Prospective PhD — Joint modeling of informative EHR collection & target-trial emulation with Neural SDEs

Dear Prof. Banerjee,

My name is Yuting Fang (Student ID: 5518340). I am currently completing a Bachelor of Mathematics & Statistics (Honours) at UNSW Sydney, supervised by Professor Quoc Le Gia (School of Maths&Stats) and Professor Flora Salim(School of Computer Science and Engineering). My thesis centers on Neural Stochastic Differential Equations (Neural SDEs) for continuous-time dynamics, with applications to irregularly sampled medical and neuroimaging time series. I’m aiming to pursue a computational neuroscience PhD (2026 entry) and would be honoured to explore opportunities in your group.

I’ve been following your recent directions on (i) EHR data-generation/continuity and the bias it introduces in pharmacoepidemiology, and (ii) trial methodology for complex endpoints (e.g., win-ratio for hierarchical endpoints in deprescribing). Your paper on a data continuity prediction algorithm for EHR-based studies highlighted how modeling “whether/when records appear” improves validity, and your team’s win-ratio work underscores pragmatic trial designs aligned to clinical priorities.

1. Informative collection × continuous-time joint modeling.  
   Develop a Neural SDE that jointly models the latent clinical state and the measurement process (who gets which test, and when) via a coupled point-process. By letting “tested/not-tested & timing” inform drift/diffusion and observation noise, we aim to improve calibration, counterfactual robustness, and decision utility under NMAR/informative sampling.
2. Causal dynamic effects: target-trial emulation × SDE states.  
   Emulate trials for drugs or digital interventions from EHR, estimating time-varying treatment effects with g-methods (AIPW/TMLE/IPTW) while representing patient trajectories with Neural SDE latent states; report transportability across sites/ancestries with standardized diagnostics.

My background equips me to contribute effectively to these directions: I have a strong foundation in mathematics and statistics, hands-on experience with Neural SDE implementation and time-series inference, and a track record of applying rigorous evaluation methods—including held-out prediction, counterfactual validation, and calibrated uncertainty analysis. Beyond technical skills, I bring genuine enthusiasm for developing high-quality, falsifiable models that are tightly integrated with experimental design and validation.

I am confident that my proficiency in PyTorch, torchsde/torchdyn, probabilistic modeling, and data analysis (including tools like Suite2p and CaImAn) will allow me to make meaningful contributions to these research topics. I would be very grateful for the opportunity to discuss potential PhD projects with you—whether as a thesis student, visiting researcher, or intern. I am happy to arrange an online discussion via any online platform at your convenience, and I can also share my CV, academic transcript, and a detailed research outline beforehand for your reference.

Thank you for your time and consideration. I greatly admire the interdisciplinary rigor of your lab and would be honored to contribute to your research.

Warm regards,  
Yuting Fang

Bachelor of Science with Honors

Major in Mathematics, Minor in Statistics  
Honours in Statistics, UNSW Sydney  
Student Email: [z5518340@ad.unsw.edu.au](mailto:z5518340@ad.unsw.edu.au)

Personal Email: ytfang23@gmail.com