Initial-DMP

Project Name My plan (C1-C2-IDN DMP) - Initial-DMP

Project Identifier C14/21/083

Grant Title C14/21/083

Principal Investigator / Researcher Michel Houssa

Project Data Contact Lino Pereira, +32 16 32 01 76, lino.pereira@kuleuven.be

Description This project consists of theoretical and experimental research, in the field of condensed matter physics. It involves the design, fabrication and characterization of solid-state systems, aiming to create and understanding non-trivial band topology and electron correlation phenomena in 2-dimensional and van der Waals materials. During the project, experimental and simulation data will be collected and analyzed in order to answer the following research questions: Which electron correlation phenomena and non-trivial band topology (e.g. ferromagnetism and quantum anomalous Hall effect) emerge in these materials? Which phenomena (e.g. Majorana states) emerge from interfacing such materials? Can these systems be driven through different quantum states in device architectures?

Institution KU Leuven

1. General Information
Name of the project lead (PI)

Michel Houssa (KU Leuven)

Internal Funds Project number & title

ZKE0454 - C14/21/083, Electron correlations and non-trivial band topology in 2D materials

2. Data description
2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

Primary data (generated by researcher and research group):

1) Analogue data:

1.1) Samples (bulk solids and films on a substrate).
1.2) Laboratory logbooks.
1.3) Researchers' notebooks.

2) Digital data (total estimated volume of tens to hundreds of GB/year):

2.1) Experimental data (as directly measured): ASCII data files containing numerical data (measurement output) and text metadata (describing experimental/instrumental parameters). Different formats (.txt, .dat and .cvs), depending on instrument used for the measurements.
2.2) Processed and analyzed experimental data: text documents, spreadsheets and graphical representation of data, used for (or resulting from) data analysis.
2.3) Electronic laboratory logbooks.

2.4) Computational simulation data, generated through scripts and codes. Intermediate raw data will generally not be saved. Instead, the scripts from which the data was generated (and can be regenerated) will be saved. I particular, the following data types will be collected: original input data (as scripts used by various comptational softwares); essential output data (atomic configurations, simulation outputs, etc.); list of computational settings in order to enable reproduction of the calculations.

3. Ethical and legal issues
3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

No personal data will be used in this project.

3.2. 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 ethical issues are encountered in this project.

3.3. Does your research 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?

Although tech transfer and valorisation are not direct goals of this project, if such an opportunity presents itself, it will be discussed among the staff and collaborators involved in the project. The conclusions of that discussion will be appended to this data management plan.

3.4. 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 regarding reuse and sharing are in place?

No 3rd party restriction.

4. Documentation and metadata

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

During each experiment, a detailed logbook will be used (different logbooks for the different experimental setups). These logbooks will contain the date, a brief description of the performed experiment, the parameters used for each measurement, as well as the names of all the saved files.

The names of the files will be structured in a comprehensible way:
system studied/date/main parameters used

In addition, data will be stored in a folder per experimental setup, the type of investigated system and the corresponding date. In this way, by tracking the corresponding logbook notes, each file can be easily found on the local computers controlling the setup and on the server of the laboratory.

The analysis files will contain notes describing the analysis procedure and mention which original data files are included. A readme file describing the goal of the experiment and the analysis procedure will be stored in the folder where the data is saved.

For the computational simulation data, descriptions of the scripts/codes will be provided. Where needed, tutorials will be developed in order to present newly developed codes.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

Regarding the experimental data, in the relevant research field, there is no formal metadata standard. However, the standardized steps described in this section 4 will ensure that the data is easy to find and reuse.

Regarding the computational simulation data, most of the meta-data will be related to entries in the Zenodo repository.
The following metadata will be stored for each entry in the Zenodo repository:

Metadata	Category	Additional comments
Type of data	Mandatory	
DOI	Mandatory	If not filled, Zenodo will assign an automatic DOI. We will keep the same DOI if the document is already identified with a DOI.
Responsible / author(s)	Mandatory	
Title	Mandatory	
Publication date	Mandatory	
Date of repository submission	Mandatory	
Version	Mandatory	
Description	Mandatory	
Keywords	Mandatory	Frequently used keywords.
Size	Mandatory	The approximate size.
Access rights	Mandatory	Open Access. Other permissions can be applied, when appropriate.
Terms of Access Rights	Optional	We will open the data under Attribution, ShareAlike and Non Commercial Licenses (Creative Common Licenses).
Funding	Mandatory	The KU Leuven grant will be acknowledged.

5. Data storage and backup during the project
5.1. Where will the data be stored?

For the experimental data, all physical logbooks will be stored in a cabinet next to the respective experimental setups. Once the each logbook is full, it will be stored in a central storage cabinet of the institute, for a period of 15 years after the last registered experiment. The electronic data, including the logbook copies, raw data and analysis files, as well as reports, will be saved on local computers. The data stored on these computers is backed-up via the MS OneDrive Business service and on local servers. All researchers involved in the project will have access to the corresponding folder.

For the compuational simulation data, all the data (digital) will be stored on local computers and backed up on local servers and via the MS OneDrive Business service.

5.2. How will the data be backed up?

The data will be stored on local computers and servers with automatic back-up procedures on local servers and via the MS OneDrive Business service.

5.3. 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. The amount of data that will be generated in this project should not exceed a few hundred GBs, which is small enough to be stored in local computers, on local servers and via the MS OneDrive Business service. In particular, for the compuational simulations, larger intermediate raw files are typically removed after the post-processing and/or the publication of the results. Therefore, the storage requirements related to this project will not be extensive.

5.4. What are the expected costs for data storage and backup during the project? How will these costs be covered?

The involved research groups have their own data servers and backup systems, which is complemented by the MS OneDrive Business service. The costs are minimal, related to the

replacement of faulty hard-disks or small upgrades to the current local storage facilities. The costs related to data storage are envisioned in the budget (0.8euro/GB per year), and will be covered from the available consumables budget.

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

The experimental data generated during the project will be systematically transferred to the local servers and MS OneDrive Business, with restricted access (managed by the local IT team). Only the (co-)promotors and involved researchers have access to the shared folders where the data, analysis files and reports will be stored. Also, credentials are required to log in to local computers in the laboratories. The data will be shared within the campus network only between PIs and members of the research team (PhD students and postdocs). If the need to share with external collaborators arises, we will use the OneDrive Business tool as a secure sharing tool.

6. Data preservation after the end of the project
6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

All the generated data will be stored for at least 10 years after the project ends, with the exception of the intermediate raw data genated during the computational simualtions, which will generally not be retained. Instead, we will retain the scripts from which the data has been generated.

6.2. Where will these data be archived (= stored for the long term)?

All data will be stored on the local servers back-up servers owned by and located at the KU Leuven, for at least 10 years.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

The system and cost allocation described in section 5.4 will also be used for data preservation after the project ends. Costs are expected to be negligible and will be covered by other running projects from the involved groups (in case of no running projects, accumulated reserves will be used).

7. Data sharing and re-use
7.1. 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 or because of IP potential)?

There are no legal restrictions or restriction related to IP potential.

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

All data will be made available, for example if requested by the editor or publisher of a scientific journal or upon request of an individual (e.g. a researcher who intends to reproduce an experiment).

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

- In an Open Access repository
- In a restricted access repository
- Upon request by mail

Upon request and after the agreement of the projects promotors, all data can be made available on an open repository (e.g. Zenodo for data), for example if requested by the editor or publisher of a scientific journal or via restricted access upon request of an individual (e.g. a researcher who intends to reproduce an experiment).

7.4. When will the data be made available?

- Upon publication of the research results

After the research results have been published and after agreement of the (co-)promotors (with a possible embargo time no longer than one year after the publication of the research)
Source codes/scripts will be made available over the duration of the project, after an internal

evaluation on the possible valorization of newly developed codes/methodologies.

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

All involved researchers will have access to all the data. Upon request and after agreement of the (co-)promotors, access to the data can be granted to other individuals/parties.

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

There will be no costs for data sharing. Since the volume of most files is not too large, the OneDrive Business tool provided by the KU Leuven, or otherwise freeware, will be used to transfer and share the files.

8. Responsibilities
8.1. Who will be responsible for the data documentation & metadata?

The promoter and co-promoters of the project.

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

The promotor and co-promoters of the project, together with the IT responsibles of the research groups, who are responsible for the implementation of the storage and back up.

8.3. Who will be responsible for ensuring data preservation and sharing?

The promotor and co-promoters of the project

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

The end responsibility for updating and implementing the DMP is with the promotor.