DMP

Project Name Magnetoseismology of accretion disks: towards the full spectrum for turbulent disks - DMP
Grant Title 11J2622N
Principal Investigator / Researcher Nicolas Brughmans
Institution KU Leuven

1. General Information

Name applicant

Nicolas Brughmans

FWO Project Number & Title

11J2622N - Magnetoseismology of accretion disks: towards the full spectrum for turbulent disks

Affiliation

- KU Leuven

2. Data description

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

- Generate new 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).

Type of data	Format	Volume	Origin
LEGOLAS spectra	.dat (binary)	10 - 50 GB	Solving large eigenvalue problem of 1D equilibria
FINESSE equilibria	Binary	100 - 500 GB	Finite element solutions to initial value problems
PHOENIX spectra	Binary	10 - 50 GB	Finite element and spectral methods, solving large eigenvalue problem of 2D equilibria
MPI-AMRVAC simulations	.dat (binary), .vtu	500 - 1000 GB	MHD simulations of accretion disks

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

Privacy Registry Reference:
Short description of the kind of personal data that 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

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?

- No

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?

- No

4. Documentation and metadata

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

LEGOLAS/MPI-AMRVAC: the data and output types of both codes are well documented on their respective websites. The data generated in every simulation will be labeled per title according to the set-up/model used. A .txt file describes the labels more clearly, also linking to the .par file containing all information regarding the initial set-up, numerical methods, ... All post-processing done on the data will be clear and repeatable with the use of Jupyter-Notebook.

FINESSE/PHOENIX: the data types of these codes are as of now not well described and usable apart from the original authors. We will document them with dedicated websites (following the example of LEGOLAS/MPI-AMRVAC) and perhaps link to a more user-friendly code as Python to allow for post-processing. Furthermore, the data will also be labeled according to the set-up/model used, which will be explained more in detail in a joint .txt file.

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.

- No

.txt files containing information (set-up, model, numerical techniques) on the generated output, which is labeled by title, by linking to the .par file used for each simulation. The .dat file is explained on the respective website of the code.

5. Data storage and backup during the FWO project

Where will the data be stored?

Data will be stored on the group's desktops, which are managed by ESAT, and hard-drives.

How is backup of the data provided?

The ESAT-run system automatically backs up files up to 200 GB, which would include all work performed with LEGOLAS and FINESSE/PHOENIX. Larger simulations performed with MPI-AMRVAC can be backed-up on our hard-disks. Moreover, by storing the .par and mod_usr.t files used for our simulations, every result could be replicated exactly. These set-up files will be stored on the system managed by ESAT and after publication of the results will become available for the general public on the MPI-AMRVAC GitHub website, which contains many other set-ups for test problems and previously published results. The new equilibria implemented in LEGOLAS can similarly be merged into the open-source code. It is the intention to undertake a similar effort in turning the pair FINESSE/PHOENIX open-source, where the set-ups implemented in this project would then be included as test cases in the code.

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 assigned desktop has more than 3 TB of dedicated memory available, part of which is

backed-up automatically by ESAT. Back-up of simulation files exceeding the ESAT-limit will be performed on the available external hard-drives of the department.

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

If needed, the bench fee can be used for the purchase of additional external hard drives. The costs connected to the system maintenance and back-up by ESAT are covered by the department.

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

The desktops connected to the ESAT system and their contents stored on the servers, are only accessible by authorised persons through SSH and two-factor authentication. Simulation set-up files stored on the GitHub can only be modified by individuals with access, which would be people contributing to the MPI-AMRVAC and LEGOLAS codes.

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, ...).

All data will be stored for five years. Data will be archived on external hard drives so that the internal memory of the desktop can be reused by new team members after the project has been concluded. The simulation set-up files will be stored even longer as tests in the GitHub repository of their respective codes.

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

Set-up files stored on GitHub suffice to exactly recreate the simlulations performed during the project. By maintining them as test cases in the open-source codes, they will effectively be archived on the longer term.

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

The costs of hard-disks to store data on the longer term (if additional ones need to be purchased) will be covered from the project bench fee during the project. GitHub freely hosts the open-source codes.

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)?

- No

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

The set-ups for all simulations and calculations will be made open-source, from which anyone could exactly rerun the original simulations. The datafiles used to generate figures for our publications are stored so that these could be shared upon request, together with selected snapshots and videos (initial and fininal snapshots, remarkable features). Moreover, the post-processing routines containing image/video settings and calculations involving datasets are maintained as well to be able to exactly recreate any result featured in our papers.

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

- In an Open Access repository
- Upon request by mail

The files stored in the codes will be Open Access on GitHub. Other files used in our publications will be made available when requested.

When will the data be made available?

- Upon publication of the research results

After each planned phase in the research project has lead to an accepted publication, which has been mage public (e.g. on Arxiv) we will upload the set-up files to GitHub.

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

Anyone interested is able to inspect and replicate the simulations from the files on GitHub. Anyone contacting the authors of each publication could furthermore gain access to selected snapshots used in the publication and tools that were used to perform analysis of the results.

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

There are no expected costs.

8. Responsibilities

Who will be responsible for data documentation & metadata?

The grantee of this fundamental research project will be responsible for documenting and structuring the gathered data.

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

The grantee of this fundamental research project will be responsible for storage and back-up for the generated data during this project, together with the system managers of ESAT for data uploaded to the servers.

Who will be responsible for ensuring data preservation and reuse ?

The grantee and his supervisor will be responsible.

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

The grantee is responsible for updating and implementing this DMP.