Towards a global theory of orthogonal polynomials and correlation kernels for non-Hermitian random matrices

A Data Management Plan created using DMPonline.be

Creator: Aron Wennman

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / **URL:** G0DDD23N

ID: 207849

Start date: 01-01-2024

End date: 31-12-2028

Project abstract:

Orthogonal polynomials are classes of polynomials subject to certain orthogonality relations, often in weighted L2 spaces on the line, the plane or along the unit circle. Just as Fourier series may be used to express periodic functions in a simple way, orthogonal polynomials are used to describe the solutions to a variety of problems in mathematics.

I am interested in orthogonal polynomials which appear in non- Hermitian random matrix theory, where they describe the correlation structure of an important instance of the 2D Coulomb gas. The Coulomb gas is fundamental in statistical mechanics, where it constitutes one of the simplest "toy models" of matter which is not confined to a lattice. This random matrix model also appears naturally in quantum mechanics, data science, and in the stability analysis of ecosystems, motivating its study also from an applied perspective. While some aspects of 2D Coulomb gases are well understood, intriguing questions concerning tunneling phenomena, emergence and resolution of singularities, and crystallization in low- temperature regimes remain mysterious.

The goal of my project is to answer open questions about beautiful patterns formed by the zeros of the orthogonal polynomials, and about the behavior of the Coulomb gas in cases where particle clouds form singularities or collide with obstacles. The research will be carried out at KU Leuven, where many break-throughs have been made in this research field.

Last modified: 11-06-2024

Towards a global theory of orthogonal polynomials and correlation kernels for non-Hermitian random matrices
DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

Towards a global theory of orthogonal polynomials and correlation kernels for non-Hermitian random matrices
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Towards a global theory of orthogonal polynomials and correlation kernels for non-Hermitian random matrices
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

Within the scope of the project, we will generate three types of data:
1) manuscripts containing new mathematical theorems and proofs (.tex and .pdf-documents),
2) computer code for numerical simulations (Mathematica, Python, C),
3) output from the computer code, primarily in the form of figures (.png, .eps, etc)
We will not collect any types of data, nor make use of data generated by others.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

1. Responsible person is the PI (Aron Wennman)
2. During preparation of manuscripts, the data will be stored both locally on the storage solutions of the research unit, and on KU Leuven OneDrive. Following completion of the manuscripts, they will be published in peer reviewed journals and made publicly available on an open access preprint repository (arXiv.org, the discipline-specific standard).
3. The discipline standard is not to publish computer code used for numerical simulations and generation of figures. If deemed of scientific value, such code will instead be kept and stored on the PI's OneDrive, and in addition be made available on his personal website (both during and after the project).
4. Since we only expect a few megabytes of text data/code and a relatively small number of images, the standard allocated space on OneDrive will be sufficient for the storage purposes.

What’s the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

I do not intend to deviate from the principle.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

N/A

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

N/A

Towards a global theory of orthogonal polynomials and correlation kernels for non-Hermitian random matrices
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.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
Manuscripts	Computer code for typesetting of manuscripts	Generate new data	Digital	Written text	.tex, pdf	<1GB	N/A
Numerical simulation	Code for numerical simulation	Generate new data	Digital	written code	.m, .py, .txt, .nb	<100MB	N/A
Generation of figures	Code for generation of figures	Generate new data	Digital		.m, .py, .txt, .nb	<100MB	N/A
Figures	Output from Numerical simulation and Generation of figures	Generate new data	Digital		.eps, .png, .ipe,	<1GB	N/A

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:

N/A

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

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

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

None foreseen

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

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

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).

N/A

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.

- No

N/A

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

Where will the data be stored?

Prior to publication, our data (consisting of manuscripts, code and images) will be stored on KU Leuven OneDrive

Permanent storage: on discipline-specific preprint repository
For unpublished material: long-term (5-year) storage on KU Leuven OneDrive

How will the data be backed up?

We will use the standard daily backup provided by KU Leuven ICTS.

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

N/A

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

None of our data contain sensitive information. However, security is guaranteed through the use of the KU Leuven ICTS data storage facility.

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

Due to the low data volume, we do not foresee any costs associated with data storage and backup. The necessary solutions are already provided by KU Leuven ICTS.

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 data which is deemed of scientific value but which are not published with open access (OA) in journals or on OA repositories will be kept indefinitely. Should the data (e.g., code for simulations/numerical experiments) be deemed of public interest, we will in addition strive to make the data freely available on the PI's personal website.

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

Primarily on open access repositories (arXiv), and on KU Leuven ICTS data storage solutions, i.e., KU Leuven One Drive.

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

Due to the small volume, no costs are foreseen.

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
- Other, please specify:

The data will primarily be posted to the arXiv (subject specific standard in mathematics). Parts of the data that do not reach sufficient scientific completeness to warrant publication shall either be available upon request, or posted to the PIs personal webpage.

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

N/A

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

N/A

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

The data will be made available on arXiv.org, the subject-specific standard for open access publication.

When will the data be made available?

Upon publication of the research results.

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

CC BY 4.0

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

A DOI will be added. This is done automatically for anything posted to the arXiv repository, which we will use.

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

None foreseen.

6. Responsibilities

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

The corresponding authors for each paper generated within the project will be responsible for data documentation and metadata. The PI (Aron Wennman) is responsible for informing the researchers of our principles, as set out in this DMP.

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

The corresponding authors for each paper generated within the project will be responsible for data storage and backup.

Who will manage data preservation and sharing?

The corresponding authors for each paper generated within the project will be responsible for data preservation and sharing.

Who will update and implement this DMP?

The PI (Aron Wennman) will be responsible for implementing and updating this DMP.