Advances and applications of symmetric tensor decomposition

A Data Management Plan created using DMPonline.be

Creator: Daniele Taufer

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / **URL:** 12ZZC23N

ID: 193891

Start date: 01-10-2022

End date: 30-09-2025

Project abstract:
My project revolves around algebraic and numerical techniques applied to symmetric tensor decomposition. These objects have been attracting considerable attention in the last decades as they provide efficient tools for managing large amounts of data. Their decomposition is often employed for optimizing their computational use or for extracting hidden information from the considered systems.

The approach I intend to pursue is twofold.
From a geometrical perspective, I aim at addressing such a decomposition problem by means of apolarity techniques, which ultimately rely on the solution of ideals with a prescribed structure. On the other hand, I target the tensor decomposition over discrete structures, such as finite fields, where efficient local techniques may serve the purpose of eventually describing the global structure of the considered tensors.

The main goal of this research line is to produce novel effective results for tensor decomposition that may impact both the theoretical and the applied communities. Furthermore, such results are planned to be concretely implemented and distributed for providing the tensor community with stable and reliable routines.

Last modified: 17-01-2023

Created using DMPonline.be. Last modified 17 January 2023

1 of 5

Advances and applications of symmetric tensor decomposition FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or reused	Digital or Physical	Only for digital data Digital Data Type	Only for digital data Digital Data format	Only for digital data Digital data volume (MB/GB/TB)	Only for physical data Physical volume
Papers	Research published papers and preprints	Generate new data	Digital	Compiled/Aggregated data	.pdf	< 1GB	NA
Programs	Mathematical programs for computational testing or formal verification	Generate new data	Digital	Software	other: .m, .m2, .jl	< 1GB	NA

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

NA

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

NA

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

NA

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

NA

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

NA

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

NA

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

The research paper will be produced and published after private investigation and computational testing, and will always enclose either theoretical or computational proofs. In the first case, no documentation is needed or provided.

The relevant computational results or proof obtained in mathematical computational softwares (such as Magma, Macaulay2 and Julia) will be pubblicly distributed and will always involve proper code explanation (in the paper or directly in the code file), together with an overview file (README.txt).

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

The papers will enclose standard keywords and MSC classification codes, when appropriate.

3. Data storage & back-up during the research project

Where will the data be stored?

The data will be stored and shared online, on the appropriate platforms (ArXiv, journal websites, GitHub & GitLab), and KU Leuven repositories (Lirias).

How will the data be backed up?

The data will be backed up within the research group, on personal computers and physical memories.

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.

Created using DMPonline.be. Last modified 17 January 2023

3 of 5

- Yes

The amount of produced data is low, and will be easily handled by private machines and public storages.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

All the data will publicly shared in a read-only form, while the access of the source data will only be granted to the researchers involved in the project.

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

No costs of such a type are expected.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All the data will be stored for 10 years according to KU Leuven RDM policy, and publicly distributed online indefinitely.

Where will these data be archived (stored and curated for the long-term)?

The data will be stored and shared online on platforms such as ArXiv, journal websites, GitHub & GitLab and KU Leuven repositories (Lirias).

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

No such costs are expected.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

NA

If access is restricted, please specify who will be able to access the data and under what conditions.

NA

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

NA

Where will the data be made available? If already known, please provide a repository per dataset or data type.

arxiv.org and journal websites

When will the data be made available?

Immadiately after their publications

Which data usage licenses are you going to provide? If none, please explain why.

They will usually be granted CC BY: Creative Commons Attribution

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

All the published papers will be granted a DOI.

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

There might be minor publishing costs. If it is the case, they will be covered by the personal bench fee

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Daniele Taufer

Who will manage data storage and backup during the research project?

Daniele Taufer

Who will manage data preservation and sharing?

Daniele Taufer

Who will update and implement this DMP?

Daniele Taufer