Saifuddin Syed

saifsyed.com | saifuddin.syed@stats.ox.ac.uk | +44 7467 304999

Dear IAIFI selection committee, October 11, 2022

I am writing to apply for the IAIFI fellowship program starting in September 2023. I am a postdoctoral research assistant in computational statistics and statistical machine learning at the University of Oxford supervised by Arnaud Doucet. Before this, I completed a doctorate in statistics at the University of British Columbia in February 2022 supervised by Alexandre Bouchard-Côté. I am interested in becoming an IAIFI fellow as my multidisciplinary research aims to blur the boundary between computational statistics and computational physics literature.

My research focuses on designing Markov Chain Monte Carlo (MCMC) motivated by physics for scalable Bayesian posterior inference. MCMC methods are among the most widely used tools in computational science. Despite their ubiquity, they fail to converge reliably when simulating distributions with multiple, well-separated modes. Parallel tempering (PT) was developed independently by computational physicists and statisticians in the 1990s to address this issue. PT involves running multiple MCMC chains in parallel with respect to tempered versions of the distribution of interest where at ``higher temperatures'' convergence is reliable.

Traditionally PT relied on a ``reversibility'' assumption, which meant PT deteriorated in performance if too many parallel chains were used. This made the performance of PT fragile for challenging problems and limited its scalability. My doctoral thesis coalesced the statistics and physics literature. It led to the development of a ``non-reversible'' variant of PT (NRPT) that provably dominates the performance of its reversible counterpart. We optimally tuned NRPT and showed that the performance improves with the number of parallel chains, making it scalable to GPUs. Our work led to publications in the Journal of the Royal Statistics Society Series B and top machine learning conferences such as NeurIPS and ICML.

NRPT is currently the state-of-the-art MCMC algorithm for Bayesian posterior inference and is the backbone of the inference engine for large-scale international research projects. For example, the BC Cancer research center and Memorial Sloan Kettering Cancer Center have adopted our methodology to model the evolution of single-cell cancer genomes (published in Nature). In collaboration with Google research, TAE Technologies Inc used NRPT to infer plasma dynamics inside the world's largest FRC nuclear fusion reactor. Most notably, the computational advancements from NRPT were essential for the Event Horizon Telescope (EHT) collaboration to discover magnetic polarization in the M87 photograph in 2021 and generate the recent picture of Sagittarius A*.

Following the success of NRPT with the EHT, I was asked to join the algorithms and inference working group in the Next Generation Event Horizon Telescope (ngEHT) collaboration as the team's only statistician. In line with the goals of IAIFI, I am currently collaborating with cosmologists and computational physicists at the Black Hole Institute and Center for Astrophysics based in Harvard to develop methodology to model and image active galactic nuclei (AGN). This collaboration has already led to the development of a model for the time evolution of AGN. We are applying these methods to EHT data and will publish our preliminary results in the coming months.

For my Postdoc with the IAIFI institute, I would like to continue working on developing physics-inspired MCMC methodology and work with the ngEHT and other experimental physicists to help address their computational challenges. I appreciate your consideration and look forward to hearing from you soon.

Sincerely,
Dr. Saifuddin Syed