The Dark Matter Data Centre

An ORIGINS project

N. Ferreiro Iachellini | iDMEu | 12.05.2021

The ORIGINS Cluster

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2094 – 390783311

- Devoted to the investigation of the Universe from the Big Bang to the emergence of life
- Interdisciplinary cluster that combines astrophysics, biophysics and particle physics
- Universities: Ludwig-Maximilians-Universität Munich
  Technische Universität München
- Other institutions: ESO – European Southern Observatory
  LRZ – Leibniz Supercomputing Centre
  Max Planck Institute for Astrophysics
  Max Planck Institute for Extraterrestrial Physics
  Max Planck Institute for Physics
  Max Planck Institute for Biochemistry
  Max Planck Institute for Plasma Physics

ODSL – The Origins Data Science Laboratory

- Data Science division of the Origins Cluster
- Development of analysis methods, algorithms and computational tools to fully exploit high-dimensional, complex data sets

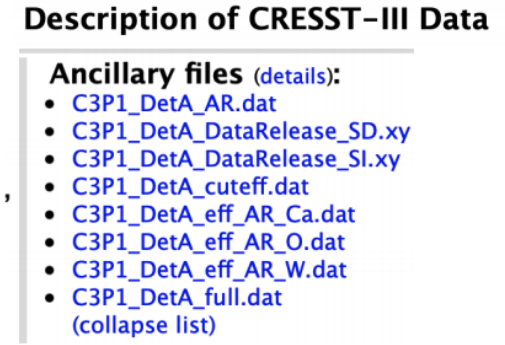
The Dark Matter Data Center: Our goal

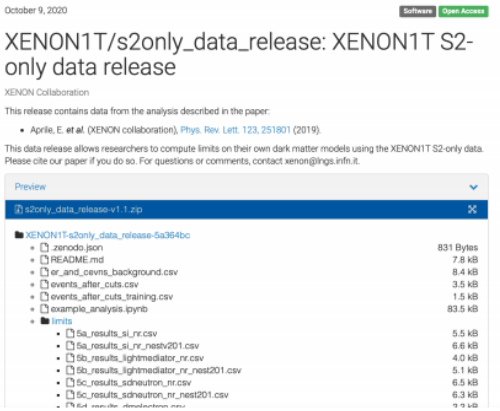
- Serve the community fostering a culture of open data and methods
- Ease and boost the (proper) usage of the available (published) experimental data
- Support for scientists for rigorous usage of the data
- Stimulate the circulation of theoretical predictions and their impacts in the various research fields
- Offer an open space constantly updated where the most up-to-date results are reported

From Laura Baudis on Monday

COMPARISON WITHIN OUR FIELD: DATA SHARING

- Data sharing: non-uniform in terms of content, location, format (Zenodo, GitHub, ancillary files on arXiv, ...)
- **Examples:**
  - CRESST-III data description, arXiv:1905.07335
  - XENON1T S2-only data: https://doi.org/10.5281/zenodo.3982637 and https://github.com/XENON1T/s2only_data_release
  - SuperCDMS, arXiv:2005.14067, ancillary files
- **Challenge: good documentation of released data**





From Paolo Privitera on Monday

Outlook

- Dark Matter direct detection with electron recoils is very young!
  noble liquids (2018), Skipper CCD (2017), HVeV (2018)
  New experimental techniques may pop out
  A large space to cover, a decade long program ahead

Community efforts as iDMEu are important for a successful program

- Theoretical efforts, both in DM model building and novel detection approaches
- Common framework: theoretical models, calculation of rates, ionization (fluctuations, Fano factor); repository?
- Complementarity/comparison with accelerator based searches
- Complementarity with indirect searches (astrophysics/cosmology)
- Experiments share many challenges in background control, also common to experiments with nuclear recoils
- Underground infrastructure (Radon-free, large clean rooms for detector assembly/test, e.f. copper in-situ)
- Many common challenges with the nuclear recoil direct detection

22

Already some attempts up to now

COMPARISON WITHIN OUR FIELD: STATISTICAL FRAMEWORK

- So far: no consensus on the statistical inference of DDNR data, different approaches used by various collaboration (see, e.g., talks at PHYSTAT-DM 2019 https://indico.cern.ch/event/769726/overview)
- New White Paper (arXiv: 2105.00599, emerged out of PHYSTAT-DM 2019): to establish set of recommended conventions for reporting results
  - e.g., profile likelihood-based test statistics -> Likelihood Ratio (PLR) (following Cowan, Cranmer, Gross, Vitells, EPJ-C 71, 2011)
  - to asses discovery significance & construct confidence intervals
- See White Paper for full list of recommendations (also on astrophysical parameters & astrophysical neutrinos)

Recommended conventions for reporting results from direct dark matter searches

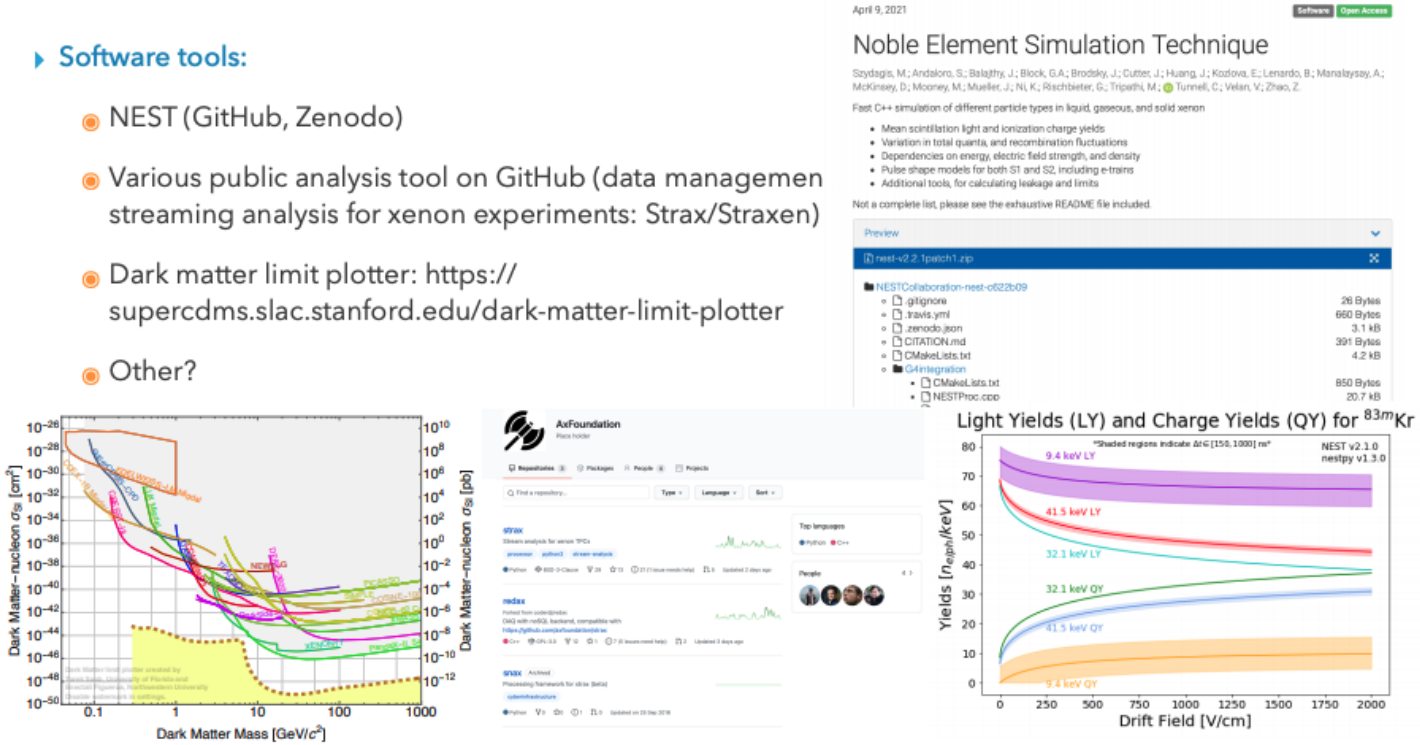
D. Baxter, I. M. Bloch, E. Bodnia, X. Chen, J. Conrad, P. Di Gangi, J. E. Y. Dobson, D. Durnford, S. J. Haselschwardt, A. Kaboth, R. F. Lang, Q. Lin, W. H. Lippincott, J. Liu, A. Manalaysay, C. McCabe, K. D. Mora, D. Naim, R. Neilson, I. Olcina, M.-C. Piro, M. Selvi, B. von Krosigk, S. Westerdale, Y. Yang, N. Zhou

PHYSTAT Dark Matter 2019

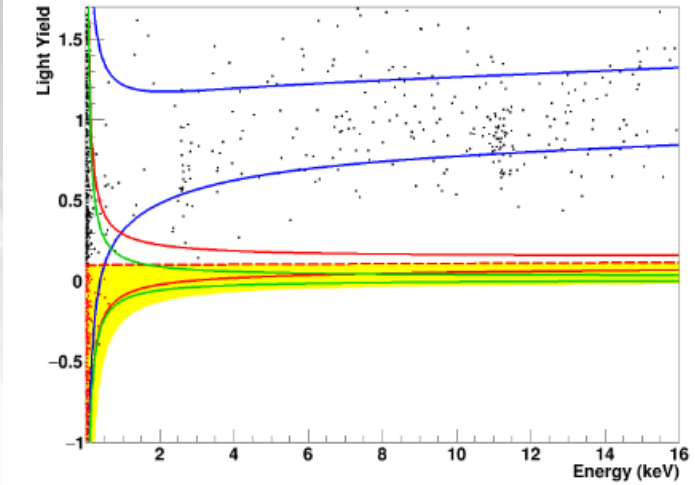
Already some attempts up to now

COMPARISON WITHIN OUR FIELD: SOFTWARE TOOLS

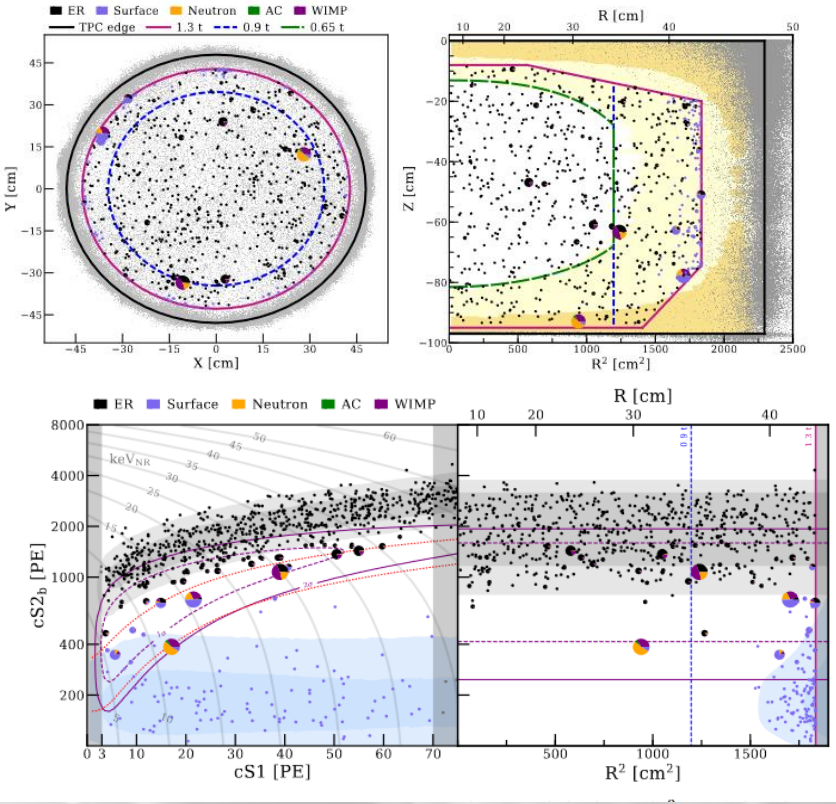
- Software tools:
  - NEST (GitHub, Zenodo)
  - Various public analysis tool on GitHub (data managemen streaming analysis for xenon experiments: Strax/Straxen)
  - Dark matter limit plotter: https://supercdms.slac.stanford.edu/dark-matter-limit-plotter
  - Other?



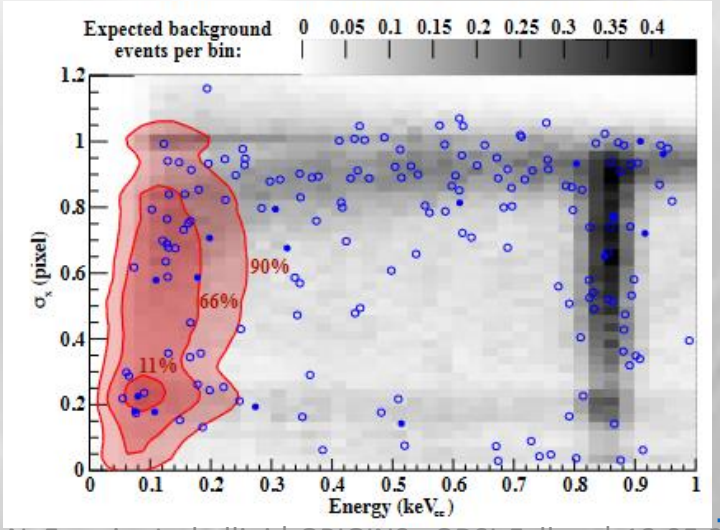
Large variety of technologies and data requirements



arXiv:1904.00498

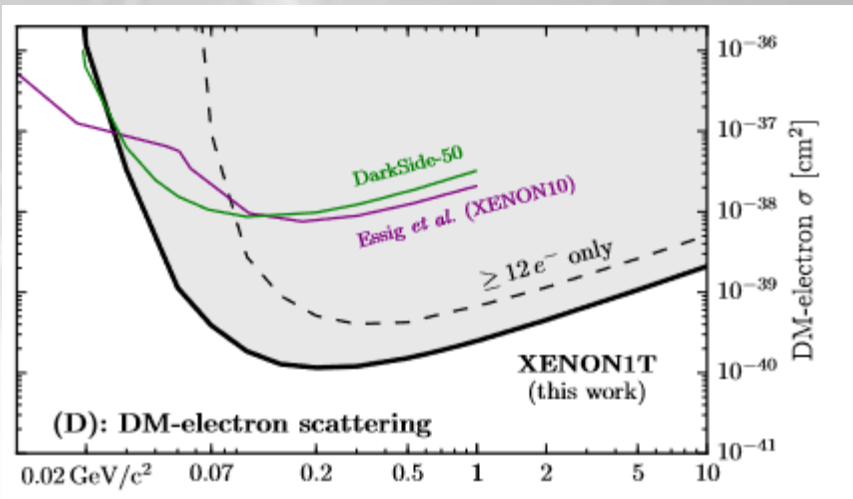


arXiv:1805.12562

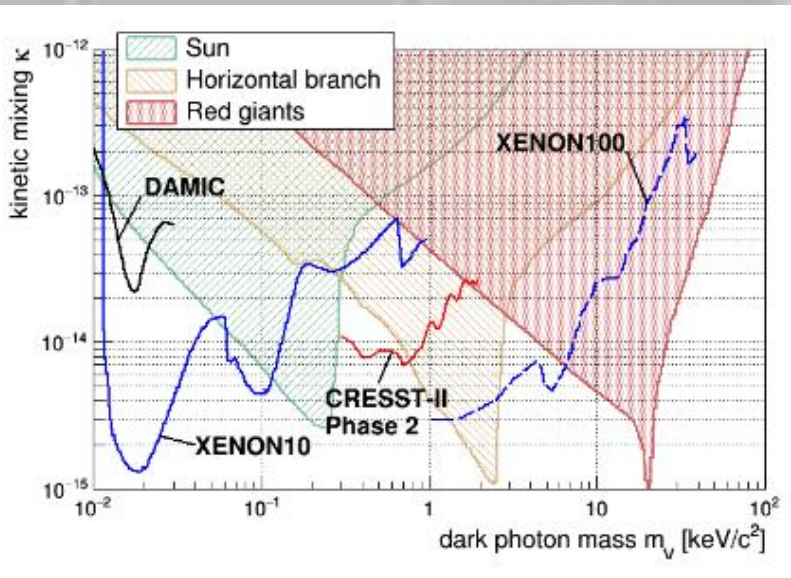


arXiv:2007.15622

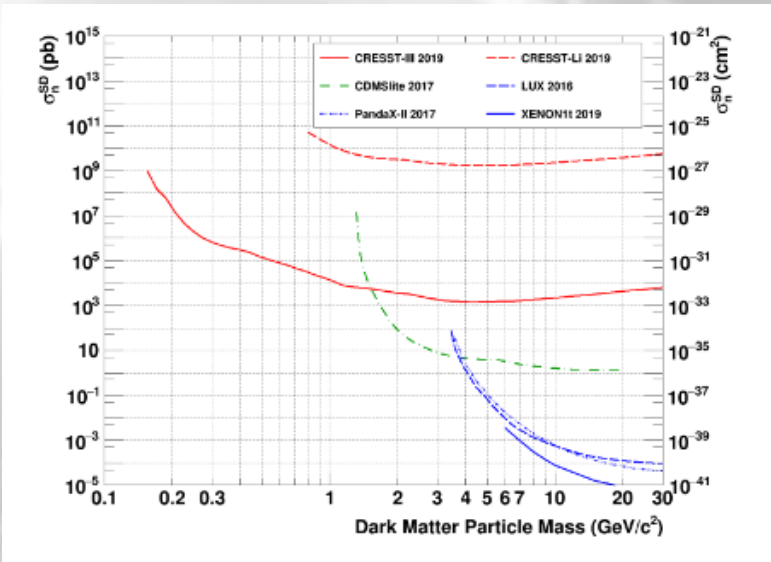
Even the parameter space is not uniquely defined

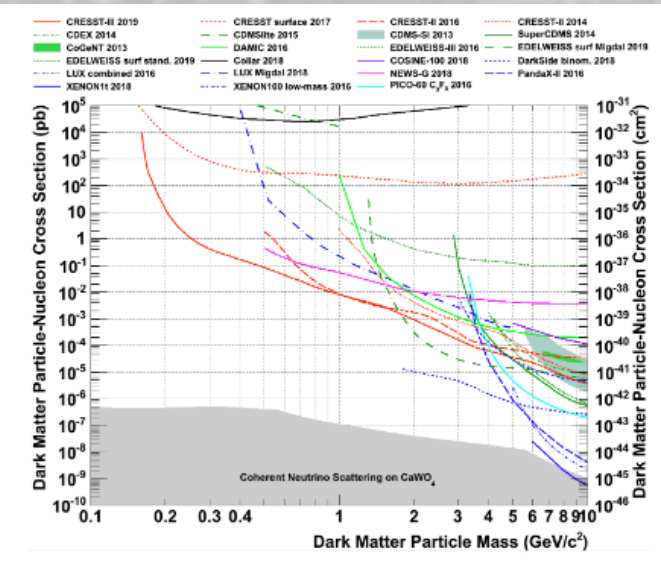


arXiv:1907.11485



arXiv:1612.07662



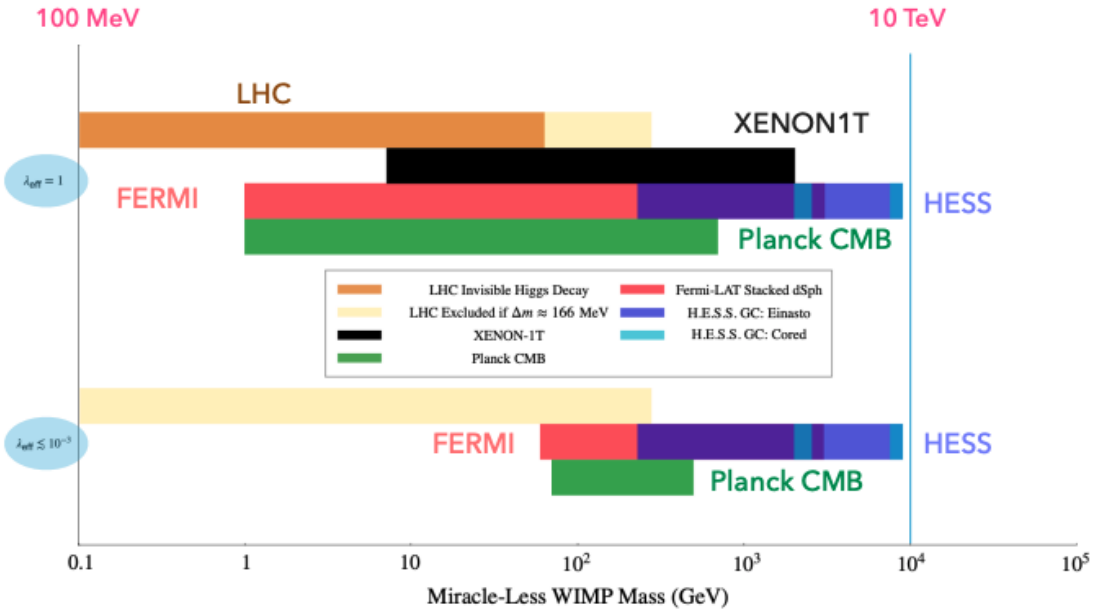


arXiv:1904.00498

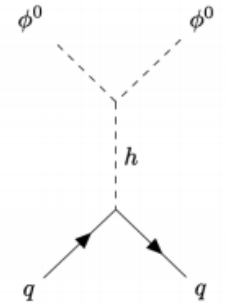
Including cross-disciplinary works makes complexity explode

COMPARISON WITH OTHER FIELDS: EXAMPLE

- Is Mirac-Less WIMP dark matter (abundance not set by freeze out during a standard cosmology) ruled out? Jason Arakawa, Tim Tait, arXiv:2101.11031



EW triplet scalar field



DD constraints depend of λ_{eff} (effective coupling through Higgs exchange)

The DMDC

- Provide support to the DM community to make the best and broadest use of experimental data and theory predictions.
- Provide support to the DM community for the development of software and methods
- Provide support to colleagues in the development of the most up-to-date view in the field.

The DMDC:
What we want to offer

- Long term data storage and catalog for dark matter data (with DOI)

- Support for collaborations to publish data in a usable and well documented format

- Support for phenomenologists for the usage of the data, algorithms execution and long term reproducibility of the results (DOI for the software and long term storage of the docker)

The DMDC:
How we want to realise

- Web interface for querying data
  (It comes together with the right BibTex)

- Web interface for uploading the data
  (Together with all the relevant literature and the option to generate a DOI if not already available)

- HPC backend for model testing with public code, DOI assignment and long term storage of software containers for reproducibility

- User‘s support for execution of newly developed methods

The DMDC Conclusions

There is an utter desire for sharing of data and knowledge

We want to offer a platform and the relevant technical infrastructure/manpower to make it happen and support a virtuous circle