l

|  |
| --- |
| *(Arial, font size 10)* |

|  |  |
| --- | --- |
| **Funding programme:** | **Programmes for Project-Related Personal Exchange (PPP) from 2022 with** Brazil |
| **Programme objective/s (outcomes of the funding programme)[[1]](#footnote-2):** | |
| **Objective 1** | **Junior scientists have gained international research experience and undergone further training at an international level** |
| **Objective 2** | **Binational research cooperation has been promoted and can be used as a starting point for future cooperations** |
| **Results of the measures/activities of the programme (outputs of the funding programme)[[2]](#footnote-3):** | |
| **Result 1** | **There are joint research results** |
| **Result 2** | **International joint publications have been created** |

|  |  |  |  |
| --- | --- | --- | --- |
| **General information** | | | |
| Project name | Efficient statistical tools for networks and their applications | | |
| Applicant institution | Universität Leipzig | | |
| Those responsible for the project | Peter F. Stadler | | |
| Those responsible for the project (outside Germany) | André Fujita | | |
| Partner country/countries | Brazil | | |
| Partners (within and outside Germany) | University of São Paulo and Universität Konstanz | | |
| For follow-up application[[3]](#footnote-4):  Approval period of the last/current funding | From: Please specify a date  To: Please specify a date | | |
| Are there parallel funding streams and/or applications under other DAAD programmes in the context of this project application? | | Yes | No |
| If yes, under which? | Please specify | | |
| Are there any parallel funding streams and/or applications under any other funding programme provided by another funding organisation in the context of this project application? | | Yes | No |
| If yes, under which? | Please specify | | |

|  |
| --- |
| **For follow-up applications: Previous project progress** |
| Please describe the previous project progress (implementation of measures/activities and achievement of objectives). |
| Please specify |

|  |
| --- |
| **Project objectives, detailed project description and reference to results logic** |
| 1. State your project objectives (outcomes), which must be consistent with the programme objectives (outcomes) mentioned above, and describe the specialised content of the project. Explain with reference to the results logic which specific project results (outputs or results of the measures/activities) are used to achieve these project objectives (outcomes)[[4]](#footnote-5). 2. Touch upon the relevance of your project and ensure that you address all selection criteria in the programme description, which are listed again here:  * Relationship of the project to the programme objectives (as per the impact analysis structure) and results-oriented planning using indicators that meet the SMART criteria[[5]](#footnote-6). * The quality of the project (clarity of project objectives and methods) and scientific relevance of the project (topical nature of the subject matter and the project’s degree of innovativeness). * Appropriate involvement of junior scientists * Transfer of knowledge between the groups of researchers,   Value (subject-specific, institutional, interdisciplinary) created through the cooperation for both groups of researchers,  Scientific and, if applicable, industrial usability of the project results   * Feasibility of the research project (in particular: financial backing, preliminary work and further plans, adequate planning for trips abroad),   Project-related competence of both groups of researchers,  Complementarity of the groups of researchers in relation to the joint project (methodically, content-related, instrumentally, etc.)   1. Describe any potential risks in relation to the success of the overall project and how you will handle them.   Note:  The project objectives (outcomes) and intended results of the project’s measures/activities (outputs) must be entered in the project planning overview table in the form of results-oriented project planning. |
| **Outcomes**  This proposal aims to consolidate the relationship between all partner groups. New collaboration opportunities will be identified, especially in regard to junior scientists research network expansion. We will specifically initiate the development of computationally efficient statistical tools to analyze extensive empirical networks from a spectral distribution as well as a cycle-base angle.    **Outputs**  1. In person meetings in Brazil and Germany will enable all involved scientists to engage in topic specific discussions as well as get first hand impressions of the work and living environments for future collaborations and research stays. We plan to organize talks every year and two short courses/workshops to also allow other students and researchers to participate.  2. The methods developed here will impact several fields of science. As aforementioned, random networks are ubiquitous. Thus, they will be helpful to analyze chemical compounds, social interactions, metabolic pathways, neural networks, and the internet. We expect that the works generated in this proposal will have a high impact, given the widespread interest in random networks. Thus we expect that joint publications will strengthen the scientific footprint for all participating researches, especially young scientists, which will be advantageous for funding applications of their own.  3. All developed algorithms will be implemented as reference software packages e.g. in R or Python libraries. Collaboration in that regard will help junior scientists to develop a set of skills, from collaborative software development via open source platforms like GitHub to planning skills for software and project development as well as distribution and maintenance of open source software.  4. As the funding periods of DAAD ends after 2 years, while the PROBAL funding ends after 4, we will apply for a second round of funding towards the end of the German project time. Junior scientists will be tightly integrated in both, report composition as well as reapplication for DAAD funding, thus providing them with hands-on experience in funding acquisition and project execution.  WP1: Data consolidation  Initial effort will be focused on the curation of datasets. The latter build the basis for collaborative  investigation of properties which can be used to characterize large empirical networks. The partners will then work in parallel on complementary measures to quantify how close a graph is to exhibit a specific property.  WP2: Cycle distributions.  Cycles encapsulate semi-local information in a graph. Cycle bases provide well-defined, manageable cycle sets that can be computed efficiently. The length distribution of cycle sets such as the relevant cycles, i.e., those that are contained in at least one minimum cycle basis can be computed efficiently even without enumerating the sometimes exponentially large cycle sets. We therefore plan to use cycle distributions as complementary source of information. In particular we will investigate the relationships between Laplacian eigenvalues and cycle distribution and explore to what extent and which graph classes they can be used for alternative classification tasks.  Properties of cycle bases also characterize planarity and potentially other embedding properties. The latter is likely of particular relevance to application in brain-neworks and other networks that are embedded into low-dimensional (Euklidean) spaces. In the 2nd period we will therefore systematically investigate the contraints of embeddings on the cycle distributions.  WP3: Sampling-based approaches  Very large networks with often millions of nodes cannot by analyzed completely by ``expensive'' algorithms with running times that substantially exceed linear or N\log N time complexity. This is in particular the case for the computation of cycle bases. This can be remedied by two approaches: (i) If available fast heuristics can be used to compute approximations. (ii) Distributions of properties can be estimated by evaluating small samples of the graph. We will therefore devise systematic ways to sample suitable cycles sets. This is not trivial because cycles are not completely local. We will therefore define compact local induced substraphs whose cycle bases are informative on the global cycle distribution. The initial empirical study will inform a more in-depth mathematical analysis geared towards determining which graph classes are amendable to such a sampling approach. Again, low-dimensionally embedded graphs are good candidates, because subgraph with fixed diameter remain finite even then total network size diverges.  WP4: Asymptotic classification.  Similar to limits of spectral distributions in limit of graph size approaching infinite, limiting distribitions can also be defined for families of graphs. Expanding on previous work [Gu et al paper] we will explore the existence of limits for cycle distributions. Of particular interest is the question whether graph classes with the same limits of their spectral distribution can distinguished based on the distribution of cycles.  **Project description**  Unlike deterministic graphs, empirical networks are stochastic, either by the underlying processes that generate them or the measurement procedures. For example, brain networks are different even among healthy individuals.  Thus, many typical properties used to characterize graphs do not apply to large empirical networks. The reason is that they are not robust against the insertion or deletion of a small number of vertices or edges. Therefore, we need measures that quantify how close a graph is to exhibit a specific property, rather than the strict notion of isomorphism, which we rarely, if ever, be attained.  A common approach is determining graph invariants such as centrality measures (van den Heuvel and Sporns, 2013). However, graphs generated by different models may present similar centralities. Conversely, graphs generated by the same set of parameters may present a vastly different centrality measure. Thus, the analyses of empirical networks using methods grounded on deterministic graph theory seem to be inappropriate.  A natural solution would be to assume that graphs are generated by probabilistic processes and then develop statistical methods. Statistical approaches for random graphs are new, with few reports in the literature (Asta and Shalizi, 2015; Ginestet et al., 2017; Tang et al., 2017; Cerqueira et al., 2017; Ghoshdastidar et al., 2017; Schieber et al., 2017, Kolaczyk et al., 2019). One of the reasons is that graphs are challenging to study from a statistical viewpoint: graphs are objects composed of vertices and edges, i.e., they are not numbers. Our solution to this problem was to analyze the graph spectrum, which “codifies” information about the graph structure. First, we define the concept of graph spectral entropy. Then, based on it, we propose a parameter estimator (Takahashi et al., 2012), a model selection approach (Takahashi et al., 2012), tests to compare graphs (Takahashi et al., 2012, Fujita et al., 2017a, 2019), and a concept of correlation between vectors of graphs (Fujita et al., 2017b). They showed to be helpful in better understand new biological mechanisms, identify biomarkers, and find differences between controls and patients.  Here we will involve junior scientists that will gain international research experience and training in this theme. Dr. Fujita coordinates a FAPESP thematic project in network statistics, including dozens of graduate students and postdocs. We will focus on the training and internship of these people. Thus, this proposal complements the FAPESP thematic project.  The Stadler Lab in Leipzig has worked on several aspects of graph theory (BrianDavies et al., 2001; Gu et al., 2016; Hellmuth et al., 2009; Fritz et al., 2020). Thus, both teams are complementary: Fujita’s team specialized in statistics, in contrast, Stadler’s team in spectral analysis. Besides, this proposal involves two neuroscience teams, one in each country. We plan to use the methods developed here to analyze their datasets.  *Problems*  1. Empirical networks are usually massive. For example, it is estimated that the brain is composed of approximately 100 billion neurons. Thus, we cannot use current statistical approaches to analyze big data. The main reason is that we need to calculate the graph's spectrum, which is computationally expensive. Suppose a network is composed of n nodes. Then, the computational cost of naïve approaches, such as the diagonalization method, is O(n^3). Recently, Cantwell and Newman (2019) introduced a message-passing approach for the normalized Laplacian spectral density. Still, it requires computing matrix inversions and matrix-vector multiplications, which are computationally expensive.  2. We know that the spectrum has codified structural characteristics of the network. For example, by analyzing the Laplacian spectrum, we can obtain its diameter (Chung et al., 1989), the number of spanning trees (Bollobás, 1998), vertex covers (Chen and Jost, 2012), Kemeny's constant (Pan et al., 2018), and chromatic number (Sun and Das, 2020). However, we do not know the contribution of a node to a network's spectral distribution. In other words, although we identify differences in the networks' spectra, we cannot associate these differences with the networks' structures. Therefore, we cannot interpret them.  3. Studying the dynamic brain interaction network is vital to understand the brain's role in behavior. In the last decade, we witnessed the introduction of many methods to measure the presence or absence of dynamic interaction between brain parts. Nevertheless, little progress has been made on developing strategies to interpret and rigorously test the characteristics of the entire inferred interaction brain networks. Using the newly introduced graph spectra, we will develop rigorous statistical methods to compare ultra-high-dimensional networks. We expect that these methods will allow us to correctly interpret the results of massive interaction brain networks obtained with state-of-art methods.  **Relevance**  Social relevance: our algorithms will allow us to identify genes or brain regions associated with diseases, abnormal connectivity structures, and changes over time, space, and subjects. Then, it will lead to the development of drugs for treatment, biomarkers for diagnosis and prognosis, and a better understanding of the biological mechanisms.  Scientific relevance: our algorithms will be helpful in computer science, engineering, physics, and chemistry. E.g., for network feature extraction (Newman, 2018), low-rank approximation (Le et al., 2016; Luo et al., 2018), spectral clustering, and community detection (Newman, 2006). It also has applications in the dynamical systems theory (Porter and Gleeson, 2014), including structural phase transitions, such as percolation (Bollobás et al., 2010), localization (Martin et al., 2014), and detectability (Nadakuditi and Newman, 2012).  **Actions that will contribute to reaching the goals of this proposal**  The problem of the network’s spectrum-based tools is the computational complexity. We compute the spectrum using numerical methods such as the QR decomposition (Gander, 1980), the power iteration method (Booth, 2006), and the Jacobi eigenvalue algorithm (Sleijpen and van der Vorst, 2000). However, these approaches are O(n^3) (where n is the number of vertices of the network), making their applications limited to small networks.  For that reason, approximation methods have been proposed (Chung et al., 2004; Semerjian and Cugliandolo, 2004; Rogers et al., 2008; Metz et al., 2011; Nadakuditi and Newman, 2013; Newman et al., 2019; Newman, 2019; Cantwell and Newman, 2019). Recently, Cantwell and Newman (2019) introduced a message-passing approach for the (normalized) Laplacian spectral density using the local neighborhood of each node. Still, it requires computing matrix inversions and matrix-vector multiplications, which are computationally expensive (Cantwell and Newman, 2019). A variation of Newman’s algorithm is to compute the number of eigenvalues in an interval. We will extend both methods so that we will not compute matrix inversions and matrix-vector multiplications and use the fact that real-life networks are locally tree-like networks. Thus, we expect to obtain linear algorithms to approximate the spectral density. Then, we will plug in these algorithms to the parameter estimator, model selection, and tests to compare networks to improve the computational complexities.  To identify the nodes associated with a specific network’s eigenvalue, we will do as follows. The spectral density can be defined using a Dirac’s delta function. We plan to re-write the Dirac’s delta function as the limit of a Lorentzian distribution and obtain the contribution of each node to the spectral density. Then, we will perform some simulation studies using some already known eigenvalues (e.g., the eigenvalues -1, 0, and 1). After that, we will develop post hoc analyses to interpret the results obtained by our network comparison tests based on the differences among network spectra.  For all tools designed in this project, we have applications to functional brain networks. For example, we will identify which brain regions are associated with a disorder (e.g., autism) and how this brain region connectivity is altered to controls.  For all goals proposed in this project, we already supervise Ph.D. students and post-doc candidates for the internship in Brazil/Germany.  **Planned joint actions with the partners**  We plan to send Ph.D. students and post-docs in all four years of the project to maintain constant communication. PIs will interact mostly via videoconference over the year and visits once a year. PIs will discuss manuscript and other proposals design during the scientific missions every year. We also plan talks in every Ph.D., post-doc, PIs visit. We will organize short courses/workshops and invite students/researchers of other universities to participate remotely (via videoconference) in the second and fourth years. We plan to submit a proposal to the Research Group Linkage Programme (https://bit.ly/3f1zS4k) of the Alexander von Humboldt Foundation for further interaction between Brazilian and German groups. Dr. Fujita is an Alexander von Humboldt Fellow; then, he satisfies the minimum requirement to submit a proposal to this call. We also plan to build a mutual agreement between the University of São Paulo and Leipzig University to facilitate student internships.  **Forms of knowledge dissemination**  We will organize two short courses/workshops, one in each country, and talks every year. Students/researchers will be able to attend physically or via videoconference. We will publish scientific results in journals and conferences.  **Available infrastructure**  Prof. Fujita is the Brazilian coordinator. He has a fully equipped IT laboratory composed of dozens of high-performance workstations and computer servers. All of them sum up approximately 2TB RAM, 11TB HD, and 400 cores, which guarantees the development of this project.  Together with the Interdisciplinary Center for Bioinformatics, the Stadler group at Leipzig University has sufficient computing power for all high performance-computing tasks associated with the proposed research: 230 CPUs with a 6 TB RAM and 350 TB disk storage. We are currently expanding these resources further. In addition, the group has access to the High-Performance Computer Center in Dresden and the de.NBI cloud, maintained by the German Network for Bioinformatics Infrastructure  Takahashi's lab is in the Brain Institute. The institute has a state-of-art primate facility, a primate surgery room, level 2 bio-security rooms, two-photon microscopy, molecular biology, and viral core facilities. The Brain Institute also has access to a supercomputer. Takahashi's lab has access to a fully trained veterinarian, animal welfare specialist, and husbandry team.  Dr. El Hady is affiliated with Universität Konstanz and the Max Planck Institute (MPI) of Animal Behavior. The MPI has one of the most advanced facilities to study animal behavior in the world. It is equipped with virtual reality arenas where researchers can change the environment in real-time. We can record animal behavior using multiple sensors (high time-of-flight cameras, ultrasound microphones) simultaneously.  **Expected results**  *Academic results*  We will organize two workshops, one in each country. Also, we plan to build an academic agreement between the Bioinformatics graduate program at USP and Leipzig Universität.    *Bibliographic*  We intend to publish at least four papers in journals/conferences, one for each goal. We expect to publish other articles with applications of the developed tools with our collaborators in neuroscience.    *Scientific*  We will obtain faster (linear) algorithms to estimate the spectral distributions and statistical tools to analyze and interpret massive networks. We will compare ultra-high-dimensional brain networks between different subjects/treatment groups/experimental conditions.    *Formation*  We expect to train at least four Ph.D. and two post-docs. Also, three master students will defend their dissertations soon and will become Ph.D. students. They are also potential candidates to participate in this program.    *Technical*  We will incorporate the designed algorithms in the software statGraph. Consequently, statGraph will be able to analyze high-dimensional networks.  **Risk management**  For goals 1 and 2, if we find problems during algorithm development, we have collaborators that may help us. For example, Prof. Marie-France Sagot (Université de Lyon / INRIA) is a specialist in graph theory and its applications in biological data. Prof. Eric Kolaczyk (Boston University) is a specialist in statistics on networks. For goal 3, we will contact Prof. Janaina Mourao-Miranda (University College London), a computer scientist specializing in machine learning applications on neuroscience data. Profs. Takahashi and El Hady have already collected both neural signals and behavioral data.  **References**  ALEX, P, et al.  Partitioning sparse matrices with eigenvectors of graphs, **SIAM Journal on matrix analysis and applications**, 11, p. 430 – 452, 1990.  ASTA, DM, and SHALIZI, CR. Geometric network comparisons, **Proceedings of the Thirty- First Conference on Uncertainty in Artificial Intelligence. AUAI Press**, p. 102-110, 2015.  BOLLOBÁS, B. Random graphs: Modern graph theory, **Springer**, p. 215 – 252, 1998.  BOLLOBÁS, B, et al. Percolation on dense graph sequences, **The Annals of Probability**, 38, p. 150 – 183, 2010.  BOOTH, TE. Power iteration method for the several largest eigenvalues and eigenfunctions, **Nuclear science and engineering**, 154, p. 48 – 62, 2006.  BRIANDAVIES, E, et al. Discrete nodal domains theorems, Linear Algebra and its Applications, 336, p. 51 - 60, 2001  CAN, ML, et al. Optimization via low-rank approximation for community detection in networks, **The Annals of Statistics**, 44, p. 373 – 400, 2016.  CANTWELL, GT and NEWMAN MEJ. Message passing on networks with loops, **Proceedings of the National Academy of Sciences**, 116, p. 23398 - 23403, 2019.  CERQUEIRA, A, et al. A test of hypotheses for random graph distributions built from EEG data, **IEEE Transactions on Network Science and Engineering**, 4, p. 75-82, 2017.  CHEN, H and JOST, J. Minimum vertex covers and the spectrum of the normalized Laplacian on trees, **Linear algebra and its applications**, 437, 1089 - 1101, 2012.  CHUNG, FRK. Diameters and eigenvalues, **Journal of the American Mathematical Society**, 2, p. 187 – 196, 1989.  CHUNG, F, et al. The spectra of random graphs with given expected degrees, **Internet Mathematics**, 1, p. 257 – 275, 2004.  FA-YUEH, W. Theory of resistor networks: the two-point resistance, **Journal of Physics A: Mathematical and General**, 37, p. 6653, 2004.  FRITZ, A, et al. Cograph editing: Merging modules is equivalent to editing P\_4s, The Art of Discrete and Applied Mathematics, 3, 2, 2020.  FUJITA, A, et al. A statistical method to distinguish functional brain networks, **Frontiers in Neuroscience**, 11, 1, 2017a.  FUJITA, A, et al. Correlation between graphs with an application to brain network analysis, **Computational Statistics & Data Analysis**, 109, p. 76 – 92, 2017b.  FUJITA, A, et al. A semi-parametric statistical test to compare complex networks, **Journal of Complex Networks**, 8, p. cnz028, 2020.  GANDER, W. Algorithms for the QR decomposition, **Research Report**, 80, p. 1251 – 1268, 1980.  GHOSHDASTIDAR, D, et al. Two-sample tests for large random graphs using network statistics, **Conference on Learning Theory. PMLR**, 2017.  GINESTET, CE, et al. Hypothesis testing for network data in functional neuroimaging, **Ann. Appl. Stat.,** 11, p. 725-750, 2017.  GU, J, et al. Spectral classes of regular, random, and empirical graphs, **Linear Algebra and its Applications**. 489, p. 30 - 49, 2016.  HAIYAN, C and FUJI Z. Resistance distance and the normalized Laplacian spectrum, **Discrete Applied Mathematics**, 155, p. 654 – 661, 2007.  HAO, C and JÜRGEN, J. Minimum vertex covers and the spectrum of the normalized Laplacian on trees, **Linear algebra and its applications**, 437, p. 1089 – 1101, 2012.  HELLMUTH, M, et al. Approximate graph products, **European Journal of Combinatorics**, 30, p. 1119 - 1133, 2009.  HUIMIN, L, et al. Computational drug repositioning using low-rank matrix approximation and randomized algorithms, **Bioinformatics**, 34, p. 1904 – 1912, 2018.  KOLACZYK, ED, et al. Averages of unlabeled networks: Geometric characterization and asymptotic behavior, **Annals of Statistics**, 48, p. 514-538, 2020.  LE, CM, et al.  Optimization via low-rank approximation for community detection in networks, **The Annals of Statistics**, 44, p. 373 – 400, 2016.  LUO, H, et al. Computational drug repositioning using low-rank matrix approximation and randomized algorithms, **Bioinformatics**, 34, 1904 – 1912, 2018.  MARTIN, T, et al. Localization and centrality in networks, **Physical Review E**, 90, p. 052808, 2014.  MASON, P, and GLEESON, JP. Dynamical systems on networks: A tutorial, arXiv preprint arXiv:1403.7663, 2014.  METZ, FL, et al. Spectra of sparse regular graphs with loops, **Physical Review E**, 84, p. 055101, 2011.  MIROSLAV, F. Algebraic connectivity of graphs, **Czechoslovak Mathematical Journal**, 23, p. 298 – 305, 1973.  NADAKUDITI, RR and NEWMAN, MEJ. Graph spectra and the detectability of community structure in networks, **Physical Review Letters**, 108, p. 188701, 2012.  NADAKUDIITI, RR and NEWMAN, MEJ. Spectra of random graphs with arbitrary expected degrees, **Physical Review E**, 87, p. 012803, 2013.  NEWMAN, MEJ. Modularity and community structure in networks, **Proceedings of the National Academy of Sciences**, 103, p. 8577 – 8582, 2006.  NEWMAN, M.  Networks. **Oxford university press**, 2018.  NEWMAN, MEJ. Spectra of networks containing short loops, **Physical Review E**, 100, p. 012314, 2019.  NEWMAN, MEJ, et al. Spectra of random networks with arbitrary degrees, **Physical Review E**, 99, p. 042309, 2019.  PAN, Y, et al. On the normalized Laplacians with some classical parameters involving graph transformations, **Linear and Multilinear Algebra**, p. 1 - 23, 2018.  PHILLIP, B. Power and centrality: A family of measures, **American journal of sociology**, 92, p. 1170 – 1182, 1987.  PORTER, MA and GLEESON, JP. Dynamical systems on networks: A tutorial, **arXiv preprint arXiv:1403.7663**, 2014.  POTHEN, A, et al. Partitioning sparse matrices with eigenvectors of graphs, **SIAM Journal on matrix analysis and applications**, 11, 430 – 452, 1990.  ROGERS, T, et al. Cavity approach to the spectral density of sparse symmetric random matrices, **Physical Review E**, 78, p. 031116, 2008.  SEMERJIAN G and CUGLIANDOLO, LF. Sparse random matrices: the eigenvalue spectrum revisited, **Journal of Physics A: Mathematical and General**, 35, p. 4837, 2002.  SHAOWEI, S, KINKAR CH, D. Normalized Laplacian eigenvalues with chromatic number and independence number of graphs, **Linear and Multilinear Algebra**, 68, p. 63 – 80, 2020.  SIN, S et al. Parallel clustered low-rank approximation of graphs and its application to link prediction, **International Workshop on Languages and Compilers for Parallel Computing**, Springer, p. 76 – 95, 2012.  SLEIJPEN, GLG and VAN DER VORST HA. A Jacobi--Davidson iteration method for linear eigenvalue problems, **SIAM Review**, 42, p. 267 – 293, 2000.    SUN, S and DAS, KC. Normalized Laplacian eigenvalues with chromatic number and independence number of graphs, **Linear and Multilinear Algebra**, 68, 63 – 80, 2020.  TAKAHASHI, DY, et al. Discriminating different classes of biological networks by analyzing the graphs spectra distribution, **PLoS ONE**, 7, p. e49949, 2012.  TANG, M, et al. A nonparametric two-sample hypothesis testing problem for random graphs, **Bernoulli**, 23.3, p. 1599-1630, 2017.  VAN DEN HEUVEL, MP and SPORNS, O. Network hubs in the human brain, **Trends in Cognitive Sciences**, 17, p. 683 – 696, 2013.  YINGUI, P, JIANPING, L, SHUCHAO, L, WENJUN, L. On the normalized Laplacians with some classical parameters involving graph transformations, **Linear and Multilinear Algebra**, p. 1 – 23, 2018. |

|  |  |
| --- | --- |
| **Measures/activities planning** | |
| **Description of the measures/activities**  Describe the planned measures/activities (also see the category 'Measures/activities eligible for funding' in the programme call for applications).  Explain the extent to which the measures/activities and expenditure are necessary and appropriate to achieving the objectives. *(Keep the description of the measures as brief as necessary).*  Insert new rows in the table for further planned measures/activities.  Note:  The measures/activities must also be entered in the project planning overview table in the sense of results-oriented project planning, and must be assigned to the project objectives (outcomes).  When describing the measures, you should also indicate, which work step will be performed by which groups of researchers, using which method and **where**. | |
| **Title of measure/activity 1:** | Development of an algorithm to estimate the Laplacian spectral distribution of a network with O(n) computational complexity and space (where n is the number of nodes in the network). |
| Description: | Groups or researchers: Stadler’s and Fujita’s groups. Method: Cantwell and Newman (2019) introduced a message-passing approach for the (normalized) Laplacian spectral density using the local neighborhood of each node. Still, it requires computing matrix inversions and matrix-vector multiplications, which are computationally expensive (Cantwell and Newman, 2019). A variation of Newman’s algorithm is to compute the number of eigenvalues in an interval. We will extend both methods so that we will not compute matrix inversions and matrix-vector multiplications and use the fact that real-life networks are locally tree-like networks. |
| Place/time frame | Brazil: 01/2022 – 03/2022 / Germany: 04/2022 – 09/2022 |
| **Title of measure/activity 2:** | Computational simulations of the algorithm to estimate the Laplacian spectral distribution |
| Description: | Groups or researchers: Stadler’s and Fujita’s groups. Method: we will plug in these algorithms to the parameter estimator, model selection, and tests to compare networks to improve the computational complexities. |
| Place/time frame | Germany: 10/2022 – 12/2022 |
| **Title of measure/activity 3:** | Development of a method to identify the nodes’ contribution to a specific eigenvalue density. |
| Description: | Groups or researchers: Stadler’s and Fujita’s groups. Method: We plan to re-write the Dirac’s delta function as the limit of a Lorentzian distribution and obtain the contribution of each node to the spectral density. |
| Place/time frame | Brazil: 01/2023 – 03/2023 / Germany: 04/2023 – 09/2023 |
| **Title of measure/activity 4:** | Computational simulations of the algorithm to identify the node’s contribution to a specific eigenvalue density. |
| Description: | Groups of researchers: Stadler’s and Fujita’s group. Method: We will perform some simulation studies using some already known eigenvalues (e.g., the eigenvalues -1, 0, and 1). After that, we will develop post hoc analyses to interpret the results obtained by our network comparison tests based on the differences among network spectra. |
| Place/time frame | Germany: 10/2023 – 12/2023 |
| **Title of measure/activity 5:** | Application of the developed methods in functional brain networks |
| Description: | Groups or researchers: El Hady’s, Takahashi’s and Fujita’s groups. Method: we will identify which brain regions are associated with a disorder and how this brain region connectivity is altered to controls. |
| Place/time frame | Brazil: 12/2022 – 12/2023 / Germany: 12/2022 – 12/2023 |
| **Title of measure/activity 6:** | Please specify |
| Description: | Please specify |
| Place/time frame | Please specify |
| **Title of measure/activity 7:** | Please specify |
| Description: | Please specify |
| Place/time frame | Please specify |
| **Title of measure/activity 8:** | Please specify |
| Description: | Please specify |
| Place/time frame | Please specify |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Planned international mobility of the groups of researchers**  Please enter the planned stays at the respective partner institute abroad of both groups of researchers during the funding period in the tables in chronological order. | | | | |
| **German project participant performing the stay** | **Academic status/ position** | **Research task to be performed** | **Duration in days** | **Date of the stay (MMYYYY)** |
| Peter F. Stadler | Professor | Review of progress; Discussion of further directions | 8 | 11/2022 |
| Peter F. Stadler | Professor | Review of progress; Manuscript writing | 8 | 11/2023 |
| Jörg Fallmann | PostDoc | Discussion of results; Adaptation of workflows; Manuscript and follow-up grant proposal preparation | 10 | 11/2022 |
| Jörg Fallmann | PostDoc | Manuscript and follow-up grant proposal finalization | 10 | 11/2023 |
| NN (Thomas Gatter) | PostDoc | Discussion of results; Adaptation of workflows; Manuscript and follow-up grant proposal preparation | 10 | 11/2022 |
| NN (Thomas Gatter) | PostDoc | Manuscript and follow-up grant proposal finalization | 10 | 11/2023 |
| Thomas Gatter (NN) | Doktorand | Presentation of results; work on manuscripts; exchange with students from Brazil | 21 | 11/2022 |
| Thomas Gatter (NN) | Doktorand | Presentation of results; work on manuscripts; exchange with students from Brazil | 21 | 11/2023 |
| David Schaller | Doktorand | Presentation of results; work on manuscripts; exchange with students from Brazil | 21 | 11/2022 |
| David Schaller | Doktorand | Presentation of results; work on manuscripts; exchange with students from Brazil | 21 | 11/2023 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Non-German project participant performing the stay** | **Academic status/ position** | **Research task to be performed** | **Duration in days** | **Date of the stay (MMYYYY)** |
| André Fujita | Associate Professor | Workshop and proposal writing to be submitted to AvH | 10 | 07/2022 |
| André Fujita | Associate Professor | Manuscript writing | 10 | 01/2023 |
| Daniel Y. Takahashi | Assistant Professor | Dr. El Hady’s data analysis and discussions | 10 | 01/2023 |
| Eduardo Lira | Graduate student | Develop algorithm described in goal 1 | 180 | 04/2022 |
| Vinicius Jardim Carvalho | Graduate student | Analyze Dr. El Hady’s data | 180 | 04/2022 |
| Heitor Baldo | Graduate student | Develop algorithm described in goal 2 | 180 | 04/2023 |
| Victor Chavauty Villela | Graduate student | Analyze Dr. El Hady’s data | 180 | 04/2023 |
| Grover E.C. Guzman | Postdoc | Develop algorithm described in goal 1 | 270 | 04/2022 |
| Diogo Costa | Postdoc | Develop algorithm described in goal 2 | 270 | 04/2023 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Further programme-specific information** | | | | | |
| **Roles in the project** List the project participants in Germany and outside Germany and state the tasks for which they are responsible in the project. | | | | | |
| Peter F. Stadler (German coordinator) and André Fujita (Brazilian coordinator) will develop the methods/algorithms and analyze the empirical data.  Daniel Y. Takahashi (Brazilian collaborator) and Ahmed El Hady (German collaborator) will provide the biological data to be analyzed and interpret the results. | | | | | |
| **Structure of the group of researchers and role of project participants**  Explain the structure of the group of researchers and the criteria based on which you selected the project participants. | | | | | |
| Our proposal ranges from theoretical/methodology development to application in neuroscience. Thus, this proposal comprises two groups of researchers, one of mathematics/computer science and one of neuroscience. Each group is composed of two labs. Mathematics/computer science: Dr. Stadler’s and Dr. Fujita’s labs. Neuroscience: Dr. El Hady’s and Dr. Takahashi’s labs. We based the participants selection criteria on the fitness for our problems treated in this proposal. Thus, participants should have background in at least one of the following areas: mathematics, theoretical computer science, statistics, neuroscience. | | | | | |
| Will third-party funds be introduced? | | Yes |  | No |  |
| Has the third-party funder provided a legally binding declaration / commitment? | | Yes |  | No |  |
| Reason: | Please specify | | | | |

|  |  |
| --- | --- |
| **Commitment to comply with the recommendations for good scientific practice** | |
| Project coordinator: | Peter F. Stadler |
|  | If my research project is accepted for the DAAD Programme for Project-Related Personal Exchange, I undertake to comply with the rules of good scientific practice.[[6]](#footnote-7)  Scientific misconduct is given if false statements are made in a context of scientific importance either intentionally or by gross negligence, if intellectual property rights of others are violated, or if the research activities of others are otherwise affected. The circumstances of the individual case are decisive. |

|  |  |  |
| --- | --- | --- |
| **Application checklist** | | |
|  | | |
| **Application documents relevant to selection** | |  |
| 1 | Project application (in the DAAD portal) |  |
| 2 | Financing plan (in the DAAD portal) |  |
| 3 | Project description |  |
| 4 | Project planning overview |  |
| 5 | Research profile/CVs of the German project coordinator(s) |  |
| 6 | List of the German project coordinators’ publications in the past 5 years that are relevant to the project |  |
| 7 | Research profile/CVs of the project coordinator(s) abroad |  |
| 8 | List of the non-German project coordinators’ publications in the past 5 years that are relevant to the project |  |
| 9 | Brief CVs of any other project participants that have already been selected at the time of application |  |
| 10 | For applications for PPP Canada and PPP USA: a confirmation letter by the cooperation partner in addition to attachments 1-8 |  |

1. The project does not necessarily need to aim at achieving all programme objectives (outcomes of the funding programme). 'Funding programme' and 'programme' are used synonymously. [↑](#footnote-ref-2)
2. Only the results of the measures/activities (outputs of the funding programme) which are relevant for the selected programme objectives (outcomes of the funding programme) must be taken into account. [↑](#footnote-ref-3)
3. Follow-up application: Application for a project which immediately follows on from funding for the previous year in the same funding programme. [↑](#footnote-ref-4)
4. For the definitions of 'Outcomes' and 'Outputs', please refer to the ‘Guide to Results-oriented Monitoring’. [↑](#footnote-ref-5)
5. See ‘Guide to Results-oriented Project Planning and Monitoring’, Chapter 2. [↑](#footnote-ref-6)
6. The rules of good scientific practice are detailed in the memorandum ‘Safeguarding Good Scientific Practice’ (WILEY-VCH Verlag) and in the Guidelines for the Use of Funds – DFG templates 2.01 and 2.02 – (available on the DFG website: http://www.dfg.de – ‘Proposals’ section). This version is based on the suggestions of the international commission for self-regulation in science and it corresponds to a resolution passed by the General Assembly of the DFG on 17 June 1998 in coordination with the HRK. [↑](#footnote-ref-7)