Template-Free MCMC Reconstruction of Single Molecule Imaging Data with Multiple Distinct Structures

Supervisors : Dr Hannah Mitchell and STFC

Single Molecule Super-Resolution imaging is revolutionizing our understanding of biological processes at sub-100nm scales. Techniques like FLImP (Fluorescence Localisation Imaging with Photobleaching) and MINFLUX (MINimal emission FLUorescence miCroscopy) offer sub-nm localization precision. However, challenges remain in fully utilizing these technologies to unravel the molecular structure of heterogeneous membrane protein systems.

Key issues in interpreting single molecule imaging data include incomplete labelling, spurious detections (clutter), repeated localizations of the same molecule, and structural heterogeneity, where multiple distinct or nested structures, or subtle variations within a population, are present. While template-led approaches have made significant strides in structural interpretation, they risk producing false patterns in noisy, incomplete data—a phenomenon known as “Einstein from noise”. Early attempts using deep learning, while promising, lack statistical robustness and explain-ability.

To address these challenges, the team at STFC have developed a Markov Chain Monte Carlo (MCMC) algorithm to identify discrete separations between EGFR oligomers at the plasma membrane (Iyer et al., 2024). The STFC team have extensive experience with EGFR in both ligand-bound (Needham et al., Nat Comms 2016) and ligand-unbound (Zanetti-Domingues et al., Nat Comms 2018) states, providing insight into EGFR auto-inhibition, phosphorylation, and signalling. Extending this method to 2D and 3D molecular reconstruction, where data consist of complex sets of separations, is challenging due to the presence of multiple, often nested, structures within heterogeneous protein systems.

This project extends our 1D MCMC algorithm to 2D and 3D reconstruction using simulations, DNA origami, and established membrane protein systems (e.g., Nuclopores, EGFR). For example, by incorporating Hierarchical and Reversible Jump MCMC into the assignment of separation sets, we aim to create a more robust solution for reconstructing molecular structures in 2D and 3D. This advancement could transform single molecule image analysis, moving beyond heuristic template matching.

Beyond its immediate application to EGFR oligomerization, this work will enhance statistical tools for the wider Single Molecule Imaging user community at STFC, advancing molecular imaging analysis techniques.