

Virtual Brain Inference (VBI) toolkit on EBRAINS

Abolfazl Ziaemehr¹, Lia Domide², Marmaduke Woodman¹, Spase Petkoski¹, Viktor Jirsa¹, and Meysam Hashemi¹

¹ Aix Marseille Univ, INSERM, INS, Institut de Neurosciences des Systèmes, Marseille, France ² Codemart, Cluj-Napoca, Romania [✉] Corresponding author

DOI: [10.21105/joss.10612](https://doi.org/10.21105/joss.10612)

Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Ujjwal Karn](#)

Reviewers:

- [@sampottinger](#)
- [@tradertanmay](#)

Submitted: 14 January 2026

Published: 27 August 2026

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Understanding the principles and causal mechanisms underlying complex brain (dys)function is essential for advances in precision medicine. The virtual brain modeling represents a major step forward, combining computational models of brain dynamics with personalized anatomical data to simulate at scales commensurate with neuroimaging modalities. However, there is a lack of widely available tools to address the inverse problem of inferring neural mechanisms from observed brain activity. This work introduces Virtual Brain Inference (VBI), a flexible, scalable, and integrative Python toolkit designed for efficient probabilistic inference on virtual brain models. VBI is now available as an open-source service on the cloud research platform EBRAINS (ebraains.eu). It offers fast simulations of popular whole-brain models, a taxonomy of feature extraction methods, efficient data storage, and advanced probabilistic machine learning algorithms for causal inference from non-invasive and invasive recordings. VBI cloud services on EBRAINS, including demo applications, enable reproducible online collaboration and seamless integration with scalable and secure workflows. This fosters research on large cohort datasets, improving generalizability, and potentially clinical translation.

Statement of Need

VBI (<https://github.com/ins-amu/vbi>) is a Python-based toolkit tailored for probabilistic inference at the whole-brain scale. A lightweight version with minimal dependencies is available to run online in JupyterLab ([Kluyver et al., 2016](#)). It also leverages the flexibility of Python while harnessing the high performance of C++ for optimized computation and massive parallelization using CUDA on GPUs. VBI seamlessly integrates structural and functional neuroimaging data, providing space-efficient storage and memory-optimized batch processing. With its user-friendly API supporting:

- **Whole-brain models:** Wilson-Cowan ([Wilson & Cowan, 1972](#)), Jansen-Rit ([Jansen & Rit, 1995](#)), Stuart-Landau ([Cabral et al., 2011](#)), Montbrió ([Montbrió et al., 2015](#)), Wong-Wang ([Deco et al., 2013](#)), and Epileptor ([Jirsa et al., 2014](#)).
- **Fast simulation:** Just-in-time compilation of models across Python/C++ and CPU/GPU devices.
- **Feature extraction:** Functional connectivity (FC), functional connectivity dynamics (FCD), and power spectral density (PSD).
- **Deep neural density estimators:** Masked autoregressive flows (MAFs; [Papamakarios et al., 2017](#)), and neural spline flows (NSFs; [Durkan et al., 2019](#)).

Traditional methods such as Markov Chain Monte Carlo (MCMC) and Approximate Bayesian Computation (ABC) face significant challenges in this context ([Sisson et al., 2007](#)). Gradient-free MCMC struggles in high-dimensional settings. Gradient-based approaches, meanwhile, often have difficulty converging for bistable systems with switching dynamics, often requiring

extensive tuning (Betancourt & Girolami, 2013) and computational resources (Hashemi et al., 2020). ABC suffers from the curse of dimensionality and relies on predefined thresholds for sample acceptance, leading to inefficiencies and potential biases when rejecting samples that fall outside narrow criteria (Jha et al., 2022). VBI leverages advanced Simulation-based Inference (SBI; Cranmer et al., 2020), which sidesteps these issues by using only forward simulations and training deep neural density estimators to directly approximate posterior distributions (Gonçalves et al., 2020; Hashemi et al., 2023). This approach enhances efficiency, scalability, and robustness for inverting complex whole-brain models (Hashemi et al., 2024).

Designed for researchers and clinical applications, VBI enables personalized simulations of normal and pathological brain activity, aiding in distinguishing healthy from diseased states, and potentially informing clinical decision making. By addressing the efficient and scalable probabilistic inverse problem, VBI leverages high-performance computing for parallel processing of large-scale datasets (Ziaemehr et al., 2025).

VBI is intended for computational neuroscientists, clinical researchers, and method developers who need to fit whole-brain network models to neuroimaging data (fMRI, EEG, MEG) at scale, without having to implement custom simulators, feature extractors, or deep-learning-based inference pipelines themselves.

State of the Field

No single existing tool combines the components needed for scalable, amortized inference on whole-brain models. The Virtual Brain (TVB) (Sanz Leon et al., 2013) provides mature forward simulators that many of VBI's models build upon, but it lacks simulation-based inference with deep neural density estimators or amortized posteriors. General-purpose SBI libraries (Cranmer et al., 2020; Gonçalves et al., 2020) offer flexible density estimation but lack domain-specific whole-brain simulators, neuroimaging feature extractors, and GPU/C++ acceleration. Related model-specific Bayesian approaches, such as self-tuning Hamiltonian Monte Carlo (Jha et al., 2022) and earlier amortized SBI for epilepsy models (Hashemi et al., 2023, 2024), demonstrate the need but are not general, reusable toolkits. VBI fills this gap, integrating whole-brain simulators, neuroimaging-specific feature extraction, and amortized inference in one GPU-accelerated, open-source package interoperable with TVB, rather than duplicating existing efforts.

Software Design

Design goals and trade-offs. VBI was designed to solve the inverse problem for whole-brain models under realistic computational constraints. The primary goals were scalability to large simulation budgets, support for amortized probabilistic inference, and modularity across models, features, and inference algorithms. A key trade-off concerned flexibility versus performance: while pure Python implementations simplify extensibility, they are insufficient for large-scale simulation-based inference. VBI therefore adopts a hybrid design, exposing a Python API while relying on just-in-time compilation and GPU acceleration for performance-critical components. This choice enables rapid prototyping while achieving up to 100× speedups compared to CPU-based single core implementations, as benchmarked in (Ziaemehr et al., 2025).

Architecture. The toolkit is organized around three decoupled layers: (1) simulators for whole-brain neural mass models, (2) feature-extraction operators for neuroimaging-derived statistics, and (3) inference backends based on simulation-based inference. Clear interfaces between these layers allow users to independently exchange models, features, or density estimators. Deployment on EBRAINS further motivated containerized, dependency-light components and a JupyterLab-based execution model, which was later extended with a graphical workflow editor for non-programmatic interaction.

Research Impact Statement

VBI enables a class of large-scale, amortized inference workflows for whole-brain models that were previously impractical using MCMC or ABC approaches. The methodology and core software have already supported peer-reviewed research demonstrating reliable parameter recovery and mechanistic insights in whole-brain dynamics, including a full validation study published in eLife (Ziaemehr et al., 2025), methodological work on simulation-based inference for virtual brain models (Hashemi et al., 2024), applications to personalized disease modeling and brain-network reconfiguration (Angiolelli et al., 2025; Mazzara et al., 2025; Rabuffo et al., 2025; Sorrentino et al., 2024), and broader virtual-brain workflows in probabilistic modeling and clinical neuroscience (Baldy et al., 2025; Hashemi et al., 2025; Monteverdi et al., 2026). This constitutes realized impact beyond a purely methodological contribution.

Near-term impact is reinforced by the EBRAINS integration, which provides ready-to-run environments, curated demonstrations, and reproducible workflows accessible to external users. Additional community-readiness signals include open-source licensing, documentation and example notebooks, GPU-accelerated benchmarks for commonly used models, and a defined extension interface. The addition of a graphical user interface further broadens accessibility to experimental and clinical researchers, positioning VBI as an enabling infrastructure for reproducible and scalable whole-brain inference.

AI Usage Disclosure

Generative AI tools, specifically Claude Sonnet 4.5 (Anthropic, 2025) and GPT-5.1 (OpenAI, 2025), were used mainly for language polishing and documentation drafting of this manuscript, and for minor code refactoring assistance. All scientific content, software design decisions, implementation, and validation results were conceived, written, and verified by the authors; the majority of the codebase was authored directly by the authors rather than produced through autonomous or “vibe coding” workflows.

Methods

Whole-brain network dynamics are modeled as coupled neural mass models interacting through structural connectivity (Hashemi et al., 2025):

$$\dot{\psi}_i(t) = \mathcal{N}(\psi_i) + G \sum_{j=1}^N w_{ij} \mathcal{H}(\psi_i, \psi_j(t - \tau_{ij})) + z(\psi_i) \xi_i(t),$$

where $i = 1, 2, \dots, N$, and N is the number of brain regions. Here, $\psi_i(t)$ represents local neural activity in region i , governed by the nonlinear function $\mathcal{N}(\psi_i)$ when uncoupled. Interactions between regions are mediated by the coupling function $\mathcal{H}(\psi_i, \psi_j(t - \tau_{ij}))$, weighted by the structural connectivity w_{ij} (derived from diffusion-weighted MRI tractography) and delayed by axonal transmission times τ_{ij} . A noise term, $z(\psi_i) \xi_i(t)$, is added as Gaussian noise with:

$$\langle \xi_i(t) \rangle = 0, \quad \langle \xi_i(t) \xi_j(t') \rangle = 2D \delta(t - t') \delta_{i,j},$$

where D is the noise strength. The system's operating regime emerges from the interplay of global coupling G , local bifurcation parameters, and noise, with connectivity structure shaping macroscopic brain activity through delays and weights (Ghosh et al., 2008; Petkoski & Jirsa, 2019; Sanz Leon et al., 2013; Ziaemehr et al., 2020).

VBI relies on three main components:

1. A prior distribution, $p(\vec{\theta})$, describing the possible range of parameters i.e., $\vec{\theta} \sim p(\vec{\theta})$.

2. A whole-brain simulator, $p(\vec{x} \mid \vec{\theta})$, generating data $\vec{x}$ given parameters $\vec{\theta}$.
3. Low-dimensional features for training deep neural density estimators.

Together, these yield a training dataset $\{(\vec{\theta}_i, \vec{x}_i)\}_{i=1}^{N_{sim}}$, where N_{sim} is the simulation budget. After training, we can efficiently estimate the posterior distribution given observed data, i.e., $p(\vec{\theta} \mid \vec{x}_{obs})$.

The VBI workflow comprises:

1. **Fast simulation:** generating data associated with various neuroimaging recordings.
2. **Feature extraction:** computing statistical, spatiotemporal, and other features.
3. **Efficient inference:** amortized training of deep neural density estimators.

Evaluation of Posterior Fit

Posterior reliability is assessed using posterior z-scores (denoted by z) and shrinkage (denoted by s) (Betancourt et al., 2017):

$$z = \left| \frac{\bar{\theta} - \theta^*}{\sigma_{post}} \right|, \quad s = 1 - \frac{\sigma_{post}^2}{\sigma_{prior}^2},$$

where $\bar{\theta}$ is the posterior mean, θ^* is the true parameter, and σ_{post} and σ_{prior} are posterior and prior standard deviations, respectively. High shrinkage indicates well-identified posteriors, while low z-scores confirm accurate capture of true values. Through in silico testing, these metrics indicated the accuracy and reliability of inference for commonly used whole-brain network models and their associated neuroimaging data (Ziaemehr et al., 2025).

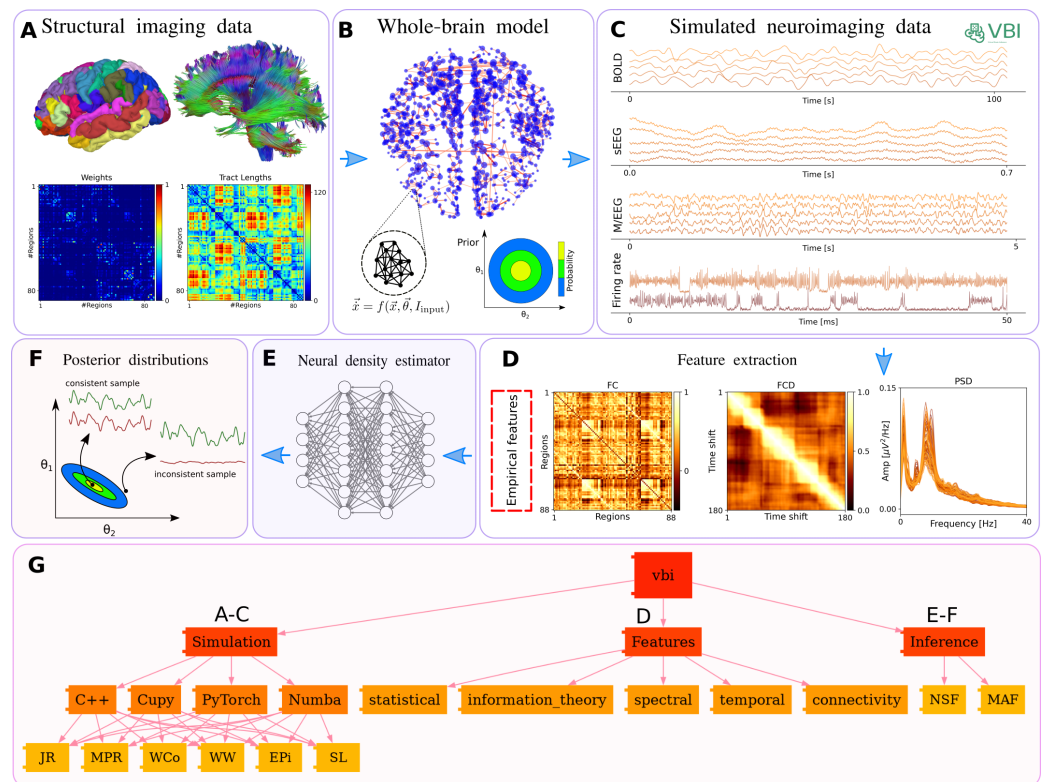


Figure 1: Overview of the VBI workflow: **(A)** A personalized connectome is built using diffusion tensor imaging and a brain parcellation atlas. **(B)** This forms the basis to build a virtual brain model, given control parameters sampled from a prior distribution. **(C)** VBI simulates time series data associated with neuroimaging recordings. **(D)** Data features (such as FC, FCD, PSD) are extracted from simulations. **(E)** Deep neural density estimators are trained on parameter-feature pairs to learn the family of posterior distributions over model parameters. **(F)** The trained network rapidly approximates the posterior for any new observation. **(G)** A flowchart of the VBI modules illustrates the pipeline's modularity and flexibility, allowing seamless integration of different simulators, feature extraction, and inference tools.

EBRAINS Integration

EBRAINS is an open research infrastructure that integrates data, tools, and computational resources for brain-related research, with interoperability as a central design principle. From the perspective of the VBI project, EBRAINS serves as a deployment environment in which the VBI toolkit can be installed after each release and subsequently shared with the broader research community. This mode of deployment enhances reproducibility and lowers the barrier to entry for users without programming expertise, as it enables direct access to the tools through a running JupyterLab environment at lab.ebrains.eu. To further improve usability, we are also developing a graphical user interface (GUI) for VBI components, allowing users to configure simulation workflows through drag-and-drop interactions. This interface is implemented as a JupyterLab extension.

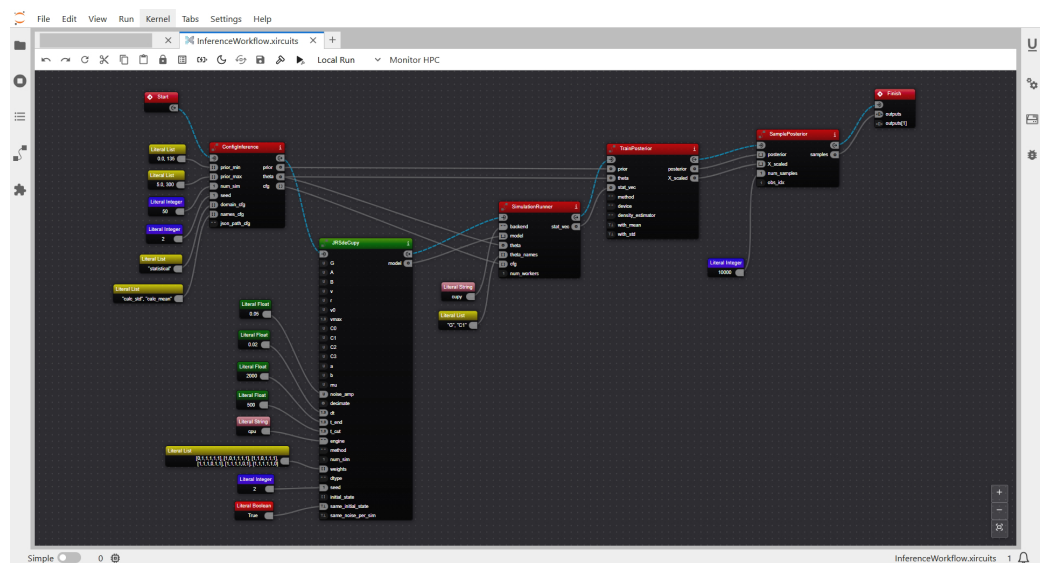


Figure 2: VBI components exposed as graphical boxes in tvb-ext-xircuits.

Acknowledgements

This project/research has received funding from the European Union's Horizon Europe Programme under the Specific Grant Agreement No. 101147319 (EBRAINS 2.0 Project), No. 101137289 (Virtual Brain Twin Project), No. 101057429 (project environMENTAL), and a government grant managed by the Agence Nationale de la Recherche reference ANR-22-PESN-0012 (France 2030 program). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Angiolelli, M., Depannemaecker, D., Agouram, H., Régis, J., Carron, R., Woodman, M., Chiodo, L., Triebkorn, P., Ziaemehr, A., Hashemi, M., & others. (2025). The virtual Parkinsonian patient. *Npj Systems Biology and Applications*, 11(1), 40. <https://doi.org/10.1038/s41540-025-00516-y>
- Anthropic. (2025). *Claude's system card*. <https://www.anthropic.com/system-cards>.
- Baldy, N., Woodman, M., Jirsa, V. K., & Hashemi, M. (2025). Dynamic causal modelling in probabilistic programming languages. *Journal of the Royal Society Interface*, 22(227). <https://doi.org/10.1098/rsif.2024.0880>
- Betancourt, M., Byrne, S., Livingstone, S., & Girolami, M. (2017). The geometric foundations of Hamiltonian Monte Carlo. *Bernoulli*, 23(4A), 2257–2298. <https://doi.org/10.3150/16-BEJ810>
- Betancourt, M., & Girolami, M. (2013). Hamiltonian Monte Carlo for hierarchical models. *arXiv:1312.0906*. <https://doi.org/10.48550/arXiv.1312.0906>
- Cabral, J., Hugues, E., Sporns, O., & Deco, G. (2011). Role of local network oscillations in resting-state functional connectivity. *NeuroImage*, 57(1), 130–139. <https://doi.org/10.1016/j.neuroimage.2011.04.010>
- Cranmer, K., Brehmer, J., & Louppe, G. (2020). The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48), 30055–30062. <https://doi.org/10.1073/pnas.1912789117>

- Deco, G., Ponce-Alvarez, A., Hagmann, P., Romani, G. L., Mantini, D., & Corbetta, M. (2013). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience*, 33(27), 11239–11252. <https://doi.org/10.1523/JNEUROSCI.1091-13.2013>
- Durkan, C., Bekasov, A., Murray, I., & Papamakarios, G. (2019). Neural spline flows. *Advances in Neural Information Processing Systems*, 32, 7511–7522. <https://doi.org/10.48550/arXiv.1906.04032>
- Ghosh, A., Rho, Y., McIntosh, A. R., Kötter, R., & Jirsa, V. K. (2008). Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS Computational Biology*, 4(10), e1000196. <https://doi.org/10.1371/journal.pcbi.1000196>
- Gonçalves, P. J., Lueckmann, J.-M., Deistler, M., Nonnenmacher, M., Öcal, K., Bassetto, G., Chintaluri, C., Podlaski, W. F., Haddad, S. A., Vogels, T. P., & others. (2020). Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife*, 9, e56261. <https://doi.org/10.7554/eLife.56261>
- Hashemi, M., Depannemaecker, D., Saggio, M., Triebkorn, P., Rabuffo, G., Fousek, J., Ziaemehr, A., Sip, V., Athanasiadis, A., Breyton, M., Woodman, M., Wang, H., Petkoski, S., Sorrentino, P., & Jirsa, V. (2025). Principles and operation of virtual brain twins. *IEEE Reviews in Biomedical Engineering*, 1–29. <https://doi.org/10.1109/RBME.2025.3562951>
- Hashemi, M., Vattikonda, A. N., Jha, J., Sip, V., Woodman, M. M., Bartolomei, F., & Jirsa, V. K. (2023). Amortized Bayesian inference on generative dynamical network models of epilepsy using deep neural density estimators. *Neural Networks*, 163, 178–194. <https://doi.org/10.1016/j.neunet.2023.03.040>
- Hashemi, M., Vattikonda, A., Sip, V., Guye, M., Bartolomei, F., Woodman, M. M., & Jirsa, V. K. (2020). The Bayesian virtual epileptic patient: A probabilistic framework designed to infer the spatial map of epileptogenicity in a personalized large-scale brain model of epilepsy spread. *NeuroImage*, 217, 116839. <https://doi.org/10.1016/j.neuroimage.2020.116839>
- Hashemi, M., Ziaemehr, A., Woodman, M. M., Fousek, J., Petkoski, S., & Jirsa, V. K. (2024). Simulation-based inference on virtual brain models of disorders. *Machine Learning: Science and Technology*, 5(3), 035019. <https://doi.org/10.1088/2632-2153/ad6230>
- Jansen, B. H., & Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73(4), 357–366. <https://doi.org/10.1007/BF00199471>
- Jha, J., Hashemi, M., Vattikonda, A. N., Wang, H., & Jirsa, V. (2022). Fully Bayesian estimation of virtual brain parameters with self-tuning Hamiltonian Monte Carlo. *Machine Learning: Science and Technology*, 3(3), 035016. <https://doi.org/10.1088/2632-2153/ac9037>
- Jirsa, V. K., Stacey, W. C., Quilichini, P. P., Ivanov, A. I., & Bernard, C. (2014). On the nature of seizure dynamics. *Brain*, 137(8), 2210–2230. <https://doi.org/10.1093/brain/awu133>
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J. B., Grout, J., Corlay, S., & others. (2016). Jupyter notebooks—a publishing format for reproducible computational workflows. *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, 87–90.
- Mazzara, C., Ziaemehr, A., Troisi Lopez, E., Cipriano, L., Angiolelli, M., Sparaco, M., Quarantelli, M., Granata, C., Sorrentino, G., Hashemi, M., & others. (2025). Mapping brain lesions to conduction delays: The next step for personalized brain models in multiple sclerosis. *Human Brain Mapping*, 46(7), e70219. <https://doi.org/10.1002/hbm.70219>
- Montbrió, E., Pazó, D., & Roxin, A. (2015). Macroscopic description for networks of spiking neurons. *Physical Review X*, 5(2), 021028. <https://doi.org/10.1103/PhysRevX.5.021028>

- Monteverdi, A., Cotta Ramusino, M., Conca, F., Augello, A., Totaro, C., Grasso, P. A., Castelnovo, A., Lorenzi, R. M., Terzaghi, M., Farina, L. M., Costa, A., Pichiecchio, A., & others. (2026). Virtual brain and electroencephalography explain the variance of memory alterations in mild cognitive impairment. *Alzheimer's Research & Therapy*. <https://doi.org/10.1186/s13195-026-02114-4>
- OpenAI. (2025). *GPT-5 system card*. <https://openai.com/index/gpt-5-system-card/>.
- Papamakarios, G., Pavlakou, T., & Murray, I. (2017). Masked autoregressive flow for density estimation. *Advances in Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.1705.07057>
- Petkoski, S., & Jirsa, V. K. (2019). Transmission time delays organize the brain network synchronization. *Philosophical Transactions of the Royal Society A*, 377(2153), 20180132. <https://doi.org/10.1098/rsta.2018.0132>
- Rabuffo, G., Lokossou, H.-A., Li, Z., Ziaee-Mehr, A., Hashemi, M., Quilichini, P. P., Ghestem, A., Arab, O., Esclapez, M., Verma, P., & others. (2025). Mapping global brain reconfigurations following local targeted manipulations. *Proceedings of the National Academy of Sciences*, 122(16), e2405706122. <https://doi.org/10.1073/pnas.2405706122>
- Sanz Leon, P., Knock, S. A., Woodman, M. M., Domide, L., Mersmann, J., McIntosh, A. R., & Jirsa, V. (2013). The virtual brain: A simulator of primate brain network dynamics. *Frontiers in Neuroinformatics*, 7, 10. <https://doi.org/10.3389/fninf.2013.00010>
- Sisson, S. A., Fan, Y., & Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6), 1760–1765. <https://doi.org/10.1073/pnas.0607208104>
- Sorrentino, P., Pathak, A., Ziaee-mehr, A., Troisi Lopez, E., Cipriano, L., Romano, A., Sparaco, M., Quarantelli, M., Banerjee, A., Sorrentino, G., & others. (2024). The virtual multiple sclerosis patient. *iScience*, 27(7), 110101. <https://doi.org/10.1016/j.isci.2024.110101>
- Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1), 1–24. [https://doi.org/10.1016/S0006-3495\(72\)86068-5](https://doi.org/10.1016/S0006-3495(72)86068-5)
- Ziaee-mehr, A., Woodman, M., Domide, L., Petkoski, S., Jirsa, V., & Hashemi, M. (2025). Virtual Brain Inference (VBI), a flexible and integrative toolkit for efficient probabilistic inference on whole-brain models. *eLife*, 14, RP106194. <https://doi.org/10.7554/eLife.106194.1>
- Ziaee-mehr, A., Zarei, M., Valizadeh, A., & Mirasso, C. R. (2020). Frequency-dependent organization of the brain's functional network through delayed-interactions. *Neural Networks*, 132, 155–165. <https://doi.org/10.1016/j.neunet.2020.08.003>