

Machine and Human Intelligence research group ([website](#))
Department of Computer Science, University of Helsinki
Finnish Center for Artificial Intelligence FCAI

The Machine and Human Intelligence research group focuses on **probabilistic machine learning** applied to statistical model fitting of complex models in the computational sciences. We are interested in smart probabilistic algorithms that are robust and sample-efficient. Our research is regularly published in top-tier machine learning conferences, such as *NeurIPS* and *ICML*, and in renowned computational journals, such as *PLoS Computational Biology* (see [Publications](#)). The group is part of the [Department of Computer Science](#) of the University of Helsinki and of the [Finnish Center for Artificial Intelligence FCAI](#).

- **Skills required:** Linear algebra, (Bayesian) statistics and probability, coding (Python or MATLAB).
- **Supervisor:** Luigi Acerbi ([website](#)); **Contact:** luigi.acerbi@helsinki.fi

Research projects in sample-efficient probabilistic machine learning:

We develop probabilistic machine learning methods to perform *optimization* and *approximate Bayesian inference* with complex scientific models, with applications largely in (but not limited to) computational and cognitive neuroscience. To work with real models and data, our algorithms are *robust* to noise (e.g., due to Monte Carlo approximations or simulations) and *sample-efficient* in that they require a relatively small number of function evaluations with respect to traditional methods. Our algorithms are released as well-documented toolboxes (see our [Resources](#) page and the [lab GitHub](#)).

Potential research projects are related to extensions and applications of our algorithms:

1. **Variational Bayesian Monte Carlo (VBMC).** [VBMC](#) is a new approach to Bayesian inference that obtains good approximations of the posterior and model evidence with a small number of likelihood evaluations ([Acerbi, 2018; NeurIPS](#)). The framework is currently being expanded in various directions, notably with the addition of support for noisy log-likelihood evaluations, such as those estimated via simulation ([Acerbi, 2020; NeurIPS](#)). We have also worked on applying surrogate modelling and active learning to the “embarrassingly parallel” setting ([de Souza, Mesquita, Kaski & Acerbi, 2022; AISTATS](#)). A Python version (PyVBMC) is about to be released ([docs](#)).
2. Fast hybrid Bayesian optimization for model fitting via **Bayesian Adaptive Direct Search (BADS)**. BADS is a fast optimization algorithm that combines a model-free approach (mesh-adaptive direct search) with a strong model-based algorithm (Bayesian optimization), achieving the best of both worlds, that is sample-efficiency and robustness to noise ([Acerbi & Ma, 2017; NeurIPS](#)). [BADS](#) is currently used by dozens of labs across the world. A Python version (PyBADS) is about to be released.
3. Estimation of log-likelihoods for simulator-based models via **inverse binomial sampling (IBS)**. [IBS](#) is an efficient statistical technique to estimate the log-likelihood when the log-likelihood is unavailable, but we can generate simulated data from the model. Unlike other “likelihood-free” methods, IBS does not use summary statistics, but computes the log-likelihood of the full data set ([van Opheusden*, Acerbi* & Ma, PLoS Computational Biology, 2020](#)). Moreover, IBS combines very well with BADS and VBMC.