

DMP title

Project Name Environment-Assisted Electron Capture in Bose-Einstein Condensates - DMP title

Project Identifier 1232922N

Grant Title 1232922N

Principal Investigator / Researcher Axel MOLLE

Project Data Contact axel.molle@kuleuven.be

Description This project seeks to calculate interatomic coulombic electron capture by a barium (II) cation in a Bose-Einstein condensate of neutral rubidium atoms. Obj. 1: introduce a new robust computational approach to many-partner-assisted electron capture, Obj. 2: quantify the aspect of large interatomic distance (characteristic experimental condition), Obj. 3: quantify the effect of coherence (characteristic experimental condition), Obj. 4: quantify the yet unknown process contribution of bound-bound excitations (new process), and Obj. 5: quantify the environmental assistance by polarisation (new process).

Institution KU Leuven

1. General Information

Name applicant

Axel Molle

FWO Project Number & Title

Environment-Assisted Electron Capture in Bose-Einstein Condensates, 1232922N

Affiliation

- KU Leuven

In collaboration with Colorado School of Mines, Golden, USA

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data
- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

This research will use proprietary and open-source scientific software and will develop open-source

software. The scientific software uses and generates mostly ASCII files as input and output files which contain program-specific keywords and text and tables of numbers. Some computational results may be visualized in digital images or videos. The relevant results of this project will be presented in manuscripts which are submitted for publication in leading research journals. Raw data

of calculations may be published as supplementary material to scientific articles. Preprints of all publications will be deposited into open-access archives. No personal data is collected or stored.

Type of data	Format	Volume	How created
data tables	text files, csv, txt	1 GB	extracted from software output or other works
software specific input/output files	text files, csv, txt, py, sh, awk	100GB	with appropriate scientific software
scientific software source code	text files in C/C++, Python, Fortran	1GB	written and tested source code
scientific manuscripts	text files, .tex, .bib, pdf	1GB	written with aim of publication of research data
documentation	text files, .tex, .html	1GB	
scientific graphs and illustrations	eps, png	100 GB	created with python or tex from data tables or source code
scientific illustrative videos	.mov, .mp4	100 GB	created with python or tex from data tables or source code

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

No personal data will be used

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

No ethical issues are expected to arise

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- Yes

Q-Chem is a proprietary quantum chemistry software.

Development of software source code within the Q-Chem software may have potential for tech transfer and valorisation

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- Yes

Non-disclosure agreement on Q-Chem program source code.

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

created data tables used for scientific graphs will be documented in a ReadMe file. If processed from raw output data files specific to used scientific software, the method of extraction of data will be sufficiently documented within the associated published manuscript or manuscript's supplementary information.

data tables created by other authors outside the project will be referenced with appropriate reference in public domain, including bibliographical references of original method of data creation or with contact details of original author who provided the data upon request.

software source code , software tests will be documented within source code files. The doxygen standard will allow to provide exporting documentation into html or pdf.

manuscripts and associated supplemental information will state the methods, formulae and programs and their versions used clearly in a fashion that ensures reproducibility.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- Yes

the KU Leuven Research Data Repository

<https://rdr.kuleuven.be/>

will be used and its associated metadata standard.

Where applicable, data repositories provided with the publisher of a related scientific article may be used and their respective metadata standards.

5. Data storage and backup during the FWO project

Where will the data be stored?

The software, manuscripts, and associated supplemental information and notes being in process of development are being stored on personal hard drives and backed-up in a personal private revision-controlled repository. To allow co-authors to contribute, a revision-control is duplicated through subscription with Overleaf.com for scientific manuscripts. The newly-developed software can be distributed additionally via open-source archives similar to QChem Quantum Chemistry modules and the Open Source Matrix Product States code. Particularly through the AMOS gateway for Atomic Molecular and Optical Scientific codes. Completed Manuscripts are stored as preprints in open-access archives, published manuscripts are stored with the publishing journal. Cleaned-up raw output and input from calculations is to be archived for ten years at the Department of Chemistry at the KU Leuven in accord with their Research Data Management policy and with the KU Leuven Research Data Repository RDR.

Simulation data is being stored on the respective high performance computer cluster used (KU Leuven, Vlaamse SuperComputer, Colorado School of Mines)

How is backup of the data provided?

Regular back up on personal hard drives, revision-controlled backup of development data on private repository and Overleaf or QChem server. Simulation data is being stored with automatic daily backup on respective high performance computer clusters used for their production with related input files (KU Leuven, Vlaamse SuperComputer, Colorado School of Mines).

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

What are the expected costs for data storage and back up during the project? How will these costs be covered?

Overleaf.com 168.00 EUR/year

Private Online Repository 29.88 EUR/year
External hard drive 70 EUR/TB
covered from bench fee
KU Leuven RDR free of charge
QChem development repository free of charge
AMOS gateway free of charge
university servers free of charge during project

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

No personal data is being used. No other data requiring extra security protocols is being used. Data is secured with password protection and controlled access permissions. Remote access is secured through Virtual Peer Networks and Secure SHell access as well as Secure File Transfer Protocols.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

acquired existing data from other projects and authors or the public domain will potentially not be archived.

Intermediate temporary simulation or analysis data will not be archived, neither will temporary or intermediate files of manuscripts or other documents of temporary nature.

Raw input and raw output files used to extract information given in scientific manuscripts will be retained and archived. Extracted data tables directly necessary to produce scientific graphs for published manuscripts will be retained and archived, if possible also stored and provided in the public domain or with the published manuscript.

Documentation and metadata will be retained and archived with the project

Unused raw output files of completed work packages deemed immaterial to further investigation and not used during the work package may be destroyed if deemed easily or low-cost reproducible.

Where will the data be archived (= stored for the longer term)?

The data will be stored on the university's central servers DIRAC and RDR (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

The investigators will work with staff of the KU Leuven Libraries to determine what to archive.

Essential data will be archived additionally if possible with the publisher of the related manuscript and/or with the preprint on a self-archiving server like arxiv.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

The Dirac cluster and RDR repository, the main data storage sites are expected to be usable free of charge.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- Yes. Specify:

Data acquired from third parties may be subject to sharing restrictions and/or copyright restrictions.

Program code developed within proprietary software may be subject to non-disclosure and other agreements restricting or preventing direct sharing.

Which data will be made available after the end of the project?

The documented cleaned-up dataset related to published work of the project will be made available in the public domain primarily through KU Leuven RDR and/or with the manuscript as indicated in section 6.

Where/how will the data be made available for reuse?

- In an Open Access repository
- In a restricted access repository

Data will be made available with the published manuscript, with the preprint server (arxiv) and with KU Leuven RDR.

Source code will be made available with the AMOS gateway and with GitHub.

When will the data be made available?

- Upon publication of the research results

The data will be made available upon publication of research results.

Prior sharing is possible upon request.

Who will be able to access the data and under what conditions?

Open access (CC BY-SA, or AGPL-3.0) to the data and software is planned where not prohibited by third party agreements.

What are the expected costs for data sharing? How will the costs be covered?

No expected extra costs for data sharing.

8. Responsibilities

Who will be responsible for data documentation & metadata?

The PI, Axel Molle, will be responsible for data documentation & metadata.

Who will be responsible for data storage & back up during the project?

The PI, Axel Molle, will be responsible for data storage & back up during the project.

The IT responsible for the respective high-performance computer cluster will be responsible for daily backup of data on the cluster.

Who will be responsible for ensuring data preservation and reuse ?

The PI, Axel Molle, will be responsible for ensuring data preservation and reuse.

Data published with the KU Leuven Research Data Repository will become subject to responsibility of the KU Leuven library.

Who bears the end responsibility for updating & implementing this DMP?

The PI bears the end responsibility of updating & implementing this DMP.