseries of planets and hold the key to understanding their formation and evolution. Disk winds, planet formation, and their mutual interactions are crucial mechanisms that drive the morphological transformation of protoplanetary disks and define the environment in which planets form and evolve. Making heavy use of the ORIGINS infrastructure C2PAP, ORIGINS scientists have been able to carry out extensive three-dimensional simulations of planet-hosting protoplanetary disks with disk winds. They discovered a strong interaction between disk winds and the substructures generated by planets with significant consequences for many aspects of planet formation.

THE STORY

In recent years, thousands of exoplanets and hundreds of systems with multiple exoplanets, showcasing a remarkable variety, have been discovered by various observational techniques. Their diversity highlights the importance of understanding the many different processes that can be at play during the formation and evolution of planets in protoplanetary disks. Two of the most critical processes that affect the structure and evolution of protoplanetary disks and the fate of their newly formed planets are planet-disk interactions and disk winds.

Through the gravitational influence on its surrounding gas, a planet still embedded in its protoplanetary disk can generate complex, asymmetric substructures, such as spiral arms or even gaps if the planet is massive enough. Astronomers have recently observed many such substructures, which affect the structure and dynamics of gas and dust in the disk, as discussed in the 2021 publication involving ORIGINS collaborators Barbara Ercolano, Anna Miotello, Leonardo Testi and Evine van Dishoeck. In turn, the disk can affect its planets, for example, by exerting a gravitational torque that leads to the migration of a planet to a different orbital separation. Modeling these planet-disk interactions requires resolving the asymmetric substructures in the plane perpendicular to the axis of rotation.

Disk winds are responsible for the quick dispersal of protoplanetary disks within a few million years and, therefore, for setting an upper limit for the time within which planets can form or stay embedded in the disk. They are launched either through magnetic effects or thermally when energetic irradiation, for example, from the central star, heats the disk surface sufficiently for the resulting pressure gradient to drive a wind – a process called photoevaporation. Modeling thermal disk winds requires accurately resolving the temperature structure in the meridional plane.

Using the ORIGINS infrastructure C2PAP, ORIGINS-funded Ph.D. student Michael Weber (LMU) in the group of Barbara Ercolano (LMU) was able to carry out computationally demanding three-dimensional radiation-hydrodynamic simulations of planet-hosting disks undergoing X-ray photoevaporation, capturing both planet-disk interactions and disk winds in the model. The team discovered strong interactions between the disk substructures generated by a massive planet and the disk wind. Spiral structures leave an imprint in the wind that could be observed in forbidden line diagnostics, helping to detect embedded planets. Planetary gaps in the disk take away the wind's pressure support, allowing wind material to fall back into the gap. Consequently, the amount of gas a newly formed planet has available for accretion increases significantly, whereas the recycling of wind material slows the dispersal of the inner disk. The redistribution of disk material towards the planetary gap by the wind thus opens new pathways for accretion onto a massive planet. This is visualized in the figure, which shows the gas density in the model and different possible accretion pathways highlighted by green tubes.

Moreover, it can modify the atmospheric compositions of massive planets or help explain the dichotomy in the isotopic ratios between non-carbonaceous and carbonaceous meteoritic records found in our solar system.

Constraints on the up-quark valence distribution in the proton

SUMMARY

The valence quarks that define the flavor of hadrons are embedded in a sea of virtual quarks and gluons coupling strongly to the valence quarks. This structure is described by the parton distribution functions (PDFs). These PDFs are of fundamental importance to our understanding of QCD and can be extracted from measurements at particle accelerators when hadrons are probed through collisions with electrons or positrons. Within a collaboration of ORIGINS scientists led by Allen Caldwell, ODSL's coordinator from 2019-2022, the distribution function for up-type quarks in the proton could be constrained with new precision.

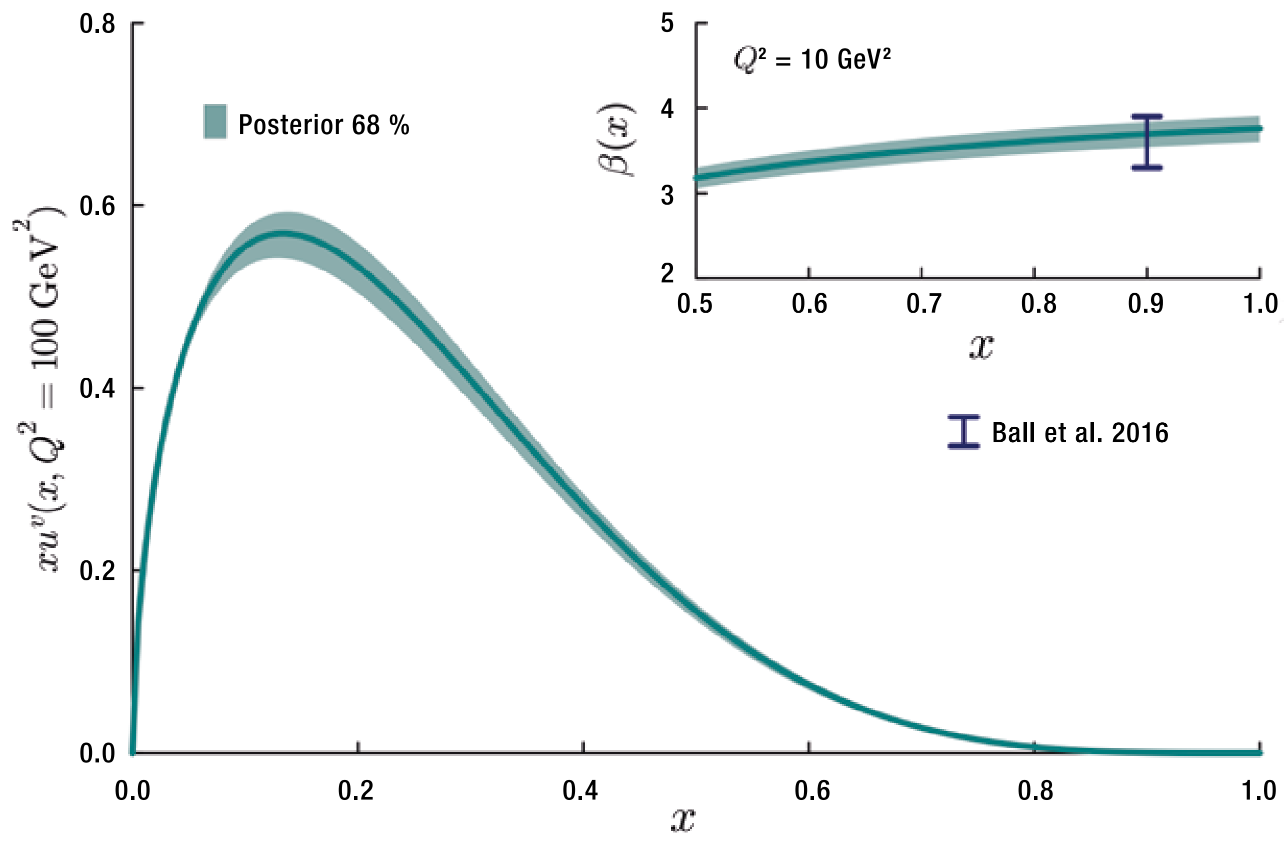
THE STORY

Building on the broad expertise on both hadron structure (PDF) and novel statistical analysis techniques (ODSL) present in the ORIGINS cluster, Allen Caldwell (MPP) and Francesca Capel (MPP, formerly ODSL) developed a new analysis framework in collaboration with Oliver Schulz (ODSL fellow). The project profited from Francesca's broad science background as she applied similar analysis methods to open questions in high-energy astrophysics. This project thus demonstrates the underlying connections between astro- and particle physics in terms of shared data analysis challenges that can be forged through the ORIGINS/ODSL environment, enabling exciting new results.

Within this project, the researchers developed a new approach to extract PDFs using a full forward model starting with input model-PDFs predicting the number of events to be observed in various kinematic bins of the actual measurement. Using Bayesian methods, this model is then fit to infer the parameters of the PDFs from the data. This approach is particularly interesting for the kinematic domain where the valence quarks studied carry a large fraction of the proton momentum (a variable called x_{Bj}). The so-called "high-x_{Bj}" data from the ZEUS experiment carried out at DESY more than a decade ago, suffering from the low event numbers, is particularly suited for this ansatz as assumptions used in existing analysis methods are no longer valid. Precise knowledge of the PDFs at high values of x_{Bj}, which corresponds to high energies in the center of mass of the struck quark and the probing electron, might offer insights into physics beyond the Standard Model.

The extraction of the PDF uses the QCDNUM software package for fast QCD evolution and convolution in combination with Bayesian inference via the BAT (Bayesian Analysis Tool developed at MPP by A. Caldwell and co-workers) and numerous additional technical developments. These tools were joined into the open-source Julia framework called PartonDensity.jl (https://github.com/cescalara/PartonDensity.jl).

Applying their method to data from the ZEUS experiment, they extracted the PDFs at very high values of x_{Bj} and Q^2 (the negative square of 4-momentum-transfer in the electron scattering process) regime for the first time. The published results provide unprecedented constraints on the valence up-quark distribution (see Figure). A further publication is currently in preparation detailing the robustness of the statistical approach and the impact of different PDF parametrizations.



The valence up-quark distribution and the effective (1-x) power, β, from this analysis at a value of Q^2 = 100 GeV². β summarizes the asymptotic behavior of the parton distribution at large $x = x_{Bj}$ values and is useful to enable comparison with theoretical predictions. The 68% smallest probability contour is depicted. The vertical range plotted in the inset summarizes relevant previous results. © Aggarwal et al. (2023)

PUBLICATIONS

* O. Schulz et al., "BAT.jl: A Julia-Based Tool for Bayesian Inference," SN Computer Science, 2(3):210 (2021)
* R. Aggarwal, M. Botje, A. Caldwell, F. Capel, and O. Schulz "Constraints on the up-quark valence distribution in the proton," Phys. Rev. Lett., 130:141901 (2023)

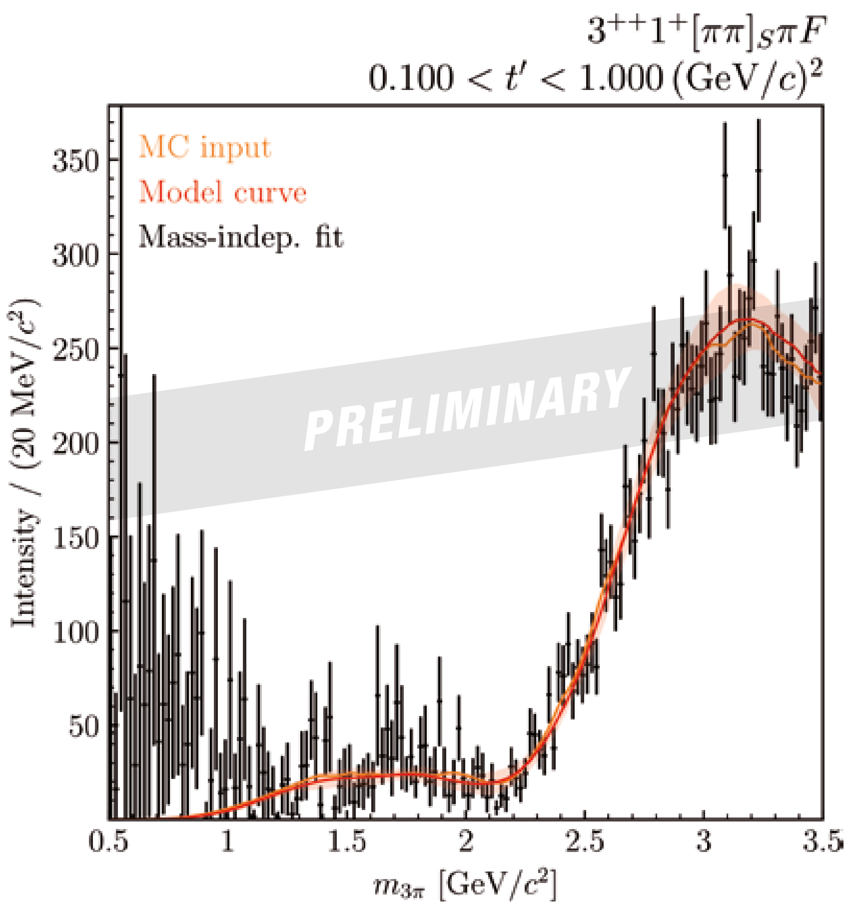
FUNDING INFORMATION

* Allen Caldwell, MPP director and ODSL coordinator from 2019-2022
* Francesca Capel, Minerva fellow at MPP, ODSL postdoc from 2020-2022
* Oliver Schulz is an ODSL fellow

Information field theory for particle physics

SUMMARY

Inverse problems are a challenge familiar to many areas of physics. Examples include determining the three-dimensional distribution of interstellar dust in the Milky Way from stellar spectra or disentangling the quantum mechanical subsystems produced in a well-defined particle physics reaction. This ORIGINS collaboration between astrophysicists and particle physicists has successfully applied information field theory, developed for astrophysical problems, to amplitude analysis in hadron physics to determine the excitation spectrum of light mesons.



Example from a simulation study: The black dots represent the results from a traditional analysis (black) performed independently in each mass bin (horizontal axis). The results of the new continuous method are shown by the red line together with the known input (orange line). The uncertainties depicted by error bars for individual fits have shrunk to the shaded uncertainty band around the red solid curve. © F. Kaspar

PUBLICATIONS

- **T. Enßlin, "Information theory for Fields," Annalen der Physik special issue on "Physics of Information," 531, 3, (2019)**
- **R. H. Leike, et al. "The Galactic 3D large-scale dust distribution via Gaussian process regression on spherical coordinates," arXiv:2204.11715 (2022)**
- **J. Beckers, F. Kaspar, J. Knollmüller, "Progress in the Partial-Wave Analysis Methods at COMPASS," Proceedings for the Meson 2023 and HADRON23 Conferences**

FUNDING INFORMATION

- **F. Kaspar, ORIGINS funded Ph.D. student at TUM**
- **J. Knollmüller, ORIGINS funded ODSL postdoc**
- **P. Frank, funded through the German Federal Ministry of Education and Research for the project ErUM-IFT: Informationsfeldtheorie für Experimente an Großforschungsanlagen (Förderkennzeichen: 05D23E01).**

THE STORY

The excitation spectrum of light hadrons is typically studied using scattering reactions with well-defined kinematics or heavy particle decays. In such reactions, all intermediate quantum states that can be produced must be considered. In a quantum mechanical system made from e.g. three particles, all intermediate states coherently overlap (sum of their amplitudes) when producing the same final state. While the mass of the intermediate state can be measured externally with high precision, the contribution of individual intermediate quantum states (their amplitudes) can only be inferred from the pattern of observed kinematic distributions of the final-state particles (this technique is called "partial-wave analysis" – PWA). Forward modeling and Bayesian methods provide a statistical model for the individual quantum states and the interference pattern. Physical states connected to individual quantum numbers span across the mass of the final-state particles. Since the shape of this mass dependence is a prior unknown, such analyses are typically done in independent bins of the final state mass. Performing fits mass-independently to determine the various amplitudes is equivalent to using complex-valued step functions for the individual amplitudes across the mass values of the final state. This is an alternative to using a model that parametrizes the mass dependence of amplitudes of physical resonances. While this provides us with a quasi-model-independent way of obtaining the mass dependence of the amplitude, the small datasets available for individual mass bins lead to large statistical fluctuations. They may even introduce systematic uncertainties due to multiple local fit optima. However, physics across the mass bins is smooth (and even unlikely discontinuities may be smoothed by resolution effects). Including this prior knowledge should thus improve the quality of the fit results but requires simultaneously using all n bins of the mass distributions. Precisely for this reason, the interdisciplinary cooperation has led to a new development promising a strong reduction of systematic uncertainties in light meson spectroscopy.

Astrophysicist Torsten Enßlin (MPA) and his group developed information field theory (IFT) (www.mpa.mpa-garching.mpg.de) to solve mathematical similar inverse problems in astrophysics, like that of tomographic reconstruction of the 3D Galactic dust distribution from its measurable absorption of starlight, or of finding signatures of an inflationary epoch of our Universe in the cosmic microwave radiation by the non-Gaussian statistics they imprinted. Since IFT addresses how to recover a field from finite data by generically exploiting its spatial correlations, it is not restricted to astrophysics but can be used in other areas as well. They also implemented the NIFTy framework for numerical information field theory, which provides us with the tools to create and fit IFT models.

In a first attempt to apply IFT-based methods to hadron spectroscopy, S. Wallner (TUM and now MPP) and F. Kaspar (TUM) in the group of S. Paul joined forces with P. Frank (MPA) and implemented the first version of such a model based on NIFTy and a ROOT-based PWA framework for partial-wave analysis. In a second attempt, J. Knollmüller (ODSL, formerly MPA) and F. Kaspar collaborated to implement a more advanced analysis framework. While J. Knollmüller provided the technical know-how of the NIFTy framework for numerical IFT, F. Kaspar developed the particle physics model and its application to data from the CERN experiment COMPASS. The result is an advanced information-theoretical model for PWA that allows to finely adjust many levels of prior information, e.g. adjusting the degree of continuity across kinematic variables or the inclusion of phase-space kinematics. This new technique's subtle but important feature is the combination of known mass dependencies of quantum states (resonance model) and unknown but coherent background contributions, which can now be extracted non-parametrically.

Using simulations, F. Kaspar proved the feasibility of this method. The figure demonstrates, for a selected partial wave, the reconstruction of the known mass dependence of the input signal (orange) using the standard (black points with uncertainty bars) and the new method (red). Including additional information in the IFT-based analysis significantly reduces the uncertainty. Currently, the new method is being applied to experimental data. Moreover, a publication about the method is in preparation.

Analysis reusability in high-energy physics

SUMMARY

The most extensive experiments worldwide provide excellent data about particle collisions at high energies. Given the complexity of the experiments, obtaining and analyzing this is a challenging task. Based on a use case from the Belle II experiment, ORIGINS scientists develop statistical tools and methods to reinterpret results in terms of new theories and pave the way for combining results within and across experiments.

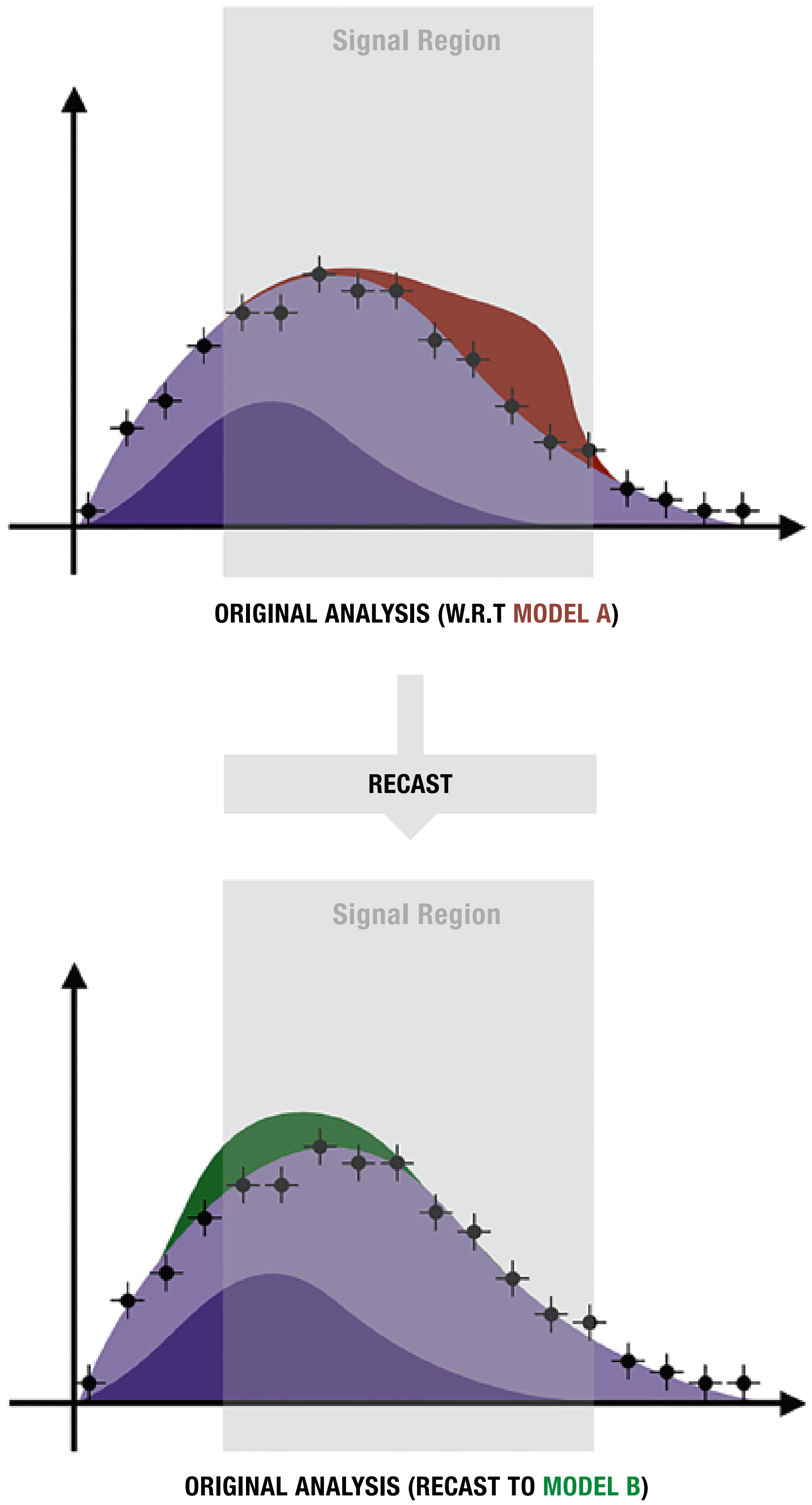
THE STORY

Rare weak decays of hadrons like B^+ –> $K^+ \nu\bar{\nu}$ searched for by the Belle II collaboration, are particularly interesting in particle physics research. Dubbed as "rare," the Standard Model marginally allows them to happen, which makes them an essential window into physics beyond the Standard Model. However, meeting the experimental challenges induced by the two undetectable neutrinos in the final state requires assumptions about the decay products' kinematic distribution. Consequently, the analysis results feature a model dependency arising from both Standard Model assumptions and the description of the pertinent hadronic matrix element (the quarks in the initial hadron must reconcile a final-state hadron despite a change of quark flavor), making reinterpretation complicated without a reanalysis of the underlying data.

ORIGINS scientists around Ph.D. student Lorenz Gärtner and supervisors Thomas Kuhr (LMU) and Lukas Heinrich (TUM) have teamed up to address the reinterpretability of data through a model-independent likelihood function. The theory space is parametrized in terms of Wilson coefficients of the weak effective theory. The simulated signal template, used as a comparison against the measured data, is reweighted according to the specific predicted kinematic signal distribution.

By extending the pyhf software, which is widely used for statistical inference in high energy physics, and interfacing it with the EOS software for flavor physics phenomenology (developed by former ORIGINS scientist Danny van Dyk, now in Durham), we can perform a computationally efficient update of the theoretical model, leading to value-constraints for the above-mentioned Wilson coefficients, characterizing effective theories.

The publication of such likelihoods strongly enhances the exploitation of experimental results. Once published, all particle physicists can use the model-independent likelihood function, who may now test existing theoretical models themselves. One can even go further and derive excluded regions of the theory parameter space, which helps to guide future research.



Given an analysis that is designed to test the prediction of a model A (red curve) against measured data (black points), our method makes it possible to exchange and test the prediction of an alternative model B (green curve). © L. Heinrich

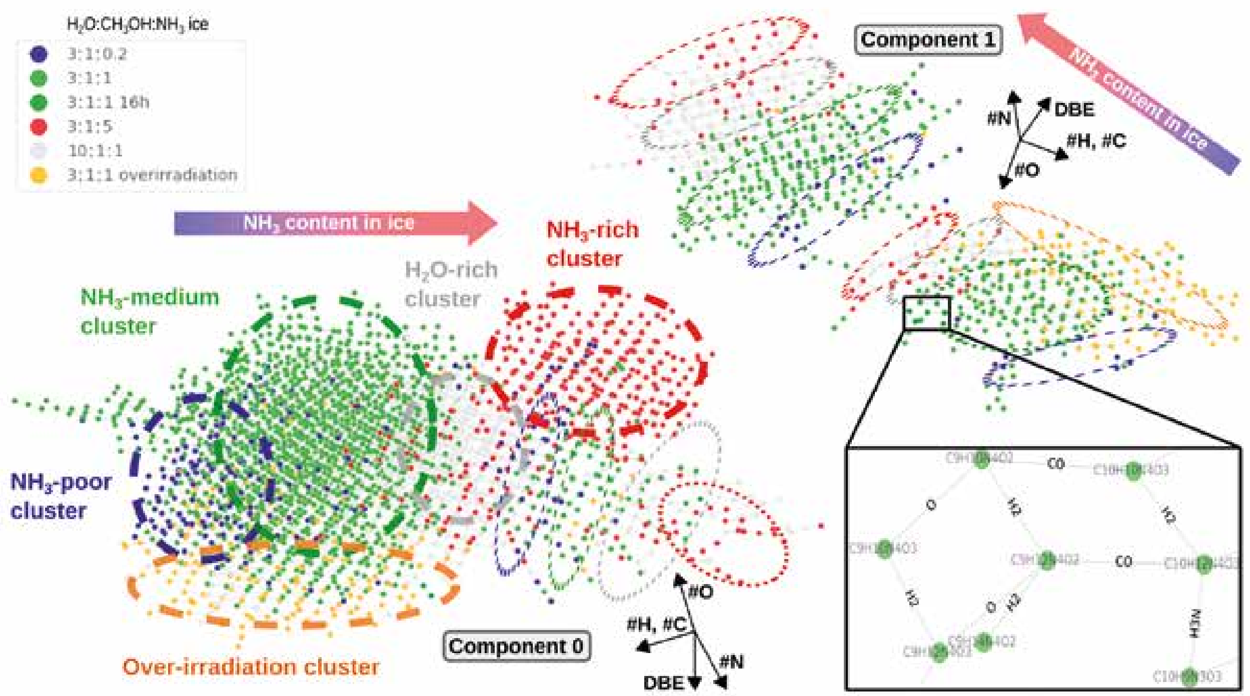
PUBLICATIONS

* **L. Heinrich, M. Feickert, G. Stark and K. Cranmer, "pyhf: pure-Python implementation of HistFactory statistical models," J. Open Source Softw. 6 (2021) 2823**
* **EOS Authors collaboration, D. van Dyk et al., "EOS: a software for flavor physics phenomenology," Eur. Phys. J. C 82 (2022) 569, arXiv:2111.15428**

FUNDING INFORMATION

* **Thomas Kuhr, LMU professor and ORIGINS PI**
* **Danny van Dyk, former TUM scientist and ORIGINS PI**
* **Lukas Heinrich is an ORIGINS-funded professor and PI (professor for data science in physics at TUM and ODSL coordinator)**

Modeling astrochemical reaction networks



mol2net molecular network of interstellar ice analogs unveiled the impact of ammonia in forming several tens of thousands of different complex organic molecules

SUMMARY

A recent interdisciplinary collaboration among astrochemists (Ruf & Bianchi) and data scientists (Würzinger & Heinrich) from the Excellence Cluster ORIGINS started a joint project to characterize path probabilities in complex molecular networks. The researchers use state-of-the-art data science methods to model mechanistic properties of complex reaction networks from a list of molecules only (without prior information about chemical reactions). This might have implications for better understanding exoplanet atmospheres or the chemistry inherited during the planet formation process.

THE STORY

The increasing number of molecular species detected in the interstellar medium and in Solar System objects, has led to the development of sophisticated astrochemical models aimed to identify the network of chemical reactions and reproduce the observed molecular abundances in an astronomical source of interest. This entails not only having prior information on the physical properties of the source but also considering, for each newly discovered molecule, a multitude of associated species and the necessary reaction pathways for their formation and destruction, both in the gas phase and on the icy dust grain mantles. Furthermore, each reaction requires to be studied theoretically and/or in laboratory experiments, demanding a significant investment of both time and financial resources. For this reason, it is fundamental to identify relations between molecular species, which could guide further astrochemical studies and steer resources toward the most critical reactions.

Graph-based Network analysis represents an excellent tool to analyze complex chemical systems as thoroughly as possible, in general or in (astro)chemical contexts. Molecular networks are characterized by molecules (detections) as nodes and transformations/reactions as edges. However, very often, the type of connections between molecules are still being determined and need to be inferred.

The mol2net method has previously been developed to characterize the ensemble properties of complex chemical systems with several tens of thousands of complex organic molecules. It has been applied to interstellar ices to unveil the role of ammonia chemistry in forming PAHs (polycyclic aromatic hydrocarbons) or N-aromatic heterocyclic molecules. Furthermore, molecular network analysis gave insights into the connection between insoluble and soluble organic matter in interstellar ice and meteorites, which shows implications for the transition between abiotic and prebiotically relevant organic molecules.

In this project, the researchers go one step further by modeling mechanistic insights of complex reaction networks from a list of molecules only, using state-of-the-art data science methods (without prior information about chemical reactions). The methods have been tested on known chemical products of irradiated methanol-ices from laboratory experiments. Different pathways for the CO –> CH3OH reaction have been analyzed. Preliminary results showed the impact of several machine learning methods, such as Gaussian processes, gradient boosting or Bayesian ridge regression.

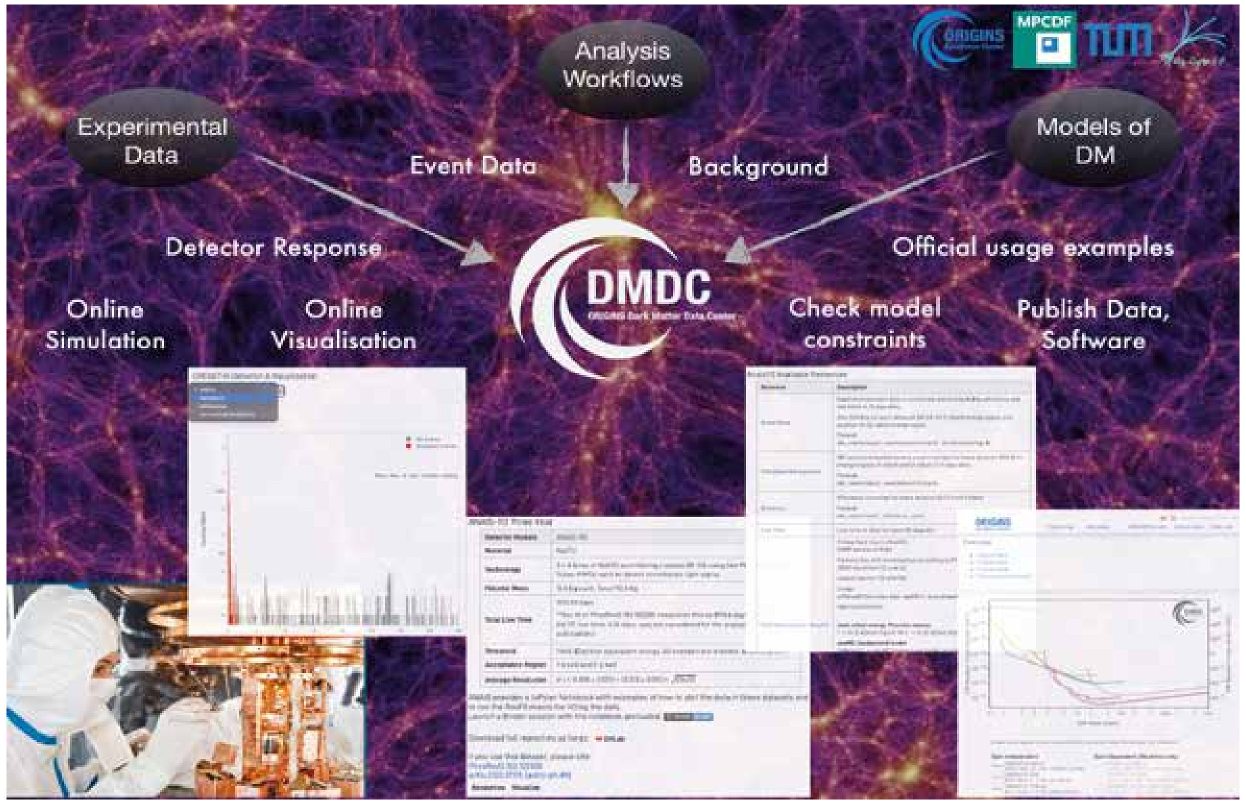
This interdisciplinary project between state-of-the-art astrochemistry and data science will enable a better understanding of complex astrochemical networks including various tens to hundred different molecular species with implications for characterizing exoplanets and planet formation processes.

PUBLICATIONS

* **A. Ruf & G. Danger, "Network Analysis Reveals Spatial Clustering and Annotation of Complex Chemical Spaces: Application to Astrochemistry," Analytical Chemistry, 94(41), 14135-14142 (2022)**
* **G. Danger, A. Ruf et al., "The transition from soluble to insoluble organic matter in interstellar ice analogs and meteorites," Astronomy & Astrophysics, 667, A120 (2022)**

FUNDING INFORMATION

* **Alexander Ruf, ORIGINS fellow, ODSL fellow, RU-E manager and PI**
* **Eleonora Bianchi, Vera Rubin fellow, RU-D manager and PI**
* **Jonas Würzinger, ODSL postdoc**
* **Lukas Heinrich, ORIGINS-funded professor and PI (professor for data science in physics and ODSL coordinator)**



Background from Springel et. al. (2005). © DMDC

A unified platform for data and information sharing within the dark matter community

SUMMARY

The Dark Matter Data Center (DMDC) brings together the large amount of recorded data and theories pertaining to dark matter (DM) research in a unified platform, making it easily accessible for the community. It is a forum where experimental collaborations can directly publish their data and workflows and phenomenologists the implementation of their models and analyses. All hosted data include clear descriptions with usage instructions and can be visualized online through interactive plots. All workflows can be run online on the platform through binders. This allows users to test their ideas before spending time and effort on rebuilding pre-implemented algorithms or even downloading the datasets. In addition to being a repository for data and workflows, the DMDC also engages in global analyses of experimental data with the central objective of enhancing experimental sensitivity through collective data analysis. This will include publication and maintenance of global constraints on the several classes of DM parameter space with exclusively open-source workflows.

PUBLICATIONS

* Heerak Banerjee et. al. (in preparation)

FUNDING INFORMATION

* Postdoc position of Heerak Banerjee is funded by ORIGINS.

THE STORY

The field of dark matter (DM) research is incredibly diverse. It encompasses a wide range of hypotheses about DM's mass and nature, from wave-like to particle-like to super-massive objects like black holes. Theories about DM's interactions with visible matter are equally varied in terms of character and strength. Efforts to study DM are correspondingly variegated. They can be categorized into three main experimental directions: direct searches, indirect searches, and collider searches, each involving various experimental technologies. For example, direct searches alone utilize at least six different detection methods, with no less than twenty experiments in the past decade alone, each employing unique technologies.

This diversity presents challenges in communicating results between theoretical and experimental researchers and even among different research groups. The lack of a unified framework for sharing information often leads to irreproducibility and unnecessary duplication of efforts. Moreover, the principles of open data and open science have become community standards in both theoretical and experimental collaborations. However, the scattered nature of data repositories, the complexity of associated workflows, and the effort required to find, download, and understand datasets hinder their accessibility.

To address these issues, the DM research community within the ORIGINS Cluster recognized the need for a centralized platform to host research data alongside clear and understandable instructions for reproduction or reinterpretation. This realization gave birth to the ORIGINS Dark Matter Data Center (DMDC), aimed at providing a single platform for hosting and publishing research data and workflows, facilitating easy exploration and comprehension.

The DMDC, currently operating within the ORIGINS domain, offers a user-friendly web page for exploring dark matter (DM) related datasets. Each dataset from a specific collaboration is listed on their dedicated page, accompanied by key information, detailed descriptions, and usage instructions for each resource. Users can visualize every dataset resource through interactive online plots, and associated workflows can be executed online via binders. Additionally, each dataset/workflow is assigned a unique DOI, simplifying dataset discovery and understanding. Users can explore, visualize, and experiment with the connected workflows without the need for downloads or installations. The computational framework, including data repositories and the servers for running the binders are provided by the Max Planck Computation and Data Framework (MPCDF) ensuring security and longevity.

Collaborations benefit from a repository to host and publish their datasets and workflows long-term, streamlining the explanation and defense of results over time. While the DMDC offers repositories, collaborations can also host their data and codes internally, with complete control over their DMDC webpage content through the metadata alone. Currently, datasets from the CRESST and ANAIS Collaborations are hosted through their repositories, and the COSINE Collaboration plans to share data through the DMDC. Public data from other collaborations is automatically included. The DMDC also currently features an interactive plot on its homepage displaying current exclusions in the DM-nucleon cross-section vs. DM mass plane. Beyond direct search, the DMDC is expanding to include indirect and collider searches. It is poised to be part of an upcoming proposal, serving as the public data repository for an experiment studying indirect signatures of multiple DM candidates using radio-frequency and polarimetric data from the LOFAR2.0 telescope.

Aside from ongoing efforts to add, maintain, and improve datasets, DMDC is actively working on a project to establish a database of current limits in the DM-SM interaction effective field theory (EFT) parameter space. This initiative will provide the community with a comprehensive resource to assess their DM models and experiment specific exclusions quickly. As an extension of this endeavor, we plan to introduce an intuitive online tool allowing users to select a specific EFT operator and experiment to access current limits on that operator. The underlying algorithm will be entirely open source, ensuring the reproducibility of published limits.

Looking ahead, our vision for the DMDC extends beyond being just a repository. We aim to evolve into an open source global DM data interpretation group. This expansion will encompass phenomenological research and data analysis, emphasizing the collective analysis of data from multiple experiments to enhance global sensitivity in detecting signals or confirming their absence. Similar to our current global limits project, all such initiatives will be open source, contributing to a database of accessible public software and code snippets making the DMDC a software development hub for the DM community.

Sequence selection in nucleic acid polymers by UV-light

SUMMARY

Using the state-of-the-art ORIGINS Ice, Dust and Sequencing Lab (IDSL), a recent collaboration of LMU and TUM biophysicists (Mast, Zinth, Braun, Bucher, Kufner) was able to measure the UV sensitivity of nucleotide polymers over an unprecedented dose and sequence range. The authors developed a new method involving high-throughput sequencing to determine the formation of UV damage for tens of thousands of different DNA sequences as a function of sequence context in a single experiment. The measurements will make it possible to model the behavior of oligonucleotides with arbitrary sequences and thus determine damaged-induced biases for large sequence pools under the influence of UV radiation. Specifically, it will allow us to simulate the influence of UV radiation on early life, which was not yet protected by an ozone layer or capable of genetic repair.

THE STORY

Various physical (e.g., radiation) or chemical (reactive compounds) processes can damage genetic material. While in the case of UV radiation, the ozone layer and the evolutionary sophistication of DNA's repair enzymes protects modern life, this was not the case for the earliest life forms. These would face a dilemma in the choice of their preferred genetic repository. While RNA is chemically unstable, a massive accumulation of light damage is likely to occur in the much longer-lived DNA. To understand the UV selection pressure on early life, damage measurements made for modern life forms are unsuitable since their highly evolved repair enzymes distort the experimental outcome. For the UV susceptibility of such "blank" DNA, however, there is hardly any literature covering the required wide dose range and the sequence context around the damage site. However, this knowledge is required to understand the selection pressures for specific sequences caused by UV radiation.

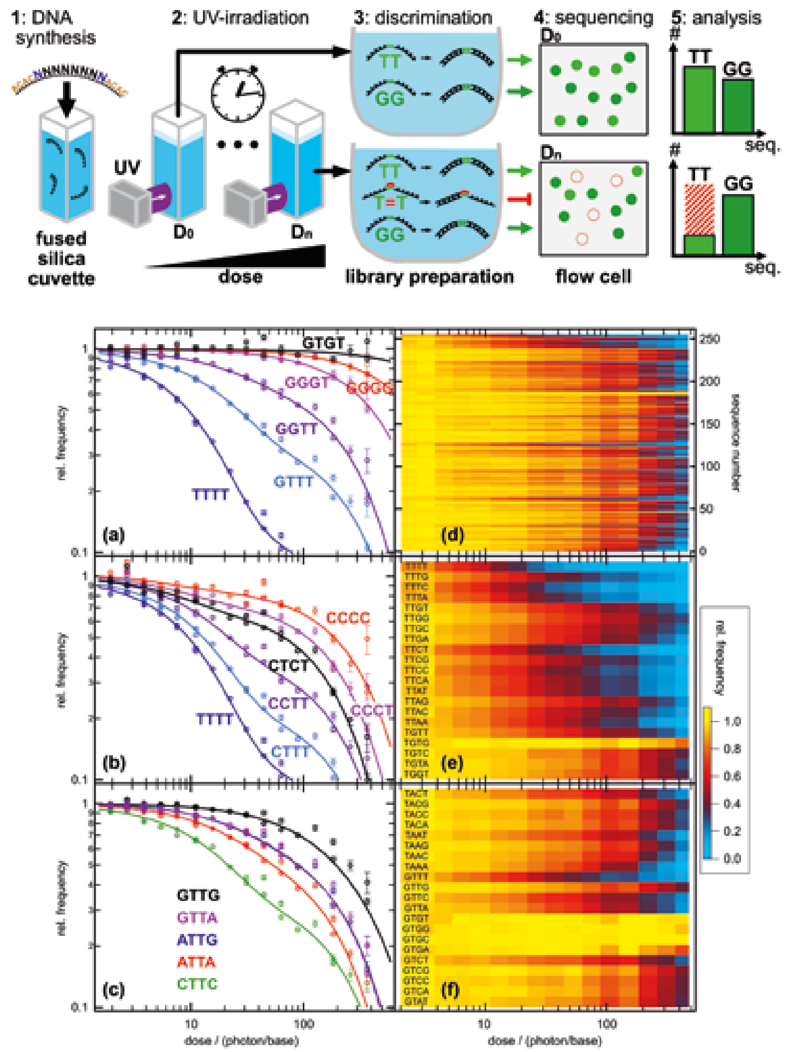
For this reason, a new method to analyze high-throughput sequencing data by a collaboration of PIs from LMU (Mast, Zinth), ORIGINS (Braun), and TUM (Bucher) was developed. They combined expertise in photo- and biophysics to determine the rates at which such photodamage occurs in non-biological DNA over a wide sequence range up to high irradiation doses. Under the first authorship of Corinna Kufner, who is now working on the photochemistry of prebiotic polymers at Harvard University and who also presented the results at the ORIGINS Science Week 2021 in Kloster Irsee, and the collaboration of Marlis Fischaleck (funded by ORIGINS), who prepared the samples for high-throughput sequencing at the Gene Center Munich, the validity of the method was demonstrated in the already known sequence and dose ranges. Following this, the development of photodamage could be quantified up to a previously unaddressed dose and sequence range. The data thus obtained allows us to accurately simulate the selection pressure on DNA by UV radiation and enables the comparison with models that predict the emergence of sequence patterns for early life forms. In a further collaboration of the participating groups, the method can also be extended to the RNA context. In summary, the methods developed in this collaboration using the IDSL lab open up new possibilities in realistically describing the molecular evolution of early life by incorporating increasingly complex and realistic non-equilibrium conditions available on the early Earth.

PUBLICATIONS

* **C. L. Kufner, S. Krebs, M. Fischaleck, J. Philippou-Massier, H. Blum, D. B. Bucher, D. Braun, W. Zinth and C. B. Mast, "Sequence dependent UV damage of complete pools of oligonucleotides," Scientific Reports doi.org/10.1038/s41598-023-29833-0 (2023)**

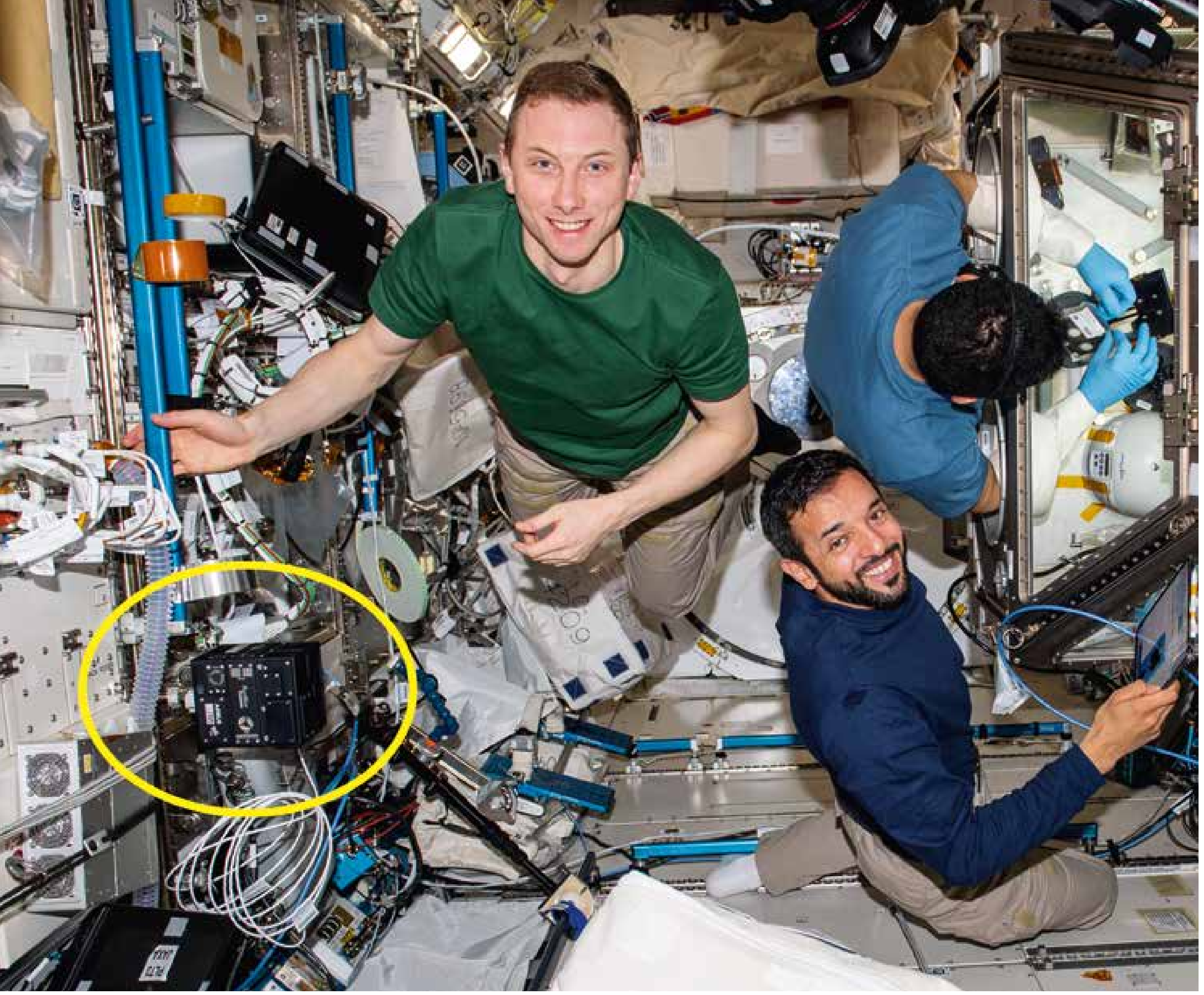
FUNDING INFORMATION

* **Marlis Fischaleck is funded by ORIGINS**
* **IDSL state of the art "ORIGINS of Life" lab mostly funded by ORIGINS**
* **The project was funded by ORIGINS 2021-2 Seed money.**



Method and results of the simultaneous measurement of sequence-dependent damage on a large number of DNA strands, allowing the accurate determination of damage frequencies over a wide dose range. © Kufner et al. (2023)

Instruments of the ORIGINS Ice, Dust and Sequencing Lab showing a high-resolution HPLC-ESI-TOF mass spectrometer, a HPLC-OrbiTrap mass spectrometer and a NT242 tunable DPSS UV-laser. The lab equipment is open to origin of life researchers in Munich and serves as a hub for manipulating and probing prebiotic molecules. It is located in downtown Physics at the LMU in the AG Braun. © AG Braun



NASA astronaut Warren Hoburg and UAE astronaut Sultan Al Neyadi in the Japanese experiment module Kibō on the International Space Station. The RadMap telescope is attached to a rack next to them. © NASA

Radiation detectors for future missions to the Moon and Mars

PUBLICATIONS

- M.J. Losekamm, S. Paul, T. Pöschl, H.J. Zachrau, "The RadMap Telescope on the International Space Station," IEEE Aerospace Conference; doi: 10.1109/AERO50100.2021.9438435 (2021)
- M.J. Losekamm et al., "Measuring Cosmic Rays with the RadMap Telescope on the International Space Station," 38th ICRC; doi: 10.22323/1.444.0099 (2023)
- L. Meyer-Hetling et al. JGR Space Phys., in preparation
- M.J. Losekamm et al. Space Sci. Rev., in preparation

FUNDING INFORMATION

- LRSM is an ORIGINS Infrastructure
- Ph.D. students Martin Losekamm, Thomas Pöschl, Peter Hinderberger, Luise Meyer-Hetling (all TUM Physics, Prof. Paul), and Sebastian Rückerl (TUM LRT, Prof. Walter) funded by ORIGINS
- The project also received funding from DFG – Projektnummer 414049180

SUMMARY

The medical consequences of long-term exposure to cosmic rays and solar particles are a major challenge for future human missions to deep-space destinations. At the LRSM, ORIGINS scientists are developing new particle detectors to help address the lack of comprehensive data on the space radiation environment beyond Earth orbit. With the RadMap Telescope, the first detector prototype has been operational on the International Space Station since April 2023.

THE STORY

The constant bombardment of spacecraft and crew by particles and nuclei of cosmic and solar origin is one of the major challenges of missions to the Moon, Mars, and other deep-space destinations. At present, uncertainties in understanding and predicting the medical consequences of exposure to the space radiation environment greatly limit the maximum permissible duration of deep-space missions – potentially putting, for example, NASA's long-term plans for crewed missions to Mars in jeopardy, at least in their currently envisaged form. This uncertainty largely stems from our incomplete understanding of the biological effects of highly ionizing radiation. It is, however, also rooted in a lack of comprehensive, time-resolved data on the radiation environment beyond Earth orbit.

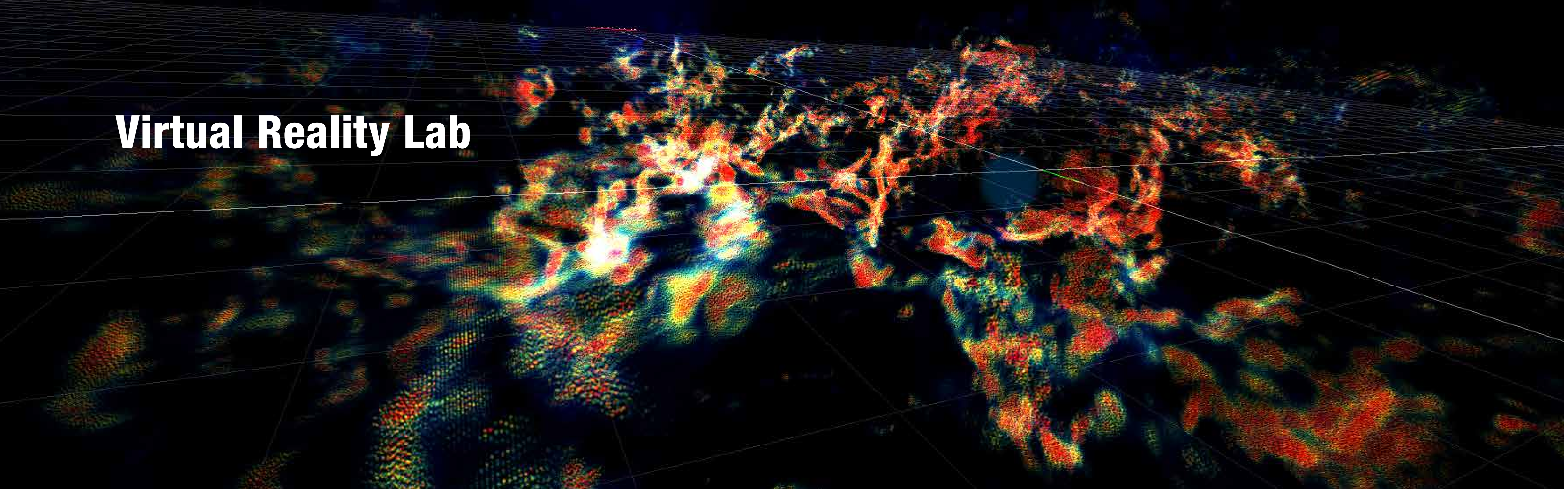
From an astrophysical perspective, the cosmic-ray spectrum at the GeV-to-TeV scale has been determined very precisely by space-based experiments like AMS-02. For radiation protection, however, this data is of limited use. To calculate meaningful dosimetric values, accurate knowledge of the short-term temporal evolution of the complex interplay between cosmic rays, solar particles, and the solar wind is required. Equally important are the particles' spectrum-altering interactions with the spacecraft and nearby planetary atmospheres, magnetospheres, and surfaces. Although robotic probes have carried radiation-monitoring instruments into deep space in the past decades, these missions provided, with few exceptions, only momentary snapshots of the radiation environment at a given location in the Solar System. For example, only a few measurements were recorded by satellites orbiting the Moon or Mars, and even fewer by probes on the lunar or Martian surface.

With partial funding from ORIGINS, scientists at the LRSM are developing compact yet powerful particle detectors that can provide more accurate dosimetric measurements than the current generation of sensors used, for example, on the International Space Station (ISS). Enabled by the advent of miniaturized yet highly sensitive silicon-based photosensors, the group was able to develop a class of conceptually simple and mechanically robust detectors that can achieve energy resolutions and particle-discrimination capabilities comparable to those of much larger systems – without the need for a heavy and power-consuming magnet. The omnidirectional sensitivity of the design also results in a very large geometrical acceptance, enabling the collection of significantly larger data samples than gathered by current-generation instruments. The simplicity of the detector design comes at the cost of a more complex data analysis – a challenge that the team is attempting to tackle by using machine-learning algorithms. If proven to work reliably, such algorithms would also allow running the data analysis on the instrument itself, paving the way for the first real-time radiation monitor with particle-discrimination capabilities.

To demonstrate the capabilities of their design, the team partnered with Houston-based Airbus US Space and Defense to test a prototype, called the RadMap Telescope, aboard the ISS. In a combined effort, the partners successfully convinced the US ISS National Laboratory to sponsor the deployment to the station. In addition to the detector developed in Munich, the instrument comprises a dosimeter developed at the German Aerospace Center (DLR) that has been part of NASA's Artemis I mission around the Moon in late 2022. Without a Rad Map-like detector in lunar orbit, data from the dosimeter will allow the scaling of measurements recorded on the ISS to the lunar environment. The instrument was launched aboard the SpaceX CRS-27 resupply flight on 15 March 2023 and successfully deployed to its initial operating location on 26 April 2023. In a period lasting at least through the summer of 2024, RadMap will move between four locations to assess the differences in shielding of American, European, and Japanese ISS modules.

Initial analyses of the recorded raw data have revealed that the instrument performs as expected. The team is currently attempting the first detailed analysis with machine-learning algorithms fine-tuned to the characteristics of the in-flight data. The experience gained with RadMap will be used to build a detector for a small-satellite mission measuring the flux of low-energy antiprotons in Earth's magnetosphere.This project profited from the cooperation of LRSM physicists and engineers from TUM's Institute of Astronautics (LRT).

Virtual Reality Lab



Example of a web visualization: the Galactic dust in the interactive Galactic Cartography Portal. The different colors indicate high (red), medium (green) and low (blue) density.
© Leike et al. (2022)

SUMMARY

Virtual Reality (VR) and Augmented Reality (AR) have recently experienced significant growth thanks to technological advancements. The ORIGINS Virtual Reality Lab explores how VR and AR technologies can be used for scientific visualization and provides it as a service for ORIGINS scientists.

THE STORY

Virtual reality has been around for decades but only recently experienced a significant resurgence due to major developments in smartphone technology. These advances have led to higher-resolution displays, lower hardware costs and lighter overall devices, all of which have helped to make VR more accessible and enjoyable. Similar techniques such as AR and holographic displays have also greatly benefited from these advances, leading to major improvements in their capabilities.

A common challenge in scientific visualization is to effectively present large amounts of 3D data simultaneously without compromising interactivity. While traditional computer displays allow some leeway in terms of fluctuating frame rates during interactive sessions, this issue becomes much more critical in the context of VR applications. Maintaining stable frame rates within the recommended range of 60 to 90 frames per second is crucial, as failure to do so can result in user discomfort ranging from nausea to stomach upset. This requirement leaves approximately 10-20 milliseconds to render each frame. To improve performance, it is important not only to implement various optimization techniques and tricks within the software, but also to rely on high-end hardware. The ORIGINS Virtual Reality Lab (VRLab) project explores and implements VR and AR technologies to enable scientific visualization for researchers within the ORIGINS cluster. The VRLab team is led by ORIGINS scientist Christian Alig (LRZ+LMU). The VRLab collaborates with the Leibniz Rechenzentrum (LRZ) Virtual Reality and Visualisation Team (V2C) led by Thomas Odaker (LRZ) as well as C2PAP which provides computational resources. The University Observatory Munich hosts the development site while a VRLab room with its user hardware and software will be installed in 2024 at the ORIGINS cluster building in Garching, making it more accessible to the ORIGINS community and closer to LRZ.

PUBLICATIONS

* **R. H. Leike, G. Edenhofer, J. Knollmüller, C. Alig, P. Frank, T. A. Enßlin, "The Galactic 3D large-scale dust distribution via Gaussian process regression on spherical coordinates," arXiv:2204.11715 (2022)**

FUNDING INFORMATION

* **Christian Alig received funding as ORIGINS postdoc**
* **Natalia Borodatchenkova was a former ORIGINS postdoc involved in the project**
* **Jakob Knollmüller is an ODSL Postdoc**
* **ORIGINS Seed Money 2021: R. Leike, T. Enßlin, J. Knollmüller, "Interactive Visualization of 3D Galactic Dust Maps"**

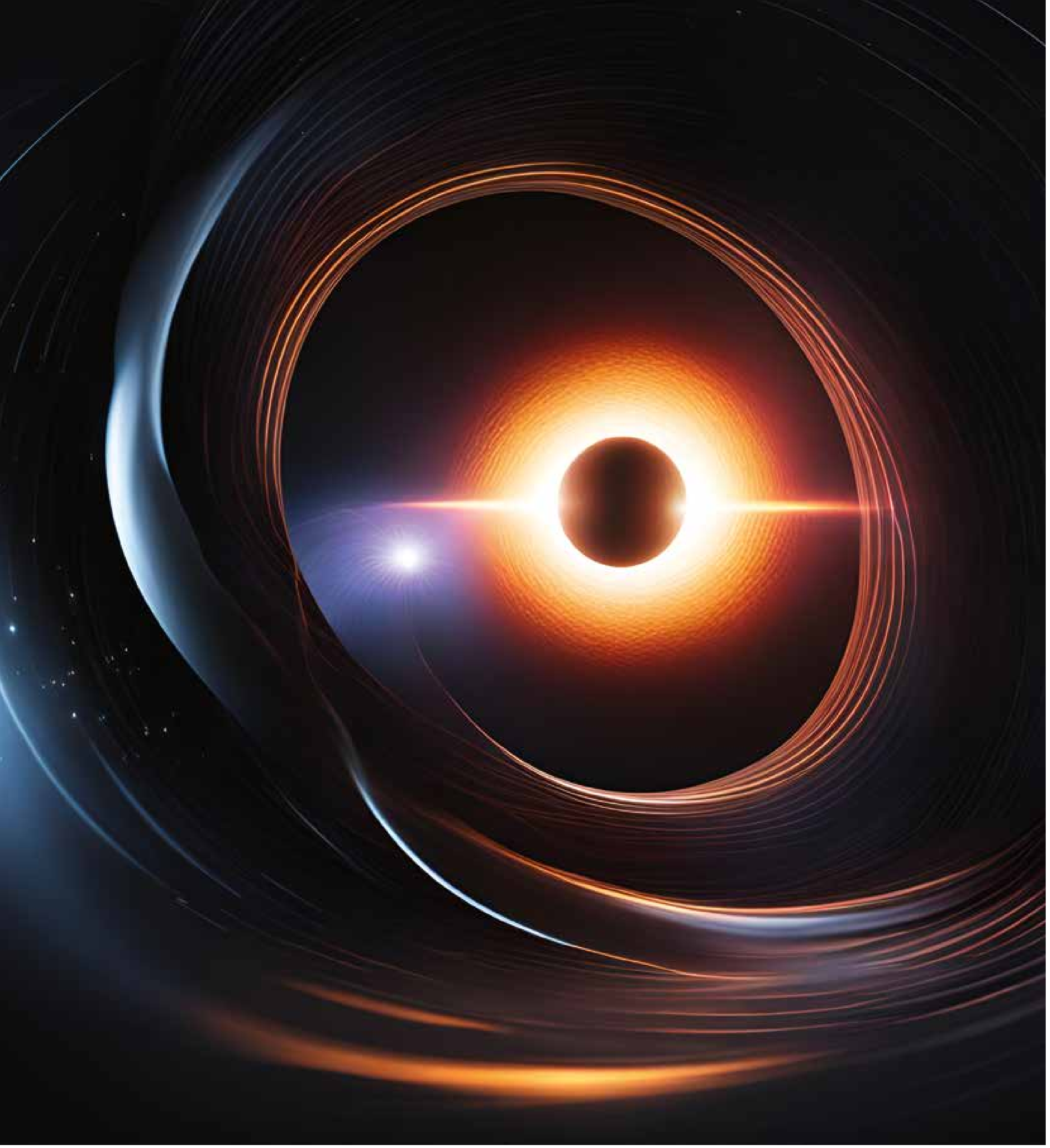
The VRLab uses Unreal Engine 5 to develop the software, an advanced platform typically used for game development. However, its versatility allows harnessing its potential for high-performance visualization. The engine easily accommodates all the different technologies, including VR headsets, mobile devices designed for AR experiences, holographic displays and more. There is no need to write code specifically for a single device; instead, Unreal provides an abstraction layer that ensures functionality across all supported platforms. In addition, Unreal Engine comes with a wide range of out-of-the-box systems such as user interfaces, user motion mechanics, camera controls, post-processing effects, lighting systems and more.

The VRLab provides the following services for ORIGINS scientists:

- VR visualization of particle and grid-based simulations, or more generally for point cloud data.
- Visualization using 3D (holographic) monitors together with hand tracking for interaction.
- Web (browser) based 3D visualization.
- AR visualization using smartphones and headsets such as HoloLens.

A recent example of a browser-based 3D visualization is the Galactic Cartography Portal: https://flute.esc.mpcdf.mpg.de. At this site, one can explore a 3D reconstruction of the Galactic dust distribution, query radial dust density slices through our Galaxy and download custom regions for scientific research. Knowing the Galactic 3D dust distribution is relevant for understanding a multitude of processes in the interstellar medium (ISM), which plays a vital role in the metabolism of our Galaxy. ORIGINS scientists Torsten Enßlin (MPA), Christian Alig (LMU) and Jakob Knollmüller (TUM), former ORIGINS researcher Reimar Leike (OpenAI), as well as Philipp Frank (MPA), reconstructed the Galactic dust distribution with an increased number of meaningful resolution elements by orders of magnitude with respect to previous reconstructions, while taking advantage of the dust's spatial correlations to inform the dust map. They use iterative grid refinement to define a log-normal process in spherical coordinates. This log-normal process assumes a fixed correlation structure, which was inferred in an earlier reconstruction of Galactic dust. The map is based on 111 million data points, combining data from PANSTARRS, 2MASS, Gaia DR2 and ALLWISE. The log-normal process is discretized to 122 billion degrees of freedom. They derive the most probable posterior map and an uncertainty estimate using natural gradient descent and the Fisher-Laplace approximation.

The dust reconstruction covers a quarter of the volume of our Galaxy, with a maximum coordinate distance of 16 kpc. Meaningful information can be found up to distances of 4 kpc, still improving upon our earlier map by a factor of 5 in the maximal distance, of 900 in volume, and of about 18 in angular grid resolution. Unfortunately, the maximum posterior approach chosen to make the reconstruction computational affordable introduces artifacts and reduces the accuracy of the uncertainty estimate. Despite the apparent limitations of the presented 3D dust map, a good part of the reconstructed structures is confirmed by independent maser observations. Thus, the map is a step towards reliable 3D Galactic cartography and can already serve for several tasks, if used with care.



Artist impression of a „Quantum black hole“ created by Nightcafe AI. The depicted swirls and illuminations have no physical origin. © V. Thoss / Nightcafe

Primordial black holes beyond the Hawking picture

SUMMARY

In a highly interdisciplinary collaboration, ORIGINS provides a new Ph.D. position to investigate whether primordial black holes could be dark matter candidates. Recent results in fundamental physics by Gia Dvali and Dieter Lüst, who developed a quantum theory of black holes that shows that light black holes do not evaporate as quickly as previously predicted by Steven Hawking, inspired this project. In a specific mass range, such primordial black holes could survive until today, which allows us to consider them as dark matter candidates. Ph.D. student Valentin Thoss investigates whether this new insight can explain the observationally inferred distribution of dark matter and how one could observationally test the conjecture that dark matter consists of primordial black holes.

THE STORY

Primordial black holes are hypothesized to have formed shortly after the Big Bang. Various formation scenarios include the collapse of inhomogeneities in the very early Universe. In principle, such objects could obtain a wide range of masses from the Planck mass (10^{-8} kg) to more than the Sun's mass (10^{30} kg). However, their potential effect on the structure and evolution of the Universe has led to strong constraints on their abundance. Comprehensive studies of scenarios could largely exclude that the entirety of dark matter comprises single mass primordial black holes. Only a narrow mass range in the ballpark of asteroids is still possible. Hawking evaporation rules out lower mass primordial black holes; they would not survive from the Big Bang until today.

However, in 2020, Gia Dvali and collaborators showed that Hawking's (semiclassical) calculation of the evaporation rate breaks down when the black hole has lost roughly half of its mass. At this point, the previously neglected quantum back-reaction of the emitted particles on the black holes becomes relevant. This back-reaction will lead to a strong suppression of the evaporation rate, as Gia Dvali and his collaborators showed. Consequently, light primordial black holes previously thought to evaporate until the present epoch of the Universe could survive until now.

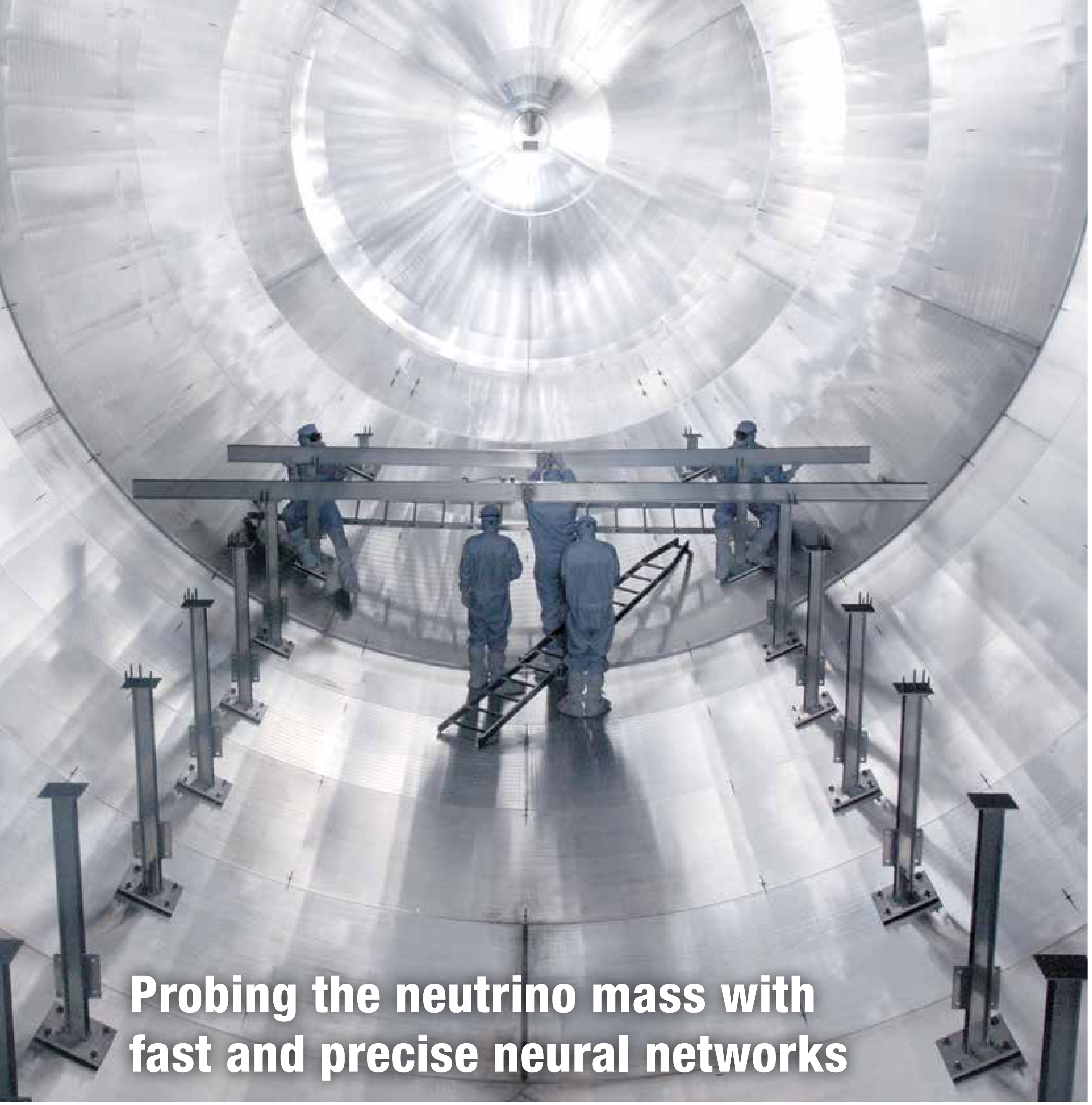
CN-1 Ph.D. student Valentin Thoss is co-supervised by MPA director Volker Springel, a leading expert on cosmic structure formation, Andreas Burkert (LMU, computational astrophysics), Dieter Lüst (LMU, mathematical physics), and MPP director Gia Dvali (LMU, theoretical particle physics). He is analyzing for which range of parameters these light primordial black holes survive until today and whether they would leave any observational signatures. Preliminary results show that if the suppression of the evaporation is strong enough, a new mass window below 10^7 kg emerges, with a lower mass limit depending on the strength of the suppression. This opens up an exciting new possibility to investigate the nature of dark matter and study the physics of these light black holes.

PUBLICATIONS

* **V. Thoss et. al. 2023, in preparation**
* **G. Dvali, “Black hole metamorphosis and stabilization by memory burden,” Phys. Rev. D, 102, 103523 (2020)**

FUNDING INFORMATION

* **ORIGINS funded Ph.D. student Valentin Thoss**



Probing the neutrino mass with fast and precise neural networks

The KATRIN main spectrometer from the inside during the installation of the inner wire electrode system. The main spectrometer acts as an electrostatic filter for the tritium beta-decay electrons. The vessel wall and the >24 000 sub-mm-thin wires define the high voltage for the filter. © KATRIN Collaboration

PUBLICATIONS

* KATRIN Collaboration, "Direct neutrino-mass measurement with sub-eV sensitivity," Nature Physics 18, 160–166 (2022)
* C. Karl, P. Eller, S. Mertens, "Fast and precise model calculation for KATRIN using a neural network," Eur. Phys. J. C 82 (2022) 5, 439
* New KATRIN data release - in preparation (to be submitted to Science)

FUNDING INFORMATION

* ODSL fellow Oliver Schulz
* ODSL fellow and previous ODSL postdoc Philipp Eller
* ODSL postdoc Jakob Knollmüller

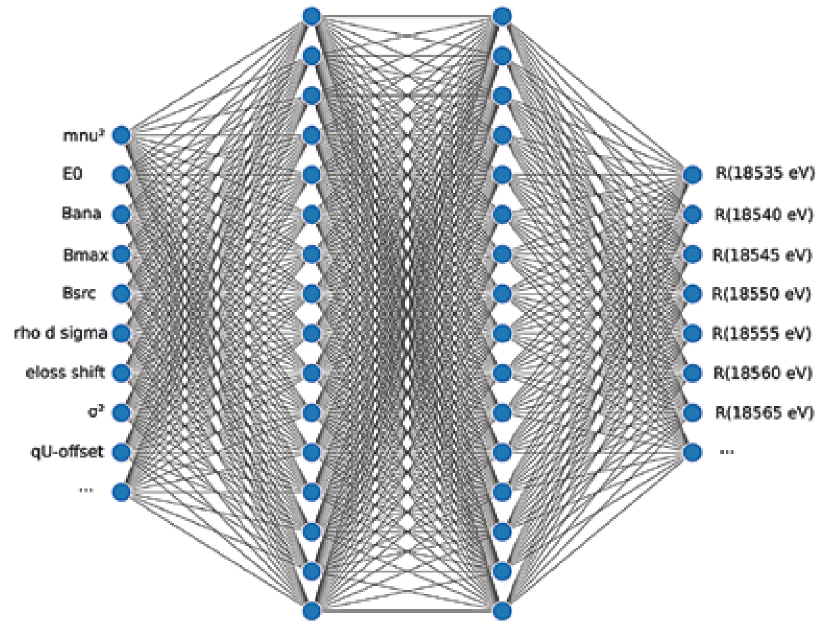
SUMMARY

The absolute value of the neutrino mass is one of the most pressing open questions in particle physics. The Karlsruhe Tritium Neutrino (KATRIN) experiment is the leading experiment to directly measure the neutrino mass with a sensitivity of 0.2 eV/c^2 at a 90% confidence level. A joint team of scientists of the ORIGINS KATRIN group and the ORIGINS Data Science Laboratory (ODSL) developed a novel technique based on machine learning to solve the challenges of the KATRIN data analysis. This new method allows for an ultra-fast and precise theoretical prediction and is now one of the two official analysis strategies employed by the KATRIN collaboration.

THE STORY

With a mass at least five orders of magnitude smaller than the mass of an electron, neutrinos are an apparent misfit in the standard model (SM) of particle physics, which may be the key to physics beyond the SM. The least model-dependent approach to assess the neutrino mass is based on the kinematics of a single beta decay. Here, the neutrino mass manifests itself as a slight distortion of the spectral shape in the close vicinity of the kinematic endpoint. The leading experiment based on this approach is the Karlsruhe Tritium Neutrino (KATRIN) experiment. The 70 meter long KATRIN beamline combines a high-luminosity molecular tritium source (10^{11} decays/s) with a high-resolution ($\Delta E/E1$ eV) spectrometer. KATRIN is taking data since 2019. In 2022 KATRIN published the world-leading upper limit on the neutrino mass of $m < 0.8$ eV. A new result, based on a six-fold increase in data sample and significantly reduced systematic uncertainties, will be released in the very near future. The ORIGINS KATRIN group is led by Susanne Mertens (TUM).

The spectral distortion caused by the neutrino mass is extremely small (10^{-3} level). Thus, an ultra-precise model is required which includes the theoretical description of molecular tritium beta-decay and all experimental effects which can alter the spectral shape. Consequently, the semi-analytical model used by the KATRIN Collaboration is computationally extremely expensive. A single fit to the KATRIN data, typically divided into hundreds of sub-data sets, can last days to weeks. The combined analysis of all KATRIN data sets may even become unfeasible using the semi-analytical model.



The conceptual design of the neural network applied in the KATRIN analysis. The input nodes represent the input parameters to initialize the theoretical tritium beta-decay spectrum, as measured by KATRIN. The output notes denote the corresponding spectrum, described by individual rates (R) at a given energy. © KATRIN Collaboration

To tackle this challenge, the ORIGINS KATRIN group joined forces with ODSL scientists Philipp Eller (TUM), who is a particle physicist specializing in neutrinos, Oliver Schulz (MPP), who works on Bayesian and numerical methods, neutrinolessdouble-beta decay and germanium detector technology, and Jakob Knollmüller (TUM), who works on Bayesian methods, approximate inference in high dimensions as well as dynamic reconstructions of black holes from VLBI data. Together, they developed a new idea: The semi-analytical model is used to train a neural network first, which can then predict the spectral shape of the measured tritium spectrum. They demonstrated that this application of the neural network leads to a speed improvement by three orders of magnitudes without any loss in precision. The new software, called NETRIUM, is now applied to the upcoming data release, which will reach a sensitivity of better than $m < 0.5$ eV (90 % CL). Moreover, the method was adapted to searches for physics beyond the SM, such as sterile neutrinos. This groundbreaking application of neural networks to KATRIN also opened up the way for a complete Bayesian analysis of the KATRIN data. This analysis is a follow-up project of the KATRIN-ODSL cooperation and is presently ongoing.

CONTACT: Excellence Cluster ORIGINS · Boltzmannstr. 2 · D-85748 Garching
Tel: +49 89 35831-7100 · info@origins-cluster.de · www.origins-cluster.de
Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2094 – 390783311.