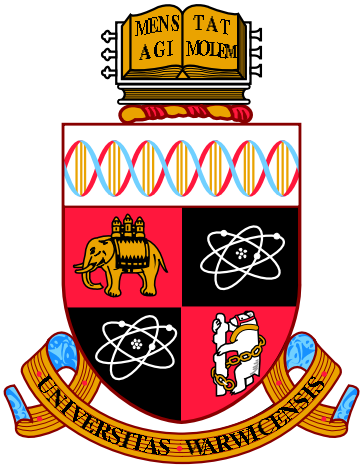

Characterising Cell Membrane Heterogeneity through Analysis of Particle Trajectories

by

Paddy J. Slator

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Systems Biology Doctoral Training Centre

December 2015

Contents

List of Tables

List of Figures

Acknowledgments

Firstly, and most importantly, thanks to Nigel Burroughs for over three years of supervision on this project. Thanks also to Christian Eggeling for many helpful discussions and suggestions, and critically reading sections of the thesis. Thanks to Chris Cairo and Philipp Kukura, both for sharing experimental data, and helping in the analysis and interpretation of that data. Thanks to Rob Cross, Rebecca Notman and Gareth Roberts for sitting on my advisory panel. Thanks to my examiners, David Wild and Andrew Golightly, for their comments and corrections. Thanks to everybody at the Warwick Systems Biology Centre, especially Anne Maynard and Brent Kiernan for always helping with various admin tasks. Finally, a big thanks to everybody in the Systems Biology DTC 2011 cohort - too many names to list here!

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented (including data generated and data analysis) was carried out by the author (PJS) except in the cases outlined below:

- Chapter 2 was written in collaboration with Chris Cairo (CWC) and Nigel Burroughs (NJB). CWC conceived, designed and performed the experiments. PJS and NJB designed the algorithm used in analysis, and PJS wrote the software. The chapter was mainly written by PJS, with contributions from NJB and CWC.

- The dataset analysed in Chapter 3 was provided by Philipp Kukura (PK). PJS and NJB designed the algorithm used in analysis, and PJS wrote the software. The chapter was mainly written by PJS, with contributions from NJB.

Parts of this thesis have been published or submitted for publication by the author:

- Chapter 2 has been previously published as: **Slator, P. J.**, Cairo, C. W. & Burroughs, N. J. Detection of Diffusion Heterogeneity in Single Particle Tracking Trajectories Using a Hidden Markov Model with Measurement Noise Propagation. PLOS ONE 10, e0140759 (2015).

- Chapter 3 has been submitted to PLOS Computational Biology as: **Slator, P. J.** & Burroughs, N. J. A Confinement Hidden Markov Model Analysis Identifies Tag-Specific Signatures within Single Particle Tracking Trajectories.

Code Availability

The relevant and applicable Matlab code developed for this thesis (e.g. algorithms for the analysis of single particle tracking data) is freely available at https://github.com/PaddySlator.

Abstract

Single particle tracking (SPT) trajectories are fundamentally stochastic, which makes the extraction of robust biological conclusions difficult. This is especially the case when trying to detect heterogeneous movement of molecules in the plasma membrane. This heterogeneity could be due to a number of biophysical processes such as: receptor clustering, traversing lipid microdomains or cytoskeletal barriers.

Working in a Bayesian framework, we developed multiple hidden Markov models for heterogeneity, such as confinement in a harmonic potential well, switching between diffusion coefficients, and diffusion in a fenced environment (or "hop" diffusion). We implement these models using a Markov chain Monte Carlo (MCMC) methodology, developing algorithms that infer model parameters and hidden states from single trajectories. We also calculate model selection statistics, to determine the most likely model given the trajectory.

For LFA-1 receptors diffusing on T cells we show that 12-26% of trajectories display clear switching between diffusive states, depending on treatment. We also demonstrated that allowing for measurement noise is essential, as otherwise false detection of heterogeneity may be observed. Analysis of the motion of GM1 lipids bound to the cholera toxin B subunit (CTxB) in model membranes confirmed transient confinement. On this dataset we also demonstrated a clear signature in the confinement shape for individual tagging molecules, and showed that confinement events are not exponentially distributed. Finally, we developed an algorithm which detects hopping diffusion, validating on simulated data.

Rather than methods which rely on generic properties of Brownian motions, our approach allows us to test which biophysical model best fits a trajectory. Using a model-based approach we can also extract biophysical parameters, segment trajectories into different motion states, and hence analyse particle motion in high detail. With the continuing improvement in spatial and temporal resolution of trajectories, these methods will be important for biological interpretation of SPT experiments.

Abbreviations

AuNP	Gold nanoparticle
CTxB	Cholera toxin B subunit
FCS	Fluorescence correlation spectroscopy
FLIP	Fluorescence loss in photobleaching
FPT	First passage time
FRAP	Fluorescence recovery after photobleaching
HMM	Hidden Markov model
HPW	Harmonic potential well
IQR	Interquartile range
iSCAT	Interferometric scattering
LFA-1	Lymphocyte function-associated antigen 1
MCMC	Markov chain Monte Carlo
MH	Metropolis-Hastings
MSD	Mean square displacement
PDF	Probability density function
PMA	Phorbol-12-myristate-13-acetate
PMF	Probability mass function
PSRF	Point scale reduction factor
Q-Q	Quantile-quantile
SDE	Stochastic differential equation
S/N	Signal to noise ratio
SPT	Single particle tracking

Chapter 1

Introduction

1.1 Background

Cell membranes are vitally important for cell function, providing the link between the cytoplasmic and extracellular environments. They are the site of many fundamental biochemical processes such as signalling, trafficking and cell adhesion. In order to facilitate these processes cells tightly regulate membrane structure, but many of the exact mechanisms remain unknown. It follows that to gain understanding of membrane function, it is necessary to understand membrane structure.

The fluid mosaic structure of the plasma membrane was proposed by Singer and Nicolson in 1972 [1], who argued that the membrane is composed of a two-dimensional phospholipid bilayer which is fluid and dynamic. Within this bilayer multiple integral proteins, which may form aggregates, are abundantly distributed, hence "mosaic" (Fig. 1.1, left panel). Whilst this view of membranes remains broadly accepted, there have been multiple refinements to the model, in particular it is clear that the lateral diffusion of proteins and lipids is not uniform [2, 3, 4].

Diffusion of membrane components has been extensively studied by population level experiments, such as fluorescence recovery after photobleaching (FRAP) and fluorescence loss in photobleaching (FLIP) [5]. These methods by definition average over any heterogeneity within the population. Other membrane techniques are highly invasive (such as detergent extraction), or focus on artificial membranes. The advent of single molecule techniques addresses many of the shortcomings of these approaches - there is no need to average over particle populations, and minimally invasive experiments are possible. For instance, methods such as single particle tracking (SPT) and fluorescence correlation spectroscopy (FCS), have allowed direct characterisation of the motion of

single molecules in live cell membranes [6, 7].

However, moving from populations to single molecules brings new challenges. Experimental measurements are no longer ensemble averaged, but rather a single realisation of a stochastic system. Thus, in order to draw robust conclusions from single molecule data, robust statistics are necessary. This thesis focuses on the analysis of SPT trajectories, for which there are a number of existing methods. Many of these, such as those based on mean square displacements (MSD), are time-averaged in order to increase statistical power. This is an attractive approach in many situations, such as calculating biophysical parameters in experiments with lots of short trajectories; yet with the increasing availability of very long trajectories at high spatial and temporal resolution, methods which can extract finer details from data (such as switches between different motion modalities) are possible. An attractive proposition is a model based approach, which allows parameterisation of an underlying mechanistic model. This brings an increase in statistical power, allowing the extraction of more information from SPT data and hence increasing biophysical knowledge. This powerful approach has been successfully demonstrated using Hidden Markov models (HMMs), where heterogeneity within a trajectory is modelled as a hidden state variable [8, 9, 10, 11, 12].

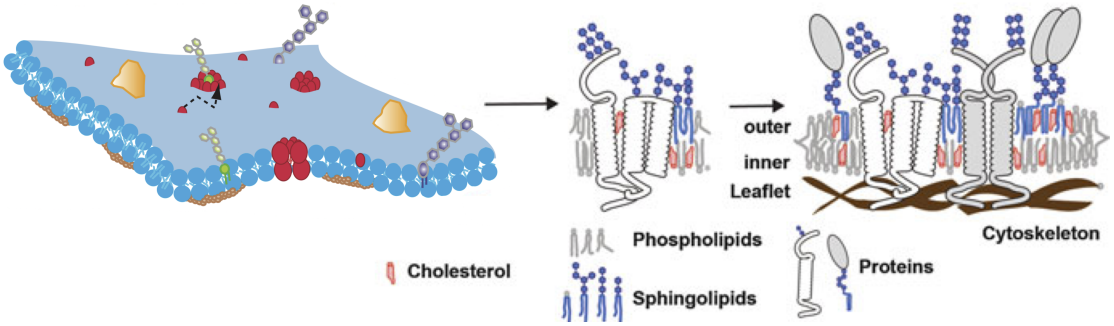

Figure 1.1: **Schematic diagram of plasma membrane structures.** Left panel: plasma membrane showing phospholipid bilayer (blue), cholesterol (red), transmembrane proteins (yellow), underlying actin cytoskeleton (brown), sphingolipids (green), and other glycolipids (purple). Right panel: Lipid microdomains or "rafts". Reproduced with permission from [13].

1.2 Membrane structure and heterogenous diffusion

The plasma membrane essentially consists of a hierarchical structure on the 2-300nm scale [4, 14]. This causes a wide variety of membrane molecules to undergo heterogeneous motion, which is observable in single molecule experiments [15, 16]. Depending on specific membrane organisation, this heterogeneous motion may take many forms. We now explore some examples of membrane structures, and their effect on the diffusion of single molecules.

1.2.1 Lipid microdomains

There are hundreds of types of lipids residing in the cell membrane, which suggests that each one may play a role in membrane function, structure and dynamics [4]. For example, lipid organisation could directly affect protein organisation through hydrophobic match sorting [17, 18]. Lipid microdomains (or "rafts") are one proposed structure for membrane organisation (Fig. 1.1, right panel). These are functional nanoscale assemblies of sphingolipids, cholesterol and proteins, and have been implicated in signalling and trafficking [19, 20, 21]. However their specific nature, including size, exact composition and lifetime are not known [18], with size estimates varying across techniques [22]. The existence of self-organising "rafts" is controversial, being in part based on the separation of saturated and unsaturated lipids into liquid-ordered and liquid-disordered phases in artificial membranes [23]. The applicability of this observation to live cell membranes is not clear, one reason being that lipids diffuse a factor of 5-100 times faster in model membranes than cell membranes [24]. Cell membranes also contain much higher protein density [25]. Other evidence for phase separation comes from highly invasive experiments, such as the extraction of detergent resistant membrane sections [26].

Cholesterol mediated trapping of sphingolipids on scales less than 20nm has been directly observed [27, 28], yet there have been alternative explanations proposed for the heterogeneous distribution of these molecules, such as membrane topography [29] or specific protein-lipid interactions [30]. Non-invasive techniques, such as SPT, will continue to be important in order to resolve these questions about lipid microdomains.

1.2.2 Receptor clustering

Receptor proteins are mobile in cell membranes [31], and generally exist in large clusters [32, 33]. This can lead to heterogeneous motion, since the radius of a cluster affects the rate of diffusion [34]. This clustering is implicated in cellular recognition, signalling [35],

cell adhesion [36, 37] and chemotaxis [38]. For example, the formation of large signalling rafts may cause signal amplification through sharing of downstream elements.

1.2.3 Membrane compartmentalisation

Direct observation has shown that some membrane molecules demonstrate so-called compartmentalised or hopping diffusion: short term confinement within a domain, and occasional moves (or hops) between adjacent compartments [39, 24, 40]. Clearly the diffusion of large transmembrane proteins could be hindered by the cytoskeleton, which resides on the cytoplasmic side of the membrane and builds up a meshwork-like or compartmentalised network immediately beneath the plasma membrane [41]. However, more surprisingly, hop diffusion of lipid molecules in the outer leaflet has also been observed [42, 24, 43]. An explanation for both observations is provided by the fences and pickets model, where diffusing molecules are trapped inside 40-300nm diameter compartments by the actin cytoskeleton "fence" or cytoskeletal anchored protein "pickets" [3, 14], Fig. 1.2. Molecules can occasionally hop between compartments, perhaps through a gap in the "fence" caused by brief separation of the cytoskeleton from the lower leaflet of the bilayer.

Compartmentalisation is not independent of other membrane mechanisms. For example an actin meshwork can induce lipid phase separation in model membranes [44], and there are also interactions between the cytoskeleton and signalling domains [45, 46].

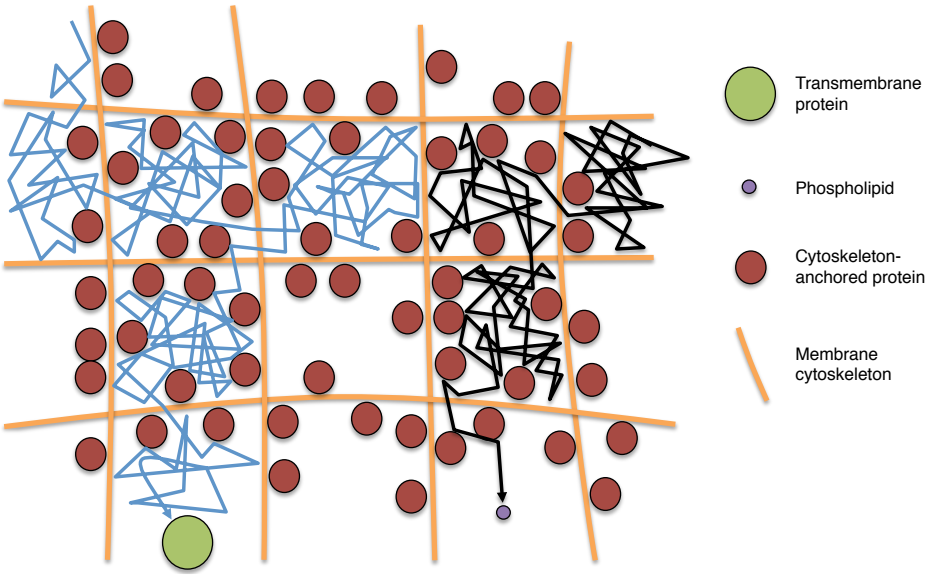

Figure 1.2: **Top-down view of the "fences and pickets" plasma membrane model.** Compartmentalised (or "hop") diffusion is caused by membrane cytoskeleton "fences" and anchored protein "pickets". Adapted from [14].

1.2.4 Role of membrane structure

There are many reasons why cells laterally organise membrane components; in short, membrane complexity allows for high level function and regulation. One example is the enrichment of specific molecules in cell membrane regions such as the neuronal [31, 47] and immunological [48] synapses. Receptor motility has evolved in order to faithfully and efficiently communicate noisy external signals [49], and compartmentalisation may align these with intracellular signalling complexes [24]. Lateral organisation may also increase the probability of reactions between less abundant molecules [13].

1.3 Single particle tracking

Single particle tracking is a powerful technique for investigating the organisation of cell membranes, through direct observation of the motion of membrane components. In an SPT experiment the molecule of interest has an observable tag attached, allowing tracking of the tag's two-dimensional position over a number of time steps. The tag is usually imaged optically with a camera or point detectors [15], although a scattering microscopy approach has been demonstrated [50, 51].

Possible tags include a gold nanoparticle (typical diameter 20-40nm) [50, 51, 39], a quantum dot (10nm) [52, 53, 54, 42, 55], a latex bead (40-1000nm) [56, 57, 58], or a single fluorescent molecule [27]. Gold nanoparticle, quantum dot, and latex bead experiments can image the particle at high temporal resolution over a long period, yet the tags are large relative to the molecules they label. There are concerns about how this affects the tag-target complex - it has been shown that nanoparticle size can affect the value of diffusion coefficients [58]. Tagging with single fluorescent molecules, such as fluorophores or fluorescent lipid analogues, is an attractive alternative, but trajectories are much shorter due to photobleaching. Quantum dots are often cited as the ideal compromise due to their small size compared to other tagging molecules, yet they are still larger than a single fluorescent molecule. They are also susceptible to temporary loss of fluorescence or "blinking", although this may be avoidable in the future [59].

1.4 Analysis of single particle tracking data

Whilst SPT can directly observe the motion of membrane molecules, this motion is stochastic; hence an SPT trajectory only provides a single realisation of the system. Thus, directly drawing conclusions from SPT data is not always possible - consideration of the stochasticity of the data is necessary. In other words, statistical methods are required to detect true heterogeneity and avoid overinterpretation of Brownian motion.

We define a single particle tracking trajectory $\mathbf{X}$. Let X_i, where $X_i = (X_{1i}, X_{2i})$ is two dimensional, be the measured particle positions in an experimental SPT trajectory (or equivalently a simulated trajectory) and t_i be the times of these measurements. In other words, we have a time series of the 2D position of the particle. For analysis we consider the displacements

$$\mathbf{X} = \{\Delta X_i, \Delta t_i\}_{i=1}^{N} \tag{1.1}$$

where $\Delta X_i = X_{i+1} - X_i$ and $\Delta t_i = t_{i+1} - t_i$.

1.4.1 Mean squared displacement analysis

A useful approach to analysing SPT trajectories is to utilise the squared displacements. Let $\Delta X_{i,n} = X_{i+n} - X_i$ and $\Delta t_{i,n} = t_{i+n} - t_i$ be displacements and time steps of length n from start point i. We also assume that all time steps are equal, and hence denote $\Delta t_n = \Delta t_{1,n}$. The mean squared displacement (MSD) function is an estimator of the

theoretical MSD value, given by

$$\mathrm{MSD}(\Delta t_n) = \frac{1}{N-n} \sum_{i=1}^{N-n} \Delta X_{i,n}^2 \tag{1.2}$$

where $\Delta X_{i,n}^2 = (X_{1i+n} - X_{1i})^2 + (X_{2i+n} - X_{2i})^2$ is the squared Euclidean distance. Thus, the MSD time averages over the trajectory length in order to improve the estimate of the average squared displacement. For an unconfined Brownian motion with diffusion coefficient D, the theoretical value of the MSD is linear as a function of time [60, 61]

$$\mathrm{MSD}(\Delta t_n) = 4D\Delta t_n. \tag{1.3}$$

The localisation accuracy, σ^2, is defined as the variance of an independent additive measurement error. If this is non-zero then there is a shift in the MSD function [62]

$$\mathrm{MSD}(\Delta t_n) = 4D\Delta t_n + 4\sigma^2. \tag{1.4}$$

Given an SPT trajectory and assuming an unconfined diffusion model, a straight line can be fitted to the MSD function. The linear fit is then analysed for interpretable features such as the linear gradient (diffusion coefficient, D) and intercept (localisation accuracy, σ^2).

Number of timepoints to include in MSD calculation

For the first timestep the MSD function is an average of $N-1$ measurements, for the second timestep $N-2$ measurements, and so on. Thus, the standard error of $\mathrm{MSD}(\Delta t_n)$ increases with increasing Δt_n. For this reason, only a proportion of the MSD function is used for the linear fit, but determining this proportion is subjective.

Michalet and Berglund provided a theoretical framework for determining the optimum number of points of the MSD function to include when estimating D [62, 63, 64]. For example, if the optimum number of MSD points is two, then the optimal D estimate comes from a linear fit to the first two MSD points. Failure to use the optimum number of MSD points can lead to D estimates which are incorrect by many orders of magnitude [62, 63].

MSD for anomalous diffusion

The MSD analysis techniques introduced thus far have assumed that the particle is freely diffusing. However, as noted earlier, many membrane components do not freely diffuse. Thus, techniques that detect deviations from Brownian motion have been developed. For anomalous diffusion the MSD has the form [65]

$$\mathrm{MSD}(\Delta t_n) = 4D\Delta t_n^{\alpha} + 4\sigma^2. \tag{1.5}$$

The anomalous exponent can be estimated from the experimentally measured MSD function [66, 67, 68], with a negative deflection ($\alpha < 1$) suggesting movement in a confined environment, and a positive deflection ($\alpha > 1$) suggesting directed motion. There are a number of possible physical models which reproduce anomalous diffusion, including fractional Brownian motion and continuous-time random walks [69, 70], but the parameter α has no direct physical meaning [42].

Fitting models to the MSD function

For a number of anomalous diffusion models, an analytical form of the MSD function has been derived [71, 72]. For example, the MSD function for diffusion confined to a circle of radius R is [73]

$$\mathrm{MSD}(\Delta t_n) = R^2\left[1 - \exp\left(-\frac{4D\Delta t_n}{R^2}\right)\right] + 4\sigma^2. \tag{1.6}$$

By comparing an experimentally measured MSD function to a theoretical curve, the model which best explains the observed trajectory's displacements can be ascertained. Clearly a statistical approach is required, since MSD functions are noisy, and generally only a small fraction of the trajectory can be used. A few approaches have been demonstrated, such as using an F-test [42] or a Bayesian approach [74] to determine which diffusion model best describes the experimental MSD function.

1.4.2 Single trajectory analysis

Time averaged methods are advantageous if the dataset contains many short trajectories, as more information can be extracted by pooling all trajectories. However, this approach averages over any heterogeneity within trajectories. If trajectories are long enough then transient, within trajectory heterogeneity can potentially be detected. Rajani *et al.* formalised this distinction by defining micro-heterogeneity (anomalies within a

single trajectory) and macro-heterogeneity (anomalies across multiple trajectories) [75]. Micro-heterogeneity may be caused by transient switching between motion models. For example, there are methods which detect changes in the diffusion coefficient within a single trajectory, either with [76], or without [77], time averaging. There are other methods which are capable of detecting other diffusion modes, such as directed motion. One approach utilises additional information such as MSD curvature and trajectory asymmetry [78], whilst another uses supervised support vector classification [79].

Confinement

Many methods for single trajectory analysis focus on detecting periods of confinement, since within this approach a number of mechanistic models are detectable. These include: changes in the diffusion coefficient (a lower diffusion coefficient causing apparent confinement), confinement due to impermeable barriers, or stationarity.

Saxton derived the form of the probability that a diffusing particle will remain in a circle of radius R for time Δt_n [80]. If Ψ is this probability, the exact relationship is

$$\log \Psi = 0.2048 - 2.5117 \frac{D\Delta t_n}{R^2} \tag{1.7}$$

if $\frac{D\Delta t_n}{R} > 0.1$ holds. A threshold α is set (by comparison with Brownian motion, such that false detection of confinement is minimised) and if $\Psi < \alpha$, then the particle is considered confined.

Simson *et al.* extended this approach [81]. The trajectory is split into overlapping windows of varying length, with each timepoint the start of windows of every possible length. For each window a probability level L, based on Equation (1.7) is calculated.

$$L = \begin{cases} -\log(\Psi) - 1 & \text{if } \Psi \leq 0.1 \\ 0 & \text{if } \Psi > 0.1 \end{cases} \tag{1.8}$$

where Ψ is calculated with $R = R_{max}$, the furthest the trajectory moves from the starting timepoint within the time window. Thus, for every timepoint there is an L value for each time window size. Averaging the L's for a timepoint gives an estimate of the confinement level at that timepoint, and thus a confinement profile for the entire trajectory can be plotted. Again there are parameters (such as the maximum window length, threshold on L for confinement, and number of consecutive timepoints above the threshold required for confinement) which are set by comparison with simulated Brownian motion to minimise false positives.

Meilhac *et al.* resolved two important limitations of the method [82]. Firstly, rather than fixing the diffusion coefficient D over the entire trajectory, it is computed locally on a time window. Secondly, the variance of displacements over the window is utilised, rather than the maximum displacement R_{max}. This is a more accurate measure of the confinement size, and hence detects more general confinements, such as in quadratic potentials.

Analysis of confinement in trajectories has also been demonstrated using a first passage time (FPT) algorithm, which detects radii at which confinement is prevalent by considering the number of displacements within a fixed radius [75].

1.4.3 Hidden Markov models

A recent development in trajectory analysis is the use of hidden Markov models (HMMs). The trajectory is modelled as a (first order) Markov chain, so that it may switch between a number of motion models, with the current motion state hidden to the observer. (See Das *et al.* [8] for a discussion justifying the use of a first order Markov process.) This hidden state can potentially parameterise any motion model. These methods have the potential to extract more information from experiments, especially micro-heterogeneity, provided they capture the true motion dynamics with sufficient accuracy.

One approach is to model the various states as having different diffusion coefficients. This has been utilised to detect heterogeneity both across multiple [8, 11], and within single [9, 83, 12] trajectories, with one multiple trajectory approach including consideration of measurement error [10]. The single trajectory HMM methods by Monnier *et al.* (which take a Bayesian approach to analysis) can also detect directed diffusion, and calculate the marginal likelihood over multiple models, and hence the most likely model given the data [83, 12].

An attractive feature of this approach is flexibility, since many relevant anomalous diffusion models can be described by a HMM. They are also suitable for analysing very long trajectories at high spatial and temporal resolution, which are becoming more prevalent as SPT technology continues to develop.

1.5 Overview of statistical techniques

1.5.1 Diffusion models

The natural model for a particle (X_t) moving in a two-dimensional fluid membrane is Brownian motion, also known as free diffusion. In stochastic differential equation (SDE)

form this is

$$dX_t = \sqrt{2D}dW_t, \tag{1.9}$$

where W_t is a two dimensional Weiner process. (See Oksendal [84] or Karatzas and Shreve [85] for an introduction to stochastic calculus.) The model has a single parameter, the diffusion coefficient D. A more general diffusion model could also include an additional term μ (known as the drift coefficient), which in this case depends on X_t and parameters θ

$$dX_t = \mu(X_t, \theta)dt + \sqrt{2D}dW_t. \tag{1.10}$$

We can simulate this SDE for discrete time points $\{t_i\}_{i=1}^N$, effectively simulating a single particle tracking experiment. Given X_1 the Euler-Maruyama approximation to Equation (1.10) for $i = 1...N-1$ is [86] (we parameterise the normal PDF with the precision here and throughout the thesis)

$$X_{i+1}|X_i \sim N(X_i + \mu(X_i, \theta)\Delta t_i, (2D\Delta t_i)^{-1}). \tag{1.11}$$

The Euler-Maruyama scheme is the natural approximation in this setting due to its simplicity. The posterior distributions derived from this approximation are tractable, and hence the subsequent statistical inference is more straightforward.

1.5.2 Bayesian inference

Equation (1.11) allows us to simulate data $\mathbf{X} = \{X_i, t_i\}_{i=1}^N$ for this probability model. Statistical inference addresses the opposite problem: given data $\mathbf{X}$, and a probability model with likelihood function $\pi(\mathbf{X}|\theta)$, what are the parameters θ? Throughout this thesis, we address this question in a Bayesian framework. We thus estimate the parameters using Bayes' rule, which (up to proportionality) is

$$\pi(\theta|\mathbf{X}) \propto \pi(\theta)\pi(\mathbf{X}|\theta). \tag{1.12}$$

The posterior distribution $\pi(\theta|\mathbf{X})$ gives the probability of the parameters given the data, and $\pi(\theta)$ is the prior distribution on the parameters.

1.5.3 Hidden Markov models

In a hidden Markov model the observation X_i depends not only on the parameters θ, but also on an unobserved state z_i. The likelihood of an observation X_i, given z_i and θ, is hence $\pi(X_i|\theta, X_{i-1}, z_i)$. The sequence of hidden states $\mathbf{z} = \{z_i\}_{i=1}^N$ is a first order

Markov chain since z_{i+1} depends only on z_i (and the state transition parameters), i.e.

$$\pi(z_{i+1}|z_i, z_{i-1}, z_{i-2}, ..., z_1, \theta) = \pi(z_{i+1}|z_i, \theta). \tag{1.13}$$

For a HMM the posterior distribution is (again using Bayes' rule)

$$\pi(\theta, \mathbf{z}|\mathbf{X}) \propto \pi(\theta, \mathbf{z})\pi(\mathbf{X}|\theta, \mathbf{z}). \tag{1.14}$$

1.5.4 Markov chain Monte Carlo

For many probability models the posterior distributions are not analytically available. Markov chain Monte Carlo (MCMC) algorithms are a means of drawing samples from an arbitrary probability distribution, which we denote $\pi(\theta)$, by constructing a Markov chain whose stationary distribution is equal to the distribution of interest. For example, in this thesis MCMC algorithms allow us, given data $\mathbf{X}$, to sample from the posterior distributions of a hidden Markov model, $\pi(\theta, \mathbf{z}|\mathbf{X})$. Given these samples, $\{\theta^{(k)}, \mathbf{z}^{(k)}\}_{k=1}^{K}$, any number of relevant statistics can be calculated.

Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm is a powerful generic MCMC algorithm [87, 88]. Here, the Markov chain is constructed as follows, given the current state of the chain θ, propose a move to a new value θ' using an arbitrary proposal distribution $q(\theta \to \theta')$, and accept this move with some probability $\alpha(\theta \to \theta')$. The equilibrium distribution of this chain is $\pi(\theta)$ if the detailed balance condition is satisfied

$$\pi(\theta)q(\theta \to \theta') = \pi(\theta')q(\theta' \to \theta). \tag{1.15}$$

The Metropolis-Hastings acceptance probability, which satisfies detailed balance, is

$$\alpha(\theta \to \theta') = \min\left\{1, \frac{\pi(\theta')q(\theta' \to \theta)}{\pi(\theta)q(\theta \to \theta')}\right\}. \tag{1.16}$$

The efficiency of the algorithm relies heavily on the choice of proposal distribution [89]. Common proposals are a random walk, where $\theta' = \theta + u$ for some random variable u, and an independence sampler, where $q(\theta \to \theta')$ is independent of θ [90].

Gibbs algorithm

Suppose we want to sample $\pi(\theta)$, where $\theta = \{\theta_1, \theta_2, ..., \theta_M\}$ is a set of M parameters. If the conditional distribution for each parameter is available for sampling, then $\pi(\theta)$ can be sampled using the Gibbs algorithm [91]. Given the current state of the Markov chain $\theta^{(k)}$ the next state is obtained by successive simulation from the conditionals

$$\begin{aligned}
\theta_1^{(k+1)} &\sim \pi(\theta_1|\theta_2^{(k)}, \theta_3^{(k)} ..., \theta_M^{(k)}) \\
\theta_2^{(k+1)} &\sim \pi(\theta_2|\theta_1^{(k+1)}, \theta_3^{(k)} ..., \theta_M^{(k)}) \\
&\vdots \\
\theta_M^{(k+1)} &\sim \pi(\theta_M|\theta_1^{(k+1)}, ..., \theta_{M-1}^{(k+1)}).
\end{aligned}$$

These updates can be performed in any order. Additional MCMC samplers can be constructed by combining Gibbs and Metropolis-Hastings moves [92, p. 280]. In fact, the Gibbs sampler is essentially a special case of the Metropolis-Hastings. For a component-wise Metropolis-Hastings algorithm on $\theta = \{\theta_1, \theta_2, ..., \theta_M\}$, the acceptance probability for θ_1 is

$$\alpha(\theta_1 \to \theta_1') = \min\left\{1, \frac{\pi(\theta_1'|\theta_2, \theta_3, ..., \theta_M) q(\theta_1' \to \theta_1)}{\pi(\theta_1|\theta_2, \theta_3, ..., \theta_M) q(\theta_1 \to \theta_1')}\right\}. \tag{1.17}$$

Using the full conditionals as proposal distributions, i.e.

$$q(\theta_1 \to \theta_1') = \pi(\theta_1'|\theta_2, \theta_3, ..., \theta_M) \tag{1.18}$$

gives the acceptance ratio 1.

Convergence

An important consideration when using MCMC algorithms is how long the Markov chains need to be run to achieve convergence, i.e. an adequate representation of the target distribution. One issue is that if the starting point is in an area of low density, then the first samples will not represent the target distribution. For this reason, a proportion of early MCMC samples are generally discarded, known as the burn in. However, determining how many iterations to discard is not obvious [93, p. 14].

There are a number of ways of assessing convergence, one simple method being to run multiple chains and compare them. If the chains have converged then their distributions should agree. The Gelman-Rubin point scale reduction factor (PSRF) is a quantitative measure of this, requiring multiple chains from overdispersed starting

points [94, 95, 92]. In order to diagnose convergence using the PSRF two conditions are required: the chains should have mixed (i.e. they should have the same distribution), and each individual chain should be stationary [92, p. 284]. The version of the PSRF utilised in this thesis is found in reference [92]. The PSRF gives an estimate of the factor by which the scale of the posterior distribution might be reduced if the sampling continued to infinity, thus a value close to 1 indicates convergence. In practice, a value less than 1.2 is suggested as a reasonable bound for convergence [94]. For a model with multiple parameters (or values of interest) the PSRF is calculated individually for each parameter, but there is also a multivariate approach [95].

1.5.5 Model selection

A common inference problem is identifying the best model given the data $\mathbf{X}$. In a Bayesian framework, we use the marginal likelihood, the probability of the data given the model,

$$\pi(\mathbf{X}|M) = \int \mathrm{d}\theta \; \pi(\mathbf{X}|\theta, M)\pi(\theta|M). \tag{1.19}$$

Or in other words, the likelihood integrated over the parameters (or the parameters and hidden states for HMMs). Given the marginal likelihoods, the Bayes factor B_{12} is a measure of the relative strength of two competing models, M_1 and M_2,

$$B_{12} = \frac{\pi(\mathbf{X}|M_1)}{\pi(\mathbf{X}|M_2)}. \tag{1.20}$$

If $B_{12} > 1$ then M_1 is the better model, and vice versa. There are a number of standard tables for assessing the strength of the evidence for the preferred model; here we followed the Kass and Raftery approach [96].

There are methods for calculating the marginal likelihood given MCMC samples. The method of Chen utilises a single MCMC run [99], and is described in detail in Section 2.2.2 (Equations (2.22)-(2.26)). Chib's methods for Gibbs and Metropolis-Hastings output require additional MCMC runs [97, 98]. Chib's methods centre on the so-called basic marginal likelihood identity, given by

$$\pi(\mathbf{X}|M) = \frac{\pi(\mathbf{X}|\theta^*)\pi(\theta^*)}{\pi(\theta^*|\mathbf{X})} \tag{1.21}$$

where θ^* is any value, but is chosen as a point of high density for numerical stability. The right hand side of Equation (1.21) includes the likelihood, prior and "posterior ordinate" for the relevant model M. Evaluating the likelihood and prior at θ^* are usually

straightforward, so estimating the marginal likelihood often reduces to estimating the posterior ordinate. This is factorised to give

$$\pi(\theta^*|\mathbf{X}) = \pi(\theta_1^*|\mathbf{X})\pi(\theta_2^*|\theta_1^*, \mathbf{X})...\pi(\theta_N^*|\theta_1^*, \theta_2^*, ..., \theta_{N-1}^*, \mathbf{X}) \tag{1.22}$$

Estimation of the density $\pi(\theta_1^*|\mathbf{X})$ is already possible given MCMC samples $\{\theta^{(k)}\}_{k=1}^K$, specifically

$$\hat{\pi}(\theta_1^*|\mathbf{X}) = \frac{1}{K}\sum_{k=1}^{K}\pi(\theta_1^*|\theta_2^{(k)}, \theta_3^{(k)}, ..., \theta_N^{(k)}, \mathbf{X}). \tag{1.23}$$

Estimation of the subsequent terms in the right hand side of Equation (1.22) is achieved by a series of reduced MCMC runs. For example, by fixing θ_1 at θ_1^*, and sampling $\{\theta_2^{(k)}, \theta_3^{(k)}, ..., \theta_N^{(k)}\}_{k=1}^K$ with a reduced MCMC sampler (requiring no additional coding), an estimate of $\pi(\theta_2^*|\theta_1^*, \mathbf{X})$ is available (similar to Equation (1.23)). We have used both methods in this thesis as a consistency check on our algorithms. Typically we used Chen's method on a full data set due to its shorter computation time, whilst applying Chib's method on selected trajectories to test that marginal likelihood estimates from these two algorithms were the same.

Chapter 2

Detection of Diffusion Heterogeneity in Single Particle Tracking Trajectories using a Hidden Markov Model with Measurement Noise Propagation

2.1 Introduction

Single particle tracking (SPT), fluorescence recovery after photobleaching (FRAP), and fluorescence correlation spectroscopy (FCS) experiments have demonstrated that rather than moving freely, molecules in the plasma membrane tend to exhibit heterogenous motion. This heterogeneity occurs on a variety of scales, and a number of potential mechanisms have been proposed to explain the behaviour. These include: lipid microdomains [20, 28], compartmentalisation by the cytoskeleton (so-called 'hop diffusion') [3, 42], protein-protein interactions [57], and inhomogeneity in the plasma membrane environment [100]. There are a number of mechanistic models which reproduce anomalous diffusion [70]. Current thinking suggests that multiple mechanisms combine to form a hierarchical structure in the plasma membrane [14].

SPT experiments can directly observe the diffusion of lipids, proteins, and other complexes in the cell membrane, providing significant insight into membrane structure. In an SPT experiment the molecule of interest has an observable tag attached, allowing

tracking of the tag's 2D position over a number of time steps. Possible tags include a gold nanoparticle [39], a quantum dot [42], a fluorophore [27], or a latex bead [57]. Gold nanoparticle, quantum dot, and latex bead experiments can image the particle at high temporal resolution (up to 40000 frames s^{-1} [3]) over a long period (seconds). However, the tags are large relative to the molecules they label, with typical diameters of 10 nm for quantum dots [42]; 40 nm, gold nanoparticles [71]; 1000 nm, latex beads [57]. Other experiments have tracked single molecules by tagging with much smaller fluorophores but, due to photobleaching, can only track for much shorter periods [27], and thus provide shorter trajectories.

An open question is the extent to which the tracked tag represents the movement of the molecule of interest. General artifacts that may be associated with the use of a tag for SPT experiments include multivalent binding, drag, interaction with the extracellular matrix, and the binding of the label itself [73]. Additionally, experimental artifacts could result from movement of the particle out of the plane of focus or from the tracking algorithm which converts the raw video data to a trajectory. There is evidence that beads affect the estimated value of diffusion coefficients [58]. For example, results from gold nanoparticle experiments by the Kusumi lab report the presence of very fast diffusion within membrane compartments [39], much faster (by around a factor of ten) than in all other studies in the field. A possible explanation is that the nanoparticles used by Kusumi make the diffusion coefficient of the tag-target complex substantially different than that of the untagged molecule [42]. The fact that the tag is diffusing in solution whilst the molecule is in the membrane also causes a concern, potentially giving a weighted average of these two diffusion coefficients. Another possible cause of bead artifacts is crosslinking of proteins due to multivalent presentation. These issues highlight the importance of decoupling the particle behaviour from that of the tag, including dealing with experimental localisation error [72, 62, 63].

There are a number of techniques for analysing SPT data, including specific methods for the detection of deviations from free diffusion. The simplest and most common approach is to use a mean squared displacement (MSD) analysis. An unconfined random walk has a cumulative MSD that is linear as a function of time [60], whilst a negative deflection in MSD (anomalous diffusion) can be caused by movement in a confined environment, and a positive deflection suggests directed motion. MSD curves are often analysed for interpretable features such as the linear gradient (diffusion coefficient) and intercept (localisation accuracy); however, the subjectivity inherent in this approach has been suggested as the cause of discrepancies between studies [62]. Alternatively, theoretical MSD curves can be fitted to the data for various physical models

(e.g. free diffusion, confined diffusion, hop diffusion, directed motion) [71, 101, 102]. Statistical analysis can be used to determine which theoretical model best describes the experimental MSD curve [42, 74]. However, these techniques are limited since they can only detect heterogeneity across multiple trajectories.

Methods for detecting heterogeneity within single trajectories (or 'microheterogeneity' [75]) have also been developed, most utilising statistics that detect deviations from random walk behaviour. For example, local periods of confinement can be detected by particles spending a significantly longer period of time within a fixed circle than a random walk [80, 81, 76], this has been utilised to detect trapping in experimental data [103, 53, 104]. This method has been further developed as a first passage time (FPT) analysis, which also gives an estimate of the size of confinement zones [75]. Other methods segment single trajectories based on transient changes in diffusion modalities, including detection of changes in the diffusion coefficient [77], local confinement and directed motion [78, 79]. Meilhac *et al.* [82] developed an algorithm which detects if a particle is moving between different confinement zones (i.e. exhibiting hop diffusion). The majority of these methods use generic properties of Brownian motion (random walks) to detect deviations, and thus, do not have an underlying mechanistic model. More information (with a corresponding increase in statistical power) can potentially be extracted by using a model that allows parametrisation of the heterogeneity and associated processes. Such models have been proposed in a hidden Markov model (HMM) framework. For instance, Das *et al.* developed a HMM for LFA-1 interacting with the actin cytoskeleton, where LFA-1 moves between "free" and "bound" states, moving with a different diffusion coefficient in each state [8]. Monnier developed a method which chooses between multiple modes of diffusion, such as directed motion, and diffusion with a variable diffusion coefficient [83, 12]. Persson *et al.* developed a HMM based method which takes multiple trajectories as input, and infers the number of diffusive states, the diffusion coefficients and the state transition rates [11].

Here we develop an improved single trajectory analysis, based on the two-state diffusion model of Das *et al.* [8]. We make two key changes to their analysis, firstly, we analyse each trajectory separately; the pooled analysis of [8] assumes homogeneity across trajectories, which we find is incorrect. This allows us to determine if individual trajectories have evidence for switching between two diffusive states, as opposed to remaining in one state throughout. Secondly, we allow for localisation accuracy. We demonstrate that a failure to do so can lead to the erroneous detection of a high degree of heterogeneity caused by structured measurement noise. We use a Bayesian analysis for both model parameter inference and model selection, using Markov chain Monte

Carlo (MCMC) algorithms for both.

We apply our methods to previously published LFA-1 SPT data [57, 8, 75], LFA-1 being a cell membrane adhesion receptor on T cells that is known to interact with the cytoskeleton and exhibits multiple states with different diffusion properties, as shown by previous SPT analysis [57, 8, 75, 37]. LFA-1 has at least two affinity states, including a low affinity closed conformation and a high affinity open conformation, which are dependent on the cytoskeletal protein talin [105]. Activation of T cells, e.g. with phorbol-12-myristate-13-acetate (PMA), causes a number of changes in the behaviour of LFA-1, including a shift from the low to the high affinity state [106, 107] with an associated change in mobility [108, 109, 57]. The protease calpain releases LFA-1 from attachment to the cytoskeleton by cleaving the talin head domain [110]. By examining 4 treatments we find multiple modes of heterogeneity are present, including switching in the diffusion coefficient within single trajectories.

2.2 Methods

Consider a single particle trajectory $\mathbf{X} = \{\Delta X_i, \Delta t_i\}_{i=1}^N$ with displacements ΔX_i at discrete time points $i = 1, 2...N$, where $\Delta X_i = (\Delta X_{i1}, \Delta X_{i2})$ is 2D. We aim to determine if a trajectory is consistent with a single diffusion process throughout, i.e. a one-state diffusion with diffusion coefficient D (to be determined), or if there is evidence of switching of the diffusion coefficient between two (again, to be determined) values, D_0 and D_1, i.e. a two-state diffusion model,

$$D_0 \underset{p_{10}}{\overset{p_{01}}{\rightleftharpoons}} D_1 \tag{2.1}$$

where p_{01}, p_{10} are the probability of switching per frame. Both these models can be considered with or without measurement noise giving 4 models. Using a Bayesian methodology, for each model we developed an MCMC algorithm to sample the posterior distribution $\pi(\theta|\mathbf{X}, M_i)$ of the model M_i and parameters θ, i.e. on individual trajectories we estimate the diffusion coefficient D for the one-state model, and the two diffusion coefficients D_0, D_1, with switching times between the two states for the two-state model. We also computed the marginal likelihood $\pi(\mathbf{X}|M_i)$ (either analytically, through MCMC sampling or, for the models with measurement noise, using an approximation). From the marginal likelihood we can compute the support for each model from the data, and thus determine the posterior model probability ratio $\pi(M_{1D}|\mathbf{X})/\pi(M_{2D}|\mathbf{X})$ for each trajectory. Under an equiprobable model prior this is equivalent to the Bayes factor $\pi(\mathbf{X}|M_{1D})/\pi(\mathbf{X}|M_{2D})$. These methods and associated algorithms are given here and in

Appendix A.3, but the Results (Section 2.3) can be read without this section.

2.2.1 One-state diffusion model without measurement noise

For a particle diffusing with a diffusion coefficient D, the log likelihood of a trajectory $\mathbf{X}$ is

$$\log_e \pi(\mathbf{X}|D) = \sum_{i=1}^{N} \log_e N(\Delta X_i; 0, 2D\Delta t_i). \tag{2.2}$$

Here and throughout we use the same notation for a probability distribution and its (joint) pdf. We use a flat prior on D, $\pi(D) = \text{Unif}(D; 0, D_{max})$, so the posterior is ($\Delta X_i^2 = \Delta X_{i,1}^2 + \Delta X_{i,2}^2$)

$$\pi(1/D|\mathbf{X}) \propto \text{Gamma}\left(1/D; N-1, \frac{1}{4}\sum_{i=1}^{N}\frac{\Delta X_i^2}{\Delta t_i}\right)\mathbb{1}_{[0,D_{max}]}(D) \tag{2.3}$$

where $\mathbb{1}_{[0,D_{max}]}(D) = 1$ if $D \in (0, D_{max})$ and 0 otherwise. We use this notation for the indicator function throughout.

Appropriate statistics can be computed from this posterior, either analytically or using a rejection sampler. For a rejection sampler the update is

$$1/D \sim \text{Gamma}_T\left(N-1, \frac{1}{4}\sum_{i=1}^{N}\frac{\Delta X_i^2}{\Delta t_i}, 1/D_{max}, \infty\right) \tag{2.4}$$

where $\text{Gamma}_T(\alpha, \beta, x_{min}, x_{max})$ denotes a truncated Gamma distribution with parameters α and β, truncated at x_{min} and x_{max}. We sample K updates from this distribution to give samples $\{D^{(k)}\}_{k=1}^{K}$, an estimate for the diffusion coefficient is then the posterior mean $\hat{D} = \frac{1}{K}\sum_{k=1}^{K} D^{(k)}$.

The marginal likelihood for this model is

$$\pi(X|M_{1D}) = \int_0^\infty \mathrm{d}D\, \pi(X|D, M_{1D})\pi(D). \tag{2.5}$$

Changing variables from D to D^{-1} and rearranging into a standard incomplete upper gamma function gives

$$\pi(\mathbf{X}|M_{1D}) = \frac{1}{D_{max}}\prod_{i=1}^{N}\frac{1}{4\pi\Delta t_i}\left(\sum_{i=1}^{N}\frac{\Delta X_i^2}{4\Delta t_i}\right)^{1-N}\Gamma\left(N-1, \frac{1}{D_{max}}\sum_{i=1}^{N}\frac{\Delta X_i^2}{4\Delta t_i}\right). \tag{2.6}$$

where Γ is the upper incomplete Gamma function, see Appendix A.1.

2.2.2 Two-state diffusion model without measurement noise

We use the hidden Markov model described by Das *et al.* [8] with four model parameters, $\theta = \{D_0, D_1, p_{01}, p_{10}\}$, two diffusion coefficients D_0, D_1 and transition probabilities p_{10}, p_{01} between the two hidden states. Denoting the hidden state by z_i at time frame i, the particle moves between $z_i = 0$ (diffusion with $D = D_0$) and $z_i = 1$ (diffusion with $D = D_1$) for N time steps, giving a trajectory $\mathbf{X}$ and hidden state sequence $\mathbf{z} = \{z_i\}_{i=1}^N$. The model can be written

$$z_i|_{z_{i-1}} \sim \text{Bernoulli}(z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01}),\ \Delta X_i|_{z_i} \sim N(0, 2D_{z_i}\Delta t_i). \tag{2.7}$$

We use conjugate priors, the full prior being

$$\pi(\theta) = \text{Unif}(D_0; 0, D_{max})\text{Unif}(D_1; 0, D_{max})\text{Beta}(p_{01}; a_0, b_0)\text{Beta}(p_{10}; a_1, b_1) \tag{2.8}$$

$$\pi(z_1|\theta) = \text{Bernoulli}\left(z_1; \frac{p_{01}}{p_{10}+p_{01}}\right). \tag{2.9}$$

The prior on the initial state is the stationary distribution for the Markov chain. The posterior distribution is then given by,

$$\begin{aligned}\pi(\theta, \mathbf{z}|\mathbf{X}) \propto \pi(\theta)\pi(z_1|\theta) &\prod_{i=1}^N N(\Delta X_i; 0, 2D_{z_i}\Delta t_i) \\ &\times \prod_{i=2}^N \text{Bernoulli}\left(z_i; z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01}\right).\end{aligned} \tag{2.10}$$

We developed an MCMC algorithm to sample this posterior, specifically we can generate samples $\{D_0^{(k)}, D_1^{(k)}, p_{01}^{(k)}, p_{10}^{(k)}, \mathbf{z}^{(k)}\}_{k=K_B}^K$ from the posterior distribution using a Gibbs sampler, see below and in Appendix A.3 as pseudocode. Here and throughout we denote the total number of MCMC steps K and the length of the burn-in K_B. The mean of these posterior samples $\{\hat{D_0}, \hat{D_1}, \hat{p_{01}}, \hat{p_{10}}, \hat{\mathbf{z}}\}$ is an estimate for the parameters and hidden state sequence.

We sample from the posterior distribution (2.10) by sampling sequentially from the conditional distributions. For D_0 and D_1 these are

$$\pi(D_0|D_1, p_{01}, p_{10}, \mathbf{z}, \mathbf{X}) \propto \text{Unif}(D_0; 0, D_{max}) \prod_{z_i=0} N(\Delta X_i; 0, 2D_{z_i}\Delta t_i) \tag{2.11}$$

$$\pi(D_1|D_0, p_{01}, p_{10}, \mathbf{z}, \mathbf{X}) \propto \text{Unif}(D_1; 0, D_{max}) \prod_{z_i=1} N(\Delta X_i; 0, 2D_{z_i}\Delta t_i). \tag{2.12}$$

Hence the updates are

$$D_0^{-1} \sim \text{Gamma}_T\left(\eta_0 - 1, \sum_{z_i=0} \frac{\Delta X_i^2}{4\Delta t_i}, \frac{1}{D_{max}}, \infty\right) \tag{2.13}$$

$$D_1^{-1} \sim \text{Gamma}_T\left(\eta_1 - 1, \sum_{z_i=1} \frac{\Delta X_i^2}{4\Delta t_i}, \frac{1}{D_{max}}, \infty\right) \tag{2.14}$$

where $\eta_0 = \sum_{z_i=0} 1$ and $\eta_1 = \sum_{z_i=1} 1$. We sample from the truncated distribution, by sampling from the full Gamma distribution, then resampling if D_0 or D_1 is bigger than D_{max}. If $\eta_0 = 0$ then $\sum_{z_i=0} \frac{\Delta X_i^2}{4\Delta t_i} = 0$, and the Gamma distribution is undefined, so we sample D_0 from the prior $\text{Unif}(0, D_{max})$. If $\eta_1 = 0$ we sample D_1 from $\text{Unif}(0, D_{max})$. For the transition probabilities let n_{jk} be the number of transitions from state j to state k in the hidden state sequence $\mathbf{z}$, i.e.

$$n_{jk} = \sum_{i|_{z_i=j, z_{i+1}=k}} 1. \tag{2.15}$$

Since we have chosen conjugate priors the updates are (ignoring the first hidden state z_1)

$$p_{01}|\mathbf{z} \sim \text{Beta}(a_0 + n_{01}, b_0 + n_{00}) \tag{2.16}$$

$$p_{10}|\mathbf{z} \sim \text{Beta}(a_1 + n_{10}, b_1 + n_{11}). \tag{2.17}$$

The hidden state $\mathbf{z}$ is updated step by step. Since $\mathbf{z}$ is a Markov chain each z_i depends only on the neighbouring points z_{i-1} and z_{i+1}, so the conditional distribution is

$$\begin{aligned}\pi(z_i|z_{i-1}, z_{i+1}, \theta, \mathbf{X}) \propto &\text{ Bernoulli}(z_i; z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01}) \\ &\times N(\Delta X_i; 0, 2D_{z_i}\Delta t_i) \\ &\times \text{Bernoulli}(z_{i+1}; z_i(1-p_{10}) + (1-z_i)p_{01}).\end{aligned} \tag{2.18}$$

By normalising (2.18) we can compute the probabilities $\pi(z_i|z_{i-1}, z_{i+1}, \theta, \mathbf{X})$ for $z_i = 0, 1$ which gives the update

$$z_i\Big|_{\theta, z_{i\pm1}} \sim \text{Bernoulli}\left(\pi(z_i = 1|z_{i-1}, z_{i+1}, \theta, \mathbf{X})\right). \tag{2.19}$$

The endpoint conditionals are slightly modified. For $i = 1$ and $i = N$ we have (again

ignoring z_1)

$$\pi(z_1|z_2,\theta,\mathbf{X}) \propto N(\Delta X_1;0,2D_{z_1}\Delta t_1)\text{Bernoulli}\left(z_2;z_1(1-p_{10})+(1-z_1)p_{01}\right) \quad (2.20)$$

$$\pi(z_N|z_{N-1},\theta,\mathbf{X}) \propto \text{Bernoulli}\left(z_N;z_{N-1}(1-p_{10})+(1-z_{N-1})p_{01}\right) N(\Delta X_N;0,2D_{z_N}\Delta t_N). \quad (2.21)$$

Thus, we can sequentially update $\mathbf{z}$ by updating each z_i for $i = 1..N$.

We also impose the condition $D_0 < D_1$, which we enforce after the MCMC run as follows: if the posterior means $\hat{D_0} > \hat{D_1}$ then we swap the D_0, D_1 chains, swap the p_{01}, p_{10} chains, and swap the 0 and 1 states in the hidden state $\mathbf{z}$ throughout the run. This is possible because although state identity switching ($0 \leftrightarrow 1$) is possible because of a permutation symmetry during a run, it isn't observed to occur.

There are a number of methods for estimating the marginal likelihood using MCMC sampling, including that of Chen [99], utilising a single MCMC chain, and Chib [97], requiring additional MCMC chains to be constructed. Typically we used both to check our algorithms, but present the simplest approach in any given case. For this model the conditional posterior $\pi(\theta|\mathbf{z},\mathbf{X})$ is normalisable; Chen's formula then reads

$$\pi(\mathbf{X}|M_{2D}) = \log_e \pi(\mathbf{X}|\theta^*) - \log_e \left[\frac{1}{K}\sum_{k=K_B}^{K} \frac{g(\theta^{(k)}|\mathbf{z}^{(k)})}{\pi(\theta^*)} \frac{\pi(\theta^*|\mathbf{z}^{(k)},\mathbf{X})}{\pi(\theta^{(k)}|\mathbf{z}^{(k)},\mathbf{X})}\right] \quad (2.22)$$

where $\theta^* = \{D_0^*, D_1^*, p_{01}^*, p_{10}^*\}$ is a suitably chosen fixed point, such as the maximum likelihood, $\theta^{(k)}$ and $\mathbf{z}^{(k)}$ are samples from the MCMC run and $g(\theta^{(k)}|\mathbf{z}^{(k)})$ is an arbitrary distribution, but its choice affects the variance of the estimate. If we choose $g(\theta^{(k)}|\mathbf{z}^{(k)}) = \pi(\theta^{(k)}|\mathbf{z}^{(k)},\mathbf{X})$ then we remove $\theta^{(k)}$ from the right hand side, obtaining

$$\log_e \pi(\mathbf{X}|M_{2D}) = \log_e \pi(\mathbf{X}|\theta^*) - \log_e \left[\frac{1}{K}\sum_{k=K_B}^{K} \frac{\pi(\theta^*|\mathbf{z}^{(k)},\mathbf{X})}{\pi(\theta^*)}\right]. \quad (2.23)$$

Thus, the sum runs over $\mathbf{z}^{(k)}$, the MCMC samples, and for each $\mathbf{z}^{(k)}$ we have to evaluate $\pi(\theta^*|\mathbf{z}^{(k)},\mathbf{X})/\pi(\theta^*)$. The log likelihood term, $\log_e \pi(\mathbf{X}|\theta^*)$, is calculated by the forward algorithm described in [8]. For the $\pi(\theta^*|\mathbf{z}^{(k)},\mathbf{X})$ term, we factorize

$$\begin{aligned}\pi(\theta^*|\mathbf{z}^{(k)},\mathbf{X}) &= \pi(D_0^*|\mathbf{z}^{(k)},\mathbf{X})\pi(D_1^*|D_0^*,\mathbf{z}^{(k)},\mathbf{X})\pi(p_{01}^*|D_0^*,D_1^*,\mathbf{z}^{(k)},\mathbf{X}) \\ &\quad \times \pi(p_{10}^*|p_{01}^*,D_0^*,D_1^*,\mathbf{z}^{(k)},\mathbf{X}) \\ &= \pi(D_0^*|\mathbf{z}^{(k)},\mathbf{X})\pi(D_1^*|\mathbf{z}^{(k)},\mathbf{X})\pi(p_{01}^*|\mathbf{z}^{(k)},\mathbf{X})\pi(p_{10}^*|\mathbf{z}^{(k)},\mathbf{X}) \quad (2.24)\end{aligned}$$

where the second line follows since the parameters are conditionally independent when $\mathbf{z}^{(k)}$ is given.

The joint pdf is thus,

$$\begin{aligned}\pi(\theta^*|\mathbf{z}^{(k)},\mathbf{X}) = &\text{Gamma}_T\left(\frac{1}{D_0};\eta_0-1,\frac{1}{4}\sum_{z_i^{(k)}=0}\frac{\Delta X_i^2}{\Delta t_i},\frac{1}{D_{max}},\infty\right)\\ &\times\text{Gamma}_T\left(\frac{1}{D_1};\eta_1-1,\frac{1}{4}\sum_{z_i^{(k)}=1}\frac{\Delta X_i^2}{\Delta t_i},\frac{1}{D_{max}},\infty\right)\\ &\times\text{Beta}(p_{01};n_{01}+a_0,n_{10}+b_0)\text{Beta}(p_{10};n_{10}+a_1,n_{11}+b_1)\end{aligned} \tag{2.25}$$

at a given value θ^*, where $\eta_0 = \sum_{z_i^{(k)}=0} 1$ and $\eta_1 = \sum_{z_i^{(k)}=1} 1$, and $z_i^{(k)}$ is the ith term in the sequence $\mathbf{z}^{(k)}$. The normalisation term for the truncated distribution is $\Gamma(\frac{1}{4D_{max}}\sum_{z_i^{(k)}=0}\frac{\Delta X_i^2}{\Delta t_i},\eta_0+1)^{-1}$, where Γ is the upper incomplete gamma function. In practice, the normalisation factor is very close to 1, since the choice of D_{max} is sufficiently high.

Equation (2.25) is valid except when $\eta_0 = 0$ or $\eta_1 = 0$, in which case we have $\pi(D_0^*|\mathbf{z}^{(k)},\mathbf{X}) = \text{Unif}(D_0;0,D_{max})$ or $\pi(D_1^*|\mathbf{z}^{(k)},\mathbf{X}) = \text{Unif}(D_1;0,D_{max})$ respectively. The prior is

$$\pi(\theta^*) = \begin{cases}\frac{D_0^2}{D_{max}}\frac{D_1^2}{D_{max}}\text{Beta}(p_{01};a_0,b_0)\text{Beta}(p_{01};a_1,b_1) & \text{if } \frac{1}{D_0},\frac{1}{D_1}\in[\frac{1}{D_{max}},\infty]\\ 0 & \text{otherwise}\end{cases} \tag{2.26}$$

which is easy to evaluate for each $\mathbf{z}^{(k)}$ from the MCMC output. Hence we can evaluate Equation (2.23).

2.2.3 One-state diffusion model with measurement noise

We now add a localisation error to the previous one-state diffusion model. The true particle position is hidden and denoted U_i, whilst the measured position is U_i up to a Gaussian noise with variance σ^2, assumed known. We attempted to construct inference schemes (for this model and the subsequent two-state version) which included a Metropolis-Hastings update for σ, but we could not find an algorithm which accurately recovered all parameters. We therefore use a fixed localisation error σ^2. In discrete time the model is

$$\Delta U_i \sim N(0,2D\Delta t_i),\ X_i\Big|_{U_i} \sim N(U_i,\sigma^2) \tag{2.27}$$

where $\Delta U_i = U_{i+1} - U_i$. In order to develop an MCMC sampler for $\pi(D, \mathbf{U}|\mathbf{X})$ we note that

$$\pi(D, \mathbf{U}|\mathbf{X}) \propto \pi(D, U_1) \prod_{i=1}^{N+1} N(X_i; U_i, \sigma^2) \prod_{i=1}^{N} N(\Delta U_i; 0, 2D\Delta t_i). \tag{2.28}$$

We select a conjugate prior $\pi(D, U_1) = \text{Unif}(D; 0, D_{max})N(U_1; \mu_{U_1}, \sigma_U^2)$, so the updates for D and $\mathbf{U}$ are Gibbs moves. The update for D is

$$1/D \sim \text{Gamma}_T\left(N - 1, \frac{1}{4}\sum_{i=1}^{N} \frac{\Delta U_i^2}{\Delta t_i}, 1/D_{max}, \infty\right). \tag{2.29}$$

The conditional distribution for U_i is a bridging distribution

$$\begin{aligned}\pi(U_i|U_{i-1}, U_{i+1}, X_i, D) \propto\ & N(\Delta U_{i-1}; 0, 2D\Delta t_{i-1}) \\ & \times N(\Delta U_i; 0, 2D\Delta t_i)N(U_i; X_i, \sigma^2)\end{aligned} \tag{2.30}$$

comprising a product of three Gaussians. The update is thus,

$$U_i\Big|_{D, U_{i\pm1}} \sim N(\mu_i, 1/\tau_i) \tag{2.31}$$

where, for $i = 2$ to $i = N$ the precision and mean are

$$\tau_i = \frac{1}{2D\Delta t_{i-1}} + \frac{1}{2D\Delta t_i} + \frac{1}{\sigma^2},\ \mu_i = \left(\frac{U_{i-1}}{2D\Delta t_{i-1}} + \frac{U_{i+1}}{2D\Delta t_i} + \frac{X_i}{\sigma^2}\right)\tau_i^{-1} \tag{2.32}$$

at the endpoints $i = 1$ (with prior $N(\mu_{U_1}, 1/\sigma_{U^2})$) and $i = N + 1$

$$\tau_1 = \frac{1}{2D\Delta t_1} + \frac{1}{\sigma_U^2},\ \mu_1 = \left(\frac{U_2}{2D\Delta t_1} + \frac{\mu_{U_1}}{\sigma_U^2}\right)\tau_1^{-1} \tag{2.33}$$

$$\tau_{N+1} = \frac{1}{2D\Delta t_N} + \frac{1}{\sigma^2},\ \mu_{N+1} = \left(\frac{U_N}{2D\Delta t_N} + \frac{X_N}{\sigma^2}\right)\tau_{N+1}^{-1}. \tag{2.34}$$

Updating a continuous-time latent state one step at a time can be very inefficient for small Δt_i. However, perhaps since each U_i depends heavily on X_i, this was not found to be a problem.

We used an approximation to compute the marginal likelihood; this involves ignoring the covariance between the displacements $\Delta X_i \sim N(\Delta U_i, 2\sigma^2)$ and $\Delta X_{i+1} \sim N(\Delta U_{i+1}, 2\sigma^2)$ that arises because of the common measurement error $X_i - U_i$ at time

point i. In this case the hidden variables U_i integrate out to give the posterior

$$\pi(D|\mathbf{X}) \propto \pi(D) \prod_{i=1}^{N} N\left(\Delta X_i; 0, 2D\Delta t_i + 2\sigma^2\right). \tag{2.35}$$

We modified the previous one-state MCMC sampler to sample from this distribution, detailed in Appendix A.1 and as pseudocode in Appendix A.3.

The sampler has a single Metropolis-Hastings move, so we calculate the marginal likelihood directly from the MCMC output, as described by Chib [98]. The log marginal identity is

$$\log_e \hat{\pi}(\mathbf{X}|M_{1D}) = \log_e \pi(\mathbf{X}|D^*) + \log_e \pi(D^*) - \log_e \hat{\pi}(D^*|\mathbf{X}) \tag{2.36}$$

where we take $D^* = \hat{D}$, the mean of the posterior samples. We can evaluate $\log_e \pi(\mathbf{X}|D^*)$ and $\log_e \pi(D^*)$ easily. We can write $\log_e \pi(D^*|\mathbf{X})$ as [98]

$$\log_e \pi(D^*|\mathbf{X}) = \log_e \left[\frac{\mathbb{E}_1 \left[\alpha(D \to D^*) q(D \to D^*)\right]}{\mathbb{E}_2 \left[\alpha(D^* \to D)\right]} \right] \tag{2.37}$$

where $\mathbb{E}_1$ is with respect to $\pi(D|\mathbf{X})$ and $\mathbb{E}_2$ is with respect to $q(D^* \to D)$. From the MCMC output we have $K - K_B$ samples from the posterior distribution $\pi(D|\mathbf{X})$, $\{D^{(k)}\}_{K=K_B}^{K}$. We then simulate $K - K_B$ samples from the proposal distribution $q(D^* \to D) \sim N(D^*, S_D)$, giving $\{\tilde{D}^{(j)}\}_{j=1}^{K-K_B}$. An estimate for $\log_e \pi(D^*|\mathbf{X})$ is then

$$\log_e \hat{\pi}(D^*|\mathbf{X}) = \log_e \left[\frac{(K-K_B)^{-1} \sum_{k=K_B}^{K} \alpha(D^{(k)} \to D^*) q(D^{(k)} \to D^*)}{(K-K_B)^{-1} \sum_{j=1}^{K-K_B} \alpha(D^* \to \tilde{D}^{(j)})} \right]. \tag{2.38}$$

Hence we can calculate $\hat{\pi}(\mathbf{X}|M_{1D})$ using Equation (2.36).

2.2.4 Two-state diffusion model with measurement noise

We now add a localisation error to the previous two-state diffusion hidden Markov model. Again, the true position is hidden and denoted U_i. The model is given by,

$$\begin{aligned} z_i|z_{i-1} &\sim \text{Bernoulli}(z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01}), \\ \Delta U_i|z_i &\sim N(0, 2D_{z_i}\Delta t_i), \\ X_i|U_i &\sim N(U_i, \sigma^2) \end{aligned} \tag{2.39}$$

i.e. there is both a continuous hidden state U_i and a discrete hidden state z_i. We developed an MCMC algorithm which samples from the full conditional distribution $\pi(\theta, \mathbf{U}, \mathbf{z}|\mathbf{X})$. Let $\theta = \{D_0, D_1, p_{01}, p_{10}\}$, the posterior for this model is ($\Delta U_i = U_{i+1} - U_i$)

$$\pi(\theta, \mathbf{U}, \mathbf{z}|\mathbf{X}) \propto \pi(\theta, U_1, z_1) \prod_{i=1}^{N+1} N(X_i; U_i, \sigma^2) \prod_{i=1}^{N} N(\Delta U_i; 0, 2D_{z_i}\Delta t_i) \times \prod_{i=1}^{N-1} \text{Bernoulli}(z_{i+1}; z_i(1-p_{10}) + (1-z_i)p_{01}). \tag{2.40}$$

The priors on θ and z_1 are the same as the two-state diffusion model without measurement noise, given in Equation (2.8), and we use a normal prior (with mean μ_{U_1}, variance σ_U^2) on U_1. The full prior is then

$$\pi(\theta, U_1, z_1) = \text{Unif}(D_0; 0, D_{max})\text{Unif}(D_1; 0, D_{max})\text{Beta}(p_{01}; a_0, b_0)\text{Beta}(p_{10}; a_1, b_1) \times N(U_1; \mu_{U_1}, \sigma_U^2)\text{Bernoulli}\left(z_1; \frac{p_{01}}{p_{10} + p_{01}}\right). \tag{2.41}$$

The MCMC updates are mostly identical to the two-state diffusion model without measurement noise, but with the observed displacements ΔX_i replaced by the hidden state displacements ΔU_i. Thus, for D_0 and D_1 we have

$$1/D_0 \sim \text{Gamma}_T\left(\eta_0 - 1, \frac{1}{4}\sum_{z_i=0} \frac{\Delta U_i^2}{\Delta t_i}, 1/D_{max}, \infty\right) \tag{2.42}$$

$$1/D_1 \sim \text{Gamma}_T\left(\eta_1 - 1, \frac{1}{4}\sum_{z_i=1} \frac{\Delta U_i^2}{\Delta t_i}, 1/D_{max}, \infty\right) \tag{2.43}$$

where $\eta_0 = \sum_{z_i=0} 1$ and $\eta_1 = \sum_{z_i=1} 1$ as before. Similarly, in the $\mathbf{z}$ update we substitute ΔX_i for ΔU_i in Equations (2.18), (2.20) and (2.21),

$$\pi(z_i|z_{i-1}, z_{i+1}, \theta, \mathbf{U}) \propto \text{Bernoulli}(z_i; z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01}) \times N(\Delta U_i; 0, 2D_{z_i}\Delta t_i) \times \text{Bernoulli}(z_{i+1}; z_i(1-p_{10}) + (1-z_i)p_{01}). \tag{2.44}$$

$$\pi(z_1|z_2, \theta, \mathbf{U}) \propto N(\Delta U_1; 0, 2D_{z_1}\Delta t_1)\text{Bernoulli}\left(z_2; z_1(1-p_{10}) + (1-z_1)p_{01}\right) \tag{2.45}$$

$$\pi(z_N|z_{N-1}, \theta, \mathbf{U}) \propto \text{Bernoulli}\left(z_N; z_{N-1}(1-p_{10}) + (1-z_{N-1})p_{01}\right) N(\Delta U_N; 0, 2D_{z_N}\Delta t_N). \tag{2.46}$$

The transition probability updates are identical to Equations (2.16) and (2.17). The update for U is almost the same as the one-state diffusion model with measurement noise. We have a Gibbs update

$$U_i\Big|_{\theta,\mathbf{z},U_{i\pm1}} \sim N(\mu_i, 1/\tau_i) \tag{2.47}$$

where, for $i = 2$ to $i = N$, the precision and mean are

$$\tau_i = \frac{1}{2D_{z_{i-1}}\Delta t_{i-1}} + \frac{1}{2D_{z_i}\Delta t_i} + \frac{1}{\sigma^2},\ \mu_i = \left(\frac{U_{i-1}}{2D_{z_{i-1}}\Delta t_{i-1}} + \frac{U_{i+1}}{2D_{z_i}\Delta t_i} + \frac{X_i}{\sigma^2}\right)\tau_i^{-1} \tag{2.48}$$

at the endpoints $i = 1$ and $i = N + 1$

$$\tau_1 = \frac{1}{2D_{z_1}\Delta t_1} + \frac{1}{\sigma_U^2},\ \mu_1 = \left(\frac{U_2}{2D_{z_1}\Delta t_1} + \frac{X_1}{\sigma_U^2}\right)\tau_1^{-1} \tag{2.49}$$

$$\tau_{N+1} = \frac{1}{2D_{z_N}\Delta t_N} + \frac{1}{\sigma^2},\ \mu_{N+1} = \left(\frac{U_N}{2D_{z_N}\Delta t_N} + \frac{X_N}{\sigma^2}\right)\tau_{N+1}^{-1}. \tag{2.50}$$

The MCMC updates for this model are given in pseudocode in Appendix A.3.

To compute the marginal likelihood we used the same approximation as the one-state diffusion model with measurement noise, ignoring the covariance between the displacements $\Delta X_i \sim N(\Delta U_i, 2\sigma^2)$ and $\Delta X_{i+1} \sim N(\Delta U_{i+1}, 2\sigma^2)$. We failed to find an efficient algorithm that could integrate over both hidden states (U_i and z_i) to allow the (exact) marginal likelihood $\pi(\mathbf{X}|M_{2D})$ to be computed. (We found implementation of Chib's method difficult on continuous hidden states, as well as very computationally expensive.) In this case the hidden variables U_i integrate out to give a posterior

$$\begin{aligned}\pi(\theta, \mathbf{z}|\mathbf{X}) \propto \pi(\theta)\pi(z_1|\theta) &\prod_{i=1}^{N} N(\Delta X_i; 0, 2(D_{z_i}\Delta t_i + \sigma^2)) \\ &\times \prod_{i=1}^{N-1} \text{Bernoulli}(z_{i+1}; z_i(1-p_{10}) + (1-z_i)p_{01}).\end{aligned} \tag{2.51}$$

We modified the two-state diffusion model sampler to incorporate the σ^2 terms, see Appendix A.1 and pseudocode in Appendix A.3. This sampler can be used with the

method of Chen, rewriting Chen's formula as

$$\log_e \pi(\mathbf{X}|M_{2D}) = \log_e \pi(\mathbf{X}|\theta^*) - \log_e \left[\frac{1}{K} \sum_{k=1}^{K} \frac{g(\theta_k|\mathbf{z}^{(k)})}{\pi(\theta^{(k)})} \frac{\pi(\mathbf{X}|\theta^*, \mathbf{z}^{(k)})}{\pi(\mathbf{X}|\theta^{(k)}, \mathbf{z}^{(k)})} \frac{\pi(\mathbf{z}^{(k)}|\theta^*)}{\pi(\mathbf{z}^{(k)}|\theta^{(k)})} \right] \tag{2.52}$$

where g is any density function, $\theta^{(k)}, \mathbf{z}^{(k)}$ are samples from the posterior distribution, and θ^* is a point of high density. If we choose $g = \pi(\theta^{(k)})$, then an estimate for the marginal likelihood is

$$\log_e \hat{\pi}(\mathbf{X}|M_{2D}) = \log_e(\mathbf{X}|\theta^*) - \log_e \left[\frac{1}{K} \sum_{k=1}^{K} \frac{\prod_{i=1}^{N} N(\Delta X_i^2; 0, 2D^*_{z_i^{(k)}} \Delta t_i + 2\sigma^2)}{\prod_{i=1}^{N} N(\Delta X_i^2; 0, 2D^{(k)}_{z_i^{(k)}} \Delta t_i + 2\sigma^2)} \right.$$
$$\left. \times \frac{\text{Beta}\left(p_{01}^*; n_{01}+1, n_{00}+1\right) \text{Beta}\left(p_{10}^*; n_{10}+1, n_{11}+1\right)}{\text{Beta}\left(p_{01}^{(k)}; n_{01}+1, n_{00}+1\right) \text{Beta}\left(p_{10}^{(k)}; n_{10}+1, n_{11}+1\right)} \right]. \tag{2.53}$$

The log likelihood, $\log_e \pi(\mathbf{X}|\theta^*)$, is calculated using a forward algorithm (Appendix A.1 and pseudocode in Appendix A.3). By computation of the marginal on multiple chains we found that its variance was small despite using the prior for the distribution g, (relative sd $< 0.0001\%$). Chib's method on selected trajectories also gave consistent findings.

2.2.5 Priors

In all algorithms we use weak priors. Specifically $D \sim \text{Unif}(0, D_{max} = 10^6 \text{nm}^2\text{s}^{-1})$ for the one-state diffusion model, and additionally $U_1 \sim N\left(\mu_{U_1} = \left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \sigma_U^2 = \left[\begin{smallmatrix}10^6\text{nm}^2 & 0\\ 0 & 10^6\text{nm}^2\end{smallmatrix}\right]\right)$ for the one-state diffusion model with measurement noise. For the two-state diffusion model we use: $D_0, D_1 \sim \text{Unif}(0, D_{max} = 10^6\text{nm}^2\text{s}^{-1})$, $p_{10}, p_{01} \sim \text{Beta}(1,1)$, with an initial ($i = 1$) hidden state, $z_1 \sim \text{Bernoulli}\left(\frac{p_{01}}{p_{01}+p_{10}}\right)$. Additionally, we used $U_1 \sim N\left(\mu_{U_1} = \left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \sigma_U^2 = \left[\begin{smallmatrix}10^6\text{nm}^2 & 0\\ 0 & 10^6\text{nm}^2\end{smallmatrix}\right]\right)$ for the two-state diffusion model with measurement noise. Our choice of weak priors has implications for model selection. Specifically the choice of D_{max} directly affects the Bayes' factor (an example of Lindley's paradox [111]). However, in Section 2.3 (and Fig. A.3 B) we demonstrate the consistency of our choice on simulated data.

2.2.6 Convergence of MCMC runs

To assess the convergence of the two-state diffusion model with measurement noise we used a multiple chain convergence diagnostic [92], specifically 12 chains with overdispersed initial values. We initialised D_0 and D_1 by sampling values u_0, u_1 from Beta$(0.1, 0.1)$, then setting $D_0 = u_0 D_{max}$ and $D_1 = u_1 D_{max}$. The transition probabilities p_{01} and p_{10} were initialised from Beta$(1, 1)$. The hidden state $\mathbf{z}$ was initialised by simulating a Markov chain using the initial transition probabilities p_{01} and p_{10}. The initial value of $\mathbf{U}$ was set to the observed trajectory $\{X_i\}_{i=1}^{N+1}$. We considered the chains converged when the Gelman-Rubin diagnostic for each parameter was less than 1.1 [95, 93].

2.2.7 Model selection between one-state and two-state diffusion models

We can calculate the log marginal likelihoods to compare the evidence for the one-state and two-state diffusion models; this can be done with or without measurement noise. Hence for each case we can calculate the log (base e) Bayes factor, $\log_e B_{1D,2D} = \log_e \pi(\mathbf{X}|M_{1D})) - \log_e \pi(\mathbf{X}|M_{2D})$. The extent to which a model is supported by the evidence (i.e. the observed trajectory $\mathbf{X}$) can then be assessed using a standard table such as in Kass *et al.*, where a log Bayes factor of 3 is considered "strong" evidence for the relevant model [96]. We hence consider a value $\log_e B_{1D,2D} > 3$ as preference for a one-state diffusion model, and $\log_e B_{1D,2D} < -3$ as preference for a two-state diffusion model. The remaining trajectories (where $-3 < \log_e B_{1D,2D} < 3$) have no strong preference for either model.

2.3 Results

Given a 2D trajectory, $\mathbf{X}$, we developed MCMC algorithms (both with and without measurement noise) for inferring the posterior distribution of the parameters and hidden states of a two-state diffusion process, $\pi(\theta, \mathbf{z}|\mathbf{X})$. The parameters, $\theta = \{D_0, D_1, p_{01}, p_{10}\}$, are the diffusion coefficients and frame transition probabilities, Equation (2.1), and $\mathbf{z}$ is the sequence of the inferred hidden diffusion state. Allowing for measurement noise propagates that uncertainty to the parameter estimates. We tested our algorithms on simulated data (full details of each simulation study are given in the corresponding figure legend); Fig. A.1 (in Section A.5 in Appendix A) shows an MCMC run of the two-state model with measurement noise, demonstrating accurate reconstruction of the parameters and hidden states. We also tested the sensitivity of the method to closely matched

diffusion states, Fig. A.2. The two-state model with measurement noise algorithm can accurately detect switching between regimes where diffusion coefficients differ by a factor of 1.5 (trajectory parameters set to those typical for the LFA-1 data). To determine whether the trajectory is better explained by this two-state model or a one-state diffusion (single diffusion coefficient D) we used the marginal likelihoods $\pi(\mathbf{X}|M)$; however, this proved difficult to calculate in our hands for the two-state model (with measurement noise). Therefore, we used an approximate likelihood (where the covariance between consecutive displacements is ignored, essentially a low measurement noise limit) where the marginals are computable for both the one-state and two-state models, see Methods. We used the Bayes factor of this approximation to determine if the two-state model is supported by the data more than a one-state diffusion process. We tested the model selection between the approximate one-state and two-state diffusion models with measurement noise; both on trajectories simulated from the full measurement noise model, and trajectories simulated without noise, Fig. A.3. The model was able to successfully discriminate between one-state and two-state simulations, with a very low false positive rate when the diffusion coefficients were separated by a factor of 5 (0.005%, using $\log_e$ Bayes factor equal to ± 3 as the threshold for model preference, see Methods). When separated by a factor of 2.5 there is a bias towards the one-state model, especially on two-state model simulations without measurement noise, Fig. A.3 B. Thus, we may fail to detect some switching events between close diffusion coefficients, underestimating the number of trajectories preferring a two-state model.

We used our algorithms to analyse SPT data sets for the LFA-1 receptor on Jurkat T cells (4 s trajectories at 1000 frames s^{-1} [57]). The receptor was tagged with 1000 nm latex beads coated with the LFA-1 binding antibody TS-1/18. This dataset has been analysed previously [57, 8, 75] demonstrating that LFA-1 diffusion is heterogeneous. We applied our methods to four treatments: control (DMSO), treated with cytochalasin D (Cyto D), treated with phorbol-12-myristate-13-acetate (PMA), and treated with PMA with calpain inhibition (PMA+Cal-I). Cytochalasin D is an inhibitor of actin polymerisation, so effects due to the cytoskeleton should be decreased, PMA is a T cell activator, moving LFA-1 to the high affinity conformational state, and calpain releases LFA-1 from attachment to the cytoskeleton by cleaving the talin head domain [110]. Thus, the first two treatments explore the effect of the cytoskeleton on the predominantly low affinity LFA-1. PMA examines dynamics of high affinity LFA-1, whilst PMA+Cal-I examines high affinity LFA-1 under conditions of enhanced interaction with the cytoskeleton.

To determine the measurement accuracy, and whether measurement noise had

to be incorporated into the model, we examined stationary beads (3 trajectories were available). These beads are attached to the surface and thus represent thermal motion and instrument noise. These beads are effectively stuck in a potential well and their movement is expected to be temporally homogeneous; thus no two-state diffusion structure should be detected. As presented below, we find that this is not the case unless a Gaussian measurement noise is incorporated into the inference. Therefore, throughout we present the analysis of LFA-1 trajectory data using the measurement noise models, comparing between the one-state diffusion and two-state diffusion model in the presence of noise. We use the approximate likelihood models for model discrimination only; all inferred parameters refer to the exact models.

2.3.1 Stationary bead analysis to determine measurement accuracy and the importance of propagating measurement noise

Trajectories of stationary beads (immobilised on glass using cell-tak, imaged using the same set up as the LFA-1 data [57]) were used to determine the signal to noise ratio (S/N) and to estimate the measurement noise (σ^2). For each trajectory (2 s at 1000 frames s^{-1}) we calculated the variance of individual displacements $\{\Delta X_i\}_{i=1}^{N}$ for both x and y directions, giving 6 estimates for the localisation accuracy (29.09 nm^2, 23.55 nm^2, 65.01 nm^2, 39.31 nm^2, 30.27 nm^2, 59.41 nm^2). This gives a mean $\sigma^2 = 41.09$ nm^2 which we used as an estimate of the localisation accuracy throughout. The variance of individual displacements, ΔX_i for the LFA-1 data are: DMSO, 133.5 nm^2 (giving S/N 3.25); Cyto D, 133.7 nm^2 (S/N 3.25); PMA, 135.1 nm^2 (S/N 3.29); PMA+Cal-I, 89.4 nm^2 (S/N 2.18), indicating that signal is present in these displacements at this resolution.

The stationary beads also provide an opportunity to check that the measurement noise does not affect model selection: stationary beads should prefer a one-state diffusion model since the time series is homogeneous. If the two-state diffusion model is preferred then measurement noise, the tracking algorithm, or instrument noise contributes to the heterogeneity in the trajectory. We applied the one-state and two-state diffusion model algorithms (without measurement noise) to the three stationary beads. The two-state diffusion model showed high frequency switching behaviour (Fig. 2.1A-C), with two distinct (well separated) diffusion coefficients, (Fig. 2.1D-F). Crucially, the two-state diffusion model is strongly preferred for all 3 trajectories (Fig. 2.2, red asterisks). Therefore there is evidence that tracked bead displacements are not unstructured and an analysis of LFA-1 trajectories using the models without allowing for measurement

noise may be unreliable, due to this inherent inhomogeneity.

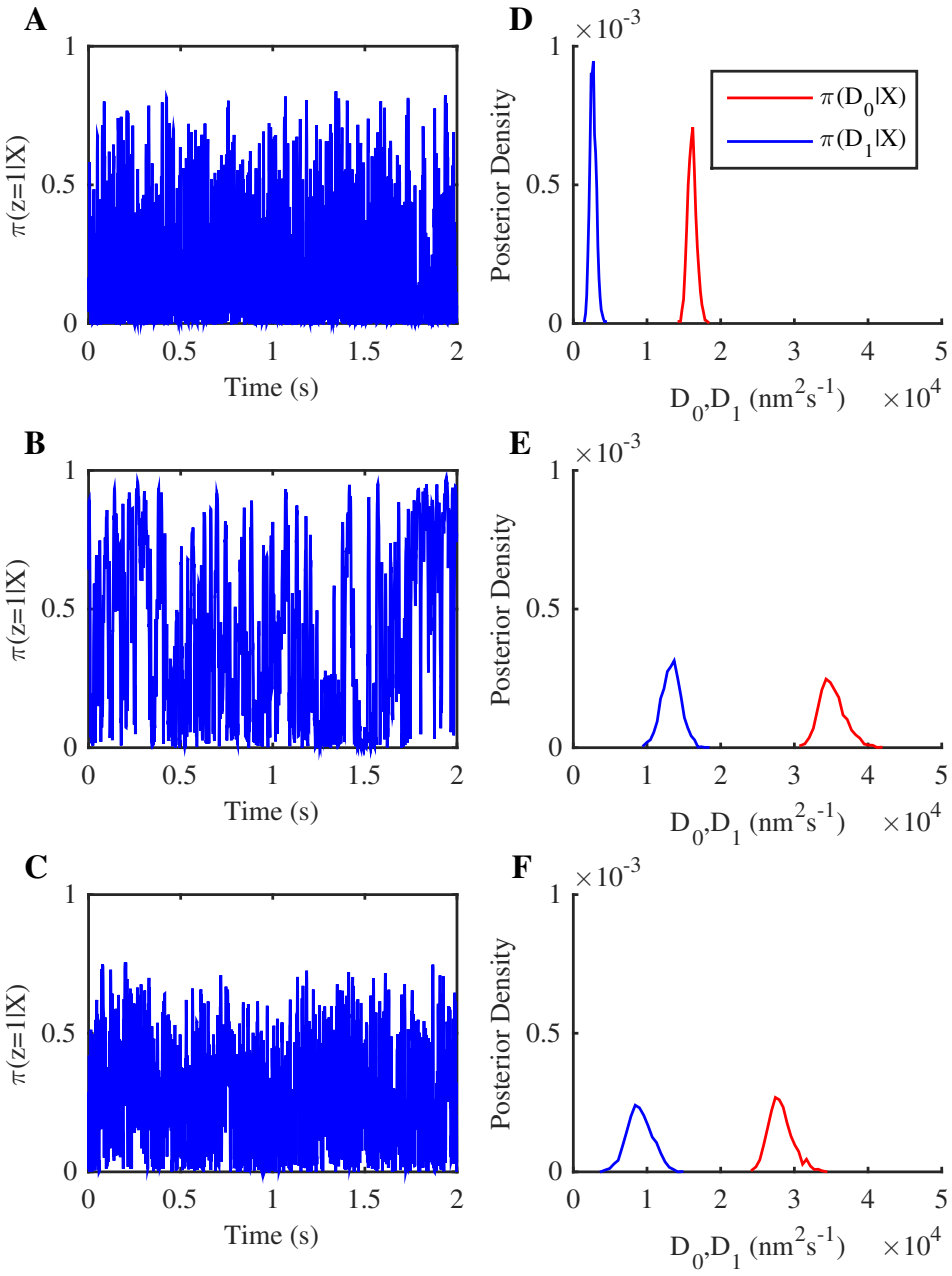

Figure 2.1: **Fit of a two-state diffusion model without measurement noise to three stationary latex bead trajectories.** MCMC output from chains of 20000 MCMC steps with a 10000 step burn-in. (A-C) Inference of the hidden state **z** shown as the probability of being in the low diffusion state. (D-F) Posterior distributions for the two diffusion coefficients: D_0 (red) and D_1 (blue). See Methods for priors and initial conditions.

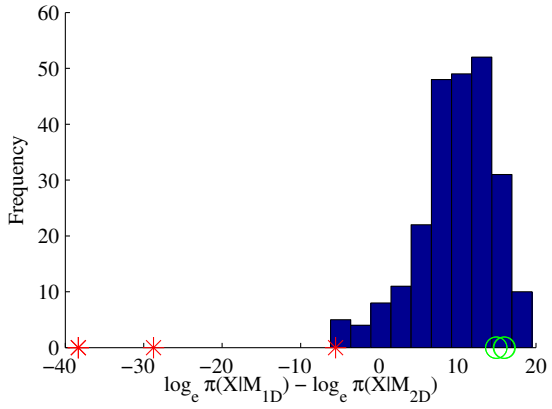

Figure 2.2: **Model selection for one-state and two-state diffusion models on simulated stationary beads and stationary latex bead trajectories.** Blue bars: Bayes factors from model selection on simulated stationary beads ($n = 240$) with added Gaussian noise ($\sigma^2 = 41.09\text{nm}^2$). Single data points on axis: Bayes factors from model selection on stationary latex bead trajectories, both without (red asterisks) and with (green circles, $\sigma^2 = 41.09\text{nm}^2$) measurement noise incorporated into the inference algorithm. Priors, see Methods.

The stationary bead data were then analysed with the approximate one-state and two-state diffusion models with measurement noise using the estimated noise variance $\sigma^2 = 41.09\text{nm}^2$ (recall the approximate models ignore the covariance between displacements since the marginal cannot be calculated for the full model). The incorporation of localisation accuracy eliminates the previous preference for a two-state diffusion model (Fig. 2.2, green circles); preference for the one-state diffusion model is in fact very strong. We also tested whether Gaussian noise can cause deterioration of the model selection accuracy. We tested the model selection analysis without measurement noise on a set ($n = 240$) of simulated stationary bead trajectories with added Gaussian noise (the localisation error in the simulations was set to 41.09nm^2). The model selection has a very strong preference for the one-state diffusion model (Fig. 2.2, blue bars), so Gaussian noise alone causes a low false detection rate. The 3 stationary bead trajectory Bayes factors (Fig. 2.2, red asterisks) are clearly not from the same distribution as the Bayes factors for the Gaussian noise trajectories (blue bars), therefore the noise in the beads cannot be Gaussian and/or independent.

There are two important conclusions: firstly, as pointed out by Michalet [62], the localisation accuracy is a potential source of bias, particularly for low signal to noise ratios, as found here. Thus, in any SPT analysis the measurement accuracy should be separately determined and an assessment made as to whether it affects the results.

Secondly, the noise of the stationary beads does not appear to be independent, having a temporal correlation. Thus, the localisation accuracy is likely not constant along a trajectory. However, as shown here, incorporating Gaussian measurement noise into the model inference removes the erroneous preference for the two-state model for the stationary beads.

2.3.2 Analysis of LFA-1 data: evidence of multiple diffusion states

We fitted the one-state and two-state diffusion models with measurement noise to each trajectory in the four treatments (36-75 trajectories depending on treatment, 4 s trajectories at 1000 frames s^{-1}), thereby estimating parameters for these models for each trajectory. Convergence was confirmed using a multiple chain protocol, see Methods. An example of a fit to the two-state diffusion model with measurement noise is shown in Fig. 2.3; inference of the hidden state shows clear evidence of state switching in this trajectory with a high probability of being in one or other of the two diffusion states and tight switching times. There is a large separation in the posterior distributions for the low and high diffusion coefficients, with the ratio of the posterior mean estimates being around 10.

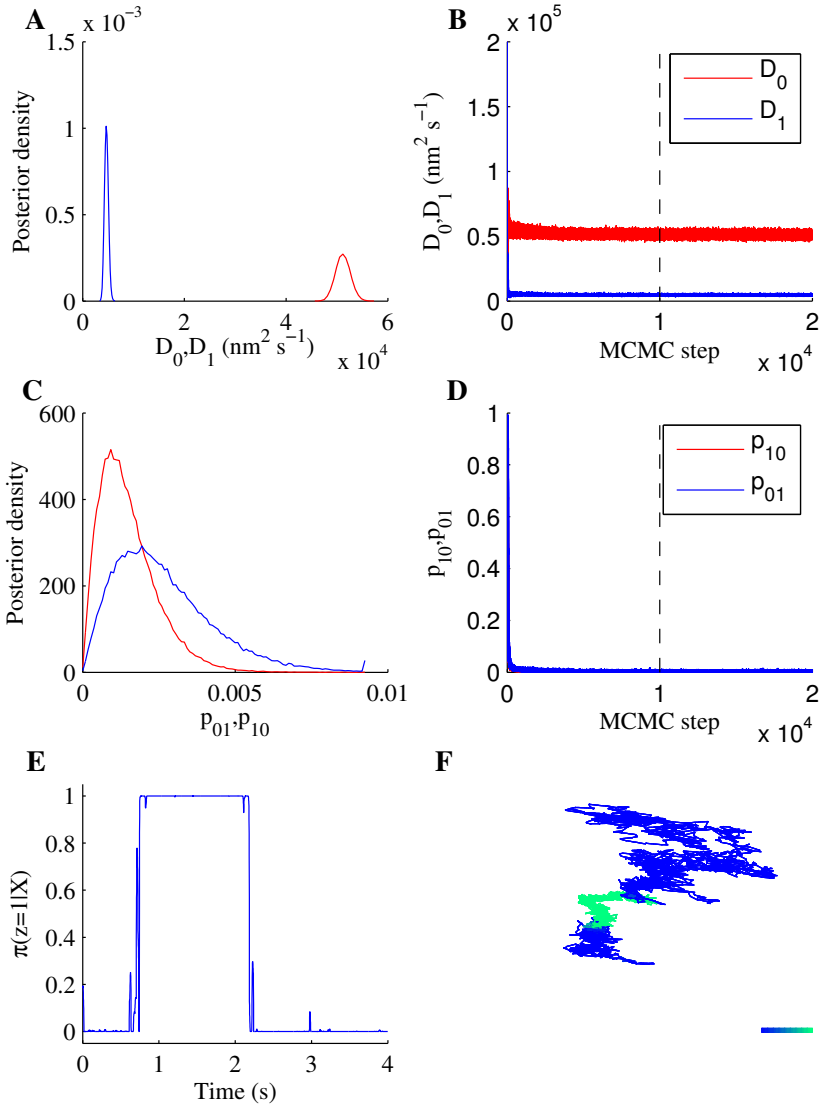

Figure 2.3: **Fit of a two-state diffusion model with measurement noise to an LFA-1 trajectory (PMA+Cal-I treatment).** MCMC output (12 independent chains of 20000 MCMC steps with a 10000 step burn-in). (A) The posteriors for the two diffusion coefficients, (B) corresponding samples (12 chains plotted in the same colour) for D_0 (red) and D_1 (blue) including burn-in (dashed line). (C) Posteriors for the switching probabilities per frame, (D) corresponding samples (12 chains) for p_{01} (red) and p_{10} (blue) including burn-in (dashed line). (E) State inference shown as the probability of being in the low diffusion state. (F) Trajectory coloured by the probability of being in the low diffusion state. Colour scale represents $\pi(\mathbf{z} = 1|\mathbf{X})$ from 0 (blue, high diffusion state) to 1 (green, low diffusion state). Colorbar length: 100nm. Priors, see Methods.

By calculating the marginal likelihood for the approximate one-state and two-state diffusion models with measurement noise, and hence the Bayes factor $B_{1D,2D} = \frac{\pi(\mathbf{X}|M_{1D})}{\pi(\mathbf{X}|M_{2D})}$, we then ascertained for each trajectory the evidence for a two-state compared to a one-state diffusion process. As described in Methods, we used fairly stringent criteria: if the log (base e) Bayes factor is smaller than -3 then we consider this preference for the two-state diffusion model, and greater than 3 as preference for the one-state diffusion model [96]. The number of trajectories with preference for each model was robust to the choice of Bayes factor threshold (Table A.1). Fig. 2.4 shows the Bayes factor estimates for each condition, and the number of trajectories which preferred each model, grouped by treatment. There are a total of 16 DMSO (out of a total of 75, 21%), 8 Cyto D (out of 36, 22%), 13 PMA (out of 19, 33%) and 8 PMA+Cal-I (out of 46, 17%) trajectories where the two-state diffusion model is preferred, Table A.1. Thus, in all treatments we detected evidence of switching within trajectories with a similar level of preference. However, a proportion of the trajectories that preferred the two-state diffusion model showed extremely fast switching; we define fast switching as either $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$, giving counts: DMSO, 3 trajectories; Cyto D, 5 trajectories; PMA, 5 trajectories; PMA+Cal-I, 2 trajectories, Table A.1. Thus, over all treatments, for trajectories where the two-state diffusion model was preferred, we saw fast switching in 33% of trajectories.

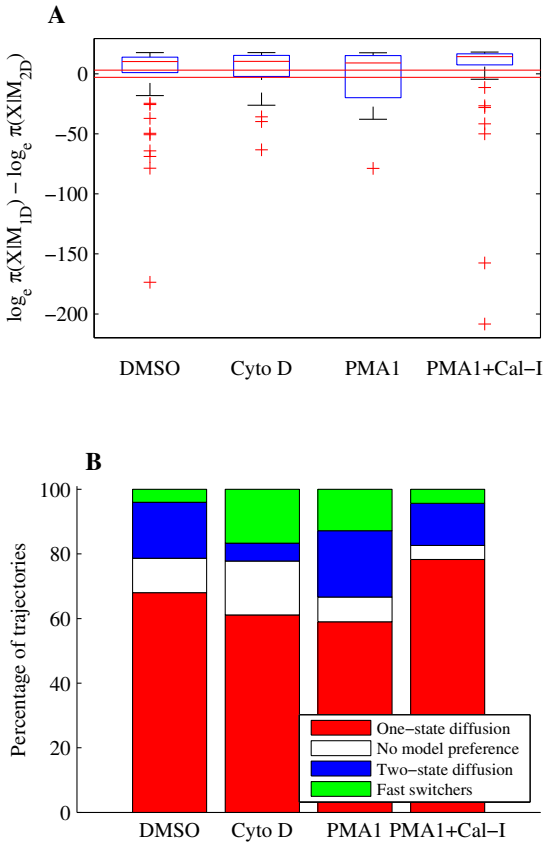

Figure 2.4: **Model selection between approximate one-state and two-state diffusion models with measurement noise on LFA-1 trajectories.** (A) Box and whisker plot of log Bayes factors by treatment, trajectories with log Bayes factor outside 1.5 times IQR are plotted as outliers (red crosses). The thresholds ±3 (red lines) are shown. (B) Stacked bar plot showing proportions for each preferred model and trajectories which demonstrate fast switching between diffusive states. A log Bayes factor of ±3 ((A), red lines) is considered preference for the relevant model. MCMC runs comprise 12 parallel chains of 20000 steps with a 10000 step burn-in. Priors, see Methods.

Table 2.1: **Model selection and proportion of time spent in the immobile state.**

Treatment	DMSO	Cyto D	PMA	PMA+Cal-I
Number of trajectories	75	36	39	46
Two-state model preferred[1]	16/67 (24%)	8/30 (27%)	13/36 (36%)	8/44 (18%)
Two-state model preferred, fast switchers removed[12]	13/64 (20%)	3/25 (12%)	8/31 (26%)	6/42 (14%)
D_1 **in immobile state**[3] **(two-state model preferred**[12]**)**	0/13 (0%)	0/3 (0%)	0/8 (0%)	1/6 (17%)
D **in immobile state**[3] **(one-state model preferred**[1]**)**	3/51 (6%)	4/22 (18%)	7/23 (30%)	19/36 (53%)
Proportion of time in immobile state[4]	0.05	0.16	0.23	0.47
Mean[5] D_0 $\times 10^4$ **nm**^{2}**/s**	9.2 ± 1.8	9.4 ± 0.6	8.9 ± 0.8	7.0 ± 1.5
Mean[5] D_1 $\times 10^4$ **nm**^{2}**/s**	3.9 ± 0.7	4.9 ± 0.5	5.0 ± 0.6	2.5 ± 1.2
Mean[6] D $\times 10^4$ **nm**^{2}**/s**	5.2 ± 0.4	5.5 ± 0.8	6.1 ± 0.9	5.1 ± 0.8
Mean[7] $D_{immobile}$ $\times 10^4$ **nm**^{2}**/s**	0.079 ± 0.025	0.057 ± 0.015	0.041 ± 0.007	0.062 ± 0.005

Diffusion coefficient units are nm^2 s^{-1} with standard error based on the number of trajectories.

1. Model selection between approximate one-state and two-state diffusion models with measurement noise, see Methods, with trajectories with no strong model preference ($-3 < \log_e B_{1D,2D} < 3$) removed.
2. Fast switching trajectories ($\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$) also removed.
3. Defined as $\log_e D < 8$ or $\log_e D_1 < 8$, for mean posterior parameters D, D_1 nm^2 s^{-1} from one-state and two-state diffusion models with measurement noise.
4. Over all trajectories with either one-state or two-state model preference, with fast switching trajectories removed (i.e. Table notes 1 and 2 apply). For one-state preference, the proportion in the immobile state is 0 if $\log_e D > 8$, 1 if $\log_e D > 8$. For two-state preference, the proportion is 0 if $\log_e D_1 > 8$, and if $\log_e D_1 < 8$ the proportion of time that was spent in the $z = 1$ state (diffusion with $D = D_1$), i.e. $1/N \sum_{i=1}^{N} \pi(z_i|\mathbf{X})$.
5. Over all posterior samples, for trajectories with two-state model preference, with fast switching trajectories removed.
6. Over all posterior samples, for trajectories with one-state model preference, restricted to $\log_e D > 8$.
7. Over all posterior samples, for trajectories with one-state model preference, restricted to $\log_e D < 8$.

This fast switching was similar to that observed for the fixed beads, questioning whether it is an experiment artifact or a true phenomena. We limited our analysis in the following to the slow (clear) switching trajectories and non-switching trajectories since the fast switchers clearly represent a different category of behaviour, irrespective of cause. Discounting those trajectories which have no strong model preference or are fast switchers, the proportion of trajectories where the two diffusion model was preferred over the one-state were: DMSO 13/64 (20%), Cyto D 3/25 (12%), PMA 8/31 (26%), PMA+Cal-I 6/42 (14%), Table A.1.

We next analysed the consistency of the diffusion coefficient estimates between trajectories. We note that the diffusion coefficients can be estimated below the measurement noise effective diffusion coefficient of $\sigma^2/(2\Delta t)$ since estimates are based on multiple time points, the error falling as $\sigma^2/(2n\Delta t)$ for n displacements. On the 4000 time points this gives a lower threshold of $\log(D) = 1.64$, so well below the lowest inferred diffusion coefficient. For both sets of trajectories, those that preferred the one-state diffusion (D) or two-state diffusion (D_0, D_1), we computed the posterior mean diffusion coefficient and pooled their posterior distributions (in the full likelihood model, Fig. 2.5, 2.6). All four conditions demonstrated similar features:

- There are two distinct clusters in the D estimates (trajectories conforming to a single homogeneous diffusion): a high (mean) diffusion coefficient greater than 3.0×10^3 nm^2 s^{-1}, and an essentially immobile state with a (mean) diffusion coefficient less than 3.0×10^3 nm^2 s^{-1}, Fig. 2.5E. Over the four conditions this split is very consistent, (Fig. A.4). We refer to these as the mobile state, with $D > 3.0 \times 10^3$ nm^2 s^{-1} ($\log_e(D) > 8$), and the immobile state, with $D < 3.0 \times 10^3$ nm^2 s^{-1} ($\log_e(D) < 8$) (classified on the posterior mean of the diffusion coefficient). The mobile state may further decompose into a 'low' and 'high' diffusing state as the pooled D distribution is bimodal with separation at 2×10^4 nm^2/s, (Fig. 2.6). The pooled distribution for D_0 also suggests a mixture distribution, although the small number of trajectories (29) makes it difficult to reliably interpret.

- Trajectories with switching of the diffusion coefficient typically exhibit two different mobile states; only 1 trajectory (out of 31) is observed to exhibit switching with the immobile state (Fig. 2.5A-D).

- The variance in the diffusion coefficient estimate for each trajectory is smaller than the variance between trajectories, (Fig. 2.6); this implies that the bimodality ('low', 'high' diffusion coefficients) is further subdivided. This explains the distinct

peaks in the pooled posterior distributions, Fig. 2.5. Thus, there is variability in the diffusion coefficient estimates suggesting the presence of a heterogeneity amongst individual trajectories.

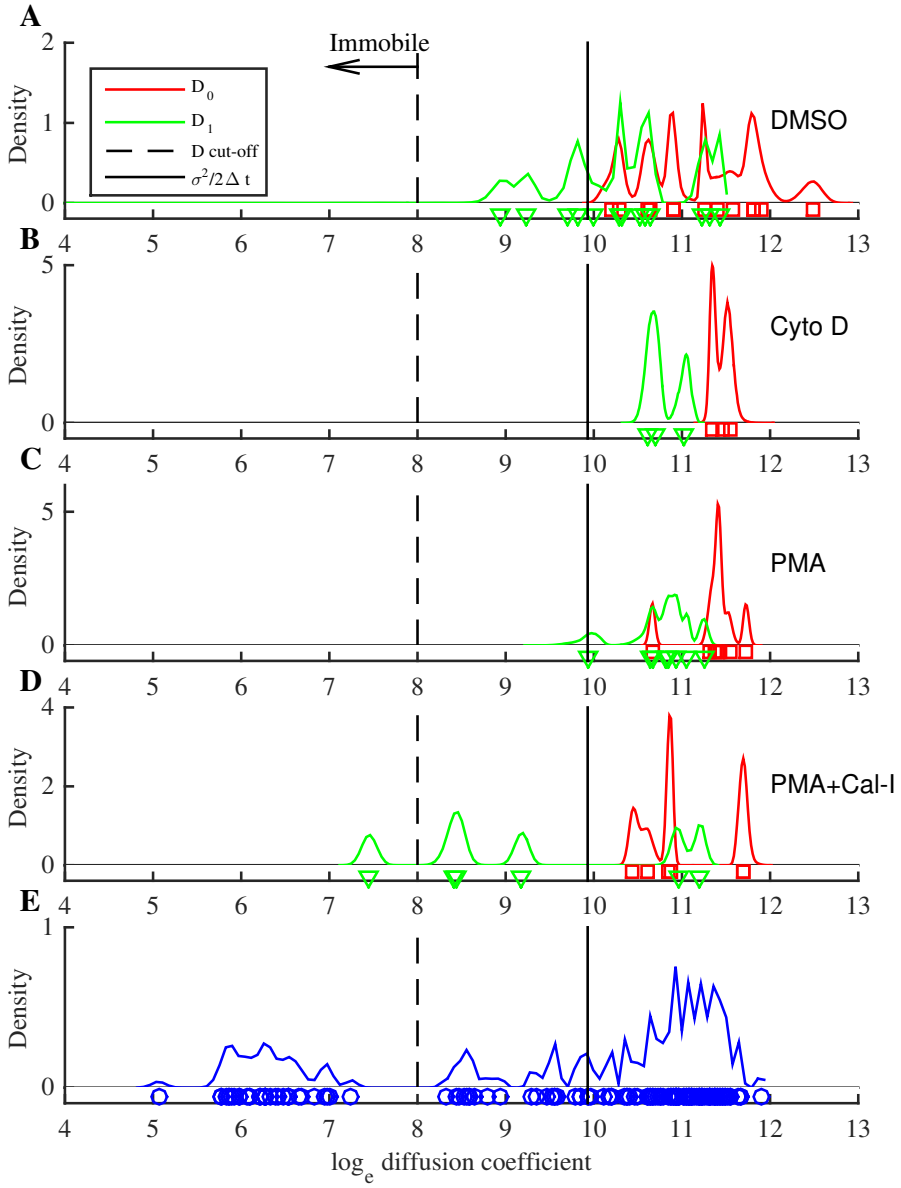

Figure 2.5: **Posterior estimates of diffusion coefficients for single LFA-1 trajectories.** (A-D) Pooled posterior samples of $\log_e D_0$ and $\log_e D_1$ for trajectories preferring the two-state diffusion model (fast switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$, trajectories removed). The posterior means for $\log_e D_0$ (red squares) and $\log_e D_1$ (green triangles), are also shown. Black line indicates value of $\sigma^2/2\Delta t$. Dashed line indicates threshold used to categorise immobile and mobile diffusion states. Treatments: (A) DMSO, two-state model preferred for 13 trajectories; (B) Cyto D, 3 trajectories; (C) PMA, 8 trajectories; (D) PMA+Cal-I, 6 trajectories. (E) Pooled $\log_e D$ estimates and posterior means (blue circles) over all treatments, for trajectories where one-state diffusion model was preferred (132 trajectories).

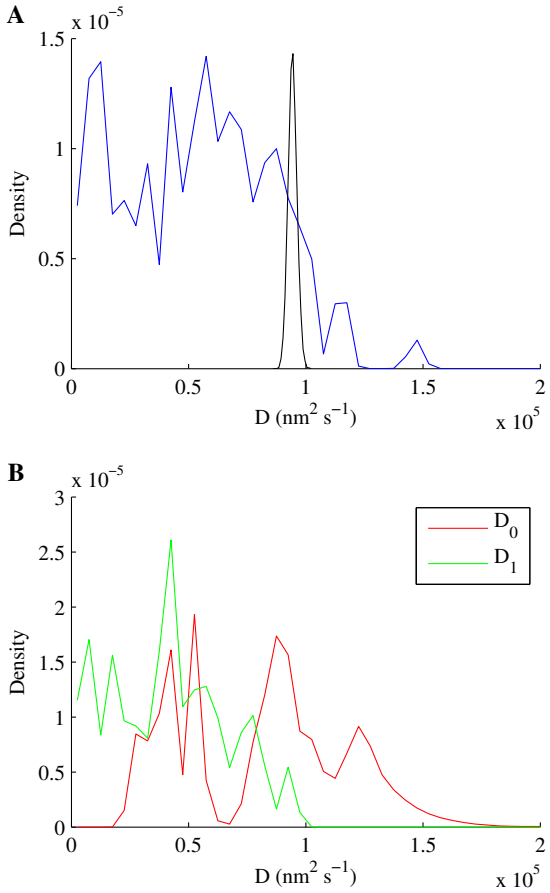

Figure 2.6: **Pooled posterior distribution of diffusion coefficients for single LFA-1 trajectories.** (A) Pooled posterior samples of D for trajectories where one-state diffusion model was preferred, restricted to $\log_e D > 8$ (99 trajectories). The posterior distribution from a single trajectory (black line, DMSO treatment) is also plotted, normalised to equal height. (B) Pooled posterior samples of D_0 and D_1 for trajectories where two-state diffusion model was preferred, restricted to $\log_e D_1 > 8$ (29 trajectories), with fast switching ($\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$) trajectories removed. One data point with $D_0 > 2 \times 10^5$ nm^2 s^{-1} not shown.

There are however differences between the four conditions. Most notably, the proportion of time in the immobile state is highest in PMA+Cal-I (47%, Table A.1). This is significantly higher than DMSO (5%, $p = 2.8 \times 10^{-8}$), Cyto D (17%, $p = 0.0091$), and PMA (23% $p = 0.031$).

For trajectories where the two-state diffusion model was preferred, (excluding the fast-switching trajectories), we examined if the diffusion coefficients between the two diffusive states are related (Fig. 2.7A). The correlation coefficient is high (r=0.84), whilst a linear relation is strongly suggested, $D_1 = 0.68D_0 - 1.5 \times 10^4$ nm^{2}s^{-1}, independent of treatment, using all points except the 2 outliers. This suggests that the switching events we are detecting are likely due to a single process. We also examined the relationship between D_0 and p_{10}, D_0 and the time in the high ($z = 0$) diffusion state and D_1 and the time in the low ($z = 1$) diffusion state, but found no correlation, Fig. 2.7B-D.

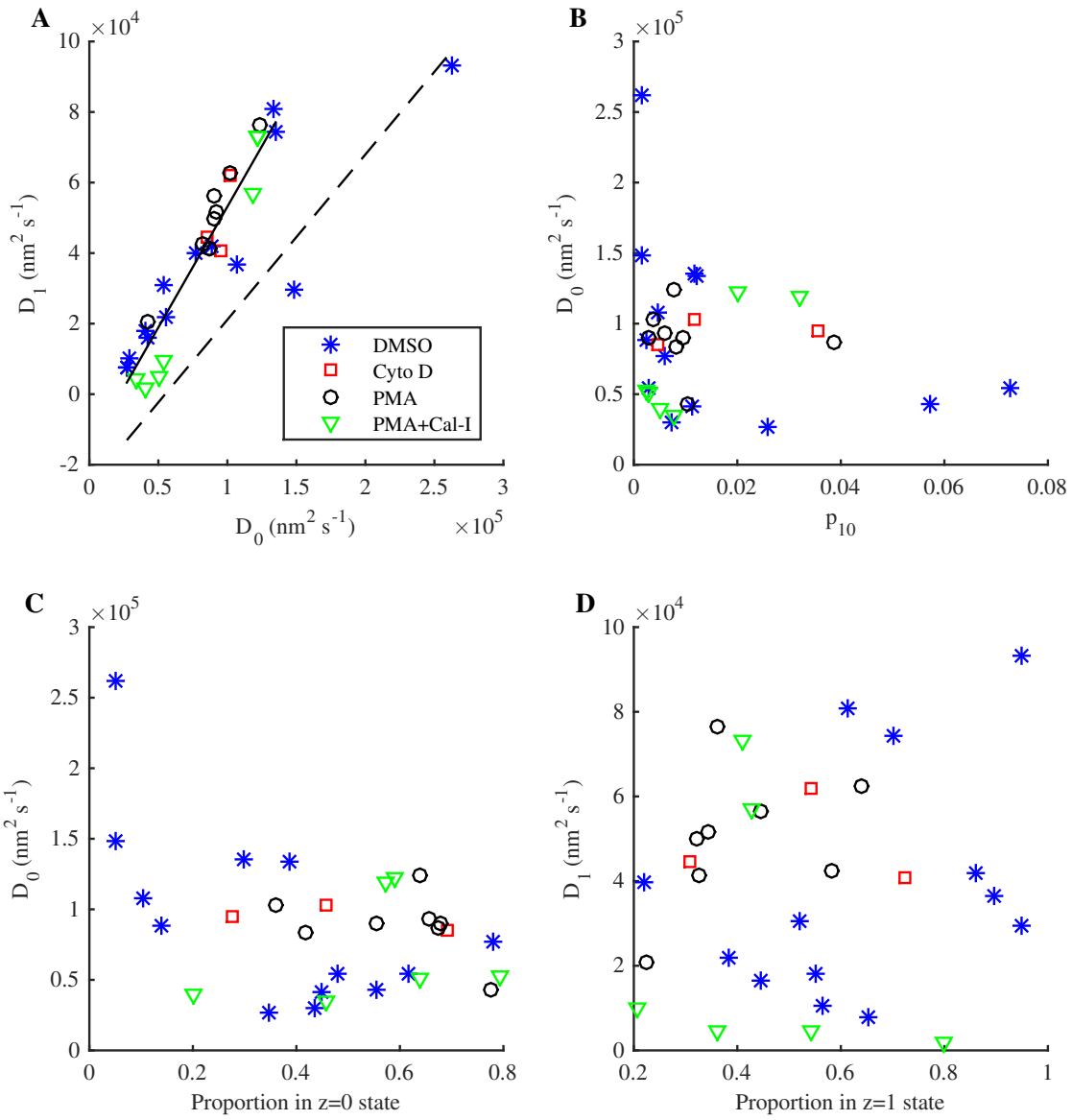

Figure 2.7: **Dependences of parameter estimates from two-state diffusion model.** (A-D) Scatter plots of posterior means for the two-state model with measurement noise, for trajectories where the approximate two-state diffusion model was preferred (fast switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$, trajectories removed). Treatments: DMSO, blue asterisks; Cyto D, red crosses; PMA, black circles; PMA+Cal-I, green triangles. In panel (A) the black solid line is a linear fit with two outlier trajectories removed, $D_1 = aD_0 + b, a = 0.68, b = -1.5 \times 10^4$ nm^2 s^{-1}; black dashed line is the double iterate, $D_1 = a(aD_0 + b) + b$.

We also examined the frequency of switching events for trajectories where the two-state diffusion model was preferred, excluding fast switching trajectories. Fig. 2.8A plots the exponentially distributed waiting times in each state (i.e. the reciprocal of the inferred transition probabilities), demonstrating a broad range of values. Some trajectories exhibit fast transient switching (Fig. 2.8A, trajectories clustered around origin, example in Fig. 2.8B), although slower than that in stationary beads. Another group of trajectories switch less frequently, with the time in a single state on the order of tenths of seconds (Fig. 2.8C-D). We also observe trajectories with very slow switching, Fig. 2.8E is an example of a trajectory with a single switch point, whilst some trajectories spend the majority of time in the $z = 0$ (fast) state, with transient switching to the $z = 1$ (slow) state (Fig. 2.8F). This variety suggests that multiple processes are affecting the waiting times since this range of behaviours would not be observed in a single exponential waiting time model.

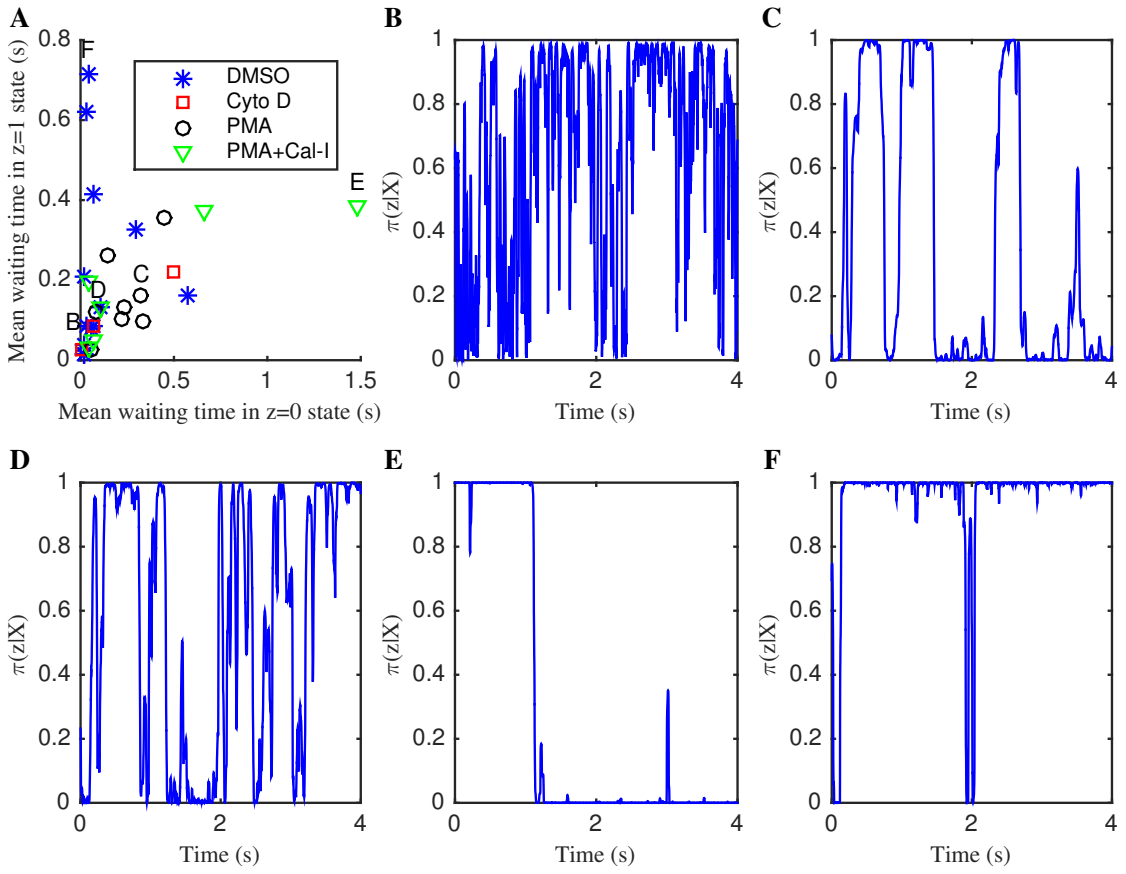

Figure 2.8: **Mean waiting times and example trajectories showing confinement for two-state diffusion model fit to LFA-1 trajectories.** (A) Mean waiting time in seconds ($1/(1000\hat{p}_{01})$ for $z = 0$ state, $1/(1000\hat{p}_{10})$ for $z = 1$ state) for trajectories where approximate two-state diffusion model was preferred (fast switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$, trajectories removed). Treatments: DMSO, blue asterisks; Cyto D, red squares; PMA black circles; PMA+Cal-I, green triangles. Labels B-F correspond to example confinement state trajectories in B-F. (B-F) Confinement state inference shown as the probability of being in the low ($z = 1$) diffusion state. (B) DMSO treatment (mean waiting time in $z = 0$ state 0.02s, in $z = 1$ state 0.04s) (C) PMA treatment ($z = 0$ state 0.32s, $z = 1$ state 0.16s) (D) PMA treatment ($z = 0$ state 0.09s, $z = 1$ state 0.12s) (E) PMA+Cal-I treatment ($z = 0$ state 1.48s, $z = 1$ state 0.39s) (F) DMSO treatment ($z = 0$ state 0.04s, $z = 1$ state 0.72s).

We examined the trajectories identified to be in the immobile state in the one-state model. These trajectories show apparent phases of linear motion in arbitrary directions, Fig. 2.9. Many trajectories have periods of consistent linear drifts in one direction, (examples in Fig. 2.10), having speeds around 110 nm s^{-1}. Some trajectories also have distinct changes in direction, (Fig. 2.10 B,C). Since the stationary beads do not show such drift, and the drift direction is variable, this is not due to microscope or sample drift (Jurkat cells in this assay being immobile [57]), but most likely reflects movements in the underlying actin cortex. These speeds are of the same order as the retrograde flow of actin in Jurkat cells (50 nm s^{-1}) [112].

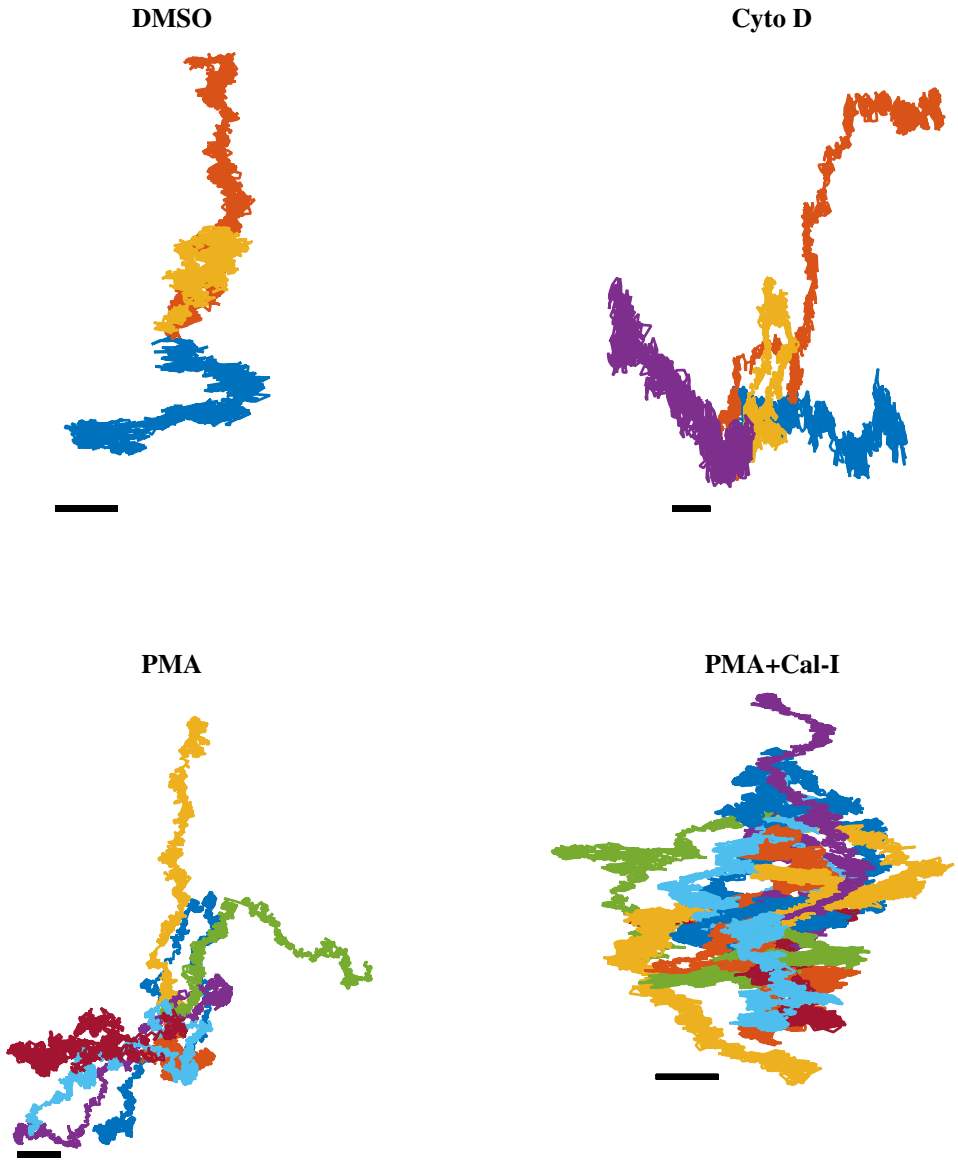

Figure 2.9: **LFA-1 trajectories categorised as immobile ($\log_e D < 8$ in the one-state model).** Trajectories are from different cells, with the first timepoints shifted to (0,0). Each trajectory is plotted in a different colour. Treatments: DMSO (3 of 75 in immobile state), Cyto D (4 of 36 in immobile state), PMA (7 of 39 in immobile state), PMA+Cal-I D (20 of 46 in immobile state). Scalebars: 50nm.

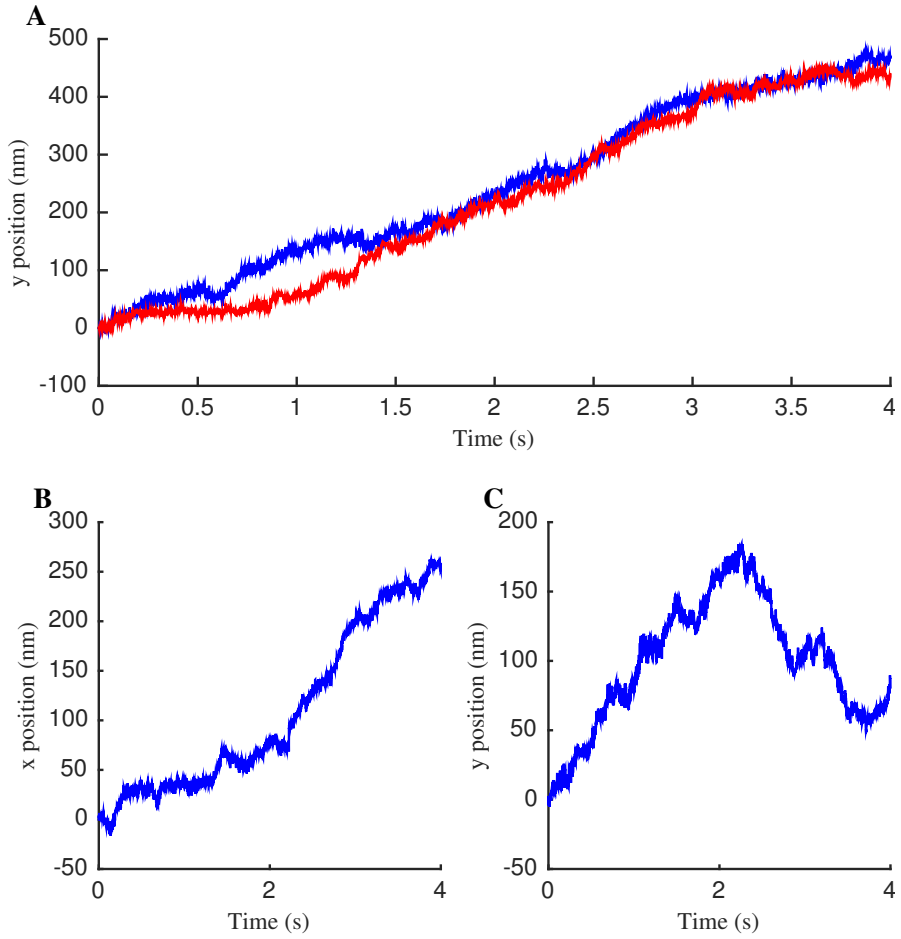

Figure 2.10: **Linear drifts in LFA-1 trajectories categorised as immobile ($\log_e D < 8$ in the one-state model).** (A) Vertical displacements for two example trajectories. Blue line: DMSO treatment, $\bar{v}_y = 117$ nm s^{-1}; red line: PMA treatment, $\bar{v}_y = 110$ nm s^{-1}. (B-C) Displacements for a trajectory (PMA treatment) with a switch in drift direction. Estimated velocities: $\bar{v}_x = 64$ nm s^{-1}, $\bar{v}_y = 80$ nm s^{-1}, (average between 0 s and 2.25 s), $\bar{v}_y = -79$ nm s^{-1} (average between 2.25 s and 3.75 s), giving an average speed of 102 nm s^{-1}.

2.3.3 Approximate versus exact models with measurement noise

We used an approximate model (low noise limit) to compute the Bayes factor to determine which of the one and two-state diffusion models are preferred by each trajectory. This approximation is justified since it gives similar results to the (exact) model on individual trajectories (Fig. A.5). On the LFA-1 trajectories that prefer the approximate model the hidden state correlation between these two algorithms is typically 80% or higher (Fig. A.6). The diffusion coefficient estimates are also highly correlated (Fig. 2.11), although they are lower under the approximation (significantly in a one-tailed Mann-Whitney test, with $p = 0.02$ for D_0 and $p = 0.001$ for D_1), indicating that failing to account for noise correlations in displacements introduces an estimation bias; this may potentially reduce the ability to detect two-state diffusion processes when the two diffusion coefficients are small (of order $\sigma^2/\Delta t$). In fact we detect no intra-trajectory switchings with both diffusion coefficients below 2×10^4 nm^{2}s^{-1}, Fig. 2.5. However, trends are similar under both analyses - in common with the one-state and two-state diffusion models with measurement noise, we also see two clear subpopulations in the posterior mean and pooled posterior samples (Fig. A.7), and a linear relationship between the D_0 and D_1 posterior means (Fig. A.8). The approximate model therefore performs well on real data although it underestimates diffusion coefficients (Fig. 2.11). Thus, we consider the approximate model sufficiently accurate for model selection SPT analysis, although parameter estimates are biased so we used the (exact) model for any estimates and interpretation after model selection.

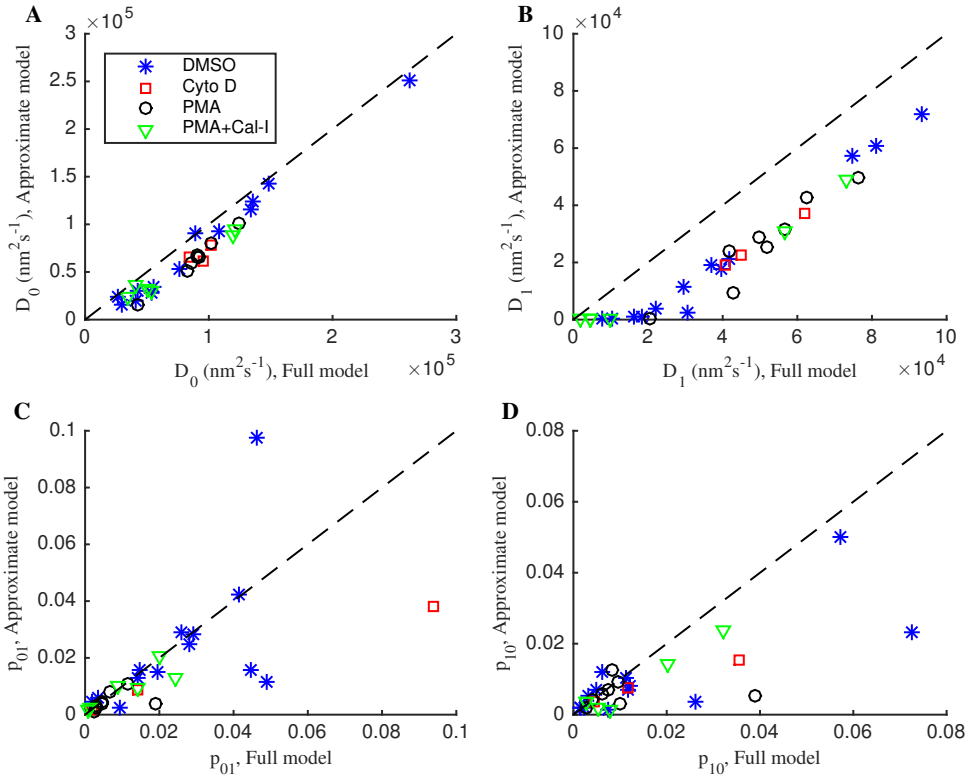

Figure 2.11: **Comparison of parameter estimates for exact and approximate two-state diffusion models with measurement noise.** (A-D) Scatter plots of two-state parameter estimates for exact model against approximate model, for 30 trajectories preferring the approximate two-state model (fast-switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$ in the exact model, trajectories removed). Line of equality is shown as dashed. Treatments: DMSO (blue asterisks), Cyto D (red squares), PMA (black circles), PMA+Cal-I (green triangles).

2.4 Discussion

We developed models and techniques for analysing single particle tracking data based on displacements between frames, including a Bayesian model selection methodology to ascertain whether the trajectory is more consistent with a one or two-state diffusion process. We confirmed the accuracy of our methods on simulated data. Two key elements of our analysis that distinguish it from other methods is the demonstration that model parameters can be estimated with high confidence from individual trajectories (1000 frames s^{-1} over 4 s), thereby not requiring trajectories to be pooled, and the inclusion of measurement noise in the trajectory inference, this propagating measurement error through to parameter estimates. We demonstrate that failure to do so leads to an inconsistency on stationary beads, (Fig. 2.1), while use of the noiseless model on the LFA-1 trajectory data results in a doubling of the detection frequency of switching within trajectories (Table A.2). In part this is a consequence of the low signal to noise ratio in this data. An alternative method to deal with this low S/N is to subsample the data so that the signal is larger (Fig. A.9). For example, modelling displacements over 4 time points reduces the effect of measurement noise. This unfortunately reduces the sample size so a balance is needed between increasing the S/N without losing too much data. We subsampled by applying a criteria per trajectory (see Appendix A.2 and Fig. A.10). This subsampling analysis gave comparative results to those obtained for the model with measurement noise on the whole data set, specifically similar numbers of trajectories showed preference for the two-state model while there is a high correlation in the model preference for each trajectory (Table A.3). This consistency between these two independent methods indicates that experimental or tracking artifacts are present, but effectively dealt with through these two alternative strategies.

Our methods were applied to single trajectories of LFA-1 tagged with latex beads under four conditions; this allowed us to show that a low but significant proportion of trajectories display within trajectory diffusion heterogeneity with switching between two distinct diffusion coefficients over a range of values ($1.6 \times 10^2 - 2.6 \times 10^5$ nm^2 s^{-1}), while the majority of trajectories conform to an homogeneous diffusion over the time scale of the trajectory. By treating each trajectory individually, rather than pooling trajectories in the analysis, we separate the heterogeneity due to this diffusive switching from a heterogeneity across trajectories, i.e. there are considerably more than two diffusive states. Previous LFA-1 studies that have pooled trajectory data miss a large component of this variability because pooling averages the heterogeneity. Basic trends are however consistent between the approaches, for instance Das *et al.*, [8] demonstrated switching

between two states which are comparable to our estimates (e.g. 8.5×10^4 nm^{2}s^{-1} and 3.1×10^4 nm^{2}s^{-1} for DMSO treatment). Three states of LFA-1 mobility have also been previously reported - "stationary", "slow" and "fast" with estimated diffusion coefficients $1.4 \pm 0.1 \times 10^4$ nm^2 s^{-1} and $5.6 \pm 0.2 \times 10^4$ nm^2 s^{-1} for the slow and fast components respectively [37]. These are broadly in agreement with the two main peaks in the diffusion coefficient distribution (Fig. 2.6A). However, our analysis demonstrates that fine detail of particle behaviour can be detected in single trajectories, in particular the diffusion coefficients can be estimated with high confidence thereby demonstrating that there is a large variability between the mobilities in individual trajectories (Fig. 2.7). The interpretation of the distribution of observed (mobile) diffusion coefficients, (Fig. 2.5E, Fig. 2.6) is subjective, for instance two Gaussians could be fitted to model the main peaks in Fig. 2.6A, thereby splitting the mobile trajectories into what could be interpreted as slow and fast populations. However, as we demonstrate here the variability is not due to measurement noise, but is intrinsic to the tagged-LFA-1 molecules, our confidence intervals per trajectory being much smaller than the range. Thus, we prefer to interpret this as a graded diffusion coefficient in a continuum. We demonstrated that for LFA-1 there is switching between diffusion states on time scales of 10-100 ms, consistent with previous analyses, [8, 75]. The former demonstrated confinement within single trajectories, corresponding to our observation of the diffusion coefficient being reduced by a factor of 1.6-23.2 under switching (Fig. A.11). However, our analysis extracts finer details than these two studies, specifically we show that there are multiple categories of behaviour, a low diffusion state consistent with immobility, and a sequence of higher diffusion states; the existence of more than two states was hinted at in the analysis of [8].

The high variability of the estimated diffusion coefficients among both fast and slow trajectories may provide biological insight into the organisation of LFA-1 in the membrane. Clustering and cytoskeletal contacts are central to the regulation of LFA-1 in the membrane [113]. Previous work has found that the movement of clusters on live cells is dependent on the conformation of the receptor [109, 57]. We propose that the multi-state diffusion observed in the current analysis is a result of changes in the size of clusters, or the number of cytoskeletal contacts for those clusters. The relationship in Fig. 2.7A suggests that the switching events we are detecting are all due to a common process. One interpretation is that we are observing diffusing aggregates of LFA-1, either in protein islands [114], or due to multiple attachments of LFA-1 molecules with the bead, a change in the aggregate size by 1 corresponding to a switch in the diffusion coefficient. We hypothesise that the diffusion coefficient reflects the size of the aggregate;

the cross section of a receptor or complex in the membrane has a predictable effect on its diffusion [34]. However, the variability in the (high) diffusion coefficients that we observe is inconsistent with this process alone. Since diffusion coefficients are observed along the straight line in Fig. 2.7A, there is an heterogeneity that determines the diffusion coefficient by smaller increments, and is presumably also responsible for the large variability in the switching frequency, (Fig. 2.8). We thus have a hierarchy of processes: on time scales less than 4 s we observe changes in the aggregate size producing large changes in the diffusion coefficient according to Fig. 2.7A, and these aggregates are also affected by a slower process that results in a finer heterogeneity (Fig. 2.6, 2.12). A potential mechanism is cytoskeletal attachment, with the number of attachments increasing with aggregate size thereby increasing the drag, and a sufficiently large number of these interactions making the receptor aggregate immobile, giving an interpretation of the non-zero intercept of the D_0/D_1 relationship in Fig. 2.7. This is consistent with calpain inhibition having the highest level of immobility, Table A.1, since calpain cleaves the talin head domain and releases LFA-1 from the cytoskeleton [110]. The fact that the mobile diffusion coefficient is reduced under calpain treatment, Table A.1, also supports the fact that cytoskeletal interactions are contributing to the aggregate drag. We also demonstrated that the immobile states (detected predominantly as immobile throughout) typically have a slow linear drift, with speeds of around 110 nm s^{-1}. We suggest that these correspond to LFA-1 (possibly clusters) strongly bound to the actin cortex, and these drift phases correspond to cortex remodelling under actin (de)polymerisation, myosin contraction or retrograde flux, [112]. Such drift was also detected by MSD analysis as super-diffusion ($\alpha > 1$) [57].

Alternative interpretations are possible, we cannot discount the possible effect of the multivalent probe on the experiment. It is possible that changes in diffusive states are the result of different numbers of contacts between the probe and receptors in the membrane. Resolving these issues will require a larger amount of data, of the order of 100s of trajectories, and ideally across different sized and variable antibody density beads.

Our analysis thus highlights the importance of large trajectory databases, with trajectory resolution and length reflecting the dynamics of the system. Quantum dots are an attractive option, since they are smaller than typical labelling molecules, and provide long trajectories [53, 55]. Ideally data on different tagging regimes is also needed to distinguish tag artifacts from molecule dynamics. With such data, sophisticated (HMM) models of temporal heterogeneity can be utilised, extending for instance to multi-diffusion states, confinement zones and drift, implemented with the algorithms and

techniques demonstrated here to analyse individual trajectories. These methods applied to large trajectory databases of long high-resolution trajectories will be an important contribution to the understanding of the complexity of membrane organisation and the multiple diffusion modalities present in cells [14].

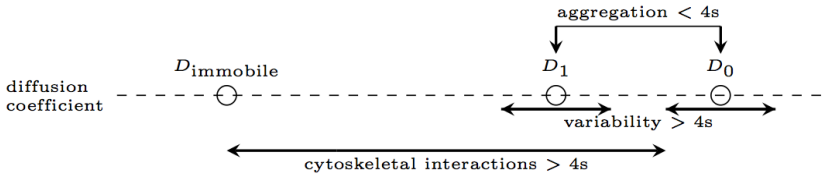

Figure 2.12: **Observed variation in the diffusion coefficient of LFA-1 in single particle tracking trajectories, with proposed mechanisms.**

Chapter 3

A Confinement Hidden Markov Model Analysis Identifies Tag-Specific Signatures within Single Particle Trajectories

3.1 Introduction

Single particle tracking (SPT) experiments offer a powerful method to analyse the membrane environment through analysis of single particle movements. However, these methods require the molecule of interest to be tagged with a (trackable) label that is imaged over a number of time steps. A number of experimental design limitations constrain the amount of information that can be extracted from such data, including spatial accuracy, temporal resolution and tracking period. New technologies are however capable of extending both the trajectory length and achieving high sampling rates and spatial resolution. A recent interferometric scattering microscopy (iSCAT) method [51, 50] can generate very long high spatiotemporal resolution trajectories ($<$ 2 nm spatial resolution, sampled at 50 kHz, tracking GM1 lipid molecules in model membranes [51]). However, a fundamental problem that impacts on interpretation is the effect of the tag itself [58]. For example, an iSCAT study demonstrated both Gaussian-like and ring-like confinement events; this was ascribed to multivalent binding of the tag [51]. Thus, a clear problem in extending this technique to *in vivo* situations is the need to separate out patterns that are due to the tag as opposed to the environment. Without achieving

this separation, the high resolution dynamics measured by these techniques may well be uninterpretable. A first step in this programme is to determine the statistical nature of the tag's signature on the trajectory.

Analysis of SPT data is not straightforward primarily because of the highly stochastic nature of diffusion, making identification of deviations from free diffusion difficult. This has led to the development of a range of statistical methods for trajectory analysis, allowing statistical significance of heterogeneities to be assessed, and avoiding over interpretation of Brownian motion. This includes a range of approaches that detect deviations from Brownian motion, such as mean square displacement (MSD) methods [60, 62, 63, 74, 42], confinement time methods [80, 81, 82, 75], and a new breed of methods implementing a model based analysis within a hidden Markov chain framework [8, 9, 10, 11, 12]. The latter methods allow for a switching of the movement dynamics to different states along the trajectory, e.g. directed versus free diffusion, so offer a powerful method to capture the detail of the different states. For high resolution data, these techniques have a clear advantage since they can utilise the high level of information within the trajectory to extract fine detail in the movement characteristics.

In this chapter, we develop a harmonic potential well confinement hidden Markov model (HMM) where the particle moves between two (hidden) states: a freely diffusing state with (to be determined) diffusion coefficient D, and confinement in a harmonic potential well (centre and strength to be determined). Working in a Bayesian framework, we derived a Markov chain Monte Carlo (MCMC) algorithm to infer model parameters and hidden states from a single trajectory. We tested the algorithm on simulated data, and then applied it to a set of experimental GM1 lipid trajectories diffusing in model membranes (generated by iSCAT microscopy). Specifically a (20nm or 40nm) gold nanoparticle (AuNP) was coated in cholera toxin B subunits (CTxB), which then bind to GM1 to form an AuNP/CTxB/GM1 complex [51], Fig. 3.1. On trajectories of 20nm AuNP/CTxB/GM1 diffusing in model membranes on a glass substrate, we detected clear periods of trapping in wells of mean radius 18nm with a mean trapping time 0.017s. However, we detected a number of deviations from homogeneity. Specifically the trapping time distribution does not fit an exponential distribution - there is a far higher frequency of longer trapping times than expected. Further, we observed that physical trapping parameters, such as the confinement radius, do not come from the same distribution across the population of trajectories. Trapping events are thus heterogeneous, suggesting different types of binding are occurring. We demonstrate that binding events within individual trajectories are more similar than between trajectories; thus indicating that the AuNP/CTxB/GM1 particles are not homogeneous and individual particles have

specific binding characteristics. We also analysed 40nm AuNP/CTxB/GM1 diffusing in model membranes on a mica substrate, and detected no trapping. This is consistent with previous analysis of this data indicating that glass substrates possess impurities where GM1 (on the lower leaflet) is immobilised, these impurities being absent on mica surfaces.

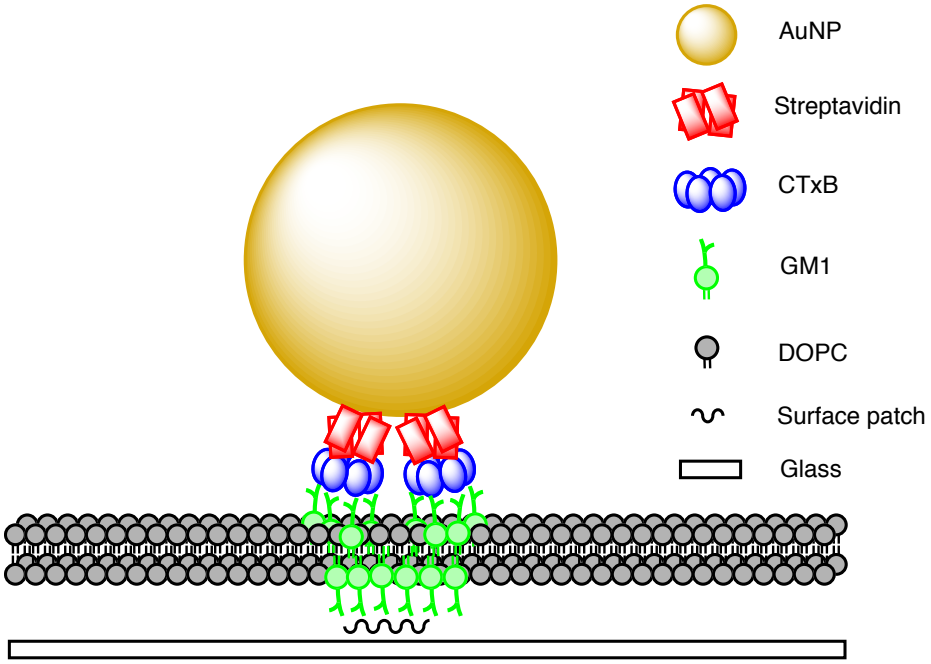

Figure 3.1: **Schematic of AuNP/CTxB/GM1 structure in Spillane *et al.* iSCAT SPT experiment.** Based on a figure in reference [51].

This chapter is organised as follows. In Results we introduce the harmonic potential well confinement model (HPW model) and associated inference (MCMC) algorithm then demonstrate accurate inference of model parameters and hidden states on simulated trajectories. We then apply the algorithm to iSCAT trajectories of AuNP/CTxB/GM1 diffusing in model membranes. The full derivation of the MCMC algorithm is then described in Methods.

3.2 Results

3.2.1 Harmonic potential well model

For a single particle tracking trajectory $\mathbf{X} = \{\Delta X_i, \Delta t_i\}_{i=1}^{N}$, where $\Delta X_i = \{\Delta X_i^{(1)}, \Delta X_i^{(2)}\}$ is 2D, we build a model where the particle switches between free and confined states. We let $z_i = 0$ if the particle is in the free state at time t_i and $z_i = 1$ for the confined state. The state z_{i+1} depends only on z_i and, assuming a constant frame rate, the transition probabilities are

$$\text{free } (z=0) \underset{p_{trap}}{\overset{p_{esc}}{\leftrightarrows}} \text{ confined } (z=1). \tag{3.1}$$

p_{trap} and p_{esc} are the probability per frame of switching to the confinement state and out of confinement respectively. The probability of being in state z_{i+1} given state z_i is therefore

$$\pi(z_{i+1}|z_i) = \text{Bernoulli}(z_{i+1}; z_i(1-p_{esc}) + (1-z_i)p_{trap}) \tag{3.2}$$

where $\text{Bernoulli}(x;p)$ denotes the Bernoulli probability distribution with variable x and parameter p. We use this notation for probability distributions throughout.

In the free state the particle diffuses freely with diffusion coefficient D. In the confined state the particle is assumed to have a directed component to its diffusive motion, proportional to the distance from the well centre $C_i = \{C_i^{(1)}, C_i^{(2)}\}$, i.e. the force is proportional to $X_i - C_i$. We refer to this as harmonic potential well (HPW) confinement. During confinement the centre diffuses much slower than the particle itself (diffusion coefficient $D_C << D$). When the particle is free C diffuses with diffusion coefficient $\hat{D}$, where $\hat{D}$ is sufficiently high that the centre can relocate between different confinement zones; the centre is thus still present even when it is not affecting the particle. In particular this means that the centre C can diffuse as fast as the particle during free diffusion; we thus use the maximum likelihood estimate of the diffusion coefficient for $\hat{D}$ estimated directly from the trajectory, although any other value can be used that is sufficiently large as it does not affect estimation of the other parameters. The stochastic differential equations (SDEs) for this model are

$$dX_t = -\kappa z_t(X_t - C_t)dt + \sqrt{2D}dW_t \tag{3.3}$$

$$dC_t = \sqrt{2\big(D_C z_t + \hat{D}(1-z_t)\big)}dW_t^{(C)}, \tag{3.4}$$

where $W_t, W_t^{(C)}$ are independent Weiner processes. The model has two hidden states to be inferred throughout the trajectory: z (the state, confined or free) and C (position

of HPW centre when confined); and five parameters: diffusion coefficients D and D_C, strength of the harmonic potential well κ, and transition probabilities p_{esc}, p_{trap}. Fig. 3.2 is a simulated HPW model trajectory, and Algorithm 1 shows the simulation algorithm in pseudocode (using the Euler approximation to the SDEs). For the simulation algorithm we used a slightly different model, where the centre C tracks X when not confined, so that when z moves from free diffusion to confinement X is close to C. (Note the difference between Equation (3.4) and the distribution of C_{i+1} in Algorithm 1.) This allows for confinement zones to be small relative to the field of view.

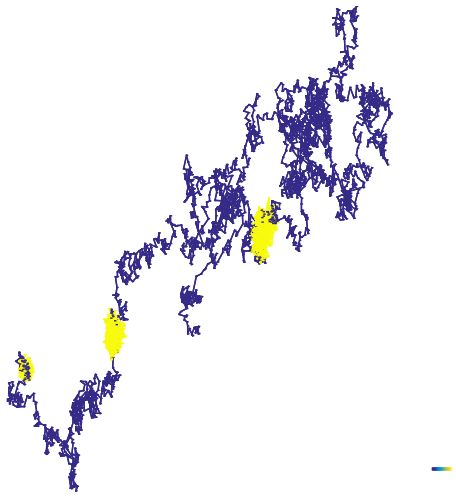

Figure 3.2: **Simulated harmonic potential well (HPW) model trajectory.** Model parameters: $D = 0.5\,\mu m^2\,s^{-1}$, $\hat{D} = 0.5\,\mu m^2\,s^{-1}$, $D_C = 0.01\,\mu m^2\,s^{-1}$, $\kappa = 3000\,s^{-1}$, $p_{esc} = 0.001$, $p_{trap} = 0.002$, time step 0.0002s. Colormap represents confinement state, with blue free diffusion and yellow confined. Colorbar length 0.1 µm.

MCMC sampler. We developed an MCMC algorithm (see Methods) to fit the HPW model to 2D trajectory data, $\mathbf{X} = \{\Delta X_i, \Delta t_i\}_{i=1}^N$. For an SPT trajectory, the algorithm samples the posterior distribution, $\pi(\theta, \mathbf{z}, \mathbf{C}|\mathbf{X})$, giving samples of the parameters $\theta^{(k)} = \{D^{(k)}, D_C^{(k)}, \kappa^{(k)}, p_{esc}^{(k)}, p_{trap}^{(k)}\}_{k=1}^K$, and the hidden states $\{\mathbf{z}^{(k)}, \mathbf{C}^{(k)}\}_{k=1}^K$, where $\mathbf{z} = \{z_i\}_{i=1}^N$ and $\mathbf{C} = \{C_i\}_{i=1}^N$. (Recall that $X_i = \{X_i^{(1)}, X_i^{(2)}\}$ and $C_i = \{C_i^{(1)}, C_i^{(2)}\}$ are 2D). We determined convergence of the MCMC sampler by calculating the Gelman point scale reduction factor (PSRF) [92].

Algorithm 1 Simulation algorithm for harmonic potential well model.

$\{D, \hat{D}, D_C, \kappa, p_{esc}, p_{trap}\} \leftarrow$ choice of model parameters
$\{\Delta t_i\}_{i=1}^{N+1} \leftarrow$ choice of time steps ($\Delta t_i = t_{i+1} - t_i$)
$X_1 \leftarrow$ initial particle position
$C_1 \leftarrow$ initial harmonic well centre position
$z_1 \leftarrow$ initial confinement state
for $i = 1$ to $i = N - 1$ **do**
    $z_{i+1} \leftarrow$ random number drawn from Bernoulli$(z_i(1 - p_{esc})) + (1 - z_i)p_{trap})$
    $C_{i+1} \leftarrow$ random number drawn from
$N\left(C_i + \kappa\Delta t_i(1 - z_i)(X_i - C_i), \left(2\Delta t_i(D_C z_i + \hat{D}(1 - z_i))\right)^{-1}\right)$
    $X_{i+1} \leftarrow$ random number drawn from $N\left(X_i - \kappa\Delta t_i z_i(X_i - C_i), (2D\Delta t_i)^{-1}\right)$
end for
$X_{N+1} \leftarrow$ random number drawn from $N\left(X_N - \kappa\Delta t_N z_N(X_N - C_N), (2D\Delta t_N)^{-1}\right)$

3.2.2 MCMC on simulated data

The HPW model sampler was extensively tested on simulated data. Figs. 3.3 and 3.4 show an MCMC run on a simulated trajectory (previously shown in Fig. 3.2). The parameter posteriors are consistent with the true (i.e. simulation) values (Fig. 3.3); the p_{trap} distribution is based on only 2 events hence explaining its poor reconstruction of the true value. Longer trajectories with more events gave consistent reconstruction in all parameters (not shown). When confined, the inferred centre closely tracks the simulated centre (Fig. 3.4A,B), and every confinement event is accurately inferred (Fig. 3.4C). During periods of free diffusion the centre diffuses; the inferred model parameters are independent of the algorithm parameter $\hat{D}$ (diffusion coefficient of the centre during free diffusion of the particle).

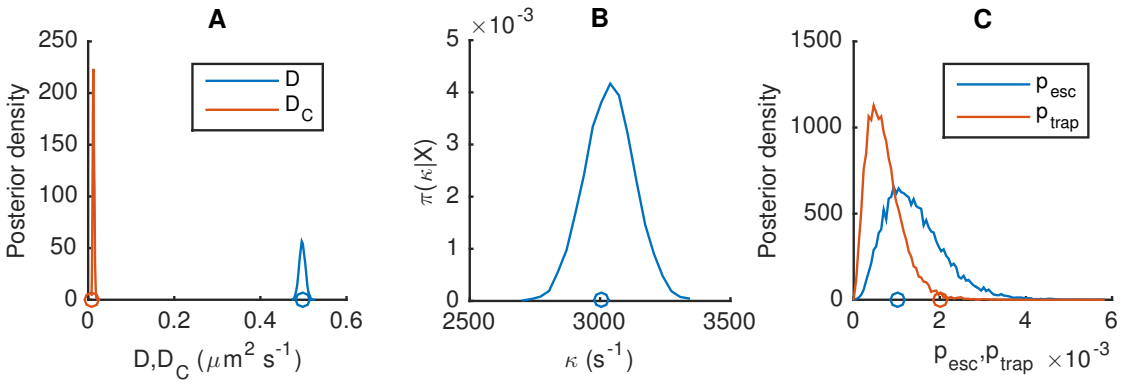

Figure 3.3: **Posterior parameter distributions of the HPW model for a simulated trajectory.** (A) Posterior distribution for D (blue) and D_C (red), with simulation values (circles), (B) corresponding MCMC values (12 independent chains of 10000 steps with a 5000 step burn in, dashed line). (C) Posterior for κ and simulation value (circle), (D) corresponding MCMC values. (E) Posterior for p_{esc} (blue) and p_{trap} (red), with simulation values (circles), (F) corresponding MCMC values. Simulation parameters: $D = 0.5\,\mu\text{m}^2\,\text{s}^{-1}$, $\hat{D} = 0.5\,\mu\text{m}^2\,\text{s}^{-1}$, $D_C = 0.01\,\mu\text{m}^2\,\text{s}^{-1}$, $\kappa = 3000\,\text{s}^{-1}$, $p_{esc} = 0.001$, $p_{trap} = 0.002$, time step 0.0002s. MCMC priors as Methods.

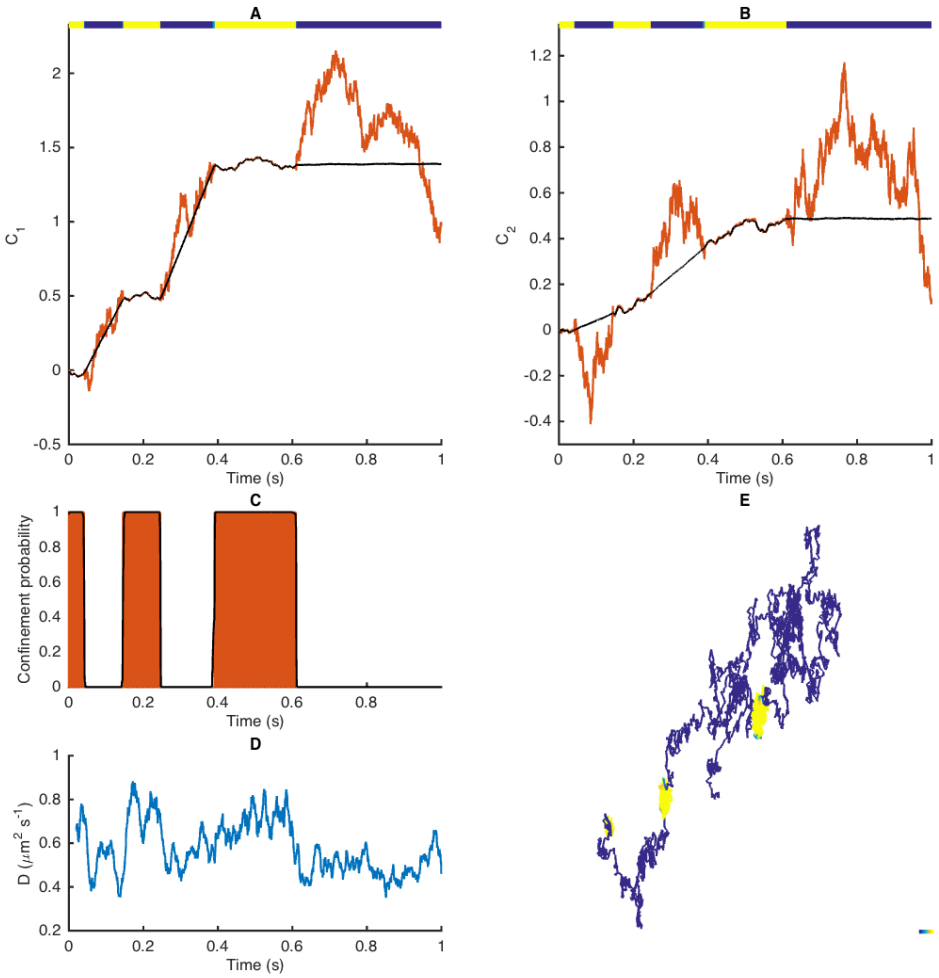

Figure 3.4: **Hidden state inference for the HPW model for a simulated trajectory.** (A-B) Mean inferred position of the harmonic potential centre in x and y directions (black), and simulated centre (red). Coloured line at the top represents particle state: free diffusion (blue) and and confinement (yellow). (C) Inferred confinement probability (black line) and simulated confinement value (red area). (D) trajectory coloured by mean inferred confinement state, from $\pi(z_i|\mathbf{X}) = 0$ (blue, free) to $\pi(z_i|\mathbf{X}) = 1$ (yellow, confined). Colorbar length 0.1μm. Trajectory and MCMC as Fig. 3.3.

3.2.3 GM1 molecules diffusing in model membranes

We applied the MCMC algorithm to previously published iSCAT SPT data [51], where CTxB coated AuNPs were introduced to a DOPC lipid bilayer containing 0.03% GM1 lipids. Confinement events were detected, which displayed two distinct spatial localisations: a Gaussian-like localisation and a ring-like confinement. The authors proposed that the CTxB attached to the AuNP (by streptavidin) binds multiple GM1 on the upper membrane leaflet; these GM1 interact with GM1 in the lower leaflet. A binding (confinement) event corresponds to an interaction with a lower leaflet GM1 that is immobilised to the glass surface; no binding events were observed for GM1 in model membranes on mica surfaces. Rocking of the AuNP around the binding site then produces a Gaussian confinement localisation. Non-Gaussian confinement distributions are explained by a second CTxB on the nanoparticle transiently binding to another, possibly diffusing GM1. These multiple-bound particles yield trajectories centered around an immobilized central CTxB that resembles a ring-like structure at the nanoscale. We investigated these confinement events in greater detail using our hidden Markov model.

The dataset includes 71 trajectories of 20nm AuNP/CTxB/GM1 diffusing in a model membrane on a glass substrate, and 18 trajectories of 40nm AuNP/CTxB/GM1 diffusing in a model membrane on a mica substrate. To determine if subsampling is necessary (to increase the signal to noise ratio), we performed an MSD analysis indicating that a subsampling of 10 is appropriate, (Appendix B.1, [62]). This analysis in fact revealed a dynamic error in the localisation efficiency at the 50 kHz sampling rate, that results in apparent superdiffusive behaviour at short times, [51]. Subsampling down to 5 kHz removes this problem. We also removed trajectory artifacts due to multiple AuNPs in the focal area (Appendix B.1).

Bayesian inference (MCMC) - fitting models to trajectories. We applied the HPW model inference algorithm individually to trajectories, thus computing samples from the parameter posterior distributions $\theta^{(k)} = \{D^{(k)}, D_C^{(k)}, \kappa^{(k)}, p_{esc}^{(k)}, p_{trap}^{(k)}\}_{k=K_B}^{K}$ and hidden state posteriors $\{\mathbf{z}^{(k)}, \mathbf{C}^{(k)}\}_{k=K_B}^{K}$, where K_B is the length of the MCMC burn in. As described in Methods, we determined convergence by calculating the Gelman point scale reduction factor (PSRF) [92], considering a trajectory converged if the PSRF was less than 1.2 for all parameters. The MCMC run length was increased for runs that failed to converge up to a maximum of 4×10^5 steps. Out of a total of 71 20nm AuNP/CTxB/GM1 trajectories, the MCMC did not converge for 1 trajectory according to this condition. (Running 4×10^5 steps for a single trajectory took around 3 days of

CPU time, hence significantly longer runs were not feasible.) The following analysis is therefore based on the set of 70 trajectories with converged MCMC runs.

3.2.4 Analysis of 20nm AuNP/CTxB/GM1 on glass trajectories

An example of the model fit is shown in Fig. 3.5, with associated (posterior) parameter estimates in Fig. 3.6. Particle state (confined or free diffusion) is well distinguished (probabilities near 0 or 1, Fig. 3.5C). The parameter estimates for D and κ are tight (low relative standard deviation), while the diffusion coefficient of the centre is very low, $D_C = 0.011 \pm 0.0012\,\mu m^2\,s^{-1}$ (mean $\pm$ s.d.) compared to $D = 0.18 \pm 0.0048\,\mu m^2\,s^{-1}$, indicating near complete immobilisation of the well. The inferred position of the well centre is also practically stationary in both coordinates during periods of confinement, Fig. 3.5A,B. As an independent measure of changes in mobility, we examined the local diffusion coefficient (Fig. 3.5D), which demonstrates a clear shift at around 0.5 s (i.e. the first inferred switch point). By colour coding the trajectory according to the probability of being confined per frame, we can also detect clear periods of confinement with approximately circular occupation profiles (Fig. 3.5E). The probability per frame of switching is reasonably well inferred (Fig. 3.6C), the broadness of the distributions coming from the small number of events. The probability of escape from a confinement zone is smaller than the probability of trapping, reflecting the short periods of time that the AuNP/CTxB/GM1 complex is undergoing free diffusion. We also observed that some confinement zones are visited multiple times, Fig. 3.7.

Since our method is applied to individual trajectories, we can examine the parameter estimates across the population of particles and determine if the population is homogeneous, Fig. 3.8. This allows quantification of the level of variability in parameters across trajectories. The marginal posterior means for D are consistent with estimates from an MSD analysis, with a mean diffusion coefficient D over all trajectories of $0.41 \pm 0.025\,\mu m^2\,s^{-1}$ ($0.34 \pm 0.03\,\mu m^2\,s^{-1}$ in the MSD analysis, Appendix B.1). However the spread between trajectories is substantially larger than the uncertainty on individual trajectories (the ratio of the mean within trajectory posterior variance to population variance of means is 0.0016 and 0.0071 for D and κ respectively). The parameter distributions would be Gaussian if the trajectories are statistically identical, i.e. derive from the same mechanistic and environmental processes. However, both D and κ show distinct deviations from being Gaussian, suggesting that binding mechanisms are not identical across AuNPs and heterogeneity is present across trajectories.

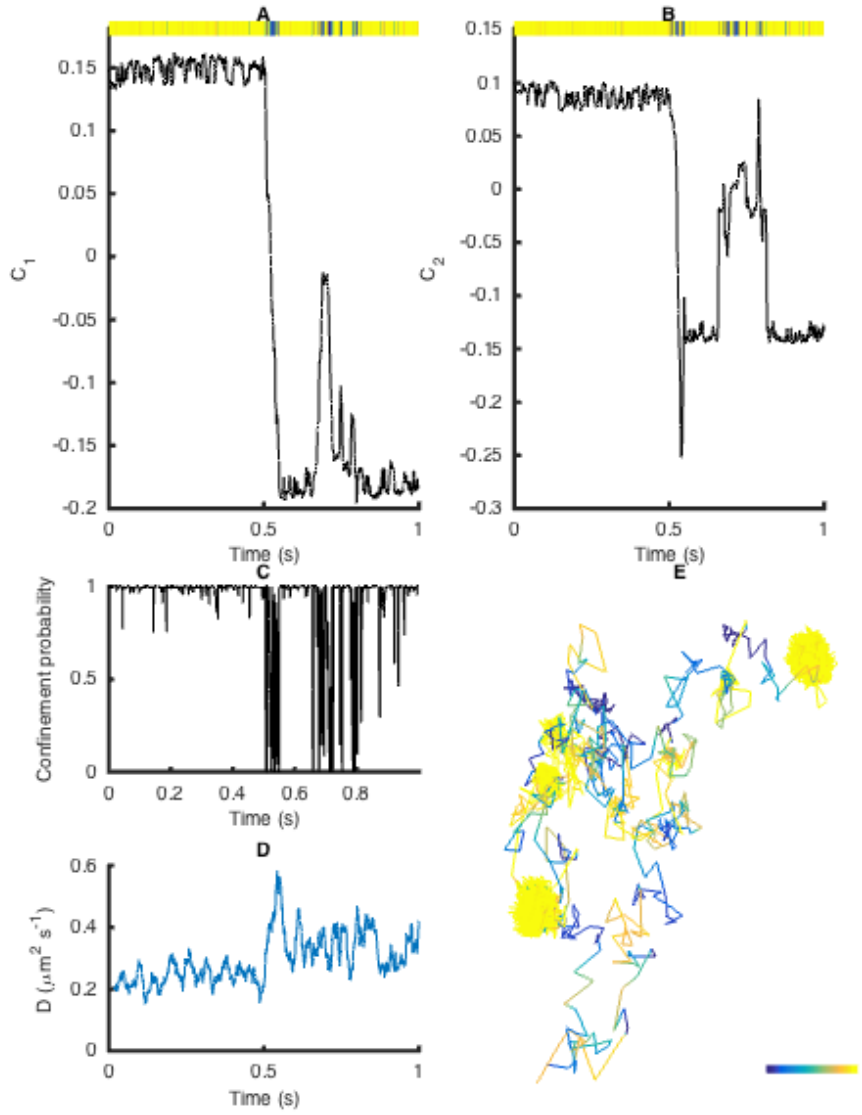

Figure 3.5: **Hidden state inference for the HPW model applied to a 20nm AuNP/CTxB/GM1 trajectory.** (A-B) Mean inferred position of the harmonic potential well centre C (x, y components) and upper coloured bar representing $\pi(\mathbf{z}|\mathbf{X})$, (colour scale goes from $\pi(z_i|\mathbf{X}) = 0$ (blue, free) to $\pi(z_i|\mathbf{X}) = 1$ (yellow, confined)). (C) Inferred mean confinement state, (D) moving average of local maximum likelihood diffusion coefficient estimate, window size 100, (E) trajectory coloured by mean inferred confinement state, colorbar length 0.1 µm. Data based on mean of 12 independent chains. MCMC priors and convergence criteria as Methods.

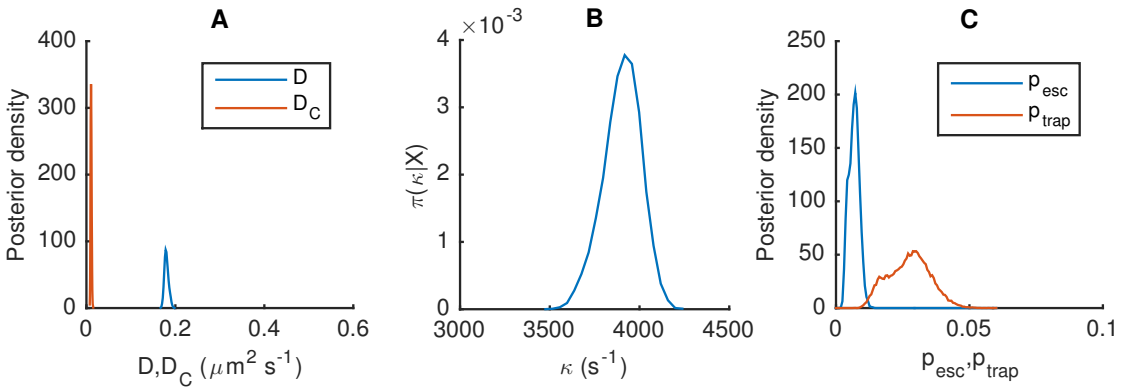

Figure 3.6: **Posterior parameters of the HPW model applied to a 20nm AuNP/CTxB/GM1 trajectory.** (A) Posterior distributions for D (blue) and D_C (red). (B) Posterior for κ. (C) Posteriors for p_{esc} and p_{trap}. Distributions consist of samples pooled from 12 independent runs. Corresponding MCMC chains shown in Fig. B.3. Trajectory and MCMC runs as Fig. 3.5.

State lifetimes. We examined the state lifetimes for confinement events and free diffusion by examining the posterior hidden state distribution of $\mathbf{z}$. For each trajectory, at each time point i, we have the probability of confinement $\pi(z_i|\mathbf{X})$. To partition the trajectory into states we turned this into a Binary signal, $\mathbf{z}^B = 0$ or $\mathbf{z}^B = 1$, for free diffusion and confinement respectively. We used a threshold of 0.5; for all i with $\pi(z_i|\mathbf{X}) < 0.5$, $\mathbf{z}^B = 0$ and for all i with $\pi(z_i|\mathbf{X}) > 0.5$, $\mathbf{z}^B = 1$. We then defined trapping events as a series of ones in the (posterior) binary state vector $\mathbf{z}^B$, and free diffusions as a series of zeros. We did not include events which contained either the first or last timepoint, as we had not witnessed the full event, and hence the exact state lifetime is unknown. Thus for some trajectories no events were included in this analysis. There is a large variation in confinement event lifetimes across trajectories (Fig. 3.9A-B), which was significant in a one-way ANOVA ($p = 6.1\times10^{-80}$, 1959 events across 59 trajectories). The lifetimes of free diffusion events were also heterogeneous, ($p = 1.2 \times 10^{-26}$, one-way ANOVA on 63 trajectories, 2011 events).

To test the assumption that switching between the two states obeys first order kinetics, we fitted an exponential distribution to all state events, and plotted a Q-Q plot of the events against these distributions, Fig. 3.9C-E. The Q-Q plot for trapping events (Fig. 3.9C) shows a distinct deviation from an exponential distribution fit (PDF: $\frac{1}{\mu}\exp -\frac{t}{\mu}$, mean event time $\mu = 0.017$s), specifically there are a far higher proportion of longer trapping events, indicative of mechanism heterogeneity. A mixture of 2 expo-

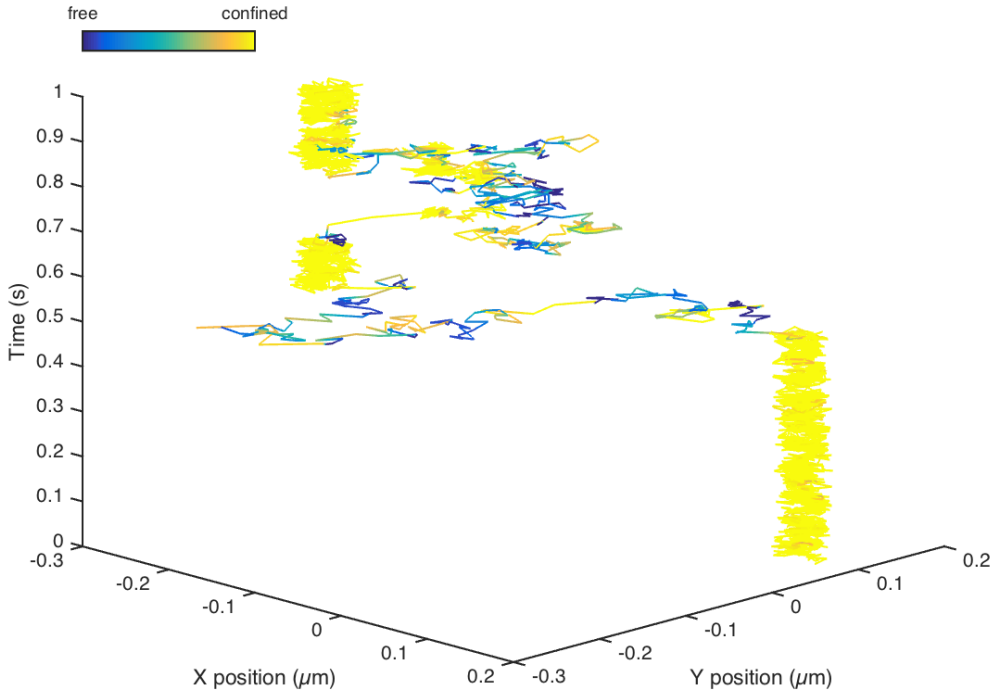

Figure 3.7: **3D representation of HPW model applied to a 20nm AuNP/CTxB/GM1 trajectory.** Colour represents the inferred hidden state. Corresponding hidden state and parameter posteriors are shown in Figs. 3.5 and 3.6 respectively.

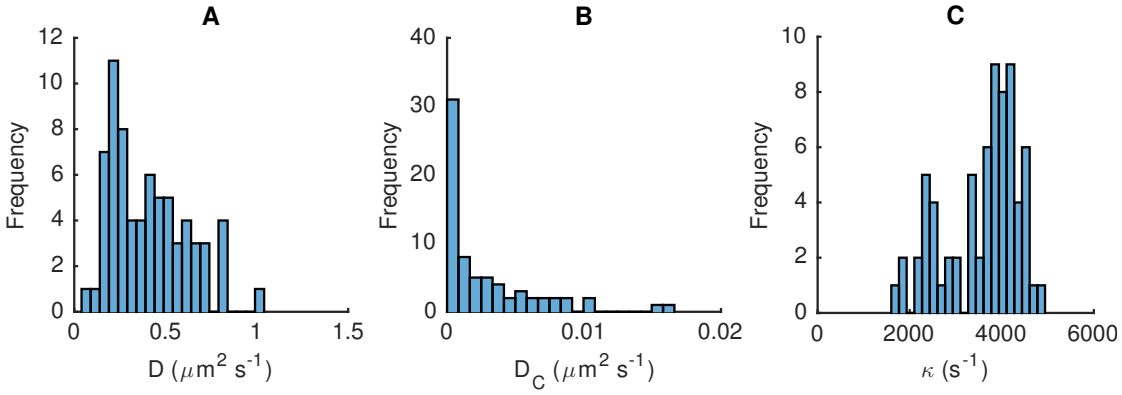

Figure 3.8: **Histogram of the mean (posterior) parameters for HPW model applied to 20nm AuNP/CTxB/GM1 on glass trajectories.** A) Particle diffusion coefficient, B) harmonic well centre diffusion coefficient, C) harmonic well strength.

nentials ($\mu_1 = 0.0034$s, $\mu_2 = 0.076$s with weights 0.81 and 0.19 respectively) is a better fit, Fig. 3.9D, suggesting that the confinement events are a heterogeneous population with at least two components. (The two-component exponential parameters are maximum likelihood estimates, for PDF $\alpha\mu_1 \exp(\mu_1 x) + \beta\mu_2 \exp(\mu_2 x)$, where α and β are the weights of the two components.) For free events the exponential fit ($\mu = 0.0023$s, mean of 11.5 frames in the subsampled data) is a much closer fit to the observed event lengths, Fig. 3.9E.

Confinement event profiling. The population heterogeneity detected above in binding time (Fig. 3.9), in the parameter estimates for κ and D (Fig. 3.8A,C), and the confinement event shape [51] may arise through a mechanism operating on indivdual events, or at the level of trajectories. To determine if there is structure in the heterogeneity we analysed trapping events restricted to events of at least 0.01s (50 subsampled timepoints). Unlike the earlier state lifetime analysis, for spatial statistics we included events which contain the first or last timepoint. We also removed events which revisited a previous trapping zone, as these events could potentially have greater similarity and dependence than spatially independent events, e.g. Fig. 3.7. We only included subsequent events if the confinement centre C was at least 30nm away from all previous event centres in that trajectory. This left a set of 325 confinement events, with the number of events within a trajectory varying from 1 (there were 6 examples where the particle remained trapped for the entire trajectory) to 12 under these restrictions. (When directly comparing spatial statistics to lifetime statistics, we clearly have to remove the events

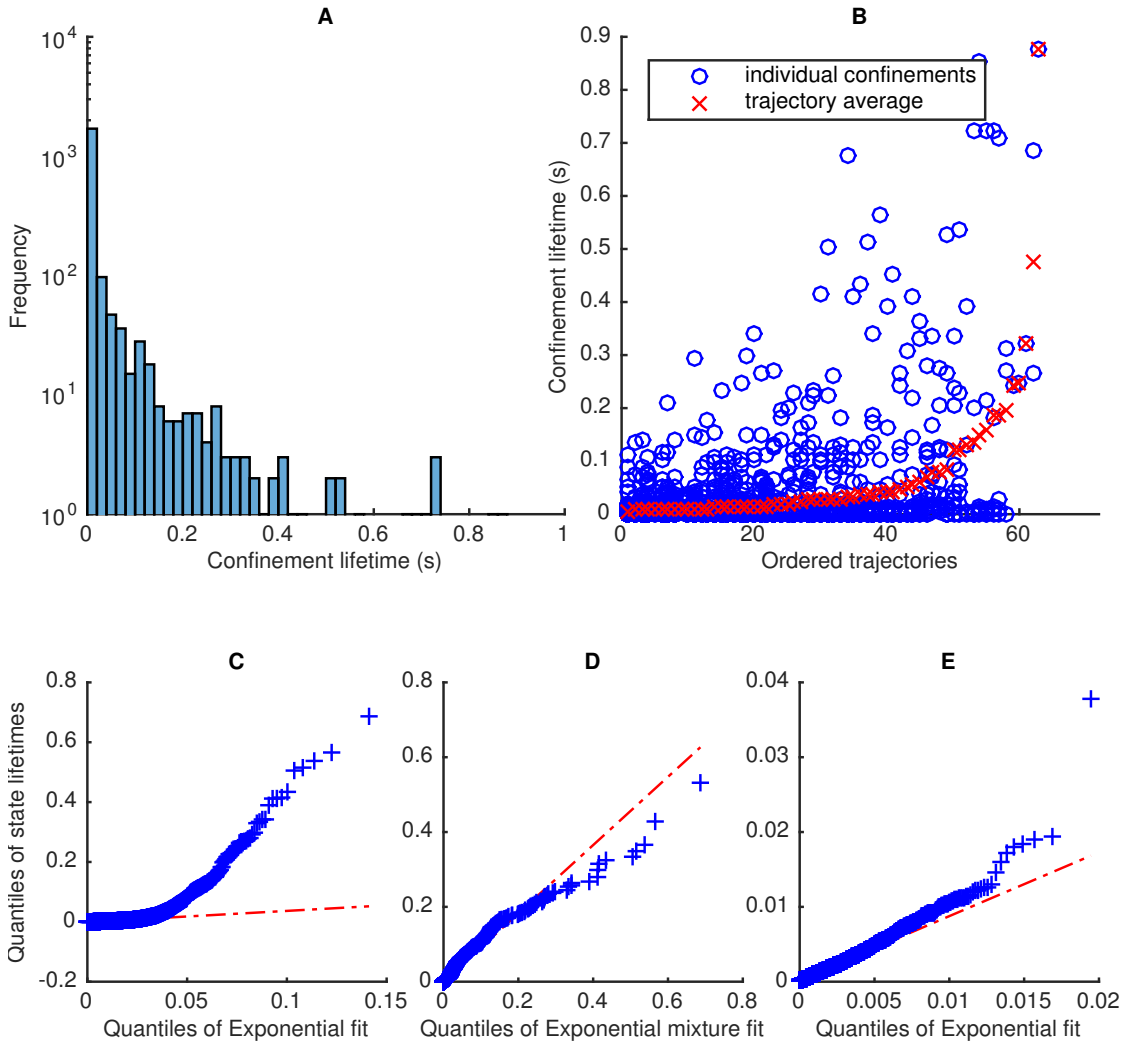

Figure 3.9: **Confinement event lifetimes are not exponentially distributed.** Plots A-E include all confinement events (no length restriction), with the exception that events which contained either the first or last timepoint are excluded. A) Histogram of all confinement lifetimes (1959 events). B) Scatterplot of confinement lifetimes against trajectories ordered by mean lifetimes. (C-E) Q-Q plots of state lifetimes against exponential fits. C) Confinement events against exponential ($\mu = 0.017$ s, $R^2 = 0$), D) confinement events against samples ($n = 10^4$) from mixture of 2 exponentials ($\mu_1 = 0.0034$ s, $\mu_2 = 0.076$ s, weights 0.81 and 0.19 respectively, $R^2 = 0.99$), E) free diffusion lifetimes (2011 events) against exponential ($\mu = 0.0023$, $R^2 = 0.98$). Red line is extrapolated linear fit to the first and third quantiles.

that include the first or last timepoint, this leaves a set of 269 events.) We denote the timepoints of the mth confinement event in the lth trajectory T_{lm}, with associated particle positions $\mathbf{X}_{lm}$ and harmonic well centre positions $\mathbf{C}_{lm}$. For each confinement event we define a number of spatial statistics, Table 3.1; the length constraint ensures that the confinement event shape can be quantified. These were based on the (Euclidean) distance between particle positions and the mean confinement centre, which we call the confinement radius. (Note that we have ignored the uncertainty in the confinement centre, and therefore this error is not propagated to the estimates of spatial statistics.) For each event we have a sample of confinement radius values, $\mathbf{R}_{lm} = \{R_i\}_{i \in T_{lm}}$, where $R_i = \|X_i - \bar{C}_{lm}\|$, from which we can compute for each event the mean confinement radius, $\bar{R}_{lm}$, and various measures of deviation of the 2D distribution from a Gaussian (Table 3.1), including the skewness of this confinement radius distribution (or "radial skewness"), S_{lm}. For a 2D Gaussian distribution the radial displacements from the mean are Rayleigh distributed, and hence have skewness $\frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}} \approx 0.63$. The mean $\bar{R}_{lm}$ over all trajectories is 18 nm (Fig. 3.10A), comparable to the size of the AuNP.

Table 3.1: **Calculation of confinement event statistics.**

Statistic	Calculation
Confinement radius	$\mathbf{R}_{lm} = \{R_i\}_{i \in T_{lm}}, R_i = \|X_i - \bar{C}_{lm}\|$
Mean confinement radius	$\bar{R}_{lm} = \frac{1}{M_{lm}} \sum_{i \in T_{lm}} R_i$
Radial skewness	$S_{lm} = \frac{\mathbb{E}[\mathbf{R}_{lm} - \bar{R}_{lm}]^3}{(\mathrm{var}[\mathbf{R}_{lm}])^{3/2}}$
Radial mean-median distance	$\|\bar{R}_{lm} - \hat{R}_{lm}\|$
Radial SD	$\sqrt{\mathrm{var}[\mathbf{R}_{lm}]}$

We denote the set of timepoints of the mth trapping event in the lth trajectory T_{lm}. Events have associated particle positions $\mathbf{X}_{lm} = \{X_i\}_{i \in T_{lm}}$ and harmonic well centre positions $\mathbf{C}_{lm} = \{C_i\}_{i \in T_{lm}}$. When calculating these statistics we restricted to trapping events of at least 0.01s. In general, we also removed events which revisit a previous trapping zone (we didn't include events if $\bar{C}_{lm}$ was within 30nm of a previous $\bar{C}_{lm}$ within trajectory l). In some cases we also included events that contained the first or last timepoint (leaving a set of 325 events). In other cases (generally when comparing to event lifetime statistics) we remove them, leaving a set of 269 events (or 486 if including events revisiting the same location). M_{lm} is the number of timepoints in T_{lm}, thus $\bar{C}_{lm} = \frac{1}{M_{lm}} \sum_{i \in T_{lm}} C_i$ is the mean posterior harmonic well centre. $\|.\|$ denotes the Euclidean distance, and $\hat{}$ denotes the median.

It is reasonable to suppose that both the non-exponential lifetime and the non-

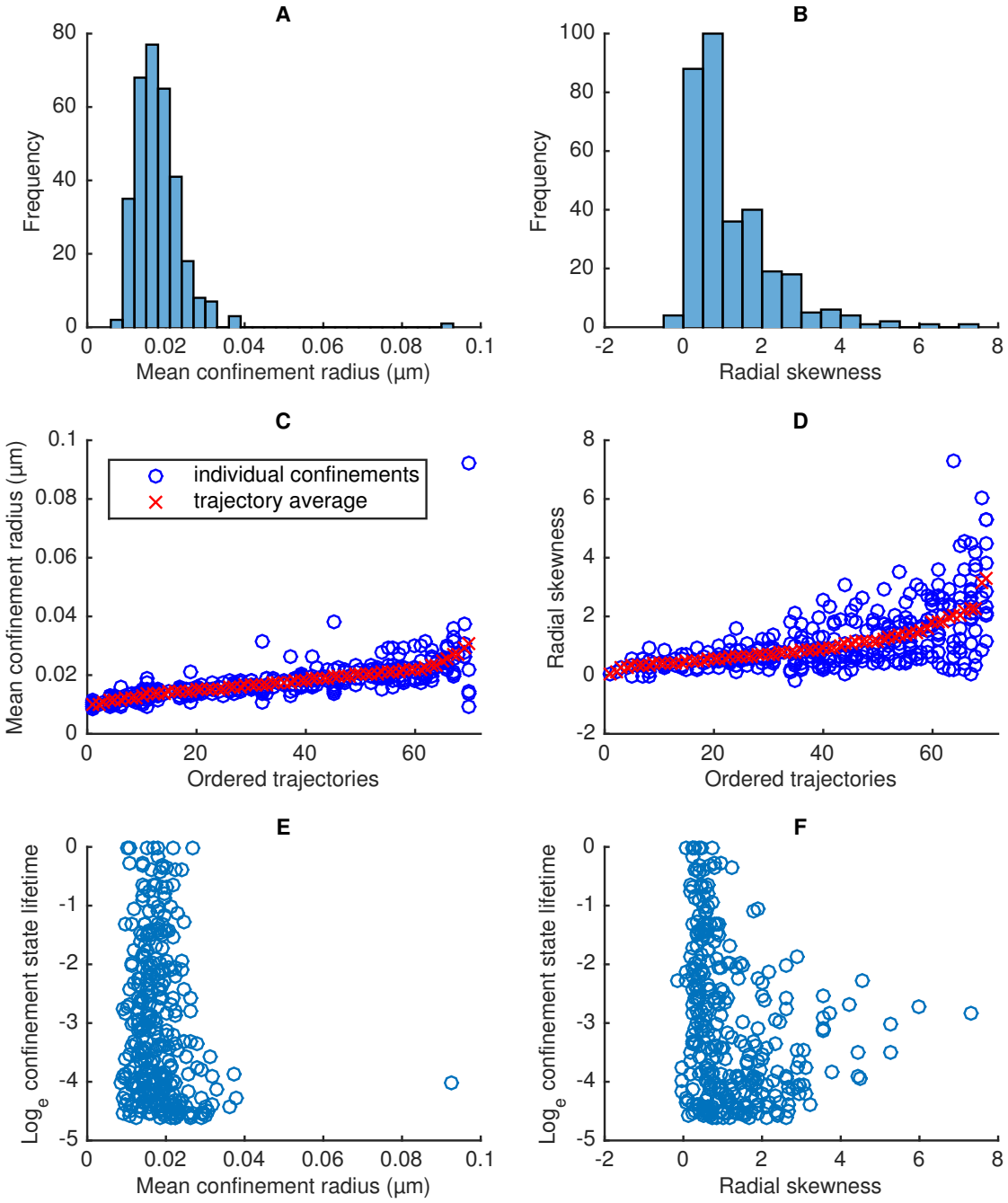

Figure 3.10: **Shape statistics for confinement events in 20nm AuNP/CTxB/GM1 trajectories.** Plots include all confinement events of at least 0.01s, with events revisiting a previous trapping zone removed (giving 325 events, as described in Table 3.1). (A-B) Histograms over all confinement events. (C-D) Scatterplots against trajectories ordered by average within trajectory value of the statistic. (E-F) Scatterplots of statistics against (log) event lifetime.

Gaussian shape of confinement events may have the same underlying cause. However, we found no correlation between confinement state lifetime and $\bar{R}_{lm}$ (Table 3.2, Fig. 3.10E), and only a weak correlation between S_{lm} and lifetime, (Table 3.2, Fig. 3.10F). This suggests that the non-exponential nature of the binding lifetime is only weakly related to the shape of the binding event.

Table 3.2: **Correlations between confinement event statistics.**

	Event lifetime	R_{lm}	S_{lm}
Mean confinement radius ($\bar{R}_{lm}$)	-0.06 ($p = 0.30$)		
Radial skewness (S_{lm})	-0.24 ($p = 9.8 \times 10^{-5}$)	-0.30 ($p = 6.0 \times 10^{-7}$)	
Radial mean-median distance	-0.27 ($p = 7.0 \times 10^{-6}$)	0.14 ($p = 0.024$)	0.31 ($p = 2.8 \times 10^{-7}$)
Radial SD	-0.21 ($p = 6.7 \times 10^{-4}$)	0.53 ($p = 1.4 \times 10^{-20}$)	0.09 ($p = 0.15$)

Pearson correlation coefficient between confinement statistics, for the set of 269 events obtained by applying the restrictions detailed in Table 3.1. Exact calculations for confinement statistics are also given in Table 3.1.

Confinement event statistics varied across the population of events (Fig. 3.10 A-B) with a large variation across trajectories (Fig. 3.10 C-D). A one-way ANOVA showed that the population was not homogeneous and within trajectory variation was lower than between trajectories for $\bar{R}_{lm}$ ($p = 2.3 \times 10^{-12}$, 325 events across 70 trajectories), and S_{lm} ($p = 3.2 \times 10^{-8}$, 325 events across 70 trajectories). (This was more significant if including events revisiting the same location: for R_{lm} $p = 5.5 \times 10^{-47}$, 542 events across 70 trajectories; for S_{lm} $p = 6.8 \times 10^{-18}$, 542 events across 70 trajectories.) This compared to confinement event lifetimes as calculated earlier ($p = 6.1 \times 10^{-80}$, one-way ANOVA), indicating that lifetimes also have greater variability between trajectories. For instance, some particles remain trapped for the entire trajectory (1 s), whereas others rapidly switch in and out of confinement. Additionally, the mean variance within trajectories was smaller than the variance across all events for $\bar{R}_{lm}$ and S_{lm} with ratios of around 0.6 (Table 3.3). Events within a trajectory are thus more similar to each other than events in different trajectories.

Table 3.3: **Comparison of within and between trajectory variances for confinement event statistics.**

Statistic	Mean within trajectory variance	Variance across all events
Mean confinement radius ($\bar{R}_{lm}$)	$2.5 \times 10^{-5}\,\mu m^2$	$4.4 \times 10^{-5}\,\mu m^2$
Radial skewness (S_{lm})	0.69	1.17
Radial mean-median distance	$7.0 \times 10^{-7}\,\mu m^2$	$1.1 \times 10^{-6}\,\mu m^2$
Radial SD	$4.5 \times 10^{-6}\,\mu m^2$	$9.0 \times 10^{-6}\,\mu m^2$
Confinement event lifetime	$0.023\,s^2$	$0.019\,s^2$

Spatial statistics calculated on the set of 325 events obtained by applying the restrictions given in Table 3.1. Lifetime statistics calculated on set of 269 events, also described in Table 3.1.

To determine which physical variable is most strongly conserved within individual trajectories, we clustered (using k-means with squared Euclidean distance) our confinement event statistics and calculated the sum of the Shannon diversity index over all trajectories (Fig. 3.11). This gives a measure of how events from the same trajectory are distributed across the clusters, a lower Shannon Index indicating that events from the same trajectories remain grouped in the same cluster. Since we are comparing spatial and lifetime statistics, we removed events that include the first or last timepoint (leaving 269 or 486 events, depending on the inclusion of events revisiting the same location). The property that is most highly preserved in trajectories is the mean confinement radius ($\bar{R}_{lm}$), whilst the other statistics all performed better than random, (Fig. 3.11). Spatial statistics performed better when repeat confinement events were included (Fig. 3.11B). This supports the conclusion above that an individual AuNP/CTxB/GM1 complex determines the characteristics of the binding events throughout the trajectory, events within a trajectory being more similar than events between trajectories.

In order to illustrate the conservation of statistics within trajectories, and their variation across trajectories, we compared two representative trajectories with 18 confinement events of at least 0.01s (11 and 7 on removal of events revisiting the same location), Fig. 3.12 A,B. A histogram of the particle positions pooled over the confinement events (as defined to give 325 events in Table 3.1) in each trajectory demonstrates a clear difference in 2D confinement shape, Fig. 3.12 C, E. The mean confinement radius ($\bar{R}_{lm}$) for events was significantly different across these trajectories ($p = 0.0019$, two-tailed t-test); the difference between the skewness ($p = 0.34$) and the confinement lifetime ($p = 0.40$) was not significant. However, the confinement event signatures are stronger if confinement events that revisit a previous trapping zone are included, Fig. 3.12 D, F. In this case, there is also a significant difference in the skewness ($p = 0.04$), yet the confinement lifetime is still insignificant ($p = 0.93$).

These trajectories highlight that a clear confinement signature is present in trajectories, thereby indicating that an individual nanoparticle determines the nature of the confinement events, each having a similar confinement time, radius of confinement and degree of non-Gaussian localisation. We also plotted the representative confinement histograms for all 70 20nm AuNP/CTxB/GM1 trajectories, which showed a wide variety of confinement shapes (ordered by the average within trajectory $\bar{R}_{lm}$, Fig.B.4; and by the average within trajectory S_{lm}, Fig. B.5).

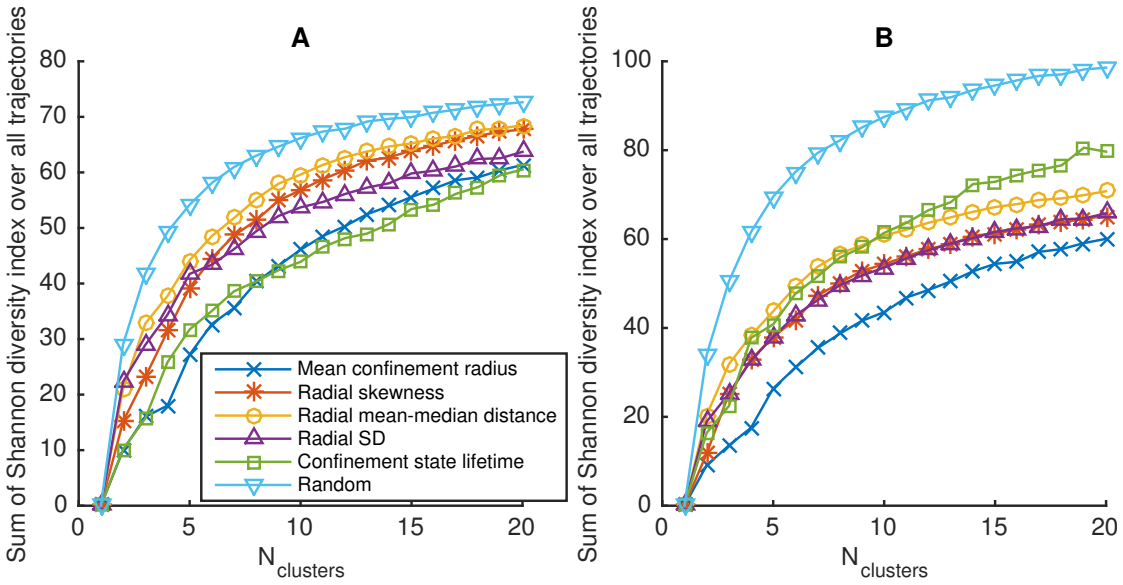

Figure 3.11: **Clustering of confinement events in 20nm AuNP/CTxB/GM1 trajectories.** Individual confinement events were clustered (k-means with squared Euclidean distance metric) based on event statistics. For each trajectory, l, the Shannon diversity index, $H_l = -\sum_{j=1}^{N_{\text{clusters}}} p_j \log p_j$, was calculated ($p_j$ is the proportion of the events in trajectory l that appeared in cluster j). The sum of the Shannon diversity index over all trajectories is then a measure of the similarity of events within trajectories. The event statistics (shown in the legend) are defined in Table 3.1. For each choice of N_{clusters}, 50 separate clusterings were performed. and the sum of the Shannon diversity index was averaged over these clusterings. A) 269 events, obtained as described in the main text and Table 3.1. B) Same as (A), except with events which revisited a previous trapping zone included (486 events).

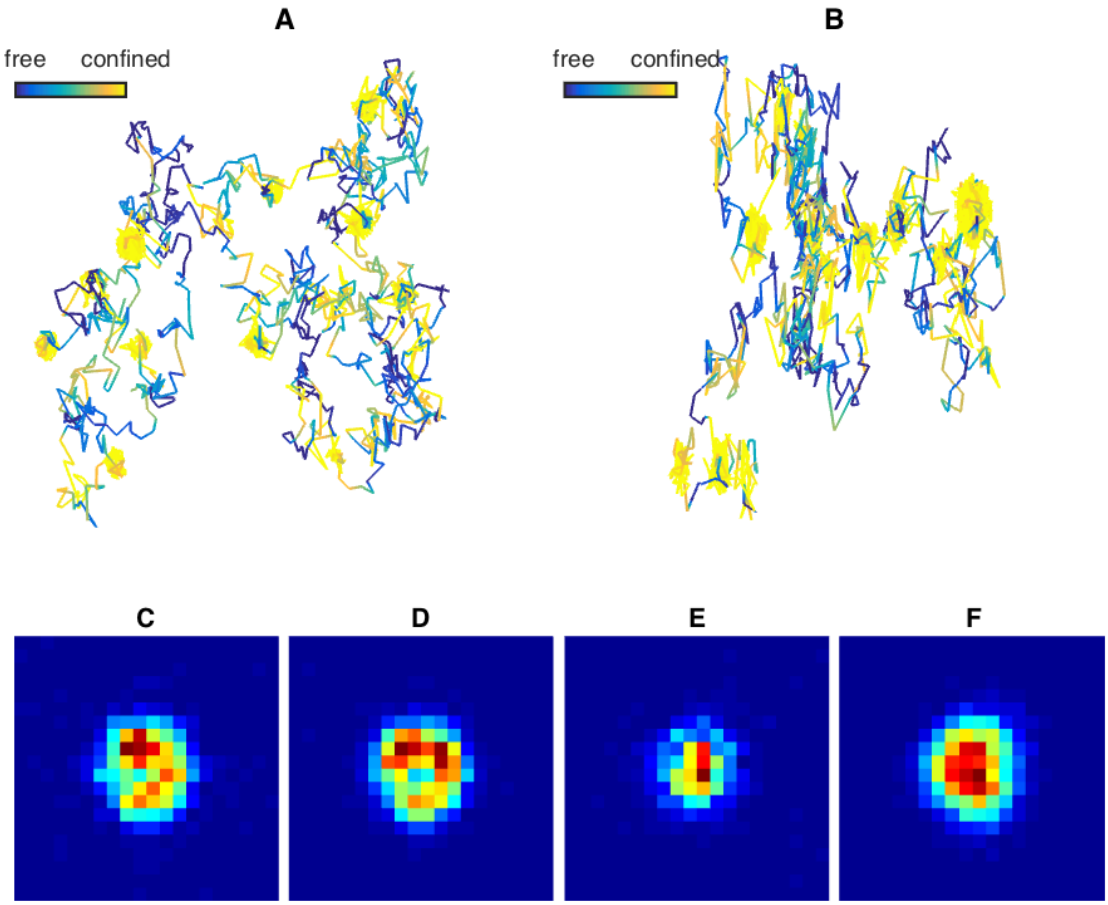

Figure 3.12: **Two example 20nm AuNP/CTxB/GM1 trajectories.** (A-B) Trajectories coloured by mean inferred confinement state, from $\pi(z_i|\mathbf{X}) = 0$ (blue, free) to $\pi(z_i|\mathbf{X}) = 1$ (yellow, confined), colorbar length 0.1 µm. (C-D) Histograms of particle positions pooled over all confinement events for trajectory A. For each event, the mean particle position was shifted to $(0, 0)$. For (C) confinement events were included based on the criteria in Table 3.1; for (D) we also included events that revisit a previous trapping zone (C: 11 events, average $\bar{R}_{lm} = 0.014\,\mu m$; D 18 events, average $\bar{R}_{lm} = 0.014\,\mu m$). (E-F) Corresponding pooled histograms for trajectory B. (E: 7 events, average $\bar{R}_{lm} = 0.012\,\mu m$; F: 18 events, average $\bar{R}_{lm} = 0.012\,\mu m$). Side length of images (C-F) is 0.1 µm.

To dissect the relationship between event lifetime and confinement shape further we split the population of events (269, Table 3.1) into two equal groups by thresholding $\bar{R}_{lm}$ and S_{lm} at the population medians. For $\bar{R}_{lm}$ this gives a value of 0.017 µm; a Q-Q plot (Fig. 3.13A) and two-tailed t-test ($p = 0.56$) are consistent with the confinement event lifetimes of these groups coming from the same distribution. This strongly suggests that the non-exponential nature of trapping lifetime (Fig. 3.9) does not have the same underlying cause as the observed heterogeneity in the confinement event radius between trajectories. We also split S_{lm} at the median (0.96) which yielded groups where event lifetimes had different distributions, Fig. 3.13B. There are clearly more events with non-Gaussian characteristics (i.e. $S_{lm} > 0.96$) and short lifetimes, Fig. 3.14A, and the difference between these groups is significant ($p = 5.8 \times 10^{-7}$, two-tailed t-test). For $S_{lm} > 0.96$ there is a reasonable fit to an exponential distribution (Fig. 3.14C, $R^2 = 0.93$). Thus there may be a common mechanism causing both the non-exponential lifetime of confinement events (Fig. 3.9) and higher radial skewness (Fig. 3.10B,D,F).

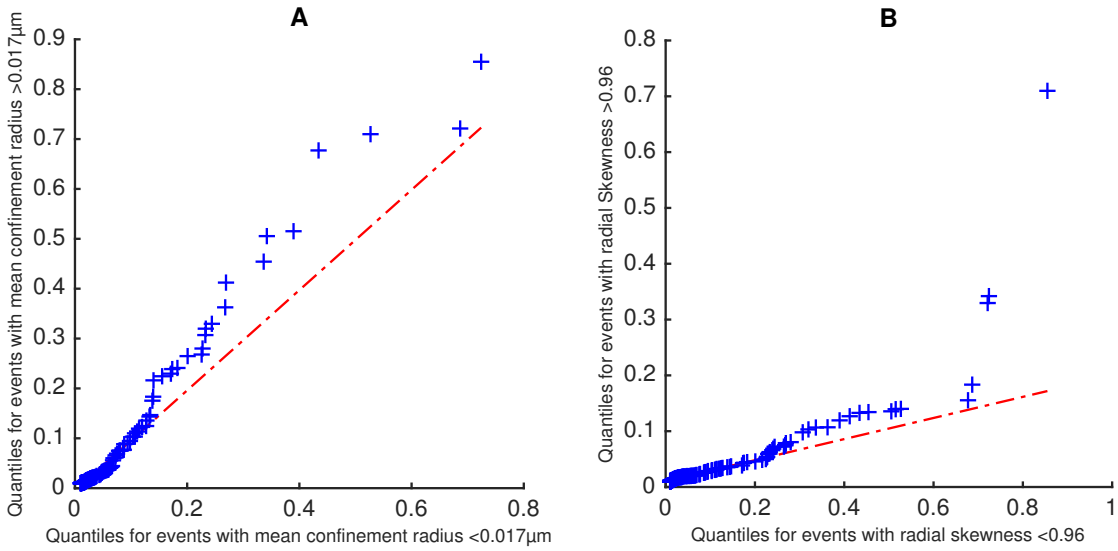

Figure 3.13: **Q-Q plots of confinement event lifetimes split by confinement shape statistics.** Lifetimes were split into two groups (134 events in each group) by thresholding at the median of the shape statistic, then the quantiles of these groups were plotted against each other. (A) Events grouped by confinement radius ($\bar{R}_{lm}$) with threshold 0.017 µm (mean event times: 0.08 s for $\bar{R}_{lm} < 0.017$ µm, 0.09 s for $\bar{R}_{lm} > 0.017$ µm). (B) Events grouped by radial skewness (S_{lm}) with threshold 0.96 (mean event times: 0.13 s for radial skewness < 0.96, 0.05 s for radial skewness > 0.96).

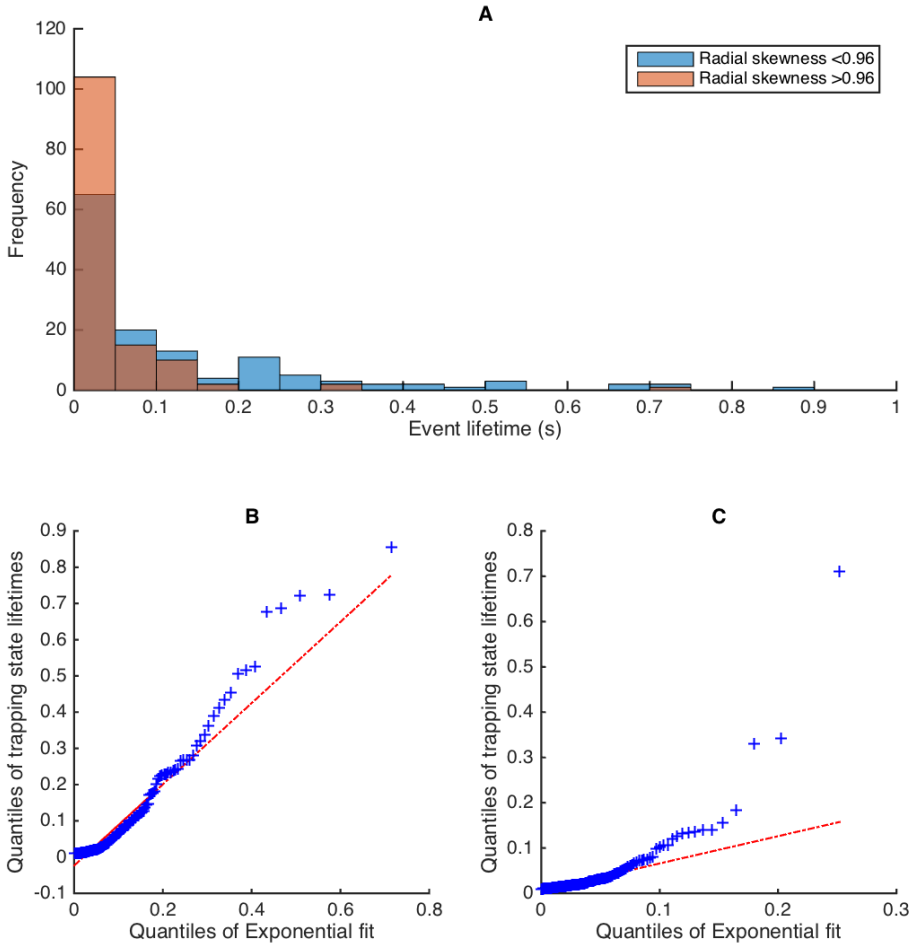

Figure 3.14: **Confinement event lifetimes are affected by radial skewness.** (A) Histogram of confinement event lifetimes, split into two groups (134 events in each group) by thresholding the median radial skewness over all events. (B-C) Q-Q plots of corresponding grouped confinement event lifetimes against exponential fit. (B) Radial skewness < 0.96 (parameter for exponential fit, $\mu = 0.13\,\mathrm{s}$; $R^2 = 0.72$), (C) radial skewness > 0.96 ($\mu = 0.05\,\mathrm{s}$, $R^2 = 0.93$).

Analysis of 40nm AuNP/CTxB/GM1 on mica trajectories. As a control, we analysed trajectories of 40nm AuNP/CTxB/GM1 diffusing in SLBs on a mica substrate. A previous analysis demonstrated that no confinement for this treatment was present [51]. We applied the HPW model to 18 40nm AuNP/CTxB/GM1 trajectories; all MCMC chains converged according to the Gelman PSRF condition (described in Methods). No confinement events were detected in any trajectory (Fig. 3.15).

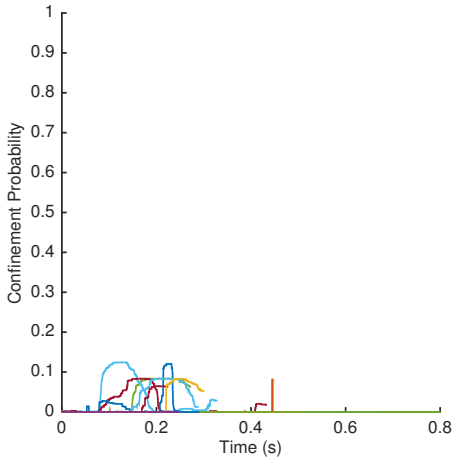

Figure 3.15: **GM1 trajectories in SLBs on a mica substrate show no confinement.** Inferred mean confinement state (average from 12 MCMC chains) from the harmonic potential well model applied to 40nm AuNP/CTxB/GM1 trajectories in SLBs on mica. 18 trajectories of varying lengths (0.16-0.8s) each plotted in a different colour. The highest confinement (posterior) probability in any trajectory, at any time was 0.12.

3.3 Discussion

We developed a harmonic potential well HMM for the partitioning of single particle tracking trajectories into periods of free diffusion and confinement within a Bayesian framework. The model determines the times for switching between the two states, the diffusion coefficient D, and characteristics of the confinement: potential well strength κ, centre C and centre diffusion coefficient D_C. On simulated data confinement zones were accurately detected, and on experimental GM1 trajectories (that are known to display immobilisation) it partitioned the trajectory with high confidence, Fig. 3.5. By examining partitioned events we also demonstrated heterogeneity in the confinement events with respect to the confinement strength (potential well strength κ, Fig. 3.8C), confinement radius (Fig. 3.10), and confinement lifetime (Fig. 3.9). Event charac-

teristics were also more similar within trajectories than between. We also used a two diffusion coefficient HMM to partition these trajectories (data not shown), specifically a model where the particle's diffusion coefficient switches between two (or more) states [8, 9, 10, 11, 12]. Using a Bayesian algorithm for inference of this model (i.e. [115]) failed to detect periods of immobility. This implies that the the diffusion coefficient of the AuNP/CTxB/GM1 complex doesn't change during confinement events; the wobble of the nanoparticle within the confinement zone allowing the diffusion coefficient to be estimated because of the high positional accuracy and high temporal sampling of these data. Subsampling could potentially lead to a difference in the (effective) diffusion coefficient between periods of confinement and free diffusion, but the short half-life of periods of free diffusion would mean that many switching events would be lost. This implies that appropriate models need to be used for analysis of high resolution data sets, models that correctly reflect the particle's behaviour. To this end, we also developed a HPW model incorporating measurement error. We did not use this model for the full analysis because of the unnecessary (we obtained almost identical results on both the 20nm glass and 40nm mica datasets when subsampling and discounting measurement error) and higher model complexity the increased computational time. These HMM methods of analysis of SPT trajectories have advantages over other methods for detecting confinement in single trajectories [80, 81, 82, 75], since they do not rely on tuning algorithm parameters by comparison with Brownian motion. Additional parameters (such as confinement strength κ and centre C) can also be extracted which allow comparison of confinement events across and within trajectories.

We demonstrated that confinement events have more similar physical characteristics (confinement time, confinement radius, radial skewness - defined in Table 3.1) within a trajectory, than between trajectories. The mean confinement radius and confinement time are the statistics most strongly conserved within trajectories (Fig. 3.11, Fig. B.4). However, it is clear that $\bar{R}_{lm}$ does not correlate with confinement event lifetime (Table 3.2, Fig. 3.10E, Fig. 3.13A). On the other hand, there is a weak dependence between radial skewness and confinement lifetime (Fig. 3.10F, Fig. 3.14); as skewness increases, the lifetime typically decreases. This indicates that the tag-target complex itself determines the characteristics of the binding events along these 3 predominantly independent factors. The non Gaussian nature of the confinement strength and the non exponential nature of the confinement time indicate heterogeneity in binding events, suggestive of higher order binding.

These dependencies are consistent with the AuNP/CTxB/GM1 particles comprising a mixed population that have various numbers of CTxB attachments to GM1 on

the surface, and these remain attached throughout the trajectory thereby giving confinement events similar characteristics (Fig. 3.16). The CTxB/GM1 bond has a high affinity - the disassociation rate in SLBs was measured as $(2.8 \pm 0.1) \times 10^{-4}$ s^{-1}, giving a mean binding lifetime of 3.6×10^4 s [116]. This is consistent with these surface attachments being present throughout the course of the trajectory. Confinement events occur when a GM1 (in the upper leaflet) attached to CTxB interacts with an immobilised GM1 in the lower leaflet, [51]. We propose that variation in the observed confinement radius is caused by differences in the geometry of bound CTxB/GM1 complexes; tightly packed (or single) CTxBs having more freedom to "wobble" (Fig. 3.16B) and broadly spaced, multiple CTxBs having less freedom (Fig. 3.16C). Non-Gaussian confinement events occur when a second (or more) CTxB/GM1 attachment is not fully immobilised allowing rotation around the immobilised binding site (Fig. 3.16D). The distance between these attachments would determine the deviation from a Gaussian and the radius of confinement, Fig. 3.10C,D. CTxB attachments have to be a sufficient distance apart for non-Gaussian profiles to be detected, whilst the CTxB molecules have to be close enough for both to bind to surface GM1 lipids. These hypotheses are consistent with the fact that there are around 25 CTxB per 20nm AuNP [51] - it would be expected that there is heterogeneity in both their number and spatial distribution. By varying the number of bound CTxB and their spatial configuration, both confinement radius and radial skewness can be altered.

We also demonstrated that confinement event lifetimes are not exponentially distributed, whilst also being conserved within trajectories. However, the fact that the mean confinement radius doesn't correlate with confinement time suggests that the area of the GM1 platform generated by the CTxB is only weakly dependent on the geometry of the bindings. The confinement time is expected to be determined by the degree of overlap of the GM1s associated with the AuNP (its GM1 platform), and the number of locally immobilised GM1 in the lower leaflet. The GM1 in the lower leaflet are immobilised by hydroxyl pinning sites on the glass surface [51], large sites aggregating a higher number of GM1 in the lower leaflet, and hence trapping the AuNP/CTxB/GM1 particle for longer. These surface hydroxyl groups have an estimated size of <10nm [51], although the aggregation of GM1 in supported lipid bilayers has been observed in AFM experiments, with domain sizes of approximately 15-60nm [117]. Larger pinning sites may trap multiple CTxB molecules on the AuNP, leading to more Gaussian behaviour as rotational degrees of freedom are lost. This would then lead to the negative correlation of non-Gaussian events and event lifetime; i.e. the lower the S_{lm} statistic (more Gaussian the confinement), the longer the typical confinement time, Fig. 3.10F. In this model,

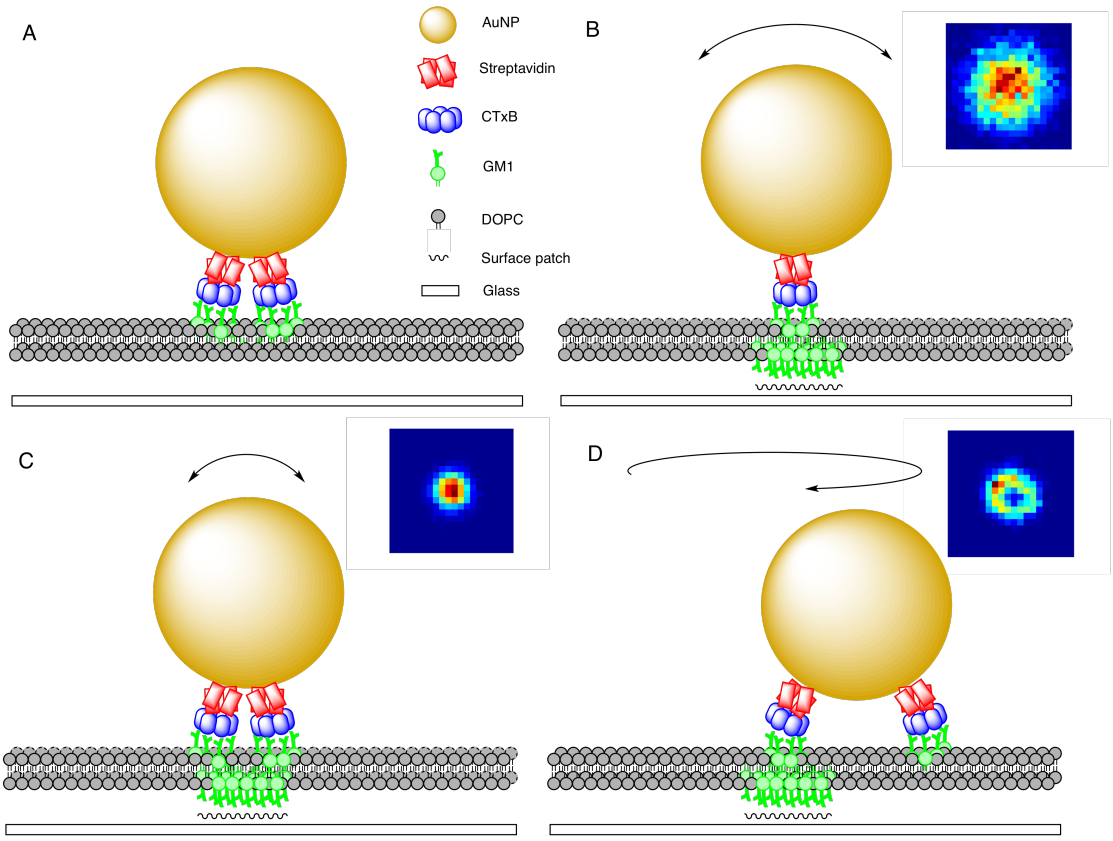

Figure 3.16: **Schematic of AuNP/CTxB/GM1 structures leading to Gaussian and non-Gaussian confinement profiles.** (A) Free diffusion, (B) wide Gaussian-like confinement, (C) narrow Gaussian-like confinement, (D) non-Gaussian confinement. Insets in (B-D) are example histograms of particle positions pooled over trajectories (e.g. Fig. 3.12, Fig. B.4, Fig. B.5). Insets have side length 0.1 µm. Schematic based on a figure in reference [51].

heterogeneity in the lifetime and shape are determined by different characteristics of the CTxB interaction with the surface, hence explaining their weak correlation.

The fact that we have demonstrated that a particular nanoparticle has a specific signature means that distinguishing the effects of this tag from other factors (such as environmental factors in cell membranes) is more difficult because of tag signature variability. Thus, using tags that have a common signature would improve extraction of environment signatures. Since the variability in the tag signature arises from the random placing of CTxB molecules on the AuNP surface, using particles with a structured surface would eliminate this problem. Virions are ideal, given their highly geometric 3D structure. Interferometric label-free tracking of virions has been demonstrated at 3s temporal resolution [118]; thus achieving the high spatial and temporal resolution of recent iSCAT microscopy (such as in the data set explored here) utilising viral particle tags is a distinct possibility. This could lead to improved deconvolution of the tagged particle signature from the environment dictated movement signal, and thus higher discrimination of different types of particle movement.

3.4 Methods

3.4.1 MCMC algorithm for harmonic potential confinement HMM

We derive an MCMC algorithm for inference of the harmonic potential well (HPW) model, as described in Results. For a 2D trajectory $\mathbf{X} = \{X_i, t_i\}_{i=1}^{N+1}$ the probability of observing $\mathbf{X}$ given the parameters (θ) and hidden states ($\mathbf{z}$, $\mathbf{C}$) is (recall that $\Delta X_i = X_{i+1} - X_i, \Delta t_i = t_{i+1} - t_i$)

$$\pi(\mathbf{X}|\theta, \mathbf{z}, \mathbf{C}) = \prod_{i=1}^{N} N(\Delta X_i; -\kappa \Delta t_i z_i (X_i - C_i), (2D\Delta t_i)^{-1}) \tag{3.5}$$

where we have used the Euler-Maruyama approximation to the SDEs (Equations (3.3) and (3.4)), although we could also solve the SDEs analytically. $N(x; \mu, \tau)$ is the normal PDF with mean μ and precision τ, we use this parameterisation and notation throughout.

Using Bayes' rule we can write the posterior ($\Delta C_i = C_{i+1} - C_i$)

$$\begin{aligned}\pi(\theta, \mathbf{z}, \mathbf{C}|\mathbf{X}) \propto \pi(\theta, z_1, C_1) \prod_{i=1}^{N-1} &\text{Bernoulli}(z_{i+1}; z_i(1-p_{esc}) + (1-z_i)p_{trap}) \\ &\times \prod_{i=1}^{N} N\left(\Delta X_i; -\kappa \Delta t_i z_i (X_i - C_i), (2D\Delta t_i)^{-1}\right) \\ &\times \prod_{i=1}^{N-1} N\left(\Delta C_i; 0, \left(2\Delta t_i (D_C z_i + \hat{D}(1-z_i))\right)^{-1}\right). \end{aligned} \tag{3.6}$$

We use conjugate priors for all parameters, specifically

$$\begin{aligned}\pi(\theta, z_1, C_1) = &\text{Unif}(D; D_{min}, D_{max})\text{Unif}(D_C; D_{C_{min}}, D_{max}/D_{ratio})\text{Unif}(\kappa; \kappa_{min}, \kappa_{max}) \\ &\times \text{Beta}(p_{esc}; a_{esc}, b_{esc})\text{Beta}(p_{trap}; a_{trap}, b_{trap}) \\ &\times \text{Bernoulli}(z_1; \pi_{trap}) N(C_1; \mu_{C_1}, \tau_{C_1}). \end{aligned} \tag{3.7}$$

We choose prior parameters which enforce model conditions, for example a large D_{ratio} implies $D >> D_C$ a priori. A conjugate prior allows us to construct Gibbs moves to sample values from the parameter posterior distributions. The updates for D and D_C are inverse Gamma

$$D^{-1}|_{D_C, \kappa, p_{esc}, p_{trap}, \mathbf{X}, \mathbf{C}, \mathbf{z}} \sim \text{Gamma}_T\left(N - 1, \frac{1}{4}\sum_{i=1}^{N} \frac{(\Delta X_i + \kappa \Delta t_i z_i (X_i - C_i))^2}{\Delta t_i}, D_{min}, D_{max}\right) \tag{3.8}$$

$$D_C^{-1}|_{D, \kappa, p_{esc}, p_{trap}, \mathbf{X}, \mathbf{C}, \mathbf{z}} \sim \text{Gamma}_T\left(-1 + \sum_{i|z_i=1} 1, \frac{1}{4}\sum_{i|z_i=1} \frac{\Delta C_i^2}{\Delta t_i}, D_{C_{min}}, D_{max}/D_{ratio}\right) \tag{3.9}$$

where $\text{Gamma}_T(\alpha, \beta, x_{min}, x_{max})$ denotes a truncated Gamma distribution with parameters α and β, truncated at x_{min} and x_{max}. We enforce the truncation by rejecting any moves which lie outside this region (Algorithm 2). If $\sum_{i|z_i=1} 1 = 0$ then the conditional for D_C reduces to the prior, so we update by sampling from $\text{Unif}(D_{C_{min}}, D_{max}/D_{ratio})$.

To calculate an update for κ we rearrange the conditional distribution into a product of Gaussian distributions in terms of κ

$$\pi(\kappa|D, D_c, p_{esc}, p_{trap}, \mathbf{X}, \mathbf{C}, \mathbf{z}) \propto \prod_{i|z_i=1} N\left(\kappa; \frac{-\Delta X_i}{\Delta t_i z_i (X_i - C_i)}, \frac{\Delta t_i z_i^2 (X_i - C_i)^2}{2D}\right).$$

This gives the update

$$\kappa|_{D,D_C,p_{esc},p_{trap},\mathbf{X},\mathbf{z},\mathbf{C}} \sim N\left(\left[-\frac{1}{2D}\sum_{i|z_i=1}\Delta X_i z_i (X_i - C_i)\right]\tau_\kappa^{-1}, \tau_\kappa\right) \tag{3.10}$$

where

$$\tau_\kappa = \frac{1}{2D}\sum_{i|z_i=1}\Delta t_i z_i^2 (X_i - C_i)^2. \tag{3.11}$$

We enforce the prior, $\text{Unif}(\kappa;\kappa_{min},\kappa_{max})$, by rejecting any values of κ lying outside this interval. If $\sum_{i|z_i=1} 1 = 0$ then $\tau_\kappa = 0$, so we update by sampling $\kappa \sim \text{Unif}(\kappa_{min},\kappa_{max})$.

The updates for the transition probabilities are

$$p_{esc}|_{D,D_C,\kappa,p_{trap},\mathbf{X},\mathbf{z},\mathbf{C}} \sim \text{Beta}\left(a_{esc}+n_{10}, b_{esc}+n_{11}\right) \tag{3.12}$$

$$p_{trap}|_{D,D_C,\kappa,p_{esc},\mathbf{X},\mathbf{z},\mathbf{C}} \sim \text{Beta}\left(a_{trap}+n_{01}, b_{trap}+n_{00}\right) \tag{3.13}$$

where n_{ml} is the number of transitions from state m to state l, i.e.

$$n_{ml} = \sum_{i|z_{i-1}=m,z_i=l} 1. \tag{3.14}$$

Gibbs update for harmonic well centre. For the centre, $\mathbf{C}$, we update using a blocked Gibbs move. We report the update here without calculation, since the full derivation (given in Appendix B.3) is somewhat lengthy. Let $\mathbf{C}_{j,n}$ be a block of length n starting at j, i.e. $\mathbf{C}_{j,n} = \{C_i\}_{i=j}^{j+n}$. The update is

$$\mathbf{C}_{j,n} \sim N(\boldsymbol{\mu}_{C_{j,n}}, \boldsymbol{\Sigma}_{C_{j,n}}^{-1}) \tag{3.15}$$

where the $n \times n$ precision matrix is

$$\boldsymbol{\Sigma}_{C_{j,n}}^{-1} = \begin{pmatrix} \Sigma_{j,j}^{-1} & \Sigma_{j,j+1}^{-1} & & & & & \\ \Sigma_{j+1,j}^{-1} & \Sigma_{j+1,j+1}^{-1} & \Sigma_{j+1,j+2}^{-1} & & & & \\ & \Sigma_{j+2,j+1}^{-1} & \ddots & & & & \\ & & & \ddots & & & \\ & & & & \ddots & \Sigma_{j+n-1,j+n}^{-1} & \\ & & & & \Sigma_{j+n,j+n-1}^{-1} & \Sigma_{j+n,j+n}^{-1} & \end{pmatrix} \tag{3.16}$$

with elements

$$\Sigma^{-1}_{i,i-1} = \Sigma^{-1}_{i-1,i} = -\frac{1}{2\Delta t_i(\hat{D}(1-z_i) + D_C z_i)} \tag{3.17}$$

$$\Sigma^{-1}_{i,i}\big|_{2\le i\le N-1} = \frac{1}{2\Delta t_{i-1}(D_C z_{i-1} + \hat{D}(1-z_{i-1}))} + \frac{(\kappa z_i)^2 \Delta t_i}{2D} + \frac{1}{2\Delta t_i(D_C z_i + \hat{D}(1-z_i))}. \tag{3.18}$$

If the block contains the first timepoint, i.e. $j = 1$, then the first element is

$$\Sigma^{-1}_{1,1} = \tau_{C_1} + \frac{(\kappa z_1)^2 \Delta t_1}{2D} + \frac{1}{2\Delta t_1(D_C z_1 + \hat{D}(1-z_1))} \tag{3.19}$$

and if the block contains the last timepoint, i.e. $j + n = N$, then the last element is

$$\Sigma^{-1}_{N,N} = \frac{1}{2\Delta t_{N-1}(D_C z_{N-1} + \hat{D}(1-z_{N-1}))} + \frac{(\kappa z_N)^2 \Delta t_N}{2D}. \tag{3.20}$$

The mean $\boldsymbol{\mu}_{C_{j,n}}$ is given by solving

$$\boldsymbol{\Sigma}^{-1}_{C_{j,n}} \boldsymbol{\mu}_{C_{j,n}} = \mathbf{b}_{j,n} \tag{3.21}$$

where $\mathbf{b}_{j,n} = \{b_i\}_{i=j}^{j+n}$ is an n-vector with elements

$$b_i\Big|_{j+1\le i\le j+n-1} = \frac{\kappa z_i}{2D}(\Delta X_i + \kappa \Delta t_i z_i X_i) \tag{3.22}$$

$$b_j\Big|_{j\ne 1} = \frac{C_{j-1}}{2\Delta t_{j-1}(D_C z_{j-1} + \hat{D}(1-z_{j-1}))} + \frac{\kappa z_j}{2D}(\Delta X_j + \kappa \Delta t_j z_j X_j) \tag{3.23}$$

$$b_{j+n}\Big|_{j+n\ne N} = \frac{C_{j+n+1}}{2\Delta t_{j+n}(D_C z_{j+n} + \hat{D}(1-z_{j+n}))} + \frac{\kappa z_{j+n}}{2D}(\Delta X_{j+n} + \kappa \Delta t_{j+n} z_{j+n} X_{j+n}). \tag{3.24}$$

If the block contains the first timepoint then

$$b_1 = \tau_{C_1}\mu_{C_1} + \frac{\kappa z_1}{2D}(\Delta X_1 + \kappa \Delta t_1 z_1 X_1) \tag{3.25}$$

and if the block contains the last timepoint,

$$b_N = \frac{\kappa z_N}{2D}(\Delta X_N + \kappa \Delta t_N z_N X_N). \tag{3.26}$$

Since $\mathbf{\Sigma}_{C_{j,n}}^{-1}$ is tridiagonal, Equation (3.21) can be efficiently solved for $\boldsymbol{\mu}_{C_{j,n}}$, for example using the left matrix division function in Matlab (with $\mathbf{\Sigma}_{C_{j,n}}^{-1}$ encoded as a sparse matrix). Hence the update can be sampled from Equation (3.15).

Metropolis-Hastings move for z and C. Since the hidden states in the model are highly correlated, we developed a joint update for $\mathbf{z}$ and $\mathbf{C}$, a blocked Metropolis-Hastings move. To simulate values from $\pi(\mathbf{z}, \mathbf{C}|\theta, \mathbf{X})$, we first propose new values of $\{\mathbf{z}, \mathbf{C}\}$ by moving a block of length n, let $\{\mathbf{z}_{j,n}, \mathbf{C}_{j,n}\} = \{z_i, C_i\}_{i=j}^{j+n}$ denote this block. We use a joint proposal distribution

$$q(\mathbf{z}'_{j,n}, \mathbf{C}'_{j,n}|\mathbf{z}_{j,n}) = q(\mathbf{z}'_{j,n}|\mathbf{z}_{j,n})q(\mathbf{C}'_{j,n}|\mathbf{z}'_{j,n}). \tag{3.27}$$

For $\mathbf{z}_{j,n}$ we use a proposal distribution $q(\mathbf{z}'_{j,n}|\mathbf{z}_{j,n})$ which is weighted towards confinement or free diffusion. Specifically, let 0^n and 1^n denote a sequence of 0's or 1's of length n, we propose $\mathbf{z}'_{j,n} = 0^n$ and $\mathbf{z}'_{j,n} = 1^n$ both with probability 1/3. We otherwise (i.e. also with probability 1/3) propose by simulating a Markov chain using the parameters p_{esc}, p_{trap}. The proposal density is hence

$$q(\mathbf{z}'_{j,n}|\mathbf{z}_{j,n})\Big|_{j\neq 1} =$$

$$\begin{cases} \frac{1}{3} + \frac{1}{3}\prod_{i=j}^{n+j+1} \text{Bernoulli}\left(z_i; z_{i-1}(1-p_{esc}) + (1-z_{i-1})p_{trap}\right) & \text{if } \mathbf{z}'_{j,n} = 0^n \\ \frac{1}{3} + \frac{1}{3}\prod_{i=j}^{n+j+1} \text{Bernoulli}\left(z_i; z_{i-1}(1-p_{esc}) + (1-z_{i-1})p_{trap}\right) & \text{if } \mathbf{z}'_{j,n} = 1^n \\ \frac{1}{3}\prod_{i=j}^{n+j+1} \text{Bernoulli}\left(z_i; z_{i-1}(1-p_{esc}) + (1-z_{i-1})p_{trap}\right) & \text{any other } \mathbf{z}'_{j,n}. \end{cases} \tag{3.28}$$

If $j = 1$, then the first term in the product in Equation (3.28) is undefined, so we propose z_1 from the prior distribution, Bernoulli$(z_1; p_{trap}/(p_{trap} + p_{esc}))$, giving

$$q(\mathbf{z}'_{1,n}|\mathbf{z}_{1,n}) = \begin{cases} \frac{1}{3} + \frac{1}{3}\left(1 - \frac{p_{trap}}{p_{trap}+p_{esc}}\right) & \\ \times \prod_{i=2}^{n+2} \text{Bernoulli}\left(z_i; z_{i-1}(1-p_{esc}) + (1-z_{i-1})p_{trap}\right) & \text{if } \mathbf{z}'_{1,n} = 0^n \\ \frac{1}{3} + \frac{1}{3}\frac{p_{trap}}{p_{trap}+p_{esc}} & \\ \times \prod_{i=2}^{n+2} \text{Bernoulli}\left(z_i; z_{i-1}(1-p_{esc}) + (1-z_{i-1})p_{trap}\right) & \text{if } \mathbf{z}'_{1,n} = 1^n \\ \frac{1}{3}\text{Bernoulli}\left(z_1; \frac{p_{trap}}{p_{trap}+p_{esc}}\right) & \\ \times \prod_{i=2}^{n+2} \text{Bernoulli}\left(z_i; z_{i-1}(1-p_{esc}) + (1-z_{i-1})p_{trap}\right) & \text{any other } \mathbf{z}'_{1,n}. \end{cases} \tag{3.29}$$

The right hand terms in the top two cases of Equations (3.28) and (3.29) are required because the sequences 0^n and 1^n are possible when simulating a Markov chain with parameters p_{esc} and p_{trap}.

We next propose a value for $\mathbf{C}_{j,n}$ using the block Gibbs update density derived earlier (Equations (3.15)-(3.26))

$$q(\mathbf{C}'_{j,n}|\mathbf{z}'_{j,n}) = N(\mathbf{C}'_{j,n}; \boldsymbol{\mu}_{C_{j,n}}, \boldsymbol{\Sigma}^{-1}_{C_{j,n}}) \tag{3.30}$$

where $\boldsymbol{\mu}_{C_{j,n}}, \boldsymbol{\Sigma}^{-1}_{C_{j,n}}$ are calculated (using Equations (3.16)-(3.21)) with the proposed value $\mathbf{z}'_{j,n}$. Up to proportionality, the density of interest is

$$\begin{aligned} \pi(\mathbf{z}, \mathbf{C}|\theta, \mathbf{X}) \propto & \prod_{i=1}^{N} N\left(\Delta X_i; -\kappa \Delta t_i z_i (X_i - C_i), (2D\Delta t_i)^{-1}\right) \\ & \times N(C_1, \mu_{C_1}, \tau_{C_1}) \prod_{i=1}^{N-1} N\left(\Delta C_i; 0, \left(2\Delta t_i \left(D_C z_i + \hat{D}(1-z_i)\right)\right)^{-1}\right) \\ & \times \text{Bernoulli}\left(z_1; \frac{p_{trap}}{p_{trap}+p_{esc}}\right) \prod_{i=1}^{N-1} \text{Bernoulli}\left(z_{i+1}; z_i(1-p_{esc}) + (1-z_i)p_{trap}\right) \end{aligned} \tag{3.31}$$

which we call $P(\mathbf{z}, \mathbf{C})$. The acceptance probability is thus

$$\alpha(\mathbf{z}', \mathbf{C}'|\mathbf{z}, \mathbf{C}) = \min\left\{1, \frac{P(\mathbf{z}', \mathbf{C}')q(\mathbf{z}_{j,n}|\mathbf{z}'_{j,n})q(\mathbf{C}_{j,n}|\mathbf{z}_{j,n})}{P(\mathbf{z}, \mathbf{C})q(\mathbf{z}'_{j,n}|\mathbf{z}_{j,n})q(\mathbf{C}'_{j,n}|\mathbf{z}'_{j,n})}\right\}. \tag{3.32}$$

By sequentially calculating these MCMC updates, we can hence sample from the posterior distribution $\pi(\theta, \mathbf{z}, \mathbf{C}|\mathbf{X})$. Algorithm 2 details this HPW model MCMC algorithm in pseudocode.

Blocking choice. The two blocked moves require a choice of block size n and starting point j. For the blocked Gibbs move for $\mathbf{C}$ we found that updating the whole Markov chain at once, i.e. $j = 1$ and $n = N - 1$ was most efficient. For the blocked Metropolis-Hastings move for $\{\mathbf{z}, \mathbf{C}\}$ we sample the block size n from $\text{Unif}(B_{min}, B_{max})$, and then sample j from $\text{Unif}(1, N - n)$. For all MCMC runs we set $B_{min} = 3$ and $B_{max} = 200$.

Initial values and priors. For all MCMC runs on simulated and experimental data the priors were as follows

$$\begin{aligned}
\pi(D) &= \text{Unif}(D; 0, 2\,\mu\text{m}^2\,\text{s}^{-1}) \\
\pi(D_C) &= \text{Unif}(D_C; 1 \times 10^{-10}\,\mu\text{m}^2\,\text{s}^{-1}, 0.04\,\mu\text{m}^2\,\text{s}^{-1}) \\
\pi(\kappa) &= \text{Unif}(\kappa; 100\,\text{s}^{-1}, 10\,000\,\text{s}^{-1}) \\
\pi(p_{esc}) &= \text{Beta}(p_{esc}; 1, 1000) \\
\pi(p_{trap}) &= \text{Beta}(p_{trap}; 1, 1000).
\end{aligned} \tag{3.33}$$

We chose non-zero $D_{C_{min}}$ because very low D_C values occasionally caused computational overflow in the blocked update covariance matrix $\mathbf{\Sigma}_{C_{k,n}}$. We chose non-zero κ_{min} because MCMC chains were occasionally very slow to converge from very low values of κ. Inferred posterior distributions for both these parameters were much higher than these minimum values. We initialised D, D_C, κ, p_{esc} and p_{trap} by sampling from the prior distributions. For $\hat{D}$ (the diffusion coefficient of C when X is confined) we used the Brownian motion maximum likelihood estimate

$$\hat{D} = \frac{1}{4N}\sum_{i=1}^{N}\frac{\Delta X_i^2}{\Delta t_i}. \tag{3.34}$$

For $\mathbf{z}$ we initialised by simulating a Markov chain using the initial p_{esc} and p_{trap}. For $\mathbf{C}$ we initialised by a Gibbs block update given $\mathbf{z}$.

Convergence diagnostics. For the parameters, we ran multiple chains and assessed convergence using the Gelman point scale reduction factor (PSRF) [92], we considered an MCMC run converged if the PSRF was less than 1.2 for all parameters. We initialised the parameter chains by sampling from the prior distributions. Under the given prior parameters the initial values were over dispersed with respect to the target distributions on all trajectories (a requirement of the PSRF method).

Algorithm 2 Metropolis-within-Gibbs sampler for harmonic potential well model

$K \leftarrow$ number of MCMC steps
$D_{min}, D_{C_{min}}, D_{max}, D_{ratio}, \kappa_{min}, \kappa_{max}, a_{esc}, b_{esc}, a_{trap}, b_{trap} \leftarrow$ prior parameters
$B_{max}, B_{min} \leftarrow$ maximum and minimum Metropolis-Hastings block sizes
$\hat{D} \leftarrow$ estimated value for D (Equation (3.34))
Choose initial values of $D^{(0)}, D_C^{(0)}, \kappa^{(0)}, p_{esc}^{(0)}, p_{trap}^{(0)}$ (e.g. sample from Equation (3.7))
Calculate $\boldsymbol{\mu}_{C_{1,N-1}}$ (mean vector) and $\boldsymbol{\Sigma}_{C_{1,N-1}}^{-1}$ (precision matrix) using Equations (3.16)-(3.21)
$\mathbf{C}^{(0)} \leftarrow$ random number drawn from $N(\boldsymbol{\mu}_{C_{1,N-1}}, \boldsymbol{\Sigma}_{C_{1,N-1}}^{-1})$
for $k = 1$ to $k = K$ **do**
  $D^{(k)} \leftarrow$ random number drawn from $\Gamma\left(N-1, \frac{1}{4}\sum_{i=1}^{N} \frac{(\Delta X_i + \kappa \Delta t_i z_i (X_i - C_i))^2}{\Delta t_i}\right)$
  if $D < D_{min}$ **or** $D > D_{max}$ **then**
    $D^{(k)} \leftarrow D^{(k-1)}$
  end if
  if $\sum_{i|z_i=1} 1 = 0$ **then**
    $D_C^{(k)} \leftarrow$ random number drawn from Unif$(D_{C_{min}}, D_{max}/D_{ratio})$
  else
    $D_C^{(k)} \leftarrow$ random number drawn from $\Gamma\left(-1 + \sum_{i|z_i=1} 1, \frac{1}{4}\sum_{i|z_i=1} \frac{\Delta C_i^2}{\Delta t_i}\right)$
  end if
  if $D_C < D_{C_{min}}$ **or** $D_C > D_{max}/D_{ratio}$ **then**
    $D_C^{(k)} \leftarrow D_C^{(k-1)}$
  end if
  Calculate τ_κ using Equation (3.11)
  if $\sum_{i|z_i=1} 1 = 0$ **then**
    $\kappa^{(k)} \leftarrow$ random number drawn from Unif$(\kappa_{min}, \kappa_{max})$
  else
    $\kappa^{(k)} \leftarrow$ random number drawn from $N\left(\left[-\frac{1}{2D}\sum_{i|z_i=1} \Delta X_i z_i (X_i - C_i)\right] \tau_\kappa^{-1}, \tau_\kappa\right)$
  end if
  if $\kappa < \kappa_{min}$ **or** $\kappa > \kappa_{max}$ **then**
    $\kappa^{(k)} \leftarrow \kappa^{(k-1)}$
  end if
  Calculate $n_{00}, n_{01}, n_{10}, n_{11}$ using Equation (3.14)
  $p_{esc}^{(k)} \leftarrow$ random number drawn from Beta$\left(a_{esc} + n_{10}, b_{esc} + n_{11}\right)$
  $p_{trap}^{(k)} \leftarrow$ random number drawn from Beta$\left(a_{trap} + n_{01}, b_{trap} + n_{00}\right)$
  Calculate $\boldsymbol{\mu}_{C_{1,N-1}}$ (mean vector) and $\boldsymbol{\Sigma}_{C_{1,N-1}}^{-1}$ (precision matrix) using Equations (3.16)-(3.21)
  $\mathbf{C} \leftarrow$ random number drawn from $N(\boldsymbol{\mu}_{C_{1,N-1}}, \boldsymbol{\Sigma}_{C_{1,N-1}}^{-1})$
  $n \leftarrow$ random integer drawn from $\{B_{min}, B_{min}+1, ..., B_{max}\}$
  $j \leftarrow$ random integer drawn from $\{1, ..., N-n\}$
  $\mathbf{z}' \leftarrow$ proposed value drawn from $q(\mathbf{z})$ (Equation (3.28))
  Calculate $\boldsymbol{\mu}_{C_{j,n}}$ (mean vector) and $\boldsymbol{\Sigma}_{C_{j,n}}^{-1}$ (precision matrix) using Equations (3.16)-(3.21)
  $\mathbf{C}'_{j,n} \leftarrow$ random number drawn from $N(\boldsymbol{\mu}_{C_{j,n}}, \boldsymbol{\Sigma}_{C_{j,n}}^{-1})$
  calculate $\alpha(\mathbf{z}', \mathbf{C}' | \mathbf{z}, \mathbf{C})$ using Equation (3.32)
  $u \leftarrow$ random number drawn from Unif$(0, 1)$
  if $u < \alpha(\mathbf{z}', \mathbf{C}' | \mathbf{z}, \mathbf{C})$ **then**
    $\mathbf{z} \leftarrow \mathbf{z}'$
    $\mathbf{C} \leftarrow \mathbf{C}'$
  end if
  $\mathbf{z}^{(k)} \leftarrow \mathbf{z}$
  $\mathbf{C}^{(k)} \leftarrow \mathbf{C}$
end for

Chapter 4

Hop Diffusion Model

4.1 Introduction

Multiple single particle tracking (SPT) studies have demonstrated that some membrane molecules exhibit so called "hop" diffusion [71, 24, 40, 39, 42, 43], characterised by short term confinement in domains, and occasional "hopping" between adjacent compartments. An explanation for hop diffusion, which has been observed for both lipids and proteins, is the fences and pickets model (Fig. 1.2).

Since individual SPT trajectories are stochastic, statistical methods are required to robustly detect hop diffusion, particularly to differentiate true inter-compartmental hopping from random fluctuations. Some methods fit experimental mean squared displacement (MSD) curves to theoretical MSD functions (as described in Section 1.4.1) [71, 72, 42]; hence deriving parameters such as the domain size and hopping frequency. Powles *et al.* derived the exact solution for particles diffusing in a meshwork of squares [101]. There are also methods which pinpoint hopping events in single trajectories. For example, data from the Kusumi lab has been segmented into confinement zones using a simple method which detects sharp increases in the apparent diffusion coefficient [24, 39]. Meilhac *et al.* developed a method for detecting hops between confinement zones [82], which detected two-thirds of jumps on example simulated trajectories. The missed jumps were double (or more) events which could only be resolved as a single jump.

4.2 Simulation

We followed existing methods [119, 120] for simulating hop diffusion by considering a particle undergoing Brownian motion in a series of square domains, of side length L.

When the particle encounters a domain barrier it crosses with some probability p_{hop}. We denote the 2D position of the particle at time t_i as X_i. At each timepoint X_i is in a square box B_i, which has side length L and (2D) centre C_{B_i},

$$B_i = \left\{ x : C_{B_i} - \begin{pmatrix} L/2 \\ L/2 \end{pmatrix} < x < C_{B_i} + \begin{pmatrix} L/2 \\ L/2 \end{pmatrix} \right\} \tag{4.1}$$

where x is also 2D. We simulate the particle position at time t_{i+1} using the Euler approximation (which is actually exact for constant D), giving $X_{i+1} \sim N(X_i, 2D\Delta t_i)$. If $X_{i+1} \notin B_i$, then it crosses the barrier with probability p_{hop}, i.e. if $u < p_{hop}$ for $u \sim \text{Unif}(0, 1)$. If $u > p_{hop}$ then we resimulate X_{i+1} until $X_{i+1} \in B_i$. An example of a simulated trajectory is shown in Fig. 4.1, and the simulation algorithm is given in pseudocode in Algorithm 3.

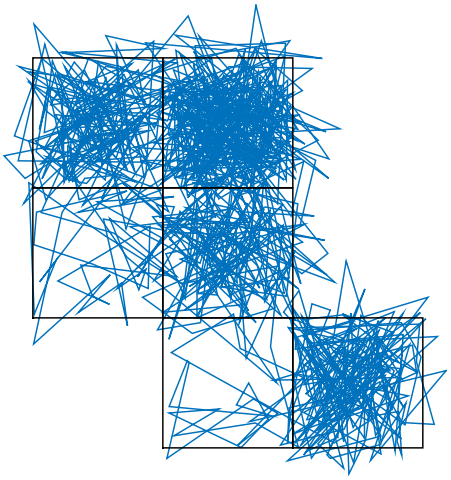

Figure 4.1: **Simulation of a molecule diffusing in a series of partially permeable square domains.** Simulation parameters: $D = 1 \times 10^5\,\text{nm}^2\,\text{s}^{-1}$, $L = 20\,\text{nm}$ and $p_{hop} = 0.005$. The time step was $0.001\,\text{s}$, and a Gaussian measurement noise with variance $\sigma^2 = 10\,\text{nm}^2$ was added.

4.3 Inference

In order to fit a model to hop diffusion data, we propose a simplified model and an MCMC algorithm for its inference. Given a single trajectory, the exact geometry of domains is often difficult to determine, so we model the confinement as a harmonic potential well

Algorithm 3 Simulation algorithm for diffusion in a meshwork of partially permeable squares.

$\{D, p_{hop}, L\} \leftarrow$ choice of model parameters
$\{\Delta t_i\}_{i=1}^{N+1} \leftarrow$ choice of time steps ($\Delta t_i = t_{i+1} - t_i$)
$X_1 \leftarrow$ initial particle position
$C_1 \leftarrow$ initial box centre position
$B_1 = \left\{ x : C_{B_1} - \begin{pmatrix} L/2 \\ L/2 \end{pmatrix} < x < C_{B_1} + \begin{pmatrix} L/2 \\ L/2 \end{pmatrix} \right\} \leftarrow$ initial confinement box
for $i = 1$ to $i = N - 1$ **do**
    $X_{i+1} \leftarrow$ random number drawn from $N\left(X_i, (2D\Delta t_i)^{-1}\right)$
    if $X_{i+1} \notin B_i$ **then**
        $u \leftarrow$ uniform random number
        if $u < p_{hop}$ **then**
            $C_{i+1} \leftarrow$ new box centre (position depends on direction of exit)
            $B_{i+1} = \left\{ x : C_{B_{i+1}} - \begin{pmatrix} L/2 \\ L/2 \end{pmatrix} < x < C_{B_{i+1}} + \begin{pmatrix} L/2 \\ L/2 \end{pmatrix} \right\} \leftarrow$ new confinement box
        else
            while $X_{i+1} \notin B_{i+1}$ **do**
                $X_{i+1} \leftarrow$ random number drawn from $N\left(X_i, (2D\Delta t_i)^{-1}\right)$
            end while
            $B_{i+1} \leftarrow B_i$
            $C_{i+1} \leftarrow C_i$
        end if
    end if
    $z_{i+1} \leftarrow$ random number drawn from Bernoulli$(z_i(1 - p_{hop})) + (1 - z_i)p_{trap})$
end for

(HPW). This effectively blurs the effect of the barriers. We propose a model with M harmonic potential wells, each well having a unique centre C and strength κ. At any one time, the particle is trapped in one of these M wells. The hidden state z_i denotes the current well at time i. The associated (2D) centre and strength for a well are $C_{z_i} = (C_{1z_i}, C_{2z_i})$ and κ_{z_i} respectively. We simplify by fixing the potential for all wells as κ. Thus, if there are M wells then the parameters are $\theta_M = \{D, C_1, ..., C_M, \kappa, P\}$. P is an M by M transition matrix, with P_{m_1,m_2} giving the probability of hopping from well m_1 to well m_2. We simplify by setting

$$P_{m_1,m_2} = \begin{cases} 1 - p_{hop} & \text{if } m_1 = m_2 \\ p_{hop} & \text{if } m_1 \neq m_2 \end{cases} \tag{4.2}$$

so that the probability of hopping between any two wells is equal. The likelihood of a trajectory, $\mathbf{X} = \{X_i\}_{i=1}^N$ where $X_i = (X_{1i}, X_{2i})$ is 2D, conditioned on $\mathbf{z} = \{z_i\}_{i=1}^N$ is

$$\pi(\mathbf{X}|\theta, \mathbf{z}) = \prod_{i=1}^{N} N(\Delta X_i; -\kappa \Delta t_i (X_i - C_{z_i}), 2D\Delta t_i). \tag{4.3}$$

The posterior distribution is hence

$$\pi(\theta, \mathbf{z}|\mathbf{X}) \propto \pi(\theta_M)\pi(z_1|\theta_M) \prod_{i=2}^{N} P_{z_{i-1},z_i} \prod_{i=1}^{N} N(\Delta X_i; -\kappa \Delta t_i (X_i - C_{z_i}), 2D\Delta t_i). \tag{4.4}$$

We now derive an MCMC algorithm for the inference of this HPW hop model.

4.3.1 Parameter updates

We use conjugate priors for the parameters, and hence Gibbs moves for their inference. Specifically, the prior distribution is

$$\pi(\theta_M) = \text{Unif}(D, D_{min}, D_{max})\text{Unif}(\kappa, \kappa_{min}, \kappa_{max})\text{Beta}(p_{hop}, a_{hop}, b_{hop}) \prod_{m=1}^{M} N(C_m; \mu_C, \tau_C). \tag{4.5}$$

The update for D is hence

$$D^{-1} \sim \text{Gamma}\left(N - 1, \frac{1}{4}\sum_{i=1}^{N} \frac{(\Delta X_i + \kappa \Delta t_i (X_i - C_{z_i}))^2}{\Delta t_i}\right) \tag{4.6}$$

with any D values outside the interval $[D_{min}, D_{max}]$ automatically rejected. For κ we have

$$\kappa \sim N(\mu_\kappa, \tau_\kappa) \tag{4.7}$$

where

$$\mu_\kappa = -\frac{1}{\tau_\kappa}\frac{1}{2D}\sum_{i=1}^{N}\Delta X_i(X_i - C_{z_i}),\ \tau_\kappa = \frac{1}{2D}\sum_{i=1}^{N}\Delta t_i(X_i - C_{z_i})^2 \tag{4.8}$$

again, any κ values outside $[\kappa_{min}, \kappa_{max}]$ are rejected. The update for p_{hop} is

$$p_{hop} \sim \text{Beta}\left(a_{hop} + n_{hop}, b_{hop} + (N - n_{hop})\right) \tag{4.9}$$

where n_{hop} is the number of hopping events, i.e. $n_{hop} = \sum_{i|z_{i-1}\neq z_i} 1$. To derive updates for the well centres, $\mathbf{C} = \{C_1, C_2, ..., C_m\}$, we note that the conditional distribution is a product of Gaussians

$$\pi(C_m|\mathbf{C}_{-m}, D, \kappa, P, \mathbf{X}) \propto \pi(C_m)\prod_{i|z_i=j} N\left(\Delta X_i; -\kappa\Delta t_i(X_i - C_{z_i}), 2D\Delta t_i\right) \tag{4.10}$$

$$= \pi(C_m)\prod_{i|z_i=j} N\left(C_{z_i}; \frac{\Delta X_i + \kappa\Delta t_i X_i}{\kappa\Delta t_i}, \frac{\kappa^2\Delta t_i}{2D}\right) \tag{4.11}$$

where $\mathbf{C}_{-m} = \mathbf{C} \setminus C_m$. This gives the update

$$C_m \sim N(\mu_{C_m}, \tau_{C_m}) \tag{4.12}$$

where

$$\mu_{C_m} = \left(\mu_C\tau_C + \sum_{i|z_i=m}\frac{\kappa}{2D}(\Delta X_i + \kappa\Delta t_i X_i)\right)\tau_{C_m}^{-1} \tag{4.13}$$

$$\tau_{C_m} = \tau_C + \sum_{i|z_i=j}\frac{\kappa^2\Delta t_i}{2D}. \tag{4.14}$$

4.3.2 Metropolis-Hastings move for shifting hop events

In order to calculate an update for $\mathbf{z}$ we choose an uninformative prior on the well at the first timepoint, z_1

$$\pi(z_1 = m|\theta_M) = \frac{1}{M}. \tag{4.15}$$

We developed a Gibbs move for single z_i's, but since consecutive values are highly correlated this was very inefficient. We thus developed two Metropolis-Hastings moves to sample $\pi(\mathbf{z}|C_1, ..., C_M, D, \kappa, P, \mathbf{X})$.

We first consider a move which shifts a hop event. Given $\mathbf{z}$, we propose $\mathbf{z}'$ by picking a hopping event k and shift length n, where n is a random integer sampled from $\{1, ..., N_{max}\}$. We then shift k, either to $k+n$ or $k-n$, each with probability 1/2. In other words we first set $\mathbf{z}' = \mathbf{z}$, then set

$$\mathbf{z}'_{k,n} = z_{k-1} \quad \text{with probability } 1/2 \tag{4.16}$$

$$\mathbf{z}'_{k-n,n} = z_k \quad \text{with probability } 1/2 \tag{4.17}$$

where $\mathbf{z}'_{k,n} = \{z'_i\}_{i=k}^{k+n}$. The proposal density is hence

$$q(\mathbf{z}'|\mathbf{z}) = \frac{1}{2N_{max}}. \tag{4.18}$$

Fig. 4.2A is a graphical representation of this proposal.

We choose N_{max} to enforce the following conditions on the position of the shifted hopping event ($k+n$ or $k-n$):

- it remains within the interval $[1, N]$
- it does not overlap the nearest hopping event to k, which we denote k'
- the proposal density for the reverse move, $q(\mathbf{z}|\mathbf{z}')$, is non-zero.

We therefore need to ensure that the shifted event is less than halfway to 1, N, and k', so we set $N_{max} = \left\lfloor \frac{1}{2} \min\{|k-k'|, k-1, N-k\} \right\rfloor$. The acceptance probability is therefore

$$\begin{aligned} \alpha(\mathbf{z} \to \mathbf{z}') &= \min\left\{1, \frac{P(\mathbf{z}')}{P(\mathbf{z})} \frac{q(\mathbf{z}|\mathbf{z}')}{q(\mathbf{z}'|\mathbf{z})}\right\} \\ &= \min\left\{1, \frac{P(\mathbf{z}')}{P(\mathbf{z})} \frac{N_{max}}{N'_{max}}\right\} \end{aligned} \tag{4.19}$$

where $P(\mathbf{z})$ is the conditional distribution given by

$$\begin{aligned} \pi(\mathbf{z}|C_1, ..., C_M, D, \kappa, P, \mathbf{X}) &\propto \pi(z_1) \prod_{i=2}^{N} P_{z_{i-1}, z_i} \prod_{i=1}^{N} N(\Delta X_i; -\kappa \Delta t_i (X_i - C_{z_i}), 2D\Delta t_i) \\ &:= P(\mathbf{z}). \end{aligned} \tag{4.20}$$

We note that this calculation is only valid if the proposed move does not remove a well, which would reduce the number of parameters in the model. However, our choice of N_{max} ensures that a well cannot be removed, and hence Equation (4.19) remains valid.

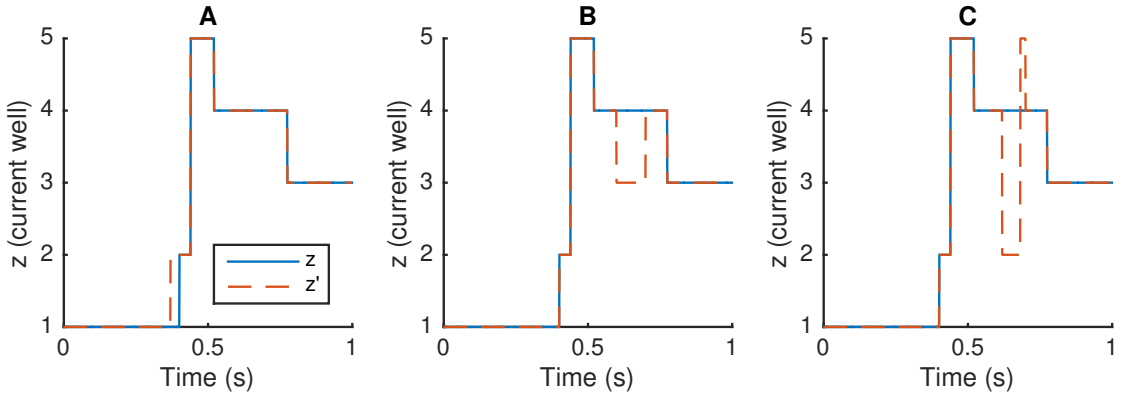

Figure 4.2: **Confinement state proposals for HPW hop MCMC algorithm.** The solid blue line is the current confinement state (z), and the dotted red line shows a possible proposed move (z'). (A) Shifted hop event (described in Section 4.3.2). (B) Moving a block to a new confinement zone (Section 4.3.3), (C) simulating a Markov chain for the new block (Section 4.3.3).

4.3.3 Metropolis-Hastings move for z, C

The hidden state $\mathbf{z}$ and well centre C are highly correlated, so we also developed a Metropolis-Hastings (MH) move to sample them concurrently. To achieve this, we first propose a new value of the confinement state $\mathbf{z}'$, then propose new centres given $\mathbf{z}'$, $\mathbf{C}' = \{C'_m\}_{m=1}^{M}$. We propose $\mathbf{z}'$ by moving a (randomly sampled) block of $\mathbf{z}$, starting at timepoint k and of length n, which we again denote $\mathbf{z}_{k,n} = \{z_i\}_{i=k}^{k+n}$. This block can contain any sequence of confinement states (i.e. all z_i within the block may be in the same well, or otherwise). Our choice of proposal distribution is ($m^{(n)}$ denote a sequence of m's of length n)

$$q(\mathbf{z}'_{k,n}|\mathbf{z}_{k,n}) = \begin{cases} \frac{1}{2M} + \frac{1}{2}P_{z'_{k-1},z'_k}(1-p_{hop})^n & \text{if } \mathbf{z}'_{k,n} = m^{(n)} \\ \frac{1}{2}\prod_{i=k}^{k+n+1} P_{z'_{i-1},z'_i} & \text{otherwise.} \end{cases} \tag{4.21}$$

In other words, we propose by either (each with probability 1/2): moving the entire block ($\mathbf{z}_{k,n}$) to a new confinement zone (Fig. 4.2B), or simulating a Markov chain for the new block ($\mathbf{z}'_{k,n}$) using the transition matrix P (Fig. 4.2C). The extra term in the $\mathbf{z}'_{k,n} = m^{(n)}$ case is required because the sequence $m^{(n)}$ is possible to simulate with a Markov chain using the matrix P.

For $\mathbf{C}$ we propose using the previously derived Gibbs density, for $m = 1...M$ we have

$$q(C_m|\mathbf{z}_{k,n}) = N(\mu_{C_m}, \tau_{C_m}) \tag{4.22}$$

where μ_{C_m}, τ_{C_m} are given in Equations (4.13) and (4.14). Up to proportionality the joint density is

$$\begin{aligned}\pi(\mathbf{z}, C_m|\mathbf{C}_{-m}, D, \kappa, P, \mathbf{X}) &\propto \pi(C_m)\pi(z_1)\prod_{i=2}^{N} P_{z_{i-1},z_i}\prod_{i=1}^{N} N(\Delta X_i; -\kappa\Delta t_i(X_i - C_{z_i}), 2D\Delta t_i) \\ &:= P(\mathbf{z}, C_m). \end{aligned} \tag{4.23}$$

Again, if the proposed move removes a well, then there is a reduction in dimension and Equation (4.23) is not valid, so if the proposed move removes a well (i.e. $m \notin \{z_i\}_{i=1}^N$ for some $m \in \{1, 2, ..., M\}$) then we set $P(\mathbf{z}', C'_m) = 0$. The acceptance probability is then

$$\alpha(\{\mathbf{z}_{k,n}, C_m\} \to \{\mathbf{z}'_{k,n}, C'_m\}) = \min\left\{1, \frac{P(\mathbf{z}'_{k,n}, C'_m)q(\mathbf{z}_{k,n})q(C_m|\mathbf{z}_{k,n})}{P(\mathbf{z}_{k,n}, C_m)q(\mathbf{z}'_{k,n})q(C'_m|\mathbf{z}'_{k,n})}.\right\} \tag{4.24}$$

The MCMC sampler for the HPW hop model is given as pseudocode in Algorithm 4.

4.3.4 Reversible jump MCMC for number of wells

All the MCMC moves so far have assumed that the number of wells, M, is known. Clearly this is not the case for most practical applications, so we now develop a move which can also sample the marginal distribution for M, $\pi(M|\mathbf{X})$. Let $\theta_M = \{D, \kappa, p_{hop}, C_1, ..., C_M\}$ be the set of parameters for the model with M wells. To switch the number of wells in the model we derive an algorithm to sample from $\pi(M, \theta_M, \mathbf{z}|\mathbf{X})$ using reversible-jump

Algorithm 4 Metropolis-within-Gibbs sampler for hop diffusion model

$K \leftarrow$ number of MCMC steps
$M \leftarrow$ number of wells
$D_{min}, D_{max}, \kappa_{min}, \kappa_{max}, a_{hop}, b_{hop}, \mu_C, \tau_C \leftarrow$ prior parameters
$D^{(0)} \leftarrow$ random number drawn from Unif(D_{min}, D_{max})
$\kappa^{(0)} \leftarrow$ random number drawn from Unif$(\kappa_{min}, \kappa_{max})$
$p_{hop}^{(0)} \leftarrow$ random number drawn from Beta(a_{hop}, b_{hop})
for $m = 1$ to $m = M$ **do**
  $C_m \leftarrow$ random number drawn from $N(\mu_C, \tau_C)$
end for
$\mathbf{C}^{(0)} \leftarrow \{C_m\}_{m=1}^{M^{(0)}}$
for $i = 1$ to $i = N$ **do**
  $z_i \leftarrow$ random integer from $\{1, ..., M\}$
end for
$\mathbf{z} \leftarrow \{z_i\}_{i=1}^{N}$
for $k = 1$ **do** to $k = K$
  $D^{(k)} \leftarrow$ random number drawn from $\Gamma\left(N+1, \frac{1}{4}\sum_{i=1}^{N} \frac{(\Delta X_i + \kappa \Delta t_i (X_i - C_{z_i}))^2}{\Delta t_i}\right)$
  if $D < D_{min}$ **or** $D > D_{max}$ **then**
    $D^{(k)} \leftarrow D^{(k-1)}$
  end if
  calculate μ_κ, τ_κ using Equation (4.8)
  $\kappa^{(k)} \leftarrow$ random number drawn from $N(\mu_\kappa, \tau_\kappa)$
  if $\kappa < \kappa_{min}$ **or** $\kappa > \kappa_{max}$ **then**
    $\kappa^{(k)} \leftarrow \kappa^{(k-1)}$
  end if
  $n_{hop} \leftarrow \sum_{i|z_{i-1} \neq z_i} 1$
  $p_{hop} \leftarrow$ random number drawn from Beta$\big(a_{hop} + n_{hop}, b_{hop} + (N - n_{hop})\big)$
  for $m = 1$ to $m = M$ **do**
    calculate μ_{C_m}, τ_{C_m} using Equations (4.13) and (4.14)
    $C_m \leftarrow$ random number drawn from $N(\mu_{C_m}, \tau_{C_m})$
  end for
  $k \leftarrow$ random hopping event drawn from $i|_{z_i \neq z_{i+1}}$
  $k' \leftarrow$ closest hopping event to k
  $N_{max} \leftarrow \min\{|k - k'|, k-1, N-k\}$
  $n \leftarrow$ random integer drawn from $\{1, ..., N_{max}\}$
  propose $\mathbf{z}'$ using Equations (4.16) and (4.17)
  calculate $\alpha(\mathbf{z} \to \mathbf{z}')$ using Equations (4.19) and(4.20)
  $u \leftarrow$ random number drawn from Unif$(0, D_{max})$
  if $u < \alpha(\mathbf{z} \to \mathbf{z}')$ **then**
    $\mathbf{z} \leftarrow \mathbf{z}'$
  end if
  propose $\mathbf{z}'_{\mathbf{k},\mathbf{n}}$ using Equation (4.21)
  for $m = 1$ to $m = M$ **do**
    propose C'_m using Equations (4.13), (4.14) and (4.22)
  end for
  calculate $\alpha(\{\mathbf{z}_{k,n}, \mathbf{C}\} \to \{\mathbf{z}'_{k,n}, \mathbf{C}'\})$ using Equation (4.24)
  $u \leftarrow$ random number drawn from Unif$(0, D_{max})$
  if $\alpha(\{\mathbf{z}_{k,n}, \mathbf{C}\} \to \{\mathbf{z}'_{k,n}, \mathbf{C}'\})$ **then**
    $\mathbf{z} \leftarrow \mathbf{z}'$
    $\mathbf{C} \leftarrow \mathbf{C}'$
  end if
end for

MCMC methodology [121]. Up to proportionality the posterior distribution is

$$\begin{aligned}\pi(M,\theta_M,\mathbf{z}|\mathbf{X}) &\propto \pi(M,\theta_M,\mathbf{z})\pi(\mathbf{X}|M,\theta_M,\mathbf{z}) \\ &= \pi(M)\pi(\theta_M|M)\pi(z_1|\theta_M,M)\prod_{i=2}^{N}P_{z_{i-1},z_i}\prod_{i=1}^{N}N(\Delta X_i;-\kappa\Delta t_i(X_i-C_{z_i}),2D\Delta t_i).\end{aligned} \tag{4.25}$$

We use a flat prior on the number of wells, $\pi(M) = \text{Unif}(M_{min}, M_{max})$, and propose well addition (move from $M \to M+1$) or well removal ($M+1 \to M$) with equal probability. Other priors are possible, such as a Poisson distribution, although there is no natural choice for the mean number of wells. For the addition of a well we propose by selecting a well m, and splitting it into two new wells (Fig. 4.3A). We denote the new wells m and $M+1$, with corresponding centres C'_m and C'_{M+1}. The proposal density for adding a well is hence

$$j(M+1,\theta_{M+1}|M,\theta_M) = \frac{1}{2}\times\frac{1}{M}. \tag{4.26}$$

In other words, the first term $(1/2)$ denotes the probability of choosing well addition, and the second term $(1/M)$ denotes the probability of choosing a specific well. For the removal of a well we pick two wells and merge them (Fig. 4.3B), for simplicity here we also denote these two wells m and $M+1$, with centres C'_m and C'_{M+1}. The proposal density for well removal is

$$j(M,\theta_M|M+1,\theta_{M+1}) = \frac{1}{2}\times\frac{1}{M(M+1)} \tag{4.27}$$

the second term denoting the probability of choosing the two wells to merge.

For addition of a well we need to define a mapping from $(\theta_M, \mathbf{z})$ to $(\theta'_{M+1}, \mathbf{z}')$. We make a simplification by assuming that there is a single hop between the proposed wells (e.g. Fig. 4.3A). Since we propose the new well centres (C'_m, C'_{M+1}) based on C_m, all other parameters and hidden states, $z_i|_{i\neq m}$, are fixed between the two models. We therefore need to construct a function which maps from $(C_m, \mathbf{z})$ to (C'_m, C'_{M+1}, H'), where H' is the timepoint of the proposed hop between wells m and $M+1$, i.e. $H' = i|_{z'_i=m,z'_{i+1}=M+1}$. Following the approach of Green [121], this mapping is a function h

$$(C'_m, C'_{M+1}, u') = h(C_m, u) \tag{4.28}$$

where $u \sim g$ is a vector of random numbers which parameterises the move between dimensions. The variables u' are required for the reverse move from M to $M+1$. We

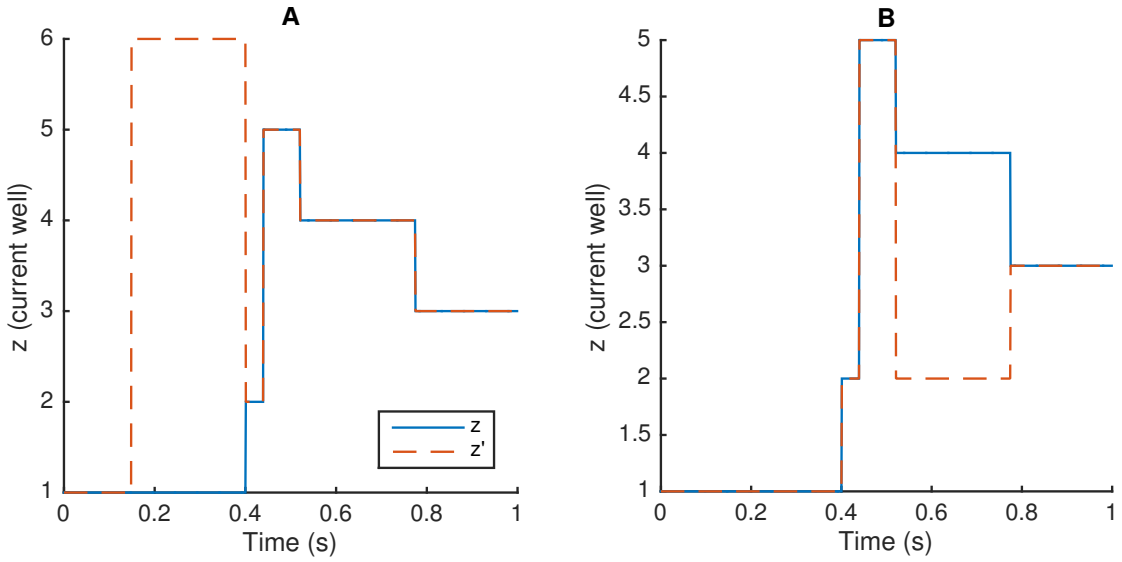

Figure 4.3: **Reversible jump move for the number of wells.** The solid blue line is the current confinement state (z), and the dotted red line shows a possible proposed move (z'), which also changes the number of wells in the model. (A) Well addition ($z = 1$ well split into two). (B) Well removal ($z = 2$ and $z = 4$ wells combined).

choose the position of the new wells by first simulating $u_1 \sim N(0, \tau_C)$, then setting $C'_m = C_m + u_1$ and $C'_{M+1} = C_m - u_1$. We choose the hop time $H' = u_2$ by simulating from the set of all timepoints where $z_i = m$, excluding the first and last timepoint. The probability of drawing u_2 is hence $\frac{1}{K-2}$ where $K = \sum_{i|z_i=m} 1$. Therefore our choice of h, mapping from (C_m, u) to (C'_m, C'_{M+1}, H'), is

$$(C'_m, C'_{M+1}, H') = h(C_m, u) = (C_m - u_1, C_m + u_1, u_2). \tag{4.29}$$

The Jacobian determinant of this transformation is

$$\left|\frac{\partial(C'_m, C'_{M+1}, H')}{\partial(C_m, u)}\right| = \begin{vmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 2. \tag{4.30}$$

And the density for $u = (u_1, u_2)$ is

$$q(u_1, u_2) = \frac{N(u_1; 0, \sigma_C^2)}{K-2}. \tag{4.31}$$

For the reverse move, we propose a move from (C'_m, C'_{M+1}, H) to (C_m, u) using h', the inverse of h

$$(C_m, u) = h'(C'_m, C'_{M+1}, H') = \left(\frac{1}{2}\left(C'_m + C'_{M+1}\right), \frac{1}{2}\left(C'_{M+1} - C'_m\right), H'\right) \tag{4.32}$$

The Jacobian determinant for the reverse transformation is

$$\left|\frac{\partial(C_m, u)}{\partial(C'_m, C'_{M+1}, H')}\right| = \begin{vmatrix} 1/2 & 1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1/2. \tag{4.33}$$

The acceptance probability for the move $M \to M+1$ is hence [121]

$$\begin{aligned} \alpha(M+1|M) &= \min\left\{1, \frac{\pi(M+1, \theta_{M+1}, \mathbf{z}'|\mathbf{X})}{\pi(M, \theta_M, \mathbf{z}|\mathbf{X})} \frac{j(M, \theta_M|M+1, \theta_{M+1})}{j(M+1, \theta_{M+1}|M, \theta_M) q(u_1, u_2)} \left|\frac{\partial h(C_m, u)}{\partial(C_m, u)}\right|\right\} \\ &= \min\left\{1, \frac{\pi(C_{M+1})\pi(\mathbf{X}|M+1, \theta_{M+1}, \mathbf{z}')}{\pi(\mathbf{X}|M, \theta_M, \mathbf{z})} \frac{K-2}{M+1} N(u_1; 0, \tau_U)^{-1} \times 2\right\}. \end{aligned} \tag{4.34}$$

And for the reverse move from $M+1 \to M$ we have

$$\alpha(M|M+1) = \min\left\{1, \frac{\pi(\mathbf{X}|M,\theta_M,\mathbf{z}')}{\pi(C_{M+1})\pi(\mathbf{X}|M+1,\theta_{M+1},\mathbf{z})}\frac{M+1}{K-2}N\left(\frac{1}{2}\left(C'_{M+1}-C'_m\right);0,\tau_U\right)\times\frac{1}{2}\right\}. \tag{4.35}$$

4.4 Application to simulated data

We tested the reversible jump MCMC algorithm for the HPW hop model on the simulated trajectory shown in Fig. 4.1. For MCMC inference we ran 12 parallel chains with priors

$$D \sim \text{Unif}(0, D_{max} = 1\times10^5\,\text{nm}^2\,\text{s}^{-1}) \tag{4.36}$$

$$\kappa \sim N(0, \kappa_{max} = 2000\,\text{s}^{-1}) \tag{4.37}$$

$$p_{hop} \sim \text{Beta}(1,1) \tag{4.38}$$

$$M \sim \text{Unif}(0,20) \tag{4.39}$$

$$C_m \sim N\left(\mu_C = \left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \tau_C = \left[\begin{smallmatrix}1\times10^{-3}\,\text{nm}^{-2} & 0\\ 0 & 1\times10^{-3}\,\text{nm}^{-2}\end{smallmatrix}\right]\right), \text{ for all } m = 1,...,M. \tag{4.40}$$

Fig. 4.4 shows the MCMC chains for the parameters and number of wells, Fig. 4.5A shows the trajectory coloured by inferred confinement well, and Fig. 4.5B plots these wells, showing accurate inference of the simulated well values. In order to estimate the domain size based on the inferred parameters, we also defined a mean confinement diameter statistic. The stationary distribution for the harmonic potential is Gaussian with PDF $\sqrt{\frac{\kappa}{2\pi D}}\exp\left(-\frac{\kappa}{2D}(X_i-C_i)^2\right)$, which has standard deviation $\sqrt{D/\kappa}$. We define the mean confinement diameter equal to $4\sqrt{D/\kappa}$, corresponding to a domain which encloses 95% of expected particle positions. For the simulations this statistic accurately estimated the size of the square domains, Fig. 4.6A. The simulation hopping probability, p_{hop}, and number of wells, M, were also well estimated, Fig. 4.6B,C.

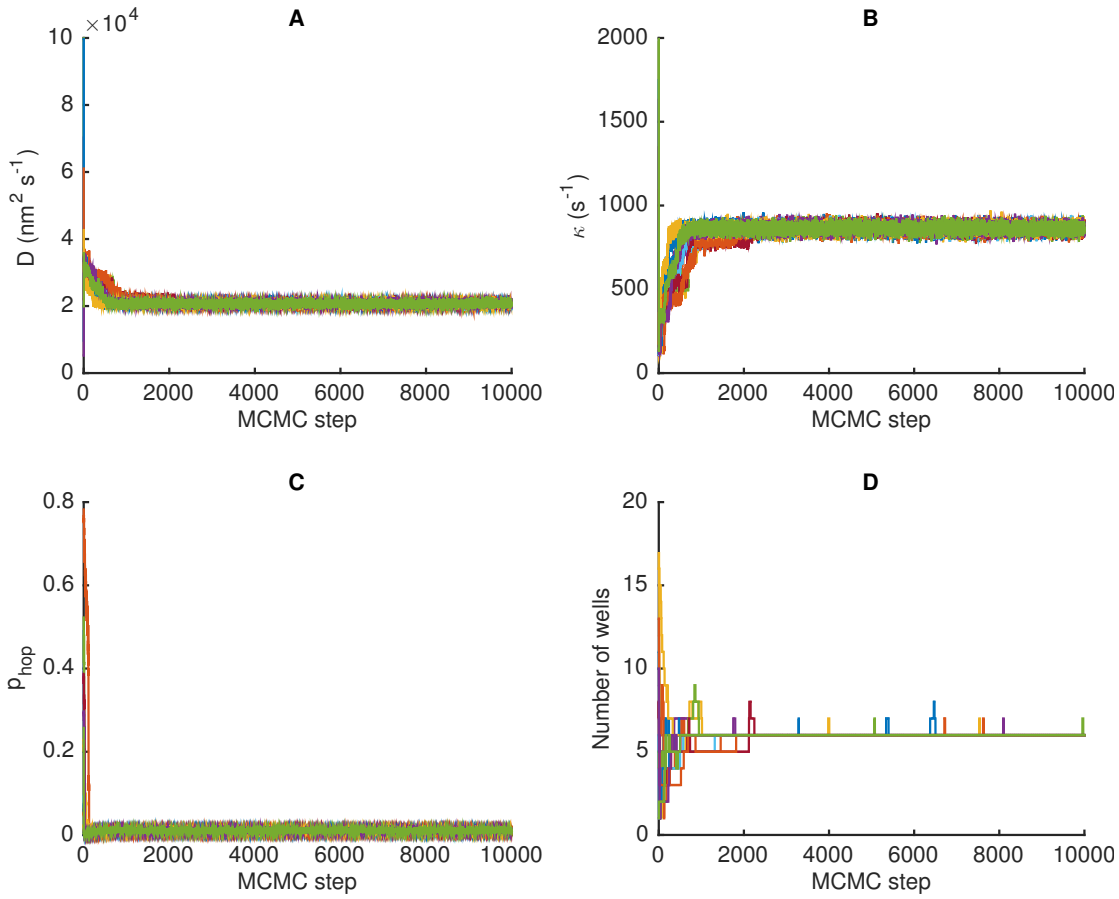

Figure 4.4: **MCMC chains from HPW hop model fit to a simulated hop diffusion trajectory.** (A) Diffusion coefficient (D), (B) harmonic well strength (κ), (C) hopping probability (p_{hop}), (D) number of wells (M). For each parameter 12 parallel MCMC chains are plotted.

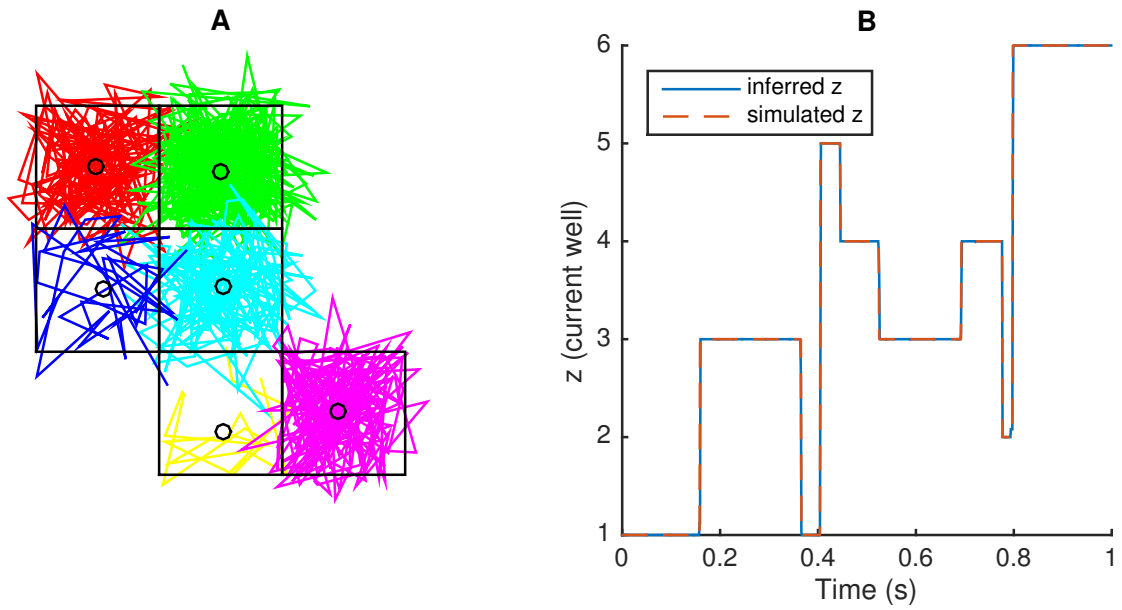

Figure 4.5: **Confinement states from HPW hop model fit to a simulated hop diffusion trajectory.** (A) Trajectory coloured by inferred confinement well, with each well plotted in a different colour. Inferred well centres are plotted as black circles. (B) Inferred (blue line) and simulated (red dashed line) confinement wells.

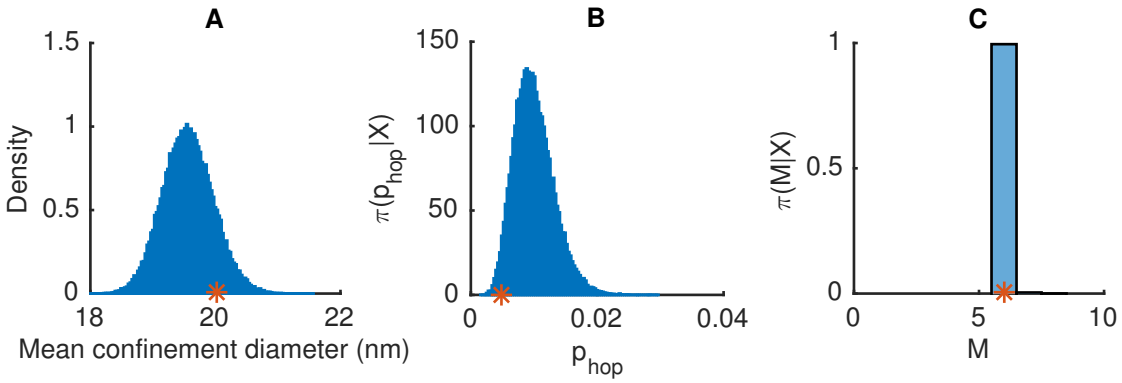

Figure 4.6: **Histograms from HPW hop model fit to a simulated hop diffusion trajectory.** (A) Mean confinement diameter, $4\sqrt{D/\kappa}$, and simulation domain side length, L (red asterisk). (B) Hopping probability per frame and simulation value (red asterisk). (C) Inferred (blue) and simulated (red asterisk) number of wells.

4.5 Conclusion

We propose a simplified hop diffusion model, where confining domains are modelled as harmonic potential wells. We developed a reversible jump MCMC algorithm for the model which (given a trajectory) infers model parameters, hidden states, and the number of confining wells. On simulated data the algorithm accurately infers the parameter values, hidden states, and number of wells. This approach can potentially extract more information from trajectories than existing methods, in part because it can resolve double events (i.e. two hops in quick succession). The method of Meilhac *et al.* could not resolve double events separated by around 100 timepoints [82], instead detecting only a single jump. In Fig. 4.5B we see that our method resolved two hopping events separated by around 20 timepoints (at around 0.8s into the trajectory). However, confirming if this algorithm is better at resolving double jumps would require a full comparison between the two methods, which is beyond the scope of this chapter.

Applying this algorithm to an experimental dataset is an obvious option for further work, since (as described in Section 4.1) hop diffusion has been observed in a number of SPT experiments. One potential problem is that in a plasma membrane partitioned by an actin meshwork (e.g. Fig. 1.2), the size of domains would not be expected to be uniform [39]. This observation could be incorporated into the model by allowing a different strength (κ_{z_i}) for each harmonic potential well (reversing a simplification we made in Section 4.3), hence enabling detection of variation in the domain size.

Another interesting avenue for future work regards artificial membranes. As shown in Chapter 3, these provide a useful model system for algorithm development, in addition to their value as a model membrane system. These experiments are especially useful for testing which components are required to reproduce effects seen in live cell membranes. For example, construction of an artificial membrane with the addition of an actin meshwork on the surface has been demonstrated [44]. An SPT experiment on this, or a similar, model membrane would be very interesting, potentially providing insight into whether hop diffusion of lipids in live cells can be reproduced by the addition of an actin meshwork to a model membrane.

Chapter 5

Discussion

5.1 Thesis summary

Working in a Bayesian framework, we have developed multiple biophysical models for the heterogeneous motion of single molecules in membranes. We implemented these models using MCMC methodology, deriving algorithms that infer model parameters and hidden states from SPT trajectories.

We first (Chapter 2) expanded an existing two-state diffusion HMM [8], where the diffusion coefficient of the particle switches between two values. We derived a Gibbs algorithm to fit the model to individual trajectories, and calculated the marginal likelihood (using an approximation) for one-state and two-state diffusion models, hence determining the most likely model given the trajectory. This approach included the effect of measurement error in the inference, and we showed (using fixed latex bead trajectories) that this is essential to prevent false detection of two-state behaviour. On SPT data of LFA-1 receptors diffusing on T cells we showed that 12-26% of trajectories show clear switching between diffusion states, depending on treatment. This within-trajectory switching was well described by a linear relationship, suggesting that the heterogeneity is due to a common mechanistic process. We also showed that there is heterogeneity in the diffusion coefficient across trajectories, with estimates being highly variable ($1.6 \times 10^2 - 2.6 \times 10^5$ nm^2 s^{-1}). Thus, LFA-1 diffusion is affected by a number of processes on a number of timescales. We hypothesise that this heterogeneity is due to clustering and cytoskeletal attachment.

We then proposed a confinement HMM (Chapter 3), where the particle switches between periods of free diffusion and confinement in a harmonic potential well (HPW). We derived an MCMC algorithm to infer model parameters and hidden states, and ap-

plied this to trajectories of GM1 lipids tagged with cholera toxin B subunit (CTxB) coated AuNPs in supported lipid bilayers (SLBs), confirming transient trapping in domains of mean radius 18 nm. The algorithm allowed us to segment trajectories into periods of free and confined motion. This revealed that confinement event shape within trajectories is more similar than across trajectories; we hence concluded that individual AuNP/CTxB/GM1 complexes have varying confinement shape signatures. We also showed that confinement event lifetimes are not exponentially distributed.

Finally (Chapter 4), we proposed a model for analysing diffusion in a series of square domains ("hop" diffusion), approximating confinement in a square as a harmonic potential well. We derived an MCMC algorithm for model inference, which contains a reversible jump move for inferring the number of wells. We demonstrated the validity of this algorithm on simulated hop diffusion trajectories.

To sum up, we have proposed multiple new HMMs, and MCMC algorithms for their inference, for the analysis of SPT trajectories. These algorithms have been applied to experimental data, extracting biophysical information which would not have been available using standard analysis techniques.

5.2 Context and further work

5.2.1 SPT in context

In addition to SPT, there are a number of other techniques for analysing membrane heterogeneity. Many of these, such as fluorescence recovery after photobleaching (FRAP) and fluorescence loss in photobleaching (FLIP), can only extract population-level information. A method which, in common with SPT, can extract information (i.e. biophysical parameters) on single particles is fluorescence correlation spectroscopy (FCS). Recent advances, such as STED-FCS [28, 30, 15], have greatly improved the characterisation of heterogeneous diffusion. These single particle methods are complementary, since SPT can track the motion of a particle within a large focal area, whereas FCS provides information on multiple particles at a single location.

5.2.2 Comparison with non-HMM approaches

As described in Chapter 1, there are a number of existing methods for detecting heterogeneity in SPT trajectories, with many approaches focusing on mean square displacements (MSD). HMM and MSD analysis have different aims and extract different (and complementary) information from trajectories. For example, MSD analysis should yield

more accurate estimates of diffusion coefficients for non-heterogeneous (i.e. Brownian) motion, as a result of pooling displacements across different timescales. Hence MSD techniques are often preferable for detecting heterogeneity across multiple trajectories. However, as shown for multiple models in this thesis, HMMs are capable of extracting fine spatiotemporal information from single trajectories, unlike an MSD analysis.

A comparison of the statistical power of various SPT analysis methods would be an interesting avenue for future work. For example, the two-state diffusion algorithm (Chapter 2) could be compared with algorithms that detect changes in the diffusion coefficient [77, 78, 79]. The HPW confinement algorithm (Chapter 3) could be compared with methods that detect spatial confinement [80, 81, 82]. This would need a careful approach since there is crossover when detecting changes in the diffusion coefficient and spatial confinement, with some methods theoretically capable of detecting both. A particularly interesting comparison could be made between the algorithm in Chapter 4 and previous methods for detecting hop diffusion [24, 39, 82]. As suggested by Meilhac *et al.* [82], these methods often cannot detect "double events" - hops between confining domains in quick succession. These events should be detectable using our algorithm.

Many existing non-HMM single trajectory analysis methods require "tuning" by comparison with Brownian motion, in order to minimise false detection of heterogeneity [80, 81, 82]. This is not necessary for the algorithms developed in this thesis.

5.2.3 HMM approach

HMM methods have been used previously to detect changes in the diffusion coefficient [8, 9, 10, 11, 12] and directional diffusion [12] in SPT trajectories. We have extended this by developing algorithms which detect transient confinement and hop diffusion. Thus a broad range of behaviour in trajectories can now theoretically be detected within a HMM approach.

The majority of previously published HMM methods segment trajectories into periods of different motion dynamics (e.g. diffusion states) along a trajectory, using either the Viterbi [9, 10, 83] or forward-backward [8] algorithms, and hence calculate the most likely value of the hidden state at each timepoint along the trajectory. This is analogous to our majority state thresholding in Chapter 3, where we approximated the posterior hidden state probabilities as a binary signal. On the other hand, the algorithms presented in this thesis calculate posterior hidden state probabilities at each timepoint, these estimates quantify the estimation uncertainty in the hidden state values. The Bayesian approach we have taken also has advantages over previous HMM algorithms,

since prior information (such as biophysical parameters obtained through previous experiments) can be incorporated into the inference. It is also a natural method for answering hypotheses (through model selection).

5.2.4 Effect of tagging molecules on particle trajectories

The extent to which a tag-target complex truly represents the motion of the target molecule in an SPT experiment is controversial [73, 58]. Additionally, as shown in Chapter 2, the tagging molecule (or tracking algorithm) can affect a single displacement analysis. In order to resolve these questions SPT data from a variety of tagging regimes (including smaller tags, such as small fluorescent molecules or quantum dots) is necessary. These trajectories will be analysable with the methods presented in this thesis.

We also showed that AuNP/CTxB/GM1 complexes have a specific confinement signature, this having implications for the reproducibility of trajectories. An attractive alternative is label-free tracking of biomolecules. For example, a viral molecule which also has GM1 as its cellular receptor is simian virus 40 (SV40). The SV40 virion is encased by a capsid containing 360 VP1 protein copies organised into an icosahedral shape [122]. Thus, unlike the random attachment of CTxB to the AuNP surface, the capsid always has the same structure, and we would not expect to see a specific confinement signature in individual virions. Interferometric, label-free, SPT of virions on supported lipid bilayers has been previously demonstrated [118], and with the subsequent improvement in the resolution of iSCAT methods, virion trajectories with high spatial and temporal resolution are theoretically obtainable.

5.2.5 Software package

A software package encompassing all the models proposed in this thesis, with some simplifications, is possible. This would be similar to the HMM-Bayes package which implements the Bayesian HMM modelling approach of Monnier *et al.* [12]; given an SPT trajectory, it would return parameter estimates for a number of models, and also the most likely model given the trajectory. In order to be attractive to experimentalists this would need to be user-friendly, requiring few input parameters. Another important feature would be a reasonable runtime on a regular desktop (as opposed to a high-performance cluster). In order to achieve this speed up model simplifications may be necessary.

A potential runtime saving could be made by using a reversible jump sampler

rather than Bayes factors. When using Bayes factors for model selection, the MCMC inference and marginal likelihood calculation has to be performed for all competing models. This computation is potentially wasteful - for example if some models have a very small marginal likelihood compared to others. A jump sampler would be preferable in this case, since it would spend almost no time inferring parameters for models with very low marginal likelihood. However, one potential drawback is that efficient proposal distributions for switching between models can be difficult to design. Additionally, since MCMC code consists mainly of loops, there is a limit to the speed that can be achieved in Matlab. Thus the algorithms, or certain sections thereof, could be improved by implementation in another language.

Another useful simplification would be to model HPW confinement zones with fixed centres, hence using a discrete (rather than continuous, as in Chapter 3) hidden state. If we denote this $z^{(C)}$ then, similarly to the hop diffusion model, it would denote the current confinement zone, i.e. $z_i^{(C)} \in \{1, .., M\}$ where M is the number of zones. If $z^{(C)} = 0$ then we have free diffusion. To develop an MCMC algorithm for this model, with a jump sampler to determine the number of confinement zones, would be a simple extension to the algorithm in Chapter 4. The jump sampler could also remove the free diffusion component, and hence switch to a hop diffusion model.

This model could also contain switches between an arbitrary number of diffusion states, via a second hidden state $z^{(D)}$ which denotes the current diffusion coefficient. A jump sampler for the number of diffusion states could then be derived. Putting these algorithms together could yield an MCMC algorithm which, provided sufficiently efficient proposal distributions could be found, infers the most likely model (and model parameters and hidden states) from the following: multi-state diffusion, HPW confinement, hop diffusion, and combinations of these models.

5.2.6 Final comments

In this thesis, we have demonstrated the utility and flexibility of using HMMs to analyse SPT data of membrane molecules. With sufficient (spatial and temporal) resolution of trajectories, this approach can extract information from experimental data that would not otherwise be available. It is reasonable to suppose that in the future, new single molecule techniques will emerge; these being capable of producing longer, higher resolution trajectories. The HMM methods presented here (and elsewhere) will be immediately applicable to these datasets. This data will also inspire the development of more intricate HMMs, enabling to the extraction of yet more detailed biophysical information. Thus,

the development of HMM methods is complementary to the continuing improvement in spatial and temporal resolution of data. This dual theoretical and experimental development will ultimately lead to more detailed biological interpretation of experiments, and hence more detailed knowledge of cell membranes.

Bibliography

[1] Singer SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. Science. 1972 Feb;175(4023):720–731.

[2] Jacobson KA, Sheets ED, Simson R. Revisiting the fluid mosaic model of membranes. Science. 1995 Jun;268(5216):1441–1442.

[3] Kusumi A, Nakada C, Ritchie K, Murase K, Suzuki K, Murakoshi H, et al. Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. Annual Review of Biophysics and Biomolecular Structure. 2005;34:351–378.

[4] Nicolson GL. The Fluid-Mosaic Model of Membrane Structure: still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years. Biochimica et Biophysica Acta. 2014 Jun;1838(6):1451–1466.

[5] Ishikawa-Ankerhold HC, Ankerhold R, Drummen GPC. Advanced Fluorescence Microscopy Techniques—FRAP, FLIP, FLAP, FRET and FLIM. Molecules. 2012 Dec;17(12):4047–4132.

[6] Levi V, Gratton E. Exploring dynamics in living cells by tracking single particles. Cell Biochemistry and Biophysics. 2007 May;48(1):1–15.

[7] Chiantia S, Ries J, Schwille P. Fluorescence correlation spectroscopy in membrane structure elucidation. Biochimica et Biophysica Acta. 2009 Jan;1788(1):225–233.

[8] Das R, Cairo CW, Coombs D. A hidden Markov model for single particle tracks quantifies dynamic interactions between LFA-1 and the actin cytoskeleton. PLOS Computational Biology. 2009 Nov;5(11):e1000556.

[9] Chung I, Akita R, Vandlen R, Toomre D, Schlessinger J, Mellman I. Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature. 2010 Jan;464(7289):783–787.

[10] Ott M, Shai Y, Haran G. Single-Particle Tracking Reveals Switching of the HIV Fusion Peptide between Two Diffusive Modes in Membranes. The Journal of Physical Chemistry B. 2013 Oct;117(42):13308–13321.

[11] Persson F, Lindén M, Unoson C, Elf J. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. Nature Methods. 2013 Mar;10(3):265–269.

[12] Monnier N, Barry Z, Park HY, Su KC, Katz Z, English BP, et al. Inferring transient particle transport dynamics in live cells. Nature Methods. 2015 Sep;12(9):838–840.

[13] Eggeling C. STED-FCS Nanoscopy of Membrane Dynamics. In: Mely Y, Duportail G, editors. Fluorescent Methods to Study Biological Membranes. Springer; 2013. p. 291–309.

[14] Kusumi A, Suzuki K, Kasai RS, Ritchie K, Fujiwara T. Hierarchical mesoscale domain organization of the plasma membrane. Trends in Biochemical Sciences. 2011 Nov;36(11):604–615.

[15] Eggeling C, Honigmann A. Molecular plasma membrane dynamics dissected by STED nanoscopy and Fluorescence Correlation Spectroscopy (STED-FCS). In: Cambi A, Lidke DS, editors. Cell Membrane nanodomains: From biochemistry to nanoscopy. CRC Press; 2014. p. 431–451.

[16] Kusumi A, Shirai YM, Koyama-Honda I, Suzuki K, Fujiwara T. Hierarchical organization of the plasma membrane: Investigations by single-molecule tracking vs. fluorescence correlation spectroscopy. FEBS Letters. 2010;584(9):1814–1823.

[17] Kaiser HJ, Orłowski A, Róg T, Nyholm TKM, Chai W, Feizi T, et al. Lateral sorting in model membranes by cholesterol-mediated hydrophobic matching. Proceedings of the National Academy of Sciences, USA. 2011 Oct;108(40):16628–16633.

[18] Sezgin E, Gutmann T, Buhl T, Dirkx R, Grzybek M, Coskun Ü, et al. Adaptive Lipid Packing and Bioactivity in Membrane Domains. PLOS ONE. 2015 Apr;10(4):e0123930.

[19] Pike LJ. The challenge of lipid rafts. The Journal of Lipid Research. 2008 Dec;50(Supplement):S323–S328.

[20] Lingwood D, Simons K. Lipid rafts as a membrane-organizing principle. Science. 2010 Jan;327(5961):46–50.

[21] Simons K, Sampaio JL. Membrane Organization and Lipid Rafts. Cold Spring Harbor Perspectives in Biology. 2011 Oct;3(10):a004697.

[22] Pike LJ. Lipid rafts: bringing order to chaos. The Journal of Lipid Research. 2003 Feb;44(4):655–667.

[23] Kraft ML. Plasma membrane organization and function: moving past lipid rafts. Molecular biology of the cell. 2013 Sep;24(18):2765–2768.

[24] Fujiwara T, Ritchie K, Murakoshi H, Jacobson KA, Kusumi A. Phospholipids undergo hop diffusion in compartmentalized cell membrane. Journal of Cell Biology. 2002 Jun;157(6):1071–1081.

[25] Jacobson KA, Mouritsen OG, Anderson RGW. Lipid rafts: at a crossroad between cell biology and physics. Nature Cell Biology. 2007 Jan;9(1):7–14.

[26] Simons K, Ikonen E. Functional rafts in cell membranes. Nature. 1997 Jul;387(6633):569–572.

[27] Sahl SJ, Leutenegger M, Hilbert M, Hell SW, Eggeling C. Fast molecular tracking maps nanoscale dynamics of plasma membrane lipids. Proceedings of the National Academy of Sciences, USA. 2010 Apr;107(15):6829–6834.

[28] Eggeling C, Ringemann C, Medda R, Schwarzmann G, Sandhoff K, Polyakova S, et al. Direct observation of the nanoscale dynamics of membrane lipids in a living cell. Nature. 2009 Feb;457(7233):1159–1162.

[29] Wüstner D. Plasma membrane sterol distribution resembles the surface topography of living cells. Molecular biology of the cell. 2007;18(1):211–228.

[30] Honigmann A, Mueller V, Ta H, Schoenle A, Sezgin E, Hell SW, et al. Scanning STED-FCS reveals spatiotemporal heterogeneity of lipid interaction in the plasma membrane of living cells. Nature Communications. 2014 Nov;5.

[31] Choquet D, Triller A. The role of receptor diffusion in the organization of the postsynaptic membrane. Nature Reviews Neuroscience. 2003 Apr;4(4):251–265.

[32] Bethani I, Skånland SS, Dikic I, Acker-Palmer A. Spatial organization of transmembrane receptor signalling. The EMBO journal. 2010 Aug;29(16):2677–2688.

[33] Garcia-Parajo MF, Cambi A, Torreno-Pina JA, Thompson NL, Jacobson KA. Nanoclustering as a dominant feature of plasma membrane organization. Journal of Cell Science. 2014 Nov;127(23):4995–5005.

[34] Saffman PG, Delbruck M. Brownian motion in biological membranes. Proceedings of the National Academy of Sciences, USA. 1975 Aug;72(8):3111–3113.

[35] Stabley D, Retterer S, Marshall S, Salaita K. Manipulating the lateral diffusion of surface-anchored EGF demonstrates that receptor clustering modulates phosphorylation levels. Integrative Biology. 2013 Mar;5(4):659–668.

[36] Cairo CW, Golan DE. T cell adhesion mechanisms revealed by receptor lateral mobility. Biopolymers. 2008 May;89(5):409–419.

[37] Bakker GJ, Eich C, Torreno-Pina JA, Diez-Ahedo R, Perez-Samper G, van Zanten TS, et al. Lateral mobility of individual integrin nanoclusters orchestrates the onset for leukocyte adhesion. Proceedings of the National Academy of Sciences, USA. 2012 Mar;109(13):4869–4874.

[38] Bray D, Levin MD, Morton-Firth CJ. Receptor clustering as a cellular mechanism to control sensitivity. Nature. 1998 May;393(6680):85–88.

[39] Suzuki K, Ritchie K, Kajikawa E, Fujiwara T, Kusumi A. Rapid hop diffusion of a G-protein-coupled receptor in the plasma membrane as revealed by single-molecule techniques. Biophysical Journal. 2005 Apr;88(5):3659–3680.

[40] Murase K, Fujiwara T, Umemura Y, Suzuki K, Iino R, Yamashita H, et al. Ultrafine Membrane Compartments for Molecular Diffusion as Revealed by Single Molecule Techniques. Biophysical Journal. 2004 Jun;86(6):4075–4093.

[41] Salbreux G, Charras G, Paluch E. Actin cortex mechanics and cellular morphogenesis. Trends in Cell Biology. 2012 Oct;22(10):536–545.

[42] Clausen MP, Lagerholm BC. Visualization of plasma membrane compartmentalization by high-speed quantum dot tracking. Nano Letters. 2013 Jun;13(6):2332–2337.

[43] Andrade DM, Clausen MP, Keller J, Mueller V, Wu C, Bear JE, et al. Cortical actin networks induce spatio-temporal confinement of phospholipids in the plasma membrane–a minimally invasive investigation by STED-FCS. Scientific Reports. 2015;5:11454.

[44] Honigmann A, Sadeghi S, Keller J, Hell SW, Eggeling C, Vink R. A lipid bound actin meshwork organizes liquid phase separation in model membranes. eLife. 2014 Mar;3.

[45] Jaqaman K, Kuwata H, Touret N, Collins R, Trimble WS, Danuser G, et al. Cytoskeletal Control of CD36 Diffusion Promotes Its Receptor and Signaling Function. Cell. 2011 Aug;146(4):593–606.

[46] Head BP, Patel HH, Insel PA. Interaction of membrane/lipid rafts with the cytoskeleton: impact on signaling and function: membrane/lipid rafts, mediators of cytoskeletal arrangement and cell signaling. Biochimica et Biophysica Acta. 2014 Feb;1838(2):532–545.

[47] Renner M, Schweizer C, Bannai H, Triller A, Lévi S. Diffusion barriers constrain receptors at synapses. PLOS ONE. 2012;7(8).

[48] Davis DM, Dustin ML. What is the importance of the immunological synapse? Trends in Immunology. 2004 Jun;25(6):323–327.

[49] Iyengar G, Rao M. A cellular solution to an information-processing problem. Proceedings of the National Academy of Sciences, USA. 2014 Aug;111(34):12402–12407.

[50] Hsieh CL, Spindler S, Ehrig J, Sandoghdar V. Tracking Single Particles on Supported Lipid Membranes: Multimobility Diffusion and Nanoscopic Confinement. The Journal of Physical Chemistry B. 2014 Jan;118(6):1545–1554.

[51] Spillane KM, Ortega-Arroyo J, de Wit G, Eggeling C, Ewers H, Wallace MI, et al. High-Speed Single-Particle Tracking of GM1 in Model Membranes Reveals Anomalous Diffusion due to Interleaflet Coupling and Molecular Pinning. Nano Letters. 2014 Sep;14(9):5390–5397.

[52] Chang YP, Pinaud F, Antelman J, Weiss S. Tracking bio-molecules in live cells using quantum dots. Journal of Biophotonics. 2008 Sep;1(4):287–298.

[53] Sergé A, Bertaux N, Rigneault H, Marguet D. Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. Nature Methods. 2008 Jul;5(8):687–694.

[54] Barroso MM. Quantum Dots in Cell Biology. Journal of Histochemistry & Cytochemistry. 2011 Mar;59(3):237–251.

[55] Clausen MP, Arnspang EC, Ballou B, Bear JE, Lagerholm BC. Simultaneous Multi-Species Tracking in Live Cells with Quantum Dot Conjugates. PLOS ONE. 2014 Jun;9(6):e97671.

[56] Wang YL, Silverman JD, Cao LG. Single particle tracking of surface receptor movement during cell division. Journal of Cell Biology. 1994 Nov;127(4):963–971.

[57] Cairo CW, Mirchev R, Golan DE. Cytoskeletal regulation couples LFA-1 conformational changes to receptor lateral mobility and clustering. Immunity. 2006 Jul;25(2):297–308.

[58] Mascalchi P, Haanappel E, Carayon K, Mazères S, Salomé L. Probing the influence of the particle in Single Particle Tracking measurements of lipid diffusion. Soft Matter. 2012;8(16):4462–4470.

[59] Ji B, Giovanelli E, Habert B, Spinicelli P, Nasilowski M, Xu X, et al. Non-blinking quantum dot with a plasmonic nanoshell resonator. Nature Nanotechnology. 2015 Feb;10(2):170–175.

[60] Qian H, Sheetz MP, Elson EL. Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. Biophysical Journal. 1991 Oct;60(4):910–921.

[61] Lee GM, Ishihara A, Jacobson KA. Direct observation of brownian motion of lipids in a membrane. Proceedings of the National Academy of Sciences, USA. 1991 Jul;88(14):6274–6278.

[62] Michalet X. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. Physical Review E. 2010 Sep;82(4 Pt 1):041914.

[63] Michalet X, Berglund AJ. Optimal diffusion coefficient estimation in single-particle tracking. Physical Review E. 2012 Jun;85(6 Pt 1).

[64] Berglund AJ. Statistics of camera-based single-particle tracking. Physical Review E. 2010 Jul;82(1):011917.

[65] Saxton MJ. Anomalous diffusion due to obstacles: a Monte Carlo study. Biophysical Journal. 1994 Feb;66(2 Pt 1):394–401.

[66] Kepten E, Bronshtein I, Garini Y. Improved estimation of anomalous diffusion exponents in single-particle tracking experiments. Physical Review E. 2013 Apr;87(5):052713.

[67] Kepten E, Weron A, Sikora G, Burnecki K, Garini Y. Guidelines for the fitting of anomalous diffusion mean square displacement graphs from single particle tracking experiments. PLOS ONE. 2015;10(2):e0117722.

[68] Burnecki K, Kepten E, Garini Y, Sikora G, Weron A. Estimating the anomalous diffusion exponent for single particle tracking data with measurement errors - An alternative approach. Scientific Reports. 2015;5:11306.

[69] Regner BM, Vučinić D, Domnisoru C, Bartol TM, Hetzer MW, Tartakovsky DM, et al. Anomalous Diffusion of Single Particles in Cytoplasm. Biophysical Journal. 2013 Apr;104(8):1652–1660.

[70] Höfling F, Franosch T. Anomalous transport in the crowded world of biological cells. Reports on Progress in Physics. 2013 Mar;76(4):046602.

[71] Kusumi A, Sako Y, Yamamoto M. Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells. Biophysical Journal. 1993;65(5):2021–2040.

[72] Wieser S, Schütz GJ. Tracking single molecules in the live cell plasma membrane—Do's and Don't's. Methods. 2008 Sep;46(2):131–140.

[73] Saxton MJ, Jacobson KA. Single-particle tracking: applications to membrane dynamics. Annual Review of Biophysics and Biomolecular Structure. 1997;26:373–399.

[74] Monnier N, Guo SM, Mori M, He J, Lénárt P, Bathe M. Bayesian approach to MSD-based analysis of particle motion in live cells. Biophysical Journal. 2012 Aug;103(3):616–626.

[75] Rajani V, Carrero G, Golan DE, de Vries G, Cairo CW. Analysis of molecular diffusion by first-passage time variance identifies the size of confinement zones. Biophysical Journal. 2011 Mar;100(6):1463–1472.

[76] Schütz GJ, Schindler H, Schmidt T. Single-molecule microscopy on model membranes reveals anomalous diffusion. Biophysical Journal. 1997 Aug;73(2):1073–1080.

[77] Montiel D, Cang H, Yang H. Quantitative Characterization of Changes in Dynamical Behavior for Single-Particle Tracking Studies. The Journal of Physical Chemistry B. 2006 Oct;110(40):19763–19770.

[78] Huet S, Karatekin E, Tran VS, Fanget I, Cribier S, Henry JP. Analysis of Transient Behavior in Complex Trajectories: Application to Secretory Vesicle Dynamics. Biophysical Journal. 2006 Nov;91(9):3542–3559.

[79] Helmuth JA, Burckhardt CJ, Koumoutsakos P, Greber UF, Sbalzarini IF. A novel supervised trajectory segmentation algorithm identifies distinct types of human adenovirus motion in host cells. Journal of Structural Biology. 2007 Sep;159(3):347–358.

[80] Saxton MJ. Lateral diffusion in an archipelago. Single-particle diffusion. Biophysical Journal. 1993 May;64(6):1766–1780.

[81] Simson R, Sheets ED, Jacobson KA. Detection of temporary lateral confinement of membrane proteins using single-particle tracking analysis. Biophysical Journal. 1995 Sep;69(3):989–993.

[82] Meilhac N, Le Guyader L, Salomé L, Destainville N. Detection of confinement and jumps in single-molecule membrane trajectories. Physical Review E. 2006 Jan;73(1 Pt 1):011915.

[83] Monnier N. Bayesian Inference Approaches for Particle Trajectory Analysis in Cell Biology. Harvard University; 2013.

[84] Øksendal BK. Stochastic Differential Equations. An Introduction with Applications. Springer; 1989.

[85] Karatzas I, Shreve S. Brownian Motion and Stochastic Calculus. Springer; 2014.

[86] Higham DJ. An algorithmic introduction to numerical simulation of stochastic differential equations. SIAM Review. 2001;43(3):525–546.

[87] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State Calculations by Fast Computing Machines. Journal of Chemical Physics. 1953;21(6):1087.

[88] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika. 1970;57(1):97–109.

[89] Roberts GO, Gelman A, Gilks WR. Weak Convergence And Optimal Scaling Of Random Walk Metropolis Algorithms. The Annals of Applied Probability. 1997;7(1):110–120.

[90] Chib S, Greenberg E. Understanding the Metropolis-Hastings Algorithm. The American Statistician. 1995;49(4):327–335.

[91] Geman S, Geman D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE transactions on pattern analysis and machine intelligence. 1984 Jun;6(6):721–741.

[92] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis, Third Edition. CRC Press; 2013.

[93] Gilks WR, Richardson S, Spiegelhalter D. Markov Chain Monte Carlo in Practice. CRC Press; 1995.

[94] Rubin DB, Gelman A. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science. 1992;7(4):457–511.

[95] Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics. 1998 Dec;7(4):434–455.

[96] Kass RE, Raftery AE. Bayes factors. Journal of the American Statistical Association. 1995;90(430):773–795.

[97] Chib S. Marginal likelihood from the Gibbs output. Journal of the American Statistical Association. 1995;90(432):1313–1321.

[98] Chib S, Jeliazkov I. Marginal likelihood from the Metropolis–Hastings output. Journal of the American Statistical Association. 2001;96(453):270–281.

[99] Chen MH. Computing marginal likelihoods from a single MCMC output. Statistica Neerlandica. 2005 Jan;59(1):16–29.

[100] Massignan P, Manzo C, Torreno-Pina JA, Garcia-Parajo MF, Lewenstein M, Lapeyre GJ. Nonergodic Subdiffusion from Brownian Motion in an Inhomogeneous Medium. Physical Review Letters. 2014 Apr;112(15):150603.

[101] Powles JG, Mallett MJD, Rickayzen G, Evans WAB. Exact Analytic Solutions for Diffusion Impeded by an Infinite Array of Partially Permeable Barriers. Proceedings of the Royal Society A. 1992 Feb;436(1897):391–403.

[102] Saxton MJ. Single-particle tracking: models of directed transport. Biophysical Journal. 1994;67(5):2110–2119.

[103] Meier J, Vannier C, Sergé A, Triller A, Choquet D. Fast and reversible trapping of surface glycine receptors by gephyrin. Nature Neuroscience. 2001 Mar;4(3):253–260.

[104] Pinaud F, Michalet X, Iyer G, Margeat E, Moore HP, Weiss S. Dynamic Partitioning of a Glycosyl-Phosphatidylinositol-Anchored Protein in Glycosphingolipid-Rich Microdomains Imaged by Single-Quantum Dot Tracking. Traffic. 2009 Jun;10(6):691–712.

[105] Simonson WTN, Franco SJ, Huttenlocher A. Talin1 Regulates TCR-Mediated LFA-1 Function. The Journal of Immunology. 2006 Nov;177(11):7707–7714.

[106] Lollo BA, Chan KW, Hanson EM, Moy VT, Brian AA. Direct evidence for two affinity states for lymphocyte function-associated antigen 1 on activated T cells. Journal of Biological Chemistry. 1993 Oct;268(29):21693–21700.

[107] Drbal K, Angelisová P, Cerný J, Hilgert I, Horejsí V. A novel anti-CD18 mAb recognizes an activation-related epitope and induces a high-affinity conformation in leukocyte integrins. Immunobiology. 2001 May;203(4):687–698.

[108] Kucik DF, Dustin ML, Miller JM, Brown EJ. Adhesion-activating phorbol ester increases the mobility of leukocyte integrin LFA-1 in cultured lymphocytes. The Journal of Clinical Investigation. 1996 May;97(9):2139–2144.

[109] Cambi A, Joosten B, Koopman M, de Lange F, Beeren I, Torensma R, et al. Organization of the integrin LFA-1 in nanoclusters regulates its activity. Molecular biology of the cell. 2006 Oct;17(10):4270–4281.

[110] Franco SJ, Rodgers MA, Perrin BJ, Han J, Bennin DA, Critchley DR, et al. Calpain-mediated proteolysis of talin regulates adhesion dynamics. Nature Cell Biology. 2004 Sep;6(10):977–983.

[111] Lindley DV. A Statistical Paradox. Biometrika. 1957 Jun;44:187–192.

[112] Hui KL, Balagopalan L, Samelson LE, Upadhyaya A. Cytoskeletal forces during signaling activation in Jurkat T-cells. Molecular biology of the cell. 2015 Feb;26(4):685–695.

[113] Hogg N, Smith A, McDowall A, Giles K, Stanley P, Laschinger M, et al. How T cells use LFA-1 to attach and migrate. Immunology Letters. 2004 Mar;92(1-2):51–54.

[114] Lillemeier BF, Mörtelmaier MA, Forstner MB, Huppa JB, Groves JT, Davis MM. TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation. Nature Immunology. 2009 Dec;11(1):90–96.

[115] Slator PJ, Cairo CW, Burroughs NJ. Detection of Diffusion Heterogeneity in Single Particle Tracking Trajectories Using a Hidden Markov Model with Measurement Noise Propagation. PLOS ONE. 2015 Oct;10(10):e0140759.

[116] Margheri G, D'Agostino R, Trigari S, Sottini S, Del Rosso M. The β-Subunit of Cholera Toxin has a High Affinity for Ganglioside GM1 Embedded into Solid Supported Lipid Membranes with a Lipid Raft-Like Composition. Lipids. 2013 Oct;49(2):203–206.

[117] Shi J, Yang T, Kataoka S, Zhang Y, Diaz AJ, Cremer PS. GM1 clustering inhibits cholera toxin binding in supported phospholipid membranes. Journal of the American Chemical Society. 2007 May;129(18):5954–5961.

[118] Ewers H, Jacobsen V, Klotzsch E, Smith AE, Helenius A, Sandoghdar V. Label-Free Optical Detection and Tracking of Single Virions Bound to Their Receptors in Supported Membrane Bilayers. Nano Letters. 2007 Aug;7(8):2263–2266.

[119] Ritchie K, Shan XY, Kondo J, Iwasawa K, Fujiwara T, Kusumi A. Detection of Non-Brownian Diffusion in the Cell Membrane in Single Molecule Tracking. Biophysical Journal. 2005 Mar;88(3):2266–2277.

[120] Niehaus AMS, Vlachos DG, Edwards JS, Plechac P, Tribe R. Microscopic simulation of membrane molecule diffusion on corralled membrane surfaces. Biophysical Journal. 2008 Mar;94(5):1551–1564.

[121] Green PJ. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika. 1995;82(4):711–732.

[122] Szklarczyk OM, González-Segredo N, Kukura P, Oppenheim A, Choquet D, Sandoghdar V, et al. Receptor Concentration and Diffusivity Control Multivalent

Binding of Sv40 to Membrane Bilayers. PLOS Computational Biology. 2013 Nov;9(11):e1003310.

Appendix A

Supporting Information for Chapter 2

A.1 Supplementary mathematical derivations

This document includes step by step calculation of likelihoods, marginal likelihoods and MCMC algorithms, for one-state and two-state diffusion models as described in the Methods section.

A.1.1 One-state diffusion model marginal likelihood calculation

The marginal likelihood is defined

$$\pi(\mathbf{X}|M_{1D}) = \int_0^\infty \mathrm{d}D\, \pi(\mathbf{X}|D, M_{1D})\pi(D)$$

since $\pi(D) = \text{Uniform}(0, D_{max})$ we can write

$$\pi(\mathbf{X}|M_{1D}) = \frac{1}{D_{max}} \int_0^{D_{max}} \mathrm{d}D\, \pi(\mathbf{X}|D, M_{1D})$$

changing variables from D to D^{-1} gives

$$\begin{aligned}\pi(\mathbf{X}|M_{1D}) &= \frac{1}{D_{max}}\int_{1/D_{max}}^{\infty} \mathrm{d}D^{-1}\, D^2\pi(\mathbf{X}|D, M_{1D}) \\ &= \frac{1}{D_{max}}\int_{1/D_{max}}^{\infty} \mathrm{d}D^{-1}\, D^2 \prod_{i=1}^{N}\frac{1}{4\pi D\Delta t_i}\exp\left(-\sum_{i=1}^{N}\frac{\Delta X_i^2}{4D\Delta t_i}\right) \\ &= \frac{1}{D_{max}}\prod_{i=1}^{N}\frac{1}{4\pi\Delta t_i}\int_{1/D_{max}}^{\infty}\mathrm{d}D^{-1}\left(\frac{1}{D}\right)^{N-2}\exp\left(-\frac{1}{D}\sum_{i=1}^{N}\frac{\Delta X_i^2}{4\Delta t_i}\right).\end{aligned}$$

In general $\int_x^\infty \mathrm{d}t\, t^{\alpha-1}\mathrm{e}^{-\beta t} = \frac{1}{\beta^\alpha}\Gamma(\alpha,\beta x)$, where $\Gamma(\alpha,\beta x)$ is an upper incomplete gamma function. Using this gives

$$\pi(\mathbf{X}|M_{1D}) = \frac{1}{D_{max}}\prod_{i=1}^{N}\frac{1}{4\pi\Delta t_i}\left(\sum_{i=1}^{N}\frac{\Delta X_i^2}{4\Delta t_i}\right)^{1-N}\Gamma\left(N-1, \frac{1}{D_{max}}\sum_{i=1}^{N}\frac{\Delta X_i^2}{4\Delta t_i}\right).$$

A.1.2 Approximate one-state diffusion model with measurement noise

We consider a trajectory $\mathbf{X}$ subject to Gaussian observation error, with fixed localisation accuracy σ^2. By a result in [62] (also see Section A.1.4) an approximation for the likelihood of $\mathbf{X}$ given D is

$$\pi(\mathbf{X}|D) = \prod_{i=1}^{N} N\left(\Delta X_i; 0, 2(D\Delta t_i + \sigma^2)\right).$$

The associated posterior is

$$\pi(D|\mathbf{X}) \propto \pi(D)\prod_{i=1}^{N} N\left(\Delta X_i; 0, 2D\Delta t_i + 2\sigma^2\right),$$

which can be sampled using a Metropolis-Hastings algorithm. We set $\pi(D) = \mathrm{Unif}(D; 0, D_{max})$, and use a random walk sampler (RW MCMC) with a symmetric Gaussian proposal, $q(D \to D') = N(D'; D, S_D)$, giving the acceptance probability

$$\alpha(D \to D') = \min\left\{1, \frac{\prod_{i=1}^{N} N\left(\Delta X_i; 0, 2(D'\Delta t_i + \sigma^2)\right)}{\prod_{i=1}^{N} N\left(\Delta X_i; 0, 2(D\Delta t_i + \sigma^2)\right)}\right\}\mathbb{1}_{[0,D_{max}]}(D).$$

Thus, any moves outside $[0, D_{max}]$ are automatically rejected. The value of S_D is tuned during the burn-in to ensure that the acceptance rate is approximately 0.25 [89]. The MCMC sampler is also given as pseudocode in Section A.3.

A.1.3 Approximate two-state diffusion model with measurement noise

We now add fixed localisation error to the previous two-state diffusion hidden Markov model. Using the same approximation to the likelihood as the approximate one-state model we can write

$$\begin{aligned} z_i|_{z_{i-1}} &\sim \text{Bernoulli}(z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01}) \\ \Delta X_i|_{z_i} &\sim N(0, 2(D_{z_i}\Delta t_i + \sigma^2)). \end{aligned}$$

Letting $\theta = \{D_0, D_1, p_{01}, p_{10}\}$ we can write the posterior as

$$\begin{aligned} \pi(\theta, \mathbf{z}|\mathbf{X}) \propto \pi(\theta)\pi(z_1|\theta) &\prod_{i=1}^{N} N(\Delta X_i; 0, 2(D_{z_i}\Delta t_i + \sigma^2)) \\ &\times \prod_{i=1}^{N-1} \text{Bernoulli}(z_{i+1}; z_i(1-p_{10}) + (1-z_i)p_{01}). \end{aligned} \tag{A.1}$$

We use the same priors on D_0, D_1, p_{01}, p_{10} and z_1 as in the two-state diffusion model without measurement noise, given in Equation (2.8), main text. D_0, D_1 are updated with Metropolis-Hastings moves. The proposals are Gaussians centred at the current value $q(D_0 \to D_0') = N(D_0'; D_0, S_{D_0}), q(D_1 \to D_1') = N(D_1'; D_1, S_{D_1})$ and the acceptance probabilities are

$$\begin{aligned} \alpha(D_0 \to D_0'|\mathbf{z}, \mathbf{X}) &= \min\left\{1, \frac{\prod_{z_i=0} N(\Delta X_i; 0, 2(D_0'\Delta t_i + \sigma^2))}{\prod_{z_i=0} N(\Delta X_i; 0, 2(D_0\Delta t_i + \sigma^2))}\right\} \mathbb{1}_{[0,D_{max}]}(D_0') \\ \alpha(D_1 \to D_1'|\mathbf{z}, \mathbf{X}) &= \min\left\{1, \frac{\prod_{z_i=1} N(\Delta X_i; 0, 2(D_1'\Delta t_i + \sigma^2))}{\prod_{z_i=1} N(\Delta X_i; 0, 2(D_1\Delta t_i + \sigma^2))}\right\} \mathbb{1}_{[0,D_{max}]}(D_1'). \end{aligned}$$

S_{D_0}, S_{D_1} are tuned during the burn-in to ensure an acceptance rate of approximately 0.25. We also impose the condition that $D_0 < D_1$, which we enforce after the MCMC run as follows: if the posterior means $\hat{D_0} > \hat{D_1}$ then we swap the D_0, D_1 chains, swap the p_{01}, p_{10} chains, and swap the 0 and 1 states in the hidden state $\mathbf{z}$ throughout the run. This is possible because although state identity switching ($0 \leftrightarrow 1$) is possible because of a permutation symmetry during a run, it isn't observed to occur. The updates for

the transition probabilities are Gibbs moves, identical to the two-state model without measurement noise, given by Equations (2.16) and (2.17), main text. The $\mathbf{z}$ update is similar to the other two-state models, the conditional is

$$\begin{aligned}\pi(z_i|z_{i-1}, z_{i+1}, ..) \propto & \text{ Bernoulli}(z_i; z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01}) \\ & \times N(\Delta X_i; 0, 2D_{z_i}\Delta t_i + 2\sigma^2) \\ & \times \text{Bernoulli}(z_{i+1}; z_i(1-p_{10}) + (1-z_i)p_{01}). \end{aligned} \tag{A.2}$$

And again the update is

$$z_i\Big|_{\theta,\mathbf{U},z_{i\pm1}} \sim \text{Bernoulli}\left(\pi(z_i = 1|z_{i-1}, z_{i+1}, \theta, \mathbf{X})\right).$$

At the endpoints $i = 1$ and $i = N$ we have

$$\pi(z_1|z_2, \theta, \mathbf{X}) \propto N(\Delta X_1; 0, 2D_{z_1}\Delta t_1 + 2\sigma^2)\text{Bernoulli}\left(z_2; z_1(1-p_{10}) + (1-z_1)p_{01}\right) \tag{A.3}$$

$$\pi(z_N|z_{N-1}, \theta, \mathbf{X}) \propto \text{Bernoulli}\left(z_N; z_{N-1}(1-p_{10}) + (1-z_{N-1})p_{01}\right) N(\Delta X_N; 0, 2D_{z_N}\Delta t_N + 2\sigma^2). \tag{A.4}$$

Pseudocode for this MCMC sampler is given in Section A.3.

A.1.4 Approximation to the likelihood for one-state diffusion model with measurement noise

(This method is mentioned in reference [62].) Consider a 2D trajectory observed with experimental noise with known localisation accuracy σ^2. Let $\{U_i\}_{i=1}^{N+1}$ be the underlying particle position and $\{X_i\}_{i=1}^{N+1}$ be the observed positions. For each time step

$$\begin{aligned} U_i - U_{i-1} &\sim N(0, 2D\Delta t_{i-1}) \\ X_i &\sim N(U_i, \sigma^2). \end{aligned}$$

Which we can write as (summing two Gaussians)

$$X_i - X_{i-1} \sim N(U_i - U_{i-1}, 2\sigma^2)$$

shifting the mean

$$X_i - X_{i-1} \sim N(0, 2\sigma^2) + U_i - U_{i-1}$$

since $U_i - U_{i-1} \sim N(0, 2D\Delta t_{i-1})$ we can write

$$X_i - X_{i-1} \sim N(0, 2D\Delta t_{i-1} + 2\sigma^2).$$

So we know that the measured displacement then satisfies $X_{i+1}|X_i \sim X_i + N(0, 2D\Delta t_i + 2\sigma^2)$, which suggests that the likelihood is given by

$$\pi(\mathbf{X}|D) = \prod_{i=1}^{N} N(\Delta X_i; 0, 2D\Delta t_i + 2\sigma^2). \tag{A.5}$$

However, this is only true if the displacements are independent, which not the case since the displacements $U_{i+1} - U_i$ and $U_i - U_{i-1}$ both depend on the measurement noise $U_i - X_i$ at time point i. However, we demonstrate that Equation (A.5) is sufficient for model selection, see Results.

A.1.5 Log likelihood for approximate two-state diffusion model with measurement noise

We use a modified version of the Das *et al.* forward algorithm [8] to calculate $\pi(\mathbf{X}|\theta)$. The initial forward probabilities in log scale are

$$\log_e \alpha_1(z_1 = 0) = \log_e \frac{p_{10}}{p_{10} + p_{01}} + \log_e \pi(\Delta X_1 | z_1 = 0, D_0, D_1)$$
$$\log_e \alpha_1(z_1 = 1) = \log_e \frac{p_{01}}{p_{10} + p_{01}} + \log_e \pi(\Delta X_1 | z_1 = 1, D_0, D_1)$$

where $\log_e \pi(\Delta X_i | z_i, D_0, D_1) = N(\Delta X_i; 0, 2(D_{z_i}\Delta t_i + \sigma^2))$ for $i = 1..N$. The recursion for $i = 2$ to $i = N$ is then

$$\log_e \alpha_i(z_i = 0) = \log_e \Big[e^{\log_e \alpha_{i-1}(z_{i-1}=0) + \log_e(1-p_{01}) + \log_e \pi(\Delta X_i | z_i=0, D_0, D_1)} + e^{\log_e \alpha_{i-1}(z_{i-1}=1) + \log_e(p_{10}) + \log_e \pi(\Delta X_i | z_i=0, D_0, D_1)} \Big]$$

$$\log_e \alpha_i(z_i = 1) = \log_e \Big[e^{\log_e \alpha_{i-1}(z_{i-1}=1) + \log_e(p_{01}) + \log_e \pi(\Delta X_i | z_i=1, D_0, D_1)} + e^{\log_e \alpha_{i-1}(z_{i-1}=0) + \log_e(1-p_{10}) + \log_e \pi(\Delta X_i | z_i=1, D_0, D_1)} \Big].$$

And the final likelihood is

$$\pi(\mathbf{X}|\theta) = \log_e \left[e^{\log_e \alpha_N(z_N=0)} + e^{\log_e \alpha_N(z_N=1)} \right].$$

A.2 Subsampling trajectories to reduce the effect of measurement noise

Here we provide justification and description of the subsampling approach described in the Discussion. We showed (Fig. 2.1) that stationary beads show fast switching between two distinct diffusion coefficients, and prefer a two-state diffusion model in a model selection analysis. To address this problem, we tried subsampling the data to increase the S/N ratio, and hence minimise bias caused by the inherent two-state diffusion behaviour observed on stationary beads. This should enable differentiation between real biological switching between states with different mobilities, and artificial two diffusion behaviour.

For a trajectory $\mathbf{X}$, we subsample by taking every nth time point to obtain a trajectory of length $\lfloor \frac{N}{n} \rfloor$. Different choices of n can greatly affect the inferred parameters. Michalet provides some theoretical reasoning for the choice of subsampling rate n by determining the optimum number of points of the mean square displacement (MSD) function to include when estimating D [62]. For example, if the optimum number of MSD points is two, then the optimal D estimate comes from a linear fit to the first two MSD points (ignoring the $(0, 0)$ point). Selecting a suboptimal number of MSD points can lead to large errors.

If the optimum number of MSD points is large then single displacement estimates for D can be out by orders of magnitude [62, 63]. The D value from a single displacement analysis is equivalent to fitting a straight line from $(0, 0)$ to the first MSD point. For the trajectories in Fig. A.10 A (fixed latex bead) and Fig. A.10 B (slow moving LFA-1) the single displacement fit is not a good approximation for an MSD fit using the optimum number of points. Fig. A.10 C shows a fast moving LFA-1 trajectory, in this case the single displacement fit is a good approximation to the MSD fit.

Intuitively, the optimum number of MSD points should be a good approximation to the best subsampling rate. To test this the one-state and two-state diffusion model and model selection analysis (without measurement noise) was run with a subsampling rate equal to the theoretical optimum number of MSD points. Table A.3 compares the preferred model from this approach to the preferred model using the one-state and two-state diffusion models with measurement noise.

A.3 Pseudocode for one-state and two-state diffusion model MCMC algorithms.

Algorithm 5 Gibbs sampler for two-state diffusion model without measurement noise

$K \leftarrow$ number of MCMC steps
$D_{max}, a_0, b_0, a_1, b_1 \leftarrow$ choice of prior parameters (Equation (2.8))
$D_0 \leftarrow$ random number drawn from Unif$(0, D_{max})$
$D_1 \leftarrow$ random number drawn from Unif$(0, D_{max})$
$p_{01} \leftarrow$ random number drawn from Beta(a_0, b_0)
$p_{10} \leftarrow$ random number drawn from Beta(a_1, b_1)
$\theta^{(1)} \leftarrow \{D_0, D_1, p_{01}, p_{10}\}$
$z_1 \leftarrow$ random number drawn from Bernoulli$\left(\frac{p_{10}}{p_{10}+p_{01}}\right)$
for $i = 2$ to $i = N$ **do**
    $z_i \leftarrow$ random number drawn from Bernoulli$(z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01})$
end for
$\mathbf{z}^{(1)} \leftarrow \{z_i\}_{i=1}^N$
for $k = 2$ to $k = K$ **do**
    $\eta_0 \leftarrow \sum_{z_i=0} 1$
    $\eta_1 \leftarrow \sum_{z_i=1} 1$
    if $\eta_0 = 0$ **then**
        $D_0 \leftarrow$ random number drawn from Unif$(0, D_{max})$
    else
        while $D_0 > D_{max}$ **do**
            $D_0^{-1} \leftarrow$ random number drawn from Gamma $\left(\eta_0 - 1, \sum_{i|z_i=0} \frac{\Delta X_i^2}{4\Delta t_i}\right)$
            $D_0 \leftarrow 1/D_0^{-1}$
        end while
    end if
    if $\eta_1 = 0$ **then**
        $D_1 \leftarrow$ random number drawn from Unif$(0, D_{max})$
    else
        while $D_1 > D_{max}$ **do**
            $D_1^{-1} \leftarrow$ random number drawn from Gamma $\left(\eta_1 - 1, \sum_{i|z_i=1} \frac{\Delta X_i^2}{4\Delta t_i}\right)$
            $D_1 \leftarrow 1/D_1^{-1}$
        end while
    end if
    Update $n_{10}, n_{11}, n_{01}, n_{00}$ using Equation (2.15)
    $p_{01} \leftarrow$ random number drawn from Beta$\big(a_0 + n_{10}, b_0 + n_{11}\big)$
    $p_{10} \leftarrow$ random number drawn from Beta$\big(a_1 + n_{01}, b_1 + n_{00}\big)$
    for $i = 1$ to $i = N$ **do**
        Calculate $\pi(z_i = 1 | z_{i-1}, z_{i+1}, D_0, D_1, p_{01}, p_{10}, \mathbf{X})$ by normalising Equation (2.18), (2.20) or (2.21)
        $z_i \leftarrow$ random number drawn from Bernoulli$(\pi(z_i = 1 | z_{i-1}, z_{i+1}, D_0, D_1, p_{01}, p_{10}, \mathbf{X}))$
    end for
    $\theta^{(k)} \leftarrow \{D_0, D_1, p_{01}, p_{10}\}$
    $\mathbf{z^{(k)}} \leftarrow \{z_i\}_{i=1}^N$
end for

Algorithm 6 Gibbs sampler for one-state diffusion model with measurement noise

$K \leftarrow$ number of MCMC steps
$D_{max}, \mu_{U_1}, \sigma_U^2 \leftarrow$ choice of prior parameters
$\sigma^2 \leftarrow$ choice of localisation accuracy
$D \leftarrow$ random number drawn from Unif$(0, D_{max})$
$D^{(1)} \leftarrow D$
$U_1 \leftarrow$ random number drawn from $N(\mu_{U_1}, \sigma_U^2)$
for $i = 2$ to $i = N + 1$ **do**
    $U_i \leftarrow$ random number drawn from $N(U_{i-1}, 2D\Delta t_{i-1})$
end for
$\mathbf{U}^{(1)} \leftarrow \{U_i\}_{i=1}^{N+1}$
for $k = 2$ to $k = K$ **do**
    $D^{-1} \leftarrow$ random number drawn from Gamma $\left(N - 1, \sum_{i=1}^{N} \frac{\Delta U_i^2}{4\Delta t_i}\right)$
    while $D > D_{max}$ **do**
        $D^{-1} \leftarrow$ random number drawn from Gamma $\left(N - 1, \sum_{i=1}^{N} \frac{\Delta U_i^2}{4\Delta t_i}\right)$
    end while
    $D^{(k)} \leftarrow D$
    for $i = 1$ to $i = N + 1$ **do**
        Calculate μ_i, τ_i from Equations (2.32), (2.33) or (2.34)
        $U_i \leftarrow$ random number drawn from $N(\mu_i, 1/\tau_i)$
    end for
    $\mathbf{U}^{(k)} \leftarrow \{U_i^{(k)}\}_{i=1}^{N+1}$
end for

Algorithm 7 Gibbs sampler for two-state diffusion model with measurement noise

$K \leftarrow$ number of MCMC steps
$D_{max}, a_0, b_0, a_1, b_1, \mu_{U_1}, \sigma_U^2 \leftarrow$ choice of prior parameters (Equation (2.41))
$\sigma^2 \leftarrow$ choice of localisation accuracy
$D_0 \leftarrow$ random number drawn from Unif$(0, D_{max})$
$D_1 \leftarrow$ random number drawn from Unif$(0, D_{max})$
$p_{01} \leftarrow$ random number drawn from Beta(a_0, b_0)
$p_{10} \leftarrow$ random number drawn from Beta(a_1, b_1)
$\theta^{(1)} \leftarrow \{D_0, D_1, p_{01}, p_{10}\}$
$z_1 \leftarrow$ random number drawn from Bernoulli$\left(\frac{p_{10}}{p_{10}+p_{01}}\right)$
for $i = 2$ to $i = N$ **do**
  $z_i^{(1)} \leftarrow$ random number drawn from Bernoulli$(z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01})$
end for
$\mathbf{z}^{(1)} \leftarrow \{z_i\}_{i=1}^N$
$U_1 \leftarrow$ random number drawn from $N(\mu_{U_1}, \sigma_U^2)$
for $i = 2$ to $i = N+1$ **do**
  $U_i \leftarrow$ random number drawn from $N(U_{i-1}, 2D_{z_{i-1}}\Delta t_{i-1})$
end for
$\mathbf{U}^{(1)} \leftarrow \{U_i\}_{i=1}^{N+1}$
for $k = 2$ to $k = K$ **do**
  $\eta_0 \leftarrow \sum_{z_i=0} 1$
  $\eta_1 \leftarrow \sum_{z_i=1} 1$
  if $\eta_0 = 0$ **then**
    $D_0 \leftarrow$ random number drawn from Unif$(0, D_{max})$
  else
    while $D_0 > D_{max}$ **do**
      $D_0^{-1} \leftarrow$ random number drawn from Gamma $\left(\eta_0 - 1, \sum_{i|z_i=0} \frac{\Delta U_i^2}{4\Delta t_i}\right)$
      $D_0 \leftarrow 1/D_0^{-1}$
    end while
  end if
  if $\eta_1 = 0$ **then**
    $D_1 \leftarrow$ random number drawn from Unif$(0, D_{max})$
  else
    while $D_1 > D_{max}$ **do**
      $D_1^{-1} \leftarrow$ random number drawn from Gamma $\left(\eta_1 - 1, \sum_{i|z_i=1} \frac{\Delta U_i^2}{4\Delta t_i}\right)$
      $D_1 \leftarrow 1/D_1^{-1}$
    end while
  end if
  Update $n_{10}, n_{11}, n_{01}, n_{00}$ using Equation (2.15)
  $p_{01} \leftarrow$ random number drawn from Beta$\big(a_0 + n_{10}, b_0 + n_{11}\big)$
  $p_{10} \leftarrow$ random number drawn from Beta$\big(a_1 + n_{01}, b_1 + n_{00}\big)$
  $\theta^{(k)} \leftarrow \{D_0, D_1, p_{01}, p_{10}\}$
  for $i = 1$ to $i = N$ **do**
    Calculate $\pi(z_i = 1 | z_{i-1}, z_{i+1}, D_0, D_1, p_{01}, p_{10}, \mathbf{U})$ by normalising Equation (2.44), (2.45) or (2.46)
    $z_i \leftarrow$ random number drawn from Bernoulli$(\pi(z_i = 1 | z_{i-1}, z_{i+1}, D_0, D_1, p_{01}, p_{10}, \mathbf{U}))$
  end for
  $\mathbf{z^{(k)}} \leftarrow \{z_i\}_{i=1}^N$
  for $i = 1$ to $i = N+1$ **do**
    Calculate μ_i, τ_i from Equation (2.48), (2.49) or (2.50)
    $U_i \leftarrow$ random number drawn from $N(\mu_i, 1/\tau_i)$
  end for
  $\mathbf{U}^{(k)} \leftarrow \{U_i\}_{i=1}^{N+1}$
end for

Algorithm 8 Metropolis-Hastings sampler for one-state diffusion model with measurement noise incorporated as independent displacements (approximate model)

$K \leftarrow$ number of MCMC steps
$D_{max} \leftarrow$ choice of prior parameters
$\sigma^2 \leftarrow$ choice of localisation accuracy
$S_D \leftarrow$ variance of proposal distribution
$D \leftarrow$ random number drawn from Unif$(0, D_{max})$
$D^{(1)} \leftarrow D$
for $k = 2 : K$ **do**
  $D' \leftarrow$ random number drawn from $N(D, S_D)$
  $\alpha(D \to D') = \min\left\{1, \frac{\prod_{i=1}^{N} N\left(\Delta X_i; 0, 2(D'\Delta t_i + \sigma^2)\right)}{\prod_{i=1}^{N} N(\Delta X_i; 0, 2(D\Delta t_i + \sigma^2))}\right\} \mathbb{1}_{[0, D_{max}]}(D)$
  $u \leftarrow$ random number drawn from Unif$(0, 1)$
  if $\alpha(D \to D') > 1$ **then**
    $D \leftarrow D'$
  else if $\alpha(D \to D') > u$ **then**
    $D \leftarrow D'$
  end if
  $D^{(k)} \leftarrow D$
end for

Algorithm 9 Metropolis-within-Gibbs sampler for two-state diffusion model with measurement noise incorporated as independent displacements (approximate model)

$K \leftarrow$ number of MCMC steps
$D_{max}, a_0, b_0, a_1, b_1 \leftarrow$ choice of prior parameters
$\sigma^2 \leftarrow$ choice of localisation accuracy
$S_{D_0}, S_{D_1} \leftarrow$ variance of proposal distributions
$D_0 \leftarrow$ random number drawn from Unif$(0, D_{max})$
$D_1 \leftarrow$ random number drawn from Unif$(0, D_{max})$
$p_{01} \leftarrow$ random number drawn from Beta(a_0, b_0)
$p_{10} \leftarrow$ random number drawn from Beta(a_1, b_1)
$\theta^{(1)} \leftarrow \{D_0, D_1, p_{01}, p_{10}\}$
$z_1 \leftarrow$ random number drawn from Bernoulli$\left(\frac{p_{10}}{p_{10}+p_{01}}\right)$
for $i = 2$ to $i = N$ **do**
    $z_i \leftarrow$ random number drawn from Bernoulli$(z_{i-1}(1-p_{10}) + (1-z_{i-1})p_{01})$
end for
$\mathbf{z} \leftarrow \{z_i\}_{i=1}^N$
for $k = 2$ to $k = K$ **do**
    $D_0' \leftarrow$ random number drawn from $N(D_0, S_{D_0})$
    $\alpha(D_0 \to D_0'|\mathbf{z}, \mathbf{X}) \leftarrow \min\left\{1, \frac{\prod_{z_i=0} N(\Delta X_i; 0, 2D_0'\Delta t_i + 2\sigma^2)}{\prod_{z_i=0} N(\Delta X_i; 0, 2D_0\Delta t_i + 2\sigma^2)}\right\} \mathbb{1}_{[0,D_{max}]}(D_0)$
    $u \leftarrow$ random number drawn from Unif$(0,1)$
    if $\alpha(D_0 \to D_0'|\mathbf{z}, \mathbf{X}) > 1$ **then**
        $D_0 \leftarrow D_0'$
    else if $\alpha(D_0 \to D_0'|\mathbf{z}, \mathbf{X}) > u$ **then**
        $D_0 \leftarrow D_0'$
    end if
    $D_1' \leftarrow$ random number drawn from $N(D_1, S_{D_1})$
    $\alpha(D_1 \to D_1'|\mathbf{z}, \mathbf{X}) \leftarrow \min\left\{1, \frac{\prod_{z_i=1} N(\Delta X_i; 0, 2D_1'\Delta t_i + 2\sigma^2)}{\prod_{z_i=1} N(\Delta X_i; 0, 2D_1\Delta t_i + 2\sigma^2)}\right\} \mathbb{1}_{[0,D_{max}]}(D_1)$
    $u \leftarrow$ random number drawn from Unif$(0,1)$
    if $\alpha(D_1 \to D_1'|\mathbf{z}, \mathbf{X}) > 1$ **then**
        $D_1 \leftarrow D_1'$
    else if $\alpha(D_1 \to D_1'|\mathbf{z}, \mathbf{X}) > u$ **then**
        $D_1 \leftarrow D_1'$
    end if
    Update $n_{10}, n_{11}, n_{01}, n_{00}$ using Equation (2.15)
    $p_{01} \leftarrow$ random number drawn from Beta$\big(a_0 + n_{10}, b_0 + n_{11}\big)$
    $p_{10} \leftarrow$ random number drawn from Beta$\big(a_1 + n_{01}, b_1 + n_{00}\big)$
    $\theta^{(k)} \leftarrow \{D_0, D_1, p_{01}, p_{10}\}$
    for $i = 1$ to $i = N$ **do**
        Calculate $\pi(z_i = 1|z_{i-1}, z_{i+1}, D_0, D_1, p_{01}, p_{10}, \mathbf{X})$ by normalising Equation (A.2), (A.3) or (A.4)
        $z_i \leftarrow$ random number drawn from Bernoulli$(\pi(z_i = 1|z_{i-1}, z_{i+1}, D_0, D_1, p_{01}, p_{10}, \mathbf{X}))$
    end for
    $\mathbf{z^{(k)}} \leftarrow \mathbf{z}$
end for

A.4 SI Tables

Table A.1: **Model selection results for different Bayes factor thresholds.**

Treatment	DMSO	Cyto D	PMA	PMA+Cal-I
Number of trajectories	75	36	39	46
Two-state model preferred, BF threshold 0[1]	18/75 (24%)	11/36 (31%)	14/39 (36%)	8/46 (17%)
Two-state model preferred, BF threshold 0[1], fast switchers removed[2]	13/70 (19%)	3/28 (11%)	8/33 (24%)	6/44 (14%)
Two-state model preferred, BF threshold ±5[3]	15/66 (23%)	7/27 (26%)	13/36 (36%)	7/42 (17%)
Two-state model preferred, BF threshold ±5[3], fast switchers removed[2]	12/63 (19%)	3/23 (13%)	8/31 (26%)	5/40 (13%)

1. Model selection between approximate one-state and two-state diffusion models with measurement noise, with $\log_e B_{1D,2D} < 0$ preference for the two-state model and $\log_e B_{1D,2D} > 0$ preference for the one-state model.
2. Fast switching trajectories ($\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$) also removed.
3. Model selection between approximate one-state and two-state diffusion models with measurement noise, with $\log_e B_{1D,2D} < -5$ preference for the two-state model and $\log_e B_{1D,2D} > 5$ preference for the one-state model.

Table A.2: **Influence of noise and S/N on model selection preferences between one-state and two-state diffusion models on LFA-1 data (157 trajectories).**

Models	**No preference**	**One-state diffusion**	**Two-state diffusion**
Approximate measurement noise	19 (9.7%)	132 (67%)	45 (23%)
No measurement noise and subsampling	15 (7.7%)	130 (66%)	51 (26%)
No measurement noise	11 (5.6%)	85 (43%)	100 (51%)

Approximate measurement noise refers to model selection between one-state and two-state diffusion models with measurement noise incorporated as Equations (2.27) and (2.39) (with $\sigma^2 = 41.09\text{nm}^2$). No measurement noise is model selection between one-state and two-state diffusion models without measurement noise, posterior as Equations (2.3) and (2.10). No measurement noise and subsampling is model selection between one-state and two-state diffusion models without measurement noise, as Equations (2.3) and (2.10), but on trajectories subsampled at a rate equal to the optimum number of MSD points for estimating D, see Section A.2. For all models the MCMC runs were 20000 steps with a 10000 step burn-in. See Section 2.2 for priors and initial conditions.

Table A.3: **Comparison of model selection for approximate measurement noise models and subsampled trajectories.**

		Approximate measurement noise		
		No preference	One-state diffusion	Two-state diffusion
No measurement noise and subsampling	No preference	5	6	4
	One-state diffusion	13	108	9
	Two-state diffusion	1	18	32

For approximate measurement noise the model selection was between one-state and two-state diffusion models with measurement noise, described in Methods. For no measurement noise and subsampling the model selection was between one-state and two-state diffusion models without measurement noise, as described in Methods but with trajectories subsampled at a rate equal to optimum number of MSD points to include when estimating D, see Section A.2. For all models the MCMC runs were 20000 steps with a 10000 step burn-in. See Section 2.2 for priors and initial conditions.

A.5 SI Figures

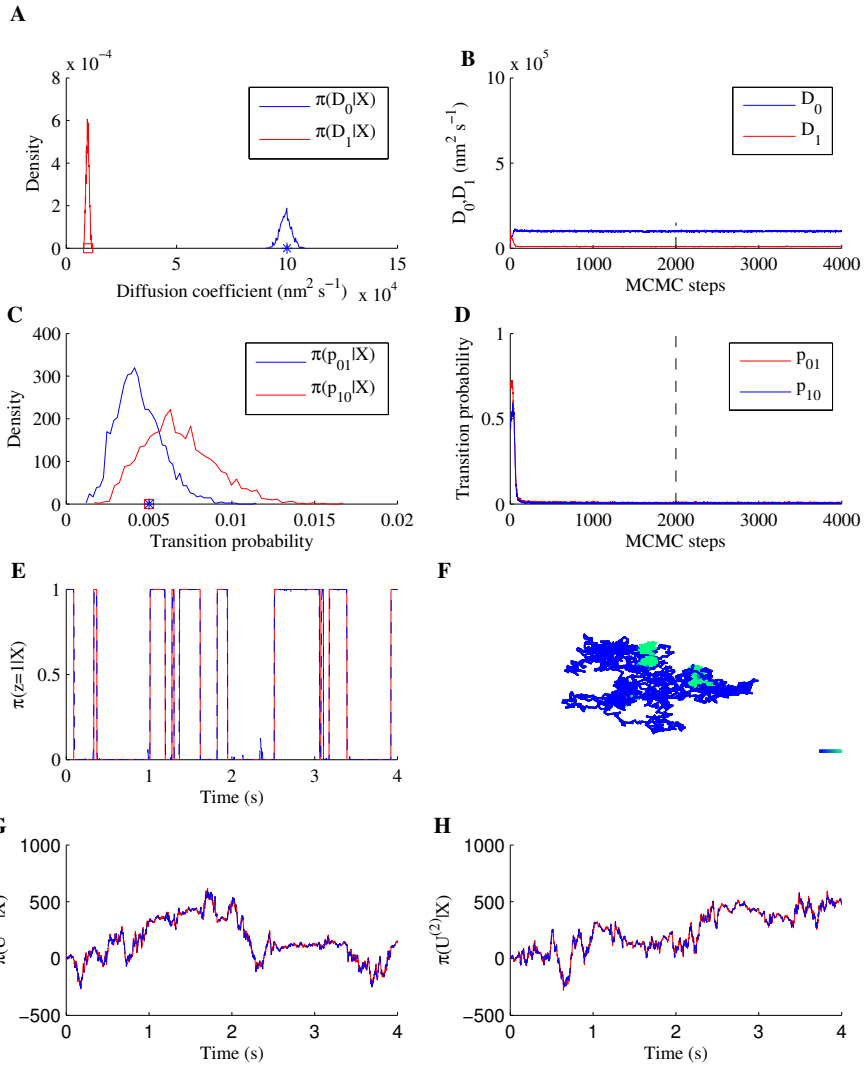

Figure A.1: **Fit of the exact two-state diffusion model with measurement noise to a simulated two-state diffusion trajectory.** (A) The posteriors for the two diffusion coefficients with true D_1 (red square) and D_0 (blue asterisk) values plotted, (B) corresponding samples for D_0 (red) and D_1 (blue) including burn-in (dashed line). (C) Posteriors for the switching probabilities per frame, with true p_{01} (blue asterisk) and p_{10} (red square) values plotted (D) corresponding samples for p_{01} (red) and p_{10} (blue) including burn-in (dashed line). (E) Diffusion state inference (blue, dashed) and true state (red) shown as the probability of being in the low diffusion state. (F) Trajectory coloured by the probability of being in the low diffusion state. Colour scale represents $\pi(\mathbf{z} = 1|\mathbf{X})$ from 0 (blue, high diffusion state) to 1 (green, low diffusion state). Colorbar length: 100nm. (G,H) Mean inferred position of $\mathbf{U}$ (blue, dashed) and true particle position (red). Trajectory was simulated as Equation (2.39). Simulated measurement noise and measurement noise for inference both set to $\sigma^2 = 41.09$nm. Data from 20000 MCMC steps with a 10000 step burn-in. See Section 2.2 for priors and initial conditions.

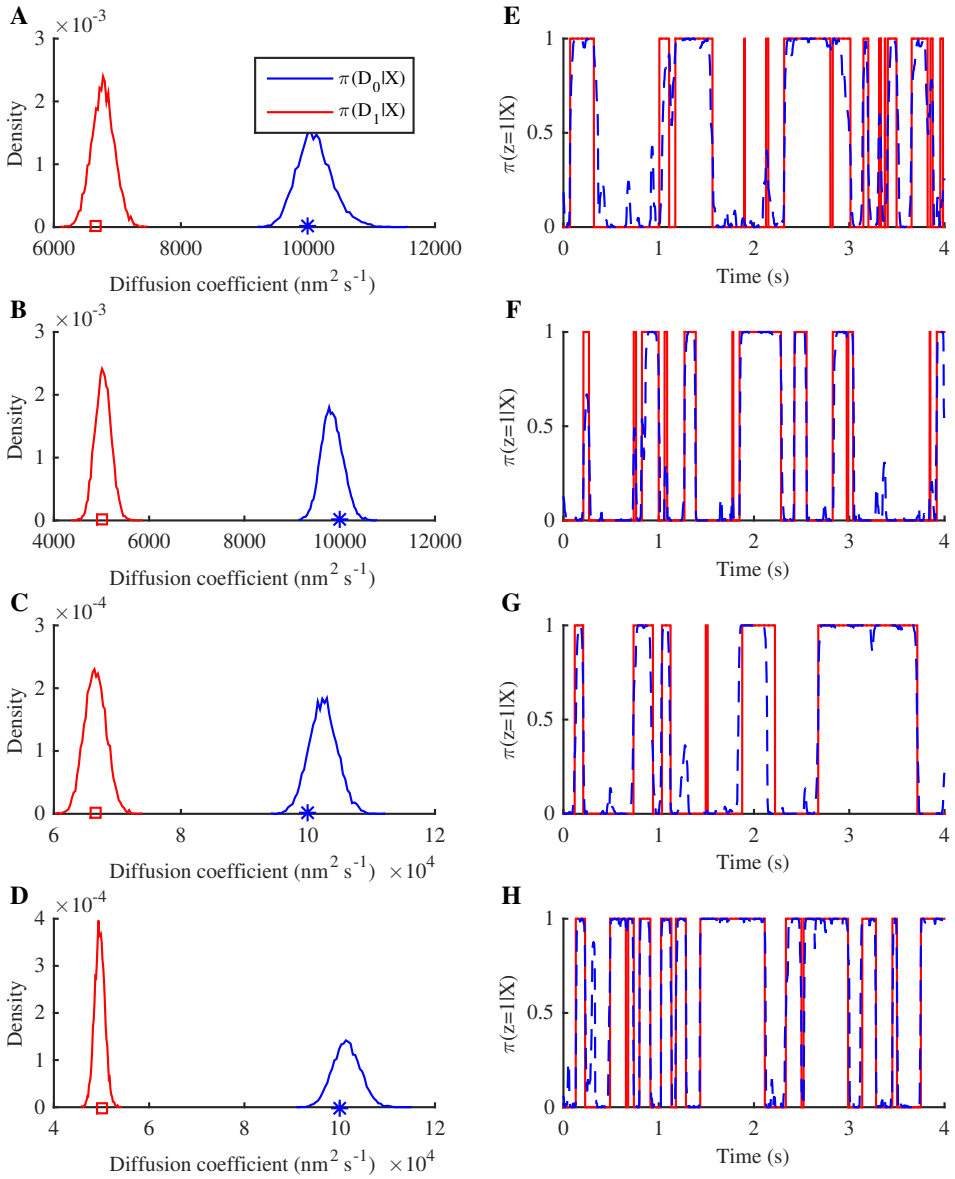

Figure A.2: **Diffusion coefficients separated by a factor of 1.5 can be detected on the exact two-state diffusion model with measurement noise.** (A-D) The posteriors for the two diffusion coefficients with true D_1 (red square) and D_0 (blue asterisk) values plotted; true D_0, D_1 differ by a factor of 1.5 (A,C) and 2 (B,D), with low (A,B) and high (C,D) diffusion coefficients. (E-H) Corresponding diffusion state inference (blue, dashed) and true state (red) shown as the probability of being in the low diffusion state. The transition probabilities for all trajectories were $p_{01} = 0.005$, $p_{10} = 0.005$. Measurement noise set to $\sigma^2 = 41.09$nm for both the simulated data and inference algorithm. Trajectories comprise 4000 frames. Data from 20000 MCMC steps with a 10000 step burn-in. See Section 2.2 for priors and initial conditions.

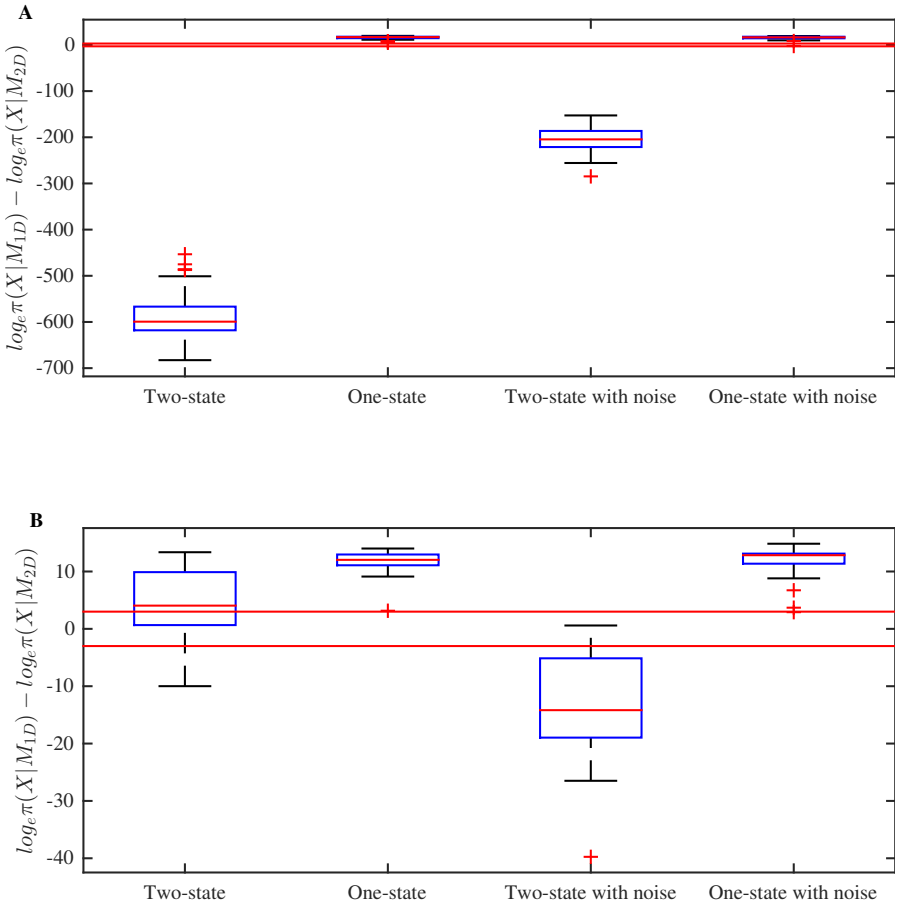

Figure A.3: **Model selection between one-state and two-state diffusion models with measurement noise on simulated trajectories.** Box and whisker plots of log Bayes factors for simulated datasets (trajectories are length 4 s with 1000 frames s^{-1}). Trajectories with log Bayes factor outside 1.5 times IQR are plotted as outliers (red crosses). Red lines correspond to the log Bayes factor thresholds of -3 and 3 (see Section 2.2.7). (A) Parameters for simulated data (50 trajectories for each model): two-state and two-state with noise, $D_0 = 10^5$ nm^2s^{-1}, $D_1 = 2 \times 10^4$ nm^2s^{-1}, $p_{01} = 0.01, p_{10} = 0.01$; one-state and one-state with noise, $D = 10^5$ nm $^2s^{-1}$. (B) Parameters for simulated data (20 trajectories for each model): two-state and two-state with noise, $D_0 = 5 \times 10^5$ nm^2s^{-1}, $D_1 = 2 \times 10^4$ nm^2s^{-1}, $p_{01} = 0.01, p_{10} = 0.01$; one-state and one-state with noise, $D = 5 \times 10^5$ nm^2s^{-1}. Measurement noise in the simulations was $\sigma^2 = 41.09$ nm. MCMC runs were 20000 steps with a 10000 step burn in, with measurement noise fixed as $\sigma^2 = 41.09$ nm.

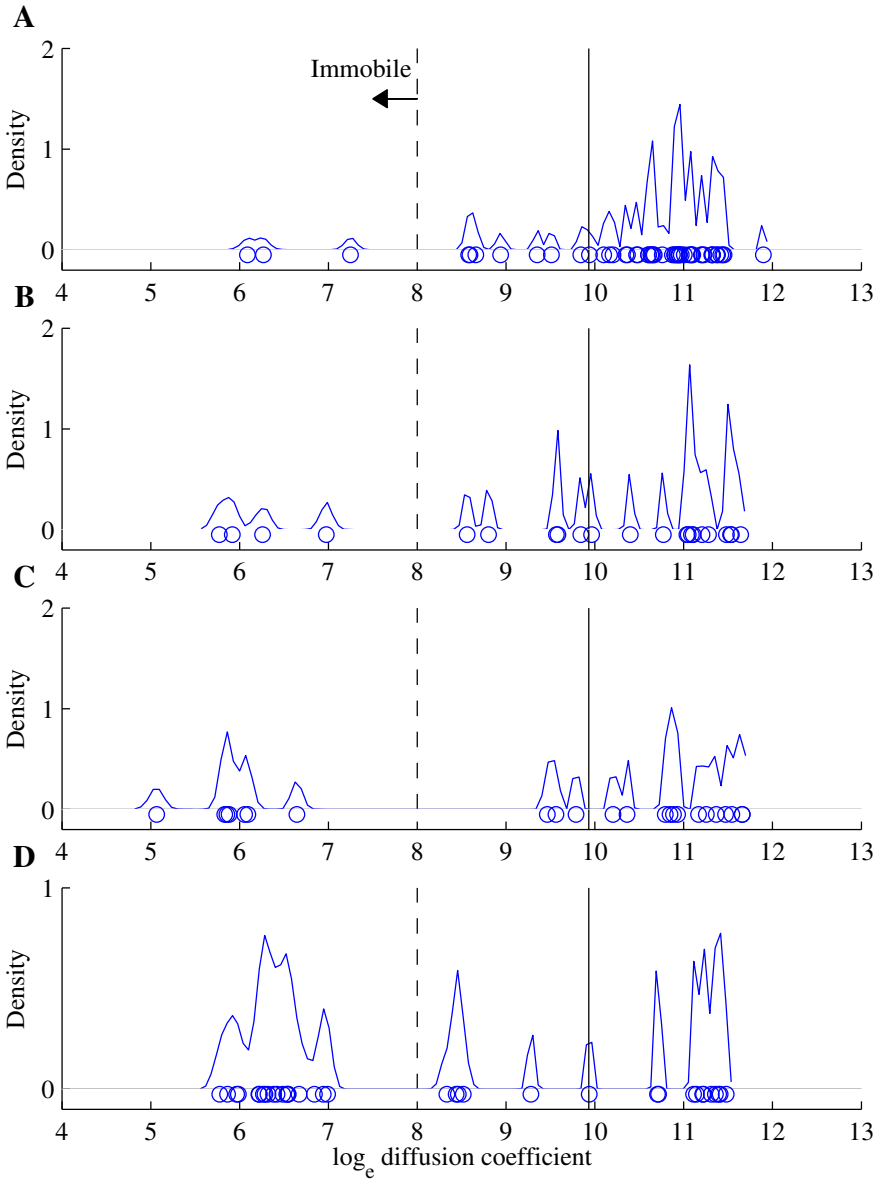

Figure A.4: **Posterior estimates of diffusion coefficients for single LFA-1 trajectories.** (A-D) Pooled posterior samples of $\log_e D$ for trajectories preferring the one-state diffusion model. The posterior means (blue circles) are also shown. Black line indicates value of $\sigma^2/2\Delta t$. Treatments: (A) DMSO, one-state model preferred for 51 trajectories; (B) Cyto D, 22 trajectories; (C) PMA, 23 trajectories; (D) PMA+Cal-I, 36 trajectories.

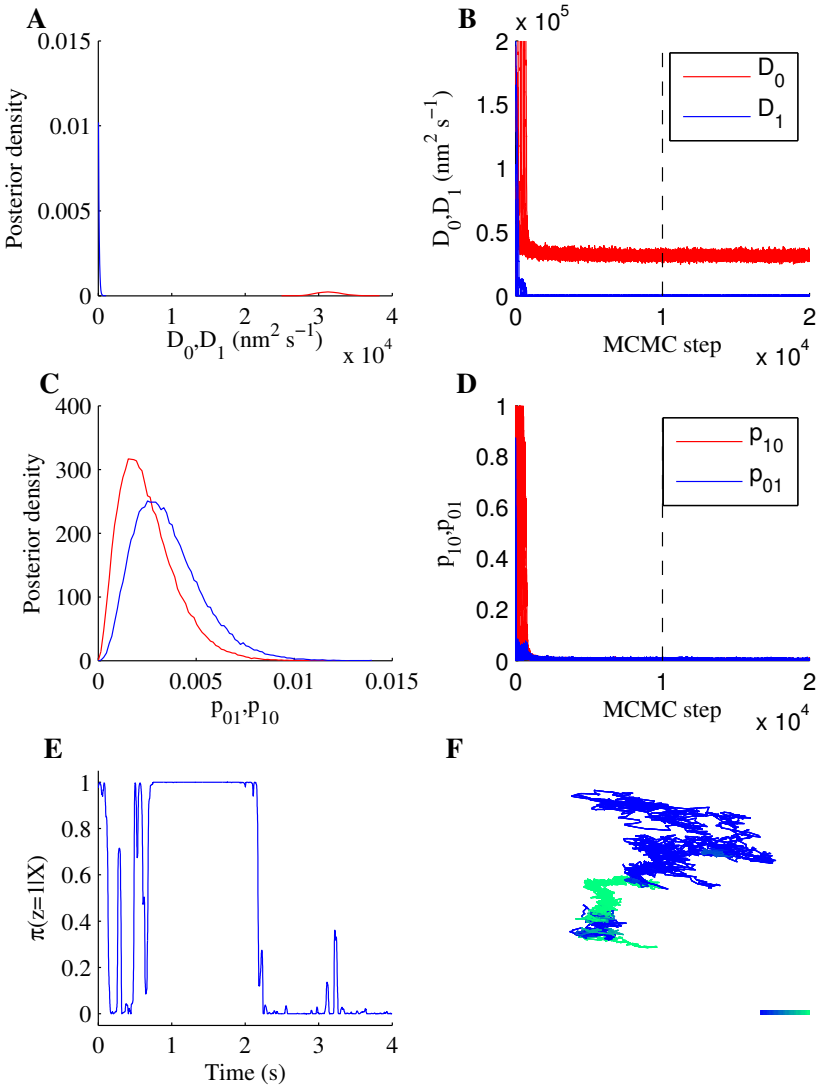

Figure A.5: **Fit of the approximate two-state diffusion model with measurement noise to an LFA-1 trajectory (PMA+Cal-I treatment).** Compare to Fig. 2.3 fitting the exact noise model to the same trajectory. (A) The posteriors for the two diffusion coefficients, (B) corresponding samples (12 independent chains plotted in the same colour) for D_0 and D_1 including burn-in (dashed line), (C) posteriors for the switching probabilities per frame, (D) corresponding samples (12 chains) for p_{01} and p_{10} including burn-in (dashed line), (E) State inference shown as the probability of being in the low diffusion state, (F) trajectory coloured by the probability of being in the low diffusion state. Colour scale represents $\pi(\mathbf{z} = 1|\mathbf{X})$ from 0 (blue, high diffusion state) to 1 (green, low diffusion state). Colorbar length: 100nm. Data from 12 parallel chains of 20000 MCMC steps with a 10000 step burn-in. Priors, see Section 2.2.

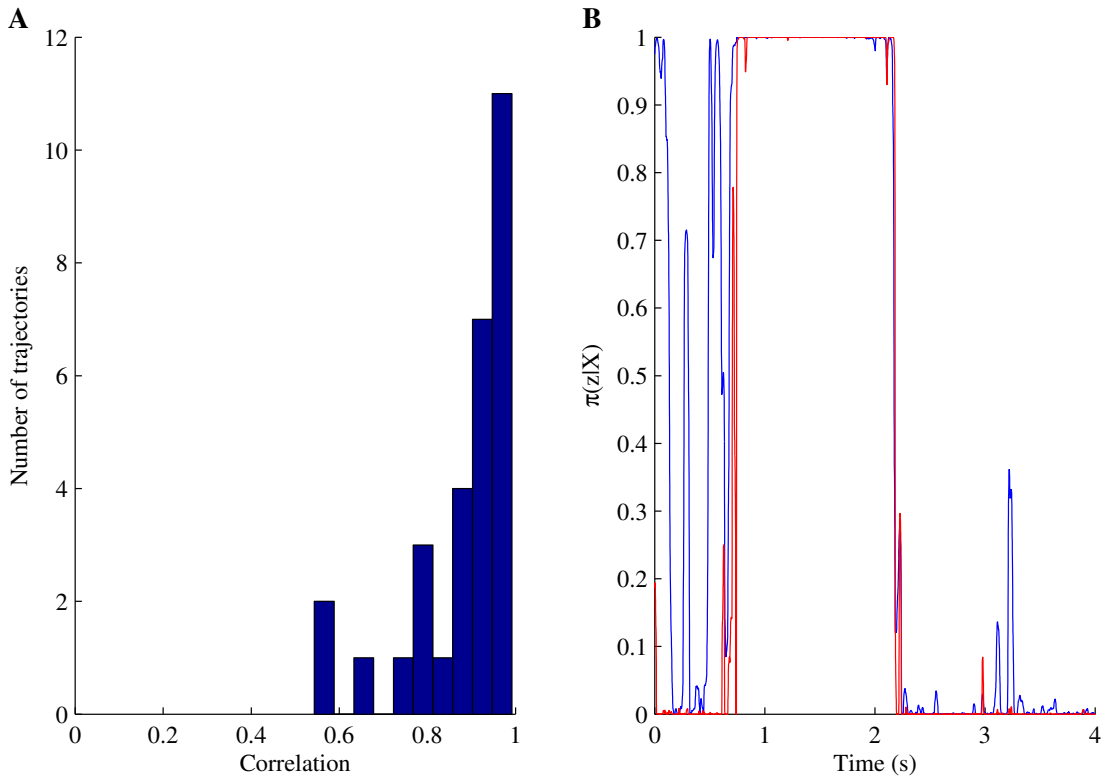

Figure A.6: **Comparison of hidden state inference for the exact and approximate two-state diffusion models with measurement noise.** (A) Correlation between inferred hidden state $\mathbf{z}$ for each model, pooled across all treatments for 30 trajectories preferring the approximate two-state model (fast-switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$ in the exact model, trajectories removed). (B) Example hidden state posterior for approximate two-state model (blue) and exact two-state model (red) for a single trajectory (PMA+Cal-I treatment).

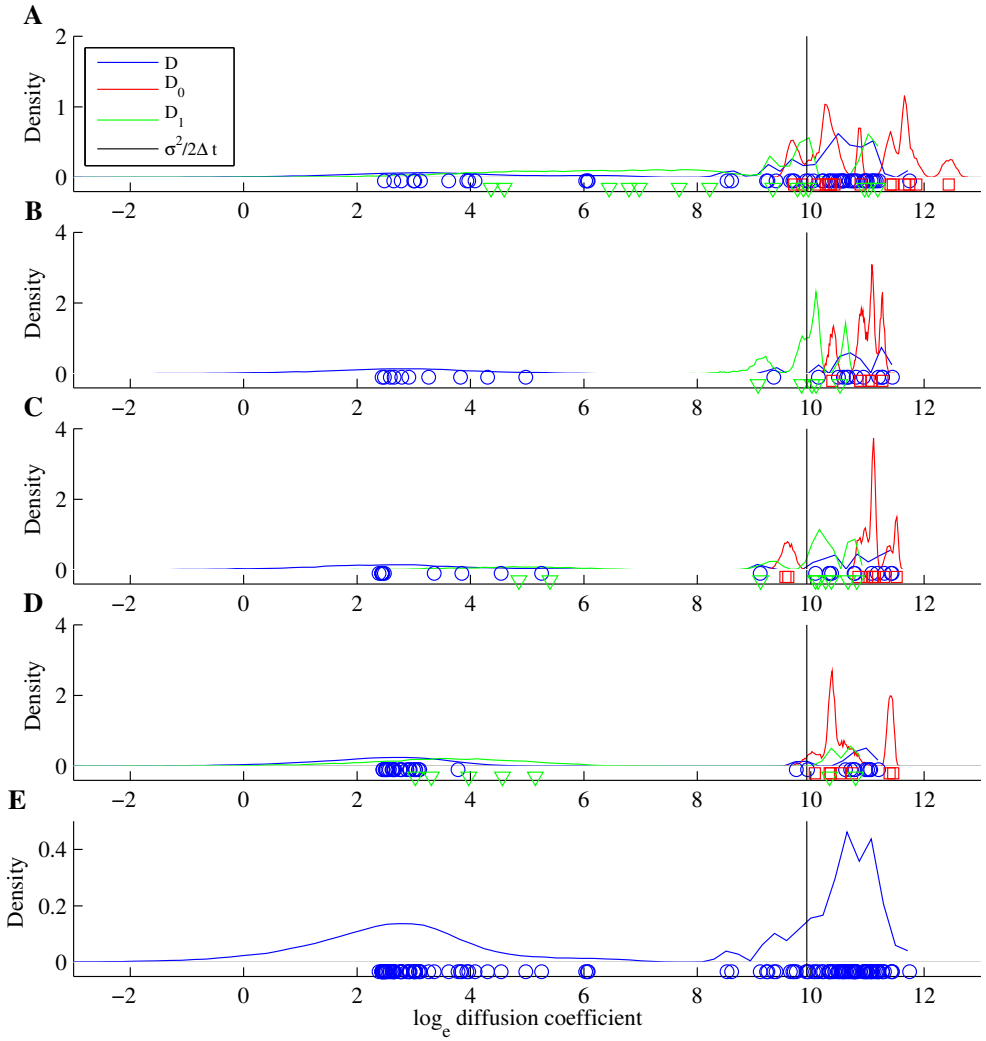

Figure A.7: **Posterior estimates of diffusion coefficients from fitting approximate two-state diffusion model with measurement noise to LFA-1 trajectories.** (A-D) Pooled posterior samples of $\log_e D$ (blue lines) for trajectories with one-state diffusion model preference, or $\log_e D_0$ (red lines) and $\log_e D_1$ (green lines) for trajectories with two-state diffusion model preference (fast switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$, trajectories removed). Also plotted are the posterior means of $\log_e D$ for each trajectory with one-state model preference (blue circles), and posterior means of $\log_e D_0$ (red squares) and $\log_e D_1$ (green triangles) for each trajectory with two-state model preference. Black line indicates value of $\sigma^2/2\Delta t$. Treatments: (A) DMSO, one-state model preferred for 51 trajectories, two-state model preferred for 14 trajectories; (B) Cyto D, 22 one-state, 5 two-state; (C) PMA 23 one-state, 14 two-state (D) PMA+Cal-I, 36 one-state, 7 two-state. (E) Pooled $\log_e D$ estimates and posterior means for each trajectory over all treatments for trajectories where one-state diffusion model was preferred.

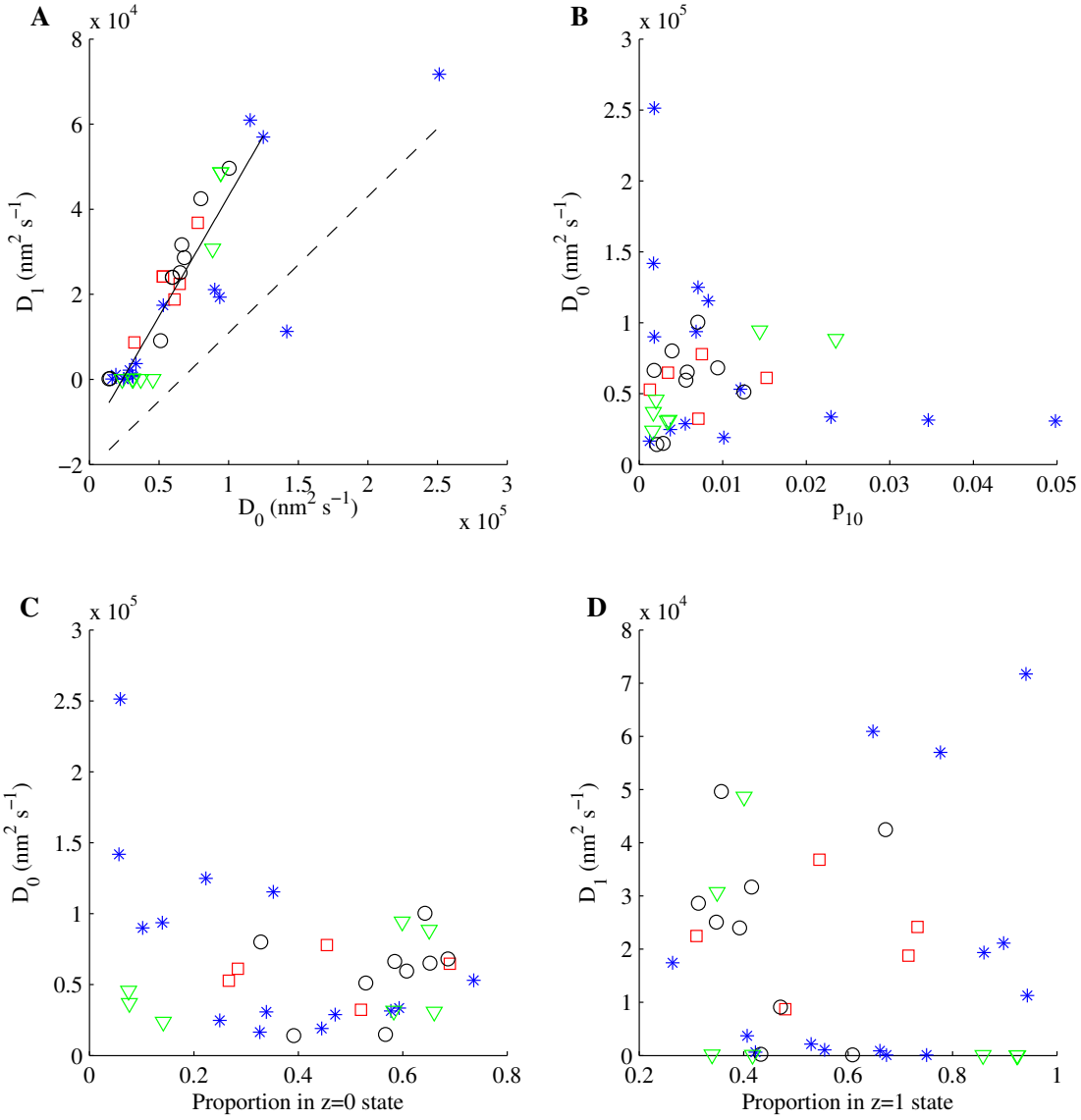

Figure A.8: **Dependences of parameter estimates from approximate two-state diffusion model with measurement noise.** (A-D) Scatter plots of posterior means of stated parameters for approximate two-state model with measurement noise inference, for trajectories where the approximate two-state diffusion model was preferred, (fast switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$, trajectories removed). Treatments: DMSO, blue asterisks; Cyto D, red squares; PMA black circles; PMA+Cal-I, green triangles. In panel (A) the black solid line is a linear fit with two outlier trajectories removed, $D_1 = aD_0 + b, a = 0.57, b = -1.3 \times 10^4$ nm^2 s^{-1}; black dashed line is double iterate, $D_1 = a(aD_0 + b) + b$.

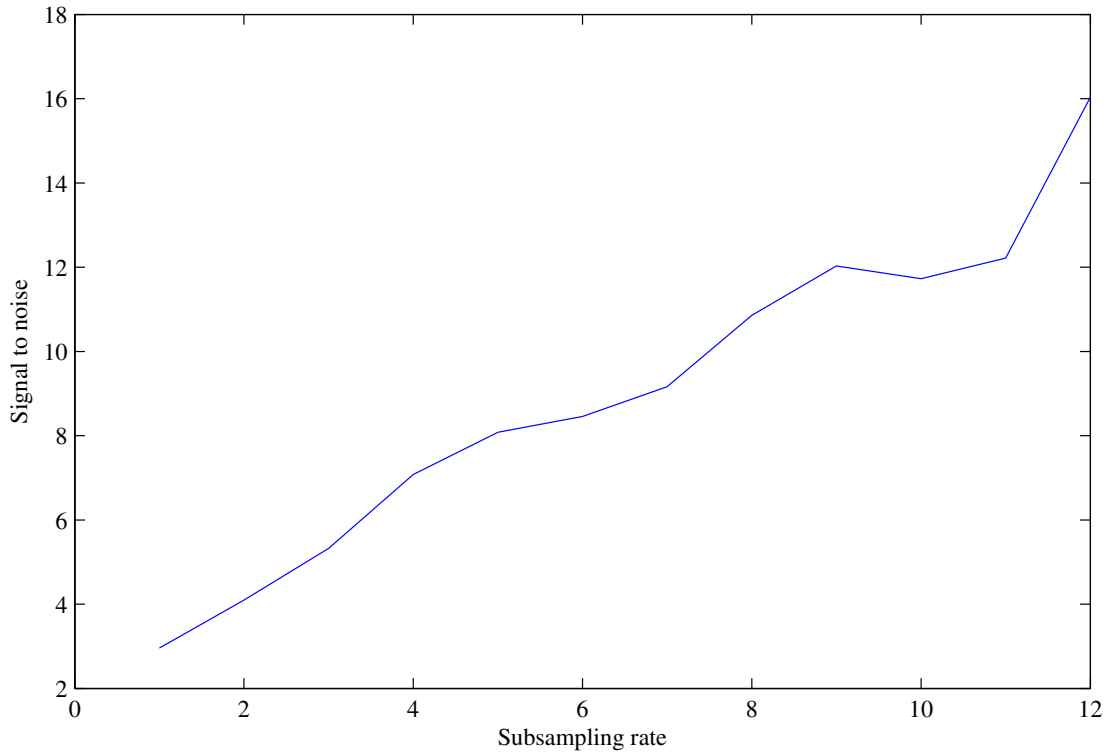

Figure A.9: **Signal to noise against subsampling rate for LFA-1 trajectories.** The "signal" is the average variance of individual displacements of LFA-1 trajectories over all treatments, and the "noise" is the average variance of individual displacements for three stationary latex bead trajectories. Trajectories were subsampled at rate n by including only every nth timepoint, giving a trajectory of length $\lfloor \frac{N}{n} \rfloor$.

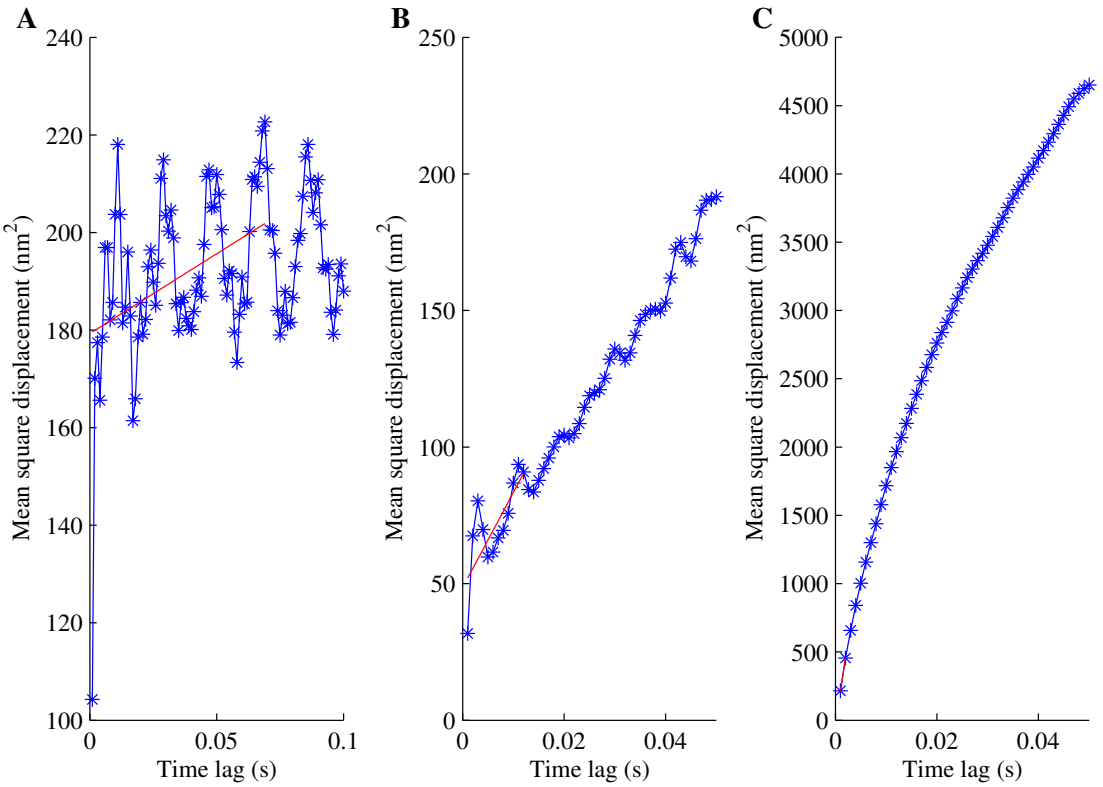

Figure A.10: **Mean square displacement plots for three SPT trajectories.** Red line is the straight line fit to the optimum number of MSD points to use when estimating the diffusion coefficient D [62]. (A) Stationary latex bead. (B) Slow diffusing LFA-1 trajectory (PMA+Cal-I treatment). (C) Fast diffusing LFA-1 trajectory (PMA+Cal-I treatment).

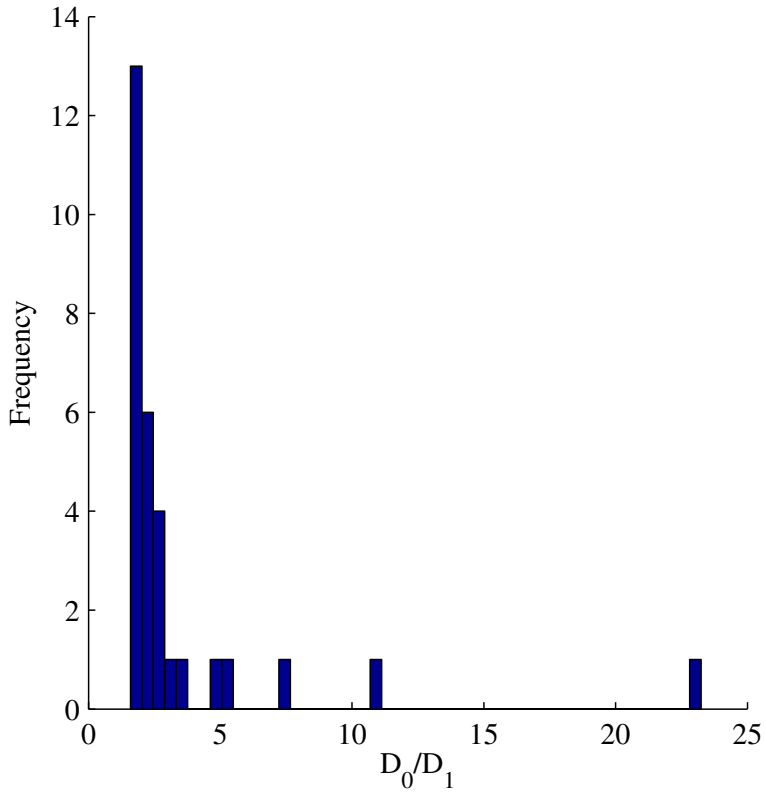

Figure A.11: **Posterior estimates of D_0/D_1 ratio for the two-state diffusion model with measurement noise fitted to LFA-1 trajectories.** Posterior mean D_0/D_1, for trajectories where two-state diffusion model was preferred, pooled across treatments (fast switching, $\hat{p}_{01} > 0.1$ or $\hat{p}_{10} > 0.1$, trajectories removed).

A.6 SI Files: Single particle tracking trajectories in MAT and HDF5 file formats

Available at http://tinyurl.com/DetectionofDiffusion.

Appendix B

Supporting Information for Chapter 3

B.1 MSD Analysis and Optimal Resampling

The iSCAT dataset contains trajectories sampled at 50 kHz. An alternative to modelling single displacements is to subsample the data, i.e. model displacements over n time steps (rather than a single time step), thereby increasing the signal to noise (S/N). Here, we use two approaches to determine a reasonable subsampling rate; both approaches suggested a subsampling rate of around 10. We also describe an additional preprocessing step; this was necessary to remove clear artifacts from a minority of trajectories.

B.1.1 Subsampling of trajectories

We used the mean square displacement (MSD) analysis of Michalet [62] to analyse each trajectory. This analysis also infers the optimum number of points to use in the MSD calculation, Table B.1. This is relevant for our methods as it determines whether modelling single displacements is reasonable, i.e. whether S/N is sufficiently high.

Table B.1: **Mean square displacement statistics for AuNP/CTxB/GM1 trajectories over two treatments.**

AuNP size (nm)	Membrane composition	Substrate	D ($\mu m^2\,s^{-1}$)	Optimum number of MSD Points
20	0.03% GM1	glass	0.34 ± 0.03	6-43
40	0.03% GM1	mica	1.05 ± 0.09	2

Over a set of 71 trajectories for 20nm AuNPs and 18 trajectories for 40nm AuNPs. The MSD displacement analysis of Michalet [62] was run on the full (i.e. not subsampled) trajectories, but with the trajectories truncated to remove clear artifactual displacements (Section B.1.3).

Trajectories with 20nm AuNP show a variety of optimum number of MSD points (Table B.1) indicating that subsampling the data is necessary. Since the number of MSD points for estimating D ranges from 6-43 with mean 11.6, for simplicity we subsampled all trajectories at rate 10. The time between observations is thus 0.0002s, which we call the "subsampled time step". We could alternatively subsample at a rate equal to the optimum number of MSD points on a trajectory by trajectory basis, although results were similar on the trajectories tested.

B.1.2 Directionality of diffusion

For trajectories with 40nm AuNP, the number of points included when estimating D is always two, i.e. a fit to the first two MSD points, not including the origin. This suggests that the S/N is high enough for a single displacement analysis. However for 40nm AuNPs, at time delays less than 10^{-4}s, superdiffusive behaviour was reported (attributed to dynamic error in measurements due to sub-nanometre localisation precision) [51].

We investigated this, and noticed that displacement angles showed a bias towards horizontal displacements (Fig. B.1). This was statistically significant in a Chi-squared test (with the null hypothesis that angular displacements follow a uniform distribution), Fig. B.2. This bias was resolved, in 40nm AuNP/CTxB/GM1 on mica trajectories, by subsampling at rate 10 (Fig. B.2A). These trajectories were previously shown to display no trapping [51]; hence the remaining bias (even when subsampling at higher rates) in displacement angles for other treatments is presumably due to directional displacements in trapping events. (We also investigated the effect of window averaging, taking the average particle position over a window size n, and found similar trends to subsampling, Fig. B.2B.) Thus, this analysis is further evidence for a subsampling rate of 10, and also suggests that we should subsample 40nm AuNP trajectories.

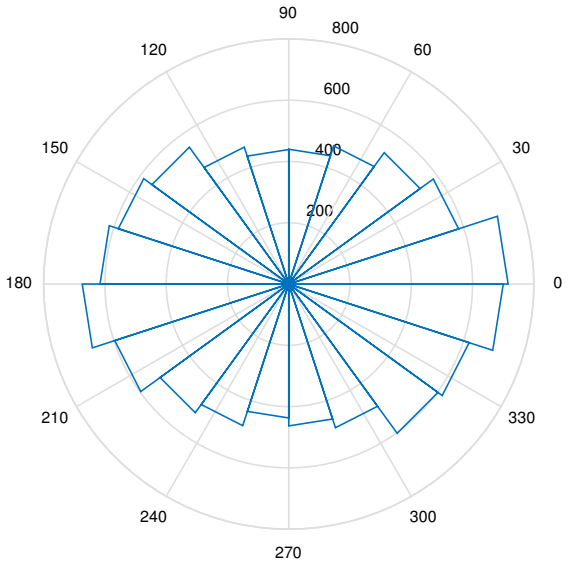

Figure B.1: **Radial histogram of angular displacements for 40nm AuNP/CTxB/GM1 on mica trajectory.** The angular displacement (plotted in degrees) in radians is $R_i = \text{atan2}(\Delta X_{1i}, \Delta X_{2i}) + \pi$, where $\Delta X_i = \{\Delta X_{1i}, \Delta X_{2i}\}$.

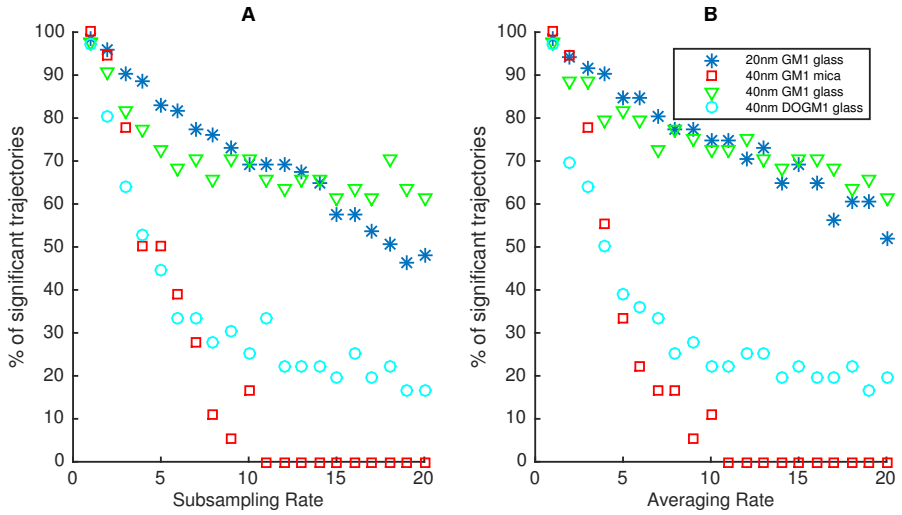

Figure B.2: **GM1 trajectory displacements have an angular bias.** Percentage of trajectories where the null hypothesis (angular displacements follow a uniform distribution) was rejected in a Chi-square test, plotted against subsampling or averaging rate. For each trajectory, the null hypothesis was rejected if $p < 0.0003$ (p=0.05 with Bonferroni correction, 169 trajectories in total). A) Subsampling, B) averaging.

B.1.3 Preprocessing of trajectories

Some 20nm AuNP/CTxB/GM1 on glass trajectories (7 out of 71) had clear artifactual displacements, probably caused by an additional AuNP in the focal area. We dealt with this by visually inspecting the trajectories, and removing (before subsampling) the section of the trajectory with the artifacts.

B.2 Gaussian and ring-like confinements

It was previously shown, using goodness of fit to a Rayleigh distribution [51], that AuNP/CTxB/GM1 trajectories had both Gaussian and ring-like confinement events. We repeated this analysis on 20nm and 40nm AuNP/CTxB/GM1 on glass trajectories, by first splitting each trajectory into 0.01s intervals. For each interval, if the particle was restricted to a region of 0.1 μm in both the x and y directions then we considered it confined. We thus split each trajectory into confined and free intervals. For each confined interval of length N, we calculated the radius, $R_i = \|\Delta X_i\|$, for all timepoints to give $\{R_i\}_{i=1}^{N}$. We then fit a Rayleigh distribution to $\{R_i\}_{i=1}^{N}$ values, and calculated the coefficient of determination, R^2. This statistic gives a measure of how Gaussian-like the confinement events are. We calculated a Gaussian-like statistic for individual trajectories by averaging R^2 values over all the confinement intervals, weighted by confinement length. The mean value of this statistic over all trajectories was 0.96 for 20nm AuNP/CTxB/GM1 on glass and 0.94 for 40nm AuNP/CTxB/GM1 on glass. (In the previous study 0.95 was used as a threshold for Gaussian confinement [51].) Our algorithm is suited to inferring Gaussian confinement events, so we thus applied the algorithm only to 20nm AuNP/CTxB/GM1 on glass trajectories in this study.

B.3 Full derivation of Gibbs move for harmonic well centre

Here we calculate a Gibbs update for the harmonic well centre ($\mathbf{C}$). As described in Methods, we update in blocks of length n, $\mathbf{C}_{j,n} = \{C_i\}_{i=j}^{j+n}$, where $1 \leq j \leq N - n$. We derive an update by comparing the conditional distribution for a block (obtained from the posterior distribution, Equation (3.6) in the main text), and the multivariate normal PDF. On the one hand we have the conditional distribution for $\mathbf{C}_{j,n}$, for which there are

two cases

$$\pi(\mathbf{C}_{j,n}|\mathbf{C}_{-j,n},\mathbf{z},\theta,\mathbf{X})\Big|_{j\neq 1} \propto \exp\left(\sum_{i=j}^{j+n}\frac{-\left(\Delta X_i+\kappa\Delta t_i z_i(X_i-C_i)\right)^2}{4D\Delta t_i}+\sum_{i=j}^{j+n+1}\frac{-\Delta C_i^2}{4\Delta t_i(D_C z_i+\hat{D}(1-z_i))}\right) \tag{B.1}$$

$$\pi(\mathbf{C}_{j,n}|\mathbf{C}_{-j,n},\mathbf{z},\theta,\mathbf{X})\Big|_{j=1} \propto \exp\left(\frac{-\tau_{C_1}}{2}\left(C_1-\mu_{C_1}\right)^2+\sum_{i=j}^{j+n}\frac{-\left(\Delta X_i+\kappa\Delta t_i z_i(X_i-C_i)\right)^2}{4D\Delta t_i} + \sum_{i=j}^{j+n+1}\frac{-\Delta C_i^2}{4\Delta t_i(D_C z_i+\hat{D}(1-z_i))}\right) \tag{B.2}$$

where $\mathbf{C}_{-j,n}=\mathbf{C}\setminus\mathbf{C}_{j,n}$. On the other hand, the multivariate normal PDF with mean $\boldsymbol{\mu}_{C_{j,n}}$ and precision matrix $\boldsymbol{\Sigma}^{-1}_{C_{j,n}}$, up to proportionality with respect to $\mathbf{C}_{j,n}$, is

$$\begin{aligned}\exp\left[-\frac{1}{2}\left((\mathbf{C}_{j,n}-\boldsymbol{\mu}_{C_{j,n}})^T\boldsymbol{\Sigma}^{-1}_{C_{j,n}}(\mathbf{C}_{j,n}-\boldsymbol{\mu}_{C_{j,n}})\right)\right] &= \exp\left[-\frac{1}{2}\sum_{l=j}^{j+n}\sum_{m=j}^{j+n}(C_l-\mu_l)(C_m-\mu_m)\Sigma^{-1}_{l,m}\right]\\ &= \exp\left[-\frac{1}{2}\sum_{l=j}^{j+n}\sum_{m=j}^{j+n}(C_lC_m-C_l\mu_m-C_m\mu_l+\mu_l\mu_m)\Sigma^{-1}_{l,m}\right]\end{aligned} \tag{B.3}$$

where $\Sigma^{-1}_{l,m}$ denotes the (l,m)th element of the precision matrix and μ_l denotes the lth element of the mean vector. Thus we can calculate a multivariate normal update by comparing the coefficients in the exponential for Equations (B.1) and (B.3) (also using Equation (B.2) if $j=1$). For the squared and cross terms this gives

$$C_i^2\Big|_{2\leq i\leq N-1}: -\frac{1}{4\Delta t_{i-1}(D_C z_{i-1}+\hat{D}(1-z_{i-1}))}-\frac{\kappa^2\Delta t_i z_i}{4D}-\frac{1}{4\Delta t_i(D_C z_i+\hat{D}(1-z_i))}=-\frac{1}{2}\Sigma^{-1}_{i,i} \tag{B.4}$$

$$C_iC_{i-1}: \frac{1}{2\Delta t_{i-1}(D_C z_{i-1}+\hat{D}(1-z_{i-1}))}=-\Sigma^{-1}_{i,i-1}=-\Sigma^{-1}_{i-1,i}. \tag{B.5}$$

And for $i = 1$ and $i = N$ we have

$$C_1^2 : -\frac{\tau_{C_1}}{2} - \frac{\kappa^2 \Delta t_1 z_i}{4D} - \frac{1}{4\Delta t_1 (D_C z_1 + \hat{D}(1 - z_1))} = -\frac{1}{2}\Sigma_{1,1}^{-1} \tag{B.6}$$

$$C_N^2 : -\frac{1}{4\Delta t_{N-1}(D_C z_{N-1} + \hat{D}(1 - z_{N-1}))} - \frac{\kappa^2 \Delta t_N z_N}{4D} = -\frac{1}{2}\Sigma_{N,N}^{-1}. \tag{B.7}$$

By solving Equations (B.4)-(B.7) we can hence write the n by n precision matrix

$$\mathbf{\Sigma}_{C_{j,n}}^{-1} = \begin{pmatrix} \Sigma_{j,j}^{-1} & \Sigma_{j,j+1}^{-1} & & & & & \\ \Sigma_{j+1,j}^{-1} & \Sigma_{j+1,j+1}^{-1} & \Sigma_{j+1,j+2}^{-1} & & & & \\ & \Sigma_{j+2,j+1}^{-1} & \ddots & & & & \\ & & & \ddots & & & \\ & & & & \ddots & \Sigma_{j+n-1,j+n}^{-1} & \\ & & & & \Sigma_{j+n,j+n-1}^{-1} & \Sigma_{j+n,j+n}^{-1} & \end{pmatrix}. \tag{B.8}$$

For $j + 1 \leq i \leq j + n - 1$, comparing the C_i coefficients gives

$$C_i\Big|_{j+1 \leq i \leq j+n-1} : \frac{\kappa z_i}{2D}(\Delta X_i + \kappa \Delta t_i z_i X_i) = \sum_{m=j}^{j+n} \mu_m \Sigma_{i,m}^{-1}. \tag{B.9}$$

And providing the block does not contain the first or last timepoint, we have

$$C_j : \frac{C_{j-1}}{2\Delta t_{j-1}(D_C z_{j-1} + \hat{D}(1 - z_{j-1}))} + \frac{\kappa z_1}{2D}(\Delta X_j + \kappa \Delta t_j z_j X_j) = \sum_{m=j}^{j+n} \mu_m \Sigma_{j,m}^{-1} \tag{B.10}$$

$$C_{j+n} : \frac{C_{j+n+1}}{2\Delta t_{j+n}(D_C z_{j+n} + \hat{D}(1 - z_{j+n}))} + \frac{\kappa z_{j+n}}{2D}(\Delta X_{j+n} + \kappa \Delta t_{j+n} z_{j+n} X_{j+n}) = \sum_{m=j}^{j+n} \mu_m \Sigma_{j+n,m}^{-1}. \tag{B.11}$$

If the block contains the first timepoint, i.e. $j = 1$ we have

$$C_1 : \tau_{C_1}\mu_{C_1} + \frac{\kappa z_1}{2D}(\Delta X_1 + \kappa \Delta t_1 z_1 X_1) = \sum_{m=j}^{j+n} \mu_m \Sigma_{1,m}^{-1} \tag{B.12}$$

and if the block contains the last timepoint, i.e. $j+n=N$, then

$$C_N : \frac{\kappa z_N}{2D}(\Delta X_N + \kappa \Delta t_N z_N X_N) = \sum_{m=j}^{j+n} \mu_m \Sigma^{-1}_{N,m}. \tag{B.13}$$

To calculate $\boldsymbol{\mu}_{C_{j,n}}$ we solve the system of linear equations

$$\boldsymbol{\Sigma}^{-1}_{C_{j,n}} \boldsymbol{\mu}_{C_{j,n}} = \mathbf{b}_{j,n} \tag{B.14}$$

where $\mathbf{b}_{j,n}$ is a column vector with elements $b_i = \left\{\frac{\kappa z_i}{2D}(\Delta X_i + \kappa \Delta t_i z_i X_i)\right\}_{i=j+1}^{j+n-1}$ and b_j, b_{j+n} from the left hand side of Equations (B.10) and (B.11). (Or Equation (B.12) or (B.13) if $j=1$ or $j+n=N$ respectively.) Since $\boldsymbol{\Sigma}^{-1}_{C_{j,n}}$ is tridiagonal this equation can be efficiently solved, for example using the left matrix division function (with $\boldsymbol{\Sigma}^{-1}_{C_{j,n}}$ as a sparse matrix) in Matlab. Given $\boldsymbol{\mu}_{C_{j,n}}$ and $\boldsymbol{\Sigma}^{-1}_{C_{j,n}}$ the Gibbs update is

$$\mathbf{C}_{j,n} \sim N(\boldsymbol{\mu}_{C_{j,n}}, \boldsymbol{\Sigma}^{-1}_{C_{j,n}}). \tag{B.15}$$

B.4 SI Figures

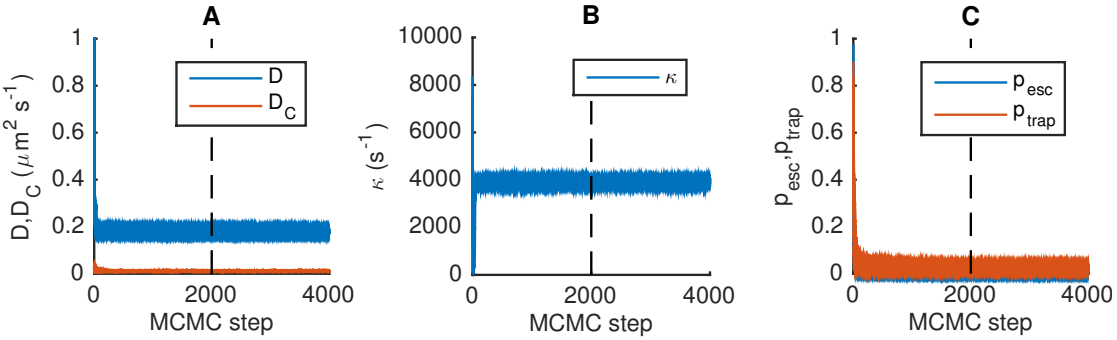

Figure B.3: **Fit of HPW model to a 20nm AuNP/CTxB/GM1 trajectory in a model membrane on glass.** (A) MCMC chains for D (blue) and D_C (red). (B) MCMC chain for κ. (C) MCMC chains for p_{esc} and p_{trap}. For each parameter 12 independent MCMC runs are shown. Corresponding parameter posterior distributions are shown in Fig. 3.6.

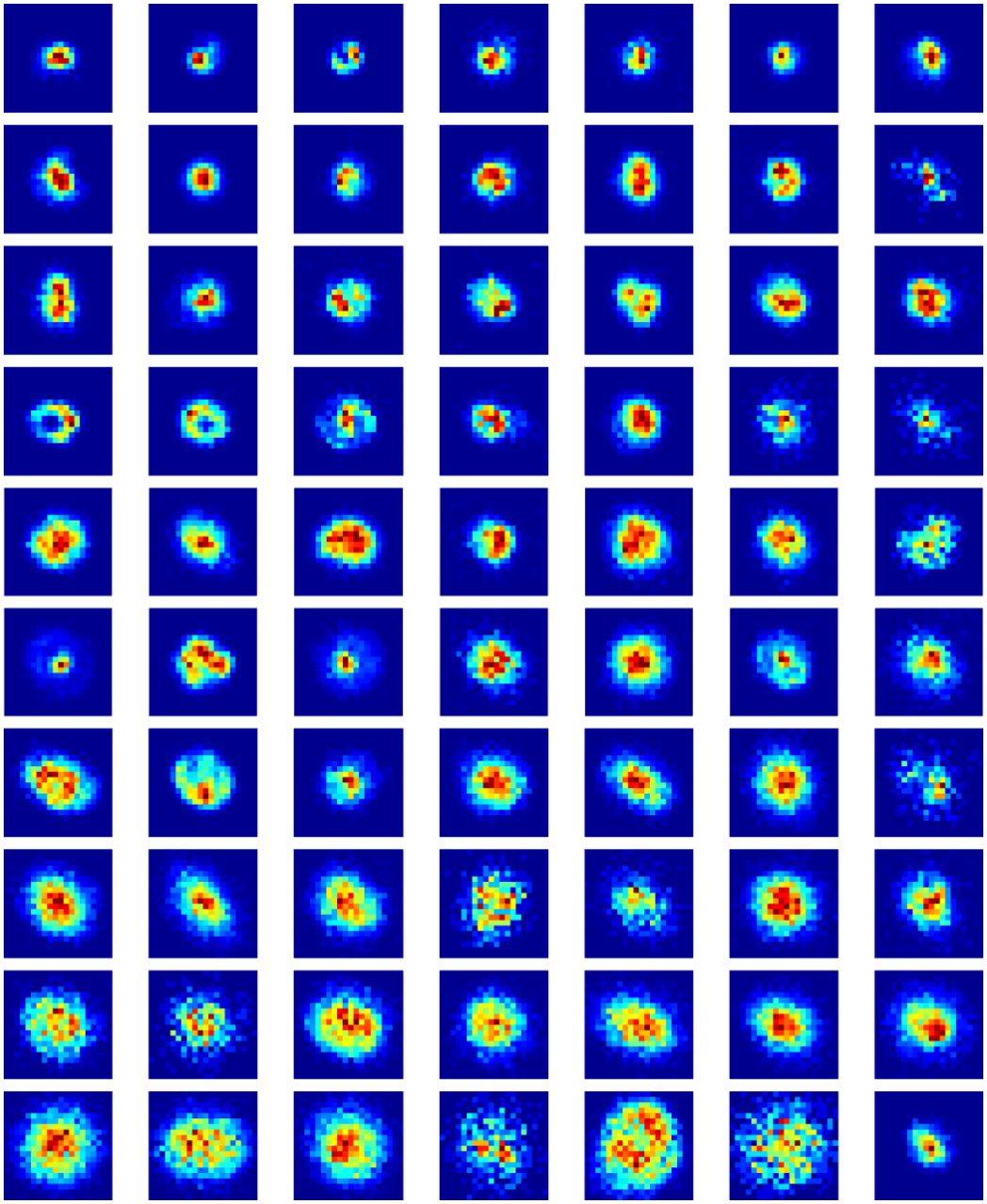

Figure B.4: **Pooled confinement histograms for all AuNP/CTxB/GM1 trajectories.** Histograms of particle positions pooled over confinement events for a trajectory. Trajectories are ordered by the average of the mean confinement radius statistic ($\bar{R}_{lm}$) over all events (i.e. the order in Fig. 3.10C). Confinement events were included based on the criteria which yield 325 events in Table 3.1, except that we also included events revisiting a previous trapping zone (i.e. we did not enforce the condition that confinement event centres have to be 30nm away from all previous centres in that trajectory). Each plot has side length 0.1 µm.

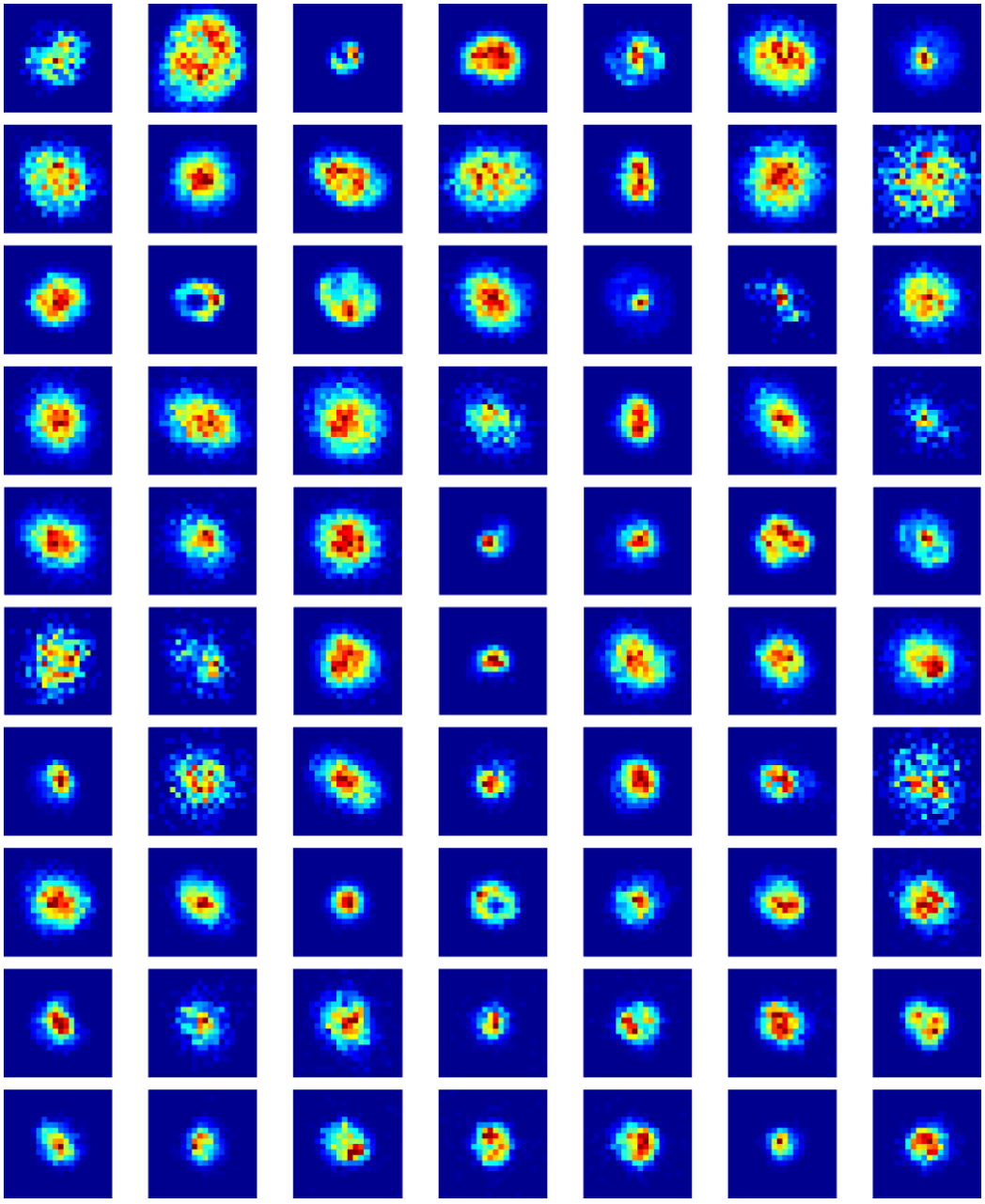

Figure B.5: **Pooled confinement histograms for all AuNP/CTxB/GM1 trajectories.** Histograms of particle positions pooled over confinement events for a trajectory. Trajectories are ordered by the average of the radial skewness statistic (S_{lm}) over all events (i.e. the order in Fig. 3.10D). Confinement events were included based on the criteria which yield 325 events in Table 3.1, except that we also included events revisiting a previous trapping zone (i.e. we did not enforce the condition that confinement event centres have to be 30nm away from all previous centres in that trajectory). Each plot has side length 0.1 µm.