# How Reliable Are Networks? A Bayesian Modeling Approach

**Anonymous Author(s)**
Affiliation
Address
`email`

## Abstract

Network analyses of white matter pathways linking brain regions—noninvasively extracted from diffusion magnetic resonance imaging—hold great clinical application promise. However, these networks display low reliability at the level of single brain connections, severely limiting inference. We present a Bayesian modeling framework to assess the reliability of network connections across repeated measurements. We model connection strength as a mixture of two probabilistic components: one representing the presence of a true connection, and its true absence. Using simulated, repeated-measures data, we estimate the posterior distribution of connection strengths and quantify the reliability by examining the spread of these distributions. The model was sensitive to connections with varying levels of reliability. However, it underestimated the probability that a connection is absent, and failed to recover the parameters after generating data with the same model.

## 1   Introduction

A wide range of conditions, including schizophrenia [Griffa et al., 2015] and bipolar disorder [Fernandes et al., 2019], are thought to arise from altered brain connectivity. However, network representations derived from diffusion MRI (dMRI) yield unreliable estimates [Maier-Hein et al., 2017, Thomas et al., 2014, Nakuci et al., 2023], hindering biomarker discovery and clinical translation. Post-processing methods such as streamline filtering [Smith et al., 2013] can improve robustness but do not quantify residual uncertainty.

Bayesian approaches offer a principled way to assess edge uncertainty [Hinne et al., 2013], and have been used to model disease-specific alterations [Peterson et al., 2020] and causal interactions [Dang et al., 2018]. Here, we validate a Bayesian framework to (1) quantify the reliability of fiber density estimates for each structural connection (SC), and (2) generate an atlas classifying connections as likely present or absent. This approach produces connectivity estimates with explicit posterior confidence for each edge, enabling more reliable interpretation of SC and, ultimately, improved clinical applicability.

## 2   Methods

### 2.1   Model specification

We model white matter track density $D$ as a mixture of two components: one for absent connections ($C = 0$) and one for present connections ($C = 1$), where $C$ is latent. Noise, motion, and processing variability can yield nonzero densities even for $C = 0$; this component is modeled with a fast-decaying exponential distribution. We model true connections ($C = 1$) to follow a normal distribution centered on the connection strength, truncated at zero since fiber densities are nonnegative. The prevalence of either component is modeled by the probability $\pi_0$ that the connection is truly absent:

$$P(D) = P(D|C = 0)P(C = 0) + P(D|C = 1)P(C = 1)$$
$$P(D) = \pi_0 \cdot \text{Exp}(\lambda) + (1 - \pi_0) \cdot \mathcal{N}^+(\mu, \sigma), \tag{1}$$

where $\text{Exp}(\lambda)$ is a decaying exponential distribution with a rate $\lambda$, and $\mathcal{N}^+(\mu, \sigma)$ is a normal distribution truncated at zero, with a mean $\mu$ and a standard deviation $\sigma$.

**Prior distributions.** We set $\pi_0 \sim \text{Beta}(2, 5)$ to model probabilities in [0,1], centering on the 5–40The rate $\lambda \sim \text{Gamma}(1, 10)$ ensures positivity with flexible deviation, and $\sigma \sim \text{HalfNormal}(0.6)$ reflects positive and small expected variability in fiber density. Because the real dataset we will fit the model on includes only 36 repeated diffusion MRI sessions, we fixed $\mu$ to the mean fiber density across repeats rather than estimating it.

**Model implementation.** We implement and fit the model using PyMC [Abril-Pla et al., 2023]. All experiments are conducted on an Intel(R) Core(TM) i9-10980XE CPU @ 3.00GHz, 36 cores, 62 GB of RAM, Ubuntu 20.04.6 LTS. For inference, we used the No-U-Turn Sampler (NUTS), PyMC's default Hamiltonian Monte Carlo algorithm, with 4 chains and 2000 posterior draws per chain following 1000 tuning steps.

## 2.2 Validation

To simulate a single subject scanned 36 times (data we will leverage to quantify within-scanner edge-wise reliability using the model), we repeated a reference SC matrix from the atlas Alemán-Gómez et al. [2022] (CC-BY-4.0 license) with different realizations of bi-modal noise. We added Gaussian noise $\mathcal{N}(0, 0.2)$ to all connections mimicking measurement noise and a stronger noise $\mathcal{N}(0, 0.5)$ to those below the $40^{\text{th}}$ fiber-density percentile ($\leq 4.76$) to reflect the empirical observation that weaker connections are less reliable. This simulation was used to validate our model through three experiments. Parameter estimates were visualized as heatmaps, with connection groups highlighted via transparency masks, and group differences tested using two-sample t-tests (`ttest_ind`, SciPy [Virtanen et al., 2020]). The model was fit assuming independent edges, using four PyMC chains per edge (`cores=1`) and up to 20 parallel `joblib` jobs (total runtime: 9 h 5 min; 107 s/edge).

**Experiment 1.** We evaluate whether the model could detect the varying levels of edge-wise reliability we injected in the simulated SC, with reliability quantified as the standard deviation $\sigma$ of the normal component in the estimated posterior distribution.

**Experiment 2.** We assess whether the estimated $\pi_0$ correctly identified connections consistently absent across all 66 subjects used to build the connectome atlas Alemán-Gómez et al. [2022] as truly absent and all other connections as truly present.

**Experiment 3.** We systematically assess the model's ability to recover known parameter values. We generate SC matrices by fixing the true parameter values in Equation (1) and sampling from the posterior distribution. For each configuration, we fit the Bayesian model to the simulated data 10-30 times and compute the relative root mean square error (RMSE) between the true values and the posterior means. The true parameter values used are listed in Equation (2):

$$\pi_0 = 0.1, \quad \lambda = 2.0, \quad \sigma = 0.5,$$
$$\mu \in [0.1, 0.2, 0.5, 0.7, 1.0, 2.0, 3.0, 5.0, 8.0, 20.0, 30.0, 1000, 10000] \tag{2}$$

Since fiber density best distinguishes true from false connections, we varied the mean connection strength $\mu$ while keeping other parameters fixed. We set $\pi_0 = 0.1$ because the model consistently estimated low $\pi_0$ across connections (Figure 2). The choices $\lambda = 2$ and $\sigma = 0.5$ match the noise characteristics from the simulated SC matrices. The $\mu$ values span the observed range of average streamline counts in the reference SC matrix. Each $\mu$ fit (30 repetitions) took about 7 min 40 s, totaling roughly 1 h 40 min for all values.

## 3 Results and Discussion

**The model is sensitive to connections with varying levels of reliability.** Experiment 1 shows that the estimated $\sigma$ is systematically higher for connections with lower fiber density—those to which we added more noise (Figure 1). The model identified the two latent $C$ groups ($p < 0.001$).

**The model underestimates the probability of absence.** Figure 2 shows that, as expected, $\pi_0$ is low for consistent connections (Panel B), but for truly absent ones (Panel A), it remains lower than expected (max = 0.16 instead of $\approx 1$).
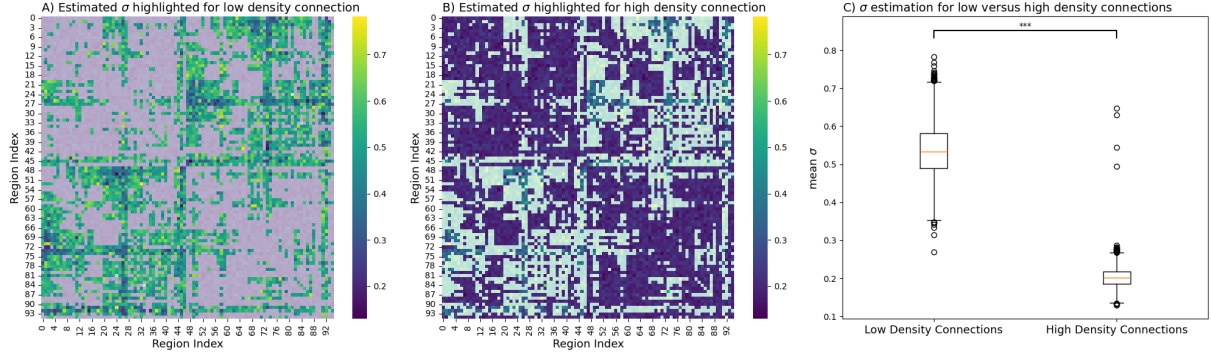
**Figure 1:** The model is sensitive to connections with varying levels of reliability.
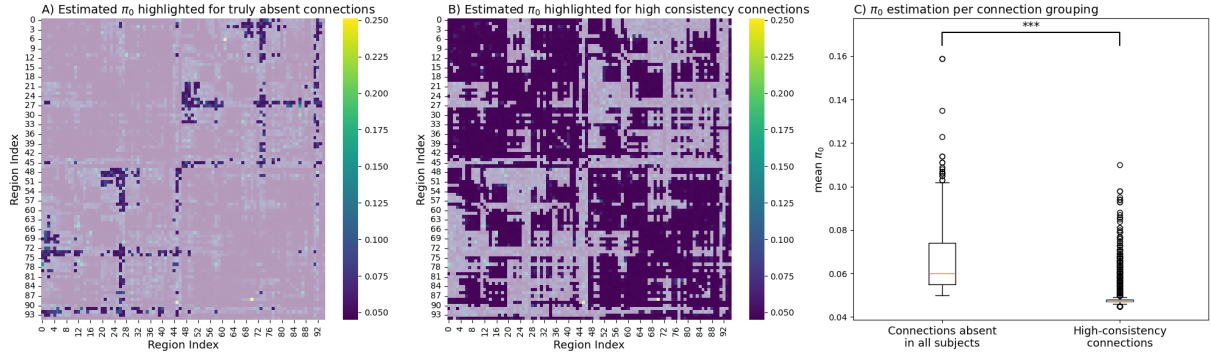


**Figure 2:** The model correctly predicts low $\pi_0$ for connections consistent across subjects, but underestimates $\pi_0$ when connections are truly absent.
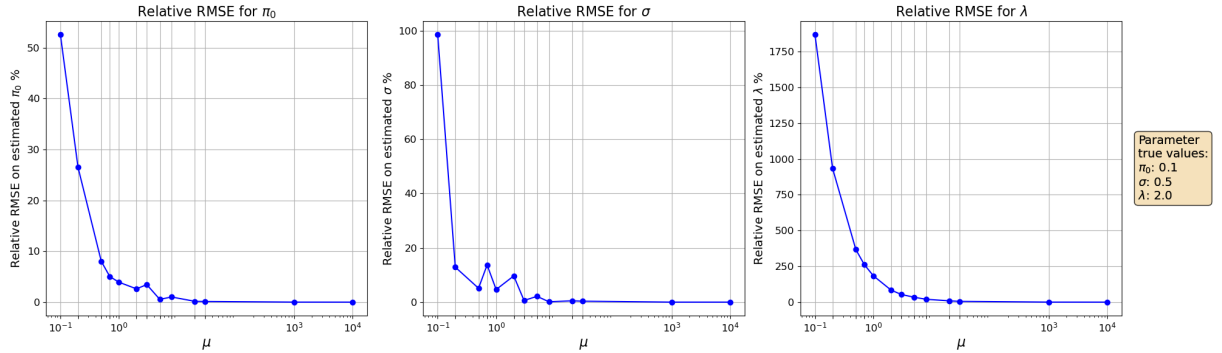


**Figure 3:** The model does not accurately estimate parameter values when the normal component of the posterior distribution is centered near zero.

**The model does not accurately estimate parameter values when the normal component of the posterior distribution is centered near zero.** Figure 3 shows that relative RMSE is high when $\mu$ is near zero, indicating poor parameter recovery, but estimation improves markedly as $\mu$ increases. To avoid misestimating $\pi_0$ for truly absent connections, this limitation should be addressed before applying the model to the real SC matrices.

## 4   Conclusion

This study establishes a foundation for embedding uncertainty into network representations, in particular, structural brain connectivity extracted with dMRI.

3

# References

Oriol Abril-Pla, Virgile Andreani, Colin Carroll, Larry Dong, Christopher J. Fonnesbeck, Maxim Kochurov, Ravin Kumar, Junpeng Lao, Christian C. Luhmann, Osvaldo A. Martin, Michael Osthege, Ricardo Vieira, Thomas Wiecki, and Robert Zinkov. PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science*, 9:e1516, September 2023. ISSN 2376-5992. doi: 10.7717/peerj-cs.1516. URL `https://peerj.com/articles/cs-1516`. Publisher: PeerJ Inc.

Yasser Alemán-Gómez, Alessandra Griffa, Jean-Christophe Houde, Elena Najdenovska, Stefano Magon, Meritxell Bach Cuadra, Maxime Descoteaux, and Patric Hagmann. A multi-scale probabilistic atlas of the human connectome. *Scientific Data*, 9(1):516, August 2022. ISSN 2052-4463. doi: 10.1038/s41597-022-01624-8. URL `https://www.nature.com/articles/s41597-022-01624-8`. Publisher: Nature Publishing Group.

Shilpa Dang, Santanu Chaudhury, Brejesh Lall, and Prasun K. Roy. Tractography-Based Score for Learning Effective Connectivity From Multimodal Imaging Data Using Dynamic Bayesian Networks. *IEEE Transactions on Biomedical Engineering*, 65(5):1057–1068, 2018. doi: 10.1109/TBME.2017.2738035.

Henrique M. Fernandes, Joana Cabral, Tim J. van Hartevelt, Louis-David Lord, Carsten Gleesborg, Arne Møller, Gustavo Deco, Peter C. Whybrow, Predrag Petrovic, Anthony C. James, and Morten L. Kringelbach. Disrupted brain structural connectivity in Pediatric Bipolar Disorder with psychosis. *Scientific Reports*, 9(1):13638, September 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-50093-4. URL `https://www.nature.com/articles/s41598-019-50093-4`. Publisher: Nature Publishing Group.

Alessandra Griffa, Philipp Sebastian Baumann, Carina Ferrari, Kim Quang Do, Philippe Conus, Jean-Philippe Thiran, and Patric Hagmann. Characterizing the connectome in schizophrenia with diffusion spectrum imaging. *Human Brain Mapping*, 36(1):354–366, January 2015. ISSN 1097-0193. doi: 10.1002/hbm.22633.

Max Hinne, Tom Heskes, Christian F. Beckmann, and Marcel A. J. van Gerven. Bayesian inference of structural brain networks. *NeuroImage*, 66:543–552, February 2013. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2012.09.068. URL `https://www.sciencedirect.com/science/article/pii/S105381191200986X`.

Klaus H. Maier-Hein, Peter F. Neher, Jean-Christophe Houde, Marc-Alexandre Côté, Eleftherios Garyfallidis, Jidan Zhong, Maxime Chamberland, and others. The challenge of mapping the human connectome based on diffusion tractography. *Nature Communications*, 8:1349, 2017. doi: 10.1038/s41467-017-01285-x.

Johan Nakuci, Nick Wasylyshyn, Matthew Cieslak, James C. Elliott, Kanika Bansal, Barry Giesbrecht, Scott T. Grafton, Jean M. Vettel, Javier O. Garcia, and Sarah F. Muldoon. Within-subject reproducibility varies in multi-modal, longitudinal brain networks. *Scientific Reports*, 13(1):6699, April 2023. ISSN 2045-2322. doi: 10.1038/s41598-023-33441-3. URL `https://www.nature.com/articles/s41598-023-33441-3`. Publisher: Nature Publishing Group.

Christine B. Peterson, Nathan Osborne, Francesco C. Stingo, Pierrick Bourgeat, James D. Doecke, and Marina Vannucci. Bayesian modeling of multiple structural connectivity networks during the progression of Alzheimer's disease. *Biometrics*, 76(4):1120–1132, 2020. doi: 10.1111/biom.13235.

Robert E. Smith, Jacques-Donald Tournier, Fernando Calamante, and Alan Connelly. SIFT: Spherical-deconvolution informed filtering of tractograms. *NeuroImage*, 67:298–312, 2013. doi: 10.1016/j.neuroimage.2012.11.049.

Cibu Thomas, Frank Q. Ye, M. Okan Irfanoglu, Pooja Modi, Kadharbatcha S. Saleem, David A. Leopold, and Carlo Pierpaoli. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proceedings of the National Academy of Sciences of the United States of America*, 111(46):16574–16579, November 2014. doi: 10.1073/pnas.1405672111. URL `www.pnas.org/cgi/doi/10.1073/pnas.1405672111`. Publisher: National Academy of Sciences.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

## NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction clearly state that the paper introduces and validates a Bayesian model to quantify connection reliability and infer presence/absence in brain networks, which are exactly the contributions evaluated in the experiments.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The limitations are discussed in the Results and Discussion section, especially regarding the model's inability to clearly identify truly absent connections when $\mu$ is close to zero.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.

- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [NA]

Justification: The paper does not contain formal theoretical results or proofs; it presents a probabilistic model validated via simulation.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: All details about the model, priors, simulation setup, parameter values, hardware/software, and sampling configurations are included in Section 2. Moreover, the code is publicly available so that every detail can be retrieved.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example

(a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.

(b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.

(c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: The links to the code and data used in this study will happened to the final paper, but were removed in this submission to abide to the double blind review rules.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Section 2 of the paper details the simulation design, noise parameters, prior choices, sampler choice, and sampling hyperparameters (e.g., number of chains, draws, tuning steps), enabling reproducibility.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Statistical significance is reported for group comparisons (e.g., t-tests in Experiment 1) and estimation error is quantified using relative RMSE in Experiment 3.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
   - If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The paper details the local machine (Intel i9 CPU, 62GB RAM), software stack, runtime per job, number of chains, number of repetitions, and total compute time (Section 2.1).

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: The research is conducted using simulated data and public atlases, involves no human subject interaction, and aligns with the NeurIPS Code of Ethics.

   Guidelines:

349      • The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.

350      • If the authors answer No, they should explain the special circumstances that require a
351        deviation from the Code of Ethics.

352      • The authors should make sure to preserve anonymity (e.g., if there is a special consid-
353        eration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [NA]

Justification: The paper discusses how incorporating uncertainty into structural connectivity modeling may be a first step towards improving clinical translation and interpretation of brain networks, but is only a first step in a vast multitude of other steps needed to make a concrete impact. We anticipate the interest in this model to be very niche among researchers interested in repeated measures reliability.

Guidelines:

     • The answer NA means that there is no societal impact of the work performed.

     • If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.

     • Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

     • The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.

     • The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.

     • If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The model is trained solely on simulated data and does not involve potentially misusable assets.

Guidelines:

     • The answer NA means that the paper poses no such risks.

     • Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.

     • Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.

- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The structural connectivity atlas from [Alemán-Gómez et al., 2022]) is properly cited and its license is mentioned. PyMC and other software libraries are also acknowledged with appropriate references.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: No new datasets or trained models are released in this paper.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve any research with human subjects or crowdsourcing.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.

- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The study does not involve human subjects and thus does not require IRB approval.

    Guidelines:
    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: [NA]

    Justification: LLMs were used to help improve the text formulation and correct grammatical mistakes in the manuscript. It was also leveraged through GitHub Copilot to produce code faster, but was not involved in the core method development.

    Guidelines:
    - The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy (https://neurips.cc/Conferences/2025/LLM) for what should or should not be described.